Development of effective stochastic potential method using random matrix theory for efficient conformational sampling of semiconductor nanoparticles at non-zero temperatures

Jeremy A. Scher,1 Michael G. Bayne,1 Amogh Srihari,2 Shikha Nangia,2 and Arindam Chakraborty1

1Department of Chemistry, Syracuse University, Syracuse, New York 13244 USA
2Department of Biomedical and Chemical Engineering, Syracuse University, Syracuse, New York 13244 USA

The relationship between structure and property is central to chemistry, and enables the understanding of chemical phenomena and processes. Need for an efficient conformational sampling of chemical systems arises from the presence of solvents and the existence of non-zero temperatures. However, conformational sampling of structures to compute molecular quantum mechanical properties is computationally expensive because a large number of electronic structure calculations are required. In this work, the development and implementation of the effective stochastic potential (ESP) method is presented to perform efficient conformational sampling of molecules. The overarching goal of this work is to alleviate the computational bottleneck associated with performing a large number of electronic structure calculations required for conformational sampling. We introduce the concept of a deformation potential and demonstrate its existence by the proof-by-construction approach. A statistical description of the fluctuations in the deformation potential due to non-zero temperature was obtained using infinite-order moment expansion of the distribution. The formal mathematical definition of the ESP was derived using functional minimization approach to match the infinite-order moment expansion for the deformation potential. Practical implementation of the ESP was obtained using the random-matrix theory method. The developed method was applied to two proof-of-concept calculations of the distribution of HOMO-LUMO gap in water molecule and solvated CdSe clusters at 300K. The need for large sample size to obtain statistically meaningful results was demonstrated by performing $10^5$ ESP calculations. The results from these prototype calculations demonstrated the efficacy of the ESP method for performing efficient conformational sampling. We envision that the fundamental nature of this work will not only extend our knowledge of chemical systems at non-zero temperatures but will also generate new insights for innovative technological applications.

Keywords: conformational sampling, ensemble averages, random matrix theory, stochastic schrodinger equation, temperature effect, solvent effect
I. INTRODUCTION

The goal of this work is to develop an efficient theoretical and computational method for performing ensemble-averaged properties of semiconductor nanoparticles (NPs). Semiconductor NPs such as quantum dots and rods have important technological applications. Because of the quantum confinement effect [1], their optoelectronic properties can be tuned by adjusting parameters such as the size, shape, material composition, and ligands of the NP. Consequently, NPs have promising applications in photovoltaics, [2–6] light emitting devices, [7–9] charge and energy transfer processes [10–13], and biological labeling. [14, 15]

One of the important environmental factors that influence their optical and electronic properties is the temperature, which has been the focus of numerous theoretical and experimental studies. Early experimental investigation performed by Joshi and co-workers in 2006 looked at this temperature dependence of band gap energies for colloidal CdSe/ZnS core/shell nanocrystals of different diameters. They found that over a temperature range of 10K to 300K, the band gap energies decreased with increasing temperature for all nanocrystal sizes [16]. Recent experimental results on the temperature-effect on quantum dot (QD) optical properties have also been obtained by Savchenko and co-workers in 2017. They studied the first exciton peak shift as a function of temperature in InP/ZnS core/shell QDs in the range of 6.5K to 296K. Similarly, to Joshi et al. they found that the peak both broadens and shifts toward larger wavelengths with increasing temperature [17]. In 2011, Chen and co-workers performed nonadiabatic molecular dynamics simulations on both spherical and elongated CdSe quantum dots using time-domain Kohn-Sham theory. Their results confirm the inverse temperature dependence on band gap energy in quantum dots. They also found that the temperature dependence of the relaxation rate of excited electrons was higher in elongated quantum dots [18]. In the same year, Fischer et al. performed time-dependent density functional theory (TD-DFT) calculations on a Si20H24 quantum dot trajectory. They found that the band gap energy of the QD shifted significantly throughout the dynamics simulation [19]. More recent quantum dynamics simulations utilizing TD-DFT and the tight-binding approximation performed by High and co-workers looked at the excited state energy transfer in a porphyrin light-harvesting device. Their simulation showed that it is essential to couple the nuclear and electronic degrees of freedom to accurately simulate excited state energy transfer [20]. Liu and Jakubikova have also investigated the electron transfer rate in a pyridine-TiO2 light harvesting assembly, and found that the transfer rate was faster when sampling from thermally accessible structures at room temperature compared to a single 0K optimized structure [21].

In colloidal systems, the NPs are surrounded by solvent molecules which provide a constant-temperature heat-bath for the NP. Constant-temperature ensembles such as NVT or NPT are used extensively for calculation and analysis of equilibrium properties. In molecular dynamics simulations, various methods are used to implement the thermostat, such as Berendsen [22], Langevin [23], and Nosé-Hoover [24, 25]. Investigation of temperature dependence is complicated and computationally expensive because it requires electronic structure calculations on a large number of deformed structures. This imposes serious limitations on the size of the systems that are being investigated.

Disorder arising from temperature or solvent effects can also impact the observed properties of organic molecules and systems as well. Much theoretical and computational work has been performed in this area. In 2001, Kwasniewski and coworkers calculated UV-vis absorption spectra on various frames throughout molecular dynamics simulations of trans-stilbene using the ZINDO-CIS method. They found that thermal motion had a strong impact on the HOMO to LUMO transition energy. At temperatures less than 400K, the energy levels displayed pronounced Gaussian thermal broadening of the related band, up to 30 nm [26]. In 2003, Lewis and co-workers performed MD simulations combined with tight-binding density functional theory to calculate the electronic states of a model DNA double helix as the molecule underwent room temperature thermal fluctuations. They found that the population density of the HOMO state over time changed dramatically due to these thermal fluctuations [27]. In 2006, Morth and Autschbach investigated the temperature dependence of the optical rotation of fenchone using TD-DFT. They found that optical rotation in this system increases linearly with temperature over the range of 273K–373K, and concluded that the temperature-dependent vibrational effects were largely responsible for the temperature-dependent rotational effects in rigid organic molecules [28].

Many researchers have also investigated these solvent and temperature effects on organic chemical systems experimentally. In 2002, Ariu et al. studied the temperature dependence on photoluminescence quantum yields (PLQYs) of poly(9, 9-diocetylfluorene) films. They found that the PLQYs increased as the temperature of many of these systems were reduced. In 2005, Dauphas and co-workers investigated β-casein protein aggregation and properties as a function of concentration, temperature, and calcium content, characterized by dynamic light scattering, confocal microscopy, and fluorescence spectroscopy. Over a temperature range of 283K to 323K, they saw that the fluorescence absorption profiles of the protein were very different when the concentration and temperature were changed. In 2014, Mani et al. looked at the fluorescent properties of the salt of 2,6-diaminopyridinium with dihydrogen phosphate. They found that as a result of proton transfer from phosphoric acid in solution to the pyridine nitrogen, the band gap between the HOMO and LUMO states is smaller than compared to 2,6-diaminopyridine. This finding was confirmed by a red shift that was observed in its absorption spectrum [29].
In this work, we present the effective stochastic potential (ESP) method to address the computational bottleneck associated with performing electronic structure calculations on a large number of molecular structures needed for obtaining ensemble-averaged quantities. The central idea of the ESP method is to approximate the explicit interactions between the system and the bath degrees of freedom by a statistically equivalent effective potential. In the ESP method, the construction of the effective potential is performed using the random-matrix theory (RMT) method. Before we describe this new method, a brief background in the history and usage of RMT will be provided for context. Originally, RMT was designed by Wigner in 1951 to deal with the statistics of eigenvalues and eigenfunctions of complex many-body quantum nuclear systems [30]. Wigner later used RMT to investigate the statistical fluctuations of scattering processes. [31] After Wigner, RMT was successfully applied to the description of spectral fluctuations of atomic nuclei [32], complex atoms [33], and complex molecules [34]. The nature of the general theory greatly lends itself to the investigation of highly complex and chaotic systems. Other, more recent examples of how RMT has been applied to physical systems include equilibrium and transport properties of disordered and classically chaotic quantum systems [35], transitions between classical and quantum distributions of hydrogen energy levels [36], and spectral resonances in quartz [37]. Much work has also been done using RMT to investigate the quantum mechanical phenomenon of so-called "persistent current," the perpetual flow of interacting electrons in a static magnetic field [38–42].

There are many forms of RMT that have been used, but in general, RMT is the process of replacing the Hamiltonian of a system with an ensemble of Hamiltonians containing random matrix elements. Most popularly used in RMT is the Gaussian orthogonal ensemble (GOE), where the Hamiltonian is partitioned into two parts: a reference component and a stochastic component where the matrix elements are uncorrelated, normally distributed random numbers. These two components are then coupled to form the total Hamiltonian. Due to their highly complex nature, QDs are a natural candidate for investigation using RMT. One of the earliest investigations of QDs using RMT was done by Jalabert, Stone, and Alhassid in 1992. They developed a statistical theory of the effect of Coulomb blockade in QDs represented as a particle in a finite potential well using random matrices to determine the Coulomb blockade amplitudes. They were able to obtain good agreement with experiment. [43] In 1996, Šeba et al. used a random matrix model to describe the conductance of a model QD connected to an ideal lead. To accurately match experimental results of QDs attached to two leads, they introduced a coupling constant to their RMT Hamiltonian. [44] In that same year, Sivan and co-workers performed an experimental investigation into the disorder of ground state energies of populations of GaAs QDs. [45] More recently, in 2008 Kaplan et al. investigated the fluctuation of two-body screened Coulomb interaction in ballistic QDs using the random wave model (a model related to RMT). They were able to derive analytical expressions for these fluctuations in terms of the linear size of the QD. Despite the Gaussian random single electron basis used in their theory, they found that both two-body and one-body matrix elements followed non-Gaussian distributions. [46] Shankar in 2006, and later in 2008, has discussed a technique for studying QDs using the renormalization group method in conjunction with RMT. [47, 48] In 2017, Akimov utilized stochastic Hamiltonians to investigate excited state relaxation in nonadiabatic condensed matter systems [49].

The remaining part of this article is organized as follows. The main theoretical development of the ESP method is described in section II. The derivation includes the definition of the quantum mechanical Hamiltonian (section II.A) under investigation, the definition of the deformation potential (section II.B), and the conceptual construction of the effective stochastic potential (section II.C). The computational details for implementation of the ESP method are presented in section III. Specifically, implementation details on the choice of sampling schemes (section III.A), choice of one-particle basis functions (section III.B), and stochastic sampling using Gaussian random matrices (section III.C) are presented. The developed ESP method was used to perform proof-of-concept calculations of ensemble-averaged HOMO-LUMO gap energies in chemical systems, and results for H2O and Cd20Se19 are presented in section IV and the discussion of the results are presented in section V.

II. THEORY

A. System setup and definitions

We start by introducing key concepts and definitions that will be used repeatedly in this work. We will start by defining the effective one-particle Hamiltonian as

\[ \hat{h}(\mathbf{r}; \mathbf{R}) = \frac{-\hbar^2}{2m} \nabla^2 + v_{\text{ext}}(\mathbf{r}; \mathbf{R}) + v_{\text{eff}}(\mathbf{r}). \] (1)

The one-particle Hamiltonian is the starting point for treating electron-electron correlation in many-electron systems. In the above expression, \( v_{\text{eff}} \) is the effective one-particle operator for treating electron-electron interaction, which can
be approximated using HF, DFT, pseudopotential, MBPT, or some model potential
\[ \psi_{\text{eff}} = \{ \psi_{\text{HF}}, \psi_{\text{KS}}, \psi_{\text{ps}}, \psi_{\text{model}}, \ldots \}. \tag{2} \]
For wave function based methods, the eigenfunctions of the one-particle Hamiltonian provide the one-particle basis for performing, MP2, CI, and, CCSD calculations.
\[ \hbar \chi_p = \epsilon_p \chi_p. \tag{3} \]
In case of DFT, these one-particle basis functions are used to construct the one-particle density. The derivation presented here is very general and does not make any \textit{a priori} assumption about the form of the effective potential. Typically, \[ \psi_{\text{eff}} \] is not known in advance and has to be constructed iteratively using a self-consistent procedure. However, because of the SCF procedure, the final SCF-converged \[ \psi_{\text{eff}} \] has a functional dependence on the nuclear geometry \[ R \].
\[ \hbar \chi_p = \epsilon_p \chi_p \xrightarrow{\text{SCF convergence}} \psi_{\text{eff, SCF}}(r)[R]. \tag{4} \]
Using the SCF converged effective potential for any nuclear geometry \[ R' \], we define the following one-particle Hamiltonian
\[ f(r; R') = -\frac{\hbar^2}{2m} \nabla^2 + v_{\text{ext}}(r; R') + v_{\text{eff, SCF}}(r)[R'] \tag{5} \]
We also define a reference structure \[ R^0 \] and the effective Hamiltonian corresponding to the reference structure is given by \[ f^0 \]:
\[ f^0(r; R^0) = -\frac{\hbar^2}{2m} \nabla^2 + v_{\text{ext}}(r; R^0) + v_{\text{eff, SCF}}(r)[R^0]. \tag{6} \]
Typically, the reference structure is selected to be the minimum energy structure. However, the derivation does not impose this as a requirement.

### B. Definition of deformation potential

We define a structure \[ R' \] to be deformed if is different from the reference structure \[ R^0 \], and cannot be generated from the reference structure by simple rotational and translational transformation on \[ R^0 \]. The difference between \[ R' \] and \[ R^0 \] is a consequence of deformation in the internal coordinates of the molecular system. Associated with each deformed structure, we define the deformation potential, which is defined as
\[ v_{\text{def}}^\eta = f^\eta - f^0. \tag{7} \]
Using the above definition, the deformation potential can be interpreted as the potential which must be added to the reference Hamiltonian to generate the deformed Hamiltonian. The existence of the deformation potential is an exact condition, and can be proved using the following argument. Because \[ f^\eta \] and \[ f^0 \] exist, therefore, using Eq. (7) \[ v_{\text{def}}^\eta \] must also exist.
Up until this point, we have used abstract Dirac representation for representing the reference and deformed Hamiltonians. However, now we seek to represent the deformed potential as a matrix in a single-particle basis. The matrix representation of the above equation is given as
\[ F^\eta = F^0 + V_{\text{def}}^\eta, \tag{8} \]
where the matrix elements are defined using as set of single-particle basis functions\( \{ \phi_{\mu}^{[\eta]} \} \)
\[ F_{\mu\nu}^\eta = \langle \phi_{\mu}^{[\eta]} | r | \phi_{\nu}^{[\eta]} \rangle. \tag{9} \]
The choice of the single-particle basis functions is an important one, and in principle, may or may not depend on the nuclear geometry \[ R' \]. The choice of the single-particle basis in which the deformed potential is represented is an important topic and will be discussed in section III.
C. Conceptual construction of Effective Stochastic Potential

Calculation of the deformation potential becomes computationally expensive for a large set of structures. The effective stochastic potential method is designed to address this computational bottleneck by replacing the exact deformation potential by the effective potential that shares common identical statistical metrics with the deformation potential for a set of structures. Conceptually, the ESP potential can be defined by the following three steps. First, we define a set of all possible unique structures that a chemical system can exist in by \( \{ R_\eta \}_{\eta = 0}^{\infty} \). Second, for each of the structures, the deformation potential is determined, where the set of all deformed potentials is defined as \( S_{\text{def}}^{V} = \{ V_\eta^\text{def} \}_{\eta = 0}^{\infty} \).

For each deformed structure \( R_\eta \), we associate a corresponding probability of \( p_\eta \) for existence of that structure in a physical system. As the definition of probability implies, \( p_\eta \) should satisfy the following two mathematical relationships:

\[
p_\eta \geq 0,
\]

\[
\sum_{\eta \in S_{\text{def}}^{V}} p_\eta = 1.
\]

For the remainder of the derivation we will use the compact notation of \( v_{ij}^\eta \) to refer to the individual matrix elements of the deformation potential

\[
v_{ij}^\eta = [ V_{\text{def}}^\eta ]_{ij}.
\]

For each matrix element of \( V_{\text{def}}^\eta \) the mean \( \mu \) or the expectation value of the distribution is defined as

\[
s_{ij}^{(1)} \equiv \mu_{ij} = \sum_{\eta \in S_{\text{def}}^{V}} p_\eta v_{ij}^\eta.
\]

Here, we have used both \( s_{ij}^{(1)} \) and \( \mu_{ij} \) to refer to the mean for convenience. Analogously, the variance \( \sigma \) of the distribution is defined as

\[
s_{ij}^{(2)} \equiv \sigma_{ij} = \sum_{\eta \in S_{\text{def}}^{V}} p_\eta [ v_{ij}^\eta - \mu_{ij} ]^2.
\]

Using the mean \( \mu_{ij} \), we define the \( m \)th central moment of the distribution by \( s_{ij}^{(m)} \) which is evaluated using the following expression:

\[
s_{ij}^{(m)} = \sum_{\eta \in S_{\text{def}}^{V}} p_\eta [ v_{ij}^\eta - \mu_{ij} ]^m \quad \text{with} \quad (m > 1).
\]

Evaluation of all the moments for each matrix element of the set of deformation potentials implies the following relationship:

\[
\{ V_{\text{def}}^\eta \}_{\eta = 0}^{\infty} \rightarrow \{ s_{1,\text{def}}, s_{2,\text{def}}, \ldots, s_{\infty,\text{def}} \}.
\]

Using these statistical metrics, we are now in the position to define the effective stochastic potential in matrix representation. We start by defining a symmetric stochastic matrix whose matrix elements are random numbers \( Z \) drawn from a probability-distribution function \( f_{ij}^{\text{pdf}}(z) \):

\[
v_{ij}^{\text{sto}} = v_{ji}^{\text{sto}} = Z.
\]

The probability distribution function must satisfy the following two conditions:

\[
f_{ij}^{\text{pdf}}(z) \geq 0,
\]

\[
\int_{-\infty}^{+\infty} dz f_{ij}^{\text{pdf}}(z) = 1.
\]
The probability of finding $Z$ in the interval $[a, b]$ is given by
\[
P[a \leq Z \leq b] = \int_{a}^{b} dz \, f_{ij}^{pdf}(z). \tag{21}
\]

We define the effective stochastic potential as the stochastic potential whose central-moments are closest to the central-moments of the deformed potential. Mathematically, the ESP potential is defined using the following constrained functional minimization:
\[
\min_{f_{ij}^{pdf}} \left[ \sum_{m=1}^{\infty} \left( s^{(m, \text{def})}_{ij} - s^{(m, \text{sto})}_{ij} \right)^2 \right] \rightarrow f_{ij}^{\text{pdf,min}} \rightarrow v_{ij}^{\text{esp}}[f_{ij}^{\text{pdf,min}}],
\tag{22}
\]
where the constraints on $f_{ij}^{\text{pdf}}$ are given in Equation 20.

### III. COMPUTATIONAL AND IMPLEMENTATION DETAILS

In this section, we present the implementation details of the ESP method by introducing additional approximations needed for practical implementation of the method. Practical implementation of the ESP method requires us to work with finite sets of data. As a consequence, only a finite set of structures $\{R^{\eta}\}_{0}^{M}$ were used for evaluation of the deformation potential and the set of deformation potentials was also finite. Computational implementation also required us to limit the set of moments $(s^{(1)}, \ldots, s^{(N)})$ on the set of deformation potentials to a finite number. The various methods of sampling, choice of single-particle basis functions, and determination of the probability distribution function for ESP are presented below.

#### A. Canonical ensemble sampling using Monte Carlo and molecular dynamics

The probability $p^{\eta}$ associated with the structure $R^{\eta}$ depends on the thermodynamic conditions of the physical system, and can be chosen to describe both equilibrium and non-equilibrium conditions. In this work, we are interested in systems at thermal equilibrium, which are well-described using the constraints of canonical ensemble with constant composition, temperature, and volume (N,V,T). For canonical ensemble, the probability is the well-known Boltzmann expression as shown below:
\[
p^{\eta} = \frac{e^{-\beta(E^{\eta} - E^{0})}}{\sum_{\lambda \in S_{\text{def}}} e^{-\beta(E^{\lambda} - E^{0})}},
\tag{23}
\]
where, $\beta = (k_{B}T)^{-1}$ and $E^{\eta}$ is the energy of the deformed structure and $E^{0}$ is the energy associated with the minimum-energy structure. The first step in the determination of the deformation potential is the generation of the set of structures $R^{\eta}$. In principle, the set of structures can be achieved using either Monte Carlo (MC) by molecular dynamics (MD) simulations. The key point in both of these approaches is to generate a statistically meaningful sample of thermally accessible structures. In the MC procedure, this is achieved by randomly distorting the optimized structure, calculating its energy, and calculating the mean and central moments using the Boltzmann-weighted procedure described in Equation 23. In the MD procedure, equilibrium molecular dynamics calculations are performed at fixed N,V,T and the structures are then randomly selected from the equilibrium distribution. It is important to note that Boltzmann weights are only needed for MC sampling, and should not be included in the case of MD sampling. The relevant equations for MD sampling are presented in Equation 24
\[
\mu = \sum_{\eta \in S_{\text{MD,def}}}^{\text{MD}} V_{\eta}^{\text{def}}.
\tag{24}
\]
In the above equation, the superscript (MD) implies that the set of structures were obtained from the MD simulation.

#### B. One-particle basis for representing the deformation potential

An important feature of the ESP method is the choice of the one-particle basis functions that are used for representing the deformation potential. In this work, we will use the eigenvectors of the reference Hamiltonian for
representing both the deformed and the ESP potential. We start by defining the AO-basis representation of the reference Hamiltonian matrix

$$F^{\eta}C^{\eta} = S^{\eta}C^{\eta}\epsilon^{\eta}. \quad (25)$$

In the first step, we perform canonical orthogonalization

$$X^{\eta} F^{\eta} X^{\eta} = \tilde{F}^{\eta} \quad (26)$$

$$X^{\eta} S^{\eta} X^{\eta} = I. \quad (27)$$

In the next step, we obtain the unitary matrix that diagonalizes the reference transformed Fock matrix $\tilde{F}^{0}$

$$U^{0\dagger} \tilde{F}^{0} U^{0} = \tilde{F}^{0} = \epsilon^{0}. \quad (28)$$

Using $U^{0}$, we transformed all the deformed Fock matrices $\tilde{F}^{\eta}$ into the eigenbasis of the reference Fock matrix $\tilde{F}^{0}$

$$U^{0\dagger} \tilde{F}^{\eta} U^{0} = \tilde{F}^{\eta}. \quad (29)$$

It is important to note that we are transforming using $U^{0}$ instead of $U^{\eta}$. Finally, we obtain the matrix representation of the deformed potential in the eigenbasis of the reference Fock matrix using the following expression

$$V^{\eta}_{\text{def}} = \tilde{F}^{\eta} - \tilde{F}^{0}, \quad (30)$$

which is equivalent to subtracting the eigenvalues of the reference Fock matrix from the deformed Fock matrix.

$$V^{\eta}_{\text{def}} = \tilde{F}^{\eta} - \epsilon^{0} \quad (31)$$

### C. Stochastic sampling using Gaussian random matrices

In this work, we have restricted the sampling to only the mean and the variance

$$\{V^{\eta}_{\text{def}}\} \rightarrow \{\mu^{\text{def}}, \sigma^{\text{def}}\}. \quad (32)$$

The above sampling condition is computationally implemented using symmetric Gaussian random matrices. Each element of the ESP matrix is a Gaussian random number $Z$ as shown in Equation 33

$$[V^{\text{esp}}]_{ij} = [V^{\text{esp}}]_{ji} = \epsilon_{ij}^{\text{exp}} = Z \quad \text{where} \quad Z \in \{f_{ij}^{\text{Gauss}}(z)\}. \quad (33)$$

The function $f_{ij}^{\text{Gauss}}$ is the Gaussian or normal probability distribution function with mean $\mu$ and variance $\sigma$, and has the following form

$$f_{ij}^{\text{Gauss}}(z) = \frac{1}{\sqrt{2\pi}\sigma_{ij}^{\text{def}}} \exp[-\frac{(z-\mu_{ij}^{\text{def}})^2}{2\sigma_{ij}^{\text{def}}^2}]. \quad (34)$$

The corresponding Fock matrix associated with ESP is given by

$$\tilde{F}^{\text{esp}} = \tilde{F}^{0} + V^{\text{esp}}(Z). \quad (35)$$

### IV. RESULTS

#### A. Statistical distribution of HOMO-LUMO gap of H$_2$O at 300 K

We begin with a proof-of-concept calculation for the ESP method, with the goal of demonstrating that the properties of a distribution of thermally-accessible geometries of a chemical system can be accurately reproduced with a stochastic potential derived from that system. We use a single isolated water molecule as our test case. This serves as a useful example of a simple, well-studied polyatomic system. Starting with NVT ensemble, our goal is to calculate the frequency distribution of the HOMO-LUMO gap of water at 300K.
TABLE I. Mean (µ), standard deviation, (σ2) and the pth root of central moments (σ4) of HOMO-LUMO gap energy in eV of water for Hartree Fock (cc-pVDZ basis) and ESP method at a given number of sampling points Nsample. Water geometries were generated from a Monte Carlo procedure.

| Method | Nsample | µ   | σ2  | σ4  |
|--------|---------|-----|-----|-----|
| HF     | 1 × 10^3 | 14.85 | 0.0118 | 0.0155 |
| ESP    | 1 × 10^3 | 14.85 | 0.1023 | 0.0137 |
| ESP    | 5 × 10^3 | 14.85 | 0.1043 | 0.0146 |
| ESP    | 1 × 10^4 | 14.85 | 0.1043 | 0.0145 |
| ESP    | 5 × 10^4 | 14.85 | 0.1039 | 0.0142 |
| ESP    | 1 × 10^5 | 14.85 | 0.1040 | 0.0142 |

1. System setup and electronic structure calculations

The calculations were performed using the canonical ensemble with constant N, V, and T at 300K. The sampling of the structures for this ensemble was performed using the Monte Carlo procedure. A set of 1000 structures of water molecules were generated by introducing random distortions to the bond-distances and bond-angles of the optimized structure of water. For each of the 1000 structures, Hartree-Fock (HF) calculations were performed using cc-pVDZ basis functions, using the GAMESS software package. [50] The probability associated with each of the 1000 structures were calculated using the Boltzmann expression as shown below

\[ p_i = \frac{e^{-E_i}}{\sum_j e^{-E_j}} \text{,} \quad (36) \]

where \( E_0 \) is the energy of the minimum-energy structure. The eigenvalues of the Fock matrix from the converged HF calculations were used to calculate the HOMO-LUMO gap for the 1000 structures.

\[ \Delta \epsilon_{HF} = \epsilon_{LUMO} - \epsilon_{HOMO} \text{.} \quad (37) \]

Using the thermal probabilities \( p_i \), the thermally-weighted frequency distribution of the HOMO-LUMO gaps for 1000 structures obtained from the HF/cc-pVDZ calculation at 300K is presented in Figure 1. The mean and variance of the distribution is presented in Table I.

2. Construction of the effective stochastic potential

The ESP for water at 300K was constructed from a different set of 1000 Fock matrices obtained from the electronic structure calculations. Using Equation 25, all the Fock matrices were first transformed into the eigenbasis of the Fock matrix associated with the minimum-energy structure \( F^0 \). Then, using Equation 31, the set of 1000 deformation potential matrices were obtained. The mean and the variance associated with each matrix element of the deformation potential matrices were calculated using Equation 14 and Equation 15. The means and variances were used in the next step to calculated the matrix elements of the ESP using Gaussian random numbers as show in Equation 34. The ESP Fock matrix was constructed from the ESP matrix using Equation 35. A set of 1000 ESP Fock matrices were generated by stochastic sampling, and the eigenvalues of the ESP Fock matrices were used to calculate the ESP HOMO-LUMO gap

\[ \Delta \epsilon_{ESP} = \epsilon_{LUMO} - \epsilon_{HOMO} \text{.} \quad (38) \]

3. Comparison of the ESP and electronic structure results

The comparison of the mean, the variance and higher central moments of the frequency distributions of HOMO-LUMO gap energies obtained using both Hartree Fock and ESP methods are presented in Table I. Using the mean \( \mu \), the \( p^{th} \) root of the central moment is defined as

\[ \sigma_p = \left[ \frac{1}{N_{sample}} \sum_{i=1}^{N_{sample}} (x_i - \mu)^p \right]^{1/p} \quad (39) \]
It was found that both the mean HOMO-LUMO gap obtained from ESP were in good agreement with the HF results. The results demonstrate that the ESP can successfully capture the fluctuations present in the deformation potential at non-zero temperatures. However, comparison of higher-order moments (Table I) reveal that the two distributions have similar but not identical characteristics. In order to compare the statistics of the Hartree Fock and ESP method distributions, we used Z-score (also called standard score) \[ Z = \frac{x - \mu}{\sigma}. \]

where \( x \) is a measured value, \( \mu \) is the mean of the value over a distribution, and \( \sigma \) is the corresponding standard deviation. Z-score is a unitless quantity that is relative to the distribution’s standard deviation. A positive Z-score reflects a value greater than the mean, and negative Z-score reflects a value smaller than the mean. A Z-score of 1 corresponds to a value that is exactly one standard deviation greater than the mean of the distribution. Although, the ESP and HF distribution were found to differ in certain regions (Figure 1), the overall features of the both the distributions were found to in good agreement with each other.

4. Effect of sample size on frequency distribution

The effect of sample size on the ESP results were investigated by performing \( 10^3, 10^4, \) and \( 10^5 \) uncorrelated sets of ESP calculations, and the results are summarized in Table I and Figure 2. In all cases, the mean and higher-order moments exhibited convergence with the size of the sample. The frequency distributions in Figure 2 show that the noise in the distribution for \( N_{\text{sample}} = 1000 \) is attenuated by increasing the sampling size to \( 10^4 \).

B. HOMO-LUMO gap of Cd_{20}Se_{19} in water at 300K

We now turn the discussion to the application of the ESP method to a small CdSe QD (1.2 nm in diameter) in aqueous media at 300K. These systems are excellent candidates on which to apply the ESP method, due to the potentially useful applications of QDs, the complexity of their electronic states, and the impact of solvent environment and non-zero temperature of its optical and electronic properties.
Equilibrium molecular dynamics simulations with constant N,V,T were performed for the Cd_{20}Se_{19} cluster with explicit water molecules at 300K. The force-field developed by Rabani et al. [52] was used for the CdSe cluster, and the explicit water molecules were treated using TIP3 force. The MD calculations and analysis of the simulations were performed by the GROMACS molecular dynamics program. The initial structure for the MD calculations was obtained from the minimum-energy structure obtained from electronic structure theory at HF/LANL2Z ECP level. The initial structure was then re-optimized using the CdSe force-field in the MD program in vacuum. The simulation cell was constructed by solvating the CdSe cluster in a water bath consisting of 11763 water molecules. The entire system was equilibrated at 300K by performing MD simulation for 500 ns. The equilibration conditions were verified by monitoring the total energy, total potential energy, and the temperature of the system.

After equilibration, a set of 50 structures was randomly selected from the MD trajectories for construction of the ESP matrix. As a first step, HF calculation using LANL2Z ECP basis functions were performed for the set of structures, and the deformation potential matrices were calculated. The mean and variance of each matrix element were calculated, and the ESP matrix and the ESP Fock matrix were constructed using the procedure described in section III.

2. Predictions from the ESP calculations

The average HOMO-LUMO gap ($\mu$) and the higher-order moments (Equation 39) obtained using the ESP are presented in Table II. The results from Table II highlight the need for large sampling size to obtain statically meaningful quantities. Specifically, we find that a sample size of $N_{\text{sample}} = 50$ is inadequate for describing the distribution of the HOMO-LUMO gap for both HF and ESP calculations. The results from ESP show that sample size in excess of 1000 are needed to converge the mean and higher-order moments. Comparison between the HF and ESP results show that mean HOMO-LUMO gap is significantly over-estimated in HF set of calculations because of inadequate sample size. The distribution of the HOMO-LUMO gap from the ESP calculations are presented in Figure 3. The distribution was found to be highly asymmetric with respect to the origin with a significant positive bias towards the -Z direction and a long tail in the +Z direction. The frequency distribution from the ESP data (Figure 3) implies that at 300K, the HOMO-LUMO gap for majority of the structures are lower than the mean gap. The long tail implies, that a very small fraction of the population exhibits very large HOMO-LUMO gaps. The distribution at 300K is compared with the the HOMO-LUMO gap of the minimum-energy structure (vertical black line in Figure 3). The comparison show that the maxima of the frequency distribution is significantly different from the minimum-energy structure. This implies
TABLE II. Mean ($\mu$), standard deviation, ($\sigma_2$) and the $p^{th}$ root of central moments ($\sigma_4$) of HOMO-LUMO gap energy in eV of Cd$_{20}$Se$_{19}$ for Hartree Fock (LAN2LZ-DZ ECP basis) and ESP method at a given number of sampling points $N_{\text{sample}}$. Cd$_{20}$Se$_{19}$ geometries were generated from molecular dynamics simulation.

| Method | $N_{\text{sample}}$ | $\mu$   | $\sigma_2$ | $\sigma_4$ |
|--------|----------------------|---------|------------|------------|
| HF     | $5 \times 10^4$      | 2.0571  | 0.2530     | 0.0828     |
| ESP    | $5 \times 10^4$      | 1.9514  | 0.5745     | 0.8697     |
| ESP    | $1 \times 10^3$      | 0.9639  | 0.8087     | 1.6318     |
| ESP    | $5 \times 10^3$      | 0.9806  | 0.7429     | 1.2373     |
| ESP    | $1 \times 10^4$      | 0.9775  | 0.7431     | 1.2458     |
| ESP    | $5 \times 10^4$      | 0.9748  | 0.7419     | 1.2719     |
| ESP    | $1 \times 10^5$      | 0.9760  | 0.7446     | 1.2736     |

FIG. 3. HOMO-LUMO gap Z-scores for a Cd$_{20}$Se$_{19}$ quantum dot using increasing number of sampling points at $T = 300$K. The vertical bar at $Z=2.2$ represents the HOMO-LUMO gap energy of the 0K optimized structure with $\Delta \epsilon = 2.7$eV. [53].

that the minimum-energy structure is a poor representative descriptor for the set of thermally accessible structures at 300K. From Figure 3 we conclude that the heating the system from 0K to 300K should decrease the HOMO-LUMO gap of CdSe clusters. This observation from the ESP calculation also is supported by previous experimental [16, 17] and theoretical [18, 19] studies on CdSe and other semiconductor clusters and quantum dots.

V. DISCUSSION

Sampling constant temperature ensembles: This work demonstrates that the ESP method can be used with either Monte Carlo or molecular dynamics simulations. Specifically, the water calculations were performed using Monte Carlo sampling and the CdSe cluster calculations were performed using classical molecular dynamics simulation. In our experience, classical MD simulation provides a an efficient computational route for sampling NVT and NPT ensembles for solvated molecular systems which explicit solvent molecules. We believe that the study of CdSe cluster in water are particularly relevant to the study of semiconductor nanoparticles, because it has generated new insight into the electronic structure of clusters at non-zero temperatures in aqueous media. The results from the ESP calculations also highlighted the requirement of a large sample set for obtaining statistically meaningful results.

Future directions: Although the proof-of-concept calculations presented here were restricted to HF calculations, the ESP method is can be combined other methods for treating electron-electron correlation. The ESP method generates a manifold of single-particle states which can be used as a starting point for combining ESP with existing many-body theories such as DFT, TDDFT, and Green’s function methods. In addition, the ESP method can also
be applied to non-equilibrium or non-Boltzmann distributions by appropriate choice of the set of probabilities \( \{ p^0 \} \) (either discrete or continuous) that govern the population of molecular structures for a given chemical environment.

VI. CONCLUSIONS

This work presented the development of the effective stochastic potential method (ESP) as a route to perform efficient conformational sampling for molecular systems at non-zero temperatures. The central idea of the ESP method is to replace the fluctuations experienced by the molecular system by a stochastic potential that provides an equivalent statistical description of those fluctuations. The concept of deformation potential was introduced and the existence theorem for such a potential was presented. The formal mathematical constructions of the ESP potential was presented using a functional minimization approach to capture the fluctuations in the deformation potential. The computational implementation of the ESP was achieved using random matrix theory. Two proof-of-concept calculations were performed using the developed ESP method, by calculating the distributions of HOMO-LUMO gaps of water and solvated CdSe cluster at 300 K. The results from these calculations demonstrate the efficacy of the ESP method in providing an efficient conformational sampling of molecular systems at non-zero temperatures.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. CHE-1349892. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575. We are also grateful to Syracuse University for additional computational and financial support.

[1] Wang, Y.; Herron, N. Nanometer-sized semiconductor clusters: Materials synthesis, quantum size effects, and photophysical properties. *Journal of Physical Chemistry* 1991, 95, 525–532.

[2] Alivisatos, A. Perspectives on the physical chemistry of semiconductor nanocrystals. *J. Phys. Chem.* 1996, 100, 13226–13239.

[3] Žídek, K.; Zheng, K.; Abdellah, M.; Lenngren, N.; Chábera, P.; Pullerits, T. Ultrafast dynamics of multiple exciton harvesting in the CdSe-ZnO system: Electron injection versus auger recombination. *Nano Lett.* 2012, 12, 6393–6399.

[4] Santra, P.; Kamat, P. Mn-doped quantum dot sensitized solar cells: A strategy to boost efficiency over 5%. *J. Am. Chem. Soc.* 2012, 134, 2508–2511.

[5] Tang, J.; Liu, H.; Zhitonimisky, D.; Hoogland, S.; Wang, X.; Furukawa, M.; Levina, L.; Sargent, E. Quantum junction solar cells. *Nano Lett.* 2012, 12, 4889–4894.

[6] Salant, A.; Shalom, M.; Tachan, Z.; Buhbut, S.; Zaban, A.; Banin, U. Quantum rod-sensitized solar cell: Nanocrystal shape effect on the photovoltaic properties. *Nano Lett.* 2012, 12, 2095–2100.

[7] Boehme, S.; Wang, H.; Siebbeles, L.; Vanmaekelbergh, D.; Houtepen, A. Electrochemical charging of CdSe quantum dot films: Dependence on void size and counterion proximity. *ACS Nano* 2013, 7, 2500–2508.

[8] Jun, S.; Lee, J.; Jang, E. Highly luminescent and photostable quantum dot-silica monolith and its application to light-emitting diodes. *ACS Nano* 2013, 7, 1472–1477.

[9] Park, I.-K.; Kwon, M.-K.; Cho, C.-Y.; Kim, J.-Y.; Cho, C.-H.; Park, S.-J. Effect of InGaN quantum dot size on the recombination process in light-emitting diodes. *Appl. Phys. Lett.* 2008, 92, 253105.

[10] Huang, J.; Mulford, K.; Du, P.; Chen, L. Photodriven charge separation dynamics in CdSe/ZnS core/shell quantum dot/cobaloxime hybrid for efficient hydrogen production. *J. Am. Chem. Soc.* 2012, 134, 16472–16475.

[11] Ji, X.; Palui, G.; Avellini, T.; Na, H.; Yi, C.; Knappenberger Jr., K.; Matteossi, H. On the pH-dependent quenching of quantum dot photoluminescence by redox active dopamine. *J. Am. Chem. Soc.* 2012, 134, 6006–6017.

[12] Peng, X.; Misewich, J.; Wong, S.; Sfeir, M. Efficient charge separation in multidimensional nanohybrids. *Nano Lett.* 2011, 11, 4562–4568.

[13] Žídek, K.; Zheng, K.; Ponseca, C.; Messing, M.; Wallenberg, L.; Chbera, P.; Abdellah, M.; Sundstrm, V.; Pullerits, T. Electron transfer in quantum-dot-sensitized ZnO nanowires: Ultrafast time-resolved absorption and terahertz study. *J. Am. Chem. Soc.* 2012, 134, 12110–12117.

[14] Zhang, Y.; Ke, X.; Zheng, Z.; Zhang, C.; Zheng, Z.; Zhang, F.; Hu, Q.; He, Z.; Wang, H. Encapsulating quantum dots into enveloped virus in living cells for tracking virus infection. *ACS Nano* 2013, 7, 3896–3904.

[15] Bruchez Jr., M.; Moronne, M.; Gin, P.; Weiss, S.; Alivisatos, A. Semiconductor nanocrystals as fluorescent biological labels. *Science* 1998, 281, 2013–2016.
[16] Joshi, A.; Narsingi, K. Y.; Manasreh, M. O.; Davis, E. A.; Weaver, B. D. Temperature dependence of the band gap of colloidal CdSe/ZnS core/shell nanocrystals embedded into an ultraviolet curable resin. Applied Physics Letters 2006, 89, 131907.

[17] Savchenko, S. S.; Vokhmintsev, A. S.; Weinstein, I. A. Temperature-induced shift of the exciton absorption band in InP/ZnS quantum dots. Opt. Mater. Express 2017, 7, 354–359.

[18] Chen, L.; Bao, H.; Tan, T.; Prezhdo, O. V.; Ruan, X. Shape and Temperature Dependence of Hot Carrier Relaxation Dynamics in Spherical and Elongated CdSe Quantum Dots. The Journal of Physical Chemistry C 2011, 115, 11400–11406.

[19] Fischer, S. A.; Habenicht, B. F.; Madrid, A. B.; Duncan, W. R.; Prezhdo, O. V. Regarding the validity of the time-dependent Kohn-Sham approach for electron-nuclear dynamics via trajectory surface hopping. The Journal of Chemical Physics 2011, 134, 024102.

[20] High, J. S.; Rego, L. G. C.; Jakubikova, E. Quantum Dynamics Simulations of Excited State Energy Transfer in a Zinc-Free-Base Porphyrin Dyad. The Journal of Physical Chemistry A 2016, 120, 8075–8084.

[21] Liu, C.; Jakubikova, E. Two-step model for ultrafast interfacial electron transfer: limitations of Fermi’s golden rule revealed by quantum dynamics simulations. Chem. Sci. 2017, 8, 5979–5991.

[22] Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. The Journal of Chemical Physics 1984, 81, 3684–3690.

[23] Schlick, T. Molecular Modeling and Simulation: An Interdisciplinary Guide: An Interdisciplinary Guide

[24] Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. The Journal of Chemical Physics 1984, 81, 511–519.

[25] Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. Phys. Rev. A 1985, 31, 1695–1697.

[26] Kwansiewski, S.; Franois, J.; Deleuze, M. Temperature effects on the UV-Vis electronic spectrum of trans-stilbene. International Journal of Quantum Chemistry 2001, 85, 557–568.

[27] Lewis, J.; Cheatham III, T.; Starikov, E.; Wang, H.; Sankey, O. Dynamically amorphous character of electronic states in poly(dA)-poly(dT) DNA. Physical Biology 2003, 107, 2581–2587.

[28] Mort, B.; Autschbach, J. Temperature dependence of the optical rotation of fenchone calculated by vibrational averaging. Journal of Physical Chemistry A 2006, 110, 11381–11383.

[29] Mani, R.; Rietveld, I.; Varadharajan, K.; Louhi-Kultanen, M.; Muthu, S. Fluorescence properties reinforced by proton transfer in the salt 2,6-diaminopyridinium dihydrogen phosphate. Journal of Physical Chemistry A 2014, 118, 6883–6892.

[30] Wigner, E. P. Characteristic Vectors of Bordered Matrices With Infinite Dimensions. Annals of Mathematics 1955, 62, 548–564.

[31] Guhr, T.; Miller-Groeling, A.; Weidemann, H. Random-matrix theories in quantum physics: Common concepts. Physics Report 1998, 299, 189–425.

[32] O. Bohigas, A. P., R.U. Haq

[33] Garg, J.; Rainwater, J.; Petersen, J.; Havens, W. Neutron resonance spectroscopy. III. Th232 and U238. Nuclear Data for Science and Technology

[34] Ericson, T.; Mayer-Kuckuk, T. Fluctuations in nuclear reactions.

[35] Wintgen, D.; Friedrich, H. Classical and quantum-mechanical transition between regularity and irregularity in a Hamiltonian system. Phys. Rev. A 1987, 35, 1464–1466.

[36] Ellegaard, C.; Guhr, T.; Lindemann, K.; Nyg˚ ard, J.; Oxborrow, M. Symmetry Breaking and Spectral Statistics of Aco ustic Resonances in Quartz Blocks. Phys. Rev. Lett. 1996, 77, 4918–4921.

[37] G. Montambaux, Persistent Current of Interacting Electrons: a Simple Hartree-Fock Picture. J. Phys. I France 1996, 6, 1–4.

[38] Kato, H.; Yoshiohta, D. Suppression of persistent currents in one-dimensional disordered rings by the Coulomb interaction. Phys. Rev. B 1994, 50, 4943–4946.

[39] Bouzerar, G.; Poilblanc, D. Persistent currents in multichannel interacting systems. Phys. Rev. B 1995, 52, 10772–10775.

[40] Pascaud, M.; Montambaux, G. Magnetization of mesoscopic disordered networks. J. Phys. I France 1997, 7, 1123–1126.

[41] Jalabert, R.; Stone, A.; Alhassid, Y. Statistical theory of Coulomb blockade oscillations: Quantum chaos in quantum dots. Physical Review B 1992, 68, 3468–3471.

[42] Šeba, P. Random matrix theory and mesoscopic fluctuations. Physical Review B - Condensed Matter and Materials Physics 1996, 53, 13024–13028.

[43] Sivan, U.; Berkovits, R.; Aloni, Y.; Prus, O.; Auerbach, A.; Ben-Yoseph, G. Mesoscopic fluctuations in the ground state energy of disordered quantum dots. Physical Review Letters 1996, 77, 1123–1126.

[44] Kaplan, L.; Alhassid, Y. Interaction matrix element fluctuations in ballistic quantum dots: Random wave model. Physical Review B - Condensed Matter and Materials Physics 2008, 78.

[45] Shankar, R. Dots for dummies. Journal of Statistical Physics 2006, 125, 1173–1182.

[46] Shankar, R. Colloquium: Chaotic quantum dots with strongly correlated electrons. Reviews of Modern Physics 2008, 80, 379–394.
[49] Akimov, A. V. Stochastic and Quasi-Stochastic Hamiltonians for Long-Time Nonadiabatic Molecular Dynamics. *The Journal of Physical Chemistry Letters* **2017**, *8*, 5190–5195, PMID: 28985075.

[50] Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. General atomic and molecular electronic structure system. *Journal of Computational Chemistry* **1993**, *14*, 1347–1363.

[51] Urdan, T. *Statistics in Plain English, Third Edition*; Taylor & Francis, 2011.

[52] Rabani, E. An interatomic pair potential for cadmium selenide. *The Journal of Chemical Physics* **2002**, *116*, 258–262.

[53] Wang, L.-W.; Zunger, A. Pseudopotential calculations of nanoscale CdSe quantum dots. *Phys. Rev. B* **1996**, *53*, 9579–9582.