Improving the efficiency of hierarchical equations of motion approach and application to coherent dynamics in Aharonov–Bohm interferometers

Dong Hou, Shikuan Wang, Rulin Wang, Lvzhou Ye, RuiXue Xu, Xiao Zheng, and YiJing Yan

Citation: The Journal of Chemical Physics 142, 104112 (2015); doi: 10.1063/1.4914514
View online: http://dx.doi.org/10.1063/1.4914514
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/142/10?ver=pdfcov
Published by the AIP Publishing

Articles you may be interested in

Reduced hierarchical equations of motion in real and imaginary time: Correlated initial states and thermodynamic quantities

A hybrid stochastic hierarchy equations of motion approach to treat the low temperature dynamics of non-Markovian open quantum systems
J. Chem. Phys. 139, 134106 (2013); 10.1063/1.4822043

Electron transport through two irreducibly-coupled Aharonov-Bohm rings with applications to nanostructure quantum computing circuits
J. Appl. Phys. 110, 054315 (2011); 10.1063/1.3631782

Riccati equation and the problem of decoherence II: Symmetry and the solution of the Riccati equation

Quantum Transport in an Aharonov-Bohm Electron Interferometer
Improving the efficiency of hierarchical equations of motion approach and application to coherent dynamics in Aharonov–Bohm interferometers

Dong Hou,1,2 Shikuan Wang,3 Rulin Wang,4 LvZhou Ye,1 RuiXue Xu,1,2 Xiao Zheng,1,2,a) and YiJing Yan1,5,b)

1Hefei National Laboratory for Physical Sciences at the Microscale, University of Science and Technology of China, Hefei 230026, China
2Synergetic Innovation Center of Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, Anhui 230026, China
3Department of Physics, Hangzhou Dianzi University, Hangzhou, Zhejiang 310018, China
4Beijing Computational Science Research Center, Beijing 100084, China
5Department of Chemistry, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China

(Received 16 January 2015; accepted 2 March 2015; published online 13 March 2015)

Several recent advancements for the hierarchical equations of motion (HEOM) approach are reported. First, we propose an a priori estimate for the optimal number of basis functions for the reservoir memory decomposition. Second, we make use of the sparsity of auxiliary density operators (ADOs) and propose two ansatzs to screen out all the intrinsic zero ADO elements. Third, we propose a new truncation scheme by utilizing the time derivatives of higher-tier ADOs. These novel techniques greatly reduce the memory cost of the HEOM approach, and thus enhance its efficiency and applicability. The improved HEOM approach is applied to simulate the coherent dynamics of Aharonov–Bohm double quantum dot interferometers. Quantitatively accurate dynamics is obtained for both noninteracting and interacting quantum dots. The crucial role of the quantum phase for the magnitude of quantum coherence and quantum entanglement is revealed. © 2015 AIP Publishing LLC. [http://dx.doi.org/10.1063/1.4914514]

I. INTRODUCTION

The accurate and efficient characterization of static and dynamic properties of general open quantum systems with strong electron correlations is fundamentally important for the development of quantum computation and quantum information processing.1–4 The coherent dynamics of the double quantum dot (DQD) Aharonov–Bohm interferometer5–13 is an ideal test bed for verification of theoretical approaches.

In an Aharonov–Bohm interferometer, the electrons conducting through the two paths interfere with each other with a constant phase difference $\phi$. Experimentally, this can be realized by applying a magnetic flux $\Phi$ to a parallel DQD system, in which the electron coherence is tunable and quantum entangled states can be generated.14 A realistic simulation of the quantum coherent dynamics of an Aharonov–Bohm DQD interferometer has remained a challenging task. This is because the non-Markovian memory effect as well as the strong electronic correlations need to be accurately accounted for. Moreover, to preserve the constant phase difference $\phi$, the two dots should be coupled to the same macroscopic electron reservoir. The large physical space and the complex inter-dot interaction thus require a highly efficient theoretical approach to address the nonequilibrium dynamic process in an Aharonov–Bohm DQD system.

Various theoretical methods have been developed for the characterization of real-time dynamics of quantum dot (QD) systems. These include the time-dependent density-matrix renormalization group method,15–18 the quantum Monte Carlo approach,19,20 the iterative summation of real-time path integrals,21–23 the time-dependent Green’s function approach,24–28 and the quantum master equations.10 Despite the progress, only a few works10,11 have focused on the Aharonov–Bohm interferometers.

In the past few years, we have developed an accurate and universal hierarchical equations of motion (HEOM) approach for general open electronic systems.29 It is based on a formally exact quantum dissipation theory,30–37 and can be derived via the Feynman–Vernon path-integral formalism38,39 or by using a dissipation picture.40,41 The HEOM approach is capable of addressing the combined effects of system-reservoir dissipation, non-Markovian memory, and electron-electron interactions in a nonperturbative manner. It has been applied to investigate a wide range of equilibrium and nonequilibrium, static and dynamic properties of QD systems.42–52

In the framework of HEOM, the QDs form the system of primary interest, and the electron reservoirs are taken as the environment. The basic variables to solve are the system reduced density matrix and a hierarchical set of auxiliary density operators (ADOs). Each ADO has the same dimension as the system Hilbert space. The computational cost is thus determined by the system degrees of freedom and the total number of ADOs. The latter is controlled by the truncation tier $L$, and the number of basis functions used to expand the environment memory kernel. Conventionally, a simple truncation scheme is adopted, that is to set all the $(L + 1)$th tier ADOs to zero. The numerical results are guaranteed to be

\[ \text{xz58@ustc.edu.cn} \]
\[ \text{yyan@ust.hk} \]
quantitatively accurate as long as they converge with respect to the increase of L.

In the present implementation of the HEOM approach, all the ADOs are stored in memory. Usually, a much larger amount of memory is required to store and operate the ADOs for systems consisting of multiple dots (or orbitals), or at a lower temperature T. This is because the size of each ADO increases as $4^{N_{d}}\times 4^{N_{\mu}}$, with $N_{d}$ being the number of dots (orbitals); while at a lower T, a higher truncation tier L is needed to ensure the numerical convergence, and more basis functions are required to accurately reproduce the reservoir memory kernel. It is highly desirable to develop algorithms and techniques to improve the efficiency of HEOM.

In this paper, we present a series of novel techniques which substantially enhance the efficiency and applicability of HEOM: (1) we propose an a priori estimate for the optimal number of memory basis functions at a given temperature $T$. (2) we design a new memory storage scheme which utilizes the sparsity feature of the ADOs. (3) a new truncation scheme is developed, which makes use of the time derivatives of the $(L + 1)$th tier ADOs. These advancements greatly reduce the memory cost by as much as one order of magnitude or more. The improved HEOM approach allows for an efficient treatment of the Aharonov–Bohm DQD systems.

The remainder of this paper is organized as follows. Section II presents the theoretical advancements which improve the efficiency of HEOM. In Sec. III, the HEOM approach is employed to study the coherent dynamics of Aharonov–Bohm DQD systems. Concluding remarks are finally given in Sec. IV.

II. IMPROVING THE EFFICIENCY OF THE HEOM APPROACH

A. The HEOM formalism and its basic features

Consider a general QD system described by the Hamiltonian of

$$H_{\text{total}} = H_{\text{dot}} + H_{\text{res}} + H_{\text{dot-res}}.$$

Here, $H_{\text{dot}}$ represents the QD of primary interest, $H_{\text{res}} = \sum_{\alpha,k} \epsilon_{\alpha} a_{\alpha,k}^\dagger a_{\alpha,k}$ describes the noninteracting electron reservoirs, and $H_{\text{dot-res}} = \sum_{\alpha,k} \epsilon_{\alpha} t_{\alpha,k} a_{\alpha,k}^\dagger + H.c.$ accounts for the dot-reservoir couplings. In these terms, $a_{\alpha,k}^\dagger$ (or $a_{\alpha,k}$) creates (annihilates) an electron of spin-$\sigma$ on the dot orbital $|\mu\rangle$, and $\hat{d}_{\alpha,k}^\dagger$ ($\hat{d}_{\alpha,k}$) creates (annihilates) an electron on the $\alpha$-reservoir orbital $|k\rangle$ of energy $\epsilon_{\alpha,k}$.

In the framework of HEOM, the influence of electron reservoir is manifested through the reservoir spectral (or hybridization) functions $J_{\alpha\mu\nu}(\omega) \equiv \pi \sum_{k} t_{\alpha,k} t_{\mu,k}^\dagger \delta(\omega - \epsilon_{\alpha,k})$. For numerical convenience, we adopt a Lorentzian form of spectral functions $J_{\alpha\mu\nu}(\omega) = \Delta_{\alpha\mu\nu} \frac{W^2}{(\omega - \mu_{\alpha})^2 + W^2}$. Here, $W$ is the reservoir bandwidth, $\mu_{\alpha}$ is the chemical potential of $\alpha$-reservoir, and $\Delta_{\alpha\mu\nu} = \sqrt{T_{\alpha\mu} T_{\alpha\nu}}$, with $T_{\alpha\mu}$ being the coupling strength between dot orbital $|\mu\rangle$ and $\alpha$-reservoir.

The detailed derivation of the HEOM formalism has been given in Refs. 29, 44, and 53. Here, we emphasize some of its important features. The dot-reservoir coupling Hamiltonian $H_{\text{dot-res}}$ is accounted for exclusively through the reservoir correlation functions $C_{\alpha\mu\nu}(t,\tau)$

$$C_{\alpha\mu\nu}(t,\tau) = \exp \left\{ \beta i \int_{\tau}^{t} dt' V_{\alpha}(t') \right\} \tilde{C}_{\alpha\mu\nu}(t - \tau),$$

$$\tilde{C}_{\alpha\mu\nu}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} du e^{\sigma i u t} f_{\beta}(\omega) J_{\alpha\mu\nu}(\omega).$$

Here, $\tilde{C}_{\alpha\mu\nu}(t - \tau)$ are the reservoir correlation functions at equilibrium; $V_{\alpha}(t)$ is the time-dependent potential applied on the $\alpha$-reservoir; $J_{\alpha\mu\nu}(\omega) = J_{\nu\alpha}(\omega) = J_{\mu\alpha}(\omega)$; $J_{\alpha\mu}(\omega) = 1/[1 + e^{\sigma(\omega - \mu_{\alpha})}]$ is the Fermi function for electron ($\sigma = +$) or hole ($\sigma = -$), $\beta = 1/k_{B}T$, $\sigma = -\sigma$, and $\mu_{\alpha}$ is the equilibrium chemical potential.

The central step for constructing the HEOM is the decomposition of reservoir memory kernel $C_{\alpha\mu\nu}(t)$ by exponential basis functions, i.e.,

$$\tilde{C}_{\alpha\mu\nu}(t) = \sum_{m = 1}^{M} \eta_{\alpha\mu\nu m} e^{-\gamma_{\alpha\mu\nu m} t},$$

where $1/\text{Re}[\gamma_{\alpha\mu\nu m}]$ is the characteristic time of $m$th dissipation mode. A number of memory decomposition schemes have been proposed, such as the Matsubara spectrum decomposition (MSD) scheme,29 the hybrid spectrum decomposition and frequency dispersion scheme,42 the partial fraction decomposition scheme,34 and the Padé spectrum decomposition scheme.55–57 To the best of our knowledge, the Padé decomposition scheme is the most efficient choice among all the existing schemes.

In the Padé decomposition scheme, the Fermi function is expanded by a sum over poles

$$f_{\beta}(z) \approx \tilde{f}_{\beta}(z) = \frac{1}{\sigma} \sum_{p \in P} \left( \frac{\eta_{p}}{z + \eta_{p}} - \frac{\eta_{p}}{z - \eta_{p}} \right).$$

This expression is similar to that in the MSD scheme. The difference is in the Padé decomposition scheme, the coefficient for each Padé term $\eta_{p} \neq 1.57$ The number of Padé terms $P$ determines the accuracy of the reservoir memory decomposition. In principle, the expansion of Eq. (5) becomes exact when $P \to \infty$. In practice, a larger $P$ results in a more accurate expansion, which however leads to a more complex hierarchy. At a lower temperature $T$, a much larger number of Padé terms are needed to attain a given accuracy for the expansion.

Conventionally, the optimal $P$ required is evaluated by comparing the resulting reservoir correlation function with the exact values. Such a trial-and-error process can be quite tedious. Therefore, it would be helpful to have an a priori estimate of $P$ for a more convenient implementation of the HEOM approach.

With the exponential decomposition of reservoir memory, the HEOM are constructed and cast into a compact form as follows:29

$$\rho_{j_{1}j_{2}..j_{n}}^{(n)} = \sum_{r = 1}^{n} \gamma_{j}^{(n)} \mathcal{A}_{j} \rho_{j_{1}..j_{r}j_{r+1}..j_{n}}^{(n-1)} - \sum_{r = 1}^{n} \sum_{s = 1}^{n} (-1)^{r-s} C_{j} \rho_{j_{1}..j_{r}j_{r+1}..j_{s}j_{s+1}..j_{n}}^{(n-1)}.$$
\( \equiv C^{\alpha \mu \nu \mu \nu}_{\alpha \mu \nu} \) are defined via their actions on a fermionic/bosonic operator \( \hat{O} \) as \( \hat{\sigma}^{\mu}_{\alpha} \hat{O} = \hat{\sigma}^{\mu}_{\alpha} \hat{O} \) and \( C^{\alpha \mu \nu \mu \nu}_{\alpha \mu \nu} \hat{O} = \hat{\sigma}^{\mu}_{\alpha} (\hat{\sigma}^{\mu}_{\alpha})^{\dagger} \), respectively. The electron-electron interactions are covered by the Liouvillian \( \hat{L} \cdot \equiv \{ H_{\text{dot}}, \cdot \} \).

The number of nth tier ADOs is roughly \( \frac{1}{2} \frac{(n+K-1)!}{n!(K-1)!} \) with \( K = 4M N_{p} N_{r}. \) Here, \( M = P + 1, N_{p} \) is the number of reservoirs, and \( N_{r} \) is the number of QD orbitals that are coupled to the reservoirs. The total number of ADOs to solve is thus \( N(K,L) = \frac{1}{2} \sum_{n=1}^{K-1} \frac{(n+K-1)!}{n!(K-1)!} \). Currently, the major challenge of the HEOM approach is the rapidly increasing memory space for storing and computing these ADOs. Reducing such a memory cost will dramatically enhance the efficiency of the HEOM approach.

### B. A priori determination of the optimal number of Padé terms for a given temperature

In the Padé memory decomposition scheme of Eq. (5), the optimal number of Padé terms \( P \) varies sensitively with the temperature \( T \).

To establish a quantitative relation between the optimal \( P \) and the temperature \( T \), we define the relative error of memory decomposition as follows (the indices \( \alpha \mu \nu \) are suppressed):

\[
E_{c} = \frac{|C^{\sigma}(t = 0) - \tilde{C}^{\sigma}(t = 0)|}{|C^{\sigma}(t = 0)|} \approx \frac{\int_{-\infty}^{\infty} d\omega [f^{\sigma}_{\beta}^{\dagger}(\omega) - f^{\sigma}_{\beta}(\omega)] f^{\sigma}(\omega)}{\int_{-\infty}^{\infty} d\omega f^{\sigma}_{\beta}(\omega) f^{\sigma}(\omega)}. \quad (7)
\]

Here, \( \tilde{C}^{\sigma}(t) \) is the exact reservoir correlation function defined by Eq. (3), and \( C^{\sigma}(t) \) is evaluated by replacing \( f^{\sigma}_{\beta}(\omega) \) with the Padé approximated \( \tilde{f}^{\sigma}_{\beta}(\omega) \) of Eq. (5).

Figure 1 plots the minimal \( P \) that gives an error less than the preset \( E_{c} \) for different temperatures. Apparently, a larger \( P \) is required for a lower \( T \) or to reach a smaller error \( E_{c} \). It is found that the data fit remarkably well to a linear relation \( \ln P = a + b \cdot \ln(T/W) \). Here, the reservoir bandwidth \( W \) is taken as a reference energy scale. The parameter \( a \) depends on the preset \( E_{c} \) via \( a = 1.10 \ln(-0.28 \ln E_{c}) \), while \( b \) is almost a constant in all cases; see Fig. 1. This leads to a universal quantitative relation of

\[
\ln P = 1.10 \ln(-0.28 \ln E_{c}) - 0.50 \ln(T/W). \quad (8)
\]

Equation (8) is most reliable at \( T/W < 0.004 \). It provides an a priori estimate of \( P \), so that the somewhat tedious trial-and-error search process can be avoided. Such a quantitative relation helps to reach an optimal balance between the accuracy and efficiency of the HEOM.

### C. Reducing the cost of memory by utilizing the sparsity feature of ADOs

In the present implementation of HEOM, all matrix elements of all the ADOs are stored in physical memory. Consequently, the memory cost increases rapidly with the number of dot orbitals \( N_{p} \) or as the temperature lowers. It is important to note that a large percentage of the ADO elements are exactly zero due to physical constraints. This means that the ADOs are sparse matrices. The sparsity feature is completely determined by the interactions among electrons on different dot orbitals. In the HEOM approach, this involves the system Hamiltonian \( H_{\text{dot}} \) and the reservoir spectral functions \( J_{\alpha \mu \nu}(\omega) \).

Take a QD system consisting of only one orbital as an example. Such a system is often described by the single-impurity Anderson model (SIAM)\(^{58,59}\) with the dot Hamiltonian being \( H_{\text{dot}} = \epsilon(\hat{n}_{\uparrow} + \hat{n}_{\downarrow}) + U \hat{n}_{\uparrow} \hat{n}_{\downarrow} \). Here, \( \hat{n}_{\uparrow}\) is the on-dot electron-electron Coulomb interaction energy. The physical space of the QD is spanned by four Fock states, \( |0\rangle, |\uparrow\rangle, |\downarrow\rangle, \) and \( |\uparrow\downarrow\rangle \). Usually, the direct spin-flip term \( \hat{d}_{\uparrow}^{\dagger} \hat{d}_{\downarrow} \) or its Hermitian conjugate is absent from \( H_{\text{dot}} \). Consequently, there is no coherence between the two states \( |\uparrow\rangle \) and \( |\downarrow\rangle \). Therefore, the reduced density matrix element \( \rho_{\uparrow \downarrow} \) is exactly zero. Similarly, the coherence terms between states with different particle numbers are zero since the number of particles is conserved by the Hamiltonian. Their contributions to all the associated first tier ADOs are also zero.

In many cases, up to 90% or even more ADO elements are zero due to the intrinsic system configuration. Therefore, utilizing the sparsity of the ADOs will lead to a substantial reduction in the memory cost of HEOM. The challenge is that the zero elements locate at different positions of different ADOs. An efficient algorithm is needed to quickly identify and screen out these zero elements. We propose here an algorithm which consists of two ansatzs.

1. The sparsity of the system reduced density matrix \( \rho = \rho^{(0)} \) is determined by an effective Hamiltonian as \( \rho^{(0)} = e^{-\beta \hat{H}_{\text{eff}}} \) with \( \hat{H}_{\text{eff}} = H_{\text{dot}} + H_{\text{ext}} + H_{Z} \). Here, \( H_{\text{dot}} \) is the same dot Hamiltonian as that in Eq. (1); \( H_{\text{ext}} \) represents the interaction between an external field and the dot, which may switch on coupling between two dot orbitals; \( H_{Z} \) originates from the reservoir spectral functions, since an off-diagonal \( J_{\alpha \mu \nu}(\omega) \) may introduce a nonzero coupling between dot orbitals \( |\mu\rangle \) and \( |\nu\rangle \) mediated by a-reservoir. Note that an off-diagonal \( J_{\alpha \mu \nu}(\omega) \) leads to the off-diagonal reservoir correlation function \( \tilde{C}^{\alpha \mu \nu}_{\sigma}(t) \) via Eq. (3), and \( \tilde{C}^{\alpha \mu \nu}_{\sigma}(t) \) is equivalent to the “embedding” self-energies...
in the nonequilibrium Green’s function theory, $\tilde{C}^{\pm}_{\mu\nu}(t) = i[I_{\mu\nu}(t)]^{\pm}$ and $\tilde{C}^{-}_{\mu\nu}(t) = iI_{\mu\nu}(t)$.\(^{29}\) For any pair of $\mu$ and $\nu$ that gives $J_{\mu\nu}(\omega) \neq 0$, we add a nonzero term $(I_{\mu\nu}d^{\alpha}_{\mu}e_{\alpha} + H.c.)$ to $H_{S}$. In this way, the reservoir mediated couplings among different dot orbitals are properly accounted for.

(ii) The sparsity of any $n$th tier ADO $\rho^{(n)}_{j_{1}\cdots j_{n}}$ is completely determined by $H_{S}$ and all $(n - 1)$th tier ADOs via $[H_{S}, \rho^{(n)}_{j_{1}\cdots j_{n}}] + \sum_{j=1}^{L} CJ_{j}\rho^{(n-1)}_{j_{1}\cdots j_{n}}$. Therefore, by scanning through the whole hierarchy from zeroth tier to the terminal $L$th tier, the sparsity of all the ADOs can be deduced.

Although not rigorously proved, the above two ansatzs have been validated by our extensive numerical tests on various types of QDs with diversified $H_{S}$ and $J_{\mu\nu}(\omega)$.

Hereafter, we will denote the original implementation of HEOM as the “standard mode,” in which all the ADOs are stored in memory. An alternative and obviously more efficient implementation is to store and compute only those nonzero elements of the ADOs. This will be referred to as the “sparse mode.” Usually, the memory space required by the “sparse mode” is only $5\%$–$10\%$ of that needed in the “standard mode,” while the numerical outcomes are exactly the same. Therefore, the use of the “sparse mode” dramatically reduces the memory cost, and hence substantially enhances the efficiency of the HEOM approach.

D. Truncating the hierarchy using time derivatives of $(L + 1)$th tier ADOs

There are many ways to close a quantum master equation, such as by using a partial ordering prescription, a chronological ordering prescription,\(^{60}\) or a self-consistent Born approximation.\(^{61,62}\) The conventional way to truncate the HEOM is to set all the ADOs beyond $L$th tier to zero, i.e., $\{\rho^{(n)}_{j_{1}\cdots j_{n}} = 0\}$. In such a scheme, the influence of all the $(L + 1)$th tier ADOs is entirely neglected. It is ideal to include some influence of the higher-tier ADOs, without having to compute their values explicitly.

To this end, we write down the EOM for the $(L + 1)$th tier ADO $\rho^{(L+1)}_{j_{1}\cdots j_{L+1}}$ as follows:

$$
\rho^{(L+1)}_{j_{1}\cdots j_{L+1}} = -(i\mathcal{L} + \sum_{r=1}^{L+1} \gamma_{jr})\rho^{(L+1)}_{j_{1}\cdots j_{L+1}} - i\sum_{j} \mathcal{A}_{j}\rho^{(L+2)}_{j_{1}\cdots j_{L+2}} \tag{9}
$$

$$
- i\sum_{r=1}^{L+1} (-)^{L+1-r} CJ_{j} \rho^{(L)}_{j_{1}\cdots j_{r-1}j_{r+1}\cdots j_{L+1}}.
$$

Now, instead of having $\rho^{(L+1)}_{j_{1}\cdots j_{L+1}} = 0$, we set the time derivative of $(L + 1)$th tier ADO to be zero, i.e., $\dot{\rho}^{(L+1)}_{j_{1}\cdots j_{L+1}} = 0$. Under the presumption that all $(L + 2)$th tier ADOs are zero, the value of $\dot{\rho}^{(L+1)}_{j_{1}\cdots j_{L+1}}$ can be estimated by

$$
\dot{\rho}^{(L+1)}_{j_{1}\cdots j_{L+1}} \approx -i\left(\mathcal{L} + \sum_{r=1}^{L+1} \gamma_{jr}\right)^{-1} \cdot \left[\sum_{r=1}^{L+1} (-)^{L+1-r} CJ_{j} \rho^{(L)}_{j_{1}\cdots j_{r-1}j_{r+1}\cdots j_{L+1}}\right]. \tag{10}
$$

The expression on the right-hand side of Eq. (10) is then adopted in the EOM for $\dot{\rho}^{(L+1)}_{j_{1}\cdots j_{L+1}}$, which leads to a new truncation scheme for the hierarchy. The corresponding implementation of HEOM is referred to as the “derivative mode.” With the new scheme, the influence of higher-tier ADOs is partially included in the terminal tier of hierarchy. Consequently, the HEOM in the “derivative mode” is expected to converge more quickly with $L$ than the “standard mode.”

The new truncation scheme does not require more memory space for storing the ADOs, since no extra variables are invoked explicitly. However, the computational complexity does increase. In particular, the Liouville matrix inversion in Eq. (10) needs to be carried out for every $\rho^{(L+1)}_{j_{1}\cdots j_{L+1}}$, which is rather time-consuming. A more economic way of performing the matrix inversion is via

$$
\left(i\mathcal{L} + \gamma_{j}\right)^{-1} = -i\mathcal{X} (\mathcal{L}_{d} - i\gamma_{j})^{-1} \mathcal{X}'. \tag{11}
$$

Here, $\gamma_{j} = \sum_{r=1}^{L+1} \gamma_{jr}$ is different for each $\rho^{(L+1)}_{j_{1}\cdots j_{L+1}}$; $\mathcal{X}$ is the eigenvector matrix of $\mathcal{L}$, and $\mathcal{L}_{d}$ is the diagonal eigenvalue matrix. Since both $\mathcal{X}$ and $\mathcal{L}_{d}$ are independent of subscript $j$, the diagonalization of $\mathcal{L}$ needs to be done only once.

In the Liouville space, the terms wrapped in the square bracket of Eq. (10) form a vector $y^{(L)}$. Equation (10) is recast into a compact form of $\dot{\rho}^{(L+1)}_{j_{1}\cdots j_{L+1}} = (i\mathcal{L} + \gamma_{j})^{-1} y^{(L)}$, and its right-hand side is evaluated as follows:

1. Perform the Liouville-space matrix-vector multiplication of $u = \mathcal{X}' y^{(L)}$.
2. Scale each element of the Liouville-space vector $u$ by the corresponding diagonal element of $(\mathcal{L}_{d} - i\gamma_{j})^{-1}$. The resulting scaled vector is denoted as $\tilde{u}$.
3. Perform the Liouville-space matrix-vector multiplication of $-i\mathcal{X}\tilde{u}$ to obtain $\dot{\rho}^{(L+1)}_{j_{1}\cdots j_{L+1}}$.

Being truncated at the same $L$th tier, the HEOM in the “derivative mode” require the same amount of memory for storing the ADOs (the physical memory) as in the “standard mode.” However, the new truncation scheme requires more memory for recording the connectivity among all the ADOs (the indexing memory). Note that the cost on indexing memory can be exempted by employing an on-the-fly indexing algorithm, with which the precise location of any arbitrary ADO is determined by an analytic formula. Such an algorithm has been developed for bosonic systems,\(^{63}\) and its extension to fermionic systems is possible.

E. Numerical verification

We have introduced the “sparse mode” and “derivative mode” of HEOM in Secs. II C and II D, respectively. These two techniques can be combined together, leading to the “derivative + sparse mode.” It should be emphasized that these new modes yield exactly the same numerical results as the “standard mode,” provided that the convergence with respect to the truncation tier $L$ is reached. In the following, we present a benchmark for the performances of all the four modes of HEOM.

The benchmark is performed on a SIAM with the system Hamiltonian of $H_{\text{dot}} = e(\hat{n}_{+} + \hat{n}_{1}) + U\hat{n}_{1}\hat{n}_{+}$. The reservoir
spectral function is $J_\alpha(\omega) = \frac{\Delta}{2} \frac{W^2}{(\omega - \mu_\alpha)^2 + \omega^2}$. In the following, the hybridization strength $\Delta$ is taken as an energy unit. The parameters $\epsilon = -U/2 = -6\Delta$ are adopted, so that the dot is exactly half-filling, and the reservoir bandwidth is $W = 20\Delta$. The system is in a nonequilibrium steady state under a bias voltage $V = 2\Delta$, and the temperature is $T = 0.1\Delta$. To reach a high truncation tier, we adopt a slightly larger error tolerance for the reservoir memory decomposition, $E_c = 7.5\%$. From Eq. (8), we immediately have $P = 10$.

Table I compares the CPU time, the physical memory, and the indexing memory consumed by HEOM calculations carried out in the four different modes. The steady-state current $I$ and the diagonal elements of $\rho$ obtained at different truncation tiers are also listed.

In terms of efficiency, the “sparse mode” indeed reduces the physical memory by more than one order of magnitude. The CPU time is also greatly shortened, with the zero elements of ADOs screened out of the hierarchy. As indicated by the calculated $I$ and $\rho$, the $L$th tier truncation in the “derivative mode” amounts to the $(L + 1)$th tier truncation in the “standard mode.” This means that the “derivative mode” allows for accessing one tier up beyond the terminal tier without increasing the cost of physical memory. This is because the $(L + 1)$th ADOs are treated on-the-fly by Eq. (10). The price to pay is the somewhat enlarged indexing memory. Nevertheless, the “derivative mode” proves to be very useful, since often the saving of physical memory exceeds the increment in the indexing memory; see the example in Sec. III.

Among all the four modes, the “sparse mode” has the overall best balance between efficiency and accuracy. If an on-the-fly indexing algorithm could be developed for fermionic systems, the cost of indexing memory would become trivial, and the “derivative + sparse mode” would be even more appealing.

In terms of accuracy, all the four modes converge to the same results as $L$ increases. This highlights the important fact that, while significantly enhancing the numerical efficiency, the newly developed techniques preserve the accuracy of the HEOM approach.

### TABLE I

The upper table lists the CPU time (in unit of second), the physical memory, and indexing memory (in unit of byte) consumed by HEOM calculations in four different modes. The lower table shows the calculated steady-state current $I$, and the diagonal elements $\rho_{11}$ and $\rho_{22}$ of the steady-state reduced density matrix. The four columns in each cell correspond to the truncation tier $L = 2, 3, 4, 5$ for the standard and sparse modes, and $L = 1, 2, 3, 4$ for the derivative and derivative + sparse modes, respectively. “N/A” means the results are unavailable because the computational cost exceeds the resources at our disposal. The system is a SIAM with the parameters (in units of $\Delta$): $\epsilon = -U/2 = -6$, $W = 20$, $T = 0.1$, and $V = 2$.

<table>
<thead>
<tr>
<th>Mode</th>
<th>CPU time (s)</th>
<th>Physical memory (byte)</th>
<th>Indexing memory (byte)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>1.6</td>
<td>38</td>
<td>0.9 k</td>
</tr>
<tr>
<td>Sparse</td>
<td>1.2</td>
<td>13</td>
<td>0.3 k</td>
</tr>
<tr>
<td>Derivative</td>
<td>1.4</td>
<td>47</td>
<td>2.0 k</td>
</tr>
<tr>
<td>Derivative + sparse</td>
<td>1.4</td>
<td>33</td>
<td>1.1 k</td>
</tr>
<tr>
<td>Standard</td>
<td>1.6</td>
<td>38</td>
<td>0.9 k</td>
</tr>
<tr>
<td>Sparse</td>
<td>1.2</td>
<td>13</td>
<td>0.3 k</td>
</tr>
<tr>
<td>Derivative</td>
<td>1.4</td>
<td>47</td>
<td>2.0 k</td>
</tr>
<tr>
<td>Derivative + sparse</td>
<td>1.4</td>
<td>33</td>
<td>1.1 k</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mode</th>
<th>Current $I$ (eA/h×10^4)</th>
<th>$\rho_{11}$</th>
<th>$\rho_{22}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>4.956</td>
<td>6.980</td>
<td>7.725</td>
</tr>
<tr>
<td>Sparse</td>
<td>4.956</td>
<td>7.765</td>
<td>7.765</td>
</tr>
<tr>
<td>Derivative</td>
<td>4.956</td>
<td>7.765</td>
<td>7.765</td>
</tr>
<tr>
<td>Derivative + sparse</td>
<td>4.956</td>
<td>7.765</td>
<td>7.765</td>
</tr>
</tbody>
</table>

### III. THE COHERENT DYNAMICS IN AHARONOV–BOHM INTERFEROMETERS

We now study an Aharonov–Bohm interferometer consisting of two QDs coupled in parallel to two electron reservoirs. The system setup is sketched in Fig. 2. The system Hamiltonian is

$$H_{\text{dot}} = \sum_{\mu = 1, 2} \epsilon(\hat{n}_{\mu \uparrow} + \hat{n}_{\mu \downarrow}) + U\hat{n}_{\mu}[\hat{n}_{\mu}]_\downarrow. \quad (12)$$

The system-reservoir coupling Hamiltonian is

$$H_{\text{dot-res}} = \sum_{\alpha \neq \mu} \Gamma_{\alpha \mu} e^{i\phi_{\mu\alpha}} \hat{a}_{\alpha \uparrow}^\dagger \hat{a}_{\mu \downarrow} + \text{H.c.}, \quad (13)$$

where $\phi_{\mu\alpha}$ is the phase factor associated with the coupling between dot orbital $|\mu\rangle$ and $\alpha$-reservoir. The phase factors originate from an applied magnetic flux $\Phi$. Under a bias voltage, the electrons conducting through the two dots interfere with each other by a constant phase $\Phi \equiv 2\pi\Phi/\Phi_0 = \phi_{1L} - \phi_{2L} - \phi_{1R} + \phi_{2R}$, with $\Phi_0$ being the magnetic flux quantum.

In the framework of HEOM, $H_{\text{dot-res}}$ is accounted for by the reservoir spectral functions, which are cast into the following matrix form:

$$J_\alpha(\omega) = \begin{pmatrix} 1 & e^{i(\phi_{1\alpha} - \phi_{2\alpha})} \frac{\Delta}{2} \frac{W^2}{(\omega - \mu_{\alpha})^2 + W^2} \end{pmatrix}. \quad (14)$$

We set $\phi_{1L} - \phi_{2L} = \phi_{2R} - \phi_{1R} = \phi/2$.

![FIG. 2. Schematic diagram of an Aharonov–Bohm DQD system. The two dots are labeled by $\mu$ ($\mu = 1, 2$), and the two reservoirs are denoted by $\alpha$ ($\alpha = L, R$), respectively. The two dots have the same orbital energy $\epsilon$ and on-dot Coulomb interaction energy $U$. A bias voltage $V_L = -V_R = -V/2$ is applied across the reservoirs. A phase factor $\phi$ is associated with the dot-reservoir coupling strengths $\{\Gamma_{\alpha \mu}\}$; see the text for details.](image-url)
Although the direct inter-dot hopping and Coulomb interaction terms are absent in $H_{\text{dot}}$ of Eq. (12), the two dots are effectively coupled to each other via the nonzero off-diagonal elements of $J_{\text{dd}}(\omega)$; see Eq. (14). Clearly, the reservoir-mediated inter-dot couplings are essentially important for the occurrence of quantum interference. They give rise to the Ruderman–Kittel–Kasuya–Yosida (RKKY) type of exchange interaction, which is the physical origin of the quantum coherence and quantum entanglement between the two dots.

Controlling the dynamical quantum coherence (or decoherence) of QDs is of fundamental importance for potential application in quantum information processing and quantum computing. Experimental realization is possible by tuning the system parameters. It is however difficult to simulate accurately the real-time coherent electronic dynamics in QD systems, particularly when strong electron correlation effects are present.

Consider first the case of $\epsilon = U = 0$, and the spin-up and spin-down orbitals are degenerate. Such an Aharonov–Bohm DQD has been studied by Tu, Zhang, and Jin. In this case, the spin degrees of freedom can be averaged out, and the two dots are characterized by their charge states. The dynamic quantum coherence between these charge states is characterized by the off-diagonal element of the reduced density matrix $\rho_{12}(t)$. Here, $|1\rangle$ and $|2\rangle$ are the singly occupied states of dot 1 and 2, respectively.

Initially the DQD is empty, and a bias voltage of $V = 6\Delta$ is switched on at $t = 0$. The time evolution of $\rho_{12}(t)$ has been obtained by solving an exact quantum master equation. We apply the improved HEOM approach developed in this work to calculate $\rho_{12}(t)$. As shown in Fig. 3, the HEOM results accurately reproduce the data of Ref. 10 for various values of $\phi$. Note that for a noninteracting DQD system ($U = 0$), the hierarchy terminates automatically and exactly at $L = 2$. Therefore, even in the “standard mode,” the cost of physical memory is as small as 13.6 MB. The use of “sparse mode” further reduces the physical memory to only 3.2 MB.

Figure 4 depicts the HEOM results for $\phi = \pi/4$ calculated with the four modes presented in Sec. II. Clearly, the four lines overlap exactly with each other. This affirms that the four modes of HEOM become equivalent when the convergence with respect to $L$ is achieved.

We proceed to investigate an interacting Aharonov–Bohm DQD with a nonzero on-dot electron-electron interaction energy ($U > 0$). Specifically, we choose the parameters $\epsilon = -U/2 = -\Delta$. Such a DQD has been studied by Büscher and Heidrich-Meisner with a time-dependent density matrix renormalization group method. In the following, we demonstrate that the coherent dynamics of an interacting Aharonov–Bohm DQD can be obtained accurately with the HEOM approach.

For the interacting DQD, the spin degrees of freedom are explicitly treated. Initially the DQD is in the thermal equilibrium state at the temperature $T = 0.25\Delta$. A bias voltage of $V = 2\Delta$ is switched on at $t = 0$, which drives the DQD out of equilibrium. In this case, the dynamical spin correlation between the two dots and its dependence on the phase $\phi$ are of particular importance, since the spin correlation is directly related to the emergence of quantum entanglement.

The spin operator associated with the $\mu$th dot is constructed as $\hat{S}_{\mu} = \frac{1}{2} \sum_{\sigma} \hat{a}_{\mu}^{\dagger} \sigma_{\sigma \sigma} \hat{a}_{\mu\sigma}$, where $\sigma$ represents the vector of Pauli matrices. For the DQD, the spin correlation between the two dots is evaluated by $S_{12} = \langle \hat{S}_1 \cdot \hat{S}_2 \rangle$. The

![Figure 3](image1.png)

**FIG. 3.** (a) Real and (b) imaginary parts of $\rho_{12}(t)$ for different values of $\phi$. The system parameters are (in units of $\Delta$): $\epsilon = U = 0$, $T = 0.05$, and $V = 6$. The dotted lines are extracted from Fig. 2 of Ref. 10. The circles are the data calculated by the HEOM approach in “sparse mode” at $L = 2$.

![Figure 4](image2.png)

**FIG. 4.** Time evolution of the real and imaginary (inset) parts of $\rho_{12}$ for the phase $\phi = \pi/4$. The results obtained by all four modes of HEOM are displayed. The terminal tier is $L = 2$ for the “standard” and “sparse” modes, and $L = 1$ for the “derivative” and “derivative + sparse” modes. The parameters are (in units of $\Delta$): $\epsilon = U = 0$, $T = 0.05$, and $V = 6$. The
quantum entanglement of the two dots is usually characterized by the concurrence $C$, which is determined from the reduced density matrix $\rho$ via

$$C_{12} = \max(0, \sqrt{\lambda_1} - \sqrt{\lambda_2} - \sqrt{\lambda_3} - \sqrt{\lambda_4}).$$

(15)

Here, $\{\lambda_1, \ldots, \lambda_4\}$ are the decreasing-ordered eigenvalues of the Hermitian operator, $R = \rho_1 \tilde{\rho}_2$, $\rho_1$ is a $4 \times 4$ submatrix of $\rho$, which corresponds to the Fock states $|\uparrow\uparrow\rangle$, $|\downarrow\downarrow\rangle$, and $|\uparrow\downarrow\rangle$; and $\tilde{\rho}_2 = (\sigma_y \otimes \sigma_y) \rho_2^T (\sigma_y \otimes \sigma_y)$. Because there is no direct spin-flip term in the system Hamiltonian of Eq. (12), the concurrence of Eq. (15) can be related to $S_{12}$ and evaluated via

$$C_{12} = \max \left(0, -\frac{1}{2} - 2 \frac{S_{12}}{\langle \hat{N}_1^2 \hat{N}_2^2 \rangle} \right),$$

(16)

with $\hat{N}_\mu = \hat{n}_\mu + \hat{\bar{n}}_\mu - 2 \hat{n}_\mu \hat{\bar{n}}_\mu$. Note that to ensure the same value of $C_{12}$ is obtained by Eqs. (15) and (16), the submatrix $\rho_s$ needs to be normalized by $(\hat{N}_1^2 \hat{N}_2^2)$.70

We apply the HEOM approach to calculate the time evolution of spin correlation $S_{12}$ and concurrence $C_{12}$ of the interacting DQD system. The resulting $S_{12}(t)$ and $C_{12}(t)$ for $\phi = 0$ and $\phi = \pi$ are shown in Fig. 5. With $\phi = 0$, $S_{12}$ is always positive. This indicates that the reservoir-mediated RKKY interaction results in a ferromagnetic interaction between the spins on the two dots. Moreover, $C_{12} = 0$ at any time $t$, and thus the DQD remains in an unentangled state. In contrast, the phase $\phi = \pi$ gives rise to an antiferromagnetic spin correlation ($S_{12} < 0$). Meanwhile, $C_{12}$ becomes positive after some time, and it reaches a positive constant in the long-time limit. This indicates the emergence of entangled state in the DQD. The HEOM calculations also affirm the numerical equivalence between Eqs. (15) and (16), as they both lead to the same $C_{12}(t)$; see the inset of Fig. 5. These calculations accentuate that the phase $\phi$ has critical influence on the quantum coherence/entanglement of an Aharonov–Bohm DQD system.

The HEOM results depicted in Fig. 5 agree consistently with those reported in Ref. 11. We have verified that these HEOM results achieve quantitative accuracy, since they have already converged with respect to the truncation tier $L$.

To reach the truncation tier $L = 4$, the HEOM calculations are carried out in the “sparse mode,” because the “standard mode” is not applicable due to the excessive memory cost. For the interacting DQD, the nonzero ADO elements are only 5% of all unknowns, and thus the cost of physical memory reduces from over 200 GB required for the “standard mode” to 10.4 GB for the “sparse mode.” In order to verify the numerical convergence, the outcomes of $L = 4$ should agree quantitatively with the $L = 5$ counterparts. However, the $L = 5$ calculation is not available in the “sparse mode,” since it requires more than 200 GB of physical memory which is much beyond the computational resources at our disposal. The “derivative + sparse mode” becomes useful in this scenario, since the truncation at $L = 4$ in the “derivative + sparse mode” amounts to the $L = 5$ truncation in the “sparse mode.” Although a larger indexing memory is invoked, the substantial reduction in the required physical memory makes the calculation in “derivative + sparse mode” feasible. Such an example clearly exemplifies the usefulness of the novel techniques developed in Sec. II, which indeed dramatically enhances the efficiency of the HEOM approach.

IV. CONCLUDING REMARKS

To conclude, we develop a series of new techniques to enhance the efficiency of the HEOM approach. First, we propose an a priori estimate for the optimal number of Padé terms for the reservoir memory decomposition. Second, we make use of the sparsity of ADOs and propose two ansatzs to screen out all the intrinsic zero ADO elements. Third, we propose a new truncation scheme by utilizing the time derivatives of higher-tier ADOs. These new advancements (particularly the latter two) reduce the memory cost by more than one order of magnitude or more, and thus greatly enhance the efficiency and applicability of HEOM. It is also important to emphasize that the newly developed techniques do not compromise the accuracy of HEOM at all.

The practicality of the improved HEOM approach is demonstrated through the calculation for the coherent dynamics of the Aharonov–Bohm DQD interferometers. Quantitatively accurate dynamics is obtained for both noninteracting and interacting DQDs. The crucial role of the phase $\phi$ for the magnitude of quantum coherence and quantum entanglement is revealed.

The improved HEOM approach is potentially very useful for simulating the coherent dynamics of charge and spin states in QD systems involving strong electron-electron interactions.

ACKNOWLEDGMENTS

The support from the National Science Foundation of China (Grant Nos. 21233007, 21303175, 21322305, and 21373191), the Strategic Priority Research Program (B) of the CAS (Grant No. XDB01020000) is gratefully appreciated. The computational resources are provided by the Supercomputing Center of University of Science and Technology of China.


