JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

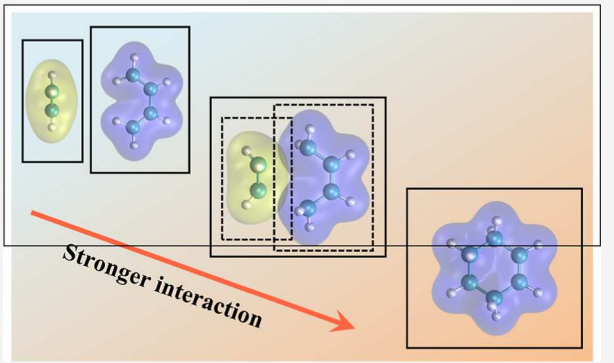
Adaptive Subsystem Density Functional Theory

Xuecheng Shao,* Andres Cifuentes Lopez, Md Rajib Khan Musa, Mohammad Reza Nouri, and Michele Pavanello*

Cite This: J. Chem. Theory Comput. 2022, 18, 6646−6655 Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Subsystem density functional theory (DFT) is emerging as a powerful electronic structure method for large-scale simulations of molecular condensed phases and interfaces. Key to its computational efficiency is the use of approximate nonadditive noninteracting kinetic energy functionals. Unfortunately, currently available nonadditive functionals lead to inaccurate results when the subsystems interact strongly such as when they engage in chemical reactions. This work disrupts the status quo by devising a workflow that extends subsystem DFT's applicability also to strongly interacting subsystems. This is achieved by implementing a fully automated adaptive definition of subsystems which is realized during geometry optimizations or ab initio molecular dynamics simulations. The new method prescribes subsystem merging and splitting events redistributing the resources (both for work and data) in an efficient way making use of modern parallelization strategies and object-oriented programming. We showcase the method with examples probing from moderate-to-strong inter-subsystem interactions, opening the door to using subsystem DFT for modeling chemical reactions in molecular condensed phases with a black box computational tool.



1. INTRODUCTION AND THEORETICAL BACKGROUND

Density functional theory (DFT), and specifically Kohn−Sham DFT (KS-DFT), is the method of choice for modeling the electronic structure of molecules and materials. The cubic scaling of algorithms resulting from practical implementations of KS-DFT becomes a major nuisance when systems of realistic (large) sizes are targeted for the simulations. The problem arises because the electronic structure in KS-DFT is found by solving the so-called KS equations[1]

$$\left[-\frac{1}{2}\nabla^2 + v_s(\mathbf{r})\right]\phi_i(\mathbf{r}) = \varepsilon_i\phi_i(\mathbf{r}) \tag{1}$$

where we introduced the KS potential, $v_s(\mathbf{r})$, and ε_i are the KS orbital energies. The KS orbitals, $\phi_i(\mathbf{r})$, are used to recover the electron density of the system

$$\rho(\mathbf{r}) = \sum_i n_i |\phi_i(\mathbf{r})|^2 \tag{2}$$

where the $\{n_i\}$ are the occupation numbers that sum up to the total number of electrons, namely, $\sum_i n_i = N$.

Equation 1 is formulated because one wishes to minimize the energy functional

$$E[\rho] = T_s[\rho] + \int v_{ext}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + E_H[\rho] + E_{xc}[\rho] \tag{3}$$

with respect to variations in the KS orbitals. $T_s[\rho] \equiv \sum_i n_i \left\langle \phi_i \middle| -\frac{1}{2}\nabla^2 \middle| \phi_i \right\rangle$ is the noninteracting kinetic energy, $E_H[\rho] = \frac{1}{2}\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}'$ is the Hartree functional, and $E_{xc}[\rho]$ is the exchange−correlation (xc) functional.

We remark that among the ingredients of the KS potential in eq 1, there is the xc potential, which is defined as the functional derivative of the xc functional, $v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[\rho]}{\delta\rho(\mathbf{r})}$. Clearly, as the xc functional needs to be approximated in practical calculations, different xc approximants will perform differently depending on the type of systems and processes the KS-DFT method is tasked to approach.

To analyze the computational cost of the algorithm, we remark that solving eq 1 is equivalent to solving an eigenvalue equation with matrices of the leading size equal to the number of basis functions used to expand the KS orbitals. These can be very large (such is the case for plane wave expansions), and thus, the

Received: July 2, 2022
Published: September 30, 2022

https://doi.org/10.1021/acs.jctc.2c00698
J. Chem. Theory Comput. 2022, 18, 6646−6655

usual $O(N^3)$ scaling of typical eigenvalue solvers can be crippling to the computational feasibility of the simulations.

The main reason for the need to diagonalize the KS Hamiltonian is because T_s can only be evaluated with a number of operations that grow linearly with the number of electrons if the KS orbitals are orthogonal to each other. Even though it is possible to reformulate mean-field methods in a basis of non-orthogonal molecular orbitals,[2−5] the cost of the resulting algorithms still scales cubically with respect to the number of electrons in the system. Therefore, an attractive strategy appears to be to find ways to improve the computational scaling for evaluating the noninteracting kinetic energy functional. Such is the strategy adopted by orbital-free DFT[6,7] as well as subsystem DFT (sDFT, also known as density embedding).[8−10]

While there are several research avenues aimed at reducing the computational cost of KS-DFT,[11−19] in this work we discuss methods based on sDFT. Similar to DFT, sDFT focuses on the electron density as the main variational function, and in the spirit of dividing and conquering the electronic structure, the density is split into subsystem contributions. Namely

$$\rho(\mathbf{r}) = \sum_{I}^{N_S} \rho_I(\mathbf{r}) \tag{4}$$

where the subsystem electron densities, $\{\rho_I(\mathbf{r})\}$, are assumed to be smooth and N-representable.[20] N_S is the total number of subsystems.

In sDFT, the energy functional is split into an additive and a nonadditive term. Namely

$$E[\{\rho_I, v_{ext}^I\}] = \sum_{I=1}^{N_S} E[\rho_I, v_{ext}^I] + E^{nad}[\{\rho_I, v_{ext}^I\}] \tag{5}$$

The subsystem energy functionals are defined in a way analogical to the KS functional

$$E[\rho_I, v_{ext}^I] = T_s[\rho_I] + E_H[\rho_I] + E_{xc}[\rho_I] + \int v_{ext}^I(\mathbf{r})\rho_I(\mathbf{r})\mathrm{d}\mathbf{r} \tag{6}$$

The nonadditive energy is defined as

$$E^{nad}[\{\rho_I, v_{ext}^I\}] = T_s^{nad}[\{\rho_I\}] + E_{xc}^{nad}[\{\rho_I\}] + E_H^{nad}[\{\rho_I\}] + \sum_{I,J\neq I} \int v_{ext}^J(\mathbf{r})\rho_I(\mathbf{r})\mathrm{d}\mathbf{r} \tag{7}$$

The nonadditive xc and nonadditive noninteracting kinetic energy functionals (hereafter NAXC and NAKE, respectively) are defined trivially as

$$F^{nad}[\{\rho_I\}] = F[\rho] - \sum_{I=1}^{N_S} F[\rho_I] \tag{8}$$

where F is any density functional.

Variational minimization with respect to the subsystem electron densities is achieved by solving coupled KS-like equations for each subsystem. The resulting KS-like equations take the form

$$\left[-\frac{1}{2}\nabla^2 + v_s^I(\mathbf{r}) + v_{emb}^I(\mathbf{r})\right]\phi_i^I(\mathbf{r}) = \varepsilon_i^I\phi_i^I(\mathbf{r}) \tag{9}$$

where the embedding potential, $v_{emb}{}^I$

$$v_{emb}^I(\mathbf{r}) = \frac{\delta T_s[\rho]}{\delta\rho(\mathbf{r})} - \frac{\delta T_s[\rho_I]}{\delta\rho_I(\mathbf{r})} + \frac{\delta E_{xc}[\rho]}{\delta\rho(\mathbf{r})} - \frac{\delta E_{xc}[\rho_I]}{\delta\rho_I(\mathbf{r})} + \sum_{J\neq I}[v_H[\rho_J](\mathbf{r}) + v_{ext}^J(\mathbf{r})] \tag{10}$$

The embedding potential can be simply derived by taking the functional derivative with respect to ρ_I of the nonadditive energy. We remark that the subsystem KS potential only contains terms depending on ρ_I and the external potential of the Ith subsystem (i.e., $v_{ext}^I(\mathbf{r})$).

sDFT is exact in principle, however, its practical implementations require the approximation of both NAXC and NAKE. For NAXC, one can benefit from the advances applicable to the functionals employed in mainstream KS-DFT implementations (such as local and semilocal functionals as well as dispersion corrections). For NAKE, the landscape is much more complicated because the literature is much less developed compared to xc approximants. In principle, NAKEs can be derived from a parent kinetic energy functional (see eq 8). However, as opposed to xc, common knowledge is that currently available kinetic energy functionals are not accurate enough to produce quantitative results when employed in an orbital-free DFT simulation for molecular systems.[21] The situation is different for bulk solids (such as metals and some semiconductors), surface slabs, and nanoparticles where orbital-free DFT has been shown to produce results close to chemical accuracy.[22−30]

However, even though the parent kinetic energy functional may not be accurate in an orbital-free DFT simulation of the system of interest, it may be accurate when employed as NAKE. In fact, there is an ample literature that generally reaches the conclusion that if the inter-subsystem interactions are weak, approximate NAKE functionals can reach accurate results,[31−37] especially when the parent functional has a nonlocal dependence on the electron densities of the subsystems.[32] Existing reviews on the subject are listed here for reference.[20,38] – We remark that when sDFT is accurate (i.e., for weak inter-subsystem interactions such as for molecular liquids and interfaces) and therefore useful for predictive simulations, linear computational scaling with system size is achieved, thanks to massive parallel implementations of sDFT and NAKE functionals.[22,24,42−44] For example, sDFT has gained grounds in recent years for replacing mainstream KS-DFT in simulations involving ab initio dynamics, such as for determining the structure of liquids[45,46] and recently also for determining free energies of solvation.[47]

The question we wish to attack in this work is: how can one retain the favorable computational scaling of sDFT while also retaining the accuracy of KS-DFT when strong inter-subsystem interactions occur? Given what we know about kinetic energy functionals, one might conclude that it is generally not possible. However, when the strong interactions occur between only a few subsystems, it may be advantageous to merge the strongly interacting subsystems into a single subsystem which then is treated as a single KS-DFT subsystem. We outline this "adaptive" sDFT in the next section paying particular attention to describing the details involved when it is translated into an efficient and scalable software. We remark that merging strongly interacting subsystems into a single subsystem is equivalent in spirit to the so-called "projection-based embedding"[48,49] for a DFT subsystem in a DFT environment.[50,51] In projection-based embedding, the partitioning takes place at the level of the occupied molecular orbitals which are kept orthogonal and

https://doi.org/10.1021/acs.jctc.2c00698
J. Chem. Theory Comput. 2022, 18, 6646−6655

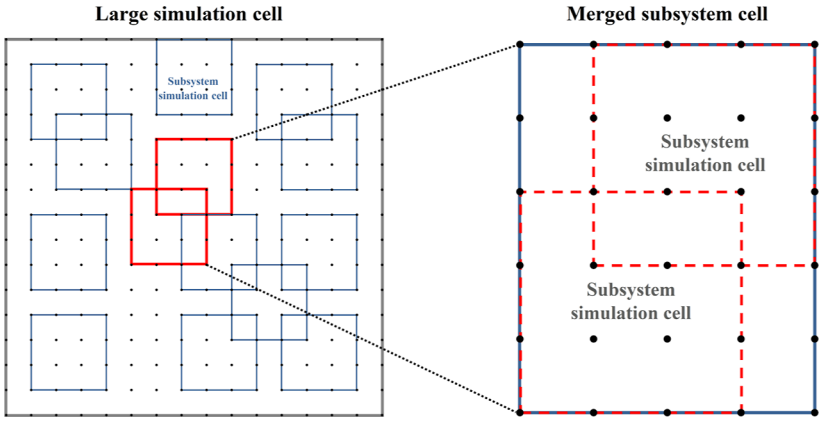


Figure 1. Depiction of the merged subsystem simulation cell. The cell of the merged subsystem has the same symmetry as the large cell with scaled lattice vectors to contain the union of the original subsystem cells.

overall reproduce the density and density matrix of the KS-DFT calculation of the combined system.

In the following section (Section 2), we enumerate the many details involved for an efficient implementation of adaptive sDFT, from parallelization to subsystem topology, to basis sets. Section 3 is devoted to the computational details. In Section 4, we present a justification of why it is advantageous to merge subsystems when their interaction becomes too strong. We also present two applications of adaptive sDFT: one regarding moderate inter-subsystem interactions such as the ones taking place between the ions and water in dissolved NaCl; and another one regarding strong inter-subsystem interactions, such as those occurring during a chemical reaction (we showcase an S_N2 reaction). Lastly, we draw the conclusions.

2. ADAPTIVE DEFINITION OF SUBSYSTEMS

When subsystems in a sDFT simulation are in proximity of each other, and if they interact too strongly, as we have explained, it is advantageous to merge them into a single subsystem. However, once the subsystems have merged, the identity of the separate subsystems is lost. A similar situation occurs when a single subsystem is found to be composed of two fragments that can be safely treated as separate subsystems. When this happens, two new subsystems come in existence.

One may think that the resulting algorithm to encode subsystem split and merge events is trivial, and that little or no complications should arise. The reality is quite the opposite as the ultimate goal is to produce a black box workflow that automatically, and with minimal user input, handles subsystem merge and split events. We explain the details of the workflow in the following sections.

2.1. Subsystem Splitting/Merging, Simulation Cells, and Plane Wave Basis Sets. The following list defines the prescription used to split and merge subsystems during an sDFT simulation:

1. Distance matrices collecting the distances between all the atoms in a subsystem (d_{ij}) and the minimum interatomic distances between different subsystems (D_{IJ}) are computed and stored.
2. Elements of D_{IJ} are compared against the sum of cutoff radii (roughly equal to the covalent radii increased by 40%), Δ, which are collected in a database. If $D_{IJ} < \Delta$, then the subsystem I and J are merged.
3. Elements of d_{ij} are compared against the sum of cutoff radii and augmented by a threshold, $\delta = 0.2$ Å. If $d_{ij} > \Delta + \delta$, then atoms i and j which belong to the same subsystem will constitute the boundary between two new subsystems. The subsystem is split into two new ones.

While points (1) and (2) simply implement a standard algorithm based on interatomic distances, the role of the additional 0.2 Å (δ) in the comparison between the interatomic distances and $\Delta + \delta$ in point (3) may seem a bit ad-hoc. Historically, the first version of the algorithm did not feature δ and the resulting ab initio dynamics and relaxations incurred into complications given by fluctuations of two subsystems between a "merged" and a "split" state. Such fluctuations would persist for several tens of MD or relaxation steps, resulting in slow timings of execution. The reason for the slow timings for relaxations is the increased number of iterations. For MD, increased timings are due to the fact that for each split event, the initial guess of the electron density needs to be generated from scratch, which leads to a slower convergence of the SCF compared to a case where the initial guess is borrowed from the previous MD step. The addition of the δ threshold cured the problem and led our simulations to be stable and timings to be optimal. Therefore, the role of δ in adaptive sDFT is similar to the role of "switching regions" in QM/MM adaptive dynamics frameworks.[52−60]

An additional complication is given by the choice of basis set to expand the molecular orbitals of the subsystems. In sDFT, the concept of using monomer basis set (i.e., a basis set localized on the single subsystem such that it does not extend to the entire system) has a long history.[61,62] For codes based on local orbitals (such as Serenity,[63] ADF[64] and CP2K[65]), employing a monomer basis set simply involves expanding the subsystem KS orbitals with atomic orbitals centered on atoms belonging to only that subsystem. In plane wave codes, employing a monomer basis is less straightforward. The plane wave basis is related to the simulation cell. Therefore, it needs to be customized for each subsystem. That is, each subsystem will use a custom simulation cell (as we have developed previously in our group[42,66,67]) large enough that the electron density of the subsystem decays to very small values before reaching the boundary of the cell.

The computational savings originate from the fact that the subsystem simulation cells are smaller in size compared to the large simulation cell chosen for the full system (see Figure 1) and hence will contain a much smaller number of plane wave basis functions and real space points. Interactions between subsystems are still fully accounted for by computing the long-

https://doi.org/10.1021/acs.jctc.2c00698
J. Chem. Theory Comput. 2022, 18, 6646−6655

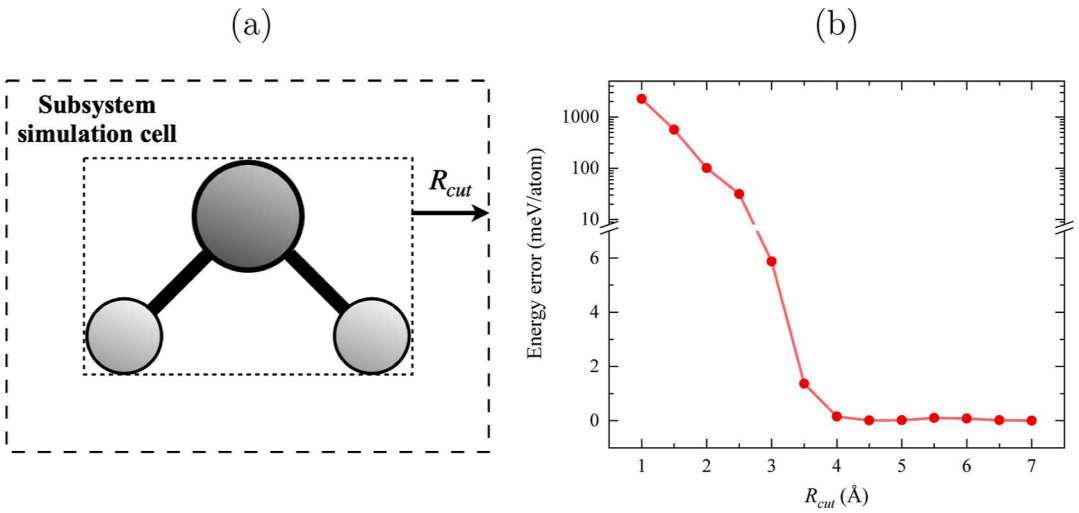


Figure 2. Formulation of a subsystem simulation cell as a function of the atomic coordinates of the subsystem. Panel (a) provides a visual for a water molecule. Panel (b) shows the effect of R_{cut} on the total energy of a water dimer. R_{cut} is the vacuum added to each lattice vector direction.

Table 1. Summary of the computational settings of merged subsystems formulated by the adaptive sDFT driver[a]

subsystem 1	subsystem 2	merge operation/choice	comments
lda_plus_u=.true.	lda_plus_u=.false.	lda_plus_u=.true.	GGA + U
smearing=.true.	smearing=.false.	smearing=.true.	smeared occupations
degauss$_1$	degauss$_2$	max[degauss$_1$, degauss$_2$]	smearing parameter
ecutrho$_1$	ecutrho$_2$	max[ecutrho$_1$, ecutrho$_2$]	density plane wave cutoff
mixing_beta$_1$	mixing_beta$_2$	min[mixing_beta$_1$, mixing_beta2]	density mixing parameter
nelec$_1$	nelec$_2$	nelec$_1$ + nelec$_2$	number of electrons
nband$_1$	nband$_2$	nband$_1$ + nband$_2$	number of KS orbitals

[a]We list a selected number of input file keywords for Quantum ESPRESSO and the corresponding operands encoding the merge.

ranged interactions on the physical, large cell, and only using the small cell computations for Hamiltonian diagonalizations and definitions of subsystem KS orbitals and electron densities. For periodic subsystems, the cell should retain periodicity. For example, for slab subsystems, the subsystem cell can be smaller than the large cell along the vertical direction but not in the *xy* plane. The typical computational savings achieved by this method are dramatic. For example, for liquid water[46] and fluid carbon dioxide,[45] we diagonalized the subsystem Hamiltonians using only about 5% of the total number of plane waves available in the large cell.

In an adaptive sDFT dynamics, where the nature and makeup of the subsystems may change during a simulation, prescriptions to combine the simulation cells of merging subsystems, and for dividing the cells of splitting subsystems, need to be formulated. We summarize them as follows:

1. Simulation cells of merging subsystems are merged by finding the smallest cell of same symmetry as the large (physical) simulation cell that contains the union of the two, original simulation cells. See a depiction in Figure 1.
2. Simulation cells of splitting subsystems are determined in two ways: (1) if the subsystems have existed before, the cells and simulation conditions are borrowed from the previous definitions; (2) if a new subsystem is formed that had not existed before, the new simulation cell is determined automatically by the adaptive sDFT driver. It first determines the smallest possible cell that contains the subsystem, then it pads the cell by R_{cut} of vacuum along each lattice vector to make sure the subsystem density decays inside the cell well before reaching the cell boundaries. Figure 2 depicts this approach showing that the error introduced in the total energy of a subsystem can be made very small (in the order of meVs) by adopting $R_{cut} \geq 3$ Å. Even though the example in the figure is specific to a single water molecule, we have tested the approach with other types of subsystems and determined that the procedure is robust.

Other quantities and input file-related settings are determined according to Table 1 and discussed in the next section.

2.2. Computational Settings of Merged/Split Subsystems. When merging two subsystems into a combined, single subsystem, the computational settings governing the electronic structure computation must be formulated automatically. It is crucial for this step to be automatic so to not require user input along a Born−Oppenheimer dynamics or geometry optimization. It would, in fact, be impractical to have to stop and restart the simulations manually. Additionally, once the simulations are stopped, they typically drop off the job scheduler of the high-performance computing system. This would further add to the frustration of the user. Therefore, we designed a completely automatic driver that combines and splits subsystems at run time. Table 1 summarizes some key rules for obtaining Quantum ESPRESSO input file parameters for a subsystem resulting from the merger of two subsystems. We recall the previous section's description of how to obtain the combined subsystem's simulation cell.

Table 1 collects a minimal set of computational settings that are determined automatically upon subsystem merge. A sensible choice for these settings has vital consequences to the simulation. For example, smearing (and the associated smearing

https://doi.org/10.1021/acs.jctc.2c00698
J. Chem. Theory Comput. 2022, 18, 6646−6655

parameter mixing_beta) determines orbital occupations which is especially important for systems with small gaps, such as metals and some spin-polarized systems. When two subsystems merge, the safe choice is to use smearing (as opposed to Aufbau occupations) if any of the merging subsystems already required smearing. Other settings are still vital but more trivially determined, such as the number of bands (i.e., computed orbitals), number of electrons, and total charge.

Less trivial would be predicting the total magnetization of the merged subsystem. It would not be generally valid to impose the total magnetization to be the sum of the merging subsystems' magnetization. Therefore, for this case, the default is to set up the merged subsystem with a vanishing total magnetization (spin unpolarized). The driver also allows the user to define custom subsystems by providing molecular formula, charge, and magnetization of specific molecules or periodic subsystems that may arise during a relaxation or a molecular dynamics simulation. While this may push away the adaptive sDFT driver from being a true "black box," it is the simplest way we found to solve the magnetization uncertainty of merging subsystems and other computational settings for *splitting* subsystems as we further explain below.

When a subsystem splits into two "children" subsystems, as mentioned in the previous section, if the subsystems were defined at the beginning of the simulation, the cells and simulation conditions are borrowed from the previous definitions. However, if the splitting event generates new subsystems, the adaptive sDFT driver assigns to them the same computational settings as the parent subsystem. Additionally, assuming the parent subsystem has neutral charge, the children subsystems are also considered to be neutral. Clearly, this is not ideal when bonds break heterolytically. If this is the case, and if the user is expecting such a possibility, then it is enough to define the molecular formula of the predicted "child" subsystem, assign to it a charge, and include it in the list of possible subsystems (a database file used by the adaptive sDFT driver). This will guide the adaptive sDFT driver in breaking bonds according to a specified, expected outcome.

2.3. Computational Resources of Merged/Split Subsystems. In addition to the input file parameters (keywords), to run a simulation one must assign computational resources to the merged subsystem. It is natural to assign a merged subsystem the sum of the resources previously employed for the original subsystems. This is straightforward in theory but less so in practice considering that the combination must be achieved at run time.

Resources are typically handled by parallel algorithms based on message passing interface (MPI). MPI communicators are used to handle the memory associated to each subsystem. A shown in Figure 3, we outline the method we implemented in eDFTpy[69] for handling the parallel resources for memory and work. When two subsystems are merged, the resources must be reconfigured. Specifically: (1) the set of *comm_sub* and *comm_region* is reset (see Figure 3 for the detailed explanations of the involved MPI communicators), assigning to the new *comm_sub* the sum of MPI tasks of the original communicators. For *comm_region*, because it depends on the spatial extension of the subsystem cell, it is redefined specifically for the merged subsystem. Along the new set of communicators, the corresponding arrays are also reallocated for the merged subsystem only.

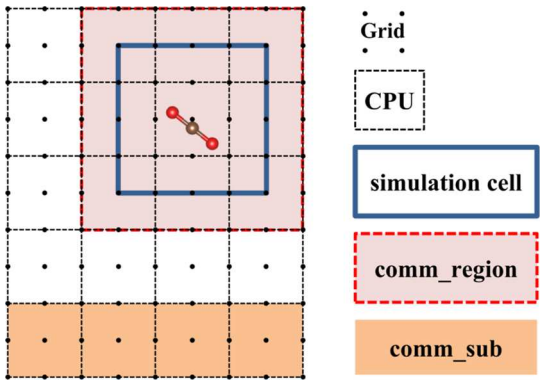


Figure 3. Parallelization of work and data in eDFTpy. MPI communicators are used for several purposes. Specifically, *comm_sub* are communicators devoted to the subsystem KS-DFT solver (in this work, Quantum ESPRESSO). Data from a subsystem (e.g., the electron density and potentials) need to be represented on the global grid for the computation of the embedding potentials. Thus, "region" communicators (*comm_region*) handle the memory when gathering and scattering a subsystem density or potential on the global grid. Data are broadcast from *comm_sub* to *comm_region* by gather-scatter on a single CPU. Given the typically reduced sizes of the subsystems, such an operation requires a small memory allocation for each subsystem. A major benefit of using *comm_region* is that it avoids gathering global data (from the large simulation grid) on a single CPU which would be unfeasible for large systems. See ref 68 for additional details.

3. COMPUTATIONAL DETAILS

All calculations have been carried out with the in-house eDFTpy code[69] which employs Quantum ESPRESSO (QE)[70] as the KS-DFT solver. In order to better incorporate QE into the workflow, we developed a Python interface to QE, called QEpy,[71] which is capable of interacting with QE from the Python side, providing QE with custom external potentials fed as NumPy arrays (gathered and distributed). Additionally, QEpy is compatible with the parallelization strategy of eDFTpy handled by mpi4py.[72] Plane wave cutoffs are set to 40 Ry for wavefunctions and 400 Ry (Ry stands for Rydberg atomic units) for the electron density, tested successfully in several past publications involving the same pseudopotentials (GBRV[73] in this work). In all simulations, we use the revAPBEK[35] NAKE and the PBE[74] xc functional. In all Born−Oppenheimer dynamics, we use a time step of 1 fs. We collect all the structures and trajectory files for all the simulations presented in the Results section and list the covalent radii used in the adaptive sDFT simulations in the supplementary materials.[75]

4. RESULTS

The results section is organized as follows. First, we present two test systems. First a water dimer and then a Diels−Alder reaction. In both cases, subsystems sample interactions from weak to strong. The test systems show that adaptive sDFT produces energies and atomic forces, in general agreement with KS-DFT of the supersystem. Then, we use NaCl in water as an example of moderate inter-subsystem interactions. We show a Born−Oppenheimer dynamics trajectory for the dissociation of the NaCl molecule to Na^+ and Cl^- solvated ions. We monitor the adaptive definition of subsystems along the molecular dynamics trajectory. Finally, we show an example of a chemical reaction in the explicit solvent. From a structure close to the transition state of an S_N2 reaction of chloroethane with bromide, we relax the geometry of the system and monitor the Br^- anion

https://doi.org/10.1021/acs.jctc.2c00698
J. Chem. Theory Comput. 2022, 18, 6646−6655

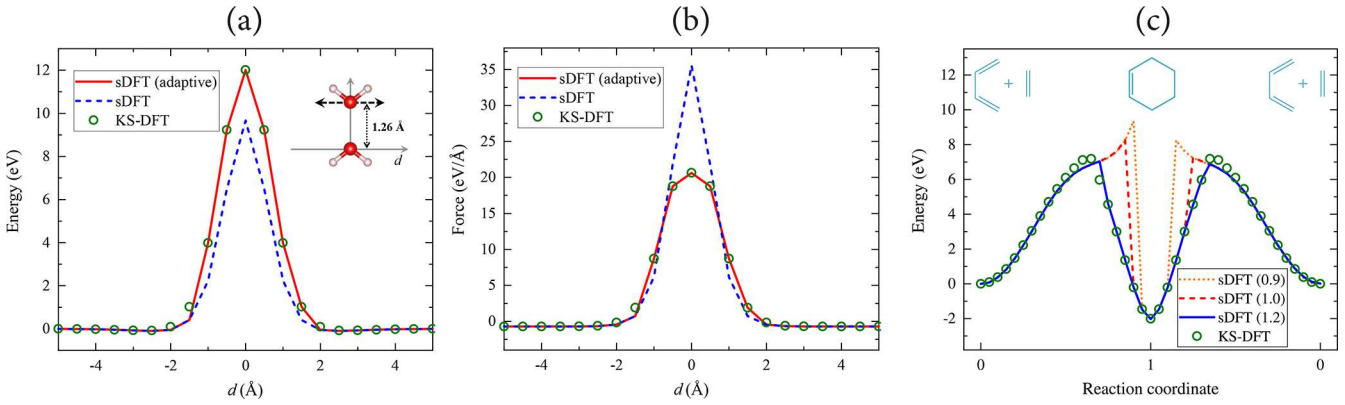


Figure 4. Interaction curves (a) and forces along the z-direction on the oxygen atom of the top water molecule (b) for a water dimer computed by KS-DFT, sDFT, and adaptive sDFT. (c) Energy curves for a Diels–Alder reaction between butadiene and ethene. See the text for additional details. (c) Value of the cutoff radius of carbon used in the simulation. R_{cut} = 4.0 Å is adopted in all simulations.

leaving the CH_3CH_2Cl molecule. We also monitor how the adaptive sDFT driver handles merging and splitting subsystems along the geometry relaxation.

4.1. Justification for Adaptive sDFT. To exemplify the need to define subsystems adaptively, we first present the water dimer as a simple yet instructive example. In panels (a) and (b) of Figure 4, we show energy curves and forces for the interaction between two water molecules as one slides past the other. When the water molecules are too close, the sDFT interaction deviates strongly from the one computed by KS-DFT of the combined system. KS-DFT is to be taken as the benchmark method because it employs the exact noninteracting kinetic energy functional.

From panel (a) of Figure 4, we notice that the errors involved in sDFT are substantial, amounting to 2 eV for the geometry where the two water molecules are closest (the O atoms are separated only by 1.26 Å). While the geometries probed by the model system in Figure 4 are unlikely to occur in room-temperature dynamics of water, including liquid water,[46] such errors in the interaction energies are completely unacceptable when chemical reactions are modeled. Merging the two water subsystems at around $d = \pm 2$ Å avoids the incorrect behavior of regular sDFT. Panel (b) of the figure shows that while sDFT forces deviate substantially from the KS-DFT reference, adaptive sDFT cures the problem providing forces that are in good agreement with the KS-DFT benchmark.

In panel (c) of Figure 4, we show the energy profiles for the simplest of Diels–Alder reactions (butadiene with ethene), which we run from reactants to products and then back to reactants along a reaction coordinate (given by a path variable, s). The reaction coordinate is defined by considering the geometry of the product molecule (cyclohexene), labeled as $\vec{g}_P$, and the geometry of the reactants (one 1,3-butadiene molecule and one ethene molecule), labelled as $\vec{g}_R$. The geometries along the reaction coordinate are labeled as $\vec{g}(s)$ and are given by this equation

$$\vec{g}(s) = s\vec{g}_P + (1 - s)\vec{g}_R, \quad 0 \le s \le 1 \tag{11}$$

Therefore, we simply vary s from $s = 0$ to $s = 1$ to drive the reaction from reactants to products and then we follow the reverse path from $s = 1$ to $s = 0$ to go back to the reactants. We carry out such simulation employing three different choices of the cutoff radius of carbon. From the figure, we see that when the cutoff radius is chosen to be 0.9 Å (the smallest considered), there are large, unacceptable energy deviations nearby the transition state. When the radius is set to 1.2 Å, the deviations between the KS-DFT and the adaptive sDFT energies are reduced to a very small and acceptable level.

It is instructive to point out that the effects of the "switching region" used in adaptive sDFT (i.e., the $\delta = 0.2$ Å threshold, see Section 2) are slightly visible in Figure 4. Specifically, the lhs of the curves (both energies and forces) slightly differ from their rhs. Of course, it is possible to further reduce the (already small) discrepancy between adaptive sDFT and the KS-DFT benchmark, by further adjusting the values of the thresholds δ and Δ.

We remark that the computational cost of the simulations after a subsystem merge event does not deteriorate dramatically unless the subsystems involved are of similar size as the full system. This is because, generally, subsystems are small in size compared to the full system, and the combination of two subsystems leads to still a small subsystem.

4.2. Moderate Interactions: NaCl in Water. In this example, see Figure 5, we carry out a computational experiment whereby a NaCl molecule is solvated by water constraining the Na–Cl distance to $d_{Na-Cl} = 2.36$ Å. After geometry relaxation of the water molecules only, the system is propagated along a Born–Oppenheimer dynamics trajectory for 12 ps. The atom's velocities are regulated by the Andersen thermostat[76] with temperature set to $T = 373$ K. The elevated temperature is chosen to allow for a faster dynamics.

Along the trajectory, we expect the NaCl molecule to dissociate to solvated ions, $Na_{(aq)}^+$ and $Cl_{(aq)}^-$. The adaptive sDFT dynamics driver initially determines that Na and Cl should belong to the same subsystem (their initial distance d_{Na-Cl} is shorter than the $\Delta + \delta$ threshold, see Section 2 for details). For similar reasons, some water molecules nearby the NaCl molecule are also included in the NaCl subsystem.

In Figure 6, we depict the variability of the number and nature of the subsystems involved. From the split/merge inset of Figure 6 is interesting to note that the sodium and chloride go back and forth belonging to the same or separate subsystems until around step 7000 (or 7 ps in time) when they finally dissociate and diffuse away from each other.

We use the NaCl system also for a secondary purpose, testing adaptive sDFT's ability to conserve energy in a 500-step *NVE* Born–Oppenheimer dynamics. Figure 7 shows the potential and total energy for the NaCl system along a *NVE* dynamics. Because the energy functional is redefined after each split and merge events, vertical jumps in the total energy are expected. It is

https://doi.org/10.1021/acs.jctc.2c00698
J. Chem. Theory Comput. 2022, 18, 6646–6655

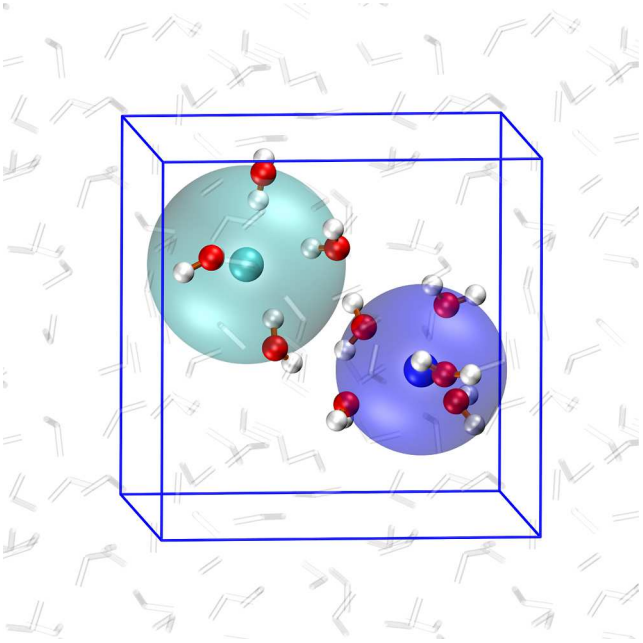


Figure 5. One snapshot of a Born−Oppenheimer dynamics of NaCl in water. The chloride and sodium subsystems are represented by the transparent green and blue spheres, respectively. The spheres indicate that several water molecules are automatically included in the subsystems (4 for Cl^- and 5 for Na^+ in this snapshot).

important to show that the energy is conserved (negligible slope) in between the split/merge events. From the figure, we can also see that the overall energy drift is still very small in the order of a few meV/atom/ps, which is acceptable.

The discussed *NVE* dynamics also shows that it is possible to employ the proposed adaptive sDFT dynamics algorithm in nuanced simulations that may rely on an accurate energy functional, such as free-energy calculations. However, caution is required as the proposed algorithm can and will be improved, specifically in the definition and handling of the switching region in future iterations.

4.3. Strong Interactions: A S_N2 Reaction in Water.

sDFT was conceived to model chemistry in explicit environments.[40,62,77−79] However, modeling chemical reactions by sDFT is difficult in practice because, as mentioned before, when two or more subsystems come too close to each other and their electron density overlap strongly (a situation needed when chemical reactions occur), the NAKE functional approximants typically employed by sDFT are not accurate enough to provide a qualitative description of the interactions. Thus, the adaptive version of sDFT presented here is expected to open the door to sDFT models of chemical reactions.

As proof of principle, we choose a simple S_N2 reaction of chloro/bromoethane interacting with chloride/bromide nucleophiles in aqueous solutions. Even though water is a bad solvent for S_N2 reactions, it has been studied nonetheless due to the large sizes of the resulting activation energies, as well as the strong solvent effect.[80] We consider Figure 8 where we present the electronic energy as a function of step number along a geometry optimization carried out with adaptive sDFT. The chosen system is the transition-state structure of the nucleophilic S_N2 reaction of chloroethane with bromide, in water with one sodium counterion. Upon optimization, the structure relaxes away from the transition state toward one of the two possible products. In the example, the products reached by the relaxation are aqueous chloroethane and aqueous bromide.

The figure shows that during the relaxation, there are four subsystem merge/split events (indicated by stars in the plot). Particularly, the chloroethane/bromide split event is the third event occurring around step 10. In that event, the first solvation shell of water is also split between the two subsystems, indicated in the figure by oxygens changing label color. It is interesting to note that the transition-state energy of about 1.3 eV matches well early predictions.[81]

5. CONCLUSIONS

In conclusion, we have developed a fully automated, adaptive definition of subsystems for subsystem DFT simulations. A new method cures a longstanding issue in practical subsystem DFT simulations, that is, the inaccurate predictions when two subsystems interact strongly. The adaptive definition is such

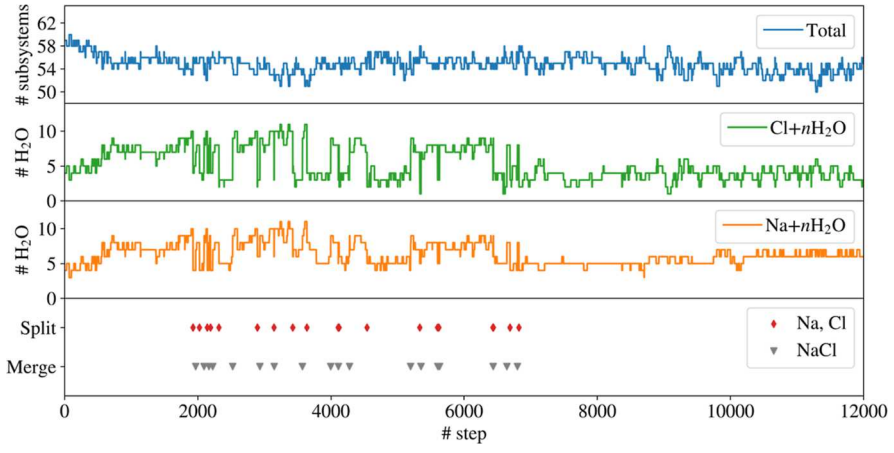


Figure 6. Variable number and nature of subsystems along a 12 ps sDFT Born−Oppenheimer trajectory of solvated, dissociating NaCl in water. Total: the (variable) total number of subsystems along the trajectory. Cl(Na)+$n$$H_2O$: number of water molecules in the subsystem that includes the Cl (Na) atom along the trajectory. Initially, Na and Cl belong to the same subsystem; therefore, the two graphs are identical until Na and Cl split into two different subsystems. Split/Merge: subsystem split and merge events for the Na and Cl atoms. After 7000 steps (7 ps), Na and Cl diffuse away as $Na_{(aq)}^+$ and $Cl_{(aq)}^-$.

https://doi.org/10.1021/acs.jctc.2c00698
J. Chem. Theory Comput. 2022, 18, 6646−6655

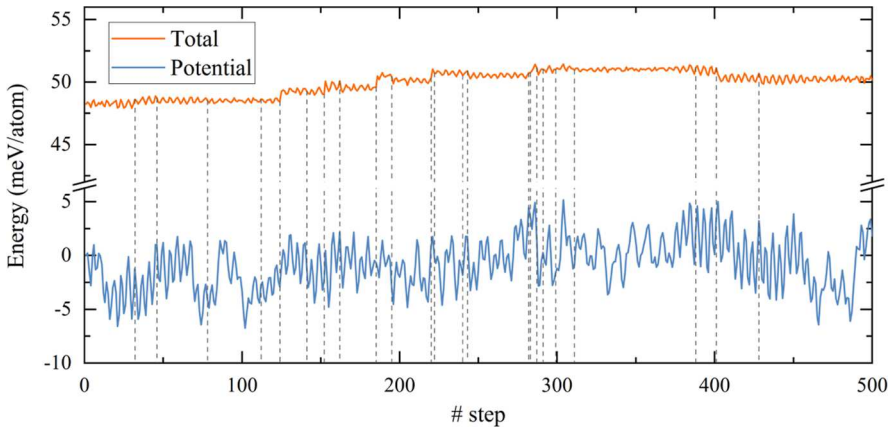


Figure 7. 500 step (0.5 ps) *NVE* Born−Oppenheimer dynamics of the water-solvated NaCl system from Figure 5. Potential: the total adaptive sDFT electronic energy. Total: the total energy (kinetic plus potential). Vertical dashed lines indicate the steps where split/merge events take place.

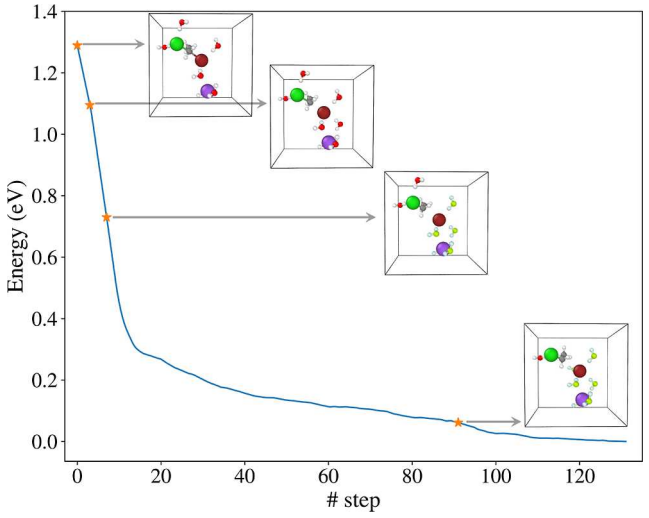


Figure 8. Adaptive sDFT geometry relaxation of the chloroethane/bromide S_N2 reaction transition state in water. The explicit water molecules shown in the figure are the ones that are included in the chloroethane (red oxygen atoms) and bromide (yellow oxygen atoms) subsystems. The points indicated by the star markers indicate the steps when a subsystem split or merge event takes place. The color labels are as follows: hydrogen: white; bromide/bromine: brown; sodium: purple; oxygen: red or yellow; and carbon: gray.

that when subsystems are spatially close (and interact strongly), they are merged into a single subsystem. Vice versa, when two fragments belonging to the same subsystem move away from each other, the method automatically splits them into two new subsystems. The method operates at run time (i.e., no executables are started and stopped during a simulation) and takes full advantage of MPI communicators. For example, when two subsystems are merged, the sum of the MPI tasks used for each of the original subsystems is assigned to the combined subsystem. "Merge" and "split" operations are encoded in a quasi-black box method that we show to be efficient and robust.

We considered two examples to showcase the usefulness of the adaptive subsystem DFT method: a 12 ps dynamics of the dissociation of NaCl in water and a geometry relaxation from the transition state down to the products of a prototypical S_N2 reaction in water. In both cases, we show that the number of subsystems in the simulation changes adaptively yielding physical predictions. A strict benchmark on a weakly-to-strongly interacting water dimer and Diels−Alder reaction model systems reveals that adaptive subsystem DFT is essentially as accurate as KS-DFT of the combined supersystem for both energies and forces.

In sum, we foresee adaptive subsystem DFT to become the new state of the art subsystem DFT simulation method, opening the door to modeling condensed phase chemical reactions with subsystem DFT and therefore increasing the applicability of DFT to even more realistic model systems than it has been possible till date.

AUTHOR INFORMATION

Corresponding Authors

Xuecheng Shao − *Department of Chemistry, Rutgers University, Newark, New Jersey 07102, United States;* orcid.org/0000-0003-2215-0926; Email: xuecheng.shao@rutgers.edu

Michele Pavanello − *Department of Physics, Rutgers University, Newark, New Jersey 07102, United States;* orcid.org/0000-0001-8294-7481; Email: m.pavanello@rutgers.edu

Authors

Andres Cifuentes Lopez − *Department of Chemistry, Rutgers University, Newark, New Jersey 07102, United States;* orcid.org/0000-0002-7861-0111

Md Rajib Khan Musa − *Department of Chemistry, Rutgers University, Newark, New Jersey 07102, United States*

Mohammad Reza Nouri − *Department of Chemistry, Rutgers University, Newark, New Jersey 07102, United States;* orcid.org/0000-0002-6046-200X

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.2c00698

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under grant numbers CHE-1553993, CHE-2154760, and OAC-1931473. We thank the Office of Advanced Research Computing at Rutgers for providing access to the Amarel cluster.

REFERENCES

(1) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133−A1138.

(2) Kolos, W.; Radzio, E. Applicaition of the statistical method in the theory of intermolecular interactions. *Int. J. Quantum Chem.* **1978**, *13*, 627−634.
(3) Farazdel, A.; Dupuis, M.; Clementi, E.; Aviram, A. Electric-field induced intramolecular electron transfer in spiro .pi.-electron systems and their suitability as molecular electronic devices. A theoretical study. *J. Am. Chem. Soc.* **1990**, *112*, 4206−4214.
(4) McWeeny, R.*Methods of Molecular Quantum Mechanics*; Academic Press: San Diego, 1992.
(5) Cembran, A.; Payaka, A.; Lin, Y.; Xie, W.; Mo, Y.; Song, L.; Gao, J. A Non-Orthogonal Block-Localized Effective Hamiltonian Approach for Chemical and Enzymatic Reactions. *J. Chem. Theory Comput.* **2010**, *6*, 2242−2251.
(6) Fermi, E. Un Metodo Statistico per la Determinazione di alcune Prioprietà dell'Atomo. *Rend. Accad. Naz. Lincei* **1927**, *6*, 602−607.
(7) Chacón, E.; Alvarellos, J. E.; Tarazona, P. Nonlocal kinetic energy functional for nonhomogeneous electron systems. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1985**, *32*, 7868−7877.
(8) Senatore, G.; Subbaswamy, K. R. Density Dependence of the Dielectric Constant of Rare-Gas Crystals. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1986**, *34*, 5754−5757.
(9) Wesolowski, T. A.; Warshel, A. Frozen Density Functional Approach for ab Initio Calculations of Solvated Molecules. *J. Chem. Phys.* **1993**, *97*, 8050.
(10) Cortona, P. Self-Consistently Determined Properties of Solids Without Band-Structure Calculations. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1991**, *44*, 8454.
(11) Yang, W. Direct Calculation of Electron Density in Density-Functional Theory. *Phys. Rev. Lett.* **1991**, *66*, 1438−1441.
(12) Zhu, T.; Pan, W.; Yang, W. Divide-and-conquer calculations for clean surfaces and surface adsorption. *Theor. Chim. Acta* **1997**, *96*, 2−6.
(13) Akama, T.; Kobayashi, M.; Nakai, H. Implementation of divide-and-conquer method including Hartree-Fock exchange interaction. *J. Comput. Chem.* **2007**, *28*, 2003−2012.
(14) Baroni, S.; Giannozzi, P. Towards Very Large-Scale Electronic-Structure Calculations. *Europhys. Lett.* **1992**, *17*, 547−552.
(15) Li, X.-P.; Nunes, R. W.; Vanderbilt, D. Density-matrix electronic-structure method with linear system-size scaling. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1993**, *47*, 10891−10894.
(16) Goedecker, S.; Colombo, L. Efficient Linear Scaling Algorithm for Tight-Binding Molecular Dynamics. *Phys. Rev. Lett.* **1994**, *73*, 122−125.
(17) Chelikowsky, J. R.; Troullier, N.; Saad, Y. Finite-difference-pseudopotential method: Electronic structure calculations without a basis. *Phys. Rev. Lett.* **1994**, *72*, 1240−1243.
(18) Kronik, L.; Makmal, A.; Tiago, M. L.; Alemany, M. M. G.; Jain, M.; Huang, X.; Saad, Y.; Chelikowsky, J. R. PARSEC - the pseudopotential algorithm for real-space electronic structure calculations: recent advances and novel applications to nano-structures. *Phys. Status Solidi B* **2006**, *243*, 1063−1079.
(19) Chen, M.; Baer, R.; Neuhauser, D.; Rabani, E. Overlapped embedded fragment stochastic density functional theory for covalently-bonded materials. *J. Chem. Phys.* **2019**, *150*, 034106.
(20) Wesolowski, T. A.*Computational Chemistry: Reviews of Current Trends*; Leszczynski, J., Ed.; World Scientific: Singapore, 2006; Vol. *10*, pp 1−82.
(21) Xia, J.; Huang, C.; Shin, I.; Carter, E. A. Can orbital-free density functional theory simulate molecules? *J. Chem. Phys.* **2012**, *136*, 084102.
(22) Shao, X.; Jiang, K.; Mi, W.; Genova, A.; Pavanello, M. DFTpy: An efficient and object-oriented platform for orbital-free DFT simulations. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2021**, *11*, No. e1482.
(23) Shao, X.; Mi, W.; Pavanello, M. Efficient DFT Solver for Nanoscale Simulations and Beyond. *J. Phys. Chem. Lett.* **2021**, *12*, 4134−4139.
(24) Shao, X.; Mi, W.; Pavanello, M. Revised Huang-Carter nonlocal kinetic energy functional for semiconductors and their surfaces. *Phys. Rev. B* **2021**, *104*, 045118.
(25) Mi, W.; Pavanello, M. Orbital-Free DFT Correctly Models Quantum Dots When Asymptotics, Nonlocality and Nonhomogeneity Are Accounted For. *Phys. Rev. B* **2019**, *100*, 041105.
(26) Mi, W.; Genova, A.; Pavanello, M. Nonlocal kinetic energy functionals by functional integration. *J. Chem. Phys.* **2018**, *148*, 184107.
(27) Huang, C.; Carter, E. A. Nonlocal orbital-free kinetic energy density functional for semiconductors. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2010**, *81*, 045206.
(28) Shin, I.; Carter, E. A. Enhanced von Weizsäcker Wang-Govind-Carter kinetic energy density functional for semiconductors. *J. Chem. Phys.* **2014**, *140*, 18A531.
(29) Luo, K.; Karasiev, V. V.; Trickey, S. A simple generalized gradient approximation for the noninteracting kinetic energy density functional. *Phys. Rev. B* **2018**, *98*, 041111.
(30) Constantin, L. A.; Fabiano, E.; Della Sala, F. Semilocal Pauli-Gaussian Kinetic Functionals for Orbital-Free Density Functional Theory Calculations of Solids. *J. Phys. Chem. Lett.* **2018**, *9*, 4385−4390.
(31) Shao, X.; Mi, W.; Pavanello, M. GGA-Level Subsystem DFT Achieves Sub-kcal/mol Accuracy Intermolecular Interactions by Mimicking Nonlocal Functionals. *J. Chem. Theory Comput.* **2021**, *17*, 3455.
(32) Mi, W.; Shao, X.; Genova, A.; Ceresoli, D.; Pavanello, M. eQE 2.0: Subsystem DFT beyond GGA functionals. *Comput. Phys. Commun.* **2021**, *269*, 108122.
(33) Mi, W.; Pavanello, M. Nonlocal Subsystem Density Functional Theory. *J. Phys. Chem. Lett.* **2020**, *11*, 272−279.
(34) Schlüns, D.; Klahr, K.; Mück-Lichtenfeld, C.; Visscher, L.; Neugebauer, J. Subsystem-DFT potential-energy curves for weakly interacting systems. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14323−14341.
(35) Laricchia, S.; Fabiano, E.; Constantin, L. A.; Della Sala, F. Generalized Gradient Approximations of the Noninteracting Kinetic Energy from the Semiclassical Atom Theory: Rationalization of the Accuracy of the Frozen Density Embedding Theory for Nonbonded Interactions. *J. Chem. Theory Comput.* **2011**, *7*, 2439−2451.
(36) Götz, A.; Beyhan, S.; Visscher, L. Performance of Kinetic Energy Functionals for Interaction Energies in a Subsystem Formulation of Density Functional Theory. *J. Chem. Theory Comput.* **2009**, *5*, 3161−3174.
(37) Kevorkyants, R.; Dulak, M.; Wesolowski, T. A. Interaction energies in hydrogen-bonded systems: A testing ground for subsystem formulation of density-functional theory. *J. Chem. Phys.* **2006**, *124*, 024104.
(38) Jacob, C.; Visscher, L.*Recent Progress in Orbital-free Density Functional Theory*; Wesolowski, T. A., Wang, Y. A., Eds.; World Scientific: Singapore, 2013; pp 299−324.
(39) Krishtal, A.; Sinha, D.; Genova, A.; Pavanello, M. Subsystem density-functional theory as an effective tool for modeling ground and excited states, their dynamics and many-body interactions. *J. Phys.: Condens. Matter* **2015**, *27*, 183202.
(40) Wesolowski, T. A.; Shedge, S.; Zhou, X. Frozen-Density Embedding Strategy for Multilevel Simulations of Electronic Structure. *Chem. Rev.* **2015**, *115*, 5891−5928.
(41) Neugebauer, J. Chromophore-Specific Theoretical Spectroscopy: From Subsystem Density Functional Theory to Mode-Specific Vibrational Spectroscopy. *Phys. Rep.* **2010**, *489*, 1−87.
(42) Genova, A.; Ceresoli, D.; Krishtal, A.; Andreussi, O.; DiStasio, R., Jr.; Pavanello, M. eQE — A Densitiy Functional Embedding Theory Code For The Condensed Phase. *Int. J. Quantum Chem.* **2017**, *117*, No. e25401.
(43) Andermatt, S.; Cha, J.; Schiffmann, F.; VandeVondele, J. Combining Linear-Scaling DFT with Subsystem DFT in Born-Oppenheimer and Ehrenfest Molecular Dynamics Simulations: From Molecules to a Virus in Solution. *J. Chem. Theory Comput.* **2016**, *12*, 3214−3227.
(44) Luber, S. Local electric dipole moments for periodic systems via density functional theory embedding. *J. Chem. Phys.* **2014**, *141*, 234110.
(45) Mi, W.; Ramos, P.; Maranhao, J.; Pavanello, M. Ab Initio Structure and Dynamics of CO2 at Supercritical Conditions. *J. Phys. Chem. Lett.* **2019**, *10*, 7554−7559.

https://doi.org/10.1021/acs.jctc.2c00698
J. Chem. Theory Comput. 2022, 18, 6646−6655

(46) Genova, A.; Ceresoli, D.; Pavanello, M. Avoiding Fractional Electrons in Subsystem DFT Based Ab-Initio Molecular Dynamics Yields Accurate Models For Liquid Water and Solvated OH Radical. *J. Chem. Phys.* **2016**, *144*, 234105.
(47) Bensberg, M.; Türtscher, P. L.; Unsleber, J. P.; Reiher, M.; Neugebauer, J. Solvation free energies in subsystem density functional theory. *J. Chem. Theory Comput.* **2022**, *18*, 723−740.
(48) Goodpaster, J. D.; Barnes, T. A.; Manby, F. R.; Miller, T. F. Density functional theory embedding for correlated wavefunctions: Improved methods for open-shell systems and transition metal complexes. *J. Chem. Phys.* **2012**, *137*, 224113.
(49) Chulhai, D. V.; Goodpaster, J. D. Projection-Based Correlated Wave Function in Density Functional Theory Embedding for Periodic Systems. *J. Chem. Theory Comput.* **2018**, *14*, 1928−1942.
(50) Chulhai, D. V.; Jensen, L. Frozen Density Embedding with External Orthogonality in Delocalized Covalent Systems. *J. Chem. Theory Comput.* **2015**, *11*, 3080−3088.
(51) Bensberg, M.; Neugebauer, J. Automatic basis-set adaptation in projection-based embedding. *J. Chem. Phys.* **2019**, *150*, 184104.
(52) Bulo, R. E.; Ensing, B.; Sikkema, J.; Visscher, L. Toward a Practical Method for Adaptive QM/MM Simulations. *J. Chem. Theory Comput.* **2009**, *5*, 2212−2221.
(53) Lambros, E.; Dasgupta, S.; Palos, E.; Swee, S.; Hu, J.; Paesani, F. General Many-Body Framework for Data-Driven Potentials with Arbitrary Quantum Mechanical Accuracy: Water as a Case Study. *J. Chem. Theory Comput.* **2021**, *17*, 5635−5650.
(54) Brunken, C.; Reiher, M. Automated Construction of Quantum-Classical Hybrid Models. *J. Chem. Theory Comput.* **2021**, *17*, 3797−3813.
(55) Zhang, B.; Altarawy, D.; Barnes, T.; Turney, J. M.; Schaefer, H. F. Janus: An Extensible Open-Source Software Package for Adaptive QM/MM Methods. *J. Chem. Theory Comput.* **2019**, *15*, 4362−4373.
(56) Watanabe, H. C.; Cui, Q. Quantitative Analysis of QM/MM Boundary Artifacts and Correction in Adaptive QM/MM Simulations. *J. Chem. Theory Comput.* **2019**, *15*, 3917−3928.
(57) Jiang, K.; Nafziger, J.; Wasserman, A. Constructing a non-additive non-interacting kinetic energy functional approximation for covalent bonds from exact conditions. *J. Chem. Phys.* **2018**, *149*, 164112.
(58) Boereboom, J. M.; Fleurat-Lessard, P.; Bulo, R. E. Explicit Solvation Matters: Performance of QM/MM Solvation Models in Nucleophilic Addition. *J. Chem. Theory Comput.* **2018**, *14*, 1841−1852.
(59) Ensing, B.; Nielsen, S. O.; Moore, P. B.; Klein, M. L.; Parrinello, M. Energy Conservation in Adaptive Hybrid Atomistic/Coarse-Grain Molecular Dynamics. *J. Chem. Theory Comput.* **2007**, *3*, 1100−1105.
(60) Csányi, G.; Albaret, T.; Payne, M. C.; De Vita, A. D. Learn on the Fly: A Hybrid Classical and Quantum-Mechanical Molecular Dynamics Simulation. *Phys. Rev. Lett.* **2004**, *93*, 175503.
(61) Jacob, C. R.; Beyhan, S. M.; Visscher, L. Exact functional derivative of the nonadditive kinetic-energy bifunctional in the long-distance limit. *J. Chem. Phys.* **2007**, *126*, 234116.
(62) Jacob, C. R.; Neugebauer, J. Subsystem density-functional theory. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 325−362.
(63) Unsleber, J. P.; Dresselhaus, T.; Klahr, K.; Schnieders, D.; Böckers, M.; Barton, D.; Neugebauer, J. Serenity: A subsystem quantum chemistry program. *J. Comput. Chem.* **2018**, *39*, 788−798.
(64) Jacob, C. R.; Neugebauer, J.; Visscher, L. A Flexible Implementation of Frozen-Density Embedding for Use in Multilevel Simulations. *J. Comput. Chem.* **2008**, *29*, 1011−1018.
(65) Iannuzzi, M.; Kirchner, B.; Hutter, J. Density Functional Embedding for Molecular Systems. *Chem. Phys. Lett.* **2006**, *421*, 16−20.
(66) Genova, A.; Pavanello, M. Exploiting the locality of periodic subsystem density-functional theory: efficient sampling of the Brillouin zone. *J. Phys.: Condens. Matter* **2015**, *27*, 495501.
(67) Mi, W.; Shao, X.; Genova, A.; Ceresoli, D.; Pavanello, M. eQE 2.0: Subsystem DFT beyond GGA functionals. *Comput. Phys. Commun.* **2021**, *269*, 108122.
(68) Shao, X.; Mi, W.; Pavanello, M. Density Embedding Method for Nanoscale Molecule-Metal Interfaces. *J. Phys. Chem. Lett.* **2022**, *13*, 7147−7154.
(69) Shao, X.; Mi, W.; Pavanello, M.eDFTpy: An object-oriented platform for density embedding simulations. 2022, http://edftpy.rutgers.edu (accessed Sept 13, 2022).
(70) Giannozzi, P.; Andreussi, O.; Brumme, T.; Bunau, O.; Buongiorno Nardelli, M. B.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Cococcioni, M.; Colonna, N.; Carnimeo, I.; Dal Corso, A. D.; de Gironcoli, S.; Delugas, P.; DiStasio, R. A. D.; Ferretti, A.; Floris, A.; Fratesi, G.; Fugallo, G.; Gebauer, R.; Gerstmann, U.; Giustino, F.; Gorni, T.; Jia, J.; Kawamura, M.; Ko, H.-Y.; Kokalj, A.; Küçükbenli, E.; Lazzeri, M.; Marsili, M.; Marzari, N.; Mauri, F.; Nguyen, N. L.; Nguyen, H.-V.; Otero-de-la-Roza, A. O.; Paulatto, L.; Poncé, S.; Rocca, D.; Sabatini, R.; Santra, B.; Schlipf, M.; Seitsonen, A. P.; Smogunov, A.; Timrov, I.; Thonhauser, T.; Umari, P.; Vast, N.; Wu, X.; Baroni, S. Advanced capabilities for materials modelling with Quantum ESPRESSO. *J. Phys.: Condens. Matter* **2017**, *29*, 465901.
(71) Shao, X.; Andreussi, O.; Ceresoli, D.; Truscott, M.; Baczewski, A.; Campbell, Q.; Pavanello, M.QEpy: Quantum ESPRESSO in Python. https://gitlab.com/shaoxc/qepy (accessed Sept 13, 2022).
(72) Dalcin, L.; Fang, Y.-L. L. mpi4py: Status update after 12 years of development. *Comput. Sci. Eng.* **2021**, *23*, 47−54.
(73) Garrity, K. F.; Bennett, J. W.; Rabe, K. M.; Vanderbilt, D. Pseudopotentials for high-throughput DFT calculations. *Comput. Mater. Sci.* **2014**, *81*, 446−452.
(74) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.
(75) Cifuentes, X. S. A. C.; Musa, R. K.; Nouri, M. R.; Pavanello, M.Supplementary materials for "Adaptive Subsystem Density Functional Theory", 2022. 10.5281/zenodo.6784575 (accessed Sept 13, 2022).
(76) Andersen, H. C. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.* **1980**, *72*, 2384−2393.
(77) Kaminski, J. W.; Gusarov, S.; Wesolowski, T. A.; Kovalenko, A. Modeling Solvatochromic Shifts Using the Orbital-Free Embedding Potential at Statistically Mechanically Averaged Solvent Density. *J. Phys. Chem. A* **2010**, *114*, 6082−6096.
(78) Neugebauer, J.; Jacob, C. R.; Wesolowski, T. A.; Baerends, E. J. An Explicit Quantum Chemical Method for Modeling Large Solvation Shells Applied to Aminocoumarin C151. *J. Phys. Chem. A* **2005**, *109*, 7805−7814.
(79) Goez, A.; Neugebauer, J.Embedding Methods in Quantum Chemistry. *Frontiers of Quantum Chemistry*; Springer: Singapore, 2017; pp 139−179.
(80) Hamlin, T. A.; Swart, M.; Bickelhaupt, F. M. Nucleophilic substitution (SN2): dependence on nucleophile, leaving group, central atom, substituents, and solvent. *ChemPhysChem* **2018**, *19*, 1315−1330.
(81) Kong, Y. S.; Jhon, M. S. Solvents effect on SN2 reactions. *Theor. Chim. Acta* **1986**, *70*, 123−131.

https://doi.org/10.1021/acs.jctc.2c00698
J. Chem. Theory Comput. 2022, 18, 6646−6655