RESEARCH ARTICLE | NOVEMBER 01 2023

Towards structural optimization of gold nanoclusters with quantum Monte Carlo

Juha Tiihonen ; Hannu Häkkinen

J. Chem. Phys. 159, 174301 (2023)
https://doi.org/10.1063/5.0174383

The Journal of Chemical Physics ARTICLE pubs.aip.org/aip/jcp

Towards structural optimization of gold nanoclusters with quantum Monte Carlo

Cite as: J. Chem. Phys. 159, 174301 (2023); doi: 10.1063/5.0174383
Submitted: 30 August 2023 • Accepted: 9 October 2023 •
Published Online: 1 November 2023

Juha Tiihonen[1,a] and Hannu Häkkinen[1,2]

AFFILIATIONS

[1] Department of Physics, Nanoscience Center, University of Jyväskylä, Jyväskylä, Finland
[2] Department of Chemistry, Nanoscience Center, University of Jyväskylä, Jyväskylä, Finland

[a] Author to whom correspondence should be addressed: tiihonen@iki.fi

ABSTRACT

We study the prospects of using quantum Monte Carlo techniques (QMC) to optimize the electronic wavefunctions and atomic geometries of gold compounds. Complex gold nanoclusters are widely studied for diverse biochemical applications, but the dynamic correlation and relativistic effects in gold set the bar high for reliable, predictive simulation methods. Here we study selected ground state properties of few-atom gold clusters by using density functional theory (DFT) and various implementations of the variational Monte Carlo (VMC) and diffusion Monte Carlo. We show that the QMC methods mitigate the exchange-correlation (XC) approximation made in the DFT approach: the average QMC results are more accurate and significantly more consistent than corresponding DFT results based on different XC functionals. Furthermore, we use demonstrate structural optimization of selected thiolated gold clusters with between 1 and 3 gold atoms using VMC forces. The optimization workflow is demonstrably consistent, robust, and its computational cost scales with n^b, where $b < 3$ and n is the system size. We discuss the implications of these results while laying out steps for further developments.

I. INTRODUCTION

Gold nanoparticles enjoy scientific and technological intrigue. Abundant and non-toxic, gold compounds are commonly used in catalysis,[1,2] biomedicine,[3,4] chemical sensing,[5] light harvesting,[6] and numerous other applications. Gold is a noble element[7] with special characteristics, such as aurofilic interaction[8] and strong relativistic effects,[9] which set the bar high for theoretical understanding.

The contemporary research on gold nanostructures leans heavily on numerical calculation methods, which aid the synthesis and engineering of technologies in all sizes.[10] The smallest of gold nanostructures gain from accurate first-principles calculation, which employ the quantum mechanical understanding of particular atomic configurations and their properties.[11–13] Prediction and synthesis of atomically precise gold nanoclusters, with all the prospects of rational engineering of superatoms,[14,15] stems from this concept. Moving on studies of large and complex structures and their thermal phenomena[16–18] is possible through force fields, at the expense of some theoretical rigor.

This work concerns the highly accurate first-principles calculation of the quantum mechanical electronic structure of gold. Over decades, the celebrated workhorse of applied calculations of gold nanoclusters (and others) has been the density functional theory (DFT).[19] The DFT calculation has successfully predicted and explained crystal structures,[20] spectroscopy,[21] optical response,[22–24] and many other properties that can substantially inform the work done in laboratories.[25,26] The success of DFT lies in simplicity and high performance, both numerically and quantitatively. Yet, the exchange-correlation (XC) functional in the heart of DFT remains an uncontrollable approximation, a seminal liability, which cannot be assessed or controlled intrinsically.[27] For this reason, numerous DFT studies exist[22–24] that benchmark and propose informed choices of the XC functionals, to grapple with their systematic errors. For instance, the local-density approximation (LDA) functional appears accurate for predicting Au–Au bonds,[23] likely due to fortuitous error cancellation. Other studies praise generalized-gradient approximation (GGA) functionals for obtaining structural and optical properties, or using LDA together with an asymptotically correct time-dependent DFT (TDDFT) functional.[22,24,28] This hidden knowledge combines experimental validation and rational assumptions of transferability of the properties of given XC functionals to new applications.

In this work, we demonstrate that the XC dependency of DFT in selected gold clusters (Au_n and $MeSAu_n$, where $n = 1, \ldots, 3$ and Me is CH_3) can be reliably controlled by using the quantum Monte Carlo (QMC) approach.[29] Monitoring DFT with QMC is not uncommon (e.g. Refs. 30 and 31) nor surprising, but to our knowledge, very few QMC studies exist on small coinage metal clusters (Ref. 32). Here we survey the variation of selected energetic and structural properties based on selected self-consistent field (SCF) methods and QMC, upon using Hartree–Fock (HF) or DFT with various XC functionals. We show that the QMC methods eliminate the XC dependency by an order of magnitude, and that the computational cost follows similar n^3 scaling with the system size n in both SCF and QMC, although the QMC is 2–3 orders of magnitude more expensive. Especially, our focus on structural optimization addresses an open challenge within the QMC community: numerous molecular ground and excited state structures and other applications of variational Monte Carlo (VMC) forces have been demonstrated with up to second row elements (e.g., Refs. 33 and 34), relaxation of transition metal compounds and beyond remains an uncharted territory. With transition metals, there is anticipation of difficulties in controlling sources of bias and statistical noise in QMC force estimators,[35,36] which we will encounter and discuss here in practical terms. We use variational Monte Carlo (VMC) and fixed-node diffusion Monte Carlo (DMC)[29] with very conventional single-determinant Slater-Jastrow wavefunctions. We use recent, high-accuracy pseudopotentials[37] in the scalar-relativistic formalism. This leaves many sophisticated techniques up for renewal, including multi-determinant wavefunctions,[38] two-component spinor.[39]

The remainder of the paper is organized as follows: In Sec. II we briefly outline the theoretical background of the numerical problems and different SCF and QMC formulations that are relevant for this work. In Sec. III we describe the respective implementation and other computational details. In Sec. IV we present and discuss results pure gold clusters, and in Sec. V thiolated gold compounds. In Sec. VI we discuss the prospects of numerical performance based on the physical results. In Sec. VII we summarize the work and discuss next steps.

II. METHODS

We consider different numerical theories for solving the non-relativistic first-principles electronic structure problem of N ions and n electrons, stated by the Schrödinger equation in the Born–Oppenheimer approximation for a given ionic geometry R:

$$H(R)|\Psi_i(R)\rangle = E_i(R)|\Psi_i(R)\rangle, \tag{1}$$

where, H, Ψ_i and E_i are, respectively, the Hamilton operator and its eigenvalues and eigenfunctions for a state i. Here we only consider scalar-relativistic calculation of the electronic structure, where the average relativistic effects are embedded[40,41] in the pseudopotential.[37]

Following from Eq. (1), the potential energy surface (PES) for a given state i is defined as

$$E_i(\mathbf{R}) = \langle \Psi_i(\mathbf{R})|H(\mathbf{R})|\Psi_i(\mathbf{R})\rangle. \tag{2}$$

Points of interest are the equilibrium geometries $\mathbf{R}_i^0$ that locally minimize the PES:

$$\mathbf{R}_i^0 = \arg\min\,[E_i(\mathbf{R})]. \tag{3}$$

The PES minimum is the best prediction of the equilibrium structure, in the absence of nuclear degrees of freedom. Thus, geometry optimization means finding the nearest local minimum of the PES (although there could be many globally), so relaxing an initial $\mathbf{R}$ to $\mathbf{R}^i$. Typical way of proceeding is minimizing the physical unsigned forces on all ions, that can be computed from the gradients of the wavefunctions. From here on we will lose the state number i and only consider the electronic ground state corresponding to $i = 0$.

To evaluate $E(\mathbf{R})$ from Eq. (2), we need to solve $|\Psi(\mathbf{R})\rangle$ from Eq. (1) using a numerically tractable method. For a given $\mathbf{R}$, the approach in this work is to solve the energy-minimizing wavefunction with a single-determinant Slater–Jastrow ansatz

$$\Psi_T(\mathbf{r}) = J(\mathbf{r})S^{\uparrow}(\mathbf{r})S^{\downarrow}(\mathbf{r}), \tag{4}$$

where $S^{\uparrow\downarrow}(\mathbf{r})$ are the up/down are Slater determinants for the respective spin species, $J(\mathbf{r})$ is a Jastrow correlation factor and $\mathbf{r}$ denotes the set of electronic coordinates. The Slater determinants $S^{\uparrow\downarrow} = \det(D^{\uparrow\downarrow})$ are taken over population coefficients in a given basis, such as a linear combination of atomic orbitals (LCAO). They can be obtained from the SCF approaches, which treat the electronic interaction based on a mean field. Here, we employ two SCF theories, namely Restricted (open-shell) Hartree–Fock (RHF/ROHF) and Kohn-Sham DFT. ROHF is used whenever the up/down spin populations are non-symmetrical. The Jastrow factor is typically expanded in terms of particle distances between distinguishable species:[42,43]

$$J(\mathbf{r}) = \sum_{ne} J_1(\mathbf{r}) + \sum_{ee} J_2(\mathbf{r}) + \sum_{nee} J_3(\mathbf{r}) + \cdots, \tag{5}$$

where ne, ee and nee refer, respectively, to nucleus-electron, electron-electron and nucleus-electron-electron correlations. Furthermore, each of J_a terms, where $a = 1, \ldots, 3$, can be broken down to two according to the electron spin alignments: nu/nd, uu/ud and nuu/nud, where u/d refer to spin up/down and uu equals dd. Higher order terms are not considered here. Different PESs ensue depending on which ansatz degrees of freedom are simultaneously optimized, for instance, whether the Jastrow factor is optimized up to J_2 or J_3, or whether $S^{\uparrow}(\mathbf{r})$ are re-optimized in the presence of $J(\mathbf{r})$.

The correlated trial wavefunction $\Psi_T(\mathbf{r})$ cannot be integrated by conventional means, but it can be evaluated by random sampling:

$$E^{\mathrm{VMC}} = \frac{\langle \Psi_T|E_L|\Psi_T\rangle}{\langle \Psi_T|\Psi_T\rangle} = \langle E_L\rangle_P \tag{6}$$

where

$$E_L(r) = \frac{H\Psi_T(r)}{\Psi_T(r)} \tag{7}$$

is the so-called local energy, and where importance sampling of r_p can be drawn from the distribution $P(r) = |\Psi_T(r)|^2/\int|\Psi_T(r)|^2$ efficiently by using the Metropolis algorithm. The coefficients of the

wavefunction Ψ_T can be optimized to match the ground state by minimization of $\langle E_L \rangle_P$, which is variational, or its variance.[29] The estimates of VMC energy are subject to finite noise, which can be controlled by adjusting the finite number of samples.

The VMC ionic forces can be estimated from sampling the analytical gradients ($d/d\lambda \equiv d_\lambda$, where λ is an ionic coordinate) of the trial wavefunction:[44]

$$d_\lambda E = \langle d_\lambda E_L + (E_L - E) d_\lambda \ln P \rangle_P, \tag{8}$$

where E is an estimate of the total energy. The estimator faces challenges at nodes, where $d_\lambda \Psi_T(r)/\Psi_T(r)$ diverges, leading to an infinite variance problem for the estimator.[45] The problem can be treated by sampling a regularized function.[44,46,47] To significantly improve the statistical performance of the estimator, a space-warp transformation is needed.[36,48] The wavefunction gradients also call for extra computation and cost-efficient algorithms.[49] Furthermore, it is important to calculate all terms that arise from $d_\lambda \ln P$, namely all the derivatives of wavefunction coefficients, because neglecting them induces estimator bias. This means that all calculation parameters contributing to P, including the orbital coefficients, must be simultaneously optimized,[35,50] which is done in the process of orbital rotations.[51–53]

Fixed-node Diffusion Monte Carlo (DMC) is a projector method, whose PES is given by

$$E^{DMC} = \frac{\langle \Psi_T | E_L | \Psi_0 \rangle}{\langle \Psi_T | \Psi_0 \rangle}, \tag{9}$$

where Ψ_0 refers to the projected ground state, given the fixed nodal surface of Ψ_T. The projection to Ψ_0 can be done according to

$$\Psi_0 = \lim_{\tau \to \infty} e^{-\tau H} |\Psi\rangle, \tag{10}$$

which eventually propagates any state Ψ into Ψ_0 that is not orthogonal to it. In practice, the propagator can be imposed to a walker population using a finite time-step $\Delta\tau$, leaving all parameters controllable for systematic accuracy. Generally, the sources of numerical error for DMC can be controlled to a high degree. The most significant practical challenge remains the fixed-node error originating from the SCF trial wavefunction. The choice of nodes conforms to the variational principle, meaning better nodes are identified by lower energy, but it also affects the evaluation of non-local projectors that are prevalent in the modern effective core potentials (ECPs).

The DMC forces face the same difficulties as VMC, but also an additional bias: the $d_\lambda \ln P$ term cannot be readily computed, because the mixed probability distribution P in DMC is not numerically attainable. Addressing the bias with the hybrid estimator[54] does not guarantee control of the bias.[35,50] Advanced DMC force estimators have been proposed in the literature[44,50,55] but they call for extra implementation and are left out of this work.

III. COMPUTATIONAL DETAILS

We perform numerical evaluation of the energies and forces, as laid out in the earlier for different approaches: SCF, VMC and DMC. First, the HF or DFT calculations are performed in GAMESS-US[56] using RHF and ROHF methods, depending on the system, and commonly referred to as the SCF or DFT results depending on whether or not HF is included. Single-determinant orbitals are considered in two different LCAO bases: cc-pVDZ and cc-pVTZ (correlation-consistent, polarized double/triple zeta). The bases are from the correlation-consistent ECPs.[37,57–59] The ECP for gold is for large pseudo-valence $Z_{\text{eff}} = 19$, which is more versatile and computationally demanding as $Z_{\text{eff}} = 11$ that is more common in applications. The orbitals are computed based on HF and selected commonly used DFT functionals, namely LDA, Perdew–Burke–Ernzerhof (PBE), B3LYP and BLYP. Nexus workflow tool[60] to manage the SCF calculations.

Based on the SCF orbitals, we prescribe various QMC PESs using ordinary workflows as implemented in the Cornell-Holland ab initio materials package (CHAMP).[61] The PESs and their selected derived properties correspond to the following ansätze: In the VMC-J12 workflow, we optimize 1 and 2-body Jastrow terms on frozen orbitals and then evaluate it using VMC. In the VMC-J123 workflow, we append and re-optimize the VMC-J12 wavefunctions with 3-body Jastrow terms, then re-evaluate using VMC. In the DMC workflow, we re-evaluate the VMC-J123 wavefunction using DMC. In the VMC-OO workflow, we re-optimize orbital coefficients in the presence of a VMC-J12 wavefunction. The total numbers of thus optimized parameters vary between 359 (Au, ROHF, cc-pVDZ) and 24 662 ($MeSAu_3$, RHF, cc-pVTZ); a full table is given in the Supplementary material. The Jastrow factors are implemented according to Ref. 42 and optimized using the stochastic reconfiguration (SR) method[62] in a conjugate gradient implementation.[63] The orbital rotations are done in the full molecular orbital bases using the SR method with $\tau = 0.01$. In DMC, the T-move scheme[64] is used to treat non-local projectors of the ECPs, with 12 randomly rotated quadrature points. A linear extrapolation to zero DMC time-step extrapolation is carried out with the following finite time-steps $\Delta\tau$: 0.008, 0.004, 0.002 Ha^{-1} (except ΔE of Au_3 systems, only based on $\Delta\tau = 0.008$ Ha^{-1}). DMC target populations of 50 walkers were simulated on 128 parallel processes.

Based on direct sampling of the PESs we estimate selected observables, namely energy differences and energy-minimizing geometries. The geometries are obtained from fitting the energy curve with a third order polynomial and locating minima. Their statistical uncertainties are estimated by statistical resampling of the fitting of the PES.

Additionally, geometry optimization was performed in CHAMP based on simultaneous orbital rotations and analytical gradients (forces) in a VMC calculation with a two-body Jastrow. The forces were estimated with an efficient estimator[49] and using a guiding function with $\varepsilon = 0.01$ to avoid an infinite variance problem.[46] The estimator is formally free of bias (see e.g., Refs. 35 and 50) when all Jastrow and orbital coefficients are simultaneously optimized, as is done in the VMC-OO workflow. Between $80, \ldots, 120$ iterations were calculated, where the geometry was updated based on conjugate gradient descent[62] with $\alpha = 0.1, \ldots, 0.4$ depending on the system. This was considered adequate to ensure statistical convergence, in the absence of stopping criteria. No symmetries or sophisticated Hessians were used in the process. Estimates of derived structural properties, such as bond lengths and dihedral angles, were calculated from geometry snapshots in the iteration process. Their statistical errorbars were estimated from the variance and sample auto-correlation of subsequent iterations of the last about 50 samples of each parameter.

The data, workflows and software are available in a data repository.[65] To control the results, we have employed additional methods and implementations (including another QMC software[66]), whose results are presented in the Supplementary material.

IV. PURE GOLD SYSTEMS

We begin by performing SCF and QMC calculations of selected few-atom gold structures, namely Au, Au_2 and Au_3. The molecules will be assessed on common merits of accuracy that are derived from the ground state energies, with the purpose of benchmarking the alternative methods and their implementations. The results with five different mean-field descriptions [R(O)HF, LDA, PBE, B3LYP and BLYP] are given in the triple-zeta basis (cc-pVTZ).

A. Au_1

The isolated gold atom, Au_1, reflects the baseline of accuracy for each of the methods. The atom is simulated in three ionization states, Au^q, where $q = -1, 0, 1$. Orbitals for the charged states, $q = \pm 1$, are obtained with RHF and the neutral state, $q = 0$, with ROHF. Correlation energies recovered from alternative QMC theories are presented in Fig. 1 for the neutral gold atom. The total binding QMC energies for the ground states are variational, and thus, differences from the R(O)HF energy are viable merits of accuracy. The VMC energies based on one and two-body Jastrows (J12) are the highest by about 1.5 eV. Upon optimization of the three-body term, the energy improves by another 0.5 eV, and by using DMC, by yet another 1.0 eV. At all stages, performance of the R(O)HF vs DFT orbitals in QMC is worse by 0.2–0.3 eV, because the lack of dynamic correlations leads to poorer nodal surfaces. The QMC data based on DFT functionals are highly consistent and the differences are indistinguishable from statistical noise. In other words, the QMC treatment of correlation in gold atom is highly independent of the DFT. Understanding differences in the DFT results would require investigation of advanced properties (see e.g., Ref. 67), whereas the use of QMC effectively lifts this requirement for numerical purposes.

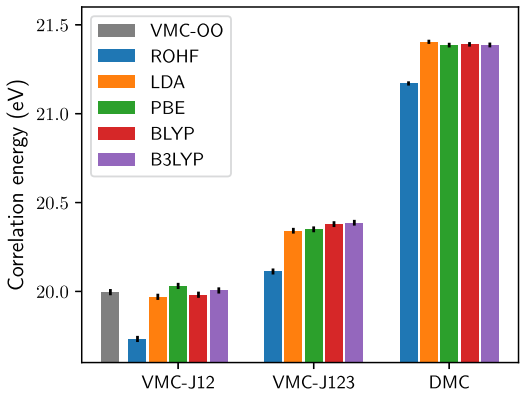

FIG. 1. Correlation energies (eV) of neutral Au_1 from different QMC implementations based on single-determinant ROHF orbitals from different SCF methods.

In Table I we report derived energetic properties of the ground state, namely the electron affinity (EA) and the ionization potential (IP) and compare them to well-known experimental values.[68] The SCF data are largely spread with the relative standard deviation (RSD) of over 32% for EA and 8% for IP, leading to high mean absolute relative errors (MARE). The MARE and RSD of the SCF data without HF, which is an outlier, are over 8% and 2%, respectively. Even so, upon introduction of the J12 wavefunctions, the orbital dependency vanishes to a high degree: The RSD drops by an order of magnitude. The RSD numbers contain statistical noise and remain upper bound estimates. Figure 2 illustrates how the spread of values from the mean-field are clustered and incrementally improved by the QMC treatments.

The MARE of the J12 wavefunction is no better than DFT, because of a (very consistent) systematic bias. The improvement of accuracy from the J123 wavefunction is modest. The DMC is better, but doesn not quite meet the experiment. Reasons for that could be in the limitations of the single-reference wavefunction, and ROHF orbitals. Also, the non-augmented basis sets are suboptimal for describing the diffuse anion, Au^-. As already demonstrated in the original publication of the ECP, single-determinant DMC does not perform any better than this even in a two-component spinor formalism.[37] This underlines the challenges of treating fully accurate dynamic correlation in coinage metals, such as gold.

B. Au_2

Next, we study the singlet ground state of the gold dimer, Au_2, where we predict the bond length d^{Au-Au} and binding energy D_e. The binding curve is computed along a regular grid of six displacements, and then fitted to a third order polynomial, where the spectroscopic properties are obtained. The molecular binding energy is estimated according to

$$D_e = E_0(Au_2) - E_v(Au_2) - 2E(Au), \tag{11}$$

where $E_0(Au_2)$ is the fitted minimum energy of the binding curve, $E_v(Au_2) = 23.7$ meV[68] is the vibrational zero-point energy, and $E(Au)$ is the total energy for a corresponding neutral atom. Statistical uncertainties of the fitted properties are estimated by random resampling.

Fitted spectroscopic data from different binding curves are presented in Table II. The systematic accuracy of predicting the bond length d^{Au-Au} with the SCF is, on average, 2.5% (MARE) with almost as high variability 1.8% (RSD). VMC with J12 optimized on frozen orbitals reduces the systematic error and XC treatment by a factor of 3. Optimizing three-body Jastrow terms only affects the results modestly. DMC has the highest accuracy with MARE of about 0.6%.

The same pattern repeats in the prediction of D_e: VMC is modestly more accurate, on average, than SCF, and much more consistent. DMC is clearly the most accurate with MARE of 9%. Our DMC results for D_e based on Eq. (11) are in a systematic disagreement with an earlier work,[37] where D_e was obtained by fitting to a Morse potential. Another outlier is J123, whose data for D_e contains a sizable inaccuracy based on estimation from Eq. (11). It turns out that J123, as implemented in 42 performs much better with Au atom than with an Au_2 dimer, hence the poor energy

TABLE I. Electronic affinities (EA; eV) and (right) the ionization potentials (IP; eV) of Au atom based on selected methods and SCF orbitals. The MARE and RSD are estimated over the sets of orbitals.

	Orbitals	SCF	J12	J123	DMC
EA	EXPT	2.31			
	RHF	0.320 06	1.67(3)	1.77(3)	2.01(2)
	LDA	2.440 45	1.78(3)	1.87(3)	2.11(2)
	PBE	2.088 36	1.75(3)	1.88(3)	2.16(2)
	BLYP	2.062 16	1.81(3)	1.86(3)	2.12(2)
	B3LYP	1.930 98	1.75(3)	1.88(3)	2.10(2)
	MARE (%)	25.7	24.04	19.88	9.09
	MARE[a] (%)	10.59	23.12	18.96	8.17
	RSD (%)	32.19	2.06	1.87	2.09
	RSD[a] (%)	8.17	1.03	0.36	1.09
IP	EXPT	9.23			
	RHF	7.693 72	8.86(3)	8.95(3)	9.11(2)
	LDA	9.906 81	8.87(3)	8.95(3)	9.18(2)
	PBE	9.537 34	8.93(3)	8.95(3)	9.14(2)
	BLYP	9.518 17	8.89(3)	9.02(3)	9.15(2)
	B3LYP	9.345 17	8.94(3)	9.03(3)	9.14(2)
	MARE (%)	6.34	3.58	2.69	0.91
	MARE[a] (%)	3.76	3.48	2.6	0.8
	RSD (%)	8.4	0.34	0.39	0.26
	RSD[a] (%)	2.22	0.31	0.39	0.18

[a]HF data excluded as an outlier.

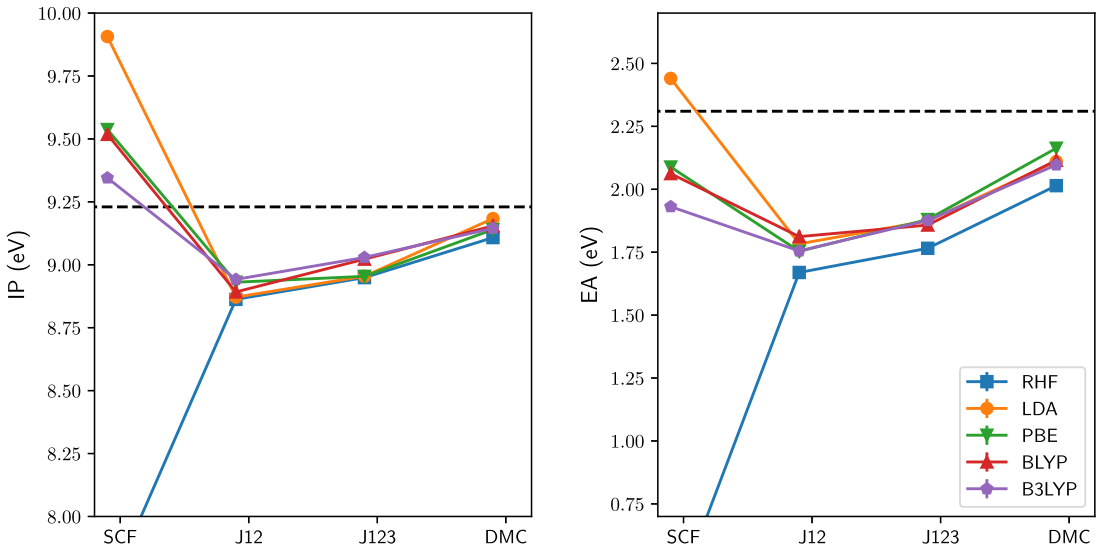

FIG. 2. Convergence of (left) the electronic affinities (EA; eV) and (right) the ionization potentials (IP; eV) of Au atom based on selected methods and SCF orbitals. The VMC methods (J12, J123) remove the orbital dependency, and increasing accuracy up to DMC brings the results closer to experiment (dashed line).

(over 50% MARE!). This is not the case with another implementation of J123,[43] as detailed in Supplementary material. Improving the systematic accuracy from here is left for another work. The values of MARE and RSD for QMC are upper bounds and could be reduced, on average, by practicing more accurate (and expensive) optimizations.

Yet, the main reason for the comparison is to set up context for the last column of Table II, VMC-OO. The data precedes full reoptimization of the J12 and orbital coefficients of the VMC wavefunction, but also the ionic geometry. The process implicitly removes dependency on the initial orbitals, and hence, in terms of RSD, the data quality far supersedes J12 and J123, let alone SCF.

TABLE II. The predicted equilibrium bond lengths (d^{Au-Au}; Å) and binding energies (D_e; eV) of Au_2 from selected methods based on variable orbitals.

	Orbitals	SCF	J12	J123	DMC	VMC-OO
d^{Au-Au}	EXPT	2.471 9				
	RHF	2.591 55	2.47(2)	2.506(12)	2.479(9)	2.503(6)
	LDA	2.455 2	2.48(2)	2.483(11)	2.476(8)	2.486(2)
	PBE	2.510 48	2.50(2)	2.488(12)	2.508(10)	2.490(2)
	BLYP	2.543 68	2.491(14)	2.484(12)	2.485(10)	2.5137(7)
	B3LYP	2.534 06	2.51(2)	2.510(13)	2.484(9)	2.495(5)
	MARE (%)	2.5	0.74	0.9	0.59	1.04
	RSD (%)	1.8	0.56	0.47	0.46	0.4
D_e	EXPT	2.27				
	RHF	0.852 13	1.71(6)	1.16(7)	1.98(4)	1.83(5)
	LDA	2.958 95	1.67(6)	1.10(5)	2.08(4)	1.81(7)
	PBE	2.336 25	1.60(6)	1.12(6)	2.14(4)	1.85(5)
	BLYP	2.146 49	1.68(7)	1.02(5)	2.09(5)	1.82(5)
	B3LYP	1.997 58	1.67(6)	1.07(6)	2.09(4)	1.82(6)
	MARE (%)	22.63	26.57	51.68	8.66	19.5
	RSD (%)	30.23	1.55	2.01	2.28	0.5

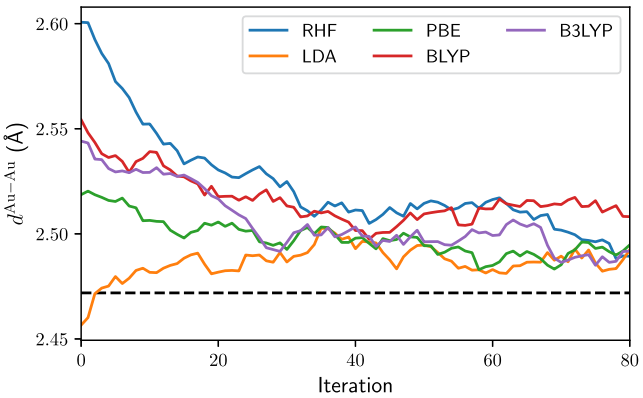

FIG. 3. The trace of convergence of d^{Au-Au} of Au_2 in VMC-OO geometry optimization based on different starting orbitals and geometries from the corresponding SCF methods. Black dashed line is the experimental reference.

The systematic accuracy (MARE) is not significantly different from frozen-orbital VMC, and it cannot par with DMC. Still, the crucial point of turning the XC functional dependency of SCF to something more robust is clearly demonstrated in the results and illustrated in Fig. 3.

C. Au_3

Next, we study the $^2B_2$ state of neutral gold trimer Au_3 in three different isomer configurations that are characterized by the isosceles triangle θ. The isomers will be referred to as I ($\theta \approx 65^\circ$), II ($\theta \approx 125^\circ$) and III ($\theta \approx 180^\circ$) and their data will compared to similar structures from an a spin-unrestricted open-shell coupled-cluster singles doubles perturbative triples [UCCDS(T)] benchmark study.[69] We perform geometry optimization of each isomer using ROHF orbitals in SCF and VMC-OO. We also compute single-shot VMC and DMC energies at the reference structures for analysis of computational cost and energy differences.

In Table III, we present bond lengths d^{Au-Au}, dihedral angles θ and energy differences ΔE to the lowest-energy isomer of Au_3. It is immediately seen that HF fails to bind the Au_3-I isomer and mistakenly relaxes it to Au_3-II. This can be attributed to failures in addressing the strong influence of d-electron correlation in these species.[69,70] Generally, in each isomer the DFT methods overestimate d^{Au-Au} by 0.02, . . . , 0.06 Å, except LDA, which underestimates it by around 0.05 Å. On the other hand, VMC-OO very strongly predicts the same bond lengths as UCCSD(T). While the bond length converges to a common regime, the dihedral angle θ is not significantly improved. A finite-difference study with DFT suggests that the force-constant matrices regarding a change in θ in isomers I and II are, respectively, more than 15 and 217 times smaller than those regarding d^{Au-Au}. Thus, the poor resolution of angles is a consequence of a poor signal-to-noise ratio, which is further elaborated in in Sec. V.

In the last column of Table III, we consider energy differences with respect to the predicted lowest-energy isomer by each method. The DFT methods (except LDA) identify Au_3-II as the lowest-energy isomer, in contrast to the UCCSD(T) benchmark suggesting Au_3-I. The VMC-OO results are not statistically significant to dictate whether I or II is the lowest-energy isomer. However, the DMC results are more clearly in agreement with UCCSD(T). The DMC runs were based on UCCSD(T) geometries instead of VMC-OO for convenience, because they were almost identical in d^{Au-Au} but resolving θ from VMC-OO was less reassuring for the reasons stated above.

TABLE III. Bond lengths d^{Au-Au}, dihedral angles θ and energy differences ΔE from SCF and VMC-OO calculations of isomers I-III of Au_3. The VMC-OO results are based on BLYP starting orbitals, and UCCSD(T) data are from Ref. 69.

Isomer	Method	d^{Au-Au} (Å)	θ (deg)	ΔE (E_H)
I	RHF	2.680 37	127.490 15	
	LDA	2.545 85	64.624 96	0.0
	PBE	2.609 06	66.574 76	0.112 55
	BLYP	2.642 46	69.732 25	0.214 28
	B3LYP	2.630 77	69.751 44	0.136 91
	VMC-OO	2.570(2)	69.43(10)	0.000(12)
	DMC	2.58	65.28	0.00(3)
	UCCSD(T)	2.58	65.28	0.0
II	RHF	2.680 81	127.378 69	0.0
	LDA	2.480 56	138.920 92	0.010 49
	PBE	2.555 04	138.209 9	0.0
	BLYP	2.594 57	139.796 69	0.0
	B3LYP	2.592 3	134.027 01	0.0
	VMC-OO	2.526 6(10)	140.88(10)	0.04(3)
	DMC	2.53	126.36	0.04(3)
	CCSD(T)	2.53	126.36	0.117 1
III	RHF	2.688 96	180.0	0.020 83
	LDA	2.504 32	180.0	0.158 1
	PBE	2.575 1	180.0	0.094 96
	BLYP	2.614 22	180.0	0.082 02
	B3LYP	2.608 59	180.0	0.067 89
	VMC-OO	2.544(2)	179.29(10)	0.084(13)
	DMC	2.55	180.0	0.12(3)
	UCCSD(T)	2.55	180.0	0.1908

V. STRUCTURAL OPTIMIZATION OF THIOLATED GOLD MOLECULES

Let us now compare DFT and VMC-OO methods in predicting the geometries of the following series of thiolated gold structures: $(MeSAu)_n$, where (Me is CH_3) and $n = 1, \ldots, 3$. Progression of the system size will help us investigate calculation merits, such as accuracy and computational cost, in view of chemically interesting applications.

In Table IV we present estimates of selected structural properties of $(MeSAu)_n$, based on the DFT and VMC-OO calculations, including a reference from an earlier PBE study.[12] The bond lengths obtained from VMC-OO are generally shorter by $0.02, \ldots, 0.1$ Å compared to PBE, BLYP and B3LYP, whereas LDA underestimates most of the bond lengths compared to VMC. While we do not present a high-accuracy control of the accuracy, it is clear that VMC-OO efficiently removes the orbital dependencies coming from the starting orbitals, as it should. We elaborate and discuss in the Supplementary material how different structural parameters converge to the same statistical locality, albeit with different statistical properties. A scheme like this alleviates one from performing an exhaustive scan of different DFT functionals to find reasonable starting orbitals for this scheme.

For dihedral angles, the performance of VMC-OO is not decisively improved over DFT. This is due to the angles being soft parameters, i.e., their force-constants are low. The stiffness (force-constant) is directly proportional to the signal (force) near the equilibrium. Therefore, the signal-to-noise ratio is much more favorable in stiff bonds, such as C–H, than in soft parameters, such as dihedral angles. Furthermore, as the geometry update is not aware of parametric definitions, the soft parameters follow a Markov chain random walk with high autocorrelation. We have estimated, approximately, that the parameter stiffnesses in $(eV/Å)^2$ for $(MeSAu)_1$ are as follows: 660 for C–H bond, 120 for S–C bond, 12 for Au–S bond, 39 for S–C–H bond angle, and 7.2 for Au–S–C bond angle. For instance, the signal-to-noise ratio is over 90 times higher for the C–H bond than for the Au–S–C angle.

Generally, the statistical performance of VMC-OO decreases approximately with $n^{-1/2}$ beyond the computational scaling ($t \sim n^3$). This has been analyzed in Table V, where we present the relative uncertainties of the VMC forces in all calculations. Moreover, the pseudo-valence Z_{eff} has an intrinsic effect on the force variance (i.e., inverse statistical efficiency) of each individual atomic species. This property has been systematically studied,[34,35] and it is generally accepted that higher Z_{eff} comes with a higher variance, because the trial wavefunction is probably less exact. Theorizing ratios of force variances in heterogeneous compounds is complex, because the elements affect each other through local energy variance. Here, based on a series of qualitative similar particles, we observe that the respective force uncertainty ratios of Au:S:C:H are fixed to ~11:3:3:1. On the other hand, the rescaled uncertainty $\sigma F n^{-1/2}$ remains effectively constant. The latter point confirms that the force variance is extensive,[71] i.e., it scales linearly with the system size n. A broad conclusion is that the statistical properties of forces can be reliably predicted in this workflow. A specific result is that coinage metals, such as gold, which are by far the hardest to converge statistically in this ensemble, might benefit from using a cheaper pseudopotential ($Z_{\text{eff}} = 11$) for geometry relaxation purposes.

VI. ACCURACY AND COMPUTATIONAL COST

Finally, we present a coarse summary of some relevant merits of accuracy and prospects of the computational cost. The comparison is made between selected DFT, VMC, and DMC data, as presented in this work, excluding R(O)HF as an outlier for practical purposes.

Thus, we consider the mean absolute deviation (MAD) and standard deviation (STD) over DFT results from LDA, PBE, BLYP and B3LYP, and their counterparts in VMC-J12, VMC-OO and DMC, where applicable. Furthermore, a basis-set effect (BSE) is computed as an absolute deviation between results from double-zeta and triple-zeta bases. This uncommon merit of accuracy is reported here as a probe for sensitivity of the basis set. The energy data contains EA/IP results from Au, D_e of Au_2 and the energy displacements of Au_3-II and Au_3-III from Au_3-I. The geometric data is based on bond length predictions in Au_2, Au_3 and $MeSAu_n$ with $n = 1, \ldots, 3$; the dihedral angles are left out, because they lack external reference data and the QMC results have poor signal-to-noise ratios. The VMC results are aggregated from both VMC-J12 and VMC-OO. The DMC data is scarce and relatively

TABLE IV. Structural properties of $MeSAu_n$ with $n = 1, \ldots, 3$ based on SCF and VMC-OO geometry optimizations, and a reference DFT study.[12]

n	Method	Au–Au	Au–S	S–C	C–H	Au–S–C	Au–S–Au
1	LDA		2.197 52	1.818 41	1.104 8	104.024 9	
	PBE		2.239 63	1.830 72	1.096 39	104.517 15	
	BLYP		2.275	1.852 28	1.093 67	104.770 2	
	B3LYP		2.264 04	1.835 29	1.088 29	104.482 72	
	VMC		2.213(2)	1.806(3)	1.077 8(5)	105.55(9)	
2	LDA	2.640 29	2.425 67	1.828 03	1.103 41	105.945 29	66.037 13
	PBE	2.714 46	2.484 58	1.840 09	1.095 05	107.776 11	66.359 83
	PBE[12]	2.76	2.50	1.83			67.1
	BLYP	2.778 61	2.537 76	1.862 49	1.092 39	108.401 67	66.554 1
	B3LYP	2.7557	2.512 74	1.8429	1.087 24	108.199 57	66.625 71
	VMC	2.677(8)	2.472(6)	1.815(3)	1.077 2(2)	110.65(2)	65.83(13)
3	LDA	2.904 72	2.320 89	1.833 19	1.102 82	107.976 51	77.445 8
	PBE	2.9919	2.363 69	1.845 02	1.094 5	108.912 35	78.537 42
	PBE[12]	3.02	2.38	1.84			78.7
	BLYP	3.075 65	2.402 76	1.867 56	1.091 78	109.223 49	79.617 85
	B3LYP	3.054 24	2.384 29	1.8461	1.086 69	109.364 76	79.685 37
	VMC	2.998(2)	2.349 0(9)	1.821 1(12)	1.076 48(7)	111.17(3)	79.24(7)

TABLE V. Relative statistical uncertainties of the VMC forces per each atomic species in $(MeSAu)_n$, where $n = 1, \ldots, 3$: the middle panels show the mean uncertainties on Au, S and C atoms per that of H, whereas the rightmost panel shows the force uncertainty for H with different n but same number of samples. The numbers demonstrate very regular scaling of the force uncertainties with the system size.

n	Au/H	S/H	C/H	H $n^{-1/2}$
1	11.3	3.0	3.4	0.0022
2	10.1	3.0	3.3	0.0022
3	10.2	3.1	3.1	0.0022

noisy, so the comparison likely undermines its prospects for making accurate predictions. Generally, the QMC data is subject to statistical fluctuation, which overestimates MAD, STD and BSE by a modest amount. The MAD are not reported without an external reference. A full list of values is available in the supplementary material.

Comparison of the energetic accuracy is presented in the top panel of Fig. 4. Both MAD and STD of the DFT results are over 0.2 eV, and the BSE is about 0.07 eV. The DFT approach is as good as it is consistent. VMC and DMC successfully mitigate the spread (STD <0.1 eV), but VMC faces trouble in predicting the right energies with MAD over 0.3 eV. DMC has the lowest MAD of 0.11 eV, which could be further by focusing the effort on more advanced wavefunctions.

Similar comparison of the geometry effects is given in the bottom panel of Fig. 4. The MAD of DFT is around 0.05 Å but the STD is only 0.03 Å, signaling systematic biases, such as overestimation of bond lengths. The VMC and DMC come out as equally good, about 0.015 Å accurate. The DMC numbers are only based on Au_2, in the absence of practical geometry relaxation techniques being studied in this work. The results from both QMC approaches are pervaded by statistical noise, which undermines their true merits but also underlines their practical challenges.

Finally, we study scaling of the computational cost with system size n based on our calculations of Au_n and $(MeSAu)_n$, where $n = 1, \ldots, 3$. The approach is empirical in the sense that we compare the actual mean running times T of various workflows on common hardware. Scaling fits of the form $T(n) = an^b$ are presented in the Supplementary Material for various energy and geometry

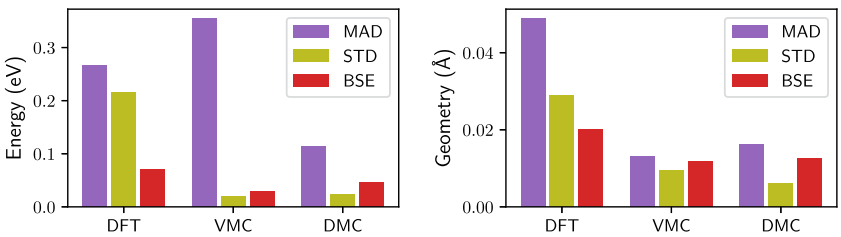

FIG. 4. Mean MAD, STD and BSE for predicting energies (top) and geometries (bottom). The data is averaged over DFT, VMC and DMC data from various gold molecules presented in this study, where applicable.

workflows. Based on the prefactors a, the SCF runs based on R(O)HF are by far the cheapest; VMC is two orders of magnitude more costly, and DMC is two more orders of magnitude costlier. Besides this order-of-magnitude ballpark, the prefactor has little importance, as it is characterized by the software, hardware, physical problem and choices of the stopping criteria. Some of this is discussed in the supplementary material. The effective scaling exponent b is consistently between 2 and 3, as also anticipated.[29,30] This is a significant property related, on one hand, to an efficient implementation of VMC-OO forces,[49] and on the other hand, because most other high-accuracy wavefunction methods have $b > 4$, such as coupled-cluster singles doubles perturbative triples [CCSD(T)] with $b = 7$.[29] Finally, increasing the basis set from double-zeta to triple-zeta effectively triples the convergence time of the SCF. The effect on VMC-OO is about 50% increase because of more gradients to compute, but in VMC-J123 and DMC the effect is almost negligible.

Many conclusions can be drawn from this section that only confirm what has been known in the past about the accuracy and scaling properties of QMC. A prospective strategy for maintaining beyond-DFT accuracy in the predictive simulation of molecular gold compounds would be through using VMC for geometry optimization, and then DMC for energies and other derived properties.

VII. SUMMARY AND OUTLOOK

We have studied first-principles calculation of small gold particles using common DFT and QMC methods. Simulating selected pure and thiolated gold molecules (Au_n and $MeSAu_n$, where $n = 1, \ldots, 3$), we have shown that QMC provides more consistent and accurate results than DFT with different XC functionals. We demonstrate that the mean spread between selected XC functionals is over 0.2 eV for energies and 0.04 Å for bond lengths, whereas the respective numbers are less than 0.03 eV and 0.02 Å for QMC (containing statistical noise). The effects due to basis set are also more modest with QMC. The QMC approach provides a consistent treatment of the dynamic correlation, but perfect agreement with the experiment remains elusive. For better predictive accuracy, we look forward to using with multi-configuration trial wavefunctions and a two-component spinor technique[39] to treat spin–orbit coupling.

Much of the work is focused on geometry optimization with the VMC-OO approach. The proof of concept is to show that the method is efficient and robust even for coinage metal compounds, albeit facing new statistical challenges. Indeed, the consistent prediction of atomic distances far exceeds that of DFT, but dihedral angles show little improvement due to poor signal-to-noise ratios. Generally, we observe that the scaling of statistical properties with the system size obeys well-known laws: the computation of VMC forces scales with n^3 (in an efficient implementation), and the force variance with n^1. Simple remedies outside of estimator techniques would be considering pseudopotentials with lower pseudo-valences Z_{eff}[35] and curve-fitting instead of gradient minimization to address soft parameters (see e.g., Refs. 72 and 73). While the prospect of robust DMC forces adds both accuracy and cost, a reasonable strategy for the meanwhile could be as follows: use VMC-OO for geometry optimization and DMC for everything else. For example, in Au_3 the QMC approaches predicted the correct structures and ordering of isomer energies, the latter of which most DFT functionals failed.

Whether admitted or not, the applied simulations of coinage metal compounds need a high-standard calculation method to control DFT, which will remain the main workhorse. We have demonstrated QMC approaches that are simple and effective in addressing the XC dependency of DFT, and also two or more orders of magnitude more expensive. This is the price to pay for higher quality.

SUPPLEMENTARY MATERIAL

Supplementary material is provided featuring figures, benchmark data and data tables that support the discussion.

ACKNOWLEDGMENTS

The authors thank Professor Claudia Filippi from her significant support of the work, providing guidance with the software and feedback on the manuscript. The work has been performed under the Project HPC-EUROPA3 (Grant No. INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme. We thank SURF (www.surf.nl) for the support in using the National Supercomputer Snellius. We acknowledge grants of computer capacity from the Finnish Grid and Cloud Infrastructure (persistent identifier Grant No. urn:nbn:fi:research-infras-2016072533). The authors also wish to acknowledge CSC–IT Center for Science, Finland, for computational resources. This work was supported by the Academy of Finland.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Juha Tiihonen: Conceptualization (equal); Data curation (lead); Formal analysis (lead); Funding acquisition (supporting); Investigation (lead); Methodology (lead); Resources (supporting); Software (lead); Visualization (lead); Writing – original draft (lead); Writing – review & editing (lead). **Hannu Häkkinen**: Conceptualization (equal); Funding acquisition (equal); Project administration (equal); Resources (equal); Supervision (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are openly available in a data repository [the final data uploaded and DOI provided after possible acceptance and revision of the manuscript].

REFERENCES

[1] L. Liu and A. Corma, "Metal catalysts for heterogeneous catalysis: From single atoms to nanoclusters and nanoparticles," Chem. Rev. **118**, 4981–5079 (2018).

[2]B. Zhang, A. Sels, G. Salassa, S. Pollitt, V. Truttmann, C. Rameshan, J. Llorca, W. Olszewski, G. Rupprechter, T. Bürgi, and N. Barrabés, "Ligand migration from cluster to support: A crucial factor for catalysis by thiolate-protected gold clusters," ChemCatChem **10**, 5372–5376 (2018).
[3]E. Boisselier and D. Astruc, "Gold nanoparticles in nanomedicine: Preparations, imaging, diagnostics, therapies and toxicity," Chem. Soc. Rev. **38**, 1759 (2009).
[4]M. Sengani, A. M. Grumezescu, and V. D. Rajeswari, "Recent trends and methodologies in gold nanoparticle synthesis—A prospective review on drug delivery aspect," OpenNano **2**, 37–46 (2017).
[5]K. Saha, S. S. Agasti, C. Kim, X. Li, and V. M. Rotello, "Gold nanoparticles in chemical and biological sensing," Chem. Rev. **112**, 2739–2779 (2012).
[6]K. G. Stamplecoskie and P. V. Kamat, "Size-dependent excited state behavior of glutathione-capped gold clusters and their light-harvesting capacity," J. Am. Chem. Soc. **136**, 11093–11099 (2014).
[7]B. Hammer and J. K. Norskov, "Why gold is the noblest of all the metals," Nature **376**, 238–240 (1995).
[8]P. Pyykkö, "Theoretical chemistry of gold," Angew. Chem., Int. Ed. **43**, 4412–4456 (2004).
[9]P. Pyykkö, "Relativistic effects in chemistry: More common than you thought," Annu. Rev. Phys. Chem. **63**, 45–64 (2012).
[10]S. Malola and H. Häkkinen, "Prospects and challenges for computer simulations of monolayer-protected metal clusters," Nat. Commun. **12**, 2197 (2021).
[11]H. Häkkinen, B. Yoon, U. Landman, X. Li, H.-J. Zhai, and L.-S. Wang, "On the electronic and atomic structures of small Au_N (N = 4–14) clusters: A photoelectron spectroscopy and density-functional study," J. Phys. Chem. A **107**, 6168–6175 (2003).
[12]H. Grönbeck, M. Walter, and H. Häkkinen, "Theoretical characterization of cyclic thiolated gold clusters," J. Am. Chem. Soc. **128**, 10268–10275 (2006).
[13]Y. Feng and L. Cheng, "Structural evolution of $(Au_2)_n$ (n = 1–8) clusters from first principles global optimization," RSC Adv. **5**, 62543–62550 (2015).
[14]H. Häkkinen, "Atomic and electronic structure of gold clusters: Understanding flakes, cages and superatoms from simple concepts," Chem. Soc. Rev. **37**, 1847 (2008).
[15]C. Zeng, "Precision at the nanoscale: On the structure and property evolution of gold nanoclusters," Pure Appl. Chem. **90**, 1409–1427 (2018).
[16]J. A. Keith, D. Fantauzzi, T. Jacob, and A. C. T. van Duin, "Reactive forcefield for simulating gold surfaces and nanoparticles," Phys. Rev. B **81**, 235404 (2010).
[17]A. Pihlajamäki, J. Hämäläinen, J. Linja, P. Nieminen, S. Malola, T. Kärkkäinen, and H. Häkkinen, "Monte Carlo simulations of $Au_{38}(SCH_3)_{24}$ nanocluster using distance-based machine learning methods," J. Phys. Chem. A **124**, 4827–4836 (2020).
[18]C. Zeni, K. Rossi, T. Pavloudis, J. Kioseoglou, S. de Gironcoli, R. E. Palmer, and F. Baletto, "Data-driven simulation and characterisation of gold nanoparticle melting," Nat. Commun. **12**, 6056 (2021).
[19]P. Makkar and N. N. Ghosh, "A review on the use of DFT for the prediction of the properties of nanomaterials," RSC Adv. **11**, 27897–27924 (2021).
[20]J. Akola, M. Walter, R. L. Whetten, H. Häkkinen, and H. Grönbeck, "On the structure of thiolate-protected Au_{25}," J. Am. Chem. Soc. **130**, 3756–3757 (2008).
[21]M. Azubel, J. Koivisto, S. Malola, D. Bushnell, G. L. Hura, A. L. Koh, H. Tsunoyama, T. Tsukuda, M. Pettersson, H. Häkkinen, and R. D. Kornberg, "Electron microscopy of gold nanoparticles at atomic resolution," Science **345**, 909–912 (2014).
[22]C. M. Aikens, "Effects of core distances, solvent, ligand, and level of theory on the TDDFT optical absorption spectrum of the thiolate-protected Au_{25} nanoparticle," J. Phys. Chem. A **113**, 10811–10817 (2009).
[23]S. A. Ivanov, I. Arachchige, and C. M. Aikens, "Density functional analysis of geometries and electronic structures of gold-phosphine clusters. The case of $Au_4(PR_3)_4^{2+}$ and $Au_4(\mu_2\text{-i})(PR_3)_4$." J. Phys. Chem. A **115**, 8017–8031 (2011).
[24]F. Muniz-Miranda, M. C. Menziani, and A. Pedone, "Assessment of exchange-correlation functionals in reproducing the structure and optical gap of organic-protected gold nanoclusters," J. Phys. Chem. C **118**, 7532–7544 (2014).
[25]G. Deng, S. Malola, P. Yuan, X. Liu, B. K. Teo, H. Häkkinen, and N. Zheng, "Enhanced surface ligands reactivity of metal clusters by bulky ligands for controlling optical and chiral properties," Angew. Chem., Int. Ed. **60**, 12897–12903 (2021).
[26]M. P. de Lara-Castells, "First-principles modelling of the new generation of subnanometric metal clusters: Recent case studies," J. Colloid Interface Sci. **612**, 737–759 (2022).
[27]K. Burke, "Perspective on density functional theory," J. Chem. Phys. **136**, 150901 (2012).
[28]C. M. Aikens, "Electronic and geometric structure, optical properties, and excited state behavior in atomically precise thiolate-stabilized noble metal nanoclusters," Acc. Chem. Res. **51**, 3065–3073 (2018).
[29]W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, "Quantum Monte Carlo simulations of solids," Rev. Mod. Phys. **73**, 33–83 (2001).
[30]N. Nemec, M. D. Towler, and R. J. Needs, "Benchmark all-electron *ab initio* quantum Monte Carlo calculations for small molecules," J. Chem. Phys. **132**, 034111 (2010).
[31]F. R. Petruzielo, J. Toulouse, and C. J. Umrigar, "Approaching chemical accuracy with quantum Monte Carlo," J. Chem. Phys. **136**, 124116 (2012).
[32]V. G. de Pina, B. G. A. Brito, G.-Q. Hai, and L. Cândido, "Quantifying electron-correlation effects in small coinage-metal clusters via *ab initio* calculations," Phys. Chem. Chem. Phys. **23**, 9832–9842 (2021).
[33]M. Barborini, S. Sorella, and L. Guidoni, "Structural optimization by quantum Monte Carlo: Investigating the low-lying excited states of ethylene," J. Chem. Theory Comput. **8**, 1260–1269 (2012).
[34]K. Nakano, T. Morresi, M. Casula, R. Maezono, and S. Sorella, "Atomic forces by quantum Monte Carlo: Application to phonon dispersion calculations," Phys. Rev. B **103**, L121110 (2021).
[35]J. Tiihonen, R. C. Clay, and J. T. Krogel, "Toward quantum Monte Carlo forces on heavier ions: Scaling properties," J. Chem. Phys. **154**, 204111 (2021).
[36]K. Nakano, A. Raghav, and S. Sorella, "Space-warp coordinate transformation for efficient ionic force calculations in quantum Monte Carlo," J. Chem. Phys. **156**, 034101 (2022).
[37]G. Wang, B. Kincaid, H. Zhou, A. Annaberdiyev, M. C. Bennett, J. T. Krogel, and L. Mitas, "A new generation of effective core potentials from correlated and spin–orbit calculations: Selected heavy elements," J. Chem. Phys. **157**, 054101 (2022).
[38]M. Dash, S. Moroni, A. Scemama, and C. Filippi, "Perturbatively selected configuration-interaction wave functions for efficient geometry optimization in quantum Monte Carlo," J. Chem. Theory Comput. **14**, 4176–4182 (2018).
[39]C. A. Melton, M. Zhu, S. Guo, A. Ambrosetti, F. Pederiva, and L. Mitas, "Spin-orbit interactions in electronic structure quantum Monte Carlo methods," Phys. Rev. A **93**, 042502 (2016).
[40]D. D. Koelling and B. N. Harmon, "A technique for relativistic spin-polarised calculations," J. Phys. C: Solid State Phys. **10**, 3107–3114 (1977).
[41]T. Takeda, "The scalar relativistic approximation," Z. Phys. B **32**, 43–48 (1978).
[42]C. Filippi and C. J. Umrigar, "Multiconfiguration wave functions for quantum Monte Carlo calculations of first-row diatomic molecules," J. Chem. Phys. **105**, 213–226 (1996).
[43]N. D. Drummond, M. D. Towler, and R. J. Needs, "Jastrow correlation factor for atoms, molecules, and solids," Phys. Rev. B **70**, 235119 (2004).
[44]J. van Rhijn, C. Filippi, S. De Palo, and S. Moroni, "Energy derivatives in real-space diffusion Monte Carlo," J. Chem. Theory Comput. **18**, 118–123 (2021).
[45]P. L. Ríos and G. J. Conduit, "Tail-regression estimator for heavy-tailed distributions of known tail indices and its application to continuum quantum Monte Carlo data," Phys. Rev. E **99**, 063312 (2019).
[46]C. Attaccalite and S. Sorella, "Stable liquid hydrogen at high pressure by a novel *ab initio* molecular-dynamics calculation," Phys. Rev. Lett. **100**, 114501 (2008).
[47]S. Pathak and L. K. Wagner, "A light weight regularization for wave function parameter gradients in quantum Monte Carlo," AIP Adv. **10**, 085213 (2020).
[48]C. Filippi and C. J. Umrigar, "Correlated sampling in quantum Monte Carlo: A route to forces," Phys. Rev. B **61**, R16291–R16294 (2000).
[49]C. Filippi, R. Assaraf, and S. Moroni, "Simple formalism for efficient derivatives and multi-determinant expansions in quantum Monte Carlo," J. Chem. Phys. **144**, 194105 (2016).
[50]S. Moroni, S. Saccani, and C. Filippi, "Practical schemes for accurate forces in quantum Monte Carlo," J. Chem. Theory Comput. **10**, 4823–4829 (2014).

[51] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, "Alleviation of the fermion-sign problem by optimization of many-body wave functions," Phys. Rev. Lett. **98**, 110201 (2007).
[52] J. Toulouse and C. J. Umrigar, "Optimization of quantum Monte Carlo wave functions by energy minimization," J. Chem. Phys. **126**, 084102 (2007).
[53] J. Toulouse and C. J. Umrigar, "Full optimization of Jastrow–Slater wave functions with application to the first-row atoms and homonuclear diatomic molecules," J. Chem. Phys. **128**, 174101 (2008).
[54] R. Assaraf and M. Caffarel, "Zero-variance zero-bias principle for observables in quantum Monte Carlo: Application to forces," J. Chem. Phys. **119**, 10536–10552 (2003).
[55] A. Badinski, P. D. Haynes, J. R. Trail, and R. J. Needs, "Methods for calculating forces within quantum Monte Carlo simulations," J. Phys.: Condens. Matter **22**, 074202 (2010).
[56] M. S. Gordon and M. W. Schmidt, "Advances in electronic structure theory," in *Theory and Applications of Computational Chemistry* (Elsevier, 2005), pp. 1167–1189.
[57] M. C. Bennett, C. A. Melton, A. Annaberdiyev, G. Wang, L. Shulenburger, and L. Mitas, "A new generation of effective core potentials for correlated calculations," J. Chem. Phys. **147**, 224106 (2017).
[58] A. Annaberdiyev, G. Wang, C. A. Melton, M. C. Bennett, L. Shulenburger, and L. Mitas, "A new generation of effective core potentials from correlated calculations: 3d transition metal series," J. Chem. Phys. **149**, 134108 (2018).
[59] M. C. Bennett, G. Wang, A. Annaberdiyev, C. A. Melton, L. Shulenburger, and L. Mitas, "A new generation of effective core potentials from correlated calculations: 2nd row elements," J. Chem. Phys. **149**, 104108 (2018).
[60] J. T. Krogel, "Nexus: A modular workflow management system for quantum simulation codes," Comput. Phys. Commun. **198**, 154–168 (2016).
[61] The Cornell-Holland *ab initio* materials package (CHAMP), 2023.
[62] S. Sorella, M. Casula, and D. Rocca, "Weak binding between two aromatic rings: Feeling the van der Waals attraction by quantum Monte Carlo methods," J. Chem. Phys. **127**, 014105 (2007).
[63] E. Neuscamman, C. J. Umrigar, and G. K.-L. Chan, "Optimizing large parameter sets in variational quantum Monte Carlo," Phys. Rev. B **85**, 045103 (2012).
[64] M. Casula, "Beyond the locality approximation in the standard diffusion Monte Carlo method," Phys. Rev. B **74**, 161102 (2006).
[65] J. Tiihonen (2023). "Towards structural optimization of gold nanoclusters with quantum Monte Carlo: Data repository (Version 1)," University of Jyväskylä. https://doi.org/10.23729/35765cd7-eab2-4b6c-9be0-cefe3b872d6b
[66] J. Kim, A. D. Baczewski, T. D. Beaudet, A. Benali, M. C. Bennett, M. A. Berrill, N. S. Blunt, E. J. L. Borda, M. Casula, D. M. Ceperley, S. Chiesa, B. K. Clark, R. C. Clay, K. T. Delaney, M. Dewing, K. P. Esler, H. Hao, O. Heinonen, P. R. C. Kent, J. T. Krogel, I. Kylänpää, Y. W. Li, M. G. Lopez, Y. Luo, F. D. Malone, R. M. Martin, A. Mathuriya, J. McMinis, C. A. Melton, L. Mitas, M. A. Morales, E. Neuscamman, W. D. Parker, S. D. Pineda Flores, N. A. Romero, B. M. Rubenstein, J. A. R. Shea, H. Shin, L. Shulenburger, A. F. Tillack, J. P. Townsend, N. M. Tubman, B. Van Der Goetz, J. E. Vincent, D. C. Yang, Y. Yang, S. Zhang, and L. Zhao, "QMCPACK: An open source *ab initio* quantum Monte Carlo package for the electronic structure of atoms, molecules and solids," J. Phys.: Condens. Matter **30**, 195901 (2018).
[67] P. Duffy, D. P. Chong, M. E. Casida, and D. R. Salahub, "Assessment of Kohn-Sham density-functional orbitals as approximate Dyson orbitals for the calculation of electron-momentum-spectroscopy scattering cross sections," Phys. Rev. A **50**, 4707–4728 (1994).
[68] M. D. Morse, "Clusters of transition-metal atoms," Chem. Rev. **86**, 1049–1109 (1986).
[69] R. R. Persaud, M. Chen, K. A. Peterson, and D. A. Dixon, "Potential energy surface of group 11 trimers (Cu, Ag, Au): Bond angle isomerism in Au_3," J. Phys. Chem. A **123**, 1198–1207 (2019).
[70] K. Balasubramanian and M. Z. Liao, "Electronic states and potential energy surfaces of gold and silver trimers," Chem. Phys. **127**, 313–324 (1988).
[71] R. Assaraf, "Computing physical properties with quantum Monte Carlo methods with statistical fluctuations independent of system size," Phys. Rev. E **90**, 063317 (2014).
[72] J. Tiihonen, P. R. C. Kent, and J. T. Krogel, "Surrogate Hessian accelerated structural optimization for stochastic electronic structure theories," J. Chem. Phys. **156**, 054104 (2022).
[73] K. K. Ly and D. M. Ceperley, "Phonons of metallic hydrogen with quantum Monte Carlo," J. Chem. Phys. **156**, 044108 (2022).