The infinite occupation number basis of bosons - solving a numerical challenge

Andreas Geißler and Walter Hofstetter

1 Institut für Theoretische Physik, Goethe-Universität, 60438 Frankfurt/Main, Germany
(Dated: April 28, 2017)

In any bosonic lattice system, which is not dominated by local interactions and thus ”frozen” in a Mott-type state, numerical methods have to cope with the infinite size of the corresponding Hilbert space even for finite lattice sizes. While it is common practice to restrict the local occupation number basis to \( N_c \) lowest occupied states, the presence of a finite condensate fraction requires the complete number basis for an exact representation of the many-body ground state. In this work we present a novel truncation scheme to account for contributions from higher number states. By simply adding a single coherent-tail state to this common truncation, we demonstrate increased numerical accuracy and the possible increase in numerical efficiency of this method for the Gutzwiller variational wave function and within dynamical mean-field theory.

PACS numbers: 05.30.Jp, 67.85.-d, 42.50.Pq, 21.60.Fw
Keywords: Gutzwiller, DMFT, bosonic, basis truncation

Applying any diagonalization-based method to bosonic lattice systems, which are not entirely in a Mott-type phase, often requires the use of a truncation scheme for the local Hilbert space. This is most evident for methods using variational wave functions, as for example the Gutzwiller state \( |\psi\rangle \) \( = \prod_i |\psi_i\rangle \), because any numerical implementation requires a finite number of variational constants, which is realized by the choice of a truncation scheme. Related examples are the density matrix renormalization group (DMRG) \[6–8\] and derived methods such as matrix product states (MPS) \[9–12\], projected entangled pair states (PEPS) \[10, 12, 13\], as well as time-evolving block decimation (TEBD) \[14–17\], which all require a truncation of the local occupation number basis to the \( N_c \) lowest number states. The same is correspondingly true for bosonic DOF. The same relation is true for the variational GS, for which \( i \) represents all sites and DOF under consideration.

Applying any diagonalization-based method to bosonic lattice systems, which are not entirely in a Mott-type phase, often requires the use of a truncation scheme for the local Hilbert space. This is most evident for methods using variational wave functions, as for example the Gutzwiller state \( |\psi\rangle \) \( = \prod_i |\psi_i\rangle \), because any numerical implementation requires a finite number of variational constants, which is realized by the choice of a truncation scheme. Related examples are the density matrix renormalization group (DMRG) \[6–8\] and derived methods such as matrix product states (MPS) \[9–12\], projected entangled pair states (PEPS) \[10, 12, 13\], as well as time-evolving block decimation (TEBD) \[14–17\], which all require a truncation of the local occupation number basis to the \( N_c \) lowest number states. The same is correspondingly true for bosonic DOF. The same relation is true for the variational GS, for which \( i \) represents all sites and DOF under consideration.

As we will show for the cases of DMFT and GS, the use of a single additional variational basis state, which we denote as coherent-tail state (CTS), can strongly increase the accuracy as compared to the common truncation scheme. Especially for DMFT the CTS is highly efficient: even strongly reduced Hilbert spaces suffice to well approximate the (quasi-)exact DMFT results, obtained by using a Hilbert space more than three times as large. Due to this reduction in computational complexity, this scheme is accompanied by a more than tenfold increase in numerical efficiency.

System — In any numerical second quantized method, utilizing the grand canonical ensemble of an interacting Bose gas on a lattice, at some point it becomes necessary to approximate the infinite local Fock basis, to allow for results within a finite algorithm. As a test case, let us consider the basic Bose-Hubbard model \[30, 32\].

\[
H = -J \sum_{\langle i,j \rangle} (\hat{b}_i^\dagger \hat{b}_j + \hat{b}_j^\dagger \hat{b}_i) + \frac{U}{2} \sum_i \hat{n}_i(\hat{n}_i - 1) - \mu \sum_i \hat{n}_i
\]  
(1)

We use the common notation, where \( \hat{b}_i \) (\( \hat{b}_i^\dagger \)) is the annihilator (creator) of a boson at site \( i \), while \( \hat{n}_i \) is the corresponding particle number operator \( \hat{n}_i = \hat{b}_i^\dagger \hat{b}_i \). The parameters are the hopping amplitudes \( J \) \[33\], the local Hubbard interaction \( U \) \[33\] – tunable by Feshbach resonances \[34, 35\] – and a chemical potential \( \mu \), determining the total particle number.

Numerous techniques have been applied to investigate this model, ranging from the Gross-Pitaevskii equation (GPE) \[37, 38\], Bogoliubov theory \[39, 41\] and variational mean-field methods such as GS \[42, 43\] to more advanced techniques including Monte Carlo methods (MC).
has the following form

\[ H_G = -J \sum_{\langle i,j \rangle} (\hat{b}_i^\dagger \phi_j + \text{h.c.}) + \frac{U}{2} \sum_i \hat{n}_i (\hat{n}_i - 1) - \mu \sum_i \hat{n}_i \]  

(3)

where \( \phi_i = \langle \hat{b}_i \rangle \). It is thus a set of local many-body problems coupled by the self-consistent fields \( \phi_i \) (commonly called condensate order parameter). The ground state energy of this simplified Hamiltonian is found by variation of these fields. In a homogeneous system, where every site has \( z \) nearest neighbours, and in the absence of spontaneous symmetry breaking, the problem reduces to a single variable \( \phi \), thus further simplifying (3):

\[ H_G^{\text{local}} = -Jz(\hat{b}^\dagger \phi + \phi^* \hat{b}) + \frac{U}{2} \hat{n}(\hat{n} - 1) - \mu \hat{n} \]  

(4)

This problem can be solved in an arbitrary local basis, but any numerical implementation requires a truncation, for example to the common finite \( N_c \)-Fock basis. In order to compare with numerical calculations in the CTS-extended basis, we furthermore need the following properties of the CTS

\[ \hat{b} |\alpha_{N_c}\rangle = c_{N_c} \frac{\alpha^{N_c}}{(N_c - 1)!} |N_c - 1\rangle + \alpha |\alpha_{N_c}\rangle, \]  

(5)

\[ \hat{b}^\dagger |\alpha_{N_c}\rangle = c_{N_c} \frac{\alpha^{N_c}}{(N_c - 2)!} |N_c - 2\rangle \]  

(6)

\[ + \alpha \left( c_{N_c} \frac{\alpha^{N_c}}{(N_c - 1)!} |N_c - 1\rangle + \alpha |\alpha_{N_c}\rangle \right) \]

which are necessary to calculate all the additional matrix elements of \( H_G^{\text{local}} \). Note that the CTS acts as the Fock state \( |N_c\rangle \) for \( \alpha \to 0 \). Now one only needs to find the minimum of \( E_{\text{tot}}^G = \langle \psi_\text{G} | H_G^{\text{local}} | \psi_\text{G} \rangle \), by simultaneous variation of both the physical parameter \( \phi \) and the non-physical CTS-parameter \( \alpha_{N_c} \). Since the final result has to be independent of the truncation scheme, a comparison for various \( N_c \) and \( \alpha_{N_c} \) at given values of \( \mu / U = J / U = 0.4 \), reveals the limited efficiency of the CTS (see Fig. 1). Thus we can now tell how a CTS-extended basis with reduced cutoff compares to a large \( N_c \)-Fock basis.

At any truncation level, if the CTS is added to the \( N_c \)-Fock basis, \( E_{\text{tot}}^G \) is improved in comparison to a single additional Fock state, corresponding to \( \alpha_{N_c} \to 0 \) in Fig. 1(a) (also Fig. 1(c,d), Fig. 1(e)) and (Fig. 1(b,d)). One even improves upon the mean-field Mott transition, for Mott lobes \( n = \langle \hat{n} \rangle = N_c \), at the limit of the Fock-basis with cutoff \( N_c + 1 \) (see Fig. 1(h)). But due to the necessary optimization of \( \alpha_{N_c} \), this comes at an additional computational cost (see Fig. 1(b)). The GS thus does not benefit much from the CTS, as far as computational effort is considered. On the other hand, as we will show, within BDMFT the CTS truncation scheme leads to a significant speed-up paired with the increased accuracy.
In the case of a homogeneous lattice gas, used is required due to the mapping onto the effective model. They are therefore treated as hard-core bosons. The additional terms including the annihilation (creation) operators \( \hat{a}_i \) (\( \hat{a}_i^\dagger \)) describe effective bath orbitals which self-consistently mimic the action of the lattice sites surrounding the given site \( j = 0 \) in the Hubbard model \( (1) \). They do so via the orbital energies \( \epsilon_i \), normal hoppings \( V_i \) and anomalous hoppings \( W_i \). For an optimal representation of this action, increasing the number of bath orbitals is favorable over increasing bath truncations. They are therefore treated as hard-core bosons. The cavity expectation value \( \langle \hat{c}_j \rangle \) is computed in a system where the impurity site has been removed, which is required due to the mapping onto the effective model \( (1) \). In the case of a homogeneous lattice gas, used here for benchmarking purposes, easily allowing for comparisons with truncations as high as \( N_c = 20 \), the term containing the self-consistent cavity order parameter simplifies to \( \sum_i \langle \hat{b}_i \rangle_C = z \cdot \phi_C \), where \( z \) is the number of nearest neighbours, and \( \phi_C \) is the cavity expectation value of the condensate order parameter.

Within this implementation, a choice of \( \alpha_{N_c} \) to compute the ground state of the full system, is efficiently obtained by minimizing the energy \( E_{\text{AIM}} = 0 (H_{\text{AIM}} | 0) \) for the lowest energy eigenstate of the self-consistent \( H_{\text{AIM}}^\text{eff} \), with regard to the variational parameter \( \alpha_{N_c} \). This yields the optimal representation for the low energy spectrum of \( H_{\text{AIM}}^\text{eff} \), which determines the interacting Green’s function in the Lehmann representation \( (20) \). Another way of optimization would be the minimization of the self-consistently converged DMFT expectation value \( E_{\text{tot}} (\alpha_{N_c}) = \langle H \rangle \) in relation to \( \alpha_{N_c} \), where we define \( E_{\text{tot}}^\text{DMFT} \) as this minimum. Let us emphasize, that the optimal \( E_{\text{tot}}(\alpha_{N_c}) \) should not depend on the choice of basis, so neither a variation in \( N_c \) nor in \( \alpha_{N_c} \) should result in a significant change of its self-consistent value, as indeed verified exemplary for \( \mu/U = J/U = 0.4 \) in Fig. 2(a). Then the optimal CTS state allows for a remarkably good approximation of the total BDMFT energy, even at a very low Fock-space truncation \( N_c \) (see Figs. 2(a) and 3(c)).

A further look at the convergence times reveals the numerical benefit of replacing a large number of Fock-states (all those with \( n \geq N_c \)) by the single variational state \( |\alpha_{N_c}\rangle \). We have simulated the Bose-Hubbard-model \( (1) \) within BDMFT using a Bethe lattice with \( z = 6 \) for \( 0 \leq \mu/U \leq 3.5 \) and \( 0 < J/U \leq 1 \). The convergence times for various truncation schemes are shown in Fig. 2(b), for \( \mu/U = 0.4 \). Note the above 10-fold decrease in convergence times when using the CTS-extended Fock basis, compared to the regular Fock basis with a high \( N_c \), used for the (quasi)-exact solution. While this speed-up is only possible over the full range of parameters, when optimizing \( E_{\text{AIM}} \), this simplified scheme leads to negligible deviations in the energy (as shown in Figs. 2(c,d)). Also note the additional time loss of the CTS scheme, compared to a truncation of equal basis size at large \( J/U \), which is due to the need to optimize \( \alpha_{N_c} \), by finding either the minimum of \( E_{\text{AIM}} \) or of \( E_{\text{tot}} \), while the latter also requires multiple runs of fully converged DMFT simulations. So due to the minor differences, we now focus on results obtained via the first scheme.

Regarding physical observables, we have calculated local observables, such as the condensate order parameter \( \phi = \langle \hat{b} \rangle \) and the occupation number \( n = \langle \hat{n} \rangle \), as well as the non-local non-condensate fluctuations \( G_n(t = 0) = \langle \langle \hat{b}_i^\dagger \hat{b}_j \rangle - \langle \hat{b}_i^\dagger \rangle \langle \hat{b}_j \rangle \rangle \), where \( i \) and \( j \) are nearest neighbours. This expression is more commonly denoted as the connected Green’s function at equal times, which we can directly extract within BDMFT. Furthermore we also ob-
tain the total energy $E_{\text{tot}}$ and the kinetic energy $E_{\text{kin}}^{\text{con}}$ due to the connected part of the Green’s function, allowing for a comparison of the quality of different truncation schemes:

$$E_{\text{kin}}^{\text{con}} = -Jz \left( \hat{b}_i^\dagger \hat{b}_j - \phi^* \phi \right) = Jz G_c(t = 0) \quad (8)$$

As is visible from the local observables as well as the total energy, replacing the highest Fock-state $|N_c\rangle$ by the CTS $|\alpha_{N_c}\rangle$ tremendously improves the results to almost the same accuracy as the (quasi-)exact result from the increased cutoff $N_c = 20$ (see Fig. 3 [a-c]). Remarkably, the CTS truncation even predicts the Mott transition for the Mott lobe $n = 4$ (for $\mu/U = 3.5$) almost exactly, as shown Fig. 3 [a], while the regular cutoff $N_c = 5$ fails to do so and both truncations also yield wrong $G_c(t = 0)$ (see Fig. 3 [d]). The high accuracy of local observables is lost, just about where the occupation number $n$ exceeds $N_c$. Differences between the three cases can be seen most clearly in the non-condensed contribution to the kinetic energy [8], due to non-local fluctuations described by the connected Green’s function (see Fig. 3 [d]). These have a monotonously decreasing tail for $J/U \to \infty$ in the exact solution. Obviously the ratio of these fluctuations to the condensate fluctuations of the BEC $\propto G_c(t = 0)/\|\phi\|^2$ vanishes in this limit.

But a hard and low truncation results in an artificially increased value of the non-local non-condensate fluctuations beyond certain values of $J/U$, while the CTS leads to the opposite behaviour, where the tail is damped more than in the exact result, thus suppressing non-condensate fluctuations early. This is likely a result of the CTS being more heavy-tailed than the Hubbard interaction would allow [17]. As non-condensate fluctuations only give a sub-leading contribution to the kinetic energy, it becomes clear why the CTS allows for the tremendous increase in accuracy and speed-up in numerical simulations compared to a simple high Fock space cutoff $N_c$, even for large $J/U$.

In conclusion, we have introduced a novel truncation scheme based on the CTS [2], for which we demonstrated an increase in the numerical accuracy and computational efficiency of GS and BDMFT simulations. This increase was shown to be especially pronounced in BDMFT. Therefore the method allows for BDMFT simulations at much larger densities than before, but with reasonable computational effort. It is thus also a very promising method for accurate simulations of systems at higher filling per site. Furthermore cluster-based methods [18, 19] should especially benefit from this softened truncation, since the size of their Fock basis scales as $N_c^L$ with cluster size $L$. The concept of softening the hard cutoff, usually applied in the number basis, should thus more generally benefit a wide range of numerical simulations of bosonic lattice systems.

We would like to thank U. Bissbort and I. Vasić.
for insightful discussions. Support by the Deutsche Forschungsgemeinschaft via DFG SPP 1929 GiRyd and the high-performance computing center LOEWE-CSC is gratefully acknowledged.

geissler@th.physik.uni-frankfurt.de

[1] M. C. Gutzwiller, Phys. Rev. Lett. 10, 159 (1963)
[2] M. C. Gutzwiller, Phys. Rev. 134, A923 (1964)
[3] M. C. Gutzwiller, Phys. Rev. 137, A1728 (1965)
[4] D. S. Rokhsar and B. G. Kotliar, Phys. Rev. B 44, 10328 (1991)
[5] W. Krauth, M. Caffarel, and J.-P. Bouchaud, Phys. Rev. B 45, 3137 (1992)
[6] S. R. White, Phys. Rev. Lett. 69, 2863 (1992)
[7] U. Schollwöck, Rev. Mod. Phys. 77, 259 (2005)
[8] S. R. White, Phys. Rev. Lett. 69, 2863 (1992)
[9] U. Schollwöck, Rev. Mod. Phys. 77, 259 (2005)
[10] W. Krauth, M. Caffarel, and J.-P. Bouchaud, Phys. Rev. B 45, 3137 (1992)
[11] S. R. White, Phys. Rev. Lett. 69, 2863 (1992)
[12] M. C. Gutzwiller, Phys. Rev. 134, A923 (1964)
[13] M. C. Gutzwiller, Phys. Rev. 137, A1728 (1965)
[14] D. S. Rokhsar and B. G. Kotliar, Phys. Rev. B 44, 10328 (1991)
[15] W. Krauth, M. Caffarel, and J.-P. Bouchaud, Phys. Rev. B 45, 3137 (1992)
[16] S. R. White, Phys. Rev. Lett. 69, 2863 (1992)
[17] M. C. Gutzwiller, Phys. Rev. Lett. 10, 159 (1963)
[18] M. C. Gutzwiller, Phys. Rev. 134, A923 (1964)
[19] M. C. Gutzwiller, Phys. Rev. 137, A1728 (1965)
[20] D. S. Rokhsar and B. G. Kotliar, Phys. Rev. B 44, 10328 (1991)
[21] W. Krauth, M. Caffarel, and J.-P. Bouchaud, Phys. Rev. B 45, 3137 (1992)
[22] M. C. Gutzwiller, Phys. Rev. 134, A923 (1964)
[23] M. C. Gutzwiller, Phys. Rev. 137, A1728 (1965)
[24] D. S. Rokhsar and B. G. Kotliar, Phys. Rev. B 44, 10328 (1991)
[25] W. Krauth, M. Caffarel, and J.-P. Bouchaud, Phys. Rev. B 45, 3137 (1992)