Teemu Peltonen

Superconductivity from Graphene with Flat Bands

UNIVERSITY OF JYVÄSKYLÄ

FACULTY OF MATHEMATICS
AND SCIENCE

Teemu Peltonen

Superconductivity from Graphene with Flat Bands

Esitetään Jyväskylän yliopiston matemaattis-luonnontieteellisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi
syyskuun 4. päivänä 2020 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Mathematics and Science of the University of Jyväskylä,
on September 4, 2020 at 12 o'clock noon.

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2020

Author   Teemu Peltonen
                Department of Physics
                Nanoscience Center
                University of Jyväskylä
                Jyväskylä, Finland

Supervisor Professor Tero Heikkilä
                Department of Physics
                Nanoscience Center
                University of Jyväskylä
                Jyväskylä, Finland

Reviewers   Associate Professor Enrico Rossi
                Department of Physics
                College of William and Mary
                Williamsburg, Virginia, USA

                Professor Hyoung Joon Choi
                Department of Physics
                Yonsei University
                Seoul, South Korea

Opponent   Professor Jörg Schmalian
                Institute for Theoretical Condensed Matter physics
                Karlsruhe Institute of Technology
                Karlsruhe, Germany

Preface

The research discussed in this thesis was carried out during the years 2015–2020 in the Condensed Matter Theory group of Prof. Tero T. Heikkilä at the department of physics at the University of Jyväskylä. I am grateful to my supervisor Prof. Tero T. Heikkilä for providing me the opportunity to pursue a PhD degree in his group, and for guiding the work towards the common goal. I feel that I have learned a lot during these years.

I am not sure if I can thank Mr. Risto Ojajärvi enough. You provided help with almost every calculation presented in this thesis. It was very easy to discuss physical, mathematical, and computational issues with you, as we really had a common language. The thesis would not have been nearly as good without you.

A big thanks goes to one of my master's thesis supervisors Dr. Pekka Koskinen, who showed me support also during the (sometimes difficult) PhD studies and provided help with some of the physics problems. Pekka also helped forming the final title of the thesis together with Risto.

I also want to thank my other coauthors, Dr. Aleksi Julku, Dr. Long Liang, and Prof. Päivi Törmä. It was great working with you: Aleksi provided continuous help with many of the computational details, Long helped in understanding the physics of the superfluid weight, and Päivi was super efficient in working with the paper. Furthermore I had the pleasure to work with my pleasant colleagues Mr. Souvik Agasti, Dr. Faluke Aikebaier, Ms. Aili Asikainen, Dr. Subrata Chakraborty, Dr. Philippe Djorwé, Dr. Timo Hyart, Mr. Kalle Kansanen, Dr. Yao Lu, Dr. Mihail Silaev, and Dr. Pauli Virtanen.

One of the greatest parts of the long journey has been the time with my closest friends at the university: Mr. Antti Hämäläinen, Mr. Otto Ikäheimonen, Mr. Ville Kivioja, Mr. Joonas Korhonen, Mr. Konsta Kurki, Ms. Laura Laulumaa, Ms. Terhi Moisala, and Mr. Risto Ojajärvi. I have studied a lot together with you, and now during the PhD studies we had some great lunch discussions varying between politics and very deep mathematics. I want to specifically thank Ville for explaining some complex but relevant (and also not so relevant) mathematical concepts, which are difficult to understand if explained by an average physicist but easy to understand if explained by a mathematician, at least to me. Laura provided help in solving some quantum-mechanical and computational problems from her cosmological point of view, which turned out to significantly guide me forward. Terhi's master's thesis really helped me get going in the beginning of the project, as did Risto's special assignment work.

Equally as important were all my friends outside the university. I would not have survived this journey without spending so much great time with you, whether it was skydiving, mountain biking, playing beach volley, some other sport, or simply hanging out. The list is long and you know who you are, but I want to highlight the people I spent probably the most time with: both Sallas, Joonas, and the wonderful family of Laura, Matti, and Elo.

Importantly, I would like to thank my parents for their constant support and interest. Talking about support, my loved fiancé Anna provided it unconditionally even during the hard times, which I am truly grateful of. A big thanks goes also to those people who showed interest in my work during the years, especially my godparents Marjatta and Pekka. Explaining my studies to people outside science made me realize how fascinated I actually was.

Jyväskylä, March 2020

Teemu Peltonen

Abstract

In this thesis superconductivity in two realistic flat-band materials is theoretically studied. In the first material, periodically strained graphene (PSG), one layer of graphene is strained in a periodic manner, creating flat electronic bands. It has been demonstrated experimentally, but no superconductivity has been observed. The second material, twisted bilayer graphene (TBG), shook the entire field of condensed matter physics after both superconductivity and correlated insulating states were reported in a breakthrough experiment in 2018, when two layers of graphene were stacked with a relative rotation angle of 1°. In TBG it is the interlayer coupling that creates flat bands when it follows the resulting periodic moiré pattern. The motivation of this thesis is to show that simple, conventional Bardeen–Cooper–Schrieffer (BCS) theory of superconductivity with spin-singlet, s-wave pairing is consistent with the many superconductivity experiments performed in TBG so far, and that the same theory predicts strongly enhanced superconductivity also in PSG. This is in contrast to the vast majority of theoretical studies suggesting an unconventional origin of superconductivity in TBG. Moreover, it is argued that superconductivity in PSG is way more tunable than in TBG, and a recipe for reaching higher superconducting critical temperatures is presented.

In the first part a generic introduction to superconductivity is given. Especially an overview of BCS theory is presented, and the following predictions of the superconducting order parameter and critical temperature are derived in a few regions of the parameter space to show why superconductivity in a flat-band system is special, possibly allowing us to realize room-temperature superconductivity in the future. Also the superfluid weight and the Berezinskii–Kosterlitz–Thouless (BKT) phase transition is introduced, because they give a more accurate description of the superconducting phase transition in two dimensions. Moreover, basics of the classification of superconductors are covered, as in the graphene literature various types of superconducting phases are identified. Lastly, the Fourier series, perhaps the most important tool in the case of periodic systems, is introduced in the elegant notation of quotient groups.

In the second part PSG of Publication I is discussed. Experimental realizations of periodic strain are presented, the continuum model of the noninteracting state is derived together with the BCS equations describing the interacting, superconducting state, and the results are discussed both in the normal and superconducting states. In the normal state flat bands are formed when the pseudomagnetic field resulting from the strain is strong enough, and the low-energy states are localized

to the extrema of this field. In the superconducting state a highly inhomogeneous order parameter is identified, with the localization pattern being determined by the local density of states. Moreover it is shown that PSG behaves like a flat-band superconductor when the interaction strength and the pseudomagnetic field are strong enough. In the end it is argued that one experiment might already be close to achieving superconductivity in this material: if the strain amplitude in those experiments could be increased by a factor of four, the transition temperature would be already of the order of $4\,\mathrm{K}$.

In the third part TBG of Publications II and III is discussed. First the basic geometry of TBG is introduced: the periodic moiré pattern, the commensurate rotation angles, and the structure of the superlattice. Then the continuum model is derived and the BCS equations are argued to be almost the same as for PSG. Results in the normal state are discussed, with emphasis given to the bandstructure and the flat bands. Also the main results regarding superconductivity are given, where it is shown that TBG is qualitatively highly similar to PSG. It is also argued that the results are consistent with the available experimental data, and that by using another kind of pairing mechanism yields an observable difference in the superfluid weight, which could be experimentally measured to distinguish between the different mechanisms. Moreover, an introduction to the vast TBG literature is given to show how rich the physics of TBG is, and to also compare to our predictions.

The MATHEMATICA notebooks used to calculate all the results of Publication I, most of the results of Publication II, and all the results of Publication III regarding the Dirac Point (DP) method are publicly available at https://gitlab.jyu.fi/jyucmt/psg-and-tbg.

Keywords: graphene, strain engineering, twisted bilayer graphene, flat bands, superconductivity, BCS theory

Tiivistelmä

Tässä väitöskirjatyössä tutkitaan teoreettisesti suprajohtavuutta kahdessa realistisessa materiaalissa. Ensimmäisessä materiaalissa, jaksollisesti venytetyssä grafeenissa (periodically strained graphene, PSG), yhtä grafeenitasoa venytetään jaksollisella tavalla, saaden aikaan elektronisia tasovöitä. Tämä on jo todennettu kokeellisesti, mutta suprajohtavuutta tässä materiaalissa ei toistaiseksi ole havaittu. Toinen materiaali, kierretty kaksikerrosgrafeeni (twisted bilayer graphene, TBG), ravisutti koko tiiviin aineen fysiikan alaa kun siinä havaittiin sekä suprajohtavuutta että korreloituneita eristäviä tiloja läpimurtokokeessa vuonna 2018, kun kaksi grafeenitasoa asetettiin päällekkäin yhden asteen verran käännettynä toisiinsa nähden. TBG:ssä tasovyöt saa puolestaan aikaan tasojen välinen kytkentä, joka seurailee jaksollista moiré-kuviota. Tämän väitöskirjatyön tarkoituksena on osoittaa, että yksinkertainen ja tavanomainen suprajohtavuuden Bardeen–Cooper–Schrieffer-teoria (BCS) spin-singletti- ja s-aaltopariutumisella on yhteensopiva useiden TBG-suprajohtavuuskokeiden kanssa, ja että sama teoria ennustaa vahvasti tehostunutta suprajohtavuutta myös PSG:ssä. Tämä poikkeaa suuresti valtaosasta muita teoreettisia tutkimuksia, jotka esittävät TBG:n suprajohtavuuden johtuvan jostakin epätavanomaisesta mekanismista. Lisäksi väitöskirjassa esitetään kuinka suprajohtavuus PSG:ssä on huomattavasti muokattavampaa kuin TBG:ssä, ja kuinka sen avulla saatetaan päästä kohti korkeampia kriittisiä lämpötiloja.

Työn ensimmäisessä osiossa esitetään yleinen johdanto suprajohtavuuteen. Erityisesti tehdään yleiskatsaus BCS-teoriasta ja johdetaan sen antamat ennusteet suprajohtavalle järjestysparametrille ja kriittiselle lämpötilalle muutamissa parametriavaruuden osissa, jotta nähdään miksi suprajohtavuus tasovyösysteemissä on erityistä, kenties mahdollistaen huoneenlämpötilan suprajohtavuuden tulevaisuudessa. Myös supranestetiheyden ja Berezinskii–Kosterlitz–Thouless-transitiolämpötilan (BKT) käsitteet määritellään, koska ne kuvaavat tarkemmin suprajohtavaa faasitransitiota kahdessa ulottuvuudessa. Lisäksi käsitellään suprajohteiden luokittelun perusteet, koska grafeenikirjallisuudessa esiintyy useita erilaisia suprajohtavia faaseja. Lopuksi esitellään Fourier-sarja, kenties tärkein yksittäinen työkalu tutkittaessa jaksollisia systeemejä, elegantissa tekijäryhmänotaatiossa.

Toisessa osiossa käsitellään Julkaisun I PSG:tä. Siinä esitellään jaksollisen venymän kokeellisia toteutuksia, johdetaan vuorovaikuttamattoman tilan jatkumomalli yhdessä vuorovaikuttavan, suprajohtavan, tilan BCS-yhtälöiden kanssa ja esitetään tulokset koskien sekä normaalia että suprajohtavaa tilaa. Havaitaan, että normaalitilassa syntyy tasovöitä kun venymästä aiheutuva pseudomagneet-

tinen kenttä on tarpeeksi vahva, ja lisäksi matalaenergiatilat ovat lokalisoituneet tämän kentän ääriarvokohtiin. Suprajohtavassa tilassa havaitaan erittäin epähomogeeninen järjestysparametri, jonka lokalisoitumisen määrää lokaali tilatiheys. Lisäksi näytetään, että PSG käyttäytyy kuten tasovyösuprajohde kun vuorovaikutuksen voimakkuus ja pseudomagneettinen kenttä ovat tarpeeksi suuria. Lopuksi perustellaan miksi eräs kokeellinen tutkimus saattaa olla jo lähellä suprajohtavuuden havaitsemista tässä materiaalissa: jos venymän amplitudia kyseisessä kokeessa pystyttäisiin kasvattamaan nelinkertaiseksi, transitiolämpötila olisi jo neljän kelvinin luokkaa.

Kolmannessa osiossa käsitellään Julkaisujen II ja III TBG:tä. Aluksi tutustutaan TBG:n geometriaan: jaksolliseen moiré-kuvioon, yhteismitallisiin kiertokulmiin ja superhilan rakenteeseen. Tämän jälkeen johdetaan eräs TBG:n jatkumomalli ja näytetään BCS-yhtälöiden olevan lähes samat kuin PSG:ssä. Normaalitilan tulokset esitellään keskittyen vyörakenteeseen ja tasovöihin. Myös suprajohtavuuden päätulokset esitellään, jolloin selviää että TBG käyttäytyy kvalitatiivisesti hyvin samankaltaisesti kuin PSG. Lisäksi osoitetaan tulosten olevan yhteensopivia lukuisten kokeiden kanssa. Edelleen huomataan, että toisenlainen pariutumismekanismi johtaa havaittavaan eroon supranestetiheydessä, joka voitaisiin havaita kokeellisesti ja näin erotella mikä mekanismi on suprajohtavuuden takana. Lopuksi näytetään kuinka rikasta TBG:n fysiikka on antamalla johdatus laajaan TBG-kirjallisuuteen, ja vertaillaan kuinka nämä vertautuvat antamiimme ennusteisiin.

MATHEMATICA-muistiot, joita on käytetty Julkaisun I kaikkien tulosten laskemiseen, Julkaisun II lähes kaikkien tulosten laskemiseen ja Julkaisun III Dirac Point (DP) -menetelmään liittyvien tulosten laskemiseen, ovat julkisesti saatavilla osoitteessa https://gitlab.jyu.fi/jyucmt/psg-and-tbg.

Asiasanat: grafeeni, jännityskenttä, kierretty kaksikerrosgrafeeni, tasovyöt, suprajohtavuus, BCS-teoria

List of Publications

This thesis consists of an introductory part and the following publications:

I Teemu J. Peltonen and Tero T. Heikkilä, Flat-band superconductivity in periodically strained graphene: mean-field and Berezinskii–Kosterlitz–Thouless transition, *Journal of Physics: Condensed Matter* **32**, 365603 (2020)

II Teemu J. Peltonen, Risto Ojajärvi, and Tero T. Heikkilä, Mean-field theory for superconductivity in twisted bilayer graphene, *Physical Review B* **98**, 220504(R) (2018)

III Aleksi Julku, Teemu J. Peltonen, Long Liang, Tero T. Heikkilä, and Päivi Törmä, Superfluid weight and Berezinskii–Kosterlitz–Thouless transition temperature of twisted bilayer graphene, *Physical Review B* **101**, 060505(R) (2020)

Author's contribution

The author developed the theory, carried out the numerical calculations, and wrote the manuscript of Publication I. The author developed most of the theory, carried out the numerical calculations, and wrote a significant part of the manuscript of Publication II. The author developed the theory and performed the numerical calculations of Publication III regarding the Dirac point (DP) method, and wrote parts of the manuscript.

Contents

Chapter 1

Opening

1.1   Introduction

Superconductivity, the lack of electrical resistance below some critical temperature T_c, has interested physicists for over a century after its discovery in 1911 by Heike Kamerlingh Onnes. Nowadays it is routinely used in some real-life applications, such as creating high magnetic fields in magnetic resonance imaging machines in hospitals or in the Large Hadron Collider in CERN. Unfortunately, the critical temperature is usually very low, a few degrees above the absolute zero, constraining the possible applications to systems where a liquid helium cooling can be arranged.

Since the finding of superconductivity, physicists have tried to find ways to increase the critical temperature, with the ultimate goal of reaching room-temperature superconductivity. Some progress was made immediately when researchers turned from elemental materials to compounds, especially niobium based, and by 1980 a critical temperature exceeding $20\,\mathrm{K}$ was already reality [1]. A true breakthrough was made in 1986, when *high-temperature superconductivity* was found in cuprates (copper oxides) [2]. The initial records were $T_c = 35\,\mathrm{K}$, but nowadays the highest recorded T_c (in ambient pressure) of cuprates is $133\,\mathrm{K}$ [3], which is already above liquid nitrogen temperatures, but still much below the room temperature. Applying immense pressure increases T_c all the way to $250\,\mathrm{K}$ in lanthanum hydride [4], but a pressure of the order of $100\,\mathrm{GPa}$, which is comparable to the pressure in the core of the earth, is not really practical in applications.

One road to a room-temperature T_c might be *flat-band superconductivity*. These are systems with one or more flat electronic bands, which means singular density of states that highly amplifies electronic correlations, being behind *e.g.* superconductivity. Flat-band superconductivity and flat-band systems in general have been studied already from the 1990s by a small group of people [5–8], but the breakthrough finding of superconductivity at $T_c \approx 1\,\mathrm{K}$ in *twisted bilayer graphene* (TBG) in 2018 [9] truly exploded the amount of interest to flat-band superconductivity. This is probably because the phase diagram of TBG highly resembles that of the high-temperature cuprate superconductors, so that researchers

are hoping for TBG to resolve the mystery of high-temperature superconductivity. A key difference is that TBG is much easier to manufacture and study: to move in the doping-axis of the phase diagram, in cuprates a new sample is needed for each doping level, whereas in TBG a simple turn of a knob controlling electrical bias voltage is enough.

In this thesis I take a different viewpoint, the viewpoint taken also in the original flat-band papers. Instead of merely trying to understand cuprates (and other known high-temperature superconductors), I try to promote flat-band superconductivity *in its own right* by showing how flat bands can be used to increase the critical temperature also in graphene-based materials. Especially I try to show that superconductivity in TBG may be *conventional*, as described by the standard (although slightly generalized) Bardeen–Cooper–Schrieffer (BCS) theory, instead of *unconventional* as in the cuprates. The unconventional origin of superconductivity is promoted by a vast majority of the theory papers and even the first experimental paper [9] itself. In Publication I we study also *periodically strained graphene* (PSG), which we show to be very similar to TBG, but being theoretically simpler and also in principle more tunable, thus probably allowing to reach a higher critical temperature.

The thesis is organized as follows. In Chapter 1 I cover a few background topics that are needed in understanding the later chapters and the presented publications. In Chapter 2 I discuss periodically strained graphene, covered in Publication I. In Chapter 3 I discuss twisted bilayer graphene, covered in Publications II and III. Finally in Chapter 4 I present conclusions of the thesis and provide an outlook to the future.

1.2   BCS theory of superconductivity

Presented not until over 50 years after the discovery of superconductivity, the Nobel-prize-worth Bardeen–Cooper–Schrieffer (BCS) theory has been successful in explaining properties of many superconductors, especially the elemental ones with a low critical temperature. In BCS theory[1] the basic ingredient is an effective *attractive* interaction between electron pairs, which in the "standard" BCS theory is seen to originate from the electron–phonon coupling.

Before the finding of high-temperature superconductors, the BCS theory was thought to explain all superconductors. Especially its prediction of exponentially suppressed critical temperature spread the belief that superconductivity is only possible with T_c of a few kelvin [12]. The finding of the high-temperature cuprate superconductors led people to search for alternative theories with a different mechanism. There is, however, also a middle ground: BCS theory that is not "standard" but slightly generalized by removing some of the original assumptions. One interesting case is to remove the assumption of a constant density of states (DOS) and replace it by the singular DOS present in a flat-band system. This significantly enhances the critical temperature and provides an alternative way

---

[1]By BCS theory I mean it in a generalized sense, in the form presented in this thesis. By "standard" BCS theory I mean the one usually presented in textbooks such as Refs. [10, 11], which is a certain limiting case as presented in Sec. 1.2.3.

towards high-temperature superconductivity, through the conventional electron–phonon coupling.

In this section I present the setup of the theory in the generalized sense. The explicit Bogoliubov–de Gennes equation and the self-consistency equation that have to be solved are derived in Chapter 2 (or in Publication I) in the case of PSG.

1.2.1   Setup of the theory

To model superconductivity, one starts from the generic second-quantized interacting Hamiltonian (in 2D)

$$H_{\mathrm{BdG}} \coloneqq H + H_{\mathrm{int}}, \tag{1.1}$$

where the noninteracting Hamiltonian is

$$H = \sum_{\sigma} \int_{\mathbb{R}^2/L_{\mathrm{BK}}} \mathrm{d}\boldsymbol{r}\, \psi_\sigma^\dagger(\boldsymbol{r}) \mathcal{H}_\sigma(\boldsymbol{r}) \psi_\sigma(\boldsymbol{r}) \tag{1.2}$$

with

$$\mathcal{H}_\sigma(\boldsymbol{r}) \coloneqq -\frac{\hbar^2 \nabla^2}{2m} + V_\sigma(\boldsymbol{r}) - \mu_\sigma(\boldsymbol{r}) \tag{1.3}$$

and the interaction Hamiltonian between pairs of electrons is

$$H_{\mathrm{int}} \coloneqq \frac{1}{2} \sum_{\sigma\sigma'} \int \mathrm{d}\boldsymbol{r} \int \mathrm{d}\boldsymbol{r}'\, \lambda_{\sigma\sigma'}(\boldsymbol{r},\boldsymbol{r}') \psi_\sigma^\dagger(\boldsymbol{r}) \psi_{\sigma'}^\dagger(\boldsymbol{r}') \psi_{\sigma'}(\boldsymbol{r}') \psi_\sigma(\boldsymbol{r}). \tag{1.4}$$

Here $\sigma \in \{\uparrow, \downarrow\}$ denotes spin, $\psi_\sigma(\boldsymbol{r})$ is the electron field operator, m is the electron mass, V is the external potential, μ is the chemical potential, $\lambda_{\sigma\sigma'}(\boldsymbol{r},\boldsymbol{r}') \coloneqq \langle \sigma\boldsymbol{r} | \hat{\lambda} | \sigma'\boldsymbol{r}' \rangle \in \mathbb{R}$ is the interaction matrix element between a σ-spin electron at position $\boldsymbol{r}$ and another σ'-spin electron at position $\boldsymbol{r}'$. In BCS theory an *attractive interaction* between electrons is the key element of superconductivity, here modeled with $\lambda < 0$. Here everything is taken to be periodic in translations of the large Born–von Kármán lattice L_{BK}, and the integral is over the Born–von Kármán unit cell, which in the quotient-group notation of Sec. 1.4.1 is written as $\mathbb{R}^2/L_{\mathrm{BK}}$.

Next in BCS theory we make the *mean-field approximation* in the Cooper channel, that is, we say that $\psi_\nu \psi_{\nu'} \approx \langle \psi_\nu \psi_{\nu'} \rangle + \delta\psi_{\nu\nu'}$ is approximated by the mean field $\langle \psi_\nu \psi_{\nu'} \rangle$ plus small fluctuations $\delta\psi_{\nu\nu'}$ around this, with $\delta\psi^2 \approx 0$. Here the angle brackets denote the thermal average [13]. This yields

$$\psi_\mu^\dagger \psi_{\mu'}^\dagger \psi_\nu \psi_{\nu'} \approx \langle \psi_{\mu'} \psi_\mu \rangle^* \psi_\nu \psi_{\nu'} + \langle \psi_\nu \psi_{\nu'} \rangle \psi_\mu^\dagger \psi_{\mu'}^\dagger - \langle \psi_{\mu'} \psi_\mu \rangle^* \langle \psi_\nu \psi_{\nu'} \rangle, \tag{1.5}$$

which for the interaction Hamiltonian means

$$H_{\mathrm{int}} \approx \frac{1}{2} \sum_{\sigma\sigma'} \int \mathrm{d}\boldsymbol{r} \int \mathrm{d}\boldsymbol{r}'\, \Delta_{\sigma\sigma'}(\boldsymbol{r},\boldsymbol{r}') \psi_\sigma^\dagger(\boldsymbol{r}) \psi_{\sigma'}^\dagger(\boldsymbol{r}') + \mathrm{h.c.} + \mathrm{const} \tag{1.6}$$

with

$$\Delta_{\sigma\sigma'}(\boldsymbol{r},\boldsymbol{r}') \coloneqq \lambda_{\sigma\sigma'}(\boldsymbol{r},\boldsymbol{r}') \langle \psi_{\sigma'}(\boldsymbol{r}') \psi_\sigma(\boldsymbol{r}) \rangle \tag{1.7}$$

being the *superconducting order parameter* determining the transition between the normal and superconducting states. The value of Δ then depends on the interaction matrix element λ and on the noninteracting Hamiltonian H through the expectation value.

The above-discussed position basis is one possibility to present BCS theory [11]. But often, for example when dealing with lattices[2] and crystals, it is convenient to change to some other basis (*e.g.* localized orbitals) $\phi_{\sigma\mu}(\boldsymbol{r})$ [14] through

$$\psi_\sigma(\boldsymbol{r}) = \sum_{\mu \in \text{states}} \phi_{\sigma\mu}(\boldsymbol{r})\psi_\mu, \tag{1.8}$$

where μ denotes the relevant states (*e.g.* lattice sites, sublattices, spins, orbitals) and ψ_μ is the annihilation operator of that state. This transforms the Hamiltonians to

$$H = \sum_{\mu\mu' \in \text{states}} \psi_\mu^\dagger \mathcal{H}_{\mu\mu'} \psi_{\mu'}, \tag{1.9}$$

$$H_\text{int} = \frac{1}{2} \sum_{\mu\mu' \in \text{states}} \psi_\mu^\dagger \Delta_{\mu\mu'} \psi_{\mu'}^\dagger + \text{h.c.} + \text{const}, \tag{1.10}$$

where the order parameter is

$$\Delta_{\mu\mu'} := \sum_{\nu\nu' \in \text{states}} \lambda_{\mu\mu'\nu\nu'} \langle \psi_{\nu'}\psi_\nu \rangle \tag{1.11}$$

and

$$\mathcal{H}_{\mu\mu'} := \sum_\sigma \int \mathrm{d}\boldsymbol{r}\, \phi_{\sigma\mu}^*(\boldsymbol{r})\mathcal{H}_\sigma(\boldsymbol{r})\phi_{\sigma\mu'}(\boldsymbol{r}), \tag{1.12}$$

$$\lambda_{\mu\mu'\nu\nu'} := \sum_{\sigma\sigma'} \int \mathrm{d}\boldsymbol{r} \int \mathrm{d}\boldsymbol{r}'\, \lambda_{\sigma\sigma'}(\boldsymbol{r},\boldsymbol{r}')\phi_{\sigma\mu}^*(\boldsymbol{r})\phi_{\sigma'\mu'}^*(\boldsymbol{r}')\phi_{\sigma'\nu'}(\boldsymbol{r}')\phi_{\sigma\nu}(\boldsymbol{r}). \tag{1.13}$$

I could in principle derive the full BCS theory from the generic representation (1.9) and (1.10), as is done in the book of Nazarov & Danon [15]. But because in all the publications we use *low-energy continuum* theory derived from the corresponding tight-binding Hamiltonian, I find it more transparent to derive the BCS equations for the continuum system directly. This we do in Chapter 2 and in Publication I in the case of PSG. For TBG the generalization is then obvious, as presented in Chapter 3.

There is one special basis requiring an extra mentioning, that is, the eigenbasis of the noninteracting Hamiltonian H. It can be diagonalized through $\mathcal{H} =: U^\dagger \epsilon U$ and $c := U\psi$ as

$$H = \sum_{\mu \in \text{eigenstates}} c_\mu^\dagger \epsilon_\mu c_\mu, \tag{1.14}$$

---

[2]Whenever I say lattice, I mean a Bravais lattice. More generally, if the system is composed of one or more intervening lattices (*i.e.* it has a "basis"), I call it a crystal.

where U is a unitary transformation and ϵ is a diagonal matrix of the eigenenergies. Then with the transformation $U \Delta U^{\mathsf{T}} \mapsto \Delta$ and a corresponding transformation for λ the interaction Hamiltonian becomes

$$H_{\text{int}} = \frac{1}{2} \sum_{\mu\mu' \in \text{eigenstates}} c_\mu^\dagger \Delta_{\mu\mu'} c_{\mu'} + \text{h.c.} + \text{const} \tag{1.15}$$

$$\Delta_{\mu\mu'} = \sum_{\nu\nu' \in \text{eigenstates}} \lambda_{\mu\mu'\nu\nu'} \langle c_{\nu'} c_\nu \rangle. \tag{1.16}$$

According to Bloch's theorem, in the case of a (periodic) crystal the eigenstates can be indexed through the crystal momentum $\boldsymbol{k}$ and a band index b, in addition to the spin σ (whenever it is a good quantum number). Then

$$H = \sum_{\sigma \boldsymbol{k} b} c_{\sigma\boldsymbol{k}b}^\dagger \epsilon_{\sigma\boldsymbol{k}b} c_{\sigma\boldsymbol{k}b}, \tag{1.17}$$

$$H_{\text{int}} = \frac{1}{2} \sum_{\sigma\sigma\boldsymbol{k}\boldsymbol{k}'bb'} c_{\sigma\boldsymbol{k}b}^\dagger \Delta_{\sigma\boldsymbol{k}b,\sigma'\boldsymbol{k}'b'} c_{\sigma'\boldsymbol{k}'b'}^\dagger + \text{h.c.} + \text{const}, \tag{1.18}$$

$$\Delta_{\sigma\boldsymbol{k}b,\sigma'\boldsymbol{k}'b'} = \sum_{\tau\tau'\boldsymbol{l}\boldsymbol{l}'aa'} \lambda_{\sigma\boldsymbol{k}b,\sigma'\boldsymbol{k}'b',\tau\boldsymbol{l}a,\tau'\boldsymbol{l}'a'} \langle c_{\tau'\boldsymbol{l}'a'} c_{\tau\boldsymbol{l}a} \rangle. \tag{1.19}$$

This form is used when deriving different limits of BCS theory in Sec. 1.2.3.

1.2.2 Classification of superconductors

In the publications we employ BCS theory with the choice of spin-singlet, s-wave pairing describing conventional superconductivity. In this section I define what these concepts and this kind of classification mean, especially because they are used a lot in the (graphene) superconductivity literature. Also in Publication III non-s-wave pairing states are identified, so it is useful to have some idea of the nomenclature. Although it might at first seem that I am only considering mean-field theory in the following classification, one can always define the order parameter by an expectation value as in Eq. (1.7) or (1.11), even without any approximations.

In general, any system (*i.e.* its Hamiltonian) has in the normal state the $U(1)$ gauge symmetry, and a nonvanishing order parameter $\Delta \in \mathbb{C}$ in Eq. (1.7) breaks this symmetry. We then define [13, 16]

- Δ breaks exactly $U(1)$ symmetry $\quad \Leftrightarrow \quad$ *conventional superconductivity*

- Δ breaks $U(1)$ symmetry and other symmetries of the system $\quad \Leftrightarrow$ *unconventional superconductivity.*

Now that $\lambda_{\mu\mu'\nu\nu'} = \lambda_{\mu'\mu\nu'\nu}$ and the fermionic operators in Eq. (1.11) anticommute, we get the constraint

$$\Delta_{\mu\mu'} = -\Delta_{\mu'\mu} \tag{1.20}$$

for the order parameter. For concreteness, and for the classification scheme [16], let us fix the so-far abstract state index to be $\mu = (\sigma, \boldsymbol{r}, b)$, where $\sigma \in \{\uparrow, \downarrow\}$

denotes spin, $\boldsymbol{r}$ is a position (a continuous variable in the case of a continuum and a discrete variable in the case of a lattice/crystal with localized orbitals) and b includes all the other quantum numbers (sublattice, layer, *etc.*, or "orbital" in general) needed to define the state. Then the constraint (1.20) reads

$$\Delta_{\sigma b,\sigma' b'}(\boldsymbol{r},\boldsymbol{r}') = -\Delta_{\sigma' b',\sigma b}(\boldsymbol{r}',\boldsymbol{r}). \tag{1.21}$$

Defining the spin-permutation operator S, the position-permutation operator P (called also the parity operator although it only inverts the relative coordinate of the pair and not the whole space), and the orbital-permutation operator O, we may also write this as[3]

$$SPO\Delta_{\sigma b,\sigma' b'}(\boldsymbol{r},\boldsymbol{r}') = -\Delta_{\sigma b,\sigma' b'}(\boldsymbol{r},\boldsymbol{r}') \quad \Leftrightarrow \quad SPO\Delta = -\Delta. \tag{1.22}$$

This clearly sets a constraint on what kind of superconductivity classes are possible, and is thus the basis of classification [16]. Note that Δ only needs to be an eigenfunction of SPO, but not necessarily of the individual operators. But if it happens to be an eigenfunction of S, we define [16, 17]

- $S\Delta = -\Delta \quad \Leftrightarrow \quad$ *spin singlet*

- $S\Delta = +\Delta \quad \Leftrightarrow \quad$ *spin triplet.*

On the other hand if Δ happens to be an eigenfunction of O we define similarly [18–21]

- $O\Delta = -\Delta \quad \Leftrightarrow \quad$ *orbital singlet*

- $O\Delta = +\Delta \quad \Leftrightarrow \quad$ *orbital triplet.*

If Δ is an eigenfunction of P, we define [13, 16, 22, 23]

- $P\Delta = -\Delta \quad \Leftrightarrow \quad$ *odd parity*

- $P\Delta = +\Delta \quad \Leftrightarrow \quad$ *even parity.*

The even-parity class includes the so-called s, d, g ...-wave superconductors while the odd parity class includes the p, f, h ...-wave superconductors [16], which I define below. Note that P exchanges the coordinates of the two particles, so that if we write Δ as a function of the center-of-mass coordinate $\boldsymbol{r}_{\mathrm{com}} \coloneqq (\boldsymbol{r} + \boldsymbol{r}')/2$ and the relative coordinate $\boldsymbol{r}_{\mathrm{rel}} \coloneqq \boldsymbol{r} - \boldsymbol{r}'$, P changes the sign of $\boldsymbol{r}_{\mathrm{rel}}$ but leaves $\boldsymbol{r}_{\mathrm{com}}$ untouched. It is thus not the true parity operator of the whole space, but instead only of the relative coordinate of the Cooper pair [16]. The parity classification thus concerns only the relative coordinate [16]. Note also that Eq. (1.22) implies that if Δ is at the same time an eigenfunction of S, P, and O, only four combinations are possible [16]: $(S, P, O) \in \{(-1, -1, -1), (-1, +1, +1), (+1, -1, +1), (+1, +1, -1)\}$.

---

[3]More generally, one should also include the permutation of the time coordinates [16]. But as we started from the time-independent Hamiltonians (1.2) and (1.4), we get no time dependence in Δ and thus discuss only *even-frequency superconductivity* in the language of Ref. [16]. This is enough for the discussion of this thesis, as odd-frequency superconductivity does not appear anywhere in the cited literature (except of Ref. [16]).
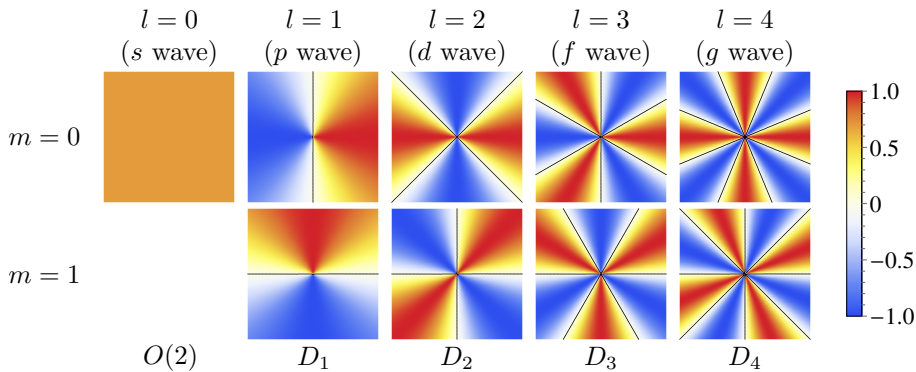
Figure 1.1: Density plot of a few of the lowest-order two-dimensional spherical harmonics Y_l^m, that is, the Fourier basis functions $Y_0^0(\theta) = 1/\sqrt{2}$, $Y_l^0(\theta) = \cos(l\theta)$, $Y_l^1(\theta) = \sin(l\theta)$ for all $l \in \{1, 2, \dots\}$, where $\theta \in [0, 2\pi[$ is the polar coordinate, and the radial coordinate is taken to be constant. The black lines are the nodal lines, *i.e.* lines at which Y_l^m changes sign. Below the plots I also show the point group of the shown pattern, showing how an "l wave" harmonics with $l \in \mathbb{N}$ has the point group D_l.

The symmetry classification (*i.e.* classification to $s, p, d, f, g, \dots$ wave) concerns the expansion of the relative coordinate part of Δ in spherical harmonics [24]. For simplicity let us consider only the two-dimensional case; after all I discuss only two-dimensional materials in this thesis. In two dimensions the spherical harmonic basis is the (sine–cosine) Fourier basis of 2π-periodic functions: $Y_0^0(\theta) := 1/\sqrt{2}$, $Y_l^0(\theta) := \cos(l\theta)$, and $Y_l^1(\theta) := \sin(l\theta)$ for all $l \in \{1, 2, \dots\}$ and[4] $\theta \in \mathbb{R}/(2\pi\mathbb{Z}) \cong [0, 2\pi[$. These are illustrated in Fig. 1.1. The spherical harmonic Y_l^m with any m is called "l wave", with $l = 0, 1, 2, 3, 4, \dots$ corresponding to $s, p, d, f, g, \dots$, with the letters continuing as in the notation of atomic orbitals.

In the spherical harmonic basis we may then expand the angular part of the relative coordinate of the $\boldsymbol{k}$-space version of Δ as[5]

$$\tilde{\Delta}(\boldsymbol{k}_{\mathrm{com}}, \boldsymbol{k}_{\mathrm{rel}}) = \tilde{\Delta}(\boldsymbol{k}_{\mathrm{com}}, \|\boldsymbol{k}_{\mathrm{rel}}\|, \theta_{\mathrm{rel}}) = \sum_{l=0}^{\infty} \sum_{m=0}^{1} b_l^m(\boldsymbol{k}_{\mathrm{com}}, \|\boldsymbol{k}_{\mathrm{rel}}\|) Y_l^m(\theta_{\mathrm{rel}}), \quad (1.23)$$

where I also define $Y_0^1 := 0$ for notational convenience. Note that here the expansion coefficient b_l^m depends on $\boldsymbol{k}_{\mathrm{com}}$ and $\|\boldsymbol{k}_{\mathrm{rel}}\|$, so that the classification may change as a function of these coordinates. If $\tilde{\Delta}$ has a component only along Y_l^m, the superconductor is called "l wave", with $l = 0, 1, 2, 3, 4, \dots$ corresponding to $s, p, d, f, g, \dots$ [13, 24, 25], similarly as in the case of spherical harmonics. Note that $\tilde{\Delta}$ of type l (or Y_l^m; see Fig. 1.1) changes sign on a π rotation exactly when l is odd, and is invariant on a π rotation when l is even. The π rotation corresponds to a sign change of $\boldsymbol{k}_{\mathrm{rel}}$, so it is the same operator as the parity operator P. Thus

---

[4] $A \cong B$ means A and B are isomorphic.

[5] I omit the σ and b indices from Δ as they play no role in this classification.

states with an odd l are called odd parity and even l even parity, consistently with what was defined already above.

In general any linear combination is possible, so if $\tilde{\Delta}$ has components only towards Y_l^m and $Y_{l'}^{m'}$ for some $l \neq l'$ it is called "$l + l'$ *wave*" [13, 24, 25], and so on for linear combinations of more components. For example if $b_0^m \neq 0$ and $b_2^{m'} \neq 0$ are the only nonvanishing components, the state is called $s + d$ *wave*.

There is still freedom in the variable m: linear combinations over $m \in \{0, 1\}$ do not affect the classification above. However, the classification is furthermore refined by defining the state to be "$l \pm \mathrm{i}l$ *wave*" [25, 26] if the only nonvanishing components are Y_l^0 and Y_l^1 and their phase difference is $\pm \pi$, *i.e.* $b_l^1 = \pm \mathrm{i}b_l^0$. For example, the linear combination $Y_2^0 + \mathrm{i}Y_2^1$ is called $d + \mathrm{i}d$ wave. This classification is useful because it defines a very special topological superconducting state [22, 26–29].

The classification above works both in the "continuum", where the non-interacting Hamiltonian H is defined on $\mathbb{R}^2$ and has its point group $G = O(2)$, and in a crystal, where G is reduced to some finite group. But in a crystal one has to be more careful. First of all, not all classes are possible due to the restrictions set by G. For example, if H is defined on a square lattice and has its point group $G = D_4$, the possible point groups for $\tilde{\Delta}$ are the subgroups $D_1, D_2, D_4 \subset D_4$, so that only the $l = 0, 1, 2, 4$ states are possible. Secondly, in this case the $l = 0$ (s-wave) state has only the point group D_4 [instead of $O(2)$], which is the same point group as that of the $l = 4$ (g-wave) state. A linear combination of these states is then called $s + g$ wave [30], similarly as above, or *extended s wave* [30–32]. This is the reason why in crystals a more elegant way to define the pairing symmetry classification would be through group representation theory [13, 24] instead of spherical harmonics.

Lastly, note that the above-used decomposition $\mu = (\sigma, \boldsymbol{r}, b)$ is not unique. For example, in the DP method of Publication III we choose $\boldsymbol{r} \in L$ to be in the (Bravais) lattice L and the orbital index $b = (i, \alpha)$ [denoted by (l, s) in the publication] includes the layer and sublattice. On the other hand, in the RM method of Publication III we choose $\boldsymbol{r} \in SL$ [denoted by i in the publication] to denote the positions of the (Bravais) superlattice SL and the orbital index b [denoted by α in the publication] to denote the sites within one superlattice unit cell, at both sublattices and layers. The corresponding classification then changes accordingly.

1.2.3 Limits of BCS theory and flat-band superconductivity

In order to understand why flat-band superconductivity is special and might help in increasing T_c, let us calculate a couple of limits for the superconducting order parameter Δ in Eq. (1.19). Let us take the interaction

$$\lambda_{\sigma \boldsymbol{k} b, \sigma' \boldsymbol{k}' b', \tau \boldsymbol{l} a, \tau' \boldsymbol{l}' a'} = \delta_{\bar{\sigma}\sigma'} \delta_{\sigma\tau} \delta_{\sigma'\tau'} \delta_{\bar{\boldsymbol{k}}\boldsymbol{k}'} \delta_{\bar{\boldsymbol{l}}\boldsymbol{l}'} \delta_{bb'} \delta_{aa'} \lambda_{\boldsymbol{k}\boldsymbol{l}}, \qquad (1.24)$$

where the bar operator is $\bar{\uparrow} = \downarrow$, $\bar{\downarrow} = \uparrow$, $\bar{\boldsymbol{k}} = -\boldsymbol{k}$. This kind of interaction can be seen to be a rough effective model for the underlying electron–phonon interaction

[15, 33]. The order parameter in Eq. (1.19) then becomes

$$\Delta_{\sigma\boldsymbol{k}b,\sigma'\boldsymbol{k}'b'} = \delta_{\bar{\sigma}\sigma'}\delta_{\bar{\boldsymbol{k}}\boldsymbol{k}'}\delta_{bb'}\Delta_{\sigma\boldsymbol{k}} \quad \text{with} \quad \Delta_{\sigma\boldsymbol{k}} \coloneqq \sum_{la}\lambda_{\boldsymbol{k}l}\left\langle c_{\bar{\sigma}\bar{l}a}c_{\sigma la}\right\rangle, \qquad (1.25)$$

leading to the interaction Hamiltonian [Eq. (1.18)]

$$H_{\text{int}} = \frac{1}{2}\sum_{\sigma\boldsymbol{k}b}c^{\dagger}_{\sigma\boldsymbol{k}b}\Delta_{\sigma\boldsymbol{k}}c^{\dagger}_{\bar{\sigma}\bar{\boldsymbol{k}}b} + \text{h.c.} + \text{const} \qquad (1.26)$$

coupling electrons with opposite spin and momenta, corresponding to the form often presented in textbooks [10, 33].

After some algebra [10, 33, 34] (details of this in real-space formalism is presented in Chapter 2 in the case of PSG), diagonalization of $H_{\text{BdG}} = H + H_{\text{int}}$ leads to

$$\left\langle c_{\bar{\sigma}\bar{l}a}c_{\sigma la}\right\rangle = -\frac{\Delta_{\sigma l}}{2VE_{\sigma la}}\tanh\frac{E_{\sigma la}}{2k_{\text{B}}T} \quad \text{with} \quad E_{\sigma la} \coloneqq \sqrt{(\epsilon_{\sigma la} - \mu)^2 + |\Delta_{\sigma l}|^2}$$
$$(1.27)$$

giving the *self-consistency equation*

$$\Delta_{\sigma\boldsymbol{k}} = -\frac{1}{2V}\sum_{la}\lambda_{\boldsymbol{k}l}\frac{\Delta_{\sigma l}}{E_{\sigma la}}\tanh\frac{E_{\sigma la}}{2k_{\text{B}}T} \approx -\frac{1}{2(2\pi)^2}\sum_a\int \mathrm{d}\boldsymbol{l}\,\lambda_{\boldsymbol{k}l}\frac{\Delta_{\sigma l}}{E_{\sigma la}}\tanh\frac{E_{\sigma la}}{2k_{\text{B}}T},$$
$$(1.28)$$

where V is the area of the system. Finally we choose the simple interaction [10, 15, 33]

$$\lambda_{\boldsymbol{k}l} = \theta(\epsilon_{\text{c}} - |\epsilon_{\boldsymbol{k}} - \mu|)\theta(\epsilon_{\text{c}} - |\epsilon_{l} - \mu|)\lambda =: \theta_{\boldsymbol{k}}\theta_{l}\lambda, \qquad (1.29)$$

restricted to a $[-\epsilon_{\text{c}}, \epsilon_{\text{c}}]$ shell from the Fermi level μ, where ϵ_{c} is a cutoff energy (*e.g.* the Debye energy $\hbar\omega_{\text{D}}$) originating from the electron–phonon interaction. This highly simplified form of the interaction is chosen mainly because it allows calculating analytical results, but still works reasonably well. This gives $\Delta_{\sigma\boldsymbol{k}} = \Delta_{\sigma}\theta_{\boldsymbol{k}}$ with

$$\Delta_{\sigma} \coloneqq -\frac{\lambda}{2(2\pi)^2}\sum_a\int \mathrm{d}\boldsymbol{l}\,\theta_{l}\frac{\Delta_{\sigma}}{E_{\sigma la}}\tanh\frac{E_{\sigma la}}{2k_{\text{B}}T}. \qquad (1.30)$$

Discarding the normal-state solution $\Delta_{\sigma} = 0$ it reads

$$1 = -\frac{\lambda}{2(2\pi)^2}\sum_a\int \mathrm{d}\boldsymbol{l}\,\theta_{l}\frac{1}{E_{\sigma la}}\tanh\frac{E_{\sigma la}}{2k_{\text{B}}T}. \qquad (1.31)$$

Equations (1.25) and (1.29) determine the type of superconductivity when both electrons are within the cutoff-energy ϵ_{c} from μ, in which case $\lambda_{\boldsymbol{k}l} = \lambda$. Firstly, it is spin singlet because $\Delta_{\sigma'\boldsymbol{k}b,\sigma\boldsymbol{k}'b'} = -\Delta_{\sigma\boldsymbol{k}b,\sigma'\boldsymbol{k}'b'}$. Secondly, $\Delta_{\sigma\boldsymbol{k}b,\sigma'\boldsymbol{k}'b'}$ depends only on the COM coordinate $(\boldsymbol{k} + \boldsymbol{k}')/2$ through $\delta_{\bar{\boldsymbol{k}}\boldsymbol{k}'}$ but is independent of the relative coordinate $\boldsymbol{k} - \boldsymbol{k}'$, implying *s*-wave pairing. Lastly, it is orbital triplet because $\Delta_{\sigma\boldsymbol{k}b',\sigma'\boldsymbol{k}'b} = \Delta_{\sigma\boldsymbol{k}b,\sigma'\boldsymbol{k}'b'}$.

"Standard" BCS-theory limit

In "standard" BCS theory, meaning the one originally presented by Bardeen, Cooper, and Schrieffer, we assume the system to have only one band and take the dispersion to be spin independent, $\epsilon_{\sigma \boldsymbol{k} b} = \epsilon_{\boldsymbol{k}}$. Further defining the density of states (per spin) $N(\epsilon) := \frac{1}{(2\pi)^2} \int d\boldsymbol{k}\, \delta(\epsilon - \epsilon_{\boldsymbol{k}})$ allows writing the momentum integral in Eq. (1.31) as an energy integral

$$1 \approx -\frac{\lambda}{2} \int_{\mu - \epsilon_c}^{\mu + \epsilon_c} d\epsilon\, N(\epsilon) \frac{1}{\sqrt{(\epsilon - \mu)^2 + |\Delta_\sigma|^2}} \tanh \frac{\sqrt{(\epsilon - \mu)^2 + |\Delta_\sigma|^2}}{2k_B T}. \quad (1.32)$$

We also assume that most of the contribution comes from close to the Fermi level and the DOS changes slowly there, *i.e.*

$$N(\epsilon) \approx N(\mu) \quad \text{for all} \quad \epsilon \in [\mu - \epsilon_c, \mu + \epsilon_c]. \quad (1.33)$$

In 2D this is satisfied for example by the quadratic dispersion $\epsilon_{\boldsymbol{k}} = \hbar^2 \|\boldsymbol{k}\|^2 / (2m)$, for which $N(\epsilon_{\boldsymbol{k}}) = m/(2\pi\hbar^2)$ is in fact constant everywhere. Further assuming

$$|\Delta_\sigma| \ll \epsilon_c \quad (1.34)$$

gives then the zero-temperature result [10, 33]

$$|\Delta_\sigma(T = 0)| = 2\epsilon_c e^{1/[\lambda N(\mu)]} \quad (1.35)$$

and the critical temperature

$$k_B T_c \approx 0.57 |\Delta_\sigma(T = 0)|, \quad (1.36)$$

which is the temperature where Δ_σ vanishes.

Antiadiabatic limit

A closely-related limit is the *antiadiabatic limit* [35, 36], where in addition to the assumptions of the "standard" BCS limit we assume $|\mu| \leq \epsilon_c$ and the DOS to be the most relevant within a $[-\mu, \mu]$ shell from μ:

$$N(\epsilon) \approx N(\mu)\theta(\mu - |\epsilon - \mu|) \quad \text{for all} \quad \epsilon \in [\mu - \epsilon_c, \mu + \epsilon_c]. \quad (1.37)$$

Further assuming

$$|\Delta_\sigma| \ll |\mu| \quad (1.38)$$

yields

$$|\Delta_\sigma(T = 0)| = 2\mu e^{1/[\lambda N(\mu)]}, \quad (1.39)$$

$$k_B T_c \approx 0.57 |\Delta_\sigma(T = 0)|. \quad (1.40)$$

The result is the same as in "standard" BCS but with ϵ_c replaced by μ. But because now $\mu \leq \epsilon_c$, it does not help in increasing T_c. The antiadiabatic limit might, however, be relevant because some studies suggest the TBG superconductivity experiments to be in that regime [36].

Graphene

Even though graphene was experimentally demonstrated already in 2004 [37] and many of its properties can be described by superlatives [38–40], it was not until 2015 when the first experiments showed superconductivity in graphene-based systems [41–44]. These experiments revealed a critical temperature of the order of 5 K after a heavy doping by lithium [41, 44] or calcium [42, 43]. Let us try to understand these experiments from the point of view of BCS theory, which was first presented by Uchoa & Castro Neto [45] and Kopnin & Sonin [34]. The derivation I present here coincides with Kopnin & Sonin.

In graphene there are two bands $b \in \{+, -\}$ and the constant-DOS approximation (1.33) is not valid, so we cannot use the "standard" BCS theory result. Instead we use directly the dispersion which, as long as ϵ_c is much smaller than the energy of the graphene saddle point, can be approximated by the linear dispersion

$$\epsilon_{\boldsymbol{k} \pm \boldsymbol{K}, b} = b \hbar v_{\mathrm{F}} \|\boldsymbol{k}\| \tag{1.41}$$

near the Dirac points $\pm \boldsymbol{K}$, where v_{F} is the graphene Fermi velocity. Shifting the integral in Eqs. (1.25) and (1.26) to the Dirac points we see that our choice of the interaction corresponds to *intervalley pairing* because it couples opposite momenta. Similarly shifting the integrals in Eq. (1.31), going to polar coordinates, and doing a change of variables through $\epsilon := b \hbar v_{\mathrm{F}} l - \mu$ gives

$$1 = -\frac{\lambda}{2\pi(\hbar v_{\mathrm{F}})^2} \sum_b \int_{-\mu}^{b\epsilon_c - \mu} \mathrm{d}\epsilon \, \frac{\epsilon + \mu}{\sqrt{\epsilon^2 + |\Delta_\sigma|^2}} \tanh \frac{\sqrt{\epsilon^2 + |\Delta_\sigma|^2}}{2k_{\mathrm{B}}T}. \tag{1.42}$$

Zero doping ($\mu = 0$) yields [34]

$$|\Delta_\sigma(T = 0, \mu = 0)| = \frac{\pi(\hbar v_{\mathrm{F}})^2}{2} \frac{\lambda_c^2 - \lambda^2}{\lambda \lambda_c^2}, \tag{1.43}$$

$$2k_{\mathrm{B}}T_c(\mu = 0) \ln \cosh \frac{\epsilon_c}{2k_{\mathrm{B}}T_c(\mu = 0)} = -\frac{\pi(\hbar v_{\mathrm{F}})^2}{\lambda}, \tag{1.44}$$

which behave very differently than the "standard" BCS result: superconductivity can only arise when the interaction strength is larger (in absolute value) than the critical strength $\lambda_c := -\pi(\hbar v_{\mathrm{F}})^2/\epsilon_c$. This explains why superconductivity has not been observed in (undoped) graphene: assuming $\epsilon_c = \hbar \omega_{\mathrm{D}} \approx 200 \, \mathrm{meV}$ [46], the interaction strength is apparently smaller than the critical value $\lambda_c \approx -7 \, \mathrm{eV} \, \mathrm{nm}^2$. On the other hand, large doping $\mu \gg |\Delta_\sigma(0)|$ yields [34]

$$|\Delta_\sigma(T = 0)| = 2|\mu| \exp\left[-\frac{\epsilon_c}{|\mu|}\left(\frac{\lambda_c}{\lambda} - 1\right) - 1\right], \tag{1.45}$$

$$k_{\mathrm{B}}T_c \approx 0.57 |\Delta_\sigma(T = 0)|, \tag{1.46}$$

which are exponential in λ as the "standard" BCS result, but relative to the doping level $|\mu|$ instead of ϵ_c. In this sense graphene is reminiscent of the antiadiabatic

limit. In this case there is no critical λ, and thus doping opens the possibility for superconductivity even for arbitrarily small λ, explaining the observed superconductivity in doped graphene.

There are also numerous other theories describing superconductivity in graphene. Uchoa & Castro Neto [45] included, in addition to the (repulsive or attractive) on-site interaction, also nearest-neighbor interactions, but used otherwise a similar mean-field theory. They found that both s and $p + \mathrm{i}p$ pairings are possible in the spin-singlet channel, with the $p + \mathrm{i}p$ state being more favorable if the on-site interaction is repulsive. Black-Schaffer & Doniach [47] used also a mean-field theory but with a resonating valence bond (RVB) interaction, which includes only the nearest-neighbor interaction. They found that both s- and d-wave states are possible, with the d-wave state having a much higher T_c for finite doping. Pathak *et al.* [48] studied a similar RVB interaction with a beyond-mean-field variational Monte Carlo method to obtain a $d + \mathrm{i}d$-wave ground state. González [49] applied Kohn–Luttinger [50, 51] (beyond-mean-field) theory to study graphene doped at the van Hove singularity (vHS), where the density of states diverges, and found a d-wave superconducting state.

Lozovik & Sokolik [52] studied electron–phonon interaction within Eliashberg theory and showed that superconductivity is possible at high doping levels, with the valley structure of Δ depending on the parameters of the electron–phonon interaction. Einenkel & Efetov [53] generalized the calculation of Lozovik & Sokolik by including Coulomb interactions on top of the electron–phonon interaction, and found the intervalley pairing to be the most favorable one and the Coulomb interaction to not be very effective in decreasing the electron–phonon interaction. Kiesel *et al.* [54] used an extended, including on-site, nearest-neighbor, and next-nearest-neighbor, Hubbard interaction together with renormalization group (RG) calculations to show that near the van Hove singularity $d + \mathrm{i}d$-wave superconductivity is favored. According to them, a spin density wave (SDW) order at the vHS and f-wave superconducting order close to the vHS are also possible, depending on the long-rangedness of the interaction. Wang *et al.* [55] used an extended Hubbard interaction up to nearest neighbors together with an RG calculation, and got a similar result that an SDW order at the vHS, surrounded by a $d + \mathrm{i}d$ superconducting order, is favored. Nandkishore *et al.* [56] used RG calculations to show that doping to the van Hove singularities yields d-wave superconductivity, and that superconductivity dominates over all competing orders for weak repulsive interactions. Nandkishore *et al.* [57] argued by Kohn–Luttinger theory that f-wave superconductivity is favored at low doping levels, while close to the vHS d-wave superconductivity emerges according to an RG calculation. Somewhat similarly to Kopnin & Sonin [34], they found out that at low doping levels superconductivity is a threshold phenomenon, emerging whenever the attraction in a certain channel is stronger than the repulsion, while at the vHS any repulsive interaction produces an instability to superconductivity. Ray *et al.* [58] showed that an extended Hubbard interaction, which is local modulo sublattice, leads to $d + \mathrm{i}d$-wave superconductivity.

The above is not a complete list, but it shows how large a set of different phases can be obtained by varying the model parameters and the method. Currently it is not clear which, if any, theory is the correct one. Thus in this thesis I stick

to the simplest one: conventional spin-singlet, s-wave superconductivity within a mean-field approximation, as presented by Kopnin & Sonin [34, 59], but with two generalizations: I allow the order parameter to be position dependent and I include the effect of the pseudo vector potential in the case of PSG or the second layer in the case of TBG.

Flat-band limit

Equations (1.35) and (1.36) are the reason why it was originally thought that superconductivity is limited to low temperatures: T_c is exponential in λ with respect to ϵ_c. Especially for the usual values $\epsilon_c = \hbar\omega_D \lesssim 500\,\text{K}$ and $|\lambda|N(\epsilon_F) \lesssim 1/3$ in metals [12], $T_c \lesssim 25\,\text{K}$. But Khodel' & Shaginyan [5] were the first to notice that if the system hosts *flat bands*, *i.e.* dispersionless energy bands in some region Ω_{FB} of momentum space, T_c becomes *linear in* λ [7, 8].

Let us go through how to arrive to this result from Eq. (1.31). Denoting by Ω the integration region with the cutoff, $\int \mathrm{d}\boldsymbol{l}\,\theta_{\boldsymbol{l}} =: \int_\Omega \mathrm{d}\boldsymbol{l}$, assuming most of the contribution to come from the flat bands,

$$\int_\Omega \mathrm{d}\boldsymbol{l}\,\frac{1}{E_{\boldsymbol{l}}}\tanh\left(\frac{E_{\boldsymbol{l}}}{2k_{\text{B}}T}\right) \approx \int_{\Omega_{\text{FB}}} \mathrm{d}\boldsymbol{l}\,\frac{1}{E_{\boldsymbol{l}}}\tanh\left(\frac{E_{\boldsymbol{l}}}{2k_{\text{B}}T}\right), \tag{1.47}$$

and assuming one flat, spin-independent, band at zero energy,

$$|\epsilon_{\boldsymbol{l}}| \ll |\Delta_\sigma| \quad \text{for all} \quad \boldsymbol{l} \in \Omega_{\text{FB}}, \tag{1.48}$$

gives

$$|\Delta_\sigma(T=0)| = \sqrt{|\Delta_\sigma(T=0,\mu=0)|^2 - \mu^2}, \tag{1.49}$$

$$|\Delta_\sigma(T=0,\mu=0)| = -\frac{\lambda\Omega_{\text{FB}}}{2(2\pi)^2}, \tag{1.50}$$

and

$$k_{\text{B}}T_{\text{c}}(\mu=0) = \frac{1}{2}|\Delta_\sigma(T=0,\mu=0)|. \tag{1.51}$$

Thus T_c at zero doping is linear in λ, giving the possibility to get a high T_c even with a small λ. Equations (1.49)–(1.51) directly tell that doping away from the flat band suppresses T_c monotonically with the critical doping level

$$\mu_{\text{c}}(T=0) = |\Delta_\sigma(T=0,\mu=0)|. \tag{1.52}$$

Khodel' & Shaginyan [5] actually studied systems where the flat band is formed by interactions. But what I mean by a flat-band system in this thesis, and what is also often meant in the literature [7, 8], is the case where the system hosts flat bands already in the noninteracting Hamiltonian [Eq. (1.48)]. These kind of flat bands are hosted for example in the kagomé [60–62] and Lieb crystals [62–66], but also in more "realistic" materials. In this thesis I concentrate on these realistic flat-band systems, namely graphene-based materials. One example is periodically strained graphene (PSG), which was studied by Kauppila *et al.* [67] and others

[68–71], and which we studied in Publication I and I discuss in Chapter 2. Another nowadays famous flat-band system is magic-angle (rotation angle between the graphene layers $\theta \approx 1°$) twisted bilayer graphene (TBG) [relevant studies are discussed in Chapter 3], which in 2018 was found to be superconducting [9], and which we studied in Publications II and III and I discuss in Chapter 3. Yet one example of a flat-band system is proposed by Ramires & Lado [72], in which flat bands are formed in tiny-angle $\theta \ll 1°$ TBG by a homogeneous electric bias. Flat bands have also been predicted [73, 74] and observed [75, 76] at the interfaces of rhombohedral graphite.

1.3   Berezinskii–Kosterlitz–Thouless transition and superfluid weight

As the Mermin–Wagner theorem [77] forbids true long-range order in two-dimensional systems, also the phase of the superconducting order parameter Δ inevitably has fluctuations, growing in space, that destroy superconductivity already below the critical temperature T_{c} defined by $\Delta(T_{\mathrm{c}}) = 0$. There can be, however, quasi-long-range order below the *Berezinskii–Kosterlitz–Thouless* (BKT) temperature $T_{\mathrm{BKT}} \leq T_{\mathrm{c}}$, below which vortices of the $\arg(\Delta)$ field bind together [78]. The pioneering work on the subject was done by Berezinskii [79, 80], Kosterlitz & Thouless [81–83], Nelson & Kosterlitz [84], and José *et al.* [85]. Good reviews on the subject can be found in Refs. [86, 87].

    We can get a basic understanding of the physics underlying the BKT transition by studying the classical xy model with N classical spins in a square lattice $L := \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$ with the primitive vectors $\boldsymbol{a}_1 := (a, 0)$ and $\boldsymbol{a}_2 := (0, a)$, where the nearest-neighbor spins $\boldsymbol{S}(\boldsymbol{r})$ and $\boldsymbol{S}(\boldsymbol{r} + \boldsymbol{a}_j)$ are coupled through $-J\boldsymbol{S}(\boldsymbol{r}) \cdot \boldsymbol{S}(\boldsymbol{r} + \boldsymbol{a}_j) = -J\cos(\theta(\boldsymbol{r}) - \theta(\boldsymbol{r} + \boldsymbol{a}_j))$. Here the angle θ of spin $\boldsymbol{S}$ is measured *e.g.* from the $\boldsymbol{a}_1$ vector. Then the Hamiltonian reads

$$H = -J \sum_{\boldsymbol{r} \in L/L_{\mathrm{BK}}} \sum_{j=1}^{2} \cos(\theta(\boldsymbol{r}) - \theta(\boldsymbol{r} + \boldsymbol{a}_j)), \tag{1.53}$$

where we take the system to be periodic in translations of the large Born–von Kármán lattice L_{BK}. Assuming θ to vary slowly in the lattice scale, we may expand the cosine to the second order. Further taking the continuum limit $a \to 0$ ($L \to \mathbb{R}^2$) we may approximate the Hamiltonian as

$$H = \frac{J}{2} \int_{\mathbb{R}^2/L_{\mathrm{BK}}} \mathrm{d}\boldsymbol{r} \, \|\nabla\theta(\boldsymbol{r})\|^2 + \mathrm{const}, \tag{1.54}$$

which constitutes the *continuum xy model*.

    To calculate the energy of creating a single vortex with the winding number n, we first note that for any closed path γ encircling the vortex core we have

$$\int_{\gamma} \mathrm{d}\boldsymbol{r} \cdot \nabla\theta_{\mathrm{vor}}(\boldsymbol{r}) = 2\pi n, \tag{1.55}$$

where θ_{vor} is the angular field of the vortex. Assuming the vortex to be spherically symmetric and γ to be a circle of radius r around the vortex core, we get on the other hand

$$\int_\gamma \mathrm{d}\boldsymbol{r} \cdot \nabla\theta_{\text{vor}}(\boldsymbol{r}) = 2\pi r \|\nabla\theta_{\text{vor}}(r)\|. \tag{1.56}$$

Combining Eqs. (1.55) and (1.56) we get

$$\|\nabla\theta_{\text{vor}}(r)\| = \frac{n}{r}, \tag{1.57}$$

which can then be used to calculate the vortex energy through Eq. (1.54) as

$$E_{\text{vor}} = \frac{J}{2} \int \mathrm{d}\boldsymbol{r} \, \|\nabla\theta_{\text{vor}}(r)\|^2 + \text{const} = J\pi n^2 \ln\left(\frac{l}{a}\right) + \text{const}, \tag{1.58}$$

where l describes the system size and where we regularized the integral by adding a lower cutoff a coming from the original lattice. Since the energy associated with higher-order vortices grows quadratically, we concentrate only on $n = 1$ vortices.

The entropy of the vortex is

$$S_{\text{vor}} = k_{\text{B}} \ln \Omega_{\text{vor}} = 2k_{\text{B}} \ln\left(\frac{l}{a}\right), \tag{1.59}$$

where $\Omega_{\text{vor}} = (l/a)^2$ is the number of possibilities to place the vortex, that is, the number of lattice sites. The most important thing is that the scaling in l is the same as for the energy, which is unique to two dimensions. The free energy of the vortex at temperature T is then

$$F_{\text{vor}} = E_{\text{vor}} - TS_{\text{vor}} = (J\pi - 2k_{\text{B}}T) \ln\left(\frac{l}{a}\right) + \text{const.} \tag{1.60}$$

We immediately see that there is critical behavior when $2k_{\text{B}}T = J\pi$: when we take $l \to \infty$, then $F \to \infty$ if $2k_{\text{B}}T < J\pi$ and $F \to -\infty$ if $2k_{\text{B}}T > J\pi$. We call this critical temperature the *Berezinskii–Kosterlitz–Thouless transition temperature*,

$$k_{\text{B}}T_{\text{BKT}} = \frac{\pi}{2}J. \tag{1.61}$$

Above this temperature the system can lower its free energy by producing vortices.

If we generalize the above calculation to an arbitrary lattice and allow the coupling to be anisotropic, *i.e.* we replace $J \mapsto J_j$ in the beginning, we get

$$k_{\text{B}}T_{\text{BKT}} = \frac{\pi}{2}\sqrt{J_1 J_2} = \frac{\pi}{2}\sqrt{\det J}. \tag{1.62}$$

Here we defined the matrix J, which in the lattice basis $\Lambda := (\boldsymbol{a}_1, \boldsymbol{a}_2)$ and natural basis $N := (\boldsymbol{e}_x, \boldsymbol{e}_y)$ can be respectively written as

$$J_\Lambda = \begin{pmatrix} J_1 & 0 \\ 0 & J_2 \end{pmatrix}, \quad J_N = \begin{pmatrix} J_{xx} & J_{xy} \\ J_{yx} & J_{yy} \end{pmatrix}, \tag{1.63}$$

where the elements $J_{\mu\nu}$ are defined through the change of basis.

It is remarkable that the above macroscopic and classical calculation is (almost) enough to capture the essential physics of the phase transition. Indeed, microscopic, quantum-mechanical calculations in superfluids [84] or superconductors [86, 88–92] yield a similar result

$$k_{\mathrm{B}}T_{\mathrm{BKT}} = \frac{\pi}{2}J(T_{\mathrm{BKT}}) \tag{1.64}$$

when J is identified with the *superfluid stiffness* $J = \hbar^2\rho^{\mathrm{s}}/(4m)$, where ρ_{s} is the superfluid mass density and m is the effective mass. The only difference to Eq. (1.61) is that J becomes temperature dependent and thus Eq. (1.64) becomes a self-consistency equation. While in the xy model the vortices are in the θ field describing the spin angles, in superconductors it is the phase θ of the superconducting order parameter $\Delta = |\Delta|\mathrm{e}^{\mathrm{i}\theta}$ that has the vortices and thus determines the transition. Since sometimes the effective mass m is not a well-defined concept [92], it is often better to use the *superfluid weight* instead, which in the case of a properly defined mass reads[6] $D^{\mathrm{s}} = \hbar^2\rho^{\mathrm{s}}/m = 4J$. Thus, when taking also the possible anisotropy into account, we get

$$k_{\mathrm{B}}T_{\mathrm{BKT}} = \frac{\pi}{8}\sqrt{\det D^{\mathrm{s}}(T_{\mathrm{BKT}})}, \tag{1.65}$$

where D^{s} is a matrix as in Eq. (1.63).

1.3.1   Superfluid weight and the London equations

Above we saw that the superfluid weight is important in determining the BKT transition temperature, but it is also important in determining other properties of superconductors, as I demonstrate below. In linear response theory, a vector potential $\boldsymbol{\mathcal{A}}$ produces a supercurrent density [11, 93]

$$\boldsymbol{j}(\boldsymbol{r},t) = -\frac{1}{VT}\int \mathrm{d}\boldsymbol{r} \int \mathrm{d}t\, K(\boldsymbol{r},t,\boldsymbol{r}',t')\boldsymbol{\mathcal{A}}(\boldsymbol{r}',t'), \tag{1.66}$$

where V is the area of the sample and T is the length of the time integral. The form of the kernel K is derived in Sec. 2.3.2 in the case of PSG in the static limit. Assuming a local kernel $K(\boldsymbol{r},t,\boldsymbol{r}',t') = \delta(\boldsymbol{r}-\boldsymbol{r}')\delta(t-t')VTK(\boldsymbol{r},t)$, the current becomes [93]

$$\boldsymbol{j}(\boldsymbol{r},t) = -K(\boldsymbol{r},t)\boldsymbol{\mathcal{A}}(\boldsymbol{r},t). \tag{1.67}$$

If K furthermore happens to be a constant, $K = \langle K \rangle$, it reads

$$\boldsymbol{j}(\boldsymbol{r},t) = -D^{\mathrm{s}}\boldsymbol{\mathcal{A}}(\boldsymbol{r},t) \quad \text{with} \quad D^{\mathrm{s}} := \langle K \rangle \tag{1.68}$$

being the superfluid weight. Here the angle bracket denotes the average over position and time.

---

[6]In this thesis and in Publications I and III we use a definition of the superfluid weight where the one in Ref. [92] is multiplied by $\hbar^2$, in order to have D^{s} in the convenient units of energy.

Taking the time derivative of Eq. (1.68) and using the Maxwell equation $\partial_t \mathcal{A} = -\boldsymbol{E}$, in the gauge where the scalar potential vanishes, yields the first London equation [10, 11, 33]

$$\partial_t \boldsymbol{j} = D^s \boldsymbol{E}. \tag{1.69}$$

Equation (1.69) can be used to experimentally determine the superfluid weight D^s by applying an external electric field $\boldsymbol{E}$ and measuring the resulting change $\partial_t \boldsymbol{j}$ in the current.

On the other hand, taking the curl of Eq. (1.68) and using the Maxwell equations $\boldsymbol{j} = \frac{1}{\mu_0} \nabla \times \boldsymbol{B}$, $\boldsymbol{B} = \nabla \times \mathcal{A}$, and $\nabla \cdot \boldsymbol{B} = 0$ yields the second London equation [10, 11, 33]

$$\nabla^2 \boldsymbol{B} = \mu_0 D^s \boldsymbol{B}, \tag{1.70}$$

where μ_0 is the vacuum permeability. In 3D systems this has only exponentially decaying (physical) solutions with the characteristic decay length, the *London penetration depth*, $\lambda := 1/\sqrt{\mu_0 D^s}$. This means that an external magnetic field $\boldsymbol{B}$ is expelled from inside the material whenever D^s is finite. This is exactly the Meissner effect, the defining property of superconductivity, so the superfluid weight D^s can actually be interpreted as the defining property of superconductivity. In thin films and 2D systems, however, a more relevant length scale is the *Pearl length* $\lambda_\perp := 2\lambda^2/l$ [10, 94], which describes how quickly the magnetic field of a (Pearl) vortex dies out in a system of thickness l.

1.4   (Discrete) Fourier series of periodic functions

Because in this thesis we are only interested in periodic systems, (discrete) Fourier series[7] plays a major role in the theory. The discrete Fourier series is needed when deriving the theory of PSG/TBG in the tight-binding level, while the Fourier series is needed when one wants to Fourier transform at the low-energy continuum level. On the other hand the summation and integration sets of the (discrete) Fourier series are written elegantly using the notation of quotient groups [95], but the notation is not generally used in the physics community, so I cover this notation first.

1.4.1   Quotient groups

Let G be a group with the group operation '+' and let H be a normal subgroup of G. Then we define the *equivalence class* of the *representative* $g \in G$ (with respect to H) as

$$[g] := g + H := \{g + h : h \in H\} \subset G. \tag{1.71}$$

This means that whenever we add an element of H, nothing changes:

$$[g + h] = g + h + H = g + H = [g] \quad \text{for all } g \in G, h \in H \tag{1.72}$$

---

[7]I use the term "series" for functions that are periodic, and the term "discrete" for functions defined on a lattice.
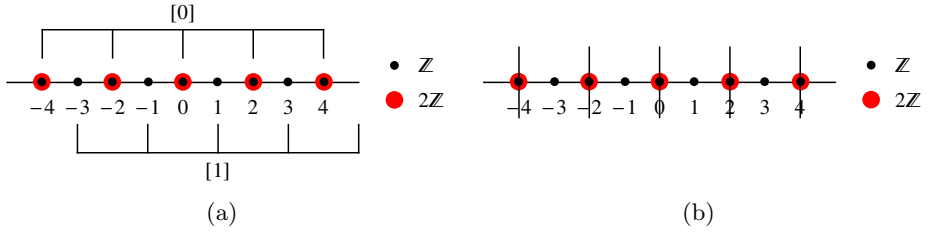
Figure 1.2: Illustration of the quotient group G/H when $G = \mathbb{Z}$ is a one-dimensional lattice and $H = 2\mathbb{Z}$ is its one-dimensional sublattice. (a) The equivalence classes $[0] = [2] = [-2] = [4] = [-4] = \ldots$ and $[1] = [-1] = [3] = [-3] = \ldots$, which compose the quotient group $\mathbb{Z}/2\mathbb{Z} = \{[0], [1]\} = \mathbb{Z}_2$. (b) One possible division to unit cells, each of which is isomorphic to $\mathbb{Z}/2\mathbb{Z}$.

because $h + H = H$. This equivalence defines a partition of G into a disjoint union, if we define the *quotient group*

$$G/H \coloneqq \{[g] : g \in G\}. \tag{1.73}$$

Especially if $G = L$ is a lattice and $H = SL \subset L$ its sublattice (or the equivalent term superlattice), $G/H = L/SL$ *identifies* all the lattice points whose difference is in SL. In other words the property (1.72) means that $[\boldsymbol{r} + \boldsymbol{t}] = [\boldsymbol{r}]$ for all $\boldsymbol{r} \in L$, $\boldsymbol{t} \in SL$, and thus it naturally describes periodicity in translations of SL. In this case L/SL is isomorphic to any unit cell of "L modulo SL", and thus L/SL can be also interpreted as that unit cell. This means that by dropping the brackets from $[\boldsymbol{r}]$, $\boldsymbol{r}$ can be interpreted either as the equivalence class, or as the representative in the unit cell. From the context it should be clear which interpretations are allowed.

As an example, let $G = \mathbb{Z}$ be the integer numbers with the group operation of addition and $H = 2\mathbb{Z}$ the even integer numbers. Then G is a one-dimensional lattice and H its one-dimensional sublattice. Then by using the definition (1.71) we see that $[0] = 0 + 2\mathbb{Z} = 2\mathbb{Z}$ is the set of even numbers, $[1] = 1 + 2\mathbb{Z}$ is the set of odd numbers, $[2] = 2 + 2\mathbb{Z} = 2\mathbb{Z} = [0]$ is again the set of even numbers, $[3] = 3 + 2\mathbb{Z} = 1 + 2\mathbb{Z} = [1]$ is again the set of odd numbers, and so on. In other words, if one adds an element of $H = 2\mathbb{Z}$, *i.e.* any multiple of 2, one stays within the same equivalence class. Thus there are exactly two equivalence classes, $[0] = [2] = [-2] = [4] = [-4] = \ldots$ and $[1] = [-1] = [3] = [-3] = \ldots$, as illustrated in Fig. 1.2a. According to the definition (1.73) the quotient group G/H is $Z/2\mathbb{Z} = \{[k] : k \in \mathbb{Z}\} = \{[0], [1]\}$, which is often denoted by $\mathbb{Z}_2$. This is clearly isomorphic to $\{0, 1\}$, which is one of the possible unit cells, as shown in Fig. 1.2b.

As a second example, take the two-dimensional lattice $G = L = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$ with the group operation of vector addition and the two-dimensional sublattice $H = SL = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{t}_1, \boldsymbol{t}_2\}$, as shown in Fig. 1.3a. In this example $\boldsymbol{t}_1 = 3\boldsymbol{a}_1 + \boldsymbol{a}_2 =: (3, 1)_L$ and $\boldsymbol{t}_2 = -\boldsymbol{a}_1 + \boldsymbol{a}_2 = (-1, 1)_L$. Then the equivalence classes are $[(0, 0)_L] = (0, 0)_L + SL = SL$, $[(0, 1)_L] = (0, 1)_L + SL$, $[(1, 1)_L] = (1, 1)_L + SL$, and $[(2, 1)_L] = (2, 1)_L + SL$. After this we start seeing periodicity: $[(1, 0)_L] = (1, 0)_L + SL = (1, 0)_L + \boldsymbol{t}_2 - \boldsymbol{t}_2 + SL = (0, 1)_L + SL = [(0, 1)_L]$
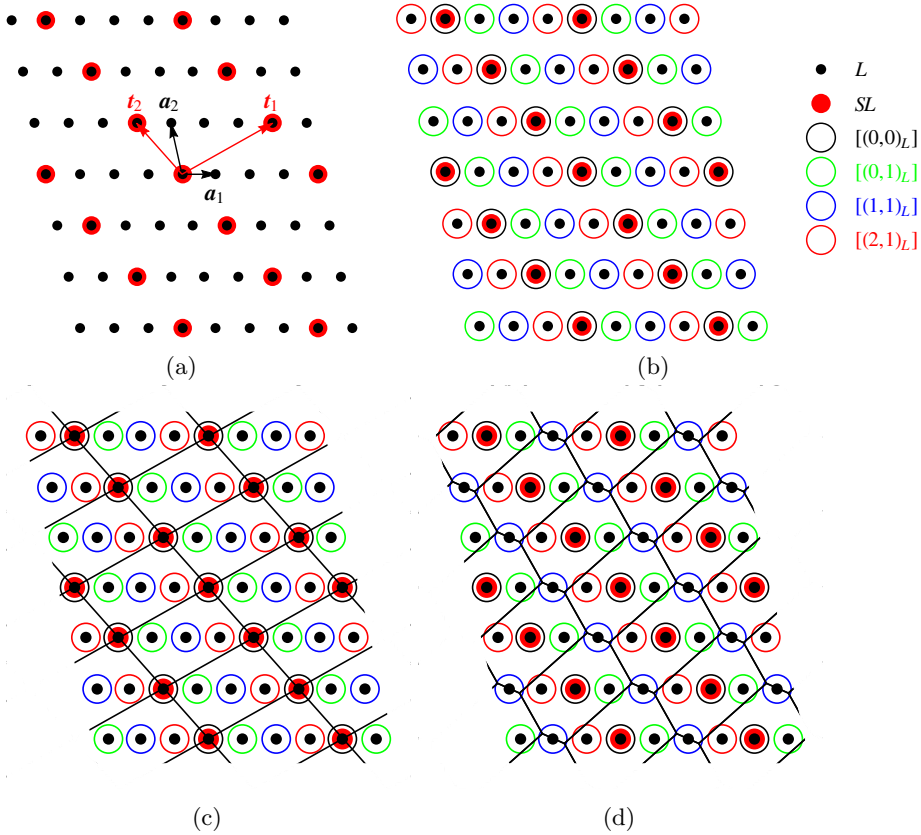
Figure 1.3: Illustration of the quotient group L/SL when $L = \operatorname{span}_{\mathbb{Z}}\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$ is a two-dimensional lattice and $SL = \operatorname{span}_{\mathbb{Z}}\{\boldsymbol{t}_2, \boldsymbol{t}_2\}$ its two-dimensional sublattice. (a) The lattices L and SL. (b) The equivalence classes that constitute the quotient group $L/SL = \{[(0,0)_L], [(0,1)_L], [(1,1)_L], [(2,1)_L]\}$. (c) One possible division to (conventional) unit cells, each of which is isomorphic to L/SL. (d) Another possible division to (Wigner–Seitz or Voronoi) unit cells, each of which is isomorphic to L/SL, and thus also to the unit cells of (c). (c–d) Each unit cell includes exactly one representative from each equivalence class. Note that an arbitrary translation of the "division" (the black lines) would not change anything: each unit cell would still include exactly one representative from each class.

because $SL - \boldsymbol{t}_2 = SL$, $[(0,-1)_L] = (0,-1)_L + SL = (2,1)_L + SL = [(2,1)_L]$ because $SL - \boldsymbol{t}_1 - \boldsymbol{t}_2 = SL$, and so on, as illustrated in Fig. 1.3b. The quotient group is then $G/H = L/SL = \{[(0,0)_L], [(0,1)_L], [(1,1)_L], [(2,1)_L]\}$ and it can be interpreted (through an isomorphism) for example as the conventional[8] unit cell $\{(0,0)_L, (0,1)_L, (1,1)_L, (2,1)_L\}$ (Fig. 1.3c) or as the Wigner–Seitz unit cell $\{(0,0)_L, (1,0)_L, (-1,-1)_L, (-1,0)_L\}$ (Fig. 1.3d).

1.4.2   Discrete Fourier series

Let $f : L/SL \to \mathbb{C}^n$ be a function *defined on a two-dimensional lattice* $L :=$ $\mathrm{span}_{\mathbb{Z}}\{\boldsymbol{a}_1, \boldsymbol{a}_2\} \subset \mathbb{R}^2$ and periodic in translations of the two-dimensional superlattice $SL := \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{t}_1, \boldsymbol{t}_2\} \subset L$. It can be shown that the two-dimensional discrete Fourier series and its coefficients can be written respectively as [95]

$$f(\boldsymbol{r}) = \sum_{\boldsymbol{G} \in SL^*/L^*} \mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} \tilde{f}(\boldsymbol{G}), \quad \tilde{f}(\boldsymbol{G}) = \frac{1}{|L/SL|} \sum_{\boldsymbol{r} \in L/SL} \mathrm{e}^{-\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} f(\boldsymbol{r}), \qquad (1.74)$$

where $L^* := \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{g}_1, \boldsymbol{g}_2\}$ is the reciprocal lattice of L defined through $\boldsymbol{a}_i \cdot \boldsymbol{g}_j = 2\pi\delta_{ij}$ and $SL^* := \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{G}_1, \boldsymbol{G}_2\}$ is the reciprocal lattice of SL defined through $\boldsymbol{t}_i \cdot \boldsymbol{G}_j = 2\pi\delta_{ij}$. Here L/SL can be interpreted as any of the superlattice unit cells, SL^*/L^* is the superlattice Brillouin zone[9] that can be interpreted as any of the unit cells of the reciprocal superlattice, and $|L/SL| = |SL^*/L^*|$ is the number of elements in the quotient group L/SL or SL^*/L^*, *i.e.* the number of lattice points in the superlattice unit cell or in the reciprocal superlattice unit cell.

1.4.3   Fourier series

Let $f : \mathbb{R}^2/SL \to \mathbb{C}^n$ be a function *defined on the continuum* $\mathbb{R}^2$ and periodic in translations of the two-dimensional lattice $SL := \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{t}_1, \boldsymbol{t}_2\} \subset \mathbb{R}^2$ with the reciprocal lattice $SL^* = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{G}_1, \boldsymbol{G}_2\}$. Assuming the one-dimensional Fourier series is known, one can write a Fourier series separately for the x_1 and x_2 components of $f(x_1, x_2)$. Then the two-dimensional Fourier series and its coefficients respectively become

$$f(\boldsymbol{r}) = \sum_{\boldsymbol{G} \in SL^*} \mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} \tilde{f}(\boldsymbol{G}), \quad \tilde{f}(\boldsymbol{G}) = \frac{1}{|\mathbb{R}^2/SL|} \int_{\mathbb{R}^2/SL} \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} f(\boldsymbol{r}), \qquad (1.75)$$

where the integral is calculated as a two-dimensional volume integral which, by interpreting $\mathbb{R}^2/SL$ as the parallelogram defined by $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$ (one of the superlattice continuum unit cells) and by change of variables, can be more concretely

---

[8]By the conventional unit cell I mean the one spanned by the primitive vectors $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$. On the other hand, by the Wigner–Seitz unit cell I mean the one that is constructed through a Voronoi decomposition of SL.

[9]Usually the superlattice Brillouin zone is defined exclusively as the Wigner–Seitz unit cell of the reciprocal superlattice [33]. But when viewed from the quotient group perspective, no unit cell is more special than any other, because they all are isomorphic to the quotient group SL^*/L^*. Thus I call both the quotient group SL^*/L^* and all the possible unit cells simply the superlattice Brillouin zone (SBZ).

written as

$$\tilde{f}(\boldsymbol{G}) = \tilde{f}(m_1\boldsymbol{G}_1 + m_2\boldsymbol{G}_2) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{d}x_1 \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{d}x_2\, \mathrm{e}^{-\mathrm{i}2\pi(m_1 x_1 + m_2 x_2)} f(x_1\boldsymbol{t}_1 + x_2\boldsymbol{t}_2).$$
(1.76)

Here $m_1, m_2 \in \mathbb{Z}$, $\left|\mathbb{R}^2/SL\right| := |\boldsymbol{G}_1 \times \boldsymbol{G}_2|$ is the area of the superlattice unit cell and '$\times$' is the two-dimensional "cross product". Writing the Fourier series this way I call the *reduced zone scheme* (RZ), for reasons explained in Sec. 2.1.1.

As a special case, if f is constant in the $\boldsymbol{t}_2$ direction, we may use the result

$$\tilde{f}(\boldsymbol{G}) = \tilde{f}(m_1\boldsymbol{G}_1 + m_2\boldsymbol{G}_2) = \delta_{m_2,0}\tilde{f}(m_1\boldsymbol{G}_1) = \delta_{m_2,0}\tilde{f}(\boldsymbol{G}) \qquad (1.77)$$

yielding the series

$$f(\boldsymbol{r}) = \sum_{\boldsymbol{G} \in SL_1^*} \mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} \tilde{f}(\boldsymbol{G}), \qquad (1.78)$$

where $SL_1^* := \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{G}_1\}$ is a one-dimensional sublattice of SL^*. If we calculate the Fourier series this way (possible only if f is constant in the $\boldsymbol{t}_2$ direction), as a sum over the one-dimensional lattice SL_1^*, we call this the *mixed zone scheme* (MZ), for reasons explained in Sec. 2.1.1.

Summarizing both the schemes together, we may write the Fourier series and its coefficients as

$$f(\boldsymbol{r}) = \sum_{\boldsymbol{G} \in SL_S^*} \mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} \tilde{f}(\boldsymbol{G}), \quad \tilde{f}(\boldsymbol{G}) = \frac{1}{|\mathbb{R}^2/SL|} \int_{\mathbb{R}^2/SL} \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} f(\boldsymbol{r}), \qquad (1.79)$$

where in the reduced zone scheme $SL_{\mathrm{RZ}}^* := SL^* = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{G}_1, \boldsymbol{G}_2\}$ and in the mixed zone scheme $SL_{\mathrm{MZ}}^* := SL_1^* = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{G}_1\}$. Although the difference between the two schemes seems minor, the difference is significant both conceptually and in numerical calculations: the meaning of a band and momentum is different when doing Bloch theory, and the dimension of the "$\boldsymbol{G}$ space" is either one or two, having a huge impact on the computational weight. While in the case of TBG (Publications II and III) the only possibility is to use the reduced zone scheme, in the case of PSG (Publications I) we employ both of them. Thus the difference is discussed more in Chapter 2.

Chapter 2

Periodically strained graphene

In this chapter I present an introduction to Publication I regarding periodically strained graphene (PSG) and summarize its main results. It has been predicted [67–70] and experimentally shown [71] that flat bands can be induced in graphene through *periodic strain*. Especially Kauppila *et al.* [67] showed that this strongly promotes superconductivity, at least when concentrating on the spin-singlet channel with *s*-wave symmetry. The motivation of Publication I is to generalize the theory and results of Kauppila *et al.*, who studied only a specific one-dimensional harmonic strain, to strains of arbitrary shape and arbitrary (1D or 2D) dimension. As a side effect of the generalization we also see that both the theory and results are extremely similar to that of twisted bilayer graphene (TBG), studied in Publications II and III and discussed in Chapter 3.

A natural question arises: is it experimentally possible to manufacture periodic strain in graphene? In fact, this has already been demonstrated in the work of Jiang *et al.* [71], where periodic secondary buckling of graphene was observed near large buckles resulting from the residual stress due to the substrate, and band flattening was observed. Many kinds of strain lattices were successfully manufactured, both 1D and 2D, as shown in Fig. 2.1. With this method the period was tunable between 8 and 25 nm and the atomic out-of-plane displacement amplitude was of the order of 1 Å.

Perhaps even better control could be achieved by optical forging [96], where arbitrary out-of-plane strain patterns are drawn by a laser, as shown in the simulation figure 2.2a. There the displacement amplitude can be even above 50 nm, but the period is often limited by the laser wavelength to some 100 nm. However, even structures below the diffraction limit can be drawn by irradiating for shorter times [97]. One disadvantage of this kind of *direct* drawing is that the buckling is probably stabilized by Stone–Wales defects [96], which makes theoretical modeling more difficult. But, as seen in Fig. 2.2a, much smaller secondary ripples are formed near the main pattern, so this kind of *indirect* drawing could be utilized. Moreover they are probably stabilized by the larger ripples and not by defects, which is also
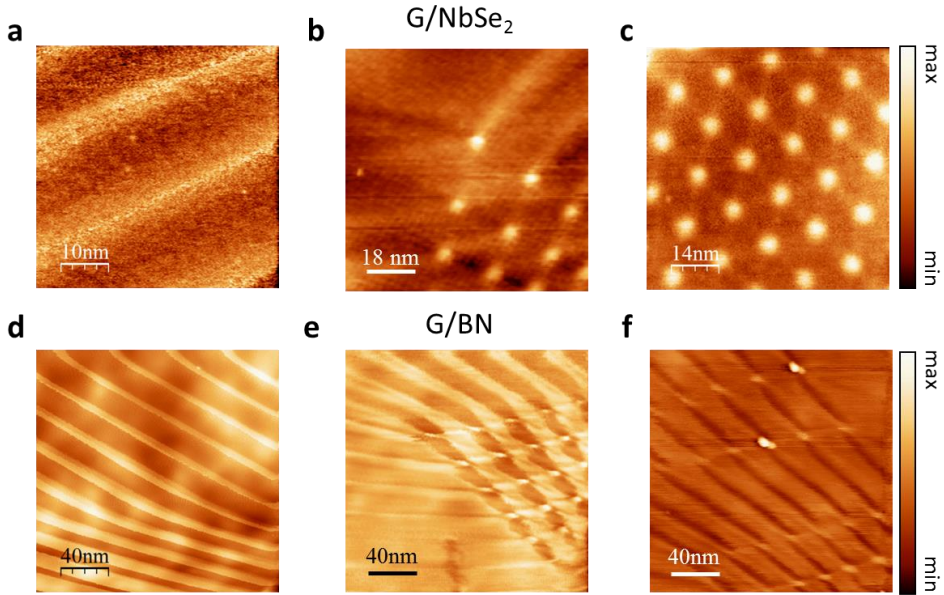
Figure 2.1: A variety of periodic buckling patterns manufactured on graphene on different substrates by Jiang *et al.* [71], as measured by STM. (a–c) Graphene on niobium diselenide. (d–f) Graphene on hexagonal boron nitride. The periodic (secondary) ripples are formed near large primary buckles resulting from the stress that remains after placing graphene on the substrate. Reprinted with permission from Ref. [71].

an advantage. The amplitude and period of these secondary ripples is still not understood, but their forming mechanism is very similar to that of Jiang *et al.* [71], so it is plausible to expect that similar band flattening could be achieved also by this method.

One further experimentally demonstrated method is to use an AFM tip to evaporate adsorbed hydrogen from a germanium surface to produce a pressurized H_2 gas under specific locations of graphene [98], as illustrated in Figs. 2.2b and 2.2c. Changing the tip voltage allowed tuning the out-of-plane displacement amplitude from 0 to roughly 10 nm, with the minimum demonstrated period of the order of 100 nm. Another option is to pattern the substrate below instead of graphene itself [99, 100], as shown experimentally by Jiang *et al.* [100] in Fig. 2.2d. They manufactured a 1 µm-periodic array of pillars, on top of which the graphene layer is buckled with an amplitude of 20 nm. On the other hand compressing graphene on the edges when it is supported on a substrate [99, 101] or fully encapsulated [102, 103] also yields (periodic) buckling, as shown in Fig. 2.2e. In the case of a substrate Aitken & Huang [99] predicted a displacement amplitude of 0.1 nm with the period 2 nm. In the encapsulated system Androulidakis *et al.* [102] predicted an amplitude of 0.6 nm with the period 1 . . . 2 nm, whereas Koukaras *et al.* [103] predicted an amplitude of the order of 0.1 nm with the period 13 . . . 24 nm,

depending on the strain strength and the distance between the encapsulating layers. In the same manner the hypothetical graphene cardboard material [104] could be manufactured. Also an ultracold atom gas in a tunable optical honeycomb lattice [105] could be used to study arbitrary strain patterns.

Periodic strain is also produced if graphene is placed on a slightly different crystal, such as hexagonal boron nitride (hBN) [106] or graphene that is twisted (*i.e.* TBG). In TBG such periodic strain has been both predicted [107–110] and observed [111], especially close to the magic angle. Constraining the relaxation within the plane, Nam & Koshino [109] predicted an in-plane displacement amplitude of 0.1 Å, as shown in Fig. 2.2f. Letting the atoms relax freely, van Wijk *et al.* [107] predicted an out-of-plane displacement amplitude of 0.1 Å, whereas that amplitude was predicted by Dai *et al.* [108] to be either 0.1 Å or 2 Å depending on the initial conditions. The period in all these studies follows the moiré period, which is 13 nm close to the magic angle.
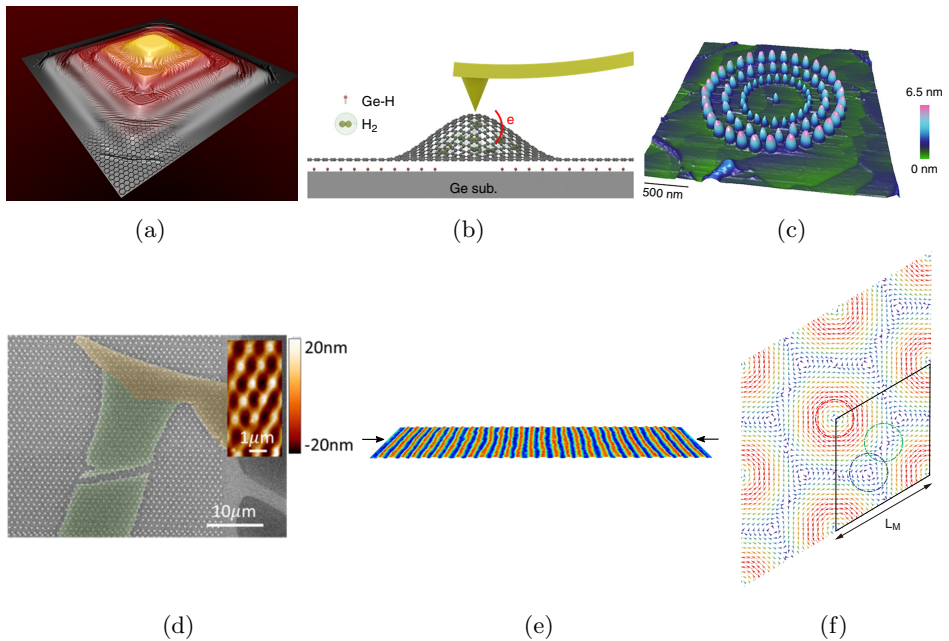
(a) (b) (c)

(d) (e) (f)

Figure 2.2: Possible ways to fabricate periodic strain in graphene. (a) Optical forging by laser irradiation [96]. Graphene bulges up at locations where a pulsed femtosecond laser is shined at, with the amplitude proportional to the irradiation time, which allows drawing arbitrary strain patterns. The yellow color of the highest point corresponds to a height of 60 nm. The figure is the result of a simulation, and close the the primary buckling pattern it shows also smaller secondary ripples, similarly as in the experiments of Jiang *et al.* [71] in Fig. 2.1. The shown graphene crystal is not to scale. Image courtesy of Pekka Koskinen. (b,c) Hydrogen pressurization by an AFM tip [98]. Hydrogen adsorbed on a germanium surface is vaporized by an AFM tip, creating pressure under specific locations of graphene, which buckles graphene. Images reprinted from Ref. [98] under the Creative Commons License CC BY 4.0, https://creativecommons.org/licenses/by/4.0/. (d) Graphene on a corrugated surface [100]. Graphene (green area) is sagged when placed on top of small pillars (white dots), made from either a lift-off resist or gold. The inset shows the measured height profile. Reprinted with permission from Ref. [100]. Copyright 2017 American Chemical Society. (e) Compressing (black arrows) graphene encapsulated within PMMA layers (not shown) [103] creates 20 nm-periodic buckling with the amplitude of roughly 2 Å. The color corresponds to height. Reprinted with permission from Ref. [103], modified by adding the black arrows. Copyright 2016 Elsevier. (f) Strain in relaxed twisted bilayer graphene [109]. The figure shows the in-plane displacement field of one of the layers when rotated at the magic angle, exhibiting an amplitude of 0.1 Å. The period is the moiré period L_M, which in the case of the magic rotation angle is 13 nm. Reprinted with permission from Ref. [109]. Copyright 2017 American Physical Society.

2.1   Continuum model

To model strained graphene we employ a widely-used continuum model. The derivation is done in detail in the Supplementary Material of Publication I and also to some extent in Ref. [112], but I summarize it here because the TBG theory presented in Chapter 3 is an extension of this. Taking the primitive vectors

$$\boldsymbol{a}_1 := \frac{a}{2}(1, \sqrt{3}), \quad \boldsymbol{a}_2 := \frac{a}{2}(-1, \sqrt{3}) \tag{2.1}$$

of the graphene lattice $L := \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$, where $a := \|\boldsymbol{a}_1\| = \|\boldsymbol{a}_2\|$ is the lattice constant, and the nearest-neighbor vectors

$$\boldsymbol{\delta}_1 := \frac{1}{3}(\boldsymbol{a}_1 + \boldsymbol{a}_2), \quad \boldsymbol{\delta}_2 := \frac{1}{3}(\boldsymbol{a}_2 - 2\boldsymbol{a}_1), \quad \boldsymbol{\delta}_3 := \frac{1}{3}(\boldsymbol{a}_1 - 2\boldsymbol{a}_2), \tag{2.2}$$

the Dirac points (where the band touchings happen) are located at

$$\boldsymbol{K} := \frac{4\pi}{3a}(1, 0) \quad \text{and} \quad \boldsymbol{K}' := -\boldsymbol{K}. \tag{2.3}$$

In the nearest-neighbor tight-binding model the Hamiltonian [at the level of Eq. (1.9)] of pristine graphene is then

$$H_{\mathrm{p}} := -t \sum_{\sigma} \sum_{j=1}^{3} \sum_{\boldsymbol{r} \in L/L_{\mathrm{BK}}} \psi_{\sigma}^{\dagger}(\boldsymbol{r})\psi_{\sigma}(\boldsymbol{r} + \boldsymbol{\delta}_j) + \text{h.c.} - \mu N, \tag{2.4}$$

where t is the nearest-neighbor hopping energy, $\sigma \in \{\uparrow, \downarrow\}$ is spin, $\psi_{\sigma}(\boldsymbol{r})$ is the electron annihilation operator at crystal site $\boldsymbol{r}$ (i.e. any sublattice) taken to be periodic in translations of the large Born–von Kármán lattice L_{BK}, μ is the chemical potential, $N := \sum_{\sigma\alpha\boldsymbol{r}} \psi_{\sigma}^{\dagger}(\boldsymbol{r} + \boldsymbol{\delta}_{\alpha})\psi_{\sigma}(\boldsymbol{r} + \boldsymbol{\delta}_{\alpha})$ is the number operator, $\alpha \in \{A, B\}$ denotes sublattice, and $\boldsymbol{\delta}_A := 0$, $\boldsymbol{\delta}_B := \boldsymbol{\delta}_1$ is the sublattice shift.

Defining the sublattice-shifted annihilation operators[1] $\psi_{\sigma,A}, \psi_{\sigma,B} : L/L_{\mathrm{BK}} \to \mathrm{Op}$, $\psi_{\sigma,A}(\boldsymbol{r}) := \psi_{\sigma}(\boldsymbol{r})$, $\psi_{\sigma,B}(\boldsymbol{r}) := \psi_{\sigma}(\boldsymbol{r} + \boldsymbol{\delta}_1)$ for all $\boldsymbol{r} \in L/L_{\mathrm{BK}}$ allows writing H_{p} as

$$H_{\mathrm{p}} = -t \sum_{\sigma} \sum_{j=1}^{3} \sum_{\boldsymbol{r} \in L/L_{\mathrm{BK}}} \psi_{\sigma,A}^{\dagger}(\boldsymbol{r})\psi_{\sigma,B}(\boldsymbol{r} + \boldsymbol{\delta}_j - \boldsymbol{\delta}_1) + \text{h.c.} - \mu N \tag{2.5}$$

with the number operator becoming $N = \sum_{\sigma\alpha\boldsymbol{r}} \psi_{\sigma,\alpha}^{\dagger}(\boldsymbol{r})\psi_{\sigma,\alpha}(\boldsymbol{r})$. Strain modifies the nearest-neighbor hoppings so that there is a position- and bond-dependent change $\delta t_j(\boldsymbol{r})$ in the hopping. The change in the Hamiltonian due to strain is then

$$\delta H_{\mathrm{s}} := -\sum_{\sigma} \sum_{j=1}^{3} \sum_{\boldsymbol{r} \in L/L_{\mathrm{BK}}} \delta t_j(\boldsymbol{r})\psi_{\sigma,A}^{\dagger}(\boldsymbol{r})\psi_{\sigma,B}(\boldsymbol{r} + \boldsymbol{\delta}_j - \boldsymbol{\delta}_1) + \text{h.c.} \tag{2.6}$$

---

[1]I denote the abstract space of annihilation/creation operators as Op. This is merely for notational convenience: its only purpose is to allow me to write the domains of maps, which in the standard $f : A \to B$ notation always needs also the codomain B.

Due to the L_{BK} periodicity there exists a discrete Fourier series which, by going to the effective low-energy theory by dropping terms far from the Dirac points, can be written as the valley expansion [40]

$$\psi_{\sigma,\alpha}(\boldsymbol{r}) = \sum_{\rho} \mathrm{e}^{\mathrm{i}\rho \boldsymbol{K}\cdot\boldsymbol{r}} \psi_{\sigma\rho,\alpha}(\boldsymbol{r}), \tag{2.7}$$

where $\rho \in \{+,-\}$ is the valley index and $\psi_{\sigma\rho,\alpha}$ is a slowly varying (*i.e.* has only small $\boldsymbol{k}$ Fourier components, compared to $\boldsymbol{K}$) valley-specific annihilation operator. Substituting the valley expansion to the Hamiltonians, linearizing $\psi_{\sigma\rho,B}(\boldsymbol{r}+\boldsymbol{\delta}_j-\boldsymbol{\delta}_1) \approx \psi_{\sigma\rho,B}(\boldsymbol{r})+(\boldsymbol{\delta}_j-\boldsymbol{\delta}_1)\cdot\nabla\psi_{\sigma\rho,B}(\boldsymbol{r})$ in the pristine Hamiltonian, and writing to zeroth order $\psi_{\sigma\rho,B}(\boldsymbol{r}+\boldsymbol{\delta}_j-\boldsymbol{\delta}_1) \approx \psi_{\sigma\rho,B}(\boldsymbol{r})$ in the strain Hamiltonian then yields the total Hamiltonian

$$H := H_{\mathrm{p}} + \delta H_{\mathrm{s}} \approx \sum_{\sigma\rho} \int_{\mathbb{R}^2/L_{\mathrm{BK}}} \mathrm{d}\boldsymbol{r}\, \psi_{\sigma\rho}^{\dagger}(\boldsymbol{r}) \mathcal{H}^{\rho}(\boldsymbol{r}) \psi_{\sigma\rho}(\boldsymbol{r}). \tag{2.8}$$

Here I also went to the continuum by approximating and redefining

$$\sum_{\boldsymbol{r}\in L/L_{\mathrm{BK}}} \approx \frac{1}{|\boldsymbol{a}_1 \times \boldsymbol{a}_2|} \int_{\mathbb{R}^2/L_{\mathrm{BK}}} \mathrm{d}\boldsymbol{r} \quad \text{and} \quad \psi_{\sigma\rho}/\sqrt{|\boldsymbol{a}_1 \times \boldsymbol{a}_2|} \mapsto \psi_{\sigma\rho}, \tag{2.9}$$

with the new operators now defined in the continuum $\psi_{\sigma\rho} : \mathbb{R}^2/L_{\mathrm{BK}} \to \mathrm{Op}^2$. Furthermore

$$\psi_{\sigma\rho} := (\psi_{\sigma\rho,A}, \psi_{\sigma\rho,B})^{\mathsf{T}} \tag{2.10}$$

is a sublattice-space vector of the field operators,

$$\mathcal{H}^{\rho}(\boldsymbol{r}) := \hbar v_{\mathrm{F}} \boldsymbol{\sigma}^{\rho} \cdot (-\mathrm{i}\nabla + \rho \boldsymbol{A}(\boldsymbol{r})) - \mu \tag{2.11}$$

is the Hamiltonian matrix element (also defined in the continuum if $\boldsymbol{A}$ is), the Fermi velocity is defined through $\hbar v_{\mathrm{F}} := \sqrt{3}at/2$, $\boldsymbol{\sigma}^{\rho} := (\rho\sigma_x, \sigma_y)$ is a vector of sublattice-space Pauli matrices, and the strain enters as a pseudo vector potential $\boldsymbol{A} = (A_x, A_y)$ with the components given by

$$\rho A_x(\boldsymbol{r}) + \mathrm{i}A_y(\boldsymbol{r}) := -\frac{\rho}{\hbar v_{\mathrm{F}}} \sum_j \mathrm{e}^{-\mathrm{i}\rho\boldsymbol{K}\cdot\boldsymbol{\delta}_j} \delta t_j(\boldsymbol{r}) \tag{2.12}$$

(also defined in continuum if δt_j is). Thus the strain behaves almost like a vector potential, but with the important difference that it changes sign between valleys, preserving time-reversal symmetry $\mathcal{H}^{\bar{\rho}*} = \mathcal{H}^{\rho}$ if $\boldsymbol{A}$ is real. Below we see that $\boldsymbol{A}$ is indeed real, at least in the case of small strain.

Equation (2.12) gives the relation between the pseudo vector potential and the modification in the hoppings, but we also want to get a relation between the hoppings and the strain itself. This is easily done by linearizing

$$\delta t_j(\boldsymbol{r}) \approx \frac{\mathrm{d}t}{\mathrm{d}a_0} \delta u_j(\boldsymbol{r}) = -\frac{t\beta_{\mathrm{G}}}{a_0} \delta u_j(\boldsymbol{r}), \tag{2.13}$$

where $\beta_{\mathrm{G}} := -\,\mathrm{d}\ln t/\mathrm{d}\ln a_0 \approx 2$ is the graphene Grüneisen parameter [113], $a_0 := \|\boldsymbol{\delta}_j\| = a/\sqrt{3}$ is the carbon–carbon bond length, and $\delta u_j(\boldsymbol{r})$ is the change in the length of the $\boldsymbol{\delta}_j$ bond. If the atom at $\boldsymbol{r}$ is displaced by $\boldsymbol{v}(\boldsymbol{r}) := (\boldsymbol{u}(\boldsymbol{r}), h(\boldsymbol{r}))$, where $\boldsymbol{u} = (u_x, u_y)$ is the in-plane and h the out-of-plane displacement field, the change in the bond length becomes

$$\delta u_j(\boldsymbol{r}) = \|\boldsymbol{r} + \boldsymbol{\delta}_j + \boldsymbol{v}(\boldsymbol{r} + \boldsymbol{\delta}_j) - [\boldsymbol{r} + \boldsymbol{v}(\boldsymbol{r})]\| - \|\boldsymbol{r} + \boldsymbol{\delta}_j - \boldsymbol{r}\| \tag{2.14}$$

$$\approx \frac{1}{a_0}\left[\boldsymbol{\delta}_j \cdot (\boldsymbol{u}(\boldsymbol{r} + \boldsymbol{\delta}_j) - \boldsymbol{u}(\boldsymbol{r})) + \frac{1}{2}(h(\boldsymbol{r} + \boldsymbol{\delta}_j) - h(\boldsymbol{r}))^2\right], \tag{2.15}$$

where in the second step I linearized in

$$\|\boldsymbol{u}(\boldsymbol{r} + \boldsymbol{\delta}_j) - \boldsymbol{u}(\boldsymbol{r})\|, \|h(\boldsymbol{r} + \boldsymbol{\delta}_j) - h(\boldsymbol{r})\| \ll \|\boldsymbol{\delta}_j\| = a_0. \tag{2.16}$$

Furthermore in the continuum we may approximate in the linear order

$$\boldsymbol{u}(\boldsymbol{r} + \boldsymbol{\delta}_j) - \boldsymbol{u}(\boldsymbol{r}) \approx (\boldsymbol{\delta}_j \cdot \nabla)\boldsymbol{u}(\boldsymbol{r}), \quad h(\boldsymbol{r} + \boldsymbol{\delta}_j) - h(\boldsymbol{r}) \approx \boldsymbol{\delta}_j \cdot \nabla h(\boldsymbol{r}), \tag{2.17}$$

and if we define the strain tensor as

$$u_{ij} := \frac{1}{2}(\partial_i u_j + \partial_j u_i) + \frac{1}{2}\partial_i h \partial_j h \tag{2.18}$$

$(i, j \in \{x, y\})$ the change in the bond length can be written as[2]

$$\delta u_j(\boldsymbol{r}) = \frac{1}{a_0}\left[\delta_{jx}^2 u_{xx}(\boldsymbol{r}) + 2\delta_{jx}\delta_{jy}u_{xy}(\boldsymbol{r}) + \delta_{jy}^2 u_{yy}(\boldsymbol{r})\right]. \tag{2.19}$$

We then find out the relation

$$\boldsymbol{A} = -\frac{\beta_{\mathrm{G}}}{2a_0}(u_{xx} - u_{yy}, -2u_{xy}) \tag{2.20}$$

between the pseudo vector potential and strain, which also means that $\boldsymbol{A}$ is indeed real.

Note that even though we are in the continuum theory, Eq. (2.20) clearly breaks the $O(2)$ rotation symmetry of the continuum. This is because the original lattice is still partly present through the $\boldsymbol{\delta}_j$'s in Eq. (2.12). So when applying Eq. (2.20), one should remember that it only applies in the coordinate system where the graphene primitive vectors are those in Eq. (2.1).

In addition to the pseudo vector potential described above, strain can have also other effects. Had we included also next-nearest-neighbor hopping, we would notice that modulations in the next-nearest-neighbor hopping leads to a scalar potential $\phi \propto u_{xx} + u_{yy}$ [40, 112–115], which is nonvanishing whenever (local) isotropic dilation or contraction is present. It adds to the Hamiltonian diagonally as $\mathcal{H}^\rho \mapsto \mathcal{H}^\rho + \phi$, which means it acts similarly as the chemical potential, and thus serves to (locally) change the charge density.

---

[2]Note that δ_{jx} and δ_{jy} are the components of $\boldsymbol{\delta}_j$, not Kronecker delta symbols.

2.1.1 Noninteracting Hamiltonian in Fourier space

At this point it is useful to assume the strain (and thus the pseudo vector potential $\boldsymbol{A}$) to be periodic, as then the diagonalization of H is easy to do in Fourier space. Taking the field operators $\psi_{\sigma\rho} : \mathbb{R}^2/L_{\mathrm{BK}} \to \mathrm{Op}^2$ to be periodic in translations of the large Born–von Kármán lattice L_{BK} and the pseudo vector potential $\boldsymbol{A} : \mathbb{R}^2/SL \to \mathbb{R}^2$ to be periodic in translations of the arbitrary superlattice $SL = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{t}_1, \boldsymbol{t}_2\}$, we may write them as the Fourier series

$$\psi_{\sigma\rho}(\boldsymbol{r}) = \sum_{\boldsymbol{k} \in L_{\mathrm{BK}}^*} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} c_{\sigma\rho}(\boldsymbol{k}), \quad \boldsymbol{A}(\boldsymbol{r}) = \sum_{\boldsymbol{G} \in SL_S^*} \mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} \tilde{\boldsymbol{A}}(\boldsymbol{G}) \qquad (2.21)$$

through Eq. (1.79). The mixed zone scheme $S = \mathrm{MZ}$ is applicable only if $\boldsymbol{A}$ is constant in the $\boldsymbol{t}_2$ direction, which we call the *1D potential* case in Publication I. Otherwise we call $\boldsymbol{A}$ a *2D potential*. Substituting these to Eq. (2.8) yields

$$H = V \sum_{\sigma\rho} \sum_{\boldsymbol{k} \in L_{\mathrm{BK}}^*} \sum_{\boldsymbol{G}' \in SL_S^*}$$
$$c_{\sigma\rho}^\dagger(\boldsymbol{k}) \left[\hbar v_{\mathrm{F}} \boldsymbol{\sigma}^\rho \cdot \left((\boldsymbol{k} - \boldsymbol{G}')\delta_{\boldsymbol{G}'0} + \rho\tilde{\boldsymbol{A}}(\boldsymbol{G}') \right) - \mu \right] c_{\sigma\rho}(\boldsymbol{k} + \boldsymbol{G}'). \qquad (2.22)$$

In the next step the chosen scheme plays a major role, so let us discuss the notion of the reduced and the mixed zone schemes. In the reduced zone scheme [33] $\boldsymbol{k} = k_1\boldsymbol{G}_1 + k_2\boldsymbol{G}_2 \in L_{\mathrm{BK}}^*/SL_{\mathrm{RZ}}^*$ is periodic both in the $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$ directions, with both $k_1, k_2 \in [-\frac{1}{2}, \frac{1}{2}[$ being periodic Bloch momenta. In the case of $\boldsymbol{A}$ being constant in the $\boldsymbol{t}_2$ direction (the 1D potential case) we are also allowed to use the mixed zone scheme, where $\boldsymbol{k} = k_1\boldsymbol{G}_1 + k_2\boldsymbol{G}_2 \in L_{\mathrm{BK}}^*/SL_{\mathrm{MZ}}^*$ is periodic only in the $\boldsymbol{G}_1$ direction but not in the $\boldsymbol{G}_2$ direction, with $k_1 \in [-\frac{1}{2}, \frac{1}{2}[$ being a periodic Bloch momentum and $k_2 \in]-\infty, \infty[$ being a nonperiodic real momentum. Thus in the traditional notion [33] the $\boldsymbol{G}_1$ direction is in the reduced zone and the $\boldsymbol{G}_2$ direction in the extended zone scheme, justifying the term mixed zone scheme.

We divide the $\boldsymbol{k}$ sum over the whole reciprocal space L_{BK}^* to the sum over the SBZ L_{BK}^*/SL_S^* plus shifted copies of this as

$$\sum_{\boldsymbol{k} \in L_{\mathrm{BK}}^*} g(\boldsymbol{k}) = \sum_{\boldsymbol{k} \in L_{\mathrm{BK}}^*/SL_S^*} \sum_{\boldsymbol{G} \in SL_S^*} g(\boldsymbol{k} + \boldsymbol{G}) \qquad (2.23)$$

(g being a test function). This division is illustrated in Fig. 2.3 in the different schemes. Then H becomes

$$H = V \sum_{\sigma\rho} \sum_{\boldsymbol{k} \in L_{\mathrm{BK}}^*/SL_S^*} \sum_{\boldsymbol{G}\boldsymbol{G}' \in SL_S^*} c_{\sigma\rho}^\dagger(\boldsymbol{k} + \boldsymbol{G}) \tilde{\mathcal{H}}_{\boldsymbol{G}\boldsymbol{G}'}^\rho(\boldsymbol{k}) c_{\sigma\rho}(\boldsymbol{k} + \boldsymbol{G}'), \qquad (2.24)$$

where the matrix element is

$$\tilde{\mathcal{H}}_{\boldsymbol{G}\boldsymbol{G}'}^\rho(\boldsymbol{k}) := \hbar v_{\mathrm{F}} \boldsymbol{\sigma}^\rho \cdot \left((\boldsymbol{k} + \boldsymbol{G})\delta_{\boldsymbol{G}\boldsymbol{G}'} + \rho\tilde{\boldsymbol{A}}(\boldsymbol{G} - \boldsymbol{G}') \right) - \mu\delta_{\boldsymbol{G}\boldsymbol{G}'}. \qquad (2.25)$$

Defining the $\boldsymbol{G}$-space matrix and vector respectively as

$$\underline{\tilde{\mathcal{H}}}^\rho(\boldsymbol{k}) := \left(\tilde{\mathcal{H}}_{\boldsymbol{G}\boldsymbol{G}'}^\rho(\boldsymbol{k}) \right)_{\boldsymbol{G}\boldsymbol{G}' \in SL_S^*}, \quad \underline{c}_{\sigma\rho}(\boldsymbol{k}) := \left(c_{\sigma\rho}(\boldsymbol{k} + \boldsymbol{G}) \right)_{\boldsymbol{G} \in SL_S^*}, \qquad (2.26)$$

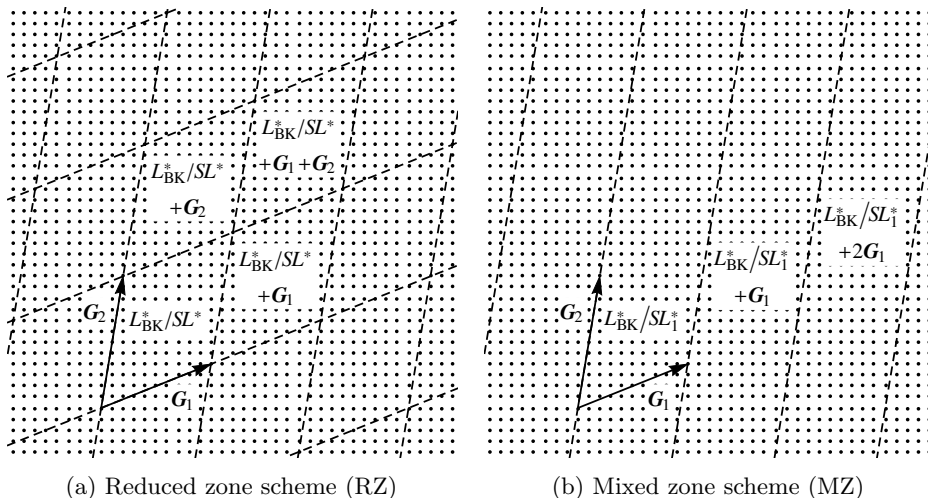(a) Reduced zone scheme (RZ)   (b) Mixed zone scheme (MZ)

Figure 2.3: Division of the dense reciprocal space L_{BK}^* (black dots) to the super-lattice Brillouin zone (SBZ) L_{BK}^*/SL_S^* plus shifted copies of this, as in Eq. (2.23). (a) In the reduced zone scheme the "$\boldsymbol{G}$ space" $SL_{\mathrm{RZ}}^* = SL^* = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{G}_1, \boldsymbol{G}_2\}$ is two dimensional while (b) in the mixed zone scheme the "$\boldsymbol{G}$ space" $SL_{\mathrm{MZ}}^* = SL_1^* = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{G}_1\}$ is one dimensional.

the Hamiltonian becomes

$$H = V \sum_{\sigma\rho} \sum_{\boldsymbol{k}\in L_{\mathrm{BK}}^*/SL_S^*} \underline{c}_{\sigma\rho}^\dagger(\boldsymbol{k}) \underline{\tilde{\mathcal{H}}}^\rho(\boldsymbol{k}) \underline{c}_{\sigma\rho}(\boldsymbol{k}). \tag{2.27}$$

Many of the electronic properties can then be calculated by diagonalizing the normal matrix $\underline{\tilde{\mathcal{H}}}^\rho(\boldsymbol{k})$. But note that the dimension of the Hamiltonian matrix $\underline{\tilde{\mathcal{H}}}^\rho(\boldsymbol{k})$ and the vector $\underline{c}_{\sigma\rho}(\boldsymbol{k})$ is countably infinite $2 \times |SL_S^*|$. Thus in numerical calculations the set SL_S^* has to be restricted to some of its finite subsets; in Publication I we take $SL_{\mathrm{MZ}} \approx \{m_1\boldsymbol{G}_1 : -n \leq m_1 \leq n\}$ in the mixed zone scheme (MZ) and $SL_{\mathrm{RZ}} \approx \{m_1\boldsymbol{G}_1 + m_2\boldsymbol{G}_2 : -n \leq m_1, m_2 \leq n\}$ in the reduced zone scheme (RZ) for some $n \in \mathbb{N}$. The needed cutoff n then depends on the form of the used potential $\boldsymbol{A}$: if its Fourier coefficients die slowly with increasing $\|\boldsymbol{G}\|$ then a larger n is needed. On the other hand in the mixed zone scheme the SBZ L_{BK}^*/SL_S^* is infinite in the $\boldsymbol{G}_2$ direction and thus the $\boldsymbol{k}$ sum also needs a cutoff in this direction. These cutoffs together roughly correspond to a cutoff in energy, which further corresponds to the cutoff in the electron–phonon interaction as in Eq. (1.29).

2.2 Electronic properties of the normal state

In Publication I we used the pseudo vector potential $\boldsymbol{A}$ as a tunable parameter and studied the effects of some simple potentials. The first, simplest, choice is

$$\boldsymbol{A}_{\cos}^{1D}(x, y) := \frac{\beta}{d}(0, \cos(2\pi x/d)), \tag{2.28}$$

which was also studied by Kauppila *et al.* [67]. Here β is a dimensionless parameter characterizing the strength of the strain and d is the period. The second potential is

$$\boldsymbol{A}_c^{1D}(x, y) := \frac{\beta}{d}(0, \text{triangleSquare}_c(x/d)), \tag{2.29}$$

which allows studying the effect of the potential shape. Here the slope parameter $c \in [4, \infty[$ can be used to interpolate between the triangle ($c = 4$) and square ($c \to \infty$) waveforms. Both of these potentials are periodic in translations of the (super)lattice $SL = \text{span}_{\mathbb{Z}}\{\boldsymbol{t}_1, \boldsymbol{t}_2\}$ with $\boldsymbol{t}_1 = (d, 0)$ and $\boldsymbol{t}_2 = (0, d)$. In fact they are both constant in the $\boldsymbol{t}_2$ direction, so they are 1D potentials in the language of Publication I. To study the effect of a potential varying in two directions, we studied also the simplest generalization of the 1D cosine potential,

$$\boldsymbol{A}_{\cos}^{2D}(x, y) := \frac{\beta}{d}(\cos(2\pi y/d), \cos(2\pi x/d)), \tag{2.30}$$

which is also periodic in translations of $SL = \text{span}_{\mathbb{Z}}\{\boldsymbol{t}_1, \boldsymbol{t}_2\}$ with $\boldsymbol{t}_1 = (d, 0)$ and $\boldsymbol{t}_2 = (0, d)$. But it is not constant in any direction, so it is a 2D potential in the language of Publication I.

From Eq. (2.20) it is easy to see that the cosine potentials can be produced *e.g.* by the in-plane displacement fields

$$\boldsymbol{u}_{\cos}^{1D}(x, y) := \frac{\beta a_0}{\beta_G \pi}(0, \sin(2\pi x/d)), \tag{2.31}$$

$$\boldsymbol{u}_{\cos}^{2D}(x, y) := \frac{\beta a_0}{\beta_G \pi}(0, \sin(2\pi x/d) + \sin(2\pi y/d)), \tag{2.32}$$

which are illustrated in Figs. 2.4a and 2.4b. But note that they are not necessarily unique: for example a field combining both in-plane and out-of-plane displacements might also be possible.

It turned out that an important parameter characterizing the flat-band physics is B, the amplitude of the pseudomagnetic field $\boldsymbol{B} := \nabla \times \boldsymbol{A} = \partial_x A_y - \partial_y A_x$. The pseudomagnetic fields produced by the cosine potentials are shown in Figs. 2.4c and 2.4d and the amplitudes for the three potentials are

$$B_{\cos}^{1D} = \frac{2\pi\beta}{d^2}, \quad B_c^{1D} = \frac{c\beta}{d^2}, \quad B_{\cos}^{2D} = \frac{4\pi\beta}{d^2}. \tag{2.33}$$

A realistic scale for β is given by the experiment of Jiang *et al.* [71] (Fig. 2.1), where a pseudomagnetic field amplitude of $\frac{\hbar}{e}B \approx 100\,\text{T}$ was observed for a strain period of $d = 14\,\text{nm}$. For the 1D cosine potential this gives $\beta \approx 5$ and for the 2D

(a) $\boldsymbol{u}_{\mathrm{cos}}^{\mathrm{1D}}$

(b) $\boldsymbol{u}_{\mathrm{cos}}^{\mathrm{2D}}$

(c) $\boldsymbol{A}_{\mathrm{cos}}^{\mathrm{1D}}$

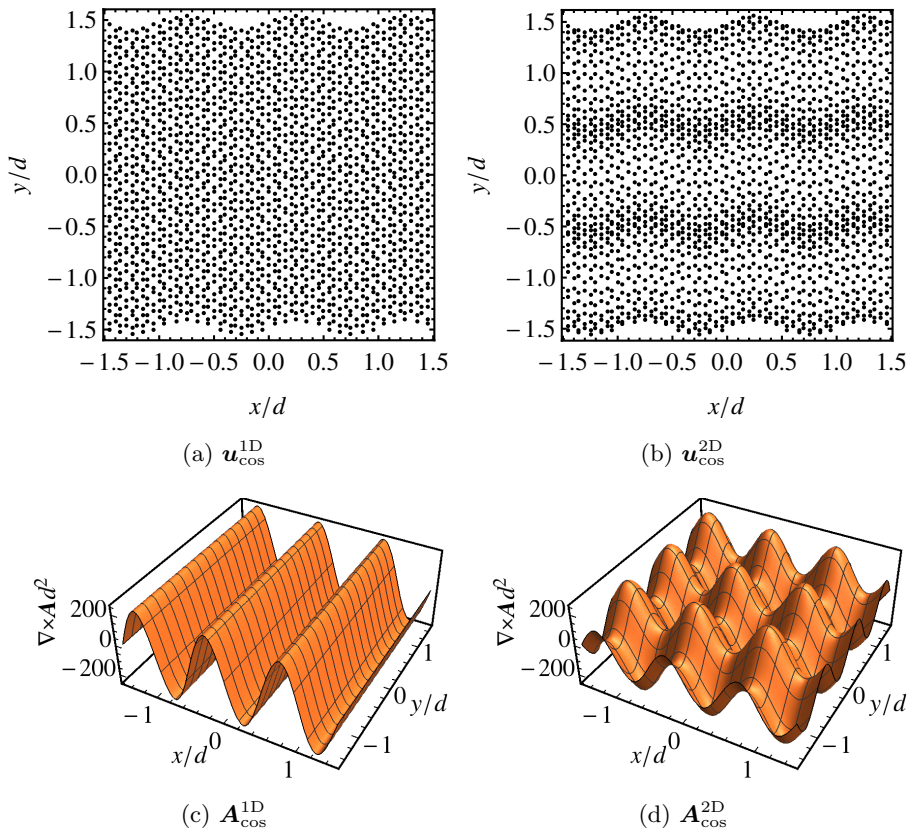(d) $\boldsymbol{A}_{\mathrm{cos}}^{\mathrm{2D}}$

Figure 2.4: (a,b) Example in-plane displacement fields $\boldsymbol{u}$ in Eqs. (2.31) and (2.32) that produce the studied cosine-type pseudo vector potentials $\boldsymbol{A}_{\mathrm{cos}}^{\mathrm{1D}}$ and $\boldsymbol{A}_{\mathrm{cos}}^{\mathrm{2D}}$ in Eqs. (2.28) and (2.30) through Eq. (2.20), shown here for exaggeratedly large β and small d. (c,d) Pseudomagnetic fields $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ produced by the pseudo vector potentials $\boldsymbol{A}_{\mathrm{cos}}^{\mathrm{1D}}$ and $\boldsymbol{A}_{\mathrm{cos}}^{\mathrm{2D}}$, shown here for $\beta = 40$ and $\beta = 20$, respectively. The amplitude B of these pseudomagnetic fields is given in Eq. (2.33). Reprinted with permission from Ref. [116]. Copyright 2020 IOP Publishing.

cosine potential $\beta \approx 2.5$. But to better be in the flat-band regime, in Publication I we used a factor of 4 to 8 times higher values of β.

The dispersions, calculated from the Hamiltonian in Eqs. (2.25)–(2.27), exhibit flat bands for high enough B, as illustrated in Fig. 2.5. The 1D potential is shown both in the mixed zone (MZ) and reduced zone (RZ) schemes, while the 2D potential is shown only in the reduced zone scheme, which is the only possibility in that case. The parameter B can be used to control the flatness of the bands; in the reduced zone scheme one can roughly say that B determines the number of flat bands.

Calculating the density of states (Fig. 2.6) shows how the flat bands emerge. The smallest B (*i.e.* β or c for the chosen potentials) is close to the conical graphene

(a) $\boldsymbol{A}_{\cos}^{1D}$, $\beta = 30$,
mixed zone scheme (MZ)

(b) $\boldsymbol{A}_{\cos}^{1D}$, $\beta = 30$,
reduced zone scheme (RZ)

(c) $\boldsymbol{A}_{\cos}^{2D}$, $\beta = 15$,
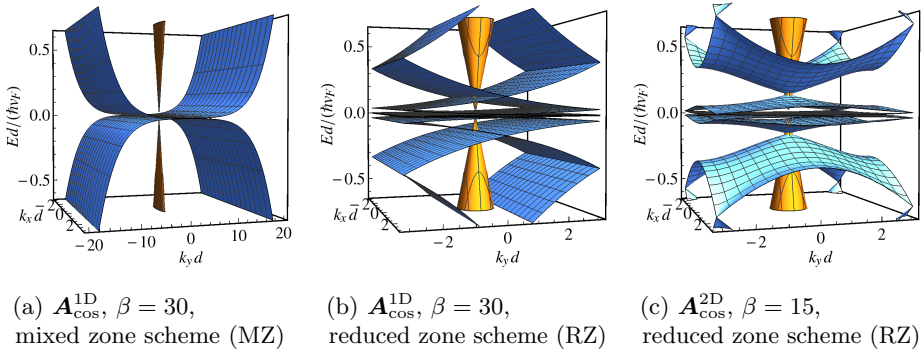reduced zone scheme (RZ)

Figure 2.5: Typical dispersions in the normal state at the valley $\rho = \pm$. The 1D cosine potential is shown both in the (a) mixed zone scheme and (b) reduced zone scheme. (c) The 2D cosine potential in the reduced zone scheme. The strained dispersions are in blue and the conical unstrained graphene dispersions in orange. Reprinted with permission from Ref. [116]. Copyright 2020 IOP Publishing.
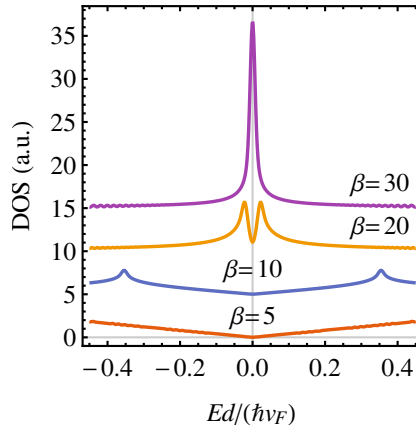

Figure 2.6: Density of states (DOS) for $\boldsymbol{A}_{\cos}^{1D}$ with increasing β. For clarity the successive curves are shifted vertically by 5 and the DOS is normalized such that the visible region integrates to unity.

dispersion and thus produces a linear DOS. Higher B produces peaks at finite energies, which approach zero energy with increasing B, with their amplitude increasing at the same time. In other words increasing B increases the flatness of the bands. Most of the results are calculated with the chemical potential μ chosen at the energy of one of the DOS peaks, which we call the *optimal doping*, because when studying superconductivity, at low interactions this amount of doping has the highest superconducting order parameter Δ.

On the other hand calculating the local density of states (LDOS, Fig. 8 of Publication I) shows that the low-energy states are localized at the extrema of the pseudomagnetic field $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ (Fig. 2.4). More specifically, the LDOS of

A (B) sublattice states are localized at the minima (maxima) of $\boldsymbol{B}$. This kind of localization was also measured by Jiang *et al.* [71].

The origin of the flat bands can be understood from many perspectives. Tang & Fu [117] discussed how they form as the Landau levels of the pseudomagnetic field $\boldsymbol{B}$. Heikkilä & Volovik [8] and Kauppila *et al.* [67] showed how PSG can be viewed as the Su–Schrieffer–Heeger model [118] with a periodic array of domain walls [119] hosting localized zero-energy states. Lastly, Tahir & Chen [70] showed how the flat bands can be understood to result from interference in a coarse-grained tight-binding model.

2.3 BCS equations for the superconducting state

2.3.1 Bogoliubov–de Gennes and self-consistency equations

To model the possible superconducting state we add an interaction Hamiltonian (1.10). But instead of choosing the eigenbasis of H as in Sec. 1.2.3, we stay in the tight-binding formalism written in real space, which is then transformed to the low-energy continuum formalism. At the tight-binding level the states are indexed by $\mu = (\sigma, \alpha, \boldsymbol{r})$, where $\sigma \in \{\uparrow, \downarrow\}$ is the spin, $\alpha \in \{A, B\}$ is the sublattice, and $\boldsymbol{r} \in L/L_{\mathrm{BK}}$ is the position in the lattice. We then choose the spin-singlet, local interaction

$$\lambda_{\mu\mu'\nu\nu'} = \lambda_{\sigma\alpha\boldsymbol{r},\sigma'\alpha'\boldsymbol{r}',\tau\beta\boldsymbol{s},\tau'\beta'\boldsymbol{s}'} = \delta_{\bar{\sigma}\sigma'}\delta_{\sigma\tau}\delta_{\sigma'\tau'}\delta_{\alpha\alpha'}\delta_{\beta\beta'}\delta_{\alpha\beta}\delta_{\boldsymbol{r}\boldsymbol{r}'}\delta_{\boldsymbol{s}\boldsymbol{s}'}\delta_{\boldsymbol{r}\boldsymbol{s}}\lambda, \quad (2.34)$$

where λ is the interaction strength (negative for attractive interaction) and the energy cutoff [120] resulting from the electron–phonon interaction [Eq. (1.29)] is left implicit. Wu *et al.* [121] actually showed that in graphene the electron–phonon interaction yields a local interaction, but only modulo sublattice, with a very specific sublattice structure that does not strictly match Eq. (2.34). We nevertheless use the simplified form (2.34) since it gives similar results in the more important (in the case of TBG) s-wave channel [121].

The choice (2.34) is more general than the interaction in Eqs. (1.24) and (1.29), since in Fourier space the position-Kronecker-deltas yield only $\delta_{\boldsymbol{k}+\boldsymbol{k}'+\boldsymbol{l}+\boldsymbol{l}',0}$, so we cannot use the results of Sec. 1.2.3 which assume more strictly $\delta_{\bar{\boldsymbol{k}}\boldsymbol{k}'}\delta_{\bar{\boldsymbol{l}}\boldsymbol{l}'}$. This more general choice allows having a (center-of-mass) position-dependent order parameter, which turns out to be important in the case of PSG and TBG. This choice yields the order parameter

$$\Delta_{\sigma\alpha\boldsymbol{r},\sigma'\alpha'\boldsymbol{r}'} = \delta_{\bar{\sigma}\sigma'}\delta_{\alpha\alpha'}\delta_{\boldsymbol{r}\boldsymbol{r}'}\Delta_{\sigma,\alpha}(\boldsymbol{r}) \quad \text{with} \quad \Delta_{\sigma,\alpha}(\boldsymbol{r}) \coloneqq \lambda\,\langle\psi_{\bar{\sigma},\alpha}(\boldsymbol{r})\psi_{\sigma,\alpha}(\boldsymbol{r})\rangle \tag{2.35}$$

and the interaction Hamiltonian

$$H_{\mathrm{int}} = \frac{1}{2}\sum_{\sigma\alpha\boldsymbol{r}}\Delta_{\sigma,\alpha}(\boldsymbol{r})\psi^{\dagger}_{\sigma,\alpha}(\boldsymbol{r})\psi^{\dagger}_{\bar{\sigma},\alpha}(\boldsymbol{r}) + \text{h.c.} + \text{const.} \tag{2.36}$$

This corresponds to spin-singlet pairing because $\Delta_{\sigma'\alpha\boldsymbol{r},\sigma\alpha'\boldsymbol{r}'} = -\Delta_{\sigma\alpha\boldsymbol{r},\sigma'\alpha'\boldsymbol{r}'}$, s-wave pairing because Δ depends only on the COM coordinate $(\boldsymbol{r} + \boldsymbol{r}')/2 = \boldsymbol{r}$

through $\delta_{\boldsymbol{rr'}}$ and is independent of the relative coordinate $\boldsymbol{r} - \boldsymbol{r'} = 0$, and orbital-triplet pairing because $\Delta_{\sigma\alpha'\boldsymbol{r},\sigma'\alpha\boldsymbol{r}} = \Delta_{\sigma\alpha\boldsymbol{r},\sigma'\alpha'\boldsymbol{r'}}$.

Making the valley expansion (2.7) gives

$$H_{\text{int}} = \frac{1}{2} \sum_{\sigma\rho\alpha\boldsymbol{r}} \Delta_{\sigma,\alpha}(\boldsymbol{r}) \psi_{\sigma\rho,\alpha}^\dagger(\boldsymbol{r}) \psi_{\bar\sigma\bar\rho,\alpha}^\dagger(\boldsymbol{r}) + \text{h.c.} + \text{const} \tag{2.37}$$

and

$$\Delta_{\sigma,\alpha}(\boldsymbol{r}) = \lambda \sum_\rho \langle \psi_{\bar\sigma\bar\rho,\alpha}(\boldsymbol{r}) \psi_{\sigma\rho,\alpha}(\boldsymbol{r}) \rangle, \tag{2.38}$$

assuming only intervalley interaction. Here $\bar\rho = -\rho$ denotes the opposite valley of ρ and the choice of intervalley is chosen because it most closely matches the "standard" BCS theory choice (1.24) of coupling opposite momenta. Defining the sublattice-space matrix

$$\Delta_\sigma := \text{diag}(\Delta_{\sigma,A}, \Delta_{\sigma,B}), \tag{2.39}$$

using the sublattice-vectors $\psi_{\sigma\rho}$ of Eq. (2.10), and going to the continuum as in Eq. (2.8) gives the total interacting *Bogoliubov–de Gennes (BdG) Hamiltonian*

$$H_{\text{BdG}} = H + H_{\text{int}} \tag{2.40}$$
$$= \sum_{\sigma\rho} \int d\boldsymbol{r} \left[\psi_{\sigma\rho}^\dagger(\boldsymbol{r}) \mathcal{H}^\rho(\boldsymbol{r}) \psi_{\sigma\rho}(\boldsymbol{r}) + \frac{1}{2} \psi_{\sigma\rho}^\dagger(\boldsymbol{r}) \Delta_\sigma(\boldsymbol{r}) \psi_{\bar\sigma\bar\rho}^{\dagger\mathsf{T}}(\boldsymbol{r}) + \text{h.c.} \right] + \text{const.}$$

The following derivation is based on the book of Nazarov & Danon [15], and its details are presented in the Supplementary Material of Publication I. But because I use many of the identities when deriving the superfluid weight in Sec. 2.3.2, I present the main points here. For diagonalizing H_{BdG} it is convenient to first write it in the *Nambu basis.* Using the symmetry

$$\int d\boldsymbol{r}\, \psi_{\sigma\rho}^\dagger(\boldsymbol{r}) \mathcal{H}^\rho(\boldsymbol{r}) \psi_{\sigma\rho}(\boldsymbol{r}) = - \int d\boldsymbol{r}\, \psi_{\sigma\rho}^\mathsf{T}(\boldsymbol{r}) \mathcal{H}^{\bar\rho}(\boldsymbol{r}) \psi_{\sigma\rho}^{\dagger\mathsf{T}}(\boldsymbol{r}) + \text{const} \tag{2.41}$$

of the noninteracting Hamiltonian, we may write the BdG Hamiltonian in the Nambu basis as

$$H_{\text{BdG}} = \frac{1}{2} \sum_{\sigma\rho} \int d\boldsymbol{r}\, \Psi_{\sigma\rho}^\dagger(\boldsymbol{r}) \mathcal{H}_{\text{BdG}}^\rho(\boldsymbol{r}) \Psi_{\sigma\rho}(\boldsymbol{r}) + \text{const.} \tag{2.42}$$

Here the spin-independent BdG-Hamiltonian matrix element in the Nambu space, the spin-independent order parameter, and the annihilation-operator Nambu-vector are

$$\mathcal{H}_{\text{BdG}}^\rho := \begin{pmatrix} \mathcal{H}^\rho & \Delta \\ \Delta^* & -\mathcal{H}^\rho \end{pmatrix}, \quad \Delta := s(\sigma)\Delta_\sigma, \quad \Psi_{\sigma\rho} := \begin{pmatrix} \psi_{\sigma\rho} \\ s(\sigma)\psi_{\bar\sigma\bar\rho}^{\dagger\mathsf{T}} \end{pmatrix}, \tag{2.43}$$

respectively, where $s(\uparrow) := 1$ and $s(\downarrow) := -1$.

H_{BdG} can be diagonalized through the resolution of identity in the position basis,

$$\delta(\boldsymbol{r} - \boldsymbol{r}') = \frac{1}{V} \sum_n w_{\rho n}(\boldsymbol{r}) w_{\rho n}^{\dagger}(\boldsymbol{r}'), \qquad (2.44)$$

where n enumerates the eigenfunctions $w_{\rho n}(\boldsymbol{r})$ of $\mathcal{H}_{\mathrm{BdG}}^{\rho}(\boldsymbol{r})$,

$$\mathcal{H}_{\mathrm{BdG}}^{\rho}(\boldsymbol{r}) w_{\rho n}(\boldsymbol{r}) = E_{\rho n} w_{\rho n}(\boldsymbol{r}). \qquad (2.45)$$

The eigenvalue equation (2.45) is called the *Bogoliubov–de Gennes equation*, and it is one of the equations that need to be solved. Writing the Nambu structure explicitly it reads

$$\begin{pmatrix} \mathcal{H}^{\rho}(\boldsymbol{r}) & \Delta(\boldsymbol{r}) \\ \Delta^*(\boldsymbol{r}) & -\mathcal{H}^{\rho}(\boldsymbol{r}) \end{pmatrix} \begin{pmatrix} u_{\rho n}(\boldsymbol{r}) \\ v_{\rho n}(\boldsymbol{r}) \end{pmatrix} = E_{\rho n} \begin{pmatrix} u_{\rho n}(\boldsymbol{r}) \\ v_{\rho n}(\boldsymbol{r}) \end{pmatrix}. \qquad (2.46)$$

Then by defining the *Bogoliubon operator* or the *Bogoliubov transformation*

$$\gamma_{\sigma \rho n} := \frac{1}{\sqrt{V}} \int \mathrm{d}\boldsymbol{r}\, w_{\rho n}^{\dagger}(\boldsymbol{r}) \Psi_{\sigma \rho}(\boldsymbol{r}) \qquad (2.47)$$

H_{BdG} is diagonalized to

$$H_{\mathrm{BdG}} = \frac{1}{2} \sum_{\sigma \rho n} E_{\rho n} \gamma_{\sigma \rho n}^{\dagger} \gamma_{\sigma \rho n} + \mathrm{const}. \qquad (2.48)$$

Equation (2.48) shows that we have diagonalized the many-body BdG Hamiltonian H_{BdG} by transforming the problem into the diagonalization of the one-body BdG Hamiltonian $\mathcal{H}_{\mathrm{BdG}}^{\rho}$, Eq. (2.45).

There is a problem in the diagonal form of Eq. (2.48): the Bogoliubon γ's are not generally fermionic operators, and thus we do not know whether "const" measures the ground state energy and $E_{\rho n}$ the excitation energies. Next we remove this problem by showing that the *positive-energy* Bogoliubons are fermionic and that there is a symmetry between the positive- and negative-energy Bogoliubons.

Taking the field operators $\psi_{\sigma \rho}$, the pseudo vector potential $\boldsymbol{A}$, and the order parameter Δ to be periodic in translations of the Born–von Kármán lattice L_{BK}, we know by Bloch's theorem that the eigenstates can be enumerated with $n = (b, \boldsymbol{k})$, where $\boldsymbol{k} \in L_{\mathrm{BK}}^*$ belongs to the reciprocal space and $b = (\eta, \nu)$ enumerates the bands for each $\boldsymbol{k}$ with $\nu \in \{+, -\}$ giving the sign of energy of this band. Using time-reversal symmetry $\mathcal{H}^{\bar{\rho}*} = \mathcal{H}^{\rho}$ implies the symmetry

$$w_{\rho \eta \bar{\nu} \boldsymbol{k}} = \mathrm{i} \tau_y w_{\bar{\rho} \eta \nu \bar{\boldsymbol{k}}}^* \qquad (2.49)$$

between the positive- and negative-energy eigenstates. This then directly gives the symmetry

$$\gamma_{\sigma \rho \eta \bar{\nu} \boldsymbol{k}} = s(\bar{\sigma}) \gamma_{\bar{\sigma} \bar{\rho} \eta \nu \bar{\boldsymbol{k}}}^{\dagger} \qquad (2.50)$$

for the Bogoliubons.

Choosing an orthogonal eigenbasis

$$\int \mathrm{d}\boldsymbol{r}\, w_{\rho n}^{\dagger}(\boldsymbol{r}) w_{\rho n'}(\boldsymbol{r}) = V \delta_{nn'} \qquad (2.51)$$

gives the first fermionic anticommutation relation $\left\{\gamma_{\sigma\rho n}, \gamma_{\sigma'\rho'n'}^{\dagger}\right\} = \delta_{\sigma\sigma'}\delta_{\rho\rho'}\delta_{nn'}$ for all n, n' and the second anticommutation relation $\left\{\gamma_{\sigma\rho\eta\nu\boldsymbol{k}}, \gamma_{\sigma'\rho'\eta'\nu'\boldsymbol{k}'}\right\} = 0$ provided $\nu = \nu'$, *i.e.* that they both are either positive- or negative-energy operators. Concentrating then only on the positive-energy ones, we get the desired result that the Bogoliubons are fermionic,

$$\left\{\gamma_{\sigma\rho\eta+\boldsymbol{k}}, \gamma_{\sigma'\rho'\eta'+\boldsymbol{k}'}^{\dagger}\right\} = \delta_{\sigma\sigma'}\delta_{\rho\rho'}\delta_{\eta\eta'}\delta_{\boldsymbol{k}\boldsymbol{k}'}, \tag{2.52}$$

$$\left\{\gamma_{\sigma\rho\eta+\boldsymbol{k}}, \gamma_{\sigma'\rho'\eta'+\boldsymbol{k}'}\right\} = 0. \tag{2.53}$$

Finally utilizing the symmetry (2.50) and the anticommutation relations (2.52) and (2.53) gives the diagonal and fermionic form

$$H_{\mathrm{BdG}} = \sum_{\sigma\rho n_+} E_{\rho n_+} \gamma_{\sigma\rho n_+}^{\dagger} \gamma_{\sigma\rho n_+} + \mathrm{const}, \tag{2.54}$$

where $n_+ = (\eta, +, \boldsymbol{k})$ is summed over the positive-energy states.

It is useful to write the definition (2.38), (2.39), and (2.43) of the order parameter Δ in the same Bogoliubon basis as the Hamiltonian in Eq. (2.54) because we know how their thermal averages are calculated. Since the positive-energy Bogoliubons are fermionic and we assume no interactions between them, they follow the Fermi–Dirac statistics

$$\left\langle \gamma_{\sigma\rho n_+} \gamma_{\sigma'\rho'n_+'} \right\rangle = 0, \tag{2.55}$$

$$\left\langle \gamma_{\sigma\rho n_+}^{\dagger} \gamma_{\sigma'\rho'n_+'} \right\rangle = \delta_{\sigma\sigma'}\delta_{\rho\rho'}\delta_{n_+ n_+'} f(E_{\rho n_+}), \tag{2.56}$$

where $f(E) = [\mathrm{e}^{E/(k_{\mathrm{B}}T)} + 1]^{-1}$ is the Fermi–Dirac distribution at the temperature T.

To write Δ in the γ basis, we need to invert the Bogoliubov transformation (2.47). Using the orthogonality condition (2.51) together with the resolution of identity (2.44), the inverse transformation can be shown to be

$$\Psi_{\sigma\rho}(\boldsymbol{r}) = \frac{1}{\sqrt{V}} \sum_n w_{\rho n}(\boldsymbol{r})\gamma_{\sigma\rho n}. \tag{2.57}$$

By using the symmetries (2.49) and (2.50), it can also be written as a sum over the positive-energy states only as

$$\Psi_{\sigma\rho}(\boldsymbol{r}) = \frac{1}{\sqrt{V}} \sum_{n_+} \left(w_{\rho n_+}(\boldsymbol{r})\gamma_{\sigma\rho n_+} + \mathrm{i}\tau_y w_{\bar{\rho}n_+}^*(\boldsymbol{r})s(\bar{\sigma})\gamma_{\bar{\sigma}\bar{\rho}n_+}^{\dagger} \right). \tag{2.58}$$

This then yields the *self-consistency equation* for the order parameter at sublattice α,

$$\Delta_\alpha(\boldsymbol{r}) = -\frac{\lambda}{V} \sum_{\rho n_+} u_{\rho n_+,\alpha}(\boldsymbol{r})v_{\rho n_+,\alpha}^*(\boldsymbol{r}) \tanh\left(\frac{E_{\rho n_+}}{2k_{\mathrm{B}}T}\right), \tag{2.59}$$

where the eigenfunction components are $w_{\rho n} = (u_{\rho n}, v_{\rho n})^{\mathsf{T}}$ in the Nambu space and $u_{\rho n} = (u_{\rho n,A}, u_{\rho n,B})^{\mathsf{T}}$, $v_{\rho n} = (v_{\rho n,A}, v_{\rho n,B})^{\mathsf{T}}$ in the sublattice space. Solving the self-consistency equation is the main part of determining the superconducting properties.

2.3.2 Supercurrent and superfluid weight

In Publications I and III we calculate the superfluid weight directly from an equation derived by Liang *et al.* [93] for a generic lattice system by linear response theory. But to get a more intuitive feeling about the superfluid weight, I derive the expression here directly for PSG by first deriving the expression of the supercurrent operator and then linearizing that expression. The linearization process is based on the calculation by de Gennes [11] and Kopnin & Sonin [59].

To describe supercurrent we add a vector potential $\mathcal{A}$ to the system by the minimal substitution principle by replacing the Hamiltonian (2.8) with

$$H_{\mathcal{A}} := \sum_{\sigma\rho} \int_{\mathbb{R}^2/L_{\mathrm{BK}}} \mathrm{d}\boldsymbol{r}\, \psi_{\sigma\rho}^{\dagger}(\boldsymbol{r})\mathcal{H}_{\mathcal{A}}^{\rho}(\boldsymbol{r})\psi_{\sigma\rho}(\boldsymbol{r}), \tag{2.60}$$

where the matrix element is

$$\mathcal{H}_{\mathcal{A}}^{\rho}(\boldsymbol{r}) := \hbar v_{\mathrm{F}}\boldsymbol{\sigma}^{\rho} \cdot (-\mathrm{i}\nabla + \rho\boldsymbol{A}(\boldsymbol{r}) + \boldsymbol{\mathcal{A}}(\boldsymbol{r})) - \mu. \tag{2.61}$$

In this case the form of the BdG Hamiltonian (2.40) stays the same,

$$H_{\mathrm{BdG},\boldsymbol{\mathcal{A}}} := H_{\boldsymbol{\mathcal{A}}} + H_{\mathrm{int}} \tag{2.62}$$

$$= \sum_{\sigma\rho} \int \mathrm{d}\boldsymbol{r} \left[\psi_{\sigma\rho}^{\dagger}(\boldsymbol{r})\mathcal{H}_{\boldsymbol{\mathcal{A}}}^{\rho}(\boldsymbol{r})\psi_{\sigma\rho}(\boldsymbol{r}) + \frac{1}{2}\psi_{\sigma\rho}^{\dagger}(\boldsymbol{r})\Delta_{\sigma}(\boldsymbol{r})\psi_{\bar{\sigma}\bar{\rho}}^{\dagger\mathsf{T}}(\boldsymbol{r}) + \mathrm{h.c.} \right] + \mathrm{const.} \tag{2.63}$$

The symmetry (2.41) becomes

$$\int \mathrm{d}\boldsymbol{r}\, \psi_{\sigma\rho}^{\dagger}(\boldsymbol{r})\mathcal{H}_{\boldsymbol{\mathcal{A}}}^{\rho}(\boldsymbol{r})\psi_{\sigma\rho}(\boldsymbol{r}) = - \int \mathrm{d}\boldsymbol{r}\, \psi_{\sigma\rho}^{\mathsf{T}}(\boldsymbol{r})\mathcal{H}_{-\boldsymbol{\mathcal{A}}}^{\bar{\rho}}(\boldsymbol{r})\psi_{\sigma\rho}^{\dagger\mathsf{T}}(\boldsymbol{r}) + \mathrm{const}, \tag{2.64}$$

which leads to the Nambu form

$$H_{\mathrm{BdG},\boldsymbol{\mathcal{A}}} = \frac{1}{2}\sum_{\sigma\rho} \int \mathrm{d}\boldsymbol{r}\, \Psi_{\sigma\rho}^{\dagger}(\boldsymbol{r})\mathcal{H}_{\mathrm{BdG},\boldsymbol{\mathcal{A}}}^{\rho}(\boldsymbol{r})\Psi_{\sigma\rho}(\boldsymbol{r}) + \mathrm{const} \tag{2.65}$$

with the Nambu matrix

$$\mathcal{H}_{\mathrm{BdG},\boldsymbol{\mathcal{A}}}^{\rho} := \begin{pmatrix} \mathcal{H}_{\boldsymbol{\mathcal{A}}}^{\rho} & \Delta \\ \Delta^* & -\mathcal{H}_{-\boldsymbol{\mathcal{A}}}^{\rho} \end{pmatrix}. \tag{2.66}$$

Note the sign-difference in $\boldsymbol{\mathcal{A}}$: a naive minimal substitution to Eq. (2.43) would have yielded a different (wrong) result. The diagonalization process goes in the same way as without the vector potential, so that the BdG equation becomes

$$\mathcal{H}_{\mathrm{BdG},\boldsymbol{\mathcal{A}}}^{\rho}(\boldsymbol{r})w_{\rho n}^{\boldsymbol{\mathcal{A}}}(\boldsymbol{r}) = E_{\rho n}^{\boldsymbol{\mathcal{A}}}w_{\rho n}^{\boldsymbol{\mathcal{A}}}(\boldsymbol{r}) \tag{2.67}$$

and the Bogoliubov transformation and its inverse transformation are

$$\gamma_{\sigma\rho n}^{\boldsymbol{\mathcal{A}}} := \frac{1}{\sqrt{V}}w_{\rho n}^{\boldsymbol{\mathcal{A}}\dagger}(\boldsymbol{r})\Psi_{\sigma\rho}(\boldsymbol{r}) \quad \Leftrightarrow \quad \Psi_{\sigma\rho}(\boldsymbol{r}) = \frac{1}{\sqrt{V}}\sum_n w_{\rho n}^{\boldsymbol{\mathcal{A}}}(\boldsymbol{r})\gamma_{\sigma\rho n}^{\boldsymbol{\mathcal{A}}}. \tag{2.68}$$

The thermal average of the current operator is given by the functional derivative [122–124]

$$\boldsymbol{j}(\boldsymbol{r}) := -\left\langle \frac{\delta H_{\text{BdG},\boldsymbol{\mathcal{A}}}}{\delta \boldsymbol{\mathcal{A}}(\boldsymbol{r})} \right\rangle = -\frac{1}{2} \hbar v_{\text{F}} \sum_{\sigma\rho} \left\langle \Psi_{\sigma\rho}^{\dagger}(\boldsymbol{r}) \boldsymbol{\sigma}^{\rho} \Psi_{\sigma\rho}(\boldsymbol{r}) \right\rangle, \qquad (2.69)$$

where in the second step I assumed that Δ is independent of $\boldsymbol{\mathcal{A}}$. Strictly speaking, generally Δ has also corrections linear in $\boldsymbol{\mathcal{A}}$, but for example making the simplified assumption of an isotropic superconductor and choosing the London gauge $\nabla \cdot \boldsymbol{\mathcal{A}} = 0$ [11, 125] renders Δ independent of $\boldsymbol{\mathcal{A}}$ in linear order.

Substituting the inverse transformation (2.68), using the symmetry (2.50) to transform all the γ's to the positive-energy ones, and using the expectation values (2.55) and (2.56) then gives the current

$$\boldsymbol{j}(\boldsymbol{r}) = -\frac{\hbar v_{\text{F}}}{V} \sum_{\rho n} w_{\rho n}^{\boldsymbol{\mathcal{A}}\dagger}(\boldsymbol{r}) \boldsymbol{\sigma}^{\rho} w_{\rho n}^{\boldsymbol{\mathcal{A}}}(\boldsymbol{r}) f(E_{\rho n}^{\boldsymbol{\mathcal{A}}}) - \text{reg}, \qquad (2.70)$$

where n is summed over both the positive- and negative-energy states. This expression (without "reg"), however, diverges according to Kopnin & Sonin [34, 59], which is why I have inserted the regularization term "reg". Had it been derived for the full lattice Hamiltonian, the expression would be convergent due to the periodicity under reciprocal lattice translations, which cancels terms far from the Dirac points. But now in the continuum theory the periodicity is lost, which causes divergence from terms far from the Dirac points. Kopnin & Sonin [34, 59] removed this problem by shifting the $\boldsymbol{k}$ sum/integral (included in the n sum) properly.[3] Liang *et al.* [93], on the other hand, did the regularization (implicitly) by forcing the periodicity by setting the surface term to zero while doing partial integration of the "diamagnetic current operator". This leads to the same result as what Kopnin & Sonin found out. I write the form of "reg" only at the end of this calculation, because only there I know its proper form.

Next I linearize the current in $\boldsymbol{\mathcal{A}}$ by using standard perturbation theory. Strictly speaking we would need degenerate perturbation theory because of the degeneracy resulting from the Nambu structure of $\mathcal{H}_{\text{BdG}}^{\rho}$. But to keep the theory simple, I use only nondegenerate perturbation theory, but modify one step so that the end result matches that of Liang *et al.* [93]. I write $w_{\rho n}^{\boldsymbol{\mathcal{A}}} = w_{\rho n} + w_{\rho n}^{(1)}$ and $E_{\rho n}^{\boldsymbol{\mathcal{A}}} = E_{\rho n} + E_{\rho n}^{(1)}$, where $w_{\rho n}$ and $E_{\rho n}$ satisfy the BdG equation (2.45) with zero vector potential and $w_{\rho n}^{(1)}$ and $E_{\rho n}^{(1)}$ are corrections proportional to the vector potential $\boldsymbol{\mathcal{A}}$. By keeping only terms linear in $\boldsymbol{\mathcal{A}}$, the BdG equation becomes

$$\left(\mathcal{H}_{\text{BdG}}^{\rho}(\boldsymbol{r}) - E_{\rho n} \right) w_{\rho n}^{(1)}(\boldsymbol{r}) + \left(\hbar v_{\text{F}} \boldsymbol{\sigma}^{\rho} \cdot \boldsymbol{\mathcal{A}}(\boldsymbol{r}) - E_{\rho n}^{(1)} \right) w_{\rho n}(\boldsymbol{r}) = 0. \qquad (2.71)$$

---

[3]Kopnin & Sonin [59] also claim that subtracting the unphysical normal-state supercurrent is an equivalent way to regularize the expression. This, however, is not true. If one did the regularization by subtracting the normal-state supercurrent one would get an extra $-|\mu|$ inside the brackets of Eq. (32) in Ref. [59]. Note also that Eq. (32) in Ref. [59] [or Eq. (C3) in Ref. [93]] is not totally correct, because at finite μ it yields a nonzero supercurrent even in the normal state $\Delta = 0$. This is a quirk of taking both the $T \to 0$ and $\Delta \to 0$ limits, as in this case the result actually measures the Drude weight instead of the superfluid weight [126]. This problem can be removed by staying at finite temperatures, in which case the finite-temperature equation (31) in Ref. [59] correctly gives (numerically) $D^{\text{s}} \to 0$ when $\Delta \to 0$.

Expanding in the $\boldsymbol{\mathcal{A}} = 0$ states through

$$w_{\rho n}^{(1)}(\boldsymbol{r}) = \frac{1}{V} \sum_m a_{nm}^\rho w_{\rho m}(\boldsymbol{r}), \tag{2.72}$$

multiplying from the left with $w_{\rho m'}^\dagger(\boldsymbol{r})$, and integrating over $\boldsymbol{r}$ then gives

$$(E_{\rho m} - E_{\rho n})a_{nm}^\rho + S_{nm}^\rho - VE_{\rho n}^{(1)}\delta_{nm} = 0, \tag{2.73}$$

where

$$S_{nm}^\rho := \int \mathrm{d}\boldsymbol{r}\, w_{\rho m}^\dagger(\boldsymbol{r})\hbar v_{\mathrm{F}}\boldsymbol{\sigma}^\rho \cdot \boldsymbol{\mathcal{A}}(\boldsymbol{r})w_{\rho n}(\boldsymbol{r}) \tag{2.74}$$

is a matrix element of the perturbation Hamiltonian. If $n = m$ this gives $E_{\rho n}^{(1)} = S_{nn}^\rho/V$, and if $n \neq m$ and $E_{\rho n} \neq E_{\rho m}$ it gives $a_{nm}^\rho = S_{nm}^\rho/(E_{\rho n} - E_{\rho m})$. Since I am using nondegenerate perturbation theory we get no requirements if $n \neq m$ and $E_{\rho n} = E_{\rho m}$, but let's assume that $[f(E_{\rho n}) - f(E_{\rho m})]a_{nm}^\rho = f'(E_{\rho n})S_{nm}^\rho$ in this case, so that the end result matches that of Liang *et al.* [93], even in the case of degeneracies. Plugging these into the expansion (2.72) and the expansion into Eq. (2.70) then gives the current

$$\boldsymbol{j}(\boldsymbol{r}) = -\frac{1}{V} \int \mathrm{d}\boldsymbol{r}'\, K(\boldsymbol{r}, \boldsymbol{r}')\boldsymbol{\mathcal{A}}(\boldsymbol{r}'), \tag{2.75}$$

which is the static limit of Eq. (1.66). Here the $\mu, \nu \in \{x, y\}$ component of the kernel is

$$K_{\mu\nu}(\boldsymbol{r}, \boldsymbol{r}') := \frac{(\hbar v_{\mathrm{F}})^2}{V} \sum_{\rho nm} \mathcal{F}_{nm}^\rho w_{\rho n}^\dagger(\boldsymbol{r})\sigma_\mu^\rho w_{\rho m}(\boldsymbol{r})w_{\rho m}^\dagger(\boldsymbol{r}')\sigma_\nu^\rho w_{\rho n}(\boldsymbol{r}') - \mathrm{reg} \tag{2.76}$$

with

$$\mathcal{F}_{nm}^\rho := \begin{cases} f'(E_{\rho n}), & \text{if} \quad E_n = E_m, \\ \frac{f(E_{\rho n}) - f(E_{\rho m})}{E_{\rho n} - E_{\rho m}}, & \text{if} \quad E_n \neq E_m. \end{cases} \tag{2.77}$$

Let us calculate only the position average $\langle \boldsymbol{j} \rangle$ of the current in response to a constant vector potential $\boldsymbol{\mathcal{A}} = \langle \boldsymbol{\mathcal{A}} \rangle$, which already yields the London equations (1.69) and (1.70) and determines the BKT transition temperature (1.65), at least in the simplest theories [93, 126]. In this case we have

$$\langle \boldsymbol{j} \rangle = -D^{\mathrm{s}}\langle \boldsymbol{\mathcal{A}} \rangle \quad \text{with} \quad D^{\mathrm{s}} = \langle K \rangle \tag{2.78}$$

being the superfluid weight. Here $\langle K \rangle = \mathrm{avg}_{\boldsymbol{r}\boldsymbol{r}'} K(\boldsymbol{r}, \boldsymbol{r}')$. In Publications I and III we only calculate the superfluid weight D^{s} and never the full kernel K.

2.3.3   Equations in Fourier space

Since we are interested in *periodic* strain, it is easiest to solve the equations in Fourier space. Taking the eigenfunctions $w_{\rho b \boldsymbol{k}'} : \mathbb{R}^2/L_{\mathrm{BK}} \to \mathbb{C}^2$ to be periodic in translations of the large Born–von Kármán lattice L_{BK}, the pseudo vector potential

$\boldsymbol{A} : \mathbb{R}^2/SL \to \mathbb{R}^2$ to be periodic in translations of the arbitrary superlattice $SL = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{t}_1, \boldsymbol{t}_2\}$, and the order parameter $\Delta : \mathbb{R}^2/SL \to \mathbb{C}_{2\times2}$ to be periodic in translations of the same superlattice SL, we may write them as the Fourier series

$$w_{\rho b\boldsymbol{k}'}(\boldsymbol{r}) = \sum_{\boldsymbol{k}\in L_{\mathrm{BK}}^*} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\tilde{w}_{\rho b\boldsymbol{k}'}(\boldsymbol{k}) \tag{2.79}$$

$$\boldsymbol{A}(\boldsymbol{r}) = \sum_{\boldsymbol{G}\in SL_S^*} \mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}}\tilde{\boldsymbol{A}}(\boldsymbol{G}) \tag{2.80}$$

$$\Delta(\boldsymbol{r}) = \sum_{\boldsymbol{G}\in SL_S^*} \mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}}\tilde{\Delta}(\boldsymbol{G}) \tag{2.81}$$

through Eq. (1.79). Substituting these to the BdG equation (2.45) and dividing the $\boldsymbol{k}$ sum as in Eq. (2.23) then yields the BdG equation in Fourier space,

$$\sum_{\boldsymbol{G}'\in SL_S^*} \tilde{\mathcal{H}}_{\mathrm{BdG},\boldsymbol{G}\boldsymbol{G}'}^{\rho}(\boldsymbol{k})\tilde{w}_{\rho b\boldsymbol{k}'}(\boldsymbol{k}+\boldsymbol{G}') = E_{\rho b\boldsymbol{k}'}\tilde{w}_{\rho b\boldsymbol{k}'}(\boldsymbol{k}+\boldsymbol{G}). \tag{2.82}$$

Here the Nambu-matrix

$$\tilde{\mathcal{H}}_{\mathrm{BdG},\boldsymbol{G}\boldsymbol{G}'}^{\rho}(\boldsymbol{k}) := \begin{pmatrix} \tilde{\mathcal{H}}_{\boldsymbol{G}\boldsymbol{G}'}^{\rho}(\boldsymbol{k}) & \tilde{\Delta}(\boldsymbol{G}-\boldsymbol{G}') \\ \tilde{\Delta}^*(\boldsymbol{G}'-\boldsymbol{G}) & -\tilde{\mathcal{H}}_{\boldsymbol{G}\boldsymbol{G}'}^{\rho}(\boldsymbol{k}) \end{pmatrix} \tag{2.83}$$

is the Fourier-space version of the BdG Hamiltonian matrix element (2.43) with the noninteracting Hamiltonian matrix element $\tilde{\mathcal{H}}_{\boldsymbol{G}\boldsymbol{G}'}^{\rho}(\boldsymbol{k})$ being as in Eq. (2.25).

Defining the $\boldsymbol{G}$-space matrix and vector respectively as

$$\tilde{\underline{\mathcal{H}}}_{\mathrm{BdG}}^{\rho}(\boldsymbol{k}) := \left(\tilde{\mathcal{H}}_{\mathrm{BdG},\boldsymbol{G}\boldsymbol{G}'}^{\rho}(\boldsymbol{k})\right)_{\boldsymbol{G}\boldsymbol{G}'\in SL_S^*}, \quad \underline{\tilde{w}}_{\rho b\boldsymbol{k}'}(\boldsymbol{k}) := \left(\tilde{w}_{\rho b\boldsymbol{k}'}(\boldsymbol{k}+\boldsymbol{G})\right)_{\boldsymbol{G}\in SL_S^*} \tag{2.84}$$

the BdG equation (2.82) becomes a matrix eigenvalue equation

$$\tilde{\underline{\mathcal{H}}}_{\mathrm{BdG}}^{\rho}(\boldsymbol{k})\underline{\tilde{w}}_{\rho b\boldsymbol{k}'}(\boldsymbol{k}) = E_{\rho b\boldsymbol{k}'}\underline{\tilde{w}}_{\rho b\boldsymbol{k}'}(\boldsymbol{k}). \tag{2.85}$$

Because $\tilde{\underline{\mathcal{H}}}_{\mathrm{BdG}}^{\rho}(\boldsymbol{k})$ depends only on $\boldsymbol{k}$, and b already enumerates all the solutions for this $\boldsymbol{k}$, we must have $\boldsymbol{k}' = \boldsymbol{k}$. And now that $\boldsymbol{k} \in L_{\mathrm{BK}}^*/SL_S^*$ belongs to the SBZ, also the original ansatz $\boldsymbol{k}' \in L_{\mathrm{BK}}^*$ gets restricted to the SBZ L_{BK}^*/SL_S^*. Thus we may set

$$\tilde{w}_{\rho b\boldsymbol{k}'}(\boldsymbol{k}+\boldsymbol{G}) = \delta_{\boldsymbol{k}\boldsymbol{k}'}\tilde{w}_{\rho b\boldsymbol{k}}(\boldsymbol{k}) \tag{2.86}$$

so that the BdG equation becomes

$$\tilde{\underline{\mathcal{H}}}_{\mathrm{BdG}}^{\rho}(\boldsymbol{k})\underline{\tilde{w}}_{\rho b\boldsymbol{k}}(\boldsymbol{k}) = E_{\rho b\boldsymbol{k}}\underline{\tilde{w}}_{\rho b\boldsymbol{k}}(\boldsymbol{k}). \tag{2.87}$$

This is clearly a separate problem for each $\boldsymbol{k} \in L_{\mathrm{BK}}^*/SL_S^*$, and for each $\boldsymbol{k}$ there are exactly $2 \times 2 \times |SL_S^*|$ (the matrix dimension) solutions labelled by the band index b.

Property (2.86) together with the division (2.23) can be used to write the Fourier series (2.79) of w in the Bloch form

$$w_{\rho b\boldsymbol{k}}(\boldsymbol{r}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \sum_{\boldsymbol{G}\in SL_S^*} \mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} \tilde{w}_{\rho b\boldsymbol{k}}(\boldsymbol{k}+\boldsymbol{G}), \qquad (2.88)$$

where the function multiplying $\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}$ is periodic in translations of the superlattice SL. Using the Fourier series (2.80) and (2.81) and the Bloch-form Fourier series (2.88) we may write the self-consistency equation (2.59) in the Fourier space as

$$\tilde{\Delta}_\alpha(\boldsymbol{G}) = -\frac{\lambda}{(2\pi)^2} \sum_{\rho b_+} \sum_{\boldsymbol{G}'\in SL_S^*} \int_{\mathbb{R}^2/SL_S^*} \mathrm{d}\boldsymbol{k}$$

$$\tilde{u}_{\rho b_+\boldsymbol{k},\alpha}(\boldsymbol{k}+\boldsymbol{G}')\tilde{v}_{\rho b_+\boldsymbol{k},\alpha}^*(\boldsymbol{k}+\boldsymbol{G}'-\boldsymbol{G})\tanh\left(\frac{E_{\rho b_+\boldsymbol{k}}}{2k_\mathrm{B}T}\right). \quad (2.89)$$

In Eq. (2.89) I also approximated the $\boldsymbol{k}$ sum as an integral

$$\sum_{\boldsymbol{k}\in L_\mathrm{BK}^*/SL_S^*} \approx \frac{V}{(2\pi)^2} \int_{\mathbb{R}^2/SL_S^*} \mathrm{d}\boldsymbol{k} \qquad (2.90)$$

by assuming the Born–von Kármán cell to be large, *i.e.* L_BK to be sparse (compared to L) or L_BK^* to be dense (compared to L^*).

The linearized supercurrent (2.75), on the other hand, becomes

$$\tilde{\boldsymbol{j}}(\boldsymbol{G}) = \sum_{\boldsymbol{G}'\in SL_S^*} \tilde{K}(\boldsymbol{G},-\boldsymbol{G}')\tilde{\boldsymbol{\mathcal{A}}}(\boldsymbol{G}'), \qquad (2.91)$$

where the Fourier components of the kernel are

$$\tilde{K}_{\mu\nu}(\boldsymbol{G},-\boldsymbol{G}') = \left(\frac{\hbar v_\mathrm{F}}{2\pi}\right)^2 \sum_{\rho bb'} \sum_{\boldsymbol{F}\boldsymbol{F}'\in SL_S^*} \int_{\mathbb{R}^2/SL_S^*} \mathrm{d}\boldsymbol{k}\,\mathcal{F}_{b\boldsymbol{k},b'\boldsymbol{k}}^\rho \times$$

$$\times \tilde{w}_{\rho b\boldsymbol{k}}^\dagger(\boldsymbol{k}+\boldsymbol{F}-\boldsymbol{G})\sigma_\mu^\rho \tilde{w}_{\rho b'\boldsymbol{k}}(\boldsymbol{k}+\boldsymbol{F})\tilde{w}_{\rho b'\boldsymbol{k}}^\dagger(\boldsymbol{k}+\boldsymbol{F}'-\boldsymbol{G}')\sigma_\nu^\rho \tilde{w}_{\rho b\boldsymbol{k}}(\boldsymbol{k}+\boldsymbol{F}') - \mathrm{reg.} \tag{2.92}$$

The superfluid weight is then

$$D_{\mu\nu}^\mathrm{s} = \tilde{K}_{\mu\nu}(0,0) = \left(\frac{\hbar v_\mathrm{F}}{2\pi}\right)^2 \sum_{\rho bb'} \int_{\mathbb{R}^2/SL_S^*} \mathrm{d}\boldsymbol{k}\,\mathcal{F}_{b\boldsymbol{k},b'\boldsymbol{k}}^\rho \times$$

$$\times \left(\underline{\tilde{w}}_{\rho b\boldsymbol{k}}^\dagger\sigma_\mu^\rho\underline{\tilde{w}}_{\rho b'\boldsymbol{k}}\underline{\tilde{w}}_{\rho b'\boldsymbol{k}}^\dagger\sigma_\nu^\rho\underline{\tilde{w}}_{\rho b\boldsymbol{k}} - \underline{\tilde{w}}_{\rho b\boldsymbol{k}}^\dagger\tau_z\sigma_\mu^\rho\underline{\tilde{w}}_{\rho b'\boldsymbol{k}}\underline{\tilde{w}}_{\rho b'\boldsymbol{k}}^\dagger\tau_z\sigma_\nu^\rho\underline{\tilde{w}}_{\rho b\boldsymbol{k}}\right), \qquad (2.93)$$

where I dropped the $\boldsymbol{k}$ argument from $\underline{\tilde{w}}_{\rho b\boldsymbol{k}}^\dagger(\boldsymbol{k})$ and also wrote explicitly the regularization term (the last term) according to the work of Liang *et al.* [93]. In the numerical calculations we then solve the Fourier-space equations (2.87), (2.89), and (2.93). The only differences between the first, the *diamagnetic current* [93] term, and the second, the *paramagnetic current* [93] term, are the Pauli-z matrices τ_z acting in the Nambu space.

2.4　Electronic properties of the superconducting state

In this section I summarize the results of Publication I regarding superconductivity. Some figures are taken directly from the publication, while others are new so that the introduction to the calculated results is as pedagogical as possible.

2.4.1　Order parameter and mean-field critical temperature

The main ingredient in determining the properties of the superconducting state is solving the self-consistency equation (2.59) [numerically the Fourier-space version (2.89)]. We solve this by the fixed-point iteration method, where one starts from an initial guess of Δ, solves the eigenenergies E and eigenvectors w from the BdG equation (2.87) with that Δ, uses these new E and w to calculate a new Δ from the self-consistency equation (2.89), and continues iterating this until convergence. But this needs an initial guess for Δ: we chose to always start from a homogeneous $\Delta_A(\boldsymbol{r}) = \Delta_B(\boldsymbol{r}) = \text{const}$. In addition to fixing the overall phase to zero, this choice also fixes the relative phase between Δ_A and Δ_B to zero, but as we showed in the Supplementary Material of Publication I, this choice leads to the lowest-energy solution under the fixed-point iteration.

　　Solving the self-consistency equation shows how the behavior of Δ is determined by the density of states. First of all, as shown in Fig. 2.7 and when compared to Figs. 2.4c and 2.4d, $\Delta_{A/B}$ is localized to minima/maxima[4] of $\boldsymbol{B}$, and is thus also periodic[5] with the period d. This is a main difference to the results of Sec. 1.2.3, where the simpler form of the interaction yields a position-independent Δ. But the position dependence of Δ raises a problem: if one wants to get simple relations between observables and Δ, such as Eqs. (1.36) and (1.51), it is not clear which number should be calculated out of $\Delta(\boldsymbol{r})$. But as we showed in Publication I, a number that well represents the strength of the superconducting state is $\max \Delta$ (maximum over $\boldsymbol{r}$), which is independent of sublattice for the chosen potentials. What I mean by this is that in the flat-band regime (large B or λ) one gets approximately linear relations between observables (such as T_{c}) and $\max \Delta$.

　　We calculate most of the results at optimal doping $\mu = \mu_{\mathrm{opt}}$, which is the position of the DOS peak (Fig. 2.6) and yields the highest Δ for small λ. Calculating Δ as a function of the interaction strength λ (Fig. 2.8a) at $T = 0$ shows how the relation becomes linear with high enough B or λ. This regime, where $\max \Delta$ is much larger than the flat-band bandwidth, we call the *flat-band regime*. The opposite regime with nonlinear dependence we call the *dispersive regime* due to the finite bandwidth contributing to Δ. In Publication I we then calculated the ratio $\zeta := \max \Delta / (-B\lambda)$ which shows that, indeed, in the flat-band regime ζ goes approximately to a constant $\zeta \approx 0.15$, depending slightly on the potential. In the

---

[4]This kind of sublattice polarization ($\Delta_A \neq \Delta_B$) was not observed by Kauppila *et al.* [67] due to a summation over sublattices in that work.

[5]When deriving the BCS equations we assumed that Δ is periodic in translations of the same lattice SL as $\boldsymbol{A}$ is. But this still allows the periodicity of Δ to be smaller than that of $\boldsymbol{A}$. However, the assumption rules out the case where Δ would have a longer period than $\boldsymbol{A}$, which is also a perfectly physical possibility.
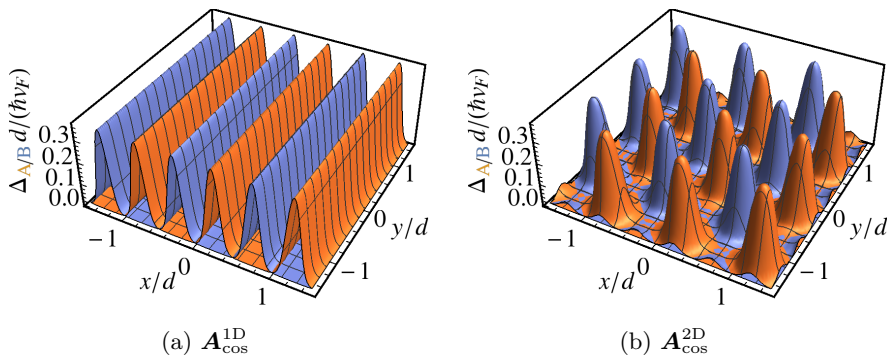
Figure 2.7: Typical profiles of the self-consistent superconducting order parameter $\Delta_{A/B}$ (A orange, B blue), which can be seen to be peaked at the minima/maxima of $\boldsymbol{B}$ when compared to Fig 2.4. Reprinted with permission from Ref. [116]. Copyright 2020 IOP Publishing.

flat-band regime we thus have the linear relation

$$\max \Delta = -\zeta B \lambda \quad \text{at} \quad \mu = \mu_{\text{opt}}, T = 0. \tag{2.94}$$

This is analogous to the exact-flat-band result (1.50) with a homogeneous Δ, which should be multiplied by n in the case of n flat bands. Now it is the amplitude B [Eq. (2.33)] of the pseudomagnetic field $\boldsymbol{B}$ that controls the flatness and the number n of approximate flat bands in the area $\Omega_{\text{FB}} = 1/d^2$ (in the reduced zone scheme). In the case of the 1D cosine potential this simply reproduces the result of Kauppila *et al.* [67], apart from the dispersive regime where the choice of optimal doping, instead of zero doping, slightly enhances Δ. In this regard the role of Publication I and Fig. 4 therein is to show that this result remains to hold also more generally: when changing the shape of the potential through $\boldsymbol{A}_c^{\text{1D}}$ and when going to 2D potentials through $\boldsymbol{A}_{\text{cos}}^{\text{2D}}$.

Another new result is the effect of doping, which I show in Fig. 2.8b in the case of the 1D cosine potential. For all the potentials see Fig. 9 in Publication I. Qualitatively, in the flat-band regime Δ is maximized at $\mu = 0$ and increasing $|\mu|$ has the effect of suppressing superconductivity, because the energy scale of Δ is larger than the bandwidth and thus Δ only "sees" one large, smeared DOS peak located at zero energy (Fig. 2.6). The "critical"[6] chemical potential where $\max \Delta$ is suppressed to $\max \Delta(\mu = 0)/2$ is

$$\mu_{50\%} = \eta \max \Delta(\mu = 0) \quad \text{at} \quad T = 0, \tag{2.95}$$

where in the flat-band regime $\eta \approx 0.7$ for all the potentials except $\boldsymbol{A}_c^{\text{1D}}$ with $c \gtrsim 10$. This is analogous to the exact-flat-band result (1.52) with a homogeneous Δ, which should be multiplied by $\sqrt{3}/2$ in the case of the half-level $\mu_{50\%}$. In the

---

[6]Due to slow convergence when Δ is close to zero it is very time-consuming to calculate the actual critical chemical potential where Δ vanishes. Thus we calculate the half-level $\mu_{50\%}$ instead, which is much faster to obtain.
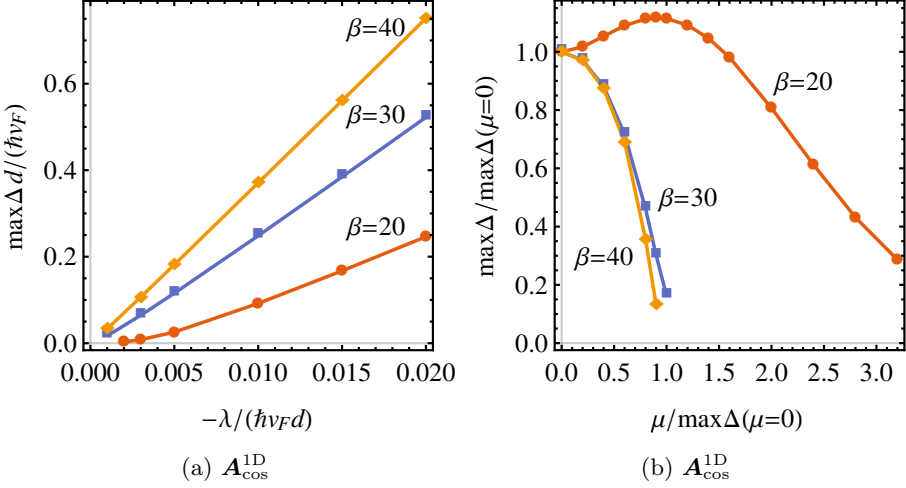
Figure 2.8: Maximum of the superconducting order parameter Δ for the 1D cosine potential $\boldsymbol{A}_{\cos}^{\mathrm{1D}}$ at $T = 0$ as a function of (a) the interaction strength λ and (b) the chemical potential μ. (a) $\max \Delta$ at optimal doping $\mu = \mu_{\mathrm{opt}}$ is approximately linear in λ in the flat-band regime, where $\max \Delta$ is much larger than the flat-band bandwidth (compare to the DOS in Fig. 2.6). (b) Doping dependence of Δ with $\lambda/(\hbar v_{\mathrm{F}} d) = -0.005$ at $T = 0$. In the flat-band regime doping monotonically suppresses superconductivity with the critical value $\mu_{\mathrm{c}} \approx \max \Delta(\mu = 0)$, whereas in the dispersive regime the optimal doping is located at a finite chemical potential at the DOS peak. Δ is symmetric with respect to the sign of μ.

dispersive regime, on the other hand, where Δ "sees" the individual DOS peaks, Δ is maximized when doped to the DOS peak (the optimal doping level).

Calculating Δ as a function of temperature T yields a similar approximate linear relation (Fig. 10 in Publication I)

$$k_{\mathrm{B}} T_{50\,\%} = \xi \max \Delta(T = 0) \quad \text{at} \quad \mu = \mu_{\mathrm{opt}} \tag{2.96}$$

for the half-level "critical" temperature in the flat-band regime. Here $\xi \approx 0.35$, which again holds for all the potentials except $\boldsymbol{A}_{\cos}^{\mathrm{1D}}$ with $c \gtrsim 10$. In the case of the 1D cosine potential this again reproduces the result of Kauppila $et\ al.$ [67], but Publication I shows that it continues to hold also for more generic potentials if $\mathrm{avg}\,\Delta$ of Kauppila $et\ al.$ is replaced by $\max \Delta$ (and the prefactor is modified accordingly). Equation (2.96) corresponds to the exact-flat-band result (1.51) with a homogeneous Δ.

2.4.2   Superfluid weight and BKT transition temperature

When Δ is known we can determine the superfluid weight D^{s} from Eq. (2.93). Shown in Fig. 2.9a for $\boldsymbol{A}_{\cos}^{\mathrm{1D}}$ at $\mu = \mu_{\mathrm{opt}}$ and $T = 0$, it shows how similar to Δ the dependence is, only with a different scale. From Fig. 11 of Publication I we can
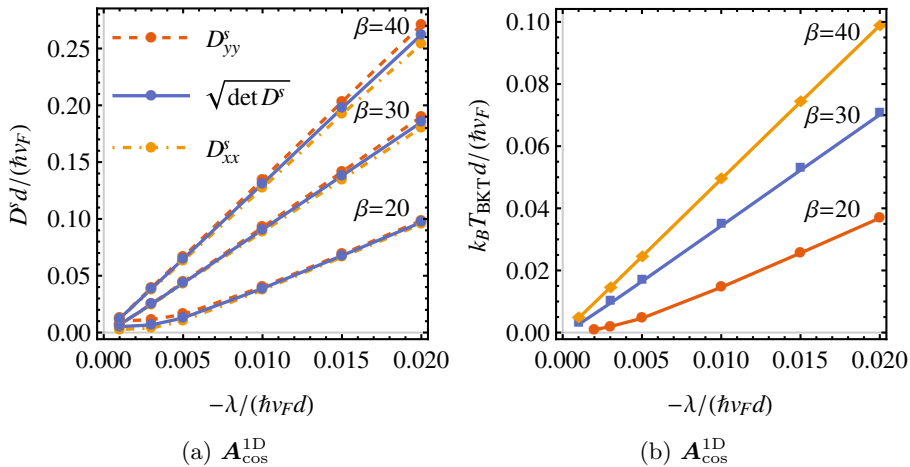
Figure 2.9: (a) The superfluid weight D^{s} and (b) the Berezinskii–Kosterlitz–Thouless transition temperature T_{BKT} at optimal doping $\mu = \mu_{\mathrm{opt}}$ as a function of the interaction strength λ in the case of the 1D cosine potential $\boldsymbol{A}_{\mathrm{cos}}^{\mathrm{1D}}$. (a) $D^{\mathrm{s}}(\lambda)$ at $T = 0$ shows very similar behavior to $\Delta(\lambda)$ in Fig. 2.8a. The superfluid weight is almost isotropic even though the potential $\boldsymbol{A}_{\mathrm{cos}}^{\mathrm{1D}}$ is highly anisotropic. Reprinted with permission from Ref. [116]. Copyright 2020 IOP Publishing. (b) $T_{\mathrm{BKT}}(\lambda)$ also shows very similar behavior both to $D^{\mathrm{s}}(\lambda)$ and $\Delta(\lambda)$.

read that in the flat-band regime we then get an approximately linear relation

$$\sqrt{\det D^{\mathrm{s}}} = \chi \max \Delta \quad \text{at} \quad \mu = \mu_{\mathrm{opt}}, T = 0, \tag{2.97}$$

where $\chi \approx 0.15 \ldots 0.4$ depending on the potential. Figure 2.9a also shows how D^{s} is almost isotropic for the 1D potentials: there is only a very slight difference between D^{s}_{xx} and D^{s}_{yy} at large or small B and λ. On the other hand the 2D potential has an isotropic D^{s} with $D^{\mathrm{s}}_{xx} = D^{\mathrm{s}}_{yy}$. For all the potentials $D^{\mathrm{s}}_{xy} = 0 = D^{\mathrm{s}}_{yx}$. Note that for this discussion the choice of the coordinate system is important; we are all the time using the same coordinate system where the graphene lattice vectors are those in Eq. (2.1).

In the case of the 1D cosine potential, Kauppila *et al.* [67] also found a linear relation[7] $D^{\mathrm{s}}_{\mu\mu} = 2a_\mu \max \Delta$ in the flat-band regime by directly calculating the supercurrent from Eq. (2.70) for small $\boldsymbol{A}$ (denoted by $\boldsymbol{k}_{\mathrm{s}}$ in that work). Their finding $a_x \approx 0.17$ and $a_y \approx 0.19$ results in $2\sqrt{a_x a_y} \approx 0.36$, which is in agreement with $\chi \approx 0.35$ found in Publication I, even though they used the different renormalization procedure of subtracting the normal-state supercurrent. Regarding the small anisotropy, the components of χ are $\chi_x \approx 0.34$ and $\chi_y \approx 0.36$ (not explicitly shown in Publication I except of Fig. 11a), again in agreement with Kauppila *et al.* [67].

---

[7]I inserted the factor of 2 because Eq. (2.70) has a sum over valleys whereas the one in Ref. [67] does not have it.

Another new result is the Berezinskii–Kosterlitz–Thouless (BKT) transition temperature T_{BKT}, which was not included in the work of Kauppila *et al.* [67]. Calculating D^{s} as a function of temperature allows obtaining T_{BKT} through Eq. (1.65). As shown in Fig. 2.9b, the behavior as a function of λ is extremely similar to that of D^{s} or Δ, with the only thing changing (in the flat-band regime) being the scale. Figure 12 of Publication I then confirms that in the flat-band regime we again have a linear relation

$$k_{\mathrm{B}}T_{\mathrm{BKT}} = \kappa \max \Delta(T=0) \quad \text{at} \quad \mu = \mu_{\mathrm{opt}} \tag{2.98}$$

with $\kappa \approx 0.05 \ldots 0.15$ depending on the potential.

To obtain realistic values for T_{BKT}, let us use the experimental strain profiles of Jiang *et al.* [71]. As one example, they manufactured a strain field with a period of $d = 14\,\mathrm{nm}$ and an accompanying pseudomagnetic field amplitude $\frac{\hbar}{e}B = 100\,\mathrm{T}$, which can be modelled by the 1D cosine potential with $\beta = 5$, assuming graphene to be suitably oriented with respect to the strain field (see the discussion in the end of Sec. 2.1). The experimental value of the interaction strength λ is not known. But because according to Publications II and III $\lambda \approx -1\,\mathrm{eV}a^2 \approx -6\,\mathrm{eV\mathring{A}}^2 \approx -0.007\hbar v_{\mathrm{F}}d$ yields the experimental [9, 127] value of $T_{\mathrm{BKT}} \approx 1\,\mathrm{K}$ in the similar flat-band system of TBG, a similar value is probably the best first guess. According to Fig. 2.9b T_{BKT} is then most likely very close to zero for such a small β and λ. If, however, one is able to increase the strain amplitude by a factor of 4 to $\beta = 20$, one would already obtain a relatively high transition temperature of $T_{\mathrm{BKT}} \approx 0.007\hbar v_{\mathrm{F}}/(dk_{\mathrm{B}}) \approx 4\,\mathrm{K}$. This, however, is not yet perfectly in the flat-band regime, so the simple linear formulas (2.94)–(2.98) cannot directly be applied to see how this number could be increased. Increasing β further or decreasing d moves us to the flat-band regime, where the linear formulas can be applied, which together with Eq. (2.33) tell us that for $\boldsymbol{A}_{\mathrm{cos}}^{\mathrm{1D}}$

$$k_{\mathrm{B}}T_{\mathrm{BKT}} = -\kappa\zeta 2\pi\beta\lambda/d^2 \tag{2.99}$$

is linearly proportional to β and inversely proportional to d^2. But at the same time one has to remember the limit (2.16) of elasticity theory, which for the in-plane strain field $\boldsymbol{u}_{\mathrm{cos}}^{\mathrm{1D}}$ is equivalent to the constraints

$$\frac{\beta}{\beta_{\mathrm{G}}} \ll \frac{d}{a_0} \quad \text{and} \quad \frac{d}{a_0} \gg 1. \tag{2.100}$$

This, on the other hand, is equivalent to constraining the strain as $|u_{ij}| \ll 1$. Thus when trying to increase T_{BKT} one should maximize β (or more generally B) and minimize d such that the constraint (2.100) still holds.

The above estimate corresponds to a displacement amplitude of $\beta a_0/(\beta_{\mathrm{G}}d) \approx 5\,\mathrm{\mathring{A}}$ if assuming the in-plane strain field (2.31) and a Grüneisen parameter of $\beta_{\mathrm{G}} = 2$. This is very large compared to the $0.1\,\mathrm{\mathring{A}}$ displacement amplitude found in TBG after structural relaxation [109], so it might be easier to stay in the out-of-plane strains. It is left for further studies to find out what kind of pseudo vector potentials realistic out-of-plane strain fields produce, whether they can be approximated by the harmonic versions $\boldsymbol{A}_{\mathrm{cos}}^{\mathrm{1D/2D}}$, and what is the optimal strain field still experimentally producible.

To provide even more measurable quantities, in Publication I we also calculated the behavior of both the total (Fig. 7) and local (Fig. 8) density of states in the superconducting state. The main conclusion is that in the flat-band regime there are peculiar multiple-peak structures in the energy dependence, as also predicted by Kauppila *et al.* [67] in the case of the 1D cosine potential, the observation of which would help confirming the used model.

2.5 Conclusions and outlook

In Publication I we applied BCS theory to study spin-singlet, *s*-wave superconductivity in periodically strained graphene (PSG) in the flat-band regime. In fact the numerical results hold also in the dispersive regime, but the linear formulas (2.94)–(2.98) work only in the flat-band regime. Within the low-energy continuum approximation the strain is modelled by a pseudo vector potential $\boldsymbol{A}$, which works like a normal (magnetic) vector potential but conserves time-reversal symmetry, thus maintaining the possibility of spin-singlet, *s*-wave superconductivity. Kauppila *et al.* [67] studied this system already in the case of a one-dimensional harmonic potential, and showed that periodic strain strongly enhances superconductivity, with the critical temperature T_{c} being linear in the attractive interaction strength λ. In Publication I we generalized the theory to include arbitrary periodic potentials. By formulating the theory both in the mixed and reduced zone schemes in the case of 1D potentials, we could bridge the gap between PSG (as formulated by Kauppila *et al.* [67]) and twisted bilayer graphene (TBG; as formulated *e.g.* in Publications II and III), which at first might seem very different systems. We furthermore extended the superconductivity theory to include also the calculation of the superfluid weight D^{s} and the Berezinskii–Kosterlitz–Thouless (BKT) transition temperature T_{BKT}.

Regarding Publication I, it is intriguing to wait when (if) the first experiments reporting superconductivity (and correlated insulating phases) in PSG appear. The striking similarity to TBG suggests that finding these phases is highly probable, if a strain with large enough amplitude and small enough period can be manufactured. There are numerous possibilities for manufacturing it, and the experiments of Jiang *et al.* [71] suggest that we are actually already close to achieving a measurable T_{BKT}. Besides, the critical temperature of PSG is (in principle) way more tunable than that of TBG through controlling the shape, period, lattice, and amplitude of the strain, so a recipe for a higher T_{c} is already available there.

In Publication I we included explicitly only the effect of attractive interaction emerging from the electron–phonon interaction. The next natural extension would be to study the effects of the repulsive Coulomb interaction, to see how large a repulsive interaction would destroy superconductivity. Within the local (Hubbard) interaction model, we would simply add another interaction Hamiltonian with the strength $u > 0$, in addition to the attractive interaction with $\lambda < 0$. But an important difference is that whereas the attractive interaction has an energy cutoff ϵ_{c} resulting from the electron–phonon interaction, the repulsive Coulomb interaction does not have a cutoff. In the case of a homogeneous Δ, the contribution from the higher energies can, however, be incorporated into a

Coulomb pseudopotential [128, 129]

$$u^* \coloneqq \frac{u}{1 + u\alpha} > 0, \tag{2.101}$$

where $\alpha > 0$ depends on the noninteracting dispersion above the cutoff ϵ_c. The resulting total interaction strength, limited to energies below ϵ_c, is then

$$\lambda_{\mathrm{eff}} \coloneqq \lambda + u^*, \tag{2.102}$$

so that $\lambda_{\mathrm{eff}} < 0$ corresponds to a superconducting state and $\lambda_{\mathrm{eff}} > 0$ to an insulating antiferromagnetic state [129]. Especially we see that due to the renormalization the Coulomb interaction is not as effective in reducing the attractive interaction as one would think by comparing the bare values λ and u, because we always have $u^* \leq u$. In the sense of Eq. (2.102) the Coulomb repulsion is already incorporated in the results of Publication I, if one simply interprets the used values of λ as the effective interaction λ_{eff}.

In the Supplementary Material of Publication II we explicitly calculated the renormalization parameter to be $\alpha \approx 3.3\,\mathrm{eV}^{-1}\,\mathrm{nm}^{-2}$ for two uncoupled graphene layers, within the simplified assumption of a homogeneous Δ. This result can be applied also to PSG if we divide the result by two, so that $\alpha \approx 1.7\,\mathrm{eV}^{-1}\,\mathrm{nm}^{-2}$. In this case the Coulomb pseudopotential has a maximum possible value of $u^* = 1/\alpha \approx 0.6\,\mathrm{eV}\,\mathrm{nm}^2$, no matter how large the bare interaction u is. This simple result (2.102) is not the whole truth about the Coulomb (modelled by the Hubbard model) interaction, as it assumes a homogeneous Δ, which clearly does not hold in PSG. The work left for future studies is to find out what kind of position-dependent pseudopotential $u^*(\boldsymbol{r})$ the position-dependent $\Delta(\boldsymbol{r})$ would yield, and how Eq. (2.102) should be generalized.

Another interesting aspect left for future studies is the possibility of correlated insulating phases, which are highly probable because they have been observed in the very similar system of twisted bilayer graphene [130]. This similarity is discussed more in Publication I. In Publication I we assumed that superconductivity has its own origin from the electron–phonon interaction, and if there are other phases, they have some other mechanism. Within this assumption one possibility to model the other phases would be through Eq. (2.102) (and its generalization to a position-dependent Δ), which models the competition between a superconductor and an antiferromagnetic insulator through the difference between λ and u^*. Then for example by taking into account screening from the electrodes, λ_{eff} could change sign for changing chemical potential μ, because screening might have a different effect on λ and u^*.

Kauppila *et al.* [67] used a Su–Schrieffer–Heeger-model argument [119] to understand the origin of the flat bands resulting from a 1D strain. One interesting future direction would be to generalize this model to understand the flat bands resulting from a 2D strain. Here one might need some kind of periodic two-dimensional array of domain wall states.

The results of Publication I are very promising when looking from the perspective of the experiments of Jiang *et al.* [71] (Fig. 2.1). Perhaps the largest uncertainty concerns the value of the interaction strength λ (or λ_{eff}), which is

not experimentally known. But because according to Publications II and III $\lambda \approx -6\,\mathrm{eV\AA}^2$ well matches the experimental [9, 127] results in the similar system of TBG, this value works as an educated guess. With this value the reported strain fields of Jiang *et al.* [131] are not yet quite enough to obtain superconductivity, but they are already close: strain with a four times larger amplitude would already yield a relatively large T_{BKT} of 4 K, assuming the resulting pseudo vector potential $\boldsymbol{A}$ is close to a harmonic form and the graphene layer is suitably oriented with respect to the strain field. To obtain more accurate predictions, one would need both the in-plane and out-of-plane strain fields, as well as the graphene orientation, to calculate a more realistic form of $\boldsymbol{A}$. Nevertheless, the realization of a new graphene-only superconductor may already be close. Not only that, superconductivity in PSG is highly tunable by the strain, especially by its amplitude, period, and shape. Moreover, when compared to TBG where the superlattice Brillouin zone is fixed by the rotation angle through the moiré pattern, in PSG one can in principle choose it arbitrarily by choosing the strain field.

Chapter 3

Twisted bilayer graphene

In this chapter I go through Publications II and III regarding twisted bilayer graphene (TBG). First I go through some general aspects of TBG, such as the arising moiré pattern (Sec. 3.1) and the used continuum model (Sec. 3.2). Then in Sec. 3.3 I summarize the normal-state results calculated in Publications II and III and in the literature. In Sec. 3.4 I present the BCS theory of TBG and finally in Sec. 3.5 I summarize the superconducting-state results calculated in Publications II and III and in the vast literature. While studying the same system and using almost the same (Dirac point, DP) model, the difference between the two publications is that in Publication II we calculated results regarding the superconducting order parameter Δ and the mean-field critical temperature T_c, and in Publication III we extended the results to the superfluid weight D^s and the Berezinskii–Kosterlitz–Thouless transition temperature T_{BKT}. For a robustness check in Publication III we employed also another model (renormalized moiré, RM), and studied how the conclusions change when using a more complicated (resonating valence bond, RVB) interaction.

3.1 Moiré superlattice

Stacking two-dimensional lattices with a misorientation angle creates so-called moiré patterns, as demonstrated in Fig. 3.1 with two graphene lattices printed on a normal and a transparency paper. The demonstrated system is called *twisted bilayer graphene* (TBG). The moiré pattern is purely of geometric origin, and in this section I review the geometric properties of the moiré pattern in the case of TBG. I follow mostly the conventions of Lopes dos Santos *et al.* [132].

3.1.1 Commensurate angles

Let us fix one graphene layer to have the lattice primitive vectors

$$\boldsymbol{a}_1 := \frac{a}{2}(1, \sqrt{3}), \quad \boldsymbol{a}_2 := \frac{a}{2}(-1, \sqrt{3}),$$
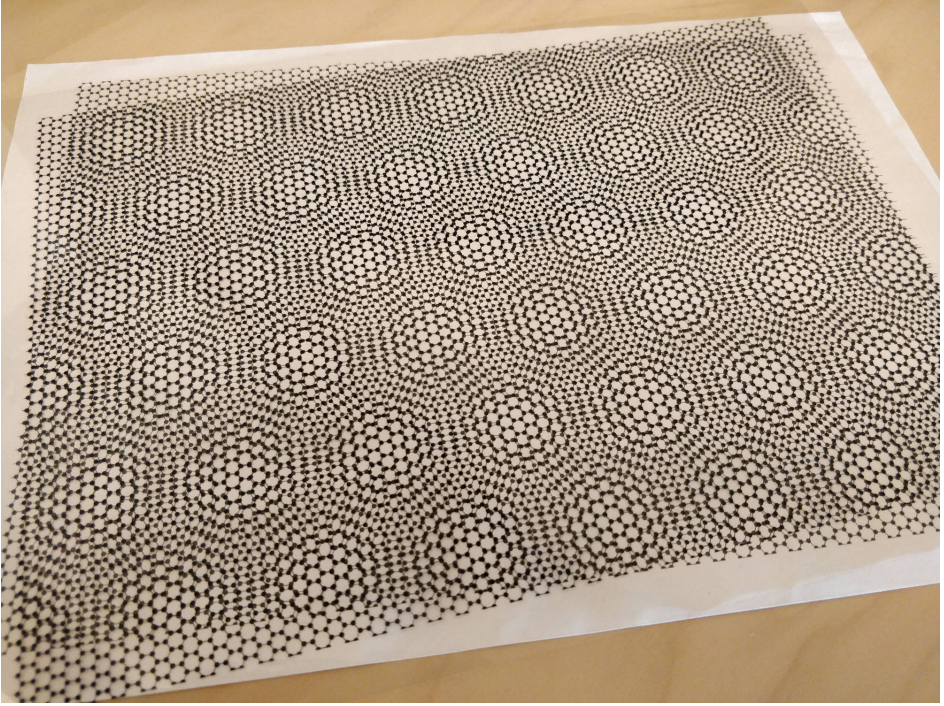
Figure 3.1: Moiré pattern in twisted bilayer graphene, as demonstrated by a honeycomb lattice printed on a normal and a transparency paper.

where $a = 2.46\,\text{Å}$ is the lattice constant. This is the same convention as for PSG in Eq. (2.1). We then add a second layer on top of the first one, rotated by an angle θ. The resulting structure is periodic, *i.e. commensurate*, if and only if more than one lattice sites match. Mathematically, the problem is to find all $m_1, m_2, n_1, n_2 \in \mathbb{Z}$ for which the *Diophantine equation*

$$m_1 \boldsymbol{a}_1 + m_2 \boldsymbol{a}_2 = n_1 R(\theta) \boldsymbol{a}_1 + n_2 R(\theta) \boldsymbol{a}_2 \tag{3.1}$$

has a solution [133]. Lopes dos Santos *et al.* [134] initially identified only a small subset ($r = 1$ in the following) of the solutions/structures by using an ansatz. Later Shallcross *et al.* [135] found all the solutions by properly solving the equation and showing (by a tedious calculation) that they have obtained all the possible structures. In short, their calculation is based on the fact that the rotation matrix $R(\theta)$ in lattice basis has to be rational, and thus the resulting structures are enumerated by a rational number. Later Lopes dos Santos *et al.* [132] also found all the solutions by a simpler calculation utilizing the symmetries of the graphene lattice. The result is that all the commensurate structures are given by two integers $m, r \in \mathbb{Z}$ for which the angle satisfies

$$\cos \theta = \frac{3m^2 + 3mr + r^2/2}{3m^2 + 3mr + r^2}. \tag{3.2}$$

Dividing by r^2, it is obvious that the commensurate angle θ depends only on the ratio $m/r \in \mathbb{Q}$, meaning that all the commensurate angles can be enumerated by a single rational number. The commensurate angles thus form only a countable set, even though it may be dense (near some angles) in the set of all angles.

Above I discuss the commensurate structures, which form only a countable subset of all the structures. What are then the rest of the structures, the *incommensurate* or aperiodic ones, which form an uncountable set in all the possible angles? They are, by definition, not periodic, yet they are clearly ordered. This means they are in fact *quasicrystals*. An interesting special case is $\theta = 30°$, which has even been demonstrated experimentally [136].

3.1.2   Superlattice primitive vectors

The solution of the Diophantine equation (3.1) yields also the primitive vectors of the moiré superlattice SL. This reveals a surprising result: there are two families of commensurate structures, depending on whether 3 divides r or not. The superlattice primitive vectors are [132, 135][1]

$$\begin{pmatrix} \boldsymbol{t}_1 \\ \boldsymbol{t}_2 \end{pmatrix} := \begin{pmatrix} m & m+r \\ -(m+r) & 2m+r \end{pmatrix} \begin{pmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \end{pmatrix}, \quad \text{if} \quad 3 \nmid r, \tag{3.3}$$

$$\begin{pmatrix} \boldsymbol{t}_1 \\ \boldsymbol{t}_2 \end{pmatrix} := \begin{pmatrix} m+r/3 & r/3 \\ -r/3 & m+2r/3 \end{pmatrix} \begin{pmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \end{pmatrix}, \quad \text{if} \quad 3 \mid r. \tag{3.4}$$

The corresponding primitive vectors of the reciprocal lattice SL^* of the superlattice (reciprocal superlattice for short) can be calculated to be [132]

$$\begin{pmatrix} \boldsymbol{G}_1 \\ \boldsymbol{G}_2 \end{pmatrix} := \frac{4\pi}{3d^2} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} \boldsymbol{t}_1 \\ \boldsymbol{t}_2 \end{pmatrix}, \tag{3.5}$$

where

$$d := \|\boldsymbol{t}_1\| = \|\boldsymbol{t}_2\| = \begin{cases} a\sqrt{3m^2 + 3mr + r^2}, & \text{if} \quad 3 \nmid r \\ a\sqrt{3m^2 + mr + r^2/9}, & \text{if} \quad 3 \mid r. \end{cases} \tag{3.6}$$

is the superlattice constant.

What is the difference between the two families of structures? Figure 3.2 shows representative structures of both families. Looking closely, one can see them differing at the C_3 symmetry points of the superlattice, marked with black circles. In one specific convention (AA stacking at $\theta = 0$ and rotation around an AA point), there is always an atom and a hexagon center coinciding at the symmetry points in the first family, $3 \nmid r$, while in the second family, $3 \mid r$, there are hexagon centers coinciding at one of the symmetry points and atoms coinciding at the other point. Mele [137] calls these structures *sublattice-exchange odd* (SE odd) when $3 \nmid r$ and *sublattice-exchange even* (SE even) when $3 \mid r$. More detailed discussion on the differences between the two structures can be found from Ref. [138].

---

[1]When comparing the superlattice primitive vectors from different references, note that there are six of them, out of which two are chosen in each reference independently of each other. Also note that the convention for the primitive vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ may be different.
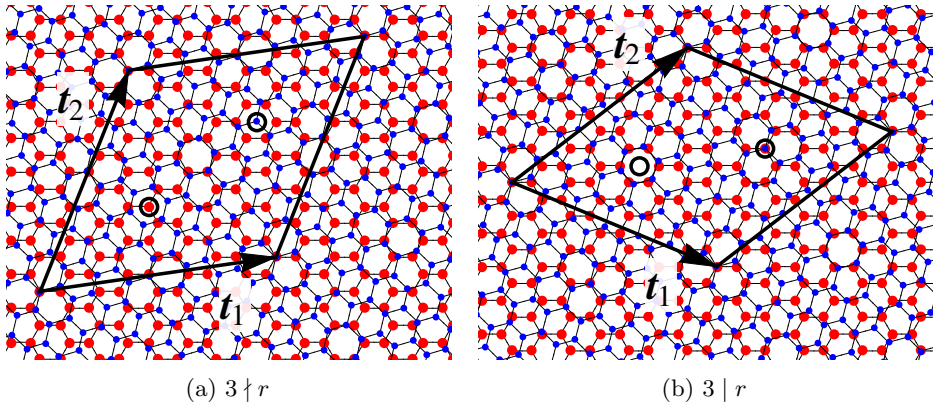
(a) $3 \nmid r$ (b) $3 \mid r$

Figure 3.2: Two families of commensurate structures differ at the C_3 symmetry points (black circles) of the lattice. Within the convention of AA stacking at $\theta = 0$ and rotation around an AA point, (a) the $3 \nmid r$ structure has an atom and a hexagon center coinciding at the symmetry points, whereas (b) the $3 \mid r$ structure has atoms coinciding at one of the symmetry points and hexagon centers coinciding at the other point. One color (red or blue) corresponds to one layer, the black parallelogram denotes the superlattice unit cell, and $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$ are the superlattice primitive vectors. The drawn structures are (a) $(m, r) = (3, 2)$ and (b) $(m, r) = (5, 3)$.

On the other hand, looking at the structures in Fig. 3.2, one might be tempted to argue that the calculated primitive superlattice vectors $\boldsymbol{t}_1, \boldsymbol{t}_2$ are not primitive, as the superlattice seems to have a smaller unit cell. However, this is not true. The primitive vectors were calculated from the condition of *exact* periodicity, while the shorter-period "unit cell" shows only *approximate* periodicity. This is further illustrated in Fig. 3.3a, where it can be clearly seen that, indeed, the only positions with perfect coincidence are those at the corners of the exact superlattice unit cell, while the top-left corner of the approximate superlattice unit cell has coinciding hexagon centers instead of coinciding lattice sites.

This curious behavior can be further examined by calculating the superlattice constant d for a finite number of (m, r) pairs, as shown in Fig. 3.3b. The exact superlattice constant d varies wildly as a function of the commensurate angle θ. Note also how the superlattice constants are symmetric with respect to inversion over $30°$, with the peculiarity that the structure changes its family in the inversion.

On the other hand it can be calculated [135, 139] that the approximate superlattice, which is seen at *any* angle $\theta \in \mathbb{R}$, has the lattice constant

$$d_{\text{approx}} = \frac{a}{2 \sin(\theta/2)} \tag{3.7}$$

as long as $\theta \in [0, 30°]$. This is a strict lower bound for the superlattice constant, as shown in Fig. 3.3b as a black dashed line. Using then Eq. (3.2) for the commensurate angles we can easily see that, in the case of commensurate structures
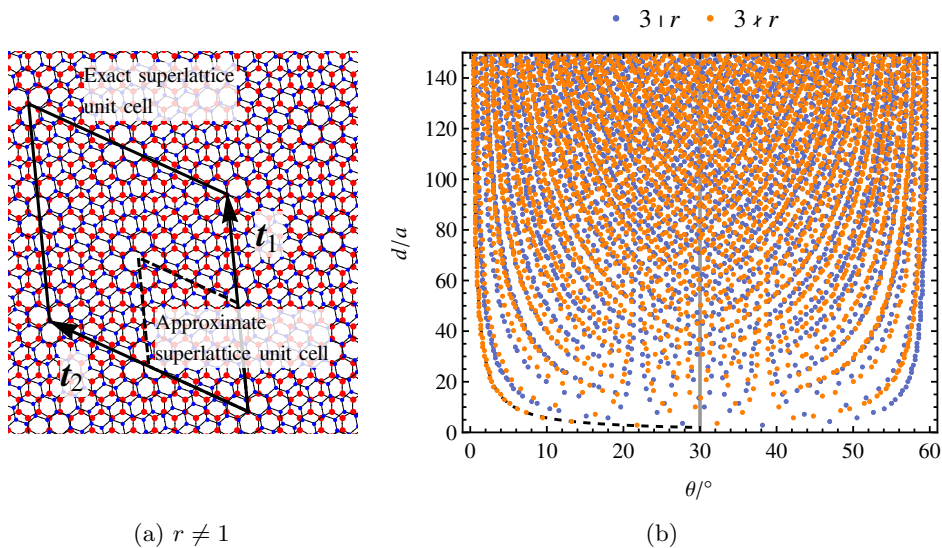
(a) $r \neq 1$ (b)

Figure 3.3: (a) Illustration of the exact (joined parallelogram) and approximate (dashed parallelogram) superlattice unit cells, which match if and only if $r = 1$. The drawn structure is $(m, r) = (5, 2)$. (b) The superlattice constant d [Eq. (3.6)] as a function of the commensurate rotation angle θ [Eq. (3.2)], shown for a finite number of (m, r) pairs, with orange corresponding to a $3 \nmid r$ structure and blue to a $3 \mid r$ structure. Also the lower bound d_{approx} [Eq. (3.7)] given by the approximate superlattice is drawn as a black dashed line, which matches with the $r = 1$ points below $30°$.

and $\theta \in [0, 30°]$, the exact superlattice matches the approximate superlattice if and only if $r = 1$, *i.e.*

$$d = d_{\text{approx}} \quad \Leftrightarrow \quad r = 1. \tag{3.8}$$

From Fig. 3.3b we can identify these structures as the dots coinciding with the dashed line showing the approximate superlattice constant. Correspondingly the reciprocal superlattice constant $G := \|\boldsymbol{G}_1\| = \|\boldsymbol{G}_2\|$ for the approximate structures becomes

$$G_{\text{approx}} = \frac{4\pi}{\sqrt{3}d_{\text{approx}}} = \frac{8\pi}{\sqrt{3}a}\sin(\theta/2), \tag{3.9}$$

which is valid again for any $\theta \in [0, 30°]$.

In Publications II and III we concentrated on the structures with $r = 1$. The used theories themselves are valid for arbitrary commensurate structures, but computationally it is convenient to study only the structures with the smallest possible unit cell. Moreover, in the small-angle limit the $r = 1$ structures are dense in all the structures, as shown in Fig. 3.3b. Any $r \neq 1$ structure between (as a function of θ) the successive $r = 1$ structures is approximately equal to them, and thus should exhibit similar physics. In the small-angle limit the $r \neq 1$ structures should thus not yield new physics, as argued also by Lopes dos Santos *et al.* [132].

Also Bistritzer & MacDonald [140] utilize this fact and model the small-angle structures with a continuous angle $\theta \in \mathbb{R}$.

In experiments (including the transparency-paper experiment in Fig. 3.1) it is always the approximate (moiré) superlattice that one observes. The first experimental evidence of the moiré pattern was provided by scanning tunneling microscopy (STM) [141–146], where the moiré pattern reveals itself as varying interlayer separation and varying local density of states. Surprisingly, it was only in 2019 when the moiré diffraction peaks in transmission electron microscopy were observed [147].

3.2   Continuum model

To calculate the electronic properties of TBG, in Publications II and III we used the continuum model of Lopes dos Santos *et al.* [132, 134]. We chose this model due to its popularity [143, 148–151] by the time we started studying TBG in 2015. I summarize the derivation of the Hamiltonian here, but with more care given for handling the rotation and sublattices correctly.

Layer 1 Hamiltonian

Taking the same conventions for layer 1 as in PSG in Sec. 2.1 its Hamiltonian reads

$$H_1 := -t \sum_{\sigma} \sum_{j=1}^{3} \sum_{\boldsymbol{r} \in L/L_{\mathrm{BK}}} \psi_\sigma^\dagger(\boldsymbol{r}) \psi_\sigma(\boldsymbol{r} + \boldsymbol{\delta}_j) + \text{h.c.} - \mu N_1, \tag{3.10}$$

where t is the interlayer nearest neighbor hopping energy, $\psi_\sigma(\boldsymbol{r})$ is the electron annihilation operator at a crystal site $\boldsymbol{r}$, periodic in translations of L_{BK}, μ is the chemical potential, $N_1 := \sum_{\sigma \alpha \boldsymbol{r}} \psi_\sigma^\dagger(\boldsymbol{r} + \boldsymbol{\delta}_{1\alpha}) \psi_\sigma(\boldsymbol{r} + \boldsymbol{\delta}_{1\alpha})$ is the number operator of layer 1, and $\boldsymbol{\delta}_{1A} := 0$, $\boldsymbol{\delta}_{1B} := \boldsymbol{\delta}_1$ is the sublattice shift. Defining the sublattice-shifted annihilation operators of layer 1 as $\psi_{\sigma,1A}, \psi_{\sigma,1B} : L/L_{\mathrm{BK}} \to \mathrm{Op}$, $\psi_{\sigma,1A}(\boldsymbol{r}) := \psi_\sigma(\boldsymbol{r})$, $\psi_{\sigma,1B}(\boldsymbol{r}) := \psi_\sigma(\boldsymbol{r} + \boldsymbol{\delta}_1)$ for all $\boldsymbol{r} \in L/L_{\mathrm{BK}}$ allows writing H_1 as

$$H_1 = -t \sum_{\sigma} \sum_{j=1}^{3} \sum_{\boldsymbol{r} \in L/L_{\mathrm{BK}}} \psi_{\sigma,1A}^\dagger(\boldsymbol{r}) \psi_{\sigma,1B}(\boldsymbol{r} + \boldsymbol{\delta}_j - \boldsymbol{\delta}_1) + \text{h.c.} - \mu N_1 \tag{3.11}$$

with the number operator becoming $N_1 = \sum_{\sigma \alpha \boldsymbol{r}} \psi_{\sigma,1\alpha}^\dagger(\boldsymbol{r}) \psi_{\sigma,1\alpha}(\boldsymbol{r})$.

Making the valley expansion

$$\psi_{\sigma,1\alpha}(\boldsymbol{r}) = \sum_{\rho} e^{i\rho \boldsymbol{K} \cdot \boldsymbol{r}} \psi_{\sigma\rho,1\alpha}(\boldsymbol{r}) \tag{3.12}$$

around the $\boldsymbol{K}$ point as in Eq. (2.7) and linearizing $\psi_{\sigma\rho,1B}(\boldsymbol{r} + \boldsymbol{\delta}_j - \boldsymbol{\delta}_1) \approx \psi_{\sigma\rho,1B}(\boldsymbol{r}) +$

$(\boldsymbol{\delta}_j - \boldsymbol{\delta}_1) \cdot \nabla \psi_{\sigma\rho,1B}(\boldsymbol{r})$ gives the linearized Hamiltonian

$$H_1 \approx \sum_{\sigma\rho} \int_{\mathbb{R}^2/L_{\mathrm{BK}}} \mathrm{d}\boldsymbol{r} \, \psi_{\sigma\rho,1A}^\dagger(\boldsymbol{r})\hbar v_{\mathrm{F}}(-\mathrm{i})(\rho\partial_x - \mathrm{i}\partial_y)\psi_{\sigma\rho,1B}(\boldsymbol{r}) + \mathrm{h.c.} - \mu N_1$$

$$= \sum_{\sigma\rho} \int \mathrm{d}\boldsymbol{r} \, \psi_{\sigma\rho,1}^\dagger(\boldsymbol{r})\mathcal{H}_1^\rho(-\mathrm{i}\nabla)\psi_{\sigma\rho,1}(\boldsymbol{r}). \tag{3.13}$$

Here I also went to the continuum as in Eq. (2.9). Furthermore

$$\mathcal{H}_1^\rho(-\mathrm{i}\nabla) := \hbar v_{\mathrm{F}}\boldsymbol{\sigma}^\rho \cdot (-\mathrm{i}\nabla) - \mu = \begin{pmatrix} -\mu & \rho\partial_x - \mathrm{i}\partial_y \\ \rho\partial_x + \mathrm{i}\partial_y & -\mu \end{pmatrix}, \tag{3.14}$$

$$\psi_{\sigma\rho,1} := (\psi_{\sigma\rho,1A}, \psi_{\sigma\rho,1B})^\mathsf{T} \tag{3.15}$$

is a matrix and a vector in sublattice space, respectively. Equation (3.13) is the same Hamiltonian as that of PSG in Eq. (2.8) without the pseudo vector potential.

Layer 2 Hamiltonian

Taking the layer 2 to be rotated by θ (as illustrated in Fig. 3.9a) such that $\theta = 0$ corresponds to Bernal (AB) stacking, we may write its Hamiltonian as

$$H_2 := -t \sum_\sigma \sum_{j=1}^3 \sum_{\boldsymbol{r} \in L/L_{\mathrm{BK}}} \psi_\sigma^\dagger(R(\theta)(\boldsymbol{r}-\boldsymbol{\delta}_1)+\boldsymbol{c})\psi_\sigma(R(\theta)(\boldsymbol{r}-\boldsymbol{\delta}_1+\boldsymbol{\delta}_j)+\boldsymbol{c})+\mathrm{h.c.}-\mu N_2. \tag{3.16}$$

Here I took the annihilation/creation operators to be periodic in translations of $R(\theta)L_{\mathrm{BK}}$ (in addition to L_{BK}), $R(\theta)$ is the rotation matrix, $-\boldsymbol{\delta}_1$ ensures Bernal stacking at $\theta = 0$, $\boldsymbol{c} := (0, 0, c_z)$ lifts the second layer by the interlayer distance c_z, $N_2 := \sum_{\sigma\beta\boldsymbol{r}} \psi_\sigma^\dagger(R(\theta)(\boldsymbol{r} + \boldsymbol{\delta}_{2\beta}))\psi_\sigma(R(\theta)(\boldsymbol{r} + \boldsymbol{\delta}_{2\beta}))$ is the number operator of layer 2, and $\boldsymbol{\delta}_{2A} := -\boldsymbol{\delta}_1$, $\boldsymbol{\delta}_{2B} := 0$ is the sublattice shift. Defining the rotated and shifted annihilation operators of layer 2 as $\psi_{\sigma,2A}, \psi_{\sigma,2B} : L/L_{\mathrm{BK}} \to \mathrm{Op}$, $\psi_{\sigma,2A}(\boldsymbol{r}) := \psi_\sigma(R(\theta)(\boldsymbol{r} - \boldsymbol{\delta}_1) + \boldsymbol{c})$, $\psi_{\sigma,2B}(\boldsymbol{r}) := \psi_\sigma(R(\theta)\boldsymbol{r} + \boldsymbol{c})$ for all $\boldsymbol{r} \in L/L_{\mathrm{BK}}$ allows writing H_2 as

$$H_2 = -t \sum_\sigma \sum_{j=1}^3 \sum_{\boldsymbol{r} \in L/L_{\mathrm{BK}}} \psi_{\sigma,2A}^\dagger(\boldsymbol{r})\psi_{\sigma,2B}(\boldsymbol{r} - \boldsymbol{\delta}_1 + \boldsymbol{\delta}_j) + \mathrm{h.c.} - \mu N_2 \tag{3.17}$$

with the number operator becoming $N_2 = \sum_{\sigma\beta\boldsymbol{r}} \psi_{\sigma,2\beta}^\dagger(\boldsymbol{r})\psi_{\sigma,2\beta}(\boldsymbol{r})$.

Making the valley expansion

$$\psi_{\sigma,2\beta}(\boldsymbol{r}) = \sum_\rho \mathrm{e}^{\mathrm{i}\rho\boldsymbol{K}\cdot\boldsymbol{r}}\psi_{\sigma\rho,2\beta}(\boldsymbol{r}) \tag{3.18}$$

around the $\boldsymbol{K}$ point[2] and linearizing $\psi_{\sigma\rho,2B}(\boldsymbol{r} - \boldsymbol{\delta}_1 + \boldsymbol{\delta}_j) \approx \psi_{\sigma\rho,2B}(\boldsymbol{r}) + R(\theta)(\boldsymbol{\delta}_j - $

---

[2]Note that the definition of $\psi_{\sigma,2\beta}$ includes the θ rotation so effectively we are expanding the original annihilation operator ψ_σ around the rotated $\boldsymbol{K}$ point $\boldsymbol{K}^\theta := R(\theta)\boldsymbol{K}$.

$\boldsymbol{\delta}_1) \cdot \nabla \psi_{\sigma\rho,2B}(\boldsymbol{r})$ gives the linearized Hamiltonian

$$
H_2 \approx \sum_{\sigma\rho} \int_{\mathbb{R}^2/L_{\mathrm{BK}}} \mathrm{d}\boldsymbol{r}\, \psi^\dagger_{\sigma\rho,2A}(\boldsymbol{r}) \hbar v_{\mathrm{F}}(-\mathrm{i}) \mathrm{e}^{\mathrm{i}\rho\theta}(\rho\partial_x - \mathrm{i}\partial_y)\psi_{\sigma\rho,2B}(\boldsymbol{r}) + \mathrm{h.c.} - \mu N_2
$$

$$
= \sum_{\sigma\rho} \int \mathrm{d}\boldsymbol{r}\, \psi^\dagger_{\sigma\rho,2}(\boldsymbol{r}) \mathcal{H}_2^\rho(-\mathrm{i}\nabla)\psi_{\sigma\rho,2}(\boldsymbol{r}). \tag{3.19}
$$

Here I also went to the continuum through Eq. (2.9) and defined the sublattice-space matrix and sublattice-space vector

$$
\mathcal{H}_2^\rho(-\mathrm{i}\nabla) := \hbar v_{\mathrm{F}} R(\theta)\boldsymbol{\sigma}^\rho \cdot (-\mathrm{i}\nabla) - \mu = \begin{pmatrix} -\mu & \mathrm{e}^{\mathrm{i}\rho\theta}(\rho\partial_x - \mathrm{i}\partial_y) \\ \mathrm{e}^{-\mathrm{i}\rho\theta}(\rho\partial_x + \mathrm{i}\partial_y) & -\mu \end{pmatrix}, \tag{3.20}
$$

$$
\psi_{\sigma\rho,2} := (\psi_{\sigma\rho,2A}, \psi_{\sigma\rho,2B})^\mathsf{T}. \tag{3.21}
$$

Interlayer Hamiltonian

A slightly generalized (compared to the original Lopes dos Santos *et al.* model [132, 134]) interlayer Hamiltonian is

$$
H_\perp := \sum_{\sigma\alpha\beta} \sum_{\boldsymbol{r}\in L_{\mathrm{BK}}} \sum_{\boldsymbol{d}^{\alpha\beta}(\boldsymbol{r})\in M^{\alpha\beta}(\boldsymbol{r})} t_\perp(\boldsymbol{d}^{\alpha\beta}(\boldsymbol{r}))\psi^\dagger_\sigma(\boldsymbol{r}+\boldsymbol{\delta}_{1\alpha})\psi_\sigma(\boldsymbol{r}+\boldsymbol{\delta}_{1\alpha}+\boldsymbol{d}^{\alpha\beta}(\boldsymbol{r})+\boldsymbol{c})
$$

$$
+ \mathrm{h.c.}, \tag{3.22}
$$

where $M^{\alpha\beta}(\boldsymbol{r})$ includes all the difference vectors between the atom at layer 1, lattice site $\boldsymbol{r}$, sublattice α, and all the atoms at layer 2, sublattice β that satisfy some criteria (*e.g.* are closer than some cutoff distance). Lopes dos Santos *et al.* [132, 134] included only the nearest neighbor, such that the set $M^{\alpha\beta}(\boldsymbol{r})$ includes exactly one element[3] $\boldsymbol{d}^{\alpha\beta}(\boldsymbol{r})$ for each $\boldsymbol{r}$, α, β. In Publications II and III we followed this model. Furthermore $t_\perp$ is the interlayer hopping energy which has to be parametrized somehow; the parametrization chosen by Lopes dos Santos *et al.* is described in the Supplementary Material of Publication III. Using the layer-specific, sublattice-shifted field operators, $H_\perp$ can be written as

$$
H_\perp = \sum_{\sigma\alpha\beta} \sum_{\boldsymbol{r}\in L/L_{\mathrm{BK}}} \sum_{\boldsymbol{d}^{\alpha\beta}(\boldsymbol{r})\in M^{\alpha\beta}(\boldsymbol{r})} t_\perp(\boldsymbol{d}^{\alpha\beta}(\boldsymbol{r})) \times
$$

$$
\times \psi^\dagger_{\sigma,1\alpha}(\boldsymbol{r})\psi_{\sigma,2\beta}(R^{-1}(\theta)(\boldsymbol{r}+\boldsymbol{\delta}_{1\alpha}+\boldsymbol{d}^{\alpha\beta}(\boldsymbol{r})) - \boldsymbol{\delta}_{2\beta}) + \mathrm{h.c.} \tag{3.23}
$$

Carrying out the valley expansions (3.12) and (3.18), the zeroth-order approximation $\psi_{\sigma\rho,2\beta}(R^{-1}(\theta)(\boldsymbol{r}+\boldsymbol{\delta}_{1\alpha}+\boldsymbol{d}^{\alpha\beta}(\boldsymbol{r})) - \boldsymbol{\delta}_{2\beta}) \approx \psi_{\sigma\rho,2\beta}(\boldsymbol{r})$ yields

$$
H_\perp \approx \sum_{\sigma\rho} \int_{\mathbb{R}^2/L_{\mathrm{BK}}} \mathrm{d}\boldsymbol{r}\, \psi^\dagger_{\sigma\rho,1}(\boldsymbol{r}) \mathrm{e}^{\mathrm{i}\rho\Delta\boldsymbol{K}\cdot\boldsymbol{r}} T_\perp^\rho(\boldsymbol{r})\psi_{\sigma\rho,2}(\boldsymbol{r}) + \mathrm{h.c.} \tag{3.24}
$$

---

[3]The single nearest-neighbor difference vector $\boldsymbol{d}^{\alpha\beta}(\boldsymbol{r}) \in M^{\alpha\beta}(\boldsymbol{r})$ included in Refs. [132, 134] and in Publications II and III is denoted by $\boldsymbol{\delta}^{\alpha\beta}(\boldsymbol{r})$ in those works.

Here $\Delta \boldsymbol{K} \coloneqq \boldsymbol{K}^\theta - \boldsymbol{K}$ is the difference vector from the original $\boldsymbol{K}$ point to the rotated $\boldsymbol{K}$ point $\boldsymbol{K}^\theta \coloneqq R(\theta)\boldsymbol{K}$, $\psi_{\sigma\rho,1}$ and $\psi_{\sigma\rho,2}$ are the sublattice-space vectors of Eqs. (3.15) and (3.21), the sublattice components of $T_\perp^\rho(\boldsymbol{r})$ are

$$T_\perp^{\rho,\alpha\beta}(\boldsymbol{r}) \coloneqq \sum_{\boldsymbol{d}^{\alpha\beta}(\boldsymbol{r}) \in M^{\alpha\beta}(\boldsymbol{r})} e^{i\rho \boldsymbol{K}^\theta \cdot (\boldsymbol{\delta}_{1\alpha} + \boldsymbol{d}^{\alpha\beta}(\boldsymbol{r}))} t_\perp(\boldsymbol{d}^{\alpha\beta}(\boldsymbol{r})), \qquad (3.25)$$

and I also went to the continuum as in Eq. (2.9). The continuum limit has to be taken carefully, however, as $T_\perp$ is only defined in the lattice L and not in $\mathbb{R}^2$, and thus $T_\perp$ has to be interpolated somehow. For commensurate structures $T_\perp : L/SL \to \mathbb{C}_{2\times 2}$ is periodic in translations of the superlattice SL, and a natural interpolation is provided by its discrete Fourier series

$$T_\perp^\rho(\boldsymbol{r}) = \sum_{\boldsymbol{G} \in SL^*/L^*} e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \tilde{T}_\perp^\rho(\boldsymbol{G}), \qquad (3.26)$$

which is trivially extended to $\mathbb{R}^2$ after the Fourier coefficients have been calculated from[4]

$$\tilde{T}_\perp^\rho(\boldsymbol{G}) = \frac{1}{|L/SL|} \sum_{\boldsymbol{r} \in L/SL} e^{-i\boldsymbol{G}\cdot\boldsymbol{r}} T_\perp^\rho(\boldsymbol{r}). \qquad (3.27)$$

In deriving Eq. (3.24) I have assumed that $T_\perp$ is slowly varying in the lattice scale, *i.e.* that its Fourier coefficients $\tilde{T}_\perp(\boldsymbol{G})$ have only components with small $\boldsymbol{G}$ (compared to $\boldsymbol{K}$); otherwise also the opposite valleys would be coupled. Also, the factor $e^{i\rho\Delta\boldsymbol{K}\cdot\boldsymbol{r}}$ is written separately of $T_\perp$ because they have generally a different periodicity.

Total Hamiltonian

The total Hamiltonian can then be written as

$$H = H_1 + H_2 + H_\perp \qquad (3.28)$$

$$= \sum_{\sigma\rho} \int d\boldsymbol{r} \begin{pmatrix} \psi_{\sigma\rho,1}(\boldsymbol{r}) \\ \psi_{\sigma\rho,2}(\boldsymbol{r}) \end{pmatrix}^\dagger \begin{pmatrix} \mathcal{H}_1^\rho(-i\nabla) & e^{i\rho\Delta\boldsymbol{K}\cdot\boldsymbol{r}} T_\perp^\rho(\boldsymbol{r}) \\ e^{-i\rho\Delta\boldsymbol{K}\cdot\boldsymbol{r}} T_\perp^\rho(\boldsymbol{r})^\dagger & \mathcal{H}_2^\rho(-i\nabla) \end{pmatrix} \begin{pmatrix} \psi_{\sigma\rho,1}(\boldsymbol{r}) \\ \psi_{\sigma\rho,2}(\boldsymbol{r}) \end{pmatrix},$$

The exponential factor can also be moved to the intralayer Hamiltonians and the basis operators by writing

$$H = \sum_{\sigma\rho} \int d\boldsymbol{r} \begin{pmatrix} \phi_{\sigma\rho,1}(\boldsymbol{r}) \\ \phi_{\sigma\rho,2}(\boldsymbol{r}) \end{pmatrix}^\dagger \begin{pmatrix} \mathcal{H}_1^\rho(-i\nabla + \rho\Delta\boldsymbol{K}/2) & T_\perp^\rho(\boldsymbol{r}) \\ T_\perp^\rho(\boldsymbol{r})^\dagger & \mathcal{H}_2^\rho(-i\nabla - \rho\Delta\boldsymbol{K}/2) \end{pmatrix} \begin{pmatrix} \phi_{\sigma\rho,1}(\boldsymbol{r}) \\ \phi_{\sigma\rho,2}(\boldsymbol{r}) \end{pmatrix}$$

$$=: \sum_{\sigma\rho} \int d\boldsymbol{r} \, \phi_{\sigma\rho}^\dagger(\boldsymbol{r}) \mathcal{H}^\rho(\boldsymbol{r}) \phi_{\sigma\rho}(\boldsymbol{r}) \qquad (3.29)$$

---

[4]$\tilde{T}_\perp$ is denoted by $\tilde{t}_\perp$ in Refs. [132, 134] and in Publications II and III. Note also that regarding the $\boldsymbol{\delta}_{1\alpha}$ term the Fourier-coefficient equation (3.27) together with Eq. (3.25) is slightly different from those works. The difference is however very small, as in the small-angle limit $\boldsymbol{K}^\theta \cdot \boldsymbol{\delta}_{1\alpha}$ is small (because $\boldsymbol{K} \cdot \boldsymbol{\delta}_1 = 0$) and with only small $\boldsymbol{G}$ components $\boldsymbol{G} \cdot \boldsymbol{\delta}_{1\alpha}$ is also small. The term $\boldsymbol{G} \cdot \boldsymbol{\delta}_{1\alpha}$ is not explicitly written in Refs. [132, 134], but it is there in the source code.

and defining the new annihilation operators

$$\phi_{\sigma\rho,1}(\boldsymbol{r}) \coloneqq \mathrm{e}^{-\mathrm{i}\rho\Delta\boldsymbol{K}\cdot\boldsymbol{r}/2}\psi_{\sigma\rho,1}(\boldsymbol{r}), \quad \phi_{\sigma\rho,2}(\boldsymbol{r}) \coloneqq \mathrm{e}^{\mathrm{i}\rho\Delta\boldsymbol{K}\cdot\boldsymbol{r}/2}\psi_{\sigma\rho,2}(\boldsymbol{r}), \qquad (3.30)$$

which is the real-space version of what Lopes dos Santos *et al.* [132, 134] also did.

3.2.1 Noninteracting Hamiltonian in Fourier space

Due to the moiré-periodicity of the commensurate structures the Hamiltonian is most conveniently diagonalized in Fourier space. The continuum "interlayer coupling" $T_\perp : \mathbb{R}^2/SL \to \mathbb{C}_{2\times2}$ is periodic in translations of the superlattice SL, so we may expand it as the Fourier series (3.26), but with the sum going over $\boldsymbol{G} \in SL^*$ instead. I take also the field operators $\phi_{\sigma\rho,1}, \phi_{\sigma\rho,2} : L/L_{\mathrm{BK}} \to \mathrm{Op}^2$ to be periodic in translations of the Born–von Kármán lattice L_{BK}, in which case there exist the Fourier series

$$\phi_{\sigma\rho,1}(\boldsymbol{r}) = \sum_{\boldsymbol{k}\in L_{\mathrm{BK}}^*} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} d_{\sigma\rho,1}(\boldsymbol{k}), \qquad (3.31)$$

$$\phi_{\sigma\rho,2}(\boldsymbol{r}) = \sum_{\boldsymbol{k}\in L_{\mathrm{BK}}^*} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} d_{\sigma\rho,2}(\boldsymbol{k}). \qquad (3.32)$$

Inserting these to Eq. (3.29) yields

$$H = N \sum_{\sigma\rho} \sum_{\boldsymbol{k}\in L_{\mathrm{BK}}^*} \Bigg[d_{\sigma\rho,1}^\dagger(\boldsymbol{k})\mathcal{H}_1^\rho(\boldsymbol{k} + \rho\Delta\boldsymbol{K}/2)d_{\sigma\rho,1}(\boldsymbol{k})$$
$$d_{\sigma\rho,2}^\dagger(\boldsymbol{k})\mathcal{H}_2^\rho(\boldsymbol{k} - \rho\Delta\boldsymbol{K}/2)d_{\sigma\rho,2}(\boldsymbol{k})$$
$$\sum_{\boldsymbol{G}'\in SL^*} d_{\sigma\rho,1}^\dagger(\boldsymbol{k})\tilde{T}_\perp^\rho(\boldsymbol{G}')d_{\sigma\rho,2}(\boldsymbol{k} - \boldsymbol{G}') + \mathrm{h.c.} \Bigg], \qquad (3.33)$$

which is the form shown also by Lopes dos Santos *et al.* [132, 134].

Dividing the $\boldsymbol{k}$ sum in the reduced zone scheme as in Eq. (2.23) yields

$$H = N \sum_{\sigma\rho} \sum_{\boldsymbol{k}\in L_{\mathrm{BK}}^*} \sum_{\boldsymbol{G}\boldsymbol{G}'\in SL^*} d_{\sigma\rho}^\dagger(\boldsymbol{k} + \boldsymbol{G})\tilde{\mathcal{H}}_{\boldsymbol{G}\boldsymbol{G}'}^\rho(\boldsymbol{k})d_{\sigma\rho}(\boldsymbol{k} + \boldsymbol{G}'). \qquad (3.34)$$

The connection between the original graphene BZ L_{BK}^*/L^* and the SBZ L_{BK}^*/SL^* is illustrated in Fig. 3.4. Note that the original graphene BZ is important in cutting off the $\boldsymbol{G}$ components when calculating the interlayer Fourier coefficients (3.27), but is otherwise lost in the continuum theory. Here

$$\tilde{\mathcal{H}}_{\boldsymbol{G}\boldsymbol{G}'}^\rho(\boldsymbol{k}) \coloneqq \begin{pmatrix} \mathcal{H}_1^\rho(\boldsymbol{k} + \boldsymbol{G} + \rho\Delta\boldsymbol{K}/2)\delta_{\boldsymbol{G}\boldsymbol{G}'} & \tilde{T}_\perp^\rho(\boldsymbol{G} - \boldsymbol{G}') \\ \tilde{T}_\perp^\rho(\boldsymbol{G}' - \boldsymbol{G})^\dagger & \mathcal{H}_2^\rho(\boldsymbol{k} + \boldsymbol{G} - \rho\Delta\boldsymbol{K}/2)\delta_{\boldsymbol{G}\boldsymbol{G}'} \end{pmatrix}, \quad (3.35)$$

$$d_{\sigma\rho} \coloneqq (d_{\sigma\rho,1}, d_{\sigma\rho,2})^\mathsf{T} \qquad (3.36)$$

is a matrix and a vector in layer space, respectively. Furthermore defining the $\boldsymbol{G}$-space matrices and vectors as in Eq. (2.26) (in the reduced zone scheme) brings
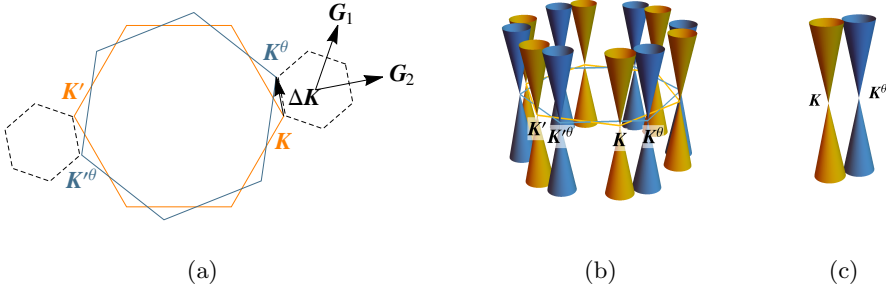
(a)   (b)   (c)

Figure 3.4: Illustration of the superlattice Brillouin zone (SBZ) and the double-cone dispersion of TBG. (a) The layer 1 BZ L^*_{BK}/L^* (orange hexagon) with the Dirac $\boldsymbol{K}$ and $\boldsymbol{K}'$ points, the rotated layer 2 BZ $R(\theta)L^*_{\mathrm{BK}}/R(\theta)L^*$ (blue hexagon) with the rotated Dirac points $\boldsymbol{K}^\theta = R(\theta)\boldsymbol{K}$ and $\boldsymbol{K}'^\theta = R(\theta)\boldsymbol{K}'$ points, and the SBZ L^*_{BK}/SL^* (black dashed hexagon) defined by the reciprocal superlattice vectors $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$ [Eq. (3.5)]. One of the SBZ edges is given by $\Delta\boldsymbol{K} = \boldsymbol{K}^\theta - \boldsymbol{K}$. (b) Illustration of the Dirac cones of the two graphene layers in the original BZ, when no interlayer coupling is present. (c) Constraining ourselves close to one pair of Dirac points effectively gives a double-cone structure for each valley $\boldsymbol{K}$ and $\boldsymbol{K}'$.

H to the form

$$H = N \sum_{\sigma\rho} \sum_{\boldsymbol{k}\in L^*_{\mathrm{BK}}/SL^*} \underline{d}^\dagger_{\sigma\rho}(\boldsymbol{k})\tilde{\underline{\mathcal{H}}}^\rho(\boldsymbol{k})\underline{d}_{\sigma\rho}(\boldsymbol{k}). \tag{3.37}$$

Now many of the electronic properties can be calculated by diagonalizing the normal matrix $\tilde{\underline{\mathcal{H}}}^\rho(\boldsymbol{k})$.

What is left is the calculation of the "interlayer coupling" Fourier coefficients $\tilde{T}_\perp(\boldsymbol{G})$ through Eq. (3.27), where one has to fix the form of the coupling $t_\perp(\boldsymbol{r})$. Using the parametrization of Lopes dos Santos *et al.* [132, 134], described in the Supplementary Material of Publication III, it can be numerically shown that the largest coefficients are those for $\boldsymbol{G} \in \{0, -\boldsymbol{G}_1, -\boldsymbol{G}_1 - \boldsymbol{G}_2\}$ (valley $\rho = +$) or $\boldsymbol{G} \in \{0, \boldsymbol{G}_1, \boldsymbol{G}_1 + \boldsymbol{G}_2\}$ (valley $\rho = -$). Their absolute value is $|\tilde{T}_\perp(\boldsymbol{G})| \approx 0.11\,\mathrm{eV}$, independent of the rotation angle up to the same precision. The rest of the components are an order of magnitude smaller. This was also shown analytically by Lopes dos Santos *et al.* [132]: the coefficients die sufficiently rapidly with increasing $\|\boldsymbol{G} + \rho\Delta\boldsymbol{K}\|$. This is why a model including only these three components in $\tilde{T}_\perp(\boldsymbol{G})$ (and setting the rest to zero) should yield the relevant physics. I call this shortly the *3G LdS model*. Opposite to this, I call the model including all the components $\boldsymbol{G} \in SL^*/L^*$ of $\tilde{T}_\perp(\boldsymbol{G})$ the *full LdS model*. The amplitude $|\tilde{T}_\perp| \approx 0.11\,\mathrm{eV}$ was experimentally confirmed in 2012 by Brihuega *et al.* [143] in scanning tunneling microscopy (STM) and spectroscopy (STS) measurements.

In the first article Lopes dos Santos *et al.* [134] included only the abovementioned three $\boldsymbol{G}$ components, but at the same time truncated also the basis of the Hamiltonian to $\{d_{\sigma+,1}(\boldsymbol{k}), d_{\sigma+,1}(\boldsymbol{k}-\boldsymbol{G}_1), d_{\sigma+,1}(\boldsymbol{k}-\boldsymbol{G}_1-\boldsymbol{G}_2), d_{\sigma+,2}(\boldsymbol{k}), d_{\sigma+,2}(\boldsymbol{k}+\boldsymbol{G}_1), d_{\sigma+,2}(\boldsymbol{k}+\boldsymbol{G}_1+\boldsymbol{G}_2)\}$ (valley $\rho = +$), which is *not* the 3G LdS model. This

works for sufficiently large angles, but it starts to break the superlattice periodicity when approaching the magic angle at $1°$. In the second article [132] they truncated the basis to much larger $\boldsymbol{G}$ vectors, determined from the saturation of the bandstructure for each angle. This is what we also did in Publications II and III. In Publication II we employed the $3G$ LdS model because it better reproduces the flat-band bandwidth [121] of the widely-used Bistritzer–MacDonald model [140], although at a slightly different (magic) angle. This difference is discussed more in Sec. 3.3. On the other hand in Publication III we used the full LdS model, because its magic angle, flat-band bandwidth, and the band gap between the flat and dispersive bands better matches many other models recently considered. While summarizing the results of Publications II and III in Secs. 3.3 and 3.5 I also discuss the key differences between the $3G$ and full LdS models.

Bistritzer & MacDonald [140] have a widely-used model for TBG which is very similar to the $3G$ LdS model. The key difference is that instead of parametrizing the interlayer hopping energy function $t_\perp$ in real space, they parametrized it directly in Fourier space. They especially used the abovementioned fact that $\tilde{T}(\boldsymbol{G})$ usually dies quickly with increasing $\|\boldsymbol{G} + \rho\Delta\boldsymbol{K}\|$, so that it is enough to choose only the values of $\tilde{T}(\boldsymbol{G})$ for three different $\boldsymbol{G}$ vectors. Moreover the different sublattice components are connected by (approximate) symmetry and they chose the three different $\boldsymbol{G}$ components to have the same amplitude, so they have only one parameter $|\tilde{T}_\perp| = 0.11\,\text{eV}$ in the interlayer coupling. This approach, where one fixes the Fourier components and assumes them to be constant as a function of the rotation angle θ, has the advantage that structures with *any* $\theta \in \mathbb{R}$ can be modelled, with no commensuration condition.

3.3 Electronic properties of the normal state

In this section I give a brief historical overview of the normal-state electronic properties that have been theoretically predicted or experimentally observed in TBG. A review on the subject is provided in Ref. [152], as well as in a few theses [153, 154]. Since in Publication III we used the 2012 Lopes dos Santos *et al.* [132] model (the full LdS model) and in Publication II a combination of the 2007 [134] and 2012 [132] models (the $3G$ LdS model), I show more detailed results of them. I also summarize the main differences between the $3G$ LdS, full LdS, and Bistritzer–MacDonald models.

3.3.1 Bandstructure and flat bands

I start by illustrating how the dispersion of TBG is formed and is usually plotted in the low-energy continuum theories within one of the valleys. Figure 3.5 shows the evolution of the dispersion when one starts from two uncoupled layers in the extended zone scheme (as in the hand-waving picture of Fig. 3.4), turns to the reduced zone scheme in the SBZ, lowers the rotation angle, and finally turns on the interaction (within the $3G$/full LdS model). This forms van Hove singularities in the spectrum, which according to Fig. 3.5d are located at $\pm 0.1\,\text{eV}$ for the example rotation of $\theta = 2.88°$. As discussed below, the bands are flattened with decreasing
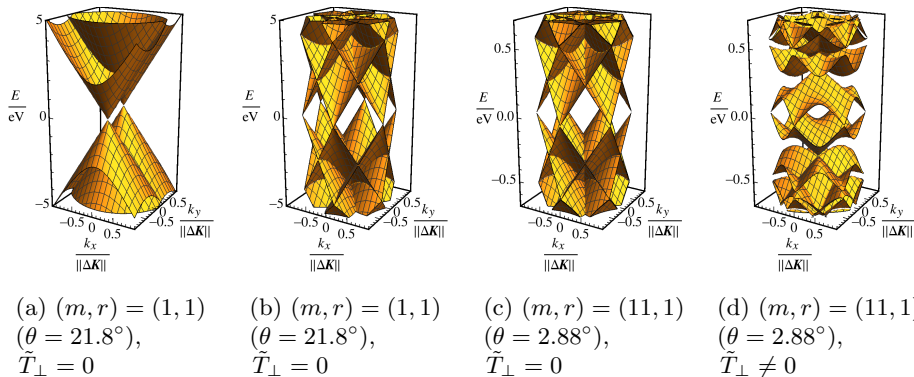
(a) $(m, r) = (1, 1)$ $(\theta = 21.8°)$, $\tilde{T}_\perp = 0$

(b) $(m, r) = (1, 1)$ $(\theta = 21.8°)$, $\tilde{T}_\perp = 0$

(c) $(m, r) = (11, 1)$ $(\theta = 2.88°)$, $\tilde{T}_\perp = 0$

(d) $(m, r) = (11, 1)$ $(\theta = 2.88°)$, $\tilde{T}_\perp \neq 0$

Figure 3.5: Illustrating the dispersion of TBG at the valley $\boldsymbol{K}$, as calculated from $\tilde{\underline{\mathcal{H}}}^+(\boldsymbol{k})$ in Eq. (3.37). (a) Two uncoupled ($\tilde{T}_\perp = 0$) layers in the extended zone scheme ($\boldsymbol{G}, \boldsymbol{G}' = 0$) gives effectively Fig. 3.4c, here shown for $m = 1$. Shifting the origin by $\boldsymbol{G}_2/2$ and going to the reduced zone scheme in the SBZ [i.e. making the superlattice-periodic copies of (a) through the $\boldsymbol{G}$ vectors, and restricting to the SBZ gives] gives (b). Lowering the rotation angle to $m = 11$ gives (c). Note the order-of-magnitude difference in the energy scale. Turning on the interlayer coupling within the $3G$/full LdS model gives (d). Further lowering the rotation angle gives Fig. 3.7.

θ, which also moves the van Hove singularities closer to zero. Some of the earliest scanning tunneling spectroscopy (STS) measurements confirming these van Hove singularities were done by Li *et al.* [141] and Brihuega *et al.* [143].

Tight-binding (TB) models are easy to develop, as demonstrated in Sec. 3.2, but without doing the low-energy continuum approximation the Hamiltonian matrices are huge. The matrix size grows linearly with the number of atoms in the superlattice unit cell, which on the other hand grows with decreasing θ through Eq. (3.7). At the interesting (magic) angle of $\theta \approx 1°$, the superlattice unit cell has some $10\,000$ atoms [155], making computations heavy. The low-energy continuum approximation greatly reduces the Hamiltonian matrix size, and in 2007 Lopes dos Santos *et al.* [134] were the first to present a continuum model of TBG (the one presented in Sec. 3.2) working in the intermediate-angle regime $3° \lesssim \theta \lesssim 15°$ [148, 156] with $r = 1$. They identified the double-Dirac-cone structure in the dispersion originating from the Dirac cones of the two layers, as illustrated in Fig. 3.5. Using perturbation theory they also found that in the intermediate-angle regime TBG behaves like two uncoupled graphene layers with a renormalized Fermi velocity v_F^* (the slope of the Dirac cone). Up to second order in α it is given by

$$\frac{v_F^*}{v_F} \approx 1 - 9\alpha^2 \quad \text{with} \quad \alpha := \frac{\tilde{T}_\perp}{\hbar v_F \|\Delta \boldsymbol{K}\|}, \tag{3.38}$$

where v_F is the Fermi velocity of graphene and $\tilde{T}_\perp \approx 0.11\,\text{eV}$ is the absolute value of the most important interlayer coupling Fourier components in Eq. (3.27), as

(a) $\frac{\tilde{T}_\perp}{\hbar v_\mathrm{F} \|\mathbf{\Delta K}\|} = 0$ (b) $\frac{\tilde{T}_\perp}{\hbar v_\mathrm{F} \|\mathbf{\Delta K}\|} = 0.1$ (c) $\frac{\tilde{T}_\perp}{\hbar v_\mathrm{F} \|\mathbf{\Delta K}\|} = 5$

(large θ)   (small θ)
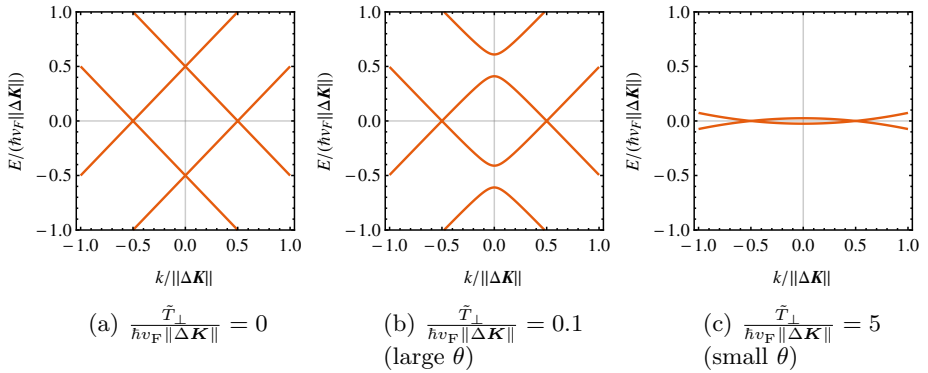
Figure 3.6: Illustration of the formation of flat bands in TBG within a simplified model of two conical intralayer dispersions and a homogeneous interlayer coupling $\tilde{T}_\perp$. (a) With no interlayer coupling, we have the double-cone dispersion of Fig. 3.4c or 3.5a. (b) Turning on the interlayer coupling $\tilde{T}_\perp$ causes an avoided crossing, the size of which is $2\tilde{T}_\perp$. The avoided crossing, on the other hand, creates van Hove singularities. (c) The size of the avoided crossing is increased with increasing interlayer coupling, which means that the van Hove singularities are moved closer to zero. Equation (3.38) says that the only relevant parameter is $\tilde{T}_\perp/\|\mathbf{\Delta K}\|$, so increasing interlayer coupling is similar to decreasing $\|\mathbf{\Delta K}\| \propto \theta$ (for small angles).

discussed in Sec. 3.2. Although not discussed in the original article, Eq. (3.38) can be seen to predict the vanishing of v_F^* at $9\alpha^2 = 1$ corresponding to $\theta \approx 1.9°$. This, however, overestimates the angle since the perturbation theory behind Eq. (3.38) breaks down when $\theta \lesssim 5°$ [132]. Vanishing of v_F^* would correspond to a flat band. The formation of the flat bands can be intuitively understood from Fig. 3.6, which shows a 1D cut of the double-cone dispersion shown in Fig. 3.4c or 3.5a. Flat bands are created with increasing (homogeneous) interlayer coupling $\tilde{T}_\perp$ through avoided crossing, which can be related to the decrease in $\|\mathbf{\Delta K}\|$ through Eq. (3.38), which is furthermore linearly proportional to the rotation angle θ in the small-angle limit. Thus the bands are flattened with decreasing θ. This simplified picture predicts a monotonous decrease of the bandwidth, but as we see below, this is not totally correct.

In the models of TBG, tight-binding and continuum models have dominated over density functional theory (DFT), which is notoriously heavy due to the huge size of the superlattice unit cell at small angles. Suárez Morell *et al.* [157] developed a hybrid method which parametrizes a TB model by a DFT calculation of large-angle structures (including lattice relaxation), and by using this method they predicted the existence of flat bands at the *magic angle*[5] $\theta \approx 1.5°$, thus confirming the "prediction" of Lopes dos Santos *et al.* [134].

Mele [137] presented a similar continuum model as Lopes dos Santos *et al.*, but

---

[5]The definition of a magic angle is a little ambiguous. For example some [157] define it through the (local) minimum bandwidth while others [140] through the (local) minimum of v_F^*. As Suárez Morell *et al.* [157] showed, these are not necessarily the same thing.

one which works for all commensurate structures. He found that at large angles the sublattice-exchange odd $(r \nmid 3)$ and even $(r \mid 3)$ structures are electronically different at large angles. Trambly de Laissardière *et al.* [156] used DFT for large angles and TB calculations for small angles and again found flat bands at the first magic angle $\theta \approx 1°$. They also confirmed that the v_{F} renormalization equation (3.38) works well when $\theta \gtrsim 3°$.

One of the most popular [121, 158–170] models nowadays is the one-parameter continuum model of Bistritzer & MacDonald [140] discussed already above, which is valid in the small-angle regime $\theta \lesssim 10°$, even for incommensurate structures. Their main result is the prediction of *multiple* magic angles: v_{F}^* vanishes at a series of angles $\theta^* \approx 1.05°, 0.5°, 0.35°, 0.24°, 0.2°, \ldots$ (described roughly by $\theta^* \approx 1.05°/n$ with $n \in \mathbb{N}$). They also presented a simplified eight-band model which is often used [171, 172] as well. Although the interlayer couplings are roughly the same $|\tilde{T}_\perp| \approx 0.11\,\mathrm{eV}$ in the Bistritzer–MacDonald and $3G$ LdS models, they produce a slightly different magic angle:[6] $\theta^* = 0.96°$ for the $3G$ LdS model (Publication II) and $\theta^* = 1.05°$ for the Bistritzer–MacDonald model. I have checked that this happens because of two reasons: (i) reducing the intralayer coupling from $t = -3.08\,\mathrm{eV}$ of the $3G$ LdS model to $t = -2.9\,\mathrm{eV}$ and (ii) making the symmetry approximation of Table I in Refs. [132, 134] between the different sublattice components of $\tilde{T}_\perp(\boldsymbol{G})$ increases the magic angle to the Bistritzer–MacDonald value of $1.05°$.

Trambly de Laissardière *et al.* [148] used a DFT-parametrized TB model to study also very small angles and found remarkable agreement with the Bistritzer–MacDonald model: flat bands at the first magic angle $\approx 1.1°$ and at the second magic angle $\approx 0.55°$. At intermediate angles they also confirmed the validity of Eq. (3.38) for $\theta \gtrsim 3°$. On the other hand they predicted that at large angles $15° \lesssim \theta \lesssim 30°$ the graphene layers are effectively decoupled.

In 2012 Lopes dos Santos *et al.* [132] refined their 2007 model [134] by using larger bases for the Hamiltonian, as discussed in Sec. 3.2.1. Using this refined continuum model they also confirmed the existence of flat bands, happening roughly at $\theta^* \approx 1.08°/n$ with $n \in \mathbb{N}$, and also the validity of Eq. (3.38) for $\theta \gtrsim 5°$. The appearance of the flat bands at $1.08°$ using this model (the full LdS model used in Publication III) is shown in the second row of Fig. 3.7, which also shows significant gaps between the flat and the dispersive bands below the magic angle. On the other hand the first row of Fig. 3.7 shows the appearance of the flat bands at $0.96°$ using the $3G$ LdS model (a combination of the 2007 [134] and 2012 [132] Lopes dos Santos *et al.* models) and the lack of band gaps. It also shows (when zoomed in) a major difference in the bandwidth: roughly $1\,\mathrm{meV}$ for the $3G$ and $15\,\mathrm{meV}$ for the full LdS model. These differences highlight the sensitivity of the electronic properties of TBG close to the magic angle. In Fig. 3.8 I further show the density of states (DOS) at the magic angle for both models. It shows that while the bandwidths are highly different, the energy range of significant DOS ($1\,\mathrm{meV}$ for the $3G$ and $3\,\mathrm{meV}$ for the full LdS model) is not. What mostly matters in Publications II and III is the DOS and not the actual bandwidth, so we use

---

[6] By "magic angle" I always mean the largest magic angle, which is also called the first magic angle.
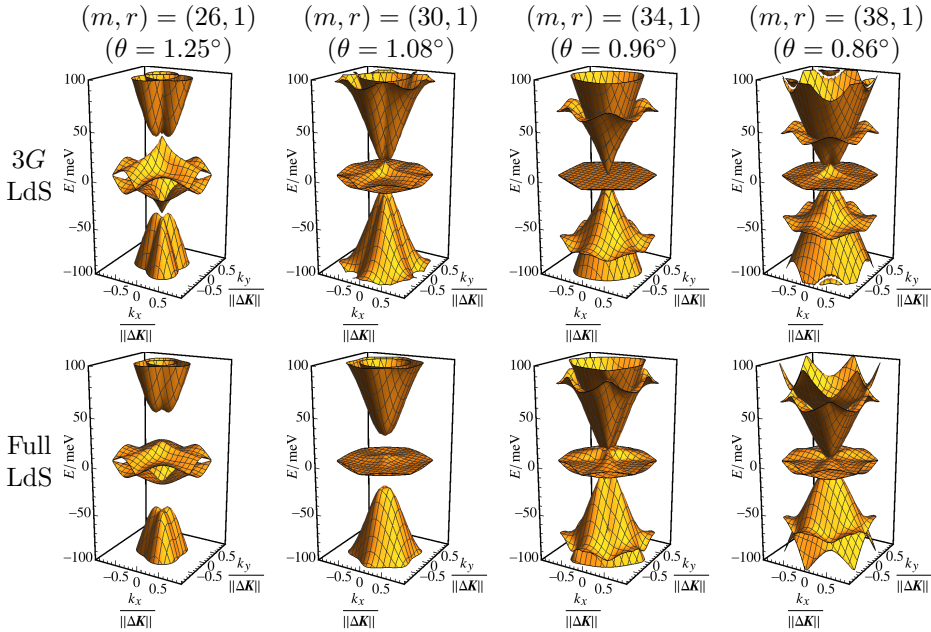
Figure 3.7: Dispersions at the valley $\rho = +$ for the $3G$ and full LdS models close to the magic angle. Out of the $r = 1$ structures, the $3G$ LdS model has the magic angle at $m = 34$ while the full LdS model at $m = 30$. The main differences at the magic angle are the larger bandwidth of the full model and the gap to the higher bands. The last three subfigures of the first row are reprinted with permission from Ref. [173]. Copyright 2018 American Physical Society. They have been modified by updating the labels.

the terms "bandwidth" and "energy range of significant DOS" interchangeably. It is notable that different STS measurements [141, 143–146] give varying values for the separation between the DOS peaks (the van Hove singularities), but most of them are much larger than the 1 meV separation in the full LdS model. On the other hand calculating the local density of states (Fig. 3.9b) along the blue line of Fig. 3.9a shows how the low-energy states are localized at the AA stacking regions [132, 140, 155, 174, 175]. This was experimentally confirmed by Luican *et al.* [142] in an STS measurement.

It is notable that the dispersions in Fig. 3.7 of the $3G$ and full LdS models are approximately electron–hole symmetric (with respect to the charge neutrality) when summed over valleys, as seen in the DOS in Fig. 3.8. The same is true for the Bistritzer–MacDonald model. However, many other models [109–111, 176, 177] predict strong electron–hole asymmetry, with much flatter bands on the hole (negative-energy) side. According to the same studies structural relaxation even amplifies this asymmetry.

By combining STM and Landau-level spectroscopy, Luican *et al.* [142] provided in 2011 an experimental verification of the three different angle regimes: (i) at large
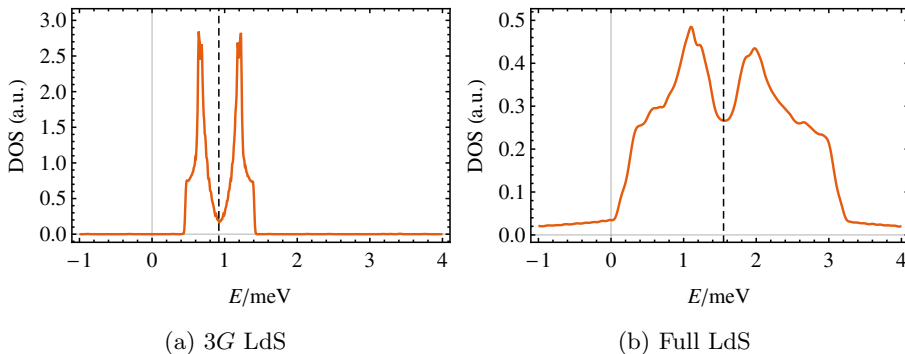
Figure 3.8: Density of states (DOS) at the magic angle for the $3G$ and full LdS models. The dashed line shows the charge neutrality $n_{\mathrm{normal}}(\mu_0) = 0$ at μ_0, where the excess charge n_{normal} in the normal state relative to the charge neutrality is calculated through Eq. (S8) of the Supplementary Material in Publication II. Each plot is normalized such that the visible area integrates to unity.

angles $15° \lesssim \theta \lesssim 30°$ the layers are effectively decoupled as predicted by Trambly de Laissardière *et al.* [156], Shallcross *et al.* [135], and others [155, 178], (ii) at the intermediate angles $3° \lesssim \theta \lesssim 20°$ electrons behave as massless Dirac fermions with a renormalized v_{F} as predicted by Lopes dos Santos *et al.* [132, 134], Bistritzer & MacDonald [140], and others [135, 155, 156], and (iii) in the small-angle regime $\theta \lesssim 3°$ the system is dominated by the van Hove singularities as predicted by the same groups.

Shallcross *et al.* [135] used Diophantine analysis to show which interlayer matrix elements vanish, and as a consequence they were able to show that the electronic structure depends only on the rotation angle θ and not on the details of the unit cell (*i.e.* not on the parameters m and r). Later it was shown [174, 178] more rigorously that the geometric momentum scale $G = \|\boldsymbol{G}_1\|$ in Eq. (3.5) (which is defined only for commensurate angles and is a highly discontinuous function of θ as shown in Fig. 3.3b) is physically irrelevant, while the effective momentum scale defining the electronic properties is G_{approx} in Eq. (3.9),[7] which is a continuous function of θ. This solved the "similar structures problem", where it was not clear why similar-looking structures with highly different reciprocal space structures should yield similar physics. This problem was also addressed by Lopes dos Santos *et al.* [132] by noting that all commensurate structures are *almost periodic* repetitions of the $r = 1$ structures. Their idea was then to approximate a given arbitrary commensurate (m, r) structure with a similar-looking $(m', 1)$ structure.[8]

Uchida *et al.* [155] performed the first full DFT calculations of TBG going all the way to the magic angle (and even below), which they observed at $1.08°$, being thus in agreement with the full LdS model and the Bistritzer–MacDonald

---

[7]Shallcross *et al.* [174] denote G_{approx} by $g^{(c)}$ and use reciprocal-space units of $2\pi/a$. This is why the effective momentum scale in that article is written as $g^{(c)} = \frac{4}{\sqrt{3}} \sin(\theta/2)$.

[8]The argumentation by Lopes dos Santos *et al.* [132] is, however, somewhat faulty: the presented method works as such only for structures with $r = 3$.
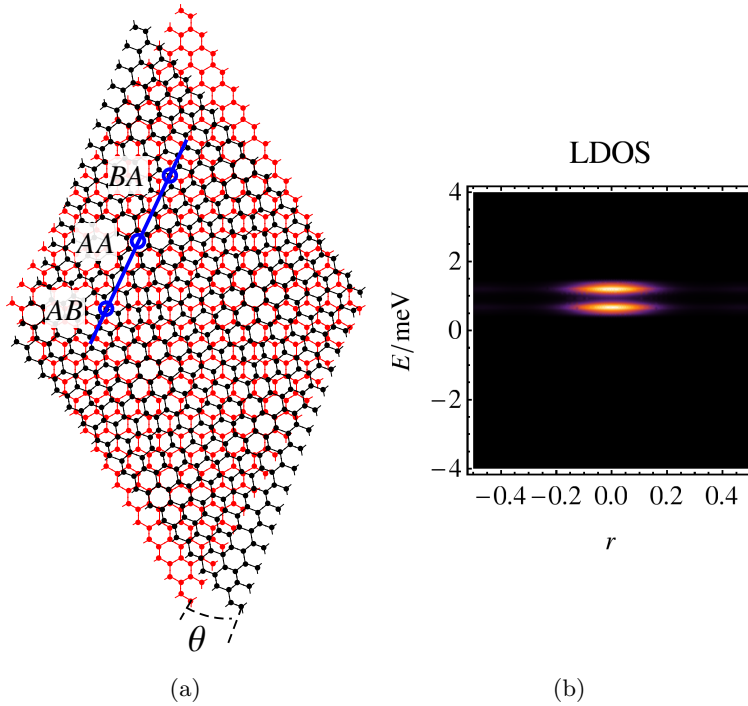
Figure 3.9: (a) Twisted bilayer graphene, where the straight layer 1 is shown in red and the θ-rotated layer 2 in black. The blue line goes through the high-symmetry points (blue circles) with AB, AA, and BA stacking. (b) Local density of states (LDOS) along the blue line of (a) at the magic angle within the $3G$ LdS model, showing how the low-energy states are localized at the AA stacking regions. Reprinted with permission from Ref. [173]. Copyright 2018 American Physical Society. Figure (a) has been modified by changing the font of the labels and increasing the thickness of the blue objects.

model. This way they were also able to study atomic corrugations due to structural relaxation, which they found to be substantial for small angles. Koshino *et al.* [179, 180] and Tarnopolsky *et al.* [169] presented a two-parameter version of the original one-parameter Bistritzer–MacDonald model to account for changes due to this structural relaxation. Lin & Tománek [181] presented a simple three-parameter tight-binding model capable of describing graphene, Bernal bilayer graphene, and TBG. More generic models for TBG have been derived *e.g.* by Fang & Kaxiras [175] and Balents [182].

Cao *et al.* [183] performed transport measurements near the magic angle in 2016 and observed the insulating states resulting from the single-particle bandgaps as predicted by Lopes dos Santos *et al.* [132] (full LdS model in Fig. 3.7), Suárez Morell *et al.* [157] and others [175, 184], but whose size is much larger than predicted. Note that many theories [158], including the 2007 Lopes dos Santos *et al.* [134] model, the $3G$ LdS model in Fig. 3.7, and the Bistritzer–MacDonald

model [140] predict no bandgaps at all. By the time Cao *et al.* did not find any further insulating (or superconducting) states.

3.3.2 Effects of strain and structural relaxation

Structural relaxation or strain in general can have a significant effect on both the atomic structure and the (flat) bands. Using STM and STS, Yan *et al.* [149] measured how a wrinkle in TBG affects the dispersion of a 5.1° TBG. Chu *et al.* [150] used a continuum model to show how tensile/compressive strain can be used to control the position of the Dirac cones and flattening of the bands. Uchida *et al.* [155] studied angles between 0.76° and 52° (although the smallest-angle structures were extrapolated) by DFT and found the interlayer distance to increase at the *AA*-stacking regions through structural relaxation. This out-of-plane atomic corrugation increased with a decreasing angle, with the maximum corrugation of 0.3 Å. van Wijk *et al.* [107] used an atomic force-field method for structural relaxation near the magic angle and found significant strain also in-plane. Dai *et al.* [185] presented a continuum model based on the Peierls–Nabarro model and calculated the energetics of dislocations in bilayer systems. Later they [108] applied this to TBG to find the equilibrium structure of TBG at different angles, and identified distinct "breathing" and "bending" modes.

Nam & Koshino [109] relaxed the structures close to the magic angle by a continuum elasticity model containing only one parameter (the binding energy difference between *AA* and *AB* stacking) and found that the *AA*-region sizes are significantly reduced. Using a tight-binding model they then observed that bandgaps are opened close to the magic angle, the flat-band bandwidth is increased at or above the magic angle, and its bandwidth is decreased below the magic angle. The decrease in the *AA*-stacking areas was experimentally confirmed by Yoo *et al.* [111]. Carr *et al.* [176] found similar results with an *ab initio* TB model: both bandgaps and flat-band bandwidth increase as a result of structural relaxation at the magic angle. Fang *et al.* [177] confirmed this with a TB calculation. On the contrary, Lucignano *et al.* [186] found by a DFT calculation that the flat-band bandwidth is decreased, but the bandgaps are formed or increased. In a comprehensive study Guinea & Walet [187] found that both the unrelaxed flat bands and the effect of relaxation on them is largely dependent on the employed model, but a common factor seems to be an increase of bandwidth as a result of relaxation.

Lin *et al.* [110] used continuum elasticity theory together with a TB model and found that magic-angle TBG is unstable to global shear, which is as important as atomic relaxation. The effect is, however, relatively small: the magic angle is changed only by 0.04°. Qiao *et al.* [188] studied the strain-engineering aspect close to the magic angle and showed both experimentally and theoretically how heterostrain (each layer strained independently) can be used to effectively tune the magic angle. Bi *et al.* [189], on the other hand, got the result that heterostrain increases the bandwidth in TBG, but can create flat bands in bilayer transition metal dichalcogenides. Luo [190] went even further and showed how a moiré pattern accompanied by flat bands can be formed in *untwisted* bilayer graphene

by applying "magical" heterostrain.

The above long list of papers suggest that strain effects are indeed significant, and explains why there can be large differences between experiments and theories (such as Lopes dos Santos *et al.* [132, 134] and Bistritzer–MacDonald [140]) including only the unrelaxed structure. Especially the increase in bandwidth and bandgap size after structural relaxation may explain why the experimentally observed values are often much larger than predicted in these simple theories. On the other hand, as shown by Koshino *et al.* [179, 180] and Tarnopolsky *et al.* [169], the structural relaxation effects can be incorporated also in the simple Bistritzer–MacDonald model by adding a second parameter that tunes the sublattice structure of the interlayer coupling.

3.3.3 Effect of pressure

External pressure can be used to decrease the interlayer distance, and thus effectively to increase the interlayer hopping strength and correspondingly the magic angle. Carr *et al.* [176] showed by an *ab initio* TB model that this is indeed the case: while the magic angle is $1.1°$ at zero pressure, for a 10% compression (a pressure of $9\,\mathrm{GPa}$) it is $2°$ and for a 20% compression ($33\,\mathrm{GPa}$) already over $3°$. Chittari *et al.* [161] used a refinement of the Bistritzer–MacDonald model, where the interlayer coupling strength is tuned by a single parameter $t_\perp$. They found that the (first) magic angle is linear in $t_\perp/t$ through

$$\theta^* = 27.8° \left| \frac{t_\perp}{t} \right|, \tag{3.39}$$

where t is the in-plane nearest-neighbor coupling strength. Taking $t = -2.6\,\mathrm{eV}$, this gives $\theta^* = 1.07°$ for the zero-pressure interlayer coupling $t_\perp = 0.1\,\mathrm{eV}$. Doubling $t_\perp$ then gives $\theta^* \approx 2°$, which through a simplified DFT calculation was found to correspond to a pressure of $8\,\mathrm{GPa}$, thus being consistent with the work of Carr *et al.* [176].

The magic-angle tuning by pressure was experimentally verified by Yankowitz *et al.* [191] by using a hydraulic press with the maximum pressure of $2.2\,\mathrm{GPa}$. A separate but closely related study was performed earlier by Yankowitz *et al.* [106], where they experimentally showed that commensurate stacking of graphene on hBN can be induced by pressure applied by an AFM tip. On the other hand, González-Tudela & Cirac [192] proposed how twisted bilayers could be realized in cold-atom experiments to tune not only the interlayer, but also the intralayer coupling.

3.3.4 Incommensurate structures

The incommensurate TBG structures, which include all the angles that are not included in Eq. (3.2), are interesting because they do not allow using the traditional momentum-space approach of condensed matter physics due to the lack of periodicity and thus need a totally different approach. Massatt *et al.* [193], however, showed how the density of states can be calculated as an integral over the Brillouin

zones of the separate (periodic) layers. This theory was generalized by Amorim & Castro [194] to arbitrary incommensurate van der Waals structures, also to calculate the local DOS. Cazeaux *et al.* [195] presented a way to do structural relaxation in incommensurate structures by performing the process in a local configuration space instead of real space. One interesting special case is $\theta = 30°$, which forms a dodecagonal quasicrystal and that has been demonstrated also experimentally by Yao *et al.* [136] and Ahn *et al.* [196]. Park *et al.* [197] showed how in this system electrons get localized in the presence of a sublattice-symmetry-breaking mass term when increasing the interlayer interaction, in contrast to commensurate TBG where the states are delocalized.

3.3.5 Other aspects

Liu *et al.* [198] provided an interesting connection between TBG and PSG: they showed that in the continuum model the position-dependent interlayer coupling can be (partly) written as a pseudo vector potential for both layers separately. But the transformation is not full, as there is still a constant interlayer coupling term left. Ramires & Lado [72], on the other hand, showed that in addition to the magic angle, tiny-angle ($\theta \ll 1°$) TBG is also interesting: in that case a perpendicular electric field is equivalent to a new kind of gauge field, which can be used to form flat bands, thus removing the need for precise angle tuning. Furthermore Carter *et al.* [199] predicted that an interlayer bias produces a new set of magic angles.

Under a magnetic field, the Landau levels of TBG were studied by de Gail *et al.* [158] and Choi *et al.* [200]. Dean *et al.* [201] observed a self-similar Hofstadter's butterfly spectrum in bilayer graphene twisted relative to the underlying hBN substrate. This is usually difficult to measure because it needs a magnetic field of the order of the magnetic flux quantum $\Phi_0 = h/(2e)$ per unit cell, and in usual materials with a lattice constant of a few ångströms this is over $10\,000\,\mathrm{T}$ [201]. But when the lattice constant is replaced by the superlattice constant of the order of $10\,\mathrm{nm}$, a magnetic field of the order of $10\,\mathrm{T}$ is already enough. Kim *et al.* [202] performed magnetotransport measurements in TBG when the chemical potential crosses the van Hove singularity and observed a topological phase transition.

He *et al.* [151] showed that Klein tunneling in TBG is tunable between perfect tunneling and perfect reflection. The optical properties were studied by Moon & Koshino [184] and Le & Do [203]. The topological properties of TBG were studied for example by San-Jose *et al.* [204], Tong *et al.* [205], and Hejazi *et al.* [163]. Linear response theory for electric current was studied by Stauber *et al.* [206, 207]. Regarding the dynamical behavior of the electrons, Do *et al.* [208] showed that electrons undergo an oscillatory motion between the layers.

3.3.6 Similar systems

Nowadays the focus is slowly starting to move towards other (twisted) multilayer van der Waals materials. Let me start from bilayers. Marchenko *et al.* [209] observed by angle-resolved photoemission spectroscopy (ARPES) that also (untwisted) bilayer graphene exhibits flat bands when relatively biasing the sublattices of the other layer. This was also explained by DFT and continuum tight-binding

calculations. The important point is that the flat band is not formed at the Fermi energy, but $0.2\,\mathrm{eV}$ higher.

Xian *et al.* [210] performed DFT calculations to show that twisted bilayer boron nitride (TBBN) hosts monotonically decreasing bandwidth with decreasing angle, having thus no magic angles. Because boron nitride (BN) is an insulator with an energy gap of $\sim 4\,\mathrm{eV}$ [210], the flattening of the bands may be at least partially a strain effect, in which case a small generalization (including the chemical potential difference between the sublattices in BN) of the PSG model in Publication I might explain the origin of the flat bands, and could be also utilized to predict superconductivity in those bands. Due to the large bandgap the flat bands are, however, not close to the Fermi energy but $2\,\mathrm{eV}$ away from it, so reaching them needs heavy chemical doping.

Staying in bilayers, Kariyado & Vishwanath [211] predicted by a continuum theory how flat bands are formed in generic twisted bilayer Bravais lattices. Crasto de Lima *et al.* [212] used a tight-binding model to predict both high- and low-energy flat bands in a twisted magic-angle bilayer kagomé crystal, where the high-energy flat bands come from the kagomé crystal and the low-energy flat bands from the twist.

Three layers allow having two separate moiré patterns which work together to form a new, larger moiré pattern. Wang *et al.* [213] verified this experimentally in a hBN/graphene/hBN system. Amorim & Castro [194] applied their generic momentum-space theory of incommensurate van der Waals structures to incommensurate twisted trilayer graphene, and found that the hybridization between the first and third layers is also significant. Mora *et al.* [214] generalized the Bistritzer–MacDonald model to study trilayer graphene with arbitrary twists and found flat bands at certain magic angles. Using symmetry arguments they also showed that the system is a metal at all energies (up to the lowest and highest bands), similarly as graphene but contrasting (twisted) bilayer graphene. This happens because the Dirac cones can only gap in pairs.

Li *et al.* [215] also used a generalized Bistritzer–MacDonald model to study trilayer graphene where one layer is twisted and found both flat and Dirac-like bands form strongly depends on the initial stacking configuration (AAA, ABA, or ABC) and on which layer is twisted. Using a simplified model with fewer bands they predicted that an "optimal" configuration is the one where every second layer is rotated with respect to the initial Bernal configuration, in the sense that then both the flat and Dirac-like bands are well formed. For this configuration they predicted the magic-angle sequences for 3, 5, 7, and 9 layers. For example the trilayer has two sequences of magic angles, one identical to TBG and one where the TBG sequence is multiplied by $\sqrt{3}$. The optimal stacking configuration is very special, but by using atomistic simulations Carr *et al.* [216] showed that the trilayer always relaxes to it when the twisting angle is fixed. Using *ab initio* tight-binding calculations they also obtained similar bandstructures to those of Li *et al.* [215]. Using a low-energy model Tsai *et al.* [217] obtained a similar bandstructure.

Four layers allow even more freedom. When the system consists of two bilayers twisted relative to each other, it is often called twisted double bilayer graphene

(TDBG). Using a tight-binding approach, Choi & Choi [218] found the first magic angle of TDBG at $1.25°$, with a similar bandwidth as in TBG, but being a much slower function of θ. Lee *et al.* [219] showed that a two-parameter Bistritzer–MacDonald model predicts a magic angle of $1.05°$ with a bandwidth of $0.25\,\mathrm{meV}$, but including further hoppings dramatically increases the bandwidth to $15\,\mathrm{meV}$. We see a similar increase in bandwidth in Fig. 3.7 in the case of TBG when including higher Fourier components of the interlayer coupling (full vs the $3G$ LdS model). Similar results were calculated by Koshino [180], who also showed that the topological properties are strongly influenced by the stacking of the bilayers.

Also twisted n-layer graphene has been studied. Khalaf *et al.* [164] showed that twisted n-layer graphene with alternating $\pm\theta$ rotations can be mapped to a set of decoupled twisted bilayers with different angles. Using this result they showed that in the twisted-graphite limit $n \to \infty$ there is a continuum of magic angles for $\theta \lesssim 2°$. Cea *et al.* [220] used a continuum model to find flat bands at small angles in three systems: (i) one rotated layer of graphene on top of untwisted graphite, (ii) graphite with a constant rotation between all the layers, and (iii) two untwisted graphite stacks rotated with respect to each other. These studies might be relevant in explaining the experiments of Scheike *et al.* [221, 222] and Ballestar *et al.* [223, 224], where high-temperature superconductivity was observed in twisted interfaces of graphite [8, 225].

3.4   BCS equations for the superconducting state

Let us move to superconductivity. Adding the assumption of intralayer interaction to the interaction (2.34) chosen for PSG, we arrive at similar equations for the superconducting state, except that now we have also the layer index (which do not couple). Explicitly, the Bogoliubov–de Gennes equation becomes

$$\begin{pmatrix} \mathcal{H}^\rho(\boldsymbol{r}) & \Delta(\boldsymbol{r}) \\ \Delta^*(\boldsymbol{r}) & -\mathcal{H}^\rho(\boldsymbol{r}) \end{pmatrix} \begin{pmatrix} u_{\rho n}(\boldsymbol{r}) \\ v_{\rho n}(\boldsymbol{r}) \end{pmatrix} = E_{\rho n} \begin{pmatrix} u_{\rho n}(\boldsymbol{r}) \\ v_{\rho n}(\boldsymbol{r}) \end{pmatrix} \tag{3.40}$$

and the self-consistency equation becomes

$$\Delta_{i\alpha}(\boldsymbol{r}) = -\frac{\lambda}{V} \sum_{\rho n_+} u_{\rho n_+, i\alpha}(\boldsymbol{r}) v^*_{\rho n_+, i\alpha}(\boldsymbol{r}) \tanh\left(\frac{E_{\rho n_+}}{2k_\mathrm{B}T}\right). \tag{3.41}$$

Now in addition to the sublattice structure, we have also the layer structure as

$$\Delta \coloneqq \mathrm{diag}(\Delta_{1A}, \Delta_{1B}, \Delta_{2A}, \Delta_{2B}), \quad u_{\rho n} \coloneqq (u_{\rho n, 1A}, u_{\rho n, 1B}, u_{\rho n, 2A}, u_{\rho n, 2B})^\mathsf{T} \tag{3.42}$$

and similarly for $v_{\rho n}$. Here the noninteracting Hamiltonian $\mathcal{H}^\rho$ is defined in Eq. (3.29) and $n_+ = (b_+, \boldsymbol{k})$ is summed over the positive-energy bands b_+ and the SBZ $\boldsymbol{k} \in L^*_\mathrm{BK}/SL^*$.

While in Publication II the chemical potential μ is used to describe doping, in Publication III we chose to use the filling[9]

$$\nu \coloneqq nA_\mathrm{moiré} - 4 \tag{3.43}$$

---

[9]The convention of filling is ambiguous; its scale depends whether the spin and valley

instead because it is approximately the one controlled in experiments by the gate voltage. Here n is the number density of electrons, $A_{\text{moiré}} \coloneqq |\boldsymbol{t}_1 \times \boldsymbol{t}_2|$ is the area of the superlattice unit cell, and -4 is used to fix the scale such that $\nu = 0$ corresponds to charge neutrality. The number density of electrons can be calculated from the thermal average

$$n \coloneqq \sum_{\sigma\rho} \int \mathrm{d}\boldsymbol{r} \left\langle \psi_{\sigma\rho}^\dagger(\boldsymbol{r})\psi_{\sigma\rho}(\boldsymbol{r}) \right\rangle \tag{3.44}$$

$$= \frac{2}{V} \sum_{\rho n_+} \int \mathrm{d}\boldsymbol{r} \left[u_{\rho n_+}^\dagger(\boldsymbol{r})u_{\rho n_+}(\boldsymbol{r})f(E_{\rho n_+}) + v_{\rho n_+}^\dagger(\boldsymbol{r})v_{\rho n_+}(\boldsymbol{r})(1 - f(E_{\rho n_+})) \right],$$

$$\tag{3.45}$$

where in the second equality we used the inverse Bogoliubov transformation (2.57) (with the w and γ replaced by those of TBG), the symmetries (2.49) and (2.50) to transform all the Bogoliubons to the positive-energy ones, and the expectation values (2.55) and (2.56).

The Fourier-space versions of the above equations are similar to those of PSG in Sec. 2.3.3, but now written only in the reduced zone scheme $\boldsymbol{G} \in SL_S^* = SL_{\text{RZ}} = SL^*$. The superfluid weight D^{s} is calculated from Eq. (2.93) but with the same replacements as above.

3.5 Electronic properties of the superconducting state

The experimental observation of correlated insulator states [130] and superconductivity [9] (Fig. 3.10) in magic-angle TBG ignited a tremendous rise of interest in this system. This is understandable, as the measured phase diagram is highly reminiscent of the high-temperature cuprate superconductors [229], but the material is much simpler, including only carbon. Moreover, simple electrical gating is enough to move between the different phases. Despite the similarity to high-temperature superconductors, unfortunately the critical temperature of magic-angle TBG turned out to be only of the order of one kelvin. Nevertheless, cracking the mechanism behind the complex phase diagram of TBG might shed light also on the mechanism behind the high-temperature superconductors. Moreover, as TBG is relatively easy to manufacture and model and its correlated phases are easily tunable, it is an interesting new material already on its own. Whether it is the first carbon-only superconductor can be argued, as similar experiments in graphite (with twisted interfaces) were done already several years ago [221–224], with critical temperatures exceeding $100\,\mathrm{K}$.

---

degeneracy is included or not. We and some others [9, 226, 227] use the scale $\nu \in \{-4, \ldots, 4\}$, where 0 means charge neutrality, 4 means all flat bands are full with electrons, and -4 means all flat bands are full with holes. In this case $\nu = \pm 2$ is often called half-filling. Some [191, 228] use the scale $\{-1, \ldots, 1\}$ instead, in which case $\pm 1/2$ is called half-filling. In both cases the filling 0 is the charge neutrality. Whenever I talk about filling, I mean the scale from -4 to 4, regardless of the article I am discussing.
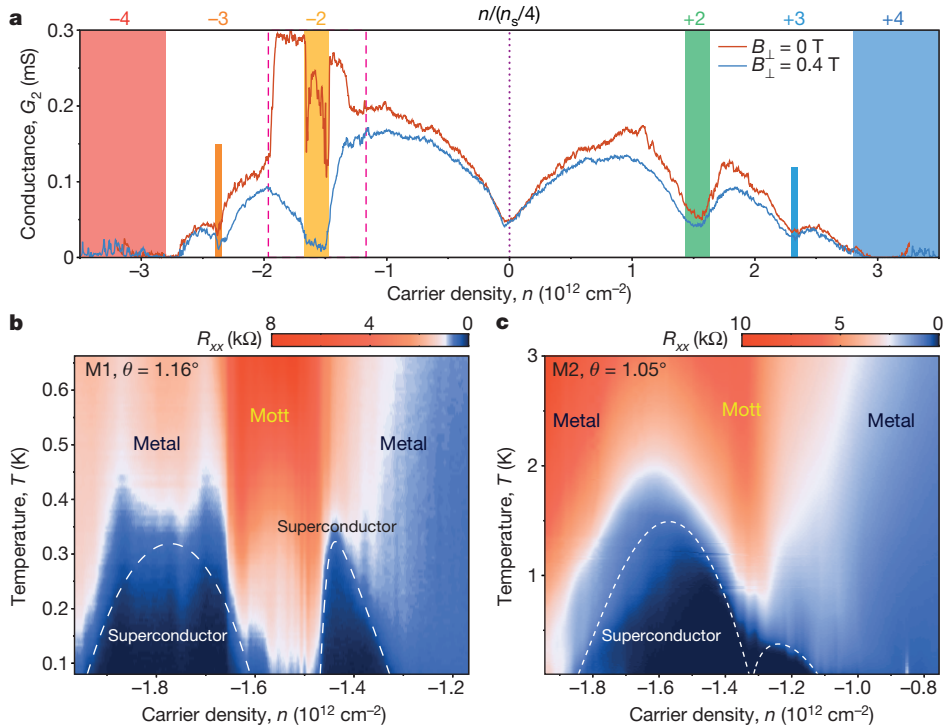
Figure 3.10: Two-terminal conductance and phase diagram measured by Cao *et al.* [9] in the breakthrough experiment, where electrically doped magic-angle TBG was shown to be superconducting. (a) Two-terminal conductance as a function of the carrier density n, which is calculated from the measured gate voltage V_g by $n = CV_g/e$, where C is the gate capacitance per unit area and e the electron charge. The insulating states at $\nu = \pm 4$ fillings (upper horizontal axis) show the bandgaps between the flat and higher bands, as shown in Fig. 3.7 within the full LdS model. (b,c) Measured phase diagram (four-terminal resistance as a function of temperature and carrier density) of two samples ($\theta = 1.16°$ and $\theta = 1.05°$) in the hole-doped region bordered by the red dashed lines in (a). The measurements reveal (correlated) insulating behavior at the integer fillings $\nu = \pm 2$, while superconductivity is observed only nearby the $\nu = -2$ filling. Reprinted with permission from Ref. [9]. Copyright 2018 Springer Nature.

In Publication II our aim was to show that the simple and well-known (slightly generalized) BCS theory with *s*-wave spin-singlet pairing, as defined by the BdG equation (3.40) and the self-consistency equation (3.41), is enough to explain (parts of) the observed *superconducting* behavior. Effectively we wanted to explain the results of Fig. 3.10: appearance of superconductivity at the magic angle close to the filling $\nu = -2$ with the critical temperature[10] $T_c \approx 1\,\text{K}$, and the vanishing

---

[10]In experiments the "critical temperature" T_c often denotes the temperature where the resistance has dropped to half of the "normal-state resistance". This is not a very accurate
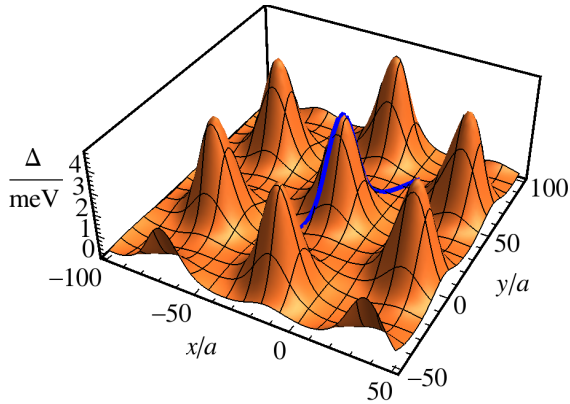
Figure 3.11: Position dependence of a typical self-consistent superconducting order parameter $\Delta = \Delta_{i\alpha}$ at the magic angle, as calculated from the self-consistency equation (3.41) within the $3G$ LdS model of Publication II. The blue line is the same as shown in Fig. 3.9a, meaning that Δ is localized around the AA stacking regions. Reprinted with permission from Ref. [173]. Copyright 2018 American Physical Society. Modified by updating the label of Δ.

of superconductivity away from the magic angle. But the results were limited to those concerning the behavior of Δ (or T_c calculated from it), which we know does not properly describe the phase transition in a 2D material, as discussed in Sec. 1.3. Thus in Publication III we expanded these results with the calculation of the superfluid weight D^s and the BKT transition temperature T_{BKT}. In both publications we concentrated only on the $r = 1$ structures because they have the smallest unit cells.

3.5.1   Order parameter and mean-field critical temperature

Let us first go to Publication II. Solving the self-consistency equation (3.41) shows [Fig. 3.11] how the order parameter[11] $\Delta := \Delta_{i\alpha}$ is localized around the AA-stacking regions, similarly as the local density of states in Fig. 3.9b. This is the same kind of behavior as what happens in PSG: localization of the low-energy states determines the localization of Δ. But while in PSG the sublattice symmetry of both the LDOS and Δ is broken by the pseudo vector potential $\boldsymbol{A}$, in TBG we only observe sublattice- and layer-symmetric LDOS and Δ.

Similarly to PSG, $\max \Delta$ (maximum over $\boldsymbol{r}$) is a number that well describes

---

definition, and is neither the T_c calculated from the vanishing of Δ nor the BKT transition temperature T_{BKT}. However, according to studies where the BKT transition temperature has been determined [9, 127], it is of the same order of magnitude as the measured T_c. For example Cao $et\ al.$ [9] cites $T_c = 1.7\,\mathrm{K}$ and $T_{BKT} = 1\,\mathrm{K}$.

[11]We use an initial guess where $\Delta_{1A}(\boldsymbol{r}) = \Delta_{1B}(\boldsymbol{r}) = \Delta_{2A}(\boldsymbol{r}) = \Delta_{2B}(\boldsymbol{r}) = \mathrm{const}$, which leads to a self-consistent solution with $\Delta_{1A} = \Delta_{1B} = \Delta_{2A} = \Delta_{2B}$. As was shown in the Supplementary Material of Publication I in the case of PSG, another kinds of initial guesses lead to solutions with higher energy. We expect the same to happen in TBG.
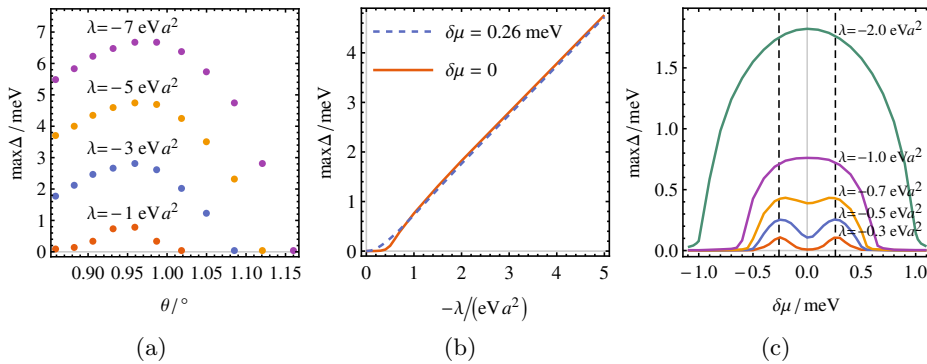
Figure 3.12: Maximum of the zero-temperature superconducting order parameter within the $3G$ LdS model of Publication II as a function of (a) the rotation angle θ at $\delta\mu = 0$, (b) the interaction strength λ at the magic angle, and (c) the chemical potential/doping $\delta\mu$ away from the charge neutrality μ_0 (*i.e.* the chemical potential entering the Hamiltonian is $\mu = \mu_0 + \delta\mu$) at the magic angle. In (c) the dashed lines denote the locations of the DOS peaks of Fig. 3.8a at $\delta\mu = \pm 0.26$ meV. Reprinted with permission from Ref. [173]. Copyright 2018 American Physical Society. Modified by updating the labels and changing the sign of λ such that $\lambda < 0$ corresponds to attractive interaction.

superconductivity in the flat-band regime, meaning that for example an approximately linear relation $T_c \propto \max\Delta$ holds there. The rest of the most important results in Publication II can be summarized by the three plots in Fig. 3.12, showing $\max\Delta$ as a function of the rotation angle θ, the interaction strength[12] λ, and the chemical potential/doping $\delta\mu$ away from the charge neutrality μ_0.

The calculated $\max\Delta(\theta)$ (Fig. 3.12a) confirms that Δ is maximized at the magic angle ($0.96°$ for the $3G$ LdS model) and is suppressed with both higher and lower angles. The experiment of Cao *et al.* [9] has only two different rotation angles, so Fig. 3.12a serves also as a prediction of the functional dependence. Figure 3.12a also shows that at the magic angle $\max\Delta$ is a linear function of the interaction strength λ. This is confirmed in Fig. 3.12b: the dependence is very linear when $\max\Delta$ exceeds 1 meV, the bandwidth in the $3G$ LdS model (*i.e.* when $|\lambda| \gtrsim 1$ eVa^2). Similarly to PSG, this is called the *flat-band regime*, because in this regime the details of the dispersion do not matter and the results [Eqs.(1.50)–(1.52)] [7, 8, 129] for exact flat bands should be valid, at least in some sense. The problem is that the cited results are derived for a homogeneous Δ_{FB}, whereas now Δ is position dependent. But, as in PSG in Publication I, we get similar linear relations if the homogeneous Δ_{FB} in those results is replaced by $\max\Delta$ of the position-dependent order parameter. For example a linear fit to the linear regime of Fig. 3.12a gives at the magic angle

$$\max\Delta = -0.2\,\mathrm{meV} - 1.0 \times 10^{-3}\lambda/a^2 \quad \mathrm{at} \quad \delta\mu = 0, T = 0, \qquad (3.46)$$

---

[12]In Publication II we had the opposite convention that $\lambda > 0$ corresponds to attractive interaction.

which is similar to the PSG result (2.94). On the other hand from Eq. (1.50) we get for 8 exactly flat bands (2 from valleys, 2 from the two bands of graphene, and 2 from the positive/negative BdG bands) with a homogeneous Δ_{FB} [7, 8] [Supplementary Material of Publication II] $\Delta_{\mathrm{FB}} = -\lambda \Omega_{\mathrm{FB}}/\pi^2 \approx -1.3 \times 10^{-3} \lambda/a^2$ at the magic angle. Here Ω_{FB} is the area of the flat bands, which is approximately the SBZ area $\Omega_{\mathrm{FB}} \approx \Omega = |\boldsymbol{G}_1 \times \boldsymbol{G}_2| = 8\pi^2/(\sqrt{3}d^2)$. This flat-band formula also predicts that a higher Δ (*i.e.* higher T_{c}) could by manufactured by increasing the SBZ area Ω, *i.e.* decreasing the superlattice constant d while still keeping the flatness of the bands. Indeed, by applying external pressure an increase of T_{c} in a higher-magic-angle (smaller d) TBG has been experimentally demonstrated [191].

On the other hand in the *dispersive regime*, where $\max \Delta$ is below the bandwidth (*i.e.* $|\lambda| \lesssim 1\,\mathrm{eV}a^2$), Δ "sees" the complicated double-peak structure in the DOS and vanishes below $|\lambda| \approx 0.5\,\mathrm{eV}a^2$ if the doping level is between the DOS peaks at $\delta\mu = 0$. Doping to the DOS peak at $\delta\mu = 0.26\,\mathrm{meV}$ prolongs the linear flat-band regime to much smaller interactions strengths.

The same critical behavior at $|\lambda| \approx 1\,\mathrm{eV}a^2$ is seen in the doping dependence in Fig. 3.12c: In the flat-band regime Δ "sees" only one large flat band and thus decreases monotonically with increasing $\delta\mu$ with the critical chemical potential

$$\delta\mu_{\mathrm{c}} \approx 0.6 \max \Delta(\delta\mu = 0) \quad \text{at} \quad T = 0 \tag{3.47}$$

at the magic angle. This again corresponds to the similar result (2.95) for PSG or to the exact-flat-band result (1.52) with a homogeneous Δ_{FB}. On the other hand in the dispersive regime Δ "sees" the two peaks of the DOS at $\delta\mu = 0.26\,\mathrm{meV}$ (dashed lines), and obtains its maximum there. Further calculating the filling from Eq. (S8) in the Supplementary Material of Publication II tells us that in the noninteracting state $\delta\mu = \pm 0.26\,\mathrm{meV}$ corresponds to the filling $\nu = \pm 2$. This very much resembles the experimental phase diagram in Fig. 3.10, except that our model disregards the correlated insulating states at $\nu = \pm 2$ and is electron–hole symmetric thus predicting superconductivity also at $\nu = +2$. While superconductivity close to $\nu = +2$ was not observed by Cao *et al.* [9], it has been later observed in other experiments [127, 191, 230], so our prediction of the symmetric doping-behavior is not really a drawback of our model. The nonmonotonic behavior suggests that the interaction strength should be in the dispersive regime $|\lambda| \lesssim 1\,\mathrm{eV}a^2$. In the view of competing phases, the correlated insulating phases at $\nu = \pm 2$ would then cut the superconducting domes in half as in Fig. 3.10. But the phase at the charge neutrality $\nu = 0$ is a normal *band insulator* instead [9, 191], so this suggests that we should actually choose $\lambda \approx -0.3\,\mathrm{eV}a^2$.

Above I discuss only the superconducting order parameter Δ, which cannot directly be compared to the experiments since it is not an observable. It creates a gap, however, as shown in Fig. 2 of Publication II, so measuring this gap by STS would create knowledge about Δ. Another observable is the critical temperature T_{c}, which in the mean-field level can be calculated from the self-consistency equation as the temperature where Δ vanishes. The temperature dependence is shown in Fig. 4 of Publication II, which shows that we get an approximately linear relation

$$k_{\mathrm{B}}T_{\mathrm{c}} \approx 0.25 \max \Delta(T = 0) \quad \text{at} \quad \delta\mu = 0 \tag{3.48}$$

in the flat-band regime at the magic angle. This is similar to the PSG result (2.96) and to the exact-flat-band result (1.51) with a homogeneous Δ_{FB}. Applying Eqs. (3.46) and (3.48) for $\lambda = -1\,\mathrm{eV}a^2$ gives $T_{\mathrm{c}} \approx 2\,\mathrm{K}$, which is in the same order of magnitude as the experiment of Cao *et al.* [9]. This is only a rough estimate, as above I argued that a more realistic value for λ is $-0.3\,\mathrm{eV}a^2$, and for this small λ Eq. (3.48) cannot be employed. However, combining this rough knowledge of T_{c} with the nonmonotonic behavior of $\Delta(\delta\mu)$ convinces us that our model is consistent with the experiment of Cao *et al.* [9].

3.5.2 Superfluid weight and BKT transition temperature

When comparing to experiments one has to remember that the T_{c} calculated from the vanishing of Δ does not correspond to the true superconducting transition measured experimentally, which is described by the BKT physics instead. The mean-field T_{c} gives only an upper bound for T_{BKT}. But with the knowledge of Δ also the superfluid weight D^{s} and the BKT transition temperature T_{BKT} can be calculated through Eqs. (2.93) and (1.65). This we did in Publication III in the case of the magic angle.

To check whether the results are sensitive to the employed model, in Publication III we used two different models for the noninteracting state: (i) the continuum full LdS model discussed in Sec. 3.2, called the *Dirac point approximation* (DP), and (ii) a tight-binding model, called the *renormalized moiré* (RM) approach, where the hoppings and lattice constants are scaled so that the magic angle increases and thus the superlattice unit cell decreases, rendering the numerical calculations much faster. These give roughly the same dispersion (Fig. 1 in Publication III) apart from the small electron–hole asymmetry of the RM method, which is amplified by the renormalization procedure. On the other hand in the superconducting state we used the same local interaction as in Publication II. But because the true form of the interaction is not known, in the case of the RM method we employed also another type of interaction that is popular in the graphene literature [47, 48, 231–234]: the nonlocal, nearest-neighbor *resonating valence bond* (RVB) interaction [229, 235]. This way we could show that the form of the interaction has an effect on the observable properties: especially the superfluid weight is either isotropic (local interaction) or anisotropic (RVB interaction), which could be used to determine experimentally the form of the interaction and thus its origin. Below I concentrate on the results calculated with the DP method and the local interaction, and discuss more briefly the results of the RM calculations.

The DP results of Publication III can be summarized in Fig. 3.13: order parameter Δ at $T = 0$, superfluid weight D^{s} at $T = 0$, and BKT transition temperature T_{BKT} as a function of the interaction strength λ and doping[13] $\delta\mu$. Because in Publication III we employed the full LdS model instead of the $3G$ LdS model in Publication II, I start by presenting $\max\Delta$ as a function of λ and $\delta\mu$ in

---

[13] In Publication III we actually present the doping behavior as a function of the filling ν, as calculated from Eq. (3.45), instead of the chemical potential $\delta\mu$. But for an easier comparison to Publication II and Fig. 3.12c I present the doping behavior here as a function of $\delta\mu$.

Figs. 3.13a and 3.13d. The qualitative behavior is very similar to that of the $3G$ model in Fig. 3.12: linear $\max \Delta(\lambda)$ in the flat band regime and a qualitatively different doping behavior in the flat-band and dispersive regimes. The critical $\max \Delta$ separating the dispersive and flat band regimes is still around $1\,\mathrm{meV}$ (the separation of the DOS peaks), which now corresponds to $\lambda \approx -2\,\mathrm{eV}a^2$. But while in Fig. 3.12b I show $\max \Delta(\lambda)$ for two *fixed chemical potentials*, in Fig. 3.13a I show it for two *fixed fillings*, as calculated from Eq. (3.43). For small λ the filling $\nu = -2$ corresponds to the DOS peak and thus slightly enhances Δ compared to $\nu = 0$, while for large λ the filling $\nu = -2$ does not anymore correspond to the DOS peak and thus suppresses Δ compared to $\nu = 0$. With the local interaction the behavior is similar with the RM method (Figs. 2c–d in Publication III), whereas the RVB interaction has a two times larger critical interaction strength needed to render Δ finite. The largest difference regarding the order parameter is in the pairing symmetry: the local interaction produces isotropic s-wave pairing, whereas the RVB yields $(d + \mathrm{i}d) + (p + \mathrm{i}p)$-wave pairing in the dispersive regime and a C_3-symmetry-breaking $s + p + d$-wave pairing in the flat-band regime. Both pair electrons in the spin-singlet channel.

The superfluid weight (Fig. 3.13b) is also linear in the flat-band regime, but, surprisingly, is already quite large even for weak interactions. This is in contrast to the PSG result in Fig. 2.9a. The RM method yields quantitatively the same result (Fig. 2c in Publication III). Perhaps a more interesting result is the fact that for the local interaction D^{s} is always isotropic, $D^{\mathrm{s}}_{xx} = D^{\mathrm{s}}_{yy}$ and $D^{\mathrm{s}}_{xy} = 0 = D^{\mathrm{s}}_{yx}$, as in the case of PSG with the 2D potential. This is contrasted with the RVB interaction (Fig. 2e in Publication III), which has an anisotropic superfluid weight in the flat-band regime. The RVB interaction thus breaks the C_3 symmetry of the superlattice, which the local interaction conserves, so this difference could serve as an important experimental fingerprint when trying to determine the origin of superconductivity. As discussed in Publication III, this anisotropy could be measured by radio-frequency impedance spectroscopy [236] in a Hall-like four-probe setup.

Calculating then the superfluid weight as a function of temperature allows to obtain the BKT transition temperature T_{BKT} through Eq. (1.65). This I show in Figs. 3.13c and 3.13f, which tell that the behavior of T_{BKT} is very similar to $\max \Delta$ apart from a different energy scale. Calculating the ratio $k_{\mathrm{B}} T_{\mathrm{BKT}} / \max \Delta(T = 0)$ (Fig. 3b in Publication III) tells that they are indeed related: in the flat-band regime we get the linear relation

$$k_{\mathrm{B}} T_{\mathrm{BKT}} = \kappa \max \Delta(T = 0) \quad \text{at} \quad \nu = -2 \qquad (3.49)$$

with $\kappa \approx 0.16 \ldots 0.2$, which holds also for the RM model with the local interaction. For the RVB interaction the proportionality factor becomes $\kappa \approx 0.09 \ldots 0.15$ instead. Since T_{BKT} can be experimentally determined by conductivity measurements and Δ by measuring the gap in the (local) density of states, this difference in the proportionality factors could also be used to determine the superconducting pairing. This, however, would need more quantitative results on the Δ dependence of the (local) energy gap, which we did not determine in Publication II or III. Equation 3.49 is again analogous to the one in PSG [Eq. (2.98)], where the
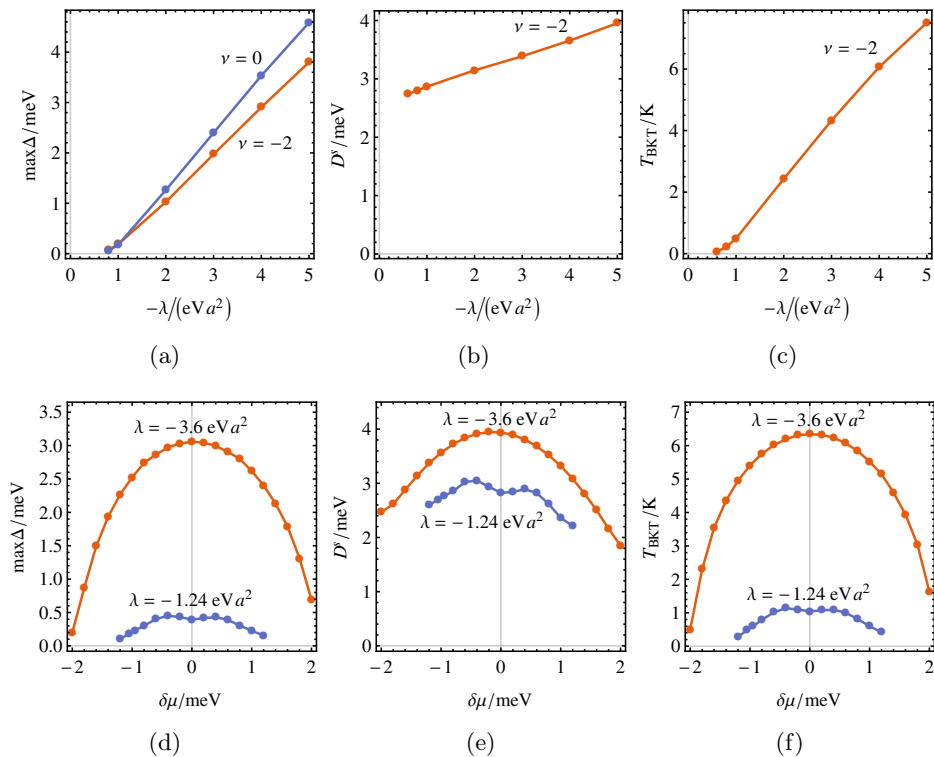
Figure 3.13: The Dirac point (DP) method (*i.e.* the full LdS model) results of Publication III using the local interaction. The superconducting order parameter max Δ at $T = 0$, the superfluid weight $D^{\mathrm{s}} = \sqrt{\det D^{\mathrm{s}}}$ (which is isotropic for the local interaction) at $T = 0$, and the BKT transition temperature T_{BKT} as a function of (a–c) the interaction strength λ and (d–f) doping $\delta\mu$. All the results are calculated at the magic angle. In (a–c) there are missing values for weak interactions due to slow convergence, but every curve vanishes at $\lambda = 0$. In (d–f) the values $\lambda = -3.6\,\mathrm{eV}a^2$ and $-1.24\,\mathrm{eV}a^2$ are chosen to yield max $\Delta \approx 3\,\mathrm{meV}$ and $0.4\,\mathrm{meV}$ at $\nu = \delta\mu = 0$, respectively, so that comparison to the renormalized moiré (RM) method in Publication III can be made.

proportionality factor is $0.05 \ldots 0.15$, depending on the pseudo vector potential. Finally combining it with Eq. (3.48) (which holds also for the full LdS model) gives the ratio $T_{\mathrm{BKT}}/T_{\mathrm{c}} \approx 0.6 \ldots 0.8$, which tells that the results of Publication II are, in fact, not that far from being "correct": in the flat-band regime the true (BKT) transition temperature is "only" some $20 \ldots 40\,\%$ lower than the calculated T_{c}. At the same time the qualitative behavior stays the same.

In the experiments the measured BKT transition temperature is roughly $T_{\mathrm{BKT}} \approx 1\,\mathrm{K}$ [9, 127]. According to Figs. 3.13c and 3.13f the value[14] $\lambda \approx -1.2\,\mathrm{eV}a^2$

---

[14]Note that this value of λ cannot be directly compared with the one in Sec. 3.5.1 as they are calculated with different models (full vs 3G LdS).

best matches this. Note that when including the Coulomb repulsion, this has to be interpreted as the effective interaction λ_{eff} [Eq. (2.102)] instead, as discussed in Sec. 2.5. This value is right at the interface between the dispersive and flat-band regimes. It is remarkable in the sense that flat-band physics already start emerging: Δ starts to become linear in λ and thus T_{BKT} can be much higher than could be anticipated from the simple "standard" BCS estimate (1.35) or the bare graphene estimates (1.43) and (1.45). On the other hand it is not large enough that the simple flat-band results (3.46)–(3.49) could be used. Luckily, there are also new experiments [191, 230] emerging that report T_{c} values much higher than in the Cao *et al.* [9] experiments (although they did not determine T_{BKT}), so that perhaps reaching the flat-band regime is possible. If one could truly go to the flat-band regime, the resulting T_{BKT} values would quickly start increasing, as the linear relations (3.46) and (3.49) become valid.

There are some aspects in the doping dependence of T_{BKT} which require further attention. For the filling dependence and comparison to the RM method with the two different interactions see Figs. 3c–d in Publication III. It is notable that in the flat-band regime ($\lambda = -3.6\,\text{eV}a^2$) the DP and RM method yield practically indistinguishable results, which is understandable by the fact that $\max \Delta$ exceeds the flat-band bandwidth of $\approx 1\,\text{meV}$ and thus the small differences in the bandstructure are not visible. In the dispersive regime ($\lambda = -1.24\,\text{eV}a^2$), however, $\max \Delta \lesssim 0.5\,\text{meV}$ is smaller than the bandwidth, and thus the small differences are visible. This highlights the fact that in the dispersive regime the TBG physics is very sensitive to the details of the equipped model. Especially the electron–hole asymmetry in the RM model is visible as a higher T_{BKT} in the hole-doped side, whereas the DP model yields an almost symmetric doping dependence. It has to be noted that the relatively high asymmetry in RM emerges from the renormalization procedure, which is unphysical. However, as discussed in Sec. 3.3, similar asymmetry is present in many models, and this might explain why T_{BKT} in experiments [9, 127, 191, 230] is often higher in the hole-doped side. A second aspect it that although the value $\lambda \approx -1.2\,\text{eV}a^2$ does not directly yield the two-dome structure often measured in experiments [191, 230], T_{BKT} is highest at $\nu = \pm 2$, which is consistent with them. The correlated insulating states may, however, change this picture a lot.

Publication III includes a few other significant results, calculated with the RM method, that require mentioning. Figure 4 in Publication III shows the separation of the superfluid weight to the *conventional* and *geometric* parts, where the former is proportional to the inverse effective mass and the latter depends also on the eigenstate overlap between different bands [93]. For perfectly flat bands the conventional contribution is thus zero, and a finite D^{s} must then be a geometric effect [92]. Indeed, Fig. 4 in Publication III shows that in the flat-band regime the geometric term becomes significant or dominating, while in the dispersive regime the main contribution comes from the conventional term, which is nonzero because the flat bands in TBG are not perfectly flat. This implies that superconductivity in TBG has at least partly *a geometric origin*. A second conclusion from the same figure is the fact that also the first dispersive bands have to be included if one wants to obtain the correct superfluid weight. This is not because the intraband

terms of the dispersive bands would be important, but because of the interband terms between the flat and dispersive bands. This is an important result, as many studies simply ignore the dispersive bands [162, 170].

3.5.3 Other experiments in TBG and related systems

The discovery of superconductivity and correlated insulating states [9, 130] in TBG have ignited a lot of related experiments confirming the results and also extending them. In the previous section I already briefly mentioned some of them, but let us discuss them in more detail here. Yankowitz et $al.$ [191] used hydrostatic pressure (of the order of a gigapascal) to decrease the interlayer separation, thus increasing the interlayer coupling, and thus increasing the magic angle. They observed both correlated insulating and superconducting behavior at a higher angle of $\theta = 1.27°$, and saw superconductivity also close to the filling $\nu = 2$ (in addition to $\nu = -2$), which Cao et $al.$ [9] did not observe. Going to a higher magic angle also increased the superconducting critical temperature to $T_c \approx 3\,\mathrm{K}$.

Before new superconductivity experiments were published, other correlated phases were found. Sharpe et $al.$ [228] observed ferromagnetism at $\nu = 3$ which, however, was attributed (though mentioned only briefly in the Supplementary Material) to the alignment between graphene and hBN. This effect was theoretically explained by Zhang et $al.$ [237] and Bultinck et $al.$ [238] by noting that the alignment between TBG and hBN breaks C_2 symmetry and the aligned hBN creates a staggering potential for the other graphene layer. Note that the hBN layer(s) are present in the vast majority of experiments, so their role can be important also in the superconductivity experiments. The problem is that usually the TBG's orientation relative to the hBN layer(s) is not known. Cao et $al.$ [226], on the other hand, observed strange-metal behavior (Planckian metal; resistivity linear in temperature) at $\nu = \pm 2$ fillings when temperature is higher than the critical temperature of the superconducting and correlated insulating phases. Polshyn et $al.$ [239] observed similar linear-in-T resistivity and attributed this to electron–phonon scattering.

Then new superconductivity and correlated insulator experiments started appearing. Lu et $al.$ [227] used a refined manufacturing method to produce cleaner (more uniform twist angle) TBG samples, and following this observed the phase diagram in Fig. 3.14: band insulators at $\nu = 0, \pm 4$, correlated insulators at all $integer$ $fillings$ $\nu = \pm 1, \pm 2, \pm 3$, and also superconducting states close to many of these fillings. The BKT transition temperature of the superconducting dome at $n \approx 0.5 \times 10^{12}\,\mathrm{cm}^{-2}$ was determined to be only $T_{\mathrm{BKT}} = 110\,\mathrm{mK}$, so according to Fig. 3.13c this suggests that in our (full LdS) model the interaction strength should be chosen as $\lambda \approx -0.7\,\mathrm{eV}a^2$. For this λ the phase diagram is also consistent with the calculated doping dependence in Fig 3.13f, as then superconductivity is probably strongest near $\nu = \pm 2$ and diminishes at charge neutrality, similarly to the measured phase diagram (disregarding the correlated insulating phases). In the view of competing phases, superconductivity would then be killed at the intermediate fillings $\nu = \pm 1, \pm 2, \pm 3$ by some separate mechanism resulting in the correlated insulating phases. Studying this mechanism is left for further studies.
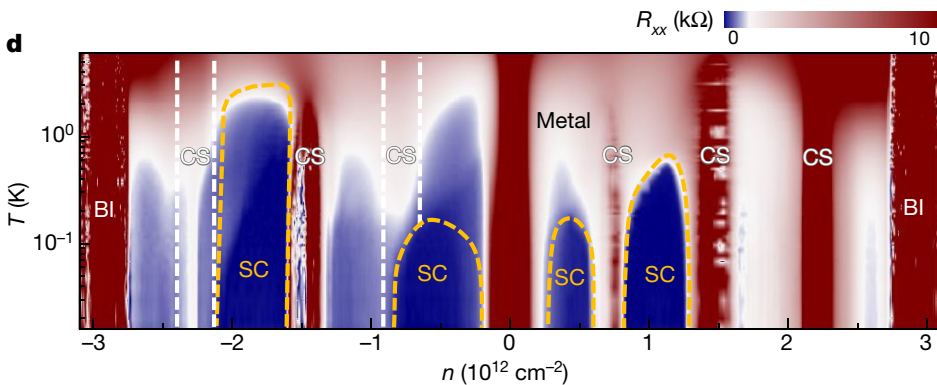
Figure 3.14: Phase diagram measured by Lu *et al.* [227] showing more correlated states as a result of a refined manufacturing process reducing disorder. Band-insulating (BI) or correlated-insulating (CS) phases appear at all integer fillings (note that the shown horizontal scale is the number density, not the filling) $\nu = -3, \ldots, 3$, while superconducting (SC) phases surround many of these. Reprinted with permission from Ref. [227]. Copyright 2019 Springer Nature.

Stepanov *et al.* [127] and Saito *et al.* [230] studied the effect of Coulomb screening by varying the thickness of the hBN layer between the graphite back gate and TBG. By increasing the screening through decreasing the hBN thickness, they could see *superconducting phases without the correlated insulating phases* when slightly deviating from the magic angle ($\theta \approx 1.15°$). The phase diagram measured by Stepanov *et al.* [127] is shown in Fig. 3.15, and a very similar phase diagram was obtained also by Saito *et al.* [230]. These are the first experiments strongly pointing towards separate mechanisms underlying the correlated insulating and superconducting phases, thus also supporting the viewpoint of Publications II and III where we assumed that superconductivity can be studied separately of the correlated insulating phase. Now the phase diagram can be readily compared to the doping-dependence figure 3.13f without resorting to *ad hoc* argumentation of the correlated insulating states. According to Stepanov *et al.* the BKT transition temperatures at the highest point of the two superconducting domes are $T_{BKT} = 710$ mK ($\nu \approx -2$) and $T_{BKT} = 410$ mK ($\nu \approx 2$), suggesting an interaction strength $\lambda = -0.9 \ldots -1.1$ eVa^2 according to Fig. 3.13c. Whether the doping dependence in Fig. 3.13f is consistent with this, should be determined by calculating it with these smaller interaction strengths and seeing if T_{BKT} vanishes at charge neutrality. The asymmetry of the domes cannot be explained by the (approximately) electron–hole symmetric DP model. But as is seen in Fig. 3c in Publication III, an electron–hole asymmetric dispersion (RM model) can reproduce this asymmetry in T_{BKT}.

Saito *et al.* had a total of five samples at different angles, allowing us to compare the θ dependence to our predicted Fig. 3.12a. But as the magic angle is slightly off with the $3G$ LdS model, I recalculate the plot with the full LdS model and show it in Fig. 3.16a in the case of zero doping $\nu = \delta\mu = 0$. With the full model the magic angle (from the viewpoint of superconductivity) is at $1.08°$ for
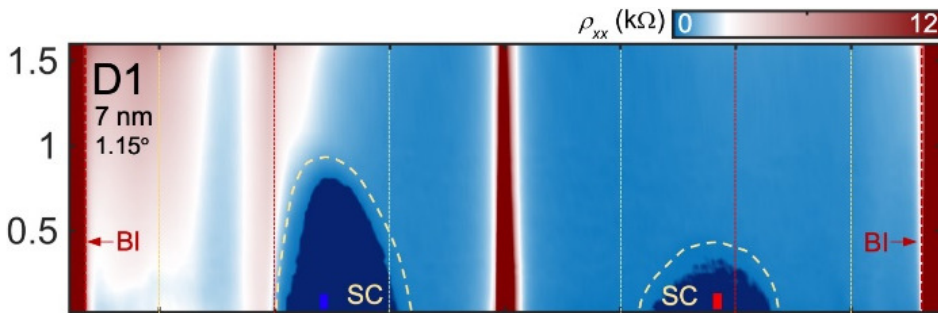
Figure 3.15: Phase diagram measured by Stepanov *et al.* [127] showing superconductivity (SC) and band insulators (BI) but no correlated insulating phases. The critical ingredient for killing the correlated insulating phases is a thinner hBN layer between the graphite back gate and the TBG sample, which enhances Coulomb screening. The vertical red dashed lines denote the integer fillings $\nu = -4, \ldots, 4$, the vertical axis is the temperature in kelvins, and the color corresponds to resistivity. Reprinted with permission from Ref. [127].

strong interactions and at $1.05°$ for weak interactions. Δ also vanishes much faster for larger angles than in the $3G$ model. To compare to the experiment data we would need the $T_{\mathrm{BKT}}(\theta)$ data, but we can calculate an estimate through Eq. (3.49) if we assume it to hold also outside the magic angle. I show this estimated T_{BKT} data in Fig. 3.16b. Unfortunately Saito *et al.* did not determine T_{BKT}, but only T_{c} which corresponds to the temperature where the resistance has dropped to half of the normal-state resistance. Thus we need to estimate also the T_{BKT} data of the experiment. According to experiments where T_{BKT} has been measured, there is roughly the relation $T_{\mathrm{BKT}} \approx 0.5 T_{\mathrm{c}}$ [9, 127]. The black points with the error bars in Fig. 3.16b are the experimental data calculated this way. We immediately see that there is a large $0.07°$ mismatch between the magic angles, and thus the calculated data does not very well match the experiment. But, as discussed in Sec. 3.3.2, the relaxation effects, which are not included in the LdS model, can be quite large. Especially according to Lin *et al.* [110] the magic angle shifts from $1.08°$ to $1.12°$ due to structural relaxation. Shifting the theory values by a slightly larger angle of $0.07°$ to match with the experimental magic angle produces Fig. 3.16c. In this case the experimental data matches the calculated $\lambda = -1 \, \mathrm{eV}a^2$ curve, which is the same value as argued already in the previous section.

Increasing the hBN thickness in the experiments of Stepanov *et al.* and Saito *et al.* to $\gtrsim 10 \, \mathrm{nm}$ produced a phase diagram that strongly resembles that of Lu *et al.* [227] in Fig. 3.14, where the thickness was also around $10 \, \mathrm{nm}$. Thus the experiments are consistent with each other, and the correlated insulators (CI) are clearly controlled by screening. Goodwin *et al.* [240] explained this theoretically by employing the tight-binding model of Trambly de Laissardière *et al.* [148] and calculating the (extended) Hubbard interaction strength resulting from the Coulomb interaction in the presence of the metallic gate(s). They found the Hubbard interaction strength, which is generally believed to underlie the CI
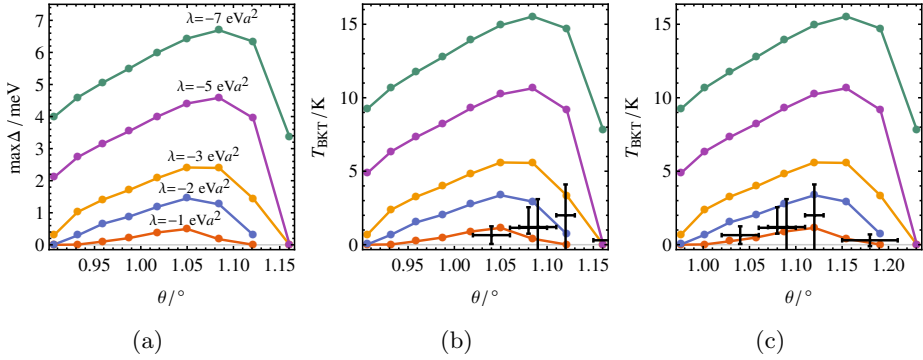
Figure 3.16: The rotation angle θ dependence of the superconducting state within the full LdS model at the charge neutrality $\nu = \delta\mu = 0$ and comparison to the experiment of Saito *et al.* [230]. (a) Maximum of the superconducting order parameter Δ. The most important difference to the $3G$ LdS model in Fig. 3.12a is the position of the magic angle, which for the full model is $1.08°$ for strong interactions and $1.05°$ for weak interactions. In the full model Δ also dies more quickly when going to larger angles. (b) The same plot as in (a), but the vertical values have been scaled by Eq. (3.49), assuming it to hold approximately also outside the magic angle. The black points with error bars show the experimental data of Saito *et al.* [230]. They reported only the T_c values, so the shown values are calculated through $T_{\mathrm{BKT}} = 0.5T_c$, which seems to roughly hold according to the experiments where the BKT transition temperature has been determined [9, 127]. The magic angle is at $1.12°$ in the experiment, so the calculated data does not match it. (c) By shifting the theory data to the magic angle of the experiment, a nice match is found with the $\lambda = -1\,\mathrm{eV}a^2$ curve.

phases, to start quickly reducing when the hBN thickness (the distance between TBG and the gate) falls below the moiré length $\approx 12\,\mathrm{nm}$. It is not clear what the critical value of the Hubbard interaction strength is to establish the CI phases, but by using rough estimates they were able to show that a critical thickness is of the order of $5\,\mathrm{nm}$, thus being consistent with the experiments. Cea *et al.* [241], on the other hand, found the bandstructure to be sensitive to screening by applying self-consistent Hartree–Fock calculations.

Codecido *et al.* [242] observed correlated insulating behavior at $\nu = \pm 2$ and superconductivity with $T_c \approx 0.4\,\mathrm{K}$ at $\nu = 2$ at an angle of $\theta = 0.93°$, which is much smaller than the first magic angle $\approx 1.1°$ but still much larger than the second magic angle $\approx 0.5°$. This can be explained by Fig. 3.16a: according to the full LdS model Δ dies slowly when going to smaller angles. However, a more accurate explanation would need the proper self-consistent calculation of T_{BKT} as a function of θ, but assuming the simple relation (3.49) to hold also outside the magic angle, the same explanation applies. Similar behavior of both T_c and T_{BKT} was predicted by Hu *et al.* [165], although in a partly simplified theory.

In Publications I and II and in Figs. 3.12 and 3.13 we mostly used a fixed chemical potential, which is not usually directly determined in experiments. Tomarken

et al. [243] performed the first capacitance measurements to obtain the "electronic compressibility" $dn/d\mu$, and obtained a qualitatively similar $n(\mu)$ dependence as we predicted in Fig. 5b in Publication II, apart from a very different energy scale. This difference is explained by the very small bandwidth of the $3G$ LdS model employed in Publication II, but the bandwidth of tens of meV observed in the experiment is not reproduced even by the full LdS model.

The TBG samples are usually produced by a "tear-and-stack" method [9, 191, 228], which does not allow controlling the rotation angle after manufacturing. Ribeiro-Palau *et al.* [244] presented a method which allows changing the rotation angle of a hBN layer on top of graphene *in situ* by using an AFM tip. While not being TBG, the principle should work for TBG also, although not with (hBN) encapsulation.

Nowadays, two years after the breakthrough experiment [9], the experiments are turning from TBG to similar systems with more freedom, such as twisted trilayer and tetralayer graphene. Moriyama *et al.* [245] observed superconductivity with $T_c = 14\,\mathrm{K}$ in *untwisted* bilayer graphene, encapsulated between hBN layers similarly as often in the TBG experiments. Chen *et al.* [246] observed correlated insulating states and superconductivity with T_c below $1\,\mathrm{K}$ in rhombohedrally stacked trilayer graphene encapsulated in hBN. This observation might be at least partly explained by the flat-band surface states in rhombohedral graphite [73, 74, 129, 247]. Tsai *et al.* [217] observed correlated insulating states and superconductivity with $T_c \approx 3.4\,\mathrm{K}$ in twisted trilayer graphene (encapsulated in hBN), at a charge density two orders of magnitude smaller than in TBG. One interesting case is the system of four layers (encapsulated in hBN), which is called twisted double bilayer graphene (TDBG) [248] or twisted bilayer-bilayer graphene (TBBG) [249], depending on which layers are rotated. Both correlated insulating states and superconductivity have been observed in this system by many groups [248–250], with $T_c \approx 7\,\mathrm{K}$ [248]. Moreover, correlated insulating states and superconductivity with $T_c \approx 6\,\mathrm{K}$ was observed in twisted double bilayer WSe_2 by An *et al.* [251].

3.5.4   Other theories in TBG and related systems

An enormous number of theory works trying to explain the various correlated phenomena in TBG has been published since the 2018 experiments [9, 130]. I present here a small subset of these works that I find somehow to be relevant to this thesis, but that also shows the wide spectrum of different models, methods, and the resulting physics.

Only superconductivity

In addition to Publications II and III, there are many other theories explaining only the superconducting phase and discarding the correlated insulator phase. Cao *et al.* [9] attributed their observations of superconductivity to *unconventional* superconductivity, as does the vast majority of theory papers, probably because many of the properties resemble the high-T_c cuprate superconductors. However, the origin of superconductivity is in fact not known, and in addition to Publications II

and III, a few other studies have shown that *conventional* theories can also explain the experiments.

Wu *et al.* [121] have very similar (and consistent) results to Publication II, except that they calculated the effective attractive electron–electron interaction strength $\lambda = -121\,\mathrm{meV\,nm^2} \approx -6\,\mathrm{eV}a^2$ resulting from the electron–phonon interaction. This, however, is suppressed by the repulsive Coulomb interaction through Eq. (2.102), the amplitude of which is much more difficult to determine. They also argued that electron–phonon interaction leads to $s + d$-wave pairing, out of which the s-wave channel has a higher T_c if discarding the Coulomb repulsion, thus supporting the chosen s-wave model of Publication II.

Lian *et al.* [171] used the 8-band Bistritzer–MacDonald model and calculated the electron–phonon coupling to be strong enough to induce conventional s-wave superconductivity with $T_c = 1\,\mathrm{K}$. Using the McMillan formula[15] they predicted that superconductivity could be observed at many other angles also if going to higher doping levels. Using an atomistic method of electrons and phonons, Choi *et al.* [36] calculated the electron–phonon coupling strength to be an order of magnitude larger than in graphene or in untwisted bilayer graphene, thus indicating that the conventional electron–phonon coupling is a viable origin of superconductivity.

The list of theory papers suggesting an unconventional mechanism is much longer. Su *et al.* [234] presented the renormalized moiré method with RVB interactions that we employed in Publication III, and found out that by increasing the interaction strength a mixed $d + \mathrm{i}d$ and $p + \mathrm{i}p$ phase evolves to a $s + p + d$ phase. Lin & Nandkishore [19] showed how d- or f-wave superconductivity can arise from weak electronic repulsion through a Kohn–Luttinger [50, 51] renormalization procedure. Wu [166] studied d-wave pairing within the model of Ref. [121] and found $(d + \mathrm{i}d) + (d - \mathrm{i}d)$-wave superconductivity with spontaneous vortices and bulk circulating supercurrents. Ray *et al.* [58] found extended s-wave superconductivity in TBG through a repulsive multiband Hubbard interaction within the Bistritzer–MacDonald model. For comparison, within the same interaction they found $d + \mathrm{i}d$-wave pairing in monolayer graphene and $p + \mathrm{i}p$-wave pairing in (untwisted) graphene on hBN. Tang *et al.* [23] used an RG approach to solve a two-orbital repulsive Hubbard model defined on the emergent superlattice, and found spin-triplet, f-wave superconductivity.

Only correlated insulator

There are also theories trying to explain only the correlated insulating phase. González-Árraga *et al.* [252] studied magnetism in a 1.5° TBG already before the Cao *et al.* experiments [9, 130], and found competing lattice-antiferromagnetism and superlattice-spiral-ferromagnetism within a repulsive Hubbard model. Xie & MacDonald [253] used self-consistent Hartree–Fock calculations to solve the

---

[15]The McMillan formula of "standard" BCS theory does not hold in the flat-band regime. In the flat-band regime one should use the modified formula of Ref. [129] instead. It is, however, not perfectly clear whether the experiments happen in the flat-band regime or not, but as discussed in Sec. 3.5 and in Publications II and III, it seems that the experiments happen on the verge of the flat-band regime. Thus it seems that neither of these McMillan equations apply.

Bistritzer–MacDonald model with Coulomb interactions and found magic-angle TBG to be (correlated) insulating at all integer fillings, out of which $\nu = \pm 1$ and $\nu = \pm 3$ may exhibit quantum anomalous Hall effect. Padhi *et al.* [254] claimed that magic-angle TBG does not satisfy a "Mott criterion" and suggested it to be a Wigner crystal instead. Ochi *et al.* [255] solved an extended Hubbard model on the effective superlattice by exact diagonalization and presented two candidates for the insulating phase: a spin- and valley-ferromagnet or a Dirac semimetal. Haule *et al.* [256] used dynamical mean-field theory (DMFT) to go beyond the mean-field and Hartree–Fock methods to show that the insulating states can be Mott insulators. Liu *et al.* [198] showed how the noninteracting bandstructure can be understood from the pseudo-Landau-level perspective by transforming the position dependence of the interlayer coupling to an intralayer pseudo vector potential, and argued that from this viewpoint the integer-filling insulators can be naturally understood when turning on the Coulomb interaction.

Superconductivity and correlated insulator

The repulsive Hubbard model is commonly used, as it can model both the correlated insulating and superconducting phases originating from the same mechanism, usually from the "doping a Mott insulator" perspective [229]. It is very common to use it on the superlattice scale, as then the unit cell contains only one or two sites, making the problem much easier. This approach is supported by studies suggesting that the Wannier states can be localized on the moiré scale [58, 179]. There are two different superlattice-scale models: in the first one the superlattice is triangular consisting of the *AA* stacking "sites" [58], whereas in the second one it is a honeycomb consisting of the *AB* and *BA* stacking "sites" [257, 258]. Some have even extended the Hubbard model to longer-range interactions, like the model developed by Yuan & Fu [259].

Let me next list a few of the studies using this effective model. Po *et al.* [258] derived effective tight-binding models starting from the symmetries of the system, and proposed that both a Mott-insulating state and a nearby *s*-wave superconducting state develop out of a state with spontaneous intervalley coherence. Guo *et al.* [28] employed a determinant quantum Monte Carlo method and found $d + \mathrm{i}d$-wave superconductivity surrounding an insulating state. Xu & Balents [22] found $d + \mathrm{i}d$-wave superconductivity near a Mott-insulator phase, as did also Fidrysiak *et al.* [18]. Liu *et al.* [29] used random phase approximation (RPA) calculations to predict a chiral spin density wave (SDW) bordered by $d + \mathrm{i}d$-wave superconductivity, driven by antiferromagnetic (AFM) fluctuations. Dodaro *et al.* [20] predicted that a single mechanism drives both nematic ferromagnetism and superconductivity surrounding it, which is mostly on-site (mostly *s* wave). Venderbos & Fernandes [260] classified the different possible superconducting, magnetic, and charge instabilities in the strong-coupling regime. Gu *et al.* [261] showed that, assuming the insulating phase to be antiferromagnetic, fluctuations in this antiferromagnet can mediate *d*-wave superconductivity. Classen *et al.* [21] used functional renormalization group calculations to predict three types of correlated phases: spin/orbital density wave, Chern insulator, and $d \pm \mathrm{i}d$- or *f*-wave superconductivity. Chen *et al.* [262] used quantum Monte Carlo method

to predict AFM-ordered Mott insulator and $d + \mathrm{i}d$-wave superconductivity.

Huang *et al.* [263] used the quantum Monte Carlo method in the lattice scale instead and predicted an AFM-ordered Mott insulator surrounded by a $d + \mathrm{i}d$-wave superconductor. Between the repulsive Hubbard models in the superlattice and lattice scales, one can use it also within the continuum model of Dirac electrons. Roy & Juričić [264] used an RG approach in this continuum model to predict competing AFM and $p + \mathrm{i}p$-wave superconducting phases.

Isobe *et al.* [265] exploited only the qualitative understanding of TBG's band-structure, wrote a general interaction Hamiltonian allowed by the symmetries, and determined by RG calculations that the leading instabilities are d- or p-wave superconductivity and charge or spin density wave as the insulating state. Laksono *et al.* [266] obtained a similar phase diagram with a totally different approach: Bistritzer–MacDonald model with a Hubbard-type repulsive interaction solved by an RPA method. Wu *et al.* [267] had a unique approach: by noting that atomic relaxation substantially enlarges the AB- and BA-stacking regions, they developed a coupled-wire model, where the AB/BA domain walls act as conducting wires between the small AA-stacking regions. Depending on the parameters, the model predicts either a correlated insulator or a superconductor with s- or $d \pm \mathrm{i}d$-wave pairing. Also standing out from the mass is the work of Chou *et al.* [268], as they predicted (by a similar phenomenological model of coupled wires) the superconducting phase to be primary, arising at generic fillings, but being interrupted by a correlated insulating state at commensurate (integer) fillings. Moreover, they are not necessarily originated from the same mechanism, which is exactly the view we took in Publications II and III (and I). Angeli *et al.* [269] showed that there are phonon modes in the superlattice scale with a flat dispersion, resulting in strong electron–phonon coupling, further yielding insulating states at integer fillings potentially together with spin-singlet, extended s-wave or $d \pm \mathrm{i}d$-wave superconductivity. González & Stauber [270] showed how p-wave superconductivity together with an SDW phase can emerge from a tight-binding model with a Coulomb repulsion through the Kohn–Luttinger instability.

Differentiating mechanisms

The list above shows that there are almost as many models, mechanisms, and correlated phases as there are studies. One problem in determining the correct pairing mechanism of superconductivity and the correlated insulators is the fact that most of the theories match (at least qualitatively) to the experiments but do not give any differentiating predictions. In Publication III we gave one, namely the isotropy/anisotropy of the superfluid weight for different pairing mechanisms. Chen *et al.* [271] used a two-band tight-binding model (in the superlattice scale) to show that different pairing symmetries could be differentiated by magnetic impurity resonance (Yu–Shiba–Rusinov) states, which can be measured by LDOS measurements near the magnetic impurity. For example, if the impurity is located at an AA site, non-s-wave pairing should yield a resonance peak close to the impurity. Using the two-parameter Bistritzer–MacDonald model Wu *et al.* [167] predicted that applying in-plane magnetic field $B_{\parallel}$ together with strain can reveal differences between spin-singlet and spin-triplet superconductivity. Specifically, $B_{\parallel}$

always suppresses T_c in spin-singlet channels whereas for spin-triplet channels a weak $B_\parallel$ increases it. On the other hand Talantsev *et al.* [272] fitted the analytical formulas of BCS and Ginzburg–Landau theory to the experimental $B_{c2}(T)$ (upper critical field) and $I_c(T)$ (self-field critical current) data of Cao *et al.* [9] and Lu *et al.* [227] and concluded that *s*- or *p*-wave pairing is most probable among the electron–phonon mechanisms. This is again consistent with Publication II.

Superfluid weight and BKT transition

There are also a handful of papers studying the superfluid weight and the BKT transition in TBG, in addition to Publication III. Hazra *et al.* [162] determined an upper bound for $D^s \lesssim 1.5\,\mathrm{meV}$ (in the units of Publication III) and T_{BKT}, which they claimed is independent of the chosen interaction mechanism, thus clearly contradicting the results of Publication III. In the Supplementary Material of Publication III we presented arguments why the claimed upper bound is not valid: neglecting all the dispersive bands is not a valid approximation when the interaction strength is strong enough, and thus the claimed "upper bound" is actually only a bound on the weak-coupling regime. Xie *et al.* [170], on the other hand, presented a topological lower bound for D^s and T_{BKT}, assuming perfectly flat bands with *s*-wave pairing. On the other hand using some estimates they found $T_{\mathrm{BKT}}/T_c \approx 0.35$, which is somewhat smaller than the value of $0.6\ldots0.8$ found in Publications II and III within the same (local) interaction model.

Hu *et al.* [165] used the Bistritzer–MacDonald model and calculated very similar results to Publication III. The results are mostly consistent with each other, except that they get somewhat larger $T_{\mathrm{BKT}}/T_c \approx 0.8\ldots0.9$ ratios. An important difference, that may explain the difference, is that in Publication III we calculated the real position-dependent Δ from the self-consistency equation and used that in calculating D^s and T_c, whereas Hu *et al.* calculated T_c from the linearized self-consistency equation and *defined a position-independent* Δ from the BCS relation $\Delta := 1.764k_B T_c$, the validity of which is questionable in TBG.

Similar systems

Theoretical explanations for the observed superconducting and correlated insulating states in TDBG *etc.* are only starting to emerge [219, 273]. Sakai *et al.* predicted a Penrose-tiling quasicrystal to exhibit an exotic superconducting phase under a magnetic field. Han *et al.* [274] predicted the flat bands of a kagomé crystal to drive a metal–correlated insulator phase transition through Wigner crystallization.

3.6 Conclusions and outlook

In Publication II we applied BCS theory to study superconductivity in twisted bilayer graphene (TBG). The motivation was to explain the 2018 breakthrough experiments of Cao *et al.* [9, 130] (and the shortly following experiment of Yankowitz *et al.* [191]), where superconductivity near a correlated insulator state was observed

when two graphene layers are relatively rotated to a magic angle $\theta \approx 1°$. The experiment has spurred a tremendous amount of interest among the field of condensed matter physics, which is understandable as TBG has turned out to be a very versatile material. Moreover, finding the mechanisms between its complex phase diagram might help solve the mystery of high-temperature superconductivity.

Our goal in Publication II was to show that simple BCS theory with spin-singlet, s-wave pairing approximately originating from electron–phonon interaction is enough to explain the observed superconducting states, whereas a vast majority of theory papers suggest a repulsive electron–electron origin. We succeeded in this goal, as we were able to reproduce the two superconducting domes at the magic angle with the critical temperature T_c in the correct order of magnitude.

The results of Publication II left, however, room for improvement, as the mean-field T_c is not a realistic description of the superconducting phase transition in 2D materials. In Publication III we filled this hole by calculating the Berezinskii–Kosterlitz–Thouless (BKT) transition temperature T_{BKT}, which properly describes the phase transition in 2D materials. It turned out that the results of Publication II were already qualitatively correct, but the true transition temperature is somewhat smaller $T_{BKT} \approx 0.6 \ldots 0.8 T_c$. Comparing to the available experimental data, we concluded that the T_{BKT} results are consistent with the experiments if the interaction strength is chosen as $\lambda \approx -1\,\mathrm{eV}a^2 \approx -6\,\mathrm{eV\mathring{A}}^2$, where a is the graphene lattice constant. We even confirmed the results by employing a second, tight-binding (called the *renormalized moiré*, RM) model in addition to the continuum (called the *Dirac point*, DP) model. Using the RM model we also showed that in calculating the superfluid weight in the flat-band regime one has to include the first dispersive bands in addition to the flat bands, contrary to many theoretical studies. Moreover, it turned out that the *geometric contribution* [93] is significant in the flat-band regime, as is expected for flat-band systems in general [93].

Often the superconducting and correlated insulating states are assumed to originate from the same mechanism, and thus according to those theories they always neighbor each other in the (doping) phase diagram. On the contrary, in Publications II and III we chose the view that they originate from different mechanisms, and can thus appear separately, allowing us to model the superconducting state by disregarding the insulating state. According to recent experiments [127, 230] this viewpoint is indeed correct: the correlated insulating states vanish and superconductivity remains when the Coulomb screening is increased by bringing the metallic gates closer.

As discussed in the case of periodically strained graphene in Sec. 2.5, also the repulsive Coulomb interaction is included in the calculation if one interprets the interaction strength as $\lambda_{\mathrm{eff}} = \lambda + u^*$ in Eq. (2.102) in the simplified view of a homogeneous Δ. Then $\lambda_{\mathrm{eff}} < 0$ corresponds to a superconductor and $\lambda_{\mathrm{eff}} > 0$ to an insulating antiferromagnet [129]. In the Supplementary Material of Publication II we estimated the parameter $\alpha \approx 3.3\,\mathrm{eV}^{-1}\,\mathrm{nm}^{-2}$ in Eq. (2.101), yielding a maximum possible value $u^* \approx 5\,\mathrm{eV}a^2$. Thus whenever the bare attractive interaction (resulting from the electron–phonon interaction) is $\lambda < -5\,\mathrm{eV}a^2$, superconductivity will occur. In reality u^* is most likely smaller, and thus a smaller λ is enough. The above, however, is a simplified picture with a homogeneous Δ,

while a more accurate description would need the generalization of Eq. (2.102) to the case of a position-dependent Δ. As discussed in Sec. 2.5, one possibility to describe the observed correlated insulators would then be through Eq. (2.102): include the screening effect of the gates as in Refs. [240, 241], calculate the doping dependence of λ and u^* in the presence of screening, and see where the sign of λ_{eff} changes.

In addition to the local s-wave-pairing model discussed above, in Publication III we employed a resonating valence bond (RVB) type of interaction, which has also been argued to be realistic in graphene. The largest difference to the local interaction was the anisotropic superfluid weight in the flat-band regime, which could be experimentally differentiated from the isotropic superfluid weight of the local interaction via radio-frequency impedance spectroscopy arranged in a Hall-like four-probe setting. The anisotropic superfluid weight actually breaks spontaneously the C_3 symmetry of the moiré superlattice, as does the order parameter Δ, which has an $s + p + d$-wave symmetry in the flat-band regime in the case of the RVB interaction.

As seen in Sec. 3.5.4, determining the origin of the superconducting and correlated insulating phases in TBG has proven a difficult problem, and it still remains debated. However, no end to the continuous flow of theory papers and new experiments is on sight, so probably we are approaching the solution. If the origin turned out to be conventional, this would simplify theories to the simple BCS type, but on the other hand if it turned out to be unconventional, it would probably help explaining the high-T_c cuprate superconductors. A vast majority of the theory papers support an unconventional mechanism which is the common origin of both the superconducting and correlated insulating phase, but on the other hand some recent experiments suggest conventional electron–phonon origin of superconductivity. Especially the experiments of Stepanov *et al.* [127] and Saito *et al.* [230] showing superconductivity without the correlated insulating phases point toward separate mechanisms. Also the experiment of Polshyn *et al.* [239] on strange-metal behavior suggests strong electron–phonon coupling. These experiments support our assumptions in Publications II and III.

There is a clear trend in the superconductivity experiments. First, superconductivity was observed [9, 190, 191] in TBG. Later, it was observed also in three [217] and four layers [248–250] of twisted graphene sheets, with a higher T_c than in TBG. Perhaps some day the experiments reach hundreds of graphene layers with controlled twist angle, and explain the high-temperature (above $100\,\text{K}$) superconductivity that was observed in graphite with random twisted interfaces already in 2012 by Scheike *et al.* [221, 222] and in 2013 by Ballestar *et al.* [223, 275], but which did not get much attention. This was probably because there was no control of the twist interfaces and angles, so that the experiments seemed nonreproducible. But now with all the new knowledge on superconductivity in twisted bilayers, trilayers, and tetralayers, those kind of graphite experiments might become fashionable soon [276].

Chapter 4

Closing

4.1 Concluding summary and outlook

In this thesis I have used (generalized) BCS theory to show that conventional spin-singlet, s-wave, flat-band superconductivity can arise in periodically strained graphene (PSG, Publication I) and in twisted bilayer graphene (TBG, Publications II and III), which both feature flat electronic bands strongly promoting superconductivity. The results are consistent with the numerous TBG superconductivity experiments performed over the last two years, while no superconductivity in PSG has been measured yet.

A huge amount of theoretical studies have emerged after the breakthrough experiments of Cao *et al.* [9, 130], trying to explain the origin of superconductivity and the correlated insulating phases, which has proven to be a complicated problem. A vast majority of them relies on an unconventional origin of superconductivity, whereas we showed that also the conventional mechanism is consistent with the experiments. Our view is supported by recent experiments suggesting strong electron–phonon coupling [239] and separate mechanisms [127, 230] for the superconducting and correlated insulating phases.

By comparing to the available experimental data, we showed that an interaction strength $\lambda \approx -6 \, \mathrm{eV \AA}^2$ is best consistent with the TBG experiments, having a Berezinskii–Kosterlitz–Thouless transition temperature of $T_{\mathrm{BKT}} \approx 1 \, \mathrm{K}$ [9, 127]. Using the same value for PSG, we calculated that the periodic-strain experiments of Jiang *et al.* [71] are already close to achieving superconductivity: if one could increase the strain amplitude by a factor of four, a relatively high $T_{\mathrm{BKT}} \approx 4 \, \mathrm{K}$ could already be measured.

The next natural extension to Publications I, II, and III would be to include the repulsive Coulomb interactions and try to model the correlated insulating state on top of superconductivity. In TBG the correlated insulators have already been observed, but because of the similarity between PSG and TBG they are very likely also in PSG. In a simplified theory the Coulomb interaction can be included through Eq. (2.102), but a more accurate description would need a generalization to the case of an inhomogeneous superconducting order parameter Δ. The competition

between the insulating (antiferromagnetic) state and superconductivity could then be modeled by the competition between the attractive interaction strength λ and the repulsive Hubbard interaction u, which is renormalized to a smaller Coulomb pseudopotential u^* through Eq. (2.101).

Regarding TBG, it remains to be seen whether at some point one is able to control the twist angles of hundreds of graphene layers, and is able to repeat in a reproducible manner the high-temperature superconductivity measurements [221–224] performed in graphite already years ago. A way towards this is already visible, as superconductivity in three [217] and four [248–250] layers of twisted graphene has now been reported, with a higher T_c than in TBG. Regarding PSG, it is intriguing to wait when (if) the first experiments reporting superconductivity (and correlated insulating phases) appear. The striking similarity to TBG suggests that finding these phases is highly probable, if a strain field with large enough amplitude and small enough period can be manufactured, simultaneously with the knowledge of the graphene lattice orientation. There are numerous possibilities for manufacturing it, and the experiments of Jiang *et al.* [71] suggest that we are actually already close to achieving a measurable T_{BKT}. Besides, the critical temperature of PSG is (in principle) way more tunable than that of TBG through controlling the shape, period, lattice, and amplitude of the strain, so a recipe for a higher T_c is already available there.

Bibliography

[1] J. R. Gavaler, Superconductivity in Nb–Ge films above 22 K, Applied Physics Letters **23**, 480 (1973).

[2] J. G. Bednorz and K. A. Müller, Possible high-T_c superconductivity in the Ba–La–Cu–O system, Zeitschrift für Physik B Condensed Matter **64**, 189 (1986).

[3] A. Schilling, M. Cantoni, J. D. Guo, and H. R. Ott, Superconductivity above 130 K in the Hg–Ba–Ca–Cu–O system, Nature **363**, 56 (1993).

[4] A. P. Drozdov, P. P. Kong, V. S. Minkov, S. P. Besedin, M. A. Kuzovnikov, S. Mozaffari, L. Balicas, F. F. Balakirev, D. E. Graf, V. B. Prakapenka, E. Greenberg, D. A. Knyazev, M. Tkacz, and M. I. Eremets, Superconductivity at 250 K in lanthanum hydride under high pressures, Nature **569**, 528 (2019).

[5] V. A. Khodel' and V. R. Shaginyan, Superfluidity in system with fermion condensate, JETP Letters **51**, 553 (1990).

[6] G. E. Volovik, Quantum Phase Transitions from Topology in Momentum Space, in *Quantum Analogues: From Phase Transitions to Black Holes and Cosmology*, Vol. 718 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2007) pp. 31–73.

[7] T. T. Heikkilä, N. B. Kopnin, and G. E. Volovik, Flat bands in topological media, JETP Letters **94**, 233 (2011).

[8] T. T. Heikkilä and G. E. Volovik, Flat Bands as a Route to High-Temperature Superconductivity in Graphite, in *Basic Physics of Functionalized Graphite*, edited by P. Esquinazi (Springer, 2016) pp. 123–143, arXiv:1504.05824 .

[9] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Unconventional superconductivity in magic-angle graphene superlattices, Nature **556**, 43 (2018).

[10] M. Tinkham, *Introduction to superconductivity*, 2nd ed., edited by J. Shira and E. Castellano (McGraw Hill, Inc., 1996).

[11] P. G. De Gennes, *Superconductivity of Metals and Alloys* (CRC Press, 2018).

[12] V. L. Ginzburg, Superconductivity: the day before yesterday — yesterday — today — tomorrow, Uspekhi Fizicheskih Nauk **170**, 619 (2000).

[13] M. Sigrist and K. Ueda, Phenomenological theory of unconventional superconductivity, Reviews of Modern Physics **63**, 239 (1991).

[14] J. Jung and A. H. MacDonald, Tight-binding model for graphene π-bands from maximally localized Wannier functions, Physical Review B **87**, 195450 (2013).

[15] Y. V. Nazarov and J. Danon, *Advanced Quantum Mechanics: A practical guide*, 1st ed. (Cambridge University Press, New York, 2013).

[16] J. Linder and A. V. Balatsky, Odd-frequency superconductivity, Reviews of Modern Physics **91**, 045005 (2019).

[17] G. E. Volovik and L. P. Gor'kov, Superconducting classes in heavy-fermion systems, JETP Letters **61**, 843 (1985).

[18] M. Fidrysiak, M. Zegrodnik, and J. Spałek, Unconventional topological superconductivity and phase diagram for an effective two-orbital model as applied to twisted bilayer graphene, Physical Review B **98**, 085436 (2018).

[19] Y.-P. Lin and R. M. Nandkishore, Kohn-Luttinger superconductivity on two orbital honeycomb lattice, Physical Review B **98**, 214521 (2018).

[20] J. F. Dodaro, S. A. Kivelson, Y. Schattner, X. Q. Sun, and C. Wang, Phases of a phenomenological model of twisted bilayer graphene, Physical Review B **98**, 075154 (2018).

[21] L. Classen, C. Honerkamp, and M. M. Scherer, Competing phases of interacting electrons on triangular lattices in moiré heterostructures, Physical Review B **99**, 195120 (2019).

[22] C. Xu and L. Balents, Topological Superconductivity in Twisted Multilayer Graphene, Physical Review Letters **121**, 087001 (2018).

[23] Q.-K. Tang, L. Yang, D. Wang, F.-C. Zhang, and Q.-H. Wang, Spin-triplet f-wave pairing in twisted bilayer graphene near $\frac{1}{4}$-filling, Physical Review B **99**, 094521 (2019).

[24] V. P. Mineev and K. V. Samokhin, *Introduction to Unconventional Superconductivity* (Gordon and Breach Science Publishers, 1999).

[25] C. C. Tsuei and J. R. Kirtley, Pairing symmetry in cuprate superconductors, Reviews of Modern Physics **72**, 969 (2000).

[26] M. Cheng, K. Sun, V. Galitski, and S. Das Sarma, Stable topological superconductivity in a family of two-dimensional fermion models, Physical Review B **81**, 024504 (2010).

[27] A. P. Schnyder, S. Ryu, A. Furusaki, and A. W. W. Ludwig, Classification of topological insulators and superconductors, Physical Review B **78**, 195125 (2008).

[28] H. Guo, X. Zhu, S. Feng, and R. T. Scalettar, Pairing symmetry of interacting fermions on a twisted bilayer graphene superlattice, Physical Review B **97**, 235453 (2018).

[29] C.-C. Liu, L.-D. Zhang, W.-Q. Chen, and F. Yang, Chiral Spin Density Wave and $d + id$ Superconductivity in the Magic-Angle-Twisted Bilayer Graphene, Physical Review Letters **121**, 217001 (2018).

[30] H. Chung, N. Kim, and H. Kim, Variation of the extended s-wave superconducting order parameter: From s-wave to g-wave, Modern Physics Letters B **29**, 1550163 (2015).

[31] L. S. Borkowski, P. J. Hirschfeld, and W. O. Putikka, Transport properties of extended- s -state superconductors, Physical Review B **52**, R3856 (1995).

[32] D. J. Van Harlingen, Phase-sensitive tests of the symmetry of the pairing state in the high-temperature superconductors — Evidence for $d_{x^2-y^2}$ symmetry, Reviews of Modern Physics **67**, 515 (1995).

[33] M. P. Marder, *Condensed Matter Physics*, 2nd ed. (John Wiley & Sons, Inc., Hoboken, New Jersey, 2010).

[34] N. B. Kopnin and E. B. Sonin, BCS Superconductivity of Dirac Electrons in Graphene Layers, Physical Review Letters **100**, 246808 (2008).

[35] D. M. Eagles, Predicted Transition Temperatures of Very Thin Films and Whiskers of Superconducting Semiconductors—Application to $SrTiO_3$, Physical Review **164**, 489 (1967).

[36] Y. W. Choi and H. J. Choi, Strong electron-phonon coupling, electron-hole asymmetry, and nonadiabaticity in magic-angle twisted bilayer graphene, Physical Review B **98**, 241412 (2018).

[37] K. S. Novoselov, Electric Field Effect in Atomically Thin Carbon Films, Science **306**, 666 (2004).

[38] X. Du, I. Skachko, A. Barker, and E. Y. Andrei, Approaching ballistic transport in suspended graphene, Nature Nanotechnology **3**, 491 (2008).

[39] M. Müller, J. Schmalian, and L. Fritz, Graphene: A Nearly Perfect Fluid, Physical Review Letters **103**, 025301 (2009).

[40] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, The electronic properties of graphene, Reviews of Modern Physics **81**, 109 (2009).

[41] B. M. Ludbrook, G. Levy, P. Nigge, M. Zonno, M. Schneider, D. J. Dvorak, C. N. Veenstra, S. Zhdanovich, D. Wong, P. Dosanjh, C. Straßer, A. Stöhr, S. Forti, C. R. Ast, U. Starke, and A. Damascelli, Evidence for superconductivity in Li-decorated monolayer graphene, Proceedings of the National Academy of Sciences **112**, 11795 (2015).

[42] J. Chapman, Y. Su, C. A. Howard, D. Kundys, A. N. Grigorenko, F. Guinea, A. K. Geim, I. V. Grigorieva, and R. R. Nair, Superconductivity in Ca-doped graphene laminates, Scientific Reports **6**, 23254 (2016).

[43] S. Ichinokura, K. Sugawara, A. Takayama, T. Takahashi, and S. Hasegawa, Superconducting Calcium-Intercalated Bilayer Graphene, ACS Nano **10**, 2761 (2016).

[44] A. P. Tiwari, S. Shin, E. Hwang, S.-G. Jung, T. Park, and H. Lee, Superconductivity at 7.4 K in few layer graphene by Li-intercalation, Journal of Physics: Condensed Matter **29**, 445701 (2017).

[45] B. Uchoa and A. H. Castro Neto, Superconducting States of Pure and Doped Graphene, Physical Review Letters **98**, 146801 (2007).

[46] T. Nihira and T. Iwata, Temperature dependence of lattice vibrations and analysis of the specific heat of graphite, Physical Review B **68**, 134305 (2003).

[47] A. M. Black-Schaffer and S. Doniach, Resonating valence bonds and mean-field d-wave superconductivity in graphite, Physical Review B **75**, 134512 (2007).

[48] S. Pathak, V. B. Shenoy, and G. Baskaran, Possible high-temperature superconducting state with a $d+id$ pairing symmetry in doped graphene, Physical Review B **81**, 085431 (2010).

[49] J. González, Kohn-Luttinger superconductivity in graphene, Physical Review B **78**, 205431 (2008).

[50] W. Kohn and J. M. Luttinger, New Mechanism for Superconductivity, Physical Review Letters **15**, 524 (1965).

[51] J. M. Luttinger, New Mechanism for Superconductivity, Physical Review **150**, 202 (1966).

[52] Y. Lozovik and A. Sokolik, Phonon-mediated electron pairing in graphene, Physics Letters A **374**, 2785 (2010).

[53] M. Einenkel and K. B. Efetov, Possibility of superconductivity due to electron-phonon interaction in graphene, Physical Review B **84**, 214508 (2011).

[54] M. L. Kiesel, C. Platt, W. Hanke, D. A. Abanin, and R. Thomale, Competing many-body instabilities and unconventional superconductivity in graphene, Physical Review B **86**, 020507 (2012).

[55] W.-S. Wang, Y.-Y. Xiang, Q.-H. Wang, F. Wang, F. Yang, and D.-H. Lee, Functional renormalization group and variational Monte Carlo studies of the electronic instabilities in graphene near $\frac{1}{4}$ doping, Physical Review B **85**, 035414 (2012).

[56] R. Nandkishore, L. S. Levitov, and A. V. Chubukov, Chiral superconductivity from repulsive interactions in doped graphene, Nature Physics **8**, 158 (2012).

[57] R. Nandkishore, R. Thomale, and A. V. Chubukov, Superconductivity from weak repulsion in hexagonal lattice systems, Physical Review B **89**, 144501 (2014).

[58] S. Ray, J. Jung, and T. Das, Wannier pairs in superconducting twisted bilayer graphene and related systems, Physical Review B **99**, 134515 (2019).

[59] N. B. Kopnin and E. B. Sonin, Supercurrent in superconducting graphene, Physical Review B **82**, 014516 (2010).

[60] I. Syôzi, Statistics of Kagome Lattice, Progress of Theoretical Physics **6**, 306 (1951).

[61] G.-B. Jo, J. Guzman, C. K. Thomas, P. Hosur, A. Vishwanath, and D. M. Stamper-Kurn, Ultracold Atoms in a Tunable Optical Kagome Lattice, Physical Review Letters **108**, 045305 (2012).

[62] D. Leykam, A. Andreanov, and S. Flach, Artificial flat band systems: from lattice models to experiments, Advances in Physics: X **3**, 1473052 (2018).

[63] E. H. Lieb, Two theorems on the Hubbard model, Physical Review Letters **62**, 1201 (1989).

[64] S. Taie, H. Ozawa, T. Ichinose, T. Nishio, S. Nakajima, and Y. Takahashi, Coherent driving and freezing of bosonic matter wave in an optical Lieb lattice, Science Advances **1**, e1500854 (2015).

[65] M. R. Slot, T. S. Gardenier, P. H. Jacobse, G. C. P. van Miert, S. N. Kempkes, S. J. M. Zevenhuizen, C. M. Smith, D. Vanmaekelbergh, and I. Swart, Experimental realization and characterization of an electronic Lieb lattice, Nature Physics **13**, 672 (2017).

[66] B. Cui, X. Zheng, J. Wang, D. Liu, S. Xie, and B. Huang, Realization of Lieb lattice in covalent-organic frameworks with tunable topology and magnetism, Nature Communications **11**, 66 (2020).

[67] V. J. Kauppila, F. Aikebaier, and T. T. Heikkilä, Flat-band superconductivity in strained Dirac materials, Physical Review B **93**, 214505 (2016).

[68] T. Wehling, A. Black-Schaffer, and A. Balatsky, Dirac materials, Advances in Physics **63**, 1 (2014).

[69] J. W. F. Venderbos and L. Fu, Interacting Dirac fermions under a spatially alternating pseudomagnetic field: Realization of spontaneous quantum Hall effect, Physical Review B **93**, 195126 (2016).

[70] M. Tahir, O. Pinaud, and H. Chen, Emergent flat band lattices in spatially periodic magnetic fields, arXiv:1808.10046 (2018).

[71] Y. Jiang, M. Anđelković, S. P. Milovanović, L. Covaci, X. Lai, Y. Cao, K. Watanabe, T. Taniguchi, F. M. Peeters, A. K. Geim, and E. Y. Andrei, Flat Bands in Buckled Graphene Superlattices, arXiv:1904.10147 (2019).

[72] A. Ramires and J. L. Lado, Electrically Tunable Gauge Fields in Tiny-Angle Twisted Bilayer Graphene, Physical Review Letters **121**, 146801 (2018).

[73] N. B. Kopnin, T. T. Heikkilä, and G. E. Volovik, High-temperature surface superconductivity in topological flat-band systems, Physical Review B **83**, 220503 (2011).

[74] N. B. Kopnin and T. T. Heikkila, Surface superconductivity in rhombohedral graphite, arXiv:1210.7075 (2012).

[75] D. Pierucci, H. Sediri, M. Hajlaoui, J.-C. Girard, T. Brumme, M. Calandra, E. Velez-fort, G. Patriarche, M. G. Silly, G. Ferro, V. Soulière, M. Marangolo, F. Sirotti, F. Mauri, and A. Ouerghi, Evidence for Flat Bands near the Fermi Level in Epitaxial Rhombohedral Multilayer Graphene, ACS Nano **9**, 5432 (2015).

[76] Y. Henni, H. P. Ojeda Collado, K. Nogajewski, M. R. Molas, G. Usaj, C. A. Balseiro, M. Orlita, M. Potemski, and C. Faugeras, Rhombohedral Multilayer Graphene: A Magneto-Raman Scattering Study, Nano Letters **16**, 3710 (2016).

[77] N. D. Mermin, Crystalline Order in Two Dimensions, Physical Review **176**, 250 (1968).

[78] Z. Hadzibabic, P. Krüger, M. Cheneau, B. Battelier, and J. Dalibard, Berezinskii–Kosterlitz–Thouless crossover in a trapped atomic gas, Nature **441**, 1118 (2006).

[79] V. L. Berezinskii, Destruction of Long-range Order in One-dimensional and Two-dimensional Systems having a Continuous Symmetry Group I. Classical Systems, Soviet Physics JETP **32**, 493 (1971).

[80] V. L. Berezinskii, Destruction of Long-range Order in One-dimensional and Two-dimensional Systems Possessing a Continuous Symmetry Group. II. Quantum Systems., Soviet Physics JETP **34**, 610 (1972).

[81] J. M. Kosterlitz and D. J. Thouless, Long range order and metastability in two dimensional solids and superfluids. (Application of dislocation theory), Journal of Physics C: Solid State Physics **5**, L124 (1972).

[82] J. M. Kosterlitz and D. J. Thouless, Ordering, metastability and phase transitions in two-dimensional systems, Journal of Physics C: Solid State Physics **6**, 1181 (1973).

[83] J. M. Kosterlitz, The critical properties of the two-dimensional xy model, Journal of Physics C: Solid State Physics **7**, 1046 (1974).

[84] D. R. Nelson and J. M. Kosterlitz, Universal Jump in the Superfluid Density of Two-Dimensional Superfluids, Physical Review Letters **39**, 1201 (1977).

[85] J. V. José, L. P. Kadanoff, S. Kirkpatrick, and D. R. Nelson, Renormalization, vortices, and symmetry-breaking perturbations in the two-dimensional planar model, Physical Review B **16**, 1217 (1977).

[86] P. Minnhagen, The two-dimensional Coulomb gas, vortex unbinding, and superfluid-superconducting films, Reviews of Modern Physics **59**, 1001 (1987).

[87] J. M. Kosterlitz, Kosterlitz-Thouless physics: a review of key issues, Reports on Progress in Physics **79**, 026001 (2016).

[88] E. Babaev and H. Kleinert, Nonperturbative XY-model approach to strong coupling superconductivity in two and three dimensions, Physical Review B **59**, 12083 (1999).

[89] V. P. Gusynin, V. M. Loktev, and S. G. Sharapov, Pseudogap phase formation in the crossover from Bose-Einstein condensation to BCS superconductivity, Journal of Experimental and Theoretical Physics **88**, 685 (1999).

[90] L. He and X.-G. Huang, BCS-BEC Crossover in 2D Fermi Gases with Rashba Spin-Orbit Coupling, Physical Review Letters **108**, 145302 (2012).

[91] L. Salasnich, P. A. Marchetti, and F. Toigo, Superfluidity, sound velocity, and quasicondensation in the two-dimensional BCS-BEC crossover, Physical Review A **88**, 053612 (2013).

[92] S. Peotta and P. Törmä, Superfluidity in topologically nontrivial flat bands, Nature Communications **6**, 8944 (2015).

[93] L. Liang, T. I. Vanhala, S. Peotta, T. Siro, A. Harju, and P. Törmä, Band geometry, Berry curvature, and superfluid weight, Physical Review B **95**, 024515 (2017).

[94] J. Pearl, Current Distribtuion In Superconducting Films Carrying Quantized Fluxoids, Applied Physics Letters **5**, 65 (1964).

[95] X. Zheng, *Efficient Fourier Transforms on Hexagonal Arrays*, PhD thesis, University of Florida (2007).

[96] A. Johansson, P. Myllyperkiö, P. Koskinen, J. Aumanen, J. Koivistoinen, H.-C. Tsai, C.-H. Chen, L.-Y. Chang, V.-M. Hiltunen, J. J. Manninen, W. Y. Woon, and M. Pettersson, Optical Forging of Graphene into Three-Dimensional Shapes, Nano Letters **17**, 6469 (2017).

[97] P. Koskinen, K. Karppinen, P. Myllyperkiö, V.-M. Hiltunen, A. Johansson, and M. Pettersson, Optically Forged Diffraction-Unlimited Ripples in Graphene, The Journal of Physical Chemistry Letters **9**, 6179 (2018).

[98] P. Jia, W. Chen, J. Qiao, M. Zhang, X. Zheng, Z. Xue, R. Liang, C. Tian, L. He, Z. Di, and X. Wang, Programmable graphene nanobubbles with three-fold symmetric pseudo-magnetic fields, Nature Communications **10**, 3127 (2019).

[99] Z. H. Aitken and R. Huang, Effects of mismatch strain and substrate surface corrugation on morphology of supported monolayer graphene, Journal of Applied Physics **107**, 123531 (2010).

[100] Y. Jiang, J. Mao, J. Duan, X. Lai, K. Watanabe, T. Taniguchi, and E. Y. Andrei, Visualizing Strain-Induced Pseudomagnetic Fields in Graphene through an hBN Magnifying Glass, Nano Letters **17**, 2839 (2017).

[101] K. Yang, Y. Chen, F. Pan, S. Wang, Y. Ma, and Q. Liu, Buckling Behavior of Substrate Supported Graphene Sheets, Materials **9**, 32 (2016).

[102] C. Androulidakis, E. N. Koukaras, O. Frank, G. Tsoukleri, D. Sfyris, J. Parthenios, N. Pugno, K. Papagelis, K. S. Novoselov, and C. Galiotis, Failure Processes in Embedded Monolayer Graphene under Axial Compression, Scientific Reports **4**, 5271 (2015).

[103] E. N. Koukaras, C. Androulidakis, G. Anagnostopoulos, K. Papagelis, and C. Galiotis, Compression behavior of simply-supported and fully embedded monolayer graphene: Theory and experiment, Extreme Mechanics Letters **8**, 191 (2016).

[104] P. Koskinen, Graphene cardboard: From ripples to tunable metamaterial, Applied Physics Letters **104**, 101902 (2014).

[105] L. Tarruell, D. Greif, T. Uehlinger, G. Jotzu, and T. Esslinger, Creating, moving and merging Dirac points with a Fermi gas in a tunable honeycomb lattice, Nature **483**, 302 (2012).

[106] M. Yankowitz, K. Watanabe, T. Taniguchi, P. San-Jose, and B. J. LeRoy, Pressure-induced commensurate stacking of graphene on boron nitride, Nature Communications **7**, 13168 (2016).

[107] M. M. van Wijk, A. Schuring, M. I. Katsnelson, and A. Fasolino, Relaxation of moiré patterns for slightly misaligned identical lattices: graphene on graphite, 2D Materials **2**, 34010 (2015).

[108] S. Dai, Y. Xiang, and D. J. Srolovitz, Twisted Bilayer Graphene: Moiré with a Twist, Nano Letters **16**, 5923 (2016).

[109] N. N. T. Nam and M. Koshino, Lattice relaxation and energy band modulation in twisted bilayer graphene, Physical Review B **96**, 075311 (2017).

[110] X. Lin, D. Liu, and D. Tománek, Shear instability in twisted bilayer graphene, Physical Review B **98**, 195432 (2018).

[111] H. Yoo, R. Engelke, S. Carr, S. Fang, K. Zhang, P. Cazeaux, S. H. Sung, R. Hovden, A. W. Tsen, T. Taniguchi, K. Watanabe, G.-C. Yi, M. Kim, M. Luskin, E. B. Tadmor, E. Kaxiras, and P. Kim, Atomic and electronic reconstruction at the van der Waals interface in twisted bilayer graphene, Nature Materials **18**, 448 (2019).

[112] H. Suzuura and T. Ando, Phonons and electron-phonon scattering in carbon nanotubes, Physical Review B **65**, 235412 (2002).

[113] M. A. H. Vozmediano, M. I. Katsnelson, and F. Guinea, Gauge fields in graphene, Physics Reports **496**, 109 (2010).

[114] J. L. Mañes, Symmetry-based approach to electron-phonon interactions in graphene, Physical Review B **76**, 045430 (2007).

[115] N. C. Yeh, C. C. Hsu, M. L. Teague, J. Q. Wang, D. A. Boyd, and C. C. Chen, Nanoscale strain engineering of graphene and graphene-based devices, Acta Mechanica Sinica **32**, 497 (2016).

[116] T. J. Peltonen and T. T. Heikkilä, Flat-band superconductivity in periodically strained graphene: mean-field and Berezinskii–Kosterlitz–Thouless transition, Journal of Physics: Condensed Matter **32**, 365603 (2020).

[117] E. Tang and L. Fu, Strain-Induced Helical Flat Band and Interface Superconductivity in Topological Crystalline Insulators, Nature Physics **10**, 964 (2014).

[118] W. P. Su, J. R. Schrieffer, and A. J. Heeger, Solitons in Polyacetylene, Physical Review Letters **42**, 1698 (1979).

[119] F. Munoz, F. Pinilla, J. Mella, and M. I. Molina, Topological properties of a bipartite lattice of domain wall states, Scientific Reports **8**, 17330 (2018).

[120] E. Langmann, C. Triola, and A. V. Balatsky, Ubiquity of Superconducting Domes in the Bardeen-Cooper-Schrieffer Theory with Finite-Range Potentials, Physical Review Letters **122**, 157001 (2019).

[121] F. Wu, A. H. MacDonald, and I. Martin, Theory of Phonon-Mediated Superconductivity in Twisted Bilayer Graphene, Physical Review Letters **121**, 257001 (2018).

[122] A. A. Abrikosov, Quantum magnetoresistance, Physical Review B **58**, 2788 (1998).

[123] H. Bruus and K. Flensberg, *Introduction to many-body quantum theory in condensed matter physics* (Oxford University Press, 2004).

[124] M. I. Katsnelson, *Graphene* (Cambridge University Press, Cambridge, 2012).

[125] J. B. Ketterson and S. N. Song, *Superconductivity* (Cambridge University Press, Cambridge, 1999).

[126] D. J. Scalapino, S. R. White, and S. Zhang, Insulator, metal, or superconductor: The criteria, Physical Review B **47**, 7995 (1993).

[127] P. Stepanov, I. Das, X. Lu, A. Fahimniya, K. Watanabe, T. Taniguchi, F. H. L. Koppens, J. Lischner, L. Levitov, and D. K. Efetov, The interplay of insulating and superconducting orders in magic-angle graphene bilayers, arXiv:1911.09198 (2019).

[128] P. Morel and P. W. Anderson, Calculation of the Superconducting State Parameters with Retarded Electron-Phonon Interaction, Physical Review **125**, 1263 (1962).

[129] R. Ojajärvi, T. Hyart, M. A. Silaev, and T. T. Heikkilä, Competition of electron-phonon mediated superconductivity and Stoner magnetism on a flat band, Physical Review B **98**, 054515 (2018).

[130] Y. Cao, V. Fatemi, A. Demir, S. Fang, S. L. Tomarken, J. Y. Luo, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, E. Kaxiras, R. C. Ashoori, and P. Jarillo-Herrero, Correlated insulator behaviour at half-filling in magic-angle graphene superlattices, Nature **556**, 80 (2018).

[131] Y. Jiang, X. Lai, K. Watanabe, T. Taniguchi, K. Haule, J. Mao, and E. Y. Andrei, Charge order and broken rotational symmetry in magic-angle twisted bilayer graphene, Nature **573**, 91 (2019).

[132] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, Continuum model of the twisted graphene bilayer, Physical Review B **86**, 155449 (2012).

[133] S. Shallcross, S. Sharma, and O. A. Pankratov, Quantum interference at the twist boundary in graphene, Physical Review Letters **101**, 056803 (2008).

[134] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, Graphene bilayer with a twist: Electronic structure, Physical Review Letters **99**, 256802 (2007).

[135] S. Shallcross, S. Sharma, E. Kandelaki, and O. A. Pankratov, Electronic structure of turbostratic graphene, Physical Review B **81**, 165105 (2010).

[136] W. Yao, E. Wang, C. Bao, Y. Zhang, K. Zhang, K. Bao, C. K. Chan, C. Chen, J. Avila, M. C. Asensio, J. Zhu, and S. Zhou, Quasicrystalline 30° twisted bilayer graphene as an incommensurate superlattice with strong interlayer coupling, Proceedings of the National Academy of Sciences **115**, 6928 (2018).

[137] E. J. Mele, Commensuration and interlayer coherence in twisted bilayer graphene, Physical Review B **81**, 161405 (2010).

[138] L. Zou, H. C. Po, A. Vishwanath, and T. Senthil, Band structure of twisted bilayer graphene: Emergent symmetries, commensurate approximants, and Wannier obstructions, Physical Review B **98**, 085435 (2018).

[139] H. Beyer, M. Müller, and T. Schimmel, Monolayers of graphite rotated by a defined angle: hexagonal superstructures by STM, Applied Physics A: Materials Science & Processing **68**, 163 (1999).

[140] R. Bistritzer and A. H. MacDonald, Moire bands in twisted double-layer graphene, Proceedings of the National Academy of Sciences **108**, 12233 (2011).

[141] G. Li, A. Luican, J. M. B. Lopes dos Santos, a. H. Castro Neto, A. Reina, J. Kong, and E. Y. Andrei, Observation of Van Hove singularities in twisted graphene layers, Nature Physics **6**, 109 (2010).

[142] A. Luican, G. Li, A. Reina, J. Kong, R. R. Nair, K. S. Novoselov, A. K. Geim, and E. Y. Andrei, Single-Layer Behavior and Its Breakdown in Twisted Graphene Layers, Physical Review Letters **106**, 126802 (2011).

[143] I. Brihuega, P. Mallet, H. González-Herrero, G. Trambly de Laissardière, M. M. Ugeda, L. Magaud, J. M. Gómez-Rodríguez, F. Ynduráin, and J.-Y. Veuillen, Unraveling the Intrinsic and Robust Nature of van Hove Singularities in Twisted Bilayer Graphene by Scanning Tunneling Microscopy and Theoretical Analysis, Physical Review Letters **109**, 196802 (2012).

[144] Q. Yao, R. Van Bremen, G. J. Slotman, L. Zhang, S. Haartsen, K. Sotthewes, P. Bampoulis, P. L. De Boeij, A. Van Houselt, S. Yuan, and H. J. Zandvliet, Spatially resolved electronic structure of twisted graphene, Physical Review B **95**, 245116 (2017).

[145] A. Kerelsky, L. J. McGilly, D. M. Kennes, L. Xian, M. Yankowitz, S. Chen, K. Watanabe, T. Taniguchi, J. Hone, C. Dean, A. Rubio, and A. N. Pasupathy, Maximized electron interactions at the magic angle in twisted bilayer graphene, Nature **572**, 95 (2019).

[146] Y. Choi, J. Kemmer, Y. Peng, A. Thomson, H. Arora, R. Polski, Y. Zhang, H. Ren, J. Alicea, G. Refael, F. von Oppen, K. Watanabe, T. Taniguchi, and S. Nadj-Perge, Electronic correlations in twisted bilayer graphene near the magic angle, Nature Physics **15**, 1174 (2019).

[147] T. Latychevskaia, C. Escher, and H.-W. Fink, Moiré structures in twisted bilayer graphene studied by transmission electron microscopy, Ultramicroscopy **197**, 46 (2019).

[148] G. Trambly de Laissardière, D. Mayou, and L. Magaud, Numerical studies of confined states in rotated bilayers of graphene, Physical Review B **86**, 125413 (2012).

[149] W. Yan, W.-Y. He, Z.-D. Chu, M. Liu, L. Meng, R.-F. Dou, Y. Zhang, Z. Liu, J.-C. Nie, and L. He, Strain and curvature induced evolution of electronic band structures in twisted graphene bilayer, Nature Communications 4, 1 (2013).

[150] Z.-D. Chu, W.-Y. He, and L. He, Coexistence of van Hove singularities and superlattice Dirac points in a slightly twisted graphene bilayer, Physical Review B 87, 155419 (2013).

[151] W. Y. He, Z. D. Chu, and L. He, Chiral tunneling in a twisted graphene bilayer, Physical Review Letters 111, 066803 (2013).

[152] A. V. Rozhkov, A. O. Sboychakov, A. L. Rakhmanov, and F. Nori, Electronic properties of graphene-based bilayer systems, Physics Reports 648, 1 (2016).

[153] G. F. S. Catarina, Twisted bilayer graphene — electronic and optical properties, PhD thesis (2017).

[154] N. Ray, Phonon-induced electron scattering and electric conductivity in twisted bilayer graphene, PhD thesis (2017).

[155] K. Uchida, S. Furuya, J.-I. Iwata, and A. Oshiyama, Atomic corrugation and electron localization due to Moiré patterns in twisted bilayer graphenes, Physical Review B 90, 155451 (2014).

[156] G. De Trambly Laissardière, D. Mayou, and L. Magaud, Localization of dirac electrons in rotated graphene bilayers, Nano Letters 10, 804 (2010).

[157] E. Suárez Morell, J. D. Correa, P. Vargas, M. Pacheco, and Z. Barticevic, Flat bands in slightly twisted bilayer graphene: Tight-binding calculations, Physical Review B 82, 121407 (2010).

[158] R. de Gail, M. O. Goerbig, F. Guinea, G. Montambaux, and A. H. Castro Neto, Topologically protected zero modes in twisted bilayer graphene, Physical Review B 84, 045436 (2011).

[159] E. J. Mele, Band symmetries and singularities in twisted multilayer graphene, Physical Review B 84, 235439 (2011).

[160] D. K. Efimkin and A. H. MacDonald, Helical network model for twisted bilayer graphene, Physical Review B 98, 035404 (2018).

[161] B. L. Chittari, N. Leconte, S. Javvaji, and J. Jung, Pressure induced compression of flatbands in twisted bilayer graphene, Electronic Structure 1, 015001 (2018).

[162] T. Hazra, N. Verma, and M. Randeria, Bounds on the Superconducting Transition Temperature: Applications to Twisted Bilayer Graphene and Cold Atoms, Physical Review X 9, 031049 (2019).

[163] K. Hejazi, C. Liu, H. Shapourian, X. Chen, and L. Balents, Multiple topological transitions in twisted bilayer graphene near the first magic angle, Physical Review B **99**, 035111 (2019).

[164] E. Khalaf, A. J. Kruchkov, G. Tarnopolsky, and A. Vishwanath, Magic angle hierarchy in twisted graphene multilayers, Physical Review B **100**, 085109 (2019).

[165] X. Hu, T. Hyart, D. I. Pikulin, and E. Rossi, Geometric and Conventional Contribution to the Superfluid Weight in Twisted Bilayer Graphene, Physical Review Letters **123**, 237002 (2019).

[166] F. Wu, Topological chiral superconductivity with spontaneous vortices and supercurrent in twisted bilayer graphene, Physical Review B **99**, 195114 (2019).

[167] F. Wu and S. Das Sarma, Identification of superconducting pairing symmetry in twisted bilayer graphene using in-plane magnetic field and strain, Physical Review B **99**, 220507 (2019).

[168] F. Wu, E. Hwang, and S. Das Sarma, Phonon-induced giant linear-in-T resistivity in magic angle twisted bilayer graphene: Ordinary strangeness and exotic superconductivity, Physical Review B **99**, 165112 (2019).

[169] G. Tarnopolsky, A. J. Kruchkov, and A. Vishwanath, Origin of Magic Angles in Twisted Bilayer Graphene, Physical Review Letters **122**, 106405 (2019).

[170] F. Xie, Z. Song, B. Lian, and B. A. Bernevig, Topology-Bounded Superfluid Weight in Twisted Bilayer Graphene, Physical Review Letters **124**, 167002 (2020).

[171] B. Lian, Z. Wang, and B. A. Bernevig, Twisted Bilayer Graphene: A Phonon-Driven Superconductor, Physical Review Letters **122**, 257002 (2019).

[172] Y. Sherkunov and J. J. Betouras, Electronic phases in twisted bilayer graphene at magic angles as a result of Van Hove singularities and interactions, Physical Review B **98**, 205151 (2018).

[173] T. J. Peltonen, R. Ojajärvi, and T. T. Heikkilä, Mean-field theory for superconductivity in twisted bilayer graphene, Physical Review B **98**, 220504(R) (2018).

[174] S. Shallcross, S. Sharma, and O. Pankratov, Emergent momentum scale, localization, and van Hove singularities in the graphene twist bilayer, Physical Review B **87**, 245403 (2013).

[175] S. Fang and E. Kaxiras, Electronic structure theory of weakly interacting bilayers, Physical Review B **93**, 235153 (2016).

[176] S. Carr, S. Fang, P. Jarillo-Herrero, and E. Kaxiras, Pressure dependence of the magic twist angle in graphene superlattices, Physical Review B **98**, 085144 (2018).

[177] S. Fang, S. Carr, Z. Zhu, D. Massatt, and E. Kaxiras, Angle-Dependent *Ab initio* Low-Energy Hamiltonians for a Relaxed Twisted Bilayer Graphene Heterostructure, arXiv:1908.00058 (2019).

[178] D. Weckbecker, S. Shallcross, M. Fleischmann, N. Ray, S. Sharma, and O. Pankratov, Low-energy theory for the graphene twist bilayer, Physical Review B **93**, 035452 (2016).

[179] M. Koshino, N. F. Yuan, T. Koretsune, M. Ochi, K. Kuroki, and L. Fu, Maximally Localized Wannier Orbitals and the Extended Hubbard Model for Twisted Bilayer Graphene, Physical Review X **8**, 031087 (2018).

[180] M. Koshino, Band structure and topological properties of twisted double bilayer graphene, Physical Review B **99**, 235406 (2019).

[181] X. Lin and D. Tománek, Minimum model for the electronic structure of twisted bilayer graphene and related structures, Physical Review B **98**, 081410 (2018).

[182] L. Balents, General continuum model for twisted bilayer graphene and arbitrary smooth deformations, SciPost Physics **7**, 048 (2019).

[183] Y. Cao, J. Y. Luo, V. Fatemi, S. Fang, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Superlattice-Induced Insulating States and Valley-Protected Orbits in Twisted Bilayer Graphene, Physical Review Letters **117**, 116804 (2016).

[184] P. Moon and M. Koshino, Optical absorption in twisted bilayer graphene, Physical Review B **87**, 205404 (2013).

[185] S. Dai, Y. Xiang, and D. J. Srolovitz, Structure and energetics of interlayer dislocations in bilayer graphene, Physical Review B **93**, 085410 (2016).

[186] P. Lucignano, D. Alfè, V. Cataudella, D. Ninno, and G. Cantele, Crucial role of atomic corrugation on the flat bands and energy gaps of twisted bilayer graphene at the "magic angle" $\theta \sim 1.08°$, Physical Review B **99**, 195419 (2019).

[187] F. Guinea and N. R. Walet, Continuum models for twisted bilayer graphene: Effect of lattice deformation and hopping parameters, Physical Review B **99**, 205134 (2019).

[188] J.-B. Qiao, L.-J. Yin, and L. He, Twisted graphene bilayer around the first magic angle engineered by heterostrain, Physical Review B **98**, 235402 (2018).

[189] Z. Bi, N. F. Q. Yuan, and L. Fu, Designing flat bands by strain, Physical Review B **100**, 035448 (2019).

[190] M. Luo, Magically strained bilayer graphene with flat bands, arXiv:1901.07435 (2019).

[191] M. Yankowitz, S. Chen, H. Polshyn, Y. Zhang, K. Watanabe, T. Taniguchi, D. Graf, A. F. Young, and C. R. Dean, Tuning superconductivity in twisted bilayer graphene, Science **363**, 1059 (2019).

[192] A. González-Tudela and J. I. Cirac, Cold atoms in twisted-bilayer optical potentials, Physical Review A **100**, 053604 (2019).

[193] D. Massatt, S. Carr, M. Luskin, and C. Ortner, Incommensurate Heterostructures in Momentum Space, Multiscale Modeling & Simulation **16**, 429 (2018).

[194] B. Amorim and E. V. Castro, Electronic spectral properties of incommensurate twisted trilayer graphene, arXiv:1807.11909 (2018).

[195] P. Cazeaux, M. Luskin, and D. Massatt, Energy Minimization of Two Dimensional Incommensurate Heterostructures, Archive for Rational Mechanics and Analysis **235**, 1289 (2020).

[196] S. J. Ahn, P. Moon, T. H. Kim, H. W. Kim, H. C. Shin, E. H. Kim, H. W. Cha, S. J. Kahng, P. Kim, M. Koshino, Y. W. Son, C. W. Yang, and J. R. Ahn, Dirac electrons in a dodecagonal graphene quasicrystal, Science **786**, 782 (2018).

[197] M. J. Park, H. S. Kim, and S. Lee, Emergent localization in dodecagonal bilayer quasicrystals, Physical Review B **99**, 245401 (2019).

[198] J. Liu, J. Liu, and X. Dai, Pseudo Landau level representation of twisted bilayer graphene: Band topology and implications on the correlated insulating phase, Physical Review B **99**, 155415 (2019).

[199] S. C. Carter, H. K. Pal, and M. Kindermann, Prediction of novel 'magic' angles and correlations for twisted bilayer graphene in a perpendicular electric field, arXiv:1810.13323 (2018).

[200] M.-Y. Choi, Y.-H. Hyun, and Y. Kim, Angle dependence of the Landau level spectrum in twisted bilayer graphene, Physical Review B **84**, 195437 (2011).

[201] C. R. Dean, L. Wang, P. Maher, C. Forsythe, F. Ghahari, Y. Gao, J. Katoch, M. Ishigami, P. Moon, M. Koshino, T. Taniguchi, K. Watanabe, K. L. Shepard, J. Hone, and P. Kim, Hofstadter's butterfly and the fractal quantum Hall effect in moiré superlattices, Nature **497**, 598 (2013).

[202] Y. Kim, P. Herlinger, P. Moon, M. Koshino, T. Taniguchi, K. Watanabe, and J. H. Smet, Charge Inversion and Topological Phase Transition at a Twist Angle Induced van Hove Singularity of Bilayer Graphene, Nano Letters **16**, 5053 (2016).

[203] H. A. Le and V. N. Do, Electronic structure and optical properties of twisted bilayer graphene calculated via time evolution of states in real space, Physical Review B **97**, 125136 (2018).

[204] P. San-Jose and E. Prada, Helical networks in twisted bilayer graphene under interlayer bias, Physical Review B **88**, 121408 (2013).

[205] Q. Tong, H. Yu, Q. Zhu, Y. Wang, X. Xu, and W. Yao, Topological mosaics in moiré superlattices of van der Waals heterobilayers, Nature Physics **13**, 356 (2017).

[206] T. Stauber, T. Low, and G. Gómez-Santos, Chiral Response of Twisted Bilayer Graphene, Physical Review Letters **120**, 046801 (2018).

[207] T. Stauber, T. Low, and G. Gómez-Santos, Linear response of twisted bilayer graphene: Continuum versus tight-binding models, Physical Review B **98**, 195414 (2018).

[208] V. N. Do, H. A. Le, and D. Bercioux, Time-evolution patterns of electrons in twisted bilayer graphene, Physical Review B **99**, 165127 (2019).

[209] D. Marchenko, D. V. Evtushinsky, E. Golias, A. Varykhalov, T. Seyller, and O. Rader, Extremely flat band in bilayer graphene, Science Advances **4**, eaau0059 (2018).

[210] L. Xian, D. M. Kennes, N. Tancogne-Dejean, M. Altarelli, and A. Rubio, Multiflat Bands and Strong Correlations in Twisted Bilayer Boron Nitride: Doping-Induced Correlated Insulator and Superconductor, Nano Letters **19**, 4934 (2019).

[211] T. Kariyado and A. Vishwanath, Flat band in twisted bilayer Bravais lattices, Physical Review Research **1**, 033076 (2019).

[212] F. Crasto de Lima, R. H. Miwa, and E. Suárez Morell, Double flat bands in kagome twisted bilayers, Physical Review B **100**, 155421 (2019).

[213] L. Wang, S. Zihlmann, M.-H. Liu, P. Makk, K. Watanabe, T. Taniguchi, A. Baumgartner, and C. Schönenberger, New Generation of Moiré Superlattices in Doubly Aligned hBN/Graphene/hBN Heterostructures, Nano Letters **19**, 2371 (2019).

[214] C. Mora, N. Regnault, and B. A. Bernevig, Flatbands and Perfect Metal in Trilayer Moiré Graphene, Physical Review Letters **123**, 026402 (2019).

[215] X. Li, F. Wu, and A. H. MacDonald, Electronic Structure of Single-Twist Trilayer Graphene, arXiv:1907.12338 (2019).

[216] S. Carr, C. Li, Z. Zhu, E. Kaxiras, S. Sachdev, and A. Kruchkov, Ultraheavy and Ultrarelativistic Dirac Quasiparticles in Sandwiched Graphenes, Nano Letters **20**, 3030 (2020).

[217] K.-T. Tsai, X. Zhang, Z. Zhu, Y. Luo, S. Carr, M. Luskin, E. Kaxiras, and K. Wang, Correlated Superconducting and Insulating States in Twisted Trilayer Graphene Moire of Moire Superlattices, arXiv:1912.03375 (2019).

[218] Y. W. Choi and H. J. Choi, Intrinsic band gap and electrically tunable flat bands in twisted double bilayer graphene, Physical Review B **100**, 201402 (2019).

[219] J. Y. Lee, E. Khalaf, S. Liu, X. Liu, Z. Hao, P. Kim, and A. Vishwanath, Theory of correlated insulating behaviour and spin-triplet superconductivity in twisted double bilayer graphene, Nature Communications **10**, 5333 (2019).

[220] T. Cea, N. R. Walet, and F. Guinea, Twists and the Electronic Structure of Graphitic Materials, Nano Letters **19**, 8683 (2019).

[221] T. Scheike, W. Böhlmann, P. Esquinazi, J. Barzola-Quiquia, A. Ballestar, and A. Setzer, Can Doping Graphite Trigger Room Temperature Superconductivity? Evidence for Granular High-Temperature Superconductivity in Water-Treated Graphite Powder, Advanced Materials **24**, 5826 (2012).

[222] T. Scheike, P. Esquinazi, A. Setzer, and W. Böhlmann, Granular superconductivity at room temperature in bulk highly oriented pyrolytic graphite samples, Carbon **59**, 140 (2013).

[223] A. Ballestar, J. Barzola-Quiquia, T. Scheike, and P. Esquinazi, Josephson-coupled superconducting regions embedded at the interfaces of highly oriented pyrolytic graphite, New Journal of Physics **15**, 023024 (2013).

[224] A. Ballestar, T. T. Heikkilä, and P. Esquinazi, Size dependence of the Josephson critical behavior in pyrolytic graphite TEM lamellae, Superconductor Science and Technology **27**, 115014 (2014).

[225] P. Esquinazi, T. T. Heikkilä, Y. V. Lysogorskiy, D. A. Tayurskii, and G. E. Volovik, On the superconductivity of graphite interfaces, JETP Letters **100**, 336 (2014).

[226] Y. Cao, D. Chowdhury, D. Rodan-Legrain, O. Rubies-Bigorda, K. Watanabe, T. Taniguchi, T. Senthil, and P. Jarillo-Herrero, Strange Metal in Magic-Angle Graphene with near Planckian Dissipation, Physical Review Letters **124**, 076801 (2020).

[227] X. Lu, P. Stepanov, W. Yang, M. Xie, M. A. Aamir, I. Das, C. Urgell, K. Watanabe, T. Taniguchi, G. Zhang, A. Bachtold, A. H. MacDonald, and D. K. Efetov, Superconductors, orbital magnets and correlated states in magic-angle bilayer graphene, Nature **574**, 653 (2019).

[228] A. L. Sharpe, E. J. Fox, A. W. Barnard, J. Finney, K. Watanabe, T. Taniguchi, M. A. Kastner, and D. Goldhaber-Gordon, Emergent ferromagnetism near three-quarters filling in twisted bilayer graphene, Science **365**, 605 (2019).

[229] P. A. Lee, N. Nagaosa, and X.-G. Wen, Doping a Mott insulator: Physics of high-temperature superconductivity, Reviews of Modern Physics **78**, 17 (2006).

[230] Y. Saito, J. Ge, K. Watanabe, T. Taniguchi, and A. F. Young, Independent superconductors and correlated insulators in twisted bilayer graphene, Nature Physics 10.1038/s41567-020-0928-3 (2020).

[231] G. Baskaran, Resonating-valence-bond contribution to superconductivity in MgB$_2$, Physical Review B **65**, 212505 (2002).

[232] A. M. Black-Schaffer, Edge Properties and Majorana Fermions in the Proposed Chiral d-Wave Superconducting State of Doped Graphene, Physical Review Letters **109**, 197001 (2012).

[233] A. M. Black-Schaffer and C. Honerkamp, Chiral d-wave superconductivity in doped graphene, Journal of Physics: Condensed Matter **26**, 423201 (2014).

[234] Y. Su and S.-Z. Lin, Pairing symmetry and spontaneous vortex-antivortex lattice in superconducting twisted-bilayer graphene: Bogoliubov-de Gennes approach, Physical Review B **98**, 195101 (2018).

[235] P. W. Anderson, The Resonating Valence Bond State in La$_2$CuO$_4$ and Superconductivity, Science **235**, 1196 (1987).

[236] F. Chiodi, M. Ferrier, K. Tikhonov, P. Virtanen, T. T. Heikkilä, M. Feigelman, S. Guéron, and H. Bouchiat, Probing the dynamics of Andreev states in a coherent Normal/Superconducting ring, Scientific Reports **1**, 3 (2011).

[237] Y.-H. Zhang, D. Mao, and T. Senthil, Twisted bilayer graphene aligned with hexagonal boron nitride: Anomalous Hall effect and a lattice model, Physical Review Research **1**, 033126 (2019).

[238] N. Bultinck, S. Chatterjee, and M. P. Zaletel, Mechanism for Anomalous Hall Ferromagnetism in Twisted Bilayer Graphene, Physical Review Letters **124**, 166601 (2020).

[239] H. Polshyn, M. Yankowitz, S. Chen, Y. Zhang, K. Watanabe, T. Taniguchi, C. R. Dean, and A. F. Young, Large linear-in-temperature resistivity in twisted bilayer graphene, Nature Physics **15**, 1011 (2019).

[240] Z. A. H. Goodwin, V. Vitale, F. Corsetti, D. K. Efetov, A. A. Mostofi, and J. Lischner, Critical role of device geometry for the phase diagram of twisted bilayer graphene, Physical Review B **101**, 165110 (2020).

[241] T. Cea, N. R. Walet, and F. Guinea, Electronic band structure and pinning of Fermi energy to Van Hove singularities in twisted bilayer graphene: A self-consistent approach, Physical Review B **100**, 205113 (2019).

[242] E. Codecido, Q. Wang, R. Koester, S. Che, H. Tian, R. Lv, S. Tran, K. Watanabe, T. Taniguchi, F. Zhang, M. Bockrath, and C. N. Lau, Correlated Insulating and Superconducting States in Twisted Bilayer Graphene Below the Magic Angle, arXiv:1902.05151 (2019).

[243] S. L. Tomarken, Y. Cao, A. Demir, K. Watanabe, T. Taniguchi, P. Jarillo-Herrero, and R. C. Ashoori, Electronic Compressibility of Magic-Angle Graphene Superlattices, Physical Review Letters **123**, 046601 (2019).

[244] R. Ribeiro-Palau, C. Zhang, K. Watanabe, T. Taniguchi, J. Hone, and C. R. Dean, Twistable electronics with dynamically rotatable heterostructures, Science **361**, 690 (2018).

[245] S. Moriyama, Y. Morita, K. Komatsu, K. Endo, T. Iwasaki, S. Naka-harai, Y. Noguchi, Y. Wakayama, E. Watanabe, D. Tsuya, K. Watan-abe, and T. Taniguchi, Observation of superconductivity in bilayer graphene/hexagonal boron nitride superlattices, arXiv:1901.09356 (2019).

[246] G. Chen, A. L. Sharpe, P. Gallagher, I. T. Rosen, E. J. Fox, L. Jiang, B. Lyu, H. Li, K. Watanabe, T. Taniguchi, J. Jung, Z. Shi, D. Goldhaber-Gordon, Y. Zhang, and F. Wang, Signatures of tunable superconductivity in a trilayer graphene moiré superlattice, Nature **572**, 215 (2019).

[247] N. B. Kopnin, Surface superconductivity in multilayered rhombohedral graphene: Supercurrent, JETP Letters **94**, 81 (2011).

[248] C. Shen, Y. Chu, Q. Wu, N. Li, S. Wang, Y. Zhao, J. Tang, J. Liu, J. Tian, K. Watanabe, T. Taniguchi, R. Yang, Z. Y. Meng, D. Shi, O. V. Yazyev, and G. Zhang, Correlated states in twisted double bilayer graphene, Nature Physics **16**, 520 (2020).

[249] Y. Cao, D. Rodan-Legrain, O. Rubies-Bigorda, J. M. Park, K. Watanabe, T. Taniguchi, and P. Jarillo-Herrero, Tunable correlated states and spin-polarized phases in twisted bilayer–bilayer graphene, Nature **583**, 215 (2020).

[250] X. Liu, Z. Hao, E. Khalaf, J. Y. Lee, Y. Ronen, H. Yoo, D. Haei Najafabadi, K. Watanabe, T. Taniguchi, A. Vishwanath, and P. Kim, Tunable spin-polarized correlated states in twisted double bilayer graphene, Nature **583**, 221 (2020).

[251] L. An, X. Cai, M. Huang, Z. Wu, J. Lin, Z. Ying, Z. Ye, X. Feng, and N. Wang, Interaction effects and superconductivity signatures in twisted double-bilayer WSe_2, arXiv:1907.03966 (2019).

[252] L. A. Gonzalez-Arraga, J. L. Lado, F. Guinea, and P. San-Jose, Electrically Controllable Magnetism in Twisted Bilayer Graphene, Physical Review Letters **119**, 107201 (2017).

[253] M. Xie and A. H. MacDonald, Nature of the Correlated Insulator States in Twisted Bilayer Graphene, Physical Review Letters **124**, 097601 (2020).

[254] B. Padhi, C. Setty, and P. W. Phillips, Doped Twisted Bilayer Graphene near Magic Angles: Proximity to Wigner Crystallization, Not Mott Insulation, Nano Letters **18**, 6175 (2018).

[255] M. Ochi, M. Koshino, and K. Kuroki, Possible correlated insulating states in magic-angle twisted bilayer graphene under strongly competing interactions, Physical Review B **98**, 081102 (2018).

[256] M. Haule, E. Y. Andrei, and K. Haule, The Mott-semiconducting state in the magic angle bilayer graphene, arXiv:1901.09852 (2019).

[257] J. Kang and O. Vafek, Strong Coupling Phases of Partially Filled Twisted Bilayer Graphene Narrow Bands, Physical Review Letters **122**, 246401 (2019).

[258] H. C. Po, L. Zou, A. Vishwanath, and T. Senthil, Origin of Mott Insulating Behavior and Superconductivity in Twisted Bilayer Graphene, Physical Review X **8**, 031089 (2018).

[259] N. F. Q. Yuan and L. Fu, Model for the metal-insulator transition in graphene superlattices and beyond, Physical Review B **98**, 045103 (2018).

[260] J. W. Venderbos and R. M. Fernandes, Correlations and electronic order in a two-orbital honeycomb lattice model for twisted bilayer graphene, Physical Review B **98**, 245103 (2018).

[261] X. Gu, C. Chen, J. N. Leaw, E. Laksono, V. M. Pereira, G. Vignale, and S. Adam, Antiferromagnetism and chiral d-wave superconductivity from an effective $t - J - D$ model for twisted bilayer graphene, Physical Review B **101**, 180506 (2020).

[262] W. Chen, Y. Chu, T. Huang, and T. Ma, Metal-insulator transition and dominant $d + id$ pairing symmetry in twisted bilayer graphene, Physical Review B **101**, 155413 (2020).

[263] T. Huang, L. Zhang, and T. Ma, Antiferromagnetically ordered Mott insulator and $d + id$ superconductivity in twisted bilayer graphene: a quantum Monte Carlo study, Science Bulletin **64**, 310 (2019).

[264] B. Roy and V. Juričić, Unconventional superconductivity in nearly flat bands in twisted bilayer graphene, Physical Review B **99**, 121407 (2019).

[265] H. Isobe, N. F. Yuan, and L. Fu, Unconventional Superconductivity and Density Waves in Twisted Bilayer Graphene, Physical Review X **8**, 041041 (2018).

[266] E. Laksono, J. N. Leaw, A. Reaves, M. Singh, X. Wang, S. Adam, and X. Gu, Singlet superconductivity enhanced by charge order in nested twisted bilayer graphene Fermi surfaces, Solid State Communications **282**, 38 (2018).

[267] X.-C. Wu, C.-M. Jian, and C. Xu, Coupled-wire description of the correlated physics in twisted bilayer graphene, Physical Review B **99**, 161405 (2019).

[268] Y.-Z. Chou, Y.-P. Lin, S. Das Sarma, and R. M. Nandkishore, Superconductor versus insulator in twisted bilayer graphene, Physical Review B **100**, 115128 (2019).

[269] M. Angeli, E. Tosatti, and M. Fabrizio, Valley Jahn-Teller Effect in Twisted Bilayer Graphene, Physical Review X **9**, 041010 (2019).

[270] J. González and T. Stauber, Kohn-Luttinger Superconductivity in Twisted Bilayer Graphene, Physical Review Letters **122**, 026801 (2019).

[271] L. Chen, H.-Z. Li, and R.-S. Han, Magnetic impurity resonance states for different pairing symmetries in twisted bilayer graphene, Journal of Physics: Condensed Matter **31**, 065601 (2019).

[272] E. F. Talantsev, R. C. Mataira, and W. P. Crump, Classifying superconductivity in Moiré graphene superlattices, Scientific Reports **10**, 212 (2020).

[273] M. S. Scheurer and R. Samajdar, Pairing in graphene-based moiré superlattices, Physical Review Research **2**, 033062 (2020).

[274] W. H. Han, S. Kim, I.-H. Lee, and K. J. Chang, A Metal-Insulator Transition via Wigner Crystallization in Boron Triangular Kagome Lattice, arXiv:1902.08390 (2019).

[275] A. Ballestar, *Superconductivity at Graphite interfaces*, PhD thesis, Universität Leipzig (2014).

[276] G. E. Volovik, Graphite, Graphene, and the Flat Band Superconductivity, JETP Letters **107**, 516 (2018).

ORIGINAL PAPERS

I

FLAT-BAND SUPERCONDUCTIVITY IN PERIODICALLY STRAINED GRAPHENE: MEAN-FIELD AND BEREZINSKII-KOSTERLITZ-THOULESS TRANSITION

by

Teemu J. Peltonen and Tero T. Heikkilä 2020

Flat-band superconductivity in periodically strained graphene: mean-field and Berezinskii–Kosterlitz–Thouless transition

Teemu J. Peltonen[1] and Tero T. Heikkilä[1]

[1]*University of Jyväskylä, Department of Physics and Nanoscience Center,*
P.O. Box 35 (YFL), FI-40014 University of Jyväskylä, Finland

In the search of high-temperature superconductivity one option is to focus on increasing the density of electronic states. Here we study both the normal and s-wave superconducting state properties of periodically strained graphene, which exhibits approximate flat bands with a high density of states, with the flatness tunable by the strain profile. We generalize earlier results regarding a one-dimensional harmonic strain to arbitrary periodic strain fields, and further extend the results by calculating the superfluid weight and the Berezinskii–Kosterlitz–Thouless (BKT) transition temperature T_{BKT} to determine the true transition point. By numerically solving the self-consistency equation, we find a strongly inhomogeneous superconducting order parameter, similarly to twisted bilayer graphene. In the flat-band regime the order parameter magnitude, critical chemical potential, critical temperature, superfluid weight, and BKT transition temperature are all approximately linear in the interaction strength, which suggests that high-temperature superconductivity might be feasible in this system. We especially show that by using realistic strain strengths T_{BKT} can be made much larger than in twisted bilayer graphene, if using similar interaction strengths. We also calculate properties such as the local density of states that could serve as experimental fingerprints for the presented model.

I. INTRODUCTION

Graphene was long waiting for superconductivity to be added to its long list of miraculous properties. It took over ten years after its discovery before superconductivity was demonstrated in chemically doped graphene [1–4] with a critical temperature T_{c} of a few kelvin. Recently the experiments on magic-angle twisted bilayer graphene (TBG) [5–7] have drawn much more attention, demonstrating superconductivity in a carbon-only material (although the role of the hexagonal boron nitride substrates is being disputed [8]) similarly with a T_{c} of a few kelvin.

Lack of superconductivity in pristine graphene can be understood from the small-ν limit of the standard Bardeen–Cooper–Schrieffer (BCS) result for the critical temperature, $T_{\mathrm{c}} \sim \omega_{\mathrm{c}} e^{-1/(|\lambda|\nu)}$ [9, 10], with $|\lambda|$ describing the strength of the attractive electron–electron interaction, ν being the density of states (DOS) at the Fermi level, and ω_{c} being the cutoff (Debye) frequency. Since for intrinsic, undoped, graphene the density of states at the Fermi level is $\nu = 0$, according to this result we have also $T_{\mathrm{c}} = 0$. The doping experiments can be understood from the same result. Since close to the Dirac point ν increases linearly with chemical potential, doping can be utilized to render T_{c} finite. But due to the exponential suppression of the critical temperature, to produce T_{c} of a few kelvin, the chemical potential shift has to be of the order of eV [1, 3], corresponding to a very heavy doping level.

TBG provides an alternative mean to render T_{c} finite: increase the density of states by flattening the electronic bands through moiré-modulated interlayer coupling. In the limit of a large ν (the flat-band limit), BCS theory gives a linear relationship $T_{\mathrm{c}} \sim |\lambda|/\Omega$ [10], where Ω is the area of the flat band, instead of the exponential one. The linear relation allows in principle to increase

T_{c} much higher even with a small interaction $|\lambda|$. Here the limiting factor seems to be the area Ω of the flat band, which in the case of TBG is roughly the superlattice (moiré) Brillouin zone (SBZ), fixed by the rotation angle θ. Since θ fixes also the interlayer coupling modulation, the whole dispersion is fixed by the rotation alone. From experiments [5, 11] and theories [12, 13] we know that in order to yield flat bands θ has to be close to the magic angle $\theta^* \approx 1°$, for which Ω is only about 0.04 % [14] of the original graphene Brillouin zone (BZ). An increase of a few kelvin in T_{c} has been successfully demonstrated [6] by applying high pressure to slightly increase θ^* and thus also Ω. In TBG the flat bands are in fact not exactly at zero energy, but of the order of meV higher and lower. But compared to chemically doped graphene where $\sim$ eV doping levels are needed, a thousand-fold reduction in the needed chemical potentials allows using much simpler and more easily tunable electrical doping.

In this paper we study yet another mechanism to produce flat bands in graphene, which is possibly free of the limitations in TBG: periodic strain [15–21]. Instead of periodically modulating interlayer hopping in TBG, we modulate the intralayer hopping in monolayer graphene by periodic strain. In this system we can, in principle, separately choose the strain period d (and thus the SBZ area $\sim \Omega$) and its strength β (and thus the flatness of the bands), potentially allowing us to increase T_{c} higher than in TBG by engineering strains with high amplitude and small period.

At low energies, near the $\boldsymbol{K}$ and $\boldsymbol{K}' = -\boldsymbol{K}$ points where graphene can be described as a Dirac material, strain is modelled by a pseudo vector potential $\boldsymbol{A}$ [16, 18, 22–25], similarly to an external magnetic field. But while the external magnetic field breaks the time-reversal symmetry and usually suppresses superconductivity, the strain-induced $\boldsymbol{A}$ has opposite signs on differ-

ent valleys, preserving time-reversal symmetry and thus preserving and even promoting spin-singlet superconductivity. Moreover, strain-induced pseudo vector potentials can easily reach an effective magnetic field strength of tens [26] or even hundreds [21, 27] of tesla, opening the possibility for extreme tuning of electronic properties.

Possibilities for experimentally producing periodic strain in graphene are numerous. In fact, flat bands have already been observed in an experiment by Jiang *et al.* [21], where both 1D and 2D periodic strains were created by boundary conditions. In this experiment the displacement amplitude was of the order of 1 Å and the period d was tunable between 8 and 25 nm. Even better control of the strain pattern could perhaps be achieved by optical forging [28], which allows drawing arbitrary out-of-plane strain patterns in graphene, even below the diffraction limit [29]. On the other hand the small secondary ripples observed in the simulations [28] could be exploited, similarly to the Jiang *et al.* experiment [21], but with better control.

Another experimentally demonstrated method is to use an AFM tip to evaporate hydrogen from a Ge-H substrate to produce a pressurized H_2 gas under specific locations of graphene [30]. One option could be graphene on a corrugated surface [31, 32]. Applying in-plane compression has been predicted to produce periodic wrinkles both in simply-supported [31, 33] and encapsulated [34, 35] graphene, with amplitude and period of the order of 0.2 Å and 2 nm, respectively. In the same spirit the proposed graphene cardboard material could be manufactured [36]. Also an ultracold atom gas in a tunable optical honeycomb lattice [37] could be used.

It has been predicted [38–42] and observed [43] that TBG exhibits moiré-periodic strain due to lattice mismatch and the following structural relaxation. The relative magnitude of the moiré and strain effects can be, however, difficult to disentangle, as superconductivity by both effects has been predicted by BCS theory [14, 19]. But if the moiré effect is enhancing for superconductivity, as it seems to be, we get a lower bound for T_c by studying the strain effects. Similarly periodic strain can be expected with other mismatch lattices, such as graphene on hBN [6].

In this work we generalize the model and results of Kauppila *et al.* [19], where both the normal and superconducting spin-singlet, s-wave state in periodically strained graphene (PSG) have been studied in the case of a cosine-like 1D potential $\boldsymbol{A}(x,y) = \frac{\beta}{d}(0, \cos(2\pi x/d))$, to arbitrary periodic pseudo vector potentials $\boldsymbol{A}$. This generalization is motivated by the experiment of Jiang *et al.* [21], where a variety of periodic strain patterns, both 1D and 2D, were manufactured. On the other hand generalizing the theory to 2D strains bridges the gap between PSG [19] and TBG [14] by showing how similar these two systems are in many aspects.

The main conclusions of Kauppila *et al.* are that (i) approximate flat bands are formed in the normal state, (ii) the superconducting order parameter $\Delta(x)$ becomes inhomogeneous and is peaked near the minima/maxima of $\nabla \times \boldsymbol{A}$, similarly to the local density of states (LDOS), (iii) magnitude of Δ can be tuned by the amplitude β, (iv) T_c is linear in λ in the flat-band regime (large λ or β), and (v) even though Δ is strongly inhomogeneous and anisotropic, supercurrent is only slightly anisotropic. We show that these results continue to hold even when we change the shape of $\boldsymbol{A}$ and move to 2D potentials. In addition we show how the shape of $\boldsymbol{A}$ and its dimensionality affect the superconducting order parameter Δ, the critical chemical potential μ_c, and the critical temperature T_c. We furthermore extend the calculations by calculating the superfluid weight [44, 45] D^s and the Berezinskii–Kosterlitz–Thouless (BKT) transition temperature T_{BKT} to determine the proper transition temperature in a 2D system.

In addition to Kauppila *et al.*, spin-singlet, s-wave superconductivity in strained graphene was studied also by Uchoa *et al.* [16]. They, however, concentrated on strain fields with a homogeneous pseudomagnetic field $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ and correspondingly to a superconducting state with a homogeneous order parameter Δ. This constraint allowed them to derive analytical formulas *e.g.* for Δ and T_c, yielding a similar T_c-linear-in-λ result as what Kauppila *et al.* found out. Here we instead focus on periodic and inhomogeneous strain accompanying also a periodic and inhomogeneous $\boldsymbol{B}$, which is probably more accessible experimentally, since a constant $\boldsymbol{B}$ might be difficult to obtain [27]. Our approach is also complementary to Uchoa *et al.* in the sense that the periodicity of $\boldsymbol{B}$ allows us to use the notion of (flat) electronic bands, whereas in the pseudo-Landau-level perspective of Uchoa *et al.* the electronic levels are the pseudo-Landau levels.

This article is organized as follows. In section II we derive the Bogoliubov–de Gennes (BdG) theory to describe the superconducting state of PSG at low energies, details of which are shown in the Supplementary Material [22]. In section III we present the results of applying some selected periodic pseudo vector potentials $\boldsymbol{A}$ by numerically solving the self-consistency equation. In section IV we summarize the main results and discuss open questions and future prospects.

II. MODEL

In the low-energy limit, after adding an in-plane displacement field $\boldsymbol{u}$ and an out-of-plane displacement field h, the graphene continuum Hamiltonian for valley $\rho \in \{+, -\}$ is

$$\mathcal{H}^\rho(\boldsymbol{r}) = \hbar v_F \boldsymbol{\sigma}^\rho \cdot (-\mathrm{i}\nabla + \rho \boldsymbol{A}(\boldsymbol{r})) - \mu, \qquad (1)$$

where the pseudo vector potential is given by [22–24]

$$\boldsymbol{A} = -\frac{\beta_G}{2a_0}(u_{xx} - u_{yy}, -2u_{xy}) \qquad (2)$$

and the strain tensor is

$$u_{ij} = \frac{1}{2}(\partial_i u_j + \partial_j u_i) + \frac{1}{2}\partial_i h \partial_j h. \qquad (3)$$

Here v_{F} is the graphene Fermi velocity, μ is the chemical potential, $\beta_{\mathrm{G}} = -\,\mathrm{d}\ln t/\mathrm{d}\ln a_0 \approx 2$ is the graphene Grüneisen parameter [24], a_0 is the carbon-carbon bond length, $\boldsymbol{\sigma}^\rho = (\rho\sigma_x, \sigma_y)$ is a vector of sublattice-space Pauli matrices, and the graphene zigzag direction is assumed to be in the x direction. Note that $\boldsymbol{A}$ works exactly like a vector potential related to an external magnetic field, but with the important difference that it changes sign on valley exchange $\rho \mapsto \bar{\rho}$, preserving time-reversal symmetry $\mathcal{H}^{\bar{\rho}*} = \mathcal{H}^\rho$. Because of the relation (2) we use the words "strain" and "pseudo vector potential" interchangeably. Note that for the linear elasticity theory to be valid we should have [22]

$$\|\boldsymbol{u}(\boldsymbol{r} + \boldsymbol{\delta}_j) - \boldsymbol{u}(\boldsymbol{r})\|, \|h(\boldsymbol{r} + \boldsymbol{\delta}_j) - h(\boldsymbol{r})\| \ll a_0, \qquad (4)$$

where $\boldsymbol{\delta}_1$, $\boldsymbol{\delta}_2$, and $\boldsymbol{\delta}_3$ are the graphene nearest neighbor vectors.

We model the possible superconducting state by a (slightly generalized) BCS theory using BdG formalism. We assume an intervalley, local (also in sublattice) interaction of strength λ (negative for attractive interaction considered here), which has been widely used in the past graphene literature [46–51] to model s-wave superconductivity. In this case the effective interacting mean-field continuum Hamiltonian can be shown to be [22]

$$H_{\mathrm{BdG}} = \sum_{\sigma\rho} \int \mathrm{d}\boldsymbol{r}\, \psi_{\sigma\rho}^\dagger(\boldsymbol{r})\mathcal{H}^\rho(\boldsymbol{r})\psi_{\sigma\rho}(\boldsymbol{r})$$
$$+ \frac{1}{2}\sum_{\sigma\rho} \int \mathrm{d}\boldsymbol{r}\, \psi_{\sigma\rho}^\dagger(\boldsymbol{r})\Delta_\sigma(\boldsymbol{r})\psi_{\bar{\sigma}\bar{\rho}}^{\dagger\mathsf{T}}(\boldsymbol{r}) + \mathrm{h.c.} + \mathrm{const}, \tag{5}$$

where $\sigma \in \{\uparrow, \downarrow\}$ denotes spin, the real-space integrals are over the Born–von Kármán cell $\mathbb{R}^2/L_{\mathrm{BK}}$, and $\psi_{\sigma\rho}(\boldsymbol{r}) = (\psi_{\sigma\rho,A}(\boldsymbol{r}), \psi_{\sigma\rho,B}(\boldsymbol{r}))^\mathsf{T}$ is a sublattice-space vector of the electron annihilation operators. Furthermore the superconducting order parameter in the sublattice space is $\Delta_\sigma(\boldsymbol{r}) = \mathrm{diag}(\Delta_{\sigma,A}(\boldsymbol{r}), \Delta_{\sigma,B}(\boldsymbol{r}))$, where

$$\Delta_{\sigma,\alpha} = \lambda \sum_\rho \langle \psi_{\bar{\sigma}\bar{\rho},\alpha} \psi_{\sigma\rho,\alpha} \rangle \tag{6}$$

with angle brackets denoting the thermal average and $\alpha \in \{A, B\}$ denoting the sublattice. Note that this kind of a local interaction corresponds to spin-singlet type of superconductivity, since from the fermionic anticommutation relations it directly follows that $\Delta_{\bar{\sigma},\alpha} = -\Delta_{\sigma,\alpha}$. Furthermore due to locality $\boldsymbol{r}$ denotes the center-of-mass coordinate of the Cooper pair, while the relative coordinate is always zero, meaning that this interaction corresponds to s-wave superconductivity.

Utilizing the fermionic anticommutation relations and by doubling the basis set we can bring H_{BdG} in (5) into

the Nambu form

$$H_{\mathrm{BdG}} = \frac{1}{2}\sum_{\sigma\rho} \int \mathrm{d}\boldsymbol{r}\, \Psi_{\sigma\rho}^\dagger(\boldsymbol{r})\mathcal{H}_{\mathrm{BdG}}^\rho(\boldsymbol{r})\Psi_{\sigma\rho}(\boldsymbol{r}) + \mathrm{const}, \tag{7}$$

where the BdG Hamiltonian in Nambu space and the Nambu-vector are

$$\mathcal{H}_{\mathrm{BdG}}^\rho = \begin{pmatrix} \mathcal{H}^\rho & \Delta \\ \Delta^* & -\mathcal{H}^\rho \end{pmatrix}, \quad \Psi_{\sigma\rho} = \begin{pmatrix} \psi_{\sigma\rho} \\ s(\sigma)\psi_{\sigma\bar{\rho}}^{\dagger\mathsf{T}} \end{pmatrix}, \tag{8}$$

respectively. Here the spin-independent order parameter is $\Delta = \Delta_\uparrow = s(\sigma)\Delta_\sigma$, $s(\uparrow) = 1$, and $s(\downarrow) = -1$.

Using the spectral theorem, a symmetry between the positive- and negative-energy states, and defining the fermionic Bogoliubon operators as

$$\gamma_{\sigma\rho b\boldsymbol{k}} = \frac{1}{\sqrt{V}} \int \mathrm{d}\boldsymbol{r}\, w_{\rho b\boldsymbol{k}}^\dagger(\boldsymbol{r})\Psi_{\sigma\rho}(\boldsymbol{r}), \tag{9}$$

we may bring H_{BdG} into the diagonal form [22]

$$H = \frac{1}{2}\sum_{\sigma\rho b\boldsymbol{k}} E_{\rho b\boldsymbol{k}} \gamma_{\sigma\rho b\boldsymbol{k}}^\dagger \gamma_{\sigma\rho b\boldsymbol{k}} + \mathrm{const}. \tag{10}$$

Here $\boldsymbol{k}$ together with the band index b enumerate the *positive-energy* solutions of the BdG equation

$$\mathcal{H}_{\mathrm{BdG}}^\rho(\boldsymbol{r})w_{\rho b\boldsymbol{k}}(\boldsymbol{r}) = E_{\rho b\boldsymbol{k}} w_{\rho b\boldsymbol{k}}(\boldsymbol{r}) \tag{11}$$

and $V = |\mathbb{R}^2/L_{\mathrm{BK}}|$ is the area of the Born–von Kármán cell. According to the calculation above, diagonalizing H_{BdG}, *i.e.* bringing it to the form (10), is equivalent to solving the BdG equation (11).

By inverting the Bogoliubov transformation (9) we may write the definition of the order parameter (6) as the self-consistency equation [22]

$$\Delta_\alpha(\boldsymbol{r}) = -\frac{\lambda}{V}\sum_{\rho b\boldsymbol{k}} u_{\rho b\boldsymbol{k},\alpha}(\boldsymbol{r})v_{\rho b\boldsymbol{k},\alpha}^*(\boldsymbol{r})\tanh\left(\frac{E_{\rho b\boldsymbol{k}}}{2k_{\mathrm{B}}T}\right), \tag{12}$$

at temperature T, where we denoted the Nambu components of w as $w = (u, v)^\mathsf{T}$. Note that Δ_α might depend on sublattice α, while Kauppila *et al.* [19] defined Δ by summing over α. As we see below, the self-consistent Δ_α is, in fact, sublattice dependent, leading to a different $\boldsymbol{r}$ dependence than in [19].

In real space the self-consistency equation (12) is local in space but the BdG equation (11) is a group of 2 difficult differential eigenvalue equations. The equations can be made easier to solve by utilizing periodicity of $\boldsymbol{A}$ and writing them in Fourier space. We assume both the pseudo vector potential $\boldsymbol{A} : \mathbb{R}^2/SL \to \mathbb{R}^2$ (and thus the strain) and the order parameter Δ to be periodic in translations of the arbitrary superlattice $SL = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{t}_1, \boldsymbol{t}_2\} \subset \mathbb{R}^2$, allowing us to use the Fourier series [22]

$$\boldsymbol{A}(\boldsymbol{r}) = \sum_{\boldsymbol{G}} \mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}}\tilde{\boldsymbol{A}}(\boldsymbol{G}), \quad \Delta(\boldsymbol{r}) = \sum_{\boldsymbol{G}} \mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}}\tilde{\Delta}(\boldsymbol{G}). \tag{13}$$

Here the sums are over SL_S^*, where $SL_{\mathrm{RZ}}^* = SL^* = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{G}_1, \boldsymbol{G}_2\}$ is the reciprocal lattice of SL, $SL_{\mathrm{MZ}}^* = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{G}_1\}$ is a one-dimensional sublattice of SL^*, and $S \in \{\mathrm{RZ}, \mathrm{MZ}\}$ denotes either the *reduced zone scheme* or the *mixed zone scheme* (the terms are justified below), the latter of which being applicable only if $\boldsymbol{A}$ and Δ are constant in the $\boldsymbol{t}_2$ direction, which we call the *1D potential* case. Otherwise we call $\boldsymbol{A}$ a *2D potential*.

Together with the assumption of the eigenfunctions $w_{\rho b \boldsymbol{k}}$ being periodic in the Born–von Kármán cell, the Fourier series (13) imply the existence of the Bloch-type Fourier series

$$w_{\rho b \boldsymbol{k}}(\boldsymbol{r}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \sum_{\boldsymbol{G}} \mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} \tilde{w}_{\rho b \boldsymbol{k}}(\boldsymbol{k}+\boldsymbol{G}) \qquad (14)$$

and the Fourier space version of the BdG equation [22]

$$\sum_{\boldsymbol{G}'} \tilde{\mathcal{H}}_{\mathrm{BdG},\boldsymbol{G}\boldsymbol{G}'}^{\rho}(\boldsymbol{k}) \tilde{w}_{\rho b \boldsymbol{k}}(\boldsymbol{k}+\boldsymbol{G}') = E_{\rho b \boldsymbol{k}} \tilde{w}_{\rho b \boldsymbol{k}}(\boldsymbol{k}+\boldsymbol{G}). \quad (15)$$

In the matrix form (15) can be written as

$$\underline{\tilde{\mathcal{H}}}_{\mathrm{BdG}}^{\rho}(\boldsymbol{k}) \underline{\tilde{w}}_{\rho b \boldsymbol{k}} = E_{\rho b \boldsymbol{k}} \underline{\tilde{w}}_{\rho b \boldsymbol{k}}, \qquad (16)$$

where the underlined variables are matrices or vectors in the $\boldsymbol{G}$ space. Here $\boldsymbol{k} \in L_{\mathrm{BK}}^*/SL_S^*$ belongs to the superlattice Brillouin zone (SBZ) in the scheme S, b enumerates the positive-energy bands for each $\boldsymbol{k}$, and the Nambu-space BdG Hamiltonian is

$$\tilde{\mathcal{H}}_{\mathrm{BdG},\boldsymbol{G}\boldsymbol{G}'}^{\rho}(\boldsymbol{k}) = \begin{pmatrix} \tilde{\mathcal{H}}_{\boldsymbol{G}\boldsymbol{G}'}^{\rho}(\boldsymbol{k}) & \tilde{\Delta}(\boldsymbol{G}-\boldsymbol{G}') \\ \tilde{\Delta}^*(\boldsymbol{G}'-\boldsymbol{G}) & -\tilde{\mathcal{H}}_{\boldsymbol{G}\boldsymbol{G}'}^{\rho}(\boldsymbol{k}) \end{pmatrix} \qquad (17)$$

with the noninteracting (normal state) Hamiltonian

$$\tilde{\mathcal{H}}_{\boldsymbol{G}\boldsymbol{G}'}^{\rho}(\boldsymbol{k}) = \qquad\qquad\qquad\qquad (18)$$
$$\hbar v_{\mathrm{F}} \boldsymbol{\sigma}^{\rho} \cdot \left[(\boldsymbol{k}+\boldsymbol{G})\delta_{\boldsymbol{G}\boldsymbol{G}'} + \rho\tilde{\boldsymbol{A}}(\boldsymbol{G}-\boldsymbol{G}') \right] - \mu\delta_{\boldsymbol{G}\boldsymbol{G}'}.$$

Note the similarity to the Dirac-point low-energy TBG model in [14, 51, 52]: while here $\tilde{\boldsymbol{A}}$ couples the sublattices and $\boldsymbol{G}$ vectors within the layer, in TBG the Hamiltonian (18) has a two-layer structure, $\tilde{\boldsymbol{A}}$ is absent, and the interlayer coupling $\tilde{t}_\perp$ couples sublattices and $\boldsymbol{G}$ vectors between the layers. As we show in this paper, the second layer is not necessary for yielding flat bands, but what seems to be enough is coupling in the $\boldsymbol{G}$ space. To generalize the theory to study the combined effect of periodic strain and moiré physics, which should yield even more pronounced flat bands, would thus be easy: add the second rotated layer to the noninteracting Hamiltonian (18) and couple the layers by $\tilde{t}_\perp(\boldsymbol{G}-\boldsymbol{G}')$.

Let us discuss the notion of the reduced and the mixed zone schemes. In the reduced zone scheme $\boldsymbol{k} = k_1\boldsymbol{G}_1 + k_2\boldsymbol{G}_2 \in L_{\mathrm{BK}}^*/SL_{\mathrm{RZ}}^*$ is periodic both in the $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$ directions, with both $k_1, k_2 \in [-\frac{1}{2}, \frac{1}{2}[$ being periodic Bloch momenta. This is also traditionally called the reduced zone (or the repeated zone) scheme. In the case of $\boldsymbol{A}$ and Δ being constant in the $\boldsymbol{t}_2$ direction (the

1D potential case) we are also allowed to use the mixed zone scheme, where $\boldsymbol{k} = k_1\boldsymbol{G}_1 + k_2\boldsymbol{G}_2 \in L_{\mathrm{BK}}^*/SL_{\mathrm{MZ}}^*$ is periodic only in the $\boldsymbol{G}_1$ direction but not in the $\boldsymbol{G}_2$ direction, with $k_1 \in [-\frac{1}{2}, \frac{1}{2}[$ being a periodic Bloch momentum and $k_2 \in]-\infty, \infty[$ being a nonperiodic real momentum. Thus in the traditional notion the $\boldsymbol{G}_1$ direction is in the reduced (or repeated) zone and the $\boldsymbol{G}_2$ direction in the extended zone scheme, justifying the term mixed zone scheme.

The reduced zone scheme is convenient if one wants to compare the effects of the 1D and 2D potentials, since the dispersions look similar and the notion of a band is the same, but the calculations are heavy due to the $\boldsymbol{G}$ space being two-dimensional. On the other hand the mixed zone scheme produces cleaner-looking dispersions and is computationally much lighter due to the $\boldsymbol{G}$ space being only one-dimensional, but with the cost of more difficult comparison between the 1D and 2D potentials. Thus in all the 1D potential calculations we use the mixed zone scheme unless otherwise stated. Also Kauppila *et al.* [19] used the mixed zone scheme in all the calculations and visualizations.

Using the Fourier series (13) and (14) in (12) and approximating the $\boldsymbol{k}$ sum as an integral (assuming the Born–von Kármán cell to be large), the Fourier-space self-consistency equation becomes [22]

$$\tilde{\Delta}_\alpha(\boldsymbol{G}) = -\frac{\lambda}{(2\pi)^2} \sum_{\rho b \boldsymbol{G}'} \int \mathrm{d}\boldsymbol{k} \tanh\left(\frac{E_{\rho b \boldsymbol{k}}}{2k_{\mathrm{B}}T}\right) \times$$
$$\times \tilde{u}_{\rho b \boldsymbol{k},\alpha}(\boldsymbol{k}+\boldsymbol{G}')\tilde{v}_{\rho b \boldsymbol{k},\alpha}^*(\boldsymbol{k}+\boldsymbol{G}'-\boldsymbol{G}), \quad (19)$$

where the integral is over the continuum superlattice Brillouin zone $\mathbb{R}^2/SL_S^*$ in the scheme S, which in the reduced zone scheme can be interpreted as the parallelogram defined by $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$, and in the mixed zone scheme as the semi-infinite parallelogram with the finite side being $\boldsymbol{G}_1$ and the infinite side being in the direction of $\boldsymbol{G}_2$.

In summary, in Fourier space we are solving the BdG equation (15) together with the self-consistency equation (19). Now the BdG equation is a normal matrix eigenvalue equation, but the price to pay is that the corresponding matrix has countably infinite dimension ($2 \times 2 \times |SL_S^*|$), and the self-consistency equation becomes nonlocal in the Fourier components. Numerically, however, they are easy to solve, provided we truncate the Fourier-component set SL_S^* and the band sum, and in the case of 1D potential add a momentum cutoff in the $\boldsymbol{k}$ integral in the $\boldsymbol{G}_2$ direction. These cutoffs we choose so large that the results (dispersion, Δ) start to become saturated, and together they correspond to the energy cutoff $\hbar\omega_{\mathrm{c}}$ introduced earlier.

In a 2D system, however, we know that the superconducting transition is not properly described by the mean-field critical temperature T_{c} determined from the order parameter Δ, but by the BKT transition temperature determined from the superfluid weight D^{s}, which describes the linearized supercurrent density response $\langle \boldsymbol{j} \rangle = (\frac{e}{\hbar})^2 D^{\mathrm{s}} \langle \boldsymbol{\mathcal{A}} \rangle$ to an external (real) vector potential

$\mathcal{A}$, where the angle brackets denote average over position. For the present model we may calculate the $\mu, \nu \in \{x, y\}$ component of the superfluid weight from [22, 45]

$$D_{\mu\nu}^{s} = \frac{(\hbar v_{\mathrm{F}})^2}{(2\pi)^2} \sum_{\rho b b'} \int d\boldsymbol{k} \, \frac{f(E_{\rho b}) - f(E_{\rho b'})}{E_{\rho b} - E_{\rho b'}} \times \qquad (20)$$

$$\times \left(\tilde{w}_{\rho b}^{\dagger} \sigma_{\mu}^{\rho} \tilde{w}_{\rho b'} \tilde{w}_{\rho b'}^{\dagger} \sigma_{\nu}^{\rho} \tilde{w}_{\rho b} - \tilde{w}_{\rho b}^{\dagger} \tau_z \sigma_{\mu}^{\rho} \tilde{w}_{\rho b'} \tilde{w}_{\rho b'}^{\dagger} \tau_z \sigma_{\nu}^{\rho} \tilde{w}_{\rho b} \right),$$

where the b, b' band sums are calculated over both the positive- and negative-energy bands, τ_z is the Pauli-z matrix in Nambu space, f is the Fermi–Dirac distribution, the difference quotient is interpreted as the derivative $f'(E_{\rho b})$ if $E_{\rho b} = E_{\rho b'}$, and where we suppressed the $\boldsymbol{k}$ dependence.

From the temperature dependence of D^s we can then calculate the BKT transition temperature T_{BKT} from the generalized KT–Nelson criterion [53–55]

$$k_{\mathrm{B}} T_{\mathrm{BKT}} = \frac{\pi}{8} \sqrt{\det D^s(T_{\mathrm{BKT}})} \qquad (21)$$

for an anisotropic superfluid weight, which also needs to be calculated self-consistently, unless $D^s(T_{\mathrm{BKT}}) \approx D^s(0)$.

III. RESULTS

We solve [56] the order parameter Δ, the superfluid weight D^s, and the Berezinskii–Kosterlitz–Thouless transition temperature T_{BKT} for a selection of periodic pseudo vector potentials $\boldsymbol{A}$ with the period d. Δ is solved from the self-consistency equation (19) by the fixed-point iteration method with the initial guess of a constant order parameter $\Delta_A = \Delta_B$ [22], D^s is calculated from (20), and T_{BKT} is calculated by interpolating (21) in a predetermined temperature mesh.

In the case of a 1D potential we concentrate on the potentials

$$\boldsymbol{A}_{\cos}^{1D}(x, y) = \frac{\beta}{d}(0, \cos(2\pi x/d)), \qquad (22)$$

$$\boldsymbol{A}_{c}^{1D}(x, y) = \frac{\beta}{d}(0, \mathrm{triangleSquare}_c(x/d)), \qquad (23)$$

both periodic in translations of the square superlattice $SL = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{t}_1, \boldsymbol{t}_2\}$ with the primitive vectors $\boldsymbol{t}_1 = (d, 0)$ and $\boldsymbol{t}_2 = (0, d)$ (or any multiple of $\boldsymbol{t}_2$). The latter utilizes the function $\mathrm{triangleSquare}_c$, shown in figure 1, which is a d-periodic waveform where the slope parameter $c \in [4, \infty[$ can be used to interpolate between the triangle and square waveforms. This allows controlling the slope $\pm\beta c/d^2$ of $\boldsymbol{A}_c^{1D}$ at the lines $x = \mp d/4$. Note that the triangle waveform $c = 4$ corresponds to the case of (nearby) islands with constant pseudomagnetic fields, as discussed qualitatively in Uchoa *et al.* [footnote 20 in Ref. 16]. Note also that the potential $\boldsymbol{A}_{2\pi}^{1D}$ has exactly the same slope as $\boldsymbol{A}_{\cos}^{1D}$ at the points $x = \pm d/4$ and also otherwise approximates that potential rather well, so all
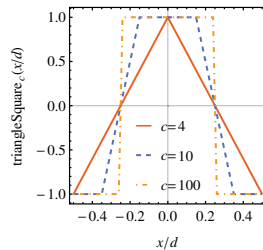

Figure 1. A plot of the d-periodic function $\mathrm{triangleSquare}_c$ used for defining the potential $\boldsymbol{A}_c^{1D}$, shown here for three values of c. The slope near the points $x = \mp d/4$ is given by $\pm c/d$.

the following results are more or less indistinguishable between these two potentials. Since both the potentials $\boldsymbol{A}_{\cos}^{1D}$ and $\boldsymbol{A}_c^{1D}$ are constant in the $\boldsymbol{t}_2$ direction, this allows us to use either the reduced zone or the mixed zone scheme in the theory.

To concretize the difference between the two schemes we write the Fourier components of the cosine potential. In the reduced zone scheme they are [22]

$$\tilde{\boldsymbol{A}}_{\cos}^{1D}(m_1 \boldsymbol{G}_1 + m_2 \boldsymbol{G}_2) =$$
$$\frac{\beta}{2d}(0, \delta_{m_1, -1} + \delta_{m_1, 1})\delta_{m_2, 0}, \qquad (24)$$

for the cosine potential and for $\boldsymbol{A}_c^{1D}$ they are given in the Supplementary Material [22]. Here $m_1 \boldsymbol{G}_1 + m_2 \boldsymbol{G}_2 \in SL_{\mathrm{RZ}}^* = SL^*$ belongs to SBZ in the reduced zone scheme, where the SBZ primitive vectors are $\boldsymbol{G}_1 = (2\pi/d, 0)$ and $\boldsymbol{G}_2 = (0, 2\pi/d)$. But since for the 1D potentials the components are multiplied by $\delta_{m_2, 0}$, we may as well use a one-dimensional Fourier series [22] and define in the mixed zone scheme

$$\tilde{\boldsymbol{A}}_{\cos}^{1D}(m_1 \boldsymbol{G}_1) = \frac{\beta}{2d}(0, \delta_{m_1, -1} + \delta_{m_1, 1}), \qquad (25)$$

where $m_1 \boldsymbol{G}_1 \in SL_{\mathrm{MZ}}^*$ belongs to SBZ in the mixed zone scheme.

On the other hand in the 2D case we concentrate on the simplest generalization of the 1D cosine-like potential $\boldsymbol{A}_{\cos}^{1D}$, the potential

$$\boldsymbol{A}_{\cos}^{2D}(x, y) = \frac{\beta}{d}(\cos(2\pi y/d), \cos(2\pi x/d)) \qquad (26)$$

with the lattice of periodicity being the square superlattice $SL = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{t}_1, \boldsymbol{t}_2\}$, with the primitive vectors $\boldsymbol{t}_1 = (d, 0)$, $\boldsymbol{t}_2 = (0, d)$. Note that we are allowed to choose a potential periodic in any superlattice, whereas in TBG the (moiré) superlattice is fixed by the rotation angle. Thus in principle the periodic strain allows much more freedom in tuning the system. The Fourier compo-
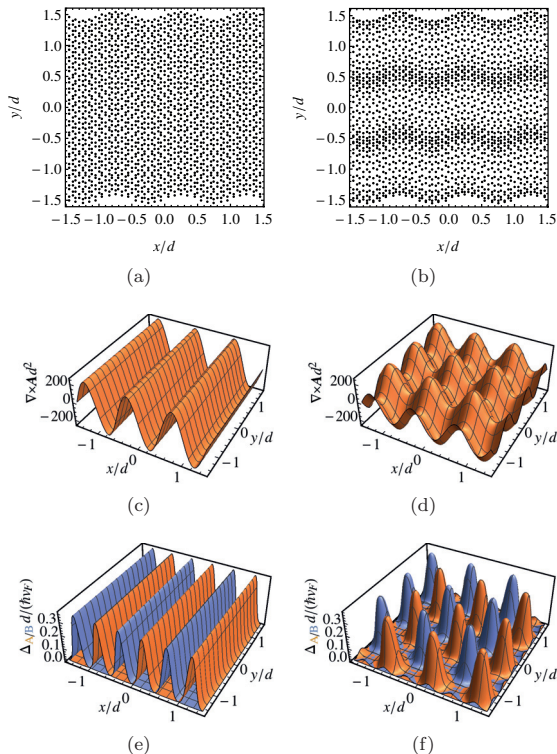
Figure 2. (a,b) Example in-plane displacement fields, defined in (28) and (29), producing the studied pseudo vector potentials $\boldsymbol{A}_{\cos}^{1D}$ and $\boldsymbol{A}_{\cos}^{2D}$ through (2) with exaggeratedly large amplitude and small period. (c,d) The corresponding pseudomagnetic fields $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ with $\beta = 40$ and $\beta = 20$, respectively. (e,f) Corresponding typical profiles of the self-consistent superconducting order parameter $\Delta_{A/B}$ (A orange, B blue), which is always peaked at the minima/maxima of $\nabla \times \boldsymbol{A}$. The parameters for calculating Δ are $T = 0$, $\lambda/(\hbar v_F d) = -0.01$, and optimal doping $\mu = \mu_{opt}$ yielding a maximal Δ ($\mu = 0$ produces the same Δ for such large λ).

nents of the 2D cosine potential are

$$
\tilde{\boldsymbol{A}}_{\cos}^{2D}(m_1\boldsymbol{G}_1 + m_2\boldsymbol{G}_2) = \\
\frac{\beta}{2d}(\delta_{m_2,-1} + \delta_{m_2,1}, \delta_{m_1,-1} + \delta_{m_1,1}) \quad (27)
$$

in the reduced zone scheme, where $m_1\boldsymbol{G}_1 + m_2\boldsymbol{G}_2 \in SL_{RZ}^*$ with $\boldsymbol{G}_1 = (2\pi/d, 0)$ and $\boldsymbol{G}_2 = (0, 2\pi/d)$.

According to (2) the potentials $\boldsymbol{A}_{\cos}^{1D}$ and $\boldsymbol{A}_{\cos}^{2D}$ can be produced for example by the in-plane displacement fields

$$
\boldsymbol{u}_{\cos}^{1D}(x,y) = \frac{\beta a_0}{\beta_G \pi}(0, \sin(2\pi x/d)), \quad (28)
$$

$$
\boldsymbol{u}_{\cos}^{2D}(x,y) = \frac{\beta a_0}{\beta_G \pi}(0, \sin(2\pi x/d) + \sin(2\pi y/d)), \quad (29)
$$

respectively. The pseudomagnetic fields $\boldsymbol{B} = \nabla \times \boldsymbol{A} =$

$\partial_x A_y - \partial_y A_x$ produced by the 1D and 2D cosine potentials, together with these example displacement fields, are shown in figures 2(a–d). The amplitude B of $\boldsymbol{B}$, which is an important factor determining the flatness of the bands and the magnitude of the superconducting order parameter $\Delta_{A/B}$, is

$$
B_{\cos}^{1D} = \frac{2\pi\beta}{d^2}, \quad B_c^{1D} = \frac{c\beta}{d^2}, \quad B_{\cos}^{2D} = \frac{4\pi\beta}{d^2} \quad (30)
$$

for the potential $\boldsymbol{A}_{\cos}^{1D}$, $\boldsymbol{A}_c^{1D}$, or $\boldsymbol{A}_{\cos}^{2D}$, respectively. To give a realistic scale for β, in the experiment by Jiang et al. [21] a pseudomagnetic field of $\frac{\hbar}{e}B \approx 100\,\text{T}$ was observed for a strain period of $d = 14\,\text{nm}$, which corresponds to $\beta \approx 5$ for the 1D cosine potential. To be better in the flat-band regime, we mostly use a factor of 4 to 8 times larger values of β in this study.

Corresponding typical profiles of $\Delta_{A/B}$ for the cosine potentials are shown in figures 2(e–f), from where it is clear that $\Delta_{A/B}$ is always peaked at the minima/maxima of the pseudomagnetic field $\boldsymbol{B}$. For comparison in TBG [14] Δ is localized around the AA stacking regions and is independent of the sublattice and layer. Note that the sublattice dependence was not present in the work by Kauppila et al. [19] due to sublattice-summation in the self-consistency equation. As we see below, it is approximately the maximum (over position $\boldsymbol{r}$) of the order parameter that is important in describing the strength of the superconducting state. As for all the studied potentials the maximum of the order parameter is independent of the sublattice, we simply denote $\max \Delta := \max \Delta_A = \max \Delta_B$.

The typical dispersion relations in the normal state are shown in figure 3 together with the conical unstrained graphene dispersions. For an easier comparison the 1D potential $\boldsymbol{A}_{\cos}^{1D}$ dispersion is shown both in the mixed zone and reduced zone schemes, while the 2D potential $\boldsymbol{A}_{\cos}^{2D}$ dispersion only in the reduced zone scheme (the only possibility in this case). We find similar-looking approximate flat bands as in TBG [14, 52], with the difference that here the number and the flatness of the flat bands can be controlled by β and c. Also all the successive bands are touching, while in TBG many models predict the flat bands to be isolated [40, 42, 52, 57].

We calculate most of the superconducting state results at optimal doping $\mu = \mu_{opt}$, which is the energy of the density of states peak as discussed in Sec. III A, and is thus the doping level with the highest Δ. We start discussing the superconducting state results by calculating $\max \Delta$ as a function of the interaction strength λ for the different potentials $\boldsymbol{A}$, as shown in figures 4(a) for the cosine potentials. The most important conclusion is that for large enough λ, β, or c, which we call the flat-band regime due to the energy scale of Δ exceeding the flat-band bandwidth, the dependence is linear in λ as we would expect for any flat-band superconductor [10]. On the other hand for small enough λ, β, and c the dispersive behavior of the lowest energy bands starts playing a role, which we call the dispersive regime. In the dispersive
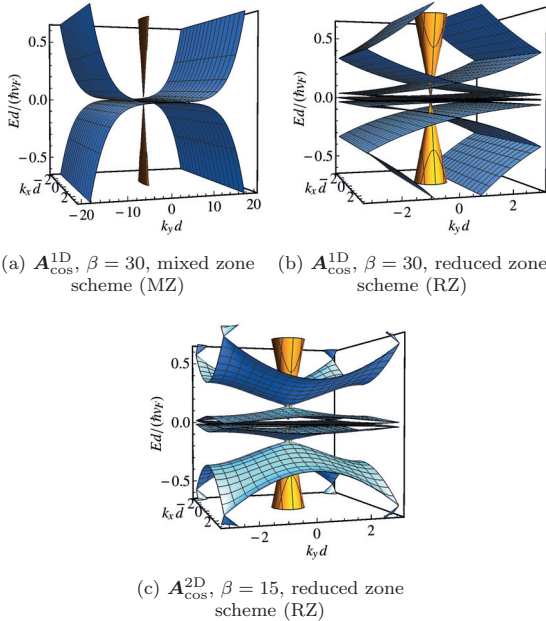
(a) $\boldsymbol{A}_{\cos}^{1\mathrm{D}}$, $\beta = 30$, mixed zone scheme (MZ)

(b) $\boldsymbol{A}_{\cos}^{1\mathrm{D}}$, $\beta = 30$, reduced zone scheme (RZ)

(c) $\boldsymbol{A}_{\cos}^{2\mathrm{D}}$, $\beta = 15$, reduced zone scheme (RZ)

Figure 3. Typical dispersions in the normal state at the valley $\rho = \pm$ with $\mu = 0$. (a,b) The 1D cosine potential $\boldsymbol{A}_{\cos}^{1\mathrm{D}}$ (shown for $\beta = 30$) (a) in the mixed zone scheme (MZ) and (b) in the reduced zone scheme (RZ). (c) The 2D cosine potential $\boldsymbol{A}_{\cos}^{2\mathrm{D}}$ (shown for $\beta = 15$) in the reduced zone scheme. The strained dispersions are shown in blue and for comparison the conical unstrained graphene dispersions in orange.

regime the order parameter is exponentially suppressed and we also start seeing quantum critical points [48]. We further see how in the flat-band regime the behavior of $\boldsymbol{A}_{\cos}^{2\mathrm{D}}$ with β is similar to that of $\boldsymbol{A}_{\cos}^{1\mathrm{D}}$ with 2β. Since in this paper we are mostly interested in the flat-band regime, we choose to calculate many of the following results at the fixed interaction strength $\lambda/(\hbar v_F d) = -0.01$, which is clearly in the flat-band regime except for $\boldsymbol{A}_{\cos}^{2\mathrm{D}}$ with $\beta = 10$, which is at the interface of the dispersive and flat-band regimes.

To further confirm that in the flat-band regime $\max \Delta$ is linear both in the interaction strength λ and the amplitude B of the pseudomagnetic field $\boldsymbol{B}$,

$$\max \Delta = -\zeta B \lambda, \tag{31}$$

we show the ratio ζ for all the potentials in figures 4(b,c) at $\mu = \mu_{\mathrm{opt}}$ and $T = 0$. In the flat-band regime ζ tends approximately to a constant $\zeta \approx 0.15$, which holds as long as $c \lesssim 20$. For $c \gtrsim 20$ we start seeing deviations from this result, with $\zeta \approx 0.05$ for the extreme case of $c = 100$. The small variation in ζ due to c even in the flat-band regime is most likely due to the fact that the maximum of Δ is not exactly the correct quantity to calculate, but it gives a very good estimate. We may compare this to the exact-flat-band result [14] with a constant Δ^{FB}, for
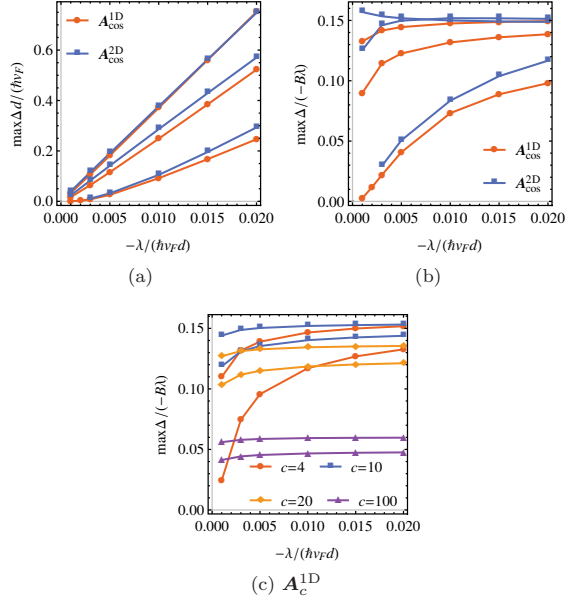

(a)

(b)

(c) $\boldsymbol{A}_c^{1\mathrm{D}}$

Figure 4. Behavior of the maximum of the superconducting order parameter Δ as a function of the interaction strength λ at optimal doping $\mu = \mu_{\mathrm{opt}}$ and $T = 0$. (a) Linearity of $\max \Delta$ in λ in the flat-band regime for the cosine potentials. Each potential has three curves corresponding to $\beta = 20, 30, 40$ (1D potential) or $\beta = 10, 15, 20$ (2D potential), from bottom to top. (b,c) The ratio $\max \Delta/(-\lambda B)$ as a function of λ for (b) the cosine potentials and (c) $\boldsymbol{A}_c^{1\mathrm{D}}$ with varying c, where B is the amplitude of the pseudomagnetic field $\boldsymbol{B}$. In (b) the curves are the same as in (a) while in (c) each c has two curves corresponding to $\beta = 30, 40$, from bottom to top. In the flat-band regime the ratio tends approximately to a constant as in (31).

which $\Delta^{\mathrm{FB}} = -\frac{1}{(2\pi)^2} n\Omega\lambda$ with $\Omega = 1/d^2$ and n being the number of flat bands. In PSG it is the amplitude B of the pseudomagnetic field $\boldsymbol{B}$ that effectively determines $n\Omega$, the number of approximate flat bands in the system with the SBZ area of $1/d^2$. A similar linear relationship was found out by Uchoa et al. [16] in the case of a homogeneous pseudomagnetic field $\boldsymbol{B}$, although with a somewhat larger prefactor ζ.

A. Order parameter profile, dispersion, and density of states

In figure 5 we show a cross section of the self-consistent $\Delta_{A/B}$ [as in figures 2(e,f)] along the line $(x, 0)$ [1D potentials] or $r(1, -1)$ [2D potential] for different potentials $\boldsymbol{A}$, strain strengths β, and slope parameters c. The effect of β is to simply linearly increase the amplitude of $\Delta_{A/B}$. On the other hand increasing c not only increases the amplitude of $\Delta_{A/B}$, but makes it also more localized. We
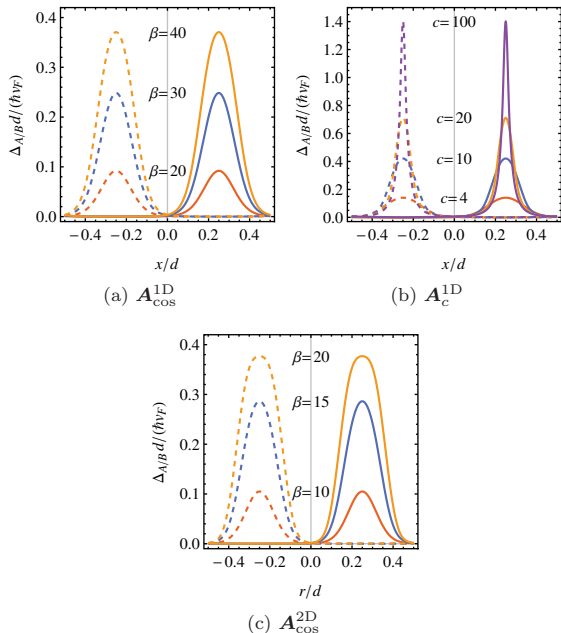
(a) $\boldsymbol{A}_{\cos}^{1D}$

(b) $\boldsymbol{A}_c^{1D}$

(c) $\boldsymbol{A}_{\cos}^{2D}$

Figure 5. Effect of the amplitude β and the slope parameter c on $\Delta_{A/B}$ (A solid, B dashed lines) at $\lambda/(\hbar v_F d) = -0.01$, and optimal doping $\mu = \mu_{\mathrm{opt}}$ ($\mu = 0$ produces the same Δ for such large λ). (a) Varying β of the 1D cosine potential $\boldsymbol{A}_{\cos}^{1D}$. (b) Varying the slope parameter c of the 1D potential $\boldsymbol{A}_c^{1D}$ with $\beta = 30$. (c) Varying β of the 2D cosine potential $\boldsymbol{A}_{\cos}^{2D}$. $\Delta_{A/B}$ is drawn along the line $(x,0)$ [1D potentials] or $r(1,-1)$ [2D potential].

also see that for the 2D potential $\boldsymbol{A}_{\cos}^{2D}$, $\Delta_{A/B}$ with the strain strength β along the diagonal behaves similarly as $\Delta_{A/B}$ in the x direction for the 1D potential $\boldsymbol{A}_{\cos}^{1D}$ with 2β.

These effects we can further see in the dispersions and densities of states in figures 6 and 7, respectively, which are plotted at $\mu = 0$ for clarity. In figure 6 we show the cross section of the dispersions in figure 3 along the line $(0, k_y)$ [1D potentials] or $k(1,1)$ [2D potential], both in the normal and superconducting states, and in the different schemes to allow for easier comparison between the 1D and 2D potentials. In figure 7 we show the corresponding densities of states (DOS). We clearly see in the normal state how increasing β and c both suppress the group velocity, thus increasing flatness of the bands. The density of states becomes correspondingly more and more peaked at zero energy. The superconducting energy gap also increases both with increasing β and c, and the peculiar double-peak structure in the superconducting DOS is also better revealed for higher β or c. In the 2D case it is notable how increasing β generates multiple peaks in the normal state DOS, and thus also in the superconducting state DOS, in a way that separates it
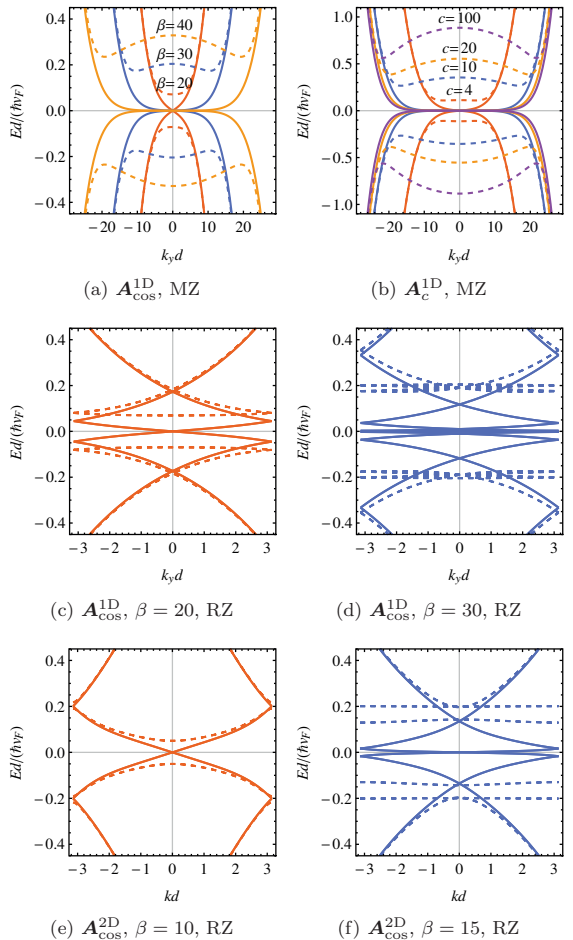

(a) $\boldsymbol{A}_{\cos}^{1D}$, MZ

(b) $\boldsymbol{A}_c^{1D}$, MZ

(c) $\boldsymbol{A}_{\cos}^{1D}$, $\beta = 20$, RZ

(d) $\boldsymbol{A}_{\cos}^{1D}$, $\beta = 30$, RZ

(e) $\boldsymbol{A}_{\cos}^{2D}$, $\beta = 10$, RZ

(f) $\boldsymbol{A}_{\cos}^{2D}$, $\beta = 15$, RZ

Figure 6. Effect of the strain strength β and the slope parameter c on the dispersion (normal state: solid, superconducting state: dashed lines) for the different potentials at $\mu = 0$. In the superconducting state $T = 0$ and $\lambda/(\hbar v_F d) = -0.01$. (a,b) Dispersions in the mixed zone scheme (MZ) along the line $(k_y, 0)$ for (a) $\boldsymbol{A}_{\cos}^{1D}$ with various β and for (b) $\boldsymbol{A}_c^{1D}$ with various c and fixed $\beta = 30$. (c,d) Corresponding dispersions in the reduced zone scheme (RZ) along the line $(0, k_y)$ for $\boldsymbol{A}_{\cos}^{1D}$ with $\beta = 20$ and 30, respectively. (e,f) Dispersions for $\boldsymbol{A}_{\cos}^{2D}$ along the diagonal line $k(1,1)$ in the reduced zone scheme for $\beta = 10$ and 15, respectively.

from the 1D potentials.

To determine more properties that could be measured $e.g.$ by STM [21, 58], we show in figure 8 the local densities of states (LDOS) along the line $(x, 0)$ [1D potentials] or $r(1, -1)$ [2D potential], which further illustrate the results discussed so far. In the normal state the overall energy dependence shows the clear peak at zero energy for the 1D potentials, as well as the multiple-peak structure for the 2D potential. In the superconducting state
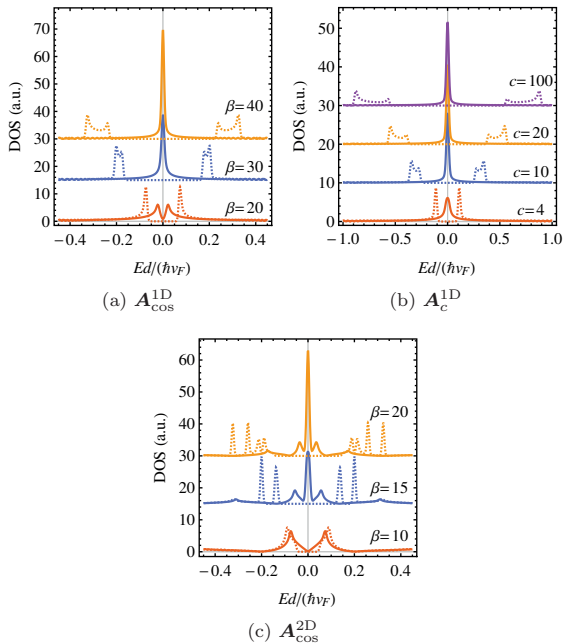
(a) $\boldsymbol{A}_{\cos}^{1D}$

(b) $\boldsymbol{A}_c^{1D}$

(c) $\boldsymbol{A}_{\cos}^{2D}$

Figure 7. Effect of the strain strength β and the slope parameter c on the density of states (DOS) at $\mu = 0$, $T = 0$, and $\lambda/(\hbar v_F d) = -0.01$ (normal state: solid, superconducting state: dashed lines) for (a) $\boldsymbol{A}_{\cos}^{1D}$, (b) $\boldsymbol{A}_c^{1D}$, and (c) $\boldsymbol{A}_{\cos}^{2D}$. For clarity the successive curves in the DOS plots are shifted vertically by 15 in (a,c) and by 10 in (b). Each curve is normalized such that the shown area integrates to unity.

the energy dependence also shows the superconducting gap, as already seen in the total DOS in figure 7. The position dependence gives us more information about the underlying strain field. They clearly show the high density of low-energy states near the points $x = \pm d/4$ (1D potentials) or $r = \pm d/4$ (2D potential), that is, points where $\boldsymbol{B}$ has extrema. Furthermore the states on the positive (negative) x or r side are those coming from the sublattice A (B), which, by comparison to Fig. 2(c–d), means that the A (B) sublattice states are localized at the minima (maxima) of $\boldsymbol{B}$. This kind of localization and sublattice polarization was also experimentally observed by Jiang et al. [21]. Since the low-energy states are the ones contributing to superconductivity, their localization explains the similar localization of the order parameter $\Delta_{A/B}$, as seen in figures 2(e,f).

In the normal state LDOS we further see the localization pattern splitting at higher energies for the 1D potentials. This is contrasted with the 2D potential, where the higher-energy peaks are separated not only in position but also in energy. Furthermore in the superconducting state LDOS we see the same behavior in the energy gaps as in the total DOS: increasing β or c leads to an increasing gap size, with the localization pattern staying

(a) $\boldsymbol{A}_{\cos}^{1D}$, $\beta = 20$, normal

(b) $\boldsymbol{A}_{\cos}^{1D}$, $\beta = 20$, SC

(c) $\boldsymbol{A}_{\cos}^{1D}$, $\beta = 30$, normal

(d) $\boldsymbol{A}_{\cos}^{1D}$, $\beta = 30$, SC

(e) $\boldsymbol{A}_{20}^{1D}$, $\beta = 30$, normal

(f) $\boldsymbol{A}_{20}^{1D}$, $\beta = 30$, SC

(g) $\boldsymbol{A}_{\cos}^{2D}$, $\beta = 10$, normal

(h) $\boldsymbol{A}_{\cos}^{2D}$, $\beta = 10$, SC

(i) $\boldsymbol{A}_{\cos}^{2D}$, $\beta = 15$, normal
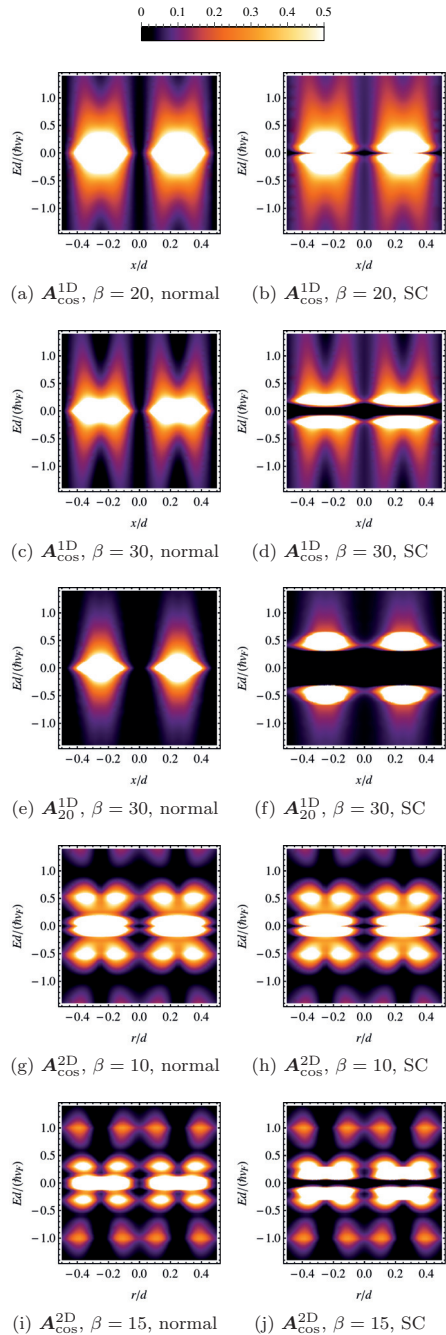
(j) $\boldsymbol{A}_{\cos}^{2D}$, $\beta = 15$, SC

Figure 8. Local density of states (LDOS) at $\mu = 0$ and $T = 0$ along the line $(x, 0)$ [1D potentials] or $r(1, -1)$ [2D potential] both in the (a,c,e,g,i) normal and (b,d,f,h,j) superconducting (SC) states. In the superconducting state $\lambda/(\hbar v_F d) = -0.01$. In each plot the states on the positive (negative) x or r side come from the sublattice A (B). Each plot is normalized such that the total visible area integrates to unity.
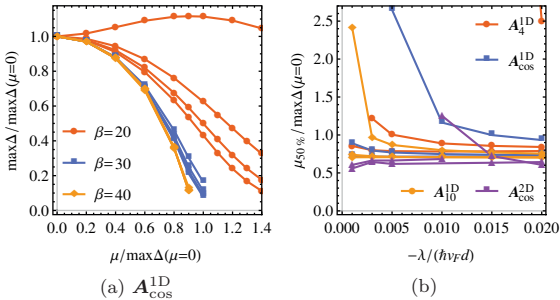
Figure 9. Solving the "critical" chemical potential $\mu_{50\%}$ at $T = 0$, where $\mu_{50\%}$ is the chemical potential where $\max \Delta$ has dropped to $\max \Delta(\mu = 0)/2$. (a) Normalized order parameter maximum $\max \Delta / \max \Delta(\mu = 0)$ as a function of the normalized chemical potential $\mu / \max \Delta(\mu = 0)$ for $\boldsymbol{A}_{\cos}^{1D}$ showing how doping away from the flat band, located at the DOS peak (which is at the zero energy in the flat-band regime and at a nonzero energy in the dispersive regime), kills superconductivity. The behavior is the same for $-\mu$. The four curves for each β are those for $-\lambda/(\hbar v_F d) = 0.005, 0.01, 0.015, 0.02$ (from top to bottom). (b) The ratio $\mu_{50\%}/\max \Delta(\mu = 0)$ as a function of λ for different potentials $\boldsymbol{A}$. Each 1D potential has three curves corresponding to $\beta = 20, 30, 40$ (from top to bottom). In the flat-band regime the ratio tends approximately to a constant as in (32).

the same. Again the 2D potential behaves slightly differently: the gap is largest at $r = \pm d/4$, while for the 1D potentials the gap at $x = \pm d/4$ is smallest.

B. Critical doping level and temperature

We can in principle calculate the critical doping level μ_c and the critical temperature T_c by solving the self-consistency equation (19) for various μ and T and by solving for the point where Δ vanishes. But since the fixed-point iteration scheme converges slowly when Δ is small, we calculate $\mu_{50\%}$ [$T_{50\%}$] instead, corresponding to the chemical potential [temperature] at which $\max \Delta$ has decreased to $\max \Delta(\mu = 0)/2$ [$\max \Delta(T = 0)/2$].

We show in figure 9(a) the μ-dependence of Δ at $T = 0$ in the case of $\boldsymbol{A}_{\cos}^{1D}$, from where $\mu_{50\%}$ is determined. We see how doping away from the flat band, which in the flat-band regime is located at zero energy, kills superconductivity. In this case $\mu_{50\%}$ approaches $\sim 0.7 \max \Delta(\mu = 0)$ in the flat-band limit. In the flat-band regime the results fit very well the relation $\max \Delta(\mu) = \sqrt{(\max \Delta(\mu = 0))^2 - (\mu/b)^2}$ with b as the fitting parameter, as compared to the result [59] $\Delta^{FB}(\mu) = \sqrt{\Delta^{FB}(\mu = 0)^2 - \mu^2}$ for exactly flat bands and homogeneous Δ^{FB}. On the other hand in the dispersive regime Δ is not maximized at zero chemical potential, but around $\mu \approx 0.9 \max \Delta(\mu = 0) \approx \frac{0.9}{1.1} \max \Delta(\mu = \mu_{opt}) \approx 0.02 \hbar v_F/d$ instead, which corresponds to the DOS peak position shown in figure 7(a). This is exactly

the same behavior as seen in TBG [14, 51]: in the flat-band regime the energy scale of Δ exceeds the DOS peak separation (the "bandwidth") and the smeared DOS is centered at zero energy, while in the dispersive regime Δ can "see" the double-peaked DOS because of the small energy scale of Δ. In TBG this might explain [14, 51] why superconductivity is observed at a nonzero doping level [5], and the same might happen also in PSG if the interaction strength λ is small enough. But note that in PSG we can in principle tune $\boldsymbol{A}$ (its functional dependence, β, c, and d) to move the interface between the flat-band and dispersive regimes so that superconductivity would be observed at zero doping.

To further verify that $\mu_{50\%}$ is linear in $\max \Delta(\mu = 0)$ in the flat-band regime,

$$\mu_{50\%} = \eta \max \Delta(\mu = 0), \qquad (32)$$

we show in figure 9(b) the ratio η at $T = 0$ for a selection of potentials. In the flat-band regime the ratio tends approximately to a constant $\eta \approx 0.7$ as long as $c \lesssim 10$. For $c \gtrsim 10$ we start seeing slight deviations from this, with $\eta \approx 0.6$ and 0.5 for $c = 20$ and 100, respectively. The critical chemical potential μ_c is slightly larger, approximately $\mu_c \approx \max \Delta(\mu = 0)$ for $\boldsymbol{A}_{\cos}^{1D}$ in the flat-band regime according to figure 9(a). This coincides with the case of perfectly flat bands and a constant Δ^{FB} for which [60, Supplemental Material of Ref. 14] $\mu_c^{FB} = \Delta^{FB}(\mu = 0)$.

In experiments the filling ν is more easily directly controlled [5] than the chemical potential μ, which we use as a parameter. While we do not calculate the filling, hints for the $\nu(\mu)$ dependence in PSG can be found from the work of Uchoa et al. [16] in the case of strained graphene with a homogeneous pseudomagnetic field [61] or from Refs. 14 and 52 in the case of TBG.

In figure 10 we show the corresponding plots for determining $T_{50\%}$ at $\mu = \mu_{opt}$. Again the ratio ξ in

$$k_B T_{50\%} = \xi \max \Delta(T = 0), \qquad (33)$$

tends approximately to a constant $\xi \approx 0.35$ in the flat-band regime as long as $c \lesssim 10$. For $c \gtrsim 10$ we start seeing deviations from this, with $\xi \approx 0.3$ for $c = 20$ and $\xi \approx 0.25$ for $c = 100$. The critical temperature T_c is slightly larger, approximately $k_B T_c \approx 0.4 \max \Delta(T = 0)$ for $\boldsymbol{A}_{\cos}^{1D}$ in the flat-band regime according to figure 10(a). For comparison, in the case of perfectly flat bands and a constant Δ^{FB} we have the result [10] $k_B T_c^{FB} = \frac{1}{2} \Delta^{FB}(T = 0)$ and in TBG [14] within the same interaction model $k_B T_c \approx 0.25 \max \Delta(T = 0)$.

C. Superfluid weight and Berezinskii–Kosterlitz–Thouless transition temperature

To determine the true superconducting transition temperature we calculate the superfluid weight D^s and the Berezinskii–Kosterlitz–Thouless transition temperature
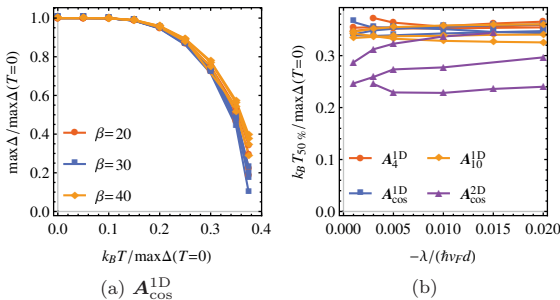
(a) $\boldsymbol{A}_{\cos}^{1D}$ (b)

Figure 10. Solving the "critical" temperature $T_{50\%}$ at optimal doping $\mu = \mu_{\mathrm{opt}}$, where $T_{50\%}$ is the temperature where $\max\Delta$ has dropped to $\max\Delta(T=0)/2$. (a) Normalized order parameter maximum $\max\Delta/\max\Delta(T=0)$ as a function of the normalized temperature $k_{\mathrm{B}}T/\max\Delta(T=0)$ for $\boldsymbol{A}_{\cos}^{1D}$. Each β has four curves corresponding to $-\lambda/(\hbar v_{\mathrm{F}}d) = 0.005, 0.01, 0.015, 0.02$. (b) The ratio $k_{\mathrm{B}}T_{50\%}/\max\Delta(T=0)$ as a function of λ for different potentials $\boldsymbol{A}$. Each $\boldsymbol{A}$ has three curves corresponding to $\beta = 20, 30, 40$ (1D potentials) or $\beta = 10, 15, 20$ (2D potential), with the outliers being those for the smallest β. In the flat-band regime the ratio tends approximately to a constant as in (33).

T_{BKT} from (20) and (21). In figure 11(a) we show the total superfluid weight $\sqrt{\det D^{\mathrm{s}}}$, together with the different components $D_{\mu\nu}^{\mathrm{s}}$, as a function of the interaction strength λ for $\boldsymbol{A}_{\cos}^{1D}$. The behavior is very similar to that of $\max\Delta$ in figure 4(a): it is linear in the flat-band regime and also increases linearly with increasing β. To further verify that $\sqrt{\det D^{\mathrm{s}}}$ is linear in $\max\Delta$,

$$\sqrt{\det D^{\mathrm{s}}} = \chi \max\Delta, \qquad (34)$$

we show the ratio χ in figure 11(c,d) at $\mu = \mu_{\mathrm{opt}}$ and $T = 0$. In the flat-band regime the ratio tends approximately to a constant $\chi \approx 0.15\ldots0.4$, which has more variation than η and ξ for $\mu_{50\%}$ and $T_{50\%}$ in the flat-band regime. For comparison, in TBG we found [52] within the same interaction model that $\chi \approx 0.35$ in the flat-band regime.

We may again compare (34) to the case of exactly flat bands and a constant Δ^{FB}. But since the superfluid weight depends heavily on the Hamiltonian itself and not only its eigenvalues, we need to specify which flat-band model to use. We take the "graphene flat-band limit", that is, graphene with $v_{\mathrm{F}} \to 0$. In this case [45, 49] $D_{\mathrm{FB}}^{\mathrm{s}} = \frac{2}{\pi}\Delta^{\mathrm{FB}}$ at $\mu = 0 \approx \mu_{\mathrm{opt}}$ and $T = 0$, which in fact holds for any v_{F}.

What is intriguing in figure 11(a) is that for the studied 1D potentials the superfluid weight is almost isotropic although the potentials are highly anisotropic. There is, however, a slight anisotropy, $D_{xx}^{\mathrm{s}} \neq D_{yy}^{\mathrm{s}}$ and $D_{xy}^{\mathrm{s}} = 0 = D_{yx}^{\mathrm{s}}$, visible for large β and λ. On the other hand the 2D potential produces an isotropic superfluid weight, $D_{xx}^{\mathrm{s}} = D_{yy}^{\mathrm{s}}$ and $D_{xy}^{\mathrm{s}} = 0 = D_{yx}^{\mathrm{s}}$. This (an)isotropy is consistent with the symmetries of the studied potentials. For comparison in TBG it was found [52] that local interaction
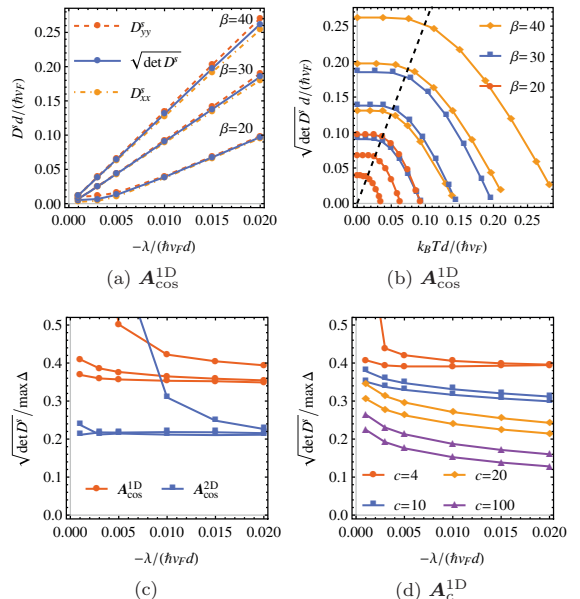

(a) $\boldsymbol{A}_{\cos}^{1D}$ (b) $\boldsymbol{A}_{\cos}^{1D}$

(c) (d) $\boldsymbol{A}_c^{1D}$

Figure 11. Behavior of the superfluid weight D^{s} at optimal doping $\mu = \mu_{\mathrm{opt}}$ and (a,c,d) $T = 0$. (a) D^{s} as a function of λ for $\boldsymbol{A}_{\cos}^{1D}$ showing linearity in the flat-band regime. The superfluid weight for the 1D potentials is slightly anisotropic, $D_{xx}^{\mathrm{s}} \neq D_{yy}^{\mathrm{s}}$, for large β and λ. For the 2D potential $\boldsymbol{A}_{\cos}^{2D}$ (not shown) the superfluid weight is isotropic, $D_{xx}^{\mathrm{s}} = D_{yy}^{\mathrm{s}}$. The off-diagonal components $D_{xy}^{\mathrm{s}} = 0 = D_{xy}^{\mathrm{s}}$ are zero for all the studied potentials. The (an)isotropy is consistent with the symmetries of the studied potentials. (b) $\sqrt{\det D^{\mathrm{s}}}$ as a function of temperature T for $\boldsymbol{A}_{\cos}^{1D}$. Each β has three curves corresponding to $-\lambda/(\hbar v_{\mathrm{F}}d) = 0.01, 0.015, 0.02$, from bottom to top. Also the dashed line $\sqrt{\det D^{\mathrm{s}}} = 8k_{\mathrm{B}}T/\pi$ is shown, from intersections of which T_{BKT} is determined through (21). (c,d) The ratio $\sqrt{\det D^{\mathrm{s}}}/\max\Delta$ as a function of the interaction strength λ for (c) the cosine potentials and (d) $\boldsymbol{A}_c^{1D}$ with varying c. In (c) each $\boldsymbol{A}$ has three curves corresponding to (from top to bottom) $\beta = 20, 30, 40$ (1D potentials) or $\beta = 10, 15, 20$ (2D potential), while in (d) each c has two curves corresponding to (from top to bottom) $\beta = 30, 40$. In the flat-band regime the ratio is approximately a constant depending slightly on the potential, as in (34).

always produces an isotropic superfluid weight, while the more complicated resonating valence bond (RVB) interaction was able to produce anisotropy through spontaneous symmetry breaking. The anisotropy could serve as one experimental signature for superconductivity described by the presented model, and it could be measured by radio frequency impedance spectroscopy [62] in a Hall-like four-probe setup [52].

Although in this work we do not separate the superfluid weight into the conventional and geometric contributions [45], from general knowledge [45] and calculations in TBG [52, 63] we expect the geometric contribution to dominate in the flat-band regime.
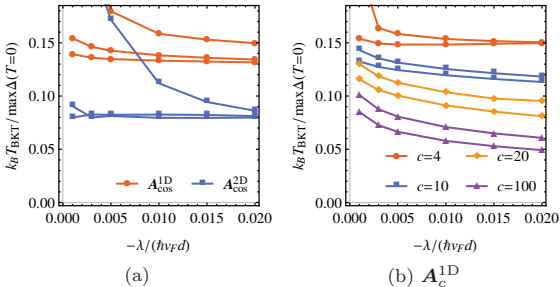
Figure 12. Behavior of the Berezinskii–Kosterlitz–Thouless transition temperature T_{BKT} at optimal doping $\mu = \mu_{\mathrm{opt}}$. The ratio $k_{\mathrm{B}}T_{\mathrm{BKT}}/\max\Delta(T=0)$ for (a) the cosine potentials and (b) $\boldsymbol{A}_c^{\mathrm{1D}}$ with varying c. In (a) both potentials have three curves corresponding to $\beta = 20, 30, 40$ (1D potential) or $\beta = 10, 15, 20$ (2D potential), from top to bottom in the flat-band regime, while in (b) each c has two curves corresponding to $\beta = 30, 40$, from top to bottom. In the flat-band regime the ratio tends approximately to a constant depending slightly on the potential, as in (35).

In figure 11(b) we further show $\sqrt{\det D^{\mathrm{s}}}$ as a function of temperature T for $\boldsymbol{A}_{\cos}^{\mathrm{1D}}$, from where T_{BKT} is determined through (21) by solving for the intersection point with the line $\sqrt{\det D^{\mathrm{s}}} = 8k_{\mathrm{B}}T/\pi$. We immediately see that in the flat-band regime $D^{\mathrm{s}}(T_{\mathrm{BKT}}) \approx D^{\mathrm{s}}(0)$ is a rather good approximation so that the self-consistency in (21) is not essential. This is very different from TBG [52], where the temperature dependence is essential due to T_{BKT} being closer to T_c. We nevertheless need to solve the full self-consistent equation for all the potentials, as the relative magnitude of T_c and T_{BKT} is not known beforehand.

The resulting ratio $k_{\mathrm{B}}T_{\mathrm{BKT}}/\max\Delta(T=0)$ is shown in figure 12 for the different potentials at $\mu = \mu_{\mathrm{opt}}$, further confirming that $D^{\mathrm{s}}(T_{\mathrm{BKT}}) \approx D^{\mathrm{s}}(0)$: apart from the different scale, the T_{BKT} plots in figure 12 are very similar to the D^{s} plots in figures 11(c,d). Furthermore in the linear relation

$$k_{\mathrm{B}}T_{\mathrm{BKT}} = \kappa \max\Delta(T=0), \qquad (35)$$

the ratio κ tends approximately to a constant $\kappa \approx 0.05\ldots0.15$ in the flat-band regime. Again in (35) we see similarity to the "graphene flat-band limit" result with a homogeneous Δ^{FB}, for which $k_{\mathrm{B}}T_{\mathrm{BKT}}^{\mathrm{FB}} = \frac{\pi}{8}D_{\mathrm{FB}}^{\mathrm{s}}(T_{\mathrm{BKT}}^{\mathrm{FB}}) \approx \frac{1}{4}\Delta^{\mathrm{FB}}(T=0)$ at $\mu = 0 \approx \mu_{\mathrm{opt}}$ if we furthermore assume $D_{\mathrm{FB}}^{\mathrm{s}}(T_{\mathrm{BKT}}^{\mathrm{FB}}) \approx D_{\mathrm{FB}}^{\mathrm{s}}(0)$.

Combining (34) and (35) we get in the flat-band regime at $\mu = \mu_{\mathrm{opt}}$ the ratio $T_{\mathrm{BKT}}/T_{50\%} = \kappa/\xi \approx 0.2\ldots0.4$ depending on the potential. For $\boldsymbol{A}_{\cos}^{\mathrm{1D}}$ this yields $T_{\mathrm{BKT}}/T_{50\%} \approx 0.4$, and within the same accuracy $T_{\mathrm{BKT}}/T_c \approx 0.4$. For comparison in TBG we found within the same interaction model in the flat-band regime $k_{\mathrm{B}}T_{\mathrm{BKT}} \approx 0.16\ldots0.2\max\Delta(T=0)$ [52] (depending slightly on λ), $k_{\mathrm{B}}T_c \approx 0.25\max\Delta(T=0)$ [14], and thus $T_{\mathrm{BKT}}/T_c \approx 0.6\ldots0.8$.

By combining (31) and (35) we get $k_{\mathrm{B}}T_{\mathrm{BKT}} = -\kappa\zeta B\lambda$ at $\mu = \mu_{\mathrm{opt}}$. Let us calculate an estimate of T_{BKT} by using $\lambda = -1\,\mathrm{eV}a^2 \approx -6\,\mathrm{eV\AA}^2$, which roughly corresponds [14, 52] to $T_{\mathrm{BKT}} \approx 1\,\mathrm{K}$ measured in TBG [5]. Here $a = \sqrt{3}a_0 \approx 2.46\,\mathrm{\AA}$ is the graphene lattice constant. For $\boldsymbol{A}_{\cos}^{\mathrm{1D}}$ we have $B = 2\pi\beta/d^2$ and in the flat-band regime $\kappa = 0.15$ and $\zeta = 0.15$, yielding a similar $T_{\mathrm{BKT}} \approx 1\,\mathrm{K}$ if we apply strain for example such that $\beta = 40$ and $d = 60\,\mathrm{nm}$ [then $\lambda/(\hbar v_F d) \approx -0.002$ if using $v_F = 1 \times 10^6\,\mathrm{m/s}$, which is in the flat-band regime according to figures 4(b) and 12(a)]. In the case of the in-plane displacement field $\boldsymbol{u}_{\cos}^{\mathrm{1D}}$ (28) this corresponds to the displacement amplitude $\beta a_0/(\beta_G\pi) \approx 1\,\mathrm{nm}$ if $\beta_G = 2$. Since in this case the elasticity theory assumes $\beta/\beta_G \ll d/a_0$ and $d/a_0 \gg 1$ [22], we are very well in the validity regime. On the other hand, if we are able to decrease the strain period to $d = 10\,\mathrm{nm}$ [then $\lambda/(\hbar v_F d) = -0.009$], we get to a high-temperature superconductor value of $T_{\mathrm{BKT}} \approx 40\,\mathrm{K}$, which is still in the validity regime. Note the optimization problem in increasing T_{BKT}: decreasing d directly enhances T_{BKT} but at the same time it makes the validity limit for β tighter, while at the same time we should have as large β as possible. But this might only be a limiting factor in our linear elasticity theory, while a more complete microscopic theory could, perhaps, yield a result that increasing β or decreasing d always increases T_{BKT}.

The experiments of Jiang et al. [21] with $\frac{\hbar}{e}B \approx 100\,\mathrm{T}$ and $d = 14\,\mathrm{nm}$ can be described by the 1D cosine potential with $\beta \approx 5$. When $\lambda = -6\,\mathrm{eV\AA}^2$, $\lambda/(\hbar v_F d) \approx -0.007$ is not in the flat-band regime. Hence T_{BKT} cannot be obtained from the simple estimate used above, and is likely much lower than $1\,\mathrm{K}$. Increasing the strain amplitude by a factor of 4, so that $\beta = 20$, would yield $\zeta = 0.05$, $\kappa = 0.17$, and thus $T_{\mathrm{BKT}} \approx 0.007\hbar v_F/(dk_{\mathrm{B}}) \approx 4\,\mathrm{K}$. Further decreasing the period to $d = 8\,\mathrm{nm}$, a period which was already observed by Jiang et al., would yield already $\lambda/(\hbar v_F d) \approx -0.01$ and thus $T_{\mathrm{BKT}} \approx 11\,\mathrm{K}$.

IV. CONCLUSIONS

We have studied both the normal and superconducting s-wave state properties of periodically strained graphene (PSG) in the continuum low-energy model. We have shown that periodic strain might be a mechanism that allows increasing the critical temperature T_c higher than a few kelvin, observed in doped graphene and in twisted bilayer graphene (TBG), or possibly even to tens of kelvins. Especially we have generalized the results of Kauppila et al. [19], where the authors studied the same problem in the case of a 1D cosine-like pseudo vector potential $\boldsymbol{A}$, to potentials with arbitrary shape and dimension. We furthermore calculated the superfluid weight and the Berezinskii–Kosterlitz–Thouless transition temperature T_{BKT} to determine the true transition temperature observed in experiments. In the normal state we

observed flat bands in the spectrum and localization of low-energy states near the extremum points of the effective pseudomagnetic field $\boldsymbol{B} = \nabla \times \boldsymbol{A}$.

We modelled the superconducting state by the Bogoliubov–de Gennes mean-field theory assuming a local interaction between the Cooper pairs leading to s-wave pairing. Because of the inhomogeneous strain field we observed a highly inhomogeneous order parameter $\Delta_{A/B}$ that is localized near the extremum points of $\boldsymbol{B}$, similarly to the localization of the low-energy states. We also noticed how the superconducting T_c or T_{BKT} can be linearly increased by increasing the strain strength β, decreasing the period d, or by increasing the slope (near the extremum points of $\boldsymbol{B}$) of the corresponding pseudo vector potential $\boldsymbol{A}$. On the other hand increasing the slope makes the order parameter also more localized.

While between the 1D potentials we observed only quantitative differences in the results, for the 2D cosine potential we saw also some qualitative differences when compared to the 1D potentials. The main differences are the localization pattern of $\Delta_{A/B}$, and the more peaked structure of the (local) density of states both in the normal and superconducting states. In the 2D case we studied only the cosine potential, but on the other hand the qualitative similarity in the results between the different 1D potentials gives us certainty that changing the shape of the potential would not change the qualitative results in the 2D case neither. However, it should be noted that it is the shape of $\boldsymbol{B}$ that matters and not that of the potential $\boldsymbol{A}$ itself, and thus even a 2D potential can produce results that are effectively those of a 1D potential.

We chose all our potentials to be periodic in a square (super)lattice, but note that any other lattice could be chosen as well, with different shapes and different periodicities in the two directions. Properties of this lattice are then directly seen in the dispersion, as well as in the localization of $\boldsymbol{B}$ and $\Delta_{A/B}$. We also observed the symmetry $\Delta_B(\boldsymbol{r}) = \Delta_A(-\boldsymbol{r})$ of the order parameter for all the chosen potentials. This is due to the inversion symmetry $\boldsymbol{A}(\boldsymbol{r}) = \boldsymbol{A}(-\boldsymbol{r})$ present in all of them. The relative magnitude between Δ_A and Δ_B can then be tuned by breaking this symmetry, *e.g.* by using a sawtooth-wave potential.

We also observed some very peculiar structures in the (local) density of states, which could serve as an experimental fingerprint of the physics described by this model. We furthermore found that in the flat-band regime the superconducting order parameter maximum $\max \Delta$ at $\mu = \mu_{opt}$ and $T = 0$, the "critical" chemical potential $\mu_{50\%}$ at $T = 0$, the "critical" temperature $T_{50\%}$ at $\mu = \mu_{opt}$, the superfluid weight $\sqrt{\det D^s}$ at $\mu = \mu_{opt}$ and $T = 0$, and the BKT transition temperature T_{BKT} at $\mu = \mu_{opt}$ are all approximately linear in the interaction strength λ. The linear relations, instead of exponential ones in usual bulk superconductors, suggest that high-temperature superconductivity might be possible in PSG.

As is known from the closely-related materials twisted bilayer graphene [5, 7, 11, 64–66], twisted trilayer graphene [67], twisted double bilayer graphene [68–70], rhombohedral graphite [71], or other graphite-based composites [72–76], also other phases like correlated insulators might be present. These are obviously excluded from the present study, but as we showed in previous studies [14, 52], the superconductivity-only model gives a plausible explanation for the observed superconducting states in TBG. This view of competing phases is supported by recent experiments where superconductivity could be seen without the correlated insulating phases [65, 66]. Thus we expect our similar model to work also in PSG when concentrating only on superconductivity. If the competing phase (if any) is magnetic, we know from a recent study [77] that in a pure flat-band system superconductivity is favored over magnetism whenever (in the weak-coupling regime) the effective attractive electron–electron interaction strength $\hat{\lambda}\hbar\omega_c = [g^2/(\hbar\omega_c)](\Omega_{FB}/\Omega_{BZ})$ is stronger than the repulsive one $u = U\Omega_{FB}/\Omega_{BZ}$. Here g is the electron–phonon coupling constant, U is the repulsive Hubbard coupling constant, $\hbar\omega_c$ is the characteristic phonon energy (in this case the Einstein energy $\hbar\omega_E$), and Ω_{FB}/Ω_{BZ} is the ratio of the flat-band area to the Brillouin zone area. The competition between superconductivity and magnetism was studied also by Löthman & Black-Schaffer [60], who showed how the two phases in the flat-band regime have somewhat different dependence on the doping.

An interesting future prospect would be to study the other phases which, by the analogue of TBG, are highly probable. Secondly the combination of moiré [14, 52] and strain [this work] physics would perhaps advance the understanding of superconductivity in TBG, where intrinsic periodic strain is inevitable. Thirdly, generalizing the Eliashberg formalism [77, 78] to the case of PSG with inhomogeneous superconductivity would make handling the electron–phonon interaction more accurate, especially in the strong-coupling limit. From the experimental point of view the challenge is to manufacture periodically strained graphene samples with large amplitudes and small periods and to perform low-temperature conductivity measurements in this (electrically doped) system to reveal the possible superconducting and/or correlated insulator states. The periodic strain and flat bands observed by Jiang *et al.* [21] are already an intriguing starting point, but according to our calculations a T_{BKT} of the order of $4\,\mathrm{K}$ would need a strain amplitude 4 times larger than in the experiment. On the other hand, further decreasing the period to $8\,\mathrm{nm}$ would yield already $T_{BKT} \approx 11\,\mathrm{K}$.

puter capacity from the Finnish Grid and Cloud Infras-

tructure (persistent identifier urn:nbn:fi:research-infras-2016072533).

[1] B. M. Ludbrook, G. Levy, P. Nigge, M. Zonno, M. Schneider, D. J. Dvorak, C. N. Veenstra, S. Zhdanovich, D. Wong, P. Dosanjh, C. Straßer, A. Stöhr, S. Forti, C. R. Ast, U. Starke, and A. Damascelli, Proc. Natl. Acad. Sci. **112**, 11795 (2015).

[2] J. Chapman, Y. Su, C. A. Howard, D. Kundys, A. N. Grigorenko, F. Guinea, A. K. Geim, I. V. Grigorieva, and R. R. Nair, Sci. Rep. **6**, 23254 (2016).

[3] S. Ichinokura, K. Sugawara, A. Takayama, T. Takahashi, and S. Hasegawa, ACS Nano **10**, 2761 (2016).

[4] A. P. Tiwari, S. Shin, E. Hwang, S.-G. Jung, T. Park, and H. Lee, J. Phys. Condens. Matter **29**, 445701 (2017).

[5] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Nature **556**, 43 (2018).

[6] M. Yankowitz, K. Watanabe, T. Taniguchi, P. San-Jose, and B. J. LeRoy, Nat. Commun. **7**, 13168 (2016).

[7] X. Lu, P. Stepanov, W. Yang, M. Xie, M. A. Aamir, I. Das, C. Urgell, K. Watanabe, T. Taniguchi, G. Zhang, A. Bachtold, A. H. MacDonald, and D. K. Efetov, Nature **574**, 653 (2019).

[8] S. Moriyama, Y. Morita, K. Komatsu, K. Endo, T. Iwasaki, S. Nakaharai, Y. Noguchi, Y. Wakayama, E. Watanabe, D. Tsuya, K. Watanabe, and T. Taniguchi, (2019), arXiv:1901.09356.

[9] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).

[10] T. T. Heikkilä, N. B. Kopnin, and G. E. Volovik, JETP Lett. **94**, 233 (2011).

[11] M. Yankowitz, S. Chen, H. Polshyn, Y. Zhang, K. Watanabe, T. Taniguchi, D. Graf, A. F. Young, and C. R. Dean, Science **363**, 1059 (2019).

[12] R. Bistritzer and A. H. MacDonald, Proc. Natl. Acad. Sci. **108**, 12233 (2011).

[13] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, Phys. Rev. B **86**, 155449 (2012).

[14] T. J. Peltonen, R. Ojajärvi, and T. T. Heikkilä, Phys. Rev. B **98**, 220504(R) (2018).

[15] F. Guinea, M. I. Katsnelson, and M. A. H. Vozmediano, Phys. Rev. B **77**, 075422 (2008).

[16] B. Uchoa and Y. Barlas, Physical Review Letters **111**, 046604 (2013).

[17] E. Tang and L. Fu, Nat. Phys. **10**, 964 (2014).

[18] J. W. F. Venderbos and L. Fu, Phys. Rev. B **93**, 195126 (2016).

[19] V. J. Kauppila, F. Aikebaier, and T. T. Heikkilä, Phys. Rev. B **93**, 214505 (2016).

[20] M. Tahir, O. Pinaud, and H. Chen, (2018), arXiv:1808.10046.

[21] Y. Jiang, M. Anđelković, S. P. Milovanović, L. Covaci, X. Lai, Y. Cao, K. Watanabe, T. Taniguchi, F. M. Peeters, A. K. Geim, and E. Y. Andrei, (2019), arXiv:1904.10147.

[22] See Supplementary Material (SM) for definition of the Fourier series in the different schemes, details on the derivation of the strained graphene Hamiltonian, details on the derivation of the superconductivity theory, and

[23] H. Suzuura and T. Ando, Phys. Rev. B **65**, 235412 (2002).

[24] M. A. H. Vozmediano, M. I. Katsnelson, and F. Guinea, Phys. Rep. **496**, 109 (2010).

[25] R. Ilan, A. G. Grushin, and D. I. Pikulin, Nat. Rev. Phys. **2**, 29 (2020).

[26] F. Guinea, M. I. Katsnelson, and A. K. Geim, Nat. Phys. **6**, 30 (2010).

[27] N. Levy, S. A. Burke, K. L. Meaker, M. Panlasigui, A. Zettl, F. Guinea, A. H. C. Neto, and M. F. Crommie, Science **329**, 544 (2010).

[28] A. Johansson, P. Myllyperkiö, P. Koskinen, J. Aumanen, J. Koivistoinen, H.-C. Tsai, C.-H. Chen, L.-Y. Chang, V.-M. Hiltunen, J. J. Manninen, W. Y. Woon, and M. Pettersson, Nano Lett. **17**, 6469 (2017).

[29] P. Koskinen, K. Karppinen, P. Myllyperkiö, V.-M. Hiltunen, A. Johansson, and M. Pettersson, J. Phys. Chem. Lett. **9**, 6179 (2018).

[30] P. Jia, W. Chen, J. Qiao, M. Zhang, X. Zheng, Z. Xue, R. Liang, C. Tian, L. He, Z. Di, and X. Wang, Nat. Commun. **10**, 3127 (2019).

[31] Z. H. Aitken and R. Huang, J. Appl. Phys. **107**, 123531 (2010).

[32] Y. Jiang, J. Mao, J. Duan, X. Lai, K. Watanabe, T. Taniguchi, and E. Y. Andrei, Nano Lett. **17**, 2839 (2017).

[33] K. Yang, Y. Chen, F. Pan, S. Wang, Y. Ma, and Q. Liu, Materials **9**, 32 (2016).

[34] C. Androulidakis, E. N. Koukaras, O. Frank, G. Tsoukleri, D. Sfyris, J. Parthenios, N. Pugno, K. Papagelis, K. S. Novoselov, and C. Galiotis, Sci. Rep. **4**, 5271 (2015).

[35] E. N. Koukaras, C. Androulidakis, G. Anagnostopoulos, K. Papagelis, and C. Galiotis, Extreme Mech. Lett. **8**, 191 (2016).

[36] P. Koskinen, Appl. Phys. Lett. **104**, 101902 (2014).

[37] L. Tarruell, D. Greif, T. Uehlinger, G. Jotzu, and T. Esslinger, Nature **483**, 302 (2012).

[38] M. M. van Wijk, A. Schuring, M. I. Katsnelson, and A. Fasolino, 2D Mater. **2**, 34010 (2015).

[39] S. Dai, Y. Xiang, and D. J. Srolovitz, Nano Lett. **16**, 5923 (2016).

[40] N. N. T. Nam and M. Koshino, Phys. Rev. B **96**, 075311 (2017).

[41] X. Lin, D. Liu, and D. Tománek, Phys. Rev. B **98**, 195432 (2018).

[42] S. Fang, S. Carr, Z. Zhu, D. Massatt, and E. Kaxiras, (2019), arXiv:1908.00058.

[43] H. Yoo, R. Engelke, S. Carr, S. Fang, K. Zhang, P. Cazeaux, S. H. Sung, R. Hovden, A. W. Tsen, T. Taniguchi, K. Watanabe, G.-C. Yi, M. Kim, M. Luskin, E. B. Tadmor, E. Kaxiras, and P. Kim, Nat. Mater. **18**, 448 (2019).

[44] S. Peotta and P. Törmä, Nat. Commun. **6**, 8944 (2015).

[45] L. Liang, T. I. Vanhala, S. Peotta, T. Siro, A. Harju, and

P. Törmä, Phys. Rev. B **95**, 024515 (2017).

[46] E. Zhao and A. Paramekanti, Phys. Rev. Lett. **97**, 230404 (2006).

[47] B. Uchoa and A. H. Castro Neto, Phys. Rev. Lett. **98**, 146801 (2007).

[48] N. B. Kopnin and E. B. Sonin, Phys. Rev. Lett. **100**, 246808 (2008).

[49] N. B. Kopnin and E. B. Sonin, Phys. Rev. B **82**, 014516 (2010).

[50] M. V. Hosseini, EPL **110**, 47010 (2015).

[51] F. Wu, A. H. MacDonald, and I. Martin, Phys. Rev. Lett. **121**, 257001 (2018).

[52] A. Julku, T. J. Peltonen, L. Liang, T. T. Heikkilä, and P. Törmä, Phys. Rev. B **101**, 060505(R) (2020).

[53] D. R. Nelson and J. M. Kosterlitz, Phys. Rev. Lett. **39**, 1201 (1977).

[54] Y. Cao, S.-H. Zou, X.-J. Liu, S. Yi, G.-L. Long, and H. Hu, Phys. Rev. Lett. **113**, 115302 (2014).

[55] Y. Xu and C. Zhang, Phys. Rev. Lett. **114**, 110401 (2015).

[56] The MATHEMATICA code we have used to compute the normal-state properties, the mean-field order parameter, the superfluid weight, and the BKT transition temperature is available at https://gitlab.jyu.fi/jyucmt/psg-and-tbg.

[57] G. Tarnopolsky, A. J. Kruchkov, and A. Vishwanath, Phys. Rev. Lett. **122**, 106405 (2019).

[58] Y. Xie, B. Lian, B. Jäck, X. Liu, C.-L. Chiu, K. Watanabe, T. Taniguchi, B. A. Bernevig, and A. Yazdani, Nature **572**, 101 (2019).

[59] T. T. Heikkilä and G. E. Volovik, in *Basic Physics of Functionalized Graphite*, edited by P. Esquinazi (Springer, 2016) pp. 123–143, arXiv:1504.05824.

[60] T. Löthman and A. M. Black-Schaffer, Phys. Rev. B **96**, 064505 (2017).

[61] When comparing to Uchoa *et al.* [16], in their language we concentrate on the zeroth pseudo-Landau level, which in Fig. 1 of Uchoa *et al.* means $-1 \leq \mu/(\hbar\omega_c) \leq 1$ or $-2 \leq \nu \leq 2$.

[62] F. Chiodi, M. Ferrier, K. Tikhonov, P. Virtanen, T. T. Heikkilä, M. Feigelman, S. Guéron, and H. Bouchiat, Sci. Rep. **1**, 3 (2011).

[63] X. Hu, T. Hyart, D. I. Pikulin, and E. Rossi, Phys. Rev. Lett. **123**, 237002 (2019).

[64] Y. Cao, V. Fatemi, A. Demir, S. Fang, S. L. Tomarken, J. Y. Luo, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, E. Kaxiras, R. C. Ashoori, and P. Jarillo-Herrero, Nature **556**, 80 (2018).

[65] P. Stepanov, I. Das, X. Lu, A. Fahimniya, K. Watanabe, T. Taniguchi, F. H. L. Koppens, J. Lischner, L. Levitov, and D. K. Efetov, (2019), arXiv:1911.09198.

[66] Y. Saito, J. Ge, K. Watanabe, T. Taniguchi, and A. F. Young, Nat. Phys. (2020).

[67] K.-T. Tsai, X. Zhang, Z. Zhu, Y. Luo, S. Carr, M. Luskin, E. Kaxiras, and K. Wang, (2019), arXiv:1912.03375.

[68] C. Shen, Y. Chu, Q. Wu, N. Li, S. Wang, Y. Zhao, J. Tang, J. Liu, J. Tian, K. Watanabe, T. Taniguchi, R. Yang, Z. Y. Meng, D. Shi, O. V. Yazyev, and G. Zhang, Nat. Phys. **16**, 520 (2020).

[69] Y. Cao, D. Rodan-Legrain, O. Rubies-Bigorda, J. M. Park, K. Watanabe, T. Taniguchi, and P. Jarillo-Herrero, Nature **583**, 215 (2020).

[70] X. Liu, Z. Hao, E. Khalaf, J. Y. Lee, Y. Ronen, H. Yoo, D. Haei Najafabadi, K. Watanabe, T. Taniguchi, A. Vishwanath, and P. Kim, Nature **583**, 221 (2020).

[71] Y. Shi, S. Xu, Y. Yang, S. Slizovskiy, S. V. Morozov, S.-K. Son, S. Ozdemir, C. Mullan, J. Barrier, J. Yin, A. I. Berdyugin, B. A. Piot, T. Taniguchi, K. Watanabe, V. I. Fal'ko, K. S. Novoselov, A. K. Geim, and A. Mishchenko, (2019), arXiv:1911.04565.

[72] R. R. da Silva, J. H. S. Torres, and Y. Kopelevich, Phys. Rev. Lett. **87**, 147001 (2001).

[73] H.-P. Yang, H.-H. Wen, Z.-W. Zhao, and S.-L. Li, Chin. Phys. Lett. **18**, 1648 (2001).

[74] S. Moehlecke, Y. Kopelevich, and M. B. Maple, Phys. Rev. B **69**, 134519 (2004).

[75] I. Felner and Y. Kopelevich, Phys. Rev. B **79**, 233409 (2009).

[76] C. Hwang, S. A. Cybart, S. J. Shin, S. Kim, K. Kim, T. G. Rappoport, S. M. Wu, C. Jozwiak, A. V. Fedorov, S.-K. Mo, D.-H. Lee, B. I. Min, E. E. Haller, R. C. Dynes, A. H. Castro Neto, and A. Lanzara, Sci. Rep. **6**, 21460 (2016).

[77] R. Ojajärvi, T. Hyart, M. A. Silaev, and T. T. Heikkilä, Phys. Rev. B **98**, 054515 (2018).

[78] G. Eliashberg, Sov. Phys. JETP **11**, 696 (1960).

[79] X. Zheng, *Efficient Fourier Transforms on Hexagonal Arrays*, PhD thesis, University of Florida (2007).

[80] Y. V. Nazarov and J. Danon, *Advanced Quantum Mechanics: A practical guide*, 1st ed. (Cambridge University Press, New York, 2013).

[81] C. W. J. Beenakker, Phys. Rev. Lett. **97**, 067007 (2006).

[82] M. Titov and C. W. J. Beenakker, Phys. Rev. B **74**, 041401(R) (2006).

Flat-band superconductivity in periodically strained graphene: mean-field and Berezinskii–Kosterlitz–Thouless transition: supplementary material

Teemu J. Peltonen[1] and Tero T. Heikkilä[1]

[1]*University of Jyväskylä, Department of Physics and Nanoscience Center,*
P.O. Box 35 (YFL), FI-40014 University of Jyväskylä, Finland

I. NOTATION

- $\alpha \in \{A, B\}$ denotes sublattice, $\sigma \in \{\uparrow, \downarrow\}$ spin, and $\rho \in \{+, -\}$ valley (*i.e.* $+\boldsymbol{K}$ or $-\boldsymbol{K} = \boldsymbol{K}'$)

- The bar operator exchanges sublattices, spins, valleys, and vectors: $\bar{A} = B$, $\bar{B} = A$, $\bar{\uparrow} = \downarrow$, $\bar{\downarrow} = \uparrow$, $\bar{\rho} = -\rho$, $\bar{\boldsymbol{k}} = -\boldsymbol{k}$

- s is a sign function for σ: $s(\uparrow) = +1$, $s(\downarrow) = -1$

- σ_x, σ_y, σ_z are the Pauli matrices in sublattice space

- $\boldsymbol{\sigma}^\rho := (\rho \sigma_x, \sigma_y)$ is a vector of Pauli matrices

- τ_x, τ_y, τ_z are the Pauli matrices in Nambu space

- $\boldsymbol{v}_1 \times \boldsymbol{v}_2$ is the 2-dimensional "cross product" of $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathbb{R}^2$ (= third component of $(\boldsymbol{v}_1, 0) \times (\boldsymbol{v}_2, 0)$ = signed area of the parallelogram defined by $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$).

- z^* is the complex conjugate of $z \in \mathbb{C}$ and L^* is the reciprocal lattice of a lattice L

- $|z|$ is the absolute value of $z \in \mathbb{C}$ and $|K|$ is the "measure" of a set K, *i.e.*

$$|K| = \begin{cases} \#K, & \text{if } K \text{ is discrete} \\ \text{length of } K, & \text{if } K \text{ is continuous and 1 dimensional} \\ \text{area of } K, & \text{if } K \text{ is continuous and 2 dimensional} \end{cases}$$

- $G/H := \{[g] : g \in G\}$ denotes the quotient group of the group G modulo a normal subgroup H, consisting of the equivalence classes $[g] := g + H = \{g + h : h \in H\}$ of the representative $g \in G$, assuming the group operation of addition. Especially if G is a lattice and H is its sublattice/superlattice, G/H identifies all the lattice points whose difference is in H, and thus G/H is isomorphic to the unit cell of G modulo H. Thus by dropping the brackets from $[g]$, under an isomorphism g might mean either the equivalence class or the representative in the unit cell, and from the context it should be clear which interpretations are allowed.

- $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are the primitive vectors of the graphene lattice $L := \text{span}_\mathbb{Z}\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$, $a := \|\boldsymbol{a}_1\| = \|\boldsymbol{a}_2\|$ is the lattice constant, the nearest-neighbor vectors are $\boldsymbol{\delta}_1$, $\boldsymbol{\delta}_2$, and $\boldsymbol{\delta}_3$, the carbon–carbon bond length is $a_0 := \|\boldsymbol{\delta}_j\| = a/\sqrt{3}$, and $\boldsymbol{\delta}_A := 0$, $\boldsymbol{\delta}_B := \boldsymbol{\delta}_1$ is the sublattice-translation vector

- L_{BK} denotes the large Born–von Kármán lattice, in translations of which the creation/annihilation operators are taken to be periodic

- $V := \left| \mathbb{R}^2/L_{\text{BK}} \right|$ is the area of the continuum Born–von Kármán unit cell

- $SL := \text{span}_\mathbb{Z}\{\boldsymbol{t}_1, \boldsymbol{t}_2\}$ denotes the superlattice created by periodic strain, $SL^* = \text{span}_\mathbb{Z}\{\boldsymbol{G}_1, \boldsymbol{G}_2\}$ is its reciprocal lattice, $SL_1 := \text{span}_\mathbb{Z}\{\boldsymbol{t}_1\}$ is a 1-dimensional sublattice of SL, $SL_1^* = \text{span}_\mathbb{Z}\{\boldsymbol{G}_1\}$ is a 1-dimensional sublattice of SL^*

- $S \in \{\text{RZ}, \text{MZ}\}$ denotes either the reduced zone scheme or the mixed zone scheme, explained in section II B, and $SL_{\text{RZ}} := SL$, $SL_{\text{MZ}} := SL_1$ are shorthand notations for writing the Fourier series in different schemes

- Op is the space of electron creation/annihilation operators (for the sake of notation)

- $\psi_\sigma : [L \cup (L + \boldsymbol{\delta}_1)]/L_{\mathrm{BK}} \to \mathrm{Op}$ is an electron annihilation operator defined in the graphene crystal (in the union of A sublattice L and B sublattice $L + \boldsymbol{\delta}_1$). In other words, if $\boldsymbol{r} \in L/L_{\mathrm{BK}}$, then $\psi_\sigma(\boldsymbol{r})$ annihilates a σ-spin electron at the (A-sub)lattice point $\boldsymbol{r}$ and $\psi_\sigma(\boldsymbol{r} + \boldsymbol{\delta}_1)$ annihilates a σ-spin electron at the B-sublattice point $\boldsymbol{r} + \boldsymbol{\delta}_1$.

- $\psi_{\sigma,A}, \psi_{\sigma,B} : L/L_{\mathrm{BK}} \to \mathrm{Op}$, $\psi_{\sigma,A}(\boldsymbol{r}) \coloneqq \psi_\sigma(\boldsymbol{r})$, $\psi_{\sigma,B}(\boldsymbol{r}) \coloneqq \psi_\sigma(\boldsymbol{r} + \boldsymbol{\delta}_1)$ are the sublattice-shifted annihilation operators and $\psi_{\sigma\rho,\alpha}$ is the corresponding valley-specific operator defined in (S24). The continuum limit is taken by replacing $L \to \mathbb{R}^2$ and $\psi/\sqrt{|\boldsymbol{a}_1 \times \boldsymbol{a}_2|} \to \psi$.

- $\psi_{\sigma\rho} \coloneqq (\psi_{\sigma\rho,A}, \psi_{\sigma\rho,B})^\mathsf{T}$ is a corresponding sublattice-space vector

II. FOURIER SERIES OF A LATTICE-PERIODIC FUNCTION

Using the notation of quotient groups, (discrete) Fourier series can be written elegantly [1]. We will be using the term "series" for functions that are periodic, and the term "discrete" for functions defined on a lattice.

A. Discrete Fourier series

Let $f : L/SL \to \mathbb{C}^n$ be a function defined on a 2-dimensional lattice $L \subset \mathbb{R}^2$ and periodic in translations of the 2-dimensional superlattice $SL \subset L$. It can be shown that the 2-dimensional discrete Fourier series and its coefficients can be written respectively as [1]

$$f(\boldsymbol{r}) = \sum_{\boldsymbol{G} \in SL^*/L^*} \mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} \tilde{f}(\boldsymbol{G}), \quad \tilde{f}(\boldsymbol{G}) = \frac{1}{|L/SL|} \sum_{\boldsymbol{r} \in L/SL} \mathrm{e}^{-\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} f(\boldsymbol{r}), \tag{S1}$$

where L/SL can be interpreted as any of the discrete superlattice unit cells and SL^*/L^* as any of the unit cells of the reciprocal superlattice (e.g. the superlattice Brillouin zone SBZ).

B. Fourier series

Let $f : \mathbb{R}^2/SL \to \mathbb{C}^n$ be a function defined on the continuum $\mathbb{R}^2$ and periodic in translations of the 2-dimensional lattice $SL = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{t}_1, \boldsymbol{t}_2\} \subset \mathbb{R}^2$ with the reciprocal lattice $SL^* = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{G}_1, \boldsymbol{G}_2\}$. Assuming the 1-dimensional Fourier series is known, it can be shown that the 2-dimensional Fourier series and its coefficients can be written respectively as

$$f(\boldsymbol{r}) = \sum_{\boldsymbol{G} \in SL^*} \mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} \tilde{f}(\boldsymbol{G}), \quad \tilde{f}(\boldsymbol{G}) = \frac{1}{|\mathbb{R}^2/SL|} \int_{\mathbb{R}^2/SL} \mathrm{d}\boldsymbol{r} \, \mathrm{e}^{-\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} f(\boldsymbol{r}), \tag{S2}$$

where the integral is calculated as a 2-dimensional volume integral which, by interpreting $\mathbb{R}^2/SL$ as the parallelogram defined by $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$ (one of the superlattice continuum unit cells) and by change of variables, can be written as

$$\tilde{f}(\boldsymbol{G}) = \tilde{f}(m_1\boldsymbol{G}_1 + m_2\boldsymbol{G}_2) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{d}x_1 \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{d}x_2 \, \mathrm{e}^{-\mathrm{i}2\pi(m_1 x_1 + m_2 x_2)} f(x_1\boldsymbol{t}_1 + x_2\boldsymbol{t}_2). \tag{S3}$$

Writing the Fourier series this way we call the *reduced zone scheme*, for reasons explained in section V A.

As a special case, if f is constant in the $\boldsymbol{t}_2$ direction, we may use the result

$$\tilde{f}(\boldsymbol{G}) = \tilde{f}(m_1\boldsymbol{G}_1 + m_2\boldsymbol{G}_2) = \delta_{m_2,0}\tilde{f}(m_1\boldsymbol{G}_1) = \delta_{m_2,0}\tilde{f}(\boldsymbol{G}) \tag{S4}$$

yielding the series

$$f(\boldsymbol{r}) = \sum_{\boldsymbol{G} \in SL_1^*} \mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} \tilde{f}(\boldsymbol{G}). \tag{S5}$$

If we calculate the Fourier series this way (possible only if f is constant in the $\boldsymbol{t}_2$ direction), as a sum over the one-dimensional lattice SL_1^*, we call this the *mixed zone scheme*, for reasons explained in section V A.

Summarizing both schemes together, we may write the Fourier series and its coefficients as

$$f(\boldsymbol{r}) = \sum_{\boldsymbol{G} \in SL_S^*} \mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} \tilde{f}(\boldsymbol{G}), \quad \tilde{f}(\boldsymbol{G}) = \frac{1}{|\mathbb{R}^2/SL|} \int_{\mathbb{R}^2/SL} \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} f(\boldsymbol{r}), \tag{S6}$$

with the mixed zone scheme $S = \mathrm{MZ}$ being applicable only in the case of f being constant in the $\boldsymbol{t}_2$ direction.

C. Fourier components of selected functions

If we calculate the Fourier components of the pseudo vector potentials chosen in the main paper by (S6) [or more explicitly, (S3)], we get

$$\tilde{\boldsymbol{A}}_{\cos}^{\mathrm{1D}}(m_1\boldsymbol{G}_1 + m_2\boldsymbol{G}_2) = \frac{\beta}{2d}(0, \delta_{m_1,-1} + \delta_{m_1,1})\delta_{m_2,0}, \tag{S7}$$

$$\tilde{\boldsymbol{A}}_{\cos}^{\mathrm{2D}}(m_1\boldsymbol{G}_1 + m_2\boldsymbol{G}_2) = \frac{\beta}{2d}(\delta_{m_2,-1} + \delta_{m_2,1}, \delta_{m_1,-1} + \delta_{m_1,1}), \tag{S8}$$

$$\tilde{\boldsymbol{A}}_{1/c}^{\mathrm{1D}}(m_1\boldsymbol{G}_1 + m_2\boldsymbol{G}_2) = \begin{cases} (0,0), & \text{if } m_1 = 0, \\ \left(0, \mathrm{e}^{\mathrm{i}\pi m_1/2}\frac{\mathrm{e}^{-\mathrm{i}2\pi m_1(1+c)}}{c(2\pi m_1)^2}(-1+\mathrm{e}^{\mathrm{i}\pi m_1})(-1+\mathrm{e}^{\mathrm{i}2\pi m_1 c})(-\mathrm{e}^{\mathrm{i}\pi m_1}+\mathrm{e}^{\mathrm{i}2\pi m_1 c})\right)\delta_{m_2,0}, & \text{otherwise} \end{cases} \tag{S9}$$

in the reduced zone scheme, where $m_1\boldsymbol{G}_1 + m_2\boldsymbol{G}_2 \in SL_{\mathrm{RZ}}^*$. As discussed in the previous section, because of the $\delta_{m_2,0}$ factor in the 1D potentials we may as well use a one-dimensional Fourier series and write

$$\tilde{\boldsymbol{A}}_{\cos}^{\mathrm{1D}}(m_1\boldsymbol{G}_1) = \frac{\beta}{2d}(0, \delta_{m_1,-1} + \delta_{m_1,1}) \tag{S10}$$

in the mixed zone scheme (and similarly for $\tilde{\boldsymbol{A}}_{1/c}^{\mathrm{1D}}$), where $m_1\boldsymbol{G}_1 \in SL_{\mathrm{MZ}}^*$.

III. STRAINED GRAPHENE LOW-ENERGY EFFECTIVE BCS HAMILTONIAN

Let us first fix the lattice vectors. Note that these are needed only when deriving the continuum theory, but after moving to the continuum theory the lattice is not anymore present, except of its orientation. We take the graphene lattice primitive vectors

$$\boldsymbol{a}_1 = \frac{a}{2}(1, \sqrt{3}), \quad \boldsymbol{a}_2 = \frac{a}{2}(-1, \sqrt{3}), \tag{S11}$$

and the nearest-neighbor vectors

$$\boldsymbol{\delta}_1 = \frac{1}{3}(\boldsymbol{a}_1 + \boldsymbol{a}_2), \quad \boldsymbol{\delta}_2 = \frac{1}{3}(\boldsymbol{a}_2 - 2\boldsymbol{a}_1), \quad \boldsymbol{\delta}_3 = \frac{1}{3}(\boldsymbol{a}_1 - 2\boldsymbol{a}_2). \tag{S12}$$

With these definitions the zigzag direction is in the x direction and the $\boldsymbol{K}$ point is located at

$$\boldsymbol{K} = \frac{4\pi}{3a^2}(\boldsymbol{a}_1 - \boldsymbol{a}_2) = \frac{4\pi}{3a}(1,0). \tag{S13}$$

In the nearest-neighbour tight-binding model the interacting Hamiltonian of strained graphene is

$$H_{\mathrm{BdG}} = H_{\mathrm{p}} + \delta H_{\mathrm{s}} + H_{\mathrm{int}} =: H + H_{\mathrm{int}}, \tag{S14}$$

where the noninteracting pristine graphene part is

$$H_{\mathrm{p}} = -t \sum_{\sigma\in\{\uparrow,\downarrow\}} \sum_{j=1}^{3} \sum_{\boldsymbol{r}\in L/L_{\mathrm{BK}}} \psi_\sigma^\dagger(\boldsymbol{r})\psi_\sigma(\boldsymbol{r}+\boldsymbol{\delta}_j) + \mathrm{h.c.} - \mu \sum_{\substack{\alpha\in\{A,B\},\,\boldsymbol{r}\in L/L_{\mathrm{BK}} \\ \sigma\in\{\uparrow,\downarrow\}}} \psi_\sigma^\dagger(\boldsymbol{r}+\boldsymbol{\delta}_\alpha)\psi_\sigma(\boldsymbol{r}+\boldsymbol{\delta}_\alpha) \tag{S15}$$

$$= -t \sum_{\sigma\in\{\uparrow,\downarrow\}} \sum_{j=1}^{3} \sum_{\boldsymbol{r}\in L/L_{\mathrm{BK}}} \psi_{\sigma,1A}^\dagger(\boldsymbol{r})\psi_{\sigma,1B}(\boldsymbol{r}+\boldsymbol{\delta}_j-\boldsymbol{\delta}_1) + \mathrm{h.c.} - \mu \sum_{\substack{\alpha\in\{A,B\},\,\boldsymbol{r}\in L/L_{\mathrm{BK}} \\ \sigma\in\{\uparrow,\downarrow\}}} \psi_{\sigma,1\alpha}^\dagger(\boldsymbol{r})\psi_{\sigma,1\alpha}(\boldsymbol{r}), \tag{S16}$$

the small change to this due to strain is

$$\delta H_{\mathrm{s}} = - \sum_{\sigma\in\{\uparrow,\downarrow\}} \sum_{j=1}^{3} \sum_{\boldsymbol{r}\in L/L_{\mathrm{BK}}} \delta t_j(\boldsymbol{r})\psi_\sigma^\dagger(\boldsymbol{r})\psi_\sigma(\boldsymbol{r}+\boldsymbol{\delta}_j) + \mathrm{h.c.} \tag{S17}$$

$$= - \sum_{\sigma\in\{\uparrow,\downarrow\}} \sum_{j=1}^{3} \sum_{\boldsymbol{r}\in L/L_{\mathrm{BK}}} \delta t_j(\boldsymbol{r})\psi_{\sigma,1A}^\dagger(\boldsymbol{r})\psi_{\sigma,1B}(\boldsymbol{r}+\boldsymbol{\delta}_j - \boldsymbol{\delta}_1) + \mathrm{h.c.}, \tag{S18}$$

and the interacting part is

$$H_{\mathrm{int}} = \frac{\lambda}{2} \sum_{\substack{\alpha\in\{A,B\},\,\boldsymbol{r}\in L/L_{\mathrm{BK}} \\ \sigma\in\{\uparrow,\downarrow\}}} \psi_\sigma^\dagger(\boldsymbol{r}+\boldsymbol{\delta}_\alpha)\psi_{\bar\sigma}^\dagger(\boldsymbol{r}+\boldsymbol{\delta}_\alpha)\psi_{\bar\sigma}(\boldsymbol{r}+\boldsymbol{\delta}_\alpha)\psi_\sigma(\boldsymbol{r}+\boldsymbol{\delta}_\alpha) \tag{S19}$$

$$= \frac{\lambda}{2} \sum_{\substack{\alpha\in\{A,B\},\,\boldsymbol{r}\in L/L_{\mathrm{BK}} \\ \sigma\in\{\uparrow,\downarrow\}}} \psi_{\sigma,\alpha}^\dagger(\boldsymbol{r})\psi_{\bar\sigma,\alpha}^\dagger(\boldsymbol{r})\psi_{\bar\sigma,\alpha}(\boldsymbol{r})\psi_{\sigma,\alpha}(\boldsymbol{r}) \tag{S20}$$

assuming local (also in sublattice) interaction of strength λ (negative for attractive interaction considered here) that is independent of sublattice, spin, and position. Here t is the graphene nearest-neighbour hopping energy, $\delta t_j(\boldsymbol{r})$ is a small change to this due to strain in the bond from $\boldsymbol{r}$ to $\boldsymbol{r}+\boldsymbol{\delta}_j$, and μ is the chemical potential.

Because of the periodicity of $\psi_{\sigma,\alpha}$ we may expand it as a discrete Fourier series

$$\psi_{\sigma,\alpha}(\boldsymbol{r}) = \sum_{\boldsymbol{k}\in L_{\mathrm{BK}}^*/L^*} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} c_{\sigma,\alpha}(\boldsymbol{k}). \tag{S21}$$

Dividing the sum in parts near and far from the Dirac points yields

$$\psi_{\sigma,\alpha}(\boldsymbol{r}) = \sum_{\rho\in\{+,-\}} \sum_{\substack{\boldsymbol{k}\in L_{\mathrm{BK}}^*/L^* \\ (\boldsymbol{k}\text{ near }\rho\boldsymbol{K})}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} c_{\sigma,\alpha}(\boldsymbol{k}) + \sum_{\substack{\boldsymbol{k}\in L_{\mathrm{BK}}^*/L^* \\ (\boldsymbol{k}\text{ not near }\boldsymbol{K},\boldsymbol{K}')}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} c_{\sigma,\alpha}(\boldsymbol{k}), \tag{S22}$$

where we can drop the last term by going into effective low-energy theory where terms far from the Dirac points are uninteresting. By defining the fermionic valley-specific annihilation operators in Fourier space,

$$c_{\sigma\rho,\alpha}(\boldsymbol{k}) := \begin{cases} c_{\sigma,\alpha}(\boldsymbol{k}+\rho\boldsymbol{K}), & \text{if } \boldsymbol{k} \text{ small}, \\ 0, & \text{otherwise}, \end{cases} \tag{S23}$$

and its corresponding discrete Fourier series

$$\psi_{\sigma\rho,\alpha}(\boldsymbol{r}) = \sum_{\substack{\boldsymbol{k}\in L_{\mathrm{BK}}^*/L^* \\ (\boldsymbol{k}\text{ small})}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} c_{\sigma\rho,\alpha}(\boldsymbol{k}), \quad c_{\sigma\rho,\alpha}(\boldsymbol{k}) = \frac{1}{|L/L_{\mathrm{BK}}|} \sum_{\boldsymbol{r}\in L/L_{\mathrm{BK}}} \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \psi_{\sigma\rho,\alpha}(\boldsymbol{r}), \tag{S24}$$

equation (S22) can be written as

$$\psi_{\sigma,\alpha}(\boldsymbol{r}) = \sum_{\rho} \sum_{\substack{\boldsymbol{k}\in L_{\mathrm{BK}}^*/L^* \\ (\boldsymbol{k}\text{ small})}} \mathrm{e}^{\mathrm{i}(\boldsymbol{k}+\rho\boldsymbol{K})\cdot\boldsymbol{r}} c_{\sigma\rho,\alpha}(\boldsymbol{k}) = \sum_{\rho} \mathrm{e}^{\mathrm{i}\rho\boldsymbol{K}\cdot\boldsymbol{r}} \psi_{\sigma\rho,\alpha}(\boldsymbol{r}) \tag{S25}$$

which is the expansion to use when we want to go to the low-energy theory and express the original operators in the valley-operator basis.

A. Strained graphene Hamiltonian

The derivation of the strained Hamiltonian has been already done in [2] in the case of carbon nanotubes, but for transparency we repeat the calculation here. Writing the annihilation/creation operators as the valley expansion

(S25) and linearizing $\psi_{\sigma\rho,B}(\boldsymbol{r} + \boldsymbol{\delta}_j) \approx \psi_{\sigma\rho,B}(\boldsymbol{r}) + (\boldsymbol{\delta}_j - \boldsymbol{\delta}_1) \cdot \nabla \psi_{\sigma\rho,B}(\boldsymbol{r})$ the pristine graphene Hamiltonian becomes

$$
H_{\mathrm{p}} = \hbar v_{\mathrm{F}} \sum_{\sigma\rho\rho'} \sum_{\boldsymbol{r}\in L/L_{\mathrm{BK}}} \mathrm{e}^{\mathrm{i}(\rho'-\rho)\boldsymbol{K}\cdot\boldsymbol{r}} \psi_{\sigma,A}^{\dagger}(\boldsymbol{r})(-\mathrm{i})(\rho'\partial_x - \mathrm{i}\partial_y)\psi_{\sigma\rho',B}(\boldsymbol{r}) + \mathrm{h.c.}
$$
$$
- \mu \sum_{\sigma\rho\rho'\alpha} \sum_{\boldsymbol{r}\in L/L_{\mathrm{BK}}} \mathrm{e}^{\mathrm{i}(\rho'-\rho)\boldsymbol{K}\cdot\boldsymbol{r}} \psi_{\sigma\rho,\alpha}^{\dagger}(\boldsymbol{r})\psi_{\sigma\rho',\alpha}(\boldsymbol{r}), \tag{S26}
$$

where we used $\sum_{j=1}^{3} \mathrm{e}^{\mathrm{i}\rho\boldsymbol{K}\cdot\boldsymbol{\delta}_j} = 0$ and defined the Fermi velocity by $\hbar v_{\mathrm{F}} \coloneqq \sqrt{3}at/2$. The exponential factor gives simply $\delta_{\rho\rho'}$. This can be seen by going into Fourier space by using (S24), after which the overall exponential gives $|L/L_{\mathrm{BK}}|\delta_{\boldsymbol{k}+\rho\boldsymbol{K},\boldsymbol{k}'+\rho'\boldsymbol{K}}$ after calculating the $\boldsymbol{r}$ sum. By using the property that $\boldsymbol{k}$ and $\boldsymbol{k}'$ are small, this is equal to $|L/L_{\mathrm{BK}}|\delta_{\boldsymbol{k}\boldsymbol{k}'}\delta_{\rho\rho'}$. Then after coming back to real space the Hamiltonian reads

$$
H_{\mathrm{p}} = \hbar v_{\mathrm{F}} \sum_{\sigma\rho} \sum_{\boldsymbol{r}\in L/L_{\mathrm{BK}}} \psi_{\sigma\rho,A}^{\dagger}(\boldsymbol{r})(-\mathrm{i})(\rho\partial_x - \mathrm{i}\partial_y)\psi_{\sigma\rho,B}(\boldsymbol{r}) + \mathrm{h.c.} - \mu \sum_{\sigma\rho\alpha} \sum_{\boldsymbol{r}\in L/L_{\mathrm{BK}}} \psi_{\sigma\rho,\alpha}^{\dagger}(\boldsymbol{r})\psi_{\sigma\rho,\alpha}(\boldsymbol{r}). \tag{S27}
$$

For the strain Hamiltonian we similarly write the sublattice-shifted annihilation operators as the valley expansion (S25) and make the zeroth-order approximation $\psi_{\sigma\rho,B}(\boldsymbol{r} + \boldsymbol{\delta}_j) \approx \psi_{\sigma\rho,B}(\boldsymbol{r})$. This yields

$$
\delta H_{\mathrm{s}} = \sum_{\sigma\rho} \sum_{\boldsymbol{r}\in L/L_{\mathrm{BK}}} \psi_{\sigma\rho,A}^{\dagger}(\boldsymbol{r})\rho\hbar v_{\mathrm{F}} A^{\rho}(\boldsymbol{r})^{*}\psi_{\sigma\rho,B}(\boldsymbol{r}) + \mathrm{h.c.}, \tag{S28}
$$

where

$$
A^{\rho}(\boldsymbol{r}) \coloneqq \rho A_x(\boldsymbol{r}) + \mathrm{i}A_y(\boldsymbol{r}) \coloneqq -\frac{\rho}{\hbar v_{\mathrm{F}}} \sum_j \mathrm{e}^{-\mathrm{i}\rho\boldsymbol{K}\cdot\boldsymbol{\delta}_j}\delta t_j(\boldsymbol{r}). \tag{S29}
$$

The strained graphene Hamiltonian then becomes

$$
H = H_{\mathrm{p}} + \delta H_{\mathrm{s}} = \sum_{\sigma\rho} \sum_{\boldsymbol{r}\in L/L_{\mathrm{BK}}} \psi_{\sigma\rho}^{\dagger}(\boldsymbol{r})\mathcal{H}^{\rho}(\boldsymbol{r})\psi_{\sigma\rho}(\boldsymbol{r}), \tag{S30}
$$

where we defined the Hamiltonian matrix element

$$
\mathcal{H}^{\rho}(\boldsymbol{r}) \coloneqq \hbar v_{\mathrm{F}}\boldsymbol{\sigma}^{\rho} \cdot (-\mathrm{i}\nabla + \rho\boldsymbol{A}(\boldsymbol{r})) - \mu \tag{S31}
$$

and the pseudo vector potential $\boldsymbol{A} \coloneqq (A_x, A_y)$.

We now know the connection (S29) between the pseudo vector potential $\boldsymbol{A}$ and the small change δt_j in the hopping energy, but we still need to find the connection between δt_j and strain. Assuming the atom at $\boldsymbol{r}$ to be displaced by a vector $\boldsymbol{v}(\boldsymbol{r}) = (\boldsymbol{u}(\boldsymbol{r}), h(\boldsymbol{r}))$, where $\boldsymbol{u} = (u_x, u_y)$ is the in-plane and h is the out-of-plane displacement field, the change in the bond length of the $\boldsymbol{\delta}_j$ bond due to strain is

$$
\delta u_j(\boldsymbol{r}) \coloneqq \|\boldsymbol{r} + \boldsymbol{\delta}_j + \boldsymbol{v}(\boldsymbol{r} + \boldsymbol{\delta}_j) - [\boldsymbol{r} + \boldsymbol{v}(\boldsymbol{r})]\| - \|\boldsymbol{r} + \boldsymbol{\delta}_j - \boldsymbol{r}\| \tag{S32}
$$
$$
\approx \frac{1}{\|\boldsymbol{\delta}_j\|}\left[\boldsymbol{\delta}_j \cdot (\boldsymbol{u}(\boldsymbol{r} + \boldsymbol{\delta}_j) - \boldsymbol{u}(\boldsymbol{r})) + \frac{1}{2}(h(\boldsymbol{r} + \boldsymbol{\delta}_j) - h(\boldsymbol{r}))^2\right], \tag{S33}
$$

where in the last step we linearized in $\|\boldsymbol{u}(\boldsymbol{r} + \boldsymbol{\delta}_j) - \boldsymbol{u}(\boldsymbol{r})\|, |h(\boldsymbol{r} + \boldsymbol{\delta}_j) - h(\boldsymbol{r})| \ll \|\boldsymbol{\delta}_j\|$. Furthermore in the linear order we may approximate [3]

$$
\boldsymbol{u}(\boldsymbol{r} + \boldsymbol{\delta}_j) - \boldsymbol{u}(\boldsymbol{r}) \approx (\boldsymbol{\delta}_j \cdot \nabla)\boldsymbol{u}(\boldsymbol{r}), \quad h(\boldsymbol{r} + \boldsymbol{\delta}_j) - h(\boldsymbol{r}) \approx \boldsymbol{\delta}_j \cdot \nabla h(\boldsymbol{r}), \tag{S34}
$$

and if we define the strain tensor as

$$
u_{ij} \coloneqq \frac{1}{2}(\partial_i u_j + \partial_j u_i) + \frac{1}{2}\partial_i h \partial_j h, \tag{S35}
$$

the change in the bond length becomes [4]

$$
\delta u_j(\boldsymbol{r}) = \frac{1}{\|\boldsymbol{\delta}_j\|}\left[\delta_{jx}^2 u_{xx}(\boldsymbol{r}) + 2\delta_{jx}\delta_{jy}u_{xy}(\boldsymbol{r}) + \delta_{jy}^2 u_{yy}(\boldsymbol{r})\right]. \tag{S36}
$$

Now that the change in the hopping energy can be linearized to

$$\delta t_j(\boldsymbol{r}) \approx \frac{\mathrm{d}t}{\mathrm{d}a_0}\delta u_j(\boldsymbol{r}) = -\frac{t\beta_{\mathrm{G}}}{a_0}\delta u_j(\boldsymbol{r}), \tag{S37}$$

where $\beta_{\mathrm{G}} := -\,\mathrm{d}\ln t/\mathrm{d}\ln a_0 \approx 2$ is the graphene Grüneisen parameter [5], equations (S29), (S36), and (S37) yield for the pseudo vector potential

$$\boldsymbol{A} = -\frac{\beta_{\mathrm{G}}}{2a_0}(u_{xx} - u_{yy}, -2u_{xy}). \tag{S38}$$

Finally we can extend the annihilation operators to the continuum $\mathbb{R}^2/L_{\mathrm{BK}}$ by the discrete Fourier series (S24), and everything else is trivially extended. By furthermore redefining the continuum annihilation operator density as $\psi_{\sigma\rho,\alpha}/\sqrt{|\boldsymbol{a}_1 \times \boldsymbol{a}_2|} \to \psi_{\sigma\rho,\alpha}$ we arrive at the continuum Hamiltonian

$$H = \sum_{\sigma\rho} \int_{\mathbb{R}^2/L_{\mathrm{BK}}} \mathrm{d}\boldsymbol{r}\, \psi_{\sigma\rho}^\dagger(\boldsymbol{r})\mathcal{H}^\rho(\boldsymbol{r})\psi_{\sigma\rho}(\boldsymbol{r}). \tag{S39}$$

Note that while a normal vector potential would break time-reversal symmetry, this Hamiltonian is time-reversal symmetric, $\mathcal{H}^{\bar\rho*} = \mathcal{H}^\rho$, because of the valley-odd pseudo vector potential.

1. Limits in the theory

The only assumptions regarding strain in deriving the Hamiltonian (S39) were $\|\boldsymbol{u}(\boldsymbol{r} + \boldsymbol{\delta}_j) - \boldsymbol{u}(\boldsymbol{r})\|, |h(\boldsymbol{r} + \boldsymbol{\delta}_j) - h(\boldsymbol{r})| \ll \|\boldsymbol{\delta}_j\|$. For the cosine displacement fields $\boldsymbol{u}_{\mathrm{cos}}^{\mathrm{1D}}$ and $\boldsymbol{u}_{\mathrm{cos}}^{\mathrm{2D}}$ of the main paper this assumption reads

$$\boldsymbol{u}_{\mathrm{cos}}^{\mathrm{1D}}: \qquad \frac{\beta}{\beta_{\mathrm{G}}} \ll \frac{d}{a_0} \quad \text{and} \quad \frac{d}{a_0} \gg 1 \tag{S40}$$

$$\boldsymbol{u}_{\mathrm{cos}}^{\mathrm{2D}}: \qquad \frac{\beta}{\beta_{\mathrm{G}}} \ll \frac{d}{2a_0} \quad \text{and} \quad \frac{d}{a_0} \gg 1. \tag{S41}$$

Note that this is equivalent to assuming that the strain is $|u_{ij}| \ll 1$.

B. Interaction Hamiltonian

The derivation regarding superconductivity is based on the book by Nazarov & Danon [6]. First making the mean-field approximation in the Cooper channel for H_{int} yields

$$H_{\mathrm{int}} \approx \frac{1}{2}\sum_{\sigma\alpha}\sum_{\boldsymbol{r}\in L/L_{\mathrm{BK}}} \Delta_{\sigma,\alpha}(\boldsymbol{r})\psi_{\sigma,\alpha}^\dagger(\boldsymbol{r})\psi_{\bar\sigma,\alpha}^\dagger(\boldsymbol{r}) + \text{h.c.} + E^0 \tag{S42}$$

where the superconducting order parameter is

$$\Delta_{\sigma,\alpha} := \lambda \langle \psi_{\bar\sigma,\alpha}\psi_{\sigma,\alpha} \rangle \tag{S43}$$

with the angle brackets denoting the thermal average and the constant term is

$$E^0 := -\frac{1}{2\lambda}\sum_{\sigma\alpha}\sum_{\boldsymbol{r}\in L/L_{\mathrm{BK}}} |\Delta_{\sigma,\alpha}(\boldsymbol{r})|^2. \tag{S44}$$

Using the valley expansion (S25) and *assuming only intervalley interaction* gives

$$H_{\mathrm{int}} = \frac{1}{2}\sum_{\sigma\rho\alpha}\sum_{\boldsymbol{r}\in L/L_{\mathrm{BK}}} \Delta_{\sigma,\alpha}(\boldsymbol{r})\psi_{\sigma\rho,\alpha}^\dagger(\boldsymbol{r})\psi_{\bar\sigma\bar\rho,\alpha}^\dagger(\boldsymbol{r}) + \text{h.c.} + E^0 \tag{S45}$$

$$= \frac{1}{2}\sum_{\sigma\rho}\sum_{\boldsymbol{r}\in L/L_{\mathrm{BK}}} \psi_{\sigma\rho}^\dagger(\boldsymbol{r})\Delta_\sigma(\boldsymbol{r})\psi_{\bar\sigma\bar\rho}^{\dagger\mathsf{T}}(\boldsymbol{r}) + \text{h.c.} + E^0, \tag{S46}$$

$$\Delta_{\sigma,\alpha} := \lambda \sum_{\rho} \langle \psi_{\bar{\sigma}\bar{\rho},\alpha} \psi_{\sigma\rho,\alpha} \rangle, \quad \Delta_{\sigma} := \mathrm{diag}(\Delta_{\sigma,A}, \Delta_{\sigma,B}) \tag{S47}$$

and

$$E^0 = -\frac{1}{2\lambda} \sum_{\sigma} \sum_{\boldsymbol{r} \in L/L_{\mathrm{BK}}} \mathrm{Tr}(\Delta_{\sigma}^*(\boldsymbol{r})\Delta_{\sigma}(\boldsymbol{r})), \tag{S48}$$

where the trace is over the sublattice structure.

Further taking the continuum limit $L \to \mathbb{R}^2$ gives

$$H_{\mathrm{int}} = \frac{1}{2} \sum_{\sigma\rho} \int_{\mathbb{R}^2/L_{\mathrm{BK}}} \mathrm{d}\boldsymbol{r}\, \psi_{\sigma\rho}^{\dagger}(\boldsymbol{r})\Delta_{\sigma}(\boldsymbol{r})\psi_{\bar{\sigma}\bar{\rho}}^{\dagger\mathsf{T}}(\boldsymbol{r}) + \mathrm{h.c.} + E^0, \tag{S49}$$

if we at the same time replace $\psi_{\sigma\rho}/\sqrt{|\boldsymbol{a}_1 \times \boldsymbol{a}_2|} \to \psi_{\sigma\rho}$ and $\lambda/|\boldsymbol{a}_1 \times \boldsymbol{a}_2| \to \lambda$. The constant term is then

$$E^0 = -\frac{1}{2\lambda} \sum_{\sigma} \int_{\mathbb{R}^2/L_{\mathrm{BK}}} \mathrm{d}\boldsymbol{r}\, \mathrm{Tr}(\Delta_{\sigma}^*(\boldsymbol{r})\Delta_{\sigma}(\boldsymbol{r})). \tag{S50}$$

IV. DIAGONALIZING THE HAMILTONIAN

A. Writing the Hamiltonian in Nambu basis

Utilizing the anticommutation relations and the identity

$$\int \mathrm{d}\boldsymbol{r}\, \psi_{\sigma\rho}^{\dagger}(\boldsymbol{r})\mathcal{H}^{\rho}(\boldsymbol{r})\psi_{\sigma\rho}(\boldsymbol{r}) = -\int \mathrm{d}\boldsymbol{r}\, \psi_{\sigma\rho}^{\mathsf{T}}(\boldsymbol{r})\mathcal{H}^{\bar{\rho}}(\boldsymbol{r})\psi_{\sigma\rho}^{\dagger\mathsf{T}}(\boldsymbol{r}) - 2\mu V \delta(0), \tag{S51}$$

where the $-2\mu V \delta(0)$ term comes from anticommuting the annihilation/creation operators in the particle number operator, we may bring the total Hamiltonian into the form

$$H_{\mathrm{BdG}} = \frac{1}{2} \sum_{\sigma\rho} \int \mathrm{d}\boldsymbol{r}\, \begin{pmatrix} \psi_{\sigma\rho}^{\dagger}(\boldsymbol{r}) & \psi_{\bar{\sigma}\bar{\rho}}^{\mathsf{T}}(\boldsymbol{r}) \end{pmatrix} \begin{pmatrix} \mathcal{H}^{\rho}(\boldsymbol{r}) & \Delta_{\sigma}(\boldsymbol{r}) \\ \Delta_{\sigma}^*(\boldsymbol{r}) & -\mathcal{H}^{\rho}(\boldsymbol{r}) \end{pmatrix} \begin{pmatrix} \psi_{\sigma\rho}(\boldsymbol{r}) \\ \psi_{\bar{\sigma}\bar{\rho}}^{\dagger\mathsf{T}}(\boldsymbol{r}) \end{pmatrix} + E^0 - 4\mu V \delta(0) \tag{S52}$$

$$= \frac{1}{2} \sum_{\sigma\rho} \int \mathrm{d}\boldsymbol{r}\, \Psi_{\sigma\rho}^{\dagger}(\boldsymbol{r})\mathcal{H}_{\mathrm{BdG}}^{\rho}(\boldsymbol{r})\Psi_{\sigma\rho}(\boldsymbol{r}) + E^0 - 4\mu V \delta(0), \tag{S53}$$

where in the last step we defined the spin-independent Bogoliubov–de Gennes Hamiltonian in Nambu space, the spin-independent order parameter, and the Nambu-vector operator respectively as

$$\mathcal{H}_{\mathrm{BdG}}^{\rho} := \begin{pmatrix} \mathcal{H}^{\rho} & \Delta \\ \Delta^* & -\mathcal{H}^{\rho} \end{pmatrix}, \quad \Delta := \Delta_{\uparrow} = -\Delta_{\downarrow} = s(\sigma)\Delta_{\sigma}, \quad \Psi_{\sigma\rho} := \begin{pmatrix} \psi_{\sigma\rho} \\ s(\sigma)\psi_{\bar{\sigma}\bar{\rho}}^{\dagger\mathsf{T}} \end{pmatrix}. \tag{S54}$$

B. Writing the Hamiltonian in eigenbasis: Bogoliubov–de Gennes equation

Simply by using the definition of the Dirac delta we may write H_{BdG} in (S53) as

$$H_{\mathrm{BdG}} = \frac{1}{2} \sum_{\sigma\rho} \int \mathrm{d}\boldsymbol{r} \int \mathrm{d}\boldsymbol{r}'\, \Psi_{\sigma\rho}^{\dagger}(\boldsymbol{r})\delta(\boldsymbol{r} - \boldsymbol{r}')\mathcal{H}_{\mathrm{BdG}}^{\rho}(\boldsymbol{r})\Psi_{\sigma\rho}(\boldsymbol{r}') + E^0 - 4\mu V \delta(0). \tag{S55}$$

Now let $\delta(\boldsymbol{r} - \boldsymbol{r}')\mathcal{H}_{\mathrm{BdG}}^{\rho}(\boldsymbol{r})$ be the representation of $\hat{\mathcal{H}}_{\mathrm{BdG}}^{\rho}$ in position space, that is,

$$\delta(\boldsymbol{r} - \boldsymbol{r}')\mathcal{H}_{\mathrm{BdG}}^{\rho}(\boldsymbol{r}) = \langle \boldsymbol{r}| \hat{\mathcal{H}}_{\mathrm{BdG}}^{\rho} |\boldsymbol{r}' \rangle. \tag{S56}$$

Since $\mathcal{H}_{\mathrm{BdG}}$ is Hermitian we may use the spectral theorem (following from the resolution of identity)

$$\mathbb{1} = \frac{1}{V} \sum_{n} |w_{\rho n}\rangle\langle w_{\rho n}| \quad \Rightarrow \quad \hat{\mathcal{H}}_{\mathrm{BdG}}^{\rho} = \frac{1}{V} \sum_{n} E_{\rho n} |w_{\rho n}\rangle\langle w_{\rho n}| \tag{S57}$$

where n enumerates all the eigenstates $|w_{\rho n}\rangle$ of $\hat{\mathcal{H}}^\rho_{\mathrm{BdG}}$, *i.e.*

$$\hat{\mathcal{H}}^\rho_{\mathrm{BdG}} |w_{\rho n}\rangle = E_{\rho n} |w_{\rho n}\rangle \quad \Leftrightarrow \quad \mathcal{H}^\rho_{\mathrm{BdG}}(\boldsymbol{r})w_{\rho n}(\boldsymbol{r}) = E_{\rho n}w_{\rho n}(\boldsymbol{r}). \tag{S58}$$

Here we fixed the normalization of the eigenstates to $\langle w_{\rho n}|w_{\rho n}\rangle = V$ (see section V B). Equation (S58) is usually called the *(Dirac–)Bogoliubov–de Gennes equation* [7–9]. Writing the Nambu structure explicitly it reads

$$\begin{pmatrix} \mathcal{H}^\rho(\boldsymbol{r}) & \Delta(\boldsymbol{r}) \\ \Delta^*(\boldsymbol{r}) & -\mathcal{H}^\rho(\boldsymbol{r}) \end{pmatrix} \begin{pmatrix} u_{\rho n}(\boldsymbol{r}) \\ v_{\rho n}(\boldsymbol{r}) \end{pmatrix} = E_{\rho n} \begin{pmatrix} u_{\rho n}(\boldsymbol{r}) \\ v_{\rho n}(\boldsymbol{r}) \end{pmatrix}. \tag{S59}$$

Using the spectral theorem (S57) in (S56) yields

$$\delta(\boldsymbol{r} - \boldsymbol{r}')\mathcal{H}^\rho_{\mathrm{BdG}}(\boldsymbol{r}) = \frac{1}{V}\sum_n E_{\rho n} \langle \boldsymbol{r}|w_{\rho n}\rangle \langle w_{\rho n}|\boldsymbol{r}'\rangle = \frac{1}{V}\sum_n E_{\rho n}w_{\rho n}(\boldsymbol{r})w^\dagger_{\rho n}(\boldsymbol{r}') \tag{S60}$$

and furthermore substituting this to (S55) brings H_{BdG} into the diagonal form

$$H_{\mathrm{BdG}} = \frac{1}{2V}\sum_{\sigma\rho n} E_{\rho n} \int \mathrm{d}\boldsymbol{r}\, \Psi^\dagger_{\sigma\rho}(\boldsymbol{r})w_{\rho n}(\boldsymbol{r}) \int \mathrm{d}\boldsymbol{r}'\, w^\dagger_{\rho n}(\boldsymbol{r}')\Psi_{\sigma\rho}(\boldsymbol{r}') + E^0 - 4\mu V\delta(0) \tag{S61}$$

$$= \frac{1}{2}\sum_{\sigma\rho n} E_{\rho n}\gamma^\dagger_{\sigma\rho n}\gamma_{\sigma\rho n} + E^0 - 4\mu V\delta(0), \tag{S62}$$

where we defined the *Bogoliubon operator* or the *Bogoliubov transformation*

$$\gamma_{\sigma\rho n} := \frac{1}{\sqrt{V}} \int \mathrm{d}\boldsymbol{r}\, w^\dagger_{\rho n}(\boldsymbol{r})\Psi_{\sigma\rho}(\boldsymbol{r}). \tag{S63}$$

The Bogoliubons are not generally fermionic operators, but as we see in section IV C, concentrating only on positive/only on negative-energy operators makes them fermionic.

C. Making the Bogoliubons fermionic

Since the noninteracting Hamiltonian has the time-reversal symmetry $\mathcal{H}^{\bar{\rho}*} = \mathcal{H}^\rho$, we have the symmetry

$$\tau_y \mathcal{H}^{\bar{\rho}*}_{\mathrm{BdG}}\tau_y = -\mathcal{H}^\rho_{\mathrm{BdG}} \tag{S64}$$

of the BdG Hamiltonian. To see what this implies for the eigenenergies and eigenfunctions, we need to identify the so-far abstract index n. Taking the annihilation operators to be periodic in translations of the Born–von Kármán lattice L_{BK} and $\boldsymbol{A}$ and Δ to be periodic in translations of the superlattice SL, let us take the ansatz that $n = (b, \boldsymbol{k})$, where $\boldsymbol{k} \in L^*_{\mathrm{BK}}/SL^*_S$ belongs to the superlattice Brillouin zone (in different schemes) and $b = (\eta, \nu)$ enumerates the bands for each $\boldsymbol{k}$ with $\nu \in \{+, -\}$ giving the sign of energy of this band. We show in section V A this ansatz to be consistent. Further assuming the bands to be ordered energy-wise such that the noninteracting energies have the symmetry $\epsilon_{\rho b\boldsymbol{k}} = \epsilon_{\bar{\rho}b\bar{\boldsymbol{k}}}$, and assuming this symmetry to be inherited to the superconducting state,

$$E_{\rho b\boldsymbol{k}} = E_{\bar{\rho}b\bar{\boldsymbol{k}}}, \tag{S65}$$

we find that

$$w_{\rho\eta\bar{\nu}\boldsymbol{k}} = \mathrm{i}\tau_y w^*_{\bar{\rho}\eta\nu\bar{\boldsymbol{k}}}. \tag{S66}$$

This then directly gives the symmetry

$$\gamma_{\sigma\rho\eta\bar{\nu}\boldsymbol{k}} = -s(\sigma)\gamma^\dagger_{\bar{\sigma}\bar{\rho}\eta\nu\bar{\boldsymbol{k}}} \tag{S67}$$

between the positive/negative-energy Bogoliubons.

Choosing an orthogonal eigenbasis

$$V\delta_{nn'} = \langle w_{\rho n}|w_{\rho n'}\rangle = \int \mathrm{d}\boldsymbol{r}\, w^\dagger_{\rho n}(\boldsymbol{r})w_{\rho n'}(\boldsymbol{r}) \tag{S68}$$

(see section V B for the chosen normalization) gives the first fermionic anticommutation relation $\left\{\gamma_{\sigma\rho n}, \gamma^\dagger_{\sigma'\rho'n'}\right\} = \delta_{\sigma\sigma'}\delta_{\rho\rho'}\delta_{nn'}$ for all n, n' and the second fermionic anticommutation relation $\left\{\gamma_{\sigma\rho\eta\nu\boldsymbol{k}}, \gamma_{\sigma'\rho'\eta'\nu'\boldsymbol{k}'}\right\} = 0$ *provided* $\nu = \nu'$ *i.e.* that they both are either positive- or negative-energy operators. Concentrating then only on the positive-energy ones, we get the desired result that the Bogoliubons are fermionic,

$$\left\{\gamma_{\sigma\rho\eta+\boldsymbol{k}}, \gamma^\dagger_{\sigma'\rho'\eta'+\boldsymbol{k}'}\right\} = \delta_{\sigma\sigma'}\delta_{\rho\rho'}\delta_{\eta\eta'}\delta_{\boldsymbol{k}\boldsymbol{k}'}, \tag{S69}$$

$$\left\{\gamma_{\sigma\rho\eta+\boldsymbol{k}}, \gamma_{\sigma'\rho'\eta'+\boldsymbol{k}'}\right\} = 0. \tag{S70}$$

Utilizing then the symmetry (S67) of the positive/negative-energy Bogoliubons, the fermionic anticommutation relations (S69) and (S70) of the positive-energy Bogoliubons, and defining

$$E^{\mathrm{gs}} := E^0 - 4\mu V \delta(0) - \frac{1}{2} \sum_{\sigma\rho n_+} E_{\rho n_+} \tag{S71}$$

allows us to finally write the Hamiltonian (S62) in the diagonal form

$$H_{\mathrm{BdG}} = \sum_{\sigma\rho n_+} E_{\rho n_+} \gamma^\dagger_{\sigma\rho n_+} \gamma_{\sigma\rho n_+} + E^{\mathrm{gs}} \tag{S72}$$

with the operators $\gamma_{\sigma\rho n_+}$ being fermionic. According to the calculation above diagonalizing H_{BdG} [*i.e.* bringing it to the form (S72)] is equivalent to solving the Bogoliubov–de Gennes equation (S58). Note that since the Hamiltonian is diagonal in the fermionic positive-energy Bogoliubons, E^{gs} measures the ground-state energy.

D. Self-consistency equation

To write the definition (S47) and (S54) of the order parameter Δ in the same Bogoliubon basis as we did for H_{BdG} in (S72), we need to invert the definition (S63). Using the orthogonality condition (S68) together with the resolution of identity (S57) the inverse transformation can be shown to be

$$\Psi_{\sigma\rho}(\boldsymbol{r}) = \frac{1}{\sqrt{V}} \sum_n w_{\rho n}(\boldsymbol{r}) \gamma_{\sigma\rho n}. \tag{S73}$$

This can also be written as a sum over only the positive-energy states as

$$\Psi_{\sigma\rho}(\boldsymbol{r}) = \frac{1}{\sqrt{V}} \sum_{n_+} \left(w_{\rho n_+}(\boldsymbol{r}) \gamma_{\sigma\rho n_+} + \mathrm{i}\tau_y w^*_{\bar{\rho} n_+}(\boldsymbol{r}) \gamma^\dagger_{\bar{\sigma}\bar{\rho} n_+} \right) \tag{S74}$$

by using the symmetries (S66) and (S67).

Since the positive-energy Bogoliubons are fermionic and we assume no interactions between them, they follow the Fermi–Dirac statistics

$$\left\langle \gamma_{\sigma\rho n_+} \gamma_{\sigma'\rho' n'_+} \right\rangle = 0, \tag{S75}$$

$$\left\langle \gamma^\dagger_{\sigma\rho n_+} \gamma_{\sigma'\rho' n'_+} \right\rangle = \delta_{\sigma\sigma'}\delta_{\rho\rho'}\delta_{n_+ n'_+} f(E_{\rho n_+}), \tag{S76}$$

where $f(E) := [\mathrm{e}^{E/(k_{\mathrm{B}}T)} + 1]^{-1}$ is the Fermi–Dirac distribution at temperature T. Substituting then the relation (S73) in the definition (S47) and (S54) of Δ, using the fermionic anticommutation relations (S69) and (S70) of the Bogoliubons, and the thermal averages (S75) and (S76) then yields the *self-consistency equation*

$$\Delta_\alpha(\boldsymbol{r}) = -\frac{\lambda}{V} \sum_{\rho n_+} u_{\rho n_+, \alpha}(\boldsymbol{r}) v^*_{\rho n_+, \alpha}(\boldsymbol{r}) \tanh\left(\frac{E_{\rho n_+}}{2k_{\mathrm{B}}T}\right) \tag{S77}$$

for the superconducting order parameter at sublattice α.

V. EQUATIONS IN FOURIER SPACE

A. Bogoliubov–de Gennes equation in Fourier space

Assuming the eigenfunctions $w_{\rho b k'} = (u_{\rho b k'}, v_{\rho b k'})^{\mathsf{T}} : \mathbb{R}^2 / L_{\mathrm{BK}} \to \mathbb{C}^2$ (with $k' \in L_{\mathrm{BK}}^* / SL_S^*$ in the superlattice Brillouin zone in different schemes) to be periodic in translations of the large Born–von Kármán lattice L_{BK}, the pseudo vector potential $\boldsymbol{A} : \mathbb{R}^2 / SL \to \mathbb{R}^2$ to be periodic in translations of the arbitrary superlattice $SL = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{t}_1, \boldsymbol{t}_2\}$, and the order parameter $\Delta : \mathbb{R}^2 / SL \to \mathbb{C}_{2 \times 2}$ to be periodic in translations of the same lattice SL, we may expand them by (S6) as the Fourier series

$$w_{\rho b k'}(\boldsymbol{r}) = \sum_{\boldsymbol{k} \in L_{\mathrm{BK}}^*} \mathrm{e}^{\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r}} \tilde{w}_{\rho b k'}(\boldsymbol{k}) \quad \text{with} \quad \tilde{w}_{\rho b k'}(\boldsymbol{k}) = \frac{1}{|\mathbb{R}^2 / L_{\mathrm{BK}}|} \int_{\mathbb{R}^2 / L_{\mathrm{BK}}} \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r}} w_{\rho b k'}(\boldsymbol{r}), \tag{S78}$$

$$\boldsymbol{A}(\boldsymbol{r}) = \sum_{\boldsymbol{G} \in SL_S^*} \mathrm{e}^{\mathrm{i}\boldsymbol{G} \cdot \boldsymbol{r}} \tilde{\boldsymbol{A}}(\boldsymbol{G}) \quad \text{with} \quad \tilde{\boldsymbol{A}}(\boldsymbol{G}) = \frac{1}{|\mathbb{R}^2 / SL|} \int_{\mathbb{R}^2 / SL} \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{G} \cdot \boldsymbol{r}} \boldsymbol{A}(\boldsymbol{r}), \tag{S79}$$

$$\Delta(\boldsymbol{r}) = \sum_{\boldsymbol{G} \in SL_S^*} \mathrm{e}^{\mathrm{i}\boldsymbol{G} \cdot \boldsymbol{r}} \tilde{\Delta}(\boldsymbol{G}) \quad \text{with} \quad \tilde{\Delta}(\boldsymbol{G}) = \frac{1}{|\mathbb{R}^2 / SL|} \int_{\mathbb{R}^2 / SL} \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{G} \cdot \boldsymbol{r}} \Delta(\boldsymbol{r}). \tag{S80}$$

Substituting these Fourier series to the BdG equation (S58), shifting the $\boldsymbol{k}$ sums properly, writing the $\boldsymbol{k}$ sum over the whole space L_{BK}^* as a sum over the superlattice Brillouin zone $L_{\mathrm{BK}}^* / SL_S^*$ (in different schemes) plus shifted copies of this,

$$\sum_{\boldsymbol{k} \in L_{\mathrm{BK}}^*} g(\boldsymbol{k}) = \sum_{\boldsymbol{k} \in L_{\mathrm{BK}}^* / SL_S^*} \sum_{\boldsymbol{G} \in SL_S^*} g(\boldsymbol{k} + \boldsymbol{G}) \tag{S81}$$

(g being a test function), and using the linear independence of the exponentials yields for all $\boldsymbol{k}, \boldsymbol{k}' \in L_{\mathrm{BK}}^* / SL_S^*$, $\boldsymbol{G} \in SL_S^*$ the BdG equation in Fourier space,

$$\sum_{\boldsymbol{G}' \in SL_S^*} \tilde{\mathcal{H}}_{\mathrm{BdG}, \boldsymbol{G}\boldsymbol{G}'}^{\rho}(\boldsymbol{k}) \tilde{w}_{\rho b k'}(\boldsymbol{k} + \boldsymbol{G}') = E_{\rho b k'} \tilde{w}_{\rho b k'}(\boldsymbol{k} + \boldsymbol{G}). \tag{S82}$$

Here the Nambu-matrix

$$\tilde{\mathcal{H}}_{\mathrm{BdG}, \boldsymbol{G}\boldsymbol{G}'}^{\rho}(\boldsymbol{k}) \coloneqq \begin{pmatrix} \tilde{\mathcal{H}}_{\boldsymbol{G}\boldsymbol{G}'}^{\rho}(\boldsymbol{k}) & \tilde{\Delta}(\boldsymbol{G} - \boldsymbol{G}') \\ \tilde{\Delta}^*(\boldsymbol{G}' - \boldsymbol{G}) & -\tilde{\mathcal{H}}_{\boldsymbol{G}\boldsymbol{G}'}^{\rho}(\boldsymbol{k}) \end{pmatrix} \tag{S83}$$

is the Fourier-space version of the BdG Hamiltonian matrix element with

$$\tilde{\mathcal{H}}_{\boldsymbol{G}\boldsymbol{G}'}^{\rho}(\boldsymbol{k}) \coloneqq \hbar v_{\mathrm{F}} \boldsymbol{\sigma}^{\rho} \cdot \left((\boldsymbol{k} + \boldsymbol{G}) \delta_{\boldsymbol{G}\boldsymbol{G}'} + \rho \tilde{\boldsymbol{A}}(\boldsymbol{G} - \boldsymbol{G}') \right) - \mu \delta_{\boldsymbol{G}\boldsymbol{G}'} \tag{S84}$$

being the Fourier-space version of the noninteracting Hamiltonian matrix element.

By collecting the $\boldsymbol{G}, \boldsymbol{G}'$ components $\tilde{\mathcal{H}}_{\mathrm{BdG}, \boldsymbol{G}\boldsymbol{G}'}^{\rho}(\boldsymbol{k})$ into a countably infinite $\boldsymbol{G}$-space matrix

$$\underline{\tilde{\mathcal{H}}}_{\mathrm{BdG}}^{\rho}(\boldsymbol{k}) \coloneqq \left(\tilde{\mathcal{H}}_{\mathrm{BdG}, \boldsymbol{G}\boldsymbol{G}'}^{\rho}(\boldsymbol{k}) \right)_{\boldsymbol{G}, \boldsymbol{G}' \in SL_S^*} \tag{S85}$$

and the $\boldsymbol{G}$ components $\tilde{w}_{\rho b k'}(\boldsymbol{k} + \boldsymbol{G})$ into a countably infinite $\boldsymbol{G}$-space vector

$$\underline{\tilde{w}}_{\rho b k'}(\boldsymbol{k}) \coloneqq \left(\tilde{w}_{\rho b k'}(\boldsymbol{k} + \boldsymbol{G}) \right)_{\boldsymbol{G} \in SL_S^*} \tag{S86}$$

the BdG equation (S82) becomes a matrix eigenvalue equation

$$\underline{\tilde{\mathcal{H}}}_{\mathrm{BdG}}^{\rho}(\boldsymbol{k}) \underline{\tilde{w}}_{\rho b k'}(\boldsymbol{k}) = E_{\rho b k'} \underline{\tilde{w}}_{\rho b k'}(\boldsymbol{k}). \tag{S87}$$

Obviously we must have $\boldsymbol{k}' = \boldsymbol{k}$, so we may set

$$\tilde{w}_{\rho b k'}(\boldsymbol{k} + \boldsymbol{G}) = \delta_{\boldsymbol{k}\boldsymbol{k}'} \tilde{w}_{\rho b k}(\boldsymbol{k} + \boldsymbol{G}) \tag{S88}$$

and the BdG equation in Fourier space reads

$$\tilde{\underline{\mathcal{H}}}^{\rho}_{\mathrm{BdG}}(\boldsymbol{k})\underline{\tilde{w}}_{\rho b\boldsymbol{k}}(\boldsymbol{k}) = E_{\rho b\boldsymbol{k}}\underline{\tilde{w}}_{\rho b\boldsymbol{k}}(\boldsymbol{k}). \tag{S89}$$

This is clearly a separate problem for each $\boldsymbol{k} \in L^*_{\mathrm{BK}}/SL^*_S$, and for each $\boldsymbol{k}$ there are exactly $2 \times 2 \times |SL^*_S|$ (the matrix size) solutions labelled by the band index b. Thus our original ansatz $n = (b, \boldsymbol{k})$ is consistent. Equation (S89) is the form of the BdG equation implemented and solved in the numerics.

Note that in the reduced zone scheme here $\boldsymbol{k} = k_1\boldsymbol{G}_1 + k_2\boldsymbol{G}_2 \in L^*_{\mathrm{BK}}/SL^*$ is periodic in both $k_1 \in [-\frac{1}{2}, \frac{1}{2}[$ and $k_2 \in [-\frac{1}{2}, \frac{1}{2}[$, so that both k_1 and k_2 are periodic Bloch momenta. In this case the $\boldsymbol{G}$ translations in (S85) in both $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$ directions are transformed to new bands. This is also traditionally called the reduced zone (or the repeated zone) scheme. However, in the case of $\boldsymbol{A}$ and Δ being constant in the $\boldsymbol{t}_2$ direction (the 1D potential case), we are also allowed to choose the mixed zone scheme, as discussed in section II B. In this case $\boldsymbol{k} = k_1\boldsymbol{G}_1 + k_2\boldsymbol{G}_2 \in L^*_{\mathrm{BK}}/SL^*_1$ is periodic in $k_1 \in [-\frac{1}{2}, \frac{1}{2}[$ but not in $k_2 \in\,]-\infty, \infty[$, so that k_1 is a periodic Bloch momentum while k_2 is a nonperiodic real momentum. In this case the $\boldsymbol{G}$ translations in (S85) are done only in $\boldsymbol{G}_1$ direction, and only these are transformed to new bands. Traditionally the k_1 direction is then called to be in the reduced zone scheme and the k_2 direction in the extended zone scheme. This is why we call this the mixed zone scheme.

Equation (S88) can be used to write the Fourier series (S78) of w in the Bloch form. Further dividing the $\boldsymbol{k}'$ sum as in (S81) gives the Fourier series in the Bloch form

$$w_{\rho b\boldsymbol{k}}(\boldsymbol{r}) = \sum_{\boldsymbol{k}' \in L^*_{\mathrm{BK}}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}} \tilde{w}_{\rho b\boldsymbol{k}}(\boldsymbol{k}') = \sum_{\boldsymbol{k}' \in L^*_{\mathrm{BK}}/SL^*_S} \sum_{\boldsymbol{G} \in SL^*_S} \mathrm{e}^{\mathrm{i}(\boldsymbol{k}'+\boldsymbol{G})} \tilde{w}_{\rho b\boldsymbol{k}}(\boldsymbol{k}' + \boldsymbol{G}) \tag{S90}$$

$$= \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \sum_{\boldsymbol{G} \in SL^*_S} \mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} \tilde{w}_{\rho b\boldsymbol{k}}(\boldsymbol{k} + \boldsymbol{G}), \tag{S91}$$

where the function multiplying $\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}$ is periodic in translations of the superlattice SL.

B. Normalization of eigenvectors

Equation (S89) is the form of the BdG equation we are solving numerically. For the eigenvectors we choose the normalization $\left\|\underline{\tilde{w}}_{\rho b\boldsymbol{k}}(\boldsymbol{k})\right\| = 1$. To see what this means for the eigenstates $|w_{\rho b\boldsymbol{k}}\rangle$, we may use Parseval's theorem

$$\sum_{\boldsymbol{k}' \in L^*_{\mathrm{BK}}} \left\|\tilde{w}_{\rho b\boldsymbol{k}}(\boldsymbol{k}')\right\|^2 = \frac{1}{|\mathbb{R}^2/L_{\mathrm{BK}}|} \int_{\mathbb{R}^2/L_{\mathrm{BK}}} \mathrm{d}\boldsymbol{r} \left\|w_{\rho b\boldsymbol{k}}(\boldsymbol{r})\right\|^2 = \frac{1}{V} \langle w_{\rho b\boldsymbol{k}}|w_{\rho b\boldsymbol{k}}\rangle. \tag{S92}$$

On the other hand dividing the $\boldsymbol{k}$ sum as in (S81), using (S88), and using the definition (S86), the l.h.s. gives

$$\sum_{\boldsymbol{k}' \in L^*_{\mathrm{BK}}} \left\|\tilde{w}_{\rho b\boldsymbol{k}}(\boldsymbol{k}')\right\|^2 = \sum_{\boldsymbol{G} \in SL^*_S} \left\|\tilde{w}_{\rho b\boldsymbol{k}}(\boldsymbol{k} + \boldsymbol{G})\right\|^2 = \left\|\underline{\tilde{w}}_{\rho b\boldsymbol{k}}(\boldsymbol{k})\right\|^2 = 1, \tag{S93}$$

which then yields the normalization

$$\langle w_{\rho b\boldsymbol{k}}|w_{\rho b\boldsymbol{k}}\rangle = V. \tag{S94}$$

C. Self-consistency equation in Fourier space

Using the Fourier series (S79), (S80), and (S91) we may write the self-consistency equation (S77) in Fourier space as

$$\tilde{\Delta}_\alpha(\boldsymbol{G}) = -\frac{\lambda}{V} \sum_{\rho b_+} \sum_{\boldsymbol{k} \in L^*_{\mathrm{BK}}/SL^*_S} \sum_{\boldsymbol{G}' \in SL^*_S} \tilde{u}_{\rho b_+\boldsymbol{k},\alpha}(\boldsymbol{k} + \boldsymbol{G}')\tilde{v}^*_{\rho b_+\boldsymbol{k},\alpha}(\boldsymbol{k} + \boldsymbol{G}' - \boldsymbol{G})\tanh\!\left(\frac{E_{\rho b_+\boldsymbol{k}}}{2k_{\mathrm{B}}T}\right) \tag{S95}$$

for all $\boldsymbol{G} \in SL^*_S$. Assuming the Born–von Kármán cell to be large, i.e. L_{BK} to be sparse or L^*_{BK} to be dense, we may approximate the $\boldsymbol{k}$ sum as an integral

$$\sum_{\boldsymbol{k} \in L^*_{\mathrm{BK}}/SL^*} \approx \frac{V}{(2\pi)^2} \int_{\mathbb{R}^2/SL^*} \mathrm{d}\boldsymbol{k}, \tag{S96}$$

yielding the self-consistency equation

$$\tilde{\Delta}_\alpha(\boldsymbol{G}) = -\frac{\lambda}{(2\pi)^2} \sum_{\rho b_+} \sum_{\boldsymbol{G}' \in SL_S^*} \int_{\mathbb{R}^2/SL_S^*} \mathrm{d}\boldsymbol{k}\, \tilde{u}_{\rho b_+ \boldsymbol{k}, \alpha}(\boldsymbol{k}+\boldsymbol{G}') \tilde{v}_{\rho b_+ \boldsymbol{k}, \alpha}^*(\boldsymbol{k}+\boldsymbol{G}'-\boldsymbol{G}) \tanh\left(\frac{E_{\rho b_+ \boldsymbol{k}}}{2 k_\mathrm{B} T}\right). \tag{S97}$$

This form, where the integration region is the rather abstract $\mathbb{R}^2/SL_S^*$, is convenient when doing analytical calculations. But in numerical calculations it is easier to integrate over simpler areas instead, which is done next by change of variables.

1. Reduced zone scheme

In the reduced zone scheme we have $SL_S = SL_\mathrm{RZ} = SL$, meaning that the integration region $\mathbb{R}^2/SL_S^* = \mathbb{R}^2/SL^*$ can be interpreted as the parallelogram defined by $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$. Making a change of variables with the function

$$\phi : \left[-\tfrac{1}{2}, \tfrac{1}{2}\right[^2 \to \mathbb{R}^2/SL^*, \quad \phi(k_1, k_2) = k_1 \boldsymbol{G}_1 + k_2 \boldsymbol{G}_2 \tag{S98}$$

the $\boldsymbol{k}$ integral in (S97) can be written as

$$\int_{\mathbb{R}^2/SL^*} \mathrm{d}\boldsymbol{k}\, g(\boldsymbol{k}) = \int_{\phi([-\frac{1}{2},\frac{1}{2}[^2)} \mathrm{d}\boldsymbol{k}\, g(\boldsymbol{k}) = \int_{[-\frac{1}{2},\frac{1}{2}[^2} \mathrm{d}\boldsymbol{k}\, (g \circ \phi)(\boldsymbol{k})|J_\phi(\boldsymbol{k})| = |\boldsymbol{G}_1 \times \boldsymbol{G}_2| \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{d}k_1 \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{d}k_2\, g(k_1 \boldsymbol{G}_1 + k_2 \boldsymbol{G}_2) \tag{S99}$$

(g being a test function) where the absolute value of the Jacobian determinant,

$$|J_\phi(\boldsymbol{k})| = |\boldsymbol{G}_1 \times \boldsymbol{G}_2| = \left|\mathbb{R}^2/SL^*\right|, \tag{S100}$$

is the area of the parallelogram defined by $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$.

The self-consistency equation (S97) then becomes

$$\tilde{\Delta}_\alpha(\boldsymbol{G}) = -\frac{\lambda}{(2\pi)^2} |\boldsymbol{G}_1 \times \boldsymbol{G}_2| \sum_{\rho b_+} \sum_{\boldsymbol{G}' \in SL^*} \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{d}k_1 \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{d}k_2$$
$$\tilde{u}_{\rho b_+, \alpha}(k_1 \boldsymbol{G}_1 + k_2 \boldsymbol{G}_2 + \boldsymbol{G}') \tilde{v}_{\rho b_+, \alpha}^*(k_1 \boldsymbol{G}_1 + k_2 \boldsymbol{G}_2 + \boldsymbol{G}' - \boldsymbol{G}) \tanh\left(\frac{E_{\rho b_+}(k_1 \boldsymbol{G}_1 + k_2 \boldsymbol{G}_2)}{2 k_\mathrm{B} T}\right), \tag{S101}$$

where we dropped $k_1 \boldsymbol{G}_1 + k_2 \boldsymbol{G}_2$ from the subscripts of $\tilde{u}$ and $\tilde{v}$ and denoted $E_{\rho b_+}(\boldsymbol{k}) := E_{\rho b_+ \boldsymbol{k}}$. In the numerics we have to make a cutoff to the countably infinite $\boldsymbol{G}'$ and b_+ sums, both corresponding to a cutoff at high energies. This cutoff can be seen to come from the electron–phonon coupling.

2. Mixed zone scheme

In the mixed zone scheme we have $SL_S = SL_\mathrm{MZ} = SL_1$, meaning that the integration region $\mathbb{R}^2/SL_S^* = \mathbb{R}^2/SL_1^*$ can be interpreted as a semi-infinite parallelogram, where the finite side is $\boldsymbol{G}_1$ and the infinite side is in the $\boldsymbol{G}_2$ direction. Making a change of variables with the function

$$\phi : \left[-\tfrac{1}{2}, \tfrac{1}{2}\right[\times \mathbb{R} \to \mathbb{R}^2/SL_1^*, \quad \phi(k_1, k_2) = k_1 \boldsymbol{G}_1 + k_2 \boldsymbol{G}_2 \tag{S102}$$

the $\boldsymbol{k}$ integral in (S97) can be written as

$$\int_{\mathbb{R}^2/SL_1^*} \mathrm{d}\boldsymbol{k}\, g(\boldsymbol{k}) = \int_{\phi([-\frac{1}{2},\frac{1}{2}[\times\mathbb{R})} \mathrm{d}\boldsymbol{k}\, g(\boldsymbol{k}) = \int_{[-\frac{1}{2},\frac{1}{2}[\times\mathbb{R}} \mathrm{d}\boldsymbol{k}\, (g \circ \phi)(\boldsymbol{k})|J_\phi(\boldsymbol{k})| \tag{S103}$$

$$= |\boldsymbol{G}_1 \times \boldsymbol{G}_2| \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{d}k_1 \int_{-\infty}^{\infty} \mathrm{d}k_2\, g(k_1 \boldsymbol{G}_1 + k_2 \boldsymbol{G}_2) \tag{S104}$$

(g being a test function) where the absolute value of the Jacobian determinant,

$$|J_\phi(\boldsymbol{k})| = |\boldsymbol{G}_1 \times \boldsymbol{G}_2| = \left|\mathbb{R}^2/SL^*\right|, \tag{S105}$$

is the area of the parallelogram defined by $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$.

The self-consistency equation (S97) then becomes

$$\tilde{\Delta}_\alpha(\boldsymbol{G}) = -\frac{\lambda}{(2\pi)^2} |\boldsymbol{G}_1 \times \boldsymbol{G}_2| \sum_{\rho b_+} \sum_{\boldsymbol{G}' \in SL^{1*}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{d}k_1 \int_{-\infty}^{\infty} \mathrm{d}k_2$$

$$\tilde{u}_{\rho b_+, \alpha}(k_1\boldsymbol{G}_1 + k_2\boldsymbol{G}_2 + \boldsymbol{G}') \tilde{v}_{\rho b_+, \alpha}^*(k_1\boldsymbol{G}_1 + k_2\boldsymbol{G}_2 + \boldsymbol{G}' - \boldsymbol{G}) \tanh\left(\frac{E_{\rho b_+}(k_1\boldsymbol{G}_1 + k_2\boldsymbol{G}_2)}{2k_\mathrm{B}T}\right), \tag{S106}$$

where we dropped $k_1\boldsymbol{G}_1 + k_2\boldsymbol{G}_2$ from the subscripts of $\tilde{u}$ and $\tilde{v}$ and denoted $E_{\rho b_+}(\boldsymbol{k}) := E_{\rho b_+\boldsymbol{k}}$. While in the reduced zone scheme in (S101) both the k_1 and k_2 momentum directions are cut off in the $\boldsymbol{G}'$ sum, in this case only the k_1 momentum direction is cut off in the $\boldsymbol{G}'$ sum while the k_2 direction is handled by a momentum cutoff in the limits of the corresponding improper integral. Also the band sum b_+ has a cutoff but it is generally different from the one in the reduced zone scheme, as the meaning of bands is different.

D. Ground-state energy expectation values

Using the Hamiltonian (S72) and equation (S76) the energy density expectation value in the ground state for the order parameter (Δ_A, Δ_B) can be shown to be

$$\frac{1}{V} \langle H_\mathrm{BdG}(\Delta_A, \Delta_B) \rangle = \frac{1}{V} \sum_{\sigma\rho b_+} \sum_{\boldsymbol{k} \in L_\mathrm{BK}^*/SL_S^*} E_{\rho b_+\boldsymbol{k}}(\Delta_A, \Delta_B) \left\langle \gamma_{\sigma\rho b_+\boldsymbol{k}}^\dagger \gamma_{\sigma\rho b_+\boldsymbol{k}} \right\rangle + \frac{1}{V} E^\mathrm{gs}(\Delta_A, \Delta_B) \tag{S107}$$

$$\approx -\frac{1}{(2\pi)^2} \sum_{\rho b_+} \int_{\mathbb{R}^2/SL_S^*} \mathrm{d}\boldsymbol{k} \, E_{\rho b_+\boldsymbol{k}}(\Delta_A, \Delta_B) \tanh\left(\frac{E_{\rho b_+\boldsymbol{k}}(\Delta_A, \Delta_B)}{2k_\mathrm{B}T}\right) + \frac{1}{V} E^0(\Delta_A, \Delta_B) - 4\mu\delta(0), \tag{S108}$$

where in the last step we also approximated the sum as an integral as in (S96). We would like to show that a ground state with a zero phase difference between the Δ components has the lowest energy, and thus we define for each $\delta \in \mathbb{R}$ the *Josephson energy density*

$$\frac{E_J(\theta)}{V} := \frac{1}{V} \left\langle H_\mathrm{BdG}(\mathrm{e}^{\mathrm{i}\theta/2}\delta, \mathrm{e}^{-\mathrm{i}\theta/2}\delta) \right\rangle - \frac{1}{V} \left\langle H_\mathrm{BdG}(\delta, \delta) \right\rangle = \tag{S109}$$

$$-\frac{1}{(2\pi)^2} \sum_{\rho b_+} \int_{\mathbb{R}^2/SL_S^*} \mathrm{d}\boldsymbol{k} \left[E_{\rho b_+\boldsymbol{k}}(\mathrm{e}^{\mathrm{i}\theta/2}\delta, \mathrm{e}^{-\mathrm{i}\theta/2}\delta) \tanh\left(\frac{E_{\rho b_+\boldsymbol{k}}(\mathrm{e}^{\mathrm{i}\theta/2}\delta, \mathrm{e}^{-\mathrm{i}\theta/2}\delta)}{2k_\mathrm{B}T}\right) - E_{\rho b_+\boldsymbol{k}}(\delta, \delta) \tanh\left(\frac{E_{\rho b_+\boldsymbol{k}}(\delta, \delta)}{2k_\mathrm{B}T}\right) \right], \tag{S110}$$

where the E^0/V and $4\mu\delta(0)$ terms cancel out because they do not contain the phases of $\Delta_{A/B}$. The integral is then calculated in the different schemes as described in section V C.

E. Superfluid weight

Writing the result of [10] for the superfluid weight D^s in the superlattice-folded picture near the Dirac points, we get for the $\mu, \nu \in \{x, y\}$ component

$$D_{\mu\nu}^\mathrm{s} = \frac{1}{V} \sum_{\rho bb'\boldsymbol{k}} \sum_{\boldsymbol{k} \in L_\mathrm{BK}^*/SL_S^*} \mathcal{F}_{\rho bb'\boldsymbol{k}} \bigg(\underline{\tilde{w}}_{\rho b\boldsymbol{k}}^\dagger(\boldsymbol{k}) \partial_\mu \underline{\tilde{\mathcal{H}}}_\mathrm{BdG}^\rho(\boldsymbol{k}) \tau_z \underline{\tilde{w}}_{\rho b'\boldsymbol{k}}(\boldsymbol{k}) \underline{\tilde{w}}_{\rho b'\boldsymbol{k}}^\dagger(\boldsymbol{k}) \tau_z \partial_\nu \underline{\tilde{\mathcal{H}}}_\mathrm{BdG}^\rho(\boldsymbol{k}) \underline{\tilde{w}}_{\rho b\boldsymbol{k}}(\boldsymbol{k})$$

$$- \underline{\tilde{w}}_{\rho b\boldsymbol{k}}^\dagger(\boldsymbol{k}) \partial_\mu \underline{\tilde{\mathcal{H}}}_\mathrm{BdG}^\rho(\boldsymbol{k}) \underline{\tilde{w}}_{\rho b'\boldsymbol{k}}(\boldsymbol{k}) \underline{\tilde{w}}_{\rho b'\boldsymbol{k}}^\dagger(\boldsymbol{k}) \partial_\nu \underline{\tilde{\mathcal{H}}}_\mathrm{BdG}^\rho(\boldsymbol{k}) \underline{\tilde{w}}_{\rho b\boldsymbol{k}}(\boldsymbol{k}) \bigg), \tag{S111}$$

where the prefactor is

$$\mathcal{F}_{\rho bb'\boldsymbol{k}} = \begin{cases} f'(E_{\rho b\boldsymbol{k}}), & \text{if } E_{\rho b\boldsymbol{k}} = E_{\rho b'\boldsymbol{k}}, \\ \frac{f(E_{\rho b\boldsymbol{k}}) - f(E_{\rho b'\boldsymbol{k}})}{E_{\rho b\boldsymbol{k}} - E_{\rho b'\boldsymbol{k}}}, & \text{otherwise}, \end{cases} \tag{S112}$$
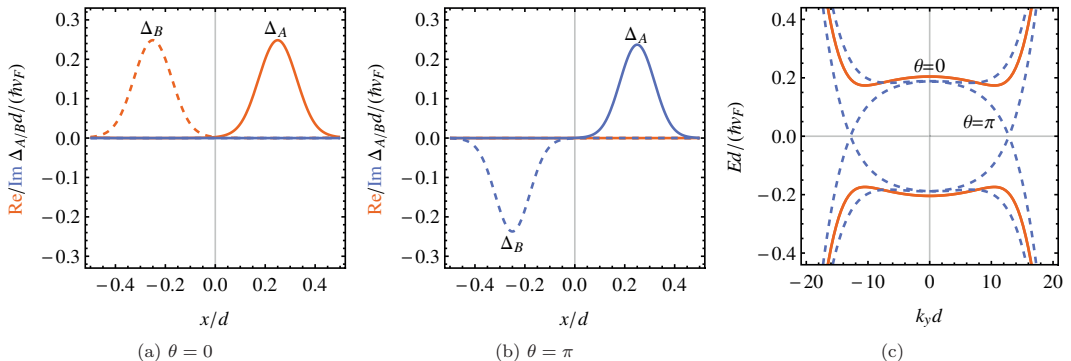
Figure S1. (a,b) Profile of the self-consistent $\Delta_{A/B}$ (A joined, B dashed lines) for the two initial guesses with the relative phases (a) $\theta = 0$ and (b) $\theta = \pi$, along the line $(x,0)$. In (a) the imaginary part is zero while in (b) the real part is zero. (c) Profile of the corresponding dispersion relation for $\theta = 0$ (joined lines) and $\theta = \pi$ (dashed lines) along the line $(0, k_y)$ in the mixed zone scheme. Here $\boldsymbol{A} = \boldsymbol{A}_{\cos}^{1D}$, $\beta = 30$, $\mu = 0$, $T = 0$, and $\lambda/(\hbar v_F d) = -0.01$.

the b, b' band sums are calculated over *both the positive- and negative-energy bands*, the partial derivatives are shortly denoted as $\partial_\mu := \partial_{k_\mu}$, and the energies and eigenvectors are calculated from the BdG equation (S89). Since Δ is $\boldsymbol{k}$ independent, we have $\partial_\mu \underline{\tilde{\mathcal{H}}}_{\mathrm{BdG}}^\rho(\boldsymbol{k}) = \tau_z \partial_\mu \underline{\tilde{\mathcal{H}}}^\rho(\boldsymbol{k}) = \hbar v_{\mathrm{F}} \tau_z \sigma_\mu^\rho$, yielding

$$D_{\mu\nu}^{\mathrm{s}} = \frac{(\hbar v_{\mathrm{F}})^2}{(2\pi)^2} \sum_{\rho b b'} \int_{\mathbb{R}^2/SL_S^*} \mathrm{d}\boldsymbol{k} \, \mathcal{F}_{\rho b b' \boldsymbol{k}} \Bigg(\underline{\tilde{w}}_{\rho b \boldsymbol{k}}^\dagger(\boldsymbol{k}) \sigma_\mu^\rho \underline{\tilde{w}}_{\rho b' \boldsymbol{k}}(\boldsymbol{k}) \underline{\tilde{w}}_{\rho b' \boldsymbol{k}}^\dagger(\boldsymbol{k}) \sigma_\nu^\rho \underline{\tilde{w}}_{\rho b \boldsymbol{k}}(\boldsymbol{k})$$

$$- \underline{\tilde{w}}_{\rho b \boldsymbol{k}}^\dagger(\boldsymbol{k}) \tau_z \sigma_\mu^\rho \underline{\tilde{w}}_{\rho b' \boldsymbol{k}}(\boldsymbol{k}) \underline{\tilde{w}}_{\rho b' \boldsymbol{k}}^\dagger(\boldsymbol{k}) \tau_z \sigma_\nu^\rho \underline{\tilde{w}}_{\rho b \boldsymbol{k}}(\boldsymbol{k}) \Bigg), \tag{S113}$$

where we also approximated the $\boldsymbol{k}$ sum as an integral.

VI. DETAILS OF THE NUMERICAL CALCULATION

A. Solving the self-consistency equation

We start by calculating analytically the Fourier coefficients of $\boldsymbol{A}$ by (S79), which are given in (S7), (S8), (S9), and (S10). The combination of the BdG equation (S89) and the self-consistency equation (S97) is then solved by the fixed-point iteration method, *i.e.* starting from an initial guess of the pair $(\tilde{\Delta}_A, \tilde{\Delta}_B)$, solving the eigenvectors $\tilde{w}_{\rho b_+ \boldsymbol{k}} = (\tilde{u}_{\rho b_+ \boldsymbol{k}}, \tilde{v}_{\rho b_+ \boldsymbol{k}})^\mathsf{T}$ and eigenvalues $E_{\rho b_+ \boldsymbol{k}}$ from the BdG equation (S89), using these eigenvectors and eigenvalues to calculate new values for $\tilde{\Delta}_A$ and $\tilde{\Delta}_B$ from (S97), and then solving the BdG equation again with these new $\tilde{\Delta}$'s. This iteration is then continued until convergence to some relative or absolute error in all of the components of $\tilde{\Delta}_{A/B}$.

B. Initial guess of the order parameter

The initial guess of the order parameter is always chosen such that both sublattice components are the same real constant in space, $\Delta_A(\boldsymbol{r}) = \Delta_B(\boldsymbol{r}) = 0.1\beta|\lambda|/d^2$. The exact value of the constant seems to have no effect on the result of the fixed-point iteration, merely affecting the speed of convergence, which is understandable from the Banach fixed-point theorem. In Fourier space the initial guess reads $\tilde{\Delta}_A(\boldsymbol{G}) = \tilde{\Delta}_B(\boldsymbol{G}) = 0.1\beta|\lambda|/d^2 \delta_{\boldsymbol{G},0}$. One should note that above we are fixing the overall phase of (Δ_A, Δ_B) to be real, since it can be shown that starting from a given overall phase, the fixed-point iteration conserves that phase at each iteration.

We are also fixing the relative phase of Δ_A and Δ_B to zero, and the justification for this is discussed next. It can be numerically shown that the initial guess $(\Delta_A, \Delta_B) = (\mathrm{e}^{\mathrm{i}\theta/2}\delta, \mathrm{e}^{-\mathrm{i}\theta/2}\delta)$ with $\theta \neq \pi$, $\delta \in \mathbb{R}$ always converges to the $\theta = 0$ solution, shown in figure S1(a) for the 1D cosine potential $\boldsymbol{A}_{\cos}^{1D}$, by the fixed-point iteration. On the other hand the $\theta = \pi$ initial guess converges to a different solution, shown in figure S1(b). The dispersion relations of these two
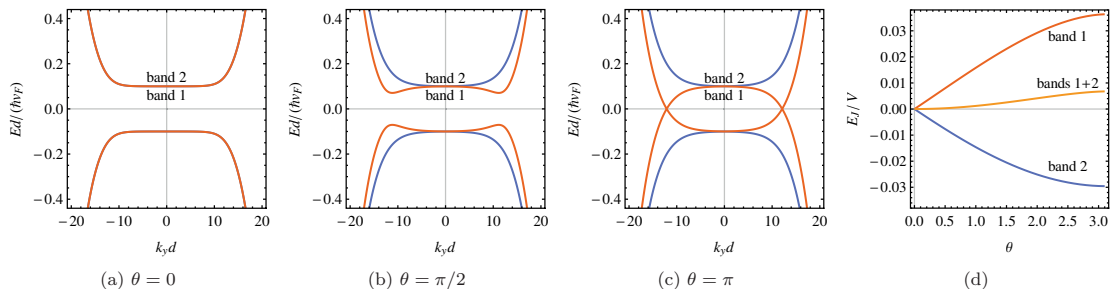
Figure S2. Effect of the relative phase θ on the non-self-consistent, constant order parameter $(\Delta_A, \Delta_B) = (e^{i\theta/2}\delta, e^{-i\theta/2}\delta)$ for $\boldsymbol{A}_{\text{cos}}^{\text{1D}}$ with $\beta = 30$, $\mu = 0$, $T = 0$, and $\delta = 0.1\hbar v_F/d$. (a–c) Dispersion of two of the lowest energy bands with increasing θ in the mixed zone scheme. (d) Ground-state energy density difference [the Josephson energy density (S110)] between the $\theta \neq 0$ and $\theta = 0$ solutions, by calculating only the contribution from the lowest band 1, the second lowest band 2, and both bands 1 and 2.

different solutions are shown in figure S1(c), showing how the degeneracy of the $\theta = 0$ state is lifted and how the gap is closed in the $\theta = \pi$ state. The dispersion relations alone can be used to calculate the ground-state energies of these states by using (S110), showing that the $\theta = 0$ solution always yields a lower energy, also at finite temperatures. This allows us to discard the $\theta = \pi$ solution and to concentrate only on the $\theta = 0$ solution, which, as discussed above, can always be reached by using the initial guess with $\theta = 0$.

The above behavior as a function of the relative phase θ can be understood by using a constant, non-self-consistent $(\Delta_A, \Delta_B) = (e^{i\theta/2}\delta, e^{-i\theta/2}\delta)$ with $\delta \in \mathbb{R}$. The dispersion of two of the lowest positive-energy bands is shown in figures S2(a–c) as a function of θ showing how the finite θ removes the degeneracy. Two competing effects are observed: energies in band 2 are slightly increased (integral-wise), while the decrease in energy in band 1 is more dominant. This can be seen also in figure S2(d), presenting the ground-state energy density difference (the Josephson energy density) between the $\theta \neq 0$ and $\theta = 0$ solutions by (S110), whose opposite value at $T = 0$ is essentially given by the difference in the $\boldsymbol{k}$ integral of the corresponding dispersions. Looking at the contributions from the individual bands 1 and 2 it is clearly seen how the contribution from band 1 is more dominant, giving the net result that the $\theta \neq 0$ solution always gives a higher ground-state energy than the $\theta = 0$ solution. At finite temperatures the behavior is otherwise the same but with smaller energy differences. The same qualitative behavior is seen for all the tested potentials.

---

[1] X. Zheng, *Efficient Fourier Transforms on Hexagonal Arrays*, PhD thesis, University of Florida (2007).
[2] H. Suzuura and T. Ando, Phys. Rev. B **65**, 235412 (2002).
[3] Suzuura & Ando [2] get a reduction factor in front due to a different definition of $\boldsymbol{u}$.
[4] Note that δ_{jx}, δ_{jy} are the components of $\boldsymbol{\delta}_j$, not Kronecker delta symbols.
[5] M. A. H. Vozmediano, M. I. Katsnelson, and F. Guinea, Phys. Rep. **496**, 109 (2010).
[6] Y. V. Nazarov and J. Danon, *Advanced Quantum Mechanics: A practical guide*, 1st ed. (Cambridge University Press, New York, 2013).
[7] C. W. J. Beenakker, Phys. Rev. Lett. **97**, 067007 (2006).
[8] M. Titov and C. W. J. Beenakker, Phys. Rev. B **74**, 041401(R) (2006).
[9] N. B. Kopnin and E. B. Sonin, Phys. Rev. B **82**, 014516 (2010).
[10] L. Liang, T. I. Vanhala, S. Peotta, T. Siro, A. Harju, and P. Törmä, Phys. Rev. B **95**, 024515 (2017).

II

MEAN-FIELD THEORY FOR SUPERCONDUCTIVITY IN TWISTED BILAYER GRAPHENE

by

Teemu J. Peltonen, Risto Ojajärvi, and Tero T. Heikkilä 2018

Physical Review B 98, 220504(R)

Mean-field theory for superconductivity in twisted bilayer graphene

Teemu J. Peltonen, Risto Ojajärvi, and Tero T. Heikkilä

Department of Physics and Nanoscience Center, University of Jyvaskyla, P.O. Box 35 (YFL), FI-40014 University of Jyvaskyla, Finland

Recent experiments show how a bilayer graphene twisted around a certain magic angle becomes superconducting as it is doped into a region with approximate flat bands. We investigate the mean-field *s*-wave superconducting state in such a system and show how the state evolves as the twist angle is tuned, and as a function of the doping level. We argue that part of the experimental findings could well be understood to result from an attractive electron-electron interaction mediated by electron-phonon coupling, but the flat-band nature of the excitation spectrum also makes the superconductivity quite unusual. For example, as the flat-band states are highly localized around certain spots in the structure, also the superconducting order parameter becomes strongly inhomogeneous.

I. INTRODUCTION

Experiments on strongly doped graphene [1–4] have shown that with proper preparations, graphene can be driven to the superconducting state. Such experiments indicate that the lack of superconductivity in undoped graphene is not necessarily due to a lack of an (effective) attractive electron-electron interaction with strength λ that would drive graphene to be superconducting, but rather the small density of states (DOS) close to the Dirac point. Technically, in contrast to the Cooper instability for metals taking place with arbitrarily small λ, superconductivity in an electron system with a massless Dirac dispersion $\epsilon_p^2 = v_F^2 p^2$ and an energy cutoff ϵ_c has a quantum critical point [5] $\lambda_c = \pi \hbar^2 v_F^2/(2\epsilon_c)$ such that for $\lambda < \lambda_c$, mean-field superconductivity does not show up at any temperature. From this perspective, doping to a potential μ leads to an increased DOS, and thereby to a nonvanishing critical temperature $T_c \approx |\mu| \exp[-(\lambda_c/\lambda - 1)\epsilon_c/|\mu| - 1]$. An alternative approach would be to change the spectrum and increase the density of states close to the Dirac point. The extreme limit would be an approximately flat band of size Ω_{FB}, where the group velocity tends to zero. In such systems the critical temperature is a linear function of the coupling strength [6,7], $T_c = \lambda \Omega_{FB}/\pi^2$, and a quite high T_c can be expected even without extra doping [8–13].

Recent observations [14] of superconductivity in twisted bilayer graphene [TBG, see Fig. 1(a)] take place in systems where theoretical studies have predicted the occurrence of asymptotically flat bands [15–25]. There have been many suggestions of an unconventional superconducting state both for regular graphene [26,27] and for TBG [24,28–36], typically directly related with the Coulomb interaction, and in some cases related with nonlocal interactions. Here, we study the mean-field theory of superconductivity in such systems, starting instead from the hypothesis that the observations could be explained with the conventional electron-phonon mechanism from the flat-band perspective [37]. This hypothesis is justified on the grounds that the relative strength and the screening of attractive and repulsive interactions are

uncertain. Furthermore, phonon-mediated attraction is considered a viable mechanism for the observed superconductivity on doped graphene [1–4,38].

In particular, we use the model of Refs. [15,20] for the spectrum of the twisted bilayer, add an on-site (leading to *s*-wave superconductivity) attractive interaction of strength λ, and evaluate the mean-field order parameter profile. We find that the order parameter, and along with it the mean-field critical temperature, have a similar nonmonotonous behavior with respect to the twist angle as in the experiments. We also predict the behavior of the density of states in the superconducting state, resulting from the peculiarities of the flat-band eigenstates and from the position dependence of the superconducting order parameter [Fig. 1(b)]. Even if our pairing interaction is quite simple, the resulting energy-dependent density
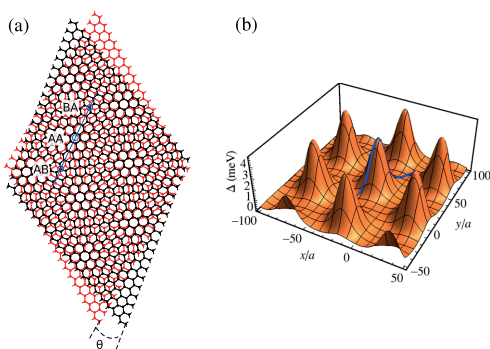
FIG. 1. (a) Twisted bilayer graphene and its moiré superlattice. The upper layer is rotated by an angle θ relative to the lower layer. (b) Position dependence of the self-consistent Δ, shown here at $T = 0$ for the magic angle $\theta = 0.96°$ and $\lambda = 5$ eV a^2. In both figures also a line passing through high-symmetry points with AB, AA, and BA stacking is shown.

of states is quite unusual. In addition, we show how doping away from the flat band eventually destroys superconductivity.

II. NORMAL STATE

We describe the normal state of TBG with the model of Refs. [15,20]. With this model, we can describe the twist angles θ at which the lattices L and L^θ of the two graphene layers are commensurate, so that the system as a whole is periodic in the moiré superlattice SL. Here, we study only the simple commensurate structures, characterized by a single rotation parameter $m \in \mathbb{N}$, for which the rotation angle is given by

$$\cos(\theta) = \frac{3m^2 + 3m + 1/2}{3m^2 + 3m + 1}. \quad (1)$$

According to Ref. [20], these structures approximate arbitrary commensurate structures. The primitive vectors of the super-lattice SL are given by $\boldsymbol{t}_1 = m\boldsymbol{a}_1 + (m+1)\boldsymbol{a}_2$, $\boldsymbol{t}_2 = -(m+1)\boldsymbol{a}_1 + (2m+1)\boldsymbol{a}_2$, and the primitive vectors of the reciprocal superlattice SL^* are $\boldsymbol{G}_1 = \frac{4\pi}{3\|\boldsymbol{t}_1\|^2}[(3m+1)\boldsymbol{a}_1 + \boldsymbol{a}_2]$, $\boldsymbol{G}_2 = \frac{4\pi}{3\|\boldsymbol{t}_1\|^2}[-(3m+2)\boldsymbol{a}_1 + (3m+1)\boldsymbol{a}_2]$, where the lattice constant of the superlattice is $\|\boldsymbol{t}_1\| = \sqrt{3m^2 + 3m + 1}\,a$ and the graphene lattice primitive vectors are $\boldsymbol{a}_1 = (1, \sqrt{3})a/2$ and $\boldsymbol{a}_2 = (-1, \sqrt{3})a/2$ with a the lattice constant [15]. In the following, we assume that $\boldsymbol{G} \in SL^*$ belongs to the reciprocal superlattice, $\boldsymbol{k} \in \mathbb{R}^2/SL^*$ to the first Brillouin zone of the superlattice, and also that the corresponding sums and integrals are restricted to these sets.

In the normal state, TBG is described by a low-energy effective Hamiltonian [15]

$$\mathcal{H}_{\rho \boldsymbol{k}}(\boldsymbol{G}, \boldsymbol{G}') = \begin{pmatrix} [\hbar v_F \boldsymbol{\sigma}^\rho \cdot (\boldsymbol{k} + \boldsymbol{G} + \rho\Delta\boldsymbol{K}/2) - \mu]\delta_{\boldsymbol{G},\boldsymbol{G}'} & t_\perp^\rho(\boldsymbol{G} - \boldsymbol{G}') \\ t_\perp^\rho(\boldsymbol{G}' - \boldsymbol{G})^\dagger & [\hbar v_F \boldsymbol{\sigma}_\theta^\rho \cdot (\boldsymbol{k} + \boldsymbol{G} - \rho\Delta\boldsymbol{K}/2) - \mu]\delta_{\boldsymbol{G},\boldsymbol{G}'} \end{pmatrix}, \quad (2)$$

where the matrix structure corresponds to the layer structure and $\rho \in \{+, -\}$ is the valley index with $+$ corresponding to $\boldsymbol{K}$ and $-$ to $\boldsymbol{K}' = -\boldsymbol{K}$. Furthermore, each entry is a 2×2 matrix due to the sublattice structure in graphene. The diagonal terms in Eq. (2) describe the Dirac dispersion in the two layers and are diagonal also in $\boldsymbol{G}$. Here, $\boldsymbol{\sigma}^\rho = (\rho\sigma_x, \sigma_y)$. For the second layer we include the rotation θ so that $\boldsymbol{\sigma}_\theta^\rho = e^{+i\theta\sigma_z/2}\boldsymbol{\sigma}^\rho e^{-i\theta\sigma_z/2}$. $\Delta\boldsymbol{K} = \boldsymbol{K}^\theta - \boldsymbol{K}$ is the relative shift of the Dirac cones between the layers. The coordinates correspond to those of layer 1 as measured from the $\boldsymbol{K}$ point, but shifted with a vector $+\Delta\boldsymbol{K}/2$ for layer 1 and $-\Delta\boldsymbol{K}/2$ for layer 2. With this choice, the relative momentum $\boldsymbol{k}$ on both layers corresponds to the same absolute momentum. Furthermore, μ is the chemical potential describing the effect of doping, here taken to be equal in both layers.

The off-diagonal terms in the Hamiltonian describe the coupling between the two layers. The matrix element at valley ρ between a state in sublattice α in layer 1 and a state in sublattice β in layer 2 is

$$t_\perp^{\rho,\alpha\beta}(\boldsymbol{G}) = \frac{1}{N}\sum_{\boldsymbol{r}} e^{-i\boldsymbol{G}\cdot(\boldsymbol{r}+\delta_{\alpha B}\delta_1)}e^{i\rho\boldsymbol{K}^\theta \cdot \boldsymbol{\delta}^{\alpha\beta}(\boldsymbol{r})}t_\perp[\boldsymbol{\delta}^{\alpha\beta}(\boldsymbol{r})], \quad (3)$$

where $\boldsymbol{\delta}^{\alpha\beta}(\boldsymbol{r})$ is the horizontal displacement vector between the site at $\boldsymbol{r}$, sublattice α in layer 1, and the nearest neighbor at sublattice β in layer 2. $\boldsymbol{\delta}_1$ denotes one of the nearest-neighbor vectors connecting the graphene A and B sublattices. The sum is over the graphene A sublattice sites in the super-lattice unit cell, and N denotes the number of these sites. For the interlayer hopping energy $t_\perp(\boldsymbol{\delta})$ we use the same Slater-Koster parametrization as in Ref. [15]. Furthermore, we approximate the interlayer coupling by only including the matrix elements with $\boldsymbol{G} \in \{0, -\boldsymbol{G}_1, -\boldsymbol{G}_1 - \boldsymbol{G}_2\}$ (valley $\boldsymbol{K}$) or $\boldsymbol{G} \in \{0, \boldsymbol{G}_1, \boldsymbol{G}_1 + \boldsymbol{G}_2\}$ (valley $\boldsymbol{K}'$), since they are an order of magnitude larger than the rest.

For $\theta \approx 1°$, the electronic dispersion becomes almost flat [19] and the group velocity $d\epsilon_p/dp$ tends towards zero. In Fig. 2 we plot the resulting normal-state dispersion

[Figs. 2(a)–2(c)] and the (local and total) density of states [Figs. 2(d)–2(i)] close to this "magic" angle. The exact value of this magic angle depends on the details of the hopping model. In our case it is around $0.96°$, i.e., somewhat lower than what was found in Ref. [19]. However, the qualitative behavior of the local density of states (LDOS) is rather similar to the previous models. In particular, there are two closely spaced DOS peaks signifying the flattening of the bands. The local density of states is plotted along the line shown in Fig. 1, including three high-symmetry points with AB, AA, and BA stacking. These correspond to $r = -1/3$, 0, and 1/3, respectively.

III. SUPERCONDUCTING STATE

We assume that there is a local attractive interaction $\lambda_{\sigma_1\sigma_2}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \delta_{\bar{\sigma}_1\sigma_2}\delta(\boldsymbol{r}_1 - \boldsymbol{r}_2)\lambda$ with strength λ, which results [7] in an order parameter $\Delta_{\alpha i}(\boldsymbol{r})$ depending only on the center-of-mass coordinate $\boldsymbol{r}$ (and sublattice α and layer i). On the other hand, the classification of the order parameter symmetries to s, d, f, etc., is based only on the relative coordinate $\boldsymbol{r}_1 - \boldsymbol{r}_2$, which in our model is always zero. Thus the symmetry is purely s wave.

We do not consider the specific nature of the pairing interaction and for the purposes of this Rapid Communication it can be mediated by phonons or other bosonic modes. This model disregards the retardation effects due to such a mechanism, but is a valid approximation to the more general Eliashberg approach for weak coupling [39,40]. That theory also shows that a direct Coulomb interaction, typically modeled via the Hubbard model, is less effective in reducing Δ than what could be naively expected, and should be included in the low-energy self-consistency equation as a Coulomb pseudopotential [7,40,41] $u^* = u/(1 + u\alpha)$, where $u = Ua^2$, U is the Hubbard interaction constant, and α is a constant measuring the amount of renormalization due to the high-energy bands above the electron-phonon cutoff frequency
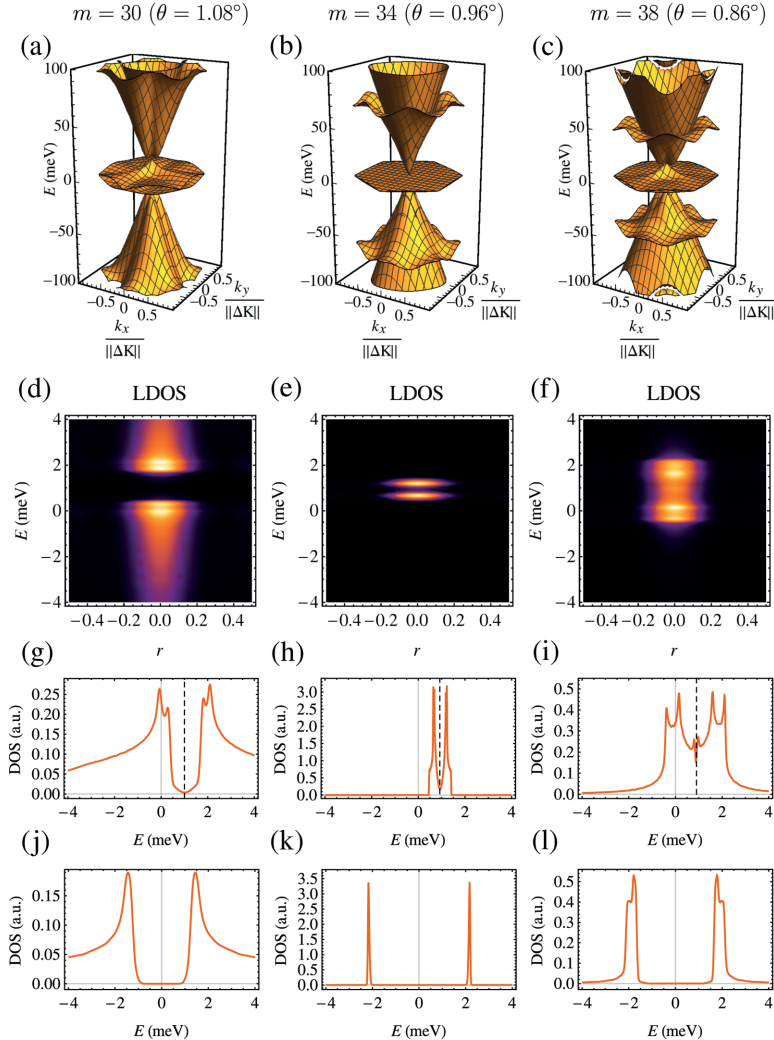
FIG. 2. (a)–(c) Normal-state dispersion, (d)–(f) local, and (g)–(i) total density of states for three different angles near the magic angle $\theta = 0.96°$ in the normal state. The bottom row (j)–(l) shows the corresponding total density of states in the superconducting state, in the case $T = 0$ and $\lambda = 5$ eV a^2 and when doped to the point μ_0 marked as a dashed line in (g)–(i).

ω_D. For TBG we find from a simplified model [7] $\alpha \approx 0.2$ eV^{-1} a^{-2}. Thus, a combination of electron-phonon and Coulomb interactions leads to an effective interaction strength $\lambda_{\text{eff}} = \lambda - u^*$. As long as $\lambda_{\text{eff}} > 0$, there is a possibility for a superconducting state even if $u > \lambda$. For example, for $U = 5$ eV, $u^* = 2.5$ eV a^2 is in the same regime as the value of λ_{eff} in Figs. 3–5. Note that in what follows, we refer to this λ_{eff} simply as λ.

Within a mean-field theory in the Cooper channel we find a self-consistency equation for a local superconducting order

parameter [7]. Assuming that this order parameter shares the periodicity of the moiré superlattice, we find the self-consistency equation

$$\Delta_{\alpha i}(\boldsymbol{G}) = \lambda \sum_{\rho, b} \sum_{\boldsymbol{G}'} \int \frac{d\boldsymbol{k}}{(2\pi)^2} \tanh\left(\frac{E_{\rho b \boldsymbol{k}}}{2 k_B T}\right)$$
$$\times \, u_{\rho b \boldsymbol{k}, \alpha i}(\boldsymbol{G}') v^*_{\rho b \boldsymbol{k}, \alpha i}(\boldsymbol{G}' - \boldsymbol{G}), \quad (4)$$

where the band sum b is calculated over the positive energy bands, $\alpha \in \{A, B\}$ is the sublattice index, $i \in \{1, 2\}$ is the
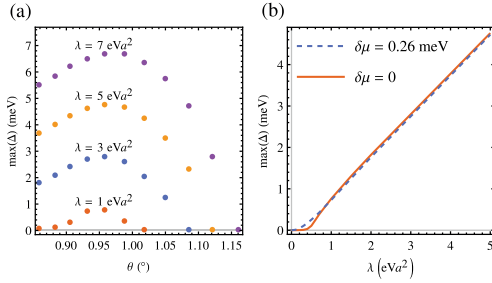
FIG. 3. Maximum of the position-dependent superconducting order parameter $\Delta(\mathbf{r})$ at $T = 0$ as a function of (a) the rotation angle and (b) the coupling strength for $\theta = 0.96°$. In (b) we also show how doping to the DOS peak affects the small-λ behavior.

FIG. 4. $\max(\Delta)$ as a function of temperature in the case $\theta = 0.96°$ for two values of λ, showing the approximate linear relation $k_B T_c \approx 0.25 \max[\Delta(T = 0)]$ for the critical temperature. The dots are the calculated values and the lines are a guide to the eye.

layer index, and $u_{\rho b \mathbf{k}}$ and $v_{\rho b \mathbf{k}}$ are the eigenvectors of the Bogoliubov–de Gennes equation,

$$\sum_{\mathbf{G}'} \begin{pmatrix} \mathcal{H}_{\rho \mathbf{k}}(\mathbf{G}, \mathbf{G}') & \Delta(\mathbf{G} - \mathbf{G}') \\ \Delta^*(\mathbf{G}' - \mathbf{G}) & -\mathcal{H}_{\rho \mathbf{k}}(\mathbf{G}, \mathbf{G}') \end{pmatrix} \begin{pmatrix} u_{\rho b \mathbf{k}}(\mathbf{G}') \\ v_{\rho b \mathbf{k}}(\mathbf{G}') \end{pmatrix}$$
$$= E_{\rho b \mathbf{k}} \begin{pmatrix} u_{\rho b \mathbf{k}}(\mathbf{G}) \\ v_{\rho b \mathbf{k}}(\mathbf{G}) \end{pmatrix}. \quad (5)$$

We solve this self-consistent order parameter with a few values of the interaction constant λ and for a few different twist angles θ close to the magic angle. We include in the sum the energy levels closest to zero energy. We have checked that the results are not sensitive to the value of the energy cutoff, which we implement as a cutoff in the b and $\mathbf{G}$ sums. For comparison between different angles, we measure the chemical potential from μ_0, corresponding to the charge neutrality and marked in Figs. 2(g)–2(i) with a dashed line, by writing $\mu = \mu_0 + \delta \mu$. The chemical potential shift μ_0 is caused by the interlayer coupling. Unless otherwise stated, all the results concern the behavior at $\delta \mu = 0$. The resulting total density of states is plotted in Figs. 2(j)–2(k), to allow for a comparison to the normal state. The corresponding local density of states (not shown) has the same localized structure as in the normal state, but the energy dependence is modified similarly as the total DOS. The effect of finite temperature on the superconducting DOS and LDOS happens solely via $\Delta(T)$, which is calculated below.

The maximum of the position-dependent Δ, which according to numerics is equal in both layers and sublattices, is plotted in Fig. 3(a) for different values of the twist angle and for four different coupling strengths. The precise angle for the maximum depends a bit on the chosen coupling strength. Moreover, $\max(\Delta)$ is almost a linear function of λ [see Fig. 3(b)], as appropriate for a flat-band superconductor [6]. This linearity is even more pronounced when the system is doped to the DOS peak at $\delta \mu \approx 0.26$ meV. Far from the magic angle, the Fermi speed $v_F(\theta)$ increases so that the chosen λ is below the critical value λ_c. This is why Δ vanishes for angles $\theta \gtrsim 1.1°$.
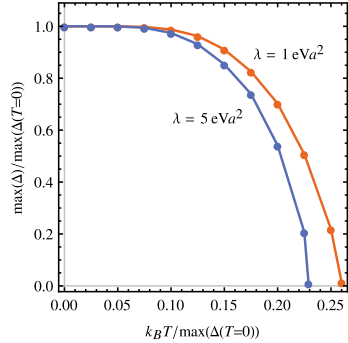
We can analyze the resulting magnitude of Δ based on a flat-band result (assuming a position-independent Δ and $E_{\rho b \mathbf{k}} \approx \Delta$ for an extreme flat band) according to which [7] $\Delta = \lambda \Omega_{FB}/\pi^2$, where $\Omega_{FB} \approx \Omega_{moiré} = 8\pi^2/(\sqrt{3}||\mathbf{t}_1||^2)$. This yields $\Delta = 1.3 \times 10^{-3} \lambda/a^2$ for $m = 34$ corresponding to the magic angle. For comparison, a linear fit to the linear region in Fig. 3(b) gives $\max(\Delta) = -0.2$ meV $+ 1.0 \times 10^{-3} \lambda/a^2$. The magnitude hence agrees very well with this simple model. Note that the precise values of these parameters especially for small λ depend on the exact value of doping as shown below.

In Fig. 4 we show the temperature dependence of Δ for $m = 34$, from which we may infer the approximate linear relation $k_B T_c \approx 0.25 \max[\Delta(T = 0)]$ for the critical temperature. The prefactor is somewhat lower than for an extreme flat band with a constant Δ, for which [7] $k_B T_c = \Delta/2$. The difference is most likely explained by the nonvanishing bandwidth and the position-dependent Δ of our model. The maximum critical temperatures for the models calculated in Fig. 3(a) range from 3 to about 20 K. The lower end of these values, calculated with $\lambda = 1$ eV a^2, is thus quite close to that found in Ref. [14].

We stress that the above result is the mean-field critical temperature; the observed resistance transition is most likely rather a Berezinskii-Kosterlitz-Thouless (BKT) transition [42,43]. Therefore, the mean-field T_c gives an upper bound for the measured transition temperature. Furthermore, even the BKT transition temperature can be calculated from the mean-field superfluid weight [44]. The mean-field results are also relevant in that the DOS and LDOS can be experimentally measured by tunneling experiments and this depends on the structure and magnitude of mean-field Δ at temperatures below the BKT transition. Note that despite the flatness of the bands, the supercurrent can be nonvanishing in the case when the eigenstate Wannier functions overlap [45], as is the case for TBG.

Besides θ dependence, we can check how doping away from the center of the two DOS peaks affects the superconducting state. In Fig. 5(a) we plot the order parameter
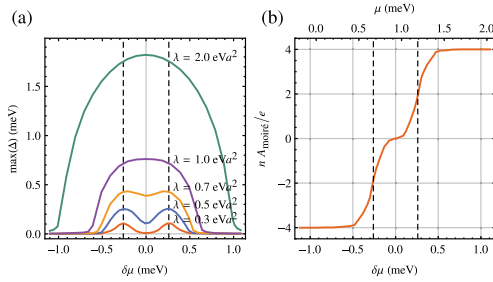
FIG. 5. Effects of electrostatic doping $\mu = \mu_0 + \delta\mu$ for $\theta = 0.96°$. (a) max(Δ) vs chemical potential for various values of λ at $T = 0$. (b) Charge density in the normal state at $T = 0$ as a function of chemical potential. The units of the charge density n are $e/A_{\text{moiré}}$, where e is the electron charge and $A_{\text{moiré}}$ is the area of the moiré unit cell. In both figures the vertical dashed lines mark the location of the DOS peaks at $\delta\mu \approx \pm0.26$ meV.

max[$\Delta(\delta\mu)$] for different values of the doping $\delta\mu$ as measured from the charge neutrality point. Close to the magic angle, for $\lambda \gtrsim 1$ eV a^2 the energy scale of superconductivity exceeds that of the normal-state dispersion, and hence the only effect of the doping is to move away from the flat-band regime, suppressing superconductivity [46]. For smaller values of λ, max(Δ) is smaller than the bandwidth, and hence doping to the DOS peaks enhances superconductivity. Especially for $\lambda \lesssim 0.3$ eV a^2 there are separate superconducting domes with doping levels close to the DOS peaks, which resembles the phase diagram in Ref. [14] for hole ($n < 0$) doping, apart from the insulating state at $n \approx -2e/A_{\text{moiré}}$. For electron doping ($n > 0$), superconductivity is absent in the experiment, whereas our model is electron-hole symmetric. Since Ref. [14] uses charge density n as a unit for the doping level while our theory is formulated in terms of the chemical potential μ, for easier comparison we show the dependence between the charge density [7] and chemical potential in Fig. 5(b). From the figure we find that the DOS peaks correspond to approximately ± 2 extra electrons per moiré unit cell.

IV. CONCLUSIONS

Concluding, we find that a BCS-type mean-field model with a relatively weak attractive interaction constant possibly even due to electron-phonon coupling can explain the occurrence of superconductivity in twisted bilayer graphene. We also make a number of predictions concerning the mean-field superconducting state, in particular, the density of states and

doping dependence. Our results form hence a checkpoint for further studies, that use a simplified picture of the TBG flat-band states or consider mechanisms beyond the one in this Rapid Communication. Our results could also have relevance in explaining the observations of superconductivity in twisted interfaces of graphite [47–49].

Our mean-field theory fails to explain the insulator state [50] found experimentally in TBG at $n \approx \pm 2e/A_{\text{moiré}}$ as well as the lack of superconductivity for electron doping [14,51]. However, the latter of these cannot be seen as a drawback of our model as in another experiment [52] some samples were found to be superconducting also on the electron-doped side, and thus it clearly depends on the samples and on the experimental setup. Regarding the insulator phase, it is plausible that the mean-field theory fails when the doping level corresponds to an integer number of electrons per superlattice unit cell. The biggest discrepancy is, however, most likely caused by the possible dependence of λ_{eff} on the charge density, because the effect of the Coulomb interaction depends on charge screening. Within the flat-band model of Ref. [40], the case $\lambda_{\text{eff}} > 0$ corresponds to a superconducting state, whereas for $\lambda_{\text{eff}} < 0$ an insulating antiferromagnetic state is realized. Thus, by taking the chemical potential dependence of λ_{eff} into account, it may be possible to describe both superconducting and insulating phases found in the experiment [14]. A detailed description would require generalizing Refs. [40,53] to the TBG case.

We point out that our simple BCS model disregards the strain effects in moiré bands, as well as the possible dependence of the interaction constant on the twist angle and doping level. Whereas such mechanisms may play a role in TBG, we believe that the simplest BCS-type mean-field superconductivity should also be considered as a viable effective model of the observations. Nevertheless, even in this case superconductivity would be highly exceptional, for example, because it can be so strongly controlled by electrostatic doping.

Note added. Recently, we became aware of Ref. [54], which addressed a similar BCS-type model as here, obtaining consistent results with this Rapid Communication. In addition to local interactions leading to s-wave superconductivity, they considered also nonlocal interactions opening the possibility to d-wave superconductivity. They found out that without including Coulomb repulsion the s-wave channel is more stable, having a higher T_c.

[1] B. M. Ludbrook, G. Levy, P. Nigge, M. Zonno, M. Schneider, D. J. Dvorak, C. N. Veenstra, S. Zhdanovich, D. Wong, P. Dosanjh *et al.*, Proc. Natl. Acad. Sci. U.S.A. **112**, 11795 (2015).

[2] A. P. Tiwari, S. Shin, E. Hwang, S.-G. Jung, T. Park, and H. Lee, J. Phys.: Condens. Matter **29**, 445701 (2017).

[3] J. Chapman, Y. Su, C. A. Howard, D. Kundys, A. N. Grigorenko, F. Guinea, A. K. Geim, I. V. Grigorieva, and R. R. Nair, Sci. Rep. **6**, 23254 (2016).

[4] S. Ichinokura, K. Sugawara, A. Takayama, T. Takahashi, and S. Hasegawa, ACS Nano **10**, 2761 (2016).

[5] N. B. Kopnin and E. B. Sonin, Phys. Rev. Lett. **100**, 246808 (2008).

[6] T. T. Heikkilä, N. B. Kopnin, and G. E. Volovik, JETP Lett. **94**, 233 (2011).

[7] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.98.220504 for details of the derivation of the self-consistency equation, for the definition of the charge density, for a simplified model of the flat-band superconducting state, and for a calculation of the Coulomb pseudopotential, which includes Refs. [55,56].

[8] V. A. Khodel' and V. R. Shaginyan, Pis'ma Zh. Eksp. Teor. Fiz. **51**, 488 (1990) [JETP Lett. **51**, 553 (1990)].

[9] N. B. Kopnin, T. T. Heikkilä, and G. E. Volovik, Phys. Rev. B **83**, 220503(R) (2011).

[10] E. Tang and L. Fu, Nat. Phys. **10**, 964 (2014).

[11] V. J. Kauppila, F. Aikebaier, and T. T. Heikkilä, Phys. Rev. B **93**, 214505 (2016).

[12] T. T. Heikkilä and G. E. Volovik, in *Basic Physics of Functionalized Graphite*, edited by P. D. Esquinazi (Springer, Berlin, 2016), pp. 123–143.

[13] T. Löthman and A. M. Black-Schaffer, Phys. Rev. B **96**, 064505 (2017).

[14] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Nature (London) **556**, 43 (2018).

[15] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, Phys. Rev. Lett. **99**, 256802 (2007).

[16] E. Suárez Morell, J. D. Correa, P. Vargas, M. Pacheco, and Z. Barticevic, Phys. Rev. B **82**, 121407(R) (2010).

[17] E. J. Mele, Phys. Rev. B **81**, 161405(R) (2010).

[18] E. J. Mele, Phys. Rev. B **84**, 235439 (2011).

[19] R. Bistritzer and A. H. MacDonald, Proc. Natl. Acad. Sci. U.S.A. **108**, 12233 (2011).

[20] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, Phys. Rev. B **86**, 155449 (2012).

[21] A. K. Geim and I. V. Grigorieva, Nature (London) **499**, 419 (2013).

[22] D. Weckbecker, S. Shallcross, M. Fleischmann, N. Ray, S. Sharma, and O. Pankratov, Phys. Rev. B **93**, 035452 (2016).

[23] S. Fang and E. Kaxiras, Phys. Rev. B **93**, 235153 (2016).

[24] A. Ramires and J. L. Lado, Phys. Rev. Lett. **121**, 146801 (2018).

[25] N. N. T. Nam and M. Koshino, Phys. Rev. B **96**, 075311 (2017).

[26] B. Uchoa and A. H. Castro Neto, Phys. Rev. Lett. **98**, 146801 (2007).

[27] R. Nandkishore, L. S. Levitov, and A. V. Chubukov, Nat. Phys. **8**, 158 (2012).

[28] H. C. Po, L. Zou, A. Vishwanath, and T. Senthil, Phys. Rev. X **8**, 031089 (2018).

[29] C. Xu and L. Balents, Phys. Rev. Lett. **121**, 087001 (2018).

[30] C.-C. Liu, L.-D. Zhang, W.-Q. Chen, and F. Yang, Phys. Rev. Lett. **121**, 217001 (2018).

[31] S. Ray and T. Das, arXiv:1804.09674.

[32] T. Huang, L. Zhang, and T. Ma, arXiv:1804.06096.

[33] J. F. Dodaro, S. A. Kivelson, Y. Schattner, X. Q. Sun, and C. Wang, Phys. Rev. B **98**, 075154 (2018).

[34] G. Baskaran, arXiv:1804.00627.

[35] B. Roy and V. Juričić, arXiv:1803.11190.

[36] H. Guo, X. Zhu, S. Feng, and R. T. Scalettar, Phys. Rev. B **97**, 235453 (2018).

[37] G. E. Volovik, JETP Lett. **107**, 516 (2018).

[38] G. Profeta, M. Calandra, and F. Mauri, Nat. Phys. **8**, 131 (2012).

[39] G. M. Éliashberg, Zh. Eksp. Teor. Fiz. **38**, 966 (1960) [Sov. Phys. JETP **11**, 696 (1960)].

[40] R. Ojajärvi, T. Hyart, M. A. Silaev, and T. T. Heikkilä, Phys. Rev. B **98**, 054515 (2018).

[41] P. Morel and P. Anderson, Phys. Rev. **125**, 1263 (1962).

[42] V. L. Berezinskii, Zh. Eksp. Teor. Fiz. **61**, 1144 (1971) [Sov. Phys. JETP **34**, 610 (1972)].

[43] J. M. Kosterlitz and D. J. Thouless, J. Phys. C: Solid State Phys. **6**, 1181 (1973).

[44] A. Julku, L. Liang, and P. Törmä, New J. Phys. **20**, 085004 (2018).

[45] S. Peotta and P. Törmä, Nat. Commun. **6**, 8944 (2015).

[46] N. B. Kopnin and T. T. Heikkilä, in *Carbon-based Superconductors: Towards High-T_c Superconductivity*, edited by J. Haruyama (CRC Press, Boca Raton, FL, 2014), Chap. 9.

[47] P. Esquinazi, Pap. Phys. **5**, 050007 (2013).

[48] A. Ballestar, J. Barzola-Quiquia, T. Scheike, and P. Esquinazi, New J. Phys. **15**, 023024 (2013).

[49] M. Stiller, P. D. Esquinazi, J. Barzola-Quiquia, and C. E. Precker, J. Low Temp. Phys. **191**, 105 (2018).

[50] B. Padhi, C. Setty, and P. W. Phillips, Nano Lett. **18**, 6175 (2018).

[51] Y. Cao, V. Fatemi, A. Demir, S. Fang, S. L. Tomarken, J. Y. Luo, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Nature (London) **556**, 80 (2018).

[52] M. Yankowitz, S. Chen, H. Polshyn, K. Watanabe, T. Taniguchi, D. Graf, A. F. Young, and C. R. Dean, arXiv:1808.07865.

[53] E. H. Hwang, R. Sensarma, and S. Das Sarma, Phys. Rev. B **82**, 195406 (2010).

[54] F. Wu, A. H. MacDonald, and I. Martin, arXiv:1805.08735.

[55] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, Rev. Mod. Phys. **81**, 109 (2009).

[56] D. K. Efetov and P. Kim, Phys. Rev. Lett. **105**, 256805 (2010).

Mean-field theory for superconductivity in twisted bilayer graphene: supplementary information

Teemu J. Peltonen,[1] Risto Ojajärvi,[1] and Tero T. Heikkilä[1]

[1]*Department of Physics and Nanoscience Center, University of Jyvaskyla,*
P.O. Box 35 (YFL), FI-40014 University of Jyvaskyla, Finland
(Dated: November 14, 2018)

Here we present the details of calculating the self-consistent order parameter of the twisted bilayer graphene. We also derive an expression that relates the charge density to the dispersion that we calculate. In addition, we present a simplified model which connects a given dispersion relation to the value of the superconducting gap. In particular, this shows why pristine graphene needs to be very strongly doped to find any signs of superconductivity, whereas a system with an approximate flat band of the size of the first Brillouin zone of the moiré superlattice can show superconductivity with the observed critical temperature even for quite weak effective attractive interaction. We furthermore present a way to include the Coulomb interactions by calculating the Coulomb pseudopotential in a simplified model.

I. DERIVATION OF THE SELF-CONSISTENCY EQUATION

The Hamiltonian for a local attractive interaction of strength $\lambda > 0$ is

$$H_{\text{int}} = -\frac{\lambda}{2} \sum_{\sigma,\alpha,i} \int d\boldsymbol{r} \, \psi_{\sigma,\alpha i}^\dagger(\boldsymbol{r}) \psi_{\bar{\sigma},\alpha i}^\dagger(\boldsymbol{r}) \psi_{\bar{\sigma},\alpha i}(\boldsymbol{r}) \psi_{\sigma,\alpha i}(\boldsymbol{r}), \tag{S1}$$

where $\psi_{\sigma,\alpha i}(\boldsymbol{r})$ is the annihilation operator for spin σ at position $\boldsymbol{r}$, layer $i \in \{1,2\}$, and sublattice $\alpha \in \{A,B\}$. Doing the mean field approximation in the Cooper channel, assuming only intervalley coupling, and transforming to the valley operators by $\psi_{\sigma,\alpha i}(\boldsymbol{r}) = \sum_\rho e^{i\rho \boldsymbol{K}\cdot\boldsymbol{r}} \psi_{\sigma\rho,\alpha i}(\boldsymbol{r})$ the interaction Hamiltonian becomes

$$H_{\text{int}} = \frac{1}{2} \sum_{\sigma,\rho,\alpha,i} \int d\boldsymbol{r} \, \Delta_{\sigma,\alpha i}(\boldsymbol{r}) \psi_{\sigma\rho,\alpha i}^\dagger(\boldsymbol{r}) \psi_{\bar{\sigma}\bar{\rho},\alpha i}^\dagger(\boldsymbol{r}) + \text{h.c.} + \frac{1}{2\lambda} \sum_{\sigma,\alpha,i} \int d\boldsymbol{r} \, |\Delta_{\sigma,\alpha i}(\boldsymbol{r})|^2, \tag{S2}$$

where the local superconducting order parameter is $\Delta_{\sigma,\alpha i}(\boldsymbol{r}) = -\lambda \sum_\rho \langle \psi_{\bar{\sigma}\bar{\rho},\alpha i}(\boldsymbol{r}) \psi_{\sigma\rho,\alpha i}(\boldsymbol{r}) \rangle$. Then by moving to the Nambu space and doing the Bogoliubov transformation we find that the self-consistency equation for the up-spin $\Delta_{\alpha i} := \Delta_{\uparrow,\alpha i}$ becomes

$$\Delta_{\alpha i}(\boldsymbol{r}) = \lambda \sum_{\rho,b} \int \frac{d\boldsymbol{k}}{(2\pi)^2} \tanh\left(\frac{E_{\rho b\boldsymbol{k}}}{2k_B T}\right) u_{\rho b\boldsymbol{k},\alpha i}(\boldsymbol{r}) v_{\rho b\boldsymbol{k},\alpha i}^*(\boldsymbol{r}), \tag{S3}$$

where $u_{\rho b\boldsymbol{k},\alpha i}$ is the (α,i)-component of the spinor $u_{\rho b\boldsymbol{k}}$ and the b sum goes over the positive energy bands. The spinors $u_{\rho b\boldsymbol{k}}$ and $v_{\rho b\boldsymbol{k}}$ are determined by solving the Bogoliubov–de Gennes equation

$$\begin{pmatrix} \mathcal{H}_\rho(\boldsymbol{r}) & \Delta(\boldsymbol{r}) \\ \Delta^*(\boldsymbol{r}) & -\mathcal{H}_\rho(\boldsymbol{r}) \end{pmatrix} \begin{pmatrix} u_{\rho b\boldsymbol{k}}(\boldsymbol{r}) \\ v_{\rho b\boldsymbol{k}}(\boldsymbol{r}) \end{pmatrix} = E_{\rho b\boldsymbol{k}} \begin{pmatrix} u_{\rho b\boldsymbol{k}}(\boldsymbol{r}) \\ v_{\rho b\boldsymbol{k}}(\boldsymbol{r}) \end{pmatrix}, \tag{S4}$$

where Δ is a diagonal 4 by 4 matrix including the components $\Delta_{\alpha i}$. Substituting the Bloch wave expansion

$$\begin{pmatrix} u_{\rho b\boldsymbol{k}}(\boldsymbol{r}) \\ v_{\rho b\boldsymbol{k}}(\boldsymbol{r}) \end{pmatrix} = e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \sum_{\boldsymbol{G}'} e^{i\boldsymbol{G}'\cdot\boldsymbol{r}} \begin{pmatrix} u_{\rho b\boldsymbol{k}}(\boldsymbol{G}') \\ v_{\rho b\boldsymbol{k}}(\boldsymbol{G}') \end{pmatrix} \tag{S5}$$

for the eigenstates into Eq. (S4) and assuming $\Delta(\boldsymbol{r})$ to be periodic in the superlattice, we find the Fourier space Bogoliubov–de Gennes equation [Eq. (5) in the main text] and the Fourier space version of the self-consistency equation [Eq. (4) in the main text].

II. CHARGE DENSITY

The non-coupled system of twisted bilayer graphene is charge neutral at the chemical potential $\mu = 0$. The charge density due to the electrons at that potential is

$$n_0 = \frac{2e}{V} \sum_{b \in B, \boldsymbol{k}} f_0(\epsilon_{0,b\boldsymbol{k}}) = \frac{2e}{V} \sum_{b \in \Omega, \boldsymbol{k}} f_0(\epsilon_{0,b\boldsymbol{k}}) + n_{\text{high}}, \quad \text{with} \quad n_{\text{high}} = \frac{2e}{V} \sum_{b \in B \backslash \Omega, \boldsymbol{k}} f_0(\epsilon_{0,b\boldsymbol{k}}) \tag{S6}$$

where e is the electron charge and the factor of 2 comes from the spin. We formulate the calculation so that the $\boldsymbol{k}$-sum goes over the superlattice Brillouin zone L_{BK}^*/SL^*, B is the set of bands and $\epsilon_{0,b\boldsymbol{k}}$ is the non-interacting dispersion. f_0 is the Fermi-Dirac distribution function at zero temperature. In the second step we introduce a cutoff by dividing the sum over the bands into two terms; to a sum over a set of low-energy bands Ω and to a sum over high-energy bands $B \backslash \Omega$.

In the presence of interlayer coupling, (normal state) dispersion changes to $\epsilon_{b\boldsymbol{k}}$. The number of bands stays constant and if the interactions, temperature and chemical potential are small compared to the energy of the lowest energy band (in absolute value) of $B \backslash \Omega$ in the non-interacting case, the index set B can be chosen so that the bands in $B \backslash \Omega$ that were full (empty) in the non-interacting case, are still full (empty) in the interacting case. The interacting charge density is

$$\tilde{n}(\mu) = \frac{2e}{V} \sum_{b \in B, \boldsymbol{k}} f(\epsilon_{b\boldsymbol{k}} - \mu) = \frac{2e}{V} \sum_{b \in \Omega, \boldsymbol{k}} f(\epsilon_{b\boldsymbol{k}} - \mu) + n_{\text{high}}, \tag{S7}$$

where f is the Fermi-Dirac distribution at temperature T and n_{high} has the the same value as in Eq. (S6). The above has been formulated in the non-linearized theory. To calculate the excess charge relative to the charge neutrality point in the linearized theory, we split the bands between the two valleys and find

$$n(\mu) := \tilde{n}(\mu) - n_0 = \frac{2e}{V} \sum_{\rho, b \in \Omega, \boldsymbol{k}} \left[f(\epsilon_{\rho b \boldsymbol{k}} - \mu) - f_0(\epsilon_{0,\rho b \boldsymbol{k}}) \right] = 2e \sum_{\rho, b \in \Omega} \int \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^2} \left[f(\epsilon_{\rho b \boldsymbol{k}} - \mu) - f_0(\epsilon_{0,\rho b \boldsymbol{k}}) \right], \tag{S8}$$

where n is the excess charge density and Ω is now the set of bands in one valley.

The charge neutrality point μ^* is determined from the equation $n(\mu^*) = 0$. It is shown for different twist angles in Figs. 2(g–i) of the main text, and is always located in the middle between the two DOS peaks.

III. SIMPLIFIED MODEL OF THE SUPERCONDUCTING STATE

The notion of weak or absent electron–phonon mediated superconductivity in pristine graphene is widely known. Here we reconcile this notion with our results claiming that a quite simple BCS-style model could describe the observations of superconductivity in twisted bilayer graphene. These results are not new, but we follow especially the treatments in Refs. 1 and 2 and adopt to the notation of the main paper, along with some estimates.

We start from the generic self-consistency equation for the mean-field order parameter Δ. If Δ is position independent, the Bogoliubov–de Gennes equation can be solved to yield

$$\Delta = 4\lambda \int^{k_c} \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^2} \frac{\Delta}{E_{\boldsymbol{k}}} \tanh\left(\frac{E_{\boldsymbol{k}}}{2 k_B T} \right), \tag{S9}$$

where the prefactor 4 comes from summation over the valley and band indices, where in the band sum we include only the doubly degenerate lowest positive energy band. The cutoff k_c is specified more below. We moreover assume that $E_{\boldsymbol{k}} = \sqrt{\epsilon_{\boldsymbol{k}}^2 + \Delta^2}$. Here and below, without loss of generality we assume $\Delta = |\Delta| \geq 0$. Our idea is to solve the self-consistency equation in three cases: (i) at the Dirac point for a Dirac dispersion $\epsilon_{\boldsymbol{k}}^2 = \hbar^2 v_F^2 k^2$, (ii) for a Dirac dispersion at non-zero doping μ, i.e., $\epsilon_{\boldsymbol{k}}^2 = (\hbar v_F k - \mu)^2$, and (iii) for a flat band with and without doping, $\epsilon_{\boldsymbol{k}} \approx \mu$ for $\boldsymbol{k} \in \Omega_{\text{FB}}$. In each case we have the normal-state solution $\Delta = 0$, which we exclude by dividing both sides in Eq. (S9) by Δ.

Note that Eq. (S9) *does not* represent the full self-consistency equation solved in the main text. Rather, we use it here simply to provide estimates of the behavior of Δ in various limits.

A. Linear dispersion, no doping

Far away from the magic angle, the twisted bilayer behaves as if the two graphene layers would be almost uncoupled. This means that the low-energy dispersion exhibits two separate copies of the graphene Dirac dispersion. Inserting an ultraviolet energy cutoff $\epsilon_c = \hbar v_F k_c$ and performing the integral for the $T = 0$ gap function, the self-consistency equation goes to the form

$$\frac{\pi \hbar^2 v_F^2}{2\lambda} = -\Delta + \sqrt{\Delta^2 + \epsilon_c^2} \tag{S10}$$

or

$$\Delta = \frac{\pi \hbar^2 v_F^2}{4} \frac{\lambda^2 - \lambda_c^2}{\lambda \lambda_c^2}, \tag{S11}$$

where $\lambda_c = \pi \hbar^2 v_F^2 / (2\epsilon_c)$. Since $\Delta \geq 0$, this solution makes sense only if $\lambda > \lambda_c$, and otherwise the only possible solution is the normal state $\Delta = 0$.

In pristine graphene, the critical interaction strength can be written also in terms of the nearest-neighbour hopping term[3] $\gamma_0 \approx 3\,\mathrm{eV}$. Namely, within a nearest-neighbour tight-binding model the Fermi speed of graphene is $v_F = \sqrt{3}\gamma_0 a/(2\hbar)$, where a is the graphene lattice constant. We hence get

$$\lambda_c = \frac{3\pi}{8} \frac{\gamma_0}{\epsilon_c} \gamma_0 a^2. \tag{S12}$$

If the attractive interaction results from electron–phonon coupling, a typical cutoff energy could be of the order of the Debye energy[4] 200 meV. In this case $\lambda_c \approx 50\,\mathrm{eV}a^2$, 5 to 50 times larger than the values of λ used in our work. 100 to 200 meV is also the range of the maximum cutoff energy that we have used in our numerical results when including the contribution from higher bands. Even if the cutoff ϵ_c would be of the order of γ_0, the resulting λ_c would be one order of magnitude larger than the smallest λ used in our results.

B. Linear dispersion, with doping

Let us try to reconcile the observations of superconductivity in Li or Ca doped graphene with the above idea. These cases are more accurately described by[5] within the Eliashberg theory. Here we just show in which sense doping fits into the above picture. Assuming $\epsilon_{\bm{k}} = \pm \hbar v_F k - \mu$ and $\Delta < \epsilon_c$, and cutting the integral at $\epsilon_{\bm{k}} = \epsilon_c$ the self-consistency equation at $T = 0$ becomes[1]

$$\frac{\pi \hbar^2 v_F^2}{2\lambda} = \sqrt{\Delta^2 + \epsilon_c^2} - \sqrt{\Delta^2 + \mu^2} + |\mu| \ln \frac{|\mu| + \sqrt{\Delta^2 + \mu^2}}{\Delta}. \tag{S13}$$

Let us assume that $\Delta \ll |\mu|, \epsilon_c$ so that we can expand the right hand side in Δ. In this case we find an analytic solution for Δ,

$$\Delta = 2|\mu| \exp\left[-\frac{\epsilon_c}{|\mu|} \left(\frac{\lambda_c}{\lambda} - 1 \right) - 1 \right]. \tag{S14}$$

Let us assume a cutoff energy $\epsilon_c = 200\,\mathrm{meV}$ and a coupling strength $\lambda = \lambda_c/22$ (corresponding to about $5\,\mathrm{eV}a^2$ with the above estimates). In this case, with $\mu = 0.7\,\mathrm{eV}$ we would get $\Delta = 1.3\,\mathrm{meV}$. This corresponds to a critical temperature of $9\,\mathrm{K}$, in the same range as the one that was measured in Li or Ca doped graphene.[6–9]

C. Flat band estimate

Let us now make similar estimates for the flat-band case of the moiré superlattice. In this case, we assume that Δ is *larger* than the bandwidth of the lowest-energy band. Within that band, we can hence approximate $E_{\bm{k}} \approx \Delta$ in Eq. (S9) and at $T = 0$ the integral is over a constant function. As a result, we get

$$\Delta_{\mathrm{FB}} = \frac{\lambda}{\pi^2} \Omega_{\mathrm{FB}} = \frac{8\lambda}{\sqrt{3}(3m^2 + 3m + 1)a^2}, \tag{S15}$$

where $\Omega_{\mathrm{FB}} = 8\pi^2/[\sqrt{3}(3m^2 + 3m + 1)a^2]$ is the area of the first Brillouin zone of the moiré superlattice. Within the model adapted in the main text, the magic angle is around $m \approx 34$, in which case we would get $\Delta_{\mathrm{FB}} = 1.3 \times 10^{-3}\lambda/a^2$. In Fig. 3b of the main text, the solid line has a slope of $1.0 \times 10^{-3}\lambda/a^2$, *i.e.*, very close to this simple estimate.

The temperature dependent Δ in the flat-band case is obtained by solving

$$\Delta = \Delta_{\mathrm{FB}} \tanh\left(\frac{\Delta}{2k_B T}\right). \tag{S16}$$

At the critical temperature, $\Delta \to 0$, and we can hence expand the right hand side to the linear order in $\Delta/(2k_B T_c)$. This directly yields $k_B T_c = \Delta_{\mathrm{FB}}/2$.

In the case of a non-zero potential μ, we can use $E_{\boldsymbol{k}} \approx \sqrt{\mu^2 + \Delta^2}$ in the self-consistency equation. It then becomes (for $\Delta > 0$)

$$\Delta = \Delta_{\mathrm{FB}} \frac{\Delta}{\sqrt{\mu^2 + \Delta^2}}, \text{ or } \Delta = \sqrt{\Delta_{\mathrm{FB}}^2 - \mu^2}. \tag{S17}$$

In this case superconductivity is hence suppressed when the absolute value of the chemical potential is larger than Δ_{FB}.

IV. SIMPLIFIED MODEL FOR COULOMB PSEUDOPOTENTIAL

Coulomb interaction differs from the electron–phonon interaction due to the fact that photons are almost instantaneous, whereas for phonon-mediated interaction we have to take the retardation into account. Usually in BCS theory, and also in our model, we approximate the retardation by imposing an energy cutoff at the maximum phonon frequency ω_D in the self-consistency equation. For Coulomb interaction there is no physical cutoff, and consequently, we cannot operate in purely low-energy regime. The high energy states do contribute logarithmically to Δ at low energies.

The proper way to formulate the low-energy theory with a cutoff which also applies to the Coulomb interaction, is to define a modified pseudopotential u^* which replaces the bare interaction in the self-consistency equation and takes the high-energy parts into account. If $\Delta(\boldsymbol{r})$ is position-dependent, the pseudopotential will be a matrix of two position coordinates $u^*(\boldsymbol{r}, \boldsymbol{r}')$. If Δ is constant in space, the pseudopotential is a scalar.

We want to consider the effect of the Hubbard interaction, described in the continuum limit by the Hamiltonian

$$H_{\mathrm{Hubbard}} = \frac{u}{2} \sum_{\sigma,\alpha,i} \int d\boldsymbol{r}\, \psi_{\sigma,\alpha i}^\dagger(\boldsymbol{r}) \psi_{\bar{\sigma},\alpha i}^\dagger(\boldsymbol{r}) \psi_{\bar{\sigma},\alpha i}(\boldsymbol{r}) \psi_{\sigma,\alpha i}(\boldsymbol{r}), \tag{S18}$$

where $u = Ua^2$ and U is the Hubbard parameter describing the on-site interaction in the tight-binding model. We assume that $U > 0$ so that the interaction is repulsive. The inclusion of such an interaction has multiple effects in a inhomogeneous system, but here we only consider the effect on the order parameter through the modification of the self-consistency equation.

As we are now not doing a low-energy calculation, separation into valleys is not useful and we cannot do the continuum approximation in which we assume that the graphene lattice L is duplicated infinitely many times in the superlattice. Therefore, in the following the sums and integrals are done over the sets $\boldsymbol{G} \in SL^*/L^*$ and $\boldsymbol{k} \in \mathbb{R}^2/SL^*$. The two graphene valleys are then separated from each other by a large, but finite $\boldsymbol{G}$-vector. The valley sum is thus included in the sum over $\boldsymbol{G}$ and there is no valley index ρ.

For simplicity, we assume $\Delta_{\alpha i}(\boldsymbol{G}) = \Delta \delta_{\boldsymbol{G},0}$ so that Δ has no position dependence and is the same on both layers and sublattices. With this simplification, we can diagonalize the Hamiltonian $\underline{\mathcal{H}}_{\boldsymbol{k}}$ and the order parameter simultaneously in the BdG equation (Eq. (5) of the main paper), which we write as

$$\begin{pmatrix} \underline{\mathcal{H}}_{\boldsymbol{k}} & \Delta\underline{1} \\ \Delta^*\underline{1} & -\underline{\mathcal{H}}_{\boldsymbol{k}} \end{pmatrix} \begin{pmatrix} \underline{u}_{\boldsymbol{k}b} \\ \underline{v}_{\boldsymbol{k}b} \end{pmatrix} = E_{\boldsymbol{k}b} \begin{pmatrix} \underline{u}_{\boldsymbol{k}b} \\ \underline{v}_{\boldsymbol{k}b} \end{pmatrix}, \tag{S19}$$

where the underlined quantities are matrices/vectors with indices $\boldsymbol{G}, \alpha$, and i. Let $\underline{\mathcal{G}}_{\boldsymbol{k}}$ be a unitary transformation which diagonalizes the normal state Hamiltonian $\underline{\mathcal{H}}_{\boldsymbol{k}}$. Then the above equation becomes

$$\begin{pmatrix} \underline{\epsilon}_{\boldsymbol{k}} & \Delta\underline{1} \\ \Delta^*\underline{1} & -\underline{\epsilon}_{\boldsymbol{k}} \end{pmatrix} \begin{pmatrix} \underline{u}'_{\boldsymbol{k}b} \\ \underline{v}'_{\boldsymbol{k}b} \end{pmatrix} = E_{\boldsymbol{k}b} \begin{pmatrix} \underline{u}'_{\boldsymbol{k}b} \\ \underline{v}'_{\boldsymbol{k}b} \end{pmatrix}, \tag{S20}$$

where $\underline{u}'_{\boldsymbol{k}b} = \underline{\mathcal{G}}_{\boldsymbol{k}}\underline{u}_{\boldsymbol{k}b}$, $\underline{v}'_{\boldsymbol{k}b} = \underline{\mathcal{G}}_{\boldsymbol{k}}\underline{v}_{\boldsymbol{k}b}$ and $\underline{\epsilon}_{\boldsymbol{k}} = \underline{\mathcal{G}}_{\boldsymbol{k}}\underline{\mathcal{H}}_{\boldsymbol{k}}\underline{\mathcal{G}}_{\boldsymbol{k}}^{\dagger}$. We now label the normal state eigenstates with band index b. With constant Δ, the positive-energy BdG eigenstates are in simple correspondence with the eigenstates (both positive and negative energy) of the normal state, and can also be labeled with the same indices. Concentrating to a single Nambu-block of the BdG-equation,

$$\begin{pmatrix} \epsilon_{\boldsymbol{k}b} & \Delta \\ \Delta^{*} & -\epsilon_{\boldsymbol{k}b} \end{pmatrix} \begin{pmatrix} u'_{\boldsymbol{k}b} \\ v'_{\boldsymbol{k}b} \end{pmatrix} = E_{\boldsymbol{k}b} \begin{pmatrix} u'_{\boldsymbol{k}b} \\ v'_{\boldsymbol{k}b} \end{pmatrix}, \tag{S21}$$

we find that the eigenenergies and eigenstates assume the usual BCS form

$$E_{\boldsymbol{k}b} = \sqrt{\epsilon_{\boldsymbol{k}b}^{2} + |\Delta|^{2}}, \tag{S22}$$

$$u_{\boldsymbol{k}b} = \frac{e^{i\phi}}{\sqrt{2}} \left(1 + \frac{\epsilon_{\boldsymbol{k}b}}{E_{\boldsymbol{k}b}} \right)^{1/2}, \tag{S23}$$

$$v_{\boldsymbol{k}b} = \frac{1}{\sqrt{2}} \left(1 - \frac{\epsilon_{\boldsymbol{k}b}}{E_{\boldsymbol{k}b}} \right)^{1/2}, \tag{S24}$$

where $\phi = \arg(\Delta)$.

The self-consistency equation [Eq. (4) in the main text with $\boldsymbol{G} = 0$ and generalized to include energy-dependent interactions] can be written in the above matrix notation as

$$\Delta_{\boldsymbol{k}b,\alpha i} = \sum_{b'} \int \frac{\mathrm{d}\boldsymbol{k}'}{(2\pi)^{2}} V_{\boldsymbol{k}\boldsymbol{k}'}^{bb'} \left(\underline{u}_{\boldsymbol{k}'b'}^{\dagger} \underline{\Pi}_{\alpha i} \underline{v}_{\boldsymbol{k}'b'} \right)^{*} \tanh\left(\frac{E_{\boldsymbol{k}'b'}}{2k_{B}T} \right), \tag{S25}$$

where $\underline{\Pi}_{\alpha i}$ is the projection operator to the sublattice α and layer i. We assume that the interaction has the simplified BCS form

$$V_{\boldsymbol{k}\boldsymbol{k}'}^{bb'} = \lambda\theta(|\epsilon_{\boldsymbol{k}b}| - \omega_{D})\theta(|\epsilon_{\boldsymbol{k}'b'}| - \omega_{D}) - u, \tag{S26}$$

with electron–phonon cutoff at Debye energy ω_{D}.

The sum of complete set of projection operators is an identity: $\sum_{\alpha,i} \underline{\Pi}_{\alpha i} = \underline{1}$. To get rid of the projection operator, we take the average over α and i. As $\Delta_{\boldsymbol{k}b,\alpha i} = \Delta_{\boldsymbol{k}b}$, we get

$$\Delta_{\boldsymbol{k}b} = \frac{1}{4} \sum_{b'} \int \frac{\mathrm{d}\boldsymbol{k}'}{(2\pi)^{2}} V_{\boldsymbol{k}\boldsymbol{k}'}^{bb'} \left(\underline{u}_{\boldsymbol{k}'b'}^{\dagger} \underline{v}_{\boldsymbol{k}'b'} \right)^{*} \tanh\left(\frac{E_{\boldsymbol{k}'b'}}{2k_{B}T} \right) \tag{S27}$$

$$= \frac{1}{4} \sum_{b'} \int \frac{\mathrm{d}\boldsymbol{k}'}{(2\pi)^{2}} V_{\boldsymbol{k}\boldsymbol{k}'}^{bb'} u'_{\boldsymbol{k}'b'} (v'_{\boldsymbol{k}'b'})^{*} \tanh\left(\frac{E_{\boldsymbol{k}'b'}}{2k_{B}T} \right). \tag{S28}$$

In the second line, we did a basis transformation with the matrix $\underline{\mathcal{G}}_{-\boldsymbol{P}}^{\dagger}$.

We now divide $\Delta_{\boldsymbol{k}b} = \Delta_{\boldsymbol{k}b}^{\lambda} + \Delta^{u}$ into two parts, with $\Delta_{\boldsymbol{k}b}^{\lambda}$ corresponding to the λ part of the interaction in the RHS of Eq. (S28) and Δ^{u} corresponding to the u part of the interaction.[10] The difference between the two terms is in the energy dependence. $\Delta_{\boldsymbol{k}b}^{\lambda}$ vanishes above the cutoff, but Δ^{u} has no energy dependence and persists at high energies. With this division, the self-consistency equation splits into two coupled equations,

$$\Delta_{\boldsymbol{k}b}^{\lambda} = \frac{\lambda}{4} \sum_{b'} \int_{|\epsilon_{\boldsymbol{k}'b'}|<\omega_{D}} \frac{\mathrm{d}\boldsymbol{k}'}{(2\pi)^{2}} u'_{\boldsymbol{k}'b'} (v'_{\boldsymbol{k}'b'})^{*} \tanh\left(\frac{E_{\boldsymbol{k}'b'}}{2k_{B}T} \right) \times \theta(|\epsilon_{\boldsymbol{k}b}| - \omega_{D}), \tag{S29}$$

$$\Delta^{u} = -\frac{u}{4} \sum_{b'} \int_{|\epsilon_{\boldsymbol{k}'b'}|<\omega_{D}} \frac{\mathrm{d}\boldsymbol{k}'}{(2\pi)^{2}} u'_{\boldsymbol{k}'b'} (v'_{\boldsymbol{k}'b'})^{*} \tanh\left(\frac{E_{\boldsymbol{k}'b'}}{2k_{B}T} \right) - \frac{u}{4} \sum_{b'} \int_{|\epsilon_{\boldsymbol{k}'b'}|>\omega_{D}} \frac{\mathrm{d}\boldsymbol{k}'}{(2\pi)^{2}} u'_{\boldsymbol{k}'b'} (v'_{\boldsymbol{k}'b'})^{*} \tanh\left(\frac{E_{\boldsymbol{k}'b'}}{2k_{B}T} \right). \tag{S30}$$

Above, we also split the sums and integrals over the eigenstates to low and high energy parts with ω_{D} as the cutoff. Assuming $\omega_{D} \gg T, \Delta^{u}$, we can approximate that for high energy states

$$u'_{\boldsymbol{k}b} (v'_{\boldsymbol{k}b})^{*} \tanh\left(\frac{E_{\boldsymbol{k}b}}{2k_{B}T} \right) \approx \frac{\Delta^{u}}{2|\epsilon_{\boldsymbol{k}b}|}. \tag{S31}$$
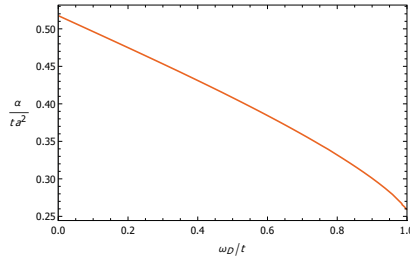
FIG. S1. Dependence of the graphene pseudopotential renormalization constant α on the electron-phonon cutoff ω_D. Pseudopotential renormalization constant of TBG can be approximated with that of graphene if $|t_\perp| \ll \omega_D$.

Inserting this into Eq. (S30), we can (partially) solve for Δ^u to obtain an equation which only refers to the low energy states,

$$\Delta^u = -\frac{u^*}{4} \sum_{b'} \int_{|\epsilon_{\mathbf{k}'b'}|<\omega_D} \frac{d\mathbf{k}'}{(2\pi)^2} u'_{\mathbf{k}'b'} (v'_{\mathbf{k}'b'})^* \tanh\left(\frac{E_{\mathbf{k}'b'}}{2k_B T}\right). \tag{S32}$$

The high energy states renormalize the interaction constant, which is replaced by the Coulomb pseudopotential

$$u^* = \frac{u}{1+u\alpha}, \quad \text{where} \quad \alpha = \frac{1}{4} \sum_b \int_{|\epsilon_{\mathbf{k}'b'}|>\omega_D} \frac{d\mathbf{k}}{(2\pi)^2} \frac{1}{2|\epsilon_{\mathbf{k}b}|}. \tag{S33}$$

The equation for the full order parameter, including both interactions, is now

$$\Delta = \frac{\lambda_{\text{eff}}}{4} \sum_{b'} \int_{|\epsilon_{\mathbf{k}'b'}|<\omega_D} \frac{d\mathbf{k}'}{(2\pi)^2} u'_{\mathbf{k}'b'} (v'_{\mathbf{k}'b'})^* \tanh\left(\frac{E_{\mathbf{k}'b'}}{2k_B T}\right) \quad \text{with} \quad \lambda_{\text{eff}} = \lambda - u^*. \tag{S34}$$

If $\Delta(\mathbf{r})$ is position dependent, the derivation becomes more complicated, and in the end, the pseudopotential becomes a matrix $u^*(\mathbf{G}, \mathbf{G}')$ instead of a scalar like above.

The pseudopotential renormalization parameter α now depends on structure of the high energy bands. It is not very sensitive to the parameters of the system and for this calculation we assume $t_\perp = 0$ so that the two graphene layers are completely independent of each other. The sums and integrals then transform as

$$\alpha = \frac{1}{4} \sum_b \int_{\substack{\mathbb{R}^2/SL^* \\ |\epsilon_{\mathbf{k}b}|>\omega_D}} \frac{d\mathbf{k}}{(2\pi)^2} \frac{1}{2|\epsilon_{\mathbf{k}b}|} \approx \frac{1}{4} \sum_i \sum_{b\in\pm 1} \int_{\substack{\mathbb{R}^2/L^* \\ |\epsilon_{\mathbf{k}b}|>\omega_D}} \frac{d\mathbf{k}}{(2\pi)^2} \frac{1}{2|\epsilon_{\mathbf{k}b}^0|} \tag{S35}$$

where $\epsilon_{\mathbf{k}b}^0$ are the graphene eigenenergies calculated from the tight binding model with only nearest neighbour hoppings.

If approximated as above, α corresponds to the pseudopotential constant for graphene. We show the dependence on the cutoff ω_D in Fig. S1. With parameters $\omega_D = 200\,\text{meV}$ and nearest neighbour hopping $t = 3\,\text{eV}$, we find that $\alpha \approx 0.2\,\text{eV}^{-1}a^{-2}$, which holds as long as $\mu \ll \omega_D$. The maximum value for the pseudopotential is thus $u^*_{\max} = 1/\alpha \approx 5\,\text{eV}a^2$, which is obtained in the limit $U \to \infty$. For $U = 5\,\text{eV}$, the effective interaction strength is reduced to half of the bare interaction strength, $u^* \approx 0.5u = 2.5\,\text{eV}a^2$.

[1] N. B. Kopnin and E. B. Sonin, Phys. Rev. Lett. **100**, 246808 (2008).
[2] N. B. Kopnin and T. T. Heikkilä, "Carbon-based Superconductors: Towards High-Tc Superconductivity," (Taylor & Francis, 2014) Chap. 9.
[3] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, Rev. Mod. Phys. **81**, 109 (2009).
[4] D. K. Efetov and P. Kim, Phys. Rev. Lett. **105**, 256805 (2010).
[5] G. Profeta, M. Calandra, and F. Mauri, Nat. Phys. **8**, 131 (2012).
[6] B. M. Ludbrook, G. Levy, P. Nigge, M. Zonno, M. Schneider, D. J. Dvorak, C. N. Veenstra, S. Zhdanovich, D. Wong, P. Dosanjh, et al., PNAS **112**, 11795 (2015).

[7] A. P. Tiwari, S. Shin, E. Hwang, S.-G. Jung, T. Park, and H. Lee, J. Phys. Cond. Matt. **29**, 445701 (2017).

[8] J. Chapman, Y. Su, C. A. Howard, D. Kundys, A. N. Grigorenko, F. Guinea, A. K. Geim, I. V. Grigorieva, and R. R. Nair, Sci. Rep. **6**, 23254 (2016).

[9] S. Ichinokura, K. Sugawara, A. Takayama, T. Takahashi, and S. Hasegawa, ACS Nano **10**, 2761 (2016).

[10] P. Morel and P. Anderson, Phys. Rev. **125**, 1263 (1962).

III

SUPERFLUID WEIGHT AND BEREZINSKII-KOSTERLITZ-THOULESS TRANSITION TEMPERATURE OF TWISTED BILAYER GRAPHENE

by

Aleksi Julku, Teemu J. Peltonen, Long Liang, Tero T. Heikkilä, and Päivi Törmä 2020

Superfluid weight and Berezinskii-Kosterlitz-Thouless transition temperature of twisted bilayer graphene

A. Julku,[1] T. J. Peltonen,[2] L. Liang,[1,3] T. T. Heikkilä,[2,*] and P. Törmä[1,†]

[1]*Department of Applied Physics, Aalto University, P.O. Box 15100, 00076 Aalto, Finland*
[2]*Department of Physics and Nanoscience Center, University of Jyväskylä, P.O. Box 35 (YFL), FI-40014, Finland*
[3]*Computational Physics Laboratory, Physics Unit, Faculty of Engineering and Natural Sciences, Tampere University,*
P.O. Box 692, FI-33014 Tampere, Finland

We study superconductivity of twisted bilayer graphene with local and nonlocal attractive interactions. We obtain the superfluid weight and Berezinskii-Kosterlitz-Thouless (BKT) transition temperature for microscopic tight-binding and low-energy continuum models. We predict qualitative differences between local and nonlocal interaction schemes which could be distinguished experimentally. In the flat-band limit where the pair potential exceeds the band width we show that the superfluid weight and BKT temperature are determined by multiband processes and quantum geometry of the band.

Recent experimental discoveries of superconductivity in bilayer graphene twisted close to a "magic angle" θ^* [1–3] call for a reconsideration of traditional theories of superconductivity [4,5], in particular because the superconductivity occurs in a regime where the noninteracting electronic states form an asymptotically flat (dispersionless) band [6–17]. As the system is two-dimensional, the transition to superconductivity is bound to occur at the Berezinskii-Kosterlitz-Thouless (BKT) temperature T_{BKT} [18–20] which can be determined from $k_B T_{\mathrm{BKT}} = \frac{\pi}{8}\sqrt{\det[D^s(T_{\mathrm{BKT}})]}$ [21,22]. Here D^s is the superfluid weight that yields the size of the supercurrent for a given phase gradient of the order parameter. In conventional theory of superconductivity [23], D^s is proportional to the group velocity of electronic bands around the Fermi level. Thus $D^s = 0$ for a flat band, and superconductivity in twisted bilayer graphene (TBG) appears puzzling. One might argue it to be due to the bands not being perfectly flat; however, we show here that a more likely explanation goes beyond the conventional theory. Here we calculate T_{BKT} for TBG as a function of the superconducting order parameter and filling. We use two models of TBG including both the flat and a number of dispersive bands and show that superconductivity in the flat-band regime has essentially a quantum geometric origin.

Recently, it was found that D^s has, in addition to the conventional contribution proportional to group velocity, a geometric contribution arising from multiband processes [24–28]. In a flat-band limit the geometric contribution dominates and is bounded from below by the band Berry curvature [27] and Chern number [24] [see also the discussion in the Supplemental Material (SM) [29] and Ref. [30] therein]. Here we show

that the geometric contribution dominates D^s and T_{BKT} in the flat-band regime of TBG. Importantly, we show that including only the few flat bands is *not* sufficient but one needs also a number of dispersive bands to correctly predict the geometric contribution. Therefore, approximate models of TBG such as those with only flat bands, as used for deriving upper [31] and lower [32] bounds of the superfluid weight and in many other works [33–45], may not be suited for quantitative predictions of TBG superconductivity. Moreover, we predict that, in the flat-band regime, local (*s*-wave) and nonlocal interactions yield distinct behavior, namely, an *anisotropic superfluid weight* in the latter case. We propose a four-terminal radio frequency spectroscopy experiment that can detect the possible anisotropy and thus distinguish between the two pairing mechanisms.

An outstanding problem in describing the TBG physics theoretically [33–44,46–62] is the fact that the unit cell of the moiré superlattice with twist angles close to θ^* contains a large amount of carbon atoms [Fig. 1(a)], and therefore TBG theory should take a stand on how to describe the interlayer couplings within this unit cell. Here we use and compare with each other two of the previously used approximation procedures: (1) the *renormalized moiré* (RM) approach [56,63], where we scale some coupling energies by a suitable scaling factor to find the flat bands at a higher θ, resulting in a smaller unit cell, and (2) the *Dirac point approximation* (DP) [11,51,64], where we make a low-energy approximation near the graphene Dirac points by linearizing the intralayer Hamiltonians and using a cutoff in the superlattice Fourier space. Both of these approaches go beyond those often used in TBG literature, either based on a single-parameter coupling model, or a vastly reduced four-band model [31,32,65–67].

Theoretical models. In the renormalized moiré lattice method (RM), we deploy the Fermi-Hubbard Hamiltonian as [29] $H = H_{\mathrm{kin}} - \mu N + H_{\mathrm{int}}$, where $H_{\mathrm{kin}} = \sum_{i\alpha j\beta\sigma} t_{i\alpha j\beta} c^{\dagger}_{i\alpha\sigma} c_{j\beta\sigma}$ is the kinetic term, N is the total particle number operator, and H_{int} is the effective

_____
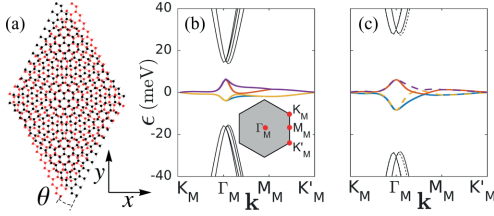
*tero.t.heikkila@jyu.fi
†paivi.torma@aalto.fi

FIG. 1. (a) The moiré superlattice of TBG depicted with a twist angle θ and the choice of the x and y axes. (b), (c) Single-particle energy band structures of the RM and DP methods, respectively, plotted within the moiré Brillouin zone along the path connecting the high symmetry points shown in the inset of (b). In the DP model (c) the bands coming from the valley **K** (**K'**) are drawn as solid (dashed) lines.

attractive interaction described below. Here $c_{i\alpha\sigma}$ annihilates a fermion in the αth lattice site of the ith moiré superlattice unit cell with spin $\sigma \in \{\uparrow, \downarrow\}$, μ is the chemical potential, and the hopping $t_{i\alpha j\beta}$ includes both the intra- and interlayer terms.

Since the type of the interaction is not currently known, we consider two different singlet pairing potentials, namely, the local pairing $H_{int} = J \sum_{i\alpha} c_{i\alpha\uparrow}^\dagger c_{i\alpha\downarrow}^\dagger c_{i\alpha\downarrow} c_{i\alpha\uparrow} \equiv H_{loc}$ and the nearest-neighbor (NN) pairing $H_{int} = \frac{J}{2} \sum_{\langle i\alpha j\beta \rangle} h_{i\alpha j\beta}^\dagger h_{i\alpha j\beta} \equiv H_{RVB}$, where $h_{i\alpha j\beta} = (c_{i\alpha\uparrow} c_{j\beta\downarrow} - c_{i\alpha\downarrow} c_{j\beta\uparrow})$ and $J < 0$ is the interaction strength. The local interaction has been used to study s-wave superconductivity, mediated by electron-phonon interaction, both in graphene [68–71] and in TBG [51,52]. The nonlocal, called resonance valence bond (RVB) interaction [72,73], has also been used both in the case of monolayer graphene [74–78] (see also SM [29] and Refs. [79–81] therein) and TBG [56]. We keep only the pairing channels by applying mean-field theory to approximate $H_{loc} \approx \Delta_{i\alpha} c_{i\alpha\uparrow}^\dagger c_{i\alpha\downarrow}^\dagger + $ H.c. and $H_{RVB} \approx \Delta_{i\alpha j\beta} h_{i\alpha j\beta}^\dagger + $ H.c., where $\Delta_{i\alpha} = J\langle c_{i\alpha\downarrow} c_{i\alpha\uparrow} \rangle$ and $\Delta_{i\alpha j\beta} = \frac{J}{2}\langle h_{i\alpha\beta} \rangle$ are the superfluid order parameters, respectively.

To reduce the number of lattice sites M within a moiré unit cell (around 12 000 for twist angle $\theta \sim 1°$), we apply a rescaling trick [56,63] under which the Fermi velocity of a monolayer graphene and the moiré periodicity remain invariant but θ becomes larger and thus reduces M. In our computations we use the rescaling such that $M = 676$ and the rescaled angle is $\theta' = 4.41°$ [29] which reproduces the four narrow bands of the bandwidth of 10 meV found experimentally with $\theta \sim 1°$ [see Fig. 1(b)].

In the Dirac point continuum method (DP) we employ the low-energy [11,29,51] Dirac point approximation for the two graphene layers as $H_{kin}^1 = \sum_{\sigma\rho\mathbf{kG}} c_{\sigma\rho,1}^\dagger(\mathbf{k}+\mathbf{G})\hbar v_F \boldsymbol{\sigma}^\rho \cdot (\mathbf{k}+\mathbf{G}) c_{\sigma\rho,1}(\mathbf{k}+\mathbf{G})$ and $H_{kin}^2 = \sum_{\sigma\rho\mathbf{kG}} c_{\sigma\rho,2}^\dagger(\mathbf{k}+\mathbf{G})\hbar v_F \boldsymbol{\sigma}_\theta^\rho \cdot (\mathbf{k}+\mathbf{G}) c_{\sigma\rho,2}(\mathbf{k}+\mathbf{G})$ and couple the layers by $H_{kin}^\perp = \sum_{\sigma\rho\mathbf{kGG'}} c_{\sigma\rho,1}^\dagger(\mathbf{k}+\mathbf{G}+\rho\frac{\Delta\mathbf{K}}{2}) t_\perp^\rho(\mathbf{G}-\mathbf{G}') c_{\sigma\rho,2}(\mathbf{k}+\mathbf{G}'-\rho\frac{\Delta\mathbf{K}}{2}) + $ H.c. Here $c_{\sigma\rho,l}(\mathbf{k}) = [c_{\sigma\rho,lA}(\mathbf{k}), c_{\sigma\rho,lB}(\mathbf{k})]^\top$ in the sublattice space, where $c_{\sigma\rho,ls}(\mathbf{k})$ is the annihilation operator for spin σ, valley $\rho \in \{+, -\}$, layer l, sublattice s, and wave vector $\mathbf{k}$, $\boldsymbol{\sigma}^\rho = (\rho\sigma_x, \sigma_y)$ is a vector of Pauli matrices in the

sublattice space, $\boldsymbol{\sigma}_\theta^\rho = R(\theta)\boldsymbol{\sigma}^\rho$ is the θ-rotated version of it, $t_\perp^\rho(\mathbf{G})$ is the Fourier component [82] of a Slater-Koster [83] parametrized interlayer potential (times an exponential factor), $\Delta\mathbf{K} = R(\theta)\mathbf{K} - \mathbf{K}$ is the difference vector from the graphene **K** point to its rotated counterpart, and v_F is the graphene Fermi velocity. The **k** sum is over the the moiré Brillouin zone and the **G**, **G'** sums are over the (truncated) reciprocal superlattice.

We then write the total Hamiltonian as $H = H_{kin}^1 + H_{kin}^2 + H_{kin}^\perp - \mu N + H_{int}$, where N is the total particle number operator. To describe the superconducting state with a local pairing interaction λ we use $H_{int} = \lambda \sum_{ls} \int d\mathbf{r}\, \mathbf{r} \psi_{\uparrow\rho,ls}^\dagger(\mathbf{r}) \psi_{\downarrow\bar\rho,ls}^\dagger(\mathbf{r}) \psi_{\downarrow\bar\rho,ls}(\mathbf{r}) \psi_{\uparrow\rho,ls}(\mathbf{r})$, which is treated in the mean-field level [29]. Here $\bar\rho$ is the opposite valley of ρ and $\psi_{\sigma\rho,ls}(\mathbf{r})$ is the continuum electron field operator.

Order parameters, superfluid weight, and pairing symmetry. In experiments [1–3] superconducting (SC) and correlated insulating states have been observed with the magic angle twist such that insulating states emerge for the flat-band fillings $\nu \in \{0, \pm1, \pm2, \pm3\}$ and SC states surround the insulating states close to $\nu \in \{0, \pm1, \pm2\}$ with the SC phase near $\nu = -2$ being observed at temperature as high as $\sim$3 K [2,3]. Here ν is the electron density per moiré unit cell so that the charge neutrality point (CNP) corresponds to $\nu = 0$ and narrow bands are empty (full) when $\nu = -4$ ($\nu = 4$).

To determine the superfluid weight D^s, we first solve order parameters from the BCS gap equations [29]. In Figs. 2(a) and 2(b) we show the spatial profiles for the local and RVB interactions computed with the RM method at $\nu \approx -2$. Here J is chosen such that the maximum value of the order parameter is max $|\Delta| \approx 3.4$ meV. From Figs. 2(c) and 2(d) we see max $|\Delta|$ depending almost linearly on the interaction constant, which is typical for generic flat-band systems [4,5,24,25,51]. From the obtained order parameter values one can compute D^s. For easier comparison between the RM and DP models, below we use max $|\Delta|$ as a "parameter." For clarity, note that the used interaction strengths are much larger than the actual energy gap max $|\Delta|$, and they are of the order of the energy separation to the dispersive bands [see Figs. 2(c) and 2(d)].

To obtain D^s we use linear response theory. In the mean-field level [84,85] the zero-frequency, long-wavelength limit of the current-current response function $K_{\mu\nu}(\mathbf{q}, \omega)$ is D^s, i.e. ($\mu, \nu \in \{x, y\}$),

$$D_{\mu\nu}^s = \lim_{\mathbf{q}\to 0}[\lim_{\omega\to 0} K_{\mu\nu}(\mathbf{q}, \omega)]. \tag{1}$$

In Ref. [27] this was computed for a generic multiorbital lattice geometry with the local interaction. The details on how $D_{\mu\nu}^s$ is obtained for our different models are discussed in the SM [29].

In Fig. 2(e) we present D^s as a function of max $|\Delta|$ at $\nu = -2$ for both local (obtained with RM and DP) and RVB interactions (only RM). Figures 2(a), 2(b), and 2(e) depict a striking distinction between the local and RVB pairing schemes related to the pairing symmetry and the resulting form of D^s. The local pairing, yielding an s-wave symmetry, conserves the underlying C_3 symmetry of the TBG lattice [Fig. 2(a)] and D^s is isotropic [29], i.e., $D_{xx}^s = D_{yy}^s$ and $D_{xy}^s = D_{yx}^s = 0$. By contrast, the RVB pairing with strong enough
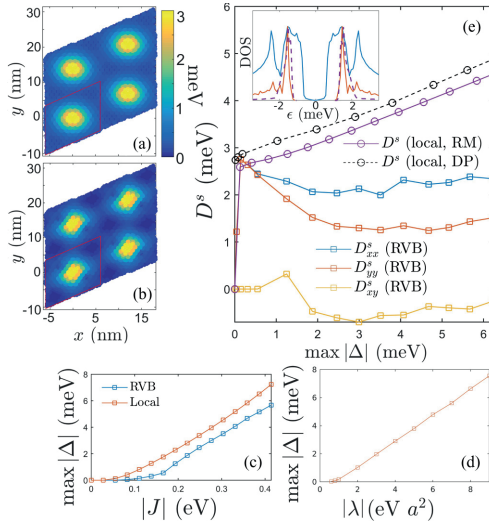
FIG. 2. (a), (b) Spatial profiles of the order parameter for local and RVB interaction schemes, respectively, computed with RM. The DP model for the local interaction yields a similar spatial distribution [51]. In the case of RVB, $\Delta_{i\alpha j\beta}$ are plotted at $\mathbf{r}_{i\alpha}$. Red parallelograms represent the moiré unit cell. The maximum order parameter in both cases is max $|\Delta| \approx 3.4$ meV. (c), (d) max $|\Delta|$ as a function of the interaction strength at $\nu \approx -2$ for the RM and DP methods, respectively. Here a is the graphene lattice constant. (e) Spatial components of D^s as a function of max $|\Delta|$ at $\nu \approx -2$ for local and RVB pairing. For local interaction $D_{xx}^s = D_{yy}^s = D^s$. Inset of (e) shows the total density of states (DOS) for RVB (blue curve) and local interaction (red) at max $|\Delta| \approx 3.4$ meV computed with RM. The dashed curve is the DOS for local interaction obtained with DP at max $|\Delta| \approx 3.5$ meV. From the DOS we see the nematic phase being gapless, while the s-wave state is gapped. The RM results are evaluated at $T \approx 0.1$ K, whereas the DP results at $T = 0$.

interaction breaks the C_3-rotational symmetry and yields a nematic pairing pattern in real space [Fig. 2(b)] which leads to an anisotropic response, i.e., $D_{xx}^s \neq D_{yy}^s$ and $D_{xy}^s = D_{yx}^s \neq 0$. The s wave is gapped, whereas the nematic phase has nodal points in the moiré Brillouin zone [see also the inset of Fig. 2(e)]. The anisotropic D^s results in an anisotropic kinetic inductance of TBG, and it can in principle be accessed via radio frequency impedance spectroscopy [86] in a Hall-like four-probe setup.

As seen from Fig. 2(e), D^s for the RVB interaction in the weak-coupling regime is still isotropic. This phase has the mixed $(d + id) + (p + ip)$ symmetry with a full energy gap, whereas the nematic phase of the flat-band regime is identified as a mixture of s-, p-, and d-wave components [56], with the d wave being the dominant symmetry. Our results for the pairing symmetry are in agreement with Ref. [56] and they differ from the topological $d + id$ symmetry predicted in many TBG studies [33,36,37,41,42,46,53,59,60,87] and also from other proposed symmetries which include s wave [50–52,55,57,59],
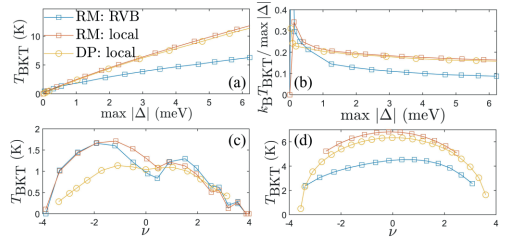
extended s wave [40,43,48,87], p wave [55,58], $p + ip$ wave [39], d wave [52,55,58], and f wave [38,41,42,55]. Apart from Ref. [56], nematic pairing has been predicted only in a few works [39,44,53,54,61,88].

BKT-transition temperature. By computing D^s, one can determine T_{BKT}. In Fig. 3(a) we show T_{BKT} as a function of max $|\Delta|$. We can distinguish two qualitatively different regimes: In the weak-coupling limit the RVB and local interactions yield similar T_{BKT}, whereas for stronger interactions T_{BKT} depends on the pairing model. Moreover, around max $|\Delta| \gtrsim 2$ meV the behavior of the T_{BKT} curves is almost linear, in accordance with previous studies [24,25] where D^s of a flat band with the local interaction was shown to depend linearly on the pairing strength. In our case the narrow bands are not exactly flat but slightly dispersive and thus their flat-band characteristics manifest only when the interaction strength is sufficiently large [51]. Because of this, we call the regime with max $|\Delta| \gtrsim 2$ meV as the flat-band limit. In this regime the DP and RM results are in agreement, whereas for weak interactions the results differ due to different band structures.

The difference of the two interaction schemes is further highlighted in Fig. 3(b) which presents the ratio $k_B T_{BKT}/\max |\Delta(T = 0)|$. At the flat-band limit this ratio approaches a constant whose value depends on the pairing potential. In experiments one can measure T_{BKT} and in principle also deduce Δ [from the local density of states (LDOS)] and thus the ratio of these two quantities can be used to characterize the SC pairing observed in experiments.

In Figs. 3(c) and 3(d) we present T_{BKT} as a function of ν. The weak-coupling regime shows a dome-shaped structure of T_{BKT} which reaches its maxima near the half-fillings of the hole- and electron-doped regimes, similar to experiments [3]. In the RM model the hole-doped region is much stronger due to higher density of states at negative energies [see Fig. 1(b)], while the DP model exhibits approximate electron-hole symmetry. The strong asymmetry of the RM model is due to the applied rescaling approximation which amplifies the finite but small asymmetry of the unscaled model [29]. In the flat-band limit, the shape of the one-particle dispersions are, except for the pronounced particle-hole asymmetry of the RM model, completely dissolved.
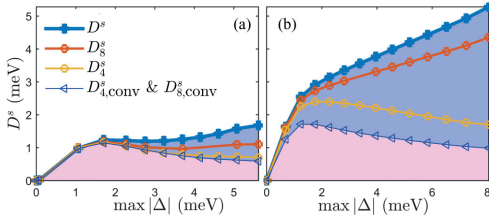

FIG. 3. (a), (b) T_{BKT} and $k_B T_{BKT}/\max |\Delta(T = 0)|$, respectively, as a function of max $|\Delta(T = 0)|$ at $\nu \approx -2$. This result is almost independent of the filling [29]. (c), (d) T_{BKT} as a function of ν at max $|\Delta| \approx 0.4$ meV and max $|\Delta| \approx 3$ meV at CNP, respectively.

FIG. 4. Various superfluid components as a function of $\max|\Delta|$ at $\nu \approx -2$ and $T = 1.5$ K for the (a) RVB and (b) local interaction obtained from the RM model. Blue curve is D^s and blue (pink) area depicts D^s_{geom} (D^s_{conv}). Results for D^s and D^s_{conv} computed by considering only four and eight Bloch bands are also shown, labeled as D^s_4, D^s_8, and $D^s_{4/8,\text{conv}}$. We have numerically checked that $D^s_{\text{conv}} \approx D^s_{4,\text{conv}}$.

Geometric contribution. One can decompose D^s to conventional, D^s_{conv}, and geometric, D^s_{geom}, parts, so that $D^s = D^s_{\text{conv}} + D^s_{\text{geom}}$ [24,27,29]. The conventional term depends on the inverse of the effective mass of the Bloch bands and is thus a single-band contribution, whereas D^s_{geom} is a multiband effect depending on the overlap of the Bloch states and their momentum derivatives of the form $\langle \partial_{\mathbf{k}} n | m \rangle$, where $|m\rangle$ are the single-particle states of the mth Bloch band and $n \neq m$ [27], i.e., $D^s_{\text{geom}} = 0$ for a single-band system. For a strictly flat band, $D^s_{\text{conv}} = 0$ so its superconductivity is purely a multiband process characterized by a finite D^s_{geom}. This raises an intriguing question related to the TBG system: How much do the interband terms between dispersive and narrow bands affect D^s via D^s_{geom}?

We study this question in Fig. 4 for RVB [Fig. 4(a)] and local [Fig. 4(b)] pairing by presenting the total D^s and its components. We further show results obtained by taking into account [29] either only the four flat bands or the eight lowest (four flat, four dispersive) bands labeled as D^s_4 and D^s_8, respectively. In both pairing cases the contribution coming from the four flat bands only is relatively small for larger interactions. The contribution of eight bands is larger due to a larger D^s_{geom}, which is caused by the interband terms between dispersive and flat bands as the terms between the dispersive bands only are negligible. The slight dispersion of the narrow bands results in a finite D^s_{conv}. Note that $D^s_{\text{conv}} \approx D^s_{4,\text{conv}}$, i.e., $D^s - D^s_{4,\text{conv}}$ gives the total D^s_{geom}. From Fig. 4 we see that, at $\max|\Delta| \sim 1 \dots 2$ meV, i.e., when the system enters the flat-band regime, D^s_{geom} surpasses D^s_{conv} for local pairing and becomes significant in the RVB case. An important implication of Fig. 4 is the importance of the dispersive bands when computing D^s and the insufficiency of four-band models even if the pairing occurs predominantly within the four narrow

bands. Also for a noninteracting system, higher bands have been argued to be necessary, but for different (symmetry) reasons [89].

Discussion. Our work shows that TBG is characterized by two distinct superconducting regimes. When Δ is much smaller than the flat-band bandwidth, the superfluid weight D^s and the BKT transition temperature T_{BKT} are well described by conventional theory of superconductivity. On the other hand, in this weak-coupling regime the results are somewhat different for the RM and DP models. This is consistent with the low-energy dispersion in TBG being very sensitive to the details of the model used [90]. In the flat-band regime where Δ is larger than the width of the significant density of states in the flat bands, a major contribution to the superfluid weight D^s originates from the geometric properties of the bands. The geometric contribution D^s_{geom} is proportional to the quantum metric [24] whose importance in physics has been recently emerging [24–28,91–104]. Moreover, in the flat-band regime, both D^s and T_{BKT} depend sensitively on the pairing mechanism, but not strongly on the employed microscopic model. In particular, for a non-local RVB interaction D^s becomes anisotropic, which could be seen in four-terminal radio frequency spectroscopy experiments to reveal information about the pairing mechanism more directly than the LDOS measurements of [105–107].

Within both of our models, at θ^* the crossover between the two regimes takes place for $\Delta = 1.5 \dots 2$ meV, implying $T_{\text{BKT}} \approx 1.5 \dots 3$ K. This is also the ballpark of the experimentally accessed critical temperatures [2,3]. Thus the geometric contribution of the superfluid weight and the dependence on the pairing mechanism should be relevant for current experiments. An interesting future direction of research is to include other interaction channels than pairing and explore the insulating states observed in TBG [2,3,108]. Based on our results, one can anticipate that quantum geometry and multiband processes are important in superconductivity and correlated states of other twisted multilayer materials [42,109–122].

Note added. Recently, a related work [123] appeared at the arXiv preprint server.

[1] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Nature (London) **556**, 43 (2018).

[2] M. Yankowitz, S. Chen, H. Polshyn, Y. Zhang, K. Watanabe, T. Taniguchi, D. Graf, A. F. Young, and C. R. Dean, Science **363**, 1059 (2019).

[3] X. Lu, P. Stepanov, W. Yang, M. Xie, M. A. Aamir, I. Das, C. Urgell, K. Watanabe, T. Taniguchi, G. Zhang, A. Bachtold, A. H. MacDonald, and D. K. Efetov, Nature (London) **574**, 653 (2019).

[4] N. B. Kopnin, T. T. Heikkilä, and G. E. Volovik, Phys. Rev. B **83**, 220503(R) (2011).

[5] R. Ojajärvi, T. Hyart, M. A. Silaev, and T. T. Heikkilä, Phys. Rev. B **98**, 054515 (2018).

[6] G. Li, A. Luican, J. M. B. Lopes dos Santos, A. H. Castro Neto, A. Reina, J. Kong, and E. Y. Andrei, Nat. Phys. **6**, 109 (2010).

[7] R. Bistritzer and A. H. MacDonald, Proc. Natl. Acad. Sci. USA **108**, 12233 (2011).

[8] S. Shallcross, S. Sharma, E. Kandelaki, and O. A. Pankratov, Phys. Rev. B **81**, 165105 (2010).

[9] E. Suárez Morell, J. D. Correa, P. Vargas, M. Pacheco, and Z. Barticevic, Phys. Rev. B **82**, 121407(R) (2010).

[10] I. Brihuega, P. Mallet, H. González-Herrero, G. Trambly de Laissardière, M. M. Ugeda, L. Magaud, J. M. Gómez-Rodríguez, F. Ynduráin, and J.-Y. Veuillen, Phys. Rev. Lett. **109**, 196802 (2012).

[11] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, Phys. Rev. B **86**, 155449 (2012).

[12] G. Trambly de Laissardière, D. Mayou, and L. Magaud, Phys. Rev. B **86**, 125413 (2012).

[13] S. Shallcross, S. Sharma, and O. Pankratov, Phys. Rev. B **87**, 245403 (2013).

[14] K. Uchida, S. Furuya, J. I. Iwata, and A. Oshiyama, Phys. Rev. B **90**, 155451 (2014).

[15] S. Fang and E. Kaxiras, Phys. Rev. B **93**, 235153 (2016).

[16] D. Weckbecker, S. Shallcross, M. Fleischmann, N. Ray, S. Sharma, and O. Pankratov, Phys. Rev. B **93**, 035452 (2016).

[17] N. N. T. Nam and M. Koshino, Phys. Rev. B **96**, 075311 (2017).

[18] V. L. Berezinskii, Zh. Eksp. Teor. Fiz. **61**, 1144 (1972) [Sov. Phys. JETP **34**, 610 (1972)].

[19] J. M. Kosterlitz and D. J. Thouless, J. Phys. C **5**, L124 (1972).

[20] J. M. Kosterlitz and D. J. Thouless, J. Phys. C **6**, 1181 (1973).

[21] D. R. Nelson and J. M. Kosterlitz, Phys. Rev. Lett. **39**, 1201 (1977).

[22] Y. Cao, S.-H. Zou, X.-J. Liu, S. Yi, G.-L. Long, and H. Hu, Phys. Rev. Lett. **113**, 115302 (2014).

[23] M. Tinkham, *Introduction to Superconductivity* (Courier Corporation, New York, 2004).

[24] S. Peotta and P. Törmä, Nat. Commun. **6**, 8944 (2015).

[25] A. Julku, S. Peotta, T. I. Vanhala, D.-H. Kim, and P. Törmä, Phys. Rev. Lett. **117**, 045303 (2016).

[26] M. Tovmasyan, S. Peotta, P. Törmä, and S. D. Huber, Phys. Rev. B **94**, 245149 (2016).

[27] L. Liang, T. I. Vanhala, S. Peotta, T. Siro, A. Harju, and P. Törmä, Phys. Rev. B **95**, 024515 (2017).

[28] P. Törmä, L. Liang, and S. Peotta, Phys. Rev. B **98**, 220511(R) (2018).

[29] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.101.060505 for details on the renormalized moiré (RM) and Dirac point (DP) models, calculation of the order parameter, calculation of the superfluid weight and its separation to conventional and geometric parts, discussion of why our calculations are not in agreement with Ref. [31],

T_{BKT} results at the CNP, and D^s results for different renormalization strengths in the RM model.

[30] C. Brouder, G. Panati, M. Calandra, C. Mourougane, and N. Marzari, Phys. Rev. Lett. **98**, 046402 (2007).

[31] T. Hazra, N. Verma, and M. Randeria, Phys. Rev. X **9**, 031049 (2019).

[32] F. Xie, Z. Song, B. Lian, and B. A. Bernevig, Phys. Rev. Lett. (to be published), arXiv:1906.02213.

[33] H. Guo, X. Zhu, S. Feng, and R. T. Scalettar, Phys. Rev. B **97**, 235453 (2018).

[34] B. Lian, Z. Wang, and B. A. Bernevig, Phys. Rev. Lett. **122**, 257002 (2019).

[35] L. Chen, H.-Z. Li, and R.-S. Han, J. Phys.: Condens. Matter **31**, 065601 (2019).

[36] C. Xu and L. Balents, Phys. Rev. Lett. **121**, 087001 (2018).

[37] C.-C. Liu, L.-D. Zhang, W.-Q. Chen, and F. Yang, Phys. Rev. Lett. **121**, 217001 (2018).

[38] Q.-K. Tang, L. Yang, D. Wang, F.-C. Zhang, and Q.-H. Wang, Phys. Rev. B **99**, 094521 (2019).

[39] B. Roy and V. Juričić, Phys. Rev. B **99**, 121407(R) (2019).

[40] Y. Sherkunov and J. J. Betouras, Phys. Rev. B **98**, 205151 (2018).

[41] Y.-P. Lin and R. M. Nandkishore, Phys. Rev. B **98**, 214521 (2018).

[42] L. Classen, C. Honerkamp, and M. M. Scherer, Phys. Rev. B **99**, 195120 (2019).

[43] S. Ray, J. Jung, and T. Das, Phys. Rev. B **99**, 134515 (2019).

[44] J. F. Dodaro, S. A. Kivelson, Y. Schattner, X. Q. Sun, and C. Wang, Phys. Rev. B **98**, 075154 (2018).

[45] J. Kang and O. Vafek, Phys. Rev. Lett. **122**, 246401 (2019).

[46] T. Huang, L. Zhang, and T. Ma, Sci. Bull. **64**, 310 (2019).

[47] V. Kozii, H. Isobe, J. W. F. Venderbos, and L. Fu, Phys. Rev. B **99**, 144507 (2019).

[48] Z. Liu, Y. Li, and Y.-F. Yang, Chin. Phys. B **28**, 077103 (2019).

[49] D. M. Kennes, J. Lischner, and C. Karrasch, Phys. Rev. B **98**, 241407(R) (2018).

[50] Y. W. Choi and H. J. Choi, Phys. Rev. B **98**, 241412(R) (2018).

[51] T. J. Peltonen, R. Ojajärvi, and T. T. Heikkilä, Phys. Rev. B **98**, 220504(R) (2018).

[52] F. Wu, A. H. MacDonald, and I. Martin, Phys. Rev. Lett. **121**, 257001 (2018).

[53] F. Wu, Phys. Rev. B **99**, 195114 (2019).

[54] F. Wu and S. Das Sarma, Phys. Rev. B **99**, 220507(R) (2019).

[55] F. Wu, E. Hwang, and S. Das Sarma, Phys. Rev. B **99**, 165112 (2019).

[56] Y. Su and S.-Z. Lin, Phys. Rev. B **98**, 195101 (2018).

[57] H. C. Po, L. Zou, A. Vishwanath, and T. Senthil, Phys. Rev. X **8**, 031089 (2018).

[58] H. Isobe, N. F. Q. Yuan, and L. Fu, Phys. Rev. X **8**, 041041 (2018).

[59] E. Laksono, J. N. Leaw, A. Reaves, M. Singh, X. Wang, S. Adam, and X. Gu, Solid State Commun. **282**, 38 (2018).

[60] M. Fidrysiak, M. Zegrodnik, and J. Spałek, Phys. Rev. B **98**, 085436 (2018).

[61] J. W. F. Venderbos and R. M. Fernandes, Phys. Rev. B **98**, 245103 (2018).

[62] J. González and T. Stauber, Phys. Rev. Lett. **122**, 026801 (2019).

[63] L. A. Gonzalez-Arraga, J. L. Lado, F. Guinea, and P. San-Jose, Phys. Rev. Lett. **119**, 107201 (2017).

[64] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, Phys. Rev. Lett. **99**, 256802 (2007).

[65] N. F. Q. Yuan and L. Fu, Phys. Rev. B **98**, 045103 (2018).

[66] J. Kang and O. Vafek, Phys. Rev. X **8**, 031088 (2018).

[67] M. Koshino, N. F. Q. Yuan, T. Koretsune, M. Ochi, K. Kuroki, and L. Fu, Phys. Rev. X **8**, 031087 (2018).

[68] E. Zhao and A. Paramekanti, Phys. Rev. Lett. **97**, 230404 (2006).

[69] B. Uchoa and A. H. Castro Neto, Phys. Rev. Lett. **98**, 146801 (2007).

[70] N. B. Kopnin and E. B. Sonin, Phys. Rev. Lett. **100**, 246808 (2008).

[71] M. V. Hosseini, Europhys. Lett. **110**, 47010 (2015).

[72] P. W. Anderson, Science **235**, 1196 (1987).

[73] P. A. Lee, N. Nagaosa, and X.-G. Wen, Rev. Mod. Phys. **78**, 17 (2006).

[74] G. Baskaran, Phys. Rev. B **65**, 212505 (2002).

[75] A. M. Black-Schaffer and S. Doniach, Phys. Rev. B **75**, 134512 (2007).

[76] S. Pathak, V. B. Shenoy, and G. Baskaran, Phys. Rev. B **81**, 085431 (2010).

[77] A. M. Black-Schaffer, Phys. Rev. Lett. **109**, 197001 (2012).

[78] A. M. Black-Schaffer and C. Honerkamp, J. Phys.: Condens. Matter **26**, 423201 (2014).

[79] B. Roy and I. F. Herbut, Phys. Rev. B **82**, 035429 (2010).

[80] M. L. Kiesel, C. Platt, W. Hanke, D. A. Abanin, and R. Thomale, Phys. Rev. B **86**, 020507(R) (2012).

[81] C. Honerkamp, Phys. Rev. Lett. **100**, 146404 (2008).

[82] In Ref. [51] we only included three of the largest Fourier components of the interlayer coupling. Now we include all of them, up to the same cutoff used also in the Hamiltonians. With this added modification the magic angle moves to $\theta = 1.08°$, the bandwidth of the flat bands increases to ~ 10 meV, and a gap of 20 meV is opened between the flat bands and the next lowest bands.

[83] J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954).

[84] D. J. Scalapino, S. R. White, and S. C. Zhang, Phys. Rev. Lett. **68**, 2830 (1992).

[85] D. J. Scalapino, S. R. White, and S. Zhang, Phys. Rev. B **47**, 7995 (1993).

[86] F. Chiodi, M. Ferrier, K. Tikhonov, P. Virtanen, T. Heikkilä, M. Feigelman, S. Guéron, and H. Bouchiat, Sci. Rep. **1**, 3 (2011).

[87] M. Angeli, E. Tosatti, and M. Fabrizio, Phys. Rev. X **9**, 041010 (2019).

[88] D. V. Chichinadze, L. Classen, and A. V. Chubukov, arXiv:1910.07379.

[89] H. C. Po, L. Zou, T. Senthil, and A. Vishwanath, Phys. Rev. B **99**, 195455 (2019).

[90] N. R. Walet and F. Guinea, arXiv:1903.00340.

[91] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States* (Cambridge University Press, Cambridge, UK, 2006).

[92] S.-J. Gu, Int. J. Mod. Phys. A **24**, 4371 (2010).

[93] T. Neupert, C. Chamon, and C. Mudry, Phys. Rev. B **87**, 245103 (2013).

[94] R. Roy, Phys. Rev. B **90**, 165139 (2014).

[95] E. Dobardžić, M. V. Milovanović, and N. Regnault, Phys. Rev. B **88**, 115117 (2013).

[96] A. Srivastava and A. Imamoğlu, Phys. Rev. Lett. **115**, 166802 (2015).

[97] Y. Gao, S. A. Yang, and Q. Niu, Phys. Rev. B **91**, 214405 (2015).

[98] F. Piéchon, A. Raoux, J.-N. Fuchs, and G. Montambaux, Phys. Rev. B **94**, 134423 (2016).

[99] T. Ozawa, Phys. Rev. B **97**, 041108(R) (2018).

[100] M. F. Lapa and T. L. Hughes, Phys. Rev. B **99**, 121111(R) (2019).

[101] M. Yu, P. Yang, M. Gong, Q. Cao, Q. Lu, H. Liu, M. B. Plenio, F. Jelezko, T. Ozawa, and N. Goldman, Natl. Sci. Rev. **nwz193** (2019).

[102] A. Gianfrate, O. Bleu, L. Dominici, V. Ardizzone, M. De Giorgi, D. Ballarini, K. West, L. N. Pfeiffer, D. D. Solnyshkov, and D. Sanvitto, arXiv:1901.03219.

[103] L. Asteria, D. T. Tran, T. Ozawa, M. Tarnowski, B. S. Rem, N. Fläschner, K. Sengstock, N. Goldman, and C. Weitenberg, Nat. Phys. **15**, 449 (2019).

[104] X. Tan, D.-W. Zhang, Z. Yang, J. Chu, Y.-Q. Zhu, D. Li, X. Yang, S. Song, Z. Han, Z. Li, Y. Dong, H.-F. Yu, H. Yan, S.-L. Zhu, and Y. Yu, Phys. Rev. Lett. **122**, 210401 (2019).

[105] Y. Jiang, X. Lai, K. Watanabe, T. Taniguchi, K. Haule, J. Mao, and E. Y. Andrei, Nature (London) **573**, 91 (2019).

[106] Y. Choi, J. Kemmer, Y. Peng, A. Thomson, H. Arora, R. Polski, Y. Zhang, H. Ren, J. Alicea, G. Refael, F. von Oppen, K. Watanabe, T. Taniguchi, and S. Nadj-Perge, Nat. Phys. (2019).

[107] A. Kerelsky, L. J. McGilly, D. M. Kennes, L. Xian, M. Yankowitz, S. Chen, K. Watanabe, T. Taniguchi, J. Hone, C. Dean, A. Rubio, and A. N. Pasupathy, Nature (London) **572**, 95 (2019).

[108] Y. Cao, V. Fatemi, A. Demir, S. Fang, S. L. Tomarken, J. Y. Luo, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, E. Kaxiras, R. C. Ashoori, and P. Jarillo-Herrero, Nature (London) **556**, 80 (2018).

[109] G. Chen, A. L. Sharpe, P. Gallagher, I. T. Rosen, E. J. Fox, L. Jiang, B. Lyu, H. Li, K. Watanabe, T. Taniguchi, J. Jung, Z. Shi, D. Goldhaber-Gordon, Y. Zhang, and F. Wang, Nature (London) **572**, 215 (2019).

[110] C. Shen, N. Li, S. Wang, Y. Zhao, J. Tang, J. Liu, J. Tian, Y. Chu, K. Watanabe, and T. Taniguchi, arXiv:1903.06952.

[111] X. Liu, Z. Hao, E. Khalaf, J. Y. Lee, K. Watanabe, T. Taniguchi, A. Vishwanath, and P. Kim, arXiv:1903.08130.

[112] Y. Cao, D. Rodan-Legrain, O. Rubies-Bigordà, J. M. Park, K. Watanabe, T. Taniguchi, and P. Jarillo-Herrero, arXiv:1903.08596.

[113] G.-Y. Zhu, T. Xiang, and G.-M. Zhang, Sci. Bull. **63**, 1087 (2018).

[114] G.-Y. Zhu, T. Xiang, and G.-M. Zhang, arXiv:1806.07535.

[115] Y.-H. Zhang and T. Senthil, Phys. Rev. B **99**, 205150 (2019).

[116] F. Wu, T. Lovorn, E. Tutuc, I. Martin, and A. H. MacDonald, Phys. Rev. Lett. **122**, 086402 (2019).

[117] Y.-H. Zhang, D. Mao, and T. Senthil, Phys. Rev. Res. **1**, 033126 (2019).

[118] Y.-H. Zhang, D. Mao, Y. Cao, P. Jarillo-Herrero, and T. Senthil, Phys. Rev. B **99**, 075127 (2019).

[119] J. Liu, Z. Ma, J. Gao, and X. Dai, Phys. Rev. X **9**, 031021 (2019).

[120] J. Y. Lee, E. Khalaf, S. Liu, X. Liu, Z. Hao, P. Kim, and A. Vishwanath, Nat. Commun. **10**, 5333 (2019).

[121] C. Schrade and L. Fu, Phys. Rev. B **100**, 035413 (2019).

[122] M. S. Scheurer, R. Samajdar, and S. Sachdev, arXiv:1906.03258.

[123] X. Hu, T. Hyart, D. I. Pikulin, and E. Rossi, Phys. Rev. Lett. **123**, 237002 (2019).

Superfluid weight and Berezinskii-Kosterlitz-Thouless transition temperature of twisted bilayer graphene: supplementary material

A. Julku,[1] T. J. Peltonen,[2] L. Liang,[1, 3] T. T. Heikkilä,[2] and P. Törmä[1]

[1]*Department of Applied Physics, Aalto University, P.O.Box 15100, 00076 Aalto, Finland*
[2]*Department of Physics and Nanoscience Center, University of Jyväskylä,*
P.O. Box 35 (YFL), FI-40014 University of Jyväskylä, Finland
[3]*Computational Physics Laboratory, Physics Unit, Faculty of Engineering and*
Natural Sciences, Tampere University, P.O. Box 692, FI-33014 Tampere, Finland
(Dated: August 26, 2019)

I. DETAILS OF THE RENORMALIZED MOIRÉ (RM) AND DIRAC POINT (DP) MODELS

In this section we provide additional information on the details of our RM and DP models and how the order parameters are solved from both models.

A. Renormalized moiré model (RM)

1. Computation of the order parameters

Let us start by writing the Fermi-Hubbard model already presented in the main text:

$$H = \sum_{i\alpha j\beta\sigma} t_{i\alpha j\beta} c^\dagger_{i\alpha\sigma} c_{j\beta\sigma} - \mu \sum_{i\alpha\sigma} c^\dagger_{i\alpha\sigma} c_{i\alpha\sigma} + H_{\text{int}}, \tag{S1}$$

where for the kinetic hopping amplitudes $t_{i\alpha j\beta}$ we use the parametrization provided by the Slater-Koster table of interatomic matrix elements [1] for p_z orbitals of the carbon atoms:

$$t_{i\alpha j\beta} = t(\mathbf{r}) = t_0 \exp\left[-\beta\frac{r-b}{b}\right]\frac{x^2+y^2}{r^2} + t_1 \exp\left[-\beta\frac{r-c_0}{b}\right]\frac{z^2}{r^2}, \tag{S2}$$

where for simplicity we have denoted the distance between $\mathbf{r}_{i\alpha}$ and $\mathbf{r}_{j\beta}$ as $\mathbf{r} = [x, y, z]$. Here the z-axis is perpendicular with respect to the graphene layers. The first term in Eq. (S2) describes the intralayer hopping processes ($z = 0$), whereas the interlayer processes are mainly described by the latter term. Here $b = a_0/\sqrt{3}$ is the distance between the nearest-neighbour carbon atoms, $a_0 = 0.246$ nm is the lattice constant of graphene, and $c_0 = 0.335$ nm is the interlayer distance. In our calculations we use parameters $t_0 = -2.7$ eV, $t_1 = 0.297$ eV, and $\beta = 7.2$. We restrict the interlayer hopping to the terms with $r < 4b$ and consider only the nearest-neighbour intralayer hopping terms.

For the interaction Hamiltonian, we use two different kinds of forms, namely the local attractive Hubbard interaction $H_{\text{int}} = J \sum_{i\alpha} c^\dagger_{i\alpha\uparrow} c^\dagger_{i\alpha\downarrow} c_{i\alpha\downarrow} c_{i\alpha\uparrow} \equiv H_{\text{loc}}$ and the resonance valence bond (RVB) type nearest-neighbour pairing potential $H_{\text{int}} = \frac{J}{2} \sum_{\langle i\alpha j\beta\rangle} h^\dagger_{i\alpha j\beta} h_{i\alpha j\beta} \equiv H_{\text{RVB}}$, in order to see how the nature of the interaction Hamiltonian affects D^s and T_{BKT}. Local interaction has been extensively used in the past graphene [2–5] and TBG studies [6, 7] to model s-wave superconductivity mediated by electron-phonon interaction. Strictly speaking, in graphene the local (meaning also intrasublattice) interaction does not capture all implications of the attractive interaction mediated by electron–phonon coupling, but also the next-nearest neighbour interactions are often also included [2, 7]. In those works, it has been shown that the intersublattice coupling results to d-wave pairing. However, s-wave pairing was found to be the dominant pairing mediated by electron–phonon coupling, in which case the mean-field results of Ref. 7 are essentially the same as in the completely local model used in e.g. Ref. 6.

In addition to local pairing potentials, also RVB and other non-local pairing schemes have been applied both in graphene [2, 8–13] and TBG studies [14]. To understand the RVB pairing scheme, let us note that it can be rewritten as $H_{\text{int}} = |J| \sum_{\langle i\alpha j\beta\rangle} \mathbf{S}_{i\alpha} \cdot \mathbf{S}_{j\beta} + 1/4n_{i\alpha}n_{j\beta}$, where $\mathbf{S}_{i\alpha}$ are spin operators, which is the usual Heisenberg antiferromagnetic Hamiltonian [15]. For example, in the usual case of large repulsive on-site Coulomb interaction, the system Hamiltonian assumes the form of the $t-J$ model which has the double occupancy excluded and the interaction term is the aforementioned Heisenberg antiferromagnet. This leads to e.g. antiferromagnetic Mott-physics in case of a undoped simple square lattice as is well known [16].

In our case we use RVB interaction but without the exclusion of the double occupancy, as in case of graphene the on-site repulsive Coulomb interaction between electrons is not necessarily large enough to justify the $t-J$ model. Early treatments of planar organic molecules of $\sigma\pi$-bonds were heavily based on the RVB approach [17] and in 2002 it was suggested [8] that the RVB approach would be a viable model to describe possible superconductivity in doped graphene and other similar organic layers. The argument for this is similar than in the case of the usual $t-J$ Hamiltonian: repulsive on-site Coulomb interaction results in two-body singlet correlations between neighbouring sites. However, as the on-site repulsion is not large enough, double occupancy is not entirely ruled out. Later in 2007 the RVB model was used in the mean-field level to show the possibility of $d+id$ superconductivity in doped graphene layers [9]. In 2010, the authors of Ref. 10, inspired by the works of Refs. 8 and 9, performed rigorous variational quantum Monte Carlo calculations by assuming local on-site Coulomb repulsion and showed that, indeed, doped graphene can support supercurrent with finite nearest-neighbour singlet pairing (RVB) correlations. Thus, non-local singlet pairing can be thought to originate from the Hubbard model of the graphene layers via the Coulomb electron-electron repulsion. Non-local singlet $d+id$ (and for some parameters also triplet f-wave) pairing in case of doped graphene was also obtained in Ref. 18 where functional renormalization group calculations were performed by assuming local and non-local repulsive Coulomb interactions. The authors of Ref. 18 showed that in graphene next-nearest-neighbour (NNN) and next-next-nearest-neighbour (NNNN) pairing schemes are possible. The nearest-neighbour RVB interaction used in our work is one of the simplest possible non-local pairing potentials but of course it would be interesting to also analyze these NNN and NNNN pairing schemes. This remains a topic of future studies.

To evaluate the chosen interaction Hamiltonians, we apply standard mean-field decoupling which yields

$$H_{\text{loc}} \approx \sum_{i\alpha} \Delta_{i\alpha} c^{\dagger}_{i\alpha\uparrow} c^{\dagger}_{i\alpha\downarrow} + \text{H.c.}, \tag{S3}$$

$$H_{\text{RVB}} \approx \sum_{\langle i\alpha j\beta \rangle} \Delta_{i\alpha j\beta}(c^{\dagger}_{i\alpha\uparrow} c^{\dagger}_{j\beta\downarrow} - c^{\dagger}_{i\alpha\downarrow} c^{\dagger}_{j\beta\uparrow}) + \text{H.c.}, \tag{S4}$$

where the order parameters are $\Delta_{i\alpha} = J\langle c_{i\alpha\downarrow} c_{i\alpha\uparrow} \rangle$ and $\Delta_{i\alpha j\beta} = \frac{J}{2}\langle c_{j\beta\downarrow} c_{i\alpha\uparrow} - c_{j\beta\uparrow} c_{i\alpha\downarrow} \rangle$.

By using mean-field interaction terms (S3) and (S4) and by exploiting the translational invariance, it is easy to rewrite the Hamiltonian (S1) in the momentum space as

$$H = \sum_{\mathbf{k}} \Psi_{\mathbf{k}} \mathcal{H}_{\mathbf{k}} \Psi_{\mathbf{k}}, \text{ where} \tag{S5}$$

$$\Psi_{\mathbf{k}} = [c_{\mathbf{k}\uparrow}, c_{-\mathbf{k}\downarrow}]^T, \tag{S6}$$

$$c_{\mathbf{k}\sigma} = [c_{\alpha=1,\mathbf{k},\sigma}, c_{\alpha=2,\mathbf{k},\sigma}, ..., c_{\alpha=M,\mathbf{k},\sigma}]^T, \tag{S7}$$

$$c_{\alpha\mathbf{k}\sigma} = \frac{1}{\sqrt{V}} \sum_{i\alpha} e^{i\mathbf{k}\cdot\mathbf{r}_{i\alpha}} c_{i\alpha\sigma}, \text{ and} \tag{S8}$$

$$\mathcal{H}_{\mathbf{k}} = \begin{bmatrix} \mathcal{H}_{\uparrow}(\mathbf{k}) - \mu_{\uparrow} & \Delta(\mathbf{k}) \\ \Delta^{\dagger}(\mathbf{k}) & -\mathcal{H}^*_{\downarrow}(-\mathbf{k}) + \mu_{\downarrow}, \end{bmatrix}, \tag{S9}$$

where the diagonal (off-diagonal) blocks are the Fourier transforms of hopping (pairing) terms. Here M is the number of lattice sites per unit cell, V is the total area of the system and $\mathbf{k}$ belong to the unit cell of the reciprocal lattice.

By solving the BdG eigenproblem $\mathcal{H}_{\mathbf{k}} |\psi_{i\mathbf{k}}\rangle = E_{i\mathbf{k}} |\psi_{i\mathbf{k}}\rangle$, we obtain the eigendecomposition $\mathcal{H}_{\mathbf{k}} = V_{\mathbf{k}} D_{\mathbf{k}} V^{\dagger}_{\mathbf{k}}$. The diagonal matrix $D_{\mathbf{k}}$ contains the eigenenergies $E_{i\mathbf{k}}$, whereas the columns of the unitary matrix $V_{\mathbf{k}}$ are the eigenstates $|\psi_{i\mathbf{k}}\rangle$. One can then write down the self-consistent gap equations for the order parameters with the aid of $D_{\mathbf{k}}$ and $V_{\mathbf{k}}$. For the local interaction these read

$$\Delta_{i\alpha} = J\langle c_{i\alpha\downarrow} c_{i\alpha\uparrow} \rangle = \frac{J}{N} \sum_{\mathbf{k}} \left[V_{\mathbf{k}} f(D_{\mathbf{k}}) V^{\dagger}_{\mathbf{k}} \right]_{\alpha, M+\beta}, \tag{S10}$$

and correspondingly for the RVB

$$\Delta_{i\alpha j\beta} = \frac{J}{2}\langle c_{j\beta\downarrow} c_{i\alpha\uparrow} - c_{j\beta\uparrow} c_{i\alpha\downarrow} \rangle = \frac{J}{2N} \sum_{\mathbf{k}} \left\{ e^{-i\mathbf{k}\cdot\mathbf{r}^{\text{rel}}_{j\beta i\alpha}} \left[V_{\mathbf{k}} f(D_{\mathbf{k}}) V^{\dagger}_{\mathbf{k}} \right]_{\alpha, M+\beta} - e^{i\mathbf{k}\cdot\mathbf{r}^{\text{rel}}_{j\beta i\alpha}} \left[V_{\mathbf{k}} \left(\mathbb{1}_{2M} - f(D_{\mathbf{k}}) \right) V^{\dagger}_{\mathbf{k}} \right]_{\beta, M+\alpha} \right\}, \tag{S11}$$

where $\mathbf{r}^{\text{rel}}_{j\beta i\alpha} \equiv \mathbf{r}_{j\beta} - \mathbf{r}_{i\alpha}$, f is the Fermi-Dirac distribution, N is the number of unit cells, and $\mathbb{1}_{2M}$ is a unity matrix

of size $2M \times 2M$. The densities for each lattice sites can be solved from the following equations,

$$n_{i\alpha\uparrow} = \frac{1}{N} \sum_{\mathbf{k}} \left[V_{\mathbf{k}} f(D_{\mathbf{k}}) V_{\mathbf{k}}^\dagger \right]_{\alpha,\alpha}, \tag{S12}$$

$$n_{i\alpha\downarrow} = \frac{1}{N} \sum_{\mathbf{k}} \left[V_{\mathbf{k}} \left(\mathbb{1}_{2M} - f(D_{\mathbf{k}}) \right) V_{\mathbf{k}}^\dagger \right]_{M+\alpha,M+\alpha}. \tag{S13}$$

The gap equations [Eqs. (S10) or (S11)] and density equations (S12) are solved iteratively with the fixed-point iteration scheme by choosing a random initial ansatz for the order parameters. The iteration is terminated when the order parameters and densities are converged to a stable solution of the gap equations. Due to the translational invariance, we can write $\Delta_{i\alpha} = \Delta_\alpha$, $\Delta_{i\alpha j\beta} = \Delta_{\alpha\beta}$ and $n_{i\alpha\sigma} = n_{\alpha\sigma}$. For local interaction there exist M order parameters (for each lattice site in the unit cell) and in case of RVB there are $3M$ order parameters to be solved (for each nearest-neighbour bond). From the obtained order parameters one can then compute the superfluid weight D^s as explained in section II.

2. Rescaling approximation

When the bilayer graphene systems are twisted close to experimentally used magic angle $\theta^* \approx 1°$, the unit cell consists of around 12000–13000 lattice sites. Such a huge problem is computationally a rather heavy task and thus we decrease the number of lattice sites per unit cell, M, by applying a rescaling trick which keeps invariant two important observables, namely the Fermi velocity of a single graphene v_F and the moiré superlattice periodicity a while increasing the twist angle θ and thus decreasing M. More specifically, the Fermi velocity is proportional to a_0 and t_0, so that $v_F \propto a_0 t_0$. On the other hand, the moiré periodicity is $a = a_0/2\sin(\theta/2)$. With this information, one can introduce the following rescaling under which a and v_F remain invariant [14, 19]:

$$t_0' = \frac{t_0}{\lambda}, \quad a_0' = \lambda a_0, \quad c_0' = \lambda c_0, \quad \lambda = \frac{\sin\theta'/2}{\sin\theta/2}, \tag{S14}$$

where the primed quantites are the ones used in computations. With this trick one can apply much larger twist angles θ' than the usual magic angle $\theta \sim 1°$ and thus have much less lattice sites per moiré unit cell than at $\theta \sim 1°$. The rescaling is characterized by the rescaling parameter $\lambda \geqslant 1$: larger λ means more aggressive rescaling and larger θ', whereas $\lambda = 1$ corresponds to the unscaled model.

Most importantly, the rescaling is able to reproduce the flat bands and dispersive bands sufficiently well near the charge neutrality point as demonstrated in Fig. S1 where the low energy band structure is depicted for three different scaling parameters λ. The unscaled angle is chosen to be $\theta = 1.0138°$ and the rescaled angles are $\theta' = 4.4085°$, $\theta' = 1.8901°$, and $\theta' = 1.0178°$, corresponding to the scaling factors of $\lambda = 4.3475$, $\lambda = 1.8643$, and $\lambda = 1.004$ (we do not use here $\lambda = 1$ as our twist $\theta = 1.0138°$ is not strictly commensurate, i.e. it strictly does not yield periodic structure, whereas our twist angles θ' used in the computations are always commensurate i.e. they strictly preserve the translational invariance.) We see that already a rather aggressive rescaling with $M = 676$ and $\lambda = 4.3475$ is able to reproduce reasonably well the low energy band structure and a less aggressive rescaling of $\lambda = 1.8643$ is in practice identical to the unscaled band structure. From Fig. S1 we also see that rescaling amplifies the small electron-hole symmetry of the unscaled system. This explains why we see a fairly non-symmetric T_{BKT} profile as a function of filling for RM model as shown in Figs. 2(c)–(d) of the main text.

The rescaling approximation can be qualitatively understood by noting that under the rescaling the intraband hopping becomes smaller, i.e. the interband hopping terms become relatively more prominent and therefore one does not need to apply such a small twist angle to obtain flat band structures near the charge neutrality point.

3. Choosing the twist angle

More aggressive rescaling, i.e. larger λ, yields a smaller amount of lattice sites per moiré unit cell. However, λ cannot be arbitrarily large as too strong rescaling cannot reproduce the original unscaled one-particle energy band structure. Feasibility of a specific rescaling λ depends on the value of θ. Some twist angles θ allow one to use more aggressive rescaling than some other twist angles.

The unscaled tight-binding model yields reasonably narrow bands near the charge neutrality point for the twist angles in the range of around $\theta \approx 0.95° \ldots 1.05°$ so that the bandwidth of the flat bands at $\theta = 1.05°$ is around $20\,\mathrm{meV}$, whereas near $0.95°$ it is less than $10\,\mathrm{meV}$. In this angle range also the band gaps between the dispersive and

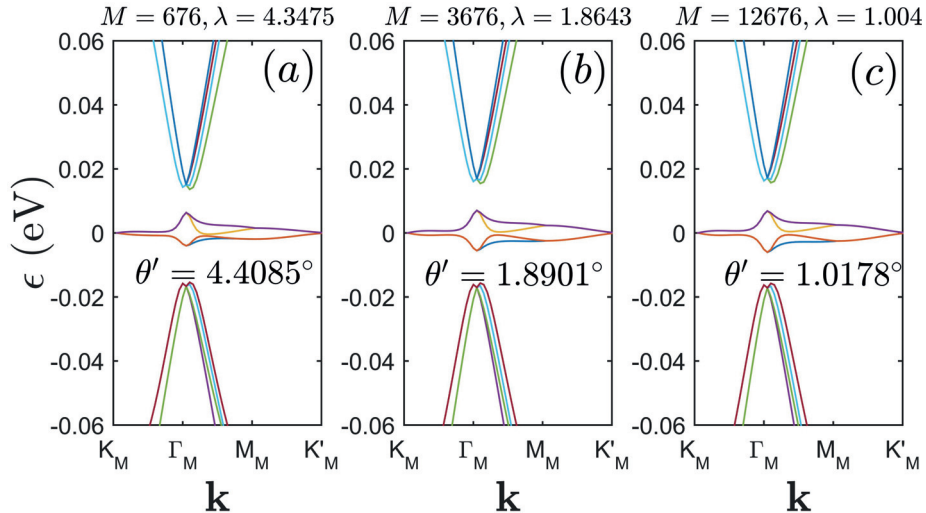FIG. S1. (a)–(c) Low energy band dispersions for $\theta = 1.0138°$ by using three different rescaling strengths. The most aggressive rescaling, namely $\lambda = 4.3475$, slightly alters the shape of the flat bands but at the same time yields a considerably easier problem to solve with only 676 lattice sites per unit cell.
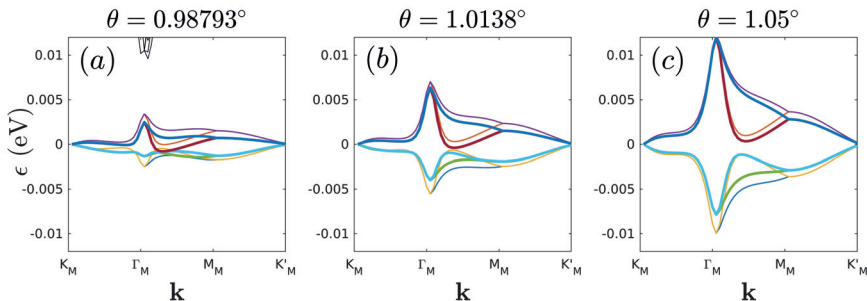

FIG. S2. Flat band dispersion for three different twist angles computed by using two different rescaling strengths. Thick lines correspond to dispersions computed with strong rescaling that yields $M = 676$, whereas thin lines correspond to rescaling of $M = 3676$.

flat bands are notable. Based on these remarks, one is tempted to use angles near $0.95°$ as there the bandwidth is at smallest. However, it turns out that to reproduce the shape of these extremely narrow flat bands near $0.95°$ requires extremely mild rescaling and so one has to deal with a large number of lattice sites within a unit cell. On the contrary, for less narrow bands one can apply a more aggressive rescaling. In Fig. S2 we show the flat band dispersions for three different twist angles: $\theta = 0.987\,93°$ [Fig. S2(a)], $\theta = 1.0138°$ [Fig. S2(b)], and $\theta = 1.05°$ [Fig. S2(c)]. The dispersions are plotted with two different rescaling angles: the thick lines correspond to the rescaling that yields $M = 676$ and the narrow lines to the rescaling with $M = 3676$. The latter rescaling is sufficient to get fairly accurate band structures compared to the unscaled model. We see that more aggressive rescaling yields smaller bandwidths and at $\theta = 0.987\,93°$ alters the shape of the bands considerably. From Fig. S2(a) one can see that with aggressive rescaling the third lowest flat band actually touches the two lowest bands. For larger θ the shapes of the flat bands remain rather invariant under the aggressive scaling. In these cases two lowest bands remain, apart at the Dirac
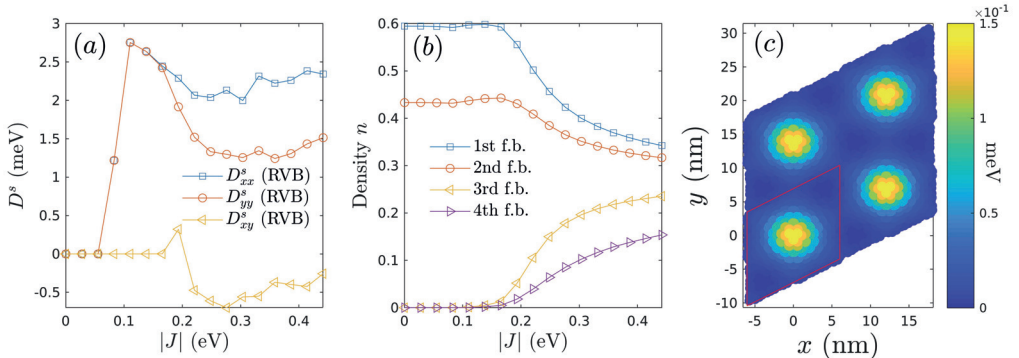
FIG. S3. (a) Different spatial components of D^s for RVB interaction at $\nu \approx -2$ as a function of the interaction strength $|J|$. (b) The corresponding densities of the flat bands. (c) The spatial profile of $\Delta_{i\alpha j\beta}$ at $|J| = 0.11\,\mathrm{eV}$.

points, isolated from the two upper flat bands and also the overall shape of the bands is fairly well reproduced. As argued in the next section, it is important to preserve the shapes of the energy bands essentially unaltered to obtain the isotropic SC state characterized by the mixed $(p + ip) + (d + id)$ pairing symmetry.

In our RM computations we use $\theta = 1.0138°$ and $\lambda \approx 4.3475$ which yields $M = 676$ and $\theta' = 4.4085°$, i.e. we use the dispersions depicted in Fig. S1(a) and S2(b). This choice is a good compromise between the bandwidth and the shape of the rescaled bands.

4. Isotropic SC state with weak RVB interaction

In the main text we showed that at the flat band regime the RVB interaction breaks the C_3-symmetry of the TBG lattice and consequently results in nematic SC states which manifest as an anisotropic superfluid weight. However, for weak enough RVB interactions, when the underlying symmetries of the lattice play a prominent role, one obtains isotropic SC states. This is shown in Fig. S3(a) where we have reproduced the superfluid weight RVB results of Fig. 2(e) of the main text but this time (for clarity) as a function of the interaction strength J. One can see in the weak coupling regime the isotropic phase for which $D_{xx}^s = D_{yy}^s$ and $D_{xy}^s = 0$. For some critical interaction strength the system then becomes nematic.

This transition to the nematic phase is visible also in Fig. S3(b) where we plot the densities of the four flat bands as a function of $|J|$. In the isotropic phase only the lower two flat bands are occupied whereas in the nematic phase the interaction is strong enough to redistribute some of the electrons to the upper flat bands. This is the reason why we are not using e.g. the twist angle $\theta = 0.988°$ depicted in Fig. S2(a) for which the third lowest flat band actually touches the lowest flat band for a chosen rescaling strength. Due to this band touching, the electrons are redistributed to the upper flat bands with a vanishingly small interaction strength which prevents one to obtain the isotropic SC state. We emphasize that this band touching is an aberration caused by the rescaling approximation: for weaker rescaling (i.e. smaller λ) one should obtain the isotropic phase in the weak coupling regime also for $\theta = 0.988°$.

For completeness, the spatial profile of the order parameters of the isotropic phase is shown in Fig. S3(c) for $|J| = 0.11\,\mathrm{eV}$. In contrast to the isotropic phase resulting from the local interaction, the pairing symmetry here is not an s wave but a mixed $(p + ip) + (d + id)$ wave as was shown in Ref. 14.

B. Dirac point model (DP)

As described in Ref. 6, the $\mathbf{G}, \mathbf{G}'$-component (vectors in the reciprocal superlattice) of the normal-state Hamiltonian matrix element $\mathcal{H}_{\rho\mathbf{k}}^0 := \mathcal{H}_{\mathrm{kin},\rho\mathbf{k}}^1 + \mathcal{H}_{\mathrm{kin},\rho\mathbf{k}}^2 + \mathcal{H}_{\mathrm{kin},\rho\mathbf{k}}^\perp - \mu\mathbb{1}$ at valley $\rho \in \{+, -\}$, $\mathbf{k} \in$ MBZ is

$$\mathcal{H}_{\rho\mathbf{k}}^0(\mathbf{G}, \mathbf{G}') = \begin{pmatrix} [\hbar v_F \boldsymbol{\sigma}^\rho \cdot (\mathbf{k} + \mathbf{G} + \rho\Delta\mathbf{K}/2) - \mu]\delta_{\mathbf{GG'}} & t_\perp^\rho(\mathbf{G} - \mathbf{G}') \\ t_\perp^\rho(\mathbf{G}' - \mathbf{G})^\dagger & [\hbar v_F \boldsymbol{\sigma}_\theta^\rho \cdot (\mathbf{k} + \mathbf{G} - \rho\Delta\mathbf{K}/2) - \mu]\delta_{\mathbf{GG'}} \end{pmatrix}, \quad (S15)$$

where the matrix structure corresponds to the layer space, $\boldsymbol{\sigma}^\rho = (\rho\sigma_x, \sigma_y)$ consists of Pauli matrices acting in the sublattice space, $\boldsymbol{\sigma}_\theta^\rho = R(\theta)\boldsymbol{\sigma}^\rho$ is the θ-rotated version of it, v_F is the Fermi velocity of monolayer graphene, $\Delta\mathbf{K} = R(\theta)\mathbf{K} - \mathbf{K}$ is a vector from the graphene $\mathbf{K}$-point to its θ-rotated counterpart, and $t_\perp^\rho(\mathbf{G})$ is a sublattice matrix containing a Fourier component of the interlayer coupling (times an exponential factor) with the elements

$$t_\perp^{\rho,ss'}(\mathbf{G}) = \frac{1}{N} \sum_{\mathbf{r}\in\text{MUC}} e^{-i\mathbf{G}\cdot(\mathbf{r}+\delta_{sB}\boldsymbol{\delta}_1)} e^{i\rho\mathbf{K}^\theta\cdot\boldsymbol{\delta}^{ss'}(\mathbf{r})} t_\perp(\boldsymbol{\delta}^{ss'}(\mathbf{r})). \tag{S16}$$

Here $\boldsymbol{\delta}^{ss'}(\mathbf{r})$ is the horizontal displacement vector between the site at $\mathbf{r}$, sublattice s in layer 1 and the nearest-neighbor at sublattice s' in layer 2. δ_1 denotes one of the nearest-neighbor vectors connecting the graphene A and B sublattices. The sum is over the graphene A sublattice sites in the superlattice unit cell (the moiré unit cell MUC), and N denotes the number of these sites. The interlayer coupling depends only on the (horizontal) distance $\boldsymbol{\delta}$ between the atoms, and is parametrized by a Slater-Koster parametrization as [20, 21]

$$t_\perp(\delta) = \frac{1}{c_0^2 + \delta^2} \left(c_0^2 V_{pp\sigma}\left(\sqrt{c_0^2 + \delta^2}\right) + \delta^2 V_{pp\pi}\left(\sqrt{c_0^2 + \delta^2}\right) \right), \tag{S17}$$

with

$$V_{pp\sigma/\pi}(r) = \alpha_1^{\sigma/\pi} f_{pp\sigma/\pi}(r), \quad f_{pp\sigma/\pi}(r) = r^{-\alpha_2^{\sigma/\pi}} \exp\left(-\alpha_3^{\sigma/\pi} r^{\alpha_4^{\sigma/\pi}}\right). \tag{S18}$$

Here $c_0 = 3.35\,\text{Å}$ is the Bernal graphite interlayer distance, $a_0 = 2.461\,\text{Å}$ is the graphene lattice constant, and the α parameters are chosen as

$$\begin{array}{llll} \alpha_1^\sigma = t_\perp^0/f_{pp\sigma}(c), & \alpha_2^\sigma = 0.7620, & \alpha_3^\sigma = 0.1624, & \alpha_4^\sigma = 2.3509, \\ \alpha_1^\pi = t^0/f_{pp\pi}(a_0/\sqrt{3}), & \alpha_2^\pi = 1.2785, & \alpha_3^\pi = 0.1383, & \alpha_4^\pi = 3.4490, \end{array}$$

where $t^0 = -3.08\,\text{eV}$ is the intralayer nearest-neighbour hopping energy and $t_\perp^0 = 0.27\,\text{eV}$ is the Bernal bilayer graphene nearest-neighbour hopping energy.

In the superconducting state we consider only the local interaction, in which case the $\mathbf{G}, \mathbf{G}'$-component of the BdG Hamiltonian reads

$$\mathcal{H}_{\rho\mathbf{k}}(\mathbf{G}, \mathbf{G}') = \begin{pmatrix} \mathcal{H}_{\rho\mathbf{k}}^0(\mathbf{G}, \mathbf{G}') & \Delta(\mathbf{G} - \mathbf{G}') \\ \Delta^*(\mathbf{G}' - \mathbf{G}) & -\mathcal{H}_{\rho\mathbf{k}}^0(\mathbf{G}, \mathbf{G}') \end{pmatrix}, \tag{S19}$$

where the matrix structure corresponds to the Nambu space, and the components of the superconducting order parameter $\Delta = \text{diag}(\Delta_{1A}, \Delta_{1B}, \Delta_{2A}, \Delta_{2B})$ are solved from the self-consistency equation

$$\Delta_{ls}(\mathbf{G}) = \lambda \sum_{\rho,b,\mathbf{G}'} \int_{\text{MBZ}} \frac{d\mathbf{k}}{(2\pi)^2} u_{\rho b\mathbf{k},ls}(\mathbf{G}') v_{\rho b\mathbf{k},ls}^*(\mathbf{G}' - \mathbf{G}) \tanh\left(\frac{E_{\rho b\mathbf{k}}}{2k_B T}\right). \tag{S20}$$

Here the band sum b is calculated over the positive energy bands, $l \in \{1, 2\}$ is the layer index, $s \in \{A, B\}$ is the sublattice index, and $|\psi_{\rho b\mathbf{k}}(\mathbf{G})\rangle = (|u_{\rho b\mathbf{k}}(\mathbf{G})\rangle, |v_{\rho b\mathbf{k}}(\mathbf{G})\rangle)^\mathsf{T}$ [in Nambu space] with $|u_{\rho b\mathbf{k}}(\mathbf{G})\rangle = (u_{\rho b\mathbf{k},1A}(\mathbf{G}), u_{\rho b\mathbf{k},1B}(\mathbf{G}), u_{\rho b\mathbf{k},2A}(\mathbf{G}), u_{\rho b\mathbf{k},2B}(\mathbf{G}))^\mathsf{T}$ and $E_{\rho b\mathbf{k}}$ are the eigenvectors and eigenenergies of the BdG equation

$$\sum_{\mathbf{G}'} \mathcal{H}_{\rho\mathbf{k}}(\mathbf{G}, \mathbf{G}') |\psi_{\rho b\mathbf{k}}(\mathbf{G}')\rangle = E_{\rho b\mathbf{k}} |\psi_{\rho b\mathbf{k}}(\mathbf{G})\rangle \quad \Leftrightarrow \quad \mathcal{H}_{\rho\mathbf{k}} |\psi_{\rho b\mathbf{k}}\rangle = E_{\rho b\mathbf{k}} |\psi_{\rho b\mathbf{k}}\rangle. \tag{S21}$$

The self-consistency equation is solved by the fixed point iteration scheme for a fixed chemical potential μ.

For the BdG Hamiltonian we can calculate the total number density for a given chemical potential μ from

$$n = 2 \sum_{\rho,b} \int_{\text{MBZ}} \frac{d\mathbf{k}}{(2\pi)^2} \left[\langle u_{\rho b\mathbf{k}}|u_{\rho b\mathbf{k}}\rangle f(E_{\rho b\mathbf{k}}) + \langle v_{\rho b\mathbf{k}}|v_{\rho b\mathbf{k}}\rangle (1 - f(E_{\rho b\mathbf{k}})) \right], \tag{S22}$$

where the factor of 2 comes from spin, f is the Fermi-Dirac distribution, and the band sum b is calculated over the positive energy bands.

II. CALCULATION OF THE SUPERFLUID WEIGHT

In this section we go through very briefly the essential equations to compute D^s in case of our RM and DP methods, show how D^s can be split to conventional and geometric terms, and discuss why our results for D^s are not in agreement with results of Ref. 22.

We compute D^s by using the linear response theory stated in Ref. 23. Therefore our starting point is the Fermi-Hubbard Hamiltonian of (S1). To probe the system current response, we apply a spatially slowly varying vector potential $\mathbf{A}$ via the Peierls substitution such that the hopping amplitude $t_{i\alpha j\beta} \equiv t_{ab}$ ($a \equiv i\alpha$, $b \equiv j\beta$) is rewritten as $t_{ab}(\mathbf{A}) = t_{ab}e^{-i\int_{\mathbf{r}_b}^{\mathbf{r}_a} \mathbf{A}(\mathbf{r}) \cdot d\mathbf{r}} \approx t_{ab}e^{-i\mathbf{A}(\mathbf{r}_{ab}^{\mathrm{CM}}) \cdot \mathbf{r}_{ab}^{\mathrm{rel}}}$ (we set the Planck constant and the elementary charge to unity, i.e. $\hbar = e = 1$), where $\mathbf{r}_{ab}^{\mathrm{CM}} = (\mathbf{r}_a + \mathbf{r}_b)/2$ and $\mathbf{r}_{ab}^{\mathrm{rel}} = \mathbf{r}_a - \mathbf{r}_b$. Then we expand the exponents up to second order so that our Hamiltonian becomes $H(\mathbf{A}) = H + \sum_\mu \sum_{ab} A_\mu(\mathbf{r}_{ab}^{\mathrm{CM}})j_\mu^p(a,b) + \frac{1}{2}\sum_{\mu\nu}\sum_{ab} A_\mu(\mathbf{r}_{ab}^{\mathrm{CM}})T_{\mu\nu}(a,b)A_\nu(\mathbf{r}_{ab}^{\mathrm{CM}})$. Here $j_\mu^p(a,b) = \sum_\sigma t_{ab}r_{ab,\mu}^{\mathrm{rel}}c_{a\sigma}^\dagger c_{b\sigma}$ is the paramagnetic current operator and $T_{\mu\nu}(a,b)A_\nu(\mathbf{r}_{ab}^{\mathrm{CM}}) = \sum_\sigma t_{ab}r_{ab,\mu}^{\mathrm{rel}}r_{ab,\nu}^{\mathrm{rel}}c_{a\sigma}^\dagger c_{b\sigma}A_\nu(\mathbf{r}_{ab}^{\mathrm{CM}})$ is the diamagnetic current operator. By using the expression for the total induced current density, $j_\mu(\mathbf{r}_{ab}^{\mathrm{CM}}) = -\delta H(\mathbf{A})/\delta A_\mu(\mathbf{r}_{ab}^{\mathrm{CM}})$, and linear response theory, we obtain in the momentum and frequency domain the relation $j_\mu(\mathbf{q}, \omega) = -K_{\mu\nu}(\mathbf{q}, \omega)A_\nu(\mathbf{q}, \omega)$, where $K_{\mu\nu}$ is the current-current response function of the form

$$K_{\mu\nu}(\mathbf{q}, \omega) = \langle T_{\mu\nu} \rangle - i \int_0^\infty dt e^{i\omega t} \langle [j_\mu^p(\mathbf{q}, t), j_\nu^p(-\mathbf{q}, 0)] \rangle, \tag{S23}$$

and

$$T_{\mu\nu} = \sum_{\mathbf{k}, \sigma} c_{\mathbf{k}\sigma}^\dagger \partial_\mu \partial_\nu \mathcal{H}_\sigma(\mathbf{k}) c_{\mathbf{k}\sigma}, \tag{S24}$$

$$j_\mu^p(\mathbf{q}) = \sum_{\mathbf{k}, \sigma} c_{\mathbf{k}\sigma}^\dagger \partial_\mu \mathcal{H}_\sigma(\mathbf{k} + \mathbf{q}/2) c_{\mathbf{k}+\mathbf{q}\sigma}, \tag{S25}$$

with $\partial_\mu \equiv \partial_{\mathbf{k}_\mu}$, are the diamagnetic and paramagnetic current parts, respectively.

The superfluid weight D^s is defined via the static Meissner effect ($\omega = 0$) and by taking the proper long wavelength limit of the transverse component of the current response function, see e.g. Refs. 23 and 24. In the mean-field level we can simply use the limit [24]

$$D_{\mu\nu}^s = \lim_{\mathbf{q} \to 0} \lim_{\omega \to 0} K_{\mu\nu}(\mathbf{q}, \omega). \tag{S26}$$

This definition is equivalent with the one defined via the change of free energy due to the phase twist applied to the superconducting order parameter which leads to the form of $D_{\mu\nu}^s \propto \frac{\partial^2 \Omega(\mathbf{A})}{\partial A_\mu \partial A_\nu}\big|_{\mathbf{A}=0}$, where Ω is the grand canonical potential [25].

Be deploying the mean-field theory and Green's function formalism, it was shown in Ref. 23 for local Hubbard interactions that Eq. (S26) leads to the following expression for D^s,

$$D_{\mu\nu}^s = \frac{1}{V}\sum_{\mathbf{k},i,j} \frac{f(E_{j\mathbf{k}}) - f(E_{i\mathbf{k}})}{E_{i\mathbf{k}} - E_{j\mathbf{k}}}(\langle\psi_{i\mathbf{k}}|\partial_\mu\mathcal{H}_{\mathbf{k}}|\psi_{j\mathbf{k}}\rangle\langle\psi_{j\mathbf{k}}|\partial_\nu\mathcal{H}_{\mathbf{k}}|\psi_{i\mathbf{k}}\rangle - \langle\psi_{i\mathbf{k}}|\partial_\mu\mathcal{H}_{\mathbf{k}}\tau_z|\psi_{j\mathbf{k}}\rangle\langle\psi_{j\mathbf{k}}|\tau_z\partial_\nu\mathcal{H}_{\mathbf{k}}|\psi_{i\mathbf{k}}\rangle), \tag{S27}$$

where the eigenstates and eigenenergies are solved from the BdG equation $\mathcal{H}_{\mathbf{k}}|\psi_{i\mathbf{k}}\rangle = E_{i\mathbf{k}}|\psi_{i\mathbf{k}}\rangle$, τ_z is a Pauli matrix acting in Nambu space, f is the Fermi-Dirac distribution, and V is the area of the sample. The difference quotient is interpreted as $-f'(E_{i\mathbf{k}})$ when $E_{i\mathbf{k}} = E_{j\mathbf{k}}$. In our TBG models the Hamiltonians are written in the superlattice-folded picture so that the $\mathbf{k}$ sum is over the moiré Brillouin zone (MBZ) and the i and j sums are over the bands enumerating the eigenstates for each $\mathbf{k}$.

In case of the local interaction used in Ref. 23, the order parameters do not have momentum dependence and thus the derivatives $\partial_\mu\mathcal{H}_{\mathbf{k}}$ are simply block diagonal matrices. However, for non-local interactions such as RVB used in our work, the order parameters depend on the momentum and thus the superfluid weight has a slightly different form,

$$D_{\mu\nu}^s = \frac{1}{V}\sum_{\mathbf{k},i,j} \frac{f(E_{j\mathbf{k}}) - f(E_{i\mathbf{k}})}{E_{i\mathbf{k}} - E_{j\mathbf{k}}}(\langle\psi_{i\mathbf{k}}|\partial_\mu\mathcal{H}_{\mathbf{k}}(\Delta = 0)|\psi_{j\mathbf{k}}\rangle\langle\psi_{j\mathbf{k}}|\partial_\nu\mathcal{H}_{\mathbf{k}}|\psi_{i\mathbf{k}}\rangle$$
$$- \langle\psi_{i\mathbf{k}}|\partial_\mu\mathcal{H}_{\mathbf{k}}(\Delta = 0)\tau_z|\psi_{j\mathbf{k}}\rangle\langle\psi_{j\mathbf{k}}|\tau_z\partial_\nu\mathcal{H}_{\mathbf{k}}(\Delta = 0)|\psi_{i\mathbf{k}}\rangle). \tag{S28}$$

The only difference compared to Eq. (S27) is the derivatives of the order parameters in the diamagnetic part. However, in our case the order parameters are always really small compared to the kinetic terms and therefore we can in

practice ignore extra terms arising from the derivatives of the order parameters. Therefore, in case of RM method we apply (S27) for both the local and RVB interaction schemes by taking $\partial_\mu \mathcal{H}_\mathbf{k} = \partial_\mu \mathcal{H}_\mathbf{k}(\Delta = 0)$.

In the DP model we assume that most of the contribution comes from states near the Dirac points, so that after writing everything in the valley-separated formalism, Eq. (S27) reads

$$D^s_{\mu\nu} = \frac{1}{V} \sum_{\rho,\mathbf{k},i,j} \frac{f(E_{\rho j \mathbf{k}}) - f(E_{\rho i \mathbf{k}})}{E_{\rho i \mathbf{k}} - E_{\rho j \mathbf{k}}} (\langle \psi_{\rho i \mathbf{k}} | \partial_\mu \mathcal{H}_{\rho \mathbf{k}} | \psi_{\rho j \mathbf{k}} \rangle \langle \psi_{\rho j \mathbf{k}} | \partial_\nu \mathcal{H}_{\rho \mathbf{k}} | \psi_{\rho i \mathbf{k}} \rangle - \langle \psi_{\rho i \mathbf{k}} | \partial_\mu \mathcal{H}_{\rho \mathbf{k}} \tau_z | \psi_{\rho j \mathbf{k}} \rangle \langle \psi_{\rho j \mathbf{k}} | \tau_z \partial_\nu \mathcal{H}_{\rho \mathbf{k}} | \psi_{\rho i \mathbf{k}} \rangle),$$

(S29)

where the eigenstates and eigenenergies are solved from the BdG equation (S21). In the case of the DP model we use Eq. (S29) to compute D^s.

A. The form of superfluid weight in the presence of C_3 rotational symmetry

It can be shown that the superfluid weight is isotropic in the presence of C_3 rotational symmetry. Let $\mathbf{e}_1$ be a unit vector, and $\mathbf{e}_2$ is obtained from $\mathbf{e}_1$ by a C_3 rotation, i.e., $\mathbf{e}_2 = R(\frac{2\pi}{3})\mathbf{e}_1$, and then $\mathbf{e}_3 = R(\frac{2\pi}{3})\mathbf{e}_2 = -\mathbf{e}_1 - \mathbf{e}_2$. We write the superfluid weight tensor by using the coordinate vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ as

$$D^s = D^s_{11}\mathbf{e}_1\mathbf{e}_1 + 2D^s_{12}\mathbf{e}_1\mathbf{e}_2 + D^s_{22}\mathbf{e}_2\mathbf{e}_2.$$

After a C_3 rotation, we get

$$C_3 D^s C_3^{-1} = D^s_{11}\mathbf{e}_2\mathbf{e}_2 + 2D^s_{12}\mathbf{e}_2(-\mathbf{e}_1 - \mathbf{e}_2) + D^s_{22}(-\mathbf{e}_1 - \mathbf{e}_2)(-\mathbf{e}_1 - \mathbf{e}_2),$$
$$= D^s_{22}\mathbf{e}_1\mathbf{e}_1 + 2(D^s_{22} - D^s_{12})\mathbf{e}_1\mathbf{e}_2 + (D^s_{11} + D^s_{22} - 2D^s_{12})\mathbf{e}_2\mathbf{e}_2.$$

Since $C_3 D^s C_3^{-1} = D^s$, we find $D^s_{11} = D^s_{22}$ and $D^s_{12} = D^s_{11}/2$. In terms of the Cartesian coordinates $\mathbf{e}_x$ and $\mathbf{e}_y$ ($\mathbf{e}_1 = \mathbf{e}_x$ and $\mathbf{e}_2 = -\frac{1}{2}\mathbf{e}_x + \frac{\sqrt{3}}{2}\mathbf{e}_y$), the superfluid weight becomes

$$D^s = D^s_{11}(\mathbf{e}_1\mathbf{e}_1 + \mathbf{e}_1\mathbf{e}_2 + \mathbf{e}_2\mathbf{e}_2),$$
$$= D^s_{11}\left[\mathbf{e}_x\mathbf{e}_x + \mathbf{e}_x\left(-\frac{\mathbf{e}_x}{2} + \frac{\sqrt{3}}{2}\mathbf{e}_y\right) + \left(-\frac{\mathbf{e}_x}{2} + \frac{\sqrt{3}}{2}\mathbf{e}_y\right)\left(-\frac{\mathbf{e}_x}{2} + \frac{\sqrt{3}}{2}\mathbf{e}_y\right)\right],$$
$$= \frac{3D^s_{11}}{4}(\mathbf{e}_x\mathbf{e}_x + \mathbf{e}_y\mathbf{e}_y),$$

(S30)

which is isotropic.

B. Geometric contribution and flat band superconductivity

As TBG is an extremely complicated multiband system, it is highly instructive to decompose D^s into the contributions of different one-particle Bloch states. We do this by using the method presented in Ref. 23, namely we expand the BdG states $|\psi_{i\mathbf{k}}\rangle$ in the basis of Bloch functions by writing

$$|\psi_{i\mathbf{k}}\rangle = \sum_{m=1}^{M} \left(w_{+,im} |+\rangle \otimes |m\rangle_\uparrow + w_{-,im} |-\rangle \otimes |m^*_-\rangle_\downarrow \right),$$

(S31)

where $|m\rangle_\uparrow$ [$|m^*_-\rangle_\downarrow$] is the eigenstate of $\mathcal{H}_\uparrow(\mathbf{k})$ [$\mathcal{H}_\downarrow^*(-\mathbf{k})$] with the eigenenergy $\epsilon_{\uparrow,m,\mathbf{k}}$ [$\epsilon_{\downarrow,m,-\mathbf{k}}$] and $|\pm\rangle$ are the eigenstates of τ_z with eigenvalues ± 1. As in our case we in practice always have $\partial_\mu \mathcal{H}_\mathbf{k} = \partial_\mu \mathcal{H}_\mathbf{k}(\Delta = 0)$, it is straightforward to rewrite D^s of Eq. (S27) in the following form:

$$D^s = 2 \sum_{\mathbf{k},i,j} \frac{f(E_{j\mathbf{k}}) - f(E_{i\mathbf{k}})}{E_{i\mathbf{k}} - E_{j\mathbf{k}}} \Bigg[\sum_{m_1,m_2} w^*_{+,im_1} w_{+,jm_2\uparrow} \langle m_1 | \partial_\mu \mathcal{H}_\uparrow(\mathbf{k}) | m_2 \rangle_\uparrow \sum_{m_3,m_4} w^*_{-,jm_3} w_{-,im_4\downarrow} \langle m^*_{-3} | \partial_\nu \mathcal{H}_\downarrow^*(-\mathbf{k}) | m^*_{-4} \rangle_\downarrow$$
$$+ (\mu \leftrightarrow \nu) \Bigg].$$

(S32)

We apply this expression when studying in Fig. 4 of the main text the superfluid weight by taking into account only the four flat bands or eight (4 flat, 4 dispersive) bands.

The matrix elements of the current operator can be further written as follows

$$[j_{\mu,\sigma}(\mathbf{k})]_{mn} = {}_\sigma \langle m| \, \partial_\mu \mathcal{H}_\sigma(\mathbf{k}) \, |n\rangle_\sigma = \partial_\mu \epsilon_{\sigma,m,\mathbf{k}} \delta_{mn} + (\epsilon_{\sigma,m,\mathbf{k}} - \epsilon_{\sigma,n,\mathbf{k}})_\sigma \, \langle \partial_\mu m| \, n\rangle_\sigma. \tag{S33}$$

From Eqs. (S32) and (S33) we see that there exist two different kinds of terms: the diagonal matrix elements of the current operator depend only on derivatives of the one-particle energy dispersions while the off-diagonal elements only on the momentum derivatives of the Bloch states. Thus D^s can be split into two terms: the conventional part D^s_{conv} that includes only the diagonal, i.e. intraband, current operator matrix elements ($m_1 = m_2$ and $m_3 = m_4$), and the geometric part D^s_{geom} that includes off-diagonal, i.e. interband, current operator matrix elements so that $D^s = D^s_{\text{conv}} + D^s_{\text{geom}}$. The conventional part consists purely of the intraband current terms and is therefore zero for a single exactly flat band (as $\partial_\mu \epsilon_{\sigma,m,\mathbf{k}} = 0$). Other way to see this is to note that D^s_{conv} is inversely proportional to the effective mass of the electrons [25] which for an exactly flat band is infinite. Therefore non-zero superconductivity of a flat band is always a multiband property involving interband current processes between the flat band and other bands, i.e. finite D^s_{geom}.

Roughly speaking, D^s_{conv} scales with the bandwidth, whereas D^s_{geom} scales with the interaction strength as larger interaction implies larger band mixing and thus more prominent interband processes. Therefore it is not surprising that we find a large geometric contribution in the flat band regime, as shown in the main text. Because of similar reasoning, it is understandable that interband terms between the flat and dispersive bands affecting D^s_{geom} for stronger interactions are important, and that at the flat band regime dispersive bands cannot be discarded when computing the total superfluid weight.

The importance of D^s_{geom} and the origin of the flat band superfluidity was for the first time addressed in Ref. 25 where generic multiband Hubbard models were studied at the mean-field level in the presence of local Hubbard interaction (characterized by the coupling strength J) and time-reversal symmetry. The authors of Ref. 25 considered an isolated flat band limit, i.e. the case where the Fermi surface lies within the flat band and interaction $|J|$ much smaller than the band gap E_{gap} between the flat and other bands, i.e. $|J| \ll E_{\text{gap}}$. As the other bands are well separated from the flat band and the interaction is weak enough not to considerably mix the bands, the Cooper pairing in practice takes place only within the flat band. One is then tempted to perceive this limit as a single-band problem for which D^s would be zero as D^s_{conv} is zero for a flat band and D^s_{geom} is zero for a single-band problem. However, from Eq. (S33), one can see that the geometric contribution actually scales as a function of E_{gap} and one has to be careful when taking the isolated flat band limit. It was shown in Ref. 25 that, indeed, the superconductivity of an isolated flat band is caused by the geometric superfluid weight term which, at low temperatures and with uniform local on-site pairing reads

$$D^s_{\text{geom},\mu\nu} \propto \Delta \int_{\text{B.Z.}} g^{\text{f.b.}}_{\mu\nu}(\mathbf{k}) \, \mathrm{d}\mathbf{k} \equiv \Delta \text{Re}[M^{\text{f.b.}}_{\mu\nu}], \text{ where} \tag{S34}$$

$$M^{\text{f.b.}}_{\mu\nu} \equiv \int_{\text{B.Z.}} B^{\text{f.b.}}_{\mu\nu}(\mathbf{k}) \, \mathrm{d}\mathbf{k}. \tag{S35}$$

Here $g^{\text{f.b.}}(\mathbf{k}) = \text{Re} \, \langle \partial_\mu n_{\text{f.b.}}(\mathbf{k})| \left(1 - |n_{\text{f.b.}}(\mathbf{k})\rangle \langle n_{\text{f.b.}}(\mathbf{k})|\right) |\partial_\nu n_{\text{f.b.}}(\mathbf{k})\rangle$ is the quantum metric of the flat band ($|n_{\text{f.b.}}\rangle$ are the Bloch states of the flat band, $\partial_\mu \equiv \partial/\partial k_\mu$) and $B^{\text{f.b.}}(\mathbf{k})$ is the corresponding quantum geometric tensor whose real (imaginary) part gives the quantum metric (Berry curvature) of the flat band [26]. Similar results can be obtained also without TRS [23]. Note that in (S34) $D^s \propto \Delta$, i.e. the superfluid weight is linearly proportional to the pairing amplitude in the isolated flat band limit. This is similar to the behaviour of D^s of TBG in the presence of local interaction as can be seen from Fig. 4 of the main text, according to which D^s grows linearly when $\max|\Delta| \gtrsim 2\,\text{meV}$, implying that D^s in this limit is dictated by the quantum metric.

An intriguing property of relation (S34) is the fact that it can be evaluated with the Bloch states of the flat band only, even if the geometric contribution is a multiband process involving the interband matrix elements of the current operator. This is because the influence of the other bands arises implicitly from the form of the flat band Bloch states which are defined by the geometric properties of the quantum states of the whole lattice structure. It should be also emphasized that the existence of the other bands are required even though the Cooper pairing essentially takes place within the flat band and the pairing within the other bands is very small compared to the flat band. This can be reflected to TBG, where in the flat band regime of $\max|\Delta| \gtrsim 2\,\text{meV}$ the existence of the dispersive bands have to be taken into account even though the Cooper pairing occurs predominantly within the narrow bands only. Furthermore, even though the geometric contribution D^s_{geom} consists of interband current terms, it does not require the interband pairing to be nonzero. Actually, in case of the Lieb lattice, the interband terms are in practice vanishingly small but D^s_{geom} is large [27]. The same applies to the TBG computations of this work: interband order parameters are in general negligible compared to the intraband order parameters but D^s_{geom} can nevertheless be the dominant contribution.
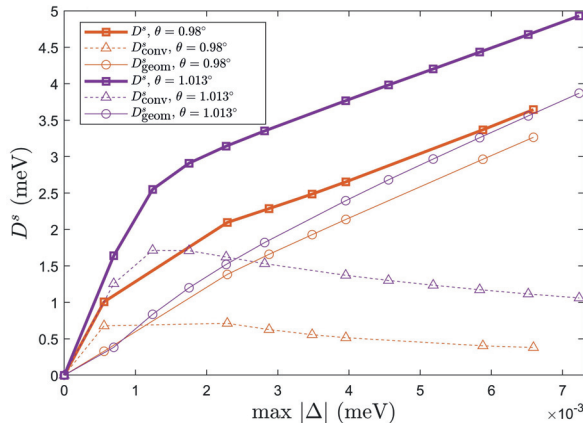
FIG. S4. D^s, D^s_{geom}, and D^s_{conv} computed with RM method for two twist angles, $\theta = 0.98°$ (red lines) and $\theta = 1.013°$ (purple lines), at $\nu \approx -2$ and $T = 1.5$ K as a function of the pairing strength in case of the local interaction. The results for $\theta = 1.013°$ are the same as those presented in Fig. 4(b) of the main text.

To give an intuitive picture to Eq. (S34) and the superfluidity of the flat bands in general, let us consider the quantum metric in a more general footing. To this end, let us introduce the infinitesimal distance between the quantum states of the nth energy band as follows

$$D_n^2(\mathbf{k}, \mathbf{k} + \mathrm{d}\mathbf{k}) = 1 - |\langle n(\mathbf{k})|n(\mathbf{k} + \mathrm{d}\mathbf{k})\rangle|^2, \tag{S36}$$

which reaches its maximum value of unity when the states are not overlapping at all. One can define the quantum metric $g_{\mu\nu}^n(\mathbf{k})$ by expanding (S36) as

$$D_{n,\mu\nu}^2(\mathbf{k}, \mathbf{k} + \mathrm{d}\mathbf{k}) = \sum_{\mu\nu} \frac{1}{2} g_{\mu\nu}^n(\mathbf{k}) \, \mathrm{d}\mathbf{k}_\mu \, \mathrm{d}\mathbf{k}_\nu \,. \tag{S37}$$

Here the higher order terms are ignored. Thus the quantum metric is related to the overlap of the quantum states. For example, in Ref. [25] it was shown that the superfluidity of a topologically non-trivial flat band is always positive as for a topological Bloch band one cannot construct the Wannier functions to be maximally localized with exponentially decaying amplitude [25, 28]. This implies finite delocalization and thus finite overlap between the Wannier functions. Therefore, superfluidity of, at least topological, flat bands can be explained by finite overlap of the Wannier functions which allows finite current transport.

However, also topologically trivial flat bands can support supercurrent which is related to the finite quantum metric. To understand this, in Ref. 29 flat band superconductivity was approached from a different angle, namely via the two-body problem. Usually, in a non-flat band system, the Fermi sea is unstable towards the formation of bound pairs. In a flat band there does not exist a well-defined Fermi surface due to the degenerate states of the flat band. If this degeneracy is preserved in the presence of interactions, the existence of a bound state is not enough for superconductivity as then the condensation of Cooper pairs to a certain momentum state and the formation of a coherent superconducting state is not probable. In other words, in the presence of degenerate bound states, the effective mass of the Cooper pairs is infinite. The authors of Ref. 29 showed that, in the presence of interactions, the degeneracy can be lifted and the mass of the Cooper pairs can be finite when the Bloch states of the flat band have finite overlap and when their spatial derivatives are non-zero. Importantly, in case of the uniform pairing, the condition of overlapping quantum states reduces to the quantum metric results of Refs. 23 and 25. This connection relates the quantum metric directly to finite overlapping of the wavefunctions of the flat band. Moreover, the quantum metric is also present in a localization functional that describes the spread of the Wannier functions [25]. Finite quantum metric integral bounds the functional from below and thus implies finite spread and hence finite overlap between the Wannier functions.

In addition to the present work, the role of the geometric superfluid weight in case of TBG was also highlighted in Ref. 30 by using a TBG continuum model developed in Ref. 31 where also other dispersive bands are taken into

account. The main finding of Ref. 30 was that when the bandwidth of the narrow bands are minimized at the magic twist angle, the geometric contribution D^s_{geom} is larger than the conventional term D^s_{conv} for the interaction strength used in Ref. 30. Correspondingly, by tuning the twist slightly off from the magic angle, the conventional term emerges as the main contribution, i.e. $D^s_{\text{conv}} > D^s_{\text{geom}}$. The conventional term was shown to depend heavily on the twist angle, whereas the geometric part was shown to be less sensitive to the twist. This is understandable as D^s_{conv} depends strongly on the bandwidth of the narrow bands, decreasing as the bandwidth becomes smaller. Furthermore, as the interaction strength chosen in Ref. 30 was rather small, T_{BKT} being around $0 - 2\,\text{K}$, it is easy to comprehend that conventional term can dominate over the geometric one, as can be also seen e.g. from Figs. 4(a) and 4(b) of the main text where, at the regime of $T_{\text{BKT}} \sim 0 - 2\,\text{K}$, D^s_{conv} indeed is prominent. To see whether the RM method yields similar behavior, D^s, D^s_{geom}, and D^s_{conv} computed with RM are plotted in Fig. S4 for two different twist angles, namely $\theta = 1.013°$ and $\theta = 0.98°$ as a function of the pairing strength at $\nu \approx -2$ in case of the local interaction. The bandwidth of the narrow band structure is slightly smaller for $\theta = 0.98°$ and, indeed, one can see from Fig. S4 that D^s_{conv} is considerably smaller for $\theta = 0.98°$, whereas D^s_{geom} is more or less the same for both angles. One can further see that there exists a weak-coupling regime where $D^s_{\text{geom}} > D^s_{\text{conv}}$ at $\theta = 0.98°$ but $D^s_{\text{geom}} < D^s_{\text{conv}}$ at $\theta = 1.013°$, reflecting the weak-coupling results of Ref. 30.

In Ref. 32 a geometric lower bound for D^s of TBG was derived in the absence of the dispersive bands and by assuming exactly flat bands. This lower bound was found to be proportional to the so-called Wilson loop winding number of the flat bands; a result that can be taken as yet another way to bound D^s from below by the geometric properties of the quantum states, in addition to the bounds defined by the Chern number [25] and the Berry curvature [23]. As the dispersive bands in Ref. 32 were ignored, the geometric contribution coming from the interband current terms between the flat and dispersive bands was not considered and the weak coupling limit was assumed. Thus the results of Ref. 32 cannot directly be related to the results presented in our work and in Ref. 30, but all three works present strong arguments that geometric properties of the quantum states play a significant role in superconductivity of TBG. Particularly, our results highlight the necessity to consider the geometric contribution, especially in case of stronger pairing interactions, when calculating the superfluid weight of TBG and other twisted multilayer systems. Therefore, being realizable in experiments, TBG can be potentially very important in terms of accessing and measuring the geometric contribution experimentally.

C. Comparison to the "upper" limit of D^s derived in Hazra *et al.*

In Ref. 22 the upper limit of D^s and T_{BKT} were computed for TBG system. The authors of Ref. 22 concluded giving an upper limit estimate of $D^s_{\text{max}} \sim 1.5\,\text{meV}$ (in our units, note that the superfluid weight definition of Ref. 22 differs from our definition by a factor of four), *regardless* of the interaction mechanism or the interaction strength. Their estimate clearly contradicts with the results obtained by our two different models. The explanation for this disagreement is the use of oversimplified approximations in Ref. 22. We go here briefly through their arguments for achieving the upper limit of D^s and we argue why their upper limit for D^s is not valid for arbitrary interaction strengths or mechanism.

The first important point is that the authors of Ref. 22 deploy an effective model, developed in Ref. 33, that consists of only four flat bands. However, we showed in Fig. 4 of the main text that especially for strong interaction strengths the geometric contribution D^s_{geom} arising from the off-diagonal matrix elements of the current operator [See Eq. (S33)] between the flat and dispersive bands is the most prominent part of the total superfluid weight. But the model used in Ref. 22 consists only of the flat bands, with dispersive bands being absent. As there are no dispersive bands implemented in their model, there cannot be any geometric contribution coming from the interband terms between flat and dispersive bands. Hence, the claim stating that the upper limit for T_{BKT} derived in Ref. 22 holds for arbitrary interaction strength or interaction mechanism is readily shown to be invalid. This is not surprising: if the interaction strength is large enough, the dispersive bands become involved to the superconducting pairing, which is manifested by our results in Fig 4. of the main text.

To further highlight that the upper limit of Ref. 22 works only on the weak coupling regime, let us write down their argument. The starting point is the expressions (S23)–(S26) which can be rewritten as

$$D^s_{\mu\nu} = D^s_{\mu\nu,\text{dia}} + D^s_{\mu\nu,\text{para}}, \tag{S38}$$

where $D^s_{\mu\nu,\text{dia}} = \langle T_{\mu\nu} \rangle$ is the diamagnetic part and correspondingly the paramagnetic contribution is $D^s_{\mu\nu,\text{para}} = \lim_{\mathbf{q}\to 0}\lim_{\omega\to 0}[-i\int_0^\infty dt e^{i\omega t}\langle[j^p_\mu(\mathbf{q},t), j^p_\nu(-\mathbf{q},0)]\rangle]$. It can be shown that the paramagnetic part is always zero or negative, thus it follows that $D^s_{\mu\nu,\text{dia}} \geqslant D^s_{\mu\nu}$ (usually in multiband systems the absolute values of dia- and paramagnetic parts are much larger than the absolute value of D^s). Therefore, the argument used by the authors of Ref. 22 is to compute the diamagnetic part $D^s_{\mu\nu,\text{dia}}$ to obtain the upper limit for the total superfluid weight D^s.

It is straightforward to rewrite the diamagnetic term in the following form,

$$D^s_{\mu\nu,\text{dia}} = \langle T_{\mu\nu}\rangle = \sum_{m,m',\mathbf{k},\sigma} M^{-1}_{mm'}(\mathbf{k},\sigma)\langle c^\dagger_{\mathbf{k}\sigma m} c_{\mathbf{k}\sigma m'}\rangle, \tag{S39}$$

where the inverse mass tensor is given by $M^{-1}_{mm'}(\mathbf{k},\sigma) = [U^\dagger(\mathbf{k})\partial_\mu\partial_\nu\mathcal{H}_\sigma(\mathbf{k})U(\mathbf{k})]_{mm'}$. Here the columns of $U(\mathbf{k})$ are the one-particle Bloch states and $c_{\mathbf{k}\sigma m'}$ is the annihilation operator for the Bloch state in the mth Bloch band of momentum $\mathbf{k}$ and spin σ.

Now let us consider a situation where we are at the hole doping regime. The authors of Ref. 22 in this case assume that the two flat bands in the electron-doped side are empty. This is already an implicit assumption about the weak-coupling regime: for stronger interaction there exists finite electron occupation also in the upper flat bands, as can be seen in Fig. S3(c) for example. Due to the Cauchy-Schwarz inequality $\langle c^\dagger_{\mathbf{k}\sigma m} c_{\mathbf{k}\sigma m'}\rangle < \sqrt{n_{\mathbf{k}\sigma m} n_{\mathbf{k}\sigma m'}}$, where $n_{\mathbf{k}\sigma m} = \langle c^\dagger_{\mathbf{k}\sigma m} c_{\mathbf{k}\sigma m}\rangle$, one can deduce then that $\langle c^\dagger_{\mathbf{k}\sigma m} c_{\mathbf{k}\sigma m'}\rangle = 0$ if the band index m or m' refer to one of the two upper flat bands. There can still exist off-diagonal term if both m and m' refer to two hole-doping regime flat bands but also these off-diagonal elements are in Ref. 22 discarded. Thus the authors ignore the interband terms and end up having the form $D^s_{\mu\nu,\text{dia}} = \sum_{m,\mathbf{k},\sigma} M^{-1}_{mm}(\mathbf{k},\sigma)\langle n_{\mathbf{k}\sigma m}\rangle$. The occupation expectation value is then evaluated by assuming the step function $\langle n_{\mathbf{k}\sigma m}\rangle = \Theta(\mu - \epsilon_m(\mathbf{k}))$, i.e. by assuming the zero temperature and non-interacting limit for the occupation numbers. Therefore, their final upper bound for the superfluid weight of TBG system is $D^s_{\max} = \sum_{m,\mathbf{k},\sigma} M^{-1}_{mm}(\mathbf{k},\sigma)\Theta(\mu - \epsilon_m(\mathbf{k}))$ and with this expression the authors obtain Fig. 1 shown in Ref. 22, where the largest value for the superfluid weight (in our units) is around $\sim 0.15\,\text{meV}$. This is of the same order of magnitude than our weak-coupling results at very low temperatures, see for example Fig. 3(e) of the main text. This is not surprising as essentially all the assumptions done in Ref. 22 implicitly require weak interaction strengths. Thus rather than calling it the generic upper limit for D^s, the result of Ref. 22 can be taken as a reasonable estimate for D^s in case of weak interactions. Note that this estimate is close to the superfluid weight of the RVB case shown in Fig. 4 of the main text but those results are evaluated at considerably higher temperatures of $T \approx 1.5\,\text{K}$.

III. SUPERFLUID WEIGHT AND T_{BKT} AT THE CHARGE NEUTRALITY POINT

In the main text we provided T_{BKT} and $T_{\text{BKT}}/\max|\Delta|$ as function of the pairing strength at half-filling of the hole-doping flat band regime, i.e. at $\nu \approx -2$. Here we provide, for completeness, similar plots for the case $\nu = 0$, i.e. at the charge neutrality point.

In Fig. S5 we show, as a function of $\max|\Delta|$, T_{BKT} [Fig. S5(a)] and the ratio $T_{\text{BKT}}/\max|\Delta|$ [Fig. S5(b)] for $\nu = 0$. For comparison, also the results of $\nu = -2$ are shown. One can see that both quantities behave very similarly for both fillings and especially in case of local interaction, $T_{\text{BKT}}/\max|\Delta|$ seems to be rather independent on the filling in the flat band regime.

IV. SUPERFLUID WEIGHT AS A FUNCTION OF THE RENORMALIZATION STRENGTH

To demonstrate the validity of the renormalization scheme, in Fig. S6 we plot D^s for both local and RVB interaction schemes at $\nu \approx 0$ and $T = 1.5\,\text{K}$ for three different renormalization strengths with $M = 676$, $M = 868$, and $M = 1324$, where M is the number of lattice sites per unit cell. In all the calculations shown in the main text one has $M = 676$. We see that the results remain more or less the same when M is increased, i.e. when the strength of the renormalization is decreased. Thus, using the renormalization of $M = 676$ in the computations presented in the main text is justified. One should also note that the DP method (which has nothing to do with the renormalization method) yields similar results for D^s than the RM method in case of the local interaction, as one can see from the results shown in the main text.

Heuristically, the renormalization method, yielding smaller amount of lattice sites within the moiré unit cell, can be thought as a coarse-grained model which has less degrees of freedom than the full microscopic model but which still features the same physics as the full model. The success of the renormalization method is not surprising as its main effect is to increase the relative strength of the interlayer coupling with respect to the intralayer coupling. Therefore, the rescaled model is still a twisted bilayer system which just has stronger interlayer coupling. Stronger interlayer hopping means that one can obtain the flat band dispersions with larger twist angles (and therefore with smaller M) than with a system of smaller interlayer coupling. The idea is the same as in the experiment conducted by Yankowitz et al. [34] where the interlayer coupling was increased by hydrostatic pressure and thus the magic angle regime was reached for larger angles than in original experiments by Cao et al. [35] where the pressure was not applied.
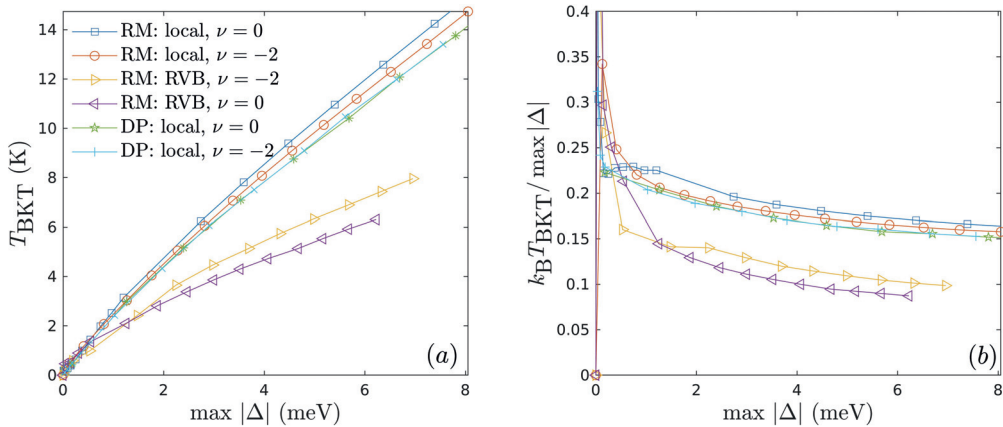
FIG. S5. (a) T_{BKT} as a function of $\max|\Delta(T=0)|$ at $\nu=0$ and $\nu=-2$. (b) The corresponding results for $T_{\mathrm{BKT}}/\max|\Delta(T=0)|$.
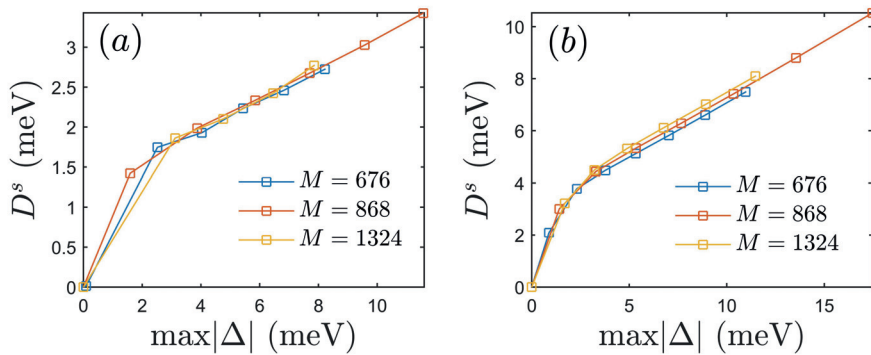

FIG. S6. D^s for three different renormalization strengths with $M=676$, $M=868$, and $M=1324$ in case of (a) RVB and (b) local interaction at $\nu\approx 0$ and $T=1.5\,\mathrm{K}$.

[1] J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954).
[2] B. Uchoa and A. H. Castro Neto, Phys. Rev. Lett. **98**, 146801 (2007).
[3] N. B. Kopnin and E. B. Sonin, Phys. Rev. Lett. **100**, 246808 (2008).
[4] E. Zhao and A. Paramekanti, Phys. Rev. Lett. **97**, 230404 (2006).
[5] M. V. Hosseini, EPL **110**, 47010 (2015).
[6] T. J. Peltonen, R. Ojajärvi, and T. T. Heikkilä, Phys. Rev. B **98**, 220504 (2018).
[7] F. Wu, A. H. MacDonald, and I. Martin, Phys. Rev. Lett. **121**, 257001 (2018).
[8] G. Baskaran, Phys. Rev. B **65**, 212505 (2002).
[9] A. M. Black-Schaffer and S. Doniach, Phys. Rev. B **75**, 134512 (2007).
[10] S. Pathak, V. B. Shenoy, and G. Baskaran, Phys. Rev. B **81**, 085431 (2010).
[11] A. M. Black-Schaffer, Phys. Rev. Lett. **109**, 197001 (2012).
[12] B. Roy and I. F. Herbut, Phys. Rev. B **82**, 035429 (2010).
[13] C. Honerkamp, Phys. Rev. Lett. **100**, 146404 (2008).
[14] Y. Su and S.-Z. Lin, Phys. Rev. B **98**, 195101 (2018).

[15] A. M. Black-Schaffer and C. Honerkamp, J. Phys. Condens. Matter **26**, 423201 (2014).
[16] P. A. Lee, N. Nagaosa, and X.-G. Wen, Rev. Mod. Phys. **78**, 17 (2006).
[17] L. Pauling, *Nature of the Chemical Bond* (Cornell University Press, 1960).
[18] M. L. Kiesel, C. Platt, W. Hanke, D. A. Abanin, and R. Thomale, Phys. Rev. B **86**, 020507 (2012).
[19] L. A. Gonzalez-Arraga, J. L. Lado, F. Guinea, and P. San-Jose, Phys. Rev. Lett. **119**, 107201 (2017).
[20] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, Phys. Rev. Lett. **99**, 256802 (2007).
[21] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, Phys. Rev. B **86**, 155449 (2012).
[22] T. Hazra, N. Verma, and M. Randeria, (2018), arXiv:1811.12428.
[23] L. Liang, T. I. Vanhala, S. Peotta, T. Siro, A. Harju, and P. Törmä, Phys. Rev. B **95**, 024515 (2017).
[24] D. J. Scalapino, S. R. White, and S. Zhang, Phys. Rev. B **47**, 7995 (1993).
[25] S. Peotta and P. Törmä, Nat. Commun. **6**, 8944 (2015).
[26] M. V. Berry, in *Geometric phases in physics*, edited by A. Shapere and F. Wilczeck (World Scientific, 1989) pp. 7–28.
[27] A. Julku, S. Peotta, T. I. Vanhala, D.-H. Kim, and P. Törmä, Phys. Rev. Lett. **117**, 045303 (2016).
[28] C. Brouder, G. Panati, M. Calandra, C. Mourougane, and N. Marzari, Phys. Rev. Lett. **98**, 046402 (2007).
[29] P. Törmä, L. Liang, and S. Peotta, Phys. Rev. B **98**, 220511 (2018).
[30] X. Hu, T. Hyart, D. I. Pikulin, and E. Rossi, (2019), arXiv:1906.07152.
[31] R. Bistritzer and A. H. MacDonald, Proc. Natl. Acad. Sci. U.S.A. **108**, 12233 (2011).
[32] F. Xie, Z. Song, B. Lian, and B. A. Bernevig, (2019), arXiv:1906.02213.
[33] M. Koshino, N. F. Q. Yuan, T. Koretsune, M. Ochi, K. Kuroki, and L. Fu, Phys. Rev. X **8**, 031087 (2018).
[34] M. Yankowitz, S. Chen, H. Polshyn, Y. Zhang, K. Watanabe, T. Taniguchi, D. Graf, A. F. Young, and C. R. Dean, Science **363**, 1059 (2019).
[35] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Nature **556**, 43 (2018).