THE JOURNAL OF PHYSICAL CHEMISTRY A
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCA
Article

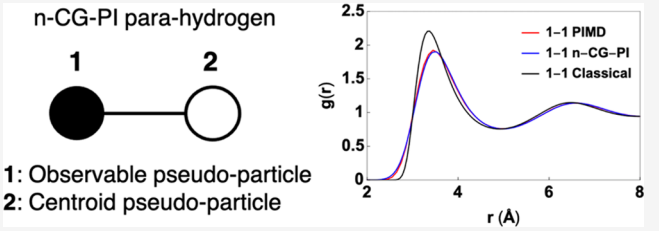
Coarse-Graining of Imaginary Time Feynman Path Integrals: Inclusion of Intramolecular Interactions and Bottom-up Force-Matching

Published as part of The Journal of Physical Chemistry virtual special issue "Jose Onuchic Festschrift".

Won Hee Ryu and Gregory A. Voth*

Cite This: *J. Phys. Chem. A* 2022, 126, 6004−6019
Read Online

ABSTRACT: Feynman's imaginary time path integral formalism of quantum statistical mechanics and the corresponding quantum-classical isomorphism provide a tangible way of incorporating nuclear quantum effect (NQE) in the simulation of condensed matter systems using well-developed classical simulation techniques. Our previous work has presented the many-body coarse-graining of path integral (CG-PI) theory that builds an isomorphism between the quantum partition function of N distinguishable particles and the classical partition function of $2N$ pseudoparticles. In this present work, we develop a generalized version of the many-body CG-PI theory that incorporates many-body interactions in the force field. Based on the new derivation, we provide a numerical CG-PI (n-CG-PI) modeling strategy parametrized from the underlying path integral molecular dynamics (PIMD) trajectories using force matching and Boltzmann inversion. The n-CG-PI models for two liquid systems are shown to capture well both the intramolecular and intermolecular structural correlations of the reference PIMD simulations. The generalized derivation of the many-body CG-PI theory and the n-CG-PI model presented in this work extend the scope of the CG-PI formalism by generalizing the previously limited theory to incorporate force fields of realistic molecular systems.



1. INTRODUCTION

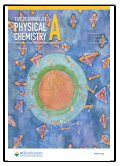
In classical molecular dynamics (MD) simulations, the equations of motion of the nuclei are described by classical mechanics. However, for systems with light nuclei and/or at low temperatures, this approximation breaks down and explicit treatment of the quantum mechanical nature of nuclei becomes necessary to correctly describe the physics of the system. The nuclear quantum effect (NQE) is essential to correctly describe the chemistry and physics of various condensed matter phenomena such as both equilibrium and dynamic quantities of water,[1−10] liquid helium and para-hydrogen,[11−15] hydrogen bonding,[16−18] and proton transport.[19−23]

An appealing method to incorporate the NQE into simulation of condensed matter systems is Feynman's imaginary time path integral formalism of quantum statistical mechanics.[24,25] An outcome of this formalism is a quantum-classical isomorphism, which states that the quantum mechanical partition function of a particle can be expressed as the classical partition function of a discrete representation of the imaginary time path, or a "ring polymer".[26,27] The isomorphism suggests that the quantum mechanical equilibrium quantities can be calculated by sampling a complex yet still classical object in the extended phase space.

Combining the quantum-classical isomorphism with the well-developed classical techniques such as MD or Monte Carlo (MC) simulations, path integral methods such as path integral MD (PIMD)[11,28−31] and path integral MC (PIMC)[11,29,32] have been used to capture the NQEs in various condensed matter systems such as molecular liquids,[33,34] clusters,[35,36] and solvated electrons.[37−39] Additionally, more recently developed path integral methods such as centroid MD (CMD)[40−45] and ring polymer MD (RPMD)[46−51] extend the PIMD method, originally aimed to capture the equilibrium quantities, to a dynamical regime and have been used to calculate the dynamical quantities such as spectra,[52−59] diffusion,[12,60−63] and reaction rates.[64,65]

However, even with the fast sampling algorithms of classical simulations applied to path integral calculations,[66,67] these simulations are more computationally demanding than their classical counterparts due to the large number of "beads" in the

Received: June 22, 2022
Revised: August 16, 2022
Published: August 25, 2022



https://doi.org/10.1021/acs.jpca.2c04349
J. Phys. Chem. A 2022, 126, 6004−6019

ring polymer which needs to be sufficiently large for a converged result, especially so for systems with large NQEs. The extended phase space of the ring polymer presents a significant sampling challenge that needs to be overcome to incorporate NQEs in the simulation of condensed matter systems with manageable computational cost. The challenge of sampling the extended phase space suggests that it would be valuable to construct an alternative isomorphism with inherently smaller number of degrees of freedom (DOF) than that of the full ring polymers, while maintaining the correct description of interaction between the ring polymers.

The approach of bottom-up coarse-graining[68,69] constructs the coarse-grained (CG) representation of the underlying fine-grained (FG) model (all-atom resolution for example) by systematically integrating out certain DOF of the fine-grained models. The CG potential of mean force (PMF) is calculated so that the CG system correctly describes the underlying physics such as structural correlations of the FG systems mapped onto the CG sites. In this spirit, the bottom-up CG method provides a way to construct an alternative isomorphism of the imaginary time paths with systematically less DOF. The coarse-graining of path integral (CG-PI) theory[70,71] combines the bottom-up CG method and the quantum-classical isomorphism to construct a reductionist representation of the off-diagonal thermal density matrix. By introducing the new CG coordinate Q, the center of mass of the intermediate beads in the ring polymer, the CG-PI theory reveals that the quantum partition function of a particle is isomorphic to the classical partition function of two (a pair of) pseudoparticles obeying an effective classical mechanics.

The CG-PI theory is not the first to combine the CG methodology with the imaginary time path integral formalism. Most notably, CMD reformulates the imaginary time path integral using the centroid (center of mass of the ring polymer) as the central CG quantity. The CMD formalism, initially intended for distinguishable particles, has further been extended for systems obeying Fermi−Dirac and Bose−Einstein statistics by mapping the system into the phase space variables.[72−74] The quantum multiscale coarse-graining (qMS-CG)[75] extends the classical MS-CG theory[76−79] to quantum Boltzmann statistics by using the path integral formalism. Several enhanced sampling methods such as replica exchange or metadynamics that combines the CG method and path integral formalism have been developed as well.[80−82]

It should be noted that while both the CMD and the CG-PI theory use CG representations of the ring polymer, there is one significant difference. The centroid used in CMD is the center of mass of the entire ring polymer, whereas the CG coordinate Q in CG-PI is the center of mass of only the intermediate beads between two end point beads of the imaginary path. Using the centroid as the central quantity significantly decreases the DOF but only enables one to calculate certain diagonal elements of the thermal density matrix. However, the mapping used in CG-PI allows explicit calculation of the off-diagonal elements which enables one to calculate expectation values of momentum-dependent operators.

Despite the appealing features of the CG-PI theory, it still lacks both theoretical and numerical formalisms to be used in simulations of realistic condensed matter systems. First, in deriving an approximation for the PMF of the pseudoparticles, the previous version of CG-PI theory assumes that the potential is pairwise decomposable. Even though the pairwise assumption has extensively been used in different simulation methods such as MS-CG to describe the effective intermolecular interactions, it cannot correctly describe the force field of a more general molecular systems with explicit intramolecular terms or three-body and higher interactions. Moreover, the previous numerical implementation of the CG-PI theory (numerical CG-PI, n-CG-PI) required an extensive set of PIMD calculations to construct a scaling relationship and determine the relevant parameters. Even though the n-CG-PI method with parameters determined in this fashion captured well the structural correlations of a simple liquid system, this process is not only computationally demanding but it is also contrary to the purpose of using a CG representation of ring polymers.

In this paper, we overcome the shortcomings of the previous CG-PI theory and n-CG-PI method. In section 2, the exact expression of the CG PMF of the pseudoparticles using a more general form of the force field is Taylor expanded up to second order, assuming that the system is close to the classical limit. The derivation shows that the previously observed quadratic coupling between the different pseudoparticles is preserved even with a more general form of the force field. More specifically, the many-body terms in the force field adds additional terms, functions of the Hessian of the many-body interactions, in the CG PMF. Moreover, based on the Taylor expanded result of the CG PMF, we construct an improved n-CG-PI modeling scheme that is parametrized from a single PIMD trajectory by using bottom-up CG methods such as force matching and Boltzmann inversion. In section 3, we revisit the two liquid systems studied in our previous work. The new n-CG-PI method is shown to capture both the intra- and inter-ring polymer structural correlations. Section 4 concludes the paper by discussing the results and concluding the findings presented.

2. THEORETICAL METHODS

2.1. Many-Body CG-PI Potential and Effective Hamiltonian. Using Feynman's imaginary time formalism of quantum statistical mechanics,[24,25] the thermal density matrix element for N distinguishable particles obeying the quantum Boltzmann statistics in one-dimension with Hamiltonian operator $\hat{H}$ can be written as

$$\langle \boldsymbol{q}|e^{-\beta\hat{H}}|\boldsymbol{q}'\rangle = \lim_{P\to\infty}\left[\left[\prod_{i=1}^{N}\left(\frac{m_i P}{2\pi\hbar^2\beta}\right)^{P/2}\right]\times\int\left[\prod_{i=1}^{N}\prod_{j=2}^{P} dq_j^{(i)}\right]e^{-\beta V_{P,N}}\right] \tag{1}$$

for a Feynman many-body isomorphic potential $V_{(P,N)}$

$$V_{P,N} = \sum_{k=1}^{P}\sum_{i=1}^{N}\left[\frac{m_i P}{2\hbar^2\beta^2}(q_k^{(i)} - q_{k+1}^{(i)})^2 + \frac{1}{P}V(q_k^{(1)}, ..., q_k^{(N)})\right] \tag{2}$$

where $q_j^{(i)}$ is the position of the jth bead in the ith ring polymer, m_i is the mass of the ith particle, and $V(q_1,\cdots,q_N)$ is the many-body potential of the system. The two N dimensional column vectors $\boldsymbol{q}$ and $\boldsymbol{q}'$ have q_i (beginning of the ith path) and q_i' (end of the ith path) as their ith elements, respectively. The boundary condition can be written as $q_1^{(i)} = q_i$ and $q_{P+1}^{(i)} = q_i'$. Equations 1 and 2 represent the quantum-classical iso-

https://doi.org/10.1021/acs.jpca.2c04349
J. Phys. Chem. A 2022, 126, 6004−6019

morphism in which N interacting particles are represented by a set of N ring polymers that obey classical mechanics. There are P coordinates or beads for each ring polymer that experiences the nearest neighbor harmonic coupling with a spring constant of $m_iP/(2\hbar^2\beta^2)$ and the physical potential scaled by $1/P$. The equality in eq 1 is exact for the limit of $P \to \infty$.

The many-body CG-PI theory[71] introduces a CG coordinate Q to construct a coarse-grained expression of the thermal density matrix element $\langle \boldsymbol{q}|e^{-\beta\hat{H}}|\boldsymbol{q}'\rangle$ with a much smaller number of DOF. The newly introduced coordinate Q_i, coarse-grained representation of the intermediate beads of the ith ring polymer, is defined as

$$Q_i = \frac{1}{P-1}(q_2^{(i)} + \ldots + q_P^{(i)}) \tag{3}$$

which is the center of mass of the intermediate beads $(q_2^{(i)},\cdots,q_P^{(i)})$ in the ith particle ring polymer. With the use of the new coordinate introduced in eq 3, the thermal density matrix element can be rewritten as

$$\langle \boldsymbol{q}|e^{-\beta\hat{H}}|\boldsymbol{q}'\rangle = \left(\prod_{i=1}^{N}\frac{m_i}{\pi\hbar^2\beta}\right)\int d\boldsymbol{Q}\; e^{-\beta V_{CG,N}(\boldsymbol{q},\boldsymbol{Q},\boldsymbol{q}')} \tag{4}$$

for N dimensional column vector $\boldsymbol{Q}$ whose ith element is Q_i, and the many-body CG potential $V_{CG,N}$ satisfies

$$V_{CG,N}(\boldsymbol{q},\boldsymbol{Q},\boldsymbol{q}') = -\frac{1}{\beta}\lim_{P\to\infty}[\ln(Z_{CG,N}^{(P)}(\boldsymbol{q},\boldsymbol{Q},\boldsymbol{q}'))] \tag{5}$$

where

$$\begin{aligned} Z_{CG,N}^{(P)}(\boldsymbol{q},\boldsymbol{Q},\boldsymbol{q}') = &\left[\prod_{i=1}^{N}\left(\frac{\pi\hbar^2\beta}{m_i}\right)\left(\frac{m_iP}{2\pi\hbar^2\beta}\right)^{P/2}\right] \\ &\times\int\left[\prod_{i=1}^{N}\prod_{j=2}^{P}dq_j^{(i)}\right]e^{-\beta V_{P,N}} \\ &\left[\prod_{i=1}^{N}\delta\left(Q_i - \frac{1}{P-1}(q_2^{(i)} + \ldots + q_P^{(i)})\right)\right] \end{aligned} \tag{6}$$

A central quantity in CG-PI theory is the effective Hamiltonian $H_{\text{eff},N}$ that provides a more classical-like expression for the expectation value of a many-body operator $\hat{A}(\{\hat{p}_i,\hat{q}_i;i = 1, \cdots, N\})$. The effective Hamiltonian is defined as

$$\langle\hat{A}\rangle = \frac{\int d\boldsymbol{P}_Q\, d\boldsymbol{Q}\, d\boldsymbol{p}\, d\bar{\boldsymbol{q}}\, \exp(-\beta H_{\text{eff},N})W_A(\boldsymbol{p},\bar{\boldsymbol{q}})}{\int d\boldsymbol{P}_Q\, d\boldsymbol{Q}\, d\boldsymbol{p}\, d\bar{\boldsymbol{q}}\, \exp(-\beta H_{\text{eff},N})} \tag{7}$$

for N dimensional column position vectors $\bar{\boldsymbol{q}} = (\boldsymbol{q} + \boldsymbol{q}')/2$, N dimensional column momenta vectors $\boldsymbol{P}_Q$ and $\boldsymbol{p}$ whose ith elements are $P_{Q,i}$ and p_i, which are conjugate momenta of $\bar{q}_i$ and Q_i, respectively. The many-body Weyl map $W_A(\boldsymbol{p},\bar{\boldsymbol{q}})$ is defined as

$$\langle\boldsymbol{q}'|\hat{A}|\boldsymbol{q}\rangle = \int(d\boldsymbol{p}/2\pi\hbar)\exp(i(\boldsymbol{p}\cdot\Delta\boldsymbol{q})/\hbar)W_A(\boldsymbol{p},\bar{\boldsymbol{q}}) \tag{8}$$

Note that the observable is taken with respect to the $\boldsymbol{p}$ and $\bar{\boldsymbol{q}}$ only. The exact expression for the effective Hamiltonian that satisfies eq 7 can be written as

$$\begin{aligned} H_{\text{eff},N}(\boldsymbol{P}_Q,\boldsymbol{p},\boldsymbol{Q},\bar{\boldsymbol{q}}) &= \frac{1}{2}\boldsymbol{P}_Q{}^T\boldsymbol{M}_Q^{-1}\boldsymbol{P}_Q - \frac{1}{\beta}\ln\Big[\text{const}\cdot\int d\Delta\boldsymbol{q} \\ &\exp(i(\boldsymbol{p}\cdot\Delta\boldsymbol{q})/\hbar - \beta V_{CG,N}(\bar{\boldsymbol{q}},\boldsymbol{Q},\Delta\boldsymbol{q}))\Big] \end{aligned} \tag{9}$$

where $\boldsymbol{M}_Q$ is the N dimensional diagonal mass matrix whose ith diagonal element is $M_{Q,i}$, a mass associated with the momenta $P_{Q,i}$, and $\Delta\boldsymbol{q}$ is the N dimensional column vector $\Delta\boldsymbol{q} = \boldsymbol{q}' - \boldsymbol{q}$.

The expression for the many-body CG potential shown in eq 5 provides a reductionist representation of N interacting ring polymers. By averaging out the intermediate bead positions using the mapping shown in eq 3, the off-diagonal thermal density matrix element can be expressed in a compact manner as shown in eq 4 with 3 DOF for each ring polymer rather than P. Moreover, only the intermediate beads are averaged and not the first and the last beads, q_i and q_i', which enables one to explicitly calculate the off-diagonal thermal density matrix elements. This choice of mapping distinguishes the CG-PI theory from other centroid-based formalisms such as CMD that coarse-grain the entire ring polymer into a centroid.

As shown in eq 9, the CG-PI theory combines the quantum-classical isomorphism and CG theory to provide a $2N$ particle representation of a system of N interacting particles obeying quantum Boltzmann distribution. The particles with position, momenta, and mass as described in the effective Hamiltonian are named "pseudo-particles" to distinguish them from quasiparticles that have distinct physical meaning. More specifically, the pseudoparticle with position Q_i and momenta $P_{Q,i}$ are named centroid pseudoparticles (cent) and the pseudoparticle with position $\bar{q}_i$ and momenta p_i are named observable pseudoparticles (obsv). The integral with respect to the vector $\Delta\boldsymbol{q}$ in eq 9 indicates that the interaction between the observable and the centroid pseudoparticles in the effective Hamiltonian are determined by the many-body CG potential $V_{CG,N}$.

2.2. Second-Order Taylor Expanded Expressions.

Despite the smaller number of DOF in eq 5 compared to that of the Feynman isomorphic potential in eq 2, the explicit functional form of $V_{CG,N}$ is nontrivial to know even for simple model systems. In our previous publications, one strategy to circumvent this issue was to specifically consider systems near the classical limit. In such cases, the exact expression of $V_{CG,N}$ can be Taylor expanded with respect to $\hbar$ and up to second-order terms would be summed to approximate $V_{CG,N}$.

We provide a similar derivation in this section. However, the new Taylor expanded result is different than that shown in the previous publication.[71] In the previous derivation, the form of the physical potential is assumed to be pairwise decomposable. Even though such an assumption has been commonly used in the development of simulation methods such as MS-CG to describe the intermolecular interactions, it lacks the specific functional forms to describe various intramolecular interactions commonly included in the force field of molecular systems. The pairwise decomposable assumption fails further as we only coarse-grain the intra-ring polymer coordinates and not more than one ring polymer, in which the many-body terms may be mapped into pair potentials. Therefore, it is beneficial to preserve the explicit intramolecular interactions within the molecule and build the CG-PI theory using a more general form of the effective force field. In this work, the form of the 1-dimensional many-body potential $V(q_1,\cdots,q_N)$ is assumed to be

https://doi.org/10.1021/acs.jpca.2c04349
J. Phys. Chem. A 2022, 126, 6004−6019

$$V(x_1, ..., x_N) = \sum_{\langle i,j \rangle} V_{ij}(x_i - x_j) + \sum_z V_z(x_1, ..., x_N) \quad (10)$$

where $V_{ij}(x_i - x_j)$ is the pair potential between the ith and jth particle as a function of distance between the two particles, $\langle i,j \rangle$ denotes a summation over all distinct pairs, $V_z(x_1,\cdots,x_N)$ is a general many-body potential which can used to describe various intramolecular interactions (bonded, angular, dihedral) or other types of many-body interactions such as the three-body Stillinger-Weber potential,[83] and Z is used to index the many-body interactions. We use x rather than q in eq 10 to indicate that the potential $V(x_1,\cdots,x_N)$ describes the interaction between N interacting classical particles, and is not a function of ring polymers or their beads. The form of potential shown in eq 10 can be used to describe any molecular system with many-body interactions.

The Taylor expanded expression of $V_{CG,N}$ up to second order using the form of potential shown in eq 10 can be written as

$$V_{CG,N}(\bar{\boldsymbol{q}}, \boldsymbol{Q}, \Delta\boldsymbol{q}) = \left[V(\boldsymbol{Q}) + (\boldsymbol{Q} - \bar{\boldsymbol{q}})^T \frac{\boldsymbol{K}_{Q\bar{q}}}{2} (\boldsymbol{Q} - \bar{\boldsymbol{q}}) + \Delta\boldsymbol{q}^T \frac{\boldsymbol{K}_{\Delta q^2}}{2} \Delta\boldsymbol{q} + K_0 \right] \times \left\{ 1 + O\left(\hbar^3 \sqrt{\frac{\beta^5}{m^3}} \{ V_{ij}''' \} \right) \right\} \quad (11)$$

for N dimensional square coupling matrices $\boldsymbol{K}_{Q\bar{q}}$ and $\boldsymbol{K}_{\Delta q^2}$ and scalar value K_0. Even though the forms of the expansion shown in eq 11 are the same as those of the previous derivation,[71] the details of the coupling matrices and scalar are different. The functional forms and the derivation of eq 11 is shown in the Supporting Information. Note that the terms in the second bracket of RHS are the third and higher order terms ($\{V_{ij}'''\}$).

In the square bracket of the Taylor expanded result, the first three terms are functions of $\bar{\boldsymbol{q}}$, $\boldsymbol{Q}$, and $\boldsymbol{\Delta q}$. The first term $V(\boldsymbol{Q})$ indicates that the centroid pseudoparticles experience the physical potential V as shown in eq 10. The second term $(\boldsymbol{Q} - \bar{\boldsymbol{q}})^T \frac{\boldsymbol{K}_{Q\bar{q}}}{2} (\boldsymbol{Q} - \bar{\boldsymbol{q}})$ states that the distance between the two pseudoparticles are quadratically coupled via the matrix $\boldsymbol{K}_{Q\bar{q}}$. Lastly, the third term $\Delta\boldsymbol{q}^T \frac{\boldsymbol{K}_{\Delta q^2}}{2} \Delta\boldsymbol{q}$ shows that the distance between the beginning and the end of the paths ($\boldsymbol{\Delta q}$) are harmonically coupled through the matrix $\boldsymbol{K}_{\Delta q^2}$ as well. The two coupling matrices $\boldsymbol{K}_{Q\bar{q}}$ and $\boldsymbol{K}_{\Delta q^2}$ are functions of mass, temperature, and second derivatives of the pair potentials V_{ij} and Hessian matrices of the many-body potentials V_z, both evaluated at $\bar{\boldsymbol{q}}$.

With the use of the second order truncated Taylor expanded result (square bracket in eq 11), the effective Hamiltonian can be written as

$$H_{eff,N}(\boldsymbol{P}_Q, \boldsymbol{p}, \boldsymbol{Q}, \bar{\boldsymbol{q}}) \cong \frac{1}{2} \boldsymbol{P}_Q{}^T \boldsymbol{M}_Q^{-1} \boldsymbol{P}_Q + \frac{1}{2\beta\hbar^2} \boldsymbol{p}^T \boldsymbol{K}_{\Delta q^2}^{-1} \boldsymbol{p} + V(\boldsymbol{Q}) + \frac{1}{2} (\boldsymbol{Q} - \bar{\boldsymbol{q}})^T \boldsymbol{K}_{Q\bar{q}} (\boldsymbol{Q} - \bar{\boldsymbol{q}}) \quad (12)$$

Equation 12 suggests that a system of N distinguishable particles obeying quantum Boltzmann distribution nearer the classical limit (small value of $\hbar$) is isomorphic to a system of $2N$ (N pairs) pseudoparticles. In this picture, the centroid pseudoparticles experience the physical potential via the $V(\boldsymbol{Q})$ term, and the two pseudoparticles are coupled to each other in a quadratic fashion through the ${}^1/_2(\boldsymbol{Q} - \bar{\boldsymbol{q}})^T \boldsymbol{K}_{Q\bar{q}} (\boldsymbol{Q} - \bar{\boldsymbol{q}})$ term. Moreover, the centroid and the observable pseudoparticles have associated momenta $\boldsymbol{P}_Q$ and $\boldsymbol{p}$. Note that in eq 12, the $\frac{1}{2\beta\hbar^2} \boldsymbol{p}^T \boldsymbol{K}_{\Delta q^2}^{-1} \boldsymbol{p}$ term suggests that the momenta of the ith and jth observable pseudoparticle are nontrivially coupled since $\boldsymbol{K}_{\Delta q^2}^{-1}$ in general will not be a diagonal matrix. This tells us that a mass cannot be straightforwardly assigned to the observable pseudoparticles due to the off-diagonal terms of the matrix $\boldsymbol{K}_{\Delta q^2}$ and its dependence on the observable pseudoparticle position $\bar{\boldsymbol{q}}$.

One can further approximate the masses associated with the observable pseudoparticles to be constant. Using the constant effective mass approximation, the many-body effective Hamiltonian can be written as

$$H_{\text{eff},N}(\boldsymbol{P}_Q, \boldsymbol{p}, \boldsymbol{Q}, \bar{\boldsymbol{q}}) \cong \frac{1}{2} \boldsymbol{P}_Q{}^T \boldsymbol{M}_Q^{-1} \boldsymbol{P}_Q + \frac{1}{2} \boldsymbol{p}^T \boldsymbol{m}_{\bar{q}}^{-1} \boldsymbol{p} + V(\boldsymbol{Q}) + \frac{1}{2} (\boldsymbol{Q} - \bar{\boldsymbol{q}})^T \boldsymbol{K}_{Q\bar{q}} (\boldsymbol{Q} - \bar{\boldsymbol{q}}) \quad (13)$$

The N dimensional diagonal mass matrix $\boldsymbol{m}_{\bar{q}}$ has constant effective mass for the ith observable pseudoparticle, $m_{\bar{q},i}$, as its ith diagonal element. In this picture, the only nontrivial coupling between the two pseudoparticles comes from the quadratic coupling term ${}^1/_2(\boldsymbol{Q} - \bar{\boldsymbol{q}})^T \boldsymbol{K}_{Q\bar{q}} (\boldsymbol{Q} - \bar{\boldsymbol{q}})$.

2.3. Challenges in Numerical Implementation. Despite the insight that the Taylor expanded expressions (eqs 11−13) gives on the nature of nuclear quantum effects for systems nearer the classical limit, these derivations cannot directly be used to model realistic molecular systems due to the complexity of calculating the coupling matrices and their inverses for every new position of the observable pseudoparticles. In the next section, we will discuss modeling strategies that are based on the functional terms shown in the Taylor expanded expression of the many-body CG potential and effective Hamiltonian.

2.4. Significance and Limitations of Previous n-CG-PI Method. In our previous work, we have used the numerical CG-PI (n-CG-PI) method to incorporate the $2N$ pseudoparticle effective Hamiltonian into simulations of realistic molecular liquid systems.[71] Throughout this paper, the previously developed n-CG-PI modeling scheme discussed in our previous publication will be referred to as the "naïve n-CG-PI."[71] The naïve n-CG-PI effective Hamiltonian can be written as

$$H_{\text{n-CG-PI}}(\boldsymbol{P}_Q, \boldsymbol{p}, \boldsymbol{Q}, \bar{\boldsymbol{q}}) = \frac{1}{2} \boldsymbol{P}_Q{}^T \boldsymbol{M}_Q^{-1} \boldsymbol{P}_Q + \frac{1}{2} \boldsymbol{p}^T \boldsymbol{m}_{\bar{q}}^{-1} \boldsymbol{p} + V(\boldsymbol{Q}) + \frac{1}{2} (\boldsymbol{Q} - \bar{\boldsymbol{q}})^T \boldsymbol{K}_{\lambda} (\boldsymbol{Q} - \bar{\boldsymbol{q}}) \quad (14)$$

where for the ith pseudoparticle pair coupling parameter λ_i ($0 \leq \lambda_i$), the ith diagonal element for the N dimensional diagonal coupling matrix $\boldsymbol{K}_{\lambda}$ is $\lambda_i \cdot 12 m_i/(\hbar^2\beta^2)$, where m_i is the physical mass of the ith particle in the system. The masses associated with the ith observable and centroid pseudoparticles were set to be m_i. One can arrive at eq 14 by making two

https://doi.org/10.1021/acs.jpca.2c04349
J. Phys. Chem. A 2022, 126, 6004−6019

approximations: (a) the off-diagonal elements of $\boldsymbol{K}_{Q\bar{q}}$ are negligible and (b) the diagonal elements of $\boldsymbol{K}_{\bar{q}\bar{q}}$, originally functions of $\bar{\boldsymbol{q}}$, can be approximated to be constants.

Despite providing a computationally tractable pseudoparticle representation, the naïve n-CG-PI method has three shortcomings to reliably model NQE in realistic molecular systems. (a) It needs an extensive set of PIMD calculations at different temperatures and masses to modulate NQEs to correctly parametrize the coupling parameter λ which in turn defeats the purpose of using the reductionist CG-PI representation. (b) It used the radial distribution function (RDF) as the criteria to parametrize λ rather than more fundamental and direct quantities from the original PIMD simulations such as pseudoparticle distributions and forces. (c) Lastly, it lacked the formalism to construct a naïve n-CG-PI model for systems with two or more different types of atoms.

To improve upon these shortcomings, we have developed a new n-CG-PI model that is directly parametrized from PIMD trajectories. However, before introducing the new modeling method, we give a brief overview of the two CG simulation methods in the context of ring polymers, Boltzmann inversion, and force matching.

2.5. Boltzmann Inversion in n-CG-PI. Boltzmann inversion is a bottom-up CG modeling method that parametrizes the bonded and nonbonded interactions in CG resolution using the distribution calculated from mapped fine-grained (FG) trajectory.[84−87] In the new n-CG-PI model, only the interaction between the centroid and observable pseudoparticles within the same pair is parametrized with Boltzmann inversion.

Using Boltzmann inversion, for the ith ring polymer, the normalized distribution of distance between the first bead and the center of mass of the $P-1$ other beads $\rho\left(q_1^{(i)} - \frac{q_2^{(i)}+\ldots+q_P^{(i)}}{P-1}\right)$ is related to the effective harmonic potential $V_{\text{cent-obsv}}(Q_i - \bar{q}_i)$ via

$$V_{\text{cent-obsv}}(Q_i - \bar{q}_i) = -k_B T \ln\left[\rho\left(q_1^{(i)} - \frac{q_2^{(i)}+\ldots+q_P^{(i)}}{P-1}\right) \Big/ \left(q_1^{(i)} - \frac{q_2^{(i)}+\ldots+q_P^{(i)}}{P-1}\right)^2\right] \tag{15}$$

The RHS of eq 15 is a function of fine-grained (PIMD) coordinates $q_j^{(i)}$ mapped into CG resolution (eq 3) and the LHS is a function of coarse-grained (CG-PI) coordinates Q_i and $\bar{q}_i$. The $1/x^2$ factor in the RHS accounts for the degeneracy of the positions. Note that the interaction term is a function of $Q_i - \bar{q}_i$, based on the Taylor expanded expression of the many-body CG potential shown in eq 11.

2.6. Force Matching in n-CG-PI. Force matching is a CG method that was initially intended to construct CG representations of classical all-atom systems.[76−79] However, it can be generalized to parametrize the interactions in the n-CG-PI model. The results shown in this section are a many-body generalization of the equations shown in Appendix 3 and Supporting Information S4.1 of one-body CG-PI theory work by Sinitskiy and Voth.[70] Note that the results shown in this section discuss the many-body CG potential for a closed ring polymer system ($q_i = q_i' = \bar{q}_i$).

The many-body CG-PI potential in eq 5 can alternatively be written as a variational equation

$$V_{CG,N}(\boldsymbol{Q}, \bar{\boldsymbol{q}}) = \arg\min_{V_{CG,N}^{\text{trial}}} \chi^2[V_{CG,N}^{\text{trial}}(\boldsymbol{Q}, \bar{\boldsymbol{q}})] + \text{const} \tag{16}$$

for the many-body trial CG potential $V_{CG,N}^{\text{trial}}$ and the variational functional χ^2

$$\chi^2[V_{CG,N}^{\text{trial}}(\boldsymbol{Q}, \bar{\boldsymbol{q}})] = \lim_{P\to\infty} \frac{\int\left[\prod_{i=1}^{N}\prod_{j=2}^{P} \mathrm{d}q_j^{(i)}\right]\Sigma e^{-\beta V_{P,N}}}{\int\left[\prod_{i=1}^{N}\prod_{j=2}^{P} \mathrm{d}q_j^{(i)}\right]e^{-\beta V_{P,N}}} \tag{17}$$

where $\sum$ is the sum of squares of deviations between the PIMD forces and CG-PI forces. It can be written as

$$\sum\left(q_j^{(i)};\ \begin{matrix} i=1,\ldots,N \\ j=1,\ldots,P\end{matrix}\right) = \sum_{i=1}^{N}[(\mathcal{F}_{\bar{q},i} - F_{q,i})^2 + (\mathcal{F}_{Q,i} - F_{Q,i})^2] \tag{18}$$

where the $\mathcal{F}$ terms are the forces experienced by the mapped ring polymer beads in the PIMD representation and the F terms are the forces experienced by the CG-PI beads by the trial potential. These terms can be expressed as

$$\mathcal{F}_{\bar{q},i} = -\frac{\partial V_{P,N}}{\partial q_1^{(i)}} \qquad \mathcal{F}_{Q,i} = -\frac{1}{P-1}\sum_{j=2}^{P}\frac{\partial V_{P,N}}{\partial q_j^{(i)}} \tag{19a}$$

$$F_{\bar{q},i} = -\frac{\partial V_{CG,N}^{\text{trial}}}{\partial \bar{q}_i} \qquad F_{Q,i} = -\frac{\partial V_{CG,N}^{\text{trial}}}{\partial Q_i} \tag{19b}$$

The $\mathcal{F}$ terms are functions of FG (PIMD) coordinates $q_j^{(i)}$, and the F terms are functions of coarse-grained (CG-PI) coordinates q_i, Q_i, and q_i'. Even though $q_1^{(i)} = q_i$ and $q_{P+1}^{(i)} = q_i'$ for all i and only the intermediate beads are coarse-grained, the two representations are distinguished for clarity.

2.7. Building the n-CG-PI Model. A wide range of molecular systems can be described with the classical force field that describes the total potential energy of the system V_{total} as

$$V_{\text{total}} = V_{\text{intra}} + V_{\text{inter}} \tag{20}$$

for intramolecular potential V_{intra} and intermolecular potential V_{inter}. Both types of potentials can further be specified as

$$V_{\text{intra}} = V_{\text{bond}} + V_{\text{angle}} + V_{\text{dihedral}} \tag{21a}$$

$$V_{\text{inter}} = V_{\text{electro}} + V_{\text{vdW}} \tag{21b}$$

for bond potential V_{bond}, angle potential V_{angle}, dihedral potential V_{dihedral}, electrostatic potential V_{electro}, and van der Waals potential V_{vdW}.

For a general molecular system with N atoms whose potential energy is described by eq 20, the proposed n-CG-PI effective Hamiltonian is written as

$$H_{\text{n-CG-PI}}(\boldsymbol{P}_Q, \boldsymbol{p}, \boldsymbol{Q}, \bar{\boldsymbol{q}}) = \frac{1}{2}\boldsymbol{P}_Q{}^T\boldsymbol{m}^{-1}\boldsymbol{P}_Q + \frac{1}{2}\boldsymbol{p}^T\boldsymbol{m}^{-1}\boldsymbol{p} + V_{\text{intra}}(\boldsymbol{Q}, \bar{\boldsymbol{q}}) + V_{\text{inter}}(\boldsymbol{Q}) \tag{22}$$

for the N dimensional diagonal mass matrix $\boldsymbol{m}$ whose ith diagonal element is mass of the ith atom m_i and the intramolecular and intermolecular potentials V_{intra} and V_{inter}. The two types of interactions are further specified as

https://doi.org/10.1021/acs.jpca.2c04349
J. Phys. Chem. A 2022, 126, 6004−6019

$$V_{\text{intra}}(\mathbf{Q}, \bar{\mathbf{q}}) = V_{\text{intra,cent-obsv}}(\mathbf{Q} - \bar{\mathbf{q}}) + V_{\text{intra,cent}}(\mathbf{Q}) \tag{23a}$$

$$V_{\text{inter}}(\mathbf{Q}) = \sum_{\langle i,j \rangle} V_{ij,\text{cent}}(Q_i - Q_j) \tag{23b}$$

for pair potential between the ith and the jth centroid pseudoparticle V_{ij}.

According to eqs 22 and 23, there are three different types of interactions in the n-CG-PI model that need to be parametrized: $V_{\text{intra,cent-obsv}}(\mathbf{Q} - \bar{\mathbf{q}})$, $V_{\text{intra,cent}}(\mathbf{Q})$, and $V_{\text{inter}}(\mathbf{Q})$. The functional terms of the intramolecular interactions can be written as

$$V_{\text{intra,cent-obsv}}(\mathbf{Q} - \bar{\mathbf{q}}) = \sum_{i=1}^{N} K_i(Q_i - \bar{q}_i)^2 \tag{24a}$$

$$V_{\text{intra,cent}}(\mathbf{Q}) = V_{\text{bond,cent}}(\mathbf{Q}) + V_{\text{angle,cent}}(\mathbf{Q}) + V_{\text{dihedral,cent}}(\mathbf{Q}) \tag{24b}$$

where the centroid intramolecular potentials ($V_{\text{intra,cent}}(\mathbf{Q})$) have the same form as that of the original force field specified in eq 21a.

The new n-CG-PI effective Hamiltonian shown in eq 22 closely resembles the naïve n-CG-PI effective Hamiltonian in eq 14. The $V_{\text{intra,cent-obsv}}(\mathbf{Q} - \bar{\mathbf{q}})$ has the same functional form as ${}^1/_2(\mathbf{Q} - \bar{\mathbf{q}})^T\mathbf{K}_\lambda(\mathbf{Q} - \bar{\mathbf{q}})$. Moreover, the centroid pseudoparticle only interactions ($V_{\text{intra,cent}}(\mathbf{Q}) + V_{\text{inter}}(\mathbf{Q})$) from eqs 24b and 23b are equivalent to the $V(\mathbf{Q})$ term in eq 14. The resemblance comes from using the same two approximations that were used to construct the naïve n-CG-PI effective Hamiltonian. The only notable difference is that in the new effective Hamiltonian, the masses associated with the pseudoparticles are set to be the physical mass of the ith particle.

However, the difference between the new n-CG-PI method and its predecessor comes in the parametrization of the potential terms. In this work, we use Boltzmann inversion to parametrize the $V_{\text{intra,cent-obsv}}(\mathbf{Q} - \bar{\mathbf{q}})$ term and force matching to parametrize the centroid pseudoparticle only potentials ($V_{\text{intra,cent}}(\mathbf{Q})$ and $V_{\text{inter}}(\mathbf{Q})$). Note that there is only one pair potential between the ith and jth centroid pseudoparticle (eq 23b), which means that the $V_{ij,\text{cent}}$ contains a contribution from both electrostatic and van der Waals interactions between the ith and jth ring polymers.

To sum up, one can construct the n-CG-PI model as follows:

(a) for a particular force field, run a PIMD simulation with P beads per ring polymer,
(b) map the PIMD trajectory positions and forces for each bead into centroid and observable pseudoparticle positions and forces using the mapping shown in eq 3 (force on the centroid pseudoparticle is calculated in a similar manner),
(c) parametrize $V_{\text{intra,cent-obsv}}(\mathbf{Q} - \bar{\mathbf{q}})$ from the mapped cent-obsv distance for each pair using eq 15,
(d) subtract the parametrized harmonic forces obtained in c) from the mapped PIMD trajectory forces,
(e) parametrize $V_{\text{intra,cent}}(\mathbf{Q})$ and $V_{\text{inter}}(\mathbf{Q})$ using the modified mapped PIMD trajectory from (d) with force matching,
(f) build the CG-PI effective Hamiltonian using the parameters from (c) and (e), and
(g) sample the effective Hamiltonian with classical molecular dynamics.

3. RESULTS AND DISCUSSION

3.1. Liquid Para-hydrogen. *3.1.1. System Description.* The liquid para-hydrogen system is modeled using the semiempirical force field by Silvera and Goldman.[88] This form of potential treats the diatomic H_2 molecule in the $J = 0$ state as a point particle and describes the interaction between the particles using a pair potential. In our previous work, we used a modified version of the semiempirical force field by systematically varying the mass and temperature of the system, and hence modulating the degree of NQE, to construct the scaling relationship. The modified para-hydrogen liquid system described in this section has 5 g/mol for the mass of the particle and temperature of 30 K.

The form of the force field is

$$V_{\text{paraH}}(x_1, \ldots, x_N) = \sum_{\langle i,j \rangle} V_{ij}(x_i - x_j) \tag{25}$$

where x_i is the coordinate of the ith particle, V_{ij} is the pair potential between the ith and the jth particle, and $\langle i,j \rangle$ is an index for summation over all distinct pairs in the system.

3.1.2. Previous Results. Previously we have used the naïve n-CG-PI method to calculate the structural correlations of the modified para-hydrogen system. The naïve n-CG-PI effective Hamiltonian for para-hydrogen can be written as

$$H_{\text{n-CG-PI,naive}}(\mathbf{P}_Q, \mathbf{p}, \mathbf{Q}, \bar{\mathbf{q}}) = \frac{1}{2}\mathbf{P}_Q{}^T\mathbf{m}^{-1}\mathbf{P}_Q + \frac{1}{2}\mathbf{p}^T\mathbf{m}^{-1}\mathbf{p} + K\sum_{i=1}^{N}(Q_i - \bar{q}_i)^2 + \sum_{\langle i,j \rangle} V_{ij}(Q_i - Q_j) \tag{26}$$

where K is the spring constant of centroid−observable bonded interaction and V_{ij} is the pair potential shown in eq 25 evaluated at centroid pseudoparticle positions.

To parametrize the spring constant K of the naïve n-CG-PI para-hydrogen model, we constructed a scaling relationship by simulating the para-hydrogen system at different values of mass and temperature using the PIMD method. For each case, the spring constant K was considered to be "fit" if the peak height and width of the first solvation shell of radial distribution function from naïve n-CG-PI simulations agreed to that of PIMD simulation. Equation 24 of our previous publication shows that the coupling strength, which can be understood to be inversely proportional to NQE, is proportional to mass and temperature squared.[71] The coupling parameter K of the naïve n-CG-PI para-hydrogen model with a mass of 5 g/mol and temperature of 30 K calculated from this scaling relationship is 2.82 kcal/mol/Å^2. This scaling relationship is an intriguing result that directly quantifies a degree of NQE as a function of mass and temperature for the CG-PI representation of a simple liquid system.

Our previous publication shows the calculated radial distribution function and nonbonded angle distribution using the PIMD, classical MD, and naïve n-CG-PI method.[71] Despite good agreement between the PIMD and the CG-PI results, the CG-PI RDFs are slightly shifted from the reference PIMD

https://doi.org/10.1021/acs.jpca.2c04349
J. Phys. Chem. A 2022, 126, 6004−6019

ones. The overall shape of the function quantitatively agree, but the CG-PI system predicts a smaller sphere of the particle compared to the PIMD system. Moreover, it shows the physical quantities (quantities that involve the correlation of observable pseudoparticles), but not the ones that involve centroid pseudoparticles, which encode the structural correlations of intra ring polymer fluctuations in the PIMD system.

3.1.3. n-CG-PI Model Parameters. For the modified para-hydrogen system, the new n-CG-PI effective Hamiltonian can be written as

$$H_{\text{n-CG-PI}}(\boldsymbol{P}_Q, \boldsymbol{p}, \boldsymbol{Q}, \bar{\boldsymbol{q}}) = \frac{1}{2}\boldsymbol{P}_Q{}^T\boldsymbol{m}^{-1}\boldsymbol{P}_Q + \frac{1}{2}\boldsymbol{p}^T\boldsymbol{m}^{-1}\boldsymbol{p} + K\sum_{i=1}^{N}(Q_i - \bar{q}_i)^2 + \sum_{\langle i,j\rangle} V_{ij,cent}(Q_i - Q_j) \quad (27)$$

for the cent-obsv spring constant K and the centroid pseudoparticle pair potential $V_{ij,\text{cent}}(r)$. Note that due to the simplicity of the system, we only need to parametrize two quantities: K and $V_{ij,\text{cent}}(r)$. We see that the functional form of the new n-CG-PI effective Hamiltonian is analogous to that shown in eq 26. This similarity suggests that the modified para-hydrogen liquid system is an adequate test case to study the differences between the naïve n-CG-PI and the new n-CG-PI methods. Figure 1 shows a single n-CG-PI para-hydrogen molecule composed of a pair of pseudoparticles.

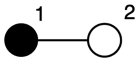


Figure 1. A schematic representation of the n-CG-PI para-hydrogen model. A ring polymer is coarse-grained into two pseudoparticles: the observable pseudoparticle (black and numbered 1) and the centroid pseudoparticle (white and numbered 2).

To build the n-CG-PI model, a 2.0 ns PIMD trajectory with a bead number of 32 was used. We use the "rerun" command in LAMMPS molecular dynamics engine[89] to subtract the harmonic interaction parametrized from Boltzmann inversion. The centroid pseudoparticle interactions were parametrized by using the in-house force matching code.[78,79]

The spring constant in the new n-CG-PI parahydrogen model parametrized via Boltzmann inversion is 1.08 kcal/mol/Å^2. The inverted distribution shown in the RHS of eq 15 had a quantitative fit to the quadratic form of the potential. Compared to the naïve n-CG-PI harmonic spring constant with a value of 2.82 kcal/mol/Å^2, the new model predicts a weaker interaction between the centroid and the observable pseudoparticle of the same pair.

Figure 2 shows two pair potentials: the original pair potential $V_{ij}(r)$ from eq 25 and the force matched centroid pseudoparticle pair potential $V_{ij,\text{cent}}(r)$ from eq 27, obtained by using force matching. The two functions quantitatively agree for $r \geq 4$ Å. For $r < 4$ Å, however, we see that the force matched pair potential is slightly shifted to the right by approximately 0.1 Å compared to the original potential. The force matched pair therefore predicts a larger radius of the centroid pseudoparticle. The combined effect of weaker harmonic interaction between the pseudoparticles and the shifted pair potential is manifested in the calculated radial

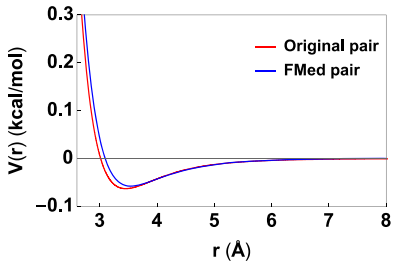


Figure 2. Original para-hydrogen pair potential (red) and the force matched centroid pseudoparticle pair potential (blue).

distribution functions. This aspect of the new n-CG-PI model will further be discussed in later sections.

3.1.4. Simulation Details. The modified para-hydrogen system is a NVT ensemble of 180 particles in a cubic box of length 19.71 Å. To modify and enhance the NQEs of the system, the mass of the particle was set to be 5 g/mol, and the temperature was set to 30 K. Three sets of methods are used to simulate the system: PIMD, n-CG-PI, and classical MD. For force matching, an in-house developed MS-CG code was used with nonbonded B-spline order and bonded B-spline order of 4. For the PIMD and classical simulations, the semiempirical para-hydrogen force field was used,[88] and for the n-CG-PI simulation, the model described in the previous section was used. For all three methods, time step of 0.5 fs was used to propagate the system. We used the Langevin thermostat with the parameter 0.01 fs^{-1} for all simulations (the centroid and the observable pseudoparticles were thermostatted equally).[90] An in-house modified LAMMPS molecular dynamics engine using the Path Integral Langevin Equation (PILE) integrator[91] was used to perform the PIMD simulation. Unmodified version of the LAMMPS molecular dynamics engine was used for n-CG-PI and classical simulations.

Additionally, the thermostatting issue discussed in our previous work has been resolved. In the naïve n-CG-PI model, the observable pseudoparticles had to be more heavily thermostatted compared to the centroid pseudoparticles. This issue arose from the fact that the observable pseudoparticle coordinates (position and force) were added into the system via the property/atom fix in LAMMPS, not as an actual atom present in the system. However, in the new n-CG-PI models, both pseudoparticles are considered as actual atoms during the simulation and no longer suffer from the thermostatting issue.

3.1.5. Structural Correlations. Three types of structural correlations calculated from the PIMD, n-CG-PI, and classical MD simulations are presented in this section: normalized intramolecular distributions $\rho(x)$, radial distribution functions $g(r)$, and three-body nonbonded angle distribution $P(\theta)$. The intramolecular distribution is normalized so that $\int dx\, \rho(x) = 1$ for variable $x = r, \theta$. The nonbonded angle distribution is given by the equation

$$P(\theta) = \sum_{i}^{N}\sum_{j>i}^{N}\sum_{k\neq i,j}^{N}\delta(\theta_{ijk} - \theta) \quad (28)$$

for the angle between particles i, j, and k centered at the ith particle θ_{ijk} with a cutoff distance of 3.5 Å. Note that the distributions calculated from the PIMD and n-CG-PI simulations that involve only the observable pseudoparticles are compared to those from classical MD simulations because

https://doi.org/10.1021/acs.jpca.2c04349
J. Phys. Chem. A 2022, 126, 6004−6019

the expectation values of operators are calculated with respect to the observable pseudoparticle coordinates as shown in eq 7. However, the structural correlations that involve the centroid pseudoparticle coordinates are included in this section for completion.

Figure 3 shows the normalized bond distribution between the centroid and the observable pseudoparticles within the ring

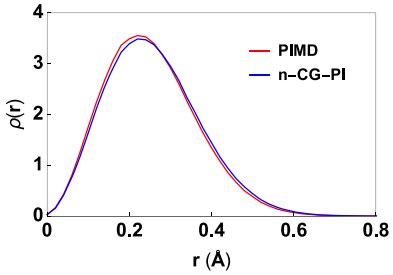


Figure 3. Calculated intramolecular distribution for the modified liquid para-hydrogen model between the centroid and the observable pseudoparticle of the same pair from the PIMD simulation (red) and from the n-CG-PI simulation (blue).

polymer (PIMD) and the same pair (n-CG-PI). The PIMD distribution shown in Figure 3 corresponds to the ρ shown in eq 15. We see that the n-CG-PI distribution well captures the height, width, and the overall shape of the PIMD distribution. However, the agreement between the two is not complete. The n-CG-PI distribution is slightly shorter and broader, which suggests that the n-CG-PI para-hydrogen model predicts a weaker effective harmonic interaction between the pseudoparticles compared to the reference PIMD distribution. The harmonic interaction parametrized from Boltzmann inversion therefore "overcorrects" the NQE in the para-hydrogen system, which is a topic that will be discussed.

The radial distribution functions from different methodologies are shown in Figure 4. A stark difference between the classical and the PIMD RDF shown in Figure 4a reflects a significant degree of NQE in the system (note that the parameters were adjusted to enhance NQEs for this example). The classical RDF shows a more structured and localized

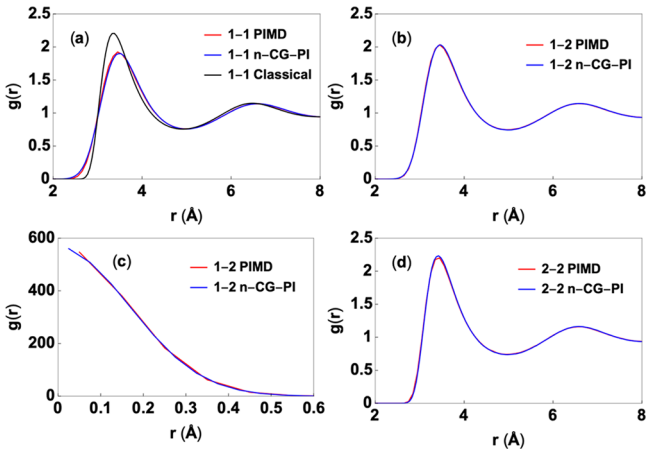


Figure 4. Calculated radial distribution function for (a) obsv−obsv (1−1), (b) obsv−cent (1−2) for 2 Å ≤ r < 8 Å, (c) obsv−cent (1−2) for r < 2 Å, and (d) cent−cent (2−2). The PIMD results are shown in red, the n-CG-PI results are shown in blue, and the classical results are shown in black when applicable.

liquid structure with a higher first solvation shell compared to the PIMD RDF. We see that for the obsv−obsv (1−1) pair correlation, the n-CG-PI result does not have the shift in the function that was observed in the naïve n-CG-PI result and has excellent agreement with the PIMD RDF. The weaker harmonic interaction and the larger sphere radius in the new n-CG-PI para-hydrogen model corrects the shift observed in the naïve n-CG-PI RDF. Note that Figure 4 panels b and c plot the same function but at different values of distance. Moreover, Figure 4c is equivalent to Figure 3, with the additional $1/r^2$ factor that accounts for degeneracies. The n-CG-PI obsv−cent (1−2) and cent−cent (2−2) RDF quantitatively agree with the PIMD ones as shown in Figure 4 panels b and d. Even though the structural correlations that involve centroid pseudoparticles do not correspond to any physical quantity, good agreement for the centroid pseudoparticle distributions suggests that the n-CG-PI para-hydrogen model captures well the overall correlations and interactions of the ring polymers in the reference PIMD para-hydrogen system.

The nonbonded angle distribution captures the three-body interactions in the liquid system. Figure 5 shows the

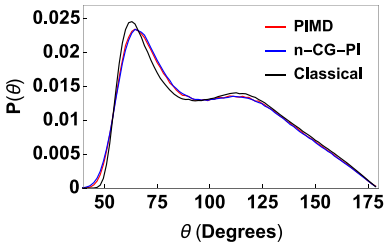


Figure 5. Calculated nonbonded angle distribution of observable pseudoparticles (PIMD and n-CG-PI) or particles (classical) with a cutoff of 3.5 Å. The PIMD results are shown in red, the n-CG-PI results are shown in blue, and the classical results are shown in black.

nonbonded angle distributions calculated from PIMD, n-CG-PI, and classical MD simulations. Similar to that in Figure 4, the difference between the classical and the PIMD $P(\theta)$ shows that the nuclear quantum effect is still manifested into higher order structural correlations. We also see that the n-CG-PI result well agrees with the PIMD distribution. This result follows naturally from the agreement shown in Figure 4 as the parahydrogen force field only contains two-body interactions.

3.2. Liquid Water. *3.2.1. System Description.* In this work, we use the qSPC-Fw water model.[3] The form of the force field can be written as

$$V_{\text{water}}(q_1, \ldots, q_N) = V_{\text{intra}} + V_{\text{inter}} \tag{29}$$

in which the intramolecular term V_{intra} and the intermolecular term V_{inter} are specified as

$$V_{\text{intra}} = \sum_{N_{\text{mol}}} k_b[(r_{\text{OH1}} - r^0_{\text{OH1}})^2 + (r_{\text{OH2}} - r^0_{\text{OH2}})^2] + k_a(\theta_{\text{HOH}} - \theta^0_{\text{HOH}})^2 \tag{30a}$$

$$V_{\text{inter}} = \sum_{\langle i,j \rangle} \left[4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{r_{ij}} \right] \tag{30b}$$

where N_{mol} is the total number of water molecules in the system, k_b and k_a are the harmonic spring constants for the

https://doi.org/10.1021/acs.jpca.2c04349
J. Phys. Chem. A 2022, 126, 6004−6019

bonded and the angle interactions, r_{OH1} and r_{OH2} are the two H−O bond distances, θ_{HOH} is the H−O−H angle, and r^0_{OH1},r^0_{OH2}, and θ^0_{HOH} are the equilibrium distances for the bonded interactions. The specific values of the intramolecular interaction parameters are k_b = 529.6 kcal/(mol rad^2), r^0_{OH1} = r^0_{OH2} = 1.0 Å, k_a = 37.95 kcal/(mol Å^2), and θ^0_{HOH} = 112°. For the H−H and O−H intermolecular pair potentials, ϵ_{HH} = ϵ_{OH} = 0 and σ_{HH} = σ_{OH} = 0. For the O−O pair, the parameters are ϵ_{OO} = 0.1554 kcal/mol, and σ_{OO} = 3.165 Å. The partial charge of the oxygen atom is −0.84, and the partial charge of the hydrogen atom is 0.42. The form of the water force field shown in eqs 29 and 30, unlike the para-hydrogen force field, contains explicit intramolecular interactions. This system provides an opportunity to showcase how the n-CG-PI modeling scheme captures NQEs of a more realistic and complicated molecular system.

3.2.2. Previous Results. Analogous to the naïve n-CG-PI para-hydrogen model, the centroid pseudoparticles in the naïve n-CG-PI water model experiences the physical potential of the water force field shown in eqs 29 and 30.[71] The chosen coupling strength between the different pseudoparticles were 265.5 kcal/(mol Å^2) and 1416 kcal/(mol Å^2) for hydrogen and oxygen, respectively. In the previous work, we did not construct a scaling relationship of the coupling parameter for the water model due to the difficulty of having more than one type of atom. Therefore, the bonding strengths were arbitrarily chosen in the previous work.

Figure 4 of our previous paper shows the calculated radial distribution function from three methods: PIMD, classical MD, and naïve n-CG-PI. We see that the agreement between the naïve n-CG-PI and the PIMD result is reasonable. The naïve n-CG-PI RDFs capture the overall shape of the pair correlation and the relative heights of the peaks. However, for all the RDFs shown, the naïve n-CG-PI predicts a more classical liquid structure with sharper and higher peaks. Moreover, for the H−H RDF at a distance larger than 2 Å, we see a shift of the naïve n-CG-PI intermolecular pair correlation compared to the PIMD one.

3.2.3. n-CG-PI Model Parameters. Figure 6 provides a schematic of the n-CG-PI water model. There are six

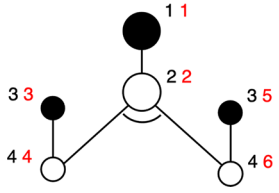


Figure 6. Schematic of the n-CG-PI water model. The black numbers represent the types of pseudoparticles: the oxygen observable pseudoparticle is numbered 1, the oxygen centroid pseudoparticle is numbered 2, the hydrogen observable pseudoparticle is numbered 3, and the hydrogen centroid pseudoparticle is numbered 4. The red numbers are indexes of the pseudoparticles ranging from 1 to 6.

pseudoparticles in a single n-CG-PI water molecule (written in red). The four types of pseudoparticles written in black are oxygen observable pseudoparticle (1), oxygen centroid pseudoparticle (2), hydrogen observable pseudoparticle (3), and hydrogen centroid pseudoparticle (4). The lines and curves in Figure 6 represent six total intramolecular interactions of four types: 1−2 bond, 3−4 bond, 2−4 bond, 4−2−4 angle. Moreover, there are three types of nonbonded pair interactions: 2−2, 2−4, and 4−4 pair. We will actively use the indexing of different pseudoparticles in describing the force field and the results for notational simplicity.

The effective Hamiltonian of the n-CG-PI water model parametrized from qSPC/Fw force field can be written as

$$H_{\text{n-CG-PI}}(\boldsymbol{P}_Q, \boldsymbol{p}, \boldsymbol{Q}, \bar{\boldsymbol{q}}) = \frac{1}{2}\boldsymbol{P}_Q{}^T\boldsymbol{m}^{-1}\boldsymbol{P}_Q + \frac{1}{2}\boldsymbol{p}^T\boldsymbol{m}^{-1}\boldsymbol{p} + V_{\text{intra,cent-obsv}}(\boldsymbol{Q} - \bar{\boldsymbol{q}}) + V_{\text{intra,cent}}(\boldsymbol{Q}) + V_{\text{inter,cent}}(\boldsymbol{Q}) \quad (31)$$

for the cent−obsv intramolecular term $V_{\text{intra,cent-obsv}}$, centroid intramolecular term $V_{\text{intra,cent}}$, and centroid intermolecular term $V_{\text{inter,cent}}$. The intramolecular terms can be specified as

$$V_{\text{intra,cent-obsv}}(\boldsymbol{Q} - \bar{\boldsymbol{q}}) = \sum_{N_{\text{mol}}} k_{12}(r_{12} - r^0_{12})^2 + k_{34}[(r_{34} - r^0_{34})^2 + (r_{56} - r^0_{34})^2] \quad (32a)$$

$$V_{\text{inter,cent}}(\boldsymbol{Q}) = \sum_{N_{\text{mol}}} k_{24}[(r_{24} - r^0_{24})^2 + (r_{26} - r^0_{24})^2] + k_{424}(\theta_{426} - \theta^0_{424})^2 \quad (32b)$$

The spring constants k_{12}, k_{34}, k_{24}, and k_{424} and equilibrium values r^0_{12}, r^0_{34}, r^0_{24}, and θ^0_{424} parametrize the four intramolecular interactions according to their indices. The distances r_{12}, r_{34}, r_{56}, r_{24}, and r_{26} are intramolecular distances between the indexed pseudoparticles, and the angle θ_{426} is an angle between pseudoparticle 4, 2, and 6 with 2 in the center. The values of these parameters are shown in Table 1. Lastly, the centroid intermolecular term can be written as

Table 1. Parametrized Spring Constants and Equilibrium Distances for the n-CG-PI Water Intramolecular Interactions

Interaction Type	k_b (kcal/mol/Å^2)	r_0 (Å)
1−2	336.4	0
3−4	24.01	0
2−4	493.0	0.988
Interaction Type	k_a (kcal/mol/rad^2)	θ_0 (Å)
4−2−4	35.42	112.3

$$V_{\text{inter,cent}}(\boldsymbol{Q}) = \sum_{\langle i,j \rangle} V_{22}(r_{22}) + V_{24}(r_{24}) + V_{44}(r_{44}) \quad (33)$$

where V_{22}, V_{24}, and V_{44} are pair potentials between the indexed pseudoparticles which are functions of the pair distance. The three force-matched pair potentials are shown in Figure 7.

There are few trends that we see in the intramolecular parameters that should be discussed. First of all, the spring constant of the 1−2 bond is an order of magnitude larger than that of the 3−4 bond. Such a difference can be explained by the fact that hydrogen has larger degrees of NQE than oxygen due to its small mass. Note that the equilibrium distances of the two obsv−cent harmonic interactions (1−2 and 3−4) are zero, which agrees with the functional form from the Taylor expanded expression of the many-body CG potential shown in eq 11. Moreover, we see that the 1−2 and 3−4 bond spring constants are much smaller compared to their naïve n-CG-PI

https://doi.org/10.1021/acs.jpca.2c04349
J. Phys. Chem. A 2022, 126, 6004−6019

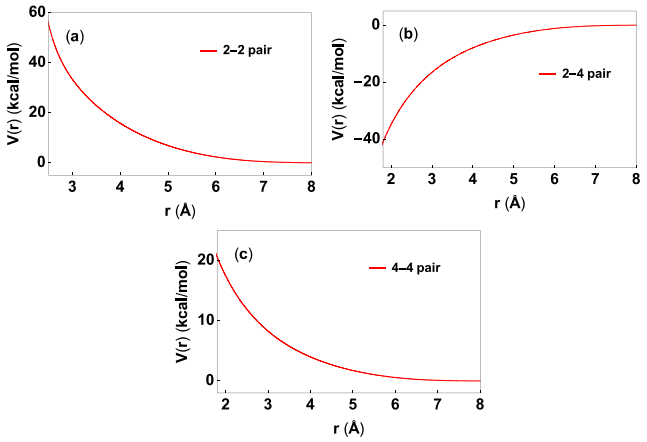


Figure 7. Force-matched (a) 2−2 pair, (b) 2−4 pair, and (c) 4−4 pair potential.

equivalents. The spring constant of other two intramolecular interactions (2−4 bond and 4−2−4 angle) are both smaller than their reference ones (O−H bond and H−O−H angle, respectively) in the original qSPC/Fw force field by 7%. The equilibrium value of the 2−4 bond and 4−2−4 interactions however does not vary much from the reference values (less than 5%) and shows no trend.

Figure 7 shows the three force-matched pair potentials included in the n-CG-PI water model. We see that the pair interactions between the same pseudoparticles are all repulsive and the pair interaction between the different pseudoparticles are all attractive. This result is somewhat surprising considering that the form of the original pair potential, for the O−O interaction at least, included both a repulsive and an attractive part. However, this result shows the same trend of the force matched pair potentials parametrized from first principle MD simulation.[8] Even though the qSPC/Fw (or SPC/Fw) water model and the first principle water model are parametrized differently, the agreement in the trend supports the validity of the force matched pair potentials shown in Figure 7.

These pair potentials will contain three types of contributions from the PIMD resolution: (a) LJ pair interaction between the beads, (b) electrostatic interaction between the beads, and (c) harmonic nearest neighbor coupling within the ring polymer (which is coarse-grained). However, using our current formalism, it is challenging to intuitively understand the trends in pair potentials. If a theory of charges in the context of CG-PI theory were to be developed, a more intuitive form of intermolecular interaction, more analogous to the form shown in eq 30b, in the n-CG-PI model can be parametrized. However, even with the current formalism, the new n-CG-PI water model using the pair potentials shown in Figure 7 provides good agreement with the reference PIMD results. The physical implications of pseudoparticle charge is left for future research.

3.2.4. Simulation Details. A NVT ensemble of 233 water molecules in a cubic box of length 19.03 Å was simulated at 300 K for 2.0 ns using three different methods: PIMD, n-CG-PI method, and classical MD. For force matching, an in-house developed MS-CG code was used with a nonbonded B-spline order of 6 and bonded B-spline order of 4. For the PIMD and classical simulations, the qSPC/Fw water model was used, and for the n-CG-PI simulation, the model described in the previous section was used. A time step of 0.5 fs and Langevin thermostat with parameter 0.05 fs^{-1} was used for all simulations.[90] The PIMD simulations were propagated by an in-house modified LAMMPS MD engine[89] using the PILE integrator.[91] The n-CG-PI and classical simulations were performed with vanilla LAMMPS without any modifications.

3.2.5. Structural Correlations. Similar to the para-hydrogen system, we calculate three types of structural correlations: intramolecular distribution (bond and angle), radial distribution function, and nonbonded angle correlation. Note that unlike the para-hydrogen system, the water n-CG-PI system has a more complicated intramolecular structures as well as pair correlations.

Figure 8 shows the 1−3 (O−H) bond and 3−1−3 (H−O−H) angle distribution calculated from PIMD, n-CG-PI, and

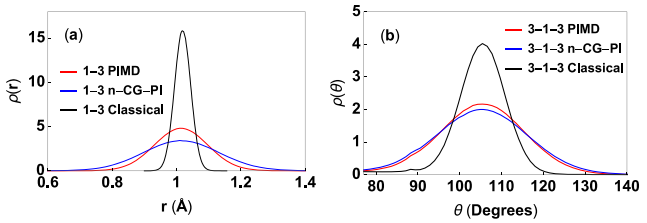


Figure 8. Calculated (a) 1−3 (O−H) bond and (b) 3−1−3 (H−O−H) angle distributions using different methods. The PIMD results are shown in red, the n-CG-PI results are shown in blue, and the classical MD results are shown in black.

classical MD simulations. It should be noted that the 1−3 bond and 3−1−3 angle are not explicitly added in the intramolecular interactions of the n-CG-PI water model. We see that for both distributions, the classical results are noticeably sharper (higher peak height and smaller width) than the PIMD ones. This suggests a large degree of NQE in the system. Even though the width of the distributions differ for all three methods, the mean of the distribution are all approximately the same for both types of distributions. Most importantly, it is intriguing to see that the two n-CG-PI distributions have larger widths than the PIMD ones. This suggests that the n-CG-PI water gives more delocalized distributions and hence has a larger degree of NQE compared to the PIMD water. We see that this overcorrection of NQE in the n-CG-PI models is a common trend.

We also investigated the intramolecular distributions that involve centroid pseudoparticle positions. These distributions are important because the corresponding intramolecular interactions are explicitly included in the n-CG-PI water model. Figure 9 shows the 1−2, 3−4, and 2−4 bond distributions and the 4−2−4 angle distribution. We see that the agreement between the n-CG-PI and the PIMD distributions are better for 1−2 (Figure 9a) than for 3−4 (Figure 9b). This comes from the fact that the n-CG-PI model is based on the Taylor expanded expression of the many-body CG potential in which we assume that the system is nearer the classical limit. Therefore, it better captures the NQE of the oxygen pseudoparticle pairs than the hydrogen ones. However, for both distributions, the agreement is still not quantitative. For both cases the n-CG-PI distributions have smaller peaks and wider widths than the PIMD ones, again signifying that the 1−2 and 3−4 bond in the n-CG-PI water model is more delocalized than those in the mapped PIMD water trajectory. Figure 9 panels c and d show the 2−4 bond and 4−2−4 angle distributions, which were interactions included in the model

https://doi.org/10.1021/acs.jpca.2c04349
J. Phys. Chem. A 2022, 126, 6004−6019

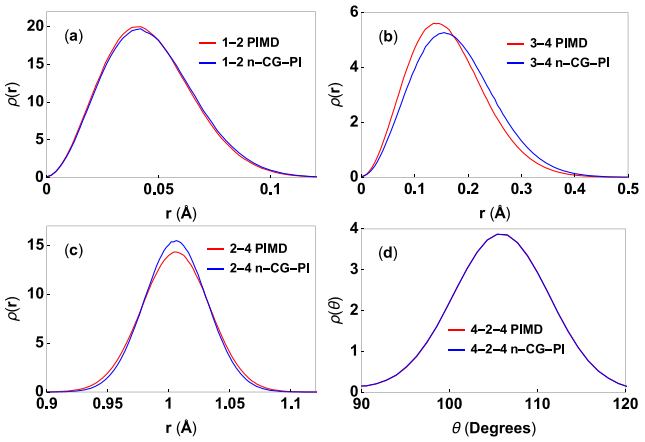


Figure 9. Calculated (a) 1−2 bond distribution, (b) 3−4 bond distribution, (c) 2−4 bond distribution, and (d) 4−2−4 angle distributions using the PIMD and the n-CG-PI methods. The PIMD results are shown in red, and the n-CG-PI results are shown in blue.

via force matching. We see that the n-CG-PI 2−4 bond distribution has a higher peak height and width, which suggests a more classical and less delocalized 2−4 bond than that in the PIMD system. Lastly, the 4−2−4 angle distributions from the two methods quantitatively agree and overlap with each other.

Radial distribution functions capture two-body correlations. By investigating the pair correlations, we are able to see how NQE plays an important role in determining the liquid structure and the performance of the new n-CG-PI water model. Figure 10 shows the three RDFs: 1−1 (O−O), 1−3

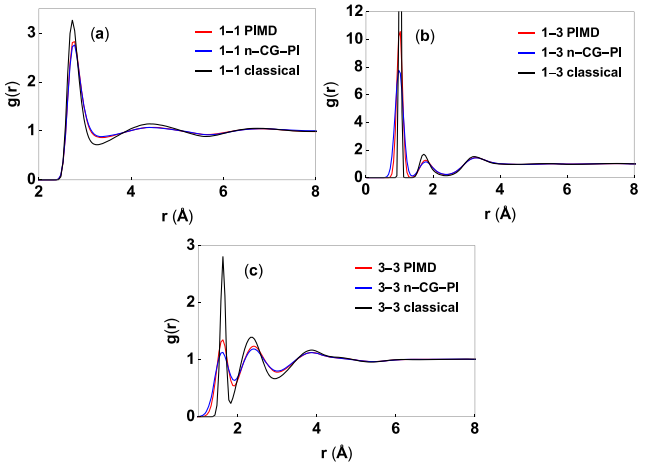


Figure 10. Calculated (a) 1−1 radial distribution function, (b) 1−3 radial distribution function, and (c) 3−3 radial distribution functions. The PIMD results are shown in red, the n-CG-PI results are shown in blue, and the classical MD results are shown in black.

(O−H), and 3−3 (H−H). Note that these pair interactions are not explicitly included in parametrizing the n-CG-PI water model. We see that for all three pair correlations the n-CG-PI well captures the overall shape of the PIMD function and its relative heights. But, there are slight differences between the n-CG-PI and PIMD results, which are more pronounced in the 3−3 RDF compared to the 1−1 and 1−3 RDF due to the larger degree of NQE in hydrogen. Moreover, the observable pair correlations from n-CG-PI method again predicts a more delocalized and quantum liquid structure compared to the PIMD results. Despite the difference, however, the n-CG-PI RDFs still capture the general features of the PIMD RDFs, especially compared to the naïve n-CG-PI RDFs.

We investigate the centroid pseudoparticle pair correlations as well. Figure 11 shows three RDFs: 2−2, 2−4, and 4−4

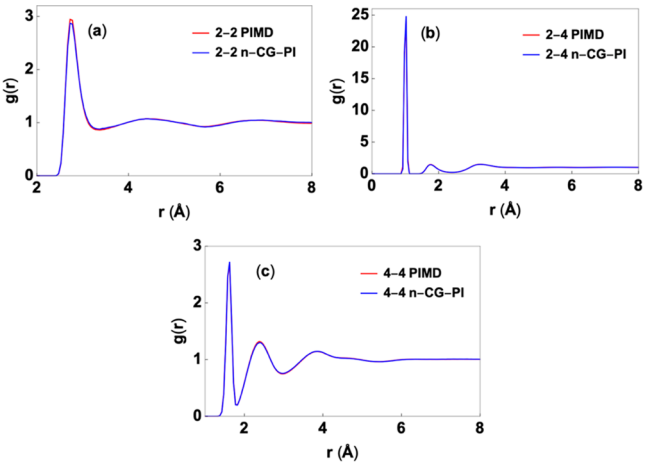


Figure 11. Calculated (a) 2−2 radial distribution function, (b) 2−4 radial distribution function, and (c) 4−4 radial distribution function. The PIMD results are shown in red and the n-CG-PI results are shown in blue.

RDFs. Note that the 2−2, 2−4, and 4−4 pair interactions are explicitly included in the n-CG-PI water model as shown in eq 33 and Figure 7. Figure 11 shows improved agreement between the n-CG-PI and PIMD RDFs, especially compared to what is shown in Figure 10. Moreover, we see good agreement even for the intramolecular O−H peaks (∼1 Å in Figure 11b) and H−H peaks (∼1.6 Å in Figure 11c). The better agreement between the two sets of RDFs shown in Figure 11 can be explained by the fact the cent−cent pair interactions are explicitly included in the n-CG-PI water model.

Since there are four different types of pseudoparticles in the mapped PIMD and the n-CG-PI water systems, there are 10 possible pair combinations. Figure 12 shows pair correlations of cent−obsv pseudoparticles. For the 1−2 and 3−4 RDFs, there are nonzero peaks centered at 0 Å, but not shown in this

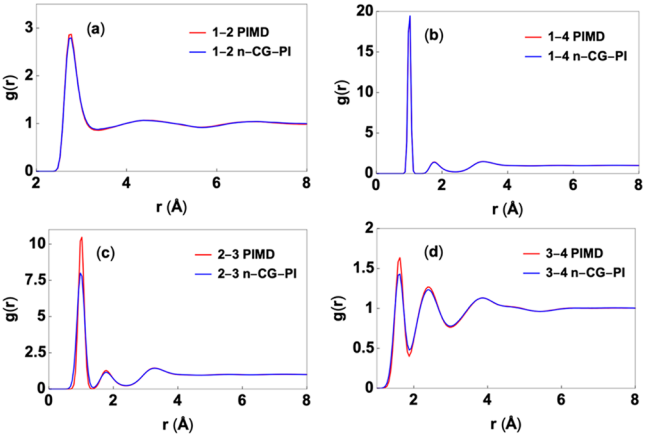


Figure 12. Calculated (a) 1−2 radial distribution function, (b) 1−4 radial distribution function, (c) 2−3 radial distribution function, and (d) 3−4 radial distribution function. The PIMD results are shown in red, and the n-CG-PI results are shown in blue.

https://doi.org/10.1021/acs.jpca.2c04349
J. Phys. Chem. A 2022, 126, 6004−6019

figure as it is redundant to Figure 9a and 9b. Similar to Figures 10 and 11, we see that the n-CG-PI RDFs agree well with the PIMD RDFs. Particularly, the 1−2 and 1−4 RDFs show almost quantitative agreement between the n-CG-PI and the PIMD results, whereas the 2−3 and 3−4 RDFs show disparities between the two. Regardless of these differences, the n-CG-PI RDFs still capture the overall shape of the PIMD pair correlations.

The nonbonded angle distribution $P(\theta)$ given in eq 28 shows three-body correlations. Even though the qSPC/Fw does not contain any many-body terms, it still is beneficial to study the three-body correlations to test how the n-CG-PI captures many-body interactions. Figure 13 shows the

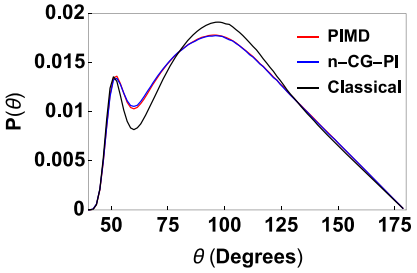


Figure 13. Calculated nonbonded angle distribution function for oxygen atoms with cutoff distance of r_0 = 3.5 Å. The PIMD results are shown in red, the n-CG-PI results are shown in blue, and the classical MD results are shown in black.

nonbonded angle distributions with respect to the oxygen observable pseudoparticle for the PIMD and the n-CG-PI simulations and oxygen atom for the classical MD simulation. We see that the n-CG-PI distribution well captures the PIMD distribution. This result suggests that the n-CG-PI correctly captures higher order water correlations by only force matching the pair intermolecular terms.

Overall, the n-CG-PI water model describes the structural correlations of the reference PIMD simulations with a good degree of accuracy. As opposed to the para-hydrogen model described earlier, the water system is more complex due to the presence of more types of atoms and extra intra- and intermolecular interactions. Different combinations of the CG methods to capture these different types of intra- and intermolecular interactions have been attempted, but the results were either numerically unstable or nonphysical. The structural correlations calculated from models that were numerically stable gave poor agreement with the PIMD results. The last thing to note is that while we were not able to find a n-CG-PI modeling scheme that performs better than the current version, such a scheme is clearly possible. For example, using iterative Boltzmann inversion (IBI)[87] to improve the agreement in intramolecular distributions could be the first step in improving the current model. Further discussion on the n-CG-PI modeling strategy is given in the next section.

4. CONCLUSIONS

Previous development of the CG-PI theory (both one and many-body) constructed an alternative isomorphism between the quantum partition function of N interacting distinguishable particles and the classical partition function of $2N$ (N pairs) pseudoparticles. By doing so, we were able to express the quantum delocalization of particles via the effective harmonic interaction between the observable and centroid pseudoparticles in the Taylor expanded expression of the CG potential $V_{CG,N}$. However, despite the success of this strategy, the previous state of CG-PI theory lacked the proper theoretical formalisms to appropriately describe realistic molecular systems that have explicit intramolecular interactions in the force field and possible many-body terms.

In this work, we presented a Taylor expanded expression of the many-body CG potential that includes intramolecular interactions as shown in eq 10. Note that even though the many-body potential V_z in this work was intended to capture the intramolecular interactions for bonded systems, it can be generalized to include any other many-body terms that may be involved in the force field. It is shown that these many-body terms in the force field contribute to quadratic coupling between the pseudoparticles analogous to the pairwise decomposable terms. The new Taylor expanded result shown in the square bracket of eq 11 suggests that the N pseudoparticle pair isomorphism of the CG-PI theory can be generalized to any form of force field and interaction between the particles. The new derivation, however, is still numerically challenging to implement directly and does not directly suggest how to implement the idea of CG-PI theory for realistic molecular systems.

Previously, we have developed a naïve n-CG-PI model by making an approximation to the Taylor expanded effective Hamiltonian: only the observable and centroid pseudoparticles in the same pair are harmonically coupled to each other, and the spring constant is parametrized by λ. We used this method to effectively capture the structural correlations of the modified liquid para-hydrogen system. However, this method still had significant shortcomings. Not only did it require an extensive set of path integral calculations to build the scaling relationship for λ but also it had no formalism to account for more than one type of atom.

To overcome these issues, we have developed a new version of the n-CG-PI model here. For a molecule described by a force field with n bonded atoms, the n-CG-PI "molecule" is composed of n centroid pseudoparticles and n observable pseudoparticles. Between the pseudoparticles, there are three types of interactions: (a) obsv−cent harmonic intramolecular, (b) cent−cent intramolecular, and (c) cent−cent intermolecular. Note that the obsv−cent intramolecular interactions are only harmonic, whereas the interactions that only involve centroid pseudoparticles have the functional form analogous to the original force field. While this coupling scheme is similar to that of the naïve n-CG-PI model, the new n-CG-PI model distinguishes itself by parametrizing the three types of interaction from a single PIMD trajectory rather than a predetermined scaling relationship. More specifically, we use force matching to parametrize the cent−cent interactions (both intra- and intermolecular) and Boltzmann inversion to parametrize the obsv−cent interactions.

With this new modeling scheme, we revisited the two liquid systems studied in our previous work: modified liquid para-hydrogen and liquid water. We studied the intra- and intermolecular structural correlations of the two systems to verify the efficacy of our new n-CG-PI models. For the liquid para-hydrogen system, the n-CG-PI model quantitatively captures the pair correlations as well as the nonbonded angle distributions calculated from PIMD simulations. For the liquid water system, the n-CG-PI model captures the general shapes of the PIMD distributions, quantitatively for some cases. The disparity between the two systems can be explained by the

https://doi.org/10.1021/acs.jpca.2c04349
J. Phys. Chem. A 2022, 126, 6004−6019

more complex intramolecular structure and greater degree of NQE for liquid water (especially internal modes). Overall, and especially compared to the naïve n-CG-PI results, the new n-CG-PI model provides a noticeable improvement in capturing the structural correlations of PIMD systems by using CG modeling techniques.

To further develop the CG-PI theory and the n-CG-PI method as a reliable and computationally manageable modeling tool to incorporate NQEs into the simulation of condensed matter systems, several aspects of the theory can be improved upon. (a) Even though water includes a more complex intramolecular system compared to para-hydrogen, it still is a relatively simple and small molecule. It would be important to test the n-CG-PI modeling scheme on more complex molecular systems and study their structural correlations. Moreover, as shown in eq 24, the obsv−cent intramolecular interaction has quadratic coupling (based on Taylor expanded result), but no such assumption of the cent−cent intramolecular interactions are made. The n-CG-PI modeling scheme should be tested for systems with more general forms of intramolecular interactions, Morse oscillator for example, to verify whether the force matched cent−cent intramolecular interactions have analogous functional forms to those in the original force field with different parameters. (b) The ring polymer contraction scheme evaluates the slowly varying long-range interactions with contracted (smaller number of beads) ring polymers, and evaluates the rapidly varying short-range interactions with noncontracted ring polymers.[67] Such a contraction scheme can accelerate the path integral simulation significantly while still capturing the delocalization of the system. Developing the CG-PI theory and n-CG-PI modeling scheme that can directly incorporate such different resolution path integral simulations will particularly be beneficial if the system of interest is large and/or has different degrees of NQE within the system. (c) As shown in eq 23b, there is only one nonbonded pair interactions between the centroid pseudoparticles. This suggests that the contribution from both the dispersion and the electrostatic interactions are accounted for in the force matched cent−cent pair potentials. This results in the parametrized pair potentials as shown in Figure 7, which correctly predicted the cent−cent pair correlations of the mapped PIMD system as shown in Figure 11. However, these pair potentials provide little physical intuition. To resolve this, the theoretical formalism of charge for different pseudoparticles might be worked out. By assigning appropriate charges to the pseudoparticles, there will be two types of nonbonded pair interactions in the n-CG-PI model which then will provide a more intuitive understanding of the intermolecular interactions between the centroid pseudoparticles. (d) Even though the n-CG-PI models contained in this work were parametrized from closed ring PIMD simulations (start of the imaginary path is identical to the end), one can directly parametrize the many-body CG potential shown in eq 11 by using open ring PIMD simulations (start of the imaginary path is not identical to the end). By parametrizing the $\boldsymbol{\Delta q}$ term, one could calculate the off-diagonal thermal density matrix element which then can be used to calculate the expectation values of momentum dependent operators. (e) Last, a set of thermodynamic estimators in the context of CG-PI theory needs to be derived. Our previous work[71] contains an expression for the CG-PI energy estimator (eq 13 from that paper). However, estimators for other thermodynamic quantities such as pressure need to be derived in order for the CG-PI theory to be used to calculate equilibrium quantities other than structural correlations.

To conclude, we present a general form of Taylor expanded many-body CG potential that explicitly includes the intramolecular interactions in the original force field. The resulting functional forms are analogous to the previous derivations with additional coupling terms that arise from the intramolecular terms. Furthermore, based on the second order Taylor expanded expression, we have constructed a n-CG-PI modeling scheme applicable to general molecular systems with more than one type of atom. Two liquid systems are studied to test the efficacy of the n-CG-PI model. It is shown that both the n-CG-PI para-hydrogen and water model can capture the structural correlations and delocalization of ring polymers shown in the full PIMD simulations. In the larger picture, the theoretical developments in this paper show that an N particle obeying quantum Boltzmann statistics can be adequately represented by a simpler system of $2N$ total particles obeying effective classical mechanics. This is a dramatic simplification of the full discretized path integral system involving NP total path integral "beads", where P is commonly a rather large number. Moreover, we have shown that with judicious parametrization of the n-CG-PI model, one can well capture the structural correlations. Overall, both the CG-PI theory and the n-CG-PI models presented in this work therefore provide a stepping stone to systematically incorporate NQEs in the modeling of condensed matter systems in a computationally and conceptually efficient manner by only including a minimal number of additional degrees of freedom compared to its classical counterpart.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpca.2c04349.

Functional forms and derivation of eq 11 (PDF)

Cent−cent pair potential for the n-CG-PI para-hydrogen model (PDF)

2−2 pair potential for the n-CG-PI water model (PDF)

2−4 pair potential for the n-CG-PI water model (PDF)

4−4 pair potential for the n-CG-PI water model (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Gregory A. Voth − *Department of Chemistry, James Franck Institute, and Institute for Biophysical Dynamics, The University of Chicago, Chicago, Illinois 60637, United States;* orcid.org/0000-0002-3267-6748; Email: gavoth@uchicago.edu

Author

Won Hee Ryu − *Department of Chemistry, James Franck Institute, and Institute for Biophysical Dynamics, The University of Chicago, Chicago, Illinois 60637, United States;* orcid.org/0000-0001-5484-9905

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpca.2c04349

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This material is based on work supported by the National Science Foundation (NSF Grant CHE-2102677). The computational resources from the University of Chicago Research Computing Center (RCC) were used for calculations presented. We thank Dr. Siyoung Kim and Dr. Jaehyeok Jin for help with the n-CG-PI water model.

■ REFERENCES

(1) Stern, H. A.; Berne, B. J. Quantum effects in liquid water: Path-integral simulations of a flexible and polarizable ab initio model. *J. Chem. Phys.* **2001**, *115*, 7622−7628.
(2) Shiga, M.; Shinoda, W. Calculation of heat capacities of light and heavy water by path-integral molecular dynamics. *J. Chem. Phys.* **2005**, *123*, 134502.
(3) Paesani, F.; Zhang, W.; Case, D. A.; Cheatham, T. E., III; Voth, G. A. An accurate and simple quantum model for liquid water. *J. Chem. Phys.* **2006**, *125*, 184507.
(4) Morrone, J. A.; Car, R. Nuclear Quantum Effects in Water. *Phys. Rev. Lett.* **2008**, *101*, 017801.
(5) Habershon, S.; Markland, T. E.; Manolopoulos, D. E. Competing quantum effects in the dynamics of a flexible water model. *J. Chem. Phys.* **2009**, *131*, 024501.
(6) Vega, C.; Conde, M. M.; McBride, C.; Abascal, J. L. F.; Noya, E. G.; Ramirez, R.; Sesé, L. M. Heat capacity of water: A signature of nuclear quantum effects. *J. Chem. Phys.* **2010**, *132*, 046101.
(7) Pamuk, B.; Soler, J. M.; Ramírez, R.; Herrero, C. P.; Stephens, P. W.; Allen, P. B.; Fernández-Serra, M. V. Anomalous Nuclear Quantum Effects in Ice. *Phys. Rev. Lett.* **2012**, *108*, 193003.
(8) Fritsch, S.; Potestio, R.; Donadio, D.; Kremer, K. Nuclear Quantum Effects in Water: A Multiscale Study. *J. Chem. Theory Comput.* **2014**, *10*, 816−824.
(9) Wang, L.; Ceriotti, M.; Markland, T. E. Quantum fluctuations and isotope effects in ab initio descriptions of water. *J. Chem. Phys.* **2014**, *141*, 104502.
(10) Litman, Y.; Donadio, D.; Ceriotti, M.; Rossi, M. Decisive role of nuclear quantum effects on surface mediated water dissociation at finite temperature. *J. Chem. Phys.* **2018**, *148*, 102320.
(11) Ceperley, D. M. Path integrals in the theory of condensed helium. *Rev. Mod. Phys.* **1995**, *67*, 279−355.
(12) Hone, T. D.; Voth, G. A. A centroid molecular dynamics study of liquid para-hydrogen and ortho-deuterium. *J. Chem. Phys.* **2004**, *121*, 6412−6422.
(13) Chen, J.; Li, X. Z.; Zhang, Q.; Probert, M. I. J.; Pickard, C. J.; Needs, R. J.; Michaelides, A.; Wang, E. Quantum simulation of low-temperature metallic liquid hydrogen. *Nat. Commun.* **2013**, *4*, 2064.
(14) Morales, M. A.; McMahon, J. M.; Pierleoni, C.; Ceperley, D. M. Nuclear Quantum Effects and Nonlocal Exchange-Correlation Functionals Applied to Liquid Hydrogen at High Pressure. *Phys. Rev. Lett.* **2013**, *110*, 065702.
(15) Herdman, C. M.; Rommal, A.; Del Maestro, A. Quantum Monte Carlo measurement of the chemical potential of 4He. *Phys. Rev. B* **2014**, *89*, 224502.
(16) Tuckerman, M. E.; Marx, D.; Klein, M. L.; Parrinello, M. On the Quantum Nature of the Shared Proton in Hydrogen Bonds. *Science* **1997**, *275*, 817−820.
(17) Walker, B.; Michaelides, A. Direct assessment of quantum nuclear effects on hydrogen bond strength by constrained-centroid ab initio path integral molecular dynamics. *J. Chem. Phys.* **2010**, *133*, 174306.
(18) Ceriotti, M.; Cuny, J.; Parrinello, M.; Manolopoulos, D. E. Nuclear quantum effects and hydrogen bond fluctuations in water. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 15591.
(19) Lobaugh, J.; Voth, G. A. The quantum dynamics of an excess proton in water. *J. Chem. Phys.* **1996**, *104*, 2056−2069.
(20) Schmitt, U. W.; Voth, G. A. The computer simulation of proton transport in water. *J. Chem. Phys.* **1999**, *111*, 9361−9381.
(21) Vuilleumier, R.; Borgis, D. Transport and spectroscopy of the hydrated proton: A molecular dynamics study. *J. Chem. Phys.* **1999**, *111*, 4251−4266.
(22) Marx, D.; Tuckerman, M. E.; Parrinello, M. Solvated excess protons in water: quantum effects on the hydration structure. *J. Condens. Matter Phys.* **2000**, *12*, A153−A159.
(23) Billeter, S. R.; Webb, S. P.; Iordanov, T.; Agarwal, P. K.; Hammes-Schiffer, S. Hybrid approach for including electronic and nuclear quantum effects in molecular dynamics simulations of hydrogen transfer reactions in enzymes. *J. Chem. Phys.* **2001**, *114*, 6925−6936.
(24) Feynman, R. P. *Quantum Mechanics and Path Integrals*; McGraw-Hill: Mineola, NY, 1965.
(25) Feynman, R. P. *Statistical Mechanics: A Set of Lectures*; W.A. Benjamin: Reading, MA, 1972.
(26) Chandler, D.; Wolynes, P. G. Exploiting the isomorphism between quantum theory and classical statistical mechanics of polyatomic fluids. *J. Chem. Phys.* **1981**, *74*, 4078−4095.
(27) Berne, B. J.; Thirumalai, D. On the Simulation of Quantum Systems: Path Integral Methods. *Annu. Rev. Phys. Chem.* **1986**, *37*, 401−424.
(28) Parrinello, M.; Rahman, A. Study of an F center in molten KCl. *J. Chem. Phys.* **1984**, *80*, 860−867.
(29) Tuckerman, M. E.; Berne, B. J.; Martyna, G. J.; Klein, M. L. Efficient molecular dynamics and hybrid Monte Carlo algorithms for path integrals. *J. Chem. Phys.* **1993**, *99*, 2796−2808.
(30) Marx, D.; Parrinello, M. Ab initio path integral molecular dynamics: Basic ideas. *J. Chem. Phys.* **1996**, *104*, 4077−4082.
(31) Tuckerman, M. E.; Marx, D.; Klein, M. L.; Parrinello, M. Efficient and general algorithms for path integral Car−Parrinello molecular dynamics. *J. Chem. Phys.* **1996**, *104*, 5579−5588.
(32) Herman, M. F.; Bruskin, E. J.; Berne, B. J. On path integral Monte Carlo simulations. *J. Chem. Phys.* **1982**, *76*, 5150−5155.
(33) Thirumalai, D.; Hall, R. W.; Berne, B. J. A path integral Monte Carlo study of liquid neon and the quantum effective pair potential. *J. Chem. Phys.* **1984**, *81*, 2523−2527.
(34) Diraison, M.; Martyna, G. J.; Tuckerman, M. E. Simulation studies of liquid ammonia by classical ab initio, classical, and path-integral molecular dynamics. *J. Chem. Phys.* **1999**, *111*, 1096−1103.
(35) Sindzingre, P.; Klein, M. L.; Ceperley, D. M. Path-integral Monte Carlo study of low-temperature 4He clusters. *Phys. Rev. Lett.* **1989**, *63*, 1601−1604.
(36) Scharf, D.; Klein, M. L.; Martyna, G. J. Path-integral Monte Carlo studies of para-hydrogen clusters. *J. Chem. Phys.* **1992**, *97*, 3590−3599.
(37) Sprik, M.; Impey, R. W.; Klein, M. L. Study of electron solvation in liquid ammonia using quantum path integral Monte Carlo calculations. *J. Chem. Phys.* **1985**, *83*, 5802−5809.
(38) Coker, D. F.; Berne, B. J.; Thirumalai, D. Path integral Monte Carlo studies of the behavior of excess electrons in simple fluids. *J. Chem. Phys.* **1987**, *86*, 5689−5702.
(39) Wallqvist, A.; Thirumalai, D.; Berne, B. J. Path integral Monte Carlo study of the hydrated electron. *J. Chem. Phys.* **1987**, *86*, 6404−6418.
(40) Cao, J.; Voth, G. A. A new perspective on quantum time correlation functions. *J. Chem. Phys.* **1993**, *99*, 10070−10073.
(41) Cao, J.; Voth, G. A. The formulation of quantum statistical mechanics based on the Feynman path centroid density. II. Dynamical properties. *J. Chem. Phys.* **1994**, *100*, 5106−5117.
(42) Cao, J.; Voth, G. A. The formulation of quantum statistical mechanics based on the Feynman path centroid density. III. Phase space formalism and analysis of centroid molecular dynamics. *J. Chem. Phys.* **1994**, *101*, 6157−6167.
(43) Cao, J.; Voth, G. A. The formulation of quantum statistical mechanics based on the Feynman path centroid density. IV. Algorithms for centroid molecular dynamics. *J. Chem. Phys.* **1994**, *101*, 6168−6183.

https://doi.org/10.1021/acs.jpca.2c04349
J. Phys. Chem. A 2022, 126, 6004−6019

(44) Jang, S.; Voth, G. A. Path integral centroid variables and the formulation of their exact real time dynamics. *J. Chem. Phys.* **1999**, *111*, 2357−2370.
(45) Jang, S.; Voth, G. A. A derivation of centroid molecular dynamics and other approximate time evolution methods for path integral centroid variables. *J. Chem. Phys.* **1999**, *111*, 2371−2384.
(46) Craig, I. R.; Manolopoulos, D. E. Quantum statistics and classical mechanics: Real time correlation functions from ring polymer molecular dynamics. *J. Chem. Phys.* **2004**, *121*, 3368−3373.
(47) Braams, B. J.; Manolopoulos, D. E. On the short-time limit of ring polymer molecular dynamics. *J. Chem. Phys.* **2006**, *125*, 124105.
(48) Richardson, J. O.; Althorpe, S. C. Ring-polymer molecular dynamics rate-theory in the deep-tunneling regime: Connection with semiclassical instanton theory. *J. Chem. Phys.* **2009**, *131*, 214106.
(49) Habershon, S.; Manolopoulos, D. E.; Markland, T. E.; Miller, T. F. Ring-Polymer Molecular Dynamics: Quantum Effects in Chemical Dynamics from Classical Trajectories in an Extended Phase Space. *Annu. Rev. Phys. Chem.* **2013**, *64*, 387−413.
(50) Jang, S.; Sinitskiy, A. V.; Voth, G. A. Can the ring polymer molecular dynamics method be interpreted as real time quantum dynamics? *J. Chem. Phys.* **2014**, *140*, 154103.
(51) Rossi, M.; Ceriotti, M.; Manolopoulos, D. E. How to remove the spurious resonances from ring polymer molecular dynamics. *J. Chem. Phys.* **2014**, *140*, 234116.
(52) Poulsen, J. A.; Rossky, P. J. Path integral centroid molecular-dynamics evaluation of vibrational energy relaxation in condensed phase. *J. Chem. Phys.* **2001**, *115*, 8024−8031.
(53) Paesani, F.; Xantheas, S. S.; Voth, G. A. Infrared Spectroscopy and Hydrogen-Bond Dynamics of Liquid Water from Centroid Molecular Dynamics with an Ab Initio-Based Force Field. *J. Phys. Chem. B* **2009**, *113*, 13118−13130.
(54) Witt, A.; Ivanov, S. D.; Shiga, M.; Forbert, H.; Marx, D. On the applicability of centroid and ring polymer path integral molecular dynamics for vibrational spectroscopy. *J. Chem. Phys.* **2009**, *130*, 194510.
(55) Paesani, F.; Voth, G. A. A quantitative assessment of the accuracy of centroid molecular dynamics for the calculation of the infrared spectrum of liquid water. *J. Chem. Phys.* **2010**, *132*, 014105.
(56) Medders, G. R.; Paesani, F. Infrared and Raman Spectroscopy of Liquid Water through "First-Principles" Many-Body Molecular Dynamics. *J. Chem. Theory Comput.* **2015**, *11*, 1145−1154.
(57) Moberg, D. R.; Straight, S. C.; Knight, C.; Paesani, F. Molecular Origin of the Vibrational Structure of Ice Ih. *J. Phys. Chem. Lett.* **2017**, *8*, 2579−2583.
(58) Reddy, S. K.; Moberg, D. R.; Straight, S. C.; Paesani, F. Temperature-dependent vibrational spectra and structure of liquid water from classical and quantum simulations with the MB-pol potential energy function. *J. Chem. Phys.* **2017**, *147*, 244504.
(59) Hunter, K. M.; Shakib, F. A.; Paesani, F. Disentangling Coupling Effects in the Infrared Spectra of Liquid Water. *J. Phys. Chem. B* **2018**, *122*, 10754−10761.
(60) Lobaugh, J.; Voth, G. A. A quantum model for water: Equilibrium and dynamical properties. *J. Chem. Phys.* **1997**, *106*, 2400−2410.
(61) Miller, T. F., III; Manolopoulos, D. E. Quantum diffusion in liquid para-hydrogen from ring-polymer molecular dynamics. *J. Chem. Phys.* **2005**, *122*, 184503.
(62) Miller, T. F., III; Manolopoulos, D. E. Quantum diffusion in liquid water from ring polymer molecular dynamics. *J. Chem. Phys.* **2005**, *123*, 154504.
(63) Medders, G. R.; Babin, V.; Paesani, F. Development of a "First-Principles" Water Potential with Flexible Monomers. III. Liquid Phase Properties. *J. Chem. Theory Comput.* **2014**, *10*, 2906−2910.
(64) Jang, S.; Voth, G. A. A relationship between centroid dynamics and path integral quantum transition state theory. *J. Chem. Phys.* **2000**, *112*, 8747−8757.
(65) Geva, E.; Shi, Q.; Voth, G. A. Quantum-mechanical reaction rate constants from centroid molecular dynamics simulations. *J. Chem. Phys.* **2001**, *115*, 9209−9222.
(66) Tuckerman, M. E.; Parrinello, M. Integrating the Car−Parrinello equations. II. Multiple time scale techniques. *J. Chem. Phys.* **1994**, *101*, 1316−1329.
(67) Markland, T. E.; Manolopoulos, D. E. An efficient ring polymer contraction scheme for imaginary time path integral simulations. *J. Chem. Phys.* **2008**, *129*, 024105.
(68) Noid, W. G. Perspective: Coarse-grained models for biomolecular systems. *J. Chem. Phys.* **2013**, *139*, 090901.
(69) Saunders, M. G.; Voth, G. A. Coarse-Graining Methods for Computational Biology. *Annu. Rev. Biophys.* **2013**, *42*, 73−93.
(70) Sinitskiy, A. V.; Voth, G. A. A reductionist perspective on quantum statistical mechanics: Coarse-graining of path integrals. *J. Chem. Phys.* **2015**, *143*, 094104.
(71) Ryu, W. H.; Han, Y.; Voth, G. A. Coarse-graining of many-body path integrals: Theory and numerical approximations. *J. Chem. Phys.* **2019**, *150*, 244103.
(72) Blinov, N.; Roy, P. N. Operator formulation of centroid dynamics for Bose−Einstein and Fermi−Dirac statistics. *J. Chem. Phys.* **2001**, *115*, 7822−7831.
(73) Blinov, N. V.; Roy, P. N.; Voth, G. A. Path integral formulation of centroid dynamics for systems obeying Bose−Einstein statistics. *J. Chem. Phys.* **2001**, *115*, 4484−4495.
(74) Roy, P. N.; Blinov, N. Centroid dynamics with quantum statistics. *Isr. J. Chem.* **2002**, *42*, 183−190.
(75) Han, Y.; Jin, J.; Wagner, J. W.; Voth, G. A. Quantum theory of multiscale coarse-graining. *J. Chem. Phys.* **2018**, *148*, 102335.
(76) Izvekov, S.; Voth, G. A. Multiscale coarse graining of liquid-state systems. *J. Chem. Phys.* **2005**, *123*, 134105.
(77) Izvekov, S.; Voth, G. A. A Multiscale Coarse-Graining Method for Biomolecular Systems. *J. Phys. Chem. B* **2005**, *109*, 2469−2473.
(78) Noid, W. G.; Chu, J.-W.; Ayton, G. S.; Krishna, V.; Izvekov, S.; Voth, G. A.; Das, A.; Andersen, H. C. The multiscale coarse-graining method. I. A rigorous bridge between atomistic and coarse-grained models. *J. Chem. Phys.* **2008**, *128*, 244114.
(79) Noid, W. G.; Liu, P.; Wang, Y.; Chu, J.-W.; Ayton, G. S.; Izvekov, S.; Andersen, H. C.; Voth, G. A. The multiscale coarse-graining method. II. Numerical implementation for coarse-grained molecular models. *J. Chem. Phys.* **2008**, *128*, 244115.
(80) Peng, Y.; Cao, Z.; Zhou, R.; Voth, G. A. Path Integral Coarse-Graining Replica Exchange Method for Enhanced Sampling. *J. Chem. Theory Comput.* **2014**, *10*, 3634−3640.
(81) Nava, M.; Quhe, R.; Palazzesi, F.; Tiwary, P.; Parrinello, M. de Broglie Swapping Metadynamics for Quantum and Classical Sampling. *J. Chem. Theory Comput.* **2015**, *11*, 5114−5119.
(82) Nava, M.; Palazzesi, F.; Perego, C.; Parrinello, M. Dimer Metadynamics. *J. Chem. Theory Comput.* **2017**, *13*, 425−430.
(83) Stillinger, F. H.; Weber, T. A. Computer simulation of local order in condensed phases of silicon. *Phys. Rev. B* **1985**, *31*, 5262−5271.
(84) Tschöp, W.; Kremer, K.; Batoulis, J.; Bürger, T.; Hahn, O. Simulation of polymer melts. I. Coarse-graining procedure for polycarbonates. *Acta Polym.* **1998**, *49*, 61−74.
(85) Müller-Plathe, F. Coarse-Graining in Polymer Simulation: From the Atomistic to the Mesoscopic Scale and Back. *ChemPhysChem* **2002**, *3*, 754−769.
(86) Reith, D.; Pütz, M.; Müller-Plathe, F. Deriving effective mesoscale potentials from atomistic simulations. *J. Comput. Chem.* **2003**, *24*, 1624−1636.
(87) Agrawal, V.; Arya, G.; Oswald, J. Simultaneous Iterative Boltzmann Inversion for Coarse-Graining of Polyurea. *Macromolecules* **2014**, *47*, 3378−3389.
(88) Silvera, I. F.; Goldman, V. V. The isotropic intermolecular potential for H2 and D2 in the solid and gas phases. *J. Chem. Phys.* **1978**, *69*, 4209−4213.
(89) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1−19.
(90) Bussi, G.; Parrinello, M. Accurate sampling using Langevin dynamics. *Phys. Rev. E* **2007**, *75*, 056707.

https://doi.org/10.1021/acs.jpca.2c04349
J. Phys. Chem. A 2022, 126, 6004−6019

(91) Ceriotti, M.; Parrinello, M.; Markland, T. E.; Manolopoulos, D. E. Efficient stochastic thermostatting of path integral molecular dynamics. *J. Chem. Phys.* **2010**, *133*, 124104.

https://doi.org/10.1021/acs.jpca.2c04349
J. Phys. Chem. A 2022, 126, 6004−6019