

## Shannon Entropy Based Time-Dependent Deterministic Sampling for Efficient “On-the-Fly” Quantum Dynamics and Electronic Structure

David Hocker,<sup>†</sup> Xiaohu Li,<sup>‡</sup> and Srinivasan S. Iyengar\*

Department of Chemistry and Department of Physics, Indiana University, 800 E. Kirkwood Ave, Bloomington, Indiana 47405, United States

Received October 14, 2010

**Abstract:** A new set of time-dependent deterministic sampling (TDDS) measures, based on local Shannon entropy, are presented to adaptively gauge the importance of various regions on a potential energy surface and to be employed in “on-the-fly” quantum dynamics. Shannon sampling and Shannon entropy are known constructs that have been used to analyze the information content in functions: for example, time-series data and discrete data sets such as amino acid sequences in a protein structure. Here the Shannon entropy, when combined with dynamical parameters such as the instantaneous potential, gradient and wavepacket density provides a reliable probe on active regions of a quantum mechanical potential surface. Numerical benchmarks indicate that the methods proposed are highly effective in locating regions of the potential that are both classically allowed as well as those that are classically forbidden, such as regions beyond the classical turning points which may be sampled during a quantum mechanical tunneling process. The approaches described here are utilized to improve computational efficiency in two different settings: (a) It is shown that the number of potential energy calculations required to be performed during on-the-fly quantum dynamics is fewer when the Shannon entropy based sampling functions are used. (b) Shannon entropy based TDDS functions are utilized to define a new family of grid-based electronic structure basis functions that reduce the computational complexity while maintaining accuracy. The role of both results for on-the-fly quantum/classical dynamics of electrons and nuclei is discussed.

### I. Introduction

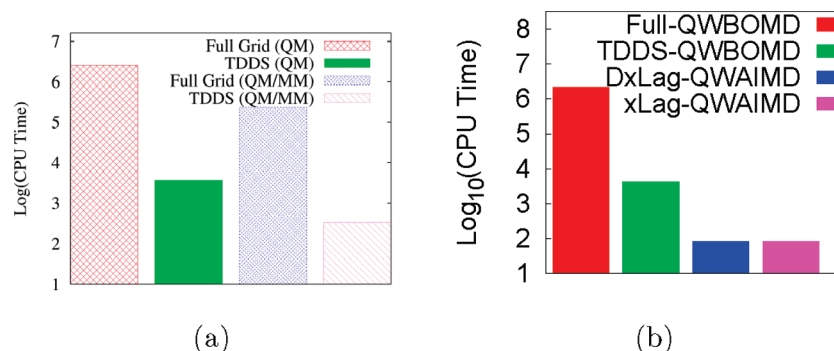
The time-dependent Schrödinger equation is the starting point for many computational methodologies employed in gas-phase<sup>1</sup> and condensed-phase chemical dynamics.<sup>2</sup> When utilized, the Born–Oppenheimer approximation allows for separation of the nuclear and electronic degrees of freedom in a system, allowing for varying treatments of the nuclei, be it classical,<sup>3–7</sup> quantum-mechanical,<sup>1,8–34</sup> or semiclassical.<sup>35–42</sup> In all cases, the nuclei are either propagated along parametrically fitted electronic surfaces known *a priori*, or along

highly accurate (and sometimes computationally expensive) electronic surfaces that require no prior knowledge of the system. Due to the large number of quantum mechanical energy and gradient calculations required by the latter approach, there has been a strong motivation toward “on-the-fly” dynamics schemes to overcome this computational barrier and potentially allow for larger, more complex systems to be studied.<sup>3–7,35,41,43–46</sup> This growing subfield of *ab initio* molecular dynamics (AIMD) approximates the electronic structure alongside the nuclei to simulate molecular dynamics. When AIMD techniques are embedded in a full quantum or semiclassical scheme, there is the potential for large systems to be accurately treated with the complete machinery of quantum dynamics. Several efforts have been made toward this goal.<sup>41,46–49</sup>

\* Corresponding author e-mail: iyengar@indiana.edu.

<sup>†</sup> Present address: Department of Chemistry, Princeton University.

<sup>‡</sup> Present address: Department of Chemistry, Northwestern University.



**Figure 1.** (a) Computational expense for QWAIMD with and without time-dependent deterministic sampling (TDDS). *Note that in all cases the vertical axis is the logarithm of CPU time.* TDDS provides an enormous reduction in computational time for two different types of embedding schemes (QM/MM and QM), with little loss in accuracy. (Reproduced with permission from ref 53. Copyright 2008, American Institute of Physics). (b) Further reduction in computation time is facilitated through the introduction of a propagation scheme that involves multiple diabatic states.<sup>54</sup> Again, accuracy in computing potential surfaces is preserved, while reducing the computational overhead substantially. (Reproduced with permission from ref 54. Copyright 2010. American Institute of Physics).

Recently,<sup>46,50–56</sup> we have introduced a methodology that accurately computes quantum dynamical effects in a subsystem while simultaneously handling the motion of the surrounding atoms and changes in electronic structure calculation. The approach is quantum-classical<sup>40,57–63</sup> and involves the synergy between a time-dependent quantum wave packet description and *ab initio* molecule dynamics. As a result, the approach is called quantum-wave packet *ab initio* molecular dynamics (QWAIMD). Since the quantum dynamics is performed on Cartesian grids, the predominant bottleneck is the computation of the grid-based, time-dependent electronic structure potential and gradients generated by the motion of the classical nuclei. This limitation is partially surmounted through the following methodological improvements:

- A time-dependent deterministic sampling (TDDS) technique was introduced in refs 51 and 52, which when combined with numerical methods such as an efficient wavelet compression scheme and low-pass filtered Lagrange interpolation,<sup>52</sup> provides computational gains of many orders of magnitude (Figure 1).
- Multiple diabatic reduced single particle electronic density matrices are propagated simultaneously with the quantum wavepacket in ref 54, and the associated diabatic states are used to construct an adiabatic surface at every instant in time using a nonorthogonal CI formalism. The diabatic approximation allows reuse of the two-electron integrals during the on-the-fly potential energy surface computation stage and leads to substantial reduction in computational costs (Figure 1).

QM/MM generalizations to QWAIMD have also been completed.<sup>53</sup> The approach is being generalized to treat extended systems<sup>64</sup> for condensed-phase simulations; a biased QWAIMD formalism to sample rare events is also currently being developed. We have utilized QWAIMD to compute vibrational properties of hydrogen-bonded clusters inclusive of quantum nuclear effects<sup>52</sup> and have also adopted the method to study hydrogen tunneling in enzyme active sites.<sup>55,65</sup> The quantum dynamics scheme in QWAIMD has also been used to develop a technique known as multistage

*ab initio* wavepacket dynamics (MSAIWD) to treat open-electronic systems.<sup>66,67</sup>

In this publication, we probe the relevant regions of a potential surface, using a new TDDS function based on the notion of Shannon entropy.<sup>68–72</sup> This paper is organized as follows: An overview of QWAIMD is presented in section II along with a discussion of time-dependent deterministic sampling and its current efficacy. The derivation and physical rationale for the sampling functions that utilize Shannon entropy are given in section III. The numerical benchmarks are arranged in a multipronged fashion. In section IV.A, we discuss the use of the Shannon entropy based TDDS functions in adaptive determination of critical regions of the potential surface during dynamics. Accuracy in computing vibrational properties is also discussed. In section IV.B, the approach is utilized to construct an accurate “grid-based” electronic basis set. This implementation of Shannon-entropy based TDDS leads to a sizable reduction in the number of electronic basis functions that need to be utilized in calculations involving hydrogen-bonded systems. Consequently, the approach is tested for accuracy and efficiency for three different kinds of hydrogen-bonded clusters. This same idea is further exploited in ref 54 to develop an implicitly time-dependent, grid-based electronic structure basis to tremendously improve the efficiency and accuracy of QWAIMD. Concluding remarks are given in section V.

## II. Main Features of Quantum Wavepacket *Ab Initio* Molecular Dynamics (QWAIMD)

As noted above, QWAIMD is based on a synergy between quantum wavepacket dynamics and *ab initio* molecular dynamics.<sup>46,50–55,64</sup> The partitioning scheme divides the system into three subsystems: subsystem A may include particles that display critical quantum dynamical effects; subsystems B and C contain the surrounding nuclei and electrons, respectively, and are treated under the AIMD formalism.<sup>46,50,51,73,74</sup> Subsystem A is propagated according to the Trotter-factorized quantum mechanical time propagator:<sup>10,75–77</sup>

$$\begin{aligned}\chi_A(x;t) &= \exp\left\{-\frac{iHt}{\hbar}\right\}\chi_A(x;t=0) \\ &= \left[\exp\left\{-\frac{iVt}{2\hbar}\right\}\exp\left\{-\frac{iKt}{\hbar}\right\}\exp\left\{-\frac{iVt}{2\hbar}\right\} + \mathcal{O}(t^3)\right]\chi_A(x;t=0),\end{aligned}\quad (1)$$

and the free-propagator,  $\exp\{-(iKt)/(\hbar)\}$ , is represented using “distributed approximating functionals” (DAF):<sup>46,50,78–81</sup>

$$\begin{aligned}\left\langle x \left| \exp\left\{-\frac{iK\Delta t}{\hbar}\right\} \right| x' \right\rangle &\equiv \tilde{K}(x, x'; \Delta t) \equiv \tilde{K}(|x - x'|, \Delta t) \\ &= \frac{(2\pi)^{-1/2}}{\sigma(0)} [e^{-(x - x')^2/[2\sigma(\Delta t)^2]}] \times \\ &\quad \sum_{n=0}^{M_{\text{DAF}}/2} \frac{(-1/4)^n}{n!} \left(\frac{\sigma(0)}{\sigma(\Delta t)}\right)^{2n+1} H_{2n}\left(\frac{x - x'}{\sqrt{2}\sigma(\Delta t)}\right)\end{aligned}\quad (2)$$

where  $\{\sigma(\Delta t)\}^2 = \sigma(0)^2 + i\Delta t\hbar/M_{\text{QM}}$  and  $M_{\text{QM}}$  is the mass of the quantum mechanical particle. Equation 2 utilizes the well-known analytical expression for the free-propagation of a Gaussian function with spread  $\sigma(0)$ ,<sup>82</sup> as well as the fact that Hermite functions,  $\{H_n(x)\}$ , are generated from Gaussian functions.<sup>50</sup> The result is a banded-Toeplitz matrix representation for the quantum propagator.<sup>78–81</sup> [The structure of a Toeplitz matrix is such that the  $(i, j)$ th element depends only on  $|i - j|$ , allowing for an efficient computational scheme that only requires the first (banded) row of the matrix to be stored. This is exploited in the “DAF” free-propagator<sup>67</sup> to reduce computational cost.] It is routine to carry out QWAIMD using Hermite functions of the order of  $M_{\text{DAF}} = 20$ –30. When using a larger number of Hermite functions, numerical stability becomes an issue, but this is surmounted through a minor modification of the recursion relation as outlined in ref 46.

**II.A. Time-Dependent Deterministic Sampling (TDDS) Based QWAIMD.** The evolution of the classical nuclei involves the wavepacket-averaged Hellmann–Feynman forces obtained from electronic structure calculations on the discrete wavepacket grid. To minimize the number of such calculations while maintaining accuracy, a time-dependent deterministic sampling (TDDS) function was introduced in refs 51 and 52. The mathematical form of the TDDS function is defined to be a function of the quantum nuclear degrees of freedom,  $R_{\text{QM}}$ , as follows: The TDDS function is chosen to be directly proportional to the wavepacket probability density,  $\rho(R_{\text{QM}})$ , and gradient of the potential,  $V'(R_{\text{QM}})$ , while being inversely proportional to the potential,  $V(R_{\text{QM}})$ , as noted in eq 3. Large values of the TDDS function represent areas where sampling should occur. The construction of TDDS has physical justifications that ensure that both classical and quantum (tunneling) regions of the dynamics are equally sampled. This gives a sampling function of the form:

$$\omega_0(R_{\text{QM}}) \propto \frac{[\tilde{\rho}(R_{\text{QM}}) + 1/I_\chi] \times [\tilde{V}'(R_{\text{QM}}) + 1/I_V]}{\tilde{V}(R_{\text{QM}}) + 1/I_V} \quad (3)$$

where  $\tilde{\rho}$ ,  $\tilde{V}'$ , and  $\tilde{V}$  are shifted, normalized, and maintained positive semidefinite<sup>51,52</sup> according to:

$$\tilde{V}(R_{\text{QM}}) \propto \frac{V(R_{\text{QM}}) - V_{\min}}{V_{\max} - V_{\min}} \quad (4)$$

and similarly for  $\tilde{\rho}(R_{\text{QM}})$  and  $\tilde{V}'(R_{\text{QM}})$ . The quantities  $V_{\max}$  and  $V_{\min}$  are the maximum and minimum values for the potential, respectively, and the overall sampling function,  $\omega_0(R_{\text{QM}})$ , is  $L^1$ -normalized according to

$$\|\omega_0(R_{\text{QM}})\|_1 = \int |\omega_0(R_{\text{QM}})| dR_{\text{QM}} = 1 \quad (5)$$

In ref,<sup>51</sup> a detailed algorithm for implementation of TDDS is provided. In addition, the stability of this algorithm is also analyzed. The choice of parameters,  $I_\chi = 1$ ,  $I_V = 3$ , and  $I_V = 1$ , retains significant distribution in both the classically allowed (minimum energy regions) and classically forbidden (classical turning point) regions of the potential and leads to a large reduction in computational cost, with little perceivable loss in accuracy. The rationale behind the choice of these parameters can be qualitatively noted from the following arguments with details in ref 51. The functions  $\tilde{\rho}$ ,  $\tilde{V}'$ , and  $\tilde{V}$  are shifted and normalized<sup>51</sup> (see eq 4), and hence, (a) minimum energy regions of the potential surface are characterized by low potential energy, low gradient, and relatively high wavepacket distribution, while (b) quantum tunneling (or classical turning point) regions of the potential are approximately characterized by moderately large values of the potential, high gradients, and smaller wavepacket values. When one enforces the condition that the TDDS function must be approximately equal in these two situations for minimal bias between the classically allowed and classically forbidden regions, it is found that  $I_\chi = 1$ ,  $I_V = 3$ , and  $I_V = 1$  provides the lowest order solution satisfying these considerations.<sup>51</sup> (Higher order solutions further increase the sampling in the classically forbidden regions.) In addition to these formal considerations, the parameters have been numerically tested in ref 51 for a set of 70 analytical and numerical potentials, and the results are found to be consistent with the above physical arguments. In ref 52, the TDDS implementation of QWAIMD has been benchmarked for accuracy in computing vibrational properties in hydrogen-bonded clusters. Specifically, the  $\text{ClHCl}^-$  system was treated since it provides significant challenges for accurate modeling of electron–nuclear coupling.<sup>52,83–85</sup> In ref 52, the TDDS implementation of QWAIMD was found to accurately reproduce the experimental spectrum at limited computational cost. The analysis of trajectories is facilitated through the introduction of a novel velocity-flux correlation function.<sup>52</sup>

The computational implementation of TDDS<sup>52</sup> is achieved as follows: For quantum dynamics beyond one dimension, the TDDS function on the full grid is evaluated at every instant in time to determine the grid points where the potential and gradients are to be obtained for the next time step. For this purpose, the TDDS function is written as a linear combination of Haar wavelets:<sup>52</sup>

$$\omega(x) \propto \frac{[\tilde{\rho} + 1/I_x] \times [\tilde{V}' + 1/I_{V'}]}{\tilde{V} + 1/I_V} = \sum_{i=0}^{N_{GEN}} \underbrace{\sum_{j_1=0}^{a^i-1} \cdots \sum_{j_{N_{Dim}}=0}^{a^i-1}}_{N_{Dim}} c_{i,\{j\}} \left\{ \prod_{k=1}^{N_{Dim}} \mathcal{H} \left( a^i x^k - \frac{j_k N_Q}{a^i} \right) \right\} \quad (6)$$

where the Haar scaling function,  $\mathcal{H}(x)$  is a square function equal to 1 for  $0 \leq x \leq 1$  and zero otherwise. The quantity  $N_{GEN}$  is the number of wavelet generations, and the underbrace below the summations is meant to indicate that there are  $N_{Dim}$  summations,  $[j_1, j_2, \dots, j_{N_{Dim}}]$ .  $c_{i,\{j\}}$  implies that the coefficients depend on  $i$  and the entire set of  $j$  indices. The Haar wavelets,  $\{\mathcal{H}(a^i x - j_k N_Q/a^i)\}$ , comprise a hierarchy of translated and dilated forms of  $\mathcal{H}(x)$ . Only the Haar scaling function is used since the Haar wavelet function is the orthogonal complement of the Haar scaling function and is not positive semidefinite, which is one of the requirements on  $\omega$ . The quantity  $x^k$ , in eq 6, is the  $k$ th component of the  $N_{Dim}$  dimensional vector, and  $a$  is chosen to be 2 or 3. That is, we employ 2- and 3-scale functions in our scheme. Once the subset of grid points for “on-the-fly” potential energy determination is computed using the TDDS function, the value of the potential at the remaining points is obtained through Hermite curve interpolation.<sup>86</sup> The forces on classical atoms are subsequently determined through a low-pass filtered Lagrange interpolation technique introduced in ref 52. Time-dependent deterministic sampling has played a pivotal role in converting QWAIMD into an efficient computational tool through reduction of computational costs by about 3 to 4 orders of magnitude.<sup>52</sup> (See Figure 1.)

It has also been numerically shown<sup>51</sup> that the TDDS function is inversely proportional to the Wentzel–Kramers–Brillouin (WKB) length scale:

$$\frac{p}{\hbar} \equiv \chi^{-1} \gg \left( \frac{1}{E - V(x)} \right) \frac{\partial V}{\partial x} \quad (7)$$

Thus, the TDDS function provides a larger sample of data points in the rapidly varying limit of the potential. Furthermore, it has been numerically shown<sup>51,52</sup> that the TDDS function is directly proportional to the Bohmian quantum potential.<sup>87–99</sup>

In addition, as noted in the Introduction, QWAIMD has been adopted to study hydrogen tunneling in enzyme active sites,<sup>55,65</sup> and QM/MM generalizations to the TDDS implementation of QWAIMD have also been completed.<sup>53</sup> In ref 100, the quantum dynamics tools from QWAIMD were used to compute the qualitative accuracy involved in classical *ab initio* molecular dynamics calculations of vibrational spectra in hydrogen bonded systems.

**II.B. Further Computational Enhancements through Diabatic Extensions to QWAIMD.** To further enhance the computational scaling of QWAIMD, in ref 54, we introduced a diabatic generalization. Essentially, multiple single particle electronic density matrices are simultaneously propagated through an extended Lagrangian scheme. Following this, the Slater determinantal wave functions associated with the density matrices are used to construct a nonorthogonal CI problem, which is computed on-the-fly to obtain the instantaneous adiabatic states. Computational efficiency arises

through the diabatic approximation for the multiple density matrices: this essentially necessitates a limited dependence of the quantum nuclear degrees of freedom on the individual electronic density matrix states. Once this condition is enforced, it is found that two-electron integrals can be reused over the entire grid, which reduces the computational complexity in determining the potential surface enormously.

As will be discussed in the next few sections, the proposed methodological extensions using Shannon’s entropy condition have multiple effects on the QWAIMD algorithm:

- An improved TDDS function is first derived and tested in section IV.A. This has direct impact on the TDDS implementation of QWAIMD.
- The TDDS functions obtained from Shannon’s entropy measure are used to locate significant regions on a potential energy surface. Grid-based electronic structure basis functions are then placed on these important regions, as discussed in section IV.B. This feature leads to two further improvements in the QWAIMD methodology:
  - The introduction of the grid based electronic structure basis functions strengthens and influences the diabatic approximation discussed above and, in further detail, in ref 54.
  - The grid-based electronic structure basis functions also reduce the computational cost of each electronic structure calculation, and this in turn has an effect on the TDDS implementation of QWAIMD.

### III. Time-Dependent Deterministic Sampling through Shannon Entropy Measure

As noted above, the physical justification for the form of TDDS is based on specific dynamical parameters (wave-packet probability density, potential, gradients), and in this section, we introduce additional sampling functions utilizing the concept of Shannon entropy. With reference to the TDDS-based implementation of QWAIMD, one particularly troublesome feature of TDDS is that sampling points can sometimes be placed in physically uninteresting regions of the potential during the dynamics simulation, in particular when both the potential energy and the gradient of the potential are high. These regions represent areas that are classically forbidden and also fail to demonstrate quantum behavior. While the TDDS function still performs remarkably well in improving efficiency with a negligible loss in accuracy,<sup>51,52</sup> the question we address in section IV.A is whether further improvements can be achieved. As will be shown in section IV.A, the new sampling functions introduced in this section provide a compressed set of sampling points and hence yield a more efficient procedure for “on-the-fly” dynamics. Furthermore, these functions also allow us to determine the positions of grid-based electronic bases in section IV.B for enhanced accuracy through diabatic extensions to QWAIMD.<sup>54</sup>



The general form of the total Shannon entropy of a system is given by<sup>68</sup>

$$-k \int_{-\infty}^{\infty} dx \rho \log(\rho) \equiv -k \int_{-\infty}^{\infty} dx S[\rho(x)] \quad (8)$$

where  $\rho$  is the probability density of the system,  $k$  is a proportionality constant, and we have defined the quantity in the integrand on the left as  $S[\rho(x)] \equiv -\rho \log(\rho)$ . Later in this publication, we refer to  $S[\rho(x)]$  as the “local Shannon entropy”. This is because, while the full sum in eq 8 is the Shannon entropy,<sup>68</sup>  $S[\rho(x)]$  implicitly depends on the local variable  $x$ . Shannon entropy has been used as a general mathematical tool to describe the information content in a system, provided there exists some probability distribution associated for the possible states of the system. A typical example of how this generic measure can apply to a physical situation is when Shannon entropy reduces to the notion of thermodynamic entropy for an ensemble of classical particles.<sup>68</sup> Given an ensemble of possible discrete microstates for a system, summation over all of these possible microstates gives the familiar thermodynamic entropy for the system, where  $k$  is now Boltzmann’s constant. Another example deals with the use of Shannon entropy in DNA and protein structure determination, and the associated definition of complexity in biological systems.<sup>70,71</sup> In this case, a certain site in a DNA sequence, or an amino acid sequence, is defined to have an entropy that reflects the probability of finding different DNA bases (or individual amino acids for proteins) at that particular site. Each site, thus, has an entropy that contributes to the complexity of the organism. Entropy here is, of course, an information entropy and not a thermodynamic entropy since it pertains to the propensity of the appearance of amino acid residues (or DNA bases) at a chosen point in the sequence. In quantum mechanics, Shannon entropy is related to von Neumann entropy, when eq 8 is rewritten as

$$-kTr[\Gamma \log(\Gamma)] = kTr[S_{vN}[\Gamma]] \quad (9)$$

where  $\Gamma \equiv |\psi\rangle\langle\psi|$  is now the density matrix associated with the system. Furthermore, in a fashion analogous to that in eq 8, we can define here a “local von Neumann entropy,”  $S_{vN}[\Gamma]$ . Along similar lines, semiclassical forms of entropy have been defined<sup>101</sup> where coherent states<sup>102</sup> have been employed for the probability function,  $\rho$ , in eq 8.

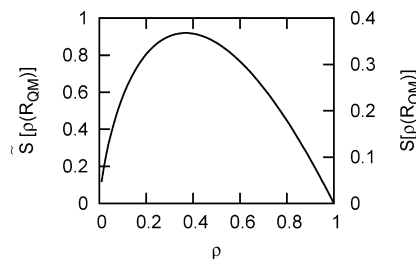
Influenced by this early work, here we utilize the local value of Shannon information entropy defined in eq 8:  $S[\rho(x)] \equiv -\rho \log(\rho)$ , where  $\rho$  is chosen as the time-dependent wavepacket density in our QWAIMD simulations, to construct suitable sampling functions of the form

$$\omega_0(R_{QM}) = \frac{(\tilde{\rho}(R_{QM}) + 1/I_\chi)(\tilde{V}'(R_{QM}) + 1/I_V)}{(\tilde{V}(R_{QM}) + 1/I_V)} \quad (10)$$

$$\omega_1(R_{QM}) = \frac{(\tilde{S}[\rho(R_{QM})] + 1/I_S)}{(\tilde{V}(R_{QM}) + 1/I_V)} \quad (11)$$

$$\omega_2(R_{QM}) = (\tilde{S}[\rho(R_{QM})] + 1/I_S) \quad (12)$$

where  $\omega_0$  is the original TDDS function and  $\omega_1$  is a composite function that utilizes the Shannon entropy as well



**Figure 2.** Behavior of  $\tilde{S}$  and  $S$  as a function of  $\rho$ .

as the potential energy. In all cases,  $\tilde{S}$  and  $\tilde{V}$  are shifted and normalized as per eq 4, and the sampling functions are scaled such that the respective values are bounded by unity (see eq 5). In addition, while the parameters  $I_\chi = 1$ ,  $I_V = 3$ , and  $I_V = 1$  define  $\omega_0$ , those for  $\omega_1$  and  $\omega_2$  are chosen as  $I_S = I_V = 1$  in this study. This allows consistent treatment of the wavepacket and the local Shannon entropy in the sampling functions. It must also be noted that the quantum mechanical Shannon entropy defined here for use in eqs 11 and 12 is a special form of the semi-classical entropy defined in ref 101. In that case, coherent states<sup>102</sup> were used to define the probability function instead of the time-dependent wavepacket density that is used here.

Before we proceed into a numerical analysis of these sampling functions, it is useful to inspect limiting cases for both  $S[\rho(R_{QM})]$  and  $\tilde{S}[\rho(R_{QM})]$ . An illustration of the behavior of  $\tilde{S}$  and  $S$  as a function of  $\rho$  is provided in Figure 2. When the probability associated with the wavepacket is high, there is information indicating the presence of the “particle” in the given region of configurational space and, hence, the entropy at that point,  $S[\rho(x)] \equiv -\rho \log(\rho) \approx 0$ . In a similar fashion, we note that when the probability is low there is information indicating the absence of the particle in the given region of configurational space, and consequently the local entropy,  $S[\rho(x)]$ , and its scaled form,  $\tilde{S}[\rho(x)]$ , are both expected to be small. Intermediate values of the probability function yield greater uncertainty in regard to the presence of the particle. This uncertainty may be qualitatively related to Shannon entropy, and as a result, the local entropy,  $S[\rho(x)]$ , and its scaled form,  $\tilde{S}[\rho(x)]$  are both higher for intermediate values of  $\rho$ .

This naturally creates the situation where a sampling function constructed from  $\tilde{S}[\rho(x)]$  alone, that is eq 12, has the effect of producing a higher distribution of sampling points in regions where the wavepacket amplitude is intermediate. While this may be desirable to represent tunneling regions, the regions that are classically populated may have larger  $\rho$  values that are not expected to be populated well enough when  $\tilde{S}[\rho(x)]$  alone is used in a sampling function. Consequently, eq 11 has been introduced as a hybrid sampling function that includes the potential to also represent the classically allowed regions. Indeed, as we will see in a later section, it is the sampling function in eq 11 that shows the best performance of the three considered above.

In the next section, we gauge the utility of these sampling functions in probing important regions of the potential

**Table 1.** Energy Conservation Data from a One-Dimensional Dynamical Treatment of the Shared Proton in  $[\text{Cl}-\text{H}-\text{Cl}]^{-a}$ 

level of theory	TDDS	$N_Q^b$	$N_E^c$	$N_Q/N_E^d$	temp (K) <sup>e</sup>	time (ps)	$\Delta E$ (kcal/mol)
HF/6-31G(d,p)	— <sup>f</sup>	101	101	1	325.26	1.0	0.03
HF/6-31G(d,p)	$\omega_0$	101	11	9.18	325.26	1.9	0.02
HF/6-31G(d,p)	$\omega_1$	101	11	9.18	318.87	1.3	0.02
HF/6-31G(d,p)	$\omega_2$	101	11	9.18	319.25	1.3	0.02
HF/6-31G(d,p)	$\omega_0$	101	9	11.22	340.85	2.5	0.13
HF/6-31G(d,p)	$\omega_1$	101	9	11.22	320.01	3.2	0.13
HF/6-31G(d,p)	$\omega_2$	101	9	11.22	337.13	3.3	0.12
HF/6-31G(d,p)	$\omega_0$	101	7	14.42	368.37	2.6	0.23
HF/6-31G(d,p)	$\omega_1$	101	7	14.42	370.04	2.7	0.30
HF/6-31G(d,p)	$\omega_2$	101	7	14.42	341.14	1.5	0.11
B3LYP/6-31+G(d,p)	—	101	101	1	258.45	1.1	0.01
B3LYP/6-31+G(d,p)	$\omega_0$	101	11	9.18	257.63	1.7	0.06
B3LYP/6-31+G(d,p)	$\omega_1$	101	11	9.18	261.08	0.4	0.00
B3LYP/6-31+G(d,p)	$\omega_2$	101	11	9.18	261.30	0.4	0.00
B3LYP/6-31+G(d,p)	$\omega_0$	101	9	11.22	261.94	2.6	0.02
B3LYP/6-31+G(d,p)	$\omega_1$	101	9	11.22	258.59	1.8	0.03
B3LYP/6-31+G(d,p)	$\omega_2$	101	9	11.22	260.45	1.8	0.02
B3LYP/6-31+G(d,p)	$\omega_1$	101	7	14.42	251.23	2.4	0.10
B3LYP/6-31+G(d,p)	$\omega_2$	101	7	14.42	256.69	4.1	0.05

<sup>a</sup> For all calculations, the quantum dynamical time step  $\Delta t_{\text{QM}} = 0.05$  fs and the classical time-step  $\Delta t_{\text{Cl}} = 0.25$  fs. <sup>b</sup> The total number of grid points. <sup>c</sup> The number of points on the grid where electronic structure calculations are performed. This set of points is obtained from TDDS and is adaptive (that is, time-dependent). <sup>d</sup> Represents the computational gain from TDDS. <sup>e</sup> Calculated from classical nuclear velocities and wavepacket kinetic energy. <sup>f</sup> No sampling. Electronic structure calculations performed on the full grid.

surface, both for quantum dynamics and electronic structure as stated above.

#### IV. Numerical Tests on Accuracy and Efficiency of the Shannon Information Entropy-Based Sampling Techniques

**IV.A. Improvements to “On-the-Fly” TDDS-Based Quantum Dynamics.** To evaluate the Shannon information entropy-based functions as effective TDDS functions, we compare the performance of  $\omega_1$  and  $\omega_2$  to that of  $\omega_0$ . QWAIMD simulations using these sampling functions were conducted on the bihalide cluster,  $[\text{Cl}-\text{H}-\text{Cl}]^{-}$ . The choice of system is based on the known challenges this system presents to accurately compute electron–nuclear coupling.<sup>52,83–85</sup> This model system has been the subject of substantial experimental and theoretical study<sup>83,103–109</sup> and has been used for previous TDDS studies under QWAIMD.<sup>51,52</sup> The bihalide system contains a shared proton undergoing exchange between donor and acceptor atoms, allowing the possibility of proton modes to couple with the other atoms in the system. Here, we utilize this system to evaluate the effectiveness of the three sampling functions presented in the previous section. For a detailed description of the vibrational properties of this system, obtained using QWAIMD, please see ref 52. The shared proton is treated using quantum dynamics, whereas all other atoms are treated with Born–Oppenheimer molecular dynamics (BOMD), as allowed within QWAIMD. The electronic structure calculations are treated with both Hartree–Fock and DFT methods. For all Hartree–Fock simulations, 6-31G(d,p) is used as the basis set, and for DFT simulations the B3LYP functional is used alongside the 6-31+G(d,p) basis set. All QWAIMD computations in this publication are performed using a development version of the Gaussian series of electronic structure codes.<sup>110</sup>

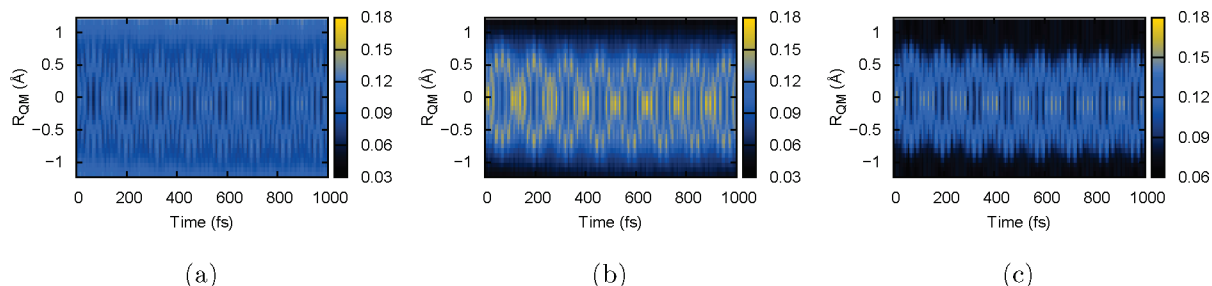
Table 1 provides a summary of energy conservation data when all three sampling functions are used with QWAIMD. While using a Hartree–Fock treatment of the electronic structure, all of the sampling functions appear to perform well with 11 sampling points per dimension leading to an order of magnitude compression of the quantum grid. But it is also noted that using seven sampling points per dimension leads to acceptable results.

Figures 3 and 4 qualitatively demonstrate the effectiveness of the Shannon entropy based sampling functions. In Figure 3, we present the evolution of all of the sampling functions computed from dynamics data calculated using  $\omega_0$ . It is already clear that there are differences in the way  $\omega_0$  samples the edges of the grid as compared to  $\omega_1$  and  $\omega_2$ . For example, note that the edges of the grid are much darker for the case of  $\omega_1$  as compared to  $\omega_0$ . This important difference is further highlighted in Figure 4, where again it is noted that  $\omega_0$  shows a higher density at the ends of the grid as compared to  $\omega_1$  and  $\omega_2$ . Furthermore, the center of the grid is sampled to a slightly greater extent by  $\omega_1$ , although all sampling functions sample this region suitably. These results are consistent with the discussions at the end of section III, where we expected  $\omega_1$  to provide a greater sample in the classically allowed regions as compared to  $\omega_2$ . However, the fact that both  $\omega_1$  and  $\omega_2$  provide a reduced sampling at the grid edges arises due to there being no functional dependence on  $V'$  in the cases of  $\omega_1$  and  $\omega_2$ .

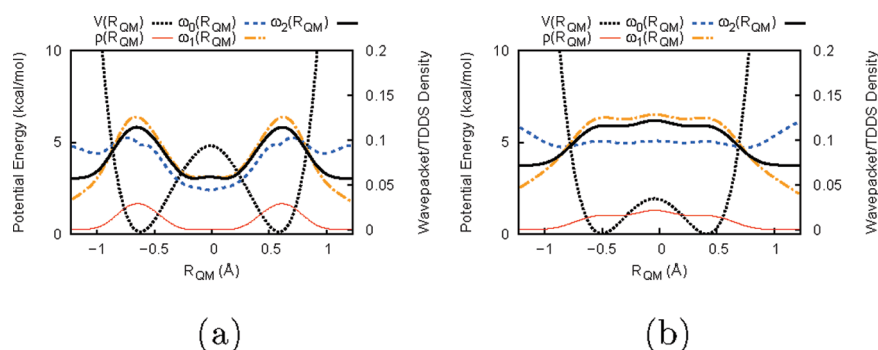
To further quantify the differences between the sampling functions, the overlapping regions between the sampling functions are calculated at each step using

$$\omega'_i(x_j;t) = \omega_i(x_j;t) \omega_0(x_j;t) \quad (13)$$

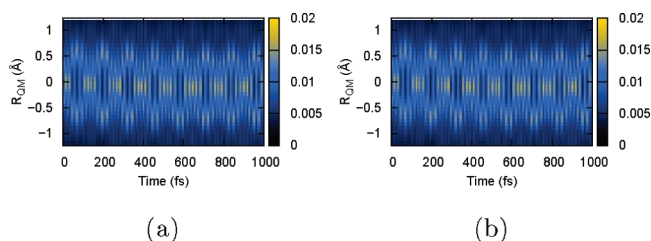
where  $x_j$  is a particular grid point, and  $i = 1$  and  $2$ ; i.e.,  $\omega_i$  above represents one of the Shannon entropy based functions. The evolution of eq 13, provided in Figure 5, shows the common and uncommon regions sampled as the functions



**Figure 3.** A comparison of the sampling functions,  $\omega_0$  (a),  $\omega_1$  (b), and  $\omega_2$  (c). The figures depict the evolution of the sampling functions during a single reference dynamics trajectory. The intensity of  $\omega_0$  is relatively high at the edges of the grid as compared to both  $\omega_1$  and  $\omega_2$ . Similarly, the intensity of  $\omega_1$  is higher in the important regions as compared to both  $\omega_0$  and  $\omega_2$ . Note that this is not a comparison of actual dynamical data.  $\omega_1$  and  $\omega_2$  were reconstructed using dynamics data performed with  $\omega_0$ .



**Figure 4.** (a) A representative time slice of the sampling functions,  $\rho$ , and the potential from a one-dimensional Hartree–Fock simulation. (b) The time-averaged behavior. Again, as already seen in Figure 3, the fact that  $\omega_0$  overestimates the significance of the edges of the grid is clearly noted. Furthermore,  $\omega_1$  has higher intensities in the important regions, consistent with Figure 3.

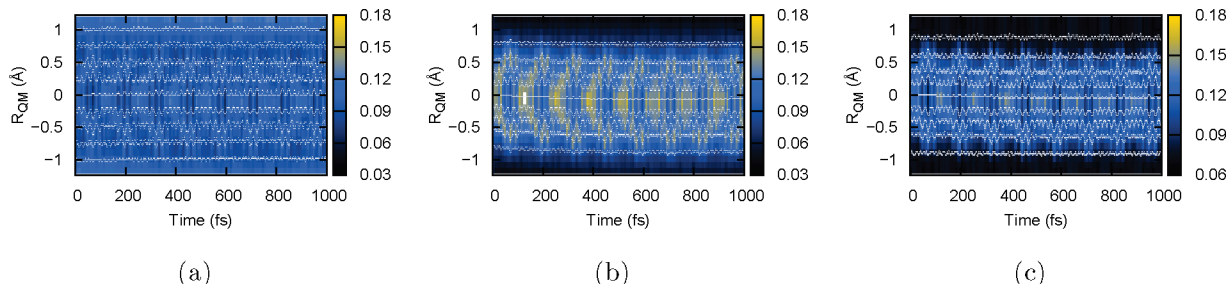


**Figure 5.** Evolution of eq 13 for  $\omega_1$  (a) and  $\omega_2$  (b). The common regions between  $\omega_0$  and the new TDDS functions are represented by regions of high intensity. These are located in the central regions of the grid. The lower intensity at the edges demonstrates the crucial difference between the two sets of sampling functions.

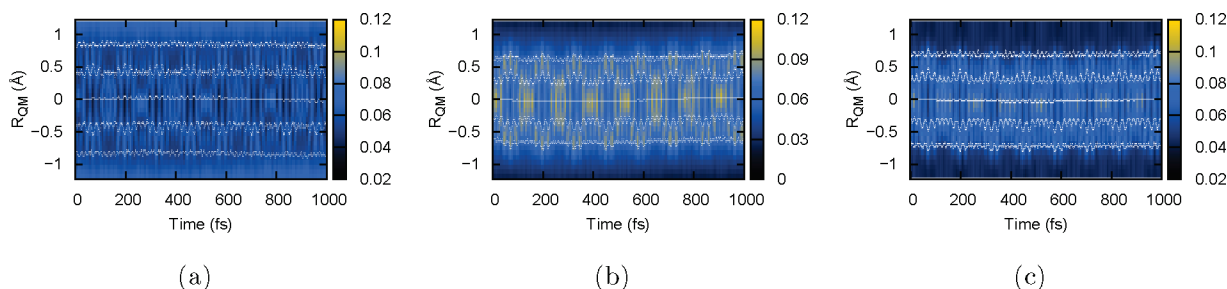
change form due to the dynamics. Regions of high density of the overlap measure correspond to regions of high commonality, while low density shows areas of divergent behavior between the two sampling functions. As seen in the time-averaged behavior of the sampling functions in Figure 4b, the common regions are contained in the center of the grid. However, now the time-dependency of this relationship is revealed, and major oscillations in these functions are preserved. Furthermore, the fact that edges of the grid are lighter in Figure 5b as compared to Figure 5a indicates a greater sampling of the grid edges for the case of  $\omega_2$  as compared to  $\omega_1$ .

In the TDDS algorithm, once the sampling function is constructed, a Haar wavelet fit of this function is discretized to obtain points in configurational space for electronic structure calculations. (The detailed algorithm is presented in refs 52 and 51.) The discrete, time-dependent set of points obtained from such an algorithm when using the functions,  $\omega_0$ ,  $\omega_1$ , and  $\omega_2$ , are shown in Figures 6 and 7. Consistent with the previous discussion,  $\omega_1$  provides the most compressed representation of the grid populating only the important regions. It is followed very closely by  $\omega_2$ , and  $\omega_0$  provides a greater sampling of points at the edges of the grid. In addition, the fundamental oscillations near the center of the grid are captured by all three functions, but these oscillations are more intense for the Shannon sampling-based functions. In section IV.B, this property is used to construct a set of potential adapted, grid-based electronic structure basis functions. That is, in section IV.B, electronic structure basis functions are to be placed along the grid lines seen in Figures 6 and 7 for potential energy calculations. Such a basis set is found to be accurate and efficient and is used in ref 54 to further enhance the computational efficiency of QWAIMD.

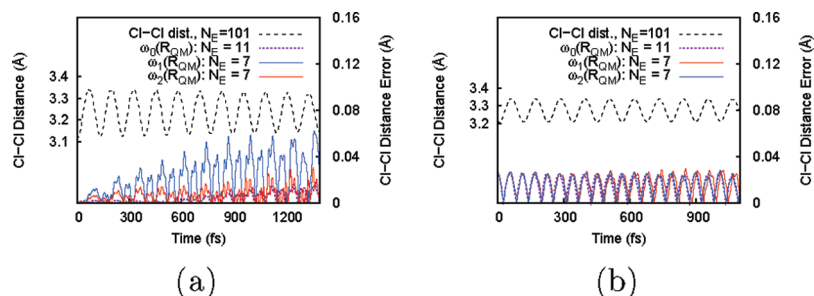
Having examined the differences between the sampling functions, it is important to see how these directly affect observables in the dynamics. Thus, we conclude this section with an analysis of the vibrational effects on the



**Figure 6.** Time-evolution of sampling points (white lines), compared to sampling function density (blue and yellow density map), for  $N_E = 11$ .  $\omega_0$  is shown in a,  $\omega_1$  in b, and  $\omega_2$  in c.



**Figure 7.** Time-evolution of sampling points (white lines), compared to sampling function density (blue and yellow density map), for  $N_E = 7$ .  $\omega_0$  is shown in a,  $\omega_1$  in b, and  $\omega_2$  in c.



**Figure 8.** Error in the Cl–Cl distance. (a) Hartree–Fock simulation. (b) DFT calculation (B3LYP). The dotted black lines in both parts display the evolution of the Cl–Cl distance (left axis). The error in the Cl–Cl distance is shown on the right axis and depicted using the red and blue lines. The Cl–Cl oscillations for Hartree–Fock have a larger amplitude due to higher temperatures for the associated simulations. (Please see Table 1.)

classical atoms. The root mean squared error in the Cl–Cl distance is shown in Figure 8. The error was referenced to a QWAIMD simulation in which no interpolation of the potential and gradients was used. In all cases, the Shannon entropy based sampling functions are able to reproduce the oscillations with fewer sampling points. This, of course, is a result of the more compact nature of these sampling functions. The oscillation frequencies are in agreement with previous calculations,<sup>52</sup> but the result in Figure 8 indicates a reduced computational cost when using  $\omega_1$  and  $\omega_2$ .

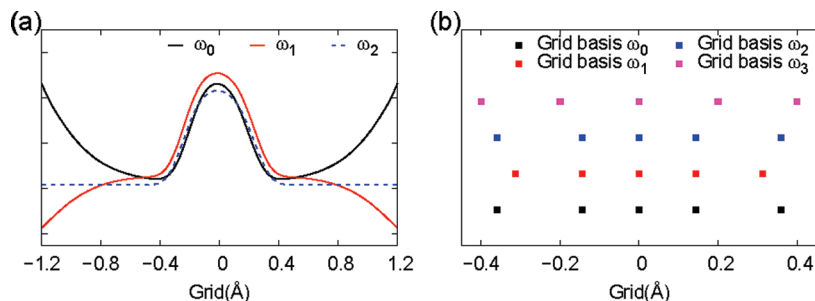
**IV.B. Locating Regions for Potential-Adapted, Grid-Based Electronic Structure Basis Functions Using the Shannon Entropy Based Sampling Functions.** In this section, we utilize the TDDS functions to obtain a grid-based description of electronic structure. This study is particularly relevant for hydrogen-bonded systems, and we show here that accurate potential energy surfaces can be obtained over a wide range of energies and nuclear geometries when grid-

based Gaussian basis functions, directed using TDDS, are utilized. Essentially, the question we pose is, if Gaussian basis functions of the kind

$$\chi_{l,m,n}^{\mathbf{R}_F}(\mathbf{r}) = (x - R_x)^l (y - R_y)^m (z - R_z)^n \exp[-\alpha(\mathbf{r} - \mathbf{R}_F)^2] \quad (14)$$

were directed such that the basis functions centers,  $\mathbf{R}_F[\equiv(R_x, R_y, R_z)]$ , were chosen to be functions of multiple classical nuclear variables according to  $\mathbf{R}_F = f(\{\mathbf{R}_C\})$  and the centers are determined using the sampling functions, can this improve efficiency while retaining the accuracy of electronic structure calculations? In eq 14, the quantities  $l$ ,  $m$ , and  $n$  are the usual orbital angular momentum indices of the basis function. The result of this discussion is a generalization of bond-centered basis functions<sup>111–114</sup> traditionally used in quantum chemistry where the positions of these Gaussian basis functions are determined using the TDDS functions discussed in section III. Furthermore, these





**Figure 9.** (a) The TDDS functions and (b) the associated origins for the grid-based electronic basis functions obtained from the TDDS functions. The system under study is  $[\text{ClHCl}]^{-1}$ , and the horizontal axis for both figures represents  $(R_{\text{H-Cl}_1} - R_{\text{H-Cl}_2})/2$ . The definitions for  $R_{\text{H-Cl}_1}$  and  $R_{\text{H-Cl}_2}$  are provided in Figure 10. Note that b shows a smaller spatial region since this is the predominant area for  $\omega_1$ . Note also that the functions  $\omega_0$  and  $\omega_2$  place a greater weight at the edges, which is consistent with our earlier discussion in section IV.A.

grid-based electronic functions are used in ref 54 to further improve the efficiency of QWAIMD.

For the case of hydrogen-bonded systems, we specialize our definition of  $\mathbf{R}_F \equiv f(\{\mathbf{R}_C\})$  to a function of the donor and acceptor coordinates:

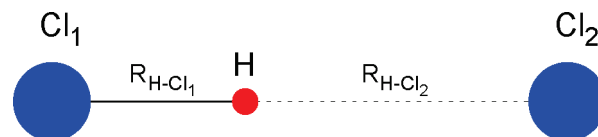
$$\mathbf{R}_F^i = \sum_j c_{ji} \mathbf{R}_C^j + \bar{\mathbf{v}}_i = a_i \mathbf{R}_A + d_i \mathbf{R}_D + \bar{\mathbf{v}}_i \quad (15)$$

where  $\mathbf{R}_A$  and  $\mathbf{R}_D$  are coordinate vectors of the donor and acceptor atoms for a hydrogen-bonded system and  $\bar{\mathbf{v}}_i$  is a uniform shift that can be used to create a three-dimensional grid of electronic basis functions. It is further important to note that the basis functions introduced in eq 15 are functions of classical nuclear coordinates. Hence, in a fashion similar to atom-centered basis functions, the centers of these functions also transform according to the classical nuclear positions. Furthermore, these grid-based functions are spread uniformly in space. But these functions differ from plane-waves<sup>115</sup> through the  $\{\mathbf{R}_F\}$  dependence of the Fourier transforms.

To choose the variables  $\{a_i, d_i, \bar{\mathbf{v}}_i\}$ , we utilize the sampling functions discussed earlier. Our test case involves three well-studied hydrogen-bonded ion clusters:<sup>83,103,107,108,116–124</sup> the bihalide cluster  $[\text{ClHCl}]^{-1}$ , the hydroxide water cluster  $[\text{OH-H}_2\text{O}]^{-1}$ , and the Zundel cation  $[\text{H}_2\text{O-H-H}_2\text{O}]^+$ . Our goal is to find the optimum number and associated positions of the grid-based basis functions in the bonding region of the transferring hydrogen. The stability and vibrational properties of the clusters discussed here are sensitive to the potential surface along the donor–acceptor axis. Thus, potential energy surfaces constructed on a one-dimensional grid were compared between grid-based basis and atom-centered basis set aug-cc-pvtz.<sup>125–127</sup> To quantify the errors, we define

$$\Delta V(\varepsilon_1, \varepsilon_2) = \sqrt{\frac{\sum_i [V_1(\mathbf{R}_C, \mathbf{R}_{QM}^i) - V_2(\mathbf{R}_C, \mathbf{R}_{QM}^i)]^2 \prod_{\varepsilon_1, \varepsilon_2} (V(\mathbf{R}_{QM}^i))}{\sum_i \prod_{\varepsilon_1, \varepsilon_2} (V(\mathbf{R}_{QM}^i))}} \quad (16)$$

where the boxcar function is defined as linear combination of Heaviside functions:  $\Pi_{\varepsilon_1, \varepsilon_2}(V) = H(V - \varepsilon_1) - H(V - \varepsilon_2)$ .



**Figure 10.** The parameters  $R_{\text{H-Cl}_1}$  and  $R_{\text{H-Cl}_2}$  are defined here and used in Figure 9.

Equation 16 allows us to inspect the accuracy in the potential surface in a tiered fashion by focusing on specific energy domains. We have utilized the three functions presented in eqs 10, 11, and 12 and compared the associated behavior with a uniform distribution function:  $\omega_3 = 1$ .

In Figure 9, distributions of the potential-adapted, grid-based basis using various TDDS schemes are presented. Compared to the uniform sampling function,  $\omega_3$ , the Shannon entropy based TDDS function reduces the population of electronic basis functions close to the edge of the grid. The standard TDDS function,  $\omega_0$ , on the contrary, places roughly equal weight at both the middle and edge of the grid. This, of course, is to be expected, since TDDS in eq 10 has been tuned such that the grid-based basis is distributed equally in both classically allowed and forbidden regions. However, bases at the edges (high gradients and large values of the potential) may not be useful during the electronic structure calculations, and hence, in practice one might expect the Shannon information based TDDS functions to be more efficient.

A detailed examination of the accuracy of various TDDS functions in obtaining good estimates for the potential surface is provided in Table 2. The error estimates utilized are those discussed in eq 14. Since a large number of grid-based basis functions are distributed in the bonding region of the hydrogen-bonded systems considered, a relatively small basis set (3-21G and STO-3G) is used at each grid point. The accuracy of the potential-adapted, grid-based basis functions is ascertained through comparison with a standard atom-centered aug-cc-pvtz basis. To perform the benchmark in a tiered fashion, we first replace the atom-centered aug-cc-pvtz on the shared proton with grid-based basis functions while retaining the aug-cc-pvtz bases on all of the other classical atoms. Following this, the aug-cc-pvtz bases on classical atoms are substituted with the

**Table 2.** Benchmarks for Grid-Based Electronic Structure Basis Functions

system	atom centered basis		grid basis			error in PES <sup>a</sup>		TDDS
	basis set	$N_{\text{basis}}^b$	quantum/classical <sup>c</sup>	$N_{\text{use}}/N_{\text{basis}}^d$	$N_{\text{grid-basis}}^e$	$\Delta V^f$	$\Delta V^g$	
[CIHCl] <sup>-1</sup> <sup>h</sup>	aug-cc-pvtz	<b>123</b>	STO-3G/aug-cc-pvtz	113/121	21	0.20	0.68	$\omega_3$
			STO-3G/aug-cc-pvtz	111/111	11	0.25	0.95	$\omega_3$
			STO-3G <sup>i</sup> /aug-cc-pvtz	109/109	9	0.44	1.22	$\omega_3$
			STO-3G/aug-cc-pvtz	111/111	11	0.26	0.92	$\omega_0$
			STO-3G/aug-cc-pvtz	110/111	11	0.23	0.88	$\omega_1$
			<b>STO-3G/aug-cc-pvtz</b>	<b>109/109</b>	<b>9</b>	<b>0.25</b>	<b>0.87</b>	<b><math>\omega_1</math></b>
			3-21G/6-31+G**	66/86	42	0.23	0.90	$\omega_3$
			3-21G/6-31+G**	61/66	22	0.22	1.06	$\omega_3$
			STO-3G/6-31+G**	57/65	21	0.20	0.84	$\omega_3$
			STO-3G/6-31+G**	55/55	11	0.20	1.15	$\omega_3$
			STO-3G <sup>j</sup> /6-31+G**	53/53	9	0.44	1.46	$\omega_3$
			STO-3G/6-31+G**	55/55	11	0.26	1.06	$\omega_0$
			STO-3G/6-31+G**	54/55	11	0.27	1.10	$\omega_1$
			<b>STO-3G/6-31+G**</b>	<b>53/53</b>	<b>9</b>	<b>0.27</b>	<b>1.11</b>	<b><math>\omega_1</math></b>
[OH-H <sub>2</sub> O] <sup>-1</sup> <sup>j</sup>	aug-cc-pvtz	<b>161</b>	STO-3G/aug-cc-pvtz	147/147	9	0.34	0.84	$\omega_3$
			STO-3G/aug-cc-pvtz	146/147	9	0.43	0.69	$\omega_0$
			STO-3G/aug-cc-pvtz	145/147	9	0.36	0.79	$\omega_1$
			<b>STO-3G/aug-cc-pvtz</b>	<b>145/145</b>	<b>7</b>	<b>0.23</b>	<b>0.85</b>	<b><math>\omega_1</math></b>
			STO-3G/6-31+G**	55/55	9	0.44	1.55	$\omega_3$
			STO-3G/6-31+G**	55/55	9	0.46	1.44	$\omega_0$
			STO-3G/6-31+G**	53/55	9	0.48	2.11	$\omega_1$
			<b>STO-3G/6-31+G**</b>	<b>53/53</b>	<b>7</b>	<b>0.51</b>	<b>2.10</b>	<b><math>\omega_1</math></b>
[H <sub>2</sub> O-H-H <sub>2</sub> O] <sup>+1</sup> <sup>j</sup>	aug-cc-pvtz	<b>207</b>	STO-3G/aug-cc-pvtz	193/193	9	0.23	0.91	$\omega_3$
			STO-3G/aug-cc-pvtz	192/193	9	0.26	0.86	$\omega_0$
			STO-3G/aug-cc-pvtz	191/193	9	0.21	0.89	$\omega_1$
			<b>STO-3G/aug-cc-pvtz</b>	<b>190/191</b>	<b>7</b>	<b>0.24</b>	<b>0.86</b>	<b><math>\omega_1</math></b>
			STO-3G/6-31+G**	65/65	9	0.41	1.56	$\omega_3$
			STO-3G/6-31+G**	64/65	9	0.32	1.42	$\omega_0$
			STO-3G/6-31+G**	63/65	9	0.35	1.84	$\omega_1$
			<b>STO-3G/6-31+G**</b>	<b>62/63</b>	<b>7</b>	<b>0.43</b>	<b>2.05</b>	<b><math>\omega_1</math></b>

<sup>a</sup> Errors (in kcal/mol) are compared between atom-centered aug-cc-pvtz and grid-based basis results using eq 16. All PESs are obtained at the Hartree–Fock level. <sup>b</sup> Number of basis functions in the atom-centered basis calculations. <sup>c</sup> The shared proton is treated quantum mechanically, and all other atoms are classical. The column represents the grid basis used (around the quantum nucleus) and the atom-centered basis placed on each classical atom. <sup>d</sup>  $N_{\text{basis}}$  is the total number of basis functions, whereas  $N_{\text{use}}$  is the number of linearly independent basis functions. <sup>e</sup> Number of basis functions used for grid-based basis. <sup>f</sup>  $\epsilon_1 = 0.0$ ,  $\epsilon_2 = 2.5$ . Unit is kcal/mol. See eq 16. <sup>g</sup>  $\epsilon_1 = 2.5$ ,  $\epsilon_2 = 15.0$ . Unit is kcal/mol. See eq 16. <sup>h</sup> Optimized geometry using MP2/aug-cc-pvtz. Cl–Cl distance is 3.13 Å. <sup>i</sup> Optimized geometry using MP2/aug-cc-pvtz. Oxygen–oxygen distance is 2.48 Å. <sup>j</sup> Optimized geometry using MP2/aug-cc-pvtz. Oxygen–oxygen distance is 2.39 Å.

relatively small double split valence 6-31+G\*\* basis. All results are summarized in Table 2.

Although all TDDS schemes give accurate results in the low energy regions, as seen from the smaller values of  $\Delta V$  in the column using  $\epsilon_1 = 0.0$  kcal/mol and  $\epsilon_2 = 2.5$  kcal/mol, the Shannon entropy based TDDS function,  $\omega_1$ , provides higher accuracy while using fewer basis functions. (These are shown in blue in Table 2.) The reduction in the number of grid-based basis functions is especially striking in this case where the number of basis functions required is reduced to roughly half in the case of [CIHCl]<sup>-1</sup> and a third in the case of the larger systems. There appears to be little loss in accuracy over the entire grid. Due to the  $O(N^3)$  scaling of the algorithms involved, this leads to a factor of 8 reduction in computation time for the smaller [CIHCl]<sup>-1</sup> system and a factor of 27 reduction in computation time for the larger systems. In ref 54, these potential-adapted grid-based electronic basis functions are utilized to facilitate an even larger reduction in computation time when employed in conjunction with new formalisms of QWAIMD.

## V. Conclusions

A new set of time-dependent deterministic sampling functions based on Shannon's entropy were introduced. These

functions were used to probe important regions of an electronic potential surface and to facilitate computational improvements in quantum-classical dynamics of electrons and nuclei. Computational gains are two-fold as discussed in the numerical results section: The direct implementation of Shannon entropy based TDDS functions reduces computational cost by eliminating the need for sampling points in physically uninteresting regions of the potential surface. In addition, when the Shannon entropy based TDDS functions are utilized to construct a potential-adapted grid-based electronic basis set, the accuracy of the electronic potential surface is well-preserved, while the computational cost is significantly lowered. This idea is further exploited in ref 54 to facilitate the development of a new QWAIMD formalism that reduces computational costs by several orders of magnitude.

**Acknowledgment.** This research is supported by the National Science Foundation, grant number CHE-0750326 to S.S.I.

## References

- (1) Wyatt, R. E.; Zhang, J. Z. H. *Dynamics of Molecules and Chemical Reactions*; Marcel Dekker Inc.: New York, 1996.

- (2) Berne, B. J.; Ciccotti, G.; Coker, D. F. *Classical and Quantum Dynamics in Condensed Phase Simulations*; World Scientific: Singapore, 1997.
- (3) Wang, I. S. Y.; Karplus, M. *J. Am. Chem. Soc.* **1973**, *95*, 8160.
- (4) Leforestier, C. *J. Chem. Phys.* **1978**, *68*, 4406.
- (5) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471.
- (6) Bolton, K.; Hase, W. L.; Peslherbe, G. H. World Scientific: Singapore, 1998; Chapter: Direct Dynamics of Reactive Systems, p 143.
- (7) Schlegel, H. B.; Millam, J. M.; Iyengar, S. S.; Voth, G. A.; Daniels, A. D.; Scuseria, G. E.; Frisch, M. J. *J. Chem. Phys.* **2001**, *114*, 9758.
- (8) Schatz, G. C.; Kupperman, A. *J. Chem. Phys.* **1976**, *65*, 4642.
- (9) Delos, J. B. *Rev. Mod. Phys.* **1981**, *53*, 287.
- (10) Feit, M. D.; Fleck, J. A. *J. Chem. Phys.* **1982**, *78*, 301.
- (11) Kosloff, R. *Annu. Rev. Phys. Chem.* **1994**, *45*, 145.
- (12) Leforestier, C.; Bisseling, R. H.; Cerjan, C.; Feit, M. D.; Freisner, R.; Guldberg, A.; Hammerich, A.; Jolicard, D.; Karrlein, W.; Meyer, H. D.; Lipkin, N.; Roncero, O.; Kosloff, R. *J. Comput. Phys.* **1991**, *94*, 59.
- (13) DeVries, P. *Atomic and molecular processes with short intense laser pulses*; Bandrauk, A. D., Ed.; Plenum Press: New York, 1988; Vol. 171 of NATO ASI Series B, Physics, p 481.
- (14) Jang, H. W.; Light, J. C. *J. Chem. Phys.* **1995**, *102*, 3262–3268.
- (15) Althorpe, S. C.; Clary, D. C. *Annu. Rev. Phys. Chem.* **2003**, *54*, 493–529.
- (16) Huang, Y.; Iyengar, S. S.; Kouri, D. J.; Hoffman, D. K. *J. Chem. Phys.* **1996**, *105*, 927.
- (17) Miller, W. H.; Schwartz, S. D.; Tromp, J. W. *J. Chem. Phys.* **1983**, *79*, 4889.
- (18) Makri, N. *Comput. Phys. Commun.* **1991**, *63*, 389–414.
- (19) Cao, J.; Voth, G. A. *J. Chem. Phys.* **1994**, *100*, 5093.
- (20) Cao, J.; Voth, G. A. *J. Chem. Phys.* **1994**, *100*, 5106.
- (21) Cao, J.; Voth, G. A. *J. Chem. Phys.* **1994**, *101*, 6168.
- (22) Jang, S.; Voth, G. A. *J. Chem. Phys.* **1999**, *111*, 2357.
- (23) Jang, S.; Voth, G. A. *J. Chem. Phys.* **1999**, *111*, 2371.
- (24) Feit, M. D.; Fleck, J. A. *J. Chem. Phys.* **1983**, *79*, 301.
- (25) Feit, M. D.; Fleck, J. A. *J. Chem. Phys.* **1984**, *80*, 2578.
- (26) Kosloff, D.; Kosloff, R. *J. Comput. Phys.* **1983**, *52*, 35.
- (27) Kosloff, D.; Kosloff, R. *J. Chem. Phys.* **1983**, *79*, 1823.
- (28) Tal-Ezer, H.; Kosloff, R. *J. Chem. Phys.* **1984**, *81*, 3967.
- (29) Hartke, B.; Kosloff, R.; Ruhman, S. *Chem. Phys. Lett.* **1986**, *158*, 223.
- (30) Iyengar, S. S.; Kouri, D. J.; Hoffman, D. K. *Theor. Chem. Acc.* **2000**, *104*, 471.
- (31) Lill, J. V.; Parker, G. A.; Light, J. C. *Chem. Phys. Lett.* **1982**, *89*, 483.
- (32) Light, J. C.; Hamilton, I. P.; Lill, J. V. *J. Chem. Phys.* **1985**, *82*, 1400.
- (33) Colbert, D. T.; Miller, W. H. *J. Chem. Phys.* **1992**, *96*, 1982–1991.
- (34) Huang, Y.; Kouri, D. J.; Arnold, M.; Thomas, L.; Marchioro, I.; Hoffman, D. K. *Comput. Phys. Commun.* **1994**, *80*, 1.
- (35) Deumens, E.; Diz, A.; Longo, R.; Öhrn, Y. *Rev. Mod. Phys.* **1994**, *66*, 917.
- (36) Hack, M. D.; Truhlar, D. G. *J. Phys. Chem. A* **2000**, *104*, 7917–7926.
- (37) Miller, W. H. *J. Phys. Chem. A* **2001**, *105*, 2942–2955.
- (38) Heller, E. J. *J. Chem. Phys.* **1975**, *62*, 1544–1555.
- (39) Fiete, G. A.; Heller, E. J. *Phys. Rev. A* **2003**, *68*, 022112.
- (40) Hammes-Schiffer, S.; Tully, J. *J. Chem. Phys.* **1994**, *101*, 4657–4667.
- (41) Martinez, T. J.; Ben-Nun, M.; Ashkenazi, G. *J. Chem. Phys.* **1996**, *104*, 2847.
- (42) Micha, D. A. *J. Phys. Chem. A* **1999**, *103*, 7562–7574.
- (43) Payne, M. C.; Teter, M. P.; Allan, D. C.; Arias, T. A.; Joannopoulos, J. D. *Rev. Mod. Phys.* **1992**, *64*, 1045.
- (44) Marx, D.; Hutter, J. John von Neumann Institute for Computing: Jülich, Germany, 2000; Chapter: Ab Initio Molecular Dynamics: Theory and Implementation, Vol. 1, pp 301–449.
- (45) Schlegel, H. B. *J. Comput. Chem.* **2003**, *24*, 1514–1527.
- (46) Iyengar, S. S.; Jakowski, J. *J. Chem. Phys.* **2005**, *122*, 114105.
- (47) Pavese, M.; Berard, D. R.; Voth, G. A. *Chem. Phys. Lett.* **1999**, *300*, 93–98.
- (48) Tuckerman, M. E.; Marx, D. *Phys. Rev. Lett.* **2001**, *86*, 4946–4949.
- (49) Chen, B.; Ivanov, I.; Klein, M. L.; Parrinello, M. *Phys. Rev. Lett.* **2003**, *91*, 215503.
- (50) Iyengar, S. S. *Theor. Chem. Acc.* **2006**, *116*, 326.
- (51) Jakowski, J.; Sumner, I.; Iyengar, S. S. *J. Chem. Theory Comput.* **2006**, *2*, 1203–1219.
- (52) Sumner, I.; Iyengar, S. S. *J. Phys. Chem. A* **2007**, *111*, 10313–10324.
- (53) Sumner, I.; Iyengar, S. S. *J. Chem. Phys.* **2008**, *129*, 054109.
- (54) Li, X.; Iyengar, S. S. *J. Chem. Phys.* **2010**, *132*, 184105.
- (55) Iyengar, S. S.; Sumner, I.; Jakowski, J. *J. Phys. Chem. B* **2008**, *112*, 7601.
- (56) Iyengar, S. S. *Int. J. Quantum Chem.* **2009**, *109*, 3798.
- (57) Tully, J. C. *Faraday Discuss.* **1998**, *110*, 407–419.
- (58) Kapral, R.; Ciccotti, G. *J. Chem. Phys.* **1999**, *110*, 8919.
- (59) Horenko, I.; Salzmann, C.; Schmidt, B.; Schutte, C. *J. Chem. Phys.* **2002**, *117*, 11075–11088.
- (60) Donoso, A.; Zheng, Y. J.; Martens, C. C. *J. Chem. Phys.* **2003**, *119*, 5010.
- (61) Brooksby, C.; Prezhdo, O. V. *Chem. Phys. Lett.* **2001**, *346*, 463–469.
- (62) Prezhdo, O. V.; Brooksby, C. *Phys. Rev. Lett.* **2000**, *86*, 3215–3219.
- (63) Gindensperger, E.; Meier, C.; Beswick, J. A. *J. Chem. Phys.* **2000**, *113*, 9369.
- (64) Li, X.; Iyengar, S. S. <http://www.indiana.edu/ssweb/papers/sadaftp.pdf>. Submitted to *J. Phys. Chem. A*.
- (65) Sumner, I.; Iyengar, S. S. *J. Chem. Theory Comput.* **2010**, *6*, 1698.

- (66) Pacheco, A.; Iyengar, S. S. *J. Chem. Phys.* **2010**, *133*, 044105.
- (67) Pacheco, A.; Iyengar, S. S. *J. Chem. Phys.* In Press. <http://www.indiana.edu/ssiweb/papers/msaiwd2.pdf> (accessed Dec 2010).
- (68) Shannon, C. *Bell Syst. Tech. J.* **1948**, *27*, 279–423.
- (69) Shannon, C. *Proc. IEEE* **1998**, *86*, 447.
- (70) C, C. A.; Ofria, C.; Collier, T. C. *Proc. Natl. Acad. Sci.* **2000**, *97*, 4463.
- (71) Schneider, T. D.; Stormo, G. D.; Gold, L.; Ehrenfeucht, A. *J. Mol. Biol.* **1986**, *188*, 415.
- (72) Wehner, S.; Winter, A. *J. Math. Phys.* **2008**, *49*, 062105.
- (73) Iyengar, S. S.; Frisch, M. J. *J. Chem. Phys.* **2004**, *121*, 5061.
- (74) Iyengar, S. S.; Schlegel, H. B.; Millam, J. M.; Voth, G. A.; Scuseria, G. E.; Frisch, M. J. *J. Chem. Phys.* **2001**, *115*, 10291.
- (75) Trotter, M. F. *Proc. Am. Math. Soc.* **1959**, *10*, 545.
- (76) Nelson, E. *J. Math. Phys.* **1964**, *5*, 332.
- (77) Strang, G. *SIAM J. Numer. Anal.* **1968**, *5*, 506–516.
- (78) Korevaar, J. *Am. Math. Soc. Trans.* **1959**, *91*, 53–101.
- (79) Kouri, D. J.; Huang, Y.; Hoffman, D. K. *Phys. Rev. Lett.* **1995**, *75*, 49–52.
- (80) Hoffman, D. K.; Nayar, N.; Sharafeddin, O. A.; Kouri, D. J. *J. Phys. Chem.* **1991**, *95*, 8299.
- (81) Yu, S.; Zhao, S.; Wei, G. W. *J. Comput. Phys.* **2005**, *206*, 727–780.
- (82) Feynman, R. P.; Hibbs, A. R. *Quantum Mechanics and Path Integrals*; McGraw-Hill Book Company: New York, 1965.
- (83) Swalina, C.; Hammes-Schiffer, S. *J. Phys. Chem. A* **2005**, *109*, 10410.
- (84) Gerber, R. B.; Ratner, M. A. *J. Chem. Phys.* **1988**, *70*, 97–132.
- (85) Matsunaga, N.; Chaban, G. M.; Gerber, R. B. *J. Chem. Phys.* **2002**, *117*, 3541.
- (86) Bartels, R. H.; Beatty, J. C.; Barsky, B. A. *An Introduction to Splines for use in computer graphics and geometric modeling*; Morgan Kaufman Publishers, Inc.: Los Altos, CA, 1987.
- (87) Sakurai, J. J. *Modern Quantum Mechanics*; Addison-Wesley Publishing Company: Reading, MA, 1994.
- (88) Madelung, E. *Z. Phys.* **1926**, *40*, 322–326.
- (89) de Broglie, L. *An introduction to the study of wave mechanics*; E. P. Dutton and Company, Inc.: New York, 1930.
- (90) Bohm, D. *Quantum Theory*; Prentice-Hall Inc.: New York, 1951.
- (91) Bohm, D. *Phys. Rev.* **1952**, *85*, 166.
- (92) Cushing, J. T.; Fine, A.; Goldstein, S. *Bohmian Mechanics: An appraisal*; Kluwer: Boston, 1996.
- (93) Holland, P. R. *The Quantum Theory of Motion*; Cambridge, New York, 1993.
- (94) Lopreore, C. L.; Wyatt, R. E. *Phys. Rev. Lett.* **1999**, *82*, 5190.
- (95) Day, B. K.; Askar, A.; Rabitz, H. A. *J. Chem. Phys.* **1998**, *109*, 8770.
- (96) Wyatt, R. E.; Kouri, D. J.; Hoffman, D. K. *J. Chem. Phys.* **2000**, *112*, 10730.
- (97) Bittner, E. R.; Wyatt, R. E. *J. Chem. Phys.* **2000**, *113*, 8888.
- (98) Wyatt, R. E.; Bittner, E. R. *Comput. Sci. Eng.* **2003**, *5*, 22–30.
- (99) Iyengar, S. S.; Schlegel, H. B.; Voth, G. A. *J. Phys. Chem. A* **2003**, *107*, 7269–7277.
- (100) Li, X.; Oomens, J.; Eyler, J. R.; Moore, D. T.; Iyengar, S. S. *J. Chem. Phys.* **2010**, *132*, 244301.
- (101) Wehrl, A. *Rev. Mod. Phys.* **1978**, *50*, 221.
- (102) Klauder, J. R.; Skagerstam, B.-S. *Coherent States: Applications in Physics and Mathematical Physics*; World Scientific Publishing Company, Inc.: River Edge, NJ, 1985.
- (103) Kawaguchi, K. *J. Chem. Phys.* **1988**, *88*, 4186–4189.
- (104) Swalina, C.; Pak, M. V.; Hammes-Schiffer, S. *Chem. Phys. Lett.* **2005**, *404*, 394.
- (105) McCoy, A. B. *J. Chem. Phys.* **1995**, *103*, 986.
- (106) Botschwina, P.; Sebal, P.; Burmeister, R. *J. Chem. Phys.* **1988**, *88*, 5246.
- (107) Metz, R. B.; Kitsopoulos, T.; Weaver, A.; Neumark, D. *J. Chem. Phys.* **1988**, *88*, 1463.
- (108) Del Popolo, M. G.; Kohanoff, J.; Lynden-Bell, R. M. *J. Phys. Chem. B* **2006**, *110*, 8798.
- (109) Mo, O.; Yanez, M.; Del Bene, J. E.; Alkorta, L.; Elguero, J. *ChemPhysChem* **2005**, *6*, 1411.
- (110) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian Development Version*, revision B.01; Gaussian, Inc.: Pittsburgh, PA, 2010.
- (111) Rothenberg, S.; Schaefer, H. F., III. *J. Chem. Phys.* **1971**, *54*, 2764.
- (112) Tao, F.-M.; Pan, Y.-K. *J. Chem. Phys.* **1992**, *97* (7), 4989–4995.
- (113) Tao, F.-M. *J. Chem. Phys.* **1993**, *98* (3), 2481–2483.
- (114) Williams, H. L.; Mas, E. M.; Szalewicz, K.; Jeziorski, B. *J. Chem. Phys.* **1995**, *103* (17), 7374–7391.
- (115) Füsti-Molnar, L.; Pulay, P. *J. Chem. Phys.* **2002**, *116*, 7795.
- (116) Kreuer, K. D.; Fuchs, A.; Ise, M.; Spaeth, M.; Maier, J. *Electrochim. Acta* **1998**, *43*, 1281–1288.



- (117) Iyengar, S. S. *J. Chem. Phys.* **2007**, *126*, 216101.
- (118) Iyengar, S. S.; Petersen, M. K.; Day, T. J. F.; Burnham, C. J.; Teige, V. E.; Voth, G. A. *J. Chem. Phys.* **2005**, *123*, 084309.
- (119) Tuckerman, M. E.; Marx, D.; Parrinello, M. *Nature* **2002**, *417*, 925–929.
- (120) Marx, D.; Tuckerman, M. E.; Hutter, J.; Parrinello, M. *Nature* **1999**, *397*, 601–604.
- (121) Asthagiri, D.; Pratt, L. R.; Kress, J. D.; Gomez, M. A. *Proc. Natl. Acad. Sci.* **2004**, *101*, 7229–7233.
- (122) Shin, J.-W.; Hammer, N. I.; Diken, E. G.; Johnson, M. A.; Walters, R. S.; Jaeger, T. D.; Duncan, M. A.; Christie, R. A.; Jordan, K. D. *Science* **2004**, *304*, 1137–1140.
- (123) Diken, E. G.; Headrick, J. M.; Roscioli, J. R.; Bopp, J. C.; Johnson, M. A.; McCoy, A. B.; Huang, X.; Carter, S.; Bowman, J. M. *J. Phys. Chem. A* **2005**, *109*, 571–575.
- (124) Hammer, N. I.; Diken, E. G.; Roscioli, J. R.; Johnson, M. A.; Myshakin, E. M.; Jordan, K. D.; McCoy, A. B.; Huang, X.; Bowman, J. M.; Carter, S. *J. Chem. Phys.* **2005**, *122* (24), 244301.
- (125) Dunning Jr, T. H. *J. Chem. Phys.* **1989**, *90* (2), 1007–1023.
- (126) Mourik, T. V.; Wilson, A. K.; Dunning, T. H., Jr. *Mol. Phys.* **1999**, *96* (4), 529–547.
- (127) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, *98* (2), 1358–1371.

CT1005856