State- and superstate-sampling in hybridization-expansion continuous-time quantum Monte Carlo

Alexander Kowalski,1 Andreas Hausoel,1 Markus Wallerberger,2 Patrik Gunacker,3 and Giorgio Sangiovanni1

1Institut für Theoretische Physik und Astrophysik, Universität Würzburg, Am Hubland, D-97074 Würzburg, Germany
2Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA
3Institute of Solid State Physics, TU Wien, 1040 Vienna, Austria
(Dated: July 31, 2018)

Due to the intrinsic complexity of the quantum many-body problem, quantum Monte Carlo algorithms and their corresponding Monte Carlo configurations can be defined in various ways. Configurations corresponding to few Feynman diagrams often lead to severe sign problems. On the other hand, computing the configuration weight becomes numerically expensive in the opposite limit in which many diagrams are grouped together. Here we show that for continuous-time quantum Monte Carlo in the hybridization expansion the efficiency can be substantially improved by dividing the local impurity trace into fragments, which are then sampled individually. For this technique, which also turns out to preserve the fermionic sign, a modified update strategy is introduced in order to ensure ergodicity. Our new (super)state-sampling is particularly beneficial to calculations with many d-orbitals and general local interactions, such as full spherically symmetric Coulomb interaction. For illustration, we reconsider the simple albeit well-known case of a degenerate three-orbital model at low temperatures. This allows us to quantify the coherence properties of the “spin-freezing” crossover, even close to the Mott transition.

I. INTRODUCTION

Continuous-time quantum Monte Carlo algorithms are state-of-the-art, numerically exact methods for the solution of the Anderson impurity model (AIM)14–16. These are widely used for the description of the physics of magnetic impurities, Kondo systems, transport through quantum junctions and are also employed as auxiliary models in Dynamical Mean Field Theory (DMFT) calculations for lattice models of correlated electron systems. Several high-level open source implementations of DMFT and of its merger with Density Functional Theory have been recently made available.5–9

One of the most successful flavors of continuous-time quantum Monte Carlo algorithms is the strong-coupling hybridization expansion (CT-HYB).10 CT-HYB is the method of choice for multi-orbital impurity models with general interactions because one observes only a moderate sign problem provided that the bath problem has sufficient symmetry. This is because CT-HYB splits each Monte Carlo configuration into a non-interacting bath part and a fully interacting impurity part, and solves the impurity part using an exact diagonalization/FullCI-type method. It thus eliminates any sign problem that would arise from the stochastic sampling of the impurity part at the cost of solving the problem explicitly. However, the dimension of the impurity Hamiltonian grows exponentially with the number of orbitals, and so does the computational effort with it. In practice, correlated d- or f-shells as well as small correlated molecules can be treated with CT-HYB.

Yet, reaching low temperatures is still challenging due to the quadratic scaling of the impurity problem with inverse temperature. This follows from the fact that the mean order of diagrammatic expansion grows linearly with inverse temperature, and both the computational cost of evaluating a single configuration as well as the observed autocorrelation time between configurations scale linearly with expansion order. While the exponential scaling with the number of orbitals and the quadratic scaling with the inverse temperature is intrinsic to the local problem, potentially model-dependent improvements to the prefactor of this overall scaling can be achieved.

Common approaches to such optimization are block-diagonalization of the local Hamiltonian using conserved quantities17 and binning, tre15 or equivalent12 algorithms in matrix-matrix implementations of CT-HYB. Additionally, with a similar motivation as for our method, outer truncation of the local trace to the few dominant contributions and calculation of those with more efficient sparse-matrix methods has been applied particularly to large systems at low temperature.13

Other more advanced strategies are local updates in imaginary time13, a fast-rejection/acceptance algorithm by calculating upper/lower boundaries of the weight11, or a partial summation of diagrams to extract more information out of one Monte Carlo configuration.15

Here we consider a matrix-vector version of the CT-HYB algorithm as implemented in the w2dynamics package9 and investigate the possibility of sampling the sum over the eigenstates of the local impurity in the Monte Carlo simulation. A hard outer truncation of high energy states has been attempted, but this is an approximation and it is unclear whether such a procedure retains ergodicity. Our approach, however, exceeds the numerical benefits of hard truncations substantially and is furthermore numerically exact.

We formulate two versions: the “superstate”-sampling
algorithm, where states grouped by the blocks of the Hamiltonian are sampled together, and the “state”-sampling algorithm, where each many-body state of the impurity is sampled individually. Conceptually, the latter can be interpreted as an equivalent of the segment-algorithm for general interactions. Our method touches the core of the exponential scaling of CT-HYB and manages to significantly reduce the computation time of a Monte Carlo weight. Furthermore, it is in principle compatible with all of the other above-mentioned algorithmic improvements. Using a five-orbital AIM with the most general form of the electron-electron Coulomb interaction as an example, we achieve speed-up factors verging on three orders of magnitude.

First, we review the basic formulas of CT-HYB in section II. In sections III and IV the superstate- and state-sampling methods are introduced. We distinguish between “General description” sections outlining the ideas and some “Implementation details” we include. Note that the latter mainly discuss technical aspects of the Monte Carlo moves and can hence be skipped if the reader wants to get an understanding of the basic ideas only. In section V we comment on the performance and the average sign, while in VI we demonstrate the capabilities of the new algorithm with a simple physical example.

II. HYBRIDIZATION EXPANSION

The hybridization expansion of the partition function of the Anderson impurity model is given by

\[ Z = \sum_C w_{\text{loc}}(C) w_{\text{bath}}(C), \]  

where

\[ w_{\text{bath}}(C) = \det F \]

is the bath part and \( F \) denotes the bath propagator. \( w_{\text{bath}}(C) \) hence represents a determinant of non-interacting hybridization functions that describe the propagation of electrons through the bath between leaving the impurity at some point and returning later. The quantity

\[ w_{\text{loc}}(C) = \sum_s \langle s | \hat{C} | s \rangle \]

\[ = \text{Tr} \left[ T_e e^{-\beta H_{\text{loc}}} \prod_{i=1}^k d_{\lambda_i}^\dagger(\tau_i) d_{\lambda_i}(\tau_i') \right] \]

is the local weight or local trace of a configuration, whose calculation will be discussed in detail in the following. \( \langle s | \rangle \) are the eigenstates of the local impurity Hamiltonian \( H_{\text{loc}} \), and \( \hat{C} = T_e e^{-\beta H_{\text{loc}}} \prod_{i=1}^k d_{\lambda_i}^\dagger(\tau_i) d_{\lambda_i}(\tau_i') \) is an abbreviation for the sequence of impurity operators. In the local trace there are \( 2k \) impurity operators of flavors \( \lambda_i \) at imaginary times \( \tau_i \), time ordered by \( T_e \).

\[ \sum_c = \int_0^\infty \prod_{i=1}^k \int_{\tau_{i-1}}^{\beta} \text{d}\tau_i \int_{\tau_{i-1}}^{\beta} \text{d}\tau_i' \sum_{\lambda_i} \sum_{\lambda_i'} \]

\[ \sum_{s \in S} \langle s | \hat{C} | s \rangle \]

denotes a sum over all internal degrees of freedom. In the conventional continuous-time hybridization expansion quantum Monte Carlo (CT-HYB) algorithm,

\[ Z = \sum_c w_{\text{loc}}(C) w_{\text{bath}}(C) \]

is sampled. Reducing the computational impact of the calculation of \( w_{\text{loc}}(C) \) is our main objective here.

III. SUPERSTATE-SAMPLING

A. General description

When non-density-density or hopping terms are present in \( H_{\text{loc}} \), the time-evolution mixes many-body states. For example, in a system with two orbitals and Kanamori interaction\(^{[16]}\), the time-evolution applied to a state \( |\uparrow, \downarrow\rangle \) yields a state \( |\uparrow, \downarrow\rangle + \beta |\downarrow, \uparrow\rangle \).

The Hamiltonian can be block-diagonalized with blocks chosen such that time-evolution does not mix states from different blocks and the application of an impurity operator does take a state from one block to more than one other block. Thus each of these blocks (called “superstates”\(^{[17]}\) by K. Haule) is characterized by a set of conserved quantum numbers. In the original algorithm, the trace is calculated straightforwardly by summing over all \( 4^{N_{\text{orb}}} \) impurity eigenstates. While this exponential growth limits the maximal number of orbitals, a local Hamiltonian \( H_{\text{loc}} \) mixing many states additionally further increases the average block size.

The main idea of this work is to transform the deterministic summation over the eigenstates of the impurity in Eq. (4) into a stochastic summation. Whilst having been proposed\(^{[18]}\), it has never been implemented to the best of our knowledge because a severe sign problem was expected.

The original Monte Carlo configuration is split into many “smaller” weights, where the summation over all superstates \( S \) is now done by Monte Carlo sampling:

\[ Z = \sum_c \sum_{S} w_{\text{loc},S}(C) w_{\text{bath}}(C), \]

with

\[ w_{\text{loc},S}(C) = \sum_{s \in S} \langle s | \hat{C} | s \rangle \]
and eigenstates $|s\rangle$ of the impurity.

This allows us to calculate one weight faster and move faster through phase space, but splitting a sum with (potentially) positive and negative terms into parts may lead to a disastrous sign problem.

In order to understand why such a sign problem may be anticipated, it is useful to reconsider the sampling of the bath weight: by combining all bath propagators into a bath determinant, a severe sign problem can be circumvented in CT-HYB\textsuperscript{18}. On the other hand, lifting the bath determinant structure and partially summing diagrams reintroduces a sign problem to the hybridization expansion algorithm\textsuperscript{19}.

As we discuss later, we do not observe any worsening of the average sign compared to the original algorithm when splitting the local weight in the way we propose. A heuristic argument to motivate the procedure can be summarized thusly: Since the mean expansion order grows linearly with the inverse temperature $\beta$\textsuperscript{18}, the average amount of superstates that have zero contribution due to the Pauli principle increases, until at a certain $\beta$, we are often left with only one outer superstate. For several metallic systems at temperatures of the order of $10^{-2}$ of the electronic bandwidth, there are many configurations with essentially one single superstate contributing to the trace. Furthermore, since in this limit of large $\beta$ the conventional sampling of the local trace and our superstate-sampling algorithm are very similar, a worsening of the sign is not to be expected.

FIG. 1. Average relative contributions of outer superstates to the local weight per configuration for a typical simulation with five orbitals and Kanamori interaction. For comparability, the superstates are ordered by their contribution (i.e. absolute value of the part of the local weight sum from all states contained in the superstate) for each individual configuration, i.e. superstate “1” does not denote one specific constant superstate, but always refers to the biggest contributor.

Since this argument does not apply to simulations at high temperatures, we have inspected the average local weight distribution of the outer superstates contributing the most per configuration in the conventional algorithm in some typical cases. The case of a five-orbital model with Kanamori interaction is shown in Fig. 1. It is clear how the local weight of each configuration is strongly dominated by the contribution from one superstate. Similar results can be obtained for a simpler two-orbital model. We can therefore expect the method to be useful for high temperatures as well if it samples configurations with their “ideal” outer superstates.

B. Implementation details: Sampling and ergodicity

When proposing a new configuration in conventional CT-HYB, it is useful to check which outer superstates violate the Pauli principle and hence do not contribute to the new local weight before calculating it explicitly. We call this procedure quantum number checking. Since each configuration in the new method contains an outer superstate, this needs to be done for the proposed outer superstate only.

To preserve ergodicity, the simulation must be able to reach all configurations with non-zero weight that differ only in the outer superstate. Simple techniques to ensure this are e.g. proposing an outer superstate randomly at each move, or a new move that only changes the outer superstate. However, such ways turn out to be inefficient. Therefore, we start from the fact that the outer superstate of a configuration uniquely determines the superstate containing an outer state evolved to any $\tau$. When making an insertion or a removal of a pair of operators at $\tau_1$ and $\tau_2$, there are two possible proposals for the outer superstate of the next configuration: one can (i) keep the old outer superstate, which changes the sequence of superstates between the two operators $\tau_1$ and $\tau_2$ (“inner pair move”), or (ii) propose the outer superstate that would only change the sequence outside of $\tau_1$ and $\tau_2$ (“outer pair move”).

Using (i) as the default local move and allowing proposals of the second kind provides a reasonable compromise. The acceptance of (ii) is however significantly smaller than that of the first kind in practice. In the next section, we show a more efficient way to ensure ergodicity with respect to the outer superstate, the so called global $\tau$-shift move.

Let us also shortly comment on how to choose the outer superstate for the initial configuration at the beginning of the simulation. While the initial choice should not influence the simulation after thermalization, for many highly excited outer superstates the local weight is close to zero. We thus select the initial outer superstates randomly with probabilities proportional to their local weights.
FIG. 2. Left: a superstate-sampling configuration for a two orbital model with Kanamori interaction. Bold horizontal lines denote a time-evolution of an eigenstate, the dotted vertical lines the operators that make transitions between superstates. Here all superstates have size 1, except the spin-flip and pair-hopping one.
Right: the configuration resulting from application of a global $\tau$-shift by $\Delta \tau$ to all operators of the one shown on the left. $\tau_f$ and $\tau_l$ denote the imaginary times of the first and last operator after the shift. In the superstate-sampling algorithm the local weight is the sum of the red and blue state as outer state of the trace, whereas in the state-sampling only one of the two is selected.

C. Implementation details: Global $\tau$-shift moves

A global “$\tau$-shift move” shifts the positions of all operators in imaginary time by a random number $\Delta \tau \in [0, \beta]$, and changes the outer superstate to the superstate that was at $\tau = \Delta \tau$ in the original configuration (see Fig. 2). Over the course of an entire simulation, the proposal probability for a specific outer superstate in this kind of move is proportional to the average relative amount of imaginary time it covers. Since the superstate sequence is cyclic and effectively also just shifted along the $\tau$-axis, there is no need to perform a quantum number check. A global move similar to our $\tau$-shift was introduced by Shinaoka for a different technical reason.

The proposal probabilities of a $\tau$-shift move and its reverse are equal. The acceptance probability of this move is 1. In appendix A we prove that the bath determinant remains unchanged, as the action of the $\tau$-shift on the hybridization matrix move effectively corresponds to a number of permutations and multiplications of rows and columns. Additionally in appendix B there is a proof that the local trace remains unchanged under a combined cyclic permutation of the operators and corresponding change of the outer superstate.

Given that we can not change the outer superstate, but only superstates of sections not including $\tau = 0 = \beta$ using inner pair moves, the sampling would not be ergodic if we only used those. Including $\tau$-shift moves, however, the section with the outer superstate (considering its position relative to the rest of the operator/superstate sequence) does not have to remain around $\tau = 0 = \beta$. A single $\tau$-shift move can shift the operator/superstate sequence relative to the $\tau$-axis in such a way that the section including the fixed outer superstate changes while the configuration remains equivalent otherwise. In combination with the ordinary inner pair moves it hence allows the superstate of any section to be changed without the need for outer pair moves. We can explicitly consider the example of one specific outer pair move, which can always be replaced by performing the corresponding inner pair move between two sufficiently big inverse $\tau$-shifts. In situations where the acceptance rate of outer pair moves is considerably smaller than that of inner pair moves, it is plausible that using the $\tau$-shift move is the more efficient alternative.

IV. STATE-SAMPLING

A. General description

The superstate-sampling method of the last section already significantly reduces the cost of computing a local weight. We wondered whether it may be possible to split the local weight further in even smaller parts, where the summation over all states $s$ within a superstate $S$ is also done as a Monte Carlo sum:

$$Z = \sum_{C} \sum_{S} \sum_{s \in S} w_{\text{loc},s}(C) w_{\text{bath}}(C),$$

with

$$w_{\text{loc},s}(C) = \langle s | \hat{C} | s \rangle.$$  

In the superstate-sampling method the average sign is not worse than in the conventional CT-HYB sampling, even at small $\beta$ when the number of superstates not violating the Pauli principle is large. As we have seen in Fig. 1, the value of the local weight is nevertheless dominated by the contribution of one superstate, and the other superstates just make exponentially small corrections. To avoid confusion, let us stress at this stage
that our method does not make any assumption on the contribution of the superstates to the local weight. It is an exact sampling with no approximation involved. Strongly peaked distributions like the ones we observe in most cases simply imply that this approach is efficient, whereas it might only perform comparably to the conventional one if many situations were characterized by much flatter state weight distributions.

For the states within one superstate we show a similar trend in Fig. 3: the contribution of one superstate is dominated by the contribution of one eigenstate. This suggests trying to apply the principle of superstate-sampling one level deeper in the form of an “(eigen)state-sampling”.

FIG. 3. Average relative contribution of outer states within one superstate, ordered by the size of contribution per configuration, normalized to the total weight of that superstate. This graph was obtained from a typical simulation of a 5-orbital system with Kanamori interaction.

B. Compatibility with implementations

Let us also discuss the applicability of state-sampling to implementations of CT-HYB. When reviewing standard CT-HYB, we simply expressed the weight of a configuration as the trace of the product of the corresponding impurity operators. Typically, there are two ways to calculate it: in a matrix-matrix implementation, the matrices representing the impurity operators are multiplied with one another and the trace is obtained from the diagonal elements of the total matrix product. In the matrix-vector flavor, we instead explicitly perform a sum over the outer states: for each outer ket-vector, we repeatedly calculate the matrix-vector product with the operators, starting from the first all the way through to the outer bra-vector, with which we finally compute the scalar product.

By decomposing the problem into superstates (i.e., block-diagonalizing the Hamiltonian and choosing the blocks such that the operators connect them in a one-to-one way), we simplify the calculation of the trace in that the outer states (per superstate) follow independent, non-crossing paths through the superstates. As a result, in the matrix-vector implementation one can just reduce the size of the initial operator matrices. In a matrix-matrix implementation the trace of the product can be decomposed into the sum of traces per outer superstate, which introduces the explicit summation whose terms can be calculated using matrices of reduced size as well.

If we consider such an implementation as the starting point, we just have to restrict the summation to one outer superstate in either case to implement superstate-sampling. While we can obviously reduce the number of needed calculations by restricting the outer sum to one state only in the matrix-vector implementation, i.e., implementing state-sampling, there is no way to beneficially implement this in a matrix-matrix algorithm. We could use only one of the diagonal elements of the resulting products, but this does not make the product calculation simpler and therefore does not improve performance, but would only waste the other contributions that could be included at negligible further cost.

C. Implementation details: choice of the outer state within a superstate

The moves used in our implementation of state-sampling are best summarized in terms of modifications to the moves of our superstate-sampling implementation, and in fact since an outer state also implies an outer superstate (the one it is contained by), internally we realize state-sampling as superstate-sampling where the outer sum is additionally restricted to just one of the contained states. In the inner pair moves, we now also keep the outer state fixed in addition to the outer superstate, and in the moves that changed the outer superstate, we now also change the outer state. If the outer superstate chosen by a superstate-sampling move contains more than one state (see Fig. 2), we now need to pick one of them as the new outer state, which we chose to do randomly as we could not see a way in which a favorable one might be implied.

Proposing one of the possible states with uniform probability, however, causes a lower acceptance rate compared to the analogous moves in superstate-sampling, where e.g., the \( \tau \)-shift move even has acceptance rate 1.

Since close to \( \tau = 0 \) there is always just the outer state \( s \) propagating with its eigenenergy, we can make the procedure more efficient by “transferring” the time-evolution of this state from the acceptance to the proposal probability; i.e., we include it in the proposal probability so that in the standard Metropolis acceptance probability formula, it cancels with the equivalent factor in the weight of the configuration. This proposal probability weighting allows us to incorporate our prior (or easier to calculate) knowledge to avoid wasting time on proposals that would likely be rejected: for a well chosen proposal probabil-
ity, i.e. one close to the actual weight distribution, we raise the acceptance probability for all outer states since those outer states that would be rejected more often with uniform proposal probability are simply proposed less often. To include the aforementioned time-evolution close to \( \tau = 0 \), we use the proposal probability

\[
p_{\text{prop}}(s) = \frac{\exp(-\tau_f - \beta - \tau_l)}{\sum_{k \in T} \exp(-E_k - E_0 - \beta)} \cdot \frac{\exp(-E_s - E_0 - \beta)}{\sum_{k \in T} \exp(-E_k - E_0 - \beta)} ,
\]

where \( T \) contains all states from all outer superstates that may be proposed and \( \tau_f \) and \( \tau_l \) are, respectively, the imaginary times of the first and last operator after the move. In this way, the acceptance rate for outer state changes is significantly increased and e.g. reaches about 50\% for the global \( \tau \)-shift move in typical calculations. The best choice of proposal probability for such an optimization depends on the weight we expect: in this case, we essentially assume that the potentially excited state at \( \tau = 0 \) will be brought closer to the ground state by the impurity operators, since if we expected the energy to stay at the level of the outer state, we could choose a better proposal probability assuming propagation over the entire imaginary time (i.e. replacing \( \tau_f + \beta - \tau_l \) by \( \beta \)).

V. SPEED-UP QUANTIFICATION

To check the correctness of the results, we use a two-orbital model with a Kanamori interaction, i.e. density-density, pair-hopping and spin-flip terms, and a finite number of bath sites for which we have a reference solution obtained using exact diagonalization. Both the reference self-energy as well as one calculated using our CT-HYB solver are shown in Fig. 11 (appendix C).

In order to quantify the performance gain, we use a five-orbital Hamiltonian modeling a realistic transition-metal impurity on the surface of a metal with both the full (spherically symmetric) Coulomb tensor as well as one derived from the same interaction matrix restricted to Kanamori-like terms only. To quantify the performance improvement we compare the sampling rates, i.e. the raw amount of generated (correlated) samples per time. For the autocorrelation time we found a minor increase of about 10\% for superstate-sampling, but only about 3\% for state-sampling as compared to the old sampling method. This can be considered negligible in comparison to the speed-up factors. The mean sign is about 1.0 for the model with Kanamori interaction using both conventional and superstate-sampling and about 0.98 using state-sampling. For the model with full Coulomb interaction the sign is significantly less than 1 in all cases, and it is slightly but significantly smaller with state-sampling than with the other methods. For the calculations used to measure the speed-up factors, the absolute values of the mean sign for all three methods can be found in Fig. 5.

FIG. 4. Comparison of Monte Carlo sampling speed, i.e. number of individual, mostly local, updates per CPU time, not including measurements, for a 5-orbital system with Kanamori or Coulomb interaction and cubic interpolation curves plotted logarithmically. Top panel: speed-up factors of the new sampling methods compared to conventional sampling for Kanamori and Coulomb interaction. Bottom panel: speed of all the sampling methods for Coulomb interaction.

FIG. 5. Absolute value of the mean sign using our new sampling methods for a 5-orbital system with Kanamori or Coulomb interaction with cubic interpolation. The sign obtained in simulations using conventional sampling is usually not better than with superstate-sampling. The graph also shows that the sign obtained in state-sampling is only slightly worse than that obtained using the other methods.
Fig. 4 shows the achieved speed-up factors (top panel) and the absolute sampling rates measured in simulations (bottom panel) of the impurity model with Coulomb interaction. We obtain a speed-up of the Monte Carlo sampling up to a factor of about 700 in the considered temperature range, depending on the used sampling method, temperature and interaction. Remarkably, the speed-up of our five-orbital example was never smaller than 100 for the most arduous case, i.e. the full interaction. The reason why the speed-up factors for the Coulomb interaction are larger under otherwise equal conditions is that a larger number of superstates contribute on average in this case. A general observation is that the speed-up decreases with decreasing temperature because the number of superstates contributing to the trace decreases with decreasing temperature as more operators tend to cause more quantum number violations. Therefore only the quantum-number checking can be avoided at lower temperatures, whereas at higher temperatures other trace contributions are present for which matrix-vector products need not be calculated any more.

Yet, there is still a noticeable speed-up even at lower temperatures as the quantum-number checking consumes a large amount of time. As the speed-up affects only the trace calculation, it will also continue to decrease with increasing temperature because the computational complexity of the bath determinant (which scales with $\beta^3$) is asymptotically larger than that of the trace calculation.

Another noticeable feature is the bigger advantage of state-sampling over superstate-sampling for Coulomb interaction, which is due to the larger average size of the outer superstates in that case. The amount of calculated matrix-vector products is reduced by approximately that factor in state-sampling compared to superstate-sampling, as only one of the outer states is chosen in the former case. This optimization is only advantageous for a matrix-vector solver like ours, as additional outer states can be included at negligible further cost if the entire product of the operator matrices for a specific outer superstate has been calculated, cf. Sec. IV C. A similar optimization is also possible without splitting configurations in the superstate-sampling method by using the cyclical invariance of the trace and starting the trace calculation at a $\tau$ where the superstate of the configuration is smaller than the outer superstate of the configuration, but this can interfere with time savings from caching of intermediate state vectors and even the size of the smallest superstate may be greater than one. More details on the methods can be found in Ref. 24.

In conclusion, we find that superstate-sampling improves performance without significant drawbacks to such an extent that it should always be preferable to the conventional sampling method. Especially in simulations with few good quantum numbers, state-sampling can provide an additional speed-up, though it can also impact the mean sign. In our examples, the speed-up in the case with full Coulomb interaction is big enough to clearly outweigh the reduced sign, but this may depend on characteristics of the system and the implementation.

VI. APPLICATION: THE SPIN-FREEZING CROSSOVER

In Ref. 24, Werner, et al. applied DMFT to a model with three degenerate orbitals and rotationally-invariant Coulomb interaction. Upon changing the filling $n$, they identified a sharp change in the qualitative behavior of the local spin susceptibility. For small fillings, the latter becomes rapidly small at large imaginary times $\tau$, as in a standard metal. Approaching the Mott transition at $n = 3$, it starts instead to closely resemble that of an atomic insulator, i.e. it seems to become essentially constant in $\tau$. This is surprising, as it happens for fillings still on the metallic side, before reaching the metal-insulator Mott-Hubbard transition. This sudden loss of coherence was interpreted as an abrupt crossover – or even a true quantum phase transition (“spin-freezing”) – to a bad metal characterized by violations of the Fermi-liquid properties, along a line in the zero-temperature $n-U$ phase diagram.

De’ Medici, Mravlje and Georges showed convincing evidence that the same model remains in a metallic phase, away from integer fillings. What changes upon getting close to the Mott transition is the coherence of the quasiparticle excitations. The “spin-freezing” is therefore a finite-temperature crossover to a “bad-metal” rather than a $T=0$ phase transition. As long as the DMFT self-consistency does not lead to a gap in the spectral function of the bath of the corresponding Anderson impurity model, the solution at zero temperature must indeed be a Fermi liquid. This conclusion was already demonstrated more rigorously in Ref. 25 and 26.

An analysis right at the spin-freezing crossover has been done later (see Fig. 7 of Ref. 26). There the focus was on the functional dependence of $\text{Im} \Sigma(i\omega_n)$, which was argued to be more compatible with that of Fermi-liquid in an extremely small region of low frequencies, at low-enough temperature. More recently, similar types of crossovers have been discussed in the presence of spin-orbit interaction always showing the change of behavior in the local spin susceptibility at a fixed temperature. Yet, one would like to unambiguously demonstrate that this is actually the physics of a crossover from a good- to a bad-metal with a coherence temperature that becomes fairly small upon approaching the Mott transition at half filling.

We therefore consider the same model of Ref. 24 as a function of doping focusing in particular on the temperature dependence. Using the CT-HYB implementation of $\text{wdynamis}$, featuring both superstate-sampling and the sliding window sampling, we are able to reach temperatures low enough to quantify the coherence scale, even deep into the “spin-freezing region”. The quantities of interest are the electronic self-energy $\Sigma(i\omega_n)$ and the...
Im(\Sigma(\iota \omega_n))/D \chi_{\mu B}^{[n]}

\begin{align*}
\chi_{\text{loc}}^{\omega=0}(T) &= \int_0^\beta d\tau \chi_{\text{loc}}(\tau),
\end{align*}

i.e. the \( \omega = 0 \)-Fourier component of the spin-spin response function \( \chi_{\text{loc}}(\tau) = g^2 \sum_{ij} \langle S_i^z(\tau) S_j^z(0) \rangle \) (with \( i \) and \( j \) running over the three orbitals). The half-bandwidth of the semi-circular non-interacting density of states of each orbital is \( D \) (corresponding to \( 2t \) in Ref. [23]). The coupling constants of the three-orbital Kanamori interaction are the usual Hund-\( J \) and Hubbard-\( U \) (with \( V = U - 2J \)) at fixed \( J/U = 1/6 \) ratio.

We first focus on the imaginary part of the Matsubara self-energy \( \text{Im}\Sigma(\iota \omega_n) \) for three different fillings at the lowest temperature \( T = 1/\beta = D/1600 \). As shown in Fig. [6] for filling \( n = 1.8 \) the extrapolation of \( \text{Im}\Sigma(\iota \omega_n) \) for \( \iota \omega_n \to 0 \) is not dramatically influenced by the temperature. This indicates that the system has reached a coherent Fermi-liquid state and further lowering \( T \) does not change the shape of \( \text{Im}\Sigma(\iota \omega_n) \) but only makes the Matsubara frequencies denser, remaining on the same “straight” line. This is a manifestation of the so-called “first-Matsubara” rule\[20\]\[32\] according to which a \( T^2 \) scattering rate characteristic of a Fermi-liquid gives rise to a linear-in-\( T \) value for \( \text{Im}\Sigma(\iota \omega_n=0) \).

The situation is drastically different for larger values of \( n \). Note that at this \( U, n = 2.35 \) and \( 2.63 \) had been already assigned to the “spin-freezing region” in the original paper by Werner, et al. (see inset to Fig. [9]). At these fillings the low-frequency part of \( \text{Im}\Sigma(\iota \omega_n) \) is highly non-linear and it is clear that to recover the “first-Matsubara” rule one has to consider the lowest temperatures (and probably even lower than \( T = D/1600 \) at \( n = 2.63 \)). This unambiguously hints at a sudden drop of the Fermi-liquid coherence temperature upon increasing the filling \( n \).

An inspection of the local spin susceptibility confirms that the physics at \( n = 2.35 \) and \( 2.63 \) is not qualitatively different from the good-metal fillings but it is just the result of a strong renormalization of the coherence properties. The results are shown in Fig. [7] For an atom, \( \chi(\tau) \) is perfectly flat independently of the temperature, so that its integral from 0 to \( \beta \) is proportional to \( \beta \) (Curie law). For a Fermi liquid, its shape instead has to change with temperature in such a way that its integral gives a constant Pauli susceptibility. Even though the speed of the decay for \( n = 2.35 \) and \( 2.63 \) is greatly reduced compared to \( n = 1.8 \), in agreement with Werner, et al.), a pronounced temperature dependence of \( \chi(\tau) \) is present also for the larger fillings, revealing the Fermi-liquid properties.

To conveniently represent the evolution with the filling we look at the value of the susceptibility at \( \tau = \beta/2 \). In the Fermi-liquid case this has to go (quadratically) to 0 upon reducing \( T \). The coherence temperature can be esti-
FIG. 8. Behavior of $\chi(\tau = \beta/2)$ for different temperatures and fillings. The lower panel contains a close-up of the three fillings shown in Fig. 7 across the spin-freezing crossover.

FIG. 9. Temperature dependence of $\chi^{\omega=0}(T)$ for different fillings (indicated by square markers of the same colors in the inset taken from Ref. [24]). The low-$T$ Pauli-like behavior is visible in the “non-frozen” regime (lower panel). At higher temperature there is a crossover to Curie-Weiss physics. The latter gets more and more dominant in “frozen regime” (upper panel).

It is too much to ask our quantitative estimate of $T_{coh}$ to reveal the precise behavior right at the spin-freezing crossover. Nevertheless, the existence of a sudden drop of $T_{coh}$ approaching half-filling, as in fact pioneered by Werner, et al., in Ref. [21] is undoubted. The reason why this crossover is so sharp, as well as its shape in the doping-$U$ diagram, are not fully clear yet.

In real materials, the position of the coherence scale can be strongly influenced by several factors, such as the non-local hybridization between orbitals, absent in the model Hamiltonian studied here. One of these factors has been also identified in the presence of sharp peaks in the non-interacting density of the states, something often coexisting with the many-body physics in strongly correlated materials.

VII. CONCLUSIONS

We have shown that the sum over all impurity eigenstates of the local problem in CT-HYB can be divided into smaller pieces, and sampled individually. This fragmentation leads to a remarkable gain in the algorithm’s efficiency, to some extent against the general intuition.
This is due to the exponential character of the imaginary time evolution $e^{-\beta n_c \tau}$, which very sensitively damps the amplitudes of high energy excitations. Touching the core of the exponential scaling in CT-HYB we manage to achieve speed-up factors of the order of $10^3$, with essentially no worsening of the average sign. Additional research has to be carried out in order to show whether the impact of the exponential scaling of the local problem can be reduced further by employing methods based on our ideas.

The speed-up figures have been obtained for a five-orbital model with full-Coulomb interaction, representing physically relevant situations such as realistic transition-metal impurities deposited on metallic substrates. We also discussed the well-known spin-freezing crossover obtained in three-orbital Hubbard-model calculations. Reaching very low temperature allows us to quantify the coherence temperature and the recovery of the Fermi-liquid properties of the self-energy, even when this physics is pushed to very low scales by the proximity to the Mott transition.

Appendix A: Invariance of the bath weight under global $\tau$-shift move

For the proof concerning the value of the bath determinant (adapted from Ref. [23]), let us consider the form of the hybridization matrix elements with time ordering along both dimensions,

$$\Delta_{ij} = \Delta \left( \tau_i - \tau_j \right), \quad (A1)$$

where $\Delta(\tau)$ is the hybridization function, an antiperiodic function with period $\beta$, $\tau_i$ the imaginary time of the $i$-th annihilator (ordered by imaginary time) and $\tau_j^\dagger$ the imaginary time of the $j$-th creator. The number of annihilators and creators shifted across $\tau = 0$ by the move will be denoted as $N_A$ and $N_C$ in the following.

Due to the $\tau$-shift move, the imaginary time of operators with $\tau < \Delta \tau$ will be transformed as $\tau \to \tau + \beta - \Delta \tau$ and that of other operators as $\tau \to \tau - \Delta \tau$. The arguments of the hybridization functions are only time differences in which the shift parameter $\Delta \tau$ always cancels, but in cases where exactly one of the two operators had a $\tau < \Delta \tau$, the corresponding matrix element changes its sign.

Additionally, since the ordering of the operators is cyclically permuted, the rows and columns are cyclically permuted such that the $N_A$ first rows become the $N_A$ last rows and the $N_C$ first columns become the $N_C$ last rows, where $N_A$ is the number of annihilators with $\tau < \Delta \tau$ and $N_C$ the number of creators with $\tau < \Delta \tau$. A cyclic permutation that moves every column or row exactly one position toward the front (“wrapping around” from the beginning to the end) is equivalent to swapping adjacent columns or rows $k - 1$ times, where $k$ is the size of the matrix in that dimension (as the hybridization matrix is a $k \times k$-matrix for hybridization expansion order $k$).

Each swap causes the determinant to change its sign, and the sign change of matrix elements where only one operator wrapped around the end is equivalent to multiplying all wrapped rows and columns by $-1$, where every multiplication of a column or row causes the determinant to change its sign as well. In total, expressed using the number of wrapped operators $N_{\text{wrap}} = N_A + N_C$, the determinant thus accumulates an additional factor of $\left( (-1)^{k-1} \right)^{N_{\text{wrap}}} \cdot (-1)^{N_{\text{wrap}}} = (-1)^{kN_{\text{wrap}}}$, where the matrix size $k$ is the expansion order.

This extra factor is compensated by the sign that is incurred due to time-ordering. That the change of this extra sign is equal to the factor acquired by the determinant may be proven by considering the amounts of permutations necessary to restore the ordering after performing a $\tau$-shift move that wraps exactly one operator around the origin. From this, the general case follows.

VIII. ACKNOWLEDGMENTS

We thank M. Aichhorn, M. Karolak, A. Toschi, L. de’ Medici, F. Assaad, M. Capone and K. Held for useful discussions. This work was supported by the DFG through SFB1170 “Tocotronics”. We gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time on the GCS Supercomputer SuperMUC at Leibniz Supercomputing Centre (www.lrz.de).
FIG. 11. Comparison of the CT-HYB self-energy calculated using the new sampling method (points) to exact diagonalization (lines) for a two-orbital system with Kanamori interaction.

Appendix B: Invariance of the local weight under global $\tau$-shift move

For the proof that the new local weight (adapted from Ref. 23) the cyclic invariance of the trace is used. Due to way superstates are chosen by definition we know that if the trace is restricted in such a way that non-zero components are left for only one superstate at any $\tau$, the result will be the same as if done so everywhere. Our local weight is effectively the conventional local weight with such a restriction applied at $\tau = 0$ and $\tau = \beta$, and if projection operators $P(x)$ are inserted onto the outer superstate $x$ at the beginning and end of the product of time-evolution, creation and annihilation operators corresponding to current configuration, it can written as a proper trace:

$$w_{\text{loc}} = \sum_{s \in x} \langle s | \hat{C} | s \rangle$$

$$= \text{tr} \left( P(x) \hat{C} P(x) \right).$$

As the time-evolution does not mix states from different superstates, the superstate projection operator commutes with all time-evolution operators,

$$\left[ \exp (-\tau \hat{H}), P(x) \right] = 0,$$

for any superstate $x$ and any $\tau$. Because the creators and annihilators map each source superstate to one unique target superstate and vice versa, a projector onto a superstate $x$ on one side of an annihilator or creator can be replaced by a projector onto the superstate $y$ which the operator maps $x$ to on the other side of the operator:

$$d(y|x) P(x) = P(y) d(y|x),$$

where the mapping of superstates relevant for the specific case is given in parentheses in superscript with the meaning that applying $d(y|x)$ to a state vector (to the right) in the subspace of superstate $x$ will produce a state vector in the subspace of superstate $y$.

Using these two relations, we can commute the projectors all the way through the product to any other point and also to any imaginary time by splitting time-evolution at that time into two consecutive time-evolutions if necessary. The projectors will not necessarily be projectors onto the old outer superstate any more, but onto the superstate that can be found at that point in the sequence for the current configuration. After commuting both projectors to the position in the product corresponding to the imaginary time $\Delta \tau$, the product can be cyclically permuted such that one of the projectors ends up at each end of the trace. It is then equivalent to the local weight of the superstate-sampling configuration after a $\tau$-shift by $\Delta \tau$. This shows that the local weight does not change after a $\tau$-shift move.

Appendix C: Exact diagonalization cross-check

Fig. II shows that the CT-HYB results with superstate-sampling agree with an exact-diagonalization benchmark.

---

1 A. N. Rubtsov, V. V. Savkin, and A. I. Lichtenstein, Phys. Rev. B 72, 035122 (2005)

2 P. Werner, A. Comanac, L. de’ Medici, M. Troyer, and A. J. Millis, Phys. Rev. Lett. 97, 076405 (2006)
Note that the number of states per superstate varies and the axis range of 10 was sufficient to capture the maximum number of contributing states over the entire course of the simulation due to many good quantum numbers.

In this work we consider the computational cost of quantum-number checking in \textit{w2dynamics} with decreasing temperature. The magnitude of this speed-up might of course be lower in other implementations.

A. Kowalski, \textit{Improved sampling in continuous-time quantum Monte Carlo algorithms for fermions}, Master’s thesis, Julius-Maximilians-Universität Würzburg, Würzburg (2017).

P. Werner, E. Gull, M. Troyer, and A. J. Millis, \textit{Physical Review Letters} \textbf{101}, 166405 (2008).

L. de Medici, J. Mravlje, and A. Georges, \textit{Physical Review Letters} \textbf{107}, 256401 (2011).

A. Georges, L. de Medici, and J. Mravlje, Annual Review of Condensed Matter Physics \textbf{4}, 137 (2013).

L. De Leo, \textit{Non-Fermi liquid behavior in multi-orbital Anderson impurity models and possible relevance for strongly correlated lattice models}, Ph.D. thesis, SISSA (2004).

L. De Leo and M. Fabrizio, \textit{Physical Review B} \textbf{69}, 245114 (2004).

A. J. Kim, H. O. Jeschke, P. Werner, and R. Valentí, \textit{Physical Review Letters} \textbf{118}, 086401 (2017).

A. V. Chubukov and D. L. Maslov, \textit{Physical Review B} \textbf{86}, 155136 (2012).

L. V. Pourovskii, J. Mravlje, A. Georges, S. I. Simak, and I. A. Abrikosov, \textit{New Journal of Physics} \textbf{19}, 073022 (2017).

A. Hausoel, M. Karolak, E. S¸a¸sıo˘ glu, A. Lichtenstein, K. Held, A. Katanin, A. Toschi, and G. Sangiovanni, \textit{Nature Communications} \textbf{8}, ncomms16062 (2017).

K. G. Wilson, \textit{Rev. Mod. Phys.} \textbf{47}, 773 (1975).

L. de’ Medici, “Hund’s metals explained,” (E. Pavarini, E. Koch, R. Scalettar, and R. Martin (eds.) The Physics of Correlated Insulators, Metals, and Superconductors Modeling and Simulation Vol. 7 Forschungszentrum Juelich, ISBN 978-3-95806-224-5, 2017).