Development of a Cyclic Periodic Wave Function Approach for the Study of Infinitely Periodic Solid-State Systems

Susanne Raynor and Hua H. Song*

ABSTRACT: The \textit{ab initio} cyclic periodic wave function (CPWF) approach is developed for the treatment of infinitely periodic systems. Using the full infinite Hamiltonian operator, as well as symmetrically identical basis set wave functions that preserve the translational symmetry of the electron density of the system, this approach can be applied at the Hartree–Fock level, or correlation can be directly included by the usual modes. In this approach, all many-body interactions are included, and no edge effects occur. Initial test calculations of the CPWF method at the \textit{ab initio} Hartree–Fock level are performed on the chains of hydrogen fluoride molecules.

1. INTRODUCTION

In the past, many models and approaches have been developed for the determination of the bulk chemical properties in solid-state systems, including some approaches which are used to study infinitely periodic molecular solids.

The simplest and most popular models for studying infinitely periodic molecular solids are cluster models\cite{1,2,3,4,5} and pair-potential, or N-body potential, methods.\cite{6} The most common approaches to their studies are through the use of Bloch orbitals\cite{7} and density function theory (DFT).\cite{8,9}

Since the 1970s, cluster models have been applied in many kinds of solids.\cite{5,6,10,11} The assumption in cluster models is that the electronic structure of the infinite crystal can be reasonably simulated by performing high-level calculations on a system consisting of a few representative molecules that are geometrically selected to imitate the overall crystal structure, as shown in Figure 1. By systematically increasing the number of molecules in the clusters, it is hoped that the properties of the clusters will smoothly converge to those of the crystal, if a large enough cluster is chosen. In a review of cluster models, Fink et al.\cite{11} pointed out that cluster termination (or edge) effects are the major deficiency of the approach, when a finite cluster is used to simulate an extended or infinite system. Adachi\cite{12} concluded that systematically increasing the cluster size in the solid will eventually yield a large enough cluster to provide an adequate model for the bulk crystal. However, this approach requires great computational effort to obtain a reasonable representation of the system studied and reduce the edge effects. Zunger proposed the cyclic cluster model (CCM) in 1970s,\cite{13,14,15} which places the atoms themselves at cyclic periodic positions in space. Jug et al.\cite{16} successfully implemented the CCM at different levels of calculations to various crystalline systems. 16\textendash;20 Janetzko et al.\cite{21} applied the CCM to different covalent periodic systems, treated by the Kohn–Sham auxiliary density function theory, which successfully retained the symmetry of the corresponding periodic systems.

In pair-potential or N-body potential methods, an important assumption\cite{6} is that the intermolecular energies of N molecules can be sufficiently well approximated by the sum of interactions between N(N − 1)/2 individual pairs. This approach works very well when the interaction energies of three-body or higher-body interactions are negligible compared to the two-body interactions.\cite{22,23} Raynor\cite{4} showed that large oscillations in error occurred in studies of molecular hydrogen solid at high pressure as higher many-body terms were added. In addition, studies on solid molecular hydrogen showed that anisotropies

Figure 1. Three different cluster models labeled with solid line\textemdash\textemdash\textemdash, dashed line\textemdash\textemdash\textemdash\textemdash, and dotted line\textemdash\textemdash\textemdash\textemdash, respectively, for an imaginary periodic system.

Received: August 24, 2020
Accepted: September 22, 2020
Published: November 24, 2020
become more important either at short intermolecular distance or at high pressure. Although those anisotropies make the pair-potential methods more difficult to use, pair potentials are still being applied to model some solid-state systems. More recently, Búcko et al. successfully used a pair interactions model with a DFT approach to treat molecular crystals, which can be applied at any level of approximations.

For the study of weak interactions in large molecular systems, it is necessary to simultaneously use quite large basis sets and reasonably high-order correlation corrections, which is still very computationally demanding. To solve some of the problems involved with these systems, we have developed a very promising new \textit{ab initio} method, called the cyclic periodic wave function (CPWF) method, which uses the complete Hamiltonian for the infinite crystal and preserves the translational symmetry of the system in the wave function. In this paper, we will show the results of some initial test calculations of our method at the \textit{ab initio} Hartree–Fock level for one-dimensional (1-D) systems. We will test and apply our CPWF approach at the AM1 and PM3 semiempirical levels of approximation to a series of two-dimensional (2-D) and three-dimensional (3-D) systems to demonstrate the viability and power of our approach, and the results will be presented in two subsequent papers.

2. COMPUTATIONAL METHODS

The novel \textit{ab initio} cyclic periodic wave function (CPWF) approach uses the full infinite Hamiltonian operator for the crystal. First, the infinitely periodic system is divided into chemically equivalent repeat units, as demonstrated by Figure 2b. Figure 2a,b shows two different methods for choosing repeat units in an infinitely periodic line of molecules. Figure 2a places the cell boundaries in positions that truncate molecules, while Figure 2b places them between molecules. The cell boundaries in Figure 2b are such that the coefficients for the molecules at odd positions remain identical to those at even positions, no matter how many cell interactions are included in the calculation. The full infinite Hamiltonian operator for the system is then be written as

\[
\hat{H} = \sum_{I} \sum_{i} \left\{ \sum_{j} \left[ -\frac{1}{2} \frac{V_j^2}{r_{ij}} + \sum_{B} \frac{Z_B}{r_{ij}} + \frac{1}{2} \sum_{j'} \frac{1}{r_{ij}} \right] \right\} + \frac{1}{2} \sum_{A} \sum_{j} \sum_{B} \sum_{B} \frac{Z_A Z_B}{R_{AB}} + \sum_{A} \frac{1}{2} \frac{V_A^2}{r_{ij}} \right\} \right\}
\]

where \(I\), \(J\), \(i\), and \(j\) denote individual repeat units, \(A\) and \(B\) denote the atomic centers within repeat units \(I\) and \(J\).

To solve the Schrödinger equation for the infinitely repeating system, we will assume that each repeat unit in the perfect crystal is electronically identical to all others, differing only perhaps by one or more symmetry transformations. Since each repeat unit therefore has a chemical environment that is identical to that of every other repeat unit in the system, the electron density distribution near repeat unit \(J\) must be kept identical to that near repeat unit \(I\). Thus, we will begin by defining the ground-state Slater determinant of repeat unit molecular orbitals to have the following form

\[
\psi = \psi_{1}^{(I)} \psi_{2}^{(I)} \psi_{3}^{(I)} ... \psi_{N}^{(I)} \psi_{1}^{(J)} \psi_{2}^{(J)} \psi_{3}^{(J)} ... \psi_{N}^{(J)}
\]

where the molecular orbitals (MOs) associated with repeat unit \(I\) must have the identical mathematical form as those associated with repeat unit \(J\)—i.e., the MOs on repeat unit \(I\) are simple translations, and perhaps also rotations or reflections, of those on repeat unit \(J\). Thus, if we can determine the “best” form for the MOs on one repeat unit, we will have determined the best form for all repeat units in the infinite system. The assumption is that a single Slater determinant ground-state wave function can be readily relaxed by incorporating Møller–Plesset perturbation theory corrections to recover the correlation energy.

To create infinitely periodic molecular orbitals which preserve the translational symmetry in the system, we follow a similar prescription to that used by us in studies of infinitely periodic molecular and ionic solids, but where the Hartree–Fock–Roothaan coefficients between neighboring repeat units are directly evaluated.

The basis set used in CPWF is constructed from a set of cyclic periodic wave functions, instead of traditional Bloch orbitals. The principal difference stems from using the individual repeat units themselves in the complex exponential expansion, rather than the unit cell, as shown in the equation

\[
\psi_{i}^{I} = \psi_{i}^{(I)} \psi_{j}^{(J)}
\]

Note that eq 3 is dependent upon the assignment of position indices \((j_x, j_y, j_z)\). Each repeat unit is first assigned a unique set of three independent position indices \((j_x, j_y, j_z)\), from which the geographic position for the center of each unit can be determined. In eq 3, \(\phi_{i}^{(I)}\) is an atomic orbital located on the monomer at position \((j_x, j_y, j_z)\). Note that this atomic orbital may be a linear combination of degenerate atomic orbitals, with the particular linear combination being determined by the orientation of the monomer at position \((j_x, j_y, j_z)\), relative to its orientation at the \((0,0,0)\) position. For example, suppose that the third and fourth atomic orbitals in the basis set on the monomer at position \((0,0,0)\) are a \(p_x\) and a \(p_y\) orbital,
respectively, and further suppose that the monomer in the
(1,0,0) position is related to that at the (0,0,0) position by a
clockwise rotation of 90° about the z-axis. Then, in order for
the third and fourth atomic orbitals at (1,0,0) to be symmetrically
identical to the third and fourth ones at (0,0,0), they should be
defined as follows

$$\phi^{(1,0,0)}_3 = -p_y^{(1,0,0)}$$ and $$\phi^{(1,0,0)}_4 = p_z^{(1,0,0)}$$

Next, the repeat length $$N_i$$ in each direction must be selected.
Choosing $$N_i = 4$$ ensures that all of the interactions for repeat
units within ±1 in that index direction are treated exactly, (these
repeat units are referred to as nearest neighbors); choosing $$N_i = 6$$
e nsures that all of the interactions for repeat units within ±2 in
that index direction are treated exactly (these repeat units are
referred to as next-nearest neighbors); and so on. Equation 3
then ensures that the cyclic periodic basis set function, $$\psi_f$$, is $$N_x$$-
fold periodic in the x-index, $$N_y$$-fold periodic in the y-index, and
$$N_z$$-fold periodic in the z-index.

Next, the periodic basis set is symmetrically orthogonalized to
give a new orthonormal basis set

$$\chi^j(K_x K_y K_z) = \frac{1}{\sqrt{\sum_{i_x=i_{x1}}^{i_x=i_{x2}} \sum_{i_y=i_{y1}}^{i_x=i_{y2}} \sum_{i_z=i_{z1}}^{i_z=i_{z2}} \Delta_i \psi^j(i_x i_y i_z)}}$$

where $$j$$ references different atomic wave functions for the atoms
inside cell ($$I_x J_x$$ $$J_x$$), and $$k$$ references atomic wave functions in
cell ($$K_x K_y K_z$$). The elements of the inverse half-root overlap
matrix, $$S^{-1/2}$$ in eq 4 are generated in the usual way, i.e., from
the eigenvalues and eigenvectors of the overlap matrix formed
from the cyclic periodic wave functions defined in eq 3. The limits
for $$I_x$$, $$J_x$$, and $$J_z$$ are defined relative to $$K_x$$, $$K_y$$, and $$K_z$$ as
follows

$$K_{x1} = K_x - \left(\frac{1}{2} N_x - 1\right)$$ and $$K_{x2} = K_x + \left(\frac{1}{2} N_x - 1\right)$$

with similar relations for the y- and z-summations.

Finally, it is these functions that form the orthonormalized
basis set that is used to determine the Hartree–Fock–Roothaan
coefficients for the molecular orbitals. Thus, the $$n$$th MO for the
repeat unit at position ($$I_x$$ $$I_y$$ $$I_z$$) is given by

$$\psi_n = \left(\begin{array}{c} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{array} \right) = \sqrt{\sum_{i_x=i_{x1}}^{i_x=i_{x2}} \sum_{i_y=i_{y1}}^{i_x=i_{y2}} \sum_{i_z=i_{z1}}^{i_z=i_{z2}} \Delta_i \psi^j(i_x i_y i_z)}$$

where $$K_x$$, $$K_y$$, and $$K_z$$ are defined relative to $$I_x$$, $$I_y$$, and $$I_z$$ as
follows

$$I_{x1} = I_x - \left(\frac{1}{2} N_x - 1\right)$$ and $$I_{x2} = I_x + \left(\frac{1}{2} N_x - 1\right)$$

with similar relations for the y- and z-summations. Because the full infinite Hamiltonian is used and the infinite
translational symmetry is preserved in both the basis functions
and the Hamiltonian, the infinite Fock matrix becomes block
diagonal. Thus, we need only diagonalize the block representing
the range of ($$K_x K_y K_z$$) = (0,0,0) to ($$N_x - 1$$ $$N_y - 1$$ $$N_z - 1$$),
since the coefficients derived from it will be identical to those
needed for all other blocks. Furthermore, even this block is
further diagonalized to produce $$N_x \times N_y \times N_z$$ individual blocks
containing complex elements, as a consequence of the fact that
all of the matrix elements between periodic basis functions at
different index positions are zero:

\[
\langle \psi_n | \hat{p}_{ij} | \psi_m \rangle = (N_x N_y N_z)^{-1} \sum_{i_x-j_x}^{N_x} \sum_{i_y-j_y}^{N_y} \sum_{i_z-j_z}^{N_z} \epsilon^{2\pi i (i_x-j_x) / N_x} \epsilon^{2\pi i (i_y-j_y) / N_y} \epsilon^{2\pi i (i_z-j_z) / N_z} f_{ij}^{(i_x, i_y, i_z)} \psi_m(i_x, i_y, i_z)
\]

where

\[
f_{ij}^{(i_x, i_y, i_z)} = \langle \psi_i^{(i_x, i_y, i_z)} | \hat{p}_{ij} | \psi_j^{(i_x, i_y, i_z)} \rangle
\]

Now, redefining the second sets of summation indices as follows

$$I_x = I_x + \Delta_x; \quad I_y = I_y + \Delta_y; \quad I_z = I_z + \Delta_z$$

where $$\Delta_x$$ varies from $$-(1/2) N_x - 1$$ to $$+(1/2) N_x - 1$$, with
similar ranges for $$\Delta_y$$ and $$\Delta_z$$. Note that any matrix element
involving an interaction that is further away than $$N/2 - 1$$ in any
position index will be 0, we get the following form for eq 8

\[
\langle \psi_n^{(i_x, i_y, i_z)} | \hat{p}_{ij} | \psi_m^{(i_x, i_y, i_z)} \rangle = (N_x N_y N_z)^{-1} \sum_{\Delta_x=0}^{N_x/2-1} \sum_{\Delta_y=0}^{N_y/2-1} \sum_{\Delta_z=0}^{N_z/2-1} \epsilon^{2\pi i (I_x-J_x) / N_x} \epsilon^{2\pi i (I_y-J_y) / N_y} \epsilon^{2\pi i (I_z-J_z) / N_z} f_{ij}^{(i_x-\Delta_x, i_y+\Delta_y, i_z+\Delta_z)}
\]

Since$$f_{ij}^{(i_x, i_y, i_z)} = f_{ij}^{(i_x-\Delta_x, i_y+\Delta_y, i_z+\Delta_z)} \equiv f_{ij}^{(0,0,0)}$$, and because each
term in the first set of summations reduces to the following,

$$\sum_{i=0}^{N_x-1} \epsilon^{2\pi i (I_x-J_x) / N_x} = \begin{cases} N_x & \text{if } I_x = J_x \\ 0 & \text{if } I_x \neq J_x \end{cases}
\]

with similar relations for the y- and z-summations. We find that
the only nonzero terms in eq 10 will be those involving center $$J =$$
center $$I$$, leading to the following form

\[
\langle \psi_n^{(i_x, i_y, i_z)} | \hat{p}_{ij} | \psi_m^{(i_x, i_y, i_z)} \rangle = \sum_{\Delta_x=0}^{N_x/2-1} \sum_{\Delta_y=0}^{N_y/2-1} \sum_{\Delta_z=0}^{N_z/2-1} \epsilon^{2\pi i (I_x-J_x) / N_x} \epsilon^{2\pi i (I_y-J_y) / N_y} \epsilon^{2\pi i (I_z-J_z) / N_z} f_{ij}^{(0,0,0)} \equiv f_{ij}^{(\Delta_x, \Delta_y, \Delta_z)}
\]

where the $$f_{ij}^{(0,0,0)}$$ elements are the Fock matrix elements for
the interaction of atomic orbital i on the monomer located at
(0,0,0) with atomic orbital j on the monomer located at
(\Delta_x, \Delta_y, \Delta_z). Thus, the Hartree–Fock problem becomes a matter
of finding the eigenfunctions of $$N_x \times N_y \times N_z$$ individual
matrices, each of which is an $$M \times M$$ matrix, where $$M$$ is
the number of atomic orbitals per monomer. However, the more
efficient eigenvalue–eigenfunction routines are designed for
matrices with real components only. By taking appropriate linear
combinations of the cyclic periodic wave functions in eq 3, these
matrices can be transformed into fully real ones. The linear
combinations needed are illustrated below in the equation

\[\])}
with similar transformations for \( K_x \) and \( K_y \). The largest block that can occur after these transformations becomes one that involves matrices with dimensions of \( 8 \times 8 \). However, for most systems, further symmetry relations aid in reducing the blocks still further, and the largest blocks that have occurred in any 3-D system that we have studied thus far had dimensions of 4 \( \times \) 4. Thus, a Hartree–Fock calculation on the full infinite solid-state system is no more computationally intense than doing a few calculations on the isolated monomer, with a basis set 4 times that used for the lone monomer.

Because the translational symmetry of the infinitely periodic crystal has been preserved in the chosen form for the wave functions, the final total electron density at each repeat unit will be identical to that at every other repeat unit. Furthermore, the total energy of the system becomes partitioned into identical components from each repeat unit. Thus, each repeat unit contributes an identical energy to the total in eq 13.

The total charge density of the crystal is obtained by summing the charge density of all repeat units along the same direction. Since for each choice of the repeat length parameters \( (N_x, N_y, N_z) \), only certain specific monomers are explicitly included in the matrix elements, the calculation of the counterpoise correction effects from neighbors further out are at least partially recovered. Thus, our method is far less sensitive to the sizes chosen for the repeat lengths, \( N_x, N_y, \) and \( N_z \) than would otherwise be expected.

Finally, using cyclic periodic wave functions, it is a straightforward process to determine basis set superposition errors (BSSE) using the Boys–Bernardi counterpoise method. Since for each choice of the repeat length parameters \( (N_x, N_y, N_z) \), only certain specific monomers are explicitly included in the matrix elements, the calculation of the counterpoise correction to the energy of the isolated monomer in each case should contain exactly the same interactions in their matrices. Thus, for example, if \( N_x \) is 4, then all atomic orbitals on positions with \( I_x = -1, 0, \) and 1 would be used in the basis set expansion relative to an isolated monomer at the position \((0,0,0)\) (i.e., with \( I_x = -1, 0, \) or 1), but integrals involving \( \Delta I_x > 1 \) would be set equal to 0. These integrals would then be partially recovered, in exactly the same way as for the infinite system, by adding in the Madelung correction terms for these interactions. Thus, the BSSE-corrected interaction energy per monomer would be determined as follows.

\[ E_{\text{int}} = E^I - E^I_{\text{CP}} + \Delta E_{\text{def}} \]  

where \( E^I \) is the energy per repeat unit for the infinite system, as given in eq 14, \( E^I_{\text{CP}} \) is the counterpoise correction to the isolated repeat unit’s energy, calculated as described above, and \( \Delta E_{\text{def}} \) is the deformation energy, i.e., the energy necessary to deform the isolated repeat unit from its equilibrium geometry to its new geometry in the infinite system.

Finally, it should be noted that although the basis set wave functions are cyclically periodic, the positions of the repeat units are not. Thus, the physical positions of each repeat unit will be translationally periodic, i.e., the molecules (and their associated atomic orbitals) are placed at their expected physical positions within the crystal.
3. RESULTS AND DISCUSSION

To gain some understanding of how sensitive the interaction energies in hydrogen-bonded systems will be to the choice of the repeat length parameter, we decided to investigate two models involving the strongest hydrogen bond—that between hydrogen and fluorine in hydrogen fluoride (HF), and perform ab initio calculations on chains of hydrogen fluoride molecules. The first model is an infinite linear chain of hydrogen fluoride molecules shown in Figure 3a, and the second one is the single infinite HF catemer chain shown in Figure 3b. Our goal is to examine how different choices of the repeat length parameter, \( N_x \), affect the predicted hydrogen-bond energy in a one-dimensional system.

It has been clearly demonstrated that the accuracy of hydrogen-bond interactions is strongly influenced by the quality of the basis set used and by the degree of correlation included. Slater-type orbitals are known to have better long-range properties than Gaussian-type orbitals, and the trends in energy as a function of repeat length \( N \) were more important to us than the attainment of high accuracy. There are some very successful software packages that have been developed using DFT so far, such as the GAUSSIAN program package, the Vienna ab initio package (VASP) for the study of molecular solid systems, and the most recent CRYSTAL and CRYSCOR programs. In our case, a software package is available that applies our cyclic periodic wave function method using Slater orbitals, so we decided to employ a simple minimum basis set of Slater orbitals for our test calculations.

3.1. Effect of Repeat Index Variation on the Linear Chains of HF Molecules. The first set of calculations was performed on a very simple model: an infinite linear chain of hydrogen fluoride molecules, as illustrated in Figure 3a. Our purpose in choosing this model was twofold: (i) to investigate the effect of different choices for the repeat length parameter, \( N_x \), on the calculated hydrogen-bond energies, and (ii) to compare the efficiency of our approach when compared to cluster-type calculations. For these initial calculations, we kept the molecular \( \text{H}^{+}\text{F}^{-} \) bond distance fixed at 1.80 bohr (0.950 Å), which was the optimal distance found for the isolated HF molecule.

First, we investigated the effect on the H-bond energy of varying the \( \text{F}^{+}\text{F}^{-} \) intermolecular distance, using three different values for \( N_x \); \( N_x = 4 \), which includes nearest-neighbor interactions exclusively; \( N_x = 6 \), which also includes next-nearest neighbors; and \( N_x = 8 \), which includes next-next-nearest neighbors, as well. In addition to varying the values of \( N_x \) we also investigated the importance of the Madelung correction term \( E_{\text{Madel}} \) in eq 14. Note that all of the curves which included this term provide a significant lowering of the overall energy per monomer, as shown in Figure 4. Since this term requires very little computation time, all further discussion will be for those curves which explicitly include the calculation of \( E_{\text{Madel}} \). As shown in Figure 4, the results for \( N_x = 6 \) and \( 8 \) are nearly indistinguishable from one another, predicting nearly identical values for the optimal \( \text{F}^{+}\text{F}^{-} \) intermolecular distance and for the H-bond energy. Although the \( N_x = 4 \) calculations predict noticeably weaker H-bond energies than the other two, they recover most of the intermolecular H-bond energy (about 80%). The \( N_x = 4 \) results also predict a slightly longer \( \text{F}^{+}\text{F}^{-} \) intermolecular distance, but it differs by only about 0.1 Å from the other two. Thus, even for a system with interaction energies nearly as strong as covalent ones, the energy converges very rapidly as the length parameter, \( N_x \), is increased.

Next, we used the same infinite linear chain model to compare the trends in the repeat parameter, \( N_x \), to the trends found using a truncated chain approach. For the cluster calculations, we calculated the H-bond energy in two different ways. The first corresponds to the difference between the average energy per monomer in the truncated chain and the energy of a lone monomer

\[
\Delta E_1 = \left( E_N/N \right) - E_1
\]

where \( E_N \) is the total energy of the \( N \)-mer and \( E_1 \) is the energy of the isolated molecule. The second method views an \( N \)-mer as being composed of \((N-1)\)-mers plus a monomer, and one hydrogen bond. The average contribution per hydrogen bond then becomes

\[
\Delta E_2 = E_N - E_{N-1} - E_1
\]

with \( E_N \) and \( E_1 \) are defined as above. Figure 5 shows the effect of cluster sizes on the predicted hydrogen-bond energy convergence for the cluster models. Also shown are the results of our cyclic periodic wave function method, with \( N_x = 4, 6, \) and 8. Since \( N_x = 4 \) corresponds to exact inclusion of nearest-neighbor interactions, its value is plotted as corresponding to the amount of computational efforts on a two-membered cluster. Similarly, our results for \( N_x = 6 \) and 8 are plotted as corresponding to three- and four-membered clusters, respectively. What is clear
from Figure 5 is that our cyclic periodic wave function method converges far more rapidly to the infinite limit than do the cluster models. Also, it is clear that $\Delta E_2$ converges more rapidly than $\Delta E_1$ for the cluster models. However, regardless of the method used to determine it, the convergence of the cluster models is extremely slow when compared to the convergence of the cyclic periodic wave function method.

The importance of rapid convergence becomes even more apparent when we consider the effect on computation time. Two factors strongly affect the amount of time needed to perform a Hartree–Fock calculation on a system: the number of unique Coulomb-type integrals that must be evaluated and the size of the Fock matrix that must be diagonalized. As a measure of the first factor, we have plotted both the $\Delta E_1$ data and the cyclic periodic wave function data as a function of the number of unique Coulomb integrals that had to be calculated. These plots are shown in Figure 6. Here, we can clearly see the benefits of using a model that preserves the symmetry of the infinite system. For the cluster models, a relatively large increase in computation time produces very little improvement in convergence. This is because in the cluster models, each molecule in the cluster occupies a different chemical environment. This asymmetry greatly increases the complexity of the calculation.

It is easy to compare these approaches with regard to the second effect, Fock matrix size. For a cluster model, the Fock matrix will have a dimension of $NM \times NM$, where $M$ is the number of basis functions used on each monomer and $N$ is the number of monomers in the cluster. Since the time needed to diagonalize matrices is on the order of $n^3$ for an $n \times n$ matrix, these times will rise very rapidly with cluster size. In our approach, we will need to perform several diagonalizations of smaller Fock matrices. For a one-dimensional chain of identical molecules (linear or zig-zag), we will need to diagonalize the following: two sets of $M \times M$ matrices and $[(1/2)N_n - 1]$ sets of $2M \times 2M$ matrices. Thus, increasing the repeat length does not affect the sizes of the matrices that are diagonalized—it only affects how many such matrices need to be diagonalized. As a result, the time dependence for the cluster models will be proportional to $(NM)^3$, but it will be proportional to only $16(2N_n - 3)M^3$ for the cyclic periodic wave function approach. This means that performing a Hartree–Fock calculation on an infinite system is only moderately more time consuming than performing the calculation on an isolated molecule.

3.2. Results for Fully Optimized Catemers at SCF and Møllet–Plesset Levels. Since our method is wave function-based, we are able to directly determine correlation corrections using Møllet–Plesset perturbation theory. These corrections are demonstrated using our second model system, which replicates more closely the geometry of HF in the solid state. In the solid state, HF crystallizes in staggered chains, as illustrated in Figure 3b. Using our cyclic periodic wave function approach, we studied the effect of varying the repeat length parameter, $N_n$, on the H-bond energy in a single infinite HF catemer chain, calculated with and without correlation. We once again employed a minimum basis set of Slater orbitals for these calculations. Since the experimental FHF angle is nearly linear at 176°, we assumed that this angle remained exactly linear and optimized the H···F and F···F bond lengths, as well as the FFF angles, at the Hartree–Fock level. We then used our best geometries while calculating the Møllet–Plesset perturbation theory corrections through MP4 for $N_n = 4$ and 6, and through MP3 for $N_n = 8$. Further corrections for $N_n = 8$ did not seem justified, since agreement was so close between the other $N_n = 8$ results with the $N_n = 6$ ones.

Our results for these studies are summarized in Table 1, where we report both the total energies for each calculation as well as the predicted hydrogen-bond energy at each level of calculation. It should be noted that although our bond angles are in excellent agreement, at the Hartree–Fock level. We then used our best geometries while calculating the Møllet–Plesset perturbation theory corrections through MP4 for $N_n = 4$ and 6, and through MP3 for $N_n = 8$. Further corrections for $N_n = 8$ did not seem justified, since agreement was so close between the other $N_n = 8$ results with the $N_n = 6$ ones.

Our results for these studies are summarized in Table 1, where we report both the total energies for each calculation as well as the predicted hydrogen-bond energy at each level of calculation. It should be noted that although our bond angles are in excellent agreement, the experimental FHF angle is nearly linear at 176°, we assumed that this angle remained exactly linear and optimized the H···F and F···F bond lengths, as well as the FFF angles, at the Hartree–Fock level. We then used our best geometries while calculating the Møllet–Plesset perturbation theory corrections through MP4 for $N_n = 4$ and 6, and through MP3 for $N_n = 8$. Further corrections for $N_n = 8$ did not seem justified, since agreement was so close between the other $N_n = 8$ results with the $N_n = 6$ ones. Our results appear to be converging on nearly equal distances of 1.10 Å for both interactions. This is most likely an artifact of using a very incomplete basis set. The shortening of the distances is probably

![Figure 5](https://dx.doi.org/10.1021/acsomega.0c04094)

**Figure 5.** Comparison of the computational efforts for the H-bond energy convergence vs the cluster length. For the truncated chain approaches, the x-axis represents the number of molecules in the truncated chain. For the cyclic periodic approach, the x-axis represents the number of nearest neighbors explicitly included in the matrix elements. The line with square symbols shows the values of $\Delta E_1$ for the truncated chain models, eq 21, the line with diamond symbols shows the values of $\Delta E_2$ for the truncated chain models, eq 22, and the line with circle symbols shows the calculated H-bond energies from the cyclic periodic model. The dashed line indicates the infinite limit for all methods.

![Figure 6](https://dx.doi.org/10.1021/acsomega.0c04094)

**Figure 6.** Comparison of the calculated energy error as a function of the number of unique Coulomb integrals that must be evaluated. The line with diamond symbols is for values of $\Delta E_1$ for the truncated chain models; the line with circle symbols is for the calculated H-bond energies from the cyclic periodic model. The dashed line shows the infinite limit for all methods.
we see that the Here, we see a very interesting trend as we add correlation. First, parentheses the % errors relative to the MP4 results for similar geometries. For the energies, we have tabulated in model, and for these, we see that all three sets predict very another with the 8, and thus that total is placed within parentheses.

8, and thus that total is placed within parentheses.

seen at the MP2 level. Clearly, care must be taken when applying toward predicting more stable H-bonds relative to the Hartree hydrogen bonds. Most interesting is the trend as higher-order any correlation corrections are to be made using the Møller Plesset perturbation theory, caution must be exhibited before accepting corrections calculated only at the MP2 level.

Since our method is based upon a closed form for the wave functions, it is a straightforward process to add in correlation through the use of the Møller–Plesset perturbation theory. For the FH–F interactions in our model studies, we found that, although the MP2 level corrections led to a correction to the hydrogen-bond energies of around 1 kcal/mol, truncating the process at the MP2 level leads to considerably larger errors than those seen in the Hartree–Fock results. The optimal level appears to be at the MP3 level for these systems, since at all levels of calculation, these results were all well within 1 kcal of the MP4 results.

4. CONCLUSIONS

There are several advantages to using our cyclic periodic wave function (CPWF) approach for the treatment of infinitely periodic systems. First, this is a wave function-based method. Thus, it can be applied at any level of approximation—from semiempirical to full ab initio, with or without the inclusion of correlation and with any desired basis set.

Second, it is a chemically intuitive approach. The repeat units are chosen to be either the individual molecules within molecular solids or the individual monomer units within polymers, and are placed at their exact translationally periodic positions, as found in the infinitely periodic crystal, without any built-in periodic boundary conditions (PBCs). For chemists, it is a more intuitive way to study large organic crystals. In this way, the CPWF approach differs from the cyclic cluster model (CCM), which places the atoms or molecules at cyclic periodic positions in space, and our approach is also different from all of the methods that apply PBCs. The zeroth-order wave function corresponds to a single isolated molecule or monomer unit. Increasing the periods for the repeat lengths (N_s = 6) leads to the exact Hartree–Fock ground state in the limit of N = ∞. Longer-range interactions are at least partially recovered by the use of our Madelung correction scheme, which corrects both the individual Fock matrix elements as well as the total energy by approximately recovering all interactions between farther neighbors. We have found that using N = 4 suffices for most solid-state systems, as long as the intermonomer interactions are weaker than covalent interactions. Even when covalent interactions occur between monomers, we have found that N = 4 often suffices, though using N = 6 in the direction of the covalent interaction will provide greater accuracy for some systems. Without the inclusion of the Madelung corrections, higher values for the repeat length would be necessary to obtain equivalent accuracy.

Next, the electron density is treated as being infinitely periodic. Thus, bulk properties of solid-state systems or infinite polymers are directly calculable from the contribution of any single repeat unit in the system. Since every repeat unit will contribute an identical electron density as every other repeat unit, the chemical equivalence of the repeat units is maintained. Note that this is not the case for calculations that use a cluster
approach. In cluster models, there will be many different chemical environments within the cluster, making it difficult, if not impossible, to determine contributions from any particular unit within the cluster.

The other advantage is that the number of integrals which must be calculated and the dimensions of matrices which must be diagonalized are tremendously reduced from those necessary using either cluster approaches or N-body approaches.

Because our wave functions use individual molecules or monomers as the repeat unit, instead of the unit cell of the crystal, and because our wave functions are cyclically periodic, instead of translationally periodic, our method is origin independent and does not depend in any way upon how the repeat unit is chosen. The CPWF approach builds the symmetry relations between identical molecules or monomers directly into the wave functions, ensuring that the electron density at each individual molecule within the infinite system is identical.

An extension of this approach allows accurate inclusion of delocalization, delocalized systems can also be treated with the same method. Although the final equations were not derived here for DFT calculations, our wave function in eq 3 can be used for DFT studies, as well as for ab initio and semiempirical calculations.

Finally, the CPWF approach is applied at the AM1 and PM3 semiempirical levels of approximation to study different three-dimensional infinitely periodic solid-state systems stabilized by different kind of weak interactions between repeat units, all intra- and intermonomer geometry parameters are optimizable in this CPWF method. Thus, our approach leads to complete flexibility in the determination of the three-dimensional geometry of the solid-state system.

## AUTHOR INFORMATION

### Corresponding Author

Hua H. Song — Department of Chemistry, Rutgers University–Newark, The State University of New Jersey, Newark, New Jersey 07102, United States; orcid.org/0000-0002-0453-0557; Phone: 973-939-0463; Email: hua.song@pfizer.com

### Author

Susanne Raynor — Department of Chemistry, Rutgers University–Newark, The State University of New Jersey, Newark, New Jersey 07102, United States

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.0c04094

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The authors would like to acknowledge many valuable discussions with Professor Roger A. Lalancette. This study was supported in part by grants from the American Chemical Society Petroleum Research Fund, the Rutgers University Research Council, and Department of Chemistry, Rutgers University – Newark. The authors would like to acknowledge the help and support from Professor Phil Huskey all the time.

## REFERENCES

1. Watkins, G. D.; Messmer, R. P. Computational Methods For Large Molecules And Localized States In Solids; Herman, F.; McLean, A. D.; Nesbet, R. K., Eds.; Published by Plenum Press: New York-London, 1973; 133–147.
2. Messmer, R. P.; Watkins, G. D. Molecular-Orbital Treatment for Deep Levels in Semiconductors. Substitutional Nitrogen and the Lattice Vacancy in Diamond. Phys. Rev. B 1973, 7, 2568–2590.
3. Turi, L.; Dannenberg, J. J. Molecular Orbital Study of Crystalline Acetic Acid. 2. Aggregates in One, Two, and Three Dimensions. J. Am. Chem. Soc. 1994, 116, 8714–8721.
4. Raynor, S. Importance of High-order Interactions in Models of Molecular Solids at High Pressure. J. Chem. Phys. 1989, 91, 7018–7023.
5. Wieczorek, R.; Dannenberg, J. J. H-Bonding Cooperativity and Energetics of α-Helix Formation of Five 17-Amino Acid Peptides. J. Am. Chem. Soc. 2003, 125, 8124–8129. and refs therein.
6. (Re), F. H.; Bender, C. F. Nonadditive Interaction in Molecular Hydrogen at High Pressure. Phys. Rev. Lett. 1974, 32, 85–88.
7. (B)loch, F. Quantum Mechanics of Electrons in Crystals. Z. Phys. 1928, 52, 555–599.
8. (H)olmberg, P.; Kohn, W. Inhomogeneous Electron Gas. Phys. Rev. 1964, 136, B864–B871.
9. (K)ohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. Phys. Rev. 1965, 140, A1133–A1138.
10. (H)errmann, J.; Bludský, O. A Novel Correction Scheme for DFT: A Combined vdW-DF/CCSD(T) Approach. J. Chem. Phys. 2013, 139, 034115.
11. (F)ink, W. H.; Banerjee, A.; Simons, J. A Multiconfiguration Self-Consistent-Field Group Function Method for Problems With Repeating Potentials. J. Chem. Phys. 1983, 79, 6104–6111.
12. (A)dachi, H. Electronic State Calculation of Transition Metal Cluster. Advances in Quantum Chemistry. 1998, 29, 49–81.
13. (Z)unger, A. A Molecular Calculation of Electronic Properties of Layered Crystals: II. Periodic Small Cluster Calculation for Graphite and Boron Nitride. J. Phys. C: Solid State Phys. 1974, 7, 96–106.
14. (Z)unger, A. Small Periodic Cluster Calculation on Point Defect Problems in Hexagonal Layered Solids. J. Chem. Phys. 1975, 62, 1861–1868.
15. (Z)unger, A.; Katzir, A. Point Defects in Hexagonal Boron Nitride. II. Theoretical Studies. Phys. Rev. B. 1975, 11, 2378–2390.
16. (B)redow, T.; Evarestov, R. A.; Jug, K. Implementation of the Cyclic Cluster Model in Hartree — Fock LCAO Calculations of Crystalline Systems. Phys. Status Solidi B 2000, 222, 495–516.
17. (B)redow, T.; Geudtner, G.; Jug, K. Development of the Cyclic Cluster Approach for Ionic Systems. J. Comput. Chem. 2001, 22, 89–101.
18. (J)anetzko, F.; Bredow, T.; Jug, K. Effects of Long-range Interactions in Cyclic Cluster Calculations of Metal Oxides. J. Chem. Phys. 2002, 116, 8994–9004.
19. (J)anetzko, F.; Evarestov, R. A.; Bredow, T.; Jug, K. First-principles Periodic and Semiempirical Cyclic Cluster Calculations for Single Oxygen Vacancies in Crystalline Al2O3. Phys. Status Solidi B 2004, 241, 1032–1040.
20. (J)ug, K.; Heidberg, B.; Bredow, T. Molecular Dynamics Study of Water Adsorption Structures on the MgO(100) Surface. J. Phys. Chem. C 2007, 111, 6846–6851.
21. (J)anetzko, F.; Küster, A. M.; Salabur, D. R. Development of the Cyclic Cluster Model Formalism for Kohn-Sham Auxiliary Density Function Theory Methods. J. Chem. Phys. 2008, 128, 024102–1–024102-14.
22. (E)tters, R. D.; Danilowicz, R.; England, W. Properties of Solid and Gaseous Hydrogen, Based upon Anisotropic Pair Interactions. Phys. Rev. A. 1975, 12, 2199–2212.
23. (L)ešar, R.; Herschbach, D. R. Likelihood of a High-Pressure Phase of Solid Hydrogen Involving Termolecular Complexes. J. Phys. Chem. A 1981, 85, 3787–3792.
24. (N)ozawa, K.; Shim, N.; Makoshi, K. Theoretical Calculations for Solid Oxygen Under High Pressure. J. Phys.: Condens. Matter 2008, 20, 335219/1–335219/6.
25. (R)aich, J. C.; Anderson, A. B.; England, W. The Anisotropic Interaction Between Hydrogen Molecules. J. Chem. Phys. 1976, 64, 5088–5094.
(26) Denton, A. R.; Kahl, G.; Hafner, J. Freezing of Simple Liquid Metals. J. Non-Cryst. Solids 1999, 250-252, 15–19.
(27) Marian, C. M.; Gastreich, M. A Systematic Theoretical Study of Molecular Si/N, B/N, and Si/B/N(H) Compounds and Parameterization of a Force-Field for Molecules and Solids. J. Mol. Struct.: THEOCHEM 2000, 506, 107–129.
(28) Somasi, S.; Khomami, B.; Lovett, R. Computer Simulation of Surface and Adatom Properties of Lennard-Jones Solids. A Comparison Between Face-Centered-Cubic and Hexagonal-Close-Packed Structures. J. Chem. Phys. 2001, 114, 6315–6326.
(29) Striolo, A.; Tavares, F. W.; Bratko, D.; Blanch, H. W.; Prausnitz, J. M. Analytic Calculation of Phase Diagrams for Charged Diplolar Colloids with Orientation-Averaged Pair Potentials. Phys. Chem. Chem. Phys. 2003, 5, 4851–4857.
(30) Morald, M. Effective Pair Potential for Solid Molecular Hydrogen at High Pressures. J. Low Temp. Phys. 2012, 168, 275–284.
(31) Bučko, T.; Hafner, J.; Lebegue, S.; Angyán, G. Improved Description of the Structure of Molecular and Layered Crystals: Ab Initio DFT Calculations with van der Waals Corrections. J. Phys. Chem. A 2010, 114, 11814–11824.
(32) Bučko, T.; Simko, F. On the Structure of Crystalline and Molten Cryolite: Insights from the ab initio Molecular Dynamics in NPT Ensemble. J. Chem. Phys. 2016, 144, 064502.
(33) Marzari, N.; Mostofi, A. A.; Yates, J. R.; Souza, I.; Vanderbilt, D. Maximally Localized Wannier Functions: Theory and Applications. Rev. Mod. Phys. 2012, 84, 1419–1475.
(34) Møller, C.; Plesset, M. S. Note on the Approximation Treatment of Many-electron Systems. Phys. Rev. 1934, 46, 618–622.
(35) Krishnan, R.; Pole, J. A. Approximate Fourth-order Perturbation Theory of the Electron Correlation Energy. Int. J. Quantum Chem. 1978, 14, 91–100. and references therein.
(36) Raynor, S. Novel ab initio Self-consistent-field Approach to Molecular Solids Under Pressure. I. Theory. J. Chem. Phys. 1987, 87, 2790–2794.
(37) Raynor, S. Novel ab initio Self-consistent-field Approach to Molecular Solids Under Pressure. II. Solid H2 Under High Pressure. J. Chem. Phys. 1987, 87, 2795–2799.
(38) Raynor, S. Novel ab initio Self-consistent-field Approach to Molecular Solids Under Pressure. III. Second-order Møller-Plesset Correlation Corrections. J. Chem. Phys. 1989, 91, 3577–3581.
(39) Raynor, S. Novel ab initio Self-consistent-field Approach to Molecular Solids Under Pressure. IV. MP3 and MP4 Correlation Corrections. J. Chem. Phys. 1990, 93, 1834–1836.
(40) Raynor, S. Avoiding Truncation Errors in Hartree–Fock Theories of Periodic Systems. J. Chem. Phys. 1991, 94, 2940–2945.
(41) Löwdin, P.-O. On the Non-Orthogonality Problem Connected with the Use of Atomic Wave Functions in the Theory of Molecules and Crystals. J. Chem. Phys. 1950, 18, 365–375.
(42) Delhalle, J.; André, J. M.; Demanet, C.; Bredas, J. L. On the Calculation of Long-range Coulombic Contributions to the Direct Space LCAO CO Matrix Elements of Model Molecules. Chem. Phys. Lett. 1978, 54, 186–190.
(43) Delhalle, J.; Piela, L.; Brédas, J. L.; André, J. M. Multipole Expansion in Tight-binding Hartree–Fock Calculations for Infinite Model Polymers. Phys. Rev. B 1980, 22, 6254–6267.
(44) Boys, S. F.; Bernardi, F. The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures With Reduced Errors. Mol. Phys. 1970, 19, 553–566.
(45) Hagnstrom, S.; Shuli, H. Nature of the Two-electron Chemical Bond. III. Natural Orbitals for H. Rev. Mod. Phys. 1963, 35, 624–629. discussion 629–630.
(46) Kolos, W.; Szalewicz, K.; Monkhorst, H. J. New Born–Oppenheimer Potential Energy Curve and Vibrational Energies for the Electronic Ground State of the Hydrogen Molecule. J. Chem. Phys. 1986, 84, 3278–3283.
(47) Hehre, W. J.; Stewart, R. F.; Pople, J. A. Self-consistent Molecular-orbital Methods. I. Use of Gaussian Expansions of Slater-type Atomic Orbitals. J. Chem. Phys. 1969, 51, 2657–2664. and the series of articles for the GAUSSIAN program package afterwards.

(48) Kresse, G.; Hafner, J. ab initio Molecular Dynamics for Open-shell Transition Metals. Phys. Rev. B 1993, 48, 13115–13118.
(49) Kresse, G.; Hafner, J. ab initio Molecular-dynamics Simulation of the Liquid-metal-amorphous-semiconductor Transition in Germanium. Phys. Rev. B 1994, 49, 14251–14269.
(50) Kresse, G.; Furthmüller, J. Efficiency of ab-initio Total Energy Calculations for Metals and Semiconductors Using a Plane-wave basis Set. Comput. Mater. Sci. 1996, 6, 15–50.
(51) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for ab initio Total-energy Calculations Using a Plane-wave Basis Set. Phys. Rev. B. 1996, 54, 11169–11186.
(52) Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-wave Method. Phys. Rev. B. 1999, 59, 1758–1775.
(53) Sun, G.; Kiürti, J.; Rajczy, P.; Kertesz, M.; Hafner, J.; Kresse, G. Performance of the Vienna ab initio Simulations Package (VASP) in Chemical Applications. J. Mol. Struct.: THEOCHEM 2008, 824, 37–45.
(54) Hafner, J. Materials Simulations Using VASP – a Quantum Perspective to Materials Science. Comput. Phys. Commun. 2007, 177, 6–13.
(55) Dovesi, R.; Saunders, V. R.; Roetti, C.; Orlando, R.; Zicovich-Wilson, C. M.; Pascale, F.; Civalleri, B.; Doll, K.; Harrison, N. M.; Bush, I. J.; D’Arco, P.; Lunell, M.; Causà, M.; Noël, Y.; Maschio, L.; Erba, A.; Rerat, M.; Casassa, S. CRYSTAL User’s Manual; University of Turin: Turin, 2017.
(56) Dovesi, R.; Erba, A.; Orlando, R.; Zicovich-Wilson, C. M.; Civalleri, B.; Maschio, L.; Rerat, M.; Casassa, S.; Baima, J.; Salustro, S.; Kirtman, B. Quantum-mechanical Condensed Matter Simulations with CRYSTAL. Wiley Interdiscip. Rev.: Comput. Mol. Sci. 2018, 8, No. e1360. and references therein.
(57) Erba, A.; Halo, M. CRYSCOR09 User’s Manual; University of Turin: www.cryscor.unito.it, 2009.
(58) Pisani, C.; Schütz, M.; Casassa, S.; Usvyat, D.; Maschio, L.; Lorenz, M.; Erba, A. CRYSCOR: a Program for the Post-Hartree–Fock Treatment of Periodic Systems. Phys. Chem. Chem. Phys. 2012, 14, 7615–7628.
(59) Günther, P.; Holm, K.; Struna, H. The Structure of Solid Hydrogen Fluoride. Z. Physikal. Chem. 1939, B43, 229–239.
(60) Raynor, S.; Song, H. H. A Cyclic Periodic Wave Function Approach for the Study of Infinitely-Periodic Solid-State Systems. I. Application to the C−H⋯π(C≡C) Hydrogen Bonding Systems ACS Omega 2020 DOI: 10.1021/acsomega.0c04095.
(61) Song, H. H.; Raynor, S. A Cyclic Periodic Wave Function Approach for the Study of Infinitely-Periodic Solid-State Systems. II. Application to Helical Polysaccharides ACS Omega 2020 DOI: 10.1021/acsomega.0c04096.