THE UNIVERSITY OF CHICAGO

REDUCTIONIST REPRESENTATION OF QUANTUM STATISTICS AND DYNAMICS USING COARSE GRAINING

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF CHEMISTRY

BY

WON HEE RYU

CHICAGO, ILLINOIS

DECEMBER 2021

For my mother, Sun Young Park (박선영), and my father, Keun Won Ryu (류근원)

Table of Contents

List of Figures

Figure 5-11 (page 150): Calculated Kubo-transformed position autocorrelation function using different methods for quartic potential with P = 32 and $\beta = 8$. In (a) and (b), the exact quantum results are shown in blue, CMD results are shown in yellow, RPMD results are shown in green, TRPMD results are shown in crimson. GLE/RPMD and GLE/TRPMD results are shown in purple in (a) and (b) respectively. Figure 5-11(c) compares GLE/RPMD and GLE/TRPMD which are shown in blue and yellow respectively.

Figure 5-12 (page 151): Calculated Kubo-transformed position autocorrelation function using different methods for quartic potential with P = 32 and $\beta = 1$. In (a) and (b), the exact quantum results are shown in blue, CMD results are shown in yellow, RPMD results are shown in green, TRPMD results are shown in crimson. GLE/RPMD and GLE/TRPMD results are shown in purple in (a) and (b) respectively. Figure 5-12(c) compares GLE/RPMD and GLE/TRPMD which are shown in blue and yellow respectively.

List of Tables

Acknowledgements

The old Korean saying goes: 잘 되면 내 탓, 안 되면 조상 탓. It badly translates to, without its original wit or rhyme, "I get the credit if things go well, and my ancestors get the blame if things go awry." I have been accused of such behavior in the past, but during graduate school I have become wiser and came to a tough personal realization that the opposite often is true. Before the verbose display of scientific accomplishments, it would be courteous to thank and appreciate those who have made it truly possible, aside from NSF, University of Chicago supercomputers, and myself.

First and foremost, none of this would have been possible without the support of Professor Gregory A. Voth. When I was but a young and helpless first year graduate student, eager to get started on a theoretical research without fully understanding what it entailed, he took me under his wings. During my time in the Voth group, I was gifted with his expertise and advice to pursue enthralling research projects on Feynman's imaginary time path integral formalism of quantum statistical mechanics. His guidance provided me with insight and opportunities to add my contribution to what seemed, and still seem, like monumental achievements made by previous generations of brilliant chemical physicists and theoretical chemists.

I also would like to thank my thesis committee members, Professor Vaikuntanathan and Professor Gagliardi for their time and advice during the dissertation process. I remember taking Professor Vaikuntanathan's statistical mechanics class during my first quarter at the University of Chicago. In his class, through various baffling thermodynamics questions I learned a valuable lesson as a theorist: the solutions to most problems in science (and in life) simply boils down to 'alge-bra.' I have not had the pleasure of talking to or taking a class from Professor Gagliardi in

real life due to the pandemic. However, she has been a delightful member of the CCTCh (Chicago Center for Theoretical Chemistry) since she has joined the faculty. I am sincerely grateful for her generous gesture of joining the thesis committee for a student who she has not met yet.

Dr. Yining Han has been a crucial part of my Ph.D. life and research. He not only guided me through my first and second in the Voth group but also mentored me through my first author paper (Chapter 3 of this dissertation), in which he is the second author. I was and still am inspired by his dedication and rigor towards research, in both the theoretical chemistry subfield and various other aspects of physics and chemistry.

Both Dr. Siyoung Kim and Dr. Jaehyeok Jin have been good friends in the Voth group, particularly before and after the group meeting presentations. I remember the strong sense of comradery that warmed me as the three of us walked amid pungent Chicago wind to the nearest café to get coffee. Both Dr. Kim and Dr. Jin would have defended their thesis before Oct 27th, and I wish them the best of luck in their defense and future career.

I also would like to show my gratitude for other Voth group members, both past and present. Their names are clumped up in a list not because I appreciate them any less but because there is not enough space. Professor Glen Hocky and Professor Alek Pak have been an inspiration of mine during their stay in the Voth group. I will always be impressed with how eloquently yet simply they presented their discoveries while maintaining scientific rigor. Dr. Thomas Dannenhoffer-Lafage and Dr. Paul Calio have guided me through working on the reactive coarse-graining project. Even though that work has not been developed enough for to be included in the dissertation, their guidance is appreciated regardless. Dr. Laura Watkins has been my friendly desk neighbor for about two years before the pandemic. Her dissertation defense is

only a few hours later than mine, and I hope her all the best. Dr. Zhefu Li and Dr. Chenghan Li are my cohorts within the Voth group whom I worked on Prof. Vaikuntanathan's statistical mechanics problem set with during first year of graduate school. Both have successfully defended their dissertations and have moved on to the next chapter of their life, and I wish them the best of luck. Lastly, I would like to thank Tim Loose and Josh Zuchniarz who I have worked with on the XSEDE proposals for the past three years. They have been meticulously and successfully keeping up with a hectic assignment and deserve a sincere gratitude for enabling a large portion of Voth group's quality scientific research.

Britt Fossum, a University of Chicago classmate of mine, has been an integral part of my early years of graduate school. I was fascinated with her character and knowledge of random trivia facts, particularly from the 80s, that I ended up learning via osmosis. She was one of few truly good friends I have made during graduate school and I hope that our friendship stands the test of time. Even though she took a leave of absence from the Ph.D. program, she is planning on starting a Master's program next winter. I wish her the best of luck during the second round of graduate school.

Dorrah Alharbi, my significant other, proudly deserves a lengthy paragraph in the acknowledgement section. Not only was she welcoming during my moments of frustration and weakness from research, but also she has provided me with seemingly endless love, pep talks, and place in her heart for me. In particular, she has helped me work through the GLE/RPMD project (Chapter 6 of this dissertation) which went through numerous revisions and alterations. She got me out of the rut when I felt like I could no longer work on the project. Dorrah herself is planning to apply to graduate school this winter. I am more than certain that she will be able to get into a program of her choice and eventually become a great scholar in art history.

Lastly, I am forever in debt to my parents, Sun Young Park and Dr. Keun Won Ryu. I was lucky enough to visit them in Korea for five weeks this summer. I left Korea in Summer of 2012 for college and have not visited them for such an extended amount of time since then. During my stay, I realized how much they work and dedication they put into in name of family and, most importantly, for the sake of my (and my brother's) education. I could not see this when I was in college, but now that I am a little bit wiser, I saw, felt, and was moved by their dedication. Granted going through graduate school and working on your doctorate degree can be challenging and frustrating at times, but one should also know that to be able to solely focus on academics without worrying about the pressures that come from other aspects of life is only possible because of someone else's sacrifice, whether that is your parents, spouse, or friends. Therefore it is by no extent an understatement that even though the doctorate degree in theoretical chemistry that I defend on Oct 27th 2021 will only be given to Won Hee Ryu, a wise person will see that underneath my name are numerous etchings of the names that made this dissertation possible in the first place.

Chapter 1: Introduction

I. Nuclear Quantum Effect and Imaginary Time Path Integral

In classical molecular dynamics simulation, a system is described as a set of atoms or point particles that interact with each other through an empirical force field. The force field is typically obtained from some higher level theory calculations and are fitted to reproduce different types of thermodynamic quantities or structural factors. Such strategy is based on the Born-Oppenheimer approximation that approximates the nuclear degrees of freedom to be decoupled from the electronic ones. As a corollary, in this context, we treat electrons with quantum mechanics and nuclei with classical mechanics. This separation with two resolution and methodologies has worked successfully to describe various condensed matter systems as large as biophysical systems such as actin filaments.[1]

However, if a system contains light nuclei and/or is at low temperature, the Born-Oppenheimer approximation can break down. For such cases, the quantum mechanics of the nuclei must be explicitly considered to accurately describe both the statistics and dynamics of the system. This manifestation of quantum character of the nuclei has been referred to as nuclear quantum effect (NQE). In the scope of computational chemistry, accounting for NQE in computer simulations has shown to be critical to correctly describe various phenomena ranging from hydrogen bonding[2-4] to reaction rates.[5-7]

The Feynman imaginary time path integral formalism of quantum statistical mechanics has been an integral theoretical formalism that enables us to incorporate NQE in computational simulations.[8, 9] In particular, the quantum-classical isomorphism, a corollary of the imaginary time path integral theory, suggests that the quantum thermal density matrix can be re-expressed as the classical thermal density matrix of a discrete representation of an imaginary time path, or a

ring polymer.[10, 11] This result builds an isomorphism between a particle that obeys quantum Boltzmann statistics and a ring polymer in the extended phase space obeying quantum mechanics. Furthermore, the corollary also suggests that one can obtain strictly quantum information by studying the classical ring polymer using well-developed classical methods such as molecular dynamics (MD) and Monte Carlo (MC) methods. Such methods, respectively named path integral MD (PIMD)[12, 13] and path integral MC (PIMC)[13, 14] have successfully been used to study the nuclear quantum effects.

One important development in the PIMD based methods has been to extend the formalism into the realm of dynamics. The original imaginary time path integral formalism does not provide any dynamical information as there is no connection between the dynamics of the ring polymer to the real-time quantum dynamics, but there have been attempts to correlate the two. Even though there have been development of various approximate quantum dynamical methods such as semiclassical initial value representation[15] and Matsubara dynamics[16], the main topic of our discussion will be on centroid molecular dynamics (CMD),[17, 18] ring polymer molecular dynamics (RPMD),[19, 20] and thermostatted RPMD (TRPMD)[21] in this dissertation. In CMD, one constructs an effective potential of mean force for the centroid (center of mass of the ring polymer) and considers the dynamics of the centroid on the effective potential surface. RPMD considers the dynamics of the ring polymers to be physical based on the symmetry considerations of the ring polymer correlation function and the real time Kubo transformed correlation functions. The TRPMD, an extension of RPMD method, attaches a Langevin thermostat to the system propagation of RPMD to address the “resonance” problem of RPMD which has been observed when the frequency of the physical potential and the frequency of the intramolecular ring polymer fluctuations comes in resonance.[22, 23]

Even though the path integral methods described above are well established with notable success, they are not without shortcomings or rooms for improvement. There are two noticeable shortcomings that will be the focus of this dissertation. When calculating quantum statistics, we are able to obtain numerically accurate result using both classical mechanics and quantum-classical isomorphism. However, in general the extended phase space of the ring polymers poses a significant sampling challenge. Moreover, the shortcomings are more serious in the realm of dynamics. So far the quantum dynamical methods have been approximations of real time dynamics and to this date there have been no satisfactory method that builds a solid connection between the two.

II. This Thesis

To further understand the behavior of the classical ring polymers in both statistics and dynamics, we will use two classical statistical mechanics methods in this dissertation: coarse graining (CG) and generalized Langevin equation (GLE). The two theories have similar philosophy which is to only focus on few collective variables when studying a highly many-body and multidimensional systems. However, the details differ slightly. Coarse-graining represents an all-atom system, or some type of more higher resolution model, with CG beads in which each CG bead corresponds to more than one atom.[24, 25] Similar to classical MD in atomistic resolution, the CG beads interact with each other in the CG PMF. By constructing the CG PMF and the CG model, one reduces the dimensionality of the original system, greatly decreasing the complexity, while capturing the essential physics of the system. On the other hand, GLE describes a dynamics of a system under the influence of external bath using an integro-differential equation in which we can determine different contributions from the bath modes to the dynamics of the

system.[26] These two classical ideas will be the theoretical foundation of the chapters described in this thesis.

The remainder of the thesis are structured as follows. In Chapter 2, we will discuss the theoretical basis for the Feynman's imaginary time path integral formalism of quantum statistical mechanics and quantum-classical isomorphism. In Chapter 3, the coarse-graining of path integral (CG-PI) theory that combines the quantum-classical isomorphism and CG mapping will be presented. In Chapter 4, we will further develop the CG-PI theory and present the numerical CG-PI method which uses classical CG modeling strategy to be used in many-body path integral systems. In Chapter 5, we will introduce the GLE/RPMD theory which uses the GLE formalism to analyze the dynamics of the ring polymer in RPMD. By dividing the ring polymer into system and bath, we are able to calculate the bath quantities and capture the essential dynamics of the full ring polymer. Chapter 6 will contain discussions and final concluding remarks.

References

1. J. W. Chu and G. A. Voth, Proc. Natl. Acad. Sci. USA. **102**, 13111 (2005).

2. M. E. Tuckerman, D. Marx, M. L. Klein and M. Parrinello, Science **275**, 817 (1997).

3. B. Walker and A. Michaelides, J. Chem. Phys. **133**, 174306 (2010).

4. M. Ceriotti, J. Cuny, M. Parrinello and D. E. Manolopoulos, Proc. Natl. Acad. Sci. USA. **110**, 15591 (2013).

5. U. W. Schmitt and G. A. Voth, J. Chem. Phys. **111**, 9361 (1999).

6. I. R. Craig and D. E. Manolopoulos, J. Chem. Phys. **122**, 084106 (2005).

7. R. Biswas, Y. L. S. Tse, A. Tokmakoff and G. A. Voth, J. Phys. Chem. B. **120**, 1793 (2016).

8. R. P. Feynman, *Quantum Mechanics and Path Integrals*. (McGraw-Hill, New York, 1965).

9. R. P. Feynman, *Statistical mechanics: A Set of Lectures* (W.A. Benjamin, Reading, Mass, 1972).

10. D. Chandler and P. G. Wolynes, J. Chem. Phys. **74**, 4078 (1981).

11. B. J. Berne and D. Thirumalai, Annu. Rev. Phys. Chem. **37**, 401 (1986).

12. M. Parrinello and A. Rahman, J. Chem. Phys. **80**, 860 (1984).

13. D. M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995).

14. M. E. Tuckerman, B. J. Berne, G. J. Martyna and M. L. Klein, J. Chem. Phys. **99**, 2796 (1993).

15. W. H. Miller, J. Phys. Chem. A. **105**, 2942 (2001).

16. T. J. H. Hele, M. J. Willatt, A. Muolo and S. C. Althorpe, J. Chem. Phys. **142**, 134103 (2015).

17. J. Cao and G. A. Voth, J. Chem. Phys. **100**, 5093 (1994).

18. J. Cao and G. A. Voth, J. Chem. Phys. **100**, 5106 (1994).

19. I. R. Craig and D. E. Manolopoulos, J. Chem. Phys. **121**, 3368 (2004).

20. B. J. Braams and D. E. Manolopoulos, J. Chem. Phys. **125**, 124105 (2006).

21. M. Rossi, M. Ceriotti and D. E. Manolopoulos, J. Chem. Phys. **140**, 234116 (2014).

22. S. Habershon, G. S. Fanourgakis and D. E. Manolopoulos, J. Chem. Phys. **129**, 074501 (2008).

23. A. Witt, S. D. Ivanov, M. Shiga, H. Forbert and D. Marx, J. Chem. Phys. **130**, 194510 (2009).

24. W. G. Noid, J. Chem. Phys. **139**, 090901 (2013).

25. M. G. Saunders and G. A. Voth, Annu. Rev. Biophys. **42**, 73 (2013).

26. R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957).

Chapter 2: Feynman's Imaginary Time Path Integral

This chapter contains a brief overview of Feynman's imaginary time path integral formalism of quantum statistical mechanics and the resulting quantum-classical isomorphism. In particular, we will discuss in detail the nomenclature of 'imaginary time' path integral by investigating the 'real time' path integral as well, an aspect of the theory that is not included in the following chapters for brevity. By no means is this a thorough overview of the topic. Instead, it will serve as a stepping stone and reference points for the formalisms developed throughout this thesis. Interested readers should visit references contained within this thesis. Lastly, it is stated here that we closely follow the organization of Chapter 12 of M. E. Tuckerman's *Statistical Mechanics: Theory and Molecular Simulation* in this chapter.

I. Imaginary Time Path Integral Expression of the Canonical Density Matrix

For a particle of mass m, Hamiltonian operator $\widehat{H} = \hat{p}^2/(2m) + \hat{V}(\hat{x})$, the canonical density matrix operator $\hat{\rho} = exp(-\beta\widehat{H})$ can be expressed in the position basis. Such thermal density matrix element is defined as

$$\rho(x, x'; \beta) \equiv \langle x' | e^{-\beta\widehat{H}} | x \rangle. \tag{1.1}$$

The general evaluation of this matrix element is nontrivial as the kinetic energy operator $\widehat{K} = \hat{p}^2/(2m)$ and the potential energy operator $\hat{V}(\hat{x})$ in general do not commute. However, one can obtain an approximation of the exact expression by using a Trotter theorem. By doing so, the density matrix operator can be expanded as

$$e^{-\beta(\widehat{K}+\widehat{V})} = \lim_{P\to\infty} \left[e^{-\beta\widehat{V}/(2P)} e^{-\beta\widehat{K}/P} e^{-\beta\widehat{V}/(2P)} \right]^P. \tag{1.2}$$

If we let

$$\widehat{\Omega} = e^{-\beta\widehat{V}/(2P)} e^{-\beta\widehat{K}/P} e^{-\beta\widehat{V}/(2P)}, \tag{1.3}$$

the thermal density matrix element can be written as

$$\rho(x,x';\beta) = \lim_{P\to\infty}\langle x'|\widehat{\Omega}^P|x\rangle. \tag{1.4}$$

To evaluate the Trotter expanded form of $\rho(x,x';\beta)$, we use an expression for the identity operator ($\hat{I}$)

$$\hat{I} = \int dx|x\rangle\langle x| \tag{1.5}$$

and insert the identity operator $P-1$ times between P number of $\widehat{\Omega}$ operators. Then, we can write

$$\rho(x,x';\beta) = \lim_{P\to\infty}\int dx_2 \dots dx_P\langle x'|\widehat{\Omega}|x_P\rangle\langle x_P|\widehat{\Omega}|x_{P-1}\rangle \dots \langle x_2|\widehat{\Omega}|x\rangle. \tag{1.6}$$

In this form, the individual elements $\langle x_{k+1}|\widehat{\Omega}|x_k\rangle$ can be analytically evaluated as

$$\langle x_{k+1}|\widehat{\Omega}|x_k\rangle = \sqrt{\frac{mP}{2\pi\beta\hbar^2}}\,exp\left(-\frac{\beta}{2P}\big(V(x_k)+V(x_{k+1})\big) - \frac{mP}{2\beta\hbar^2}(x_{k+1}-x_k)^2\right). \tag{1.7}$$

Using Eq. (1.7), the thermal density matrix element can be rewritten as

$$\rho(x,x';\beta) = \lim_{P\to\infty}\left(\frac{mP}{2\pi\beta\hbar^2}\right)^{P/2}\int dx_2 \dots dx_P$$
$$\times\, exp\left(-\frac{1}{\hbar}\sum_{k=1}^{P}\left(\frac{\beta\hbar}{2P}\big(V(x_k)+V(x_{k+1})\big) + \frac{mP}{2\beta\hbar}(x_{k+1}-x_k)^2\right)\right)\Big|_{\substack{x_{P+1}=x'\\ x_1=x}}. \tag{1.8}$$

This expression can further be simplified by noting that

$$\tfrac{1}{2}\sum_{k=1}^{P}\big(V(x_k)+V(x_{k+1})\big) = \sum_{k=1}^{P}V(x_k). \tag{1.9}$$

Using Eq. (1.9), we can simplify $\rho(x,x';\beta)$ to be

$$\rho(x,x';\beta) = \lim_{P\to\infty}\left(\frac{mP}{2\pi\beta\hbar^2}\right)^{P/2}\int dx_2 \dots dx_P$$
$$\times\, exp\left(-\frac{1}{\hbar}\sum_{k=1}^{P}\left(\frac{\beta\hbar}{P}V(x_k) + \frac{mP}{2\beta\hbar}(x_{k+1}-x_k)^2\right)\right)\Big|_{\substack{x_{P+1}=x'\\ x_1=x}}. \tag{1.10}$$

In this expression, the quantum kinetic energy is represented as a set of nearest-neighbor coupling between the intermediate beads which comes from evaluating the Trotter expansion term $\langle x_{k+1}|\widehat{\Omega}|x_k\rangle$.

From the discrete representation of $\rho(x, x'; \beta)$ in Eq. (1.10), we can derive the path integral expression for the canonical partition function Q for a particle

$$Q = \int dx\, \rho(x, x; \beta) \tag{1.11}$$

Since the integrand of Eq. (1.11) only calls for diagonal thermal density matrix element ($x = x'$), we integrate with respect to x the discrete representation of a closed path to have

$$Q = \lim_{P\to\infty} \left(\frac{mP}{2\pi\beta\hbar^2}\right)^{P/2} \int dx_1 \dots dx_P\, exp\left(-\beta V_P(x_1, \dots, x_P)\right) \tag{1.12}$$

Where $V_P(x_1, \dots, x_P)$ is the isomorphic potential

$$V_P(x_1, \dots, x_P) = \sum_{k=1}^{P} \frac{mP}{2\beta^2\hbar^2}(x_{k+1} - x_k)^2 + \frac{1}{P}V(x_k). \tag{1.13}$$

Before we move on to how to use the path integral expression of thermal density matrix element to calculate thermodynamics and expectation values, we discuss the origin of this formalism's nomenclature 'imaginary time path integral.' This comes from noticing that the canonical density matrix $\hat{\rho}(\beta) = exp\left(-\beta\hat{H}\right)$ is related to the time propagator operator $\hat{U}(t) = exp\left(-i\hat{H}t/\hbar\right)$ such that for $\hat{U}$ if we let time be imaginary ($t = -i\beta\hbar$) or temperature be imaginary ($\beta = it/\hbar$), we have equivalence between the two expressions as

$$\hat{\rho}(\beta) = \hat{U}(-i\beta\hbar) \tag{1.14}$$

and

$$\hat{U}(t) = \hat{\rho}(it/\hbar). \tag{1.15}$$

Since the operator $\hat{U}$ propagates a quantum state function in real time, the thermal density matrix can be understood to be a propagator of a quantum state in imaginary time. Moreover, this suggests that for the previously derived equations, we can obtain information about quantum dynamics by simply substituting the inverse temperature β with $it/\hbar$. Using Eq. (1.10) and letting $\beta = it/\hbar$, we have

$$U(x, x'; t)$$

$$= \lim_{P\to\infty} \left(\frac{mP}{2\pi i t\hbar}\right)^{P/2} \int dx_2 \dots dx_P exp\left(\frac{i}{\hbar}\sum_{k=1}^{P}\left(\begin{matrix}\frac{mP}{2t}(x_{k+1}-x_k)^2 \\ -\frac{t}{P}V(x_k)\end{matrix}\right)\right)\Bigg|_{\substack{x_{P+1}=x' \\ x_1=x}} \tag{1.16}$$

II. Operator Expectation Values and Thermodynamics

For a general operator $\hat{A}$, its expectation value can be written compactly as

$$\langle\hat{A}\rangle = \frac{Tr\left[\hat{A}e^{-\beta\hat{H}}\right]}{Tr\left[e^{-\beta\hat{H}}\right]} = \frac{\int dx \left\langle x\middle|\hat{A}e^{-\beta\hat{H}}\middle|x\right\rangle}{\int dx \left\langle x\middle|e^{-\beta\hat{H}}\middle|x\right\rangle}. \tag{1.17}$$

In this chapter, we will only consider three types of expectation values: of operators that only depend on position operator, of operators that only depend on momentum operator, and of thermodynamic functions.

If the operator $\hat{A}$ can be expressed as $\hat{A}(\hat{x}) = A(\hat{x})$, the position basis is an eigenvector of $\hat{A}$ which means that Eq. (17) can be written as

$$\langle\hat{A}\rangle = \frac{1}{Q}\lim_{P\to\infty}\left(\frac{mP}{2\pi\beta\hbar^2}\right)^{P/2} \int dx_1 \dots dx_P \; A_P(x_1, \dots, x_1) exp\big(-\beta V_P(x_1, \dots, x_1)\big) \tag{1.18}$$

where

$$A_P(x_1, \dots, x_1) = \frac{1}{P}\sum_{k=1}^{P} A(x_k). \tag{1.19}$$

Note that we only used the diagonal thermal density matrix elements and the isomorphic potential V_P has the boundary condition. This suggests that the isomorphic ring polymer is a closed ring polymer in which the first discrete representation of the imaginary path is same as that of the last. A_P is the estimator for the operator $\hat{A}$.

Consider an operator $\hat{B}$ defined in a similar manner as $\hat{A}$. It can be expressed as a function of momentum vector such that $\hat{B}(\hat{p}) = B(\hat{p})$. In this case, the position basis is not an

eigenvector of $\hat{B}$. For such cases, the expectation value of $\hat{B}$ using imaginary time path integral expression of $\rho(x, x'; \beta)$ can be written as

$$\langle \hat{B} \rangle = \frac{1}{Q} \lim_{P\to\infty} \left(\frac{mP}{2\pi\beta\hbar^2}\right)^{P/2} \int dx_1 \dots dx_{P+1} \left\langle x_1 \middle| \hat{B} \middle| x_{P+1} \right\rangle exp\left(-\beta V_P(x_1, \dots, x_{P+1})\right). \quad (1.20)$$

Note that this expression does not have the closed ring polymer boundary condition. Therefore, when calculating momenta dependent operator values, one needs to simulate an open ring polymer and take average value of the bra-ket quantity $\left\langle x_1 \middle| \hat{B} \middle| x_{P+1} \right\rangle$.

In general calculating the expectation values of operators that depend on both position and momenta are as complicated as the second example, which involves sampling the open ring polymer in the extended phase space. However, if the operator in question is a thermodynamic quantity, one can use thermodynamic relations to calculate expectation values while only sampling the closed ring polymers. For example, consider the expectation value of the total energy E which is $\langle \hat{p}^2/(2m) + \hat{V}(\hat{x}) \rangle$. To calculate this quantity by brute force, one would need both open and closed ring polymers. However, we can circumvent this challenge by using the thermodynamic relationship between total energy of the system and the canonical partition function which suggests

$$E = -\frac{\partial \ln Q}{\partial \beta} = \frac{1}{Q}\frac{\partial Q}{\partial \beta}. \quad (1.21)$$

Using the imaginary time path integral expression of the canonical partition function as shown in Eq. (1.12), we can write the expression for E and the energy estimator as

$$E = \frac{1}{Q} \lim_{P\to\infty} \left(\frac{mP}{2\pi\beta\hbar^2}\right)^{P/2} \int dx_1 \dots dx_P \, \epsilon_P(x_1, \dots, x_1) exp\left(-\beta V_P(x_1, \dots, x_1)\right) \quad (1.22)$$

and

$$\epsilon_P(x_1, \dots, x_1) = \frac{P}{2\beta} - \sum_{k=1}^{P} \frac{mP}{2\beta^2\hbar^2}(x_{k+1} - x_k)^2 + \frac{1}{P}V(x_k). \quad (1.23)$$

Chapter 3: Coarse-graining of Many-body Path Integrals: Theory and Numerical Approximations

This chapter is reproduced from J. Chem. Phys. **150**, 244103 (2019) with the permission of AIP Publishing with slight modification.

Abstract

Feynman's imaginary time path integral approach to quantum statistical mechanics provides a theoretical formalism for including nuclear quantum effects in simulation of condensed matter systems. Sinitskiy and Voth [J. Chem. Phys. **143**, 094104 (2015)] have presented the coarse-grained path integral (CG-PI) theory, which provides a reductionist coarse-grained representation of the imaginary time path integral based on the quantum-classical isomorphism. In this chapter, the many-body generalization of the CG-PI theory is presented. It is shown that the N interacting particles obeying quantum Boltzmann statistics can be represented as a system of N pairs of classical-like pseudo-particles coupled to each other analogous to the pseudo-particle pair of the one-body theory. Moreover, we present a numerical CG-PI (n-CG-PI) method by applying a simple approximation to the coupling scheme between the pseudo-particles due to numerical challenges of directly implementing the full many-body CG-PI theory. Structural correlations of two liquid systems are investigated to demonstrate the performance of the n-CG-PI method. Both the many-body CG-PI theory and the n-CG-PI method not only present reductionist picture of the many-body quantum Boltzmann statistics but also provide theoretical and numerical insight into how to explicitly incorporate nuclear quantum effects in condensed matter systems with minimal additional degrees of freedom.

I. Introduction

For systems involving light nuclei and/or at low temperature, nuclear quantum effects (NQEs) must be explicitly considered to accurately describe both the equilibrium and dynamic properties. The quantum mechanical treatment of the nuclei has shown to be important in describing various phenomena in chemistry and physics ranging from proton transfer,[1-4] condensed phase systems,[5-8] hydrogen bonding,[9, 10] and water[11-19] to name a few. Feynman's imaginary time path integral theory of quantum statistical mechanics[20, 21] is a powerful theoretical formalism to describe systems with such significant NQEs. It provides a basis for the quantum-classical isomorphism[22, 23] which suggests that a quantum particle can be represented as a classical ring polymer with infinite number of "beads". In other words, the properties of a quantum system can be described by those of a more complicated yet completely classical system.

Many well developed numerical path-integral techniques such as path integral molecular dynamics (PIMD)[5, 24-28] and path integral Monte Carlo (PIMC)[5, 25, 29] combine the imaginary time path integral formalism and the well-developed sampling algorithms of classical statistics. Both methods have been used to study the NQEs of different classes of systems such as liquids,[30, 31] para-hydrogen and helium clusters,[32, 33] and solvation of electrons.[34-36] Moreover, path integral based methods such as centroid molecular dynamics (CMD)[37-43] and ring polymer molecular dynamics (RPMD)[44-48] extend the imaginary time path integral to a dynamical regime and provide an approximation to the Kubo-transformed correlation functions. These methods have been used to study dynamical properties of condensed matter systems such as dynamic structure factors of liquids,[49] vibrational spectra of condensed matter systems,[50, 51] diffusion of liquids,[52, 53] and chemical reaction rates.[52-54] However, despite the conciseness of the theory and fast

sampling algorithms, path-integral based simulations are computationally more demanding than their classical counterparts due to the presence of multiple beads in the ring polymer, particularly for large systems with significant degree of NQEs. Therefore, a reductionist representation of the imaginary time path integral theory can provide not only valuable theoretical insight into the behavior of quantum Boltzmann statistics at finite temperature but also potential numerical path integral techniques with more computationally manageable costs while comparable accuracy.

Coarse-graining (CG) systematically decreases the degrees of freedom (DOF) of the system while maintaining the description of its essential physics,[55, 56] which suggests that it can be an appropriate tool in constructing a reductionist representation of the quantum-classical isomorphism. Specifically, bottom-up coarse-graining is a branch of CG methodology in which the CG models are constructed based on a more detailed fine-grained (FG), usually atomistic, models and aim to reproduce the underlying physics of the FG systems. Combining the idea of bottom-up CG and the imaginary time path integral formalism, the coarse-grained path integral (CG-PI) theory provides a reductionist CG representation of the off-diagonal thermal density matrix element for a quantum particle interacting with a potential energy source.[57] Moreover, using such representation, the CG-PI theory constructs an effective Hamiltonian that estimates the expectation value of observables at finite temperature. Compared to the conventional infinite ring polymer representation, the CG-PI theory gives an attractive reductionist perspective of the quantum-classical isomorphism. However, despite its appeal, it cannot yet be applied to realistic molecular systems as it is strictly a one-body theory.

There have been previous works that combine the CG methodology with the imaginary time path integral formalism. These works range from constructing a quantum dynamical framework of the imaginary time path integral theory to developing enhanced sampling methods

for classical simulations. The CMD method[37-43] provides a reformulated theory of quantum statistics and dynamics based on the path centroid (center of mass of the ring polymer) representation of the classical ring polymer. Additionally, centroid-based formalisms have been extended to describe the physics of both Fermi-Dirac and Bose-Einstein distributions of condensed matter systems by directly mapping the quantum system into phase space variables,[58-60] in a similar spirit to the CG-PI formalism. More recently, Han et al. have presented the quantum multiscale coarse-graining (qMS-CG) theory[61] which extends the classical MS-CG method[62-66] to quantum Boltzmann statistics. Additionally, Peng et al.[67] and Nava et al.[68] have used the idea of coarse-graining, or bead contraction, and imaginary time path integrals to develop enhanced sampling methods.

In this chapter, we generalize the one-body CG-PI theory to the case of N interacting particles obeying quantum Boltzmann statistics. It is shown that the off-diagonal thermal density matrix elements for a many-body system can be represented with $3N$-body (initial, final, and intermediate CG beads for each particle) potential term using the CG representation. Based on such representation of the thermal density matrix, we show that the N interacting particles at equilibrium can be represented using the many-body effective Hamiltonian that consists of N pseudo-particle pairs with a coupling matrix that is dependent on the masses of the particles, temperature, and the many-body potential. However, due to the complexity of numerically calculating the coupling matrix, we provide a numerical CG-PI (n-CG-PI) formalism based on a simple approximation of the coupling matrix and test it on calculating the statistics of realistic molecular systems. The analytical many-body CG-PI and the n-CG-PI method extend the original one-body CG-PI theory to provide both reductionist insight and practical algorithms to account for NQEs present in interacting many-body systems.

The remainder of this chapter is organized as follows. The one-body CG-PI theory is briefly presented in Sec. II-A, followed by the analytical many-body generalization in Sec. II-B, including the Taylor expanded expression for the many-body CG potential and the many-body effective Hamiltonian. Section III presents the n-CG-PI method, validation of the method, and a discussion on the details of its numerical implementation. Numerical examples of the n-CG-PI method applied to para-hydrogen and liquid water systems are presented in Sec. IV, with discussion and conclusions in Sec. V.

II. General Many-Body CG-PI Theory

A. One-dimensional one-body CG-PI theory

In the imaginary time path integral formalism, for a particle of mass m, inverse temperature $\beta = 1/(k_B T)$, and the Hamiltonian operator $\hat{H} = \hat{p}^2/(2m) + V(\hat{q})$, the off-diagonal thermal density matrix element is written as

$$\langle q|e^{-\beta\hat{H}}|q'\rangle = \lim_{P\to\infty}\left[\left(\frac{mP}{2\pi\hbar^2\beta}\right)^{P/2}\int\left(\prod_{i=2}^{P} dq_i\right) e^{-\beta V_P}\right] \tag{3.1}$$

in which Feynman's isomorphic potential V_P defined as

$$V_P = V_P(q, q_2, \dots, q_P, q') = \sum_{i=1}^{P}\left[\frac{mP}{2\hbar^2\beta^2}(q_i - q_{i+1})^2 + \frac{1}{P}V(q_i)\right] \tag{3.2}$$

and boundary condition of $q_1 = q$ and $q_{P+1} = q'$, $P-1$ intermediate position variables $q_2, \dots, q_P$, and the physical potential $V(q)$. The isomorphic ring polymer described in Eq. (3.2) consists of $P+1$ beads that all experience the physical potential and harmonic nearest neighbor coupling with spring constant of $mP/(\hbar^2\beta^2)$.

The one-body CG-PI theory uses the CG coordinate Q that is defined as the center of mass of all the intermediate beads $(q_2, \dots, q_P)$ of the ring polymer

$$Q = \frac{1}{P-1}(q_2 + \cdots + q_P). \tag{3.3}$$

One significant result of the one-body CG-PI theory is the construction of the effective Hamiltonian that consists of only two interacting pseudo-particles using the CG representation of the off-diagonal thermal density matrix. In this work, the word "pseudo-particle" refers to the classical-like particles in the context of imaginary time path integral whose phase space information is used to approximate the exact quantum statistics. The expectation value of an operator can thus be expressed in a classical-like fashion as shown in Eq. (24) of Ref. 57 using the effective Hamiltonian $H_{eff}\left(P_Q, p, Q, \bar{q}\right)$. In particular, the effective Hamiltonian with constant effective mass approximation can be written as

$$H_{eff}\left(P_Q, p, Q, \bar{q}\right) \cong \frac{{P_Q}^2}{2M_Q} + \frac{p^2}{2m_{\bar{q}}} + V(Q) + \frac{1}{2}k_{Q\bar{q}}(\bar{q})(Q-\bar{q})^2. \tag{3.4}$$

In this form of the effective Hamiltonian, the centroid pseudo-particle (with position Q, momenta P_Q, and associated mass M_Q) experiences the physical potential via the $V(Q)$ term, but the expectation value is taken with respect to the observable pseudo-particle (with position $\bar{q}$, momenta p, and associated mass $m_{\bar{q}}$). The two are harmonically coupled to each other with a spring constant $k_{Q\bar{q}}(\bar{q})$ which itself is a function of the observable pseudo-particle position such that

$$k_{Q\bar{q}}(\bar{q}) = 4m/(\hbar^2\beta^2) \cdot \frac{f^2\, tanh(f)}{f - tanh(f)} \tag{3.5a}$$

$$f = f(\bar{q}) = \hbar\beta\sqrt{V''(\bar{q})/(4m)}. \tag{3.5b}$$

In other words, at equilibrium and at finite temperature, the CG-PI theory shows that a minimum of *two* effective classical particles, coupled to each other, must be used to represent a single physical quantum particle.

B. One-dimensional many-body CG-PI theory

For simplicity, the results and expressions shown in this section are in one physical dimension. For N distinguishable particles obeying quantum Boltzmann statistics, the many-body generalization of Eq. (3.1) is written as

$$\langle \boldsymbol{q}|e^{-\beta\widehat{H}}|\boldsymbol{q}'\rangle = \lim_{P\to\infty}\left[\left[\prod_{i=1}^{N}\left(\frac{m_i P}{2\pi\hbar^2\beta}\right)^{P/2}\right]\times\int\left[\prod_{i=1}^{N}\prod_{j=2}^{P} dq_j^{(i)}\right] e^{-\beta V_{P,N}}\right] \tag{3.6}$$

for Feynman's many-body isomorphic potential $V_{P,N}$ of form

$$V_{P,N} = V_{P,N}\left(\left\{q_j^{(i)}; \begin{matrix} i=1,\dots,N \\ j=1,\dots,P \end{matrix}\right\}\right)$$

$$= \sum_{k=1}^{P}\sum_{i=1}^{N}\left[\frac{m_i P}{2\hbar^2\beta^2}\left(q_k^{(i)} - q_{k+1}^{(i)}\right)^2 + \frac{1}{P}V\left(q_k^{(1)},\dots,q_k^{(N)}\right)\right]. \tag{3.7}$$

The N dimensional column vectors $\boldsymbol{q}$ and $\boldsymbol{q}'$ have q_i and q_i' as their ith elements, respectively, $q_j^{(i)}$ is the position of jth bead of the ith ring polymer, m_i is the mass of the ith particle, and $V(q_1,\dots,q_N)$ is the many-body physical potential of the system. The boundary condition of $q_1^{(i)} = q_i$ and $q_{P+1}^{(i)} = q_i'$ for $i = 1,\dots,N$ is applied. Note that the nearest neighbor beads within the ith ring polymer are harmonically coupled with spring constant of $m_i P/(\hbar^2\beta^2)$ and only the same indexed beads interact with each other through the $V\left(q_k^{(1)},\dots,q_k^{(N)}\right)$ term.

The many-body CG-PI theory uses an analogous mapping of the ith CG coordinate Q_i as Eq. (3.3) such that

$$Q_i = \frac{1}{P-1}\left(q_2^{(i)} + \cdots + q_P^{(i)}\right). \tag{3.8}$$

For the N dimensional column vector $\boldsymbol{Q}$ with Q_i as its ith element, we define a many-body effective CG potential $V_{CG,N}$ that satisfies

$$\langle \boldsymbol{q}|e^{-\beta\widehat{H}}|\boldsymbol{q}'\rangle = \left(\prod_{i=1}^{N}\frac{m_i}{\pi\hbar^2\beta}\right)\int d\boldsymbol{Q}\, e^{-\beta V_{CG,N}(\boldsymbol{q},\boldsymbol{Q},\boldsymbol{q}')}. \tag{3.9}$$

The exact expression of $V_{CG,N}$ can be written as

$$V_{CG,N}(\boldsymbol{q},\boldsymbol{Q},\boldsymbol{q}') = -\frac{1}{\beta}\lim_{P\to\infty}\left[ln\left(Z_{CG,N}^{(P)}(\boldsymbol{q},\boldsymbol{Q},\boldsymbol{q}')\right)\right] \tag{3.10}$$

where

$$Z_{CG,N}^{(P)}(\boldsymbol{q},\boldsymbol{Q},\boldsymbol{q}') = \left[\prod_{i=1}^{N}\left(\frac{\pi\hbar^2\beta}{m_i}\right)\left(\frac{m_iP}{2\pi\hbar^2\beta}\right)^{P/2}\right]$$

$$\times\int\left[\prod_{i=1}^{N}\prod_{j=2}^{P}dq_j^{(i)}\right]e^{-\beta V_{P,N}}\left[\prod_{i=1}^{N}\delta\left(Q_i-\frac{1}{P-1}\left(q_2^{(i)}+\cdots+q_P^{(i)}\right)\right)\right] \tag{3.11}$$

such that Eq. (3.9) is satisfied.

It is important to discuss the importance of the representation shown in Eqs. (3.9) and (3.10). The delocalization of the quantum particle is represented by the nearest neighbor coupling strength between different beads $mP/(\hbar^2\beta^2)$ in the many-body isomorphic potential shown in Eq. (3.7). For systems with significant degrees of NQE ($\hbar\beta \gg 0$), the coupling strength is small such that the ring polymer can span a larger volume within the phase space and hence account for the delocalization. For systems near the classical limit ($\hbar\beta \to 0$) the coupling approaches infinity, which suggests that the ring polymer collapses into a point particle, equivalent to a classical description. On the other hand, the CG-PI theory explicitly treats the intermediate imaginary path (intermediate beads) by using the CG mapping as shown in Eq. (3.8). Therefore, regardless of the degree of NQE in the system, the representation of a the delocalization is always explicit and straightforward. By Taylor expanding, the CG-PI theory re-sums the nearest neighbor coupling between the beads in the limit of $P \to \infty$ into coupling terms between the two pseudo-particles. It is in these pseudo-particle coupling matrix, which are functions of mass, temperature, and potential, that NQEs manifest themselves. This choice of CG mapping and corresponding formulation of the effective CG potential sets the CG-PI theory apart

from different centroid based formalisms[21, 37-40, 58, 60] that focus on the effective centroid potential that only reproduces the physics described by closed ring polymer systems.

One can Taylor expand Eq. (3.10) with respect to $\hbar$ as a small parameter in order to obtain closed form expressions with respect to different orders of $\hbar$. In this work, upto second order Taylor expanded expressions (in which the coupling between the different pseudo-particles are quadratic) are used to study systems near the classical limit. Moreover, for simplicity, we consider a pairwise decomposable form of the many-body potential with pair potential $V_{ij}(r)$ between the *i*th and the *j*th particle. Even though such an assumption cannot describe more complex forms of interactions commonly seen in molecular systems (angular, dihedral, and other many-body intermolecular terms), it has been successfully used in different simulation methods such as MS-CG[62-66] to describe the intermolecular interactions. Derivation of the many-body CG-PI theory with more general form of potential is left for future work.

The Taylor expanded expression of $V_{CG,N}$ up to second order with respect to $\hbar$ is written as

$$V_{CG,N}(\overline{\boldsymbol{q}}, \boldsymbol{Q}, \Delta\boldsymbol{q}) = \left[V(\boldsymbol{Q}) + \frac{1}{2}(\boldsymbol{Q} - \overline{\boldsymbol{q}})^T \boldsymbol{K}_{Q\overline{q}}(\boldsymbol{Q} - \overline{\boldsymbol{q}}) + \frac{1}{2}\Delta\boldsymbol{q}^T \boldsymbol{K}_{\Delta q^2}\Delta\boldsymbol{q} + K_0\right]$$
$$\times\left\{1 + \mathcal{O}\left(\hbar^3\sqrt{\frac{\beta^5}{m^3}}\{V_{ij}'''\}\right)\right\} \tag{3.12}$$

with N dimensional column vectors $\overline{\boldsymbol{q}} = (\boldsymbol{q} + \boldsymbol{q}')/2$ and $\Delta\boldsymbol{q} = \boldsymbol{q}' - \boldsymbol{q}$, N dimensional square coupling matrices $\boldsymbol{K}_{Q\overline{q}}$ and $\boldsymbol{K}_{\Delta q^2}$, and a scalar K_0. The bracket $\{V_{ij}'''\}$ represents the set of third and higher order derivatives of all the pair potentials. The functional details of the coupling matrices $\boldsymbol{K}_{Q\overline{q}}$ and $\boldsymbol{K}_{\Delta q^2}$ and scalar K_0 are shown in Appendix 3A. The derivation of Eq. (3.12) is shown in Appendix 3B. The choice of CG mapping as shown in Eq. (3.8) and the corresponding CG expression of off-diagonal thermal density matrix as shown in Eq. (3.12) provides an

advantage over other possible CG mappings as it still retains the minimal amount of detail ($\boldsymbol{q}$ and $\boldsymbol{q}'$ are not included in the CG coordinate mapping) necessary to describe the off-diagonal elements.[57]

It is important to establish estimators for different thermodynamic properties in the context of CG-PI theory. In this work only the CG-PI estimator for total energy is included. However, a more general relationship between the CG-PI estimators will further be derived in the future. For a system of N distinguishable particles obeying quantum Boltzmann statistics, the CG-PI estimator for the energy $\varepsilon(\boldsymbol{Q}, \overline{\boldsymbol{q}})$ can be written as

$$\varepsilon(\boldsymbol{Q}, \overline{\boldsymbol{q}}) = N/\beta + \beta \frac{\partial}{\partial \beta} V_{CG,N} + V_{CG,N} \tag{3.13}$$

and is related to the total energy of the system E such that

$$E = \langle \varepsilon(\boldsymbol{Q}, \overline{\boldsymbol{q}}) \rangle_{CG,N} \tag{3.14}$$

in which the expectation value notation $\langle f \rangle_{CG,N}$ is defined as

$$\langle f \rangle_{CG,N} = \frac{\int d\boldsymbol{Q} d\overline{\boldsymbol{q}}\, f\, exp(-\beta V_{CG,N})}{\int d\boldsymbol{Q} d\overline{\boldsymbol{q}}\, exp(-\beta V_{CG,N})}. \tag{3.15}$$

The CG-PI energy estimator's derivation, physical interpretation, and equivalence to the energy estimator used in imaginary time path integral formalism are discussed in Appendix 3C.

The many-body effective Hamiltonian $H_{eff,N}$ satisfies a classical-like expression for the expectation value of a many-body operator $\hat{A}(\{\hat{p}_i, \hat{q}_i; i = 1, \dots, N\})$ at equilibrium at finite temperature such that

$$\langle \hat{A} \rangle = \frac{\int d\boldsymbol{P}_Q d\boldsymbol{Q} d\boldsymbol{p} d\overline{\boldsymbol{q}}\, exp\left(-\beta H_{eff,N}\right) W_A(\boldsymbol{p}, \overline{\boldsymbol{q}})}{\int d\boldsymbol{P}_Q d\boldsymbol{Q} d\boldsymbol{p} d\overline{\boldsymbol{q}}\, exp\left(-\beta H_{eff,N}\right)}. \tag{3.16}$$

for N dimensional momentum column vectors for centroid and observable pseudo-particles $\boldsymbol{P}_Q$ and $\boldsymbol{p}$ with $P_{Q,i}$ and p_i as their ith elements, respectively, and many-body Weyl

map $W_A(\boldsymbol{p},\overline{\boldsymbol{q}})$ defined as $\langle \boldsymbol{q}'|\hat{A}|\boldsymbol{q}\rangle = \int (d\boldsymbol{p}/2\pi\hbar)\, exp(i(\boldsymbol{p}\cdot\Delta\boldsymbol{q})/\hbar)\, W_A(\boldsymbol{p},\overline{\boldsymbol{q}})$. The exact expression for $H_{eff,N}$ is written as

$$H_{eff,N}\left(\boldsymbol{P}_Q,\boldsymbol{p},\boldsymbol{Q},\overline{\boldsymbol{q}}\right)$$
$$= \frac{1}{2}\boldsymbol{P}_Q{}^T\boldsymbol{M}_Q^{-1}\boldsymbol{P}_Q - \frac{1}{\beta}\ln\left[const\cdot\int d\boldsymbol{\Delta q}\, exp\left(i(\boldsymbol{p}\cdot\Delta\boldsymbol{q})/\hbar - \beta V_{CG,N}(\overline{\boldsymbol{q}},\boldsymbol{Q},\Delta\boldsymbol{q})\right)\right] \quad (3.17)$$

for N dimensional diagonal mass matrix $\boldsymbol{M}_Q$ with ith diagonal element as a positive mass associated with the ith centroid particle, $M_{Q,i}$.

If we are interested in calculating the expectation values of a system nearer its classical limit, we can combine the first bracket of Eqs. (3.12) and (3.17) to have a second order truncated expression of the many-body effective Hamiltonian

$$H_{eff,N}\left(\boldsymbol{P}_Q,\boldsymbol{p},\boldsymbol{Q},\overline{\boldsymbol{q}}\right)$$
$$\cong \frac{1}{2}\boldsymbol{P}_Q{}^T\boldsymbol{M}_Q^{-1}\boldsymbol{P}_Q + \frac{1}{2\beta\hbar^2}\boldsymbol{p}^T\boldsymbol{K}_{\Delta q^2}^{-1}\boldsymbol{p} + V(\boldsymbol{Q}) + \frac{1}{2}(\boldsymbol{Q}-\overline{\boldsymbol{q}})^T\boldsymbol{K}_{Q\overline{q}}(\boldsymbol{Q}-\overline{\boldsymbol{q}}). \quad (3.18)$$

However, we find that the mass associated with observable pseudo-particles have explicit position dependence via the $\boldsymbol{K}_{\Delta q^2}^{-1}$ term, which would make commonly used molecular dynamics integrators inapplicable as it means that the momentum between different particles are explicitly coupled. Therefore, another level of approximation can be made so that the observable pseudo-particles have constant effective masses.[57] For N dimensional diagonal mass matrix $\boldsymbol{m}_{\overline{q}}$ with ith diagonal element as the constant effective mass for the ith observable pseudo-particle, $m_{\bar{q},i}$, the effective Hamiltonian can be written as

$$H_{eff,N}\left(\boldsymbol{P}_Q,\boldsymbol{p},\boldsymbol{Q},\overline{\boldsymbol{q}}\right)$$
$$\cong \frac{1}{2}\boldsymbol{P}_Q{}^T\boldsymbol{M}_Q^{-1}\boldsymbol{P}_Q + \frac{1}{2}\boldsymbol{p}^T\boldsymbol{m}_{\bar{q}}^{-1}\boldsymbol{p} + V(\boldsymbol{Q}) + \frac{1}{2}(\boldsymbol{Q}-\overline{\boldsymbol{q}})^T\boldsymbol{K}_{Q\overline{q}}(\boldsymbol{Q}-\overline{\boldsymbol{q}}). \quad (3.19)$$

The derivations for Eqs. (3.17) – (3.19) are included in Appendix 3D.

The effective Hamiltonian shown in Eq. (3.19), a many-body generalization of Eq. (3.4), presents the essential CG representation of pairwise interacting many-body systems. Similar to the one-body effective Hamiltonian shown in Eq. (3.4), the many-body effective Hamiltonian consists of two types of pseudo-particles: centroid and observable. The centroid pseudo-particles interact with other centroid pseudo-particles via the many-body pairwise potential term $V(\boldsymbol{Q})$. Since the Weyl map is a function of $\boldsymbol{p}$ and $\overline{\boldsymbol{q}}$, the observable pseudo-particles' coordinates and momenta are used to calculate expectation values of operators. The difference between Eqs. (3.4) and (3.19) is that the matrix $\boldsymbol{K}_{Q\overline{q}}$ is arguably more complicated as it explicitly considers all the possible pairwise interactions and distances between the pseudo-particle pairs. Moreover, it will in general have non-zero off-diagonal elements, suggesting that the different pseudo-particle pairs are explicitly coupled to each other.

However, one faces several technical challenges when trying to use the effective Hamiltonian in Eq. (3.19) in numerical calculations. First, the calculation of coupling matrix $\boldsymbol{K}_{Q\overline{q}}$ involves several inversions of N dimensional square matrices. Considering that matrix inversion in general scales poorly with growing system size (~ O(N^3)), performing such computationally intensive operations every timestep for new observable pseudo-particle positions defeats the purpose of using the many-body CG-PI theory, especially when one could calculate equivalent information with other linear scaling numerical path-integral techniques. Moreover, calculating $\boldsymbol{K}_{Q\overline{q}}$ involves performing infinite sums which adds another complexity to numerical implementation. To deal with such issues, a development named "numerical CG-PI" (n-CG-PI) based on the approximation of $\boldsymbol{K}_{Q\overline{q}}$ is presented in the next section.

III. The n-CG-PI Method

The approximations used in developing the n-CG-PI method, its validation, and relevant discussions will be presented in this section. The equations shown in this section are in one-dimension for simplicity.

A. One-dimensional n-CG-PI method

In constructing the n-CG-PI method, we make two assumptions about the coupling matrix $\boldsymbol{K}_{Q\overline{q}}$ that will be validated shortly after. First, we assume that the matrix can be well represented by a constant matrix. Moreover, we also assume that the off-diagonal elements of $\boldsymbol{K}_{Q\overline{q}}$ are negligible. The form of the n-CG-PI Hamiltonian with such assumptions then can be written as

$$H_{n-CG-PI}\left(\boldsymbol{P}_Q, \boldsymbol{p}, \boldsymbol{Q}, \overline{\boldsymbol{q}}\right)$$
$$= \frac{1}{2}\boldsymbol{P}_Q{}^T\boldsymbol{M}_Q^{-1}\boldsymbol{P}_Q + \frac{1}{2}\boldsymbol{p}^T\boldsymbol{m}_{\overline{q}}^{-1}\boldsymbol{p} + V(\boldsymbol{Q}) + \frac{1}{2}(\boldsymbol{Q}-\overline{\boldsymbol{q}})^T\boldsymbol{K}_\lambda(\boldsymbol{Q}-\overline{\boldsymbol{q}}) \tag{3.20}$$

for N dimensional diagonal coupling matrix $\boldsymbol{K}_\lambda$ with $\lambda_i \cdot 12m_i/(\hbar^2\beta^2)$ as its ith diagonal element, and the coupling parameter for the ith pseudo-particle pair λ_i that can range from 1 to ∞. Similar to the effective Hamiltonian shown in Eq. (3.19), the centroid pseudo-particles in n-CG-PI Hamiltonian interact with each other via the potential term $V(\boldsymbol{Q})$ and the expectation values are taken with respect to the phase space positions of the observable pseudo-particles. The n-CG-PI coupling matrix $\boldsymbol{K}_\lambda$ is an approximation of the ensemble average of the full coupling matrix $\boldsymbol{K}_{Q\overline{q}}$ such that

$$\boldsymbol{K}_\lambda \cong \frac{\int d\boldsymbol{P}_Q d\boldsymbol{Q} d\boldsymbol{p} d\overline{\boldsymbol{q}}\, exp\left(-\beta H_{eff,N}\right)\boldsymbol{K}_{Q\overline{q}}(\overline{\boldsymbol{q}})}{\int d\boldsymbol{P}_Q d\boldsymbol{Q} d\boldsymbol{p} d\overline{\boldsymbol{q}}\, exp\left(-\beta H_{eff,N}\right)}. \tag{3.21}$$

For the given form of $\boldsymbol{K}_\lambda$ there only exists diagonal coupling, meaning that the different pseudo-particle pairs are independent from each other aside from indirect coupling that happens via the

centroid pseudo-particle effective potential term $V(\boldsymbol{Q})$. Moreover, the strengths of the coupling between the pseudo-particles are constant, parametrized by a set of scalar values $\{\lambda_i\}$. The physical interpretation of the coupling parameters is that they determine how delocalized the observable pseudo-particles will be from the centroid pseudo-particles and hence the degree of quantum effect for the corresponding pseudo-particle pair.

The two extreme values that λ_i can take, namely 1 and ∞, correspond to two important limits of the n-CG-PI method: free particle and classical limit. For $\lambda_i = 1, i = 1, \dots, N$, $\boldsymbol{K}_\lambda$ corresponds to $\boldsymbol{K}_{Q\overline{q}}$ for $V(q_1, \dots, q_N) = 0$. Therefore, for a nonzero form of the many-body potential, the choice of $\lambda_i = 1$ in the n-CG-PI Hamiltonian would overestimate the delocalization for all pseudo-particle pairs. As the values of λ_i increase, the delocalization of the particle is suppressed as the coupling between the pseudo-particle becomes stronger. Finally for the case of $\lambda_i \to \infty$, the spring constant between the pseudo-particle pair approaches infinity and the pair collapse into a point mass, which corresponds to the classical limit in which no delocalization takes place. Due to the trivial coupling scheme between the pseudo-particles, the simulation of a many-body system using the n-CG-PI Hamiltonian for any values of λ_i would essentially have the same cost as that of a classical simulation that scales linearly with respect to system size, giving the n-CG-PI method a substantial numerical advantage given that one can choose the appropriate set of coupling parameters.

B. Validation

The assumption that $\boldsymbol{K}_{Q\overline{q}}$ can be well represented with a constant matrix is based on revisiting the one-dimensional model studied in the original CG-PI work.[57] The Figure S2 of the Ref. 57 shows distributions of normalized coupling strength $k_{Q\overline{q}} \cdot \hbar^2\beta^2/(12m)$ for one-dimensional anharmonic potential. We assume that such distributions can be minimally

represented with a mean, meaning that one would be able to effectively capture the result sampled using the full one-body CG-PI effective Hamiltonian from Eq. (3.4) by using the one-body n-CG-PI Hamiltonian with λ as the mean of the coupling strength distribution.

To test this idea, we revisited the one-dimensional anharmonic potential of form $V(x) = 0.252x^2 - 0.570x^3 + 0.814x^4$ with x in Å and V in kcal/mol for a particle of mass 20.0 g/mol and temperature at 5 K. Both the one-body CG-PI effective Hamiltonian from Eq. (3.4) and the one-body n-CG-PI Hamiltonian of form

$$H_{n-CG-PI}\left(P_Q, p, Q, \overline{q}\right) = \frac{{P_Q}^2}{2M_Q} + \frac{p^2}{2m_{\overline{q}}} + V(Q) + \lambda \frac{12m}{\hbar^2\beta^2}\frac{\left(Q-\overline{q}\right)^2}{2} \tag{3.22}$$

were used to calculate the normalized diagonal density matrix elements in position basis. For both cases, all masses associated with the pseudo-particles were set to be 20 g/mol. The average of the distribution of normalized coupling strength shown in Figure S2 of Ref. 57 was 1.454, which was set to be the value of λ for the one-body n-CG-PI Hamiltonian. Other calculation details from Ref. 57 on were followed.

Figure 3-1 compares the normalized diagonal thermal density matrix calculated from the CG-PI effective Hamiltonian from Eq. (3.4) and from the one-body n-CG-PI Hamiltonian from Eq. (3.22) with $\lambda = 1.454$. We see from Figure 3-1 that the diagonal thermal density matrix calculated from the n-CG-PI Hamiltonian very closely captures that calculated with the CG-PI effective Hamiltonian. The only source of discrepancy between the two comes from the form of scalar coupling between the two pseudo-particles as there are no off-diagonal terms to consider in the one-body system.

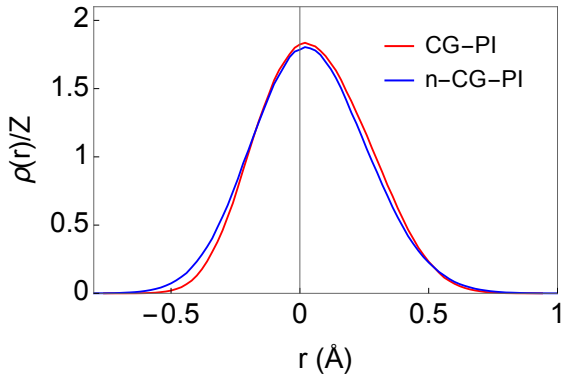

Figure 3-1. Calculated diagonal thermal density matrix elements for the one-body anharmonic potential with CG-PI (red) and n-CG-PI method (blue).

The assumption that the off-diagonal terms of $\boldsymbol{K}_{Q\overline{q}}$ do not contribute significantly to the statistics comes from investigating a system of N particles all of mass m obeying quantum Boltzmann distribution interacting with each other via harmonic pair potential of form $V_{ij}(r) = K\left(r - r_{ij}\right)^2/2$ for the ith and jth particle with spring constant K and equilibrium distance r_{ij}. Even though such potential does not correctly describe behavior of realistic molecular systems, we study it in this section because it is a nontrivial analytically solvable form of the potential that gives a physical insight into the interactions between different pseudo-particles for nonzero forms of potential. The expression of $\boldsymbol{K}_{Q\overline{q}}$ for such harmonic system has diagonal and off-diagonal elements given by the expression $12mC_{diag}/(\hbar^2\beta^2)$ and $12mC_{off-diag}/(\hbar^2\beta^2)$, respectively, in which C_{diag} and $C_{off-diag}$ are functions of K and N given by

$$C_{diag}(f,N) = \frac{3f\sqrt{N}+\left(N(N-1)f^2-3\right)tanh\left(f\sqrt{N}\right)}{3\left(fN^{3/2}-N\,tanh\left(f\sqrt{N}\right)\right)} \tag{3.23a}$$

$$C_{off-diag}(f,N) = \frac{3f\sqrt{N}+\left(Nf^2+3\right)tanh\left(f\sqrt{N}\right)}{3\left(fN^{3/2}-N\,tanh\left(f\sqrt{N}\right)\right)} \tag{3.23b}$$

for $f = \hbar\beta\sqrt{K/4m}$. Both Eqs. (3.23a) and (3.23b) are exact and constant at any temperature as the harmonic pair potential have zero third and higher order derivatives. The expressions for the coupling elements in Eqs. (3.23a) and (3.23b) provide a perspective on how the pseudo-particles are coupled in many-body systems. The $(N-1)$ term in the denominator of Eq. (3.23a), $(N(N-1)f^2-3)$, comes from considering the *i*th particle interacting with $(N-1)$ other particles in the system, whereas the expression for $C_{off-diag}$ only considers pairwise interaction between the *i*th and the *j*th pair and hence has the (Nf^2+3) factor. Moreover, one can plot C_{diag} and $C_{off-diag}$ as a function of f for different values of N as shown in Figure 3-2.

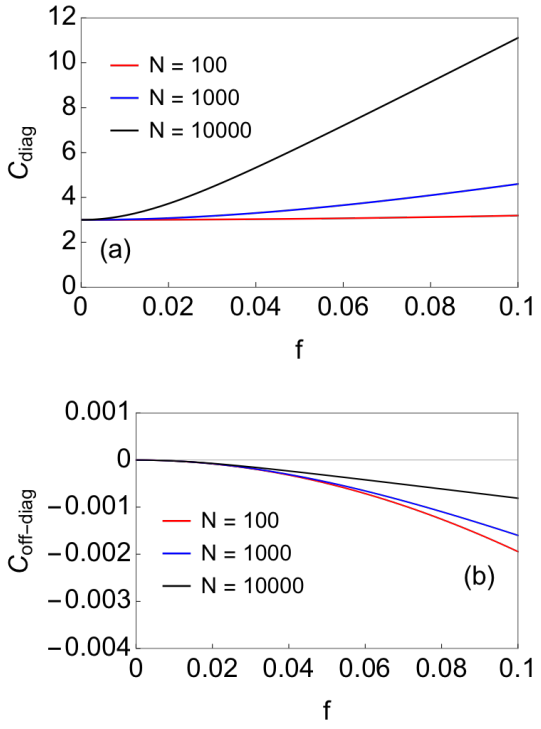

Figure 3-2. Plotted (a) diagonal and (b) off-diagonal coupling element for many-body harmonic system as a function of *f* for $N = 100$ (red), $N = 1000$ (blue), and $N = 10000$ (black).

Note that for similar ranges of f, C_{diag} is approximately three orders of magnitude larger than $C_{non-diag}$. Most importantly, they have the limiting property of $\lim_{f\to 0} C_{diag} = 3$ and $\lim_{f\to 0} C_{non-diag} = 0$, respectively. Considering that the Taylor expansion of the exact expression for the many-body CG potential assumed $\hbar$ to be a small parameter and hence a small value of f, we can assume that only using the diagonal elements of $\boldsymbol{K}_{Q\overline{q}}$ would effectively capture the full coupling scheme between the two pseudo-particles.

C. Choosing the coupling strength

From the two examples shown in the previous section, we have seen that the substitution made in Eq. (3.20) is valid for one-body systems and a special harmonic many-body system. Even though one cannot obtain an ensemble average or analytical expressions of $\boldsymbol{K}_{Q\overline{q}}$ to construct $\boldsymbol{K}_{\lambda}$ for a general case, we were able to extrapolate the optimum values of λ that best reproduces the result from exact numerical PIMD as a function of mass and temperature and obtain a closed form expression for the modified liquid para-hydrogen system. Even though more work needs to be done in predicting the value of λ without *a priori* knowledge of the full $\boldsymbol{K}_{Q\overline{q}}$ or exact quantum statistics, the result shown in later sections will provide a guideline into using the n-CG-PI method for more complicated molecular systems.

D. Thermostatting

For the canonical simulations presented in the numerical examples section, the Langevin thermostat[69] was used to maintain the constant temperature for PIMD, classical MD, and n-CG-PI simulations. In the n-CG-PI simulations the observable pseudo-particles were more strongly thermostatted than the centroid pseudo-particles. The discrepancy between the thermostatting scheme of the two pseudo-particles comes from the nature of the interaction and the potential energy surface (PES). In the n-CG-PI Hamiltonian, the centroid pseudo-particle interacts with

other centroid pseudo-particles via the physical potential term and its corresponding observable pseudo-particle through the constant coupling strength harmonic term. However, the observable pseudo-particle only has the harmonic interaction with the centroid pseudo-particle and lacks a rugged PES, and hence requires stronger thermostatting. Canonical samplings of one-body systems using the full one-body CG-PI effective Hamiltonian in Eq. (3.4) or one-body n-CG-PI Hamiltonian in Eq. (3.22) does not suffer from this discrepancy. For the n-CG-PI simulations presented in the numerical examples section, the subset of thermostatting parameters that gave a numerically stable simulation and converged statistics were used.

IV. Numerical Examples

To demonstrate the efficacy of the n-CG-PI method, modified liquid para-hydrogen and liquid water systems were studied. For both cases, the results from the n-CG-PI method were compared to numerically exact quantum results obtained from PIMD simulations and classical results obtained from classical MD simulations. For both PIMD simulations, $P = 32$ was large enough to account for the NQEs. For the n-CG-PI simulations of both systems, the associated masses of the two pseudo-particles were set to be the physical mass of the particle.

A. Modified liquid para-hydrogen

A canonical ensemble of 180 liquid para-hydrogen molecules in a cubic box of length 19.71 Å was simulated at different temperatures for 2.0 ns using different methods. A semi-empirical model potential[70] that treats the interacting diatomics as a set of spheres was used so that only one pair potential is used to describe the interaction between the molecules. The mass and the temperature of the system were varied to modulate the NQE of the system. A timestep of 0.5 fs with a Langevin thermostat were used for all three simulations.[69] The thermostatting

parameter of 0.01 fs^{-1} was used for PIMD and classical MD and 0.01 fs^{-1} and 5.0 fs^{-1} for n-CG-PI centroid and observable pseudo-particles, respectively. The PIMD simulations were propagated by an in-house modified LAMMPS molecular dynamics engine[71] using the Path Integral Langevin Equation (PILE)[72] integrator. The n-CG-PI trajectories were also generated using an in-house modified LAMMPS MD engine.[71]

In order to find the scaling relationship between λ, mass, and temperature, several systems with different masses and temperatures were investigated. For each cases, the value of λ was considered to be optimal if the first peak of the radial distribution function (RDF) calculated from n-CG-PI corresponded with that from PIMD. The details of these test cases can be found in Appendix 3E. The resulting expression of λ with the prefactor $12m/(\hbar^2\beta^2)$ multiplied as a function of mass and temperature is given as

$$\lambda(m,T)\cdot\frac{12m}{\hbar^2\beta^2} = a + b\cdot m + c\cdot T^2 \tag{3.24}$$

where *m* and *T* are mass in g/mol and temperature in K, respectively, and *a, b,* and *c* are fitted parameters with values of $-2.64\times10^{-3}\,fs^{-2}\cdot g/mol$, $5.36\times10^{-4}\,fs^{-2}$, and $1.27\times10^{-6}\,K^{-2}fs^{-2}\cdot g/mol$, respectively.

It is important to discuss the limiting behaviors of the scaling relationship shown in Eq. (3.24). In the limit of large mass and high temperature, we see that Eq. (3.24) predicts the coupling strength to approach infinity, which is to be expected as the system becomes more classical and the delocalization between the pseudo-particles becomes less favorable. On the other hand, in the limit of small mass and low temperature, we see that Eq. (3.24) predicts a negative value of the coupling strength ($a<0$), which would cause serious problems when sampling using the n-CG-PI Hamiltonian as the distance between the two pseudo-particle pair would increase throughout the simulation. However, such an artifact comes from the fact that the

n-CG-PI is based on the full many-body CG-PI effective Hamiltonian shown in Eq. (3.19) which itself is based on assuming that the system is not far from its classical limit and not too anharmonic and thus should not be considered as an issue innate to the n-CG-PI method.

We simulated the modified para-hydrogen system with $m = 5\ g/mol$ and $T = 30\ K$ in which we have $\lambda = 1.276$ according to Eq. (3.24). Figure 3-3 shows the calculated RDFs and three-body non-bonded angle distribution from three different methods to show the structural factors of the liquid para-hydrogen. The non-bonded angle distribution function $P(\theta)$ is given by

$$P(\theta) = \sum_i^N \sum_{j>i}^N \sum_{k\neq i,j}^N \delta\left(\theta_{ijk} - \theta\right) \tag{3.25}$$

for angle between particle i, j, and k centered at the ith particle θ_{ijk}.

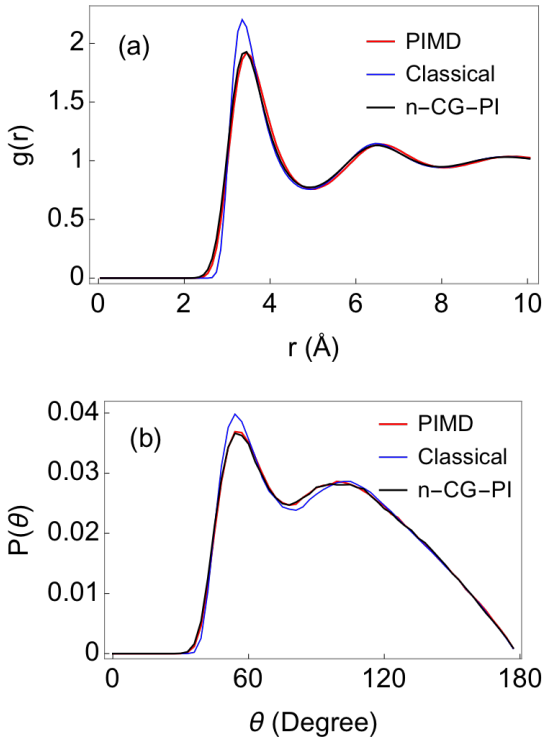

Figure 3-3. Calculated (a) radial distribution functions and (b) three-body non-bonded angles distributions obtained from PIMD simulation (red), classical MD simulation (blue), and n-CG-PI simulation (black) for the modified para-hydrogen model with mass of 5 g/mol at 30K. The cutoff distance used to calculate (b) was 4.5 Å.

As shown in Figure 3-3, there are distinctive differences between the classical and PIMD results. Therefore, the modified para-hydrogen system with mass of 5 g/mol and temperature of 30K cannot accurately be described with classical mechanics. This system is an ideal case to test the performance of the n-CG-PI method. Both Figures 3-3(a) and 3-3(b) show that the n-CG-PI model of the modified para-hydrogen quantitatively captures both the pairwise correlations and the non-bonded three-body angular distribution obtained from numerically exact PIMD simulations. Such agreement between the two results supports the validity of not only the approximation made in constructing the n-CG-PI Hamiltonian but also the scaling relationships in Eq. (3.24). The small discrepancies between the n-CG-PI and the PIMD result, particularly for the long range pair correlation, comes from both the third or higher order Taylor expansion terms that were not included in the derivation of Eq. (3.12) as well as the approximations made in constructing the coupling scheme in the n-CG-PI Hamiltonian.

B. Liquid water

Liquid water was simulated to test the performance of n-CG-PI method on more complicated molecular systems. Even though the n-CG-PI method is based on the CG-PI effective Hamiltonian shown in Eq. (3.12) and derived with the assumption of pairwise decomposable form of the many-body potential, we explored whether it can be applied to more complex molecular systems with explicit angular interactions. It is noted that unlike the previous section, we do not provide a scaling relationship of λ for liquid water due to the complex nature of the potential and the presence of more than one type of atom. Elucidating the n-CG-PI scaling behavior for more complex systems is left for future work, but the water results from the n-CG-PI method is presented as a proof of concept for the versatility of the method.

A canonical ensemble of 233 water molecules in a cubic box of length 19.03 Å was simulated at 300 K for 5.0 ns. The qSPC/Fw water force field[13] was used with particle-particle particle-mesh method with a cutoff distance of 10.0 Å to evaluate the short-range electrostatic interactions. A timestep of 0.1 fs and Langevin thermostat[69] was used for all the simulations. The thermostatting parameter of 0.1 fs^{-1} was used for classical MD and PIMD and 0.4 fs^{-1} and 5.0 fs^{-1} for centroid and observable pseudo-particles for n-CG-PI respectively. The PIMD simulations were propagated by an in-house modified LAMMPS molecular dynamics engine[71] using PILE[72] integrator. The n-CG-PI trajectories were also generated from in-house modified LAMMPS molecular dynamics engine.[71] The choices of λ for the n-CG-PI method were 6.0 and 2.0 for hydrogen and oxygen, respectively, based on the optimal fitting of the RDF as mentioned before.

Figure 3-4 shows three calculated RDFs using PIMD, classical MD, and n-CG-PI.

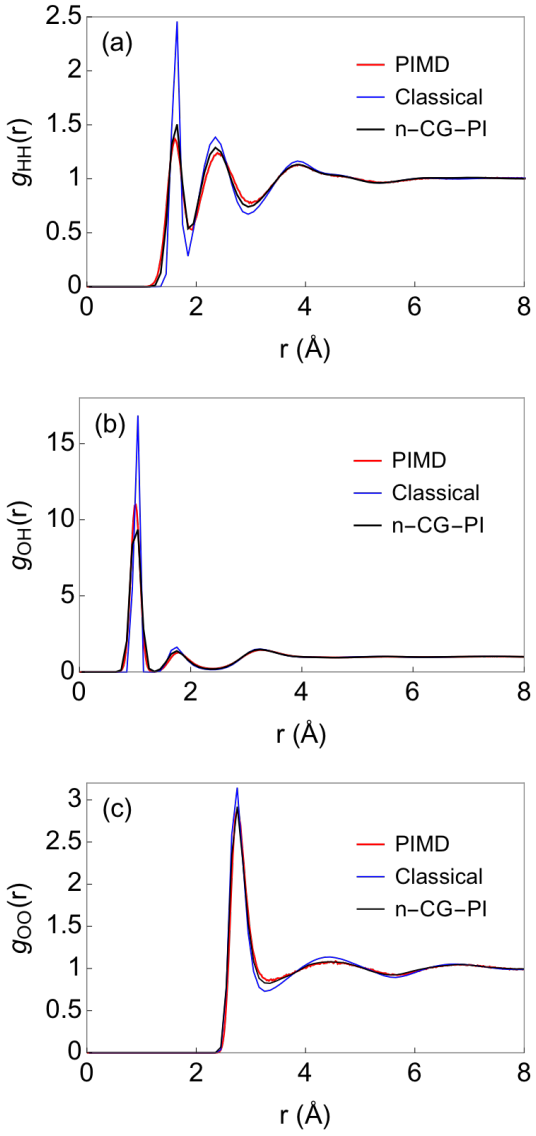

Figure 3-4. Calculated (a) H-H RDF, (b) O-H RDF, and (c) O-O RDF obtained from PIMD simulation (red), classical MD simulation (blue) and n-CG-PI simulation (black) for the qSPC/Fw water model at 300K.

Despite the ambient temperature (300K) of the system, the RDFs from PIMD and classical MD simulations differ considerably, suggesting the presence of significant NQE in the system.[11-19] Figure 3-4 shows that the results from n-CG-PI captures NQEs of the PIMD RDFs. In particular, the n-CG-PI result well reproduces the O-O pair structure from PIMD simulations,

aside from the minor discrepancies at the minima around 3.2 Å as seen in Figure 3-3(c). Due to the stronger NQE for lighter hydrogen atoms the agreement between the n-CG-PI and PIMD results is better for the O-O RDF compared to the H-H and O-H RDFs, although the n-CG-PI RDFs qualitatively captures the general features of the PIMD RDFs as shown in Figure 3-3(a) and 3-3(b). It should be noted that in addition to the reasons for discrepancies between the n-CG-PI and the PIMD result discussed in the previous section of the modified para-hydrogen model, the fact that the n-CG-PI Hamiltonian from Eq. (3.20) was based on the assumption of pairwise decomposable form of potential should be accounted for. However, despite such discrepancies, it is promising to see that n-CG-PI's simple coupling scheme can account for the essential aspects of the pairwise correlation in a more complicated molecular system with explicit angular interactions.

V. Discussion and Conclusion

The general expression for the many-body CG potential shown in Eq. (3.10) significantly reduces the number of DOF while maintaining the minimal level of detail in the imaginary time path integral expression of the off-diagonal thermal density matrix. It is important to note that the expressions for both the many-body effective CG potential in Eq. (3.10) and effective Hamiltonian in Eq. (3.17) are exact and able to describe many-body systems with arbitrarily strong NQEs. In particular, the many-body CG potential in Eq. (3.10) can be expanded with respect to $\hbar$, its effective harmonic terms re-summed, and expressed in a closed form up to second order to study systems nearer the classical limit. We note that the assumption of the pairwise decomposable form of potential is used only to simplify the derivation and the expression for the many-body CG potential. Using the effective many-body Hamiltonian with

constant mass approximation as shown in Eq. (3.19), the N interacting particles obeying quantum Boltzmann statistics are represented as N pairs of pseudo-particles in which the coupling between the pseudo-particles is determined by a complex coupling matrix $\boldsymbol{K}_{Q\overline{q}}$ which is numerically demanding and impractical to directly calculate for realistic molecular many-body systems.

To address this complication, the n-CG-PI Hamiltonian is constructed by coupling the pseudo-particle pairs with the matrix $\boldsymbol{K}_{\lambda}$, which is an approximation to the canonical ensemble average of $\boldsymbol{K}_{Q\overline{q}}$. The n-CG-PI Hamiltonian has a constant coupling between the pseudo-particle pairs parametrized by coupling parameters $\{\lambda_i\}$ and ignores any cross-coupling between the different pairs. Unfortunately, in general one cannot know the set of values of $\{\lambda_i\}$ to choose as the calculation of $\boldsymbol{K}_{Q\overline{q}}$ for general many-body systems is numerically infeasible. However, despite such shortcomings we provided a scaling relationship of λ as a function of mass and temperature of the system that provides the optimal RDFs for the case of a model para-hydrogen system. It was thus demonstrated that the n-CG-PI method with appropriately chosen coupling parameters is capable of capturing the structural correlations of a model para-hydrogen system. Moreover, despite the assumption of pairwise decomposable form of potential made in deriving Eq. (3.12), the n-CG-PI method was used to simulate a liquid water system having non-pairwise interactions as a proof of concept. It is shown that for well-chosen coupling parameters the n-CG-PI can capture quite well the essential aspects of the pairwise correlations in liquid water.

One clear direction in extending the CG-PI theory will be to derive the effective many-body CG potential with more general assumptions for the potential and corresponding approximation of the coupling matrix. Addition of angular, dihedral, and many-body intermolecular interaction terms into the potential would arguably make the derivation more

challenging as they need to be expanded with respect to more than one variable unlike the pair potential. Moreover, similar to how the n-CG-PI Hamiltonian was constructed from the many-body CG potential, some form of approximation to the coupling scheme between the pseudo-particles that explicitly includes non-pairwise interactions would have to be worked out to fully apply the idea of CG-PI for realistic molecular systems. A derivation for a more general relationship between CG-PI estimators and various thermodynamic quantities also remains to be worked out.

In terms of the n-CG-PI method, one might develop a method that calculates the value of optimum λ without *a priori* PIMD results. Based on the scaling behavior for the modified liquid para-hydrogen model, for a pairwise interacting system λ is a function of mass and temperature. Moreover, the three fitting parameters a, b, and c in Eq. (3.24) are functions of the pair potential. Lastly, further analysis on the behavior of λ would be required to extend the n-CG-PI method to more complicated molecular systems with non-pairwise form of potential and more than one type of atom.

In conclusion, we presented in this work a many-body generalization of the CG-PI theory and a corresponding numerical method that can capture the structural properties of condensed matter systems obeying quantum Boltzmann statistics. The theoretical and numerical insights introduced and discussed in this work may provide both reductionist and computationally affordable tools to study quantum systems, particularly ones not too far away from the classical limit.

Appendix 3A: Coupling Terms in the Many-Body CG Potential

The coupling matrices $\boldsymbol{K}_{Q\overline{q}}$, $\boldsymbol{K}_{\Delta q^2}$, and scalar K_0 in Eq. (3.12) are defined as

$$\boldsymbol{K}_{Q\overline{q}} = \frac{1}{\beta}\boldsymbol{L}^{-1} - \frac{1}{2}\boldsymbol{M} \tag{3.A1a}$$

$$\boldsymbol{K}_{\Delta q^2} = \frac{1}{24}\boldsymbol{M} + \frac{1}{\beta}\boldsymbol{\Sigma} - \frac{\beta}{4}\boldsymbol{C} \tag{3.A1b}$$

$$K_0 = \frac{1}{2\beta}\ln[(\prod_{i=1}^{N} 4m_i/(\hbar^2\beta))det(\boldsymbol{\Omega} + \boldsymbol{I})det(\boldsymbol{L})] \tag{3.A1c}$$

for N dimensional square matrices $\boldsymbol{L}$, $\boldsymbol{M}$, $\boldsymbol{\Sigma}$, and $\boldsymbol{C}$, the N dimensional square block matrix $\boldsymbol{\Omega}$, and the identity matrix $\boldsymbol{I}$ of same dimensionality as $\boldsymbol{\Omega}$.

We define two-particle quantities σ_{ij}, γ_{ij}, and $\overline{V_{ij}''}$ such that

$$\sigma_{ij} = \sqrt{\hbar^2\beta/\sqrt{m_i m_j}} \tag{3.A2a}$$

$$\gamma_{ij} = \left(m_j/m_i\right)^{1/4} \tag{3.A2b}$$

$$\overline{V_{ij}''} = V_{ij}''\left(\bar{q}_i - \bar{q}_j\right) + V_{ij}''\left(\bar{q}_j - \bar{q}_i\right). \tag{3.A2c}$$

The ith row jth column block element of the block matrix $\boldsymbol{\Omega}$ is $\boldsymbol{\Omega}_{ij}$, which itself is an infinite dimensional diagonal matrix. The kth row lth column matrix element of $\boldsymbol{\Omega}_{ij}$ is $\Omega_{ij,kl}$ and can be written as

$$\Omega_{ij,kl} = \begin{cases} \delta_{kl}\sum_{\substack{m=1 \\ m\neq i}}^{N}\left(\frac{\beta}{2(k\pi)^2}{\sigma_{im}}^2\overline{V_{im}''} + ({\gamma_{im}}^2 - 1)\right) & (i = j) \\ -\delta_{kl}\frac{\beta}{2(k\pi)^2}{\sigma_{ij}}^2\overline{V_{ij}''} & (i \neq j) \end{cases}. \tag{3.A3}$$

For a given $\boldsymbol{\Omega}$, another block matrix $\boldsymbol{\Lambda}$ of same dimensionality is defined as

$$\boldsymbol{\Lambda} = (\boldsymbol{\Omega} + \boldsymbol{I})^{-1} \tag{3.A4}$$

in which the notation for its block element and element of block element is analogous to $\boldsymbol{\Omega}$.

For two infinite dimensional column vectors $\boldsymbol{b}_\xi$ and $\boldsymbol{b}_\lambda$ with the kth elements $b_{\xi,k}$ and $b_{\lambda,k}$, respectively, defined as

$$b_{\xi,k} = -2\sqrt{2}\cdot odd(k)/(\pi^2 k^2) \tag{3.A5a}$$

$$b_{\lambda,k} = \sqrt{2} \cdot \big(1 - odd(k)\big)/(\pi^2 k^2) \tag{3.A5b}$$

with $odd(k) = \begin{cases} 1\ (k = odd) \\ 0\ (k = even) \end{cases}$, we define N dimensional square matrices $\boldsymbol{L}$ and $\boldsymbol{R}$ whose ith row and jth column elements L_{ij} and R_{ij}, respectively, are defined as

$$L_{ij} = {\sigma_{ij}}^2 \big({\boldsymbol{b}_\xi}^T \boldsymbol{\Lambda}_{ij} \boldsymbol{b}_\xi \big) \tag{3.A6a}$$

$$R_{ij} = {\sigma_{ij}}^2 \big({\boldsymbol{b}_\lambda}^T \boldsymbol{\Lambda}_{ij} \boldsymbol{b}_\lambda \big). \tag{3.A6b}$$

For the remaining three N dimensional square matrices $\boldsymbol{C}, \boldsymbol{\Sigma}$, and $\boldsymbol{M}$, their ith row and jth column elements are C_{ij}, Σ_{ij}, and M_{ij}, respectively, and defined as

$$C_{ij} = \sum_{m=1}^{N} \sum_{n=1}^{N} \big(R_{ij} - R_{in} - R_{mj} + R_{mn} \big) \overline{V_{im}''}\ \overline{V_{jn}''} \tag{3.A7a}$$

$$\Sigma_{ij} = \delta_{ij} / {\sigma_{ij}}^2 \tag{3.A7b}$$

$$M_{ij} = \begin{cases} \sum_{\substack{m=1 \\ m \neq i}}^{N} \overline{V_{im}''} & (i = j) \\ -\overline{V_{ij}''} & (i \neq j) \end{cases}. \tag{3.A7c}$$

Appendix 3B: Derivation of One-Dimensional Taylor Expanded Many-Body CG Potential

The d-dimensional generalization is not explicitly included in this work as the full CG-PI effective Hamiltonian shown in Eq. (3.19) is not numerically implemented. Moreover, one can easily obtain higher dimensional expressions by substituting the single and double derivative of the pair potential in the current derivation with d dimensional gradient vectors and d dimensional square Hessian matrices respectively. For a system of N particles obeying quantum Boltzmann statistics, the series representation of an imaginary path for the ith particle with mass m_i is written as

$$q_i(u) = q_i(0) + \big(q_i(1) - q_i(0) \big) u + \frac{\sqrt{2}}{\pi} \sigma_i \sum_{k=1}^{\infty} a_{i,k} sin(k\pi u)/k \tag{3.B1}$$

with $\sigma_i = \sqrt{\hbar^2\beta/m_i}$, path parameters $\{a_{i,k}\}$ in which i indexes the particle number and k indexes the Fourier decomposition of the path and ranges from 1 to ∞, and variable u that ranges from 0 to 1. It is noted here that i and j are used to index particle numbers unless there is an imaginary number i in the expression, in such cases m and n are used to avoid confusion. Moreover, k and l are used to index the Fourier decomposition of modes as shown in Eq. (3.B1). The centroid of ith particle in path representation is written as

$$\lim_{P\to\infty} \frac{1}{P-1}\left(q_2^{(i)} + \cdots + q_P^{(i)}\right) = \int_0^1 du\, q_i(u) = \overline{q}_i + \frac{2\sqrt{2}}{\pi^2}\sigma_i \sum_{k=1,3,5,\ldots}^{\infty} a_{i,k}/k^2 \qquad (3.B2)$$

in which $q_j^{(i)}$ is the position of jth bead of ith classical polymer and $\overline{q}_i = \big(q_i(1) + q_i(0)\big)/2$.

By changing the discrete form of integration to a continuous one by taking the $P \to \infty$ limit, for a general form of the many-body potential $V(q_1, \ldots, q_N)$ the exact expression for the many-body CG distribution $exp\left(-\beta V_{CG,N}(\boldsymbol{q}, \boldsymbol{Q}, \boldsymbol{q}')\right)$ can be written as

$$e^{-\beta V_{CG,N}(\boldsymbol{q},\boldsymbol{Q},\boldsymbol{q}')} = \left(\prod_{i=1}^{N} \sigma_i /\sqrt{8\pi}\right) exp\left(-\frac{1}{2}\Delta\boldsymbol{q}^T\boldsymbol{\Sigma}\Delta\boldsymbol{q}\right)$$

$$\times \int \frac{d\boldsymbol{\xi} d\boldsymbol{a}}{\sqrt{2\pi}} \begin{bmatrix} exp\left(-\frac{1}{2}\boldsymbol{a}^2\right) exp\left(-\beta \int_0^1 du\, V\big(q_1(u), \ldots, q_N(u)\big)\right) \\ exp\big(i\boldsymbol{\xi}\cdot(\boldsymbol{Q}-\overline{\boldsymbol{q}})\big)\, exp\left(-i\frac{2\sqrt{2}}{\pi^2}\sum_{m=1}^{N}\sum_{k=1,3,5,\ldots}^{\infty}\sigma_m\xi_m a_{m,k}/k^2\right) \end{bmatrix}. \qquad (3.B3)$$

The N dimensional column vectors $\boldsymbol{q}$, $\boldsymbol{Q}$, $\boldsymbol{q}'$, $\Delta\boldsymbol{q}$, and $\overline{\boldsymbol{q}}$ have $q_i(0)$, $q_i(1)$, Q_i, $\Delta q_i = q_i(1) - q_i(0)$, and $\overline{q}_i$ as their ith elements respectively. The N dimensional column vector $\boldsymbol{\xi}$ with ξ_m as its mth element is introduced by using the integral representation of the delta function for each mth centroid variable Q_m introduced by coarse-graining the intermediate beads. The N dimensional block column vector $\boldsymbol{a}$ has $\boldsymbol{a}_i$ as its ith block element, and the vector $\boldsymbol{a}_i$ is an infinite dimensional block vector with $a_{i,k}$ as its kth element. Finally, the N dimensional square matrix $\boldsymbol{\Sigma}$ is defined in the Eq. (3.A7b) of the main document.

We define pairwise quantities A_{ij} and B_{ij} as

$$A_{ij} = q_i(0) - q_j(0) \tag{3.B4a}$$

$$B_{ij} = q_i(1) - q_j(1). \tag{3.B4b}$$

Assuming a pairwise decomposable form of potential, a single pair potential V_{ij} can be Taylor expanded with respect to the weighed difference of the path parameters $\left(\gamma_{ij}a_{i,k} - \gamma_{ji}a_{j,k}\right)$ such that

$$\begin{aligned}
&\int_0^1 du\, V_{ij}\big(q_i(u) - q_i(u)\big) \\
&= \int_0^1 du\, V_{ij}\big(A_{ij} + \big(B_{ij} - A_{ij}\big)u\big) - \tfrac{1}{\beta}\textstyle\sum_{k=1}^{\infty}\big(\gamma_{ij}a_{i,k} - \gamma_{ji}a_{j,k}\big)b_{0,ij,k} \\
&+ \tfrac{1}{2\beta}\textstyle\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\big(\gamma_{ij}a_{i,k} - \gamma_{ji}a_{j,k}\big)\big(\mathcal{A}_{ij,kl} - \delta_{kl}\big)\big(\gamma_{ij}a_{i,l} - \gamma_{ji}a_{j,l}\big) + \mathcal{O}(\hbar^3)
\end{aligned} \tag{3.B5}$$

with the notation

$$\tfrac{1}{2\beta}\big(\mathcal{A}_{ij,kl} - \delta_{kl}\big) = \tfrac{1}{kl\pi^2}{\sigma_{ij}}^2 \textstyle\int_0^1 du\, V_{ij}''\big(A_{ij} + \big(B_{ij} - A_{ij}\big)u\big)\, sin(k\pi u)\, sin(l\pi u) \tag{3.B6a}$$

$$\tfrac{1}{\beta}b_{0,ij,k} = -\tfrac{\sqrt{2}}{k\pi}\sigma_{ij}\textstyle\int_0^1 du\, V_{ij}'\big(A_{ij} + \big(B_{ij} - A_{ij}\big)u\big)\, sin(k\pi u) \tag{3.B6b}$$

for σ_{ij}, γ_{ij}, $b_{\xi,k}$, and $b_{\lambda,k}$ defined according to Eqs. (3.A2a), (3.A2b), (3.A5a), and (3.A5b). Combining Eq. (3.B3) and (3.B5) gives the expression for $exp\big(-\beta V_{CG,N}\big)$ as

$$\begin{aligned}
&e^{-\beta V_{CG,N}(\boldsymbol{q},\boldsymbol{Q},\boldsymbol{q}')} \\
&= \big(\textstyle\prod_{i=1}^{N}\sigma_i/\sqrt{8\pi}\big)exp\left(-\tfrac{1}{2}\Delta\boldsymbol{q}^T\boldsymbol{\Sigma}\Delta\boldsymbol{q}\right)exp\left(-\beta\textstyle\sum_{\langle i,j\rangle}\int_0^1 du\, V_{ij}\big(A_{ij} + \big(B_{ij} - A_{ij}\big)u\big)\right) \\
&\times \textstyle\int \frac{d\boldsymbol{\xi}d\boldsymbol{a}}{\sqrt{2\pi}} \\
&\times \begin{bmatrix} exp\left(-\frac{1}{2}\boldsymbol{a}^2\right)exp\big(\sum_{\langle i,j\rangle}\sum_{k=1}^{\infty}\big(\gamma_{ij}a_{i,k} - \gamma_{ji}a_{j,k}\big)b_{0,ij,k}\big) \\ exp\left(-\frac{1}{2}\sum_{\langle i,j\rangle}\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\big(\gamma_{ij}a_{i,k} - \gamma_{ji}a_{j,k}\big)\big(\mathcal{A}_{ij,kl} - \delta_{kl}\big)\big(\gamma_{ij}a_{i,l} - \gamma_{ji}a_{j,l}\big)\right) \\ exp\big(i\boldsymbol{\xi}\cdot(\boldsymbol{Q} - \overline{\boldsymbol{q}})\big)exp\big(-i\sum_{m=1}^{N}\sum_{k=1}^{\infty}\sigma_m\xi_m a_{m,k}b_{\xi,k}\big)exp\big(-\beta\mathcal{O}(\hbar^3)\big) \end{bmatrix}
\end{aligned} \tag{3.B7}$$

where $\langle i,j\rangle$ represents a sum over all unique pair among the N particles. Note that it can alternatively be represented as $\sum_{\langle i,j\rangle} x_{ij} = \frac{1}{2}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N} x_{ij}$ for some pairwise quantity x_{ij}.

The expression shown in Eq. (3.B7) is an exact quantum result that includes the higher order terms $\mathcal{O}(\hbar^3)$. For the rest of the derivation, however, we focus on the expression of second order truncated portion of $exp\left(-\beta V_{CG,N}\right)$ following the approach of the original one-body CG-PI derivation (Ref. 57) by omitting the higher order terms. We name the second order truncated many-body CG potential $V_{CG,N,trunc}$ whose distribution can be written as

$$
\begin{aligned}
&e^{-\beta V_{CG,N,trunc}}\\
&= \left(\prod_{i=1}^{N}\sigma_i/\sqrt{8\pi}\right) exp\left(-\frac{1}{2}\Delta\boldsymbol{q}^T\boldsymbol{\Sigma}\Delta\boldsymbol{q}\right) exp\left(-\beta\sum_{\langle i,j\rangle}\int_0^1 du\, V_{ij}\left(A_{ij}+\left(B_{ij}-A_{ij}\right)u\right)\right)\\
&\times\int\frac{d\boldsymbol{\xi}d\boldsymbol{a}}{\sqrt{2\pi}}\\
&\times\begin{bmatrix} exp\left(-\frac{1}{2}\boldsymbol{a}^2\right) exp\left(\sum_{\langle i,j\rangle}\sum_{k=1}^{\infty}\left(\gamma_{ij}a_{i,k}-\gamma_{ji}a_{j,k}\right)b_{0,ij,k}\right)\\ exp\left(-\frac{1}{2}\sum_{\langle i,j\rangle}\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\left(\gamma_{ij}a_{i,k}-\gamma_{ji}a_{j,k}\right)\left(\mathcal{A}_{ij,kl}-\delta_{kl}\right)\left(\gamma_{ij}a_{i,l}-\gamma_{ji}a_{j,l}\right)\right)\\ exp\left(i\boldsymbol{\xi}\cdot\left(\boldsymbol{Q}-\overline{\boldsymbol{q}}\right)\right)exp\left(-i\sum_{m=1}^{N}\sum_{k=1}^{\infty}\sigma_m\xi_m a_{m,k}b_{\xi,k}\right)\end{bmatrix}.
\end{aligned}
\tag{3.B8}
$$

The quadratic terms with respect to $a_{i,k}$ can be simplified such that

$$
\begin{aligned}
&-\frac{1}{2}\sum_{\langle i,j\rangle}\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\left(\gamma_{ij}a_{i,k}-\gamma_{ji}a_{j,k}\right)\left(\mathcal{A}_{ij,kl}-\delta_{kl}\right)\left(\gamma_{ij}a_{i,l}-\gamma_{ji}a_{j,l}\right)\\
&= -\frac{1}{2}\boldsymbol{a}^T\boldsymbol{\Omega}\boldsymbol{a}
\end{aligned}
\tag{3.B9}
$$

where the N dimensional square block matrix $\boldsymbol{\Omega}$ has been defined in Eq. (3.A3), but the elements can alternatively be written as

$$
\Omega_{ij,kl} = \begin{cases} \sum_{\substack{m=1\\ m\neq i}}^{N}{\gamma_{im}}^2\left(\left(\mathcal{A}_{im,kl}+\mathcal{A}_{mi,kl}\right)/2-\delta_{kl}\right) & (for\ i=j)\\ \delta_{kl}-\left(\mathcal{A}_{ij,kl}+\mathcal{A}_{ji,kl}\right)/2 & (for\ i\neq j)\end{cases}. \tag{3.B10}
$$

using the notation of $\mathcal{A}_{ij,kl}$ introduced according to Eq. (3.B6a). Moreover, the linear term with respect to $a_{i,k}$ involving $b_{0,ij,k}$ also can be simplified into a vector notation. First, we have a simple index change

$$\sum_{\langle i,j\rangle}\sum_{k=1}^{\infty}\left(\gamma_{ij}a_{i,k}-\gamma_{ji}a_{j,k}\right)b_{0,ij,k}$$

$$=\frac{1}{2}\sum_{i=1}^{N}\sum_{k=1}^{\infty}a_{i,k}\left(\sum_{\substack{j=1\\ j\neq i}}^{N}\gamma_{ij}\left(b_{0,ij,k}-b_{0,ji,k}\right)\right). \tag{3.B11}$$

Then, Taylor expanding $b_{0,ij,k}$ up to second order with respect to the pair potential gives

$$b_{0,ij,k}=\beta\sigma_{ij}\left(b_{\xi,k}V_{ij}'\left(\overline{q}_i-\overline{q}_j\right)+b_{\lambda,k}V_{ij}''\left(\overline{q}_i-\overline{q}_j\right)\cdot\left(\Delta q_i-\Delta q_j\right)\right). \tag{3.B12}$$

Then, using notation of $\overline{V_{ij}'}=V_{ij}'\left(\overline{q}_i-\overline{q}_j\right)-V_{ij}'\left(\overline{q}_j-\overline{q}_i\right)$ and $\overline{V_{ij}''}$ as defined in Eq. (3.A2c), and N dimensional column vector $\boldsymbol{F}$ and $\boldsymbol{G}$ whose ith element F_i and G_i are respectively defined as

$$F_i=\sum_{\substack{j=1\\ j\neq i}}^{N}\overline{V_{ij}'} \tag{3.B13a}$$

$$G_i=\sum_{\substack{j=1\\ j\neq i}}^{N}\left(\Delta q_i-\Delta q_j\right)\overline{V_{ij}''}, \tag{3.B13b}$$

RHS of Eq. (3.B11) can be rewritten as

$$\frac{1}{2}\sum_{i=1}^{N}\sum_{k=1}^{\infty}a_{i,k}\left(\sum_{\substack{j=1\\ j\neq i}}^{N}\gamma_{ij}\left(b_{0,ij,k}-b_{0,ji,k}\right)\right)$$

$$=\frac{\beta}{2}\sum_{i=1}^{N}\sum_{k=1}^{\infty}\sigma_i a_{i,k}\left(b_{\xi,k}F_i+b_{\lambda,k}G_i\right). \tag{3.B14}$$

The linear exponent term on the RHS of Eq. (3.B14) and the linear exponent term that involves imaginary number in Eq. (3.B8) can be combined to have

$$\frac{\beta}{2}\sum_{m=1}^{N}\sum_{k=1}^{\infty}\sigma_m a_{m,k}\left(b_{\xi,k}F_m+b_{\lambda,k}G_m\right)-i\sum_{m=1}^{N}\sum_{k=1}^{\infty}\sigma_m\xi_m a_{m,k}b_{\xi,k}$$

$$=\boldsymbol{\mathcal{B}}\cdot\boldsymbol{a} \tag{3.B15}$$

where $\boldsymbol{\mathcal{B}}$ is a N dimensional block column vector and has $\boldsymbol{\mathcal{B}}_m$ as its mth block element, and the infinite dimensional column vector $\boldsymbol{\mathcal{B}}_m$ is defined as

$$\boldsymbol{\mathcal{B}}_m = \left(i\sigma_m\xi_m + \frac{\beta}{2}F_m\sigma_m\right)\boldsymbol{b}_\xi + \left(\frac{\beta}{2}G_m\sigma_m\right)\boldsymbol{b}_\lambda. \tag{3.B16}$$

where $\boldsymbol{b}_\xi$ and $\boldsymbol{b}_\lambda$ are infinite dimensional column vectors with kth element defined according to Eq (3.A5a) and (3.A5b), respectively.

Combining the vector notations of the quadratic and the linear terms we can write a more concise vector notation version of $exp\left(-\beta V_{CG,N,trunc}\right)$ such that

$$\begin{aligned} e^{-\beta V_{CG,N,trunc}} &= \left(\prod_{i=1}^{N}\sigma_i/\sqrt{8\pi}\right) exp\left(-\frac{1}{2}\Delta\boldsymbol{q}^T\boldsymbol{\Sigma}\Delta\boldsymbol{q}\right) \\ &\times\ exp\left(-\beta\sum_{\langle i,j\rangle}\int_0^1 du\, V_{ij}\left(A_{ij} + \left(B_{ij} - A_{ij}\right)u\right)\right) \\ &\times \int \frac{d\boldsymbol{\xi}d\boldsymbol{a}}{\sqrt{2\pi}} exp\left(-\frac{1}{2}\boldsymbol{a}^T(\boldsymbol{\Omega}+\boldsymbol{I})\boldsymbol{a} + \boldsymbol{\mathcal{B}}\cdot\boldsymbol{a} + i\boldsymbol{\xi}\cdot(\boldsymbol{Q}-\overline{\boldsymbol{q}})\right) \end{aligned} \tag{3.B17}$$

in which $\boldsymbol{I}$ is an identity matrix with same dimensionality as $\boldsymbol{\Omega}$. Performing the infinite dimensional Gaussian integral with respect to the vector $\boldsymbol{a}$ gives

$$\int \frac{d\boldsymbol{a}}{\sqrt{2\pi}} exp\left(-\frac{1}{2}\boldsymbol{a}^T(\boldsymbol{\Omega}+\boldsymbol{I})\boldsymbol{a} + \boldsymbol{\mathcal{B}}\cdot\boldsymbol{a}\right) = \sqrt{\frac{1}{det(\boldsymbol{\Omega}+\boldsymbol{I})}} exp\left(\frac{1}{2}\boldsymbol{\mathcal{B}}^T(\boldsymbol{\Omega}+\boldsymbol{I})^{-1}\boldsymbol{\mathcal{B}}\right). \tag{3.B18}$$

Using the RHS of Eq. (3.B18), the truncated CG distribution then can be written as

$$\begin{aligned} & e^{-\beta V_{CG,N,trunc}} \\ &= \left(\prod_{i=1}^{N}\sigma_i/\sqrt{8\pi}\right) exp\left(-\frac{1}{2}\Delta\boldsymbol{q}^T\boldsymbol{\Sigma}\Delta\boldsymbol{q}\right) exp\left(-\beta\sum_{\langle i,j\rangle}\int_0^1 du\, V_{ij}\left(A_{ij} + \left(B_{ij} - A_{ij}\right)u\right)\right) \\ &\times \sqrt{1/det(\boldsymbol{\Omega}+\boldsymbol{I})}\int d\boldsymbol{\xi}\, exp\left(\frac{1}{2}\boldsymbol{\mathcal{B}}^T(\boldsymbol{\Omega}+\boldsymbol{I})^{-1}\boldsymbol{\mathcal{B}} + i\boldsymbol{\xi}\cdot(\boldsymbol{Q}-\overline{\boldsymbol{q}})\right). \end{aligned} \tag{3.B19}$$

The matrix $\boldsymbol{\Lambda}$ is also a N dimensional square block matrix, similar to $\boldsymbol{\Omega}$. We note the ith row and jth column block element of $\boldsymbol{\Lambda}$ as $\boldsymbol{\Lambda}_{ij}$, and the block elements $\boldsymbol{\Lambda}_{ij}$ are themselves infinite

dimensional square matrices. We define N dimensional column vector $\boldsymbol{S}$ defined as $\boldsymbol{S} = \boldsymbol{LF}$ for matrix $\boldsymbol{L}$ defined as in Eq. (3.A6a).

Then, the exponent of the integrand in Eq. (3.B19) then can be written out as

$$\frac{1}{2}\boldsymbol{\mathcal{B}}^T\boldsymbol{\Lambda}\boldsymbol{\mathcal{B}} + i\boldsymbol{\xi}\cdot(\boldsymbol{Q}-\overline{\boldsymbol{q}})$$

$$= -\frac{1}{2}\boldsymbol{\xi}^T\boldsymbol{L}\boldsymbol{\xi} + i\boldsymbol{\xi}\cdot\left(\frac{\beta}{2}\boldsymbol{S} + (\boldsymbol{Q}-\overline{\boldsymbol{q}})\right) + \frac{\beta^2}{8}(\boldsymbol{F}^T\boldsymbol{L}\boldsymbol{F} + \boldsymbol{G}^T\boldsymbol{R}\boldsymbol{G}). \tag{3.B20}$$

Now, the Gaussian integral with respect to $\boldsymbol{\xi}$ gives

$$\int d\boldsymbol{\xi}\, exp\left(-\frac{1}{2}\boldsymbol{\xi}^T\boldsymbol{L}\boldsymbol{\xi} + i\boldsymbol{\xi}\cdot\left(\frac{\beta}{2}\boldsymbol{S} + (\boldsymbol{Q}-\overline{\boldsymbol{q}})\right)\right)$$

$$= \sqrt{(2\pi)^N/det(\boldsymbol{L})}$$

$$\times\ exp\left(-\frac{1}{2}\left(\frac{\beta^2}{4}\boldsymbol{F}^T\boldsymbol{L}\boldsymbol{F} + \beta\boldsymbol{F}\cdot(\boldsymbol{Q}-\overline{\boldsymbol{q}}) + (\boldsymbol{Q}-\overline{\boldsymbol{q}})^T\boldsymbol{L}^{-1}(\boldsymbol{Q}-\overline{\boldsymbol{q}})\right)\right). \tag{3.B21}$$

We can write then

$$e^{-\beta V_{CG,N,trunc}} = \left(\prod_{i=1}^{N}\sigma_i/\sqrt{8\pi}\right)\sqrt{(2\pi)^N/\left(det(\boldsymbol{\Omega}+\boldsymbol{I})det(\boldsymbol{L})\right)}$$

$$\times\ exp\left(-\frac{1}{2}\Delta\boldsymbol{q}^T\boldsymbol{\Sigma}\Delta\boldsymbol{q}\right)exp\left(-\beta\sum_{\langle i,j\rangle}\int_0^1 du\, V_{ij}\left(A_{ij} + \left(B_{ij}-A_{ij}\right)u\right)\right)$$

$$\times\ exp\left(\frac{\beta^2}{8}\boldsymbol{G}^T\boldsymbol{R}\boldsymbol{G} - \frac{\beta}{2}\boldsymbol{F}\cdot(\boldsymbol{Q}-\overline{\boldsymbol{q}}) - \frac{1}{2}(\boldsymbol{Q}-\overline{\boldsymbol{q}})^T\boldsymbol{L}^{-1}(\boldsymbol{Q}-\overline{\boldsymbol{q}})\right). \tag{3.B22}$$

Alternatively, $V_{CG,N,trunc}$ can be written as

$$V_{CG,N,trunc}$$

$$= \frac{1}{2\beta}\Delta\boldsymbol{q}^T\boldsymbol{\Sigma}\Delta\boldsymbol{q} + \sum_{\langle i,j\rangle}\int_0^1 du\, V_{ij}\left(A_{ij} + \left(B_{ij}-A_{ij}\right)u\right)$$

$$-\frac{\beta}{8}\boldsymbol{G}^T\boldsymbol{R}\boldsymbol{G} + \frac{1}{2}\boldsymbol{F}\cdot(\boldsymbol{Q}-\overline{\boldsymbol{q}}) + \frac{1}{2\beta}(\boldsymbol{Q}-\overline{\boldsymbol{q}})^T\boldsymbol{L}^{-1}(\boldsymbol{Q}-\overline{\boldsymbol{q}})$$

$$-\frac{1}{\beta}ln\left[\left(\prod_{i=1}^{N}\sigma_i/\sqrt{8\pi}\right)\sqrt{(2\pi)^N/\left(det(\boldsymbol{\Omega}+\boldsymbol{I})det(\boldsymbol{L})\right)}\right]. \tag{3.B23}$$

Rest of the section will focus on Taylor expanding and simplifying the expression shown in Eq. (3.B23).

The pair potential term in Eq. (3.B23) can be Taylor expanded up to second order as

$$\int_0^1 du\, V_{ij}\big(A_{ij} + \big(B_{ij} - A_{ij}\big)u\big)$$
$$= V_{ij}\left(\overline{q}_i - \overline{q}_j\right) + \frac{1}{24}V_{ij}''\left(\overline{q}_i - \overline{q}_j\right)\cdot\big(\Delta q_i - \Delta q_j\big)^2 + \mathcal{O}\big(V_{ij}'''\big). \tag{3.B24}$$

Combining Eq. (3.B23) and the first two terms of RHS of (3.B24), $V_{CG,N,trunc}$ can be written as

$$V_{CG,N,trunc}$$
$$= \sum_{\langle i,j\rangle} V_{ij}\left(\overline{q}_i - \overline{q}_j\right) + \frac{1}{2}\boldsymbol{F}\cdot(\boldsymbol{Q} - \overline{\boldsymbol{q}}) + \frac{1}{24}\sum_{\langle i,j\rangle}\left(V_{ij}''\left(\overline{q}_i - \overline{q}_j\right) + V_{ij}''\left(\overline{q}_j - \overline{q}_i\right)\right)\Delta {q_i}^2$$
$$-\frac{1}{12}\sum_{\langle i,j\rangle}\Delta q_i \Delta q_j V_{ij}''\left(\overline{q}_i - \overline{q}_j\right) + \frac{1}{2\beta}\Delta\boldsymbol{q}^T\boldsymbol{\Sigma}\Delta\boldsymbol{q} - \frac{\beta}{8}\boldsymbol{G}^T\boldsymbol{R}\boldsymbol{G} + \frac{1}{2\beta}\boldsymbol{X}^T\boldsymbol{L}^{-1}\boldsymbol{X}$$
$$-\frac{1}{\beta}ln\left[\left(\prod_{i=1}^N \sigma_i/\sqrt{8\pi}\right)\sqrt{(2\pi)^N/\big(det(\boldsymbol{\Omega}+\boldsymbol{I})det(\boldsymbol{L})\big)}\right]. \tag{3.B25}$$

Similar to the derivation of the original CG-PI work, the key to deriving Eq. (3.12) is Taylor expansion of $V_{ij}\big(Q_i - Q_j\big)$ at $\overline{q}_i - \overline{q}_j$. This means that for Eq. (3.B25) we make a substitution of

$$\sum_{\langle i,j\rangle} V_{ij}\left(\overline{q}_i - \overline{q}_j\right) + \frac{1}{2}\boldsymbol{F}\cdot(\boldsymbol{Q} - \overline{\boldsymbol{q}})$$
$$\rightarrow \sum_{\langle i,j\rangle}\left(V_{ij}\big(Q_i - Q_j\big) - \frac{1}{2}V_{ij}''\left(\overline{q}_i - \overline{q}_j\right)\left(Q_i - \overline{q}_i - Q_j + \overline{q}_j\right)^2\right) \tag{3.B26}$$

such that the resulting expression $V_{CG,N,trunc}$ becomes

$$V_{CG,N,trunc}$$
$$\cong \sum_{\langle i,j\rangle} V_{ij}\big(Q_i - Q_j\big) + \frac{1}{2\beta}(\boldsymbol{Q} - \overline{\boldsymbol{q}})^T\boldsymbol{L}^{-1}(\boldsymbol{Q} - \overline{\boldsymbol{q}}) - \frac{1}{4}\sum_{i=1}^N\sum_{\substack{j=1\\ j\neq i}}^N\big(Q_i - \overline{q}_i\big)^2\overline{V_{ij}''}$$
$$+\frac{1}{2}\sum_{i=1}^N\sum_{\substack{j=1\\ j\neq i}}^N\big(Q_i - \overline{q}_i\big)\overline{V_{ij}''}\left(Q_j - \overline{q}_j\right) + \frac{1}{48}\sum_{i=1}^N\Delta {q_i}^2\left(\sum_{\substack{j=1\\ j\neq i}}^N\overline{V_{ij}''}\right)$$

$$-\frac{1}{24}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N} V_{ij}''\left(\overline{q}_i-\overline{q}_j\right)\cdot\Delta q_i\Delta q_j+\frac{1}{2\beta}\sum_{i=1}^{N}\Delta q_i{}^2/\sigma_i{}^2-\frac{\beta}{8}\boldsymbol{G}^T\boldsymbol{R}\boldsymbol{G}$$

$$-\frac{1}{\beta}ln\left[\left(\prod_{i=1}^{N}\sigma_i/\sqrt{8\pi}\right)\sqrt{(2\pi)^N/\left(det(\boldsymbol{\Omega}+\boldsymbol{I})det(\boldsymbol{L})\right)}\right] \tag{3.B27}$$

To finally organize the terms in Eq. (3.B27), it is easy see that for N by N matrices $\boldsymbol{C}$ and $\boldsymbol{M}$ defined according to Eq. (3.A7a) and (3.A7c), respectively, the terms in Eq. (3.B27) can be simplified as

$$-\frac{1}{4}\left(\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\left(Q_i-\overline{q}_i\right)^2\overline{V_{ij}''}-2\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\left(Q_i-\overline{q}_i\right)\overline{V_{ij}''}\left(Q_j-\overline{q}_j\right)\right)$$

$$=-(\boldsymbol{Q}-\overline{\boldsymbol{q}})^T\frac{\boldsymbol{M}}{4}(\boldsymbol{Q}-\overline{\boldsymbol{q}}) \tag{3.B28a}$$

$$\frac{1}{48}\left(\sum_{i=1}^{N}\Delta q_i{}^2\left(\sum_{\substack{j=1\\ j\neq i}}^{N}\overline{V_{ij}''}\right)-2\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}V_{ij}''\left(\overline{q}_i-\overline{q}_j\right)\cdot\Delta q_i\Delta q_j\right)$$

$$=\Delta\boldsymbol{q}^T\frac{\boldsymbol{M}}{48}\Delta\boldsymbol{q} \tag{3.B28b}$$

$$-\frac{\beta}{8}\boldsymbol{G}^T\boldsymbol{R}\boldsymbol{G}=-\frac{\beta}{8}\Delta\boldsymbol{q}^T\boldsymbol{C}\Delta\boldsymbol{q} \tag{3.B28c}$$

to give the final truncated expression in the first square bracket of Eq. (3.12).

Appendix 3C: Derivation of the Energy Estimator

In deriving Eq. (3.13), we start with a well-known relationship between energy and logarithmic derivative of the canonical partition function Z. The energy of the system can be written as

$$E=-\frac{\partial}{\partial\beta}\ln Z. \tag{3.C1}$$

Using the CG-PI expression for the thermal density matrix as shown in Eq. (3.9), the canonical partition function can be written as

$$Z = \left(\prod_{i=1}^{N} m_i/(\pi\hbar^2\beta)\right) \int d\boldsymbol{Q} d\overline{\boldsymbol{q}}\, exp\left(-\beta V_{CG,N}\right). \tag{3.C2}$$

Note that the many-body CG potential $V_{CG,N}$ shown in this derivation has $\Delta\boldsymbol{q} = \boldsymbol{0}$ as we are only concerned with the diagonal thermal density matrix element. Combining Eqs. (3.C1) and (3.C2), it is easy to show that

$$-\frac{\partial}{\partial\beta}\ln Z = \langle\varepsilon(\boldsymbol{Q}, \overline{\boldsymbol{q}})\rangle_{CG,N} \tag{3.C3}$$

with the energy estimator $\varepsilon(\boldsymbol{Q}, \overline{\boldsymbol{q}}) = \frac{N}{\beta} + \beta\frac{\partial}{\partial\beta}V_{CG,N} + V_{CG,N}$ from Eq. (3.13) and expectation value notation $\langle f\rangle_{CG,N}$ according to Eq. (3.15) respectively.

To obtain more physical insight from the functional form of the CG-PI energy estimator, we show how $\varepsilon(\boldsymbol{Q}, \overline{\boldsymbol{q}})$ is related to the energy estimator used in in imaginary time path integrals as shown in Eq. (1.22). We can do this by expanding the second term of $\varepsilon(\boldsymbol{Q}, \overline{\boldsymbol{q}})$. Using the exact definition of $V_{CG,N}$ as shown in Eq. (3.10), we can write

$$\beta\frac{\partial}{\partial\beta}V_{CG,N} = -V_{CG,N} - \frac{\partial}{\partial\beta}\lim_{P\to\infty}\left(\ln Z_{CG,N}^{(P)}\right). \tag{3.C4}$$

The last term in Eq. (3.C4) can also be further expanded to have

$$\begin{aligned}\frac{\partial}{\partial\beta}\lim_{P\to\infty}\left(\ln Z_{CG,N}^{(P)}\right) &= \frac{N\left(1-\frac{P}{2}\right)}{\beta} \\ &+ \frac{\lim_{P\to\infty}\int\left[\prod_{i=1}^{N}\prod_{j=2}^{P} dq_j^{(i)}\right]\varphi e^{-\beta V_{P,N}}\left[\prod_{i=1}^{N}\delta\left(Q_i - \frac{q_2^{(i)}+\cdots+q_P^{(i)}}{P-1}\right)\right]}{\lim_{P\to\infty}\int\left[\prod_{i=1}^{N}\prod_{j=2}^{P} dq_j^{(i)}\right]e^{-\beta V_{P,N}}\left[\prod_{i=1}^{N}\delta\left(Q_i - \frac{q_2^{(i)}+\cdots+q_P^{(i)}}{P-1}\right)\right]}\end{aligned} \tag{3.C5}$$

in which the quantity φ is given as

$$\varphi = \sum_{k=1}^{P}\sum_{i=1}^{N}\frac{m_i P}{2\hbar^2\beta^2}\left(q_k^{(i)} - q_{k+1}^{(i)}\right)^2 - \frac{1}{P}V\left(q_k^{(1)}, \ldots, q_k^{(N)}\right). \tag{3.C6}$$

For simplicity, we will note $\langle \ldots \rangle_{Z_{CG,N}^{(P)}}$ to be taking an expectation value with respect to the partial partition function defined by Eq. (3.11) such that

$$\langle f \rangle_{Z_{CG,N}^{(P)}} = \frac{\lim_{P\to\infty} \int \left[\prod_{i=1}^{N}\prod_{j=2}^{P} dq_j^{(i)}\right] f e^{-\beta V_{P,N}} \left[\prod_{i=1}^{N} \delta\left(Q_i - \frac{q_2^{(i)}+\cdots+q_P^{(i)}}{P-1}\right)\right]}{\lim_{P\to\infty} \int \left[\prod_{i=1}^{N}\prod_{j=2}^{P} dq_j^{(i)}\right] e^{-\beta V_{P,N}} \left[\prod_{i=1}^{N} \delta\left(Q_i - \frac{q_2^{(i)}+\cdots+q_P^{(i)}}{P-1}\right)\right]}. \tag{3.C7}$$

Using this notation, we can rewrite Eq. (3.C5) as

$$\frac{\partial}{\partial\beta} \lim_{P\to\infty} \left(\ln Z_{CG,N}^{(P)}\right) = N\left(1 - \frac{P}{2}\right)/\beta + \langle\varphi\rangle_{Z_{CG,N}^{(P)}}. \tag{3.C8}$$

Finally, combining Eq. (3.C4) and (3.C8), we have

$$\varepsilon(\boldsymbol{Q}, \overline{\boldsymbol{q}}) = \frac{NP}{2\beta} + \left\langle -\sum_{k=1}^{P}\sum_{i=1}^{N} \left(\begin{array}{l} \frac{m_i P}{2\hbar^2\beta^2}\left(q_k^{(i)} - q_{k+1}^{(i)}\right)^2 \\ +\frac{1}{P} V\left(q_k^{(1)}, \ldots, q_k^{(N)}\right) \end{array} \right) \right\rangle_{Z_{CG,N}^{(P)}}. \tag{3.C9}$$

We see that the resulting expression for the CG-PI energy estimator very closely resembles that of the imaginary time path integral formalism as shown in Eq. (1.22). The only difference is that the second term in Eq. (3.C9) is pre-averaged with $\langle \ldots \rangle_{Z_{CG,N}^{(P)}}$, which is equivalent to coarse-graining the intermediate beads of the ring polymer. It should be noted that the CG-PI estimator itself is partially an expectation value of the full ring polymer energy with respect to the intermediate bead positions. The expectation value of energy is taken with respect to the distribution $exp\left(-\beta V_{CG,N}\right)$ which is a function of the coordinates of the pseudo-particles.

It is important to analyze the three energy contributions in Eq. (3.13). The first term comes from the kinetic energies of the N pseudo-particle pairs, with a contribution of $1/(2\beta)$ for each pseudo-particle. According to Eqs. (3.C4) and (3.C8), the second term can be understood as the full imaginary path integral energy of the system with the intermediate bead positions averaged out ($NP/(2\beta) - \langle\varphi\rangle_{Z_{CG,N}^{(P)}}$) without the pseudo-particle pair kinetic energy and interaction potential ($-V_{CG,N} - N/\beta$). Lastly, the third term is a contribution from the effective interaction between the pseudo-particles ($\boldsymbol{Q}, \overline{\boldsymbol{q}}$) according to the many-body effective CG potential $V_{CG,N}$. It should be noted that the discussion and derivation of estimators in this work establishes the

analytical theory for formalism purposes and does not discuss its numerical implementation, especially due to the challenges of numerically calculating $\frac{\partial}{\partial\beta}V_{CG,N}$. This issue still requires more work and will be addressed in the future works.

Appendix 3D: Derivation of the Many-Body Effective Hamiltonian

We start with the expectation value of a many-body quantum operator $\hat{A}$ such that

$$\langle\hat{A}(\{\hat{p}_i,\hat{q}_i;i=1,\dots,N\})\rangle=\frac{1}{Z}\int d\boldsymbol{q}d\boldsymbol{q}'\,\langle\boldsymbol{q}|e^{-\beta\hat{H}}|\boldsymbol{q}'\rangle\,\langle\boldsymbol{q}'|\hat{A}|\boldsymbol{q}\rangle. \tag{3.D1}$$

For $Z=\int d\boldsymbol{q}\,\langle\boldsymbol{q}|e^{-\beta\hat{H}}|\boldsymbol{q}\rangle$. Using Eq. (3.9), Eq. (3.D1) can be rewritten as

$$\langle\hat{A}\rangle=(\textstyle\prod_{i=1}^{N}m_i/(\pi\hbar^2\beta))/Z$$
$$\times\int d\boldsymbol{q}d\boldsymbol{q}'d\boldsymbol{Q}d\boldsymbol{p}/(2\pi\hbar)\;e^{-\beta V_{CG,N}(\boldsymbol{q},\boldsymbol{Q},\boldsymbol{q}')}e^{i(\boldsymbol{p}\cdot\Delta\boldsymbol{q})/\hbar}\;W_A(\boldsymbol{p},\overline{\boldsymbol{q}}) \tag{3.D2}$$

for N dimensional column momentum vector $\boldsymbol{p}$ with momentum for the ith particle p_i as its ith element and many-body Weyl map $W_A(\boldsymbol{p},\overline{\boldsymbol{q}})$. The denominator of Eq. (3.D2) can be rewritten using the identity of the delta function as

$$Z=(\textstyle\prod_{i=1}^{N}m_i/(\pi\hbar^2\beta))\int d\boldsymbol{q}d\boldsymbol{Q}\;e^{-\beta V_{CG,N}(\boldsymbol{q},\boldsymbol{Q},\boldsymbol{q})}$$
$$=(\textstyle\prod_{i=1}^{N}m_i/(\pi\hbar^2\beta))\int d\overline{\boldsymbol{q}}d\Delta\boldsymbol{q}d\boldsymbol{Q}d\boldsymbol{p}/(2\pi\hbar)\;e^{-\beta V_{CG,N}\left(\overline{\boldsymbol{q}}-\frac{\Delta\boldsymbol{q}}{2},\boldsymbol{Q},\overline{\boldsymbol{q}}+\frac{\Delta\boldsymbol{q}}{2}\right)}e^{i(\boldsymbol{p}\cdot\Delta\boldsymbol{q})/\hbar} \tag{3.D3}$$

such that $\langle\hat{A}\rangle$ can be rewritten according to Eq. (3.16). In deriving this relation, the numerator's integration variable was changed in a trivial fashion and constant term of

$$\int d\boldsymbol{P}_Q\;e^{-\frac{\beta}{2}\boldsymbol{P}_Q{}^T\boldsymbol{M}_Q^{-1}\boldsymbol{P}_Q}=\textstyle\prod_{i=1}^{N}\sqrt{2\pi M_{Q,i}/\beta} \tag{3.D4}$$

was added to both numerator and denominator for N dimensional column vector $\boldsymbol{P}_Q$ and N dimensional square diagonal matrix $\boldsymbol{M}_Q$ such that with $P_{Q,i}$ and $M_{Q,i}$ as their ith element respectively. The momenta $P_{Q,i}$ and mass $M_{Q,i}$ term corresponds to the momenta and mass associated with the ith centroid pseudo-particle.

To derive Eq. (3.18), combine the truncated term from Eq. (3.12) and Eq. (3.17) to have

$$H_{eff,N}\left(\boldsymbol{P}_Q,\boldsymbol{p},\boldsymbol{Q},\overline{\boldsymbol{q}}\right)\cong\frac{1}{2}{\boldsymbol{P}_Q}^T\boldsymbol{M}_Q^{-1}\boldsymbol{P}_Q+V(\boldsymbol{Q})+\frac{1}{2}(\boldsymbol{Q}-\overline{\boldsymbol{q}})^T\boldsymbol{K}_{Q\overline{q}}(\boldsymbol{Q}-\overline{\boldsymbol{q}})$$

$$-\frac{1}{\beta}\ln\left[const\cdot e^{-\beta K_0}\cdot\int d\boldsymbol{\Delta q}\,e^{i(\boldsymbol{p}\cdot\Delta\boldsymbol{q})/\hbar-\frac{\beta}{2}\boldsymbol{\Delta q}^T\boldsymbol{K}_{\Delta q^2}\boldsymbol{\Delta q}}\right] \tag{3.D5}$$

in which the Gaussian integral can be computed as

$$\int d\boldsymbol{\Delta q}\,exp\left(-\frac{\beta}{2}\boldsymbol{\Delta q}^T\boldsymbol{K}_{\Delta q^2}\boldsymbol{\Delta q}+\frac{i(\boldsymbol{p}\cdot\Delta\boldsymbol{q})}{\hbar}\right)$$

$$=\sqrt{\frac{(2\pi/\beta)^N}{det\left(\boldsymbol{K}_{\Delta q^2}\right)}}exp\left(-\frac{1}{2\beta\hbar^2}\boldsymbol{p}^T\boldsymbol{K}_{\Delta q^2}^{-1}\boldsymbol{p}\right). \tag{3.D6}$$

Combining Eq. (3.D5) and (3.D6), we have

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

$$+\frac{1}{2}(\boldsymbol{Q}-\overline{\boldsymbol{q}})^T\boldsymbol{K}_{Q\overline{q}}(\boldsymbol{Q}-\overline{\boldsymbol{q}})-\frac{1}{\beta}\ln\left[const\cdot e^{-\beta K_0}\cdot\sqrt{\frac{(2\pi/\beta)^N}{det\left(\boldsymbol{K}_{\Delta q^2}\right)}}\right]. \tag{3.D7}$$

For choice of the arbitrary constant as $const=e^{\beta K_0}\cdot\sqrt{\frac{det\left(\boldsymbol{K}_{\Delta q^2}\right)}{(2\pi/\beta)^N}}$, the expression for the effective Hamiltonian becomes

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

$$+V(\boldsymbol{Q})+\frac{1}{2}(\boldsymbol{Q}-\overline{\boldsymbol{q}})^T\boldsymbol{K}_{Q\overline{q}}(\boldsymbol{Q}-\overline{\boldsymbol{q}}) \tag{3.D8}$$

which corresponds to the expression shown in Eq. (3.18). One simply substitutes the observable pseudo-particle kinetic energy term in Eq. (3.D8) with $\frac{1}{2}\boldsymbol{p}^T\boldsymbol{m}_{\overline{q}}^{-1}\boldsymbol{p}$ to arrive at the constant effective mass approximated many-body effective Hamiltonian as shown in Eq. (3.19).

Appendix 3E: Constructing the Scaling Relationship for Modified Para-Hydrogen Model

To find a scaling relationship of λ of the model para-hydrogen system as a function of mass and temperature, systems with different masses and temperatures were simulated. For each cases, the values of λ were varied so that the first peak of the radial distribution function (RDF) calculated from the n-CG-PI method corresponded to that from PIMD. The best value of λ in an increment of 0.05 were chosen. The different set of masses and temperatures of the system, the fitted λ value, and the coupling strength between the pseudo-particle pair ($\lambda \cdot 12m/(\hbar^2\beta^2)$) are shown in Table 3.1. The calculation details for the PIMD and n-CG-PI simulations were same as described in the numerical examples section.

Mass (g/mol)	Temperature (K)	Fitted λ	Coupling strength (fs^{-2} g/mol)
2	50	1.45	1.49E-03
4	50	1.30	2.67E-03
8	50	1.20	4.94E-03
16	50	1.10	9.05E-03
4	40	1.25	1.65E-03
4	60	1.40	4.15E-03
4	70	1.40	5.64E-03
4	80	1.40	7.37E-03

Table 3-1. Test cases of modified para-hydrogen model at different masses and temperatures along with the values of fitted λ and corresponding coupling strength.

The set of data shown in the last column of Table 3.1 was linearly fitted with respect to mass and temperature squared to give the result shown in Eq. (3.24). It is noted that the fitted values of λ for the last four rows of Table 3.1 are all 1.4 as the functional form of the RDF becomes less sensitive to the value of λ when the system becomes more classical and we were not able to distinguish the values of λ with the grid of 0.05. A narrower grid was not used because the last four row cases have different overall coupling strength, which is the quantity of interest in

investigating the test cases. Moreover, the fitted scaling relationship gives a reasonable good agreement with the PIMD result when used on the of $m = 5\ g/mol$ and $T = 30\ K$ systems shown in the numerical result section.

References

1. U. W. Schmitt and G. A. Voth, J. Chem. Phys. **111**, 9361-9381 (1999).

2. R. Vuilleumier and D. Borgis, J. Chem. Phys. **111**, 4251-4266 (1999).

3. S. R. Billeter, S. P. Webb, T. Iordanov, P. K. Agarwal and S. Hammes-Schiffer, J. Chem. Phys. **114**, 6925-6936 (2001).

4. M. Wang, Z. Lu and W. Yang, J. Chem. Phys. **124**, 124516 (2006).

5. D. M. Ceperley, Reviews of Modern Physics **67**, 279-355 (1995).

6. J. Chen, X.-Z. Li, Q. Zhang, M. I. J. Probert, C. J. Pickard, R. J. Needs, A. Michaelides and E. Wang, Nat. Commun. **4**, 2064 (2013).

7. M. A. Morales, J. M. McMahon, C. Pierleoni and D. M. Ceperley, Phys. Rev. Lett. **110** (6), 065702 (2013).

8. C. M. Herdman, A. Rommal and A. Del Maestro, Phys. Rev. B **89**, 224502 (2014).

9. M. E. Tuckerman, D. Marx, M. L. Klein and M. Parrinello, Science **275**, 817-820 (1997).

10. B. Walker and A. Michaelides, J. Chem. Phys. **133**, 174306 (2010).

11. H. A. Stern and B. J. Berne, J. Chem. Phys. **115**, 7622-7628 (2001).

12. M. Shiga and W. Shinoda, J. Chem. Phys. **123**, 134502 (2005).

13. F. Paesani, W. Zhang, D. A. Case, T. E. CheathamIII and G. A. Voth, J. Chem. Phys. **125**, 184507 (2006).

14. J. A. Morrone and R. Car, Phys. Rev. Lett. **101**, 017801 (2008).

15. S. Habershon, T. E. Markland and D. E. Manolopoulos, J. Chem. Phys. **131**, 024501 (2009).

16. C. Vega, M. M. Conde, C. McBride, J. L. F. Abascal, E. G. Noya, R. Ramirez and L. M.

Sesé, J. Chem. Phys. **132**, 046101 (2010).

17. B. Pamuk, J. M. Soler, R. Ramírez, C. P. Herrero, P. W. Stephens, P. B. Allen and M. V. Fernández-Serra, Phys. Rev. Lett. **108**, 193003 (2012).

18. L. Wang, M. Ceriotti and T. E. Markland, J. Chem. Phys. **141**, 104502 (2014).

19. Y. Litman, D. Donadio, M. Ceriotti and M. Rossi, J. Chem. Phys. **148**, 102320 (2018).

20. R. P. Feynman, *Quantum Mechanics and Path Integrals*. (McGraw-Hill, New York, 1965).

21. R. P. Feynman, *Statistical mechanics: A Set of Lectures* (W.A. Benjamin, Reading, Mass, 1972).

22. D. Chandler and P. G. Wolynes, J. Chem. Phys. **74**, 4078-4095 (1981).

23. a. B J Berne and D. Thirumalai, Annu. Rev. Phys. Chem. **37**, 401-424 (1986).

24. M. Parrinello and A. Rahman, J. Chem. Phys. **80**, 860-867 (1984).

25. M. E. Tuckerman, B. J. Berne, G. J. Martyna and M. L. Klein, J. Chem. Phys. **99**, 2796-2808 (1993).

26. D. Marx and M. Parrinello, J. Chem. Phys. **104**, 4077-4082 (1996).

27. M. E. Tuckerman, D. Marx, M. L. Klein and M. Parrinello, J. Chem. Phys. **104**, 5579-5588 (1996).

28. O. Marsalek and T. E. Markland, J. Chem. Phys. **144**, 054112 (2016).

29. M. F. Herman, E. J. Bruskin and B. J. Berne, J. Chem. Phys. **76**, 5150-5155 (1982).

30. D. Thirumalai, R. W. Hall and B. J. Berne, J. Chem. Phys. **81**, 2523-2527 (1984).

31. M. Diraison, G. J. Martyna and M. E. Tuckerman, J. Chem. Phys. **111**, 1096-1103 (1999).

32. P. Sindzingre, M. L. Klein and D. M. Ceperley, Phys. Rev. Lett. **63**, 1601-1604 (1989).

33. D. Scharf, M. L. Klein and G. J. Martyna, J. Chem. Phys. **97**, 3590-3599 (1992).

34. M. Sprik, R. W. Impey and M. L. Klein, J. Chem. Phys. **83**, 5802-5809 (1985).

35. D. F. Coker, B. J. Berne and D. Thirumalai, J. Chem. Phys. **86**, 5689-5702 (1987).

36. A. Wallqvist, D. Thirumalai and B. J. Berne, J. Chem. Phys. **86**, 6404-6418 (1987).

37. J. Cao and G. A. Voth, J. Chem. Phys. **100**, 5093-5105 (1994).

38. J. Cao and G. A. Voth, J. Chem. Phys. **100**, 5106-5117 (1994).

39. J. Cao and G. A. Voth, J. Chem. Phys. **101**, 6157-6167 (1994).

40. J. Cao and G. A. Voth, J. Chem. Phys. **101**, 6168-6183 (1994).

41. J. Cao and G. A. Voth, J. Chem. Phys. **101**, 6184-6192 (1994).

42. S. Jang and G. A. Voth, J. Chem. Phys. **111**, 2371-2384 (1999).

43. S. Jang and G. A. Voth, J. Chem. Phys. **111**, 2357-2370 (1999).

44. I. R. Craig and D. E. Manolopoulos, J. Chem. Phys. **121**, 3368-3373 (2004).

45. B. J. Braams and D. E. Manolopoulos, J. Chem. Phys. **125**, 124105 (2006).

46. J. O. Richardson and S. C. Althorpe, J. Chem. Phys. **131**, 214106 (2009).

47. S. Habershon, D. E. Manolopoulos, T. E. Markland and T. F. M. III, Annu. Rev. Phys. Chem. **64**, 387-413 (2013).

48. M. Rossi, M. Ceriotti and D. E. Manolopoulos, J. Chem. Phys. **140**, 234116 (2014).

49. K. Kinugawa, Chem. Phys. Lett. **292**, 454-460 (1998).

50. J. A. Poulsen and P. J. Rossky, J. Chem. Phys. **115**, 8024-8031 (2001).

51. A. Witt, S. D. Ivanov, M. Shiga, H. Forbert and D. Marx, J. Chem. Phys. **130**, 194510 (2009).

52. T. F. M. III and D. E. Manolopoulos, J. Chem. Phys. **122**, 184503 (2005).

53. T. F. M. III and D. E. Manolopoulos, J. Chem. Phys. **123**, 154504 (2005).

54. E. Geva, Q. Shi and G. A. Voth, J. Chem. Phys. **115**, 9209-9222 (2001).

55. M. G. Saunders and G. A. Voth, Annu. Rev. Biophys. **42**, 73-93 (2013).

56. W. G. Noid, J. Chem. Phys. **139**, 090901 (2013).

57. A. V. Sinitskiy and G. A. Voth, J. Chem. Phys. **143**, 094104 (2015).

58. N. Blinov and P. N. Roy, J. Chem. Phys. **115**, 7822-7831 (2001).

59. N. V. Blinov, P. N. Roy and G. A. Voth, J. Chem. Phys. **115**, 4484-4495 (2001).

60. P. N. Roy and N. V. Blinov, Isr. J. Chem. **42**, 183-190 (2002).

61. Y. Han, J. Jin, J. W. Wagner and G. A. Voth, J. Chem. Phys. **148**, 102335 (2018).

62. S. Izvekov and G. A. Voth, J. Chem. Phys. **123**, 134105 (2005).

63. S. Izvekov and G. A. Voth, J. Phys. Chem. B **109**, 2469-2473 (2005).

64. W. G. Noid, J. W. Chu, G. S. Ayton and G. A. Voth, J. Phys. Chem. B **111**, 4116-4127 (2007).

65. W. G. Noid, J. W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das and H. C. Andersen, J. Chem. Phys. **128**, 244114 (2008).

66. W. G. Noid, P. Liu, Y. Wang, J. W. Chu, G. S. Ayton, S. Izvekov, H. C. Andersen and G. A. Voth, J. Chem. Phys. **128**, 244115 (2008).

67. Y. Peng, Z. Cao, R. Zhou and G. A. Voth, J. Chem. Theory Comput. **10**, 3634-3640 (2014).

68. M. Nava, R. Quhe, F. Palazzesi, P. Tiwary and M. Parrinello, J. Chem. Theory Comput. **11**, 5114-5119 (2015).

69. G. Bussi and M. Parrinello, Phys. Rev. E. **75**, 056707 (2007).

70. I. F. Silvera and V. V. Goldman, J. Chem. Phys. **69**, 4209-4213 (1978).

71. S. Plimpton, J. Comput. Phys. **117**, 1-19 (1995).

72. M. Ceriotti, M. Parrinello, T. E. Markland and D. E. Manolopoulos, J. Chem. Phys. **133**, 124104 (2010).

Chapter 4: Coarse-graining of Many-body Path Integrals: Inclusion of Intramolecular Interactions and Bottom-up Force-matching

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 simulation of condensed matter systems using well-developed classical simulation techniques. The previous chapter 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 $2N$ pseudo-particles. In this chapter, we present 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 well capture 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 CG-PI theory to incorporate force fields of realistic molecular systems.

I. 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-hydrogn,[11-15] hydrogen bonding,[16-18] and proton transport.[19-23]

An appealing method to incorporate 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 the quantum-classical isomorphism, which states that the quantum canonical partition function of a particle can be expressed as the classical equivalent but of a discrete representation of the imaginary time path, or a 'ring polymer' in the extended phase space.[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 widely used to capture the NQEs of 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)[40-44] and ring polymer MD (RPMD)[45-49] extend the PIMD method, originally aimed to only capture equilibrium quantities, to a dynamical regime and have been used to calculate the dynamical quantities of the system such as vibrational spectra,[50, 51] diffusion,[52, 53] and reaction rates.[54]

However, even with the fast sampling algorithms of classical simulations applied to path integral calculations,[55, 56] proper sampling of ring polymer systems are more computationally demanding than their classical counterparts due to the large number of beads in the ring polymer.

In particular, the simulation becomes more challenging for systems with pronounced NQEs because the size of the ring polymer needs to be sufficiently large for a numerically converged result. The extended phase space of the ring polymer presents a significant sampling challenge that needs to be overcome to incorporate NQEs with manageable computational costs. This suggests that it would be valuable to construct an alternative isomorphism with inherently smaller DOF than that of the ring polymers while maintaining the correct description of the quantum partition function and thermal density matrix.

The method of bottom-up coarse graining (CG)[57, 58] constructs the coarse-grained representation of the underlying fine-grained model (all-atom resolution for example) by systematically integrating the DOFs of the fine-grained models. The CG potential of mean force (PMF) are calculated so that the CG system correctly describes the underlying physics such as structural correlations of the mapped FG systems. 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[59, 60] 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, which is the center of mass of the intermediate beads in the ring polymer, the CG-PI theory shows that the quantum partition function of a particle is isomorphic to the classical partition function of two (a pair) pseudo-particles obeying classical mechanics.

The CG-PI theory is not the first theoretical formalism 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 quantity. The CMD formalism, initially for distinguishable particles, has further been extended to study

many-body systems obeying Fermi-Dirac and Bose-Einstein statistics by mapping the system into the phase space variables.[61-63] The quantum multiscale coarse-graining (qMS-CG)[64] extends the classical MS-CG theory[65-68] to quantum Boltzmann statistics by using the path integral formalism. Not only that, several enhanced sampling methods such as replica exchange or metadynamics that combine the CG method and path integral formalism has been developed as well.[69-71]

It should be noted that while both the CMD and the CG-PI theory use CG representations of the ring polymer, there is one small yet 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. Using the centroid as the central quantity significantly decreases the DOF but only enables us to calculate the diagonal elements of the thermal density matrix. However, the mapping used in CG-PI allows us to explicitly calculate the off-diagonal elements and hence expectation values of momenta 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 CG PMF, the previous rendition assumed that the potential is pairwise decomposable. Even though the pairwise assumption has extensively been used in various simulation methods such as MS-CG to describe the intermolecular interactions, it cannot correctly describe the force field of a more general molecular systems with explicit intramolecular terms or general many-body interactions. Moreover, the previous numerical implementation of the CG-PI theory named numerical CG-PI or n-CG-PI needed an extensive set of *a priori* PIMD calculations to construct a scaling relationship and determine the model

parameters. Even though the n-CG-PI method with parameters determined in this fashion well captured the structural correlations of a simple liquid system, this process is not only computationally demanding but also counters the purpose of using the CG representation of ring polymers.

In this paper, we overcome the shortcomings of the previous development of the CG-PI theory and the n-CG-PI method. In Sec. II, we focus on developing the extended formalism of CG-PI and n-CG-PI method. The exact expression of the CG PMF of the pseudo-particles 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 quadratic form of the coupling between different pseudo-particles is preserved and the many-body terms in the force field simply adds additional terms which are functions of the Hessian of the many-body interactions. 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 Sec. III, we revisit the two liquid systems studied in our previous work. The new n-CG-PI method is shown to well capture both the intra and inter ring polymer structural correlations. Finally, Sec. IV contains discussion, conclusion, and future work so that the CG-PI formalism can be fully extended to be used as a reliable simulation tool to study NQEs of condensed matter systems.

II. Theoretical Methods

A. Many-body CG-PI potential and effective Hamiltonian

Using Feynman's imaginary time path integral formalism of quantum statistical mechanics,[24, 25] the canonical thermal density matrix element for N distinguishable particles

obeying the quantum Boltzmann statistics in one-dimension with Hamiltonian operator $\widehat{H}$ can be written as

$$\langle \boldsymbol{q} | e^{-\beta \widehat{H}} | \boldsymbol{q}' \rangle = \lim_{P\to\infty} \left[\left[\prod_{i=1}^{N} \left(\frac{m_i P}{2\pi\hbar^2\beta} \right)^{P/2} \right] \times \int \left[\prod_{i=1}^{N} \prod_{j=2}^{P} dq_j^{(i)} \right] e^{-\beta V_{P,N}} \right] \tag{4.1}$$

for a Feynman many-body isomorphic potential $V_{P,N}$ whose form is

$$V_{P,N} = \sum_{k=1}^{P} \sum_{i=1}^{N} \left[\frac{m_i P}{2\hbar^2\beta^2} \left(q_k^{(i)} - q_{k+1}^{(i)} \right)^2 + \frac{1}{P} V\left(q_k^{(1)}, \dots, q_k^{(N)} \right) \right] \tag{4.2}$$

where $q_j^{(i)}$ is the position of the *j*th bead in the *i*th ring polymer, m_i is the mass of the *i*th particle, and $V(q_1, \dots, q_N)$ 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 *i*th imaginary path) and q_i' (end of the *i*th imaginary path) as their *i*th elements, respectively. The boundary condition satisfies $q_1^{(i)} = q_i$ and $q_{P+1}^{(i)} = q_i'$. Eqs. (4.1) and (4.2) represent the quantum-classical isomorphism in which the thermal density matrix element of *N* interacting particles are represented by using a set of *N* ring polymers that obey classical mechanics. There are *P* coordinates or beads in the ring polymer that experiences the nearest neighbor harmonic coupling with spring constant of $m_i P/(2\hbar^2\beta^2)$ and the physical potential scaled by 1/P. The equality in Eq. (4.1) is exact in the limit of $P \to \infty$.

The many-body CG-PI theory[60] introduces a CG coordinate Q to construct a coarse-grained expression of the thermal density matrix element $\langle \boldsymbol{q} | e^{-\beta \widehat{H}} | \boldsymbol{q}' \rangle$ with a much smaller number of DOF. The newly introduced coordinate Q_i is defined as

$$Q_i = \frac{1}{P-1} \left(q_2^{(i)} + \cdots + q_P^{(i)} \right), \tag{4.3}$$

which is the center of mass of the intermediate beads ($q_2^{(i)}$,…, $q_P^{(i)}$) in the ith ring polymer. Using the new coordinate introduced in Eq. (4.3), the thermal density matrix element can be rewritten as

$$\langle \boldsymbol{q}|e^{-\beta\hat{H}}|\boldsymbol{q}'\rangle = \left(\prod_{i=1}^{N}\frac{m_i}{\pi\hbar^2\beta}\right)\int d\boldsymbol{Q}\, e^{-\beta V_{CG,N}(\boldsymbol{q},\boldsymbol{Q},\boldsymbol{q}')} \tag{4.4}$$

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

$$V_{CG,N}(\boldsymbol{q},\boldsymbol{Q},\boldsymbol{q}') = -\frac{1}{\beta}\lim_{P\to\infty}\left[ln\left(Z_{CG,N}^{(P)}(\boldsymbol{q},\boldsymbol{Q},\boldsymbol{q}')\right)\right] \tag{4.5}$$

where

$$Z_{CG,N}^{(P)}(\boldsymbol{q},\boldsymbol{Q},\boldsymbol{q}') = \left[\prod_{i=1}^{N}\left(\frac{\pi\hbar^2\beta}{m_i}\right)\left(\frac{m_iP}{2\pi\hbar^2\beta}\right)^{P/2}\right]$$

$$\times\int\left[\prod_{i=1}^{N}\prod_{j=2}^{P}dq_j^{(i)}\right]e^{-\beta V_{P,N}}\left[\prod_{i=1}^{N}\delta\left(Q_i-\frac{1}{P-1}\left(q_2^{(i)}+\cdots+q_P^{(i)}\right)\right)\right]. \tag{4.6}$$

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

$$\langle\hat{A}\rangle = \frac{\int d\boldsymbol{P}_Q d\boldsymbol{Q} d\boldsymbol{p} d\overline{\boldsymbol{q}}\, exp(-\beta H_{eff,N})W_A(\boldsymbol{p},\overline{\boldsymbol{q}})}{\int d\boldsymbol{P}_Q d\boldsymbol{Q} d\boldsymbol{p} d\overline{\boldsymbol{q}}\, exp(-\beta H_{eff,N})} \tag{4.7}$$

for N dimensional column position vector $\overline{\boldsymbol{q}} = (\boldsymbol{q}+\boldsymbol{q}')/2$, 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 $\overline{q}_i$ and Q_i, respectively. The many body Weyl map $W_A(\boldsymbol{p},\overline{\boldsymbol{q}})$ is defined as

$$\langle\boldsymbol{q}'|\hat{A}|\boldsymbol{q}\rangle = \int(d\boldsymbol{p}/2\pi\hbar)\, exp(i(\boldsymbol{p}\cdot\Delta\boldsymbol{q})/\hbar)\, W_A(\boldsymbol{p},\overline{\boldsymbol{q}}). \tag{4.8}$$

Note that the observable is taken with respect to the $\boldsymbol{p}$ and $\overline{\boldsymbol{q}}$ only. The exact expression for the effective Hamiltonian that satisfies Eq. (4.7) can be written as

$$H_{eff,N}\left(\boldsymbol{P}_Q, \boldsymbol{p}, \boldsymbol{Q}, \overline{\boldsymbol{q}}\right)$$

$$= \frac{1}{2}\boldsymbol{P}_Q{}^T \boldsymbol{M}_Q^{-1}\boldsymbol{P}_Q - \frac{1}{\beta}\ln\left[const \cdot \int d\boldsymbol{\Delta q}\, exp\left(i(\boldsymbol{p}\cdot\Delta\boldsymbol{q})/\hbar - \beta V_{CG,N}(\overline{\boldsymbol{q}}, \boldsymbol{Q}, \Delta\boldsymbol{q})\right)\right] \quad (4.9)$$

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 the N dimensional column vector $\boldsymbol{\Delta q} = \boldsymbol{q}' - \boldsymbol{q}$.

The expression for the many-body CG potential shown in Eq. (4.5) provides a reductionist representation of N interacting ring polymers. By averaging out the intermediate bead positions using the mapping shown in Eq. (4.3), the off-diagonal thermal density matrix element can be expressed in a compact manner as shown in Eq. (4.4). Moreover, only the intermediate beads are averaged and not the first and the last beads, q_i and q_i', which enables us 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. (4.9), 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 which have distinct physical meaning. More specifically, the pseudo-particle with position Q_i and momenta $P_{Q,i}$ are named centroid pseudo-particles (cent) and the pseudo-particle with position $\overline{q}_i$ and momenta p_i are named observable pseudo-particles (obsv). The integral with respect to the vector $\boldsymbol{\Delta q}$ in Eq. (4.9) indicates that the interactions between the observable and the centroid pseudo-particles in the effective Hamiltonian are determined by the many-body CG potential $V_{CG,N}$.

B. Second order Taylor expanded expressions

Despite the simplicity of the expressions shown in Eq. (4.5) compared to that of the Feynman isomorphic potential from Eq. (4.2), the explicit functional form of $V_{CG,N}$ is nontrivial even for simple model systems. In our previous work, 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 up to provide an approximation for $V_{CG,N}$.

We give a similar formalism in this section. However, the new Taylor expanded result is different than Eq. (3.12). In the previous derivation, the form of the physical potential is assumed to be pairwise decomposable. Even though such assumption has been commonly used in development of simulation methods such as MS-CG to describe the intermolecular interactions, it lacks the specific functional form to describe various intramolecular interactions commonly included in the molecular force field. 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 force field. In this chapter, the form of the 1-dimensional many-body potential $V(q_1, \dots, q_N)$ is assumed to be

$$V(q_1, \dots, q_N) = \sum_{\langle i,j \rangle} V_{ij}(q_i - q_j) + \sum_z V_z(q_1, \dots, q_N) \tag{4.10}$$

where $V_{ij}(q_i - q_j)$ is the pair potential between the *i*th and *j*th particle as a function of distance between the two particles, $\langle i, j \rangle$ denotes a summation over all distinct pairs, $V_z(q_1, \dots, q_N)$ 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 Stillinger-Weber

potential,[72] and Z is used to index the many-body interactions. The form of potential shown in Eq. (4.10) can be used to describe any molecular systems with many-body interactions.

The Taylor expanded expression of $V_{CG,N}$ upto second order using the form of potential shown in Eq. (4.10) can be written as

$$V_{CG,N}(\overline{\boldsymbol{q}}, \boldsymbol{Q}, \Delta\boldsymbol{q}) = \left[V(\boldsymbol{Q}) + (\boldsymbol{Q} - \overline{\boldsymbol{q}})^T \frac{\boldsymbol{K}_{Q\bar{q}}}{2} (\boldsymbol{Q} - \overline{\boldsymbol{q}}) + \Delta\boldsymbol{q}^T \frac{\boldsymbol{K}_{\Delta q^2}}{2} \Delta\boldsymbol{q} + K_0\right]$$
$$\times \left\{1 + \mathcal{O}\left(\hbar^3 \sqrt{\frac{\beta^5}{m^3}} \{V_{ij}'''\}\right)\right\} \tag{4.11}$$

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. (4.11) is same as that of Eq. (3.12), the details of the coupling matrices and scalar differ. The functional forms of these quantities are shown in the Appendix 4A, and the derivation of Eq. (4.11) are shown in Appendix 4B. 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 $\overline{\boldsymbol{q}}$, $\boldsymbol{Q}$, and $\Delta\boldsymbol{q}$. The first term $V(\boldsymbol{Q})$ indicates that the centroid pseudo-particles experience the physical potential V as shown in Eq. (4.10). The second term $(\boldsymbol{Q} - \overline{\boldsymbol{q}})^T \frac{\boldsymbol{K}_{Q\bar{q}}}{2} (\boldsymbol{Q} - \overline{\boldsymbol{q}})$ states that the distance between the two pseudo-particles 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 ($\Delta\boldsymbol{q}$) 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 evaluated with respect to $\overline{\boldsymbol{q}}$.

Using the second order truncated Taylor expanded result (square bracket in Eq. (4.11)), the effective Hamiltonian can be written as

$$H_{eff,N}\left(\boldsymbol{P}_Q, \boldsymbol{p}, \boldsymbol{Q}, \overline{\boldsymbol{q}}\right)$$

$$\cong \frac{1}{2}\boldsymbol{P}_Q{}^T \boldsymbol{M}_Q^{-1}\boldsymbol{P}_Q + \frac{1}{2\beta\hbar^2}\boldsymbol{p}^T \boldsymbol{K}_{\Delta q^2}^{-1}\boldsymbol{p} + V(\boldsymbol{Q}) + \frac{1}{2}(\boldsymbol{Q}-\overline{\boldsymbol{q}})^T \boldsymbol{K}_{Q\overline{q}}(\boldsymbol{Q}-\overline{\boldsymbol{q}}). \quad (4.12)$$

Eq. (4.12) 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) pseudo-particles. In this picture, the centroid pseudo-particles experience the physical potential via the $V(\boldsymbol{Q})$ term, and the two pseudo-particles are coupled to each other in a quadratic fashion through the $\frac{1}{2}(\boldsymbol{Q}-\overline{\boldsymbol{q}})^T \boldsymbol{K}_{Q\overline{q}}(\boldsymbol{Q}-\overline{\boldsymbol{q}})$ term. Moreover, the centroid and the observable pseudo-particles have associated momenta $\boldsymbol{P}_Q$ and $\boldsymbol{p}$. Note that in Eq. (4.12), 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 pseudo-particle are nontrivially coupled since $\boldsymbol{K}_{\Delta q^2}^{-1}$ in general will not be a diagonal matrix. This tells that a mass cannot be straightforwardly assigned to the observable pseudo-particles due to the off-diagonal terms of the matrix $\boldsymbol{K}_{\Delta q^2}$ and its dependence on the observable pseudo-particle position $\overline{\boldsymbol{q}}$.

One can further approximate the masses associated with the observable pseudo-particles to be constant. Using such approximation, the effective Hamiltonian can be written as

$$H_{eff,N}\left(\boldsymbol{P}_Q, \boldsymbol{p}, \boldsymbol{Q}, \overline{\boldsymbol{q}}\right)$$

$$\cong \frac{1}{2}\boldsymbol{P}_Q{}^T \boldsymbol{M}_Q^{-1}\boldsymbol{P}_Q + \frac{1}{2}\boldsymbol{p}^T \boldsymbol{m}_{\bar{q}}^{-1}\boldsymbol{p} + V(\boldsymbol{Q}) + \frac{1}{2}(\boldsymbol{Q}-\overline{\boldsymbol{q}})^T \boldsymbol{K}_{Q\overline{q}}(\boldsymbol{Q}-\overline{\boldsymbol{q}}). \quad (4.13)$$

The N dimensional diagonal mass matrix $\boldsymbol{m}_{\overline{q}}$ has constant effective mass for the ith observable pseudo-particle, $m_{\bar{q},i}$, as its ith diagonal element. In this picture, the only nontrivial coupling between the two pseudo-particles comes from the quadratic coupling term $\frac{1}{2}(\boldsymbol{Q}-\overline{\boldsymbol{q}})^T \boldsymbol{K}_{Q\overline{q}}(\boldsymbol{Q}-\overline{\boldsymbol{q}})$.

C. Challenges in numerical implementation

Despite the insight that the Taylor expanded expressions (Eqs. (4.11) – (4.13)) 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 for new observable pseudo-particle positions. In the next section, we will discuss modeling strategies that is based off of the functional terms shown in the Taylor expanded expression of the many-body CG potential and effective Hamiltonian.

D. Significance and limitation of previous n-CG-PI method

In our previous work, we have presented the numerical CG-PI (n-CG-PI) method as a way of incorporating the 2*N* pseudo-particle effective Hamiltonian representation into simulations of realistic molecular liquid systems.[60] Throughout this paper, the previously developed n-CG-PI modeling scheme discussed in Chapter 3 of this dissertation will be referred to as the "naïve n-CG-PI." The naïve n-CG-PI effective Hamiltonian can be written as

$$H_{n-CG-PI}\left(\boldsymbol{P}_Q, \boldsymbol{p}, \boldsymbol{Q}, \overline{\boldsymbol{q}}\right)$$
$$= \frac{1}{2}{\boldsymbol{P}_Q}^T \boldsymbol{M}_Q^{-1} \boldsymbol{P}_Q + \frac{1}{2}\boldsymbol{p}^T \boldsymbol{m}_{\overline{q}}^{-1} \boldsymbol{p} + V(\boldsymbol{Q}) + \frac{1}{2}(\boldsymbol{Q} - \overline{\boldsymbol{q}})^T \boldsymbol{K}_\lambda (\boldsymbol{Q} - \overline{\boldsymbol{q}}) \quad (4.14)$$

where for the *i*th pseudo-particle pair coupling parameter λ_i ($0 \leq \lambda_i$), the *i*th diagonal element for the *N* dimensional diagonal coupling matrix $\boldsymbol{K}_\lambda$ is $\lambda_i \cdot 12m_i/(\hbar^2\beta^2)$ where m_i is the physical mass of the *i*th particle in the system. The masses associated to the *i*th observable and centroid pseudo-particles were set to be m_i. One can arrive at Eq. (4.14) by making two key assumptions: a) the off-diagonal elements of $\boldsymbol{K}_{Q\overline{q}}$ are negligible and b) the diagonal elements of $\boldsymbol{K}_{Q\overline{q}}$ which originally are functions of $\overline{\boldsymbol{q}}$ can be approximated to be constants.

Despite a computationally tractable pseudo-particle representation that the naïve n-CG-PI method provides, it still had three shortcomings to reliably model NQE in realistic molecular

systems. a) It needed an extensive set of PIMD calculations at different temperatures and masses to correctly parametrize the coupling parameter λ which 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 pseudo-particle distributions and forces. c) Lastly, it lacked the formalism to construct a naïve n-CG-PI model for system 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 a single PIMD trajectory. However, before introducing the new modeling method, we give a brief overview of the two bottom-up CG simulation methods used in this work, Boltzmann inversion and force matching.

E. Boltzmann inversion in n-CG-PI

Boltzmann inversion is a bottom-up CG modeling method that parametrizes the bonded and nonbonded interactions in the CG resolution using the distribution calculated from mapped FG resolution.[73-76] In the new n-CG-PI model, only the interaction between the centroid and observable pseudo-particles within the same pair is treated with Boltzmann inversion.

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

$$V_{cent-obsv}\left(Q_i - \overline{q}_i\right) = -k_B T \ln\left[\rho\left(q_1^{(i)} - \frac{q_2^{(i)}+\cdots+q_P^{(i)}}{P-1}\right) \Big/ \left(q_1^{(i)} - \frac{q_2^{(i)}+\cdots+q_P^{(i)}}{P-1}\right)^2\right]. \quad (4.15)$$

The RHS of Eq. (4.15) is a function of fine-grained PIMD coordinates $q_j^{(i)}$ mapped into CG resolution (Eq. (4.3)) and the LHS is a function of coarse-grained (CG-PI) coordinates Q_i and $\overline{q}_i$.

The $1/x^2$ factor in the RHS accounts for the degeneracy of the positions. Note that the interaction term is a function of $Q_i - \overline{q}_i$ based on the Taylor expanded expression of the many-body CG potential shown in Eq. (4.11).

F. Force matching in n-CG-PI

Force matching is a CG method that was initially intended for CG representation of classical systems.[65-68] However, it can easily be generalized to parametrize the interactions in the n-CG-PI model. The results shown in this section is a many-body generalization of the equations shown in Appendix 3 and supplementary material S4.1 of Ref. 59. Note that the results shown in this section discuss the many-body CG potential for a closed ring polymer system ($q_i = q_i' = \overline{q}_i$).

The many-body CG-PI potential in Eq. (4.5) can alternatively be written as a variational equation

$$V_{CG,N}(\boldsymbol{Q}, \overline{\boldsymbol{q}}) = \underset{V_{CG,N}^{trial}}{\arg\min}\, \chi^2\left[V_{CG,N}^{trial}(\boldsymbol{Q}, \overline{\boldsymbol{q}})\right] + const \tag{4.16}$$

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

$$\chi^2\left[V_{CG,N}^{trial}(\boldsymbol{Q}, \overline{\boldsymbol{q}})\right] = \lim_{P\to\infty} \frac{\int\left[\prod_{i=1}^{N}\prod_{j=2}^{P} dq_j^{(i)}\right] \Sigma\, e^{-\beta V_{P,N}}}{\int\left[\prod_{i=1}^{N}\prod_{j=2}^{P} dq_j^{(i)}\right] e^{-\beta V_{P,N}}} \tag{4.17}$$

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

$$\Sigma\left(q_j^{(i)}; \begin{matrix} i = 1, \dots, N \\ j = 1, \dots, P \end{matrix}\right) = \sum_{i=1}^{N}\left[\left(\mathcal{F}_{\overline{q},i} - F_{q,i}\right)^2 + \left(\mathcal{F}_{Q,i} - F_{Q,i}\right)^2\right] \tag{4.18}$$

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

$$\mathcal{F}_{\overline{q},i} = -\frac{\partial V_{P,N}}{\partial q_1^{(i)}} \quad \mathcal{F}_{Q,i} = -\frac{1}{P-1}\sum_{j=2}^{P}\frac{\partial V_{P,N}}{\partial q_j^{(i)}} \tag{4.19a}$$

$$F_{\overline{q},i} = -\frac{\partial V_{CG,N}^{trial}}{\partial \overline{q}_i} \quad F_{Q,i} = -\frac{\partial V_{CG,N}^{trial}}{\partial Q_i}. \tag{4.19b}$$

The $\mathcal{F}$ terms are functions of FG PIMD coordinates $q_j^{(i)}$ the F terms are functions of CG (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.

G. Building the 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

$$V_{total} = V_{intra} + V_{inter} \tag{4.20}$$

for intramolecular potential V_{intra} and intermolecular potential V_{inter} which can further be specified by

$$V_{intra} = V_{bond} + V_{angle} + V_{dihedral} \tag{4.21a}$$

$$V_{inter} = V_{electro} + V_{vdW} \tag{4.21b}$$

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. (4.20), the proposed n-CG-PI effective Hamiltonian is written as

$$H_{n-CG-PI}\left(\boldsymbol{P}_Q, \boldsymbol{p}, \boldsymbol{Q}, \overline{\boldsymbol{q}}\right) = \frac{1}{2}{\boldsymbol{P}_Q}^T \boldsymbol{m}^{-1}\boldsymbol{P}_Q + \frac{1}{2}\boldsymbol{p}^T\boldsymbol{m}^{-1}\boldsymbol{p} + V_{intra}(\boldsymbol{Q}, \overline{\boldsymbol{q}}) + V_{inter}(\boldsymbol{Q}) \tag{4.22}$$

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

$$V_{intra}(\boldsymbol{Q}, \overline{\boldsymbol{q}}) = V_{intra,cent-obsv}(\boldsymbol{Q} - \overline{\boldsymbol{q}}) + V_{intra,cent}(\boldsymbol{Q}) \tag{4.23a}$$

$$V_{inter}(\boldsymbol{Q}) = \sum_{\langle i,j \rangle} V_{ij,cent}\left(Q_i - Q_j\right) \tag{4.23b}$$

for pair potential between the *i*th and the *j*th centroid pseudo-particle $V_{ij,cent}$.

According to Eqs. (4.22) and (4.23), there are three different types of interactions in the n-CG-PI model that need to be parametrized: $V_{intra,cent-obsv}(\boldsymbol{Q} - \overline{\boldsymbol{q}})$, $V_{intra,cent}(\boldsymbol{Q})$, and $V_{inter}(\boldsymbol{Q})$. The functional terms of the intramolecular interactions can be written as

$$V_{intra,cent-obsv}(\boldsymbol{Q} - \overline{\boldsymbol{q}}) = \sum_{i=1}^{N} K_i\left(Q_i - \overline{q}_i\right)^2 \tag{4.24a}$$

$$V_{intra,cent}(\boldsymbol{Q}) = V_{bond,cent}(\boldsymbol{Q}) + V_{angle,cent}(\boldsymbol{Q}) + V_{dihedral,cent}(\boldsymbol{Q}) \tag{4.24b}$$

where the centroid intramolecular potentials have the same functional form as that of the original force field shown in Eq. (4.21a).

In a way, the new n-CG-PI effective Hamiltonian shown in Eq. (4.22) closely resembles the naïve n-CG-PI effective Hamiltonian in Eq. (4.14). The $V_{intra,cent-obsv}(\boldsymbol{Q} - \overline{\boldsymbol{q}})$ has the same functional form as $\frac{1}{2}(\boldsymbol{Q} - \overline{\boldsymbol{q}})^T \boldsymbol{K}_\lambda (\boldsymbol{Q} - \overline{\boldsymbol{q}})$. Moreover, the centroid pseudo-particle only interactions ($V_{intra,cent}(\boldsymbol{Q}) + V_{inter}(\boldsymbol{Q})$) from Eqs. (4.24b) and (4.23b) are equivalent to the $V(\boldsymbol{Q})$ term in Eq. (4.14). This comes from the fact that we still use the same two approximations that was 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 pseudo-particles are set to be the physical mass of the *i*th 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 $V_{intra,cent-obsv}(\boldsymbol{Q} - \overline{\boldsymbol{q}})$ term and force matching to parametrize the centroid pseudo-particle only potentials ($V_{intra,cent}(\boldsymbol{Q})$ and $V_{inter}(\boldsymbol{Q})$). Note that there are only one pair potential between the *i*th and *j*th centroid pseudo-particle, which means that the V_{ij} contains

contribution from both electrostatic and van der Waals interactions between the *i*th and *j*th ring polymers.

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

a) For a particular set of 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 psdueo-particle position and forces using the mapping shown in Eq. (4.3) (force on the centroid pseudo-particle is calculated in a similar manner),

c) Parametrize $V_{intra,cent-obsv}(\boldsymbol{Q}-\overline{\boldsymbol{q}})$ from the mapped cent – obsv distance for each pair using Eq. (4.15),

d) Subtract the parametrized harmonic forces obtained in c) from the mapped PIMD trajectory forces,

e) Parametrize $V_{intra,cent}(\boldsymbol{Q})$ and $V_{inter}(\boldsymbol{Q})$ using the modified mapped PIMD trajectory from d) using force matching,

f) Build the CG-PI effective Hamiltonian using the parameters from c) and e), and

g) Sample the n-CG-PI effective Hamiltonian using classical molecular dynamics.

As a last note, other combinations of various CG methods were attempted (for example, parametrizing $V_{intra,cent}(\boldsymbol{Q})$ with BI or $V_{intra,cent-obsv}(\boldsymbol{Q}-\overline{\boldsymbol{q}})$ with force matching) but with little success. This aspect of the new n-CG-PI model will further be discussed in the conclusion section.

III. Results and Discussion

A. Liquid para-hydrogen

1. System description

The liquid para-hydrogen system is modeled using the semiempirical force field by Silvera and Goldman.[77] This form of potential treats the diatomic molecule as a point particle and describes the interaction between the particles using a pair potential. In Chapter 3, mass and temperature of the system were systematically varied to modulate the degree of NQE in the system and 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

$$V_{paraH}(q_1, \dots, q_N) = \sum_{\langle i,j \rangle} V_{ij}(q_i - q_j) \tag{4.25}$$

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

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

$$H_{n-CG-PI,naive}(\boldsymbol{P}_Q, \boldsymbol{p}, \boldsymbol{Q}, \overline{\boldsymbol{q}})$$
$$= \frac{1}{2}{\boldsymbol{P}_Q}^T \boldsymbol{m}^{-1} \boldsymbol{P}_Q + \frac{1}{2}\boldsymbol{p}^T \boldsymbol{m}^{-1} \boldsymbol{p} + K \sum_{i=1}^{N} (Q_i - \overline{q}_i)^2 + \sum_{\langle i,j \rangle} V_{ij}(Q_i - Q_j) \tag{4.26}$$

where *K* is the spring constant of centroid – observable bonded interaction and V_{ij} is the pair potential shown in Eq. (4.25) evaluated at centroid pseudo-particle 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 cases, the spring constant K was considered to be 'fit' if peak height and width of the first solvation shell of radial distribution function from naïve n-CG-PI simulations agreed to that of the PIMD simulation. Eq. (3.24) shows that the coupling strength, which can be understood to be inversely proportional to NQE, is proportional to mass and temperature squared. The coupling parameter K of the naïve n-CG-PI para-hydrogen model with mass of 5 g/mol and temperature of 30K calculated from this scaling relationship is 2.82 kcal/mol/Å^2. This scaling relationship is an intriguing result thar directly quantifies a degree of NQE as a function of mass and temperature for a simple liquid system.

Figure 3-3 shows the calculated radial distribution function and nonbonded angle distribution using the PIMD, classical MD, and naïve n-CG-PI method. Despite good agreement between the PIMD and the n-CG-PI results, there is a slight shift between the PIMD and n-CG-PI RDFs. The overall shape of the function quantitatively agree, but the n-CG-PI system predicts a smaller sphere of the particle compared to the PIMD system. Moreover, the results shown in Chapter 3 only show the physical quantities (correlation of observable pseudo-particles), but not the ones that involve centroid pseudo-particles, which encode the structural correlations of intra ring polymer fluctuations.

3. n-CG-PI model parameters

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

$$H_{n-CG-PI}\left(\boldsymbol{P}_Q, \boldsymbol{p}, \boldsymbol{Q}, \overline{\boldsymbol{q}}\right)$$

$$= \frac{1}{2}{\boldsymbol{P}_Q}^T \boldsymbol{m}^{-1} \boldsymbol{P}_Q + \frac{1}{2}\boldsymbol{p}^T \boldsymbol{m}^{-1} \boldsymbol{p} + K \sum_{i=1}^{N}\left(Q_i - \overline{q}_i\right)^2 + \sum_{\langle i,j \rangle} V_{ij,cent}\left(Q_i - Q_j\right) \qquad (4.27)$$

for the cent – obsv spring constant K and the centroid pseudo-particle pair potential $V_{ij,cent}(r)$. Note that due to the simplicity of the system, we only need to parametrize two

quantities: K and $V_{ij,cent}(r)$. We see that the functional form of the new n-CG-PI effective Hamiltonian is analogous to that shown in Eq. (4.26). This similarity suggests that the modified para-hydrogen liquid system is an ideal test case to study the differences between the naïve n-CG-PI and the new n-CG-PI methods. Figure 4-1 shows a single n-CG-PI para-hydrogen molecule, composed of a pair of pseudo-particles.

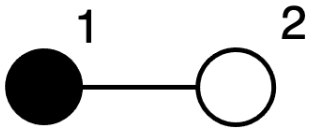

Figure 4-1. A schematic representation of the n-CG-PI para-hydrogen model. A ring polymer is coarse-grained into two pseudo-particles: the observable pseudo-particle (black and numbered 1) and the centroid pseudo-particle (white and numbered 2).

To build the n-CG-PI water model, a 2.0 ns PIMD trajectory with bead number of 32 was used. We use the 'rerun' command in LAMMPS molecular dynamics engine[78] to subtract the harmonic interaction parametrized from Boltzmann inversion. The centroid pseudo-particle interactions were parametrized by using the in-house force matching code.[67,68]

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 RHS of Eq. (4.15) 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 within the same pair.

Figure 4-2 shows two pair potentials: the original pair potential $V_{ij}(r)$ from Eq. (4.25) and the force matched centroid pseudo-particle pair potential $V_{ij,cent}(r)$ from Eq. (4.27) obtained by using force matching. The two functions quantitatively agree for $r \geq 4\text{Å}$. For $r < 4\text{Å}$,

however, we see that the force matched pair potential is slightly shifted to the right by approximately 0.1Å compared to the original potential. The force matched pair therefore predicts a larger radius of the centroid pseudo-particle. The combined effect of weaker harmonic interaction between the pseudo-particles and the shifted pair potential is manifested in the calculated radial distribution functions. This aspect of the new n-CG-PI model will further be discussed in latter sections.

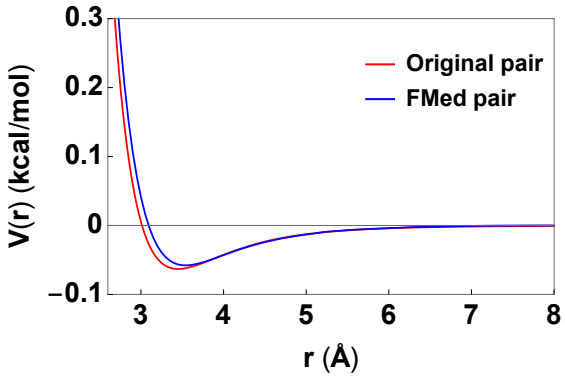

Figure 4-2. The original para-hydrogen pair potential (red) and the force matched centroid pseudo-particle pair potential (blue).

4. Simulation details

The modified para-hydrogen system is a canonical ensemble of 180 particles in a cubic box of length 19.71 Å. To modify the NQE of the system, the mass of the particle was set to be 5 g/mol and the temperature 30K. 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 in Ref. 77 was used, and for the n-CG-PI simulation, the model described in the previous section was used. For all three methods, timestep 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 pseudo-particles were thermostatted equally).[79] An in-house modified LAMMPS molecular dynamics engine using the Path Integral Langevin Equation (PILE) integrator[80] was used to perform the PIMD simulation. Unmodified version of LAMMPS molecular dynamics engine was used for n-CG-PI and classical simulations.

Additionally, the thermostatting issue discussed in the previous chapter has been resolved. In the naïve n-CG-PI model, the observable pseudo-particles had to be more heavily thermostatted compared to the centroid pseudo-particles. This issue arose from the fact that the observable pseudo-particle infromation (position, velocity, and force) was 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 pseudo-particles are considered as actual atoms during the simulation and no longer suffer from the thermostatting issue.

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 distribution ρ, radial distribution function g(r), and three-body nonbonded angle distribution P(θ). 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

$$P(\theta) = \sum_i^N \sum_{j>i}^N \sum_{k\neq i,j}^N \delta\left(\theta_{ijk} - \theta\right) \tag{4.28}$$

for angle between particle 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 pseudo-particles are compared to those from classical MD simulations because the expectation values of operators are calculated with respect to the observable pseudo-

particle coordinates as shown in Eq. (4.7). However, the structural correlations that involve the centroid pseudo-particle coordinates are included in this section for completion.

Figure 4-3 shows the normalized bond distribution between the centroid and the observable pseudo-particles within the same pair. The PIMD distribution shown in Figure 4-3 corresponds to the ρ shown in Eq. (4.15). 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 harmonic interaction between the pseudo-particles 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 in later sections.

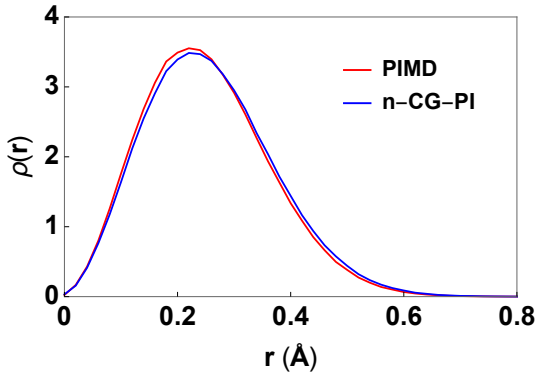

Figure 4-3. Calculated intramolecular distribution for the modified liquid para-hydrogen model between the centroid and the observable pseudo-particle from the PIMD simulation (red) and from the n-CG-PI simulation (blue).

The radial distribution functions from different methodologies are shown in Figure 4-4. A stark difference between the classical and the PIMD RDF shown in Figure 4-4(a) hints at a significant degree of NQE in the system. The classical RDF shows a more structured and

localized 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 to 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 4-4(b) and 4-4(c) plot the same function but at different values of distance. Moreover, Figure 4-4(c) is equivalent to Figure 4-3, 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 4-4(b) and 4-4(d). Even though the structural correlations that involve centroid pseudo-particles does not correspond to any physical quantity, good agreement for the centroid pseudo-particle distributions suggest that the n-CG-PI para-hydrogen model well captures the overall correlations and interactions of the ring polymers in the reference PIMD system.

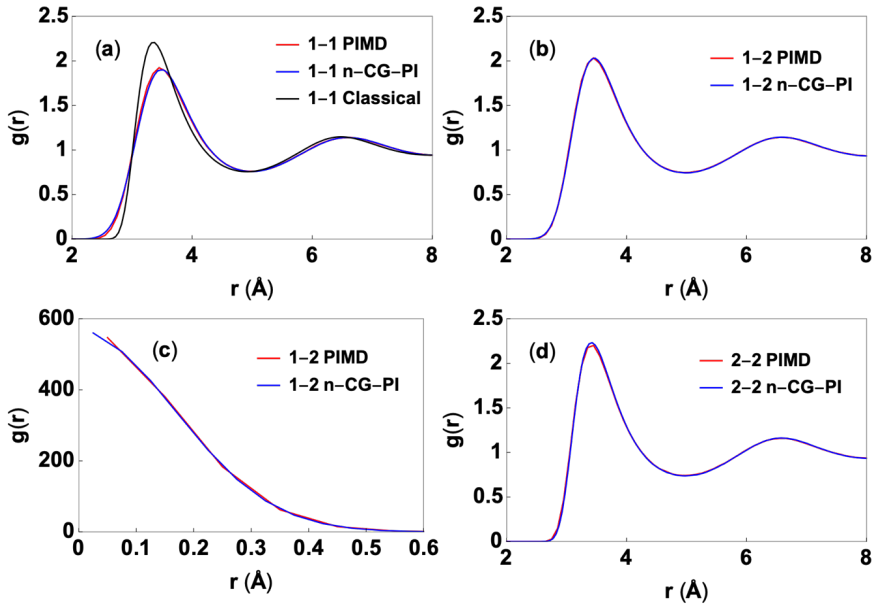

Figure 4-4. Calculated radial distribution function for (a) obsv – obsv (1 – 1), (b) obsv – cent (1 – 2) for $2\text{Å} \leq r < 8\text{Å}$, (c) obsv – cent (1 – 2) for $r < 2\text{Å}$, and (d) cent – cent (2 – 2). The PIMD results are shown in red, the n-CG-PI results are shown in blue, and the classical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

The nonbonded angle distribution captures the three-body interactions in the liquid system. Figure 4-5 shows the nonbonded angle distributions calculated from PIMD, n-CG-PI, and classical MD simulations. Similar as in Figure 4-4, 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 4-4 as the original force field only contains two-body interactions.

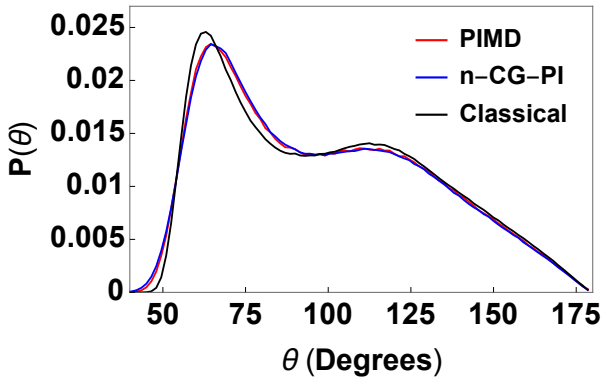

Figure 4-5. Calculated nonbonded angle distribution of observable pseudo-particles (PIMD and n-CG-PI) or particles (classical) with a cutoff of 3.5 Å. The PIMD results are shown in red, the n-CG-PI results are shown in blue, and the classical results are shown in black.

B. Liquid water

1. System description

In this work, we use the qSPC-Fw water model.[3] The form of its force field can be written as

$$V_{water}(q_1, \dots, q_N) = V_{intra} + V_{inter} \tag{4.29}$$

for the intramolecular term V_{intra} and the intermolecular term V_{inter} specified as

$$V_{intra} = \sum_{N_{mol}} k_b[(r_{OH1} - r_{OH1}^0)^2 + (r_{OH2} - r_{OH2}^0)^2] + k_a(\theta_{HOH} - \theta_{HOH}^0)^2 \tag{4.30a}$$

$$V_{inter} = \sum_{\langle i,j \rangle} \left[4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{r_{ij}} \right] \tag{4.30b}$$

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 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_{OH1}^0, r_{OH2}^0, and θ_{HOH}^0 are the equilibrium distances for the bonded interactions. The specific values of the intramolecular interaction parameters are: $k_b = 529.6 \frac{kcal}{mol\,Å^2}$, $r_{OH1}^0 = r_{OH2}^0 = 1.0$ Å, $k_a = 37.95 \frac{kcal}{mol\,rad^2}$, and $\theta_{HOH}^0 = 112°$. For the

H-H and O-H intermolecular pair potentials, $\varepsilon_{HH} = \varepsilon_{OH} = 0$ and $\sigma_{HH} = \sigma_{OH} = 0$. For the O-O pair, we have $\epsilon_{OO} = 0.1554\ kcal/mol$ and $\sigma_{OO} = 3.165$ 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. (4.29) and (4.30), unlike the para-hydrogen force field, contains explicit intramolecular interactions. This system is an opportunity to showcase how well the n-CG-PI modeling scheme captures NQEs of a more realistic and complicated molecular systems with nontrivial intramolecular structure.

2. Previous results

Analogous to the naïve n-CG-PI para-hydrogen model, the centroid pseudo-particles in the naïve n-CG-PI water model experiences the physical potential of the water force field shown in Eqs. (4.29) and (4.30).[60] The chosen coupling strength between the different pseudo-particles 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 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 3-4 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 are decent. 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 distance larger than 2Å, we see a shift of the naïve n-CG-PI intermolecular pair correlation compared to the PIMD one.

3. n-CG-PI model parameters

Figure 4-6 shows a schematic of the n-CG-PI water model. There are six pseudo-particles (written in red) and four types of pseudo-particles (written in black) in a single n-CG-PI water molecule. The four types of pseudo-particles: oxygen observable pseudo-particle (1), oxygen centroid pseudo-particle (2), hydrogen observable pseudo-particle (3), and hydrogen centroid pseudo-particle (4). The lines and curves in Figure 4-6 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 pseudo-particles in describing the force field and the results for notational simplicity.

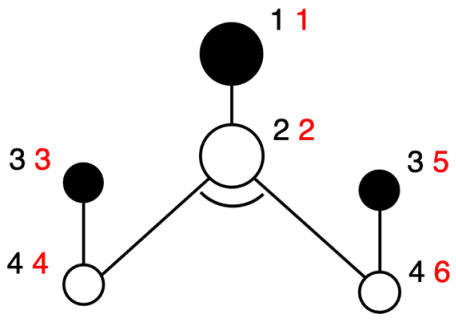

Figure 4-6. Schematic of the n-CG-PI water model. The black numbers represent the types of pseudo-particles: the oxygen observable pseudo-particle is numbered 1, the oxygen centroid pseudo-particle is numbered 2, the hydrogen observable pseudo-particle is numbered 3, and the hydrogen centroid pseudo-particle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

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

$$H_{n-CG-PI}(\boldsymbol{P}_Q, \boldsymbol{p}, \boldsymbol{Q}, \overline{\boldsymbol{q}}) = \frac{1}{2}\boldsymbol{P}_Q{}^T \boldsymbol{m}^{-1}\boldsymbol{P}_Q + \frac{1}{2}\boldsymbol{p}^T\boldsymbol{m}^{-1}\boldsymbol{p} + V_{intra,cent-obsv}(\boldsymbol{Q} - \overline{\boldsymbol{q}}) + V_{intra,cent}(\boldsymbol{Q}) + V_{inter,cent}(\boldsymbol{Q}) \quad (4.31)$$

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

$$V_{intra,cent-obsv}(\boldsymbol{Q} - \overline{\boldsymbol{q}}) = \sum_{N_{mol}} k_{12}(r_{12} - r_{12}^0)^2 + k_{34}[(r_{34} - r_{34}^0)^2 + (r_{56} - r_{34}^0)^2] \quad (4.32a)$$

$$V_{inter,cent}(\boldsymbol{Q}) = \sum_{N_{mol}} k_{24}[(r_{24} - r_{24}^0)^2 + (r_{26} - r_{24}^0)^2] + k_{424}(\theta_{426} - \theta_{424}^0)^2. \quad (4.32b)$$

The spring constants k_{12}, k_{34}, k_{24}, and k_{424} and equilibrium values r_{12}^0, r_{34}^0, r_{24}^0, and θ_{424}^0 parametrize the four intramolecular interactions according to their indices. The distances $r_{12}, r_{34}, r_{56}, r_{24}, r_{26}$ are intramolecular distances between the indexed pseudo-particles

and the angle θ_{426} is an angle between pseudo-particle 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

$$V_{inter,cent}(\boldsymbol{Q}) = \sum_{\langle i,j \rangle} V_{22}(r_{22}) + V_{24}(r_{24}) + V_{44}(r_{44}) \tag{4.33}$$

where V_{22}, V_{24}, and V_{44} are pair potentials between the indexed pseudo-particles which are functions of the pair distance. The three force-matched pair potentials are shown in Figure 4-7.

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

Table 4-1. Parametrized spring constants and equilibrium distances for the n-CG-PI water intramolecular interactions.

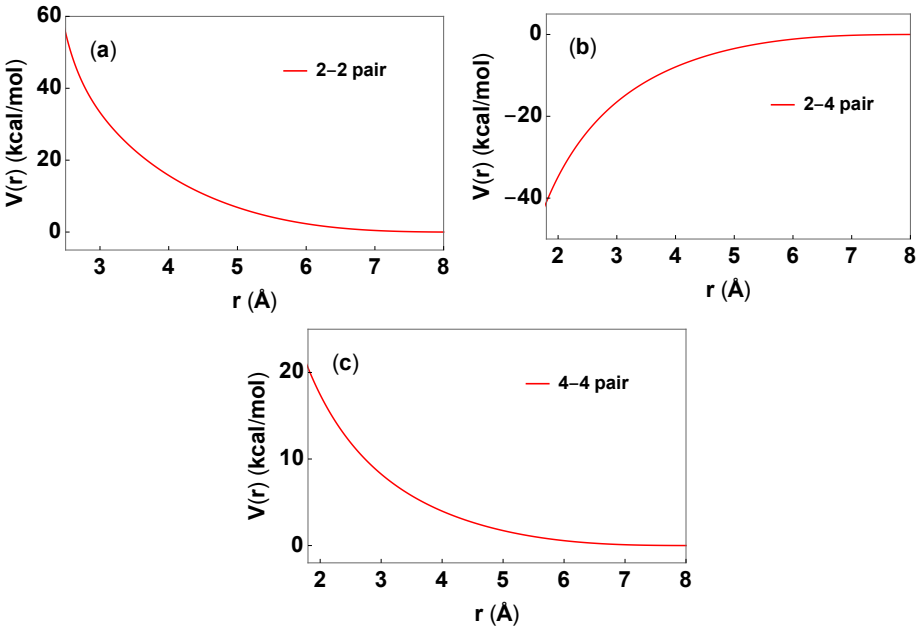

Figure 4-7. Force-matched (a) 2-2 pair, (b) 2-4 pair, and (c) 4-4 pair potential.

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 difference can be explained by the fact that hydrogen has larger degrees of NQE than oxygen due to its small mass and hence a weak harmonic interaction to account for the large degree of delocalization. Note that the equilibrium distances of the two obsv – cent harmonic interactions (1-2 and 3-4) are zero, which agrees with the functional from Taylor expanded expression of the many-body CG potential shown in Eq. (4.11). Moreover, we see that the 1-2 and 3-4 bond spring constants are much smaller compared to their naïve n-CG-PI 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 show no trend.

Figure 4-7 shows the three force-matched pair potentials included in the n-CG-PI water model. We see that the pair interactions between the same pseudo-particles are all repulsive and the pair interaction between the different pseudo-particles is all attractive. This result is somewhat surprising considering 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 principled MD simulation as shown in Figure 1 of Ref. 8. Even though the qSPC/Fw (or SPC/Fw) water model and the water model used in Ref. 8 are parametrized in different fashion, the agreement in the trend supports the validity of the force matched pair potentials shown in Figure 4-7.

The force matched pair potentials 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 intermolecular interactions, more analogous to the form shown in Eq. (4.30b), 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 4-7 provides a good agreement with the reference PIMD results.

4. Simulation details

A canonical ensemble of 233 water molecules in a cubic box of length 19.03 Å was simulated at 300K 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 nonbonded B-spline order of 6 and bonded B-spline order of 4. For the PIMD and classical simulations, qSPC/Fw water model was used, and for the n-CG-PI simulation, the model described in the previous section was used. A timestep of 0.5 fs and Langevin thermostat with parameter 0.05 fs^{-1} was used for all simulations.[79] The PIMD simulations were propagated by an in-house modified LAMMPS MD engine[78] using PILE integrator.[80] The n-CG-PI and classical simulations were performed with LAMMPS MD engine without any modifications.

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 case, water n-CG-PI system has a more complicated intramolecular structures as well as pair correlations.

Figure 4-8 shows the 1-3 (O-H) bond and 3-1-3 (H-O-H) angle distribution calculated from PIMD, n-CG-PI, and 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 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 have a larger degree of NQE compared to the PIMD water. We see that this over-correction of NQE in the n-CG-PI models is a common trend observed in other structural correlations.

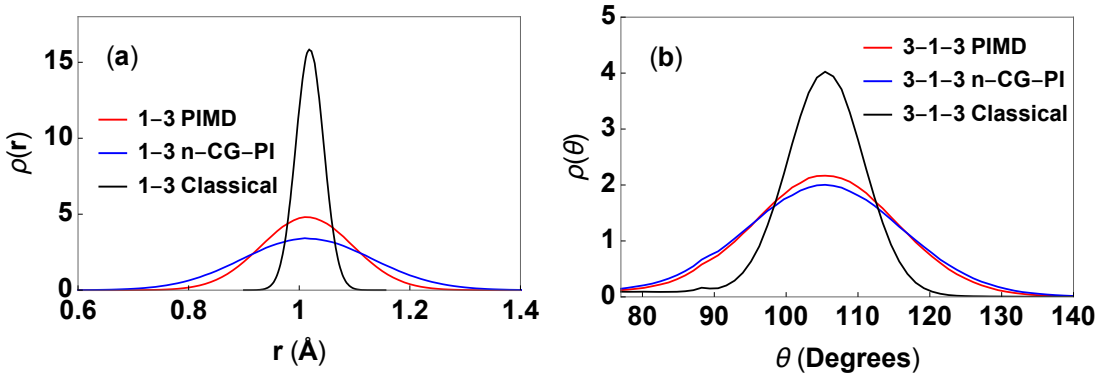

Figure 4-8. Calculated (a) 1-3 (O-H) bond and (b) 3-1-3 (H-O-H) angle distributions using different methods. The PIMD results are shown in red, the n-CG-PI results are shown in blue, and the classical MD results are shown in black.

We also investigate the intramolecular distributions that involve centroid pseudo-particle positions. These distributions are important because they are explicitly included in the n-CG-PI water model. Figure 4-9 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 4-9(a)) than for 3-4 (Figure 4-9(b)). 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 pseudo-particle 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 trajectory. Figure 4-9(c) and 4-9(d) show the 2-4 bond and 4-2-4 angle distributions, which were interactions included in the model via force matching. We see that the n-CG-PI 2-4 bond distribution has a higher peak height and width, which eludes to a more classical and less

delocalized 2-4 bond than that in the PIMD system. Lastly, the 4-2-4 angle distribution from the two methods quantitatively agree and overlap with each other.

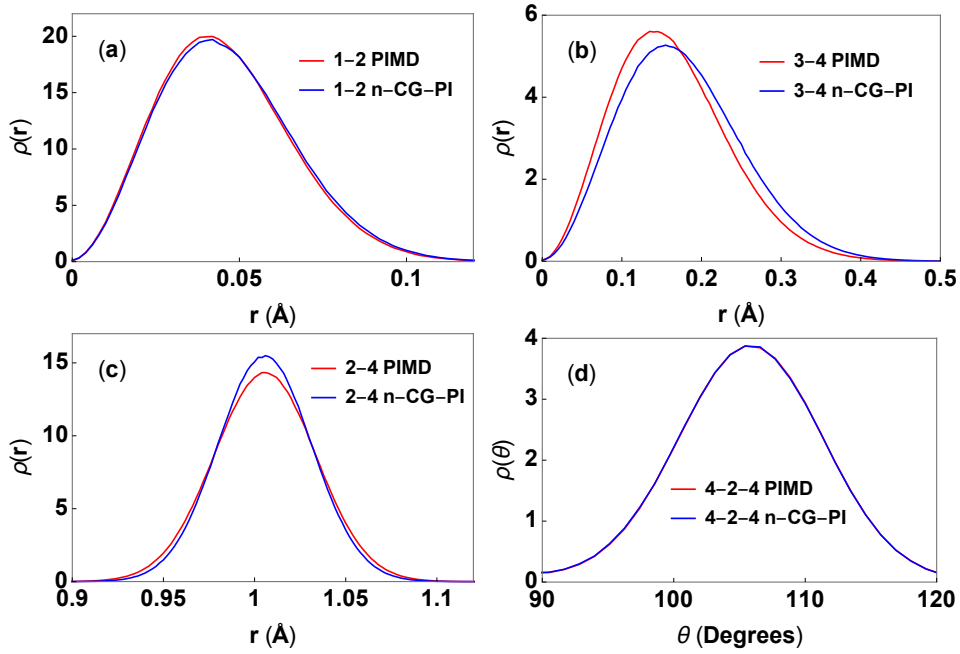

Figure 4-9. Calculated (a) 1-2 bond distribution, (b) 3-4 bond distribution, (c) 2-4 bond distribution, and (d) 4-2-4 angle distributions using the PIMD and the n-CG-PI methods. The PIMD results are shown in red and the n-CG-PI results are shown in blue.

Radial distribution functions capture the two-body correlations. By investigating the pair correlations, we are able to see how NQE plays an important role in determining the liquid structure. Figure 4-10 shows the three RDFs: 1-1 (O-O), 1-3 (O-H), and 3-3 (H-H). Note that these pair interactions are not explicitly included in 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. However, 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. The observable pair correlations n-CG-PI again predicts a more

delocalized and ‘quantum’ liquid structure compared to the PIMD results. However, the n-CG-PI RDFs still capture the general features of the PIMD RDFs, especially compared to the naïve n-CG-PI RDFs.

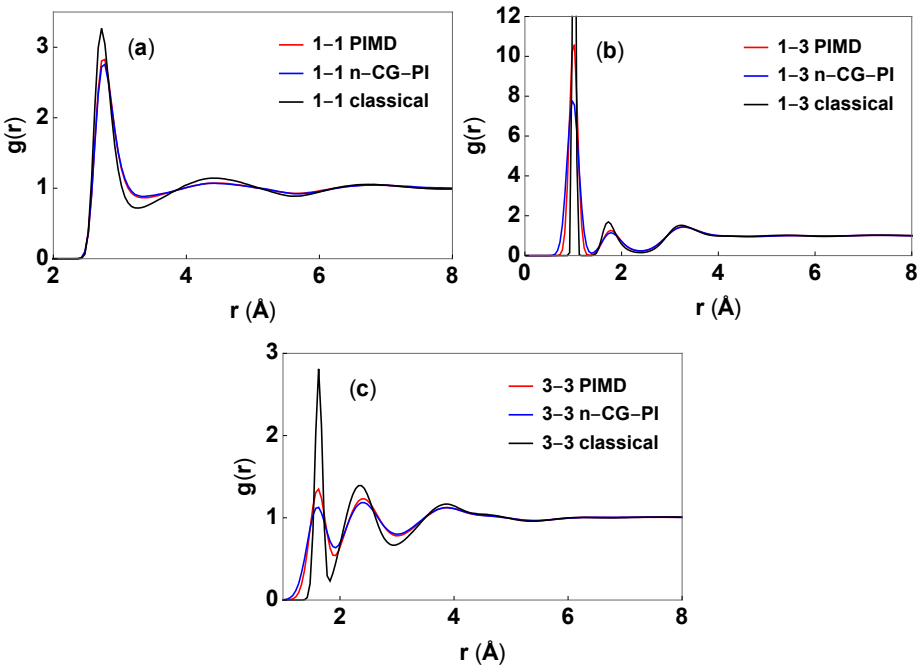

Figure 4-10. Calculated (a) 1-1 radial distribution function, (b) 1-3 radial distribution function, and (c) 3-3 radial distribution functions. The PIMD results are shown in red, the n-CG-PI results are shown in blue, and the classical MD results are shown in black.

We investigate the centroid pseudo-particle pair correlations as well. Figure 4-11 shows three RDFs: 2-2, 2-4, and 4-4 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. (4.33) and Figure 4-7. The agreement between the n-CG-PI and PIMD RDFs are much better in Figure 4-11 compared to what is shown in Figure 4-10. Moreover, we see good agreement even for the intramolecular O-H peaks (~ 1 Å in Figure 4-11(b)) and H-H peaks (~ 1.6 Å in Figure 4-11(c)). The better

agreement between the two sets of RDFs shown in Figure 4-11 can be explained by the fact the cent – cent pair interactions are explicitly included in the n-CG-PI water model.

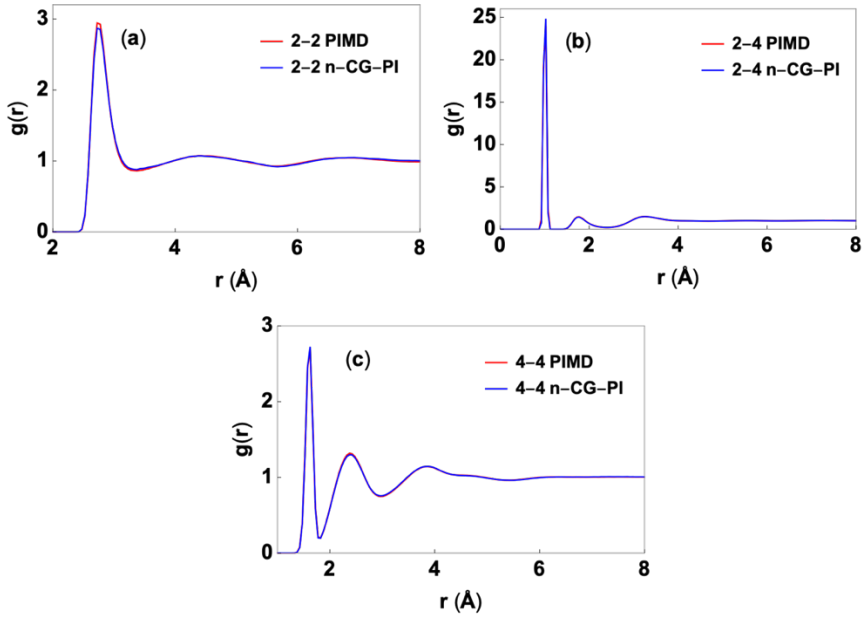

Figure 4-11. Calculated (a) 2-2 radial distribution function, (b) 2-4 radial distribution function, and (c) 4-4 radial distribution function. The PIMD results are shown in red and the n-CG-PI results are shown in blue.

Since there are 4 different types of pseudo-particles in the mapped PIMD and the n-CG-PI water systems, there are 10 possible pair combinations. Figure 4-12 shows pair correlations of cent – obsv pseudo-particles. For the 1-2 and 3-4 RDFs, there are nonzero peaks centered at 0Å, but they were not shown in this figure as it is redundant to Figure 4-9(a) and 4-9(b). Similar to Figures 4-10 and 4-11, 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 well capture the overall shape of the PIMD pair correlations.

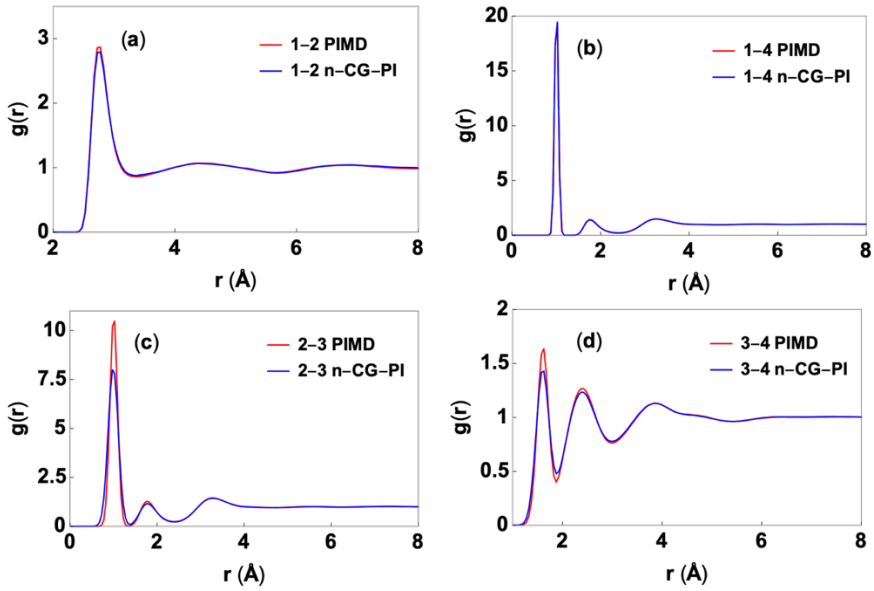

Figure 4-12. Calculated (a) 1-2 radial distribution function, (b) 1-4 radial distribution function, (c) 2-3 radial distribution function, and (d) 3-4 radial distribution function. The PIMD results are shown in red and the n-CG-PI results are shown in blue.

The nonbonded angle distribution $P(\theta)$ 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 correlations. Figure 4-13 shows the nonbonded angle distributions with respect to the oxygen observable pseudo-particle 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. Even though there are no explicit many-body intermolecular interactions in the original qSPC/Fw force field and any higher order correlations present in the system can be decomposed into combination of two-body pair interactions, this result suggests that the n-CG-PI can correctly captures higher order correlations by only force matching the pair intermolecular terms despite small disparities seen in Figure 4-10.

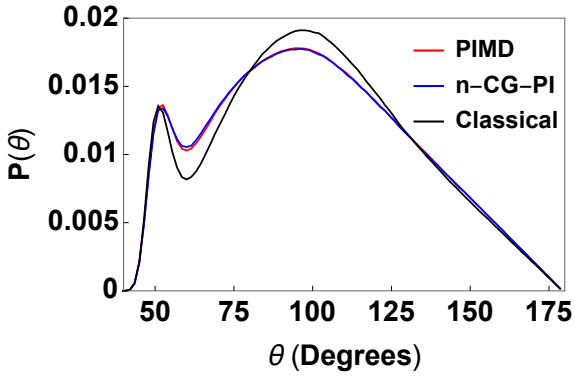

Figure 4-13. Calculated nonbonded angle distribution function for oxygen observable pseudo-particles (for PIMD and n-CG-PI) or oxygen atoms (classical) with cutoff distance of r_0 = 3.5 Å. The The PIMD results are shown in red, the n-CG-PI results are shown in blue, and the classical MD results are shown in black.

Overall, the n-CG-PI water model well captures the structural correlations of the reference PIMD simulations. Different from the previous para-hydrogen model, the water system is more complex due to the presence of more types of atoms and extra intra and intermolecular interactions. Different combinations of 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 to the PIMD results. Last thing to note is that while the authors were not able to find a better n-CG-PI modeling scheme that performs better than the current version, it is possible to do so. For example, using an iterative Boltzmann inversion (IBI)[76] to improve the agreement in intramolecular distributions could be the first step in further developing the current model. Further discussion on the n-CG-PI modeling strategy is contained in the next section.

IV. Conclusion

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

In this work, we present a Taylor expanded expression of the many-body CG potential that considers intramolecular interactions. Note that even though the many-body potential V_z in this work was aimed 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 pseudo-particles analogous to the pairwise decomposable terms. The new Taylor expanded result shown in the square bracket of Eq. (4.11) suggest that the N pseudo-particle pair isomorphism of the CG-PI theory can be generalized to any forms of force field and interaction between the particles. The new derivation of the CG potential, however, is still numerically challenging to implement directly and does not hint 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 pseudo-particles in the same pair are harmonically coupled to each other and the spring constant is parametrized

by λ. We have 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 it needed an extensive set of path integral calculations to build the scaling relationship for λ but it had no formalism to account for more than one types of atoms.

To overcome these issues, we have developed a new version of the n-CG-PI model. For a molecule described in the force field with n bonded atoms, the n-CG-PI molecule is composed of $2n$ pseudo-particles. Between the pseudo-particles, 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 pseudo-particles have the functional form used in 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 pre-determined scaling relationship using classical CG methods. More specifically, we use force matching to parametrize the cent – cent interactions (both intra and intermolecular) and Boltzmann inversion for the obsv – cent interactions.

With this new modeling scheme, we revisit the two liquid systems studied in Chapter 3: modified liquid para-hydrogen and liquid water. We have 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 well captures the general shapes of the PIMD distributions, quantitatively for some cases. The disparity between the two systems can be explained by a more complex intramolecular structure and larger degrees of NQE for liquid water. 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. Not only that, the new n-CG-PI method gives a guideline on how to construct the n-CG-PI model given a certain type of molecular all-atom force field.

As shown in Figures 4-3, 4-8, and 4-9, there are three types of intramolecular distributions calculated: obsv – cent, cent – cent, and obsv – obsv. The obsv – obsv intramolecular interactions are not explicitly included in the n-CG-PI model. We see that the n-CG-PI cent – cent intramolecular distributions (Figures 4-9(c) and 4-9(d)) are more localized than the PIMD equivalents. The n-CG-PI obsv – cent intramolecular distributions (Figures 4-3, 4-9(a), and 4-9(b)), on the other hand, are more delocalized than the PIMD results. Lastly, the obsv – obsv n-CG-PI distributions (Figure 4-8) are also more delocalized than the PIMD ones, to a larger extent than the obsv – cent distributions.

Moreover, another trend can be identified from analyzing a set of trends by looking at the calculated RDFs from the water model. In the n-CG-PI water model, there are three types of cent – cent nonbonded pair potentials: 2-2, 2-4, and 4-4. Among the three sets of RDFs (obsv – obsv as shown in Figure 4-10, cent – cent as shown in Figure 4-11, and obsv – cent as shown in Figure 4-12), we see that only the cent – cent n-CG-PI RDFs give quantitative agreement with the PIMD RDFs. Even though the other n-CG-PI RDFs qualitatively agree with the PIMD results and well predict the relative shape of the liquid structure, they all are more delocalized and give more ‘quantum’ liquid structure. This difference comes from the fact that the nonbonded pair correlations of observable pseudo-particles is determined by cent – cent pair interactions and obsv – cent bonded interactions rather than directly by obsv – obsv pair interactions.

Shortcomings of the current n-CG-PI model suggests that it may be improved upon by including additional intra or intermolecular interactions in the n-CG-PI model. We have made different models by varying three main aspects but had little success other than the presented result.

a) Different combinations of CG methods were attempted. For example, we have tried using only force-matching or only Boltzmann inversion to parametrize the two types of intramolecular interactions. However, the force-matching only models suggested noticeably weaker intramolecular interactions and hence more delocalized intramolecular distributions. The Boltzmann inversion only models showed the opposite trend in which the intramolecular distributions were more localized than the PIMD ones. This difference not only implies the different scopes in which the two CG methods should be used but also suggests that the two types of intramolecular interactions (obsv – cent and cent – cent) are discernable from one another and should be treated with different methodologies.

b) Adding different combinations of intramolecular interactions were tried. For instance, since the *j*th bead of one ring polymer experiences nonboned pair interactions from the *j*th bead of another ring polymer as shown in Eq. (4.2), there may be a noticeable intramolecular interaction between the observable pseudo-particles that the model needs to incorporate to correctly capture the structural correlations. To address this, we have tried adding the 1-3 bond and/or 3-1-3 angle interactions for the water model. However, these water models predicted a more rigid and localized intramolecular structures. This result further solidifies the understanding of observable pseudo-particle as a representation of delocalization in quantum statistical mechanics and how it only should harmonically interact with classically behaving centroid pseudo-particles and not with other observable pseudo-particles.

c) Lastly, different combinations of intermolecular interactions were attempted. In a similar reasoning for adding different types of intramolecular interactions, one can assume that there can be nonbonded pair interactions between the observable pseudo-particles. Note that since only the same indexed beads in the ring polymers interact with each other, it is safe to assume that the interaction between the *i*th observable pseudo-particle and the *j*th centroid pseudo-particle is negligible. To test this idea, we have added 1-1 pair term in the n-CG-PI para-hydrogen model and 1-1, 1-3, 3-3 pair terms in the n-CG-PI water model. These models not only predicted overstructured RDFs for both liquid systems but also did not show diffusivity and resulted in anomalous liquid structures. This suggests that any intermolecular interactions present in the PIMD system is and should be accounted into cent – cent pair interactions only.

Two strategies can improve upon the current n-CG-PI model as shown in Eq. (4.22). a) The n-CG-PI model can benefit from using different types of CG methods. As described before, the current modeling scheme has been determined empirically, which suggests that there can be better modeling strategies to be used in the n-CG-PI model. For example, the use of iterative Boltzmann inversion (IBI)[76] for the obsv – cent intramolecular interaction may improve upon the agreement between the n-CG-PI and PIMD intramolecular distributions. This will eventually improve the obsv – obsv intra and intermolecular correlations, which are important quantities that correspond to expectation values of operators. b) The shortcomings observed in the n-CG-PI water model may arise from the functional form of the n-CG-PI model showed in Eq. (4.22). Similar to the naïve n-CG-PI method, the coupling matrix between the observable and the centroid pseudo-particles is still approximated to be diagonal and constant. Even though this approximation worked for simple systems with little intramolecular structure such as para-hydrogen, it may need to be generalized and include off-diagonal coupling terms for systems

with nontrivial intramolecular structures. Alternatively, for systems with higher degrees of NQE, the higher order terms in the Taylor expanded expression of many-body CG potential can be explicitly included in the n-CG-PI model. For one-body system, the third Taylor expanded terms are functions of $(Q-\bar{q})^3$, $\Delta q^2(Q-\bar{q})^2$, and $(Q-\bar{q})$ as shown in Eq. (A15) of Ref. 59.

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, four 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 out the n-CG-PI modeling scheme on a more complex molecular systems and study their structural correlations. Moreover, as shown in Eq. (4.24), 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 non-contracted ring polymers.[56] Such a contraction scheme can accelerate the path integral simulation significantly while 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. (4-23(b)), there is only one nonbonded pair interactions

between the centroid pseudo-particles. 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 4-7, which correctly predicted the cent – cent pair correlations of mapped PIMD system as shown in Figure 4-11. However, these pair potentials provide little physical intuition. To resolve this, the theoretical formalism of charge for different pseudo-particles can be worked out. By assigning appropriate charges to the pseudo-particles, 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 pseudo-particles. 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. (4.11) by using open ring PIMD simulations (start of the imaginary path is not identical to the end). By parametrizing the $\Delta \boldsymbol{q}$ 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) Lastly, a set of thermodynamic estimators in the context of CG-PI theory need to be derived. Chapter 3 contains an expression for the CG-PI energy estimator in Eq. (3.13). 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 well captures the structural correlations and delocalization of ring polymers observed in PIMD simulations. Overall, both the CG-PI theory and the n-CG-PI models presented in this work provide a stepping stone to systematically incorporate NQEs of condensed matter systems in a computationally efficient manner by only including minimal additional degrees of freedom.

Appendix 4A: Functional Details of the Coupling Terms in Taylor Expanded CG Potential

The coupling matrices and scalar from Eq. (4.11) can be written as

$$\boldsymbol{K}_{Q\bar{q}} = \frac{\mathbb{L}^{-1}}{\beta} - \frac{\boldsymbol{M}}{2} - \sum_Z D^2 V_z(\overline{\boldsymbol{q}}) \tag{4.A1a}$$

$$\boldsymbol{K}_{\Delta q^2} = \frac{\boldsymbol{M}}{24} + \frac{\boldsymbol{\Sigma}}{\beta} - \frac{\beta}{4}\boldsymbol{C} + \frac{1}{12}\sum_Z D^2 V_z(\overline{\boldsymbol{q}}) \tag{4.A1b}$$

$$K_0 = \frac{1}{2\beta} ln\left[\left(\prod_{i=1}^{N} \frac{4m_i}{\beta\hbar^2}\right) det(\boldsymbol{\Omega} + \boldsymbol{\Omega_2} + \boldsymbol{I}) det(\boldsymbol{L})\right]. \tag{4.A1c}$$

For two-body terms

$$\sigma_{ij} = \sqrt{\frac{\hbar^2\beta}{\sqrt{m_i m_j}}} \tag{4.A2a}$$

$$\sigma_i = \sqrt{\frac{\hbar^2\beta}{m_i}} \tag{4.A2b}$$

$$\gamma_{ij} = \left(\frac{m_j}{m_i}\right)^{\frac{1}{4}}, \tag{4.A2c}$$

we define two N dimensional square block matrices whose matrix elements are infinite dimensional diagonal matrix $\boldsymbol{\Omega}_1$ and $\boldsymbol{\Omega}_2$. The indices i and j are for block elements and the indices k and l are for within the element matrices. Their elements are written as

$$(\boldsymbol{\Omega}_1)_{ij,kl} = \begin{cases} \delta_{kl} \sum_{m\neq i}^{N} \left(\frac{\beta {\sigma_{im}}^2}{2(k\pi)^2} \overline{V_{im}''} \right) & (i = j) \\ +({\gamma_{im}}^2 - 1) & \\ -\delta_{kl} \frac{\beta {\sigma_{ij}}^2}{2(k\pi)^2} \overline{V_{ij}''} & (i \neq j) \end{cases} \tag{4.A3}$$

for

$$\overline{V_{ij}''} = V_{ij}''\left[\bar{q}_i - \bar{q}_j\right] + V_{ij}''\left[\bar{q}_j - \bar{q}_i\right] \tag{4.A4}$$

and

$$(\boldsymbol{\Omega}_2)_{ij,kl} = \sum_z \frac{2\beta}{\pi^2} \frac{{\sigma_{ij}}^2}{kl} \int_0^1 du \left(D^2 V_{z,ij}\big|_c \right) sin(k\pi u) sin(l\pi u) \tag{4.A5}$$

for $D^2 V_{z,ij}\big|_c$ which is the *i*th row *j*th column element of the Hessian matrix of V_z evaluated at the classical path (linear line between the beginning and the end of the path).

For given $\boldsymbol{\Omega}_1$ and $\boldsymbol{\Omega}_2$, and the identity matrix $\boldsymbol{I}$ of same dimensionality, we define a matrix $\boldsymbol{\Lambda}$ given by

$$\boldsymbol{\Lambda} = (\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2 + \boldsymbol{I})^{-1} \tag{4.A6}$$

with *i*th row *j*th column block element $\boldsymbol{\Lambda}_{ij}$. Moreover, for infinite dimensional column vectors $\boldsymbol{b}_\xi$ and $\boldsymbol{b}_\lambda$ with elements

$$\left(\boldsymbol{b}_\xi\right)_k = -\frac{2\sqrt{2}}{\pi^2} \frac{odd(k)}{k^2} \tag{4.A7a}$$

$$(\boldsymbol{b}_\lambda)_k = \frac{\sqrt{2}}{\pi^2} \frac{1 - odd(k)}{k^2} \tag{4.A7b}$$

$$odd(k) = \begin{cases} 1 \ (k = odd) \\ 0 \ (k = even) \end{cases} \tag{4.A7c}$$

we can define two N dimensional square matrices $\boldsymbol{L}$ and $\boldsymbol{R}$. Their *i*th row and *j*th column elements are specified by

$$(\boldsymbol{L})_{ij} = \sigma_i \sigma_j \left({\boldsymbol{b}_\xi}^T \boldsymbol{\Lambda}_{ij} \boldsymbol{b}_\xi \right) \tag{4.A8a}$$

$$(\boldsymbol{R})_{ij} = \sigma_i \sigma_j \left({\boldsymbol{b}_\lambda}^T \boldsymbol{\Lambda}_{ij} \boldsymbol{b}_\lambda \right). \tag{4.A8b}$$

The *i*th row and *j*th column elements for the *N* dimensional square matrices $\boldsymbol{\Sigma}$ and $\boldsymbol{M}$ are written as

$$(\boldsymbol{\Sigma})_{ij} = \frac{\delta_{ij}}{{\sigma_i}^2} \tag{4.A9a}$$

$$(\boldsymbol{M})_{ij} = \begin{cases} \sum_{m\neq i}^{N} \overline{V_{im}''} \; (for\; i = j) \\ -2V_{ij}''\left(\bar{q}_i - \bar{q}_j\right) \; (for\; i \neq j) \end{cases}. \tag{4.A9b}$$

The *N* dimensional square matrix $\sum_Z D^2 V_z(\overline{\boldsymbol{q}})$ is a Hessian matrix of V_z evaluated at $\overline{\boldsymbol{q}}$. Lastly, the *N* dimensional square matrix $\boldsymbol{C}$ satisfies the relation

$$\Delta\boldsymbol{q}^T\boldsymbol{C}\Delta\boldsymbol{q} = \left(\boldsymbol{G} + 2\mathcal{E}^{(2)}\right)^T \boldsymbol{R}\left(\boldsymbol{G} + 2\mathcal{E}^{(2)}\right) \tag{4.A10}$$

for *N* dimensional column vectors $\boldsymbol{G}$ and $\mathcal{E}^{(2)}$ with *i*th row element

$$(\boldsymbol{G})_i = \sum_{\substack{j=1 \\ j\neq i}}^{N} \left(\Delta q_{\boldsymbol{i}} - \Delta q_{\boldsymbol{j}}\right)\overline{V_{ij}''} \tag{4.A11a}$$

$$\left(\mathcal{E}^{(2)}\right)_i = \sum_Z D^2 V_{z,i}\big|_{\overline{\boldsymbol{q}}} \cdot \Delta\boldsymbol{q} \tag{4.A11b}$$

where $D^2 V_{z,i}\big|_{\overline{\boldsymbol{q}}}$ is the *i*th row of the Hessian matrix of V_z evaluated with respect to $\overline{\boldsymbol{q}}$.

Appendix 4B: Derivation of Taylor Expanded CG Potential

We present a derivation of Eq. (4.11). This appendix is a generalized version of derivation for Eq. (3.12) of Chapter 3. Throughout the current derivation, we will refer to the equations from Appendix 3B.

An alternative representation of the imaginary path using Fourier decomposition can be written as

$$q_i(u) = q_i(0) + \left(q_i(1) - q_i(0)\right)u + \frac{\sqrt{2}}{\pi}\sigma_i \sum_{k=1}^{\infty} a_{i,k} sin(k\pi u)/k \tag{4.B1}$$

in which $\sigma_i = \sqrt{\hbar^2\beta/m_i}$ where m_i is the mass associated with the *i*th particle or ring polymer, $a_{i,k}$ are the path parameter where *i* indexes the ring polymer number and *k* indexes the

Fourier decomposition. The variable u ranges from 0 to 1. $q_i(0)$ is the beginning of the path and $q_i(1)$ is the end of the path. For notational simplicity, we decompose Eq. (4.B1) into two parts: the classical ($q_{C,i}(u)$) and the quantum path ($q_{Q,i}(u)$) such that

$$q_{C,i}(u) = q_i(0) + \big(q_i(1) - q_i(0)\big)u \tag{4.B2a}$$

$$q_{Q,i}(u) = \frac{\sqrt{2}\sigma_i}{\pi}\sum_{n=1}^{\infty} a_{i,n}\sin(n\pi u)/n. \tag{4.B2b}$$

Note that the classical path is a linear line between the beginning and the end of the imaginary path while the quantum path represents a delocalization from the said classical path as an infinite sum of sine functions. Using Eq. (4.B1), the exact expression of the many-body CG potential $V_{CG,N}(\boldsymbol{q}, \boldsymbol{Q}, \boldsymbol{q}')$ can be written as

$$e^{-\beta V_{CG,N}(\boldsymbol{q},\boldsymbol{Q},\boldsymbol{q}')} = \Big(\prod_{i=1}^{N}\sigma_i/\sqrt{8\pi}\Big)exp\left(-\frac{1}{2}\Delta\boldsymbol{q}^T\boldsymbol{\Sigma}\Delta\boldsymbol{q}\right)$$

$$\times\int\frac{d\boldsymbol{\xi}d\boldsymbol{a}}{\sqrt{2\pi}}\begin{bmatrix} exp\left(-\frac{1}{2}\boldsymbol{a}^2\right)exp\left(-\beta\int_0^1 du\, V\big(q_1(u),\dots,q_N(u)\big)\right) \\ exp\big(i\boldsymbol{\xi}\cdot(\boldsymbol{Q}-\overline{\boldsymbol{q}})\big)\, exp\left(-i\frac{2\sqrt{2}}{\pi^2}\sum_{m=1}^{N}\sum_{k=1,3,5,\dots}^{\infty}\sigma_m\xi_m a_{m,k}/k^2\right)\end{bmatrix} \tag{4.B3}$$

for a general many-body potential $V(q_1, \dots, q_N)$. The N dimensional column vectors of pseudo-particle positions $\boldsymbol{q}$, $\boldsymbol{Q}$, $\boldsymbol{q}'$, $\Delta\boldsymbol{q}$, and $\overline{\boldsymbol{q}}$ are defined in the main text. The N dimensional column vector $\boldsymbol{\xi}$, whose mth element is ξ_m, is introduced by using the integral form of delta function in coarse-graining the intermediate beads. The N dimensional block column vector $\boldsymbol{a}$ has $\boldsymbol{a}_m$ as its mth block element, and the kth element of the infinite dimensional column vector $\boldsymbol{a}_m$ is $a_{m,k}$. Lastly, the N dimensional matrix $\boldsymbol{\Sigma}$ is defined in Eq. (4.A9a).

As shown in Eq. (4.10), we assume that there are two terms in the potential: pairwise decomposable interactions V_{ij} and many-body interactions V_z. The second term in the integral on the RHS of Eq. (4.B3) then can be written as

$$exp\left(-\beta\int_0^1 du\, V\big(q_1(u),\dots,q_N(u)\big)\right)$$

$$= exp\left(-\beta \sum_{\langle i,j\rangle} \int_0^1 du\, V_{ij}\big(q_i(u) - q_i(u)\big)\right) exp\left(-\beta \sum_Z \int_0^1 du\, V_z\big(\boldsymbol{q}(u)\big)\right) \tag{4.B4}$$

Eq. (3.B5) of Appendix 3B shows the Taylor expanded result of the first term on the RHS of Eq. (4.B4). In a similar manner, we Taylor expand the many-body term. We can write

$$\int_0^1 du\, V_z\big(\boldsymbol{q}(u)\big) = \int_0^1 du\, V_z\big(\boldsymbol{q}(0) + \big(\boldsymbol{q}(1) - \boldsymbol{q}(0)\big)u\big)$$

$$+ \int_0^1 du\, \boldsymbol{q}_Q(u) \cdot (DV_z|_c) + \frac{1}{2}\int_0^1 du\, \boldsymbol{q}_Q^T(u)(D^2V_z|_c)\boldsymbol{q}_Q(u) + \int_0^1 du\, \mathcal{O}(D^3V_z) \tag{4.B5}$$

where $DV_z|_c$ is the transposed Jacobian matrix of V_z evaluated at the classical path and $D^2V_z|_c$ is the Hessian matrix of V_z evaluated at the classical path.

For simplicity, we define a N dimensional column vector $\boldsymbol{D}_{Z,c}^{(1)}$ and a N dimensional square matrix $\boldsymbol{D}_{Z,c}^{(2)}$ such that $DV_z|_c = \boldsymbol{D}_{Z,c}^{(1)}$ and $D^2V_z|_c = \boldsymbol{D}_{Z,c}^{(2)}$. The ith element of $\boldsymbol{D}_{Z,c}^{(1)}$ is noted as $\left(\boldsymbol{D}_{Z,c}^{(1)}\right)_i$ and the ith row and jth column element of $\boldsymbol{D}_{Z,c}^{(2)}$ is noted as $\left(\boldsymbol{D}_{Z,c}^{(2)}\right)_{ij}$. Using this notation, integrands of the second and third term on RHS of Eq. (4.B5) can be written as

$$\boldsymbol{q}_Q(u) \cdot (DV_z|_c) = \sum_{i=1}^N \frac{\sqrt{2}\sigma_i}{\pi}\left(\sum_{n=1}^\infty \frac{a_{i,n}\sin(n\pi u)}{n}\right)\left(\boldsymbol{D}_{Z,c}^{(1)}\right)_i \tag{4.B6a}$$

$$\boldsymbol{q}_Q^T(u)(D^2V_z|_c)\boldsymbol{q}_Q(u)$$

$$= \sum_{i=1}^N \sum_{j=1}^N \frac{2\sigma_i\sigma_j}{\pi^2}\left(\sum_{n=1}^\infty \sum_{m=1}^\infty \frac{a_{i,n}a_{j,m}\sin(n\pi u)\sin(m\pi u)}{nm}\right)\left(\boldsymbol{D}_{Z,c}^{(2)}\right)_{ij}. \tag{4.B6b}$$

Before evaluating and simplifying Eq. (4.B5), we define a N dimensional block vector $\boldsymbol{b}_Z$ with infinite dimensional vector as the block element and a N dimensional square block matrix $\boldsymbol{A}_Z$ with infinite dimensional square matrix as the block element . The nth element of the ith block element of $\boldsymbol{b}_Z$ and the nth row and mth column element of the ith row and jth column element of $\boldsymbol{A}_Z$ can be written as

$$(\boldsymbol{b}_Z)_{i,n} = -\beta\frac{\sqrt{2}\sigma_i}{n\pi}\int_0^1 du\left(\boldsymbol{D}_{Z,c}^{(1)}\right)_i \sin(n\pi u) \tag{4.B7a}$$

$$(\boldsymbol{A}_Z)_{ij,nm} = \beta \frac{2\sigma_i\sigma_j}{nm\pi^2} \int_0^1 du \left(\boldsymbol{D}_{Z,c}^{(2)}\right)_{ij} \sin(n\pi u) \sin(m\pi u). \tag{4.B7b}$$

Lastly, we define a vector $\boldsymbol{B}_2$ and a matrix $\boldsymbol{\Omega}_2$ with same dimensionality as $\boldsymbol{b}_Z$ and $\boldsymbol{A}_Z$ such that $\boldsymbol{B}_2 = \sum_Z \boldsymbol{b}_Z$ and $\boldsymbol{\Omega}_2 = \sum_Z \boldsymbol{A}_Z$. Note that $\boldsymbol{b}_Z$ and $\boldsymbol{A}_Z$ are two quantities for a signle many-body potential V_z and $\boldsymbol{B}_2$ and $\boldsymbol{\Omega}_2$ sum over all many-body interactions present in the force field.

Using Eqs. (4.B6a) and (4.B6b) along with the newly defined vectors, we can evaluate the integral with respect to u for the many-body potential V_Z. Eq. (4.B5) can be rewritten as

$$\int_0^1 du\, V_z\big(\boldsymbol{q}(u)\big) = \int_0^1 du\, V_z\big(\boldsymbol{q}(0) + \big(\boldsymbol{q}(1) - \boldsymbol{q}(0)\big)u\big)$$

$$-\frac{1}{\beta}\boldsymbol{a}\cdot\boldsymbol{b}_Z + \frac{1}{2\beta}\boldsymbol{a}^T\boldsymbol{A}_Z\boldsymbol{a} + \int_0^1 du\, \mathcal{O}(D^3V_z). \tag{4.B8}$$

Using both Eq. (4.B8) and Eq. (3.B5) to rewrite Eq. (4.B3), we have

$$e^{-\beta V_{CG,N,trunc}} = \left(\prod_{i=1}^N \frac{\sigma_i}{\sqrt{8\pi}}\right) exp\left(-\frac{1}{2}\Delta\boldsymbol{q}^T\boldsymbol{\Sigma}\Delta\boldsymbol{q}\right)$$

$$\times exp\left(-\beta\int_0^1 du\, V\big(\boldsymbol{q}(0) + \big(\boldsymbol{q}(1) - \boldsymbol{q}(0)\big)u\big)\right)$$

$$\times \int \frac{d\boldsymbol{\xi}d\boldsymbol{a}}{\sqrt{2\pi}} exp\left[-\frac{1}{2}\boldsymbol{a}^T(\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2 + \boldsymbol{I})\boldsymbol{a} + (\boldsymbol{B}_1 + \boldsymbol{B}_2)\cdot\boldsymbol{a} + i\boldsymbol{\xi}\cdot(\boldsymbol{Q} - \overline{\boldsymbol{q}})\right]. \tag{4.B9}$$

Note that third and higher order Taylor expansion terms are not included in Eq. (4.B9). We name the second order truncated expression of the many-body CG potential as $V_{CG,N,trunc}$. In Eq. (4.B9), the matrix $\boldsymbol{\Omega}_1$ has the same dimensionality as $\boldsymbol{\Omega}_2$ and is the matrix defined in Eq. (3.B10), and the vector $\boldsymbol{B}_1$ has the same dimensionality as $\boldsymbol{B}_2$ as the vector defined in Eq. (3.B16). Nomenclature of both quantities were changed for notational convenience.

The Gaussian integral with respect to $\boldsymbol{a}$ in Eq. (4.B9) can be evaluated as

$$\int \frac{d\boldsymbol{a}}{\sqrt{2\pi}} exp\left[-\frac{1}{2}\boldsymbol{a}^T(\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2 + \boldsymbol{I})\boldsymbol{a} + (\boldsymbol{B} + \boldsymbol{B}_2)\cdot\boldsymbol{a}\right]$$

$$= \sqrt{\frac{1}{det(\boldsymbol{\Omega}_1+\boldsymbol{\Omega}_2+\boldsymbol{I})}}\, exp\left(\frac{1}{2}(\boldsymbol{B} + \boldsymbol{B}_2)^T\boldsymbol{\Lambda}(\boldsymbol{B} + \boldsymbol{B}_2)\right) \tag{4.B10}$$

for $\boldsymbol{\Lambda} = (\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2 + \boldsymbol{I})^{-1}$. Eq. (4.B9) can then be simplified as

$$e^{-\beta V_{CG,N,trunc}} = \left(\prod_{i=1}^{N} \frac{\sigma_i}{\sqrt{8\pi}}\right) exp\left(-\frac{1}{2}\Delta \boldsymbol{q}^T \boldsymbol{\Sigma} \Delta \boldsymbol{q}\right)$$

$$\times exp\left(-\beta \int_0^1 du\, V\big(\boldsymbol{q}(0) + \big(\boldsymbol{q}(1) - \boldsymbol{q}(0)\big)u\big)\right)$$

$$\times \sqrt{\frac{1}{det(\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2 + \boldsymbol{I})}} \int d\boldsymbol{\xi}\, exp\left[\frac{1}{2}(\boldsymbol{B}_1 + \boldsymbol{B}_2)^T \boldsymbol{\Lambda}(\boldsymbol{B}_1 + \boldsymbol{B}_2) + i\boldsymbol{\xi} \cdot (\boldsymbol{Q} - \overline{\boldsymbol{q}})\right]. \quad (4.B11)$$

Before we evaluate the integral with respect to $\boldsymbol{\xi}$ in Eq. (4.B11), we define matrices to simplify the notation. For the two N dimensional square matrices $\boldsymbol{L}$ and $\boldsymbol{R}$ defined in Eqs. (3.A8a) and (3.A8b), we define two N dimensional column vector $\boldsymbol{U}$ and $\boldsymbol{D}$ whose mth element are written as

$$\boldsymbol{U}_m = \sigma_m \sum_{n=1}^{N} \boldsymbol{b}_\xi{}^T \boldsymbol{\Lambda}_{mn} \boldsymbol{B}_2 \quad (4.B12a)$$

$$\boldsymbol{D}_m = \sigma_m \sum_{n=1}^{N} \boldsymbol{b}_\lambda{}^T \boldsymbol{\Lambda}_{mn} \boldsymbol{B}_2. \quad (4.B12b)$$

Using the new notation, we can expand the exponent in Eq. (4.B11) to have

$$\frac{1}{2}(\boldsymbol{B}_1 + \boldsymbol{B}_2)^T \boldsymbol{\Lambda}(\boldsymbol{B}_1 + \boldsymbol{B}_2) + i\boldsymbol{\xi} \cdot (\boldsymbol{Q} - \overline{\boldsymbol{q}})$$

$$= -\frac{1}{2}\boldsymbol{\xi}^T \boldsymbol{L} \boldsymbol{\xi} + i\boldsymbol{\xi}\left(\frac{\beta}{2}\boldsymbol{S} + \boldsymbol{U} + (\boldsymbol{Q} - \overline{\boldsymbol{q}})\right)$$

$$+ \frac{\beta^2}{8}(\boldsymbol{F}^T \boldsymbol{L} \boldsymbol{F} + \boldsymbol{G}^T \boldsymbol{R} \boldsymbol{G}) + \frac{\beta}{2}(\boldsymbol{F} \cdot \boldsymbol{U} + \boldsymbol{G} \cdot \boldsymbol{D}) + \frac{1}{2}\boldsymbol{B}_2{}^T \boldsymbol{\Lambda} \boldsymbol{B}_2 \quad (4.B13)$$

where the two N dimensional square matrices $\boldsymbol{F}$ and $\boldsymbol{G}$ are defined according to Eqs. (3.B13a) and (3.B13b). and the N dimensional column vector $\boldsymbol{S}$ satisfies $\boldsymbol{S} = \boldsymbol{L}\boldsymbol{F}$. Combine Eqs. (4.B11) and (4.B13) to have

$$e^{-\beta V_{CG,N,trunc}} = \left(\prod_{i=1}^{N} \frac{\sigma_i}{\sqrt{8\pi}}\right) exp\left(-\frac{1}{2}\Delta \boldsymbol{q}^T \boldsymbol{\Sigma} \Delta \boldsymbol{q}\right)$$

$$\times exp\left(-\beta \int_0^1 du\, V\big(\boldsymbol{q}(0) + \big(\boldsymbol{q}(1) - \boldsymbol{q}(0)\big)u\big)\right)$$

$$\times \sqrt{\frac{1}{det(\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2 + \boldsymbol{I})}}\, exp\left(\frac{\beta^2}{8}(\boldsymbol{F}^T \boldsymbol{L} \boldsymbol{F} + \boldsymbol{G}^T \boldsymbol{R} \boldsymbol{G}) + \frac{\beta}{2}(\boldsymbol{F} \cdot \boldsymbol{U} + \boldsymbol{G} \cdot \boldsymbol{D}) + \frac{1}{2}\boldsymbol{B}_2{}^T \boldsymbol{\Lambda} \boldsymbol{B}_2\right)$$

$$\times \int d\boldsymbol{\xi}\, exp\left[-\frac{1}{2}\boldsymbol{\xi}^T \boldsymbol{L} \boldsymbol{\xi} + i\boldsymbol{\xi}\left(\frac{\beta}{2}\boldsymbol{S} + \boldsymbol{U} + (\boldsymbol{Q} - \overline{\boldsymbol{q}})\right)\right]. \quad (4.B14)$$

Evaluating the Gaussian integral in Eq. (4.B14) gives

$$e^{-\beta V_{CG,N,trunc}} = \left(\prod_{i=1}^{N} \frac{\sigma_i}{\sqrt{8\pi}}\right) exp\left(-\frac{1}{2}\Delta \boldsymbol{q}^T \boldsymbol{\Sigma} \Delta \boldsymbol{q}\right)$$

$$\times exp\left(-\beta \int_0^1 du\, V\big(\boldsymbol{q}(0) + \big(\boldsymbol{q}(1) - \boldsymbol{q}(0)\big)u\big)\right) \sqrt{\frac{(2\pi)^N}{det(\boldsymbol{\Omega}_1+\boldsymbol{\Omega}_2+\boldsymbol{I})det(\boldsymbol{L})}}$$

$$\times exp\left(\begin{matrix} \frac{\beta^2}{8}\boldsymbol{G}^T\boldsymbol{R}\boldsymbol{G} + \frac{\beta}{2}\boldsymbol{G}\cdot\boldsymbol{D} + \frac{1}{2}\boldsymbol{B}_2{}^T\boldsymbol{\Lambda}\boldsymbol{B}_2 - \frac{1}{2}(\boldsymbol{Q}-\overline{\boldsymbol{q}})^T\boldsymbol{L}^{-1}(\boldsymbol{Q}-\overline{\boldsymbol{q}}) \\ -\frac{1}{2}\boldsymbol{U}^T\boldsymbol{L}^{-1}\boldsymbol{U} - \frac{\beta}{2}\boldsymbol{F}\cdot(\boldsymbol{Q}-\overline{\boldsymbol{q}}) - \boldsymbol{U}^T\boldsymbol{L}^{-1}(\boldsymbol{Q}-\overline{\boldsymbol{q}}) \end{matrix}\right). \tag{4.B15}$$

Lastly, we can evaluate the integral with respect to u in Eq. (4.B7a) by Taylor expanding the classical with respect to $\overline{\boldsymbol{q}}$ and retain up to second derivative (Hessian) terms such that

$$(\boldsymbol{b}_Z)_{i,n} = -\beta \frac{\sqrt{2}\sigma_i}{n\pi} \int_0^1 du \left(\boldsymbol{D}_{Z,c}^{(1)}\right)_i \sin(n\pi u)$$

$$\cong \beta\sigma_i \left(b_{\xi,n}\left(\boldsymbol{D}_{Z,\overline{\boldsymbol{q}}}^{(1)}\right)_i + b_{\lambda,n}\left(\boldsymbol{D}_{Z,\overline{\boldsymbol{q}}}^{(2)}\right)_i \cdot \Delta\boldsymbol{q}\right) \tag{4.B16}$$

The term $\left(\boldsymbol{D}_{Z,\overline{\boldsymbol{q}}}^{(1)}\right)_i$ is the ith element of the transposed Jacobian matrix of V_Z evaluated at the position $\overline{\boldsymbol{q}}$. Similarly, the term $\left(\boldsymbol{D}_{Z,\overline{\boldsymbol{q}}}^{(2)}\right)_i$ is the ith row of the Hessian matrix of V_Z evaluated at the position $\overline{\boldsymbol{q}}$. Using the Taylor expanded result in Eq. (4.B16), we can rewrite the ith block element of $\boldsymbol{B}_2$ such that

$$(\boldsymbol{B}_2)_i = \beta\sigma_i\left(\mathcal{E}_i^{(1)}\boldsymbol{b}_\xi + \mathcal{E}_i^{(2)}\boldsymbol{b}_\lambda\right) \tag{4.B17}$$

for two N dimensional column vectors $\boldsymbol{\mathcal{E}}^{(1)}$ and $\boldsymbol{\mathcal{E}}^{(2)}$ whose ith elements $\mathcal{E}_i^{(1)}$ and $\mathcal{E}_i^{(2)}$ are written as

$$\mathcal{E}_i^{(1)} = \sum_Z \left(\boldsymbol{D}_{Z,\overline{\boldsymbol{q}}}^{(1)}\right)_i \tag{4.B18a}$$

$$\mathcal{E}_i^{(2)} = \sum_Z \left(\boldsymbol{D}_{Z,\overline{\boldsymbol{q}}}^{(2)}\right)_i \cdot \Delta\boldsymbol{q}. \tag{4.B18b}$$

Using this notation, we can rewrite the two vectors $\boldsymbol{U}$ and $\boldsymbol{D}$ as well such that

$$\boldsymbol{U} = \beta\boldsymbol{L}\boldsymbol{\mathcal{E}}^{(1)} \tag{4.B19a}$$

$$\boldsymbol{D} = \beta \boldsymbol{R} \boldsymbol{\mathcal{E}}^{(2)}. \tag{4.B19b}$$

Using the new notations, we can simplify Eq. (4.B15) to have

$$e^{-\beta V_{CG,N,trunc}} = \left(\prod_{i=1}^{N} \frac{\sigma_i}{\sqrt{8\pi}}\right) exp\left(-\frac{1}{2}\Delta \boldsymbol{q}^T \boldsymbol{\Sigma} \Delta \boldsymbol{q}\right)$$

$$\times exp\left(-\beta \int_0^1 du\, V\big(\boldsymbol{q}(0) + \big(\boldsymbol{q}(1) - \boldsymbol{q}(0)\big)u\big)\right) \sqrt{\frac{(2\pi)^N}{det(\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2 + \boldsymbol{I}) det(\boldsymbol{L})}}$$

$$\times exp\begin{pmatrix} \frac{\beta^2}{8}\big(\boldsymbol{G} + 2\boldsymbol{\mathcal{E}}^{(2)}\big)^T \boldsymbol{R}\big(\boldsymbol{G} + 2\boldsymbol{\mathcal{E}}^{(2)}\big) \\ -\frac{1}{2}(\boldsymbol{Q} - \overline{\boldsymbol{q}})^T \boldsymbol{L}^{-1}(\boldsymbol{Q} - \overline{\boldsymbol{q}}) \\ -\frac{\beta}{2}\big(\boldsymbol{F} + 2\boldsymbol{\mathcal{E}}^{(1)}\big) \cdot (\boldsymbol{Q} - \overline{\boldsymbol{q}}) \end{pmatrix}. \tag{4.B20}$$

Or alternatively, the truncated many-body CG potential $V_{CG,N}$ can be written as

$$V_{CG,N,trunc} = \frac{1}{2\beta}\Delta \boldsymbol{q}^T \boldsymbol{\Sigma} \Delta \boldsymbol{q} + \int_0^1 du\, V\big(\boldsymbol{q}(0) + \big(\boldsymbol{q}(1) - \boldsymbol{q}(0)\big)u\big)$$

$$-\frac{\beta}{8}\big(\boldsymbol{G} + 2\boldsymbol{\mathcal{E}}^{(2)}\big)^T \boldsymbol{R}\big(\boldsymbol{G} + 2\boldsymbol{\mathcal{E}}^{(2)}\big) + \frac{1}{2\beta}(\boldsymbol{Q} - \overline{\boldsymbol{q}})^T \boldsymbol{L}^{-1}(\boldsymbol{Q} - \overline{\boldsymbol{q}})$$

$$+\frac{1}{2}\big(\boldsymbol{F} + 2\boldsymbol{\mathcal{E}}^{(1)}\big) \cdot (\boldsymbol{Q} - \overline{\boldsymbol{q}}) - \frac{1}{\beta}\ln\left[\left(\prod_{i=1}^{N} \frac{\sigma_i}{\sqrt{8\pi}}\right) \sqrt{\frac{(2\pi)^N}{det(\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2 + \boldsymbol{I}) det(\boldsymbol{L})}}\right] \tag{4.B21}$$

The integral term in Eq. (4.B21) can be decomposed and Taylor expanded to have

$$\int_0^1 du\, V_{ij}\big(A_{ij} + \big(B_{ij} - A_{ij}\big)u\big)$$

$$= V_{ij}\left(\overline{q}_i - \overline{q}_j\right) + \frac{1}{24} V_{ij}''\left(\overline{q}_i - \overline{q}_j\right) \cdot \left(\Delta q_i - \Delta q_j\right)^2 + \mathcal{O}\big(V_{ij}'''\big) \tag{4.B22}$$

and

$$\int_0^1 du\, V_z\big(\boldsymbol{q}(0) + \big(\boldsymbol{q}(1) - \boldsymbol{q}(0)\big)u\big) = V_z(\overline{\boldsymbol{q}}) + \frac{1}{24}\Delta \boldsymbol{q}^T \boldsymbol{D}_{Z,\overline{\boldsymbol{q}}}^{(2)} \Delta \boldsymbol{q} + \mathcal{O}(D^3) \tag{4.B23}$$

where $A_{ij} = q_i(0) - q_j(0)$ and $B_{ij} = q_i(1) - q_j(1)$ and $\boldsymbol{D}_{Z,\overline{\boldsymbol{q}}}^{(2)}$ is the Hessian of V_z evaluated at $\overline{\boldsymbol{q}}$. Omitting the third and higher order terms in Eqs. (4.B22) and (4.B23), we can rewrite Eq. (4.B21) as

$$V_{CG,N,trunc} = \frac{1}{2\beta}\Delta\boldsymbol{q}^T\boldsymbol{\Sigma}\Delta\boldsymbol{q} + V(\overline{\boldsymbol{q}}) + \frac{1}{24}\sum_{\langle i,j\rangle} V_{ij}''\left(\overline{q}_i - \overline{q}_j\right)\cdot\left(\Delta q_i - \Delta q_j\right)^2 +$$

$$\frac{1}{24}\sum_{\boldsymbol{Z}}\Delta\boldsymbol{q}^T\boldsymbol{D}_{Z,\overline{\boldsymbol{q}}}^{(2)}\Delta\boldsymbol{q} - \frac{\beta}{8}\left(\boldsymbol{G} + 2\boldsymbol{\mathcal{E}}^{(2)}\right)^T\boldsymbol{R}\left(\boldsymbol{G} + 2\boldsymbol{\mathcal{E}}^{(2)}\right) + \frac{1}{2\beta}(\boldsymbol{Q} - \overline{\boldsymbol{q}})^T\boldsymbol{L}^{-1}(\boldsymbol{Q} - \overline{\boldsymbol{q}}) +$$

$$\frac{1}{2}\left(\boldsymbol{F} + 2\boldsymbol{\mathcal{E}}^{(1)}\right)\cdot(\boldsymbol{Q} - \overline{\boldsymbol{q}}) - \frac{1}{\beta}\ln\left[\left(\prod_{i=1}^{N}\frac{\sigma_i}{\sqrt{8\pi}}\right)\sqrt{\frac{(2\pi)^N}{det(\boldsymbol{\Omega}_1+\boldsymbol{\Omega}_2+\boldsymbol{I})det(\boldsymbol{L})}}\right]. \tag{4.B24}$$

Lastly, we re-express the $V(\overline{\boldsymbol{q}})$ term by expanding $V(\boldsymbol{Q})$ at $\overline{\boldsymbol{q}}$ by making the two substitutions shown as

$$\sum_{\langle i,j\rangle} V_{ij}\left(\overline{q}_i - \overline{q}_j\right) + \frac{1}{2}\boldsymbol{F}\cdot(\boldsymbol{Q} - \overline{\boldsymbol{q}})$$

$$\rightarrow \sum_{\langle i,j\rangle} V_{ij}\left(Q_i - Q_j\right) - \frac{1}{2}V_{ij}''\left(\overline{q}_i - \overline{q}_j\right)\left(\left(Q_i - \overline{q}_i\right) - \left(Q_j - \overline{q}_j\right)\right)^2 \tag{4.B25}$$

and

$$\sum_{\boldsymbol{Z}} V_Z(\overline{\boldsymbol{q}}) + \boldsymbol{\mathcal{E}}^{(1)}\cdot(\boldsymbol{Q} - \overline{\boldsymbol{q}}) \rightarrow \sum_{\boldsymbol{Z}} V_Z(\boldsymbol{Q}) - \frac{1}{2}(\boldsymbol{Q} - \overline{\boldsymbol{q}})^T\boldsymbol{D}_{Z,\overline{\boldsymbol{q}}}^{(2)}(\boldsymbol{Q} - \overline{\boldsymbol{q}}). \tag{4.B26}$$

The LHS of Eqs. (4.B25) and (4.B26) are already included in Eq. (4.B24). Combining the substitution, the second order truncated Taylor expanded expression of the many-body CG potential can be written as

$$V_{CG,N,trunc} = V(\boldsymbol{Q}) + \frac{1}{2}(\boldsymbol{Q} - \overline{\boldsymbol{q}})^T\boldsymbol{K}_{Q\bar{q}}(\boldsymbol{Q} - \overline{\boldsymbol{q}}) + \frac{1}{2}\Delta\boldsymbol{q}^T\boldsymbol{K}_{\Delta q}\Delta\boldsymbol{q} + K_0 \tag{4.B27}$$

where

$$\boldsymbol{K}_{Q\bar{q}} = \frac{1}{\beta}\boldsymbol{L}^{-1} - \frac{1}{2}M - \sum_{\boldsymbol{Z}}\boldsymbol{D}_{Z,\overline{\boldsymbol{q}}}^{(2)} \tag{4.B28a}$$

$$\boldsymbol{K}_{\Delta q} = \frac{1}{24}M + \frac{1}{\beta}\boldsymbol{\Sigma} - \frac{\beta}{4}\boldsymbol{C} + \frac{1}{12}\sum_{\boldsymbol{Z}}\boldsymbol{D}_{Z,\overline{\boldsymbol{q}}}^{(2)} \tag{4.B28b}$$

$$K_0 = \frac{1}{2\beta}\ln\left[\left(\prod_{i=1}^{N}\frac{4m_i}{\hbar^2\beta}\right)det(\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2 + \boldsymbol{I})det(\boldsymbol{L})\right] \tag{4.B28c}$$

for the N dimensional square matrices $\boldsymbol{\Sigma}$ defined in Eq. (4.A9a), M defined in Eq. (4.A9b), and $\boldsymbol{C}$ defined in Eq. (4.A10).

References

1. H. A. Stern and B. J. Berne, J. Chem. Phys. **115**, 7622 (2001).

2. M. Shiga and W. Shinoda, J. Chem. Phys. **123**, 134502 (2005).

3. F. Paesani, W. Zhang, D. A. Case, T. E. CheathamIII, and G. A. Voth, J. Chem. Phys. **125**, 184507 (2006).

4. J. A. Morrone and R. Car, Phys. Rev. Lett. **101**, 017801 (2008).

5. S. Habershon, T. E. Markland, and D. E. Manolopoulos, J. Chem. Phys. **131**, 024501 (2009).

6. C. Vega, M. M. Conde, C. McBride, J. L. F. Abascal, E. G. Noya, R. Ramirez, and L. M. Sesé, J. Chem. Phys. **132**, 046101 (2010).

7. B. Pamuk, J. M. Soler, R. Ramírez, C. P. Herrero, P. W. Stephens, P. B. Allen, and M. V. Fernández-Serra, Phys. Rev. Lett. **108**, 193003 (2012).

8. S. Fritsch, R. Potestio, D. Donadio, and K. Kremer, J. Chem. Theory Comput. **10**, 816 (2014).

9. L. Wang, M. Ceriotti, and T. E. Markland, J. Chem. Phys. **141**, 104502 (2014).

10. Y. Litman, D. Donadio, M. Ceriotti, and M. Rossi, J. Chem. Phys. **148**, 102320 (2018).

11. D. M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995).

12. T. D. Hone and G. A. Voth, J. Chem. Phys. **121**, 6412 (2004).

13. J. Chen, X. Z. Li, Q. Zhang, M. I. J. Probert, C. J. Pickard, R. J. Needs, A. Michaelides, and E. Wang, Nat. Commun. **4**, 2064 (2013).

14. M. A. Morales, J. M. McMahon, C. Pierleonim and D. M. Ceperley, Phys. Rev. Lett. **110**, 065702 (2013).

15. C. M. Herdman, A. Rommal, and A. Del Maestro, Phys. Rev. B. **89**, 224502 (2014).

16. M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello, Science. **275**, 817 (1997).

17. M. Ceriotti, J. Cuny, M. Parrinello, and D. E. Manolopoulos, Proc. Natl. Acad. Sci. U. S. A., **110**, 15591 (2013).

18. B. Walker and A. Michaelides, J. Chem. Phys. **133**, 174306 (2010).

19. J. Lobaugh and G. A. Voth, J. Chem. Phys. **104**, 2056 (1996).

20. U. W. Schmitt and G. A. Voth, J. Chem. Phys. **111**, 9361 (1999).

21. R. Vuilleumier and D. Borgis, J. Chem. Phys. **111**, 4251 (1999).

22. D. Marx, M. E. Tuckerman, and M. Parrinello, J. Condens. Matter. Phys. **12**, A153 (2000).

23. S. R. Billeter, S. P. Webb, T. Iordanov, P. K. Agarwal, and S. Hammes-Schiffer, J. Chem. Phys. **114**, 6925 (2001).

24. R. P. Feynman, *Quantum Mechanics and Path Integrals*. (McGraw-Hill, New York, 1965).

25. R. P. Feynman, *Statistical mechanics: A Set of Lectures* (W.A. Benjamin, Reading, Mass, 1972).

26. D. Chandler and P. G. Wolynes, J. Chem. Phys. **74**, 4078 (1981).

27. B. J. Berne and D. Thirumalai, Annu. Rev. Phys. Chem. **37**, 401 (1986).

28. M. Parrinello and A. Rahman, J. Chem. Phys. **80**, 860 (1984).

29. M. E. Tuckerman, B. J. Berne, G. J. Martyna, and M. L. Klein, J. Chem. Phys. **99**, 2796 (1993).

30. D. Marx and M. Parrinello, J. Chem. Phys. **104**, 4077 (1996).

31. M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello, J. Chem. Phys. **104**, 5579 (1996).

32. M. F. Herman, E. J. Bruskin, and B. J. Berne, J. Chem. Phys. **76**, 5150 (1982).

33. D. Thirumalai, R. W. Hall, and B. J. Berne, J. Chem. Phys. **81**, 2523 (1984).

34. M. Diraison, G. J. Martyna, and M. E. Tuckerman, J. Chem. Phys. **111**, 1096 (1999).

35. P. Sindzingre, M. L. Klein, and D. M. Ceperley, Phys. Rev. Lett. **63**, 1601 (1989).

36. D. Scharf, M. L. Klein, and G. J. Martyna, J. Chem. Phys. **97**, 3590 (1992).

37. M. Sprik, R. W. Impey, and M. L. Klein, J. Chem. Phys. **83**, 5802 (1985).

38. D. F. Coker, B. J. Berne, and D. Thirumalai, J. Chem. Phys. **86**, 5689 (1987).

39. A. Wallqvist, D. Thirumalai, and B. J. Berne, J. Chem. Phys. **86**, 6404 (1987).

40. J. Cao and G. A. Voth, J. Chem. Phys. **100**, 5093 (1994).

41. J. Cao and G. A. Voth, J. Chem. Phys. **100**, 5106 (1994).

42. J. Cao and G. A. Voth, J. Chem. Phys. **101**, 6157 (1994).

43. J. Cao and G. A. Voth, J. Chem. Phys. **101**, 6168 (1994).

44. J. Cao and G. A. Voth, J. Chem. Phys. **101**, 6184 (1994).

45. I. R. Craig and D. E. Manolopoulos, J. Chem. Phys. **121**, 3368 (2004).

46. B. J. Braams and D. E. Manolopoulos, J. Chem. Phys. **125**, 124105 (2006).

47. J. O. Richardson and S. C. Althorpe, J. Chem. Phys. **131**, 214106 (2009).

48. S. Habershon, D. E. Manolopoulos, T. E. Markland, and T. F. M. III, Annu. Rev. Phys. Chem. **64**, 387 (2013).

49. M. Rossi, M. Ceriotti, and D. E. Manolopoulos, J. Chem. Phys. **140**, 234116 (2014).

50. J. A. Poulsen and P. J. Rossky, J. Chem. Phys. **115**, 8024 (2001).

51. A. Witt, S. D. Ivanov, M. Shiga, H. Forbert, and D. Marx, J. Chem. Phys. **130**, 194510 (2009).

52. T. F. M. III and D. E. Manolopoulos, J. Chem. Phys. **122**, 184503 (2005).

53. T. F. M. III and D. E. Manolopoulos, J. Chem. Phys. **123**, 154504 (2005).

54. E. Geva, Q. Shi, and G. A. Voth, J. Chem. Phys. **115**, 9209 (2001).

55. M. E. Tuckerman and M. Parrinello, J. Chem. Phys. **101**, 1316 (1994).

56. T. E. Markland and D. E. Manolopoulos, J. Chem. Phys. **129**, 024105 (2008).

57. W. G. Noid, J. Chem. Phys. **139**, 090901 (2013).

58. M. G. Saunders and G. A. Voth, Annu. Rev. Biophys. **42**, 73 (2013).

59. A. V. Sinitskiy and G. A. Voth, J. Chem. Phys. **143**, 094104 (2015).

60. W. H. Ryu, Y. Han and G. A. Voth, J. Chem. Phys. **150**, 244103 (2019).

61. N. Blinov and P. N. Roy, J. Chem. Phys. **115**, 7822 (2001).

62. N. V. Blinov, P. N. Roy, and G. A. Voth, J. Chem. Phys. **115**, 4484 (2001).

63. P. N. Roy and N. Blinov, Isr. J. Chem. **42**, 183 (2002).

64. Y. Han, J. Jin, J. W. Wagner, and G. A. Voth, J. Chem. Phys. **148**, 102335 (2018).

65. S. Izvekov and G. A. Voth, J. Chem. Phys. **123**, 134105 (2005).

66. S. Izvekov and G. A. Voth, J. Phys. Chem. B. **109**, 2469 (2005).

67. W. G. Noid, J. W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, and H. C. Andersen, J. Chem. Phys. **128**, 244114 (2008).

68. W. G. Noid, P. Liu, Y. Wang, J. W. Chu, G. S. Ayton, S. Izvekov, H. C. Andersen, and G. A. Voth, J. Chem. Phys. **128**, 244115 (2008).

69. Y. Peng, Z. Cao, R. Zhou, and G. A. Voth, J. Chem. Theory. Comput. **10**, 3634 (2014).

70. M. Nava, R. Quhe, F. Palazzesi, P. Tiwary, and M. Parrinello, J. Chem. Theory. Comput. **11**, 5114 (2015).

71. M. Nava, F. Palazzesi, C. Perego, and M. Parrinello, J. Chem. Theory. Comput. **13**, 425 (2017).

72. F. H. Stillinger and T. A. Weber, Phys. Rev. B. **31**, 5262 (1985).

73. W. Tschöp, K. Kremer, J. Batoulis, T. Bürger, and O. Hahn, Acta. Polym. **49**, 61 (1998).

74. F. Müller-Plathe, Chem. Phys. Chem **3**, 754 (2002).

75. D. Reith, M. Pütz, and F. Müller-Plathe, J. Comput. Chem. **24**, 1624 (2003).

76. V. Agrawal, G. Arya, and J. Oswald, Macromolecules. **47**, 3378 (2014).

77. I. F. Silvera and V. V. Goldman, J. Chem. Phys. **69**, 4209 (1978).

78. S. Plimpton, J. Comput. Chem. **117**, 1 (1995).

79. G. Bussi and M. Parrinello, Phys. Rev. E. **75**, 056707 (2007).

80. M. Ceriotti, M. Parrinello, T. E. Markland, and D. E. Manolopoulos, J. Chem. Phys. **133**, 124104 (2010).

Chapter 5: Generalized Langevin Equation Analysis of Ring Polymer Molecular Dynamics

Abstract

The Feynman's imaginary time path integral provides a theoretical basis for calculating numerically exact quantum statistics using classical methods such as molecular dynamics or Monte Carlo. Both path integral molecular dynamics (PIMD) and path integral Monte Carlo (PIMC) methods have been widely used to calculate equilibrium properties of many condensed matter systems. Extending the path integral methods into a dynamical realm, however, has been a less straightforward problem. Two methods, centroid molecular dynamics (CMD) and ring polymer molecular dynamics (RPMD), have been used to study approximate quantum dynamics of systems with distinguished degree of nuclear quantum effect. Recently, the thermostatted RPMD (TRPMD) has been proposed to improve upon the "curvature problem" and "resonance problem" of CMD and RPMD respectively by propagating the ring polymer using Langevin dynamics. In this work we attempt to recapitulate the RPMD dynamics using the generalized Langevin equation (GLE) analysis based on the observation that the P beads in a ring polymer can be understood as a particle interacting with a bath of $P - 1$ beads. A mapping scheme between the ring polymer Hamiltonian and the system and bath Hamiltonian is presented. We then derive the GLE from the system and the bath form. To check the validity of the analysis, we calculate the Kubo-transformed position autocorrelation functions of simple systems. It is shown that the GLE picture of RPMD dynamics not only qualitatively captures the oscillation of the full RPMD dynamics but also highlights the importance of higher order interactions for the bath modes.

I. Introduction

Feynman's formalism of the imaginary time path integral provides an elegant recapitulation of quantum statistical mechanics without using operators or matrices.[1, 2] One key result of the theory is the isomorphism[3] between the quantum canonical partition function of a particle and the classical canonical partition function of a ring polymer with $P + 1$ "beads" or pseudo-particles that obeys classical mechanics. The ring polymers are characterized by the nearest neighbor interactions and physical potential applied to each beads. The isomorphism suggests that the thermal density matrix element can be obtained by sampling the classical ring polymer in its extended phase space. Many path integral methods that use well-developed classical sampling methods such as molecular dynamics (path integral molecular dynamics, PIMD) [4, 5] and Monte Carlo sampling (path integral Monte Carlo, PIMC)[4, 6] have been widely developed.

There are two successors of the PIMD method that extend the imaginary time path integral formalism to the realm of quantum dynamics: centroid molecular dynamics (CMD)[7-11] (CMD) and ring polymer molecular dynamics (RPMD).[12, 13] CMD uses the imaginary time path centroid (center of mass of the ring polymer) as the central quantity in describing the equilibrium and dynamical quantities in quantum statistical mechanics.[14, 15] In principle, when running CMD calculation, one obtains the centroid potential of mean force (PMF) from the physical potential by sampling the available configuration of non-centroid modes for a fixed centroid position. Then, the classical Hamiltonian dynamics of the centroid on the calculated PMF describes the time evolution of the system. Moreover, a more computationally efficient variants of CMD named adiabatic CMD[11] (ACMD) and partially adiabatic CMD[16] (PA-CMD) use an adiabatic decoupling of centroid and non-centroid modes to calculate the centroid PMF

on the fly. On the other hand, RPMD directly considers the dynamics of the P beads of the ring polymer, all of same mass as the original particle. The correlation functions of the beads are then related to Kubo-transformed real time correlation functions based on the symmetry properties of the two functions. Both methods have been extensively used to study many dynamical problems such as calculation of vibrational spectra,[17-21] diffusion of liquids,[22, 23] dynamics of water, [24] and chemical reaction rates in complex systems.[25-28]

The two methods approximate the Kubo transformed correlation functions with the correlation functions of the classical systems (the centroid and the ring polymer).[9, 12] Even though both inherently are approximation to the real-time quantum dynamics, they are accurate in several important limits. CMD and RPMD give exact time correlation functions in the harmonic limit ($V(x) = kx^2$), the classical limit ($\hbar\beta \to 0$), and short time limit ($t \to 0$).[9, 13] Accuracy in the short time limit is especially important due to the quick decay of the quantum interference effects for many condensed matter systems.[12] The centroid dynamics and the approximations that are used in CMD have been studied using quantum dynamical formalisms.[14, 15] However, to this date no specific method has provided a tractable way to provide a solid connection between the ring polymer dynamics and quantum dynamical equations.[29] More recently, a theoretical scheme named Matsubara Dynamics[30, 31], a classical technique that preserves the quantum Boltzmann distribution, has been developed that incorporates both CMD and RPMD based on the level of approximations used in the imaginary time path integral formalism.

Despite their popularity and success, both CMD and RPMD have limitations in which they can give unphysical results: “curvature problem”[32, 33] and “resonance problem,”[32, 34] respectively. Both problems arise from inherent multidimensionality of the ring polymer for

large values of P and are documented from simulation of vibration spectroscopy.[32] CMD's curvature problem comes into play for low temperature systems in which the nearest neighbor coupling strength of the ring polymer decreases and the ring polymer gets "floppier". In such cases, the ring polymer stretches along the angular coordinate and the calculated centroid PMF is determined to have a more shallow minima or lower frequency than that of the actual potential. The calculated vibration spectra with such centroid potentials give broader and red shifted peaks.[32, 33] However the curvature problem appears to be rather insignificant in realistic molecular systems such as liquid water at room temperature.[22, 33-35] Moreover, a recent study that addresses this issue of curvature problem has been developed.[36, 37] RPMD's resonance problem causes spurious splitting of the physical peaks and the appearance of unphysical ones in the calculated vibrational spectra when the intrinsic physical frequencies of the system come into resonance to those of the higher order imaginary time ring polymer normal modes, which is to be expected from classical mechanics.[32, 34] In an attempt to circumvent this problem, Rossi et al.[38] proposed a thermostatted RPMD (TRPMD) approach to be an approximate quantum dynamics method immune to both the curvature and resonance problems by propagating the higher order normal modes of the ring polymer using Langevin dynamics. Moreover, a recent work by Hele has made a connection between Matsubara dynamics and TRPMD by showing that the TRPMD force on the centroid is identical to the Matsubara force on the centroid.[39] Overall, the generalization of the imaginary time path integral methods into a dynamical one is a challenging problem and is an active field of theoretical research.[29, 31, 40, 41]

It should be noted that TRPMD is not the first imaginary time path integral method to incorporate the classical motif of Langevin equation or its more general version, generalized Langevin equation (GLE), to study NQEs. Most notably, the GLE formalism, combined with

imaginary time path integral methods, has been used to mimic quantum effects in condensed matter systems. The PI+GLE method,[42] devised by Ceriotti et al., accelerates the convergence of PIMD simulations by propagating the system in the extended momentum space using Markovian dynamics. Moreover, work by Ceriotti and Manolopoulos has extended the PI+GLE method to calculate quantum kinetic energy and momentum distribution of liquid water.[43] The subsequent work presents how the GLE thermostats can be used to control the dynamical properties such as vibrational spectra.[44] More recently, the work by Kapil et al. introduces a post-processing scheme that recovers dynamic correlation properties from a thermostatted trajectory.[45]

In a similar spirit of combining the quantum-classical isomorphism with GLE, we provide here a perspective on the ring polymer dynamics of both RPMD and TRPMD in the context of GLE in attempt to construct an alternate formalism of the approximate quantum dynamics.[46] Unlike the previously mentioned methods that use a form GLE to propagate the ring polymer dynamics (GLE thermostat), we show here a different use of GLE formalism. In this work, the RPMD Hamiltonian itself is directly mapped into a system and bath Hamiltonian in which the bath coordinates are explicitly treated as set of harmonic oscillators that bilinearly interacts with the system coordinate. The important difference is that the coordinates of the ring polymer become both system and bath during the mapping process. Using the system and bath form of Hamiltonian, we then derive the generalized Langevin equation in the context of ring polymer by solving the system and bath equations of motion using Laplace transform. Moreover, we also present how the Langevin damping term in TRPMD can be modeled in the context of GLE by attaching an exponential decay term in both the friction kernel and the random force. In particular, our formalism will focus on the effect of the bath coordinates to the

system dynamics by comparing the calculated correlation functions to the full RPMD ones as well as exact quantum variational method.

The basics of RPMD and TRPMD and the details of their approximation are presented in Sec. II-A,B. Sec. II-C briefly describes the canonical form of the GLE, followed by Sec. II-D,E that describes the mapping of the ring polymer Hamiltonian into the system and bath Hamiltonian and deriving the generalized Langevin equation for the system and bath form. Sec. II-F presents how the TRPMD dynamics can be modeled using the GLE picture. Sec. II-G discusses the Liouvillian operator formalism of ring polymer dynamics. Numerical examples of calculated bath properties, GLE friction kernel and random force, and finally the position autocorrelation function are presented in Sec. III, with discussion and concluding remarks in Sec. IV.

II. Theory

A. RPMD statistics and dynamics

For a particle in one dimension with mass m and the Hamiltonian operator $\hat{H} = \hat{p}^2/(2m) + \hat{V}(\hat{q})$, the canonical partition function is written as

$$Z = Tr\left[e^{-\beta\hat{H}}\right] \tag{5.1}$$

for inverse temperature $\beta = 1/(k_B T)$. The RPMD formalism uses the isomorphism between the diagonal thermal density matrix element of imaginary time path integral formalism and a classical fictitious ring polymer with P beads to numerically calculate statistics of the system.[3, 12] The canonical partition function of the ring polymer Z_P can be written as

$$Z_P = \frac{1}{(2\pi\hbar)^P}\int d\boldsymbol{q}d\boldsymbol{p}\, exp\left(-\frac{\beta}{P}H_{RPMD}(\boldsymbol{q},\boldsymbol{p})\right) \tag{5.2}$$

in which the P dimensional position and momenta column vectors $\boldsymbol{q}$ and $\boldsymbol{p}$ have position and momenta of the ith ring polymer beads, q_i and p_i respectively, as their ith element and the RPMD Hamiltonian H_{RPMD} can compactly be written in vector and matrix notation as

$$H_{RPMD}(\boldsymbol{q},\boldsymbol{p}) = \frac{1}{2m}\boldsymbol{p}^2 + \frac{1}{2}\boldsymbol{q}^T\boldsymbol{K}\boldsymbol{q} + \sum_{i=1}^{P} V(q_i). \tag{5.3}$$

The ring polymer beads interact with each other and the potential according to the second and the third term in Eq. (5.3). The second term represents the nearest neighbor coupling in which the ith row jth column coupling matrix element of $\boldsymbol{K}$ can be written as

$$K_{ij} = \frac{mP^2}{\beta^2\hbar^2}\left(2\delta_{i,j} - \delta_{i,j-1} - \delta_{i,j+1}\right) \tag{5.4}$$

with $i, j = 1, \ldots, P$ and cyclic property of $q_{P+1} = q_1$.[12] Note that the ith and the i+1th beads are harmonically coupled with the coupling strength of $(mP^2)/(\beta^2\hbar^2)$. The third term suggests that each bead experiences on its own the physical potential of the system $V(q_i)$. The isomorphism itself can be succinctly stated as

$$Z = \lim_{P\to\infty} Z_P. \tag{5.5}$$

An expectation value of a quantum operator also can be estimated with the equivalent classical expectation value of the ring polymer. The thermal expectation value of a position dependent operator $\hat{A}(\hat{q})$ is written as

$$\langle A\rangle = \frac{1}{Z} Tr\left[e^{-\beta\hat{H}}\hat{A}\right]. \tag{5.6}$$

Analogous to Eq. (5.2), the classical expectation value for the position dependent quantity $A(q)$ is written as

$$\langle A\rangle_P = \frac{1}{(2\pi\hbar)^P Z_P}\int d\boldsymbol{q}\, d\boldsymbol{p}\, exp\left(-\frac{\beta}{P}\, H_{RPMD}(\boldsymbol{q},\boldsymbol{p})\right) A_P(\boldsymbol{q}) \tag{5.7}$$

with $A_P(\boldsymbol{q})$ defined as $A_P(\boldsymbol{q}) = \frac{1}{P}\sum_{i=1}^{P} A(q_i)$. Similar to the limiting properties of the partition function, $\langle A\rangle_P$ also satisfies $\lim_{P\to\infty}\langle A\rangle_P = \langle A\rangle$.[12]

The connection between real time quantum dynamics and the RPMD dynamics of the ring polymer starts with equations of motion generated by the classical Hamiltonian defined in Eq. (5.3)

$$\dot{\boldsymbol{p}} = -\frac{\partial}{\partial \boldsymbol{x}} H_{RPMD}(\boldsymbol{q},\boldsymbol{p}) = \boldsymbol{f}(\boldsymbol{q}) - \boldsymbol{K}\boldsymbol{q} \tag{5.8a}$$

$$\dot{\boldsymbol{q}} = \frac{\partial}{\partial \boldsymbol{p}} H_{RPMD}(\boldsymbol{q},\boldsymbol{p}) = \boldsymbol{p}/m \tag{5.8b}$$

in which $\boldsymbol{f}(\boldsymbol{q})$ is a P dimensional column vector with ith element $f_i = -V'(q_i)$ for $i = 1, \dots, P$.[12] The two operator generalization of Eq. (5.7) with explicit time evolution of the phase space variables $\boldsymbol{q}(t)$ and $\boldsymbol{p}(t)$ according to Eqs. (5.8a) and (5.8b) can be written as

$$\langle A(0)B(t)\rangle_P = \frac{1}{(2\pi\hbar)^P Z_P}$$

$$\int d\boldsymbol{q}(0)\, d\boldsymbol{p}(0)\; exp\left(-\frac{\beta}{P}\, H_{RPMD}(\boldsymbol{q},\boldsymbol{p})\right) A_P\big(\boldsymbol{q}(0)\big) B_P\big(\boldsymbol{q}(t)\big) \tag{5.9}$$

in which B_P for a position dependent operator $\hat{B}(\hat{q})$ is analogously defined as A_P. The time dependent function is determined by integrating the bracketed expression with respect to the initial phase space variables $\{\boldsymbol{q}(0), \boldsymbol{p}(0)\}$.[12] The small time limit value of Eq (5.9) corresponds to the $t \to 0$ limit of the Kubo-transformed real time correlation function[12, 47] of the operators $\hat{A}$ and $\hat{B}$

$$\tilde{C}_{AB}(t) = \frac{1}{\beta Z}\int_0^\beta d\lambda\, Tr\left[e^{-(\beta-\lambda)\hat{H}}\, \hat{A}\, e^{-\lambda\hat{H}}\, e^{+i\hat{H}\frac{t}{\hbar}}\, \hat{B}\, e^{-i\hat{H}\frac{t}{\hbar}}\right]. \tag{5.10}$$

The RPMD method proposes $\langle A(0)B(t)\rangle_P$ to be a quasiclassical approximation of $\tilde{C}_{AB}(t)$ for sufficiently large value of P based on the symmetry conditions and the correspondence in the

short time limit of the two functions.[12] Both functions $\langle A(0)B(t)\rangle_P$ and $\tilde{C}_{AB}(t)$ are real-valued and even with respect to time. The real time correlation function $C_{AB}(t)$

$$C_{AB}(t) = \frac{1}{Z} Tr\left[e^{-\beta\hat{H}}\,\hat{A}\,e^{i\hat{H}\frac{t}{\hbar}}\,\hat{B}\,e^{-i\hat{H}\frac{t}{\hbar}}\right] \tag{5.11}$$

is related to $\tilde{C}_{AB}(t)$ in the frequency domain via relationship

$$G_{AB}(\omega) = \left(\beta\hbar\omega/\left(1 - e^{-\beta\hbar\omega}\right)\right)\tilde{G}_{AB}(\omega) \tag{5.12}$$

in which $G_{AB}(\omega)$ is a Fourier transforms of $C_{AB}(t)$ defined as

$$\tilde{G}_{AB}(\omega) = \int_{-\infty}^{\infty} dt\, e^{-i\omega t}\tilde{C}_{AB}(t) \tag{5.13}$$

with analogous Fourier transform relation between $G_{AB}(\omega)$ and $C_{AB}(t)$.[12, 47]

Lastly, we conclude summary of the RPMD theory by introducing a normal mode representation of the ring polymer Hamiltonian.[4, 48] It is more common to use such representation to avoid slow convergence problem that may arise due to the different time scales in the ring polymer dynamics. More importantly, it is a more natural representation to use in the context of the present work as we will use similar coordinate transform techniques when going from ring polymer bead position to system and bath representation.

The coordinate transform between the bead representation and the normal mode representation is defined as

$$\tilde{\boldsymbol{p}} = \boldsymbol{C}\boldsymbol{p} \tag{5.14a}$$

$$\tilde{\boldsymbol{q}} = \boldsymbol{C}\boldsymbol{q} \tag{5.14b}$$

for P dimensional column vectors $\tilde{\boldsymbol{q}}$ and $\tilde{\boldsymbol{p}}$. Their ith elements $\tilde{q}_i$ and $\tilde{p}_i$ represent the ith normal mode position and momenta, respectively. The ith row jth column element of the P dimensional square coordinate transform matrix[49] $\boldsymbol{C}$ can be written as

$$C_{ij} = \begin{cases} \sqrt{1/P}\ for\ j = 1 \\ \sqrt{2/P}\, cos(2\pi ij/P)\ for\ 2 \le j \le P/2 \\ \sqrt{1/P}\,(-1)^{i+1}\ for\ j = P/2 + 1 \\ \sqrt{2/P}\, sin(2\pi ij/P)\ for\ \ P/2 + 2 \le j \le P \end{cases} \tag{5.15}$$

for an even integer P and $i = 1, \dots, P$. For a ring polymer in one dimension with Hamiltonian in bead representation as shown in Eq. (5.3), the equivalent Hamiltonian in normal mode representation is written as

$$H_{RPMD}(\tilde{\boldsymbol{q}}, \tilde{\boldsymbol{p}}) = \frac{1}{2m}\tilde{\boldsymbol{p}}^2 + \frac{m}{2}\tilde{\boldsymbol{q}}^T \boldsymbol{\alpha}\tilde{\boldsymbol{q}} + \sum_{i=1}^{P} V[q_i(\tilde{\boldsymbol{q}})] \tag{5.16}$$

for P dimensional diagonal normal mode frequency matrix $\boldsymbol{\alpha}$ with ith diagonal element ${\alpha_i}^2 = \frac{4P^2}{(\beta\hbar)^2} sin^2((i-1)\pi/P)$ with $i = 1, \dots, P$. Note that the ith bead coordinate q_i in the potential term is expressed as a function of normal mode coordinates.

B. TRPMD statistics and dynamics

Unlike the purely Hamiltonian equations of motion as shown in Eqs. (5.8a) and (5.8b) used to propagate the system in RPMD, the TRPMD method uses Langevin dynamics to propagate the ring polymer. Following the notation of Rossi et al.[38], the TRPMD equations of motion is written as

$$\dot{\boldsymbol{p}} = \boldsymbol{f}(\boldsymbol{q}) - \boldsymbol{K}\boldsymbol{q} - \boldsymbol{\gamma}\boldsymbol{p} + \sqrt{2mP\boldsymbol{\gamma}/\beta}\,\boldsymbol{\zeta}(t) \tag{5.17a}$$

$$\dot{\boldsymbol{q}} = \boldsymbol{p}/m \tag{5.17b}$$

for P dimensional square friction matrix $\boldsymbol{\gamma}$ and P dimensional column vector $\boldsymbol{\zeta}(t)$ whose elements are P uncorrelated Gaussian random numbers with zero mean and unit variance. The friction matrix $\boldsymbol{\gamma}$ is real, symmetric, positive semi-definite and detached from the centroid dynamics. The proposed choice of friction matrix[38] $\boldsymbol{\gamma}$ is parametrized by a parameter λ and can be written as

$$\boldsymbol{\gamma} = \lambda \boldsymbol{\gamma}_c \tag{5.18a}$$

$$\boldsymbol{\gamma}_c = 2\boldsymbol{C}\sqrt{\boldsymbol{\alpha}}\boldsymbol{C}^T \tag{5.18b}$$

with coordinate transform matrix $\boldsymbol{C}$ as specified in Eq. (5.15) and normal mode frequency matrix $\boldsymbol{\alpha}$ as in Eq. (5.16). It has been shown that adding Langevin thermostatting to the bead dynamics as shown in Eqs. (5.18a) and (5.18b) does not affect the calculated equilibrium quantities.[38] Moveover, the thermostatting also does not affect the short time accuracy of the RPMD approximation of the Kubo-transformed time correlation functions.[38]

C. Generalized Langevin equation

The generalized Langevin equation (GLE) for a particle (system) of mass μ in one dimension interacting with the surrounding (bath) is written as

$$\mu\ddot{q}(t) = -\frac{d}{dx}W[q(t)] - \int_0^t d\tau\, \dot{q}(t)\xi(t-\tau) + R(t) \tag{5.19}$$

with the time dependent coordinate of the system $q(t)$, potential of mean force $W[q(t)]$, friction kernel $\xi(t)$, and random force $R(t)$ that satisfies the second fluctuation dissipation theorem (FDT).[50] The integro-differential equation shown in Eq. (5.19) describes an equation of motion of a system interacting with a bath in which the detail of the interaction between the two is defined via the form of the friction kernel.

D. Mapping of RPMD Hamiltonian into the system and bath form

Before we present the GLE analysis of the RPMD dynamics, we note that the derivation presented in this section follow that shown in Chapter 15.2 and 15.7 of Ref.[51]

We begin the mapping from the RPMD Hamiltonian to the system and bath Hamiltonian by observing the RPMD Hamiltonian shown in Eq. (5.3). For a more intuitive understanding, we present it again without using the compact form

$$H_{RPMD}(\boldsymbol{p},\boldsymbol{q}) = \sum_{k=1}^{P}\left[\frac{{p_k}^2}{2m} + \frac{m}{2}\left(\frac{P}{\beta\hbar}\right)^2 (q_k - q_{k+1})^2 + V(q_k)\right]. \tag{5.20}$$

In the context of GLE and system and bath, the system coordinate and the bath coordinates need to be chosen. For the ring polymer system, with P degrees of freedom, the system coordinate or the coordinate of interest will be the first bead position (q_1) and the bath coordinates will be positions of the second to Pth beads ($q_2, \dots, q_P$). This mapping corresponds to the fact that expectation values can be taken with respect to any single ring polymer beads according to Eq. (5.7) and the definition of $A_P(\boldsymbol{q})$. For the closed ring polymer, since the ring polymer is cyclic and no bead is distinct from one another, this divide is arbitrary and made to simplify the derivation.

With this distinction, we see that the full RPMD Hamiltonian can be divided into two parts: the system Hamiltonian $H_{system}(p_1, q_1)$ and the bath Hamiltonian $H_{bath}(\boldsymbol{p}, \boldsymbol{q})$. Note that the system Hamiltonian is a function of only the first bead momentum and position but the bath Hamiltonian is a function of momenta and position of the entire ring polymer. We have

$$H_{system}(p_1, q_1) = \frac{{p_1}^2}{2m} + V(q_1) \tag{5.21}$$

and

$$H_{bath}(\boldsymbol{p}, \boldsymbol{q}) = \sum_{k=2}^{P} \left[\frac{{p_k}^2}{2m} + V(q_k)\right] + \sum_{k=1}^{P} \left[\frac{m}{2}\left(\frac{P}{\beta\hbar}\right)^2 (q_k - q_{k+1})^2\right]. \tag{5.22}$$

The system Hamiltonian is same as that of a classical particle of mass m and potential $V(q)$ as shown in Eq. (5.21). The bath Hamiltonian, however, has more involved interaction scheme. As shown in Eq. (5.22), the bath Hamiltonian describes a set of P – 1 particles as shown in the first summation term, also with mass m and potential $V(q)$, but with additional harmonic coupling with each other and to the system coordinate as shown in the second summation.

The first step in the mapping process is modifying the bath Hamiltonian to be a set of P – 1 harmonic oscillators. However, we see that from Eq. (5.22) that the bath coordinates, due to the potential term ($\sum_{k=2}^{P} V(q_k)$) which may not necessarily be harmonic, cannot directly be

assumed to be a set of harmonic oscillators. Therefore, we Taylor expand the bath potential ($V_{bath} = \sum_{k=2}^{P}[V(q_k)] + \sum_{k=1}^{P}\left[\frac{m}{2}\left(\frac{P}{\beta\hbar}\right)^2 (q_k - q_{k+1})^2\right]$) upto second order in order to make all the interactions present in it harmonic. Let the potential satisfy $V'(0) = 0$. Then, we can Taylor expand V_{bath} with respect to $\boldsymbol{q} = \mathbf{0}$ (where $\mathbf{0}$ is a P dimensional column of zeros) upto second order to have

$$V_{bath}(\boldsymbol{q}) \cong V_{bath}(\mathbf{0}) + \sum_{k=1}^{P} q_k \left.\frac{\partial V_{bath}}{\partial q_k}\right|_{\{q_k=0\}} + \frac{1}{2}\sum_{k=1}^{P}\sum_{l=1}^{P} q_k \left.\frac{\partial^2 V_{bath}}{\partial q_k \partial q_l}\right|_{\substack{\{q_k=0\} \\ \{q_l=0\}}} q_l. \quad (5.23)$$

Note that the first two terms of the expansion can be omitted: the first term is a constant, and the second term goes to zero as we defined the derivative of the bath potential evaluated at zero to be zero. Therefore, Eq. (5.23) can be rewritten in a more compact way by defining the P dimensional square Hessian matrix $\boldsymbol{H}$ whose kth row and lth column matrix element can be written as

$$H_{kl} = \left.\frac{\partial^2 V_{bath}}{\partial q_k \partial q_l}\right|_{\substack{\{q_k=0\} \\ \{q_l=0\}}}. \quad (5.24)$$

Using this notation, the second order term of the bath potential can be written in the vector and matrix notation as

$$V_{bath}(\boldsymbol{q}) = \frac{1}{2}\boldsymbol{q}^T \boldsymbol{H} \boldsymbol{q}. \quad (5.25)$$

In the current form, by omitting the third and higher polynomial terms, we can understand the P coordinates in the bath potential to experience only harmonic interactions. However, since the goal is to divide the entire ring polymer Hamiltonian into distinct terms that contain the system and bath coordinates, we divide Eq. (5.25) into three parts.

$$V_{bath}(\boldsymbol{q}) = H_{11}{q_1}^2 + q_1 \sum_{k=2}^{P} q_k H_{1k} + \frac{1}{2}\sum_{k=2}^{P}\sum_{l=2}^{P} q_k H_{kl} q_l \quad (5.26)$$

Note that the P^2 terms in Eq. (5.25) have been divided into three parts in Eq. (5.26), with 1, 2(P – 1), and $(P-1)^2$ terms, respectively. The first term (${H_{11}q_1}^2$) is a function of only the system coordinate and therefore can be added into the system Hamiltonian. The second term ($q_1 \sum_{k=2}^{P} q_k H_{1k}$) suggests that the second order Taylor approximation to the bath Hamiltonian gives bilinear coupling between the system and bath coordinates. Lastly, the third term ($\frac{1}{2}\sum_{k=2}^{P}\sum_{l=2}^{P} q_k H_{kl} q_l$) suggests that by separating the first two terms, the P – 1 bath coordinates are now effectively harmonic without external interaction to the system coordinate. To finish the mapping, we diagonalize the third term to have

$$V_{bath}(\boldsymbol{q}) = H_{11}{q_1}^2 + q_1 \sum_{k=1}^{P-1} g_k x_k + \frac{1}{2}\sum_{k=1}^{P-1} m{\omega_k}^2 {x_k}^2. \tag{5.27}$$

For the P – 1 dimensional square truncated Hessian matrix $\widetilde{\boldsymbol{H}}$ obtained by omitting the first row and column of $\boldsymbol{H}$, the orthogonal matrix $\boldsymbol{R}$ satisfies $\widetilde{\boldsymbol{H}}_{diag} = \boldsymbol{R}^T \widetilde{\boldsymbol{H}} \boldsymbol{R}$ where $\widetilde{\boldsymbol{H}}_{diag}$ is the diagonalized $\widetilde{\boldsymbol{H}}$ whose kth diagonal element or eigenvalue is $m{\omega_k}^2$. The new bath coordinate x is defined by the relation $q_{l+1} = \sum_{k=1}^{P-1} R_{lk} x_k$. Lastly, the bilinear coupling constant g is given by $g_a = \sum_{b=1}^{P-1} H_{1,b+1} R_{ba}$.

Using the expression for the new bath potential as shown in Eq. (5.27), the total system-bath Hamiltonian now can be written as

$$H_{P,system-bath} = \left\{\frac{{p_1}^2}{2m} + V(q_1) + H_{11}{q_1}^2\right\}$$
$$+ \{q_1 \sum_{k=1}^{P-1} g_k x_k\} + \left\{\sum_{k=1}^{P-1} \frac{{p_k}^2}{2m} + \frac{1}{2} m{\omega_k}^2 {x_k}^2\right\}. \tag{5.28}$$

Several aspects of the system-bath Hamiltonian should be discussed before we derive the GLE using this system. First of all, the system and bath coordinates (q_1 and $x_k, k = 1, \ldots, P-1$) are distinct from the normal mode coordinates. The system and bath coordinates are obtained by

diagonalizing the truncated Hessian whereas the normal mode coordinates are obtained by diagonalizing the coupling matrix $\boldsymbol{K}$ whose elements are shown in Eq. (5.4).

In addition, it is important to account for how the two types of interaction in the original RPMD Hamiltonian, nearest neighbor coupling and potential for each bead, have been accounted for in the system and bath Hamiltonian. The nearest neighbor harmonic coupling, since it is entirely harmonic, has all been accounted for in the Hessian. As for the physical potential applied to each bead position, the potential applied to the first bead has not been modified and it shows up in the first bracket of Eq. (5.28). The third and higher polynomial potential terms applied to the second to *P*th bead, however, has been truncated in the Taylor expansion process. Therefore, the only difference between the full RPMD Hamiltonian (Eq. (5.3)) and the system and bath Hamiltonian (Eq. (5.28)) are the higher order potential terms applied on the second to *P*th ring polymer beads.

Lastly, we see that aside from mass, physical potential, *P*, the only external parameters that fully describe the system and bath Hamiltonian shown in Eq. (5.28) are the bilinear coupling coefficients g_k and the characteristic bath frequency ω_k. These bath parameters are calculated from the original ring polymer parameters and are complex functions of mass, temperature, *P*, and second order derivative of the potential evaluated at the minima. In the following section, we will see how the GLE derived from equations of motion of the system and bath Hamiltonian are uniquely described by these two sets of parameters.

E. Ring polymer GLE from Laplace transform

Based on the system-bath Hamiltonian shown in Eq. (5.28), the equations of motion for the system and the bath can be written as

$$m\ddot{q}_1 = -\frac{\partial V}{\partial q_1} - 2H_{11}q_1 - \sum_{k=1}^{P-1} g_k x_k \tag{5.29a}$$

$$m\ddot{x}_k = -m{\omega_k}^2 - g_k q_1. \tag{5.29b}$$

We solve the equations of motion using Laplace transform. Solving for $m\ddot{x}_k$ gives

$$x_k(t) = -\frac{g_k}{m{\omega_k}^2} q_1(t) + \left(x_k(0) + \frac{g_k}{m{\omega_k}^2} q_1(0) \right) \cos(\omega_k t)$$

$$+ \frac{\dot{x}_k}{\omega_k} \sin(\omega_k t) + \frac{g_k}{m{\omega_k}^2} \int_0^t d\tau \, \dot{q}_1(t) \cos(\omega_k t). \tag{5.30}$$

Using Eq. (5.30) to solve for $m\ddot{q}_1$ in Eq. (5.29a), we have the generalized Langevin equation for the first bead of the ring polymer

$$m\ddot{q}_1 = -\frac{d}{dq_1} W(q_1) - \int_0^t d\tau \, \dot{q}_1(t) \xi(t-\tau) + R(t) \tag{5.31}$$

Where the system PMF $W(u_1)$, friction kernel $\xi(t)$, and random force $R(t)$ are defined as

$$W(q_1) = V(q_1) - \frac{1}{2} \left[\left(\sum_{k=1}^{P-1} \frac{{g_k}^2}{m{\omega_k}^2} \right) - 2H_{11} \right] {q_1}^2 \tag{5.32a}$$

$$\xi(t) = \sum_{k=1}^{P-1} \frac{{g_k}^2}{m{\omega_k}^2} \cos(\omega_k t) \tag{5.32b}$$

$$R(t) = -\sum_{k=1}^{P-1} g_k \left[\left(x_k(0) + \frac{g_k}{m{\omega_k}^2} q_1(0) \right) \cos(\omega_k t) + \frac{\dot{x}_k}{m\omega_k} \sin(\omega_k t) \right]. \tag{5.32c}$$

For the system and bath Hamiltonian described in Eq. (5.28), the time evolution of the system coordinate q_1 is entirely described by the generalized Langevin equation as shown in Eq. (5.31). Note that there are three aspects of the equation as shown in Eqs. (5.32a), (5.32b), and (5.32c). The first term describes the effective PMF that the system coordinate experiences. Note that it includes the contribution from the physical potential, $V(q_1)$, as well as harmonic additional term with harmonic strength of $2H_{11} - \left(\sum_{k=1}^{P-1} \frac{{g_k}^2}{m{\omega_k}^2} \right)$. We can understand the additional harmonic term to arise from the harmonic nearest neighbor coupling. The second term, $\xi(t)$, is the friction kernel. The convolution term in Eq. (5.31) arises from the Laplace transform solution of the bilinear coupling between the system and the bath coordinates. We see

that this term is a complete sum of cosine function with different magnitude and frequencies for each bath coordinate, which are all determined by the bath parameters. Lastly, the $R(t)$ term corresponds to the random force. The random force is similar in construct to the friction kernel, but it has a contribution from the sine functions. Also, the magnitudes of the oscillation is not only a function of bath parameters but also initial positions of the system and bath position and momenta. The random force also is related to the friction kernel via the famous second fluctuation dissipation theorem which can eloquently be stated as

$$\langle R(0)R(t)\rangle = \xi(t)/\beta. \tag{5.33}$$

Few aspects of the theory should be clarified before further developments of the GLE picture and numerical examples are presented. Within the level of approximation that have been made, Eq. (5.31) completely describes the dynamics of the system coordinate. However, even though the derivation of GLE and its physical implications are conceptually fascinating, in general the numerical integration of GLE to directly calculate the dynamics of the system is challenging. The derivations and the numerical results of the bath parameters and GLE quantities shown later should be taken to be a phenomenological theory in which we can derive the parameters from the original RPMD system. Alternatively, rather than numerically integrating GLE, it is common to simply simulate the entire system rather than explicitly calculating the bath parameters for larger systems, and this is what we will do in later sections We simply sill propagate the RPMD Hamiltonian with truncated physical potential applied to the second to *P*th ring polymer bead positions. Truncating the potential for the beads in the regular RPMD Hamiltonian gives the same effective dynamics as described in Eqs. (5.29a) and (5.29b). For the rest of the chapter, the theory described in this section will be referred to as "GLE/RPMD theory" or "GLE picture of RPMD dynamics."

F. Modeling TRPMD with GLE

In this section, we will construct a GLE picture that models the TRPMD dynamics. As shown in Eqs. (5.17a) and (5.17b), the TRPMD dynamics is described by the Langevin dynamics with friction and random force contribution to the force. This suggests that if we want to derive the GLE picture of TRPMD dynamics, the system and bath coordinates of the Hamiltonian shown in Eq. (5.28) need to be propagated using Langevin dynamics. However, this suggests that we need to Laplace transform the time dependent random force $\boldsymbol{\zeta}(t)$ which is both conceptually baffling and outside the scope of this work.

An alternative that we propose in this work is simply to add an exponential damping term to the friction kernel and random force. The modified functions are

$$\xi(t) = \sum_{a=1}^{P-1} \frac{{g_a}^2}{m{\omega_a}^2} \cos(\omega_a t)\, e^{-\kappa\omega_a t/2} \tag{5.34a}$$

$$R(t) = -\sum_{a=1}^{P-1} g_a \left[\begin{matrix} \left(x_a(0) + \frac{g_a}{m{\omega_a}^2} u_1(0) \right) \cos(\omega_a t) \\ + \frac{\dot{x}_a}{\omega_a} \sin(\omega_a t) \end{matrix} \right] e^{-\kappa\omega_a t/2} \tag{5.34b}$$

Note that the two functions shown in Eqs. (5.34a) and (5.34b) largely resemble the original GLE forms shown in Eqs. (5.32b) and (5.32c). The only difference is the additional damping term $e^{-\kappa\omega_a t/2}$ parametrized by a single unitless parameter κ. Note that each bath mode gets damped to the same degree as the damping term is a function of the individual bath frequency. The GLE parameter κ can be understood to be this theory's analogue to the TRPMD's λ. Lastly, the modified friction kernel and random force satisfies the second FDT as we only add the simple exponential term in the functional form. Later in the numerical example section, we see how the damping parameter κ affects the overall shape and behavior of the calculated GLE quantities.

Similar to the GLE/RPMD theory, for the TRPMD version of the GLE picture, we will propagate the truncated RPMD Hamiltonian using Langevin dynamics EOM (Eqs. (5.17a) and (5.17b)) in the numerical example section. For the remainder of this work, the theory described in this section will be named "GLE/TRPMD theory" or "GLE picture of TRPMD dynamics."

G. Liouvillian description using system and bath Hamiltonians

Before we move on to test the validity of the GLE/RPMD theory, we present the rigorous derivation of the ring polymer GLE using Liouvillian operator formalism. The original and rigorous derivation of the GLE picture comes from the projection operator formalism.[52] In this section, we do not use the exact projection operator formalism but use operator calculus and algebra. We will see how both the original GLE quantities and the second FDT arise in the operator equivalent. Moreover, we will also see how the additional commutator operators comes up in the equation due to non-commutivity between the system Hamiltonian Liouvillian and the bath Hamiltonian Liouvillian.

We start by considering the classical time propagator e^{iLt} where iL is the classical Liouvillian operator associated with the ring polymer Hamiltonian as shown in Eq. (5.3). As shown in Eqs. (5.21) and (5.22), the RPMD Hamiltonian can be divided into two contributions, system and bath. Similarly, we can divide the Liouvillian into two contributions, from the system and the bath Hamiltonians.

$$iL = iL_s + iL_b \tag{5.35}$$

In Eq. (5.35), iL_s is the Liouvillian associated with system Hamiltonian and iL_b is the Liouvillian associated with bath Hamiltonian.

For the system variable vector $\boldsymbol{A} = \begin{pmatrix} q_1 \\ p_1 \end{pmatrix}$, its time evolution can be expressed as

$$\boldsymbol{A}(t) = e^{iLt}\boldsymbol{A}(0). \tag{5.36}$$

We can take a derivative of both sides with respect to time to have

$$\frac{d\boldsymbol{A}}{dt} = e^{iLt} iL\boldsymbol{A}(0) = e^{iLt}(iL_s + iL_b)\boldsymbol{A}(0). \tag{5.37}$$

In the time propagation of the zero time system variable vector $\boldsymbol{A}(0)$ there are two contributions, from the system Liouvillian and the bath Liouvillian. We solve for the first term to have

$$e^{iLt} iL_s\boldsymbol{A}(0) = iL_s\boldsymbol{A}(t) - \left[e^{iLt}, iL_s\right]\boldsymbol{A}(0). \tag{5.38}$$

Note that the commutator term arises because the two operators iL and iL_s do not commute. The second term can be solved using the Laplace identity such that

$$e^{iLt} iL_b\boldsymbol{A}(0) = e^{iL_b t} iL_b\boldsymbol{A}(0) + \int_0^t d\tau\, e^{iL(t-\tau)} iL_s e^{iL_b\tau} iL_b\boldsymbol{A}(0) \tag{5.39}$$

The integrand in Eq. (5.39) can be written as

$$e^{iL(t-\tau)} iL_s e^{iL_b\tau} iL_b\boldsymbol{A}(0) = -B(\tau)B(0)\boldsymbol{A}(t-\tau) + C(t,\tau)\boldsymbol{A}(0) \tag{5.40}$$

where $B(t)$ and $C(t,\tau)$ are operators defined as

$$B(t) = e^{iL_b t} iL_b \tag{5.41a}$$

$$C(t,\tau) = e^{iL(t-\tau)} iLB(\tau) + \left[e^{iL(t-\tau)}, B(\tau)B(0)\right]. \tag{5.41b}$$

We can then combine Eqs. (5.37) – (5.40) to have

$$\frac{d\boldsymbol{A}}{dt} = iL_s\boldsymbol{A}(t) - \int_0^t d\tau\, B(\tau)B(0)\boldsymbol{A}(t-\tau) + B(t)\boldsymbol{A}(0)$$

$$-\left[e^{iLt}, iL_s\right]\boldsymbol{A}(0) - \int_0^t d\tau C(t,\tau)\boldsymbol{A}(0). \tag{5.42}$$

Eq. (5.42) shows the time evolution of the system vector $\boldsymbol{A}$ analogous to GLE form as shown in Eq. (5.19). Few things to note in the operator formalism is that there should be a clear distinction between e^{iLt}, $e^{iL_s t}$, and $e^{iL_b t}$ due to commutivity between iL_s and iL_b. The true and full time evolution of the entire ring polymer is only given by e^{iLt} , and

both $e^{iL_s t}$ and $e^{iL_b t}$ should be understood to be partial time evolution associated with system and bath Hamiltonians.

The first term of Eq. (5.42), $iL_s \boldsymbol{A}(t)$, shows the contribution of the system Hamiltonian on the full propagation by the full Liouvillian. This term is analogous to the system PMF as shown in Eq. (5.31). The third term, $B(t)\boldsymbol{A}(0)$, shows the effect of the bath Hamiltonian on the partial propagation of $\boldsymbol{A}(0)$, analogous to the random force. Note that based on the definition of $B(t)$ we are propagating $\boldsymbol{A}(0)$ but not with respect to the full Liouvillian. The second convolution term, $\int_0^t d\tau\, B(\tau)B(0)\boldsymbol{A}(t-\tau)$, shows the effect of the bath Hamiltonian on the full propagation and is analogous to the friction term in the originally stated GLE. Note that the similar operator defined in the third term, B, shows up in the integrand of the convolution term. We see that this is the equivalent second FDT that manifests itself in the operator formalism of ring polymer dynamics. Lastly, the fourth and fifth terms arise from commutation relations between the system and the bath Liouvillian and encode higher order interactions between the two.

Note that the quantities in Eq. (5.42) that correspond to regular functions in Eq. (5.28) are operators that act on the initial condition of the system variable $\boldsymbol{A}(0)$. Moreover, since this is an operator formalism, the second FDT shows up as the memory kernel operator being $B(\tau)B(0)$.

Even though the operator formalism shed light on how the system and bath Hamiltonian contributions are related to each other in time evolution of the system coordinate, the numerical implementation of operator algebra is not straightforward. Therefore as mentioned before, the numerical results shown in the next section only uses the previous

formalism developed in which the ring polymer Hamiltonian with truncated third and higher order potential terms. This operator derivations, for now, remain a theoretical exercise.

III. Numerical Examples

To analyze the dynamics of the methods discussed in previous sections, the Kubo-transformed position autocorrelation function for the one dimensional anharmonic potential of form

$$V(x) = 0.5\,x^2 + 0.1\,x^3 + 0.01\,x^4 \tag{5.43}$$

and quartic potential of form

$$V(x) = 0.25\,x^4 \tag{5.44}$$

with $m = 1$, $\hbar = 1$, and at two different temperatures, β = 8 and 1, are presented. The potentials shown in Eqs. (5.43) and (5.44) have been widely studied as simple one-dimensional models to compare the performances of different approximate quantum dynamics methods such as CMD and RPMD to that of the numerically exact quantum result.[9, 12]

A. Bath parameters

We have calculated the bath parameters for the ring polymer with three different parameters: temperature (β), number of beads (P), and the quadratic constant (harmonic spring constant, k) of the physical potential. Note that regardless of the form of the potential, the only aspect that affects the bath parameters are the second order Taylor expansion coefficients. Moreover, for convenience, the bath parameters are plotted as a function of bath index. The baths were indexed in two different ways: from low to high frequency when plotting bath frequency vs. bath index and from low to high coupling coefficient when plotting bath

magnitude vs. bath index. This was to better understand the behavior of the bath parameters and for clearer representation of the plot.

Figure 5-1 show how both the bath frequency and coupling magnitude change with respect to temperature. We see that both frequency and magnitude increase with temperature. Moreover, we see that in Figure 5-1(b) that the distribution of the coupling magnitude is roughly symmetric in which similar number of bath modes have coupling magnitude of similar degree but with opposite sign. Interestingly, we also observe that about half of the coupling magnitude is zero, which means that these bath modes with zero magnitude comes up as the truncated Hessian is diagonalized but are not effectively interacting with the system coordinate.

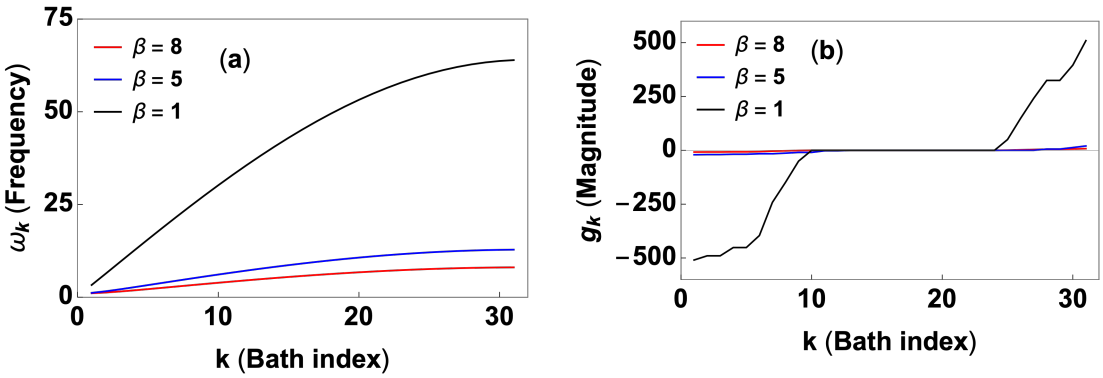

Figure 5-1. Bath frequency and coupling magnitude with $P = 32$ and $k = 0.5$ for different temperatures plotted against the bath index. $\beta = 8$ results are plotted in red, $\beta = 5$ results are plotted in blue, and $\beta = 1$ results are plotted in black. For (a), the bath modes are indexed from low to high frequency. For (b), the bath modes are indexed from low to high magnitude.

Another important variable to consider is the total number of beads in the ring polymer, *P*. This not only determines the number of bath coordinates but also the value and size of the truncated Hessian and hence the bath parameters directly. Figure 5-2 shows how the bath parameters are affected by the number of beads. Note that the index of the $P = 64$ results were

compressed for direct comparison with the P = 32 result. We see that again, similar to the temperature, both the bath frequencies and magnitude increase with increasing P. Note that for both curves in Figure 5-2(a), the low index region contains slight curve which was not visible in Figure 5-1(a). This most likely is the effect of the temperature, which also suggests that the temperature affects the frequency distribution of the low frequency bath modes. Figure 5-2(b) shows the coupling magnitude. Note that aside from the increase of the overall magnitude of the curve, there are shift in the flat region of the curve, suggesting that for the larger P system, higher indexed bath have zero magnitudes.

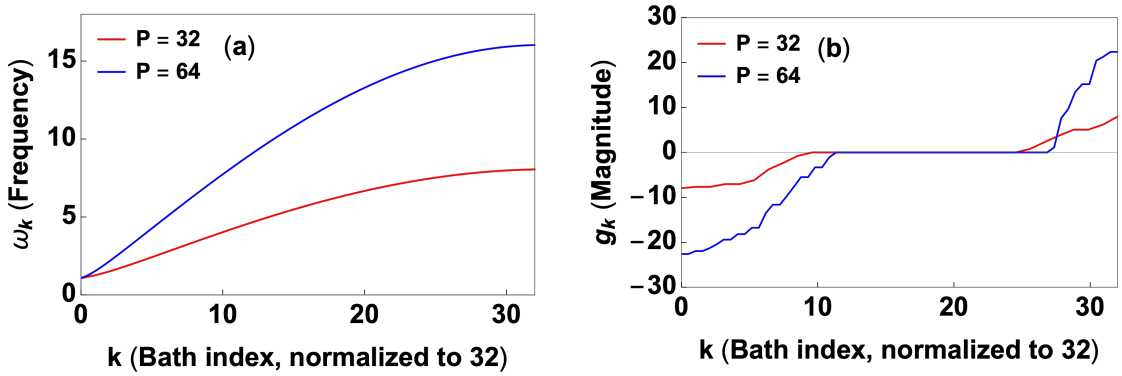

Figure 5-2. Bath frequency and coupling magnitude with $\beta = 8$ and $k = 0.5$ for different number of beads plotted against the bath index. P = 64 results are plotted in blue and P = 32 results are plotted in red. For (a), the bath modes are indexed from low to high frequency. For (b), the bath modes are indexed from low to high magnitude.

The last variable to consider that affects the bath parameters is the harmonic constant of the potential. Figure 5-3 shows the effect of the harmonic strength in the bath parameters. Note that $k = 0.5$ casecorresponds to a system that contains the $0.5\,x^2$ term in its potential whereas the $k = 0$ case corresponds to a system that does not have any quadratic contribution in the potential. We see that there are little differences between the two. As seen in Figure 5-3(a), the

overall trend of the two curves correspond to each other. However, we see that the main difference occurs in the low index region in which the $k = 0.5$ case has a higher lowest bound of frequency. Moreover, Figure 5-3(b) show that the coupling magnitude plotted against the bath index are the same, which suggests that the harmonic strength of the potential does not affect the coupling magnitude.

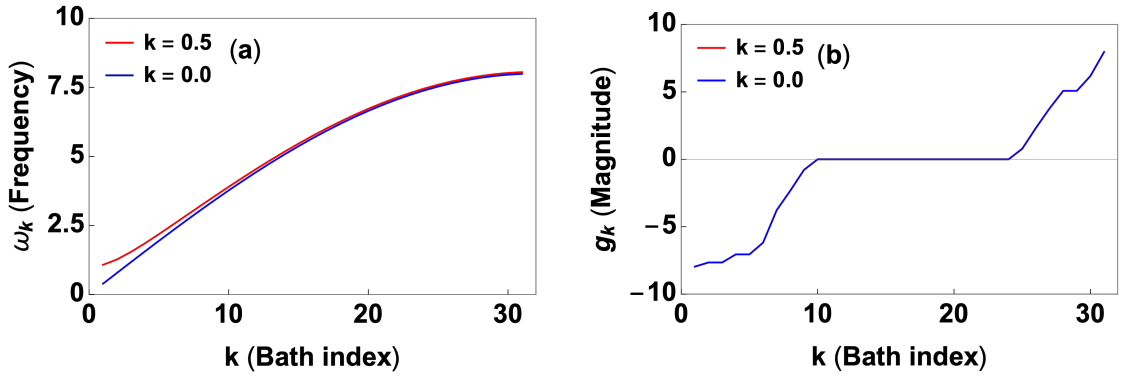

Figure 5-3. Bath frequency and coupling magnitude with $P = 32$ and $\beta = 8$ for different values of quadratic constants plotted against the bath index. $k = 0.5$ results are plotted in red and $k = 0$ results are plotted in blue. For (a), the bath modes are indexed from low to high frequency. For (b), the bath modes are indexed from low to high magnitude.

B. Friction kernel and random force

Using the calculated bath parameters presented in the previous section and Eqs. (5.32b) and (5.32c), we can plot analytically defined friction kernel and random force. Similar to how the previous section was organized, we will observe how temperature, number of ring polymer beads, and quadratic term of the potential affect both GLE quantities. However, it should be noted that we will only present the plots of the random force for varying temperature as the trend observed in friction kernel and random force are similar and it is hard to understand the behavior of the random force aside from its magnitude.

Figure 5-4 shows two friction kernels with different temperatures $\beta = 8$ and 5. We see that the overall magnitude of the oscillation increase with increasing temperature. This trend is consistent with what was observed in Figure 5-1(b). Moreover, we also see that the frequency in the friction kernel increase with respect to increasing temperature as well and observe more oscillations within the same time.

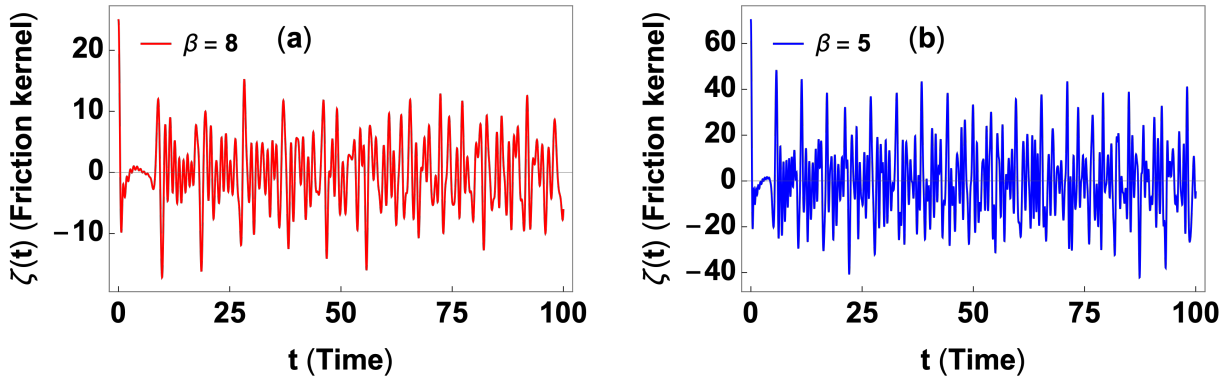

Figure 5-4. Two friction kernels with $P = 32$ and $k = 0.5$ and different temperatures plotted. (a) Low temperature case with $\beta = 8$ shown in red and (b) high temperature case with $\beta = 5$ shown in blue.

Using Eq. (5.32c), we can calculate the random force term for a predefined bath parameters. Initial value terms in the expression of random force was sampled according to the distribution determined by the system and bath Hamiltonian. Figure 5-5 shows how the random forces differ for different temperature. Similar to the trend observed in Figure 5-1 and Figure 5-4, we see that the overall magnitude of the random force increase with temperature. However, unlike the friction kernel, it is hard to tell what the characteristic frequency of the random force is, which is what we would expect from truly random force, even though the functions presented are deterministic. Instead, the only quantity that can be distinguished are the magnitudes of the fluctuation from the zero mean.

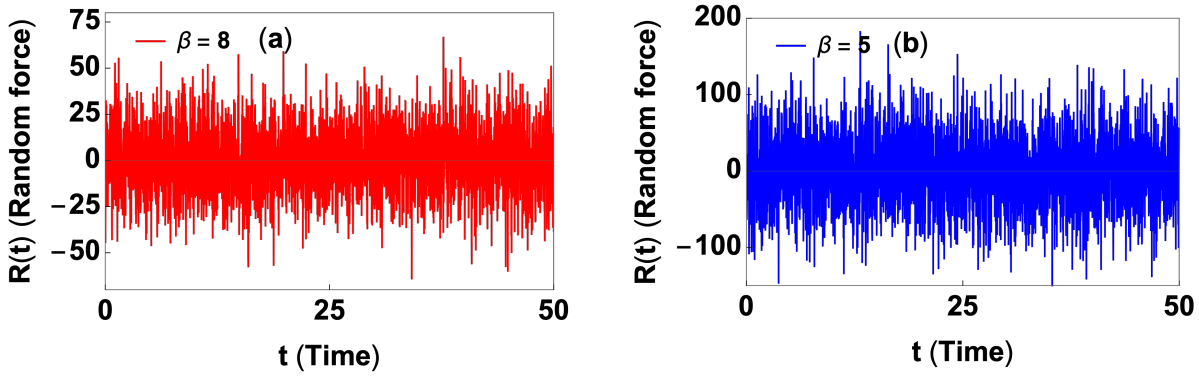

Figure 5-5. Two random forces with P = 32 and $k = 0.5$ and different temperatures plotted. (a) Low temperature case with $\beta = 8$ shown in red and (b) high temperature case with $\beta = 5$ shown in blue.

The next variable to consider is the number of beads. The two friction kernel calculated with different number of beads are shown in Figure 5-6. We see that, consistent with previous observation as shown in Figure 5-2, the overall magnitude of the friction kernel increases with number of beads. Moreover, since the P = 64 result has more number of beads, it has more oscillations with higher frequency. We see this by observing additional high frequency oscillations of the P = 64 case compared to the P = 32 one. However, we also observe that the characteristic low frequency oscillation of the P = 32 result is preserved in the P = 64 one. This suggests that even though increasing the number of beads adds additional bath coordinates with higher frequency, the two systems have similar characteristic frequencies for the low frequency bath modes.

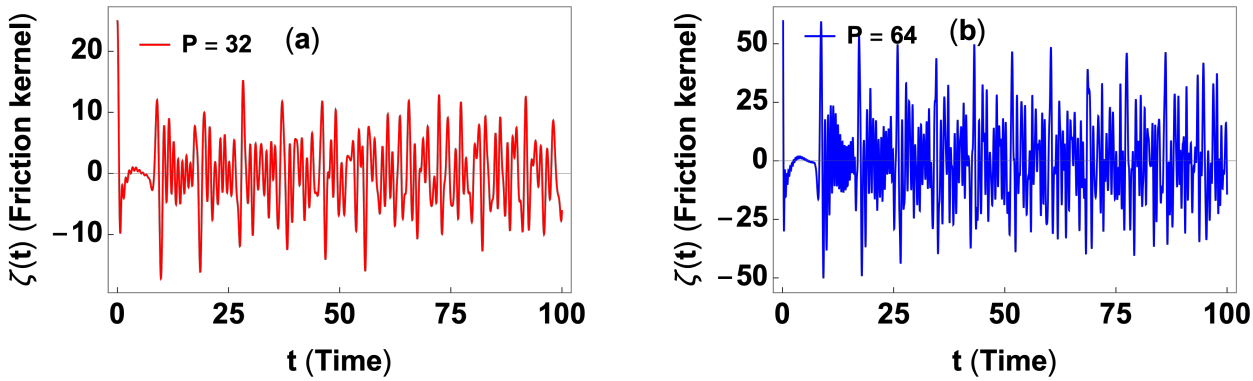

Figure 5-6. Two friction kernels with $\beta = 8$ and $k = 0.5$ and different number of ring polymer beads plotted. (a) $P = 32$ case in red and (b) $P = 64$ case in blue.

Figure 5-7 shows how the friction kernel responds to changes in the harmonic strength of the physical potential. Analogous to the trend observed in Figure. 5-3(b), we see that the overall magnitude of the two friction kernels are similar to each other with no distinguishable difference. However, we do see that the specific oscillation profile of the two still differ. Qualitatively, we see that the $k = 0.5$ results have more high frequency oscillations. This can be explained by increase in the lower bound of the frequencies of the bath modes as shown in Figure. 5-3(a), which leads to the low frequency modes for the $k = 0.5$ case to have higher frequencies than the low frequency modes for the $k = 0$ case.

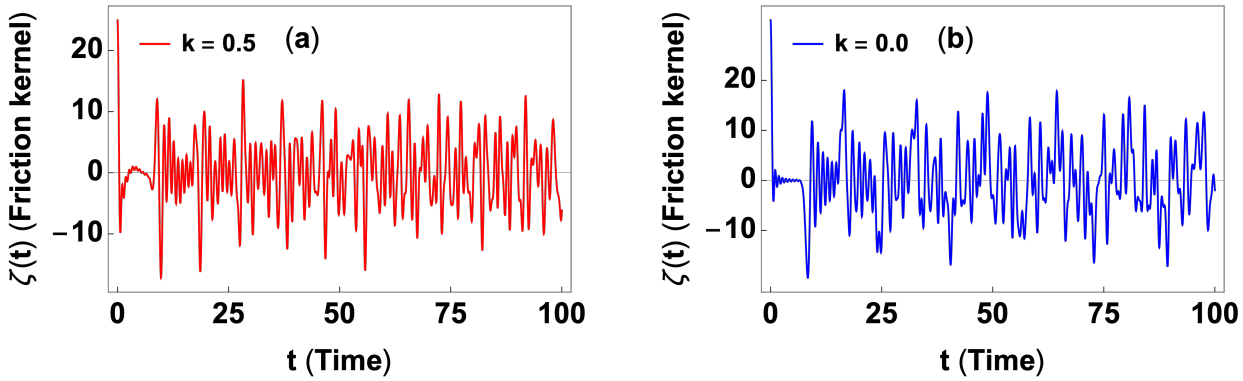

Figure 5-7. Two friction kernels with $\beta = 8$ and $P = 32$ and different harmonic strength plotted. (a) $k = 0.5$ case in red and (b) $k = 0$ case in blue.

Lastly, in Figure 5-8 we see how the damping coefficient κ described in Eqs. (5.34a) and (5.34b) affect the GLE quantities. As expected, with increasing damping parameter, the functions decay faster with respect to time. For the plotted time, the $\kappa = 0$ case, analogous to the GLE quantities shown in Figures 5-4 to 5-7, shows no decay. However, for the $\kappa = 0.5$ case we see that the both functions decay to zero after $t = 10$ whereas the $\kappa = 0.05$ case shows intermediate level of decaying behavior.

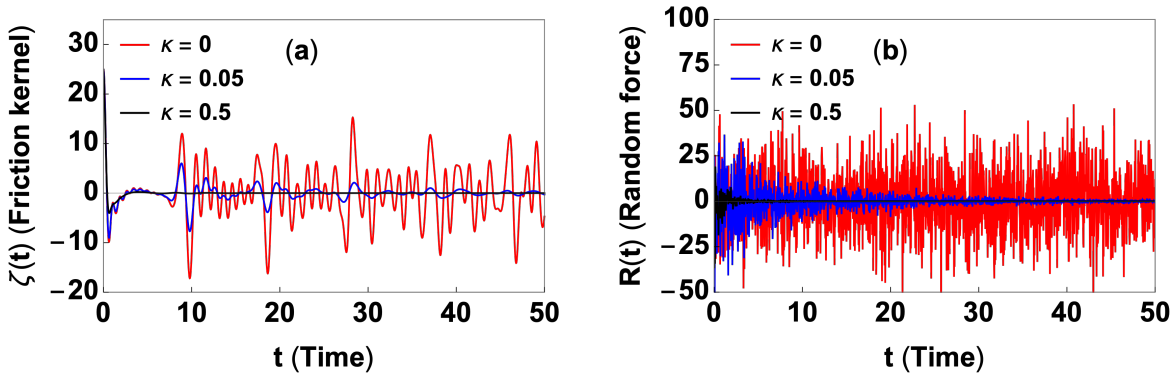

Figure 5-8. (a) Friction kernel and (b) random force, both with $\beta = 8$, $P = 32$, and $k = 0.5$ with different damping parameters plotted. $\kappa = 0$ results are shown in red, $\kappa = 0.05$ results shown in blue, and $\kappa = 0.5$ results shown in black.

C. Position autocorrelation function

In this section, we present the Kubo-transformed position autocorrelation functions calculated from exact, CMD, RPMD, TRPMD, GLE/RPMD, and GLE/TRPMD method. The exact and CMD correlation functions were calculated according to numerical procedures described in Ref. 7, both using an in-house code. The RPMD and TRPMD correlation function was calculated according to numerical procedures described in Ref. 12 using the one-dimensional code provided by M. Rossi. Finally, the GLE/RPMD and GLE/TRPMD calculation details followed those of RPMD and TRPMD. A modified version of the (T)RPMD code was used for GLE calculations.

Figure 5-9 shows Kubo-transformed position autocorrelation function for anharmonic potential at $\beta = 8$. For clarity, any 'correlation function' mentioned for the rest of this chapter refers to Kubo-transformed correlation functions. First of all, the exact quantum result is completely oscillatory in which the frequencies of the oscillations are determined by the thermally weighed eigenvalue of the Hamiltonian whereas other PI based methods, classical in nature, show classical behavior in which the correlation function decays with respect to time. The degree of decay for the PI methods (CMD, RPMD, and TRPMD) are: CMD < RPMD < TRPMD. Also compared to the exact result, the other PI correlation functions are blue shifted with the degrees of shift from low to high are: CMD < RPMD = TRPMD < GLE/RPMD = GLE/TRPMD. Moreover, we also observe that GLE/RPMD and GLE/TRPMD does not capture the small time limit correctly. This is due to the fact that in deriving the GLE picture of RPMD and TRPMD we have truncated the higher order interactions in the initial Hamiltonian, meaning that we are not completely sampling the entire extended phase space of the ring polymer. However, both the GLE methods provide qualitative agreement of the overall magnitude and

frequency of the oscillation in the correlation function oscillation. We see that in Figure. 5-9(c) there are no frequency difference between the GLE/RPMD and GLE/TRPMD but only a magnitude difference in which the GLE/TRPMD result has larger oscillation than the GLE/RPMD one.

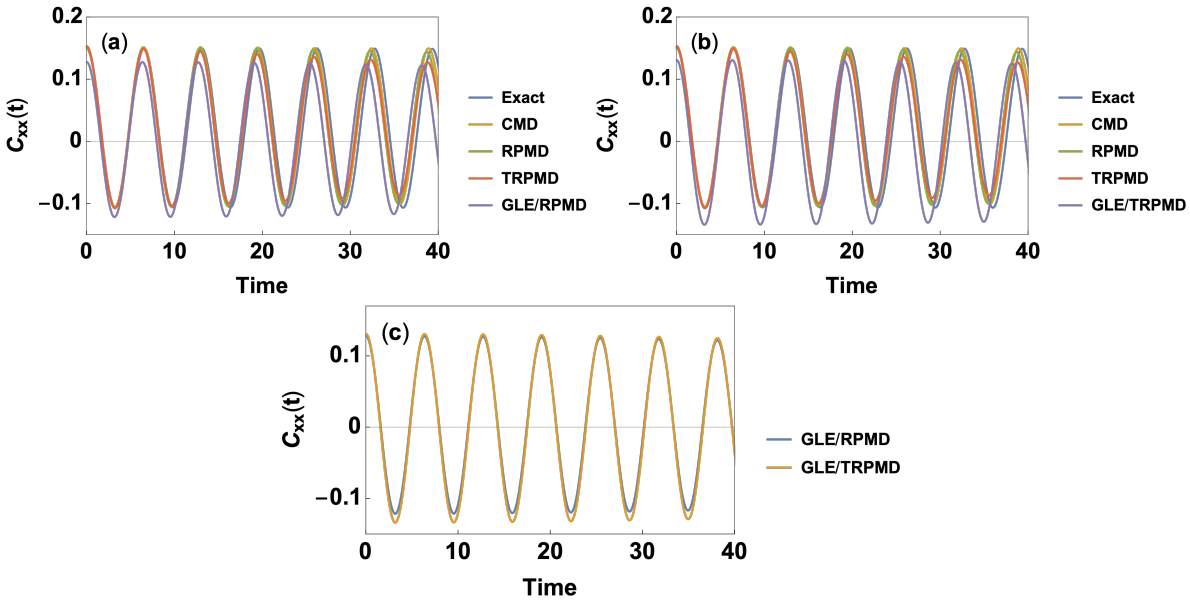

Figure 5-9. Calculated Kubo-transformed position autocorrelation function using different methods for anharmonic potential with $P = 32$ and $\beta = 8$. In (a) and (b), the exact quantum results are shown in blue, CMD results are shown in yellow, RPMD results are shown in green, TRPMD results are shown in crimson. GLE/RPMD and GLE/TRPMD results are shown in purple in (a) and (b) respectively. Figure 5-9(c) compares the GLE/RPMD and GLE/TRPMD results which are shown in blue and yellow respectively.

Figure 5-10 shows the position autocorrelation function for anharmonic potential at $\beta = 1$. Compared to the $\beta = 8$ case, we see that the overall agreement between the exact result and the other methods have improved. This is because all the results give the correct classical limit in which for high enough temperature they converge to the classical autocorrelation functions.

However, we see that even though the PI methods well capture the oscillations of the exact result, GLE methods still differ more distinguishably from the exact ones. In this case, the GLE methods capture the correct oscillation frequency but not the magnitude. Again, this is due to the fact that we have omitted parts of the original RPMD Hamiltonian in constructing the GLE picture. We also see that there are no distinguishable difference between RPMD and TRPMD and GLE/RPMD and GLE/TRPMD.

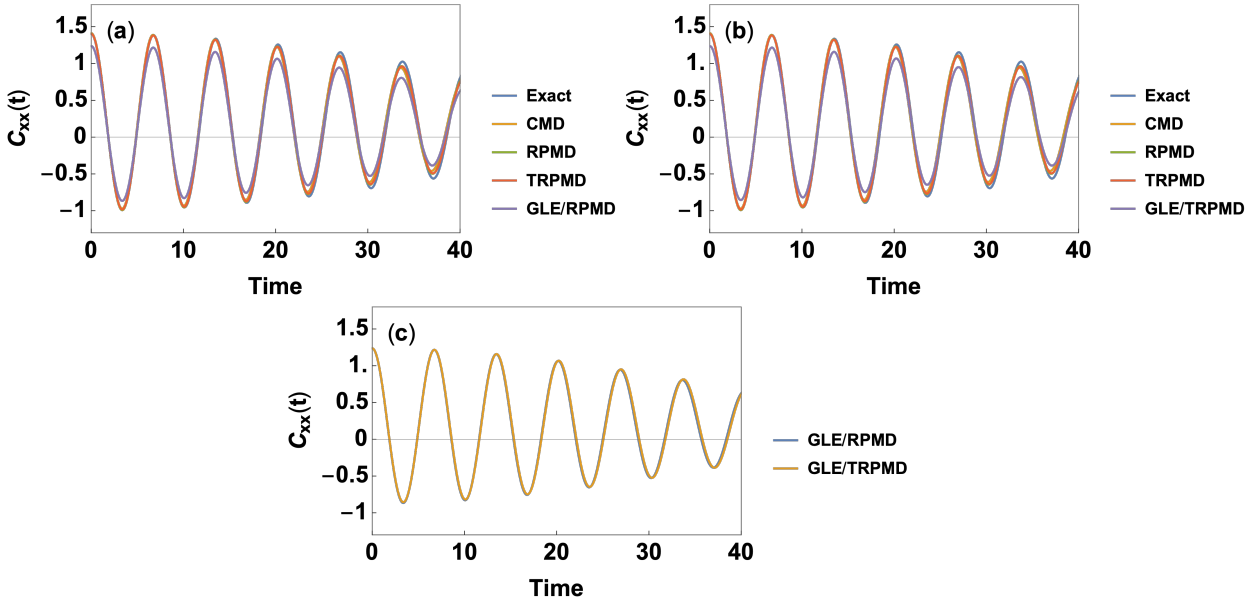

Figure 5-10. Calculated Kubo-transformed position autocorrelation function using different methods for anharmonic potential with $P = 32$ and $\beta = 1$. In (a) and (b), the exact quantum results are shown in blue, CMD results are shown in yellow, RPMD results are shown in green, TRPMD results are shown in crimson. GLE/RPMD and GLE/TRPMD results are shown in purple in (a) and (b) respectively. Figure 5-9(c) compares GLE/RPMD and GLE/TRPMD which are shown in blue and yellow respectively.

It should be noted that compared to the anharmonic potential shown in Eq. (5.43), the quartic potential shown in Eq. (5.44) is in general more challenging as it is strongly anharmonic,

and in the context of the GLE picture, does not have any quadratic contributions. Therefore, the Hessian matrix in the GLE formalism for the quartic potential does not have any contributions from the potential and only contains the nearest neighbor harmonic interactions.

Figure 5-11 shows the position autocorrelation function for quartic potential at $\beta = 8$. We see that between the PI methods, the degree of decay has shifted in which it's CMD < TRPMD < RPMD. All three methods provide good agreement to the frequency of the exact result. However, for the quartic case, both GLE/RPMD and GLE/TRPMD agrees very poorly with the exact result. They not only get the initial value of the correlation function wrong, but also does not capture the correct oscillation frequency. Again, this is due to fact that the quartic potential is strongly anharmonic and has no harmonic contribution. Hence, the bath parameters fail to capture any contribution from the potential in constructing the system and bath Hamiltonian. We see that it is similar, but not exactly the same, as the classical results shown in Figure 2 of Ref. 12. This suggests that in the absence of bath mode interactions that comes from the harmonic term of the potential, the GLE formalism provides similar dynamics to classical mechanics. Lastly, we see that there are slight difference between the GLE/RPMD and GLE/TRPMD correlation function in which the GLE/TRPMD result is slightly shifted to the left.

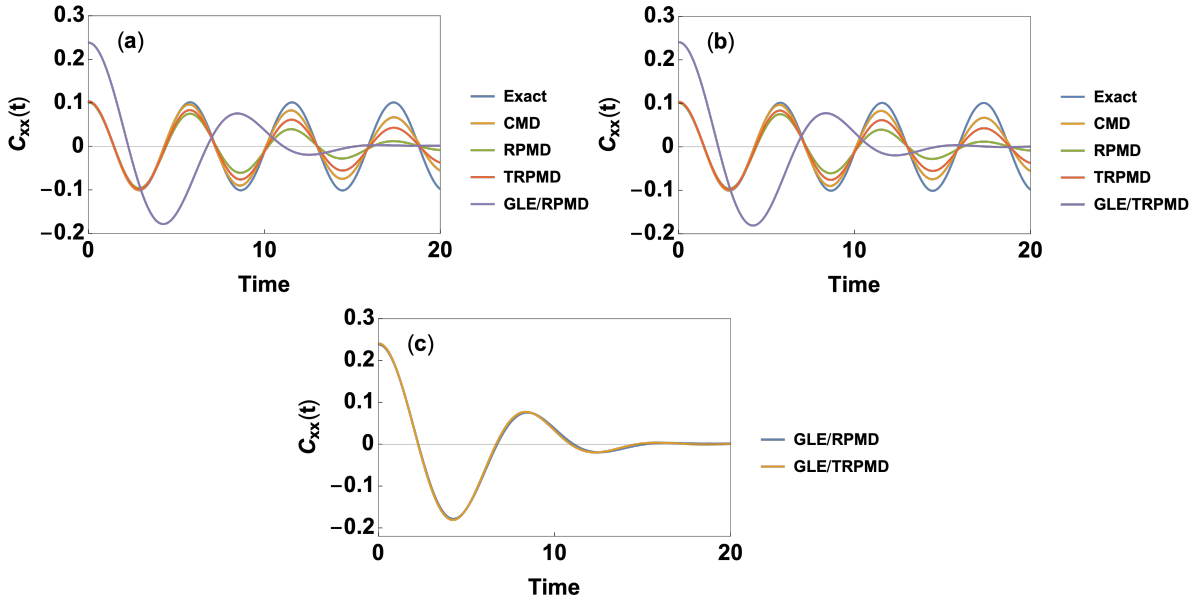

Figure 5-11. Calculated Kubo-transformed position autocorrelation function using different methods for quartic potential with $P = 32$ and $\beta = 8$. In (a) and (b), the exact quantum results are shown in blue, CMD results are shown in yellow, RPMD results are shown in green, TRPMD results are shown in crimson. GLE/RPMD and GLE/TRPMD results are shown in purple in (a) and (b) respectively. Figure 5-11(c) compares GLE/RPMD and GLE/TRPMD which are shown in blue and yellow respectively.

Similar to the trend observed for high temperature case in Figure 5-10, Figure 5-12 shows the position autocorrelation functions for the quartic potential at $\beta = 1$. Compared to Figure 5-11, the agreement between the exact and other methods have improved, but not as good as in the anharmonic case. In this high temperature case, we see that the CMD, RPMD, and TRPMD results only get the frequency of the initial oscillation right but displays a classical dissipation after $t = 5$. Similar trend is observed for both GLE formalism results. However, the GLE methods capture the correct oscillation frequency of the initial oscillation. Additionally, GLE methods also capture correctly the position of the second peak in exact result compared to

the other methods. Lastly, we see that the difference between the two GLE formalism is more pronounced.

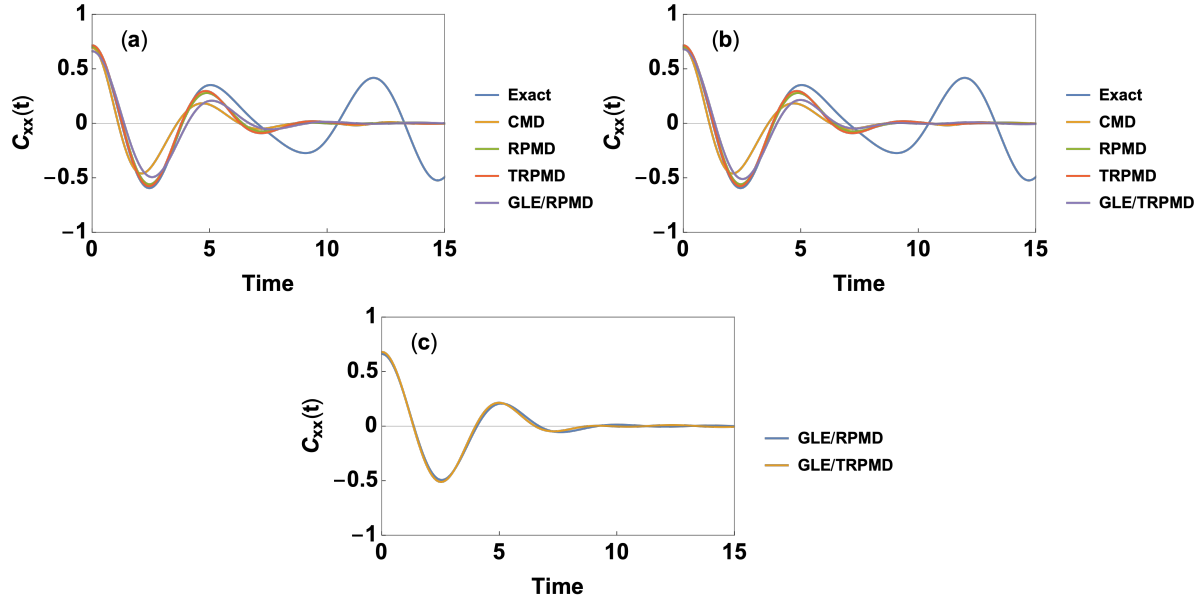

Figure 5-12. Calculated Kubo-transformed position autocorrelation function using different methods for quartic potential with $P = 32$ and $\beta = 1$. In (a) and (b), the exact quantum results are shown in blue, CMD results are shown in yellow, RPMD results are shown in green, TRPMD results are shown in crimson. GLE/RPMD and GLE/TRPMD results are shown in purple in (a) and (b) respectively. Figure 5-12(c) compares GLE/RPMD and GLE/TRPMD which are shown in blue and yellow respectively.

IV. Discussion and Conclusion

The initial motivation for developing the GLE/RPMD theory is based on the observation that in the original RPMD formulation the expectation values of the system are taken with respect to the single bead coordinate. This inspires us to only focus on the single bead rather than the entire ring polymer. It also intrigues the use of the classical motif of GLE in which the entire set of DOFs is divided into two parts: the system (single DOF of interest) and

the bath (the other DOFs). In our case, the single (first) bead of the ring polymer is set to be the system and the P-1 other beads correspond to the bath. We use the GLE formalism to analyze the RPMD dynamics so that the different contributions from the system and the bath can be analyzed.

To use the GLE formalism on the RPMD system, we start with the original RPMD Hamiltonian. The mapping process has two steps: a) truncating the RPMD Hamiltonian to the system and bath Hamiltonian and b) deriving GLE from the Hamiltonian EOM of the system and bath Hamiltonian. In the first step, we Taylor expand the bath Hamiltonian which contains all the nearest neighbor harmonic coupling and physical potential applied to the 2 to Pth bead coordinates and omit the third and higher order polynomial terms. In this process, we fully capture the nearest neighbor coupling as they are entirely harmonic. Then we diagonalize the truncated Hessian matrix of the bath Hamiltonian to project the $P - 1$ coordinates of the 2 to Pth beads in the ring polymer to $P - 1$ number of bath coordinates which can be considered to be a set of harmonic oscillators with characteristic frequency that corresponds to the eigenvalue of the diagonalized truncated Hessian. We also calculate the coupling magnitudes between the system and the bath coordinates. The two sets of bath parameters, characteristic bath frequency and the coupling magnitude, entirely characterize the bath. The next step in the GLE/RPMD theory is deriving the GLE from the system and bath Hamiltonian. The Hamiltonian EOM of the system and bath is analytically solved using Laplace transform. In this context, GLE is an exact solution to the system dynamics. The resulting GLE, similar to the classical form, is characterized by three components: system PMF, friction kernel, and random force, which themselves are determined by the bath parameters. Moreover, we can also model the TRPMD using the GLE picture by including an arbitrary damping term in both the friction kernel and the

random force. This ad hoc measure is an alternative to a rigorous derivation of the TRPMD version of the theory as calculating the Laplace transform of a random force term (originally from the TRPMD dynamics) is beyond the scope of this work.

It should be noted that even though the GLE picture of both RPMD and TRPMD provides how the system and bath mapped RPMD Hamiltonian can be divided into two contributions, when performing numerical calculations, we did not directly integrate the GLE as it is numerically challenging. Instead, we simulated a modified version of the RPMD Hamiltonian in which the third an higher order terms are omitted. This form effectively samples the system and bath Hamiltonian as shown in Eq. (5.28). For the GLE/TRPMD, we used the same truncated potential as in numerical GLE/RPMD simulation but propagated it with Langevin equation dynamics shown in Eqs. (5.17a) and (5.17b) using damping parameter of $\lambda = 0.5$.

The bath parameters, the bath frequencies and bilinear coupling constants, are determined as we diagonalize the bath Hamiltonian Hessian evaluated at the minimum. They are also functions of mass, temperature, *P*, and harmonic term of the physical potential. This contrasts to the nearest neighbor harmonic coupling strength in the RPMD Hamiltonian, which encodes the quantum delocalization of the particle and is a function of mass, temperature, and *P*. Note that the nearest neighbor harmonic strength is not a function of potential because all the bead coordinates directly experience the physical potential through a separate term.

We have presented the calculated bath parameters with varying temperature, P, and harmonic strength as shown in Figure 5-1, 5-2, and 5-3. We observed that all three variables affect the distribution of bath parameters in different ways. First, the overall magnitude of the parameters increased with increasing temperature. More interestingly, with decreasing

temperature, larger number of the bath modes had low characteristic frequency compared to the high temperature case. The overall magnitude of the parameters also increased with increasing values of P. However, increasing the value of P displayed a shift in the index of the bath modes with zero coupling magnitudes. The harmonic strength did not drastically change the bath parameter as temperature and P. In particular, varying the harmonic strength did not affect the coupling magnitude. However, increasing harmonic strength increased the lower bound of the characteristic frequency.

Using the bath parameters calculated from before, we calculated the GLE quantities, friction kernel and random force according to Eqs. (5.32b) and (5.32c). Since the GLE quantities are exact and deterministic equations parametrized by the bath quantities, they show similar trend to what was seen in the bath parameters. For the friction kernel, we were able to observe the effect of the coupling magnitude, and more importantly the delicate change of the bath frequencies. However, when looking at the random force, we only were to observe the effect of the coupling magnitude. Random forces displayed fluctuations with no characteristic frequency, which suggests that it successfully captures the random fluctuations of the bath even though they are deterministic functions.

Using the numerical version of the GLE/RPMD and GLE/TRPMD theory, we calculated the position ACF of both anharmonic and quartic potential at two different temperatures. The anharmonic potential is designed to be a model system with small degrees of anharmonicity. This suggests that the bare harmonic term included in the Hessian of the bath Hamiltonian will capture majority of the interactions that the 2 to Pth beads experience. If we look at the calculated position ACFs for both temperatures shown in Figures 5-9 and 5-10, the GLE/RPMD and GLE/TRPMD ACFs well captures both the magnitude and frequency of the

oscillation of both exact and other PI methods, even though the GLE/(T)RPMD frequencies are slightly blue shifted. Moreover, it should be noted that the GLE formalisms do not capture the correct initial value of the Kubo transformed correlation functions, which are equilibrium quantities, because we directly omit the higher order terms in the original RPMD Hamiltonian. However, we see that the zero time values from both exact and PI methods does not vary too much from the GLE ones because the anharmonicity in this case is not too large. In comparing GLE/RPMD and GLE/TRPMD, we only observe difference in magnitude only, but not frequency. This distinction between the two GLE formalism is consistent with that between RPMD and TRPMD. Overall, we see that even though the GLE picture, which applies the full physical potential to the system coordinate and only harmonic term to the bath, gives qualitatively correct oscillations of both the full RPMD. To correctly capture the RPMD oscillation, we would need to include the higher order terms that we omitted during the mapping process.

Quartic potential, compared to the anharmonic potential, is aimed to model a strongly anharmonic system in which there is little to no harmonic contributions in the system. Moreover, since there are no quadratic term in the expression of the potential, when calculating the bath parameters, we are only including the contribution from the nearest neighbor coupling in the system and bath Hamiltonian. For both temperatures, we see that both GLE/RPMD and GLE/TRPMD fail to qualitatively capture the oscillation of the RPMD result and the exact quantum one, even though the agreement improves with increasing temperature. Moreover, we observe a shift between the GLE/RPMD and GLE/TRPMD correlation functions but not in magnitude of the oscillation. This dissonance is expected as the GLE picture works best for

systems with low degrees of anharmonicity. This is an inherent limitation of the theory as we approximate the bath modes to be a set of harmonic oscillators.

Moving forward, we wish to improve upon the inherent shortcoming of the current GLE formalism. In current development, we omit the potential terms in the original RPMD Hamiltonian. By doing so, when we run dynamics of the system and bath Hamiltonian, we do not obtain the correct equilibrium values. However, it would be beneficial to formalize the GLE/RPMD theory so that it can capture the numerically exact equilibrium quantities like the numerical PI methods. Moreover, we can envision a form of GLE/RPMD theory that parametrizes the bath parameters so they can reproduce the dynamic quantities calculated from higher level calculations such as RPMD, TRPMD, and exact. Another improvement that can be made is to incorporate formalism that can deal with more complex forms of potential as current version assumes that the form of the potential is simple with only one minima. Finally, the current form of the GLE/RPMD theory is an inherently one-body theory. It would be greatly beneficial to construct a many-body equivalent of the GLE picture to be used to analyze the dynamics of a many-body condensed matter systems.

In this work we have presented a GLE analysis of RPMD dynamics. We hope to have provided an alternative picture of approximate quantum dynamics method using the GLE formalism to describe the ring polymer dynamics which will trigger further studies to solve the challenging problem of real-time quantum dynamics.

References

1. R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*. (McGraw-Hill, New York, 1965).

2. R. P. Feynman, *Statistical Mechanics: A Set of Lectures*. (W. A. Benjamin, Reading, MA,

1972).

3. D. Chandler and P. G. Wolynes, J. Chem Phys. **74**, 4078 (1981).

4. M. E. Tuckerman, B. J. Berne, G. J. Martyna and M. L. Klein, J. Chem Phys. **99**, 2796 (1993).

5. D. Marx and M. Parrinello, J. Chem Phys. **104**, 4077 (1996).

6. D. M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995).

7. J. Cao and G. A. Voth, J. Chem. Phys. **99**, 10070 (1993).

8. J. Cao and G. A. Voth, J. Chem. Phys. **100**, 5093 (1994).

9. J. Cao and G. A. Voth, J. Chem. Phys. **100**, 5106 (1994).

10. J. Cao and G. A. Voth, J. Chem. Phys. **101**, 6157 (1994).

11. J. Cao and G. A. Voth, J. Chem. Phys. **101**, 6168 (1994).

12. I. R. Craig and D. E. Manolopoulos, J. Chem. Phys. **121**, 3368 (2004).

13. B. J. Braams and D. E. Manolopoulos, J. Chem. Phys. **125**, 124105 (2006).

14. S. Jang and G. A. Voth, J. Chem. Phys. **111**, 2357 (1999).

15. S. Jang and G. A. Voth, J. Chem. Phys. **111**, 2371 (1999).

16. T. D. Hone, P. J. Rossky and G. A. Voth, J. Chem. Phys. **124**, 154103 (2006).

17. M. Shiga and A. Nakayama, Chem. Phys. Lett. **451**, 175 (2008).

18. G. R. Medders and F. Paesani, J. Chem. Theory Comput. **11**, 1145 (2015).

19. S. K. Reddy, D. R. Moberg, S. C. Straight and F. Paesani, J. Chem. Phys. **147**, 244504 (2017).

20. D. R. Moberg, S. C. Straight, C. Knight and F. Paesani, J. Phys. Chem. Lett. **8**, 2579 (2017).

21. K. M. Hunter, F. A. Shakib and F. Paesani, J. Phys. Chem. B. **122**, 10754 (2018).

22. T. F. M. III and D. E. Manolopoulos, J. Chem. Phys. **123**, 154504 (2005).

23. T. F. M. III and D. E. Manolopoulos, J. Chem. Phys. **122**, 184503 (2005).

24. J. Lobaugh and G. A. Voth, J. Chem. Phys. **106**, 2400 (1997).

25. U. W. Schmitt and G. A. Voth, J. Chem. Phys. **111**, 9361 (1999).

26. I. R. Craig and D. E. Manolopoulos, J. Chem. Phys. **122**, 084106 (2005).

27. I. R. Craig and D. E. Manolopoulos, J. Chem. Phys. **123**, 034102 (2005).

28. R. Biswas, Y. L. S. Tse, A. Tokmakoff and G. A. Voth, J. Phys. Chem. B. **120**, 1793 (2016).

29. S. Jang, A. V. Sinitskiy and G. A. Voth, J. Chem. Phys. **140**, 154103 (2014).

30. T. J. H. Hele, M. J. Willatt, A. Muolo and S. C. Althorpe, J. Chem. Phys. **142**, 191101 (2015).

31. T. J. H. Hele, M. J. Willatt, A. Muolo and S. C. Althorpe, J. Chem. Phys. **142**, 134103 (2015).

32. A. Witt, S. D. Ivanov, M. Shiga, H. Forbert and D. Marx, J. Chem. Phys. **130**, 194510 (2009).

33. S. D. Ivanov, A. Witt, M. Shiga and D. Marx, J. Chem. Phys. **132**, 031101 (2010).

34. S. Habershon, G. S. Fanourgakis and D. E. Manolopoulos, J. Chem. Phys. **129**, 074501 (2008).

35. F. Paesani and G. A. Voth, J. Chem. Phys. **132**, 014105 (2010).

36. G. Trenins and S. C. Althorpe, J. Chem. Phys. **149**, 014102 (2018).

37. G. Trenins, M. J. Willatt and S. C. Althorpe, J. Chem. Phys. **151**, 054109 (2019).

38. M. Rossi, M. Ceriotti and D. E. Manolopoulos, J. Chem. Phys. **140**, 234116 (2014).

39. T. J. H. Hele, Mol. Phys. **114**, 1461 (2016).

40. K. Kreis, M. E. Tuckerman, D. Donadio, K. Kremer and R. Potestio, J. Chem. Theory Comput. **12**, 3030 (2016).

41. J. R. Cendagorta, Z. Bačić and M. E. Tuckerman, J. Chem. Phys. **148**, 102340 (2018).

42. M. Ceriotti, D. E. Manolopoulos and M. Parrinello, J. Chem. Phys. **134**, 084104 (2011).

43. M. Ceriotti and D. E. Manolopoulos, Phys. Rev. Lett. **109**, 100604 (2012).

44. M. Rossi, V. Kapil and M. Ceriotti, J. Chem. Phys. **148**, 102301 (2018).

45. V. Kapil, D. M. Wilkins, J. Lan and M. Ceriotti, J. Chem. Phys. **152**, 124104 (2020).

46. D. Chandler, *Introduction to Modern Statistical Mechanics*. (Oxford University Press, New York, 1987).

47. R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957).

48. R. D. Coalson, J. Chem. Phys. **85**, 926 (1986).

49. M. Ceriotti, M. Parrinello, T. E. Markland and D. E. Manolopoulos, J. Chem. Phys. **133**, 124104 (2010).

50. R. Kubo, Rep. Prog. Phys. **29**, 255 (1966).

51. M. E. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation.* (Oxford University Press, New York, 2010).

52. R. Zwanzig, Phys. Rev. **124**, 983 (1961).

Chapter 6: Conclusion and Future Directions

I. Introduction

The projects presented in this dissertation address challenges of quantum statistics and dynamics in Feynman's imaginary time formalism using classical coarse-graining techniques. In Chapter 3, we have presented a primitive version of the CG-PI theory by constructing a simplified representation of the discrete imaginary path, or ring polymer, using an intermediate CG coordinate and tested how well it can quantitatively capture quantum structural correlations of realistic condensed matter systems at classical cost. In Chapter 4, we have developed further the CG-PI theory and n-CG-PI methodology by explicitly including the many-body interactions in the original force field when constructing the many-body CG potential. Moreover, we have constructed a robust n-CG-PI modeling scheme applicable to general molecular system. In this formalism, the n-CG-PI model can conveniently be constructed from a single PIMD trajectory. Lastly in Chapter 5, we have extended the scope of imaginary time path integral to quantum dynamics by using the well-known generalized Langevin equation formalism to analyze the dynamics of a single ring polymer bead under the influence of the other beads. The three research projects open up exciting possibilities of alternate representation of the quantum-classical isomorphism using classical techniques and therefore possibilities for future developments of imaginary time path integral theories.

II. Remaining Challenges

Even though the CG-PI theory and the n-CG-PI method succeeds in effectively, and quantitatively for various structural correlations, capturing NQE with minimal additional

computational cost, there are various shortcomings that need to be addressed in order for them to be used as a reliable modeling technique.

The most notable limitation to the CG-PI theory is that the development of the theory is heavily based on the near classical limit assumption. Even though the general derivation of exact CG-PI quantities such as CG potential and effective Hamiltonian are provided, these are in general too complicated for direct numerical implementation. To circumvent this issue, we only have dealt with systems nearer the classical limit and assumed that the third and higher order Taylor expanded terms of the many-body CG potential can be omitted. However, little work has done to apply ideas of the CG-PI theory to systems further away from the classical limit in which h can no longer be considered a small parameter. Even though the original work by Sinitskiy and Voth presents a variational equation for the exact CG potential, the ansatz of the functional form has been based on the expressions derived from Taylor expansion.

Moreover, even though the current n-CG-PI method gives good to quantitative agreement to the exact quantum results, this result is only based on comparison of structural correlations as CG-PI theory at the current stage lacks formalism on how to calculate thermodynamic quantities such as pressure or heat capacities. Even though the expression of an energy estimator is provided in Eq. (3.13), the derivative with respect to beta is nontrivial to calculate, especially on the fly during the numerical sampling of the n-CG-PI effective Hamiltonian.

The GLE/RPMD theory succeeds in understanding the behavior of the ring polymer by mapping it into the system and bath Hamiltonian and subsequently solving its equation of motion using Laplace transform to derive the GLE. However, the theory still lacks a way to capture the true real time quantum dynamics as it is an approximation of RPMD, which itself is an

approximate method. Therefore, even though the original motivation for developing the theory was to understand the RPMD dynamics, it should not be taken to capture any real quantum dynamics and the exact results shown in Chapter 5 should be seen as reference only.

The GLE/RPMD theory omits the third and higher order interactions for the $P - 1$ beads in the ring polymer. Such approximation only succeeds for weakly anharmonic systems where the potential is dominated by the harmonic term. However, it is obvious that such approximation breaks down for strongly anharmonic systems as we have observed in the quartic potential case.

III. Future Directions

In order to finalize the CG-PI theory formalism, several theoretical aspects should be worked out. To begin with, it would be beneficial to define the charge of the pseudo-particles in the context of CG-PI theory. In the Feynman's isomorphic potential, the beads within a ring polymer have the same charge as that of the original particle in the force field. However, there are no intra-ring polymer pair interactions as only the same indexed beads of the ring polymer experience the scaled potential of the system. The CG-PI theory has an analogous interaction scheme in which the harmonic nearest neighbor interactions in the ring polymer are coarse-grained into a single harmonic interactions within the pseudo-particle pair. However, since we calculate a new set of cent-cent intra and intermolecular interactions (that are analogous to the original force field), it is not straightforward how to cast the intermolecular electrostatic interactions in the new set of centroid pseudo-particle force fields. The author suggests the charge of the centroid pseudo-particle to be equal to that of the original particle. This way, when subtracting the Boltzmann inversion parametrized intramolecular interactions from the mapped PIMD trajectory, the electrostatic interactions between the centroid pseudo-particles can also be

subtracted. However, this hypothesis should be tested and verified. Moreover, a general derivation of the CG-PI thermodynamic estimators is needed to calculate the thermodynamics of the system of interest aside from structural correlations. Even though the relation between the PIMD and CG-PI energy estimators is presented in Appendix 3C, a more general discussion is required.

To further improve the n-CG-PI method, we would need to apply the current modeling strategy to a more complicated systems. The more complicated system should have a more complex intramolecular structure with more bonded/angular/dihedral/improper interactions and/or more complex functional form in the force field, especially non-harmonic terms. The n-CG-PI modeling strategy discussed in Chapter 5 has successfully been implemented to a nontrivial molecular system (qSPC/Fw water) with intramolecular interactions. However, it still would be beneficial to test the efficacy of this modeling strategy on a system with more intricate structure. Moreover, the assumption in which force-matched centroid intramolecular interactions have the functional form of the original force field must be verified as well with a more general form of bonded interaction terms (polynomial or Morse for example).

In the n-CG-PI water model shown in Chapter 4, we have observed that the obsv-obsv and cent-obsv radial distribution functions (Figure 4.10 and 4.12) predict a more delocalized and soft liquid structure for the n-CG-PI water model compared to the PIMD one. Note that this was not the case for the para-hydrogen model, or at least such 'over-correction' of NQE was minimal due to para-hydrogen force field's simple form of pair interaction and lack of intramolecular structure. The dissonance between the n-CG-PI and the PIMD result in the water system comes from the series of approximations that we have made in constructing the n-CG-PI effective Hamiltonian. First, we approximated that the coupling matrix is constant with respect to time.

Second, we omitted the off-diagonal terms in the matrix as well. We believe that improving upon the second approximation can help improve the result of the n-CG-PI model. It would be worth adding constant harmonic off-diagonal intramolecular term in the n-CG-PI Hamiltonian. Such interactions will not be only easy to parametrize from the same PIMD trajectory that has previously been used but also only marginally more expensive to implement in numerical sampling as we would only add simple harmonic term. The addition of these off-diagonal terms in the n-CG-PI effective Hamiltonian will in turn make the system more 'classical' and hence correct the over-correction of NQE in the current n-CG-PI model.

The current form of the GLE/RPMD theory starts with the original RPMD Hamiltonian, map it into a system and bath form, and calculate the resulting equations of motion. In going from the original RPMD Hamiltonian to the system and bath form, we are omitting the third and higher order polynomial terms in the potential energy term for the $P - 1$ beads. This suggests that we are omitting the interactions in the original Hamiltonian and hence not fully capturing the position distribution of the ring polymer or system and bath coordinates. Even though the GLE/RPMD theory is an approximate theory aimed to understand how the bath coordinates affect the system dynamics for the ring polymer, it still would be beneficial for it to fully capture the equilibrium properties. Moreover, it would be interesting to obtain bath parameters from RPMD trajectories so that it reproduces such higher level calculation results (either correlation functions or other types of dynamic quantities).

Lastly, one limitation of the GLE/RPMD theory is that it is inherently an one-body theory. Not only is it an one-body theory, but it also assumes a very simple form of potential in which there are only one minima so that the second order polynomial term evaluated at the minima is used to construct the bath Hamiltonian Hessian and the bath parameters. The author

believes that in order for the GLE/RPMD theory to become more versatile, such generalized case needs to be worked out and not only for a simplistic 1D toy models.

IV. Final Thoughts

Due to the extensive development of path integral techniques, calculating quantum statistics of complex molecular and many-body systems has become routine. However, these calculations still provide a significant sampling challenge in which the extended phase space of the ring polymer needs to be properly sampled in order to get true quantum statistics. Furthermore, the various dynamic path-integral methods such as CMD, RPMD, and TRPMD are still only an approximation to the true quantum dynamics. To the author, it is not only fascinating but also relieving and delightful that the classical ideas, primarily approximating things to be sets of harmonic oscillators, provide both new insights and ideas to topics that have been extensively studied in the past. The author believes that such neverending chain of thesis, antithesis, and then synthesis, just like the ring polymer, pushes the frontier of our theoretical understanding of the world.