Simplicity of the Hamiltonian and complexity of the behaviour is the main characteristic of the Hubbard model. Although similar electronic models had existed for decades \[1\], the breakthrough of the 1960s \[2–5\] was to observe that only two parameters are sufficient to describe a rich spectrum of phenomena \[6–8\] that occurs in correlated electron systems. To go from a qualitative description of these phenomena to quantitative predictions for real materials, it is necessary to determine the parameters of the Hubbard model that correspond to any given material. The essential point is that comparatively simple methods can be used for the many weakly correlated bands, and to identify from this the appropriate correlated subspace with the corresponding Hubbard parameters. The state-of-the-art methods to do this are GW $+$ DMFT \[9\] and cRPA \[10\]. The former determines both the kinetic and the interaction terms of the Hubbard model explicitly in a Green’s function approach. In the latter approach, the kinetic term is extracted from the DFT dispersion, whereas the interaction is determined using the cRPA approximation. The cRPA part could in the future be replaced by more sophisticated methods such as cFRG \[11, 12\].

In general, these methods provide an interaction which is not local. The bare Coulomb interaction is long-ranged. Although screening by the weakly interacting bands reduces the effective range of the interaction, some spatial character generally remains, especially in two-dimensional compounds. The Hubbard model, however, contains only a local interaction and, importantly, the often used dynamical mean-field theory \[13–15\] for the solution of the Hubbard model is also restricted to local interactions. One way out is to extend the Hubbard model and to extend DMFT. This is under active investigation \[16, 17\], but it is complicated both conceptually and numerically. An alternative in the spirit of simple models is to try to capture the intersite interactions in the effective Hubbard parameters. This approach was first studied in graphene, where it has been shown that screening due to the intersite interactions is
crucial to keep graphene in its metallic state [18, 19]. It has been used further to study the effect of intersite interactions on the order of the metal-insulator transition [20]. In these works, only the effective interaction was renormalized.

In addition to screening the local interaction, another important physical effect of intersite interactions is the renormalization of the bandwidth [21]. Recent work [22] in the context of EDMFT [23–28] has emphasized the importance of this effect. There, a perturbative treatment (the Fock diagram) of this effect was proposed. In strongly correlated systems, it is questionable if such a perturbative treatment that uses Wick’s theorem is valid and sufficient. This question is particularly relevant now that non-equilibrium applications of EDMFT are appearing [29, 30].

In this work, we use a variational approach to study the bandwidth renormalization due to intersite interactions. A related physical effect, the bandwidth renormalization due to phonons, has previously been studied [31]. The variational approach is not fundamentally limited to weakly correlated systems, so that we can use it to judge when a perturbative (Fock) treatment is reasonable. We compare bandwidth renormalization, interaction renormalization and a combined variational approach and explain how the physical regime determines which scheme is most effective.

To study the variational principle, we consider situations where the Hubbard model is exactly solvable: small systems where exact diagonalization is feasible and half-filled bipartite systems where QMC does not suffer from the sign problem. In the former, even the extended Hubbard model can be solved exactly, so that the applicability of the variational approach can be established. The latter scenario is directly relevant for graphene, which can be described as a half-filled single-band Hubbard model on the honeycomb lattice, and we use cRPA values for the intersite interaction to determine the bandwidth renormalization in graphene.

1. Variational principle

In this section, we provide a short overview of the variational principle, more details can be found in appendix A and in [19, 32]. The idea of the variational principle is to describe an effective Hubbard model, $H^*$, with parameters $t^*$ and $U^*$, in formulas:

\[ H = -t \sum_{\langle ij \rangle, \sigma} c_{i \sigma}^+ c_{j \sigma} + U \sum_i n_{i \uparrow} n_{i \downarrow} + \frac{V}{2} \sum_i n_i n_j \]

\[ H^* = -t^* \sum_{\langle ij \rangle, \sigma} c_{i \sigma}^+ c_{j \sigma} + U^* \sum_i n_{i \uparrow} n_{i \downarrow}. \]

(1)

Here $t$ is the hopping parameter, $U$ is the on-site Coulomb interaction, $V$ is the nearest-neighbor Coulomb interaction, $c_{i \sigma}^+$ and $c_{i \sigma}$ are creation and annihilation operators for an electron on site $i$ with spin $\sigma$ and $n_{i \sigma} = c_{i \sigma}^+ c_{i \sigma}$ is the corresponding number operator. The total density on site $i$ is equal to $n_i = n_{i \uparrow} + n_{i \downarrow}$. In the sum, $(i, j)$ denotes pairs of nearest-neighbor sites $i, j$ counted twice, i.e. both $ij$ and $ji$.

We study this system either in the canonical ensemble (section 2.1) or in the grand-canonical ensemble at particle-hole symmetry (section 2.2). The latter corresponds to $\mu = U/2$ and an average density of one electron per site and $n_i$ is replaced by $n_i - 1$ in the intersite interaction term. For simplicity, we restrict ourselves to translationally invariant systems.

The true free energy of the extended Hubbard model is lower than the variational free energy [33–35] of the effective Hubbard model for any choice of $t^*$ and $U^*$,

\[ F \leq F_v = F^* + \langle H - H^* \rangle_s. \]

(2)

Here $\langle \cdot \rangle_s$ denotes the expectation value of $\cdot$ with respect to the Hamiltonian $H^*$. This inequality means that $t^*$ and $U^*$ can be used as variational parameters to minimize the variational free energy and to get an estimate of the true free energy $F$ of the original extended Hubbard model. Furthermore, one might expect at least some observables of this effective Hubbard model to be close to the values of the original extended Hubbard model.

In previous works [19, 20, 32], only a renormalization of the interaction strength $U^*$ was considered and the hopping parameter was kept at its original value $t^* = t$. In this work, on the other hand, we focus exactly on the bandwidth renormalization. We will consider both bandwidth renormalization only, keeping $U^* = U$, which we will call the $t^*$ scheme, and the simultaneous variation of $t^*$ and $U^*$, which we call the $(U^*, t^*)$ scheme. In the former case, the bandwidth renormalization is determined by the minimum of the variational free energy and can be written as

\[ t^* = t + \alpha(t^*)V, \]

(3)

\[ \alpha(t^*) = \frac{1}{4} \frac{\partial^2 \langle n_0 n_1 \rangle_s}{\partial t \frac{\partial}{\partial t^*}} G_{01}, \]

(4)

with $G_{01} = \langle c_{1, \sigma}^+ c_{0, \sigma} \rangle_s$ the equal-time nearest-neighbor Green’s function, which is proportional to the kinetic energy per site. We note that in (3) $\alpha$ depends on $t^*$ since the expectation values are those of the effective Hubbard model, $\alpha$ also implicitly depends on the temperature via the expectation values.

An alternative to the variational principle is to use Hartree–Fock to determine the bandwidth renormalization. As shown in appendix B, this corresponds to

\[ \alpha^{HF} = -G_{01}. \]

(5)

This result is derived from (4) by assuming that the effective Hubbard model $H^*$ is uncorrelated so that Wick’s theorem can be applied. This assumption is not true for $U^* \neq 0$ and we will study how applicable the Hartree–Fock approximation is in section 3.3.

The variational principles give upper bounds for the free energy. Even when the exact free energy of the extended Hubbard model is not known, as is the case for the two-dimensional systems studied below, the methods can be compared since a lower upper bound is better. Variation over both $t^*$ and $U^*$ is obviously at least as good as varying over one of these parameters.
For small values of $V$, the situation simplifies and some exact relations can be obtained by expanding the free energy as a power series in $V$. The single parameter schemes give the exact value of their respective conjugate observables $[36]$ to linear order in $V$, as shown in appendix D. The variational principle also gives the exact free energy to linear order in $V$. Looking at $F_{\text{variational}} - F_{\text{true}}$ as function of $V$, there cannot be a linear term since this difference has to be positive for both signs of $V$. Since the two-parameter scheme gives a free energy that is lower than the single-parameter schemes, it also gives the exact linear order in $V$ of the free energy. One might expect that the additional variational parameter also allows the exact second-order coefficient to be captured, but we will see that this is generally not the case. The quadratic order in $V$ of the variational free energy is different in the two schemes, for our purposes it is sufficient to calculate how much free energy is gained by the variational principle, $\Delta F = F_v(U^*, t^*) - F_v(U, t)$, without needing to calculate the free energy of an extended Hubbard model. The result of this derivation, shown in appendix E, is

$$F_v(U, t^*) - F_v(U, t) = \frac{zV^2}{16} \left( \frac{\partial}{\partial U} \left( \langle n_{0\uparrow} n_{0\downarrow} \rangle \right) \right)^2,$$

$$F_v(U^*, t) - F_v(U, t) = \frac{zV^2}{8} \left( \frac{\partial}{\partial U} \left( \langle n_{0\uparrow} n_{0\downarrow} \rangle \right) \right)^2.$$  

Here, $z$ is the coordination number and $D = \langle n_{0\uparrow} n_{0\downarrow} \rangle$ stands for the double occupancy, the probability of having two electrons at the same site. The sign of the difference $F_v(U, t^*) - F_v(U^*, t)$ tells us which variational parameter works best.

2. Solving the Hubbard model

The variational treatment of $V$ requires an exact solution of the Hubbard model, the reference system. In this section we explain the methods we use for this purpose.

2.1. Exact diagonalization

If the number of lattice sites is small, the system can be solved exactly by diagonalizing the Hamiltonian. This can be done for the effective Hubbard model, but also for the original extended Hubbard model. In this way, it is possible to compare the free energy (and other observables) of the effective system to see how well the variational principle works.

Exact diagonalization is limited to finite and rather small systems. We consider a one-dimensional chain of six atoms (a ‘benzene ring’). Here, this system is used simply as a toy model, the relation to actual benzene is discussed in section 3.6. The total number of electrons is either 6 (half-filling) or 5 (17% doping), and we use the canonical ensemble.

Since the ED is both exact and computationally light, we can calculate even the numerical derivatives with high accuracy.

For the nearest-neighbor interaction $V$, we consider both positive and negative values. Positive values correspond to repulsive interaction and are most physically intuitive for the Coulomb interaction between electrons. We find the study of negative $V$ useful for the determination of linear coefficients. Attractive $V$ also serves as an interesting test case $[37]$ since the physics of phase separation starts to become relevant. We restrict ourselves to $|V| < U$. At $V = U/2$, a non-uniform phase with alternating high and low density sites becomes favorable in terms of the potential energy. In an infinitely large system, this could lead to a phase transition to a charge-ordered phase $[38, 39]$. There are no phase transitions in finite systems, so the exact free energy is a smooth function. However, quite sharp features foreshadow phase transitions in the infinite volume limit. Approximate methods can (and do) result in discontinuities in finite systems.

2.2. Determinant quantum Monte Carlo

The exact diagonalization is restricted to rather small systems. To study larger systems, we consider bipartite (honeycomb and square) lattices at half-filling using determinant quantum Monte Carlo methods $[40]$. In this scenario the Hubbard model is free from the fermionic sign problem. We use the open source quest code (‘QUantum Electron Simulation Toolbox’ QUEST 1.3.0 Tomas et al (http://quest.ucdavis.edu/)) to perform the simulations.

Regarding lattice sizes, for the honeycomb lattice we used lattices up to 16 × 16 unit cells (512 atoms) and comparing lattices with linear dimension $L = 4, 8, 12, 16$ we found relatively good convergence for lattices larger than 4 × 4 unit cells (32 atoms). We used 8 × 8 unit cells for all honeycomb lattice results shown here. A more detailed analysis of finite-size effects can be found in the appendix of $[20]$.

The determination of the effective Hubbard parameters only needs local and nearest-neighbor observables, which converge quickly with lattice size, and in this work we mostly stay away from phase transitions where careful extrapolation is needed.

As opposed to the exact diagonalization results, the QMC is computationally quite expensive and all results suffer from numerical noise due to a finite amount of MC steps. The simulations are generally more difficult, and the noise is worse, at low temperatures. The variational principle requires us to take numerical derivatives of noisy data, we use the Savitzky–Golay $[41]$ filter to improve the stability of this numerical derivative.

For these systems, we only perform calculations for the effective Hubbard model, determining the effective Hubbard parameters corresponding to a particular value of $V$. We cannot solve the extended Hubbard model exactly and thus we can also not compare the observables with the exact results. Therefore, we will focus on comparing the effective Hubbard parameters. The renormalization of the interaction in the square and honeycomb lattice has been studied previously $[19, 20, 32]$, here we focus on the renormalization of the bandwidth and on the combined scheme.

Since the determination of the effective $t^*$ requires a derivative only with respect to $t^*$, the honeycomb lattice simulations have been performed for constant $U$ and $\beta$ and varying
3. Results

3.1. Free energy

We start our analysis with the free energy, since the variational principle is based on minimizing this quantity. We compare the variational schemes and come back to the Fock diagram in section 3.3. We only study the square lattice and the chain in this section.

3.1.1. Square lattice. As a starting point, we briefly consider the square lattice, based on the simulations of [42]. These differ from the other simulations in this work by the fact that finite-size extrapolation to the thermodynamic limit has been performed. This is of no concern to the current study since the variational principle is applicable both to finite and infinite systems. Large scale scans over both \( t^* \) and \( U^* \) were performed, so that the two variational schemes can be compared.

We use the expansion of the variational free energy for small \( V \), as given in (7) and appendix E, to find out which variational method works best, bandwidth renormalization or interaction renormalization. Figure 1 shows the difference in free energy between the \( t^* \) and \( U^* \) optimizations. To facilitate the comparison, following the discussion in appendix E we have divided the free energy difference by \( V^2 \). We find that the bandwidth renormalization performs better at larger \( U/t \) (red) and worse at small \( U/t \) (blue).

This observation might come as a surprise, but it can be understood in terms of energetics. When \( U/t \) is large, both the extended and the reference Hubbard model have minimized their potential energy. The remaining low-energy degrees of freedom differ in their kinetic energy and this is captured well by \( t^* \). When \( U/t \) is small, the system first minimizes its kinetic energy and the remaining low-energy degrees of freedom are governed by the potential energy, so the \( U^* \) variational principle works better [42]. This is somewhat reminiscent of the Slater [43] and Heisenberg [44] mechanisms of antiferromagnetism in the Hubbard model [45–50], with kinetic energy driven Heisenberg antiferromagnetism at large \( U \) and potential energy driven Slater antiferromagnetism at small \( U \).

3.1.2. Chain. We now turn to the chain model, where figure 2 again compares the free energy gain of the two variational schemes. This time, it turns out that the \( U^* \) scheme performs better across the parameter space, although the difference is negligible for large \( U \) and/or \( t \). Unlike for the square lattice, here we did not find any situation where the \( t^* \) scheme works better than the \( U^* \) scheme. Physically, an important difference is that the six-site chain has discrete spectra. In the Heisenberg limit of the square lattice (where the \( t^* \) scheme works well), there is a continuum of low-energy spin excitations down to zero energy. On the other hand, in the finite system the spin spectrum is gapped and consists of discrete peaks. We speculate that this is the physical origin for the difference between figures 1 and 2.

In the chain system, we can also solve the original extended Hubbard model to assess how well the variational schemes are working. Figure 3 shows the deviation between the variational free energy and the true free energy of the extended Hubbard model for a six-site chain with five electrons.

Our first observation is that the variational free energy is indeed always larger than the true free energy of the system, which shows that the variational principle does not yield the exact solution of the extended Hubbard model. Furthermore, the variational free energy obtained from varying either of \( t^* \) and \( U^* \) is always larger or equal than the one obtained by
varying both. For small $V$, the deviation from the true free energy is proportional to $V^2$ (all curves approach a finite number for $V \to 0$), showing that the variational principle does capture the linear term correctly.

At the lowest temperatures, we see that there is hardly any difference between varying $t^*$, $U^*$ or both. In fact, as shown in appendix $F$, at $T = 0$ these variational principles are completely equivalent since there is only a single physical parameter $U^*/t^*$. The numerical observations here show that this result also holds approximately at sufficiently low temperature and that $\beta = 10$ is indeed sufficiently low, in this chain model.

For panel (a), where $U^*/t^*$ is very small, the lines are almost completely flat. This shows that the parabolic approximation of the variational free energy works very well in this regime. For panel (b), deviations are visible albeit small.

Panel (c) shows a more elevated temperature, where the performance of the $t^*$ and $U^*$ schemes is different. For almost all values of $V$, the $U^*$ scheme gives a significantly lower free energy than varying $t^*$. There is little gain from varying both simultaneously. Although the observables will be studied in more detail later on, it is already useful to discuss the poor performance of the $t^*$ scheme here. It turns out that it also poorly predicts observables like the nearest-neighbor spin correlation (see, e.g. figure 8). Looking at the variational parameters, the $U^*$ scheme finds $U^* \approx 0.9 \approx U - V$ for $V = 0.1$ (see figure 5), so that $U^*/t^* = 0.9$, while $t^*/t^* \approx 1.05$ gives $U^*/t^* = 0.95$. The $t^*$ scheme overestimates the effective interaction and this leads to the poor overall results.

Moving on to panel (d), $U/t = 10$, the physics starts to be dominated by the potential energy. In particular, this means that there is a sharp crossover at $V = U/2 = 5$ when the alternating high and low density sites become favorable. A zoomed in version of the free energy at small $V$ is given in figure 4.

The figures shows that varying both $U^*$ and $t^*$ gives a better result than just varying either parameter in this situation.

The results of figure 3 correspond to a doped system. The situation in a half-filled system at these parameters is very similar.

### 3.2. Effective parameters

The variational principle minimizes the variational free energy by changing the effective parameters, so these are a natural second step in our investigation. The bandwidth renormalization corresponds to the value of $t^*$ that is found in the variational scheme. The variational free energy landscape provides insight into the overall structure of the three variational schemes. We only show results for the chain model, where exact results are available.

Figure 5 shows how the variational free energy depends on the effective parameters. At $\beta = 10$, $U = 1$ and $V = 0.1$, the top panel, the free energy surface consists of curves of constant energy that are locally almost parallel lines. The $t^*$ and $U^*$ schemes are restricted to the dashed blue lines, with the associated minimum indicated by the blue dot. The global minimum found by the $(U^*, t^*)$-scheme is given by the black star. All renormalization schemes improve significantly on the original parameters $(U, t)$, i.e. on completely ignoring the intersite interaction.

At $T = 0$, the variational free energy landscape would be constant along lines of constant $U^*/t^*$. There would be one line with the lowest free energy and all three variational schemes would find the same free energy on this line. That the free energy landscape at $\beta = 10$ still resembles the line picture of zero temperature supports the previous notion that it can be considered as almost zero temperature.

At higher temperature $\beta = 3$, shown in the second panel of figure 5, the free energy surfaces are ellipses. We find that the global minimum is quite close to $t^* = 1$, so that changing only $U^*$ gives a better energy than changing only $t^*$, as was also discussed for figure 3.

The third panel of figure 5 shows a rather different situation, $U = 10$ and $V/U = 0.82 > 1/2$. Here, the nearest-neighbor repulsion is so strong that the electrons order themselves into alternating empty and doubly occupied sites [38]. This
ordering is not energetically favorable in the effective Hubbard model, so the variational principle cannot capture it properly. This alternating pattern is even more favorable at half-filling, visible in figure 6, where there are two local minima along the line of constant $t^* = t$ and $U^* = U$ corresponding to single parameter variational schemes. Corresponding results at half-filling are given in figure 6. All results have been obtained at $t = 1$.

A comparison of figures 5 and 6, which differ only in the filling, shows that only the charge ordering physics depends strongly on the density, the free energy surfaces for the other scenarios look very similar with five and with six electrons. These parameters are rather extreme and not necessarily what one would expect in realistic scenarios (see section 3.6), we show them here to clearly illustrate how discontinuities coming from charge-density waves occur. They are associated with a change of sign in the effective interaction $\tilde{U}$. This attractive interaction favors a high double occupancy. This is the effective Hubbard model’s way of describing the charge ordered phase, which indeed has many doubly occupied sites. What is lost in this effective local description is the alternating spatial character of the charge ordered state, which comes directly from the shape of the interaction $V$.

Figure 5. Variational free energy surfaces for five electrons in a six atom chain. The color scale and the contour lines give the variational free energy functional $F_v(U^*, t^*)$. The star is the absolute minimum of this functional and the two blue dots the minima along the blue lines $t^* = t$ and $U^* = U$ corresponding to single parameter variational schemes. Corresponding results at half-filling are given in figure 6. All results have been obtained at $t = 1$.

Figure 6. Variational free energy surfaces for six electrons in a six atom chain. The same results for five electrons are given in figure 5.
makes them look different from discontinuities associated with the metal-insulator transition [42].

The free energy surfaces for the half-filled and the doped system look very similar for both $\beta = 3$ and $\beta = 10$, as can be seen by comparing figures 5 and 6.

Physically, what we see in these figures is that $V > 0$ effectively makes the electrons more delocalized and more likely to doubly occupy a site. This is captured either by reducing the effective local interaction (lowering the potential energy penalty) or by increasing the hopping amplitude (increasing the potential energy gain from delocalization). Both correspond to reducing $U/t$, the balance between potential and kinetic energy in the Hubbard model.

### 3.3. Applicability of Fock bandwidth renormalization

Let us come back to one of our main questions, does Hartree–Fock describe the bandwidth renormalization due to $V$? The Hartree–Fock theory is perturbative in $V$ and can only be expected to work for small $V$, but even there the applicability depends on an approximation: using Wick’s theorem for a correlated starting point.

#### 3.3.1. Chain

For the six-site Hubbard chain, the results of Fock bandwidth renormalization are shown in figure 3. We observe that the Fock result gives a free energy that is strictly larger than the variational approaches. Only at $t = 1, U = 1$ and $\beta = 3$ is the Fock result comparable with the $t^*$ scheme. This is also exactly in the regime where the Fock diagram could be expected to be reliable, $U/t$ is small and the temperature is high so that correlations should be moderate and Wick’s theorem applies, so that the derivation of appendix B is valid.

However, exactly in this regime the bandwidth renormalization does not work very well at capturing the effect of $V$: the $t^*$ and Fock results are equally far off. In general, even though the Fock scheme gives considerably larger free energy deviations than the variational schemes, they are still clearly of the same order of magnitude, usually only 30% larger and they also show the qualitative trends about the role of $V$. This indicates that, although the variational principle performs better, the Fock estimate is still useful as a poor man’s approach to bandwidth renormalization, better than ignoring $V$ completely, especially in the moderately correlated regime.

#### 3.3.2. Honeycomb lattice

Figure 7 shows the effective renormalization $\alpha(t^*)$ for various values of $U$ and $t^*$ (red symbols). We remind the reader that $\alpha$ is calculated in the effective Hubbard model and therefore only depends on the parameters of the effective model, in this case $t^*$, $U$ and $\beta$. As in the exact diagonalization results, we find that $\alpha > 0$, i.e. the bandwidth gets wider due to intersite Coulomb repulsion $V > 0$. At small $t^*$, increasing $U$ leads to a monotonous decrease of $\alpha$. This

![Figure 8. Nearest-neighbor spin correlation for a six-site chain with five electrons. The gray dashed line is the exact result for the extended Hubbard model, the other lines are the predictions of the variational approaches.](image)
can already be understood from the Hartree–Fock perspective, since $U$ reduces $G_{01}$. At larger $t^*$, the situation changes and $\alpha$ is no longer a monotonous function of $U$. For some values of $U$ ($U = 0$, $U = 32$), $\alpha$ seems to approach a constant for large $t^*$ whereas for intermediate values of $U$ the slope appears to stay non-zero within the studied range of $t^*$.

The figure also shows that as $t^*$ goes to 0, $\alpha(t^*)$ also goes to 0. At $t = 0$, the electrons cannot move and there is no dispersion, regardless of the value of $V$. In particular, the nearest-neighbor Green’s function is zero in the $t = 0$ extended Hubbard model. We know that the variational principle must reproduce this, and the only way to do that is to have $t^* = t = 0$. Conversely, any effective model with $t^* = 0$ has $G_{01} = 0$ and can only correspond to $t = 0$. The renormalization of the dispersion does not actually contain any information about $V$ in this extreme situation and the $t^*$ variational scheme is not applicable.

Figure 7 also compares the renormalization factor $\alpha$ obtained with the $t^*$ variational principle (red symbols) to the Hartree–Fock result (blue lines). As expected, there is an exact match at $U = 0$. This happens because the Wick decoupling that is used to derive the Hartree–Fock expression is exact at $U = 0$. At small $U$, the Hartree–Fock results initially match the variational results but subsequently deviate at larger $t^*$. Larger $t^*$ corresponds to larger $\beta t^*$, i.e., it effectively means that the temperature is lower. Two-particle correlations are typically more important at low temperature, so that Hartree–Fock becomes less appropriate. The results at $U = 4$ and $U = 8$ show that Hartree–Fock deviates from the variational results in both directions. It initially overestimates the bandwidth renormalization and then saturates to a constant value of $\alpha$ that is too low.

Quantitatively, we observe deviations as large as a factor of 2 at intermediate and large interaction strengths. This clearly shows the limits to the quantitative usefulness of the Hartree–Fock approach for incorporating bandwidth renormalization. Whether Hartree–Fock overestimates or underestimates the bandwidth renormalization seems to depend quite sensitively on the value of $t^*$ and $U$. However, in all cases studied, the Fock diagram at least produces the correct sign and order of magnitude of the bandwidth renormalization.

### 3.4. Observables

For the exact diagonalization results, we have access to all observables of the extended Hubbard model. This allows us to see how well the observables of the reference system match those of the true extended Hubbard model. Theoretically, variational principles only make statements about (free) energies and there is no ground to identify observables of the reference system with those of the original system. Practically, such an identification is still regularly made.

A previous study of the variational principle for the effective interaction [32] found two main conclusion regarding observables: first, using the effective interaction leads to the exact double occupancy at small $V$, since the double occupancy and the interaction strength are conjugate variables. Second, the variational principle does not predict the (momentum-resolved) charge susceptibility very well, since it depends explicitly on $V$ even in the weakly interacting limit. On the other hand, the prediction for many other observables is quite reasonable even though the variational principle technically only deals with the free energy.

Regarding the first point, according to the same conjugate variable argument (see appendix D), the $t^*$ scheme gives the exact value of the nearest-neighbor Green’s function $\langle \hat{c}_0 \hat{c}_1 \rangle$, to linear order in $V$. This exact statement is borne out by the numerical results shown in appendix D. In this case, the $U^*$ scheme actually only shows rather small deviations.

For a generic observable not linked to any of the variational parameters, such as the nearest-neighbor spin correlation shown in figure 8, none of the schemes necessarily predicts the correct linear coefficient in $V$. In fact, even in the moderately correlated regime at $t = 10$, $U = 1$, $\beta = 3$ (top left), all variational schemes deviate from the exact solution already in linear order, although the deviations are not very large in an absolute sense.

Overall, the best predictions for the free energy were given by the $U^*$ and ($U^*, t^*$) variational schemes and these also perform best at predicting the spin correlation. However, noticeable deviations from the exact result occur in all four panels, at large $V$.

Coming to the second point, the nearest-neighbor correlation function $\langle n_0 n_1 \rangle$ is a clear example of a quantity that depends explicitly on $V$. Indeed, we find that the variational principle does not capture this observable accurately, as shown in figure 9. The variational principle underestimates how much the nearest-neighbor correlation function depends on $V$. This underestimation happens for all variational schemes and at almost all parameters we studied.
variational formula (3) to determine which values of the inverse temperature $1/\beta = U/\varepsilon_{\text{F}} = 3.63$. This entire process was done at several fixed $1/\beta = U/\varepsilon_{\text{F}} = 3.63$. This entire process was done at several fixed

$$\beta \alpha$$

3.5. Graphene

Table 1 shows the bandwidth renormalization in graphene. This is based on the graphene parameters of [19]. As in the rest of this paper, we only consider the nearest-neighbor interaction $V$, with fixed $V/U = 2.03/3.63$. These results have been obtained from simulations at fixed interaction $U = 1$, where the parameter $r^*$ is varied to determine the curve $\alpha(r^*)$, as in figure 7. We then use the variational formula (3) to determine which $r^*$ corresponds to $U/U = 1/\beta = 1 / 3.63$. This entire process was done at several fixed values of the inverse temperature $\beta$, as indicated in the table.

Comparing the bandwidth renormalization with the interaction renormalization, here we find $U/r^* \approx 1.8$ for the bandwidth renormalization at the lowest temperature, which is comparable to the $U^*/t \approx 1.6 \pm 0.2$ obtained for interaction renormalization [19]. That result included interactions beyond nearest-neighbors, but for graphene the interaction renormalization based only on nearest-neighbors agrees within error bars, as can be seen in the supplemental material of [19].

Table 1 shows significant temperature dependence, with the bandwidth renormalization getting stronger as the temperature is lowered. Here we should note that all temperatures listed here are at least one order of magnitude above room temperature when the graphene parameters are filled in, so that these results have to be extrapolated to make predictions about room temperature. Given the temperature dependence observed, the bandwidth renormalization at room temperature is expected to be at least as strong as it is at the lowest temperatures studied here. We also note that the temperatures investigated here are already small compared to the intersite interaction, $\beta V \gg 1$.

The bandwidth renormalization comes from the intersite part of the Coulomb interaction $V$. In an experimental set-up, this implies that control over the intersite interaction $V$ gives control over the effective bandwidth and the amount of correlation in the system. For example, reducing $V$ should lead to a reduction of the bandwidth renormalization, so to a smaller bandwidth and a more correlated system.

3.6. Benzene

We have used a periodic chain with six sites to illustrate the mapping of the extended Hubbard model onto an effective Hubbard model. A chain of six sites, with one orbital per site, is also a toy model for a benzene molecule [51–53]. Ab initio density matrix downfolding (AIDMD) has been used [54] to study the appropriateness of such a downfolding from 30 electrons (5 per site) to 6 (1 per site) in a single-orbital model. They present a downfolding both to a Hubbard model and to an extended Hubbard model, with the idea that both have a density matrix similar to the original benzene system. Using diffusion Monte Carlo (DMC), they find an extended Hubbard model with nearest-neighbor hopping $t = 2.76$ eV, local interaction $U = 10.92$ eV, nearest-neighbor interaction $V_{01} = 7.13$ eV and next-nearest-neighbor interaction $V_{02} = 5.41$ eV. In the same scheme, when downfolding to a Hubbard model with local interactions only, the optimal parameters are $r^* = 2.80$ eV and $U^* = 3.99$ eV, as illustrated by the green square in figure 10. We have used the variational principle to map the extended Hubbard model for benzene onto an effective Hubbard model and find comparable values, as shown by the black star in figure 10. The similar results of both techniques for mapping onto a Hubbard model, the variational principle and AIDMD, are a sign that these approaches give reasonable physical results.

The main difference between the variational method and AIDMD is the temperature. The AIDMD is fundamentally based on zero temperature DMC calculations (although their scheme also involved solving the benzene molecule at a low temperature $T = 1/20$ eV), whereas the variational principle is naturally formulated in terms of the free energy at finite temperature. In particular, according to the variational principle the parameters of the optimal Hubbard model can change as a function of temperature. For the benzene molecule, this temperature dependence is only relatively weak. In other situations, especially close to a metal-insulator transition, the temperature is more important [42].

Looking at observables, Changlani et al [54] also found that the Hubbard description is reasonable for many aspects except for the nearest-neighbor density correlation function.

Table 1. Effective hopping parameters for graphene, using the parameters of [19], $t = 2.8$ eV, $U/t = 3.63$ and $V/t = 2.03$.

| $\beta U$ | $\beta$ (eV$^{-1}$) | $t'/t$ | $U/t'$ | $\alpha$ |
|-----------|-------------------|--------|--------|---------|
| 1.0       | 0.1               | 1.13   | 3.33   | 0.06    |
| 4.0       | 0.4               | 1.30   | 2.86   | 0.14    |
| 9.0       | 0.9               | 1.59   | 2.22   | 0.30    |
| 16.0      | 1.6               | 2.07   | 1.82   | 0.51    |

Figure 10. Variational free energy for the benzene model. The black star corresponds to the optimal Hubbard model in our method, the green square to the optimal Hubbard model according to the DMC results of [54]. The black and green lines denote constant $U/t$, which would be the only relevant parameter at zero temperature. The DMC values of the parameters are used, in particular $t = 2.76$ and $U = 10.92$ (blue dashed lines), all energies in units of eV.

* $U^*/t^* = \text{const.}$
This conclusion is consistent with the results of figure 9 and of [32].

The magnitude of the non-local Coulomb interactions in benzene is considerable compared to the on-site interaction. In fact, even though $V > U/2$, charge-ordering does not occur due to the large next-nearest-neighbor interaction. The competition between local and nonlocal interactions leads to a considerably smaller value of $U^*/t^*$ in the effective Hubbard model. In this example, the renormalization of the interaction $U^*$ is large and the bandwidth renormalization is relatively small. The predominance of interaction renormalization over bandwidth renormalization in the six site chain is consistent with our results presented here.

4. Conclusions and discussion

Bandwidth renormalization is one of the most direct effects of intersite interactions. We have used the variational principle to study how the nearest-neighbor interaction in particular renormalizes the electronic bandwidth. We have found that the bandwidth of graphene can be widened by as much as 50%–100%.

Determining the bandwidth renormalization is useful when a material is studied using a computational method that can only deal with the Hubbard model without intersite interactions, such as dynamical mean-field theory. The variational principle gives a good estimate for how much intersite interactions change properties of the system such as the free energy and the kinetic energy. Of course, this estimate does not work very well for observables that depend explicitly on the spatial interactions, such as the charge-charge correlation function. The variational principle also breaks down when the intersite interactions are so strong that they qualitatively change the physics of the system away from what can be in expected in a Hubbard model. The charge order physics visible in the third panel of figure 5 is a good example of this.

Previously, it has been suggested [55] that a Hartree–Fock expression could be used to determine the bandwidth renormalization. This approach assumes that the underlying Hubbard model is uncorrelated. This assumption works well at small interaction strengths, where the system is only moderately correlated. On the other hand, it performs poorly at large $U$ or low temperature, as we have illustrated by comparing its results with the variational principle. Even then, qualitatively, it at least predicts the correct sign of the bandwidth renormalization, but more detailed aspects such as the non-monotonous dependence on $U$ (figure 7) are not reproduced.

We have also identified some slightly pathological situations where the Hartree–Fock approach fails, mostly in few-electron systems where (spatial) correlations are naturally large, as illustrated in appendix G. Charge conservation induces these spatial two-particle correlations. Computational approaches for nanoscopic systems [56–58] need to take spatial correlations into account.

A perhaps somewhat surprising conclusion is that in our study, relatively high temperatures turn out to be more interesting than very low temperatures (and that in the benzene system, $\beta t = 10$ is already a low temperature when looking at charge fluctuations). The origin is that at higher temperatures, entropy competes with kinetic and potential energy and two relevant dimensionless parameters can be formed from $t$, $U$ and $T$. At low temperature, only the ratio $U/t$ matters.

The energy scale of charge fluctuations relevant in this study is given by $t$, $U$ and $V$ themselves, for solid state applications these are typically on the eV scale. In contrast to this large energy scale for charge fluctuations, the spin (and superconducting) physics typically plays out at smaller emergent temperature scales, such as those given by the Heisenberg exchange interaction and by the Kondo temperature. Because of this, a low temperature description of $V$ can be appropriate even when the temperature is high compared to magnetic scales.

In experiments, Hubbard physics can be realized on various energy and temperature scales, depending on the bandwidth of the material in question. Traditional condensed matter applications have $t$ on the eV scale so that $\beta t \gg 10$ and the low temperature limit applies to the charge excitations. Recent experiments modelling the Hubbard model (without $V$) using fermions in optical lattices [59, 60] were performed down to temperatures corresponding to $\beta t \approx 1.6$, similar to the values studied here. Current dipolar fermion experiments [61] reach $\beta t = 10^{-2}$. These are clearly not in the $\beta t \gg 1$ regime. As another example, magic-angle twisted bilayer graphene [62, 63] has a bandwidth on the meV scale, so that temperature effects come into play at a few to a dozen Kelvin.

From a computational point of view, there are advantages of studying charge fluctuations at temperatures comparable to the hopping. Cluster approaches to the extended Hubbard model [64, 65] are computationally lighter at high temperature and the phase diagrams of [65] go down to $\beta t \approx 5$. For diagrammatic extensions of DMFT [17], at higher temperatures, the spin fluctuations are much less important allowing for a clearer vision on the charge fluctuations themselves and validating ladder approaches without feedback between the channels.

This suggests that future comparisons of computational approaches to the extended Hubbard model should occur not just at low temperatures but also at $T \approx t, U, V$. In fact, the observation that similar phase boundaries [22, 66] are found in a method based on bandwidth renormalization [22] and in a method based on vertex corrections [66], more associated with interaction renormalization, might have to do with the fact that these comparisons were made at relatively low temperature ($\beta t = 12.5$) where the Hubbard model has only a single parameter $U/t$ (especially since both studies [22, 66] were DMFT based and did not specifically deal with spin excitations that could introduce an additional small energy scale). This single effective parameter might explain why two very different approaches ended up with the same numerical results.
Acknowledgments

The authors thank Alexander Lichtenstein for useful discussion. MIK and EGCPV/L acknowledge support from ERC Advanced Grant 338957 FEMTO/NANO. The authors acknowledge the North-German Supercomputing Alliance (HLRN) for providing HPC resources under project hbp00046.

Appendix A. Derivation of variational approaches

In this appendix, we give an overview of the derivations of the variational approaches dealing with the intersite interaction $V$. We restrict our analysis to nearest-neighbor interactions, since this situation is easiest to interpret as a renormalization of the bandwidth [55]. In general, $V$ can induce a change in electron density in addition to renormalization of the bandwidth and interaction. Within a variational framework, this can be described by using the chemical potential $\mu$ as an additional variational parameter. In this work, we only consider situations where this density change does not occur: finite systems in the canonical ensemble and a bipartite (graphene) lattice at particle-hole symmetry. Comparisons with experiment are also usually made by fixing the density. The Hubbard model only provides a description of the low energy physics and going from band structure to Hubbard model involves shifts in the chemical potential anyway.

A.1. $t^*$

The derivation of the variational formula for the bandwidth renormalization is similar to that of the effective interaction [19]. The Hamiltonians of the original and the effective system are

$$H = -t \sum_{\langle ij \rangle, \sigma} c^\dagger_{i \sigma} c_{j \sigma} + U \sum_{i} n_{i \uparrow} n_{i \downarrow} + \frac{V}{2} \sum_{\langle ij \rangle} n_{i} n_{j}$$

$$H^* = -t^* \sum_{\langle ij \rangle, \sigma} c^\dagger_{i \sigma} c_{j \sigma} + U \sum_{i} n_{i \uparrow} n_{i \downarrow}.$$ 

To simplify the notation, we introduce the bandwidth renormalization as $\Delta t = t^* - t$, and calculate the required difference of the Hamiltonians.

$$\frac{1}{N} \langle H - H^* \rangle_\ast = \Delta t \sum_{\langle ij \rangle, \sigma} \langle c^\dagger_{i \sigma} c_{j \sigma} \rangle_\ast + \frac{V}{2} \sum_{\langle ij \rangle} \langle n_{i} n_{j} \rangle_\ast$$

$$= -2z \Delta t \cdot G_{01} + \frac{V}{2} \cdot z \langle n_{01} \rangle_\ast.$$

Here $z$ is the coordination number of the lattice and $G_{01} = -\frac{1}{2} \left( c^\dagger_{1 \uparrow} c_{0 \uparrow} + c^\dagger_{1 \downarrow} c_{0 \downarrow} \right)_\ast$ is the Green’s function averaged over spin. We continue with the derivative with respect to $t^*$.

$$\frac{1}{N} \partial_{t^*} \langle H - H^* \rangle_\ast = -2z G_{01} - 2z \Delta t \cdot \partial_{t} G_{01} + \frac{V}{2} z \partial_{t^*} \langle n_{01} \rangle_\ast,$$

and the derivative of the free energy of the reference system $F^*$ with respect to $t^*$,

$$\frac{1}{N} \partial_{t^*} F^* = -\frac{1}{N} \sum_{\langle ij \rangle, \sigma} \langle c^\dagger_{i \sigma} c_{j \sigma} \rangle_\ast = 2z G_{01}.$$

Combining the two gives,

$$\frac{1}{N} \partial_{t^*} (F^* + \langle H - H^* \rangle_\ast) = -2z \Delta t \cdot \partial_{t^*} G_{01} + \frac{V}{2} z \partial_{t^*} \langle n_{01} \rangle_\ast,$$

(A.1)

so that the minimum is found when

$$\Delta t = \frac{V \partial_{t^*} \langle n_{01} \rangle_\ast}{4 \partial_{t^*} G_{01}}.$$ (A.2)

The factor of 4 essentially comes from two factors of 2, namely the sum over spins for the Green’s function and the fact that the hopping term is directional (every bond counts twice) and the intersite interaction is not.

A.2. $(U^*, t^*)$

The derivation proceeds in a somewhat similar way when both the hopping and the interaction are taken as variational parameters.

$$H = -t \sum_{\langle ij \rangle, \sigma} c^\dagger_{i \sigma} c_{j \sigma} + U \sum_{i} n_{i \uparrow} n_{i \downarrow} + \frac{V}{2} \sum_{\langle ij \rangle} n_{i} n_{j}$$

$$H^* = -t^* \sum_{\langle ij \rangle, \sigma} c^\dagger_{i \sigma} c_{j \sigma} + U^* \sum_{i} n_{i \uparrow} n_{i \downarrow}.$$ (A.3)

We introduce $\Delta t = t^* - t$ and $\Delta U = U - U^*$ (note the different sign).

$$\langle H - H^* \rangle_\ast = -2z \Delta t \cdot G_{01} + \Delta U \cdot D + \frac{V}{2} \cdot \langle n_{01} \rangle_\ast,$$

with $D = \langle n_{01} \rangle_\ast$, and finally

$$F_v = F^* + \langle H - H^* \rangle_\ast.$$ (A.4)

One could now proceed in the same way as before and calculate the derivatives of $F_v$ with respect to $U^*$ and $t^*$, obtain the coefficient $\alpha$ for both parameters and so on. In practice, for the two-parameter optimization, we did not explicitly calculate numerical derivatives of the variational free energy, instead we simply calculated all necessary observables on a grid in $t^*$-$U^*$ space and subsequently determine the minimum and location of the minimum $(t^*, U^*)$ of the variational free energy $F_v$ for any desired combination of original variables $(t, U, V)$.

Appendix B. Hartree–Fock

The central idea of this work is to write the Hamiltonian of the extended Hubbard model as the sum of the Hubbard Hamiltonian and the intersite interaction,

$$H_{\text{ext.}} = H_{\text{Hub.}} + VH_{\text{nl}}.$$ (B.1)
The idea of this decomposition is that the first term on the left-hand side is solvable, either exactly as in the situations studied in this work or approximately using DMFT in realistic materials, whereas the combination is not. Instead of the variational approach, it is possible to consider perturbation theory in $V$ to deal with $H_{\text{H}}$.

If the Hubbard Hamiltonian would have been a non-interacting system, Wick’s theorem would apply and the Hartree and the Fock terms would be the lowest order (linear in $V$) contributions in this perturbation theory. The Hartree term renormalizes the chemical potential, the Fock term can be interpreted as a renormalization of the hopping [55],

$$\Delta t_{ij} = -V_{ij}G_{ij}. \quad (B.2)$$

It is important to note here that $G_{ij}$ corresponds to the exact Green’s function of the interacting Hubbard model, which plays the role of $H_0$ in equation (B.1). Because the Hubbard Hamiltonian describes a correlated system, even the linear in $V$ term in perturbation theory involves additional, more complicated terms that contain higher-order correlation functions (see appendix C). Hartree–Fock only captures the correct linear in $V$ behaviour when Wick’s theorem applies, i.e. for $U = 0$.

The variational principle reproduces the Hartree–Fock expression in the uncorrelated limit, which shows that the variational principle also has the exact linear in $V$ terms at $U = 0$. As mentioned before, this limit corresponds to the Wick decomposition of correlation functions, which states that higher-order correlations are simply a combination of single-particle Green’s functions. This allows us to simplify the variational formula. It is most convenient to use the notation of the nearest-neighbor interaction in terms of density fluctuations,

$$\langle VH_{ab} \rangle = \frac{eV}{2} \left( \langle n_0 n_1 \rangle - \langle n_0 \rangle \langle n_1 \rangle \right), \quad (B.3)$$

where $\epsilon$ is the number of nearest-neighbors. The Wick decomposition of the correlation function gives

$$\langle n_0 n_1 \rangle \overset{\text{Wick}}{=} G_{00}G_{11} - G_{01}G_{10} \quad (B.4)$$

$$\langle n_0 n_1 \rangle - \langle n_0 \rangle \langle n_1 \rangle \overset{\text{Wick}}{=} -G_{01}G_{10}. \quad (B.5)$$

This final expression contains the square of the nearest-neighbor Green’s function and taking the derivative is simple,

$$\partial_r \left( \langle n_0 n_1 \rangle - \langle n_0 \rangle \langle n_1 \rangle \right) = -2G_{01}\partial_r G_{01}. \quad (B.6)$$

Inserting this into the variational expression, the derivative of the Green’s function cancels and we are left with the Hartree–Fock formula, (B.2). Note that this cancellation does not work when $\partial_r G = 0$, as in appendix G.

The first term in (B.4) corresponds to the Hartree shift in the chemical potential, which is relevant when the Hamiltonian is written in terms of the density instead of the density fluctuations,

$$\delta \mu = V G_{00}. \quad (B.7)$$

To summarize, the Hartree–Fock expression for the band-width renormalization corresponds to two assumptions: small $V$ so that first-order perturbation theory in $V$ is justified and the Wick decomposition of expectation values of the associated Hubbard model without $V$. The variational principle does not have the second assumption, in the sense that it is applicable at arbitrary interaction strength $U$ and, for some observables, even exact to first order in $V$, as shown in appendix D.

**Appendix C. Diagrammatic interpretation of the variational formula**

The previous Appendix shows how the (diagrammatic) Hartree–Fock expression arises from the variational formula by assuming Wick’s theorem. A diagrammatic interpretation of the complete variational formula is also possible, as we show in this Appendix. In this case, the diagrammatic expressions involve vertex corrections and higher-order correlators. These are typically not accessible in actual calculations, so that the diagrammatic expression provided in this Appendix are not meant for computational purposes. Because of this, and to simplify the notation, we will proceed with the equations symbolically and will drop all numerical prefactors (including temperature), spin labels and site indices/moments. Regarding the last point, these equations should be understood as matrix equations in real space.

As in the Hartree–Fock theory, we again interpret $\Delta t = \delta V$ as a self-energy (contribution) $\Sigma V$. This is the linear contribution in $V$ towards the self-energy, starting from an interacting starting point $t_0, U_0$. All expectation values, Green’s functions and correlation functions given below are those of the interacting starting point. Since the diagrammatic interpretation depends on the smallness of $V, t' \approx t$ and we drop all * labels to simplify the notation further.

Derivatives with respect to $t$ are equivalent to higher-order correlation functions. The derivative in the denominator of (4) gives

$$\partial_r \langle c_i^\dagger c_i \rangle = \langle c_i^\dagger c_i c_i^\dagger c_i \rangle - \langle c_i^\dagger c_i \rangle \langle c_i^\dagger c_i \rangle \quad (C.1)$$

$$= GG\gamma GG - GG \quad (C.2)$$

$$= -GG(1 - \gamma GG), \quad (C.3)$$

where $\gamma$ denotes the two-particle vertex, a connected amputated two-particle correlation function. Here, the labels need to be treated with some care, since the two-particle vertex $\gamma$ is obtained from the correlation function by subtraction of two disconnected contributions. If we use labels 1, 2, 3, 4 to denote the momentum quantum number, for which the single-particle Green’s function is diagonal, $\gamma$ is defined as

$$G_1G_2G_3G_4\gamma_{1234} = \left\langle c_i^\dagger c_j^\dagger c_k^\dagger c_l^\dagger \right\rangle - G_1G_3(\delta_{12}\delta_{34} - \delta_{14}\delta_{32})$$

and we find...
contribution of site labels. Looking at the vertex $A_{abcd}$, for example, there are four site labels occurring in two pairs, $A_{abcd} = \frac{\partial}{\partial a_{ab}} \langle n_{ab} \rangle$. With both the interaction $V$ and the hopping $t$ restricted to nearest-neighbors, the labels of the vertex $A$ are two pairs of nearest-neighbors. In the DMFT spirit, we can then assume that the dominant correlations are between electrons on the same site, so that we only take the $\delta_{at} \delta_{bd}$ contribution of $A_{abcd}$ into account. The second mean-field assumption is that $A_{abcd}$ then factorizes into $\lambda_0 G_{ab} \lambda_b$, where $\lambda$ denotes a local fermion-boson vertex (see figure C2). This brings us back to the kind of diagrams regularly drawn in, for example, the dual boson theory [66]).

This construction suggests that if one is interested dealing with the bandwidth renormalization beyond Fock, in an impurity model based method, then the simplest approach might be to simply attach the equal-time fermion-boson vertex on both ends of the Fock diagram. For a single-band system, the equal-time impurity fermion-boson vertex is simply a number related to the double occupancy and the density. Verifying such a simple scaling relation for the Fock energy could be done in a cluster approach to the extended Hubbard model [64, 65], since it has direct access to the self-energy and the double occupancy of the extended Hubbard model. For a multiband system, the equal-time fermion-boson vertex is a matrix in orbital space.

**Appendix D. Exact linear dependence of observables**

The variational approach is non-perturbative and can in principle be applied at arbitrary values of $V$. As usual with variational approaches, it is in general not guaranteed that the resulting observables correspond to the true value of the system. Formally, the variational approach just provides an upper bound for the free energy of the system. However, at least at small $V$ it can be shown that the variational approach reproduces the exact linear in $V$ term of the observable conjugate to the variational parameter. This has been shown [32] for the double occupancy when the Hubbard interaction $U$ is the variational parameter. A similar result holds for the hopping term when $t$ is the variational parameter.

The conjugate variable of $t$ is $-\left\langle c_i^\dagger c_i \right\rangle = G_{01}$, the equal-time nearest-neighbor Green’s function or the off-diagonal element of the density matrix. It is the first derivative of the free energy with respect to $t$, with a factor of two coming from the number of spin flavors,

$$G_{01} = \frac{1}{2} \partial_t F.$$  \hspace{1cm} (D.1)

This makes the derivative of the Green’s function with respect to $V$ a second derivative of the free energy,

$$\partial_V G_{01} = \frac{1}{2} \frac{\partial^2 F}{\partial V \partial t} \hspace{1cm} (D.2)$$

$$= \frac{1}{4} \partial_t \langle n_0 n_1 \rangle,$$  \hspace{1cm} (D.3)

where the second line is obtained by interchanging the derivative operators. If the Green’s function depends smoothly on $V$, we find

$$G_{01}(t, V) \equiv G_{01}(t, V = 0) + V \cdot \frac{1}{4} \partial_t \langle n_0 n_1 \rangle.$$  \hspace{1cm} (D.4)
The variational principle predicts

\[ G_{01}(t, V) = G_{01}(t^*, V = 0) \]

\[ = G_{01}(t, V = 0) + \partial_t G_{01} \cdot \Delta t \]

\[ \approx G_{01}(t, V = 0) + V \cdot \frac{1}{4} \partial_t^2 \langle n_0 n_1 \rangle, \]  

(D.5)

and (D.4) and (D.5) are clearly identical. Note that this proof assumes the smallness of \( V \), which implies smallness of \( \Delta t \) so that it does not matter if the observables are calculated at \( t \) or at \( t^* \).

This proof is based on the interchangeability of second derivatives and does not necessarily apply at phase transitions, where the free energy is not smooth.

We illustrate these exact statements with results for the extended Hubbard model with six sites. In particular, figure D1 shows that the \( U^* \) scheme captures the exact linear coefficient in the double occupancy. The \( t^* \) and the Hartree–Fock schemes show significant deviations in the double occupancy at small \( V \), as is visible in the \( t = 1, U = 1, \beta = 3 \) results (bottom left).

For the \( t^* \) scheme, the nearest-neighbor Green’s function is the conjugate variable. Figure D2 shows that this scheme indeed matches the exact nearest-neighbor Green’s function to linear order in \( V \). The \( U^* \) scheme does not capture the linear coefficient exactly, but it is sufficiently close that the deviation is not visible in the figure.

**Figure D1.** Double occupancy for a six-site chain with five electrons. The gray dashed line is the exact result for the extended Hubbard model, the other lines are the predictions of the variational approaches.

**Figure D2.** Nearest-neighbor Green’s function for a six-site chain with five electrons. The gray dashed line is the exact result for the extended Hubbard model, the other lines are the predictions of the variational approaches.

**Appendix E. Comparison of variational approaches**

With two variational parameters, \( t^* \) and \( U^* \), the question arises which variational parameter is best, that is, which parameter leads to the lowest free energy. As in the proof of the exact observables of appendix D, this question can be addressed for small \( V \) simply by taking derivatives of the free energy.

Let us start with the variation of the bandwidth. We know that there should be a minimum in the variational free energy at some \( t^* \), and since \( V \) is small we know that \( t^* - t = \Delta t \) is small and that we can expand the variational free energy only up to second order in \( t \) to find the location of the minimum. To simplify the notation, we write the variational free energy functional around the original value \( t_0 \) as

\[ \Phi(x) = \Phi(0) + \partial_x \Phi \mid_{x=0} \cdot x + \frac{1}{2} \partial_x^2 \Phi \mid_{x=0} \cdot x^2, \]

with minimum at

\[ \Delta t = x_{min} = -\partial_x \Phi \mid_{x=0} + \frac{1}{2} \partial_x^2 \Phi \mid_{x=0}. \]

The value of the free energy functional at its minimum \( x_{min} \) corresponds to the variational free energy.

\[ \Phi(\Delta t) = \Phi(0) - \frac{1}{2} \partial_x^2 \Phi \mid_{x=0}. \]  

(E.1)
Now, let us evaluate these derivatives as in (A.1).
\[
\partial_t \Phi = x \cdot 2z \partial_x \mathcal{G} + \frac{z V}{2} \partial_x \langle n_{n_1} \rangle \\
\partial^2_t \Phi = x \cdot 2z \partial_x^2 \mathcal{G} + 2z \partial_x \mathcal{G} + \frac{z V}{2} \partial^2_x \langle n_{n_1} \rangle
\]
We should note that we are evaluating these derivatives at the original value of \( t \), i.e. at \( t = 0 \) so that the first term in these equations vanishes,
\[
\partial_t \Phi \big|_{t=0} = -\frac{z V}{2} \partial_x \langle n_{n_1} \rangle \\
\partial^2_t \Phi \big|_{t=0} = 2z \partial_x \mathcal{G} + \frac{z V}{2} \partial^2_x \langle n_{n_1} \rangle.
\]
The lowest (quadratic) order in \( V \) is obtained by neglecting the second term in the second equation. Going back to (E.1) and changing the derivatives back from \( t \) to \( t^* \), we obtain
\[
\Phi(\Delta t) - \Phi(0) = -\frac{z V^2}{16} \left( \frac{\partial_{r^*}}{\partial_{r^*}} - \mathcal{G} \right).
\] (E.2)
The expansion in \( U \) is very similar and gives
\[
\Phi(\Delta U) - \Phi(0) = -\frac{z V^2}{8} \left( \frac{\partial_{r^*}}{\partial_{r^*}} - \langle n_{n_1} \rangle \mathcal{G} \right).
\] (E.3)
Together, we can determine which variational principle works best by looking at \( \frac{1}{\Delta t} \left[ \Phi(\Delta U) - \Phi(\Delta t) \right] \), as is done in the main text. For the exact diagonalization results, we have access to the true free energy and can use it as an absolute reference point, as is done in figure 3.

Appendix F. Zero temperature

At zero temperature, there is effectively only a single parameter \( U/t \) in the Hubbard model. This means that varying \( t \) while keeping \( U \) constant and vice versa results in the same set of density operators (ground states) \( \rho_{n_1 \uparrow} \). Since both variational approaches have the same variational space, they find the same optimal density operator and corresponding observables. This shows that both variational approaches are equivalent at \( T = 0 \). We expect this equivalence to still hold approximately at finite but very low temperature, if \( T \) is smaller than any relevant effective energy scale in the problem.

Appendix G. Single electron and dimer

Some small systems can be understood analytically. If there is only a single electron (in a lattice of arbitrary size), there is no electron-electron interaction and both \( U \) and \( V \) do not change the properties of the system. In particular, this means that there is no renormalization of the bandwidth by \( V \).

This statement holds in the canonical ensemble, where the particle number can indeed be fixed to a single electron and the ensemble averages over configurations with one spin up electron and those with one spin down electron, so that the Green’s function retains its spin symmetry. On the other hand, in the grand canonical ensemble at finite temperature, the ensemble averages always includes configurations with more than one electron, where there is interaction so the bandwidth renormalization due to \( V \) is finite.

The variational principle correctly captures the lack of self-interaction in the canonical ensemble. The numerator of the variational formula contains \( \partial_x \langle n_{n_1} \rangle \), and \( \langle n_{n_1} \rangle = 0 \) since the single electron cannot be at site 0 and site 1 simultaneously. This results in \( \Delta t = 0 \).

The derivation of the Hartree–Fock method at finite temperature naturally works in the grand canonical ensemble, where the density is not fixed and where the bandwidth renormalization is finite. Indeed, the Green’s function in the Hartree–Fock expression does not vanish.

At zero temperature, on the other hand, the density does not fluctuate and we are back to the situation where there should be no renormalization of the bandwidth. In fact, \( t^* \) is the only dimensionful parameter in this case and as a result the dimensionless quantity \( G_{01} \propto E_{ka}/t^* \) is independent of \( t^* \). This means that both the numerator and the denominator of the variational formula are zero and that they cannot be divided out to derive the Hartree–Fock formula. So in this case, the Wick decoupling still works but does not lead to the Hartree–Fock expression.

We note that although there is correlation even without electronic interaction in the single-electron system: \( 0 = \langle n_{n_1 \uparrow} \rangle \neq \langle n_{n_1 \downarrow} \rangle = 1/4N^2 \), with \( N \) the number of lattice sites.

As a second example, the Hubbard dimer, a system consisting of two Hubbard atoms, provides perhaps the simplest tractable example of intersite Coulomb interactions. In this case, the variational mapping \( U^* = U - V \) is exact [19]. The \( t^* \) variational principle is not exact.

In the dimer, the Hartree–Fock approach performs poorly at any \( U > 0 \). With only two sites in the system, any on-site correlation automatically corresponds to spatial correlation: if the second electron is not on site 1 it is always on site 2.

ORCID iDs

Malte Schüler https://orcid.org/0000-0002-8472-3745
Tim O Wehling https://orcid.org/0000-0002-5579-2231
Mikhail I Katsnelson https://orcid.org/0000-0001-5165-7553
Erik GCP van Loon https://orcid.org/0000-0002-5436-290X

References

[1] Schubin S and Wonsowsky S 1934 On the electron theory of metals Proc. R. Soc. A 145 159–80
[2] Hubbard J 1963 Electron correlations in narrow energy bands Proc. R. Soc. A 276 238–53
[3] Gutzwiller M C 1963 Effect of correlation on the ferromagnetism of transition metals Phys. Rev. Lett. 10 159–62
[4] Kanamori J 1963 Electron correlation and ferromagnetism of transition metals Prog. Theor. Phys. 30 275–89
[5] Hubbard J 1964 Electron correlations in narrow energy bands. III. An improved solution Proc. R. Soc. A 281 401–19
[6] Nagaoka Y 1966 Ferromagnetism in a narrow, almost half-filled e band Phys. Rev. 147 392–405
[7] Imada M, Fujimori A and Tokura Y 1998 Metal-insulator transitions Rev. Mod. Phys. 70 1039–263
[8] Scalapino DJ 2012 A common thread: the pairing interaction for unconventional superconductors Rev. Mod. Phys. 84 1383–417
[9] Biermann S, Aryasetiawan F and Georges A 2003 First-principles approach to the electronic structure of strongly correlated systems: combining the GW approximation and dynamical mean-field theory Phys. Rev. Lett. 90 086402
[10] Aryasetiawan F, Imada M, Georges A, Kotliar G, Biermann S and Lichtenstein A I 2004 Frequency-dependent local interactions and low-energy effective models from electronic structure calculation Phys. Rev. B 70 195104
[11] Kinza M and Honerkamp C 2015 Low-energy effective interactions beyond the constrained random-phase approximation by the functional renormalization group Phys. Rev. B 92 045113
[12] Honerkamp C, Shinaoka H, Assaad F F and Werner P 2018 Limitations of constrained random phase approximation downward Phys. Rev. B 98 235151
[13] Metzner W and Vollhardt D 1989 Correlated lattice fermions in d = ∞ dimensions Phys. Rev. Lett. 62 324–7
[14] Georges A, Kotliar G, Krauth W and Rozenberg M J 1996 Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions Rev. Mod. Phys. 68 13–125
[15] Kotliar G, Savrasov S Y, Haule K, Oudovenko V S, Parcollet O and Marianetti C A 2006 Electronic structure calculations with dynamical mean-field theory Rev. Mod. Phys. 78 865–951
[16] Maier T, Jarrell M, Pruschke T and Hettler M H 2005 Quantum cluster theories Rev. Mod. Phys. 77 1027–80
[17] Rohringer G, Hafermann H, Toschi A, Katkin A A, Antipov A E, Katsnelson M I, Lichtenstein A I, Rubtsov A N and Held K 2018 Diagrammatic routes to nonlocal correlations beyond dynamical mean-field theory Rev. Mod. Phys. 90 025003
[18] Ulybyshev M V, Buividovich P V, Katsnelson M I and Polikarpov M I 2013 Monte Carlo study of the semimetal-insulator phase transition in monolayer graphene with a realistic interelectron interaction potential Phys. Rev. Lett. 111 056801
[19] Schüler M, Rössner M, Wehling T O, Lichtenstein A I and Katsnelson M I 2013 Optimal Hubbard models for materials with nonlocal coulomb interactions: graphene, silicene, and benzene Phys. Rev. Lett. 111 036601
[20] Schüler M, van Loo E G C P, Katsnelson M I and Wehling T O 2018 First-order metal-insulator transitions in the extended Hubbard model due to self-consistent screening of the effective interaction Phys. Rev. B 97 165153
[21] Ulstrup S, Schüler M, Bianchi M, Fromm F, Raidel C, Seyller T, Wehling T and Hofmann P 2016 Manifestation of nonlocal electron-electron interaction in graphene Phys. Rev. B 94 081403
[22] Ayral T, Biermann S, Werner P and Boehnke L 2017 Influence of fock exchange in combined many-body perturbation and dynamical mean field theory Phys. Rev. B 95 245130
[23] Sengupta A M and Georges A 1995 Non-fermi-liquid behavior near a T = 0 spin-glass transition Phys. Rev. B 52 10295–302
[24] Si Q and Smith J L 1996 Kosterlitz–Thouless transition and short range spatial correlations in an extended Hubbard model Phys. Rev. Lett. 77 3391–4
[25] Kajueter H 1996 Interpolating perturbation scheme for correlated electron systems PhD Thesis Rutgers University
[26] Smith J L and Si Q 2000 Spatial correlations in dynamical mean-field theory Phys. Rev. B 61 5184–93
[27] Chitra R and Kotliar G 2000 Effect of long range coulomb interactions on the mott transition Phys. Rev. Lett. 84 3678–81
[28] Chitra R and Kotliar G 2001 Effective-action approach to strongly correlated fermion systems Phys. Rev. B 63 115110
[29] Golež D, Boehnke L, Strand H U R, Eckstein M and Werner P 2017 Nonequilibrium GW + EDMFT: antiscreening and inverted populations from nonlocal correlations Phys. Rev. Lett. 118 246402
[30] Golež D, Boehnke L, Eckstein M and Werner P 2018 Dynamics of photo-doped charge transfer insulators Phys. Rev. B 100 041111
[31] Casula M, Werner P, Vaugier L, Aryasetiawan F, Miyake T, Millis A J and Biermann S 2012 Low-energy models for correlated materials: bandwidth renormalization from coulombic screening Phys. Rev. Lett. 109 126408
[32] van Loo E G C P, Schüler M, Katsnelson M I and Wehling T O 2016 Capturing nonlocal interaction effects in the Hubbard model: Optimal mappings and limits of applicability Phys. Rev. B 94 165141
[33] Pieper R 1938 On a minimum property of the free energy Phys. Rev. 54 918–9
[34] Bogoliubov N N 1958 Dokl. Akad. Nauk SSSR 119 244
[35] Feynman R P 1972 Statistical Mechanics (Reading, MA: Benjamin)
[36] van Loo E G C P, Krien F, Hafermann H, Stepnov E A, Lichtenstein A I and Katsnelson M I 2016 Double-occupancy in dynamical mean-field theory and the dual boson approach Phys. Rev. B 93 155162
[37] van Loo E G C P and Katsnelson M I 2018 The extended Hubbard model with attractive interactions J. Phys.: Conf. Ser. 1136 012006
[38] Vonsovsky S V and Katsnelson M I 1979 Some types of instabilities in the electron energy spectrum of the polar model of the crystal. I. The maximum-polarity state J. Phys. C: Solid State Phys. 12 2043
[39] Vonsovsky S V and Katsnelson M I 1979 Some types of instabilities in the electron energy spectrum of the polar model of the crystal. II. The criterion of stability of a metallic sta J. Phys. C: Solid State Phys. 12 2055
[40] Blankenbecler R, Scalapino D J and Sugar R L 1981 Monte Carlo calculations of coupled boson-fermion systems. I Phys. Rev. D 24 2278–86
[41] Savitzky A and Golay M J E 1964 Smoothing and differentiation of data by simplified least squares procedures Anal. Chem. 36 1627–39
[42] Schüler M, van Loo E G C P, Katsnelson M I and Wehling T O 2019 Thermodynamics of the metal-insulator transition in the extended Hubbard model SciPost Phys. 6 67
[43] Slater J C 1951 Magnetic effects and the Hartree–Fock equation Phys. Rev. 82 538–41
[44] Anderson P W 1959 New approach to the theory of superexchange interactions Phys. Rev. 115 2–13
[45] Moukouri S and Jarrell M 2001 Absence of a Slater transition in the two-dimensional Hubbard model Phys. Rev. Lett. 87 167010
[46] Kyung B, Landry J S, Poulin D and Tremblay A-M S 2003 Comment on ‘Absence of a Slater transition in the two-dimensional Hubbard model’ Phys. Rev. Lett. 90 099702
[47] Gull E, Werner P, Wang X, Troyer M and Millis A J 2008 Local order and the gapped phase of the Hubbard model: a plaquette dynamical mean-field investigation Europhys. Lett. 84 37009
[48] Taranto C, Sangiovanni G, Held K, Capone M, Georges A and Toschi A 2012 Signature of antiferromagnetic long-range order in the optical spectrum of strongly correlated electron systems Phys. Rev. B 85 085124
[49] Fratino L, Séémon P, Charlebois M, Sordi G and Tremblay A-M S 2017 Signatures of the Mott transition in the antiferromagnetic state of the two-dimensional Hubbard model Phys. Rev. B 95 235109

[50] van Loon E G C P, Hafermann H and Katsnelson M I 2018 Precursors of the insulating state in the square-lattice Hubbard model Phys. Rev. B 97 085125

[51] Pariser R and Parr R G 1953 A semi-empirical theory of the electronic spectra and electronic structure of complex unsaturated molecules. J. Chem. Phys. 21 466–71

[52] Pople J A 1953 Electron interaction in unsaturated hydrocarbons Trans. Faraday Soc. 49 1375–85

[53] Bursill R J, Castleton C and Barford W 1998 Optimal parametrisation of the pariser-parr-pople model for benzene and biphenyl Chem. Phys. Lett. 294 305–13

[54] Changlani H J, Zheng H and Wagner L K 2015 Density-matrix based determination of low-energy model hamiltonians from ab initio wavefunctions J. Chem. Phys. 143 102814

[55] Ayral T, Vaučičević J and Parcollet O 2017 Fierz convergence criterion: a controlled approach to strongly interacting systems with small embedded clusters Phys. Rev. Lett. 119 166401

[56] Valli A, Sangiovanni G, Gunnarsson O, Toschi A and Held K 2010 Dynamical vertex approximation for nanoscopic systems Phys. Rev. Lett. 104 246402

[57] Valli A, Schäfer T, Thunström F, Rohringer G, Andergassen S, Sangiovanni G, Held K and Toschi A 2015 Dynamical vertex approximation in its parquet implementation: application to Hubbard nanorings Phys. Rev. B 91 115115

[58] Schäfer M, Barthel S, Wehling T, Karolak M, Valli A and Sangiovanni G 2017 Realistic theory of electronic correlations in nanoscopic systems Eur. Phys. J. Spec. Top. 226 2615–40

[59] Cocchi E, Miller L A, Drewes J H, Koschorreck M, Pertot D, Brennecke F and Kölh M 2016 Equation of state of the two-dimensional Hubbard model Phys. Rev. Lett. 116 175301

[60] Cocchi E, Miller L A, Drewes J H, Chan C F, Pertot D, Brennecke F and Kölh M 2017 Measuring entropy and short-range correlations in the two-dimensional Hubbard model Phys. Rev. X 7 031025

[61] Baier S, Petter D, Becher J H, Patscheider A, Natale G, Chomaz L, Mark M J and Ferlaino F 2018 Realization of a strongly interacting fermi gas of dipolar atoms Phys. Rev. Lett. 121 093602

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

[63] Cao Y et al 2018 Correlated insulator behaviour at half-filling in magic-angle graphene superlattices Nature 556 80

[64] Terletska H, Chen T and Gull E 2017 Charge ordering and correlation effects in the extended Hubbard model Phys. Rev. B 95 115149

[65] Terletska H, Chen T, Paki J and Gull E 2018 Charge ordering and nonlocal correlations in the doped extended Hubbard model Phys. Rev. B 97 115117

[66] van Loon E G C P, Lichtenstein A I, Katsnelson M I, Parcollet O and Hafermann H 2014 Beyond extended dynamical mean-field theory: dual boson approach to the two-dimensional extended Hubbard model Phys. Rev. B 90 235135