



This is the accepted manuscript made available via CHORUS. The article has been published as:

## Accurate Kohn-Sham auxiliary system from the ground-state density of solids

Ayoub Aouina, Matteo Gatti, Siyuan Chen, Shiwei Zhang, and Lucia Reining

Phys. Rev. B **107**, 195123 — Published 12 May 2023

DOI: [10.1103/PhysRevB.107.195123](https://doi.org/10.1103/PhysRevB.107.195123)

# Accurate Kohn-Sham auxiliary system from the ground state density of solids

Ayoub Aouina,<sup>1,2</sup> Matteo Gatti,<sup>1,2,3</sup> Siyuan Chen,<sup>4</sup> Shiwei Zhang,<sup>5,4</sup> and Lucia Reining<sup>1,2</sup>

<sup>1</sup>*LSI, CNRS, CEA/DRF/IRAMIS, École Polytechnique, Institut Polytechnique de Paris, F-91120 Palaiseau, France*

<sup>2</sup>*European Theoretical Spectroscopy Facility (ETSF)*

<sup>3</sup>*Synchrotron SOLEIL, L'Orme des Merisiers, Saint-Aubin, BP 48, F-91192 Gif-sur-Yvette, France*

<sup>4</sup>*Department of Physics, College of William & Mary, Williamsburg, Virginia 23185, USA*

<sup>5</sup>*Center for Computational Quantum Physics, Flatiron Institute, New York, New York 10010, USA*

(Dated: April 25, 2023)

The Kohn-Sham (KS) system is an auxiliary system whose effective potential is unknown in most cases. It is in principle determined by the ground state density, and it has been found numerically for some low-dimensional systems by inverting the KS equations starting from a given accurate density. For solids, only approximate results are available. In this work, we determine accurate exchange correlation (xc) potentials for Si and NaCl using the ground state densities obtained from Auxiliary Field Quantum Monte Carlo calculations. We show that these xc potentials can be rationalized as an ensemble of a few local functions of the density, whose form depends on the specific environment and can be well characterised by the gradient of the density and the local kinetic energy density. The KS band structure can be obtained with high accuracy. The true KS band gap turns out to be larger than the prediction of the local density approximation, but significantly smaller than the measurable photoemission gap, which confirms previous estimates. Finally, our findings show that the conjecture that very different xc potentials can lead to very similar densities and other KS observables is true also in solids, which questions the meaning of details of the potentials and, at the same time, confirms the stability of the KS system.

## I. INTRODUCTION

Density Functional Theory [1, 2] (DFT) is undoubtedly one of the biggest success stories of condensed matter theory, since it has made realistic electronic structure calculations possible for a huge range of materials, and since it has led to numerous insights [3, 4]. Two main problems had to be overcome in order to make DFT applicable in practice: first, it was necessary to find reliable approximations for the total energy as functional of the ground state density; second, an efficient way to determine the ground state density itself was needed. The solution to both problems relies on the approach of Kohn and Sham [2], where the interacting system is mapped onto an auxiliary system of non-interacting electrons with an effective Kohn-Sham (KS) potential that is designed to yield the ground state density. The exchange-correlation (xc) contribution to this potential,  $v_{xc}(\mathbf{r})$ , and to the xc energy density per particle,  $\epsilon_{xc}(\mathbf{r})$ , is unknown in most systems. The initial breakthrough came with the Local Density Approximation [2] (LDA). This approximation takes the energy density, and hence  $v_{xc}(\mathbf{r})$ , locally from the homogeneous electron gas (HEG), where it was calculated using Quantum Monte Carlo [5] (QMC). However, to find approximations that are systematically better than the LDA has turned out to be exceedingly difficult [6, 7]. Today, in spite of the developments of successful gradient corrections and sophisticated approximations tailored by exact constraints [8–11], one may say that there is still no generally established multi-purpose approximation beyond the LDA. One of the difficulties is that it is not easy to benchmark  $v_{xc}$ . First, the Kohn-Sham potential is not an observable by itself, which means, there are no experimental data to compare

with. Second, since  $v_{xc}$  is the potential of an *auxiliary* system, besides the density, any other observables calculated in the KS system can in principle be arbitrarily far from measured values. The prototype example for this dilemma is the KS eigenvalue band gap [12–14]. For example in the LDA, this KS gap is in general much smaller than the fundamental **electron addition-removal gap that can be extracted** from direct and inverse photoemission [15, 16]. It would, however, be too simple to just blame the LDA. While the exact direct gap of the auxiliary KS system equals the optical excitation gap in the limit of a single electron, this is only approximately true in real molecules or materials. Moreover, the fundamental ~~electron addition-removal gap, which can be extracted from direct and inverse photoemission measurements~~, is in general larger than the optical gap. The exact fundamental gap could in principle be determined as a difference of total energies with varying particle number. The exact KS eigenvalue gap should be smaller than the exact fundamental gap by a constant called derivative discontinuity. Therefore, there is no reason for the eigenvalue gap of the auxiliary KS system to equal the photoemission gap of the true material [2, 17]. However, the difference, i.e., the derivative discontinuity, is in general unknown. Indeed, the respective errors of the approximate functionals and of the KS system itself have been a matter of debate for many years. Results derived from many-body perturbation theory to first order in the screened Coulomb interaction [18–26] gave evidence that the error in simple semiconductors is mainly due to the auxiliary nature of the KS system, i.e., due to the missing derivative discontinuity [12–14], rather than due to the LDA for  $v_{xc}$ . However, these are merely estimates based on perturbation theory, and the numerically exact KS potential

and KS band gap of solids remain to date unknown.

More information is available in low-dimensional, often finite, systems, where ways have been proposed to invert the KS equations and find the KS potential starting from a given density [27–30]. This density could be determined by analytical or numerical methods. This has given precious insight about the potential and observables in the KS system [31–46]. For example, in the helium atom the exact highest occupied molecular orbital (HOMO) lies about 25 eV below the vacuum level, and an additional electron is unbound. The exact KS eigenvalue gap, instead, turns out to be only 20.3 eV, since KS binds the lowest **unoccupied** molecular orbital (LUMO)[36, 47, 48]. On top of this underestimate, the LDA reduces the HOMO-LUMO gap further, yielding 15.85 eV. The inversion of the KS system is not an easy task, though, and in particular a finite basis set may lead to drastically modified results [49–53]. Moreover, small changes in the density can yield large differences in the potential [54, 55]. Altogether, a reliable inversion of the KS equations remains a difficult task even for finite systems, and while various methods have been proposed to overcome the problems, research in this direction is still ongoing [54–65].

In realistic three-dimensional periodic systems, the density of beryllium obtained from x-ray diffraction experiments has been used to determine an auxiliary non-interacting system[66]. However, to the best of our knowledge, no results for  $v_{xc}$  obtained directly from a numerically exact density are available. This has several reasons, including the fact that data for numerically exact densities of solids were not available in the literature, and that the inversion in extended three-dimensional systems may bear new technical difficulties. Therefore, to date a series of important fundamental questions remain to be answered, in particular: *How can we adapt an inversion approach designed for finite systems to the case of solids, and which kind of precision can be obtained in solids? How different is the resulting  $v_{xc}$  from standard approximations for solids, such as the LDA or Perdew-Burke-Ernzerhof (PBE) generalised gradient approximation[9] (GGA)? What about observables in this numerically exact KS system, and in particular, the band gap? How much does  $v_{xc}$  depend on details of the density? And if it depends significantly, do the resulting changes have an impact on other KS observables?* Starting from nearly numerical exact densities[67] for the simple semiconductor bulk silicon and insulating sodium chloride obtained by the Auxiliary Field (AF) QMC method [68, 69] in Ref. [70], in the present work we answer these long-standing questions.

## II. HOW TO INVERT THE KS PROBLEM IN INFINITE SYSTEMS

The probably simplest algorithm to obtain the KS potential from a given density  $n_{\text{ref}}$  has been proposed for

finite systems by van Leeuwen and Baerends[27]. In its original form it was derived by solving the KS equations for the KS potential  $v_{\text{KS}}$ . The result was then translated into an iteration procedure which relates a potential  $v^{i+1}$  at step  $i + 1$  to the potential  $v^i$  at step  $i$  by the ratio of the target density  $n_{\text{ref}}$  and the density  $n^i$  at step  $i$ . As pointed out in [30], the best use of this ratio depends on the sign of the potential that is updated: for example,  $v$  may be either the usually negative total  $v_{\text{KS}}$ , or its rather positive interaction part  $v_H + v_{xc}$  with  $v_H$  the Hartree potential. In the present work we use

$$v_{xc}^{i+1}(\mathbf{r}) = \frac{n_{\text{ref}}(\mathbf{r}) + a}{\tilde{n}^i(\mathbf{r}) + a} v_{xc}^i(\mathbf{r}), \quad (1)$$

where  $a$  is a parameter that avoids instabilities in regions of very low density as suggested in [27], and the mixing  $\tilde{n}^i = \alpha n^{i-1} + (1 - \alpha)n^i$ , with  $0 < \alpha < 1$ , is introduced to smooth the convergence. This density  $\tilde{n}^i$  is also used to update the Hartree potential at each iteration. Eq. (1) is clearly a good strategy if  $v_{xc}$  is negative, and if the density at a point  $\mathbf{r}$  is determined only by the KS potential at that same point. Suppose that at a given iteration  $\tilde{n}^i(\mathbf{r})$  is larger than  $n_{\text{ref}}(\mathbf{r})$ . The algorithm then decreases the absolute value of  $v_{xc}(\mathbf{r})$ . If the exchange-correlation potential is negative, this step makes the potential more shallow, and less density will be attracted to the point  $\mathbf{r}$  in the next iteration, which pushes the solution in the good direction. Of course, it is not true that  $n(\mathbf{r})$  depends only on the KS potential in the same point  $\mathbf{r}$ , and it has to be seen to which extent the relation is near-sighted enough to make the algorithm work in a solid.

The negative sign of the potential that is updated in (1) is crucial for the algorithm to work, because a positive sign would drive the result in the wrong direction. However, contrary to the HEG, a real system can also exhibit regions of positive  $v_{xc}$ . Moreover, while in a low-dimensional system one can impose that the potential tends to zero at large distances, in a three-dimensional solid the zero of the potential is not defined. One cannot even use the ionization potential theorem, which would force the eigenvalue of the highest occupied state to be minus the ionization energy, since this theorem does not hold in extended systems[71]. The arbitrary energy scale represents both an advantage and a drawback. On the upside, it allows us to introduce a rigid negative shift such that the potential remains negative throughout the iteration. This shift is arbitrary within reasonable limits: if it is too small, positive regions may appear and become an obstacle for convergence. If it is too large, the algorithm becomes unstable, as the shift is multiplied at every step by the density ratio. Reasonable values lie within the maximum amplitude of the potential. On the downside, iteration of (1) yields  $v_{xc}$  only up to a constant. This is not due to our introduction of a shift, but to the fact that the density does not contain information about the absolute value of the potential[1]. Therefore, this limitation cannot be avoided. The resulting potential can, however, be used to calculate a well defined density and

KS observables such as the KS band structure (besides the meaningless constant shift).

We have tested the algorithm using known functionals. As documented in Appendix B, the inversion works straightforwardly for the LDA, but two further aspects have to be verified. On one side, the true functional, unlike the LDA, is non-local, which might influence the behavior of the algorithm. This is sorted out in Appendix C. Second, the QMC data contains statistical noise; this aspect is deepened in Appendix D. Finally, Appendix E shows that the inversion starting from the QMC density behaves as expected. It also demonstrates that the final results do not depend on the starting point of the iterative procedure, including starting points as far from the final result as, e.g.,  $0.1 \times v_{xc}^{LDA}$ . For the results shown in the following, we use as starting point  $0.3 \times v_{xc}^{LDA}$  with a rigid downwards shift of 0.2 Hartree for silicon and 0.4 Hartree for NaCl. In the following, we report results in atomic units for both densities (expressed in bohr<sup>-3</sup>) and xc potentials (in Hartree).

### III. RESULTS

#### A. Kohn-Sham potential of silicon and sodium chloride

We have applied the algorithm to obtain the exchange-correlation potential from the charge density obtained by AFQMC calculations. For silicon, we have used the results of Ref. [70]. For NaCl, we have applied additional symmetry operations to the density from the same Ref. [70]. Ideally, from the inversion procedure for iteration  $i \rightarrow \infty$  one should find  $n^{\text{QMC},i}(\mathbf{r}) \rightarrow n_{\text{QMC}}(\mathbf{r})$ . However, since the QMC data contains statistical noise, the inversion has a more limited precision than in the case of, e.g., clean LDA data (see App. D). The Mean Absolute value of the Percentage Error (MeAPE) of the density of silicon  $n^{\text{QMC},i}$  at iteration  $i$  compared to  $n_{\text{QMC}}$ ,  $100 \times \text{mean}_{\mathbf{r}} |1 - n^{\text{QMC},i}(\mathbf{r})/n_{\text{QMC}}(\mathbf{r})|$ , does not fall below 0.02%, while the Maximum (over the unit cell) of the Absolute value of the Percentage Error (MaAPE) of the density,  $100 \times \max_{\mathbf{r}} |1 - n^{\text{QMC},i}(\mathbf{r})/n_{\text{QMC}}(\mathbf{r})|$ , decreases to 0.38% at  $i = 20$  iterations. This is in any case sufficient to make significant distinctions between different densities and potentials. The upper panel of Fig. 1 shows the Local Percentage Difference (LPD) of the iterative density with respect to the QMC density after 20 iterations (blue line),  $100 \times (n^{\text{QMC},20}(\mathbf{r})/n_{\text{QMC}}(\mathbf{r}) - 1)$ , along a path through the unit cell (the same as in Ref. [70], see the inset to the second panel of Fig. 1). The result stays well within the stochastic error bar of the QMC calculation (grey area). For comparison, we also show the LPD of the LDA and PBE densities (dot-dashed orange and dashed green lines, respectively), with respect to the QMC. As also shown in Ref. [70], differences between LDA, PBE and the QMC densities are largest on the atoms and also in other regions of low density[72]

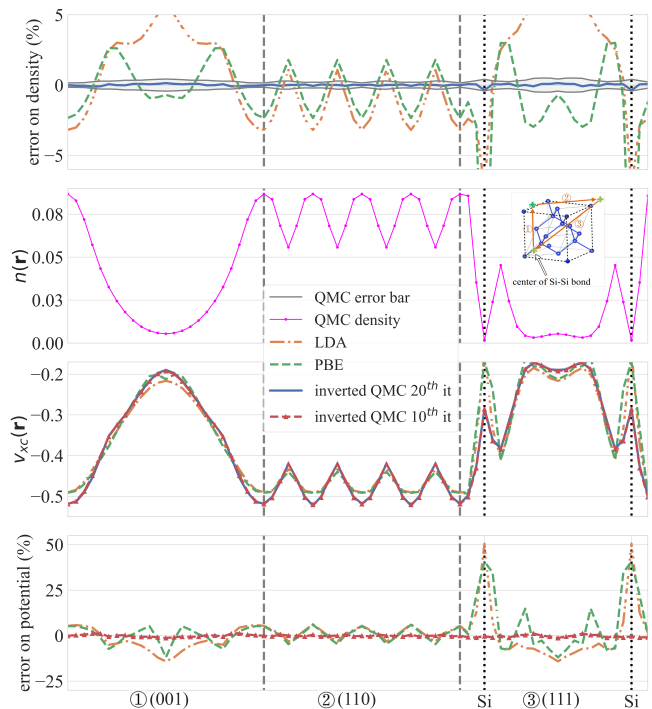


FIG. 1. Density and xc potential of bulk silicon along the same path across the unit cell as in Ref. [70]. The positions of atoms are indicated by dotted vertical lines. The iterative inversion follows Eq. (1) with the QMC density  $n_{\text{QMC}}$  of silicon as reference density. The potential  $v_{xc}^{\text{QMC},20}$  is obtained after  $i = 20$  iterations. The density  $n^{\text{QMC},20}$  is calculated using  $v_{xc}^{\text{QMC},20}$  in the KS equation. The MaAPE at  $i = 20$  compared to  $n_{\text{QMC}}$  is 0.38 %, and the MeAPE is 0.04%. Top panel: LPD of  $n^{\text{QMC},20}$  (blue), self-consistent LDA  $n_{\text{LDA}}$  (orange), and PBE  $n_{\text{PBE}}$  (green) densities with respect to  $n_{\text{QMC}}$ . The grey area is the stochastic error bar of the QMC density. Second panel: The QMC density  $n_{\text{QMC}}$  (magenta line). The inset shows the chosen path across the crystal from Ref. [70]. Third panel:  $v_{xc}^{\text{QMC},20}$  (blue),  $v_{xc}^{\text{LDA}}$  (orange),  $v_{xc}^{\text{PBE}}$  (green), and  $v_{xc}^{\text{QMC},10}$  (red). Note that the two QMC potentials (blue and red lines) are almost indistinguishable. The average potentials are aligned. Bottom panel: LPD of xc potentials with respect to  $v_{xc}^{\text{QMC},20}$  for LDA (orange), PBE (green), and QMC at  $i = 10$  (red).

(see the magenta line in the second panel of Fig. 1), but they are still significant in regions of higher density, along the (110) direction, where LDA and PBE are very similar, but differ from the QMC result. Most importantly, the differences between different densities are much larger than the error due to the inversion of the QMC density: while the MeAPE at  $i = 20$  is 0.04%, the mean absolute relative difference between the LDA and QMC densities is 1.93 %, and it is 1.07 % between the PBE and QMC densities.

The xc potentials are compared in the third panel of Fig. 1. For this comparison, the potentials are aligned at their average value. Our numerically determined and supposedly most accurate KS xc potential, obtained from

the QMC density, is similar to the local and semi-local approximations. This result is stable: the QMC result obtained at  $i = 10$ , where the MaAPE and MeAPE on the density are 0.90 % and 0.09 %, respectively, is almost indistinguishable from the one at  $i = 20$ . The differences between QMC on one side, and LDA and PBE on the other side, can be appreciated in the bottom panel, which shows the LPD of LDA and PBE with respect to the QMC xc potential obtained at  $i = 20$ . These differences are similar for LDA and PBE along most of the path. The MeAPE with respect to the QMC result for potentials is 3.90 % for the LDA and 3.88 % for the PBE: of similar order, though larger, than the MeAPE of the densities. Instead, the LPD of the QMC potential at  $i = 10$  with respect to the potential obtained at  $i = 20$  can hardly be seen. We have hence reached sufficient precision on the density, which lies within the QMC error bar, and the xc potential, which shows some differences with respect to common functionals. The effect of iterating further using the noisy QMC data is discussed in the App. D.

Results for sodium chloride show similar trends, with even better convergence properties of the potential due to the fact that our QMC density for NaCl is less noisy than the one of silicon in the important regions of high density (see also App. D). For the density, we obtain a MeAPE of 0.03% and a MaAPE of 0.29% at  $i = 200$ . Here, analogous to Fig. 1 for silicon, in Fig. 2 we show the LPD of the density and of the xc potential along a path (see the inset to the second panel). The QMC-derived xc potential differs from the LDA and PBE especially on the sodium atoms, where the density shows rapid changes. At first sight, however, and as in the case of silicon, it is difficult to rationalize the differences between the three potentials. While it is exciting to see the numerically exact xc potential for real semiconductors and insulators, it is useful to switch to a representation that highlights the essence of the difference, in order to gain more insight.

### B. Non-local dependence of the KS exchange-correlation functional on the density

In order to appreciate the non-locality of the QMC derived potential, we will again compare it to known functionals. It is interesting to add the TranBlaha modified Becke-Johnson potential[73] to the more common functionals LDA and PBE, since this functional is designed with a different purpose, namely, to yield an eigenvalue gap closer to the measurable fundamental gap than the exact KS potential. This requires a compromise concerning the resulting density. The deviation of the latter from the QMC one is shown in the top and bottom panels of Fig. 3 and for silicon and NaCl, respectively. The density errors have opposite sign with respect to those of the LDA throughout, and they are of similar order of magnitude in silicon, and about 4-5 times larger in NaCl. This raises the question how the potential and KS observables will compare.

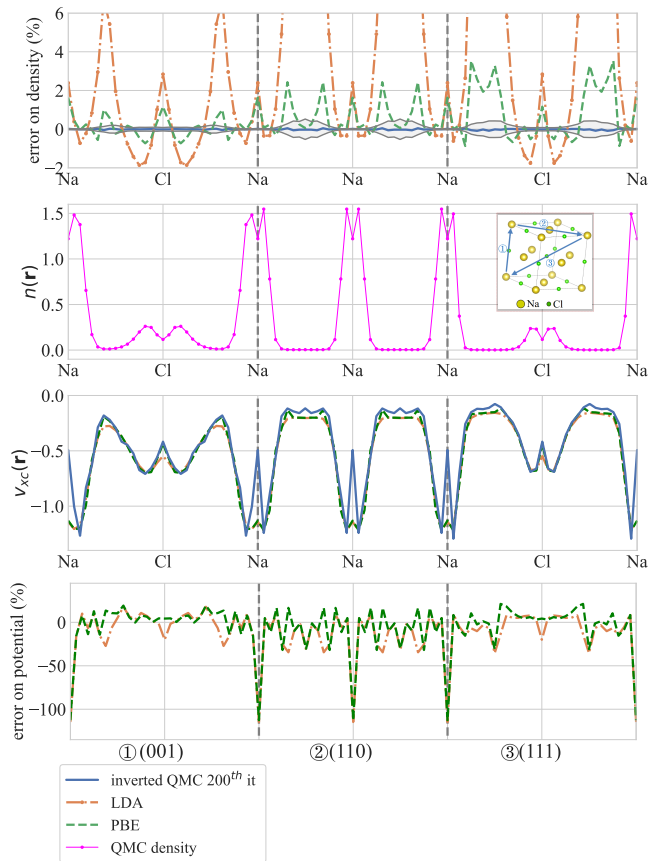


FIG. 2. Density and xc potential of NaCl along the same path across the unit cell as in Ref. [70]. The positions of atoms are indicated by dotted vertical lines. The iterative inversion follows Eq. (1) with the QMC density  $n_{\text{QMC}}$  of NaCl as reference density. The result of the QMC inversion is shown at  $i = 200$ . The MaAPE on the density at  $i = 200$  compared to  $n_{\text{QMC}}$  is 0.29 %, and the MeAPE is 0.03%. Top panel: LPD of  $n_{\text{QMC},200}$  (blue), self-consistent LDA  $n_{\text{LDA}}$  (orange), and PBE  $n_{\text{PBE}}$  (green) densities with respect to  $n_{\text{QMC}}$ . The grey area in the top panel indicates the stochastic error bar of the QMC density. Second panel: The QMC density  $n_{\text{QMC}}$  (thin magenta line). The inset shows the chosen path across the crystal from Ref. [70]. Third panel:  $v_{\text{xc}}^{\text{QMC},200}$  (blue),  $v_{\text{xc}}^{\text{LDA}}$  (orange) and  $v_{\text{xc}}^{\text{PBE}}$  (green). The averages of the potentials are aligned. Bottom panel: LPD of xc potentials with respect to  $v_{\text{xc}}^{\text{QMC},200}$  for LDA (orange) and PBE (green).

In the LDA,  $v_{\text{xc}}(\mathbf{r})$  is a monotonic function of  $n(\mathbf{r})$ . The exact KS potential is a functional of the density everywhere, which means that it can take different values in different points  $\mathbf{r}$  where the density, instead, is the same. This expresses the fact that  $v_{\text{xc}}(\mathbf{r})$  depends not just on the local density, but also on the environment. In order to highlight how the true  $v_{\text{xc}}$  differs from a function of the local density, we create a map of  $v_{\text{xc}}(\mathbf{r})$  with respect to  $n(\mathbf{r})$ : for each point  $\mathbf{r}$  in real space, we add a point  $[v_{\text{xc}}(\mathbf{r}) \leftrightarrow n(\mathbf{r})]$  to Fig. 4 and Fig. 5 for silicon and NaCl, respectively. In the case of the LDA, this plot shows the universal function  $v_{\text{xc}}(n)$ , the same for silicon and NaCl,

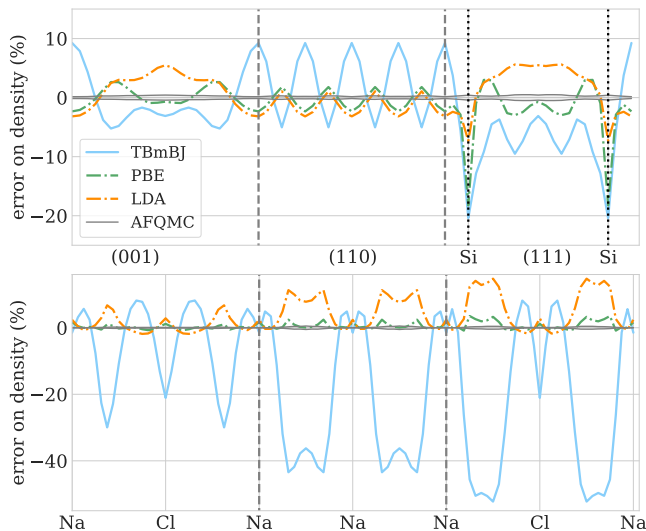


FIG. 3. LPD of self-consistent LDA  $n_{\text{LDA}}$  (orange), PBE  $n_{\text{PBE}}$  (green) and TBmBJ (light blue) densities with respect to  $n_{\text{QMC}}$ . The grey area is the stochastic error bar of the QMC density. (Top panel) Silicon. (Bottom panel) NaCl.

which is identical to the function in the HEG. Beyond the LDA, different environments may change this function, such that it is different for different materials. Moreover, in one and the same material the presence of different environments may lead to the presence of more than one function, and finally, if the result is very sensitive to details, one might find it difficult to detect anything like a limited number of functions. All these effects are possible signatures of the non-local dependence of  $v_{\text{xc}}(\mathbf{r})$  on the density, and being able to discern them, and to characterize different environments, may give precious input for further modeling of  $v_{\text{xc}}$ .

The maps in Fig. 4 and 5 contain the results of LDA, PBE, TBmBJ and QMC. The universal LDA result is given for reference in all panels of Fig. 4 and 5. LDA, PBE and TBmBJ potentials are directly obtained from the KS calculation; only the QMC result stems from the inversion. In silicon, the PBE result (upper panels of Fig. 4) is dominated by a simple monotonic function, but it is slightly steeper than the LDA function. Moreover, the result appears to be a little more scattered. Finally, a new branch appears at low densities. For more insight, the colors reflect, respectively, the modulus of the local gradient of the density (left panels) and the KS kinetic energy density defined as  $\tau(\mathbf{r}) = \frac{1}{2} \sum_i^{\text{occ}} |\nabla \phi_i(\mathbf{r})|^2$ , where  $\phi_i$  are KS orbitals (right panels). The TBmBJ potential in the bottom panels is similar to the PBE, with an extra upwards branch at low densities, some blurring at higher densities and a tendency to be lower than the LDA result at high density, but the latter two features appear to be much stronger than in the PBE.

The QMC xc potential in the middle panels is also blurred. However, this cannot be interpreted as a reliable feature of the true KS potential, since the QMC data is

noisy, which may also cause blurring, as we have demonstrated in App. D. The overall shape and branches of the QMC xc potential, instead, are significant. Similarly to the PBE and TBmBJ, one can identify a dominant curve, and, with respect to the LDA, two main changes are seen: the curve is steeper than that of the LDA, and at low densities an additional branch appears. The change in slope of the main branch with respect to the LDA goes in the same direction as in the PBE and TBmBJ results, and it is more pronounced than in the case of PBE, similarly to TBmBJ. Also the branching happens in a similar region as in the case of PBE and TBmBJ. However, the branch departs in the opposite direction.

In all cases, the extra branch is characterized by a very different gradient and kinetic energy density with respect to the main branch at the same local density. Indeed, the region in space where the potential lies on the extra branch is close to the center of the atoms, where the density varies rapidly. It should therefore be noted that it will be particularly sensitive to details of the pseudopotential. This, together with the fact that the inversion error on the density is largest on the atoms (see Fig. 11 in the Appendix), which then also influences the large density gradients in the vicinity, suggests that the different directions of the branches observed here would deserve more studies including many more QMC datasets using different pseudopotentials and including different materials, which is beyond the scope of the present work. The changed slope of the main branch, instead, happens over the complete range of densities and should be a feature of silicon independent of the pseudopotential and other ingredients of the calculation.

The modifications of the different branches with respect to the LDA  $v_{\text{xc}}^{\text{LDA}}$  may be translated in different ways, for example:  $v_{\text{xc}}^e(\mathbf{r}) = F^e(n(\mathbf{r}))v_{\text{xc}}^{\text{LDA}}(n(\mathbf{r}))$  with a correction factor  $F^e$  that depends on the local density and on an environment  $e$ , which must be characterized. Another possibility would be  $v_{\text{xc}}^e(\mathbf{r}) = v_{\text{xc}}^{\text{LDA}}(\mathcal{F}^e(n(\mathbf{r})) \times n(\mathbf{r}))$ . The GGAs, for example, are an attempt to characterize the environment by the local gradient of the density (see, e.g., [74]). Our results motivate further search for improved approximations of the true *functional* that can be expressed as *functions* of a limited number of parameters, such as the local density and its gradients.

Consistently with the fact that the QMC density for NaCl is less noisy than in the case of silicon, the map for NaCl in Fig. 5 (middle panels) shows less scattering. As for silicon, we find a main branch that corresponds to a modified LDA. Moreover, there is an additional branch at low densities and another branch at high density, both characterized by differences in the gradient or kinetic energy density. The analogous secondary branches for PBE (see upper panels) are less pronounced whereas these features are much stronger in TBmBJ. The inset in the middle right panel also shows the QMC density along part of the path. Numbers indicate to which locations selected data points correspond. For example, data point 1 on the

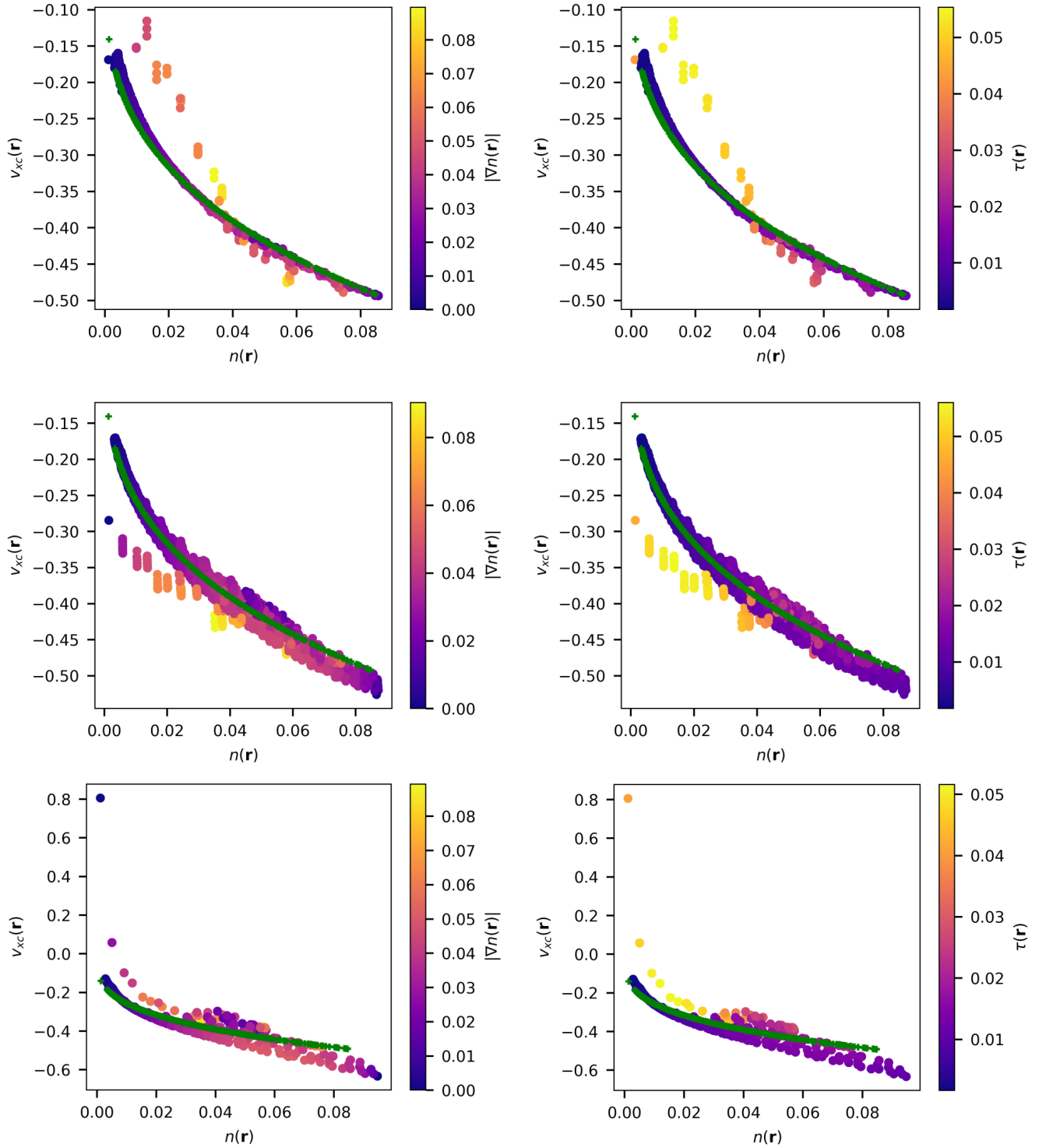


FIG. 4. Map of the xc potential of silicon with respect to the local density at all points in the unit cell. Color codes reflect the modulus of the local gradient of the density (left column) or the local kinetic energy density (right column). Upper figures are for PBE, middle figures for QMC, and bottom figures for TBmBJ. The LDA is shown in green.

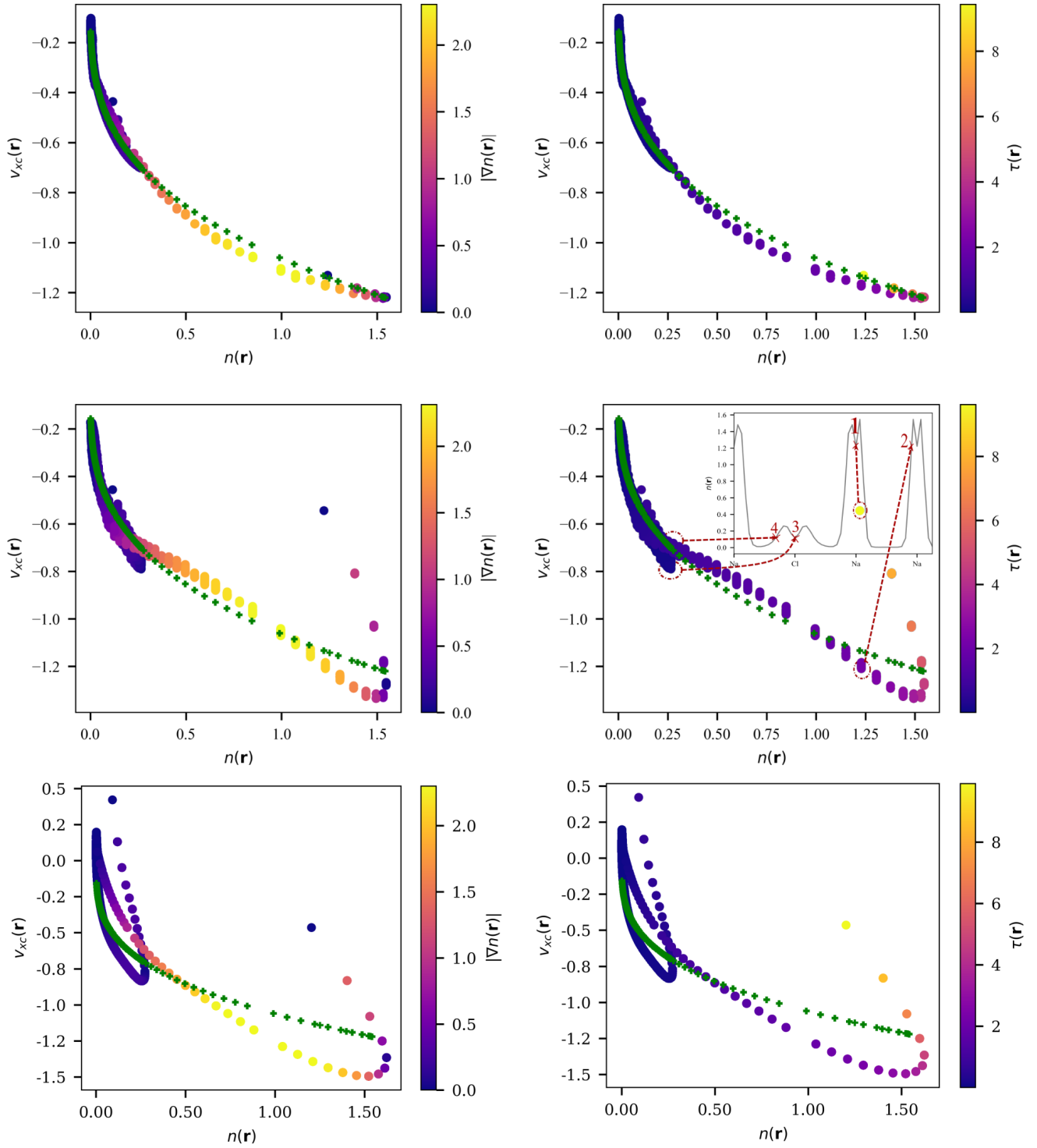


FIG. 5. Same as Fig. 4, for NaCl. Note the much larger range for  $n(\mathbf{r})$  here. Upper panels are for PBE, middle panels for QMC, and bottom panels for TBmBJ. The LDA is shown in green. The inset in the middle right panel shows the QMC density along a part of the path across the unit cell. Note that a data point in the main graph (marked by 1) falls into the inset area.



additional high-energy branch corresponds to the potential on the sodium atom, with an environment where the density is very quickly varying, which explains why the LDA completely fails. Data point 2, instead, corresponds to a place with similar density but located in a more gentle environment, although the gradient of the density is significant. As expected, in this point we are on the main branch, which is, however, modified with respect to the LDA.

Similarly, points 3 and 4 on the chlorine atom explain the extra branch at lower density. These results show that the potential-density relation presented as maps such as the one in Figs. 4 and 5 may give further insight about the most efficient way to introduce correction factors, and about the most important features distinguishing different environments, which could eventually be combined with machine learning approaches [75].

### C. Kohn-Sham approximation to the band gap

It is most interesting to look at KS observables other than the density, in particular, KS eigenvalues: even though, as discussed above, these do not by themselves have direct physical meaning in an exact sense, they can still be seen as an approximation to the physical quantities [76, 77], and they are frequently used as starting points for calculations in a more appropriate framework, such as many-body perturbation theory[78]. In particular, the study of the KS band gap is interesting by itself, since it is a matter of a long-standing debate. In absence of knowledge of the exact Kohn-Sham potential, it was not possible to distinguish between the discrepancies due to approximations of the functional, and those due to the difference between the (even exact) Kohn-Sham system itself and the real material. Precious hints were already given by work on model systems; for example, Knorr and Godby [34, 35] determined the xc potential by inversion from the density of a finite one-dimensional model semiconducting wire that was then extrapolated to infinite length. Most of the band gap error was shown to be due to the fact that the exact KS eigenvalue gap differs from the fundamental electron addition and removal gap, and not due to approximations. Indeed, the KS eigenvalue gap calculated at fixed particle number disregards the derivative discontinuity of the exact xc potential upon change of particle number [12–14]. Since the numerically exact density and/or xc potential could be obtained only for very few, low-dimensional, systems, several studies used the link between the xc potential and the self-energy given by the Sham-Schlüter equation [13] in order to extract  $v_{xc}$  from the self-energy. These include work on a two-plane wave model [13, 79], the surface barrier for semi-infinite jellium [80], finite systems [21, 81–83], and the study of several real semiconductors and insulators [18, 20–26]. These studies confirmed that the error inherent in using Kohn-Sham eigenvalues instead of true electron addition and removal energies is

significant. However, the approaches used to determine the potential involved themselves approximations whose quantitative impact on the findings are not known: first, the Sham-Schlüter equation was linearized in all studies; second, the self-energy itself was approximated in many-body perturbation theory, mostly on the GW[84] level. With the present work, we finally do have an almost numerically exact Kohn-Sham potential at hand for real materials, and we can therefore draw definite conclusions concerning the band structure, and in particular the band gap, of standard semiconductors and insulators.

With the fact in mind that errors of the potential can be much larger than errors of the density, the quality of the gap resulting from inversion has to be checked separately. To this end, we show in Fig. 6 the direct KS band gap of silicon (upper panel) and sodium chloride (lower panel) at  $\Gamma$  as a function of the number of iterations at which the KS potential and corresponding density were calculated. For all functionals, also when noise is included, the result converges very rapidly and remains stable, within 1 meV, over a wide range of iterations even after the potential has developed huge spikes (in the case of noisy density for silicon). This means that the results for the band gaps are reliable with high precision.

Results for the converged band gaps of silicon and NaCl are shown in Table I. For silicon, our numerically exact minimum indirect KS band gap is 0.69 eV, about 40 % larger than the KS gap of 0.49 eV calculated in LDA, and significantly smaller than the experimental gap [85] of 1.17 eV. The PBE gap of 0.66 eV is close to the QMC-derived value. The direct band gap opening of QMC with respect to LDA is analogously 0.17 eV at  $\Gamma$  and 0.12 eV at X. TBmBJ yields a direct gap of 3.09 eV, close to experiment. In NaCl, the situation is similar, with the QMC-derived gap about 14% larger than the LDA one, and only 3% larger than the PBE gap. The 5.25 eV QMC-derived KS gap is again much smaller than the 8.5 eV experimental gap [86], while TBmBJ is again close to experiment, with a direct gap of 8.93 eV. The QMC bandwidths do not change in a noteworthy way with respect to LDA (or PBE): in silicon the QMC valence bandwidth is reduced by 0.1 eV compared to LDA (and 0.05 eV compared to PBE), while in NaCl the QMC bandwidth is 0.15 eV smaller than LDA and 0.04 eV larger than PBE.[87]

Our QMC derived KS gaps confirm the conclusion of Ref.[18, 20] and thus definitely highlight the fact that the true multiplicative KS potential does not yield a “good” eigenvalue band gap in solids. Overall, the band gap is an excellent illustration for the fact that the exact Kohn-Sham system is an auxiliary system designed to yield the density in principle exactly, but for other observables, it can only give an approximation. At the same time, the TBmBJ results illustrate that by enhancing the deviations of the true potential from the LDA one can design a multiplicative potential which overcorrects the LDA errors on the density and on the KS eigenvalue gap, bringing the latter close to the experimental fundamental gap.

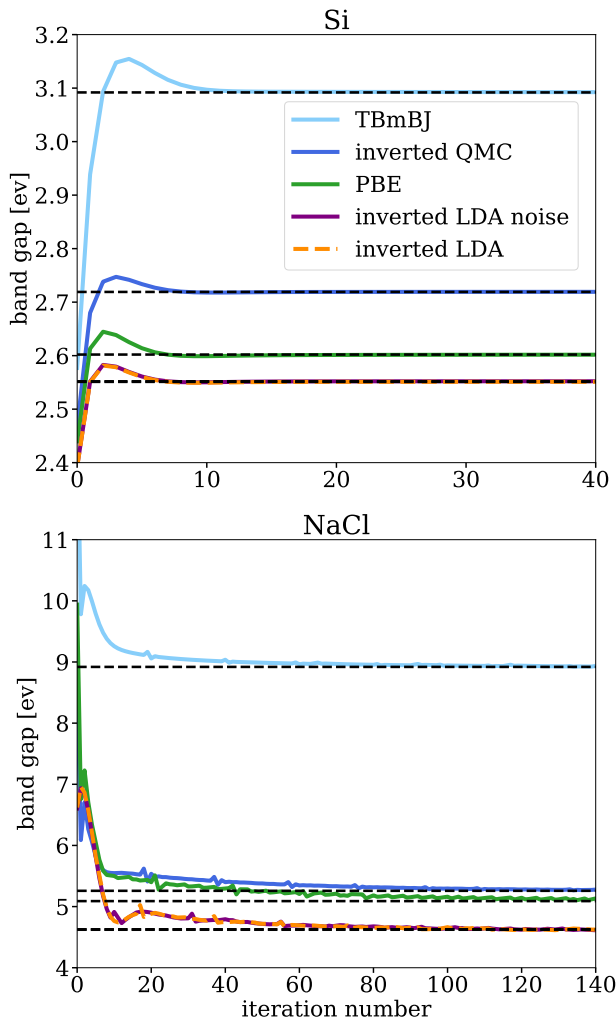


FIG. 6. Convergence with the number of iterations of the direct band gap at  $\Gamma$  obtained from the inversion, for silicon (upper panel) and NaCl (lower panel). Shown are the results of LDA (orange), noisy LDA (purple), PBE (green), TBmBJ (light blue) and QMC (blue). The horizontal continuous lines are the reference results, which are obtained directly from the KS calculation except for the noisy LDA where the reference is the result at  $i_0 = 24$ , and for QMC, where the reference is the result at  $i_0 = 20$ . The clean and noisy LDA results are almost entirely overlapping and converge to the reference result of clean LDA, 2.55 eV. The QMC result converges to 2.72 eV.

This may yield a compromise between the accuracy of the density and the search for a simple estimate of the fundamental gap.

There is another interesting aspect to this study of the band gap: while in certain cases (in particular, the noisy LDA) the potential can develop huge spikes during the iterations, the gap, similarly to the density, remains close to the reference value. This means that very different xc potentials can yield not only very similar densities, but also very similar KS observables more in general. Fig. 6

	Si		NaCl
	indirect	direct at $\Gamma$	direct at $\Gamma$
QMC derived	0.69	2.72	5.25
PBE	0.66	2.60	5.08
LDA	0.49	2.55	4.59
TBmBJ	1.19	3.09	8.93
Exp.	1.17 [85]	3.05[88] 3.40[85]	8.5[86]

TABLE I. KS minimum band gaps and direct band gaps at  $\Gamma$  (eV) in comparison with experimental photoemission gaps from Refs. [85, 86, 88].

also shows that the gaps corresponding to clean and noisy LDA densities are almost indistinguishable, i.e., the noise does not affect KS observables. The band gap results confirm the statement, mostly based on findings from low-dimensional systems, that examining the xc potential alone is not sufficiently meaningful[89–91]. Our study also suggests that an effort is needed to distinguish in the KS potential crucial features, which must be contained in good functionals, from others that may be quantitatively strong in the potential, but insignificant for its effects.

#### IV. CONCLUSION

In conclusion, we have shown that a simple algorithm allows one to obtain the Kohn-Sham xc potential for periodic semiconductors and insulators, given their ground state density. The precision that can be obtained is limited by the quality of the input data. Here, we use densities taken from AFQMC calculations, and the limiting factor is the stochastic noise. Nevertheless, meaningful results are obtained, with an error bar smaller than the difference between the resulting potentials and their LDA PBE, or TBmBJ counterparts, which allows us to safely draw conclusions. In particular, for the materials studied here, namely bulk silicon and NaCl, the xc potential *functional of the density everywhere* can be represented in terms of two or three *functions of the local density*, each of which is determined by a specific environment. These environments appear to be characterized by the local gradient of the density or, even more clearly, by the local kinetic energy density. The function that represents most of the data points is close to the LDA, but with slight material-dependent deviations. PBE also predicts deviations and the existence of the additional functions, although it does not always describe them well. On the other hand, our results clearly illustrate that very different potentials may lead to very similar densities, and more generally, to very similar KS observables. In particular, the KS band gap converges rapidly with the number of iterations of the inversion process, while the xc potential still undergoes violent modifications. More work is needed to discern important features of the xc potential

from those that do not influence KS observables; sum rules and other exact constraints may be helpful for this [8–11]. Our results for the KS band gap confirm previous conjectures based on model systems and/or many-body perturbation theory, which predict that the exact KS band gap is closer to the LDA one than to the measurable electron addition and removal gap, in other words, that the derivative discontinuity of the true xc potential is sizable. Still, the LDA error is non negligible, whereas PBE predicts the exact KS gap with an error of less than 5% for the materials studied here. Our work highlights directions for the improvement of density functionals, stressing the need for, and usefulness of, QMC calculations of the density in many more materials.

## V. ACKNOWLEDGEMENTS

Fruitful discussions with Kieron Burke and Rex Godby are acknowledged. S.C. was supported by the U.S. Department of Energy (DOE) under Grant No. DE-SC0001303, and by the Center for Computational Quantum Physics, Flatiron Institute. The Flatiron Institute is a division of the Simons Foundation.

### Appendix A: Computational details

We have adopted the same computational parameters (lattice constants, cutoff energies and k-point grids) and pseudopotentials as in Ref. [70]. Following Ref. [70], all the LDA, PBE, TBmBJ and QMC results have been obtained with the same optimized norm-conserving LDA pseudopotentials[92]. We have employed the Abinit code[93] and verified that it gives the same numerical results as the Quantum Espresso code[94] used in Ref. [70] for LDA and PBE. For the calculations with the TBmBJ functional, composed of a modified version of the Becke-Johnson exchange potential[95] and a LDA correlation part, we have used the Abinit implementation[96] with the libxc library[97]. The KS inversion algorithm has been implemented in our own KS python code, which is interfaced with the Abinit code.

### Appendix B: Accuracy of the algorithm: inverting the LDA

To illustrate the reliability of the inversion algorithm, it is instructive to examine a case where the density-potential relation is well known; as a start, we choose the LDA [98]. This means that in Eq. (1),  $n_{\text{ref}} = n_{\text{LDA}}$  is the density obtained in a standard LDA Kohn-Sham self-consistent calculation with exchange-correlation potential  $v_{\text{xc}}^{\text{LDA}}$  at convergence. Ideally, from the iteration procedure for  $i \rightarrow \infty$  we should find  $v_{\text{xc}}^{\text{LDA},i} \rightarrow v_{\text{xc}}^{\text{LDA}}$  and  $n^{\text{LDA},i} \rightarrow n_{\text{LDA}}$ . Fig. 7 and 8 show results for silicon.

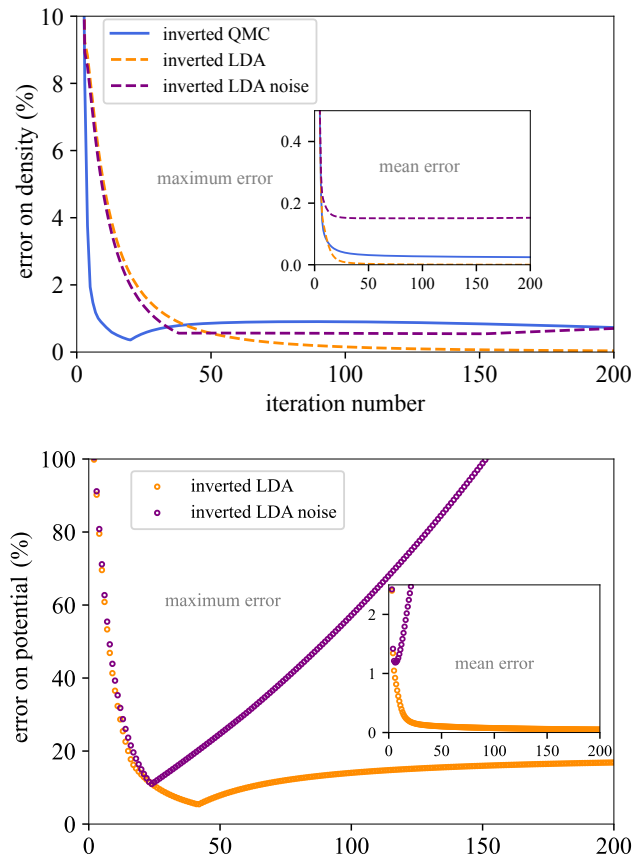


FIG. 7. Errors of the iteration procedure as a function of number of iterations  $i$  in silicon. Top panel: MaAPE of the density for the inverted LDA (orange), inverted noisy LDA (purple) and QMC (blue) xc potentials. Each inverted density is compared to its corresponding reference result. In the inset: MeAPE of the density for the inverted LDA, noisy LDA and QMC xc potentials. Bottom panel: MaAPE of the xc potential for the inverted LDA (orange) and inverted noisy LDA (purple). In these cases the error is defined with respect to  $v_{\text{xc}}$  of KS LDA, since the xc potential that yields the noisy density is unknown. In the inset: MeAPE of the xc potential for the inverted LDA and inverted noisy LDA potentials.

In Fig. 7 (upper panel) the Maximum (over the unit cell) of the Absolute value of the Percentage Error (MaAPE) of the density  $n^{\text{LDA},i}$  compared to the LDA one,  $100 \times \max_{\mathbf{r}} |1 - n^{\text{LDA},i}(\mathbf{r})/n_{\text{LDA}}(\mathbf{r})|$ , is shown as a function of the iteration number  $i$ . It decreases smoothly and very fast. The same is true for the Mean Absolute (over the unit cell) Percentage Error (MeAPE), given in the inset.

In Fig. 8 snapshots for the errors on density and  $v_{\text{xc}}$  are presented. The upper panel gives  $100 \times (n^{\text{LDA},i}(\mathbf{r})/n_{\text{LDA}}(\mathbf{r}) - 1)$ , the Local Percentage Difference (LPD) along a path through the unit cell (the same as in Ref. [70]) of the density at  $i = 500$  iterations (orange dashed line) with respect to the LDA one. In the scale of the figure, it is close to zero everywhere: it is

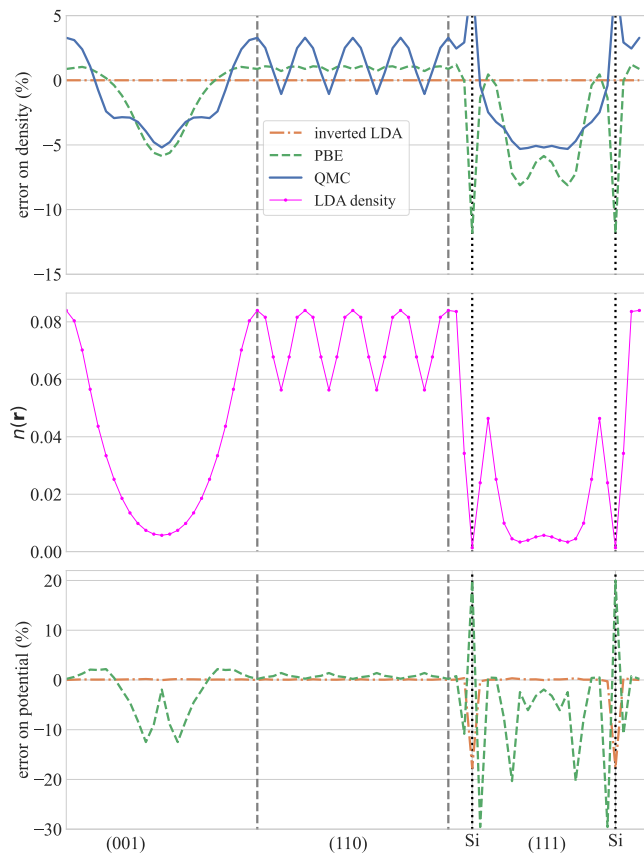


FIG. 8. Upper panel: The LPD of the density: in orange, LPD of  $n^{\text{LDA},500}$  with respect to  $n_{\text{LDA}}$  (where the MaAPE is  $6.55 \times 10^{-4}\%$ ); in green LPD of the self-consistent PBE and in blue QMC densities with respect to  $n_{\text{LDA}}$ . Middle panel: Self-consistent LDA density  $n_{\text{LDA}}$  (thin magenta line). Lower panel: LPD of  $v_{\text{xc}}^{\text{LDA},500}$  (orange) and self-consistent PBE potential  $v_{\text{xc}}^{\text{PBE}}$  (green) with respect to  $v_{\text{xc}}^{\text{LDA}}$ .

largest, with a maximum of  $6.55 \times 10^{-4}\%$ , in places of low LDA density, shown by the thin magenta line in the middle panel. Note that the density is close to zero on the silicon atoms. The LPD of the potential is shown in the bottom panel of Fig. 8. Because of the arbitrary shift, only a comparison of the variations of the potential is pertinent. Indeed, during the iterations the average potential continuously moves upwards. The figure has been obtained by re-aligning at the end of the iterations the average potentials. This requires a final downwards shift of the iterative potential of 0.06 Hartree. Again, the maximum error is found in places of lower density. The only significant error is found in the point exactly on the silicon atoms, where the density is almost vanishing. As we will also see below for other functionals, the result in this specific point cannot be considered to be reliable. Also for other regions, however, it is true that the LPD is significantly larger for the potential than for the density. This can also be appreciated in Fig. 7 (bottom panel), where the open circles in the main panel give the MaAPE and MeAPE on the xc potential as a function

of iterations. In order to illustrate that the remaining errors are small enough to make discussions meaningful, the errors in Fig. 8 are compared to the difference between two different functionals obtained directly from the KS calculation, here, between LDA and PBE (see green lines). The top panel contains the LPD of the PBE density with respect to the LDA self-consistent one,  $100 \times (n_{\text{PBE}}(\mathbf{r})/n_{\text{LDA}}(\mathbf{r}) - 1)$ . This difference can be seen very clearly, since it is more than  $10^4$  times larger than the inversion error. The LPD of the PBE with respect to the LDA xc potential,  $100 \times (v_{\text{xc}}^{\text{PBE}}(\mathbf{r})/v_{\text{xc}}^{\text{LDA}}(\mathbf{r}) - 1)$ , can be found in the bottom panel. Differences can be seen throughout the path, although regions of lower density show larger differences. These differences are, though slightly larger, of the same order as the differences in the density. Except for immediate vicinity of the atoms, they are much larger than the error of the inversion, like in the case of the density. This demonstrates that the inversion yields meaningful results, with an error bar that is much smaller than the differences of interest, except for few points of very small density.

### Appendix C: Inversion of non-local functionals

The true xc functional is expected to be non-local, and testing the LDA alone is therefore not sufficient. For a conclusive test, we should invert functionals with a similar degree of non-locality as the QMC results. Figs. 4 and 5 show that the deviation from a local potential-density relation of the QMC-derived potential is larger than for the PBE and smaller than for the TBmBJ functional. This means that PBE and TBmBJ cover the range of non-locality of the true KS potential. We have therefore also performed the inversion tests with these two functionals, with the results given in Figs. 9 and 10.

As one can see, except for the points of lowest density the agreement between the reference result and the result of the inversion is again very good, and definitely accurate enough to distinguish the main features of the potentials. We conclude that non-locality in the range of that of the true KS potential does not hamper the use of our inversion algorithm.

### Appendix D: Inversion starting from noisy densities

To conclude on the reliability of the results, we also have to take into account the fact that the QMC results have statistical errors. In order to elucidate the influence of the QMC stochastic noise on the results, we have taken the LDA density  $n(\mathbf{r})$  and added point-wise Gaussian noise, which is obtained from a normal distribution with mean zero and standard deviation given by the characteristic AFQMC statistical error bar scaled by  $n(\mathbf{r})$ . To appreciate what this means, the yellow curve in the upper panel gives the relative difference of the noisy LDA density with respect to the clean one. Fig.

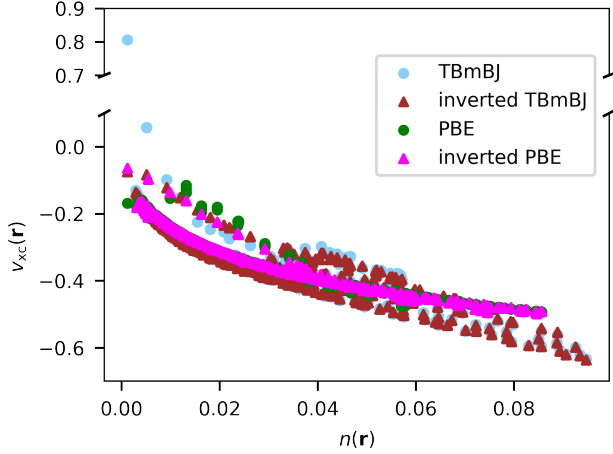


FIG. 9. Map of the xc potential of silicon with respect to the local density at all points in the unit cell. The results of the KS cycle (circles) are compared to the result of inversion (triangles) for PBE (green) and TBmBJ (light blue).

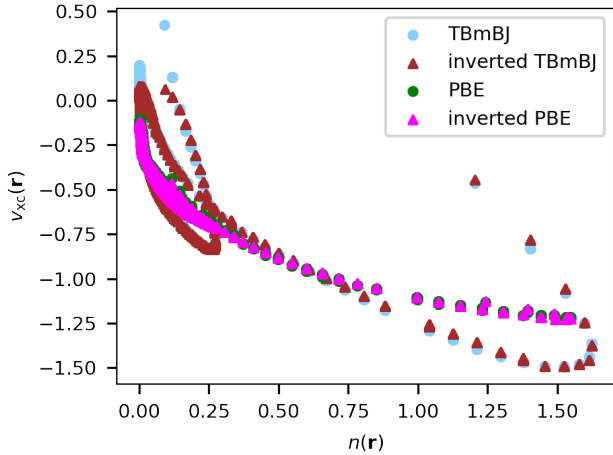


FIG. 10. Map of the xc potential of NaCl with respect to the local density at all points in the unit cell. The results of the KS cycle (circles) are compared to the result of inversion (triangles) for PBE (green) and TBmBJ (light blue).

11 shows results for the inversion starting from this new reference density. The error of the inversion of the noisy LDA data is displayed in Fig. 7 (upper panel, purple curve). It behaves similarly to the QMC inversion error: the MaAPE decreases rapidly and reaches a minimum, from whereon a slight increase followed by a decrease is noted. The MeAPE, instead, reaches a plateau. The inversion error on the density is given by the purple curve in the upper panel of Fig. 11, representing the LPD  $100 \times (n^{\text{LDA}+\text{noise},39}(\mathbf{r})/n^{\text{LDA}+\text{noise}}(\mathbf{r}) - 1)$  at 39 iterations. It is of similar magnitude as the noise itself, as in the QMC case. Iterating further to  $i = 150$  (red), only a slight smoothing of the error on the density is observed.

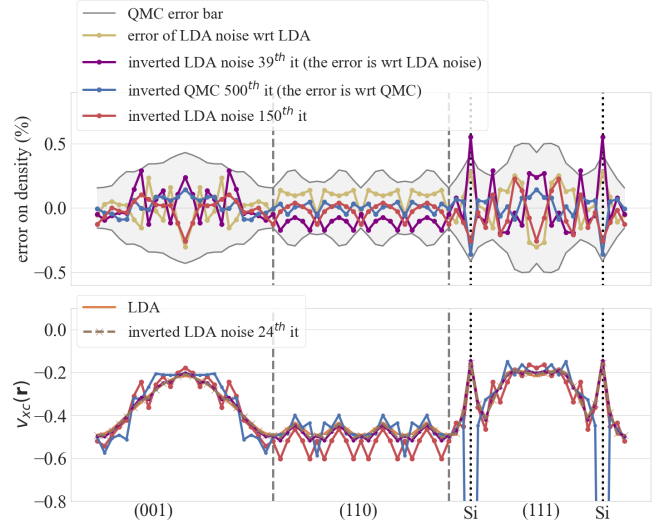


FIG. 11. Relation between noise and errors of the iterative procedures that use  $n_{\text{QMC}}$  or  $n_{\text{LDA}+\text{noise}}$  as reference densities. Top panel: In yellow, LPD of the LDA density from the KS calculation decorated with a point-wise Gaussian noise  $n_{\text{LDA}+\text{noise}}$  with respect to the clean LDA density  $n_{\text{LDA}}$ . The noise lies within the stochastic error bar of the QMC calculation (grey area). In purple, LPD of  $n^{\text{LDA}+\text{noise},39}$  at the iteration  $i = 39$ , where the MaAPE on the density with respect to  $n_{\text{LDA}+\text{noise}}$  has its first minimum. Here, the MaAPE is 0.56%, and the MeAPE is 0.15%. In red, LPD of  $n^{\text{LDA}+\text{noise},150}$  at the iteration  $i = 150$ , where the MaAPE is 0.54%. In blue, LPD of  $n^{\text{QMC},500}$  at the iteration  $i = 500$ , where the MaAPE is 0.21% and the MeAPE is 0.02%, with respect to  $n_{\text{QMC}}$ . Bottom panel: the LDA  $v_{\text{xc}}^{\text{LDA}}$  potential (orange) is compared to the xc potentials obtained by inversion of the QMC density  $v_{\text{xc}}^{\text{QMC},500}$  (blue), and by inversion of the noisy LDA density  $v_{\text{xc}}^{\text{LDA}+\text{noise},24}$  (brown dashed line),  $v_{\text{xc}}^{\text{LDA}+\text{noise},39}$  (purple), and  $v_{\text{xc}}^{\text{LDA}+\text{noise},150}$  (red) at  $i = 150$ .

The bottom panel of Fig. 11 shows xc potentials: the red curve is the xc potential resulting from inversion of the noisy LDA data at  $i = 150$ . It has spikes that are of the same order as those of  $v_{\text{xc}}^{\text{QMC},500}$  (blue curve) and that are in percentage orders of magnitude larger than the noise of the density, again as in the case of  $v_{\text{xc}}^{\text{QMC},500}$ . With such an error bar, one would not be able to distinguish the LDA and QMC potentials. By way of contrast, the xc potential resulting from the noisy LDA data but at only  $i = 39$  iterations, where the MaAPE on the density has its minimum, shows only very small spikes (purple curve). The result is stable in the range of iterations preceding that minimum: the bottom panel also shows the result for  $i = 24$  (brown dashed curve), with a virtually indistinguishable potential. Moreover, this potential is close to the clean LDA potential, given by the orange curve. In this range of iterations, we can consider the resulting potential to be reliable. The spikes that develop by iterating further, instead, may suffer from the fact that the noisy density and the KS LDA hamiltonian are not completely consistent, which means that a higher

precision cannot be reached.

The observations concerning the behavior of the noisy LDA are strictly analogous to our QMC-based results, as we will also discuss in the next section. This gives strong evidence for the fact that the inversion problem of the QMC data after a certain number of iterations is indeed due to the stochastic noise of the QMC. Moreover, it suggests that a sufficiently reliable xc potential is obtained by taking the result in the range where a stable and relatively smooth potential is obtained, and before the MeAPE on the density stops to decrease. In the present case, this confirms the choice  $i = 20$ , for which the QMC xc potential is given in Fig. 1. In other words, this xc potential is, to the best of our knowledge, today's most precise estimate for the true xc potential of bulk silicon.

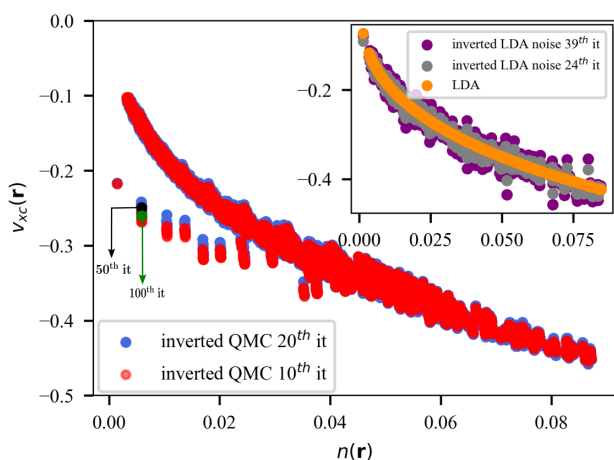


FIG. 12. Map of the xc potential versus the local density in silicon. Main panel: Results of the inversion of QMC for different  $i$ . Each  $v_{xc}$  is plotted against its own density. QMC inversion results are shown at 10 (smooth potential) and 20 (minimum of MaAPE of the density) iterations. For one density below 0.01, additional results at  $i = 50$  and  $i = 100$  are given, to illustrate the stability of the extra branch. In the inset, KS LDA is compared to the inversion of noisy LDA at 24 (smooth potential) and 39 (minimum of MaAPE of the density) iterations.

Finally, Fig. 12 confirms that the noise is at least partially responsible for the blurring of the QMC result observed in the map  $[v_{xc}(\mathbf{r}) \leftrightarrow n(\mathbf{r})]$  in Fig. 4: inversion of the noisy LDA data leads to a more scattered potential, as shown in the inset of Fig. 12. The comparison of the result for  $i = 24$  and  $i = 39$  also shows that it remains essentially a scattered version of the clean LDA result, with a blurring according to the number of iterations, whereas no additional features are caused by the noise. In the main panel, we compare different iterations of the QMC result. Also in this case higher  $i$  leads to stronger blurring, but the extra branch is confirmed to be a stable feature. Note that the point at lowest density on the extra branch undergoes large oscillations with the

number of iterations. All other points, instead, are well behaved, as it is illustrated on the figure for the point of second-lowest density.

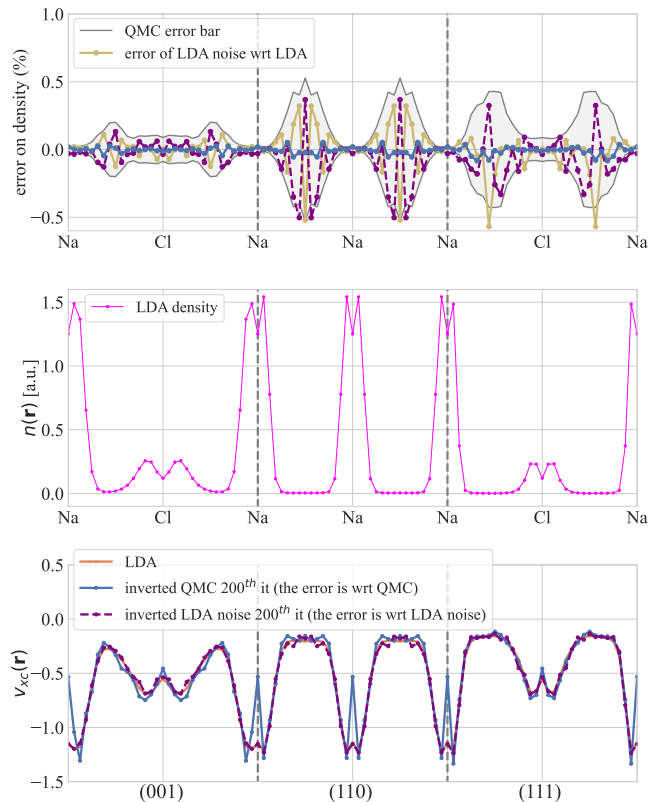


FIG. 13. Analog of Fig. 11, but for NaCl. Here spikes in noisy LDA and QMC are less developed than in the case of silicon. Note that the QMC error bar is smaller than in the case of silicon, in particular in regions of significant density.

For completeness, we investigate the issue of noise also in NaCl. Details are displayed in Fig. 13. As one can see, the QMC error bar is smaller in this case, especially in the important region of high density. Therefore, although qualitatively the noise has similar effects as in the case of silicon, quantitatively the effect on the results is negligible, as one can see in the local potential-density relation displayed in Fig. 14.

#### Appendix E: More about the inversion starting from the QMC density

Here, we will examine the consequences of noise on the inversion of the QMC results of silicon, where the effect is more pronounced than in NaCl.

The first interesting fact is that the error of the QMC inversion behaves similarly to that of the noisy LDA. This can be appreciated in Fig. 7 (upper panel), where the MaAPE of the density as a function of the number of iterations  $i$  is given by the blue curve. It shows an overall decrease, but with a pronounced minimum at  $i = 20$ .

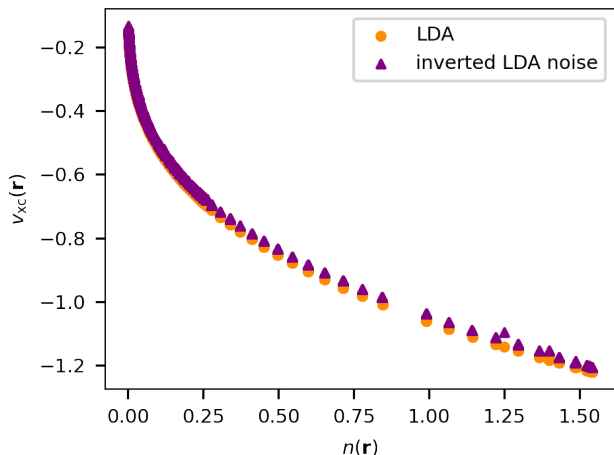


FIG. 14. Map of the xc potential versus the local density in NaCl. The result of the inversion of LDA with gaussian noise for  $i = 200$  is compared to the clean LDA xc potential resulting from the KS calculation.

At this point, it has decreased to 0.38%. The minimum is followed by a modest increase, after which the error decreases again monotonously. Instead, as the inset in Fig. 7 shows, in correspondence to the minimum the MeAPE reaches a plateau of about 0.02% and a better precision cannot be reached. For this reason, the inversion error remains larger than what we obtained in the case of the inverted clean LDA, by about a factor 500 for the MaAPE and 1000 for the MeAPE.

Of course, one could think to continue the iterations, since Fig. 7 shows that the MaAPE could be decreased further. However, the fact that a plateau is reached in the MeAPE anticipates that one might encounter problems when doing so. The blue curve in the upper panel of Fig. 11 shows the density from the QMC inversion at  $i = 500$ , where the MaAPE has decreased from 0.38% to 0.21%. Indeed, the error is now further away from the QMC error bar in the most critical points along the path, with respect to the  $i = 20$  result shown in Fig. 8. Also the MeAPE has decreased from 0.04% at  $i = 20$  to 0.02% at  $i = 500$ . However, the xc potential obtained from the QMC inversion, shown in the bottom panel of Fig. 11 (blue curve), is no longer smooth. It develops spikes that become even more pronounced when one iterates further, while still decreasing the MaAPE on the density, but with a quite constant MeAPE, which points to a mere redistribution of errors (see Fig. 15 for illustration).

Visibly, the algorithm does not succeed in improving the result any further and introduces unexpected features when trying to do so. Difficulties with the inversion procedure have also been reported for finite systems, and they have been attributed to the finite basis set which may introduce an inconsistency between density and external potential [49–53]. In the present work, as we have verified, the results are sufficiently well converged to ex-

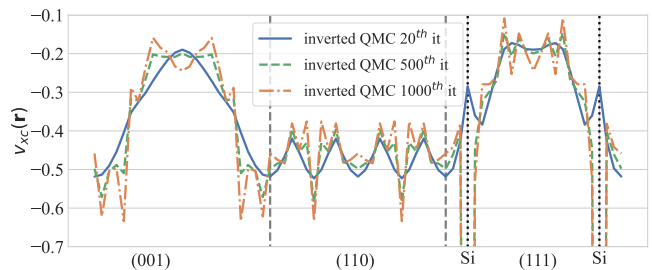


FIG. 15. Exchange-correlation potentials obtained by inversion of the QMC density of bulk silicon, at three different iteration steps. All three potentials yield a very similar, accurate density. The MeAPE of the density is 0.04% at  $i = 20$ , 0.02% at  $i = 500$ , and the same value at  $i = 1000$ . For the MaAPE, we get 0.38% at  $i = 20$ , 0.21% at  $i = 500$ , and 0.18% at  $i = 1000$ .

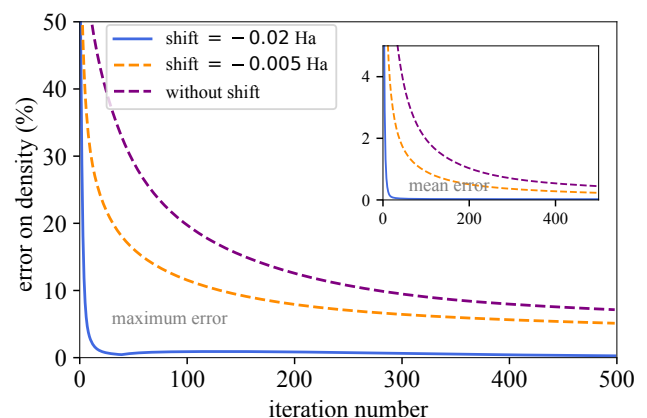


FIG. 16. Silicon: error of the inverted QMC density as a function of the iteration number for three different choices of the shift of the starting point. The starting point is  $0.1v_{xc}^{LDA} + \text{shift}$ .

clude basis set problems. Instead, the strict analogy to the behavior observed for the noisy LDA is strong evidence for the fact that here the QMC stochastic noise is the limiting factor. We stress again that noise leads to blurring, not to the appearance of spurious features, and its effect can therefore be detected and tested.

Finally, it is important to note that the results of the inversion do not depend on the starting point. This is true both for the shape and for the average value of the starting potential. The latter is important for the convergence behavior: Fig. 16 shows that a value of the initial shift which guarantees that the potential does not change sign improves the convergence for silicon significantly. Fig. 17 shows results obtained for a starting point corresponding to different fractions of  $v_{xc}^{LDA}$ : the scatter plot of silicon shows stability both of the main and of the extra branch, except for the point of very low density.

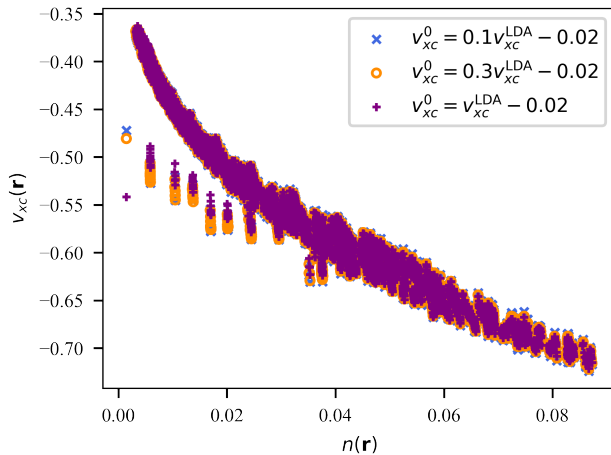


FIG. 17. Map of the xc potential of silicon with respect to the local density for three different choices of the starting point. The three xc potentials have the same MeAPE of 0.02 %. This precision is reached at  $i = 11, 20$  or  $38$  when the starting point is a shifted  $v_{xc}^{LDA}$ ,  $0.3v_{xc}^{LDA}$  or  $0.1v_{xc}^{LDA}$ , respectively.

- 
- [1] P. Hohenberg and W. Kohn, Inhomogeneous electron gas, *Phys. Rev.* **136**, B864 (1964).
- [2] W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, *Phys. Rev.* **140**, A1133 (1965).
- [3] W. Kohn, Nobel lecture: Electronic structure of matter—wave functions and density functionals, *Rev. Mod. Phys.* **71**, 1253 (1999).
- [4] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, 2004).
- [5] D. M. Ceperley and B. J. Alder, Ground state of the electron gas by a stochastic method, *Phys. Rev. Lett.* **45**, 566 (1980).
- [6] M. G. Medvedev, I. S. Bushmarinov, J. Sun, J. P. Perdew, and K. A. Lyssenko, Density functional theory is straying from the path toward the exact functional, *Science* **355**, 49 (2017).
- [7] A. D. Becke, Density-functional theory vs density-functional fits, *The Journal of Chemical Physics* **156**, 214101 (2022).
- [8] J. P. Perdew and W. Yue, Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation, *Phys. Rev. B* **33**, 8800 (1986).
- [9] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized gradient approximation made simple, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [10] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, Restoring the density-gradient expansion for exchange in solids and surfaces, *Phys. Rev. Lett.* **100**, 136406 (2008).
- [11] J. Sun, A. Ruzsinszky, and J. P. Perdew, Strongly constrained and appropriately normed semilocal density functional, *Phys. Rev. Lett.* **115**, 036402 (2015).
- [12] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, Density-functional theory for fractional particle number: Derivative discontinuities of the energy, *Phys. Rev. Lett.* **49**, 1691 (1982).
- [13] L. J. Sham and M. Schlüter, Density-functional theory of the energy gap, *Phys. Rev. Lett.* **51**, 1888 (1983).
- [14] J. P. Perdew and M. Levy, Physical content of the exact kohn-sham orbital energies: Band gaps and derivative discontinuities, *Phys. Rev. Lett.* **51**, 1884 (1983).
- [15] J. P. Perdew and A. Zunger, Self-interaction correction to density-functional approximations for many-electron systems, *Phys. Rev. B* **23**, 5048 (1981).
- [16] M. van Schilfhaarde, T. Kotani, and S. Faleev, Quasi-particle self-consistent  $gw$  theory, *Phys. Rev. Lett.* **96**, 226402 (2006).
- [17] L. J. Sham and W. Kohn, One-particle properties of an inhomogeneous interacting electron gas, *Phys. Rev.* **145**, 561 (1966).
- [18] R. W. Godby, M. Schlüter, and L. J. Sham, Accurate exchange-correlation potential for silicon and its discontinuity on addition of an electron, *Phys. Rev. Lett.* **56**, 2415 (1986).
- [19] R. W. Godby, M. Schlüter, and L. J. Sham, Trends in self-energy operators and their corresponding exchange-correlation potentials, *Phys. Rev. B* **36**, 6497 (1987).
- [20] R. W. Godby, M. Schlüter, and L. J. Sham, Self-energy operators and exchange-correlation potentials in semiconductors, *Phys. Rev. B* **37**, 10159 (1988).
- [21] Y. M. Niquet and X. Gonze, Band-gap energy in the random-phase approximation to density-functional the-



- ory, *Phys. Rev. B* **70**, 245115 (2004).
- [22] M. Grüning, A. Marini, and A. Rubio, Effect of spatial nonlocality on the density functional band gap, *Phys. Rev. B* **74**, 161103 (2006).
- [23] M. Grüning, A. Marini, and A. Rubio, Density functionals from many-body perturbation theory: The band gap for semiconductors and insulators, *The Journal of Chemical Physics* **124**, 154108 (2006).
- [24] T. Kotani, An optimized-effective-potential method for solids with exact exchange and random-phase approximation correlation, *J. Phys.: Condens. Matter* **10**, 9241 (1998).
- [25] J. Klimeck and G. Kresse, Kohn-sham band gaps and potentials of solids from the optimised effective potential method within the random phase approximation, *The Journal of Chemical Physics* **140**, 054516 (2014).
- [26] S. Riemelmoser, M. Kaltak, and G. Kresse, Optimized effective potentials from the random-phase approximation: Accuracy of the quasiparticle approximation, *The Journal of Chemical Physics* **154**, 154103 (2021).
- [27] R. van Leeuwen and E. J. Baerends, Exchange-correlation potential with correct asymptotic behavior, *Phys. Rev. A* **49**, 2421 (1994).
- [28] Q. Zhao, R. C. Morrison, and R. G. Parr, From electron densities to kohn-sham kinetic energies, orbital energies, exchange-correlation potentials, and exchange-correlation energies, *Phys. Rev. A* **50**, 2138 (1994).
- [29] Q. Wu and W. Yang, A direct optimization method for calculating density functionals and exchange-correlation potentials from electron densities, *The Journal of Chemical Physics* **118**, 2498 (2003).
- [30] K. Peirs, D. Van Neck, and M. Waroquier, Algorithm to derive exact exchange-correlation potentials from correlated densities in atoms, *Phys. Rev. A* **67**, 012505 (2003).
- [31] C. O. Almbladh and A. C. Pedroza, Density-functional exchange-correlation potentials and orbital eigenvalues for light atoms, *Phys. Rev. A* **29**, 2322 (1984).
- [32] F. Aryasetiawan and M. J. Stott, Effective potentials in density-functional theory, *Phys. Rev. B* **38**, 2974 (1988).
- [33] A. Görling, Kohn-sham potentials and wave functions from electron densities, *Phys. Rev. A* **46**, 3753 (1992).
- [34] W. Knorrand R. W. Godby, Investigating exact density-functional theory of a model semiconductor, *Phys. Rev. Lett.* **68**, 639 (1992).
- [35] W. Knorrand R. W. Godby, Quantum monte carlo study of density-functional theory for a semiconducting wire, *Phys. Rev. B* **50**, 1779 (1994).
- [36] C. J. Umrigar and X. Gonze, Accurate exchange-correlation potentials and total-energy components for the helium isoelectronic series, *Phys. Rev. A* **50**, 3827 (1994).
- [37] O. V. Gritsenko, R. van Leeuwen, and E. J. Baerends, Molecular kohn-sham exchange-correlation potential from the correlated ab initio electron density, *Phys. Rev. A* **52**, 1870 (1995).
- [38] D. J. Tozer, V. E. Ingamells, and N. C. Handy, Exchange-correlation potentials, *The Journal of Chemical Physics* **105**, 9200 (1996).
- [39] N. Helbig, I. V. Tokatly, and A. Rubio, Exact kohnsham potential of strongly correlated finite systems, *The Journal of Chemical Physics* **131**, 224105 (2009).
- [40] E. M. Stoudenmire, L. O. Wagner, S. R. White, and K. Burke, One-dimensional continuum electronic structure with the density-matrix renormalization group and its implications for density-functional theory, *Phys. Rev. Lett.* **109**, 056402 (2012).
- [41] D. Varsano, M. Barborini, and L. Guidoni, Kohn-sham orbitals and potentials from quantum monte carlo molecular densities, *The Journal of Chemical Physics* **140**, 054102 (2014).
- [42] T. W. Hollins, S. J. Clark, K. Refson, and N. I. Gidopoulos, A local fock-exchange potential in kohn-sham equations, *Journal of Physics: Condensed Matter* **29**, 04LT01 (2016).
- [43] M. J. P. Hodgson, J. D. Ramsden, and R. W. Godby, Origin of static and dynamic steps in exact kohn-sham potentials, *Phys. Rev. B* **93**, 155146 (2016).
- [44] M. J. P. Hodgson, E. Kraisler, A. Schild, and E. K. U. Gross, How interatomic steps in the exact kohnsham potential relate to derivative discontinuities of the energy, *The Journal of Physical Chemistry Letters* **8**, 5974 (2017).
- [45] J. Wetherell, M. J. P. Hodgson, L. Talirz, and R. W. Godby, Advantageous nearsightedness of many-body perturbation theory contrasted with kohn-sham density functional theory, *Phys. Rev. B* **99**, 045129 (2019).
- [46] S. Nam, S. Song, E. Sim, and K. Burke, Measuring density-driven errors using kohnsham inversion, *Journal of Chemical Theory and Computation* **16**, 5014 (2020).
- [47] A. Savin, C. Umrigar, and X. Gonze, Relationship of kohnsham eigenvalues to excitation energies, *Chemical Physics Letters* **288**, 391 (1998).
- [48] J. Li, N. D. Drummond, P. Schuck, and V. Olevano, Comparing many-body approaches against the helium atom exact solution, *SciPost Phys.* **6**, 40 (2019).
- [49] P. R. T. Schipper, O. V. Gritsenko, and E. J. Baerends, Kohn-sham potentials corresponding to Slater and gaussian basis set densities, *Theoretical Chemistry Accounts* **98**, 16 (1997).
- [50] M. E. Mura, P. J. Knowles, and C. A. Reynolds, Accurate numerical determination of kohn-sham potentials from electronic densities: I. two-electron systems, *The Journal of Chemical Physics* **106**, 9659 (1997).
- [51] T. Heaton-Burgess, F. A. Bulat, and W. Yang, Optimized effective potentials in finite basis sets, *Phys. Rev. Lett.* **98**, 256401 (2007).
- [52] C. R. Jacob, Unambiguous optimization of effective potentials in finite basis sets, *The Journal of Chemical Physics* **135**, 244102 (2011).
- [53] A. P. Gaiduk, I. G. Ryabinkin, and V. N. Staroverov, Removal of basis-set artifacts in kohnsham potentials recovered from electron densities, *Journal of Chemical Theory and Computation* **9**, 3959 (2013).
- [54] D. S. Jensen and A. Wasserman, Numerical methods for the inverse problem of density functional theory, *International Journal of Quantum Chemistry* **118**, e25425 (2018).
- [55] Y. Shi and A. Wasserman, Inverse kohnsham density functional theory: Progress and challenges, *The Journal of Physical Chemistry Letters* **12**, 5308 (2021).
- [56] I. G. Ryabinkin and V. N. Staroverov, Determination of kohnsham effective potentials from electron densities using the differential virial theorem, *The Journal of Chemical Physics* **137**, 164113 (2012).
- [57] Q. Ou and E. A. Carter, Potential functional embedding theory with an improved kohnsham inversion algorithm, *Journal of Chemical Theory and Computation* **14**, 5680 (2018), PMID: 30216062.

- [58] B. Kanungo, P. M. Zimmerman, and V. Gavini, Exact exchange-correlation potentials from ground-state electron densities, *Nature Communications* **10**, 4497 (2019).
- [59] A. Kumar, R. Singh, and M. K. Harbola, Universal nature of different methods of obtaining the exact kohn-sham exchange-correlation potential for a given density, *Journal of Physics B: Atomic, Molecular and Optical Physics* **52**, 075007 (2019).
- [60] A. Kumar and M. K. Harbola, A general penalty method for density-to-potential inversion, *International Journal of Quantum Chemistry* **120**, e26400 (2020).
- [61] A. Kumar, R. Singh, and M. K. Harbola, Accurate effective potential for density amplitude and the corresponding kohn-sham exchange-correlation potential calculated from approximate wavefunctions, *Journal of Physics B: Atomic, Molecular and Optical Physics* **53**, 165002 (2020).
- [62] T. J. Callow, N. N. Lathiotakis, and N. I. Gidopoulos, Density-inversion method for the kohnsham potential: Role of the screening density, *The Journal of Chemical Physics* **152**, 164114 (2020).
- [63] S. Nam, R. J. McCarty, H. Park, and E. Sim, Ks-pies: Kohnsham inversion toolkit, *The Journal of Chemical Physics* **154**, 124122 (2021).
- [64] J. Erhard, E. Trushin, and A. Grling, Numerically stable inversion approach to construct kohn-sham potentials for given electron densities within a gaussian basis set framework, *The Journal of Chemical Physics* **0**, null (0).
- [65] Y. Shi, V. H. Chvez, and A. Wasserman, n2v: A density-to-potential inversion suite. a sandbox for creating, testing, and benchmarking density functional theory inversion methods, *WIREs Computational Molecular Science* **n/a**, e1617.
- [66] D. Jayatilaka, Wave function for beryllium from x-ray diffraction data, *Phys. Rev. Lett.* **80**, 798 (1998).
- [67] Given the pseudopotential. In other words, all discussions are valid because we have used the same hamiltonian, with a fixed LDA pseudopotential, for the valence electron problem throughout, including for the QMC calculations. A full all-electron  $v_{xc}$  would of course look different.
- [68] S. Zhang and H. Krakauer, Quantum monte carlo method using phase-free random walks with slater determinants, *Phys. Rev. Lett.* **90**, 136401 (2003).
- [69] M. Motta and S. Zhang, Ab initio computations of molecular systems by the auxiliary-field quantum monte carlo method, *WIREs Computational Molecular Science* **8**, e1364 (2018).
- [70] S. Chen, M. Motta, F. Ma, and S. Zhang, Ab initio electronic density in solids by many-body plane-wave auxiliary-field quantum monte carlo calculations, *Phys. Rev. B* **103**, 075138 (2021).
- [71] V. U. Nazarov, Breakdown of the ionization potential theorem of density functional theory in mesoscopic systems, *The Journal of Chemical Physics* **155**, 194105 (2021).
- [72] Note that we have defined the error with opposite sign with respect to Ref. [70].
- [73] F. Tran and P. Blaha, Accurate band gaps of semiconductors and insulators with a semilocal exchange-correlation potential, *Phys. Rev. Lett.* **102**, 226401 (2009).
- [74] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation, *Phys. Rev. B* **46**, 6671 (1992).
- [75] B. Kalita, L. Li, R. J. McCarty, and K. Burke, Learning to approximate density functionals, *Accounts of Chemical Research* **54**, 818 (2021).
- [76] D. P. Chong, O. V. Gritsenko, and E. J. Baerends, Interpretation of the kohnsham orbital energies as approximate vertical ionization potentials, *The Journal of Chemical Physics* **116**, 1760 (2002).
- [77] C. Filippi, C. J. Umrigar, and X. Gonze, Excitation energies from density functional perturbation theory, *The Journal of Chemical Physics* **107**, 9994 (1997).
- [78] R. Martin, L. Reining, and D. Ceperley, *Interacting Electrons: Theory and Computational Approaches* (Cambridge University Press, 2016).
- [79] M. Lannoo, M. Schlüter, and L. J. Sham, Calculation of the kohn-sham potential and its discontinuity for a model-semiconductor, *Phys. Rev. B* **32**, 3890 (1985).
- [80] A. G. Eguluz, M. Heinrichsmeier, A. Fleszar, and W. Hanke, First-principles evaluation of the surface barrier for a kohn-sham electron at a metal surface, *Phys. Rev. Lett.* **68**, 1359 (1992).
- [81] M. Hellgren and U. von Barth, Correlation potential in density functional theory at the gwa level: Spherical atoms, *Phys. Rev. B* **76**, 075107 (2007).
- [82] M. Hellgren and U. von Barth, Correlation energy functional and potential from time-dependent exact-exchange theory, *The Journal of Chemical Physics* **132**, 044101 (2010).
- [83] P. Bleiziffer, A. Heilmann, and A. Grling, Efficient self-consistent treatment of electron correlation within the random phase approximation, *The Journal of Chemical Physics* **139**, 084113 (2013).
- [84] L. Hedin, New method for calculating the one-particle green's function with application to the electron-gas problem, *Phys. Rev.* **139**, A796 (1965).
- [85] Landolt-Börnstein, *Numerical Data and Functional Relationships in Science and Technology*, New Series Group III, Subvolume B: II-VII and I-VII Compounds, Vol. 17 Pt A (Springer-Verlag, Berlin-Heidelberg, New York, 1982).
- [86] R. T. Poole, J. G. Jenkin, J. Liesegang, and R. C. G. Leckey, Electronic band structure of the alkali halides. i. experimental parameters, *Phys. Rev. B* **11**, 5179 (1975).
- [87] Note that the difference between the exact KS band structure and the true quasiparticle one is not limited to a rigid shift of conduction states, since the KS potential is only supposed to correctly describe the highest occupied state of the N-electron and (when the derivative discontinuity is taken into account) N+1-electron systems.
- [88] J. E. Ortega and F. J. Himpsel, Inverse-photoemission study of ge(100), si(100), and gaas(100): Bulk bands and surface states, *Phys. Rev. B* **47**, 2130 (1993).
- [89] A. Savin, F. Colonna, and R. Pollet, Adiabatic connection approach to density functional theory of electronic systems, *International Journal of Quantum Chemistry* **93**, 166 (2003).
- [90] M.-C. Kim, E. Sim, and K. Burke, Understanding and reducing errors in density functional calculations, *Phys. Rev. Lett.* **111**, 073003 (2013).
- [91] A. Wasserman, J. Nafziger, K. Jiang, M.-C. Kim, E. Sim, and K. Burke, The importance of being inconsistent, *Annual Review of Physical Chemistry* **68**, 555 (2017).
- [92] D. R. Hamann, Optimized norm-conserving vanderbilt

- pseudopotentials, *Phys. Rev. B* **88**, 085117 (2013).
- [93] X. Gonze, B. Amadon, G. Antonius, F. Arnardi, L. Baguet, J.-M. Beuken, J. Bieder, F. Bottin, J. Bouchet, E. Bousquet, N. Brouwer, F. Bruneval, G. Brunin, T. Cavignac, J.-B. Charraud, W. Chen, M. Ct, S. Cottelier, J. Denier, G. Geneste, P. Ghosez, M. Giantomassi, Y. Gillet, O. Gingras, D. R. Hamann, G. Hautier, X. He, N. Helbig, N. Holzwarth, Y. Jia, F. Jollet, W. Lafargue-Dit-Hauret, K. Lejaeghere, M. A. Marques, A. Martin, C. Martins, H. P. Miranda, F. Nacarrato, K. Persson, G. Petretto, V. Planes, Y. Pouillon, S. Prokhorenko, F. Ricci, G.-M. Rignanese, A. H. Romero, M. M. Schmitt, M. Torrent, M. J. van Setten, B. Van Troeye, M. J. Verstraete, G. Zrah, and J. W. Zwanziger, *The abinitproject: Impact, environment and recent developments*, *Computer Physics Communications* **248**, 107042 (2020).
- [94] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. D. Corso, S. de Gironcoli, P. Delugas, R. A. DiStasio, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Kçkbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. O. de-la Roza, L. Paulatto, S. Ponc e, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, and S. Baroni, *Advanced capabilities for materials modelling with quantum ESPRESSO*, *Journal of Physics: Condensed Matter* **29**, 465901 (2017).
- [95] A. D. Becke and E. R. Johnson, *A simple effective potential for exchange*, *The Journal of Chemical Physics* **124**, 221101 (2006).
- [96] D. Waroquiers, A. Lherbier, A. Miglio, M. Stankovski, S. Ponc e, M. J. T. Oliveira, M. Giantomassi, G.-M. Rignanese, and X. Gonze, *Band widths and gaps from the tran-blaha functional: Comparison with many-body perturbation theory*, *Phys. Rev. B* **87**, 075121 (2013).
- [97] M. A. Marques, M. J. Oliveira, and T. Burnus, *Libxc: A library of exchange and correlation functionals for density functional theory*, *Computer Physics Communications* **183**, 2272 (2012).
- [98] One might think that inverting the LDA is trivial, since the LDA xc potential is a monotonic function of the density, and the density is therefore a function (and not a non-local functional) of the LDA xc potential. This might suggest that our algorithm should obviously work well for the LDA. However, one has to consider that at the start of the iteration procedure the xc potential is not the LDA one, so the density-potential relation is not local. Therefore, the algorithm could in principle be non-convergent or lead to a wrong solution, even for the LDA.