PHYSICAL REVIEW E **95**, 031301(R) (2017)

Recommender engine for continuous-time quantum Monte Carlo methods

Li Huang,[1] Yi-feng Yang,[2,3,4,*] and Lei Wang[2,†]

[1]*Science and Technology on Surface Physics and Chemistry Laboratory, P.O. Box 9-35, Jiangyou 621908, China*

[2]*Beijing National Laboratory for Condensed Matter Physics and Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China*

[3]*Collaborative Innovation Center of Quantum Matter, Beijing 100190, China*

[4]*School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100190, China*

(Received 17 December 2016; published 29 March 2017)

Recommender systems play an essential role in the modern business world. They recommend favorable items such as books, movies, and search queries to users based on their past preferences. Applying similar ideas and techniques to Monte Carlo simulations of physical systems boosts their efficiency without sacrificing accuracy. Exploiting the quantum to classical mapping inherent in the continuous-time quantum Monte Carlo methods, we construct a classical molecular gas model to reproduce the quantum distributions. We then utilize powerful molecular simulation techniques to propose efficient quantum Monte Carlo updates. The recommender engine approach provides a general way to speed up the quantum impurity solvers.

DOI: 10.1103/PhysRevE.95.031301

At the heart of every quantum Monte Carlo (QMC) method is a quantum to classical mapping. One has to find a classical representation of the quantum system to program it into the classical computers [1]. There are various QMC methods [2–14] since the mapping is not unique. The classical representations behind these QMC methods can be classical spins, particles, or polymers, etc. In this unified point of view, various QMC methods differ by the implementation details, but all share the same principle.

Besides finding the suitable classical representations, another key ingredient of the QMC algorithms is to design efficient strategies to sample the configurations. Successful algorithms [5–7] typically identify the collective modes of the effective statistical mechanics problem and make proposals accordingly. However, it is not always easy to devise these efficient updates for generic quantum many-body systems. For example, most of the QMC simulations of the fermionic systems in condensed matter physics still use simple local updates [2,11]. These updates can be inefficient due to high rejection rate and long autocorrelation times.

The idea of a "recommender system" points to a general route to accelerate the quantum Monte Carlo simulations. The recommender system is a broad and active research field [15] in machine learning. One can build a probabilistic model based on the users' past behavior and suggest favorable products back with high acceptance rates. Similarly, one can model the probability distribution of the QMC configurations with machine learning techniques and propose new efficient Monte Carlo updates accordingly. This core idea has been presented in Refs. [16–18]. Recommender systems built on the restricted Boltzmann machine [19,20] and the classical spin systems do speed up the Monte Carlo sampling without introducing any bias to the physical results.

Howevere, it is not obvious how to extend these ideas to a broader class of modern QMC methods [3–14] in which the number of random variables can fluctuate in the simulation. This Rapid Communication presents several new ingredients to achieve speedup in these, and in particular, the continuous-time quantum Monte Carlo (CT-QMC) methods. The CT-QMC methods have revolutionized the study of quantum impurity models [11] since their invention a decade ago [9,10]. The latest developments [13,14] built on Ref. [12] further extend these successes to lattice fermions [21,22]. There were dedicated efforts to optimize the sampling of the CT-QMC methods [22–26] because of their broad impacts. Our innovations include reverse engineering a classical molecular gas model from the CT-QMC configurations and leveraging mature molecular simulation techniques to propose efficient updates for the CT-QMC simulation. This approach provides a systematic and principled approach to improve the efficiency of the CT-QMC methods. This progress has an immediate impact on the realistic simulation of correlated materials [27].

To make our discussion concrete, we consider the single impurity Anderson model. The action reads [28]

$$S = -\int_0^\beta d\tau \int_0^\beta d\tau' \sum_{\sigma=\{\uparrow,\downarrow\}} c_\sigma^\dagger(\tau)\mathcal{G}_0^{-1}(\tau-\tau')c_\sigma(\tau') + U\int_0^\beta d\tau\left(n_\uparrow-\frac{1}{2}\right)\left(n_\downarrow-\frac{1}{2}\right). \quad (1)$$

The model describes a quantum impurity embedded in a noninteracting environment illustrated in Fig. 1(a). β is the inverse temperature and U is the on-site interaction strength. $\mathcal{G}_0$ is the noninteracting Green's function of the impurity. In the Matsubara frequency, it reads $\mathcal{G}_0^{-1}(i\omega_n) = i\omega_n + \varepsilon - \lambda^2\Delta(i\omega_n)$, where ε is the local chemical potential of the impurity and λ is the hybridization strength between the impurity and the noninteracting bath. In the following, we consider a bath with a semicircular density of states $\Delta(i\omega_n) = 2/(i\omega_n + \sqrt{(i\omega_n)^2 - D^2})$ [28] and set the half bandwidth $D = 2$ as the energy unit. Physically, the single impurity Anderson model (1) is relevant to cases of magnetic atoms hosted in a metal or quantum dots coupled to the leads. The model captures rich physical phenomena including local moments, Coulomb blockade, and the Kondo effect [29]. Moreover, solving the quantum impurity model is the computational

*yifeng@iphy.ac.cn
†wanglei@iphy.ac.cn

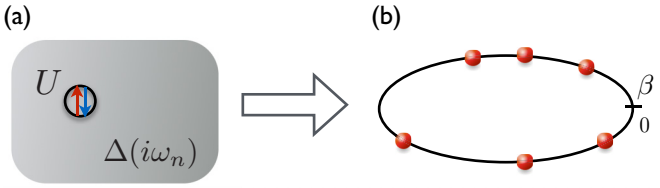

FIG. 1. (a) The quantum impurity problem consists of an impurity with local interaction embedded in a bath of non-interacting fermions. (b) The CT-QMC method maps the quantum impurity model to a one-dimension classical molecular gas model. Each red dot represents an interaction vertex in the interaction expansion Eq. (2). The length of periodic imaginary time is the inverse temperature β. Configurational bias Monte Carlo simulation of the molecular gas recommends efficient updates to the CT-QMC simulation.

engine in the dynamical mean-field theory studies of correlated lattice models [28].

The interaction expansion CT-QMC impurity solver [9,11] performs diagrammatic expansion of the partition function in terms of the interaction strength

$$Z/Z_0 = \sum_{k=0}^{\infty} \int_0^\beta d\tau_1 \int_{\tau_1}^\beta d\tau_2 \cdots \int_{\tau_{k-1}}^\beta d\tau_k \, (-U)^k |\det(G)|^2, \tag{2}$$

where Z_0 is the noninteracting partition function. Introducing the configuration $C = \{\tau_1, \tau_2, \ldots, \tau_k\}$ and the weight $w(C) = (-U)^k |\det(G)|^2$, we rewrite Eq. (2) as $Z/Z_0 = \sum_C w(C)$. Here the summation over configurations denotes the discrete summation over the expansion order and the time-ordered integrations over the imaginary times. G is a $k \times k$ matrix whose matrix elements are given by the noninteracting Green's function $G_{ij} = \mathcal{G}_0(\tau_i - \tau_j) - \delta_{ij}/2$. The CT-QMC simulation of the single impurity Anderson model is not hindered by the fermion sign problem in general [30]. For simplicity, in the following we consider $U < 0$ so that we can directly interpret $w(C)$ as a positive Boltzmann weight [31]. Accelerating the CT-QMC methods with a recommender engine is nevertheless detached from the issue of the sign problem because one can always model the probability distribution $|w(C)|$ and achieve speedups.

The expansion Eq. (2) formally maps the zero-dimensional quantum impurity model (1) to a one-dimensional "classical molecular gas" model, shown in Fig. 1(b) [21,32]. The molecular gas is in the grand canonical ensemble, where each molecule represents an interaction vertex residing in the continuous imaginary-time axis. Conventional updates of the CT-QMC methods [9,11] indeed resemble the grand canonical Monte Carlo simulation of the molecular gases [33,34]. In these simulations, one attempts to insert or remove a vertex according to a uniform probability distribution and accepts or rejects the move according to the change of the Monte Carlo weight. However, these simple updates ignore the correlations between the vertices and can suffer from low acceptance rates and long autocorrelation times. Curing such inefficiency requires a better analytical understanding of the correlations in the Monte Carlo configuration and designing suitable updates correspondingly. This is, however, a nontrivial task because of the determinant in the CT-QMC weight (2).

To address these problems we exploit the aforementioned quantum to classical mapping explicitly. First, we distill the correlations in the QMC configurations into a classical reference system. Then, we use the reference system as a recommender engine to guide future QMC sampling. In line with the mapping of Fig. 1(b), we write the partition function of the molecular gas as $Z_{\text{gas}} = \sum_C e^{-E(C)}$ and assume an explicit energy function form

$$E(C) = -\sum_{i=1}^{k} \mathcal{V}^{(2)}(\tau_{i+1} - \tau_i) - \sum_{i=1}^{k} \mathcal{V}^{(3)}(\tau_{i+1} - \tau_{i-1}) - \mu k - b. \tag{3}$$

$\mathcal{V}^{(2)}$ is a two-body interaction potential depending on the time difference of the two adjacent vertices. Here the subscripts of the imaginary times and the time differences all take into account the periodic boundary condition of the imaginary time axis. The second term of (3) is a three-body interaction where the two vertices interact with $\mathcal{V}^{(3)}$ only if there is the third vertex in between. The effective chemical potential μ term controls the average molecule number in the grand canonical ensemble. Finally, b is an energy offset which controls the relative magnitude of Z/Z_0 and Z_{gas}. Equation (3) defines an energy-based model for the CT-QMC probability distribution. To determine its exact form, we adopt a data-driven approach and train the model parameters using collected CT-QMC configuration data. Note that the model (3) does not need to reproduce the probabilities exactly. Capturing the crucial correlations in the original CT-QMC configurations is already good enough to be a useful recommender engine.

To determine the interaction potentials, we parametrize these continuous functions using the Legendre polynomials $\mathcal{V}^{(2,3)}(\tau) = \sum_{\ell=1}^{L} P_\ell[x(\tau)] \mathcal{V}_\ell^{(2,3)}$, where $P_\ell[x]$ is the ℓth-order Legendre polynomial and $x(\tau) = 2\tau/\beta - 1$ maps the continuous time differences τ to the region $[-1,1]$. We keep the expansion coefficients $\mathcal{V}_\ell^{(2,3)}$ up to Lth order [35]. Requiring Z_{gas} to match the expansion of Z/Z_0 term by term, we have

$$\ln[w(C)] = \sum_{\ell=1}^{L} \left\{ \sum_{i=1}^{k} P_\ell[x(\tau_{i+1} - \tau_i)] \right\} \mathcal{V}_\ell^{(2)} + \sum_{\ell=1}^{L} \left\{ \sum_{i=1}^{k} P_\ell[x(\tau_{i+1} - \tau_{i-1})] \right\} \mathcal{V}_\ell^{(3)} + \mu k + b. \tag{4}$$

Equation (4) defines a linear regression problem for the parameters $\{\mathcal{V}_\ell^{(2)}, \mathcal{V}_\ell^{(3)}, \mu, b\}$. From a machine learning perspective, the Legendre polynomials $P_\ell[x]$ and the expansion order k are the features we manually extracted from the CT-QMC configuration $C = \{\tau_1, \tau_2, \ldots, \tau_k\}$. This feature engineering is motivated by the physical considerations based on the molecular gas model (3). The quantum to classical mapping naturally solves the problem of modeling a fluctuating number of continuous random variables.

To collect the training data, we perform CT-QMC simulations with conventional random insertion and removal updates [9]. For each update whether it is accepted and

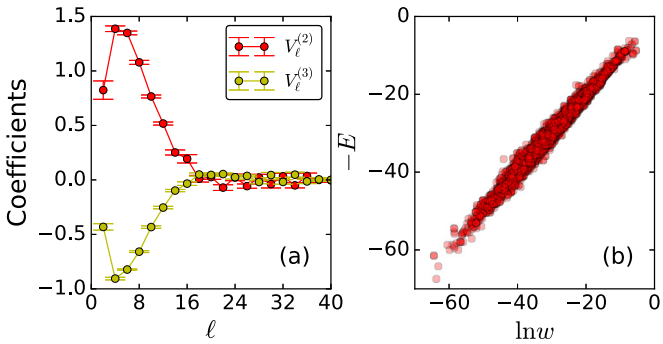

FIG. 2. (a) The Legendre coefficients $\mathcal{V}_\ell^{(2,3)}$ of the two- and three-body interactions in Eq. (3). (b) The log weight of the CT-QMC, Eq. (2), and of the molecular gas model, Eq. (3). Each red dot represents a test sample which is not used for training. The physical parameters of the quantum impurity model (1) are $\beta = 100$, $U = -2$, $\varepsilon = 0.2$, and $\lambda = 1.0$.

rejected we extract the features in the right-hand side of Eq. (4) and compute the log weight as the regression target. After collecting around 20 000 samples we perform the ridge regression [36] for the fitting parameters, where we use an L_2 regularization of the strength 10^{-3} for the coefficients $\{\mathcal{V}_\ell^{(2)}, \mathcal{V}_\ell^{(3)}, \mu\}$ to prevent overfitting. Figure 2(a) shows the fitted Legendre coefficients. The error bars are estimated using eight independent runs. The coefficients vanish for large ℓ, justifying the truncation of the Legendre expansion [37]. To verify success of the fitting, Fig. 2(b) shows the exact log weight of the CT-QMC and the predicted log weight of the classical gas model (3) on the test samples. Strong positive correlation indicates that the fitting indeed captures the distribution of the CT-QMC configurations well.

Figure 3 shows the effective interaction potentials of the molecular gas model (3). For weak hybridization strength, the molecules are effectively noninteracting because there is very little correlation in the imaginary time for an almost isolated quantum impurity. As the hybridization strength λ increases,

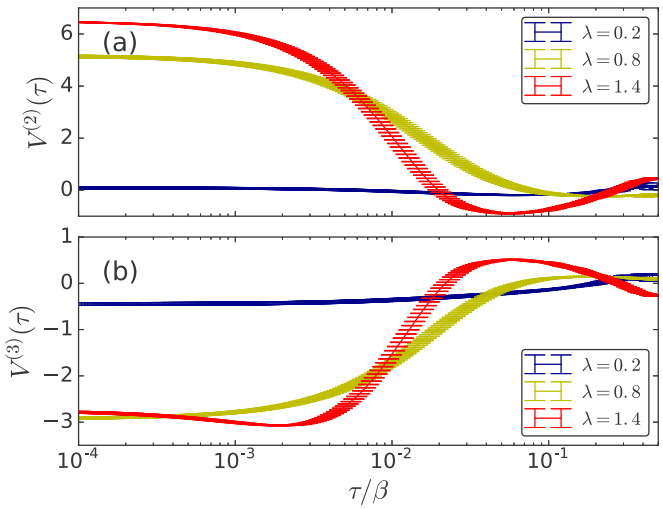

FIG. 3. (a) The two-body and (b) the three-body interaction potentials for various hybridization strengths λ. Positive (negative) value means attractive (repulsive) interactions due to the minus sign in Eq. (3). The physical parameters are identical to Fig. 2.

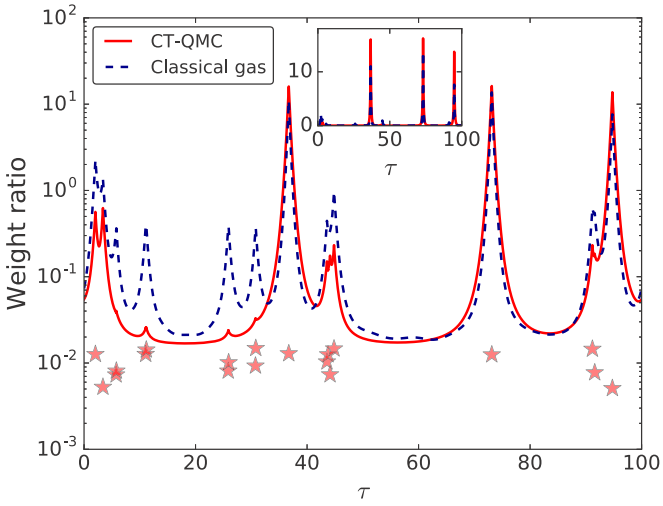

FIG. 4. The weight ratio of adding a vertex at the imaginary time τ to the configuration C. The red solid line is $w(C \cup \{\tau\})/w(C)$ and the blue dashed line is $e^{-E(C\cup\{\tau\})+E(C)}$. The red stars indicate the location of the existing vertices in C. For visibility we offset their vertical positions in the graph. Inset shows the same plot in the linear scale which highlights the three dominant peaks. The physical parameters β, U, ε are identical to Fig. 2.

the two-body interaction potential $\mathcal{V}^{(2)}$ becomes attractive while the three-body interaction $\mathcal{V}^{(3)}$ becomes repulsive near short time differences. The combined interaction effects will favor configurations with bounded pairs of vertices. Physically, various impurity quantum phase transitions and crossovers manifest themselves in the classical gas model [21,32,38]. Computationally, knowing the effective interactions between the vertices can help us sample them more efficiently.

To better understand how well the classical molecular gas model captures the correlation between the interaction vertices, Fig. 4 shows the weight ratio of adding a vertex at the imaginary time τ. This ratio determines the acceptance rate of the insertion update. The red stars indicate the locations of the existing vertices C. The weight ratio exhibits peaks around them; some peaks are more pronounced than others. This is because the effective two-body attractive interactions enhance the probability of adding the new vertex near the existing vertices, while the effective three-body repulsive interaction suppresses the probability of adding the third vertex in the vicinity of two already paired-up vertices. For example, the CT-QMC weight ratio only exhibits a small peak around $\tau \sim 11$ in Fig. 4. It is crucial to include the three-body interaction term in the classical model Eq. (3). Without them the vertices will collapse into clusters which are certainly not favored in the original CT-QMC simulation.

The overall effect of the combined interactions is that there are only three dominant peaks in the weight ratio around the three isolated vertices (see the inset of Fig. 4). Randomly inserting a vertex in $[0,\beta)$ without taking into account this highly nonuniform distribution will have poor acceptance rate.

Having trained the classical molecular gas model as a proxy of the original probability distribution, we simulate it using efficient classical Monte Carlo methods and recommend the update back to the original QMC simulation. Assuming the simulation of the classical gas model satisfies the detailed

balance condition, we accept the recommended move from the configuration C to C' with the acceptance probability [16,18],

$$A(C \to C') = \min\left\{1, \frac{e^{-E(C)}}{e^{-E(C')}} \frac{w(C')}{w(C)}\right\}. \tag{5}$$

Equation (5) guarantees an unbiased simulation with improved acceptance ratio. This approach boosts the overall performance because of proposing more favorable updates.

There are various ways that the CT-QMC simulation can benefit from the recommender system [16–18]. First, the updates can be nonlocal, in which case C' can differ drastically from C while still having a high acceptance rate with an accurate fitting in Eq. (4). Furthermore, even without the luxury of performing global updates for the reference system, one can still afford to accumulate many local updates before recommending a nonlocal update to the CT-QMC simulation. Because the simulation of the molecular gas is much cheaper than the CT-QMC [$O(1)$ versus $O(k^2)$ operations per local update step] [39], the recommendation step has little overhead. Finally, as long as the classical molecular gas model captures correlations in the CT-QMC configurations, it is already beneficial since it suggests better update proposals by exploiting the correlations. In the last case, the recommended update can be local, but has an improved acceptance rate and enjoys the advantage of the $O(k^2)$ fast update schemes in the CT-QMC [9,11]. Using the recommender engine in this way, one can always speed up the simulation compared to the original case because the factor $e^{E(C')-E(C)}$ in Eq. (5) tends to increase the acceptance ratio by compensating the fluctuation of the Boltzmann weight ratio.

We employ the last approach and simulate the molecular gas model (3) using the configurational bias Monte Carlo (CBMC) method [40]. CBMC is an efficient molecular simulation technique [33,34], which is particularly useful for growing long molecular chains. The basic idea of CBMC is to probe the landscape around the current configuration and find a move with higher acceptance rate. To achieve this goal, we perform N independent trial updates from the old configuration and select an actual one according to their relative probabilities [41]. We then propose this CBMC update to the CT-QMC and accept it with probability (5). The number of trial steps N in the CBMC controls how much information we would like to extract from the molecular gas recommender system.

As a relevant measure of the improvement in the Monte Carlo sampling, Fig. 5 shows the autocorrelation times of the expansion order [41] measured in the unit of 10^3 CT-QMC update steps. Taking the recommendations from the CBMC simulation greatly reduces the autocorrelation time, especially in the challenging parameter regions with strong hybridization strength and at low temperature. It is encouraging to see that a few trial steps ($N = 10$) in the CBMC already significantly improves the efficiency of the CT-QMC. Moreover, one can even afford a larger number of CBMC trial steps because computing the weight ratio of Eq. (3) is much cheaper than Eq. (2). Increasing N further improves the autocorrelation time because it finds more probable updates in accordance to the correlation of the CT-QMC configurations. Improved autocorrelation time means one can collect a larger number

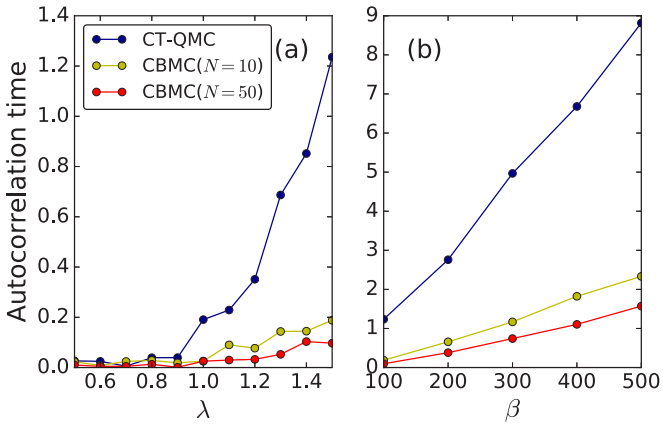

FIG. 5. The autocorrelation times of the total expansion order. The improvement of the ordinary CT-QMC results (blue dots) [9,11] is more significant in the difficult parameter region where the hybridization strength λ or the inverse temperature β is large. Increasing the trial steps N in the CBMC simulation of the molecular gas model (3) further reduces the autocorrelation time. In (a) $\beta = 100$ and in (b) $\lambda = 1.5$, the other physical parameters U, ε are identical to Fig. 2.

of independent samples in the same Monte Carlo steps and reduce the statistical uncertainty of the physical results [42].

Since the expansion order is a global property of the CT-QMC configuration, its autocorrelation time is typically larger than other physical observables. Greatly reducing this global autocorrelation time in the CT-QMC gives the hope of computing the fidelity susceptibility of correlated fermions at substantially lower temperature and larger system sizes [21,32,38].

Explicitly constructing the effective classical model and using it as a recommender system is a general approach to speed up the QMC simulations. Identifying the classical molecular gases as the recommender engines for CT-QMC methods brings a large number of powerful and mature molecular simulation techniques [33,34] into the game. In this Rapid Communication we employ the CBMC method [40] to efficiently explore the probability landscape of the classical molecular gas model. One may consider using other successful molecular simulation approaches other than the CBMC method [40], such as the hybrid Monte Carlo [43] or geometric cluster algorithms [44], for further improvements.

For other variants of the CT-QMC methods with auxiliary fields in the configurations [9,45,46], it is natural to generalize the effective interaction potentials in Eq. (3) to be dependent on the auxiliary fields in addition to the imaginary times. The present recommender engine strategy can be generalized to the hybridization expansion [10] and the Kondo coupling expansion [47] CT-QMC algorithms. In the latter case, the spin-flip events in the imaginary time map to classical Coulomb charges [32] according to the seminal work of Anderson and Yuval [48]. Going beyond the quantum impurity models, the "recommender engine" approach can also benefit a broad range of modern QMC methods for interacting bosons [4,5,7,8], quantum spins [3,6], and lattice fermions [13,14,22]. In those cases, the classical molecules will carry additional indices to indicate the spatial location of the interaction events.

Besides serving as a recommender engine to speed up the QMC simulations, the classical reference system does capture physical information of the original quantum problem. For example, the average particle number is related to the average interaction energy, while the classical compressibility is related to the second-order derivative of the quantum free energy; and the bipartite particle fluctuation [49] of the classical system corresponds to the fidelity susceptibility of the quantum system [21]. These discussions indicate that the phase transitions of the quantum system will always manifest themselves in the corresponding classical representations. The correspondence calls for special attention in designing the classical reference system. For example, it is known that a one-dimensional model with the nearest-neighbor interaction only has no phase transition [50]. Therefore, to ensure the classical reference system has enough descriptive power one may need to allow additional internal degrees of freedom or longer ranged interactions between the molecules.

The recommender engine approach works as well even if the sign of the weight is not positive definite. In this case, sampling according to the absolute value of the weight still defines a legitimate statistical mechanics problem. However, the physics of the classical reference model may be detached from the original quantum system [51,52].

L.W. is supported by the Ministry of Science and Technology of China under Grant No. 2016YFA0302400 and the start-up grant of IOP-CAS. L.H. is supported by the Natural Science Foundation of China under Grant No. 11504340 and the Foundation of President of China Academy of Engineering Physics (Grant No. YZ2015012). Y.Y. is supported by the National Natural Science Foundation of China (Grant No. 11522435) and the Strategic Priority Research Program (B) of the Chinese Academy of Sciences (Grant No. XDB07020200). We use the ALPS library [53] for the Monte Carlo data analysis.

APPENDIX: CONFIGURATIONAL BIAS MONTE CARLO SIMULATION FOR THE MOLECULAR GAS MODEL

We present details of the configurational bias Monte Carlo [40] simulation of the classical molecular gas model $Z_{\text{gas}} = \sum_C e^{-E(C)}$. Pedagogical introductions about the CBMC method can be found in the textbooks of the molecular simulations [33,34].

For the insertion update, we randomly generate N imaginary times in the range $\tau_i \in [0,\beta)$ and compute the corresponding Boltzmann weight ratios $r_i = e^{-E(C\cup\{\tau_i\})+E(C)}$. We then select a τ_i according to the discrete distribution r_i/W, where $W = \sum_{i=1}^{N} r_i$. Assuming the current configuration C contains k imaginary times, the acceptance rate of the insertion update is

$$A(C \to C\cup\{\tau_i\}) = \min\left\{1, \frac{\beta W}{(k+1)N}\right\}. \tag{A1}$$

For the removal update, we randomly select a vertex from the k existing vertices in the current configuration C. Suppose its imaginary time is τ; we compute $W = e^{-E(C)+E(C\setminus\{\tau\})} + \sum_{i=2}^{N} e^{-E(C\setminus\{\tau\}\cup\{\tau_i\})+E(C\setminus\{\tau\})}$ with $N-1$ random numbers drawn uniformly from $\tau_i \in [0,\beta)$. The acceptance rate reads

$$A(C \to C\setminus\{\tau\}) = \min\left\{1, \frac{kN}{\beta W}\right\}. \tag{A2}$$

In the case of $N = 1$, the above algorithm reduces to the ordinary single-particle insertion or removal update of a molecular gas in a grand canonical ensemble. With increasing N, one will find a more probable update according to the probability distribution $e^{-E(C)}$. Finally, we recommend the configuration from the CBMC update to the CT-QMC simulation. Notice that the update remains local for any choice of N.

[1] J. Gubernatis, N. Kawashima, and P. Werner, *Quantum Monte Carlo Methods: Algorithms for Lattice Models* (Cambridge University Press, Cambridge, UK, 2016).
[2] R. Blankenbecler, D. J. Scalapino, and R. L. Sugar, Phys. Rev. D **24**, 2278 (1981).
[3] A. W. Sandvik and J. Kurkijärvi, Phys. Rev. B **43**, 5950 (1991).
[4] B. B. Beard and U.-J. Wiese, Phys. Rev. Lett. **77**, 5130 (1996).
[5] N. V. Prokof'ev, B. V. Svistunov, and I. S. Tupitsyn, J. Exp. Theor. Phys. **87**, 310 (1998).
[6] A. W. Sandvik, Phys. Rev. B **59**, R14157 (1999).
[7] H. G. Evertz, Adv. Phys. **52**, 1 (2003).
[8] N. Kawashima and K. Harada, J. Phys. Soc. Jpn. **73**, 1379 (2004).
[9] A. N. Rubtsov, V. V. Savkin, and A. J. Lichtenstein, Phys. Rev. B **72**, 035122 (2005).
[10] P. Werner, A. Comanac, L. de' Medici, M. Troyer, and A. J. Millis, Phys. Rev. Lett. **97**, 076405 (2006).
[11] E. Gull, A. J. Millis, A. I. Lichtenstein, A. N. Rubtsov, M. Troyer, and P. Werner, Rev. Mod. Phys. **83**, 349 (2011).
[12] S. M. A. Rombouts, K. Heyde, and N. Jachowicz, Phys. Rev. Lett. **82**, 4155 (1999).
[13] M. Iazzi and M. Troyer, Phys. Rev. B **91**, 241118 (2015).
[14] L. Wang, M. Iazzi, P. Corboz, and M. Troyer, Phys. Rev. B **91**, 235151 (2015).
[15] C. C. Aggarwal, *Recommender Systems: The Textbook* (Springer, New York, 2016).
[16] L. Huang and L. Wang, Phys. Rev. B **95**, 035105 (2017).
[17] J. Liu, Y. Qi, Z. Y. Meng, and L. Fu, Phys. Rev. B **95**, 041101 (2017).
[18] J. Liu, H. Shen, Y. Qi, Z. Y. Meng, and L. Fu, arXiv:1611.09364.
[19] P. Smolensky, in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (MIT Press, Cambridge, MA, 1986), Vol. 1, pp. 194–281.
[20] G. E. Hinton, Neural Comput. **14**, 1771 (2002).
[21] L. Wang, Y.-H. Liu, J. Imriška, P. N. Ma, and M. Troyer, Phys. Rev. X **5**, 031007 (2015).
[22] Y.-H. Liu and L. Wang, Phys. Rev. B **92**, 235129 (2015).
[23] E. Bourovski, N. Prokof'ev, and B. Svistunov, Phys. Rev. B **70**, 193101 (2004).
[24] E. Burovski, N. Prokof'ev, B. Svistunov, and M. Troyer, New J. Phys. **8**, 153 (2006).
[25] H. Shinaoka, M. Dolfi, M. Troyer, and P. Werner, J. Stat. Mech.: Theory Exp. (2014) P06012.

[26] H. Shinaoka, Y. Nomura, S. Biermann, M. Troyer, and P. Werner, Phys. Rev. B **92**, 195126 (2015).
[27] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti, Rev. Mod. Phys. **78**, 865 (2006).
[28] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996).
[29] H. R. Krishna-murthy, J. W. Wilkins, and K. G. Wilson, Phys. Rev. B **21**, 1003 (1980).
[30] J. Yoo, S. Chandrasekharan, R. K. Kaul, D. Ullmo, and H. U. Baranger, J. Phys. A: Math. Gen. **38**, 10307 (2005).
[31] The repulsive case can be handled with the trick of Refs. [9,45]. Note that we also use a finite local chemical potential ε to tune away from the special particle-hole symmetric point, where all the odd-order contributions in Eq. (2) vanish [9].
[32] L. Wang, H. Shinaoka, and M. Troyer, Phys. Rev. Lett. **115**, 236601 (2015).
[33] A. R. Leach, *Molecular Modeling: Principles and Applications* (Pearson Education, Upper Saddle River, NJ, 2001).
[34] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Academic Press, London, UK, 2002).
[35] We only keep the even ℓ terms in the Legendre expansion because the periodicity of the imaginary-time axis implies $\mathcal{V}^{(2,3)}(\tau) = \mathcal{V}^{(2,3)}(\beta - \tau)$. We also exclude the constant term $\ell = 0$ in the expansion since it can be absorbed into the chemical-potential term.
[36] T. Hastie, R. J. Tibshirani, and J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer, New York, 2009).
[37] Higher-order Legendre polynomials fit the detailed oscillatory behavior of the interaction potential and potentially lead to overfitting.
[38] L. Huang, Y. Wang, L. Wang, and P. Werner, Phys. Rev. B **94**, 235110 (2016).
[39] After the fitting Eq. (4) is done we precompute the interaction potentials $\mathcal{V}^{(2,3)}(\tau)$ on a fine mesh and use linear interpolations to obtain their values at other imaginary times in the CBMC simulation.
[40] J. I. Siepmann and D. Frenkel, Mol. Phys. **75**, 59 (1992).
[41] See the Appendix for details of the CBMC simulation.
[42] V. Ambegaokar and M. Troyer, Am. J. Phys. **78**, 150 (2010).
[43] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, Phys. Lett. B **195**, 216 (1987).
[44] J. Liu and E. Luijten, Phys. Rev. Lett. **92**, 035504 (2004).
[45] F. F. Assaad and T. C. Lang, Phys. Rev. B **76**, 035116 (2007).
[46] E. Gull, P. Werner, O. Parcollet, and M. Troyer, Europhys. Lett. **82**, 57003 (2008).
[47] J. Otsuki, H. Kusunose, P. Werner, and Y. Kuramoto, J. Phys. Soc. Jpn. **76**, 114707 (2007).
[48] P. W. Anderson and G. Yuval, Phys. Rev. Lett. **23**, 89 (1969).
[49] S. Rachel, N. Laflorencie, H. F. Song, and K. Le Hur, Phys. Rev. Lett. **108**, 116401 (2012).
[50] H. Takahashi, Proc. Phys.-Math. Soc. Jpn. **24**, 60 (1942).
[51] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst, arXiv:1608.07848.
[52] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, arXiv:1609.02552.
[53] B. Bauer *et al.*, J. Stat. Mech.: Theory Exp. (2011) P05001.