

Quantitative electron delocalization in solids from maximally localized Wannier functions

A. Otero-de-la-Roza,^{1,*} Ángel Martín Pendás,¹ and Erin R. Johnson²

¹*Departamento de Química Física y Analítica, Facultad de Química, Universidad de Oviedo, 33006 Oviedo, Spain*

²*Department of Chemistry, Dalhousie University,
6274 Coburg Road, Halifax, Nova Scotia, Canada B3H 4R2*

(Dated: June 4, 2018)

Electron delocalization is the quantum-mechanical principle behind chemical concepts such as aromaticity, resonance, and bonding. A common way to measure electron delocalization in the solid state is through the visualization of maximally localized Wannier functions, a method similar to using localized orbitals in molecular quantum chemistry. Although informative, this method can only provide qualitative information, and is essentially limited by the arbitrariness in the choice of orbital rotation. Quantitative orbital-independent interatomic delocalization indices can be calculated by integration inside atomic regions of probability densities obtained from the system’s wavefunction. In particular, Bader’s delocalization indices are very informative, but typically expensive to calculate. In this article, we present a fast method to obtain the localization and delocalization indices in a periodic solid under the plane-wave/pseudopotential approximation. The efficiency of the proposed method hinges on the use of grid-based atomic integration techniques and maximally localized Wannier functions. The former enables the rapid calculation of all atomic overlap integrals required in the construction of the delocalization indices. The latter allows discarding the overlaps between maximally localized Wannier functions whose centers are far enough apart. Using the new method, all localization and delocalization indices in solids with dozens of atoms can be calculated in hours on a desktop computer. Illustrative examples are presented and studied: some simple and molecular solids, polymeric nitrogen, intermolecular delocalization in ten phases of ice, and the self-ionization of ammonia under pressure. This work is an important step towards the quantitative description of chemical bonding in solids under pressure.

INTRODUCTION

The interpretation of quantum chemical results is fraught with difficulty because traditional chemical concepts are not readily obtainable from the outcome of a quantum mechanical calculation. Chemical models developed before the advent of quantum mechanics (aromaticity, bond order, etc.) are not uniquely defined in terms of the system’s wavefunction, yet they are invaluable to understand and predict chemical phenomena. Over the years, many methods have been developed to extract chemical information from computed wavefunctions. A common technique in the solid state is the interpretation of the band structure, and of the projections of Bloch states onto localized atomic-like orbitals. The equivalent of Bloch states in molecular quantum chemistry—orbitals—have long been used to predict chemical reactivity. An alternative to orbital-based descriptors of chemical bonding is Bader’s Quantum Theory of Atoms in Molecules[1–5] (QTAIM), in which atomic regions (basins) are defined based on physical observables such as the electron density.

An important chemical concept is interatomic electron delocalization. In quantum mechanical terms, delocalization between two atoms occurs when there is a statistical correlation between their electron populations. An electron delocalized between A and B contributes to the average electron population of both. It also imposes a negative correlation between the two populations, since

the delocalized electron can be observed on A or on B, but not on both. Therefore, interatomic electron delocalization is directly calculated from the system’s wavefunction as the (negative) covariance between the population distributions of two given atoms. In addition to being related to the traditional chemical concepts of electron sharing and covalent bonding, electron delocalization is also behind physical phenomena such as magnetic exchange and electrical conductivity.

A number of local properties, the most popular of which is the electron localization function[6, 7] (ELF), have been developed to interpret chemical bonding. However, these quantities are only tangentially related to spatial delocalization, and their interpretation is often not straightforward. Other approaches are based on examining the shape of localized orbitals, calculated using some ultimately arbitrary localization procedure.[8–11] In contrast to these methods, the calculation of the population covariances in the QTAIM basins, known as localization and delocalization indices (DIs),[12] gives a direct quantitative measure of interatomic electron delocalization. DIs have found ample use in the literature, with applications such as the calculation of bond orders and the position of electron pairs[13–18], NMR coupling constants[19], and aromaticity[20], to name a few. Some of us recently used DIs to detect delocalization error from density functional approximations in halogen-bonded systems.[21]

The downside of DIs is that their computation is significantly more complex than the aforementioned alterna-

tives. At the simplest level (Hartree-Fock or Kohn-Sham density-functional theory), DIs require the calculation of the overlaps between all pairs of occupied orbitals, integrated in each atomic basin. While manageable in gas-phase molecules, the sheer number of atomic overlap integrals in infinite periodic solids makes the calculation of DIs much more challenging. To our knowledge, the only practical approach to the calculation of DIs in condensed systems has been proposed by Baranov and Kohout[22] in the context of the augmented-plane-wave (APW) method, and it requires a relatively large amount of computational resources (specifically, memory).

In this work, we present an alternative method to calculate DIs in periodic solids using the plane-wave/pseudopotential approach. The efficiency of the atomic basin integrations is greatly enhanced by using grid-specific integration methods, particularly Henkelman et al.’s[23–25] and the Yu-Trinkle method.[26, 27] In addition, the use of maximally-localized Wannier functions (MLWF) allows bypassing the calculation of all atomic overlaps involving functions whose centers are sufficiently far apart. The combination of these two ideas results in a method that efficiently calculates all the DIs in a periodic solid, and can be applied to systems with dozens of atoms in the unit cell using a desktop computer. The application of our new method to the calculation of DIs in solids under pressure will allow the quantitative description of bonding in such cases, where traditional chemical assumptions often do not apply. A few illustrative examples in this regard are provided in the last section.

THEORY

Localization and delocalization indices

The average electron population, N_A , of an atom A is given by integrating the electron density $\rho(\mathbf{r})$ over its basin:

$$N_A = \langle \hat{n}_A \rangle = \int_A \rho(\mathbf{r}) d\mathbf{r}, \quad (1)$$

In this work, we use the QTAIM atomic basins, which are defined as the region enclosed by zero-flux surfaces of the electron density.[4] Localization and delocalization are related to the variance and covariance of the atomic populations. In particular, we define the localization (λ_A) and delocalization (δ_{AB}) indices[12] as:

$$\lambda_A = N_A - \text{Var}(n_A) = \langle \hat{n}_A \rangle - (\langle \hat{n}_A^2 \rangle - \langle \hat{n}_A \rangle^2) \quad (2)$$

$$\delta_{AB} = -2 \text{Cov}(n_A, n_B) = -2(\langle \hat{n}_A \hat{n}_B \rangle - \langle \hat{n}_A \rangle \langle \hat{n}_B \rangle) \quad (3)$$

where $N_A \geq \lambda_A \geq 0$ and $\delta_{AB} \geq 0$. Inside a region where electrons are perfectly localized, $\lambda_A = N_A$ and $\delta_{AB} =$

0 for all B . The properties of variance and covariance ensure that:

$$N_A = \lambda_A + \frac{1}{2} \sum_{B \neq A} \delta_{AB} \quad (4)$$

for every atom A in the system. Thus, the average number of electrons in atom A can be effectively partitioned into “localized” (first term) and “shared” (second term).

The average values in the preceding equations are calculated by integrating the relevant one- and two-particle probability densities. Namely, the electron density ($\rho(\mathbf{r})$) and the pair density ($\pi(\mathbf{r}_1, \mathbf{r}_2)$). The latter is written in terms of an exchange-correlation density:

$$\pi(\mathbf{r}_1, \mathbf{r}_2) = \rho(\mathbf{r}_1)\rho(\mathbf{r}_2) - \rho_{xc}(\mathbf{r}_1, \mathbf{r}_2) \quad (5)$$

that measures the deviation of the pair density from the independent-electron distribution. In a Hartree-Fock calculation, the exchange-correlation density contains only the exchange contribution:

$$\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2) = \gamma(\mathbf{r}_1; \mathbf{r}_2)\gamma(\mathbf{r}_2; \mathbf{r}_1) \quad (6)$$

where $\gamma(\mathbf{r}_1; \mathbf{r}_2)$ is the one-electron density matrix:

$$\gamma(\mathbf{r}_1; \mathbf{r}_2) = \sum_i^{\text{occ}} \psi_i^*(\mathbf{r}_1)\psi_i(\mathbf{r}_2) \quad (7)$$

In a Kohn-Sham density-functional theory (DFT) calculation, only the electron density is available, but the exchange-correlation density is approximated by replacing ψ_i with the Kohn-Sham orbitals. Since the exchange part of the exchange-correlation density contains most of its features (and the correct normalization) and Kohn-Sham and Hartree-Fock orbitals are similar, these assumptions yield delocalization indices that are still chemically meaningful,[28–30] at least in cases where strong correlation effects can be ignored. All equations and results in the rest of the article use Kohn-Sham DFT.

The localization and delocalization indices can be shown to be[12]:

$$\lambda_A = F_{AA}^\alpha + F_{AA}^\beta \quad (8)$$

$$\delta_{AB} = 2(F_{AB}^\alpha + F_{AB}^\beta) \quad (9)$$

where the F_{AB}^σ are integrals of the exchange-correlation density over the basins of A and B :

$$F_{AB}^\sigma = \int_A \int_B \rho_{xc}(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \quad (10)$$

$$= \sum_{ij} S_{ji}^{A\sigma} S_{ij}^{B\sigma} \quad (11)$$

and the atomic overlap matrices $S_{ij}^{A\sigma}$ are defined as:

$$S_{ij}^{A\sigma} = \int_A \psi_i^{\sigma*}(\mathbf{r})\psi_j^\sigma(\mathbf{r}) d\mathbf{r} \quad (12)$$

The $S_{ij}^{A\sigma}$ are complex Hermitian matrices whose value depends on the particular set of orbitals used to describe the system. In contrast, F_{AB}^σ is a real symmetric matrix that does not depend on the orbital rotation and, in consequence, neither do the localization and delocalization indices.

In molecular solids, which can be partitioned into discrete fragments, we group the atomic λ_A and δ_{AB} to define molecular localization and delocalization indices:

$$\lambda_M = \sum_{A \in M} \lambda_A + \frac{1}{2} \sum_{\substack{B \in M \\ B \neq A}} \delta_{AB} \quad (13)$$

$$\delta_{MN} = \sum_{A \in M} \sum_{B \in N} \delta_{AB} \quad (14)$$

Finally, we define the valence delocalization of an atom as the variance in its electron population:

$$\Lambda_A = \text{Var}(A) = N_A - \lambda_A = \frac{1}{2} \sum_{B \neq A} \delta_{AB} \quad (15)$$

and its molecular counterpart is defined as:

$$\Lambda_M = \text{Var}(M) = N_M - \lambda_M = \frac{1}{2} \sum_{\substack{A \in M \\ B \notin M}} \delta_{AB} \quad (16)$$

where $N_M = \sum_{A \in M} N_A$.

Maximally Localized Wannier Functions

In periodic solids, the orbitals that enter the calculation of the atomic overlap matrices (Eq. 12) can take several forms. Baranov and Kohout,[22] for instance, use Bloch states:

$$\psi_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}} \quad (17)$$

written as a truncated expansion in augmented plane-waves. In this equation, \mathbf{k} is a vector in the first Brillouin zone and n is the band index, and the Bloch states are separated into a periodic part ($u_{n\mathbf{k}}(\mathbf{r})$) and a phase ($e^{i\mathbf{k}\cdot\mathbf{r}}$) by virtue of Bloch's theorem. A real-space alternative are Wannier functions[31, 32], which are obtained from Bloch states via the transformation:

$$w_{n\mathbf{R}}(\mathbf{r}) = \frac{V}{(2\pi)^3} \int \psi_{n\mathbf{k}}(\mathbf{r})e^{-i\mathbf{k}\cdot\mathbf{R}}d\mathbf{k} \quad (18)$$

where the integration goes over the first Brillouin zone and \mathbf{R} is a real-space lattice vector. Wannier functions constitute a representation of a periodic system's electronic wavefunction that is completely equivalent to the one obtained using extended Bloch states.

In a typical periodic crystal calculation, a \mathbf{k} -point sampling is chosen in order to perform the required quadrature integrations over the first Brillouin zone. A uniform sampling that uses n_1 , n_2 , and n_3 points along

the \mathbf{a}^* , \mathbf{b}^* , and \mathbf{c}^* reciprocal directions (respectively) is equivalent to assuming periodic boundary conditions for all one-electron states over a $n_1 \times n_2 \times n_3$ supercell in real space (in the following, the supercell). This results in $N = n_1 n_2 n_3$ different Wannier functions per band (Eq. 18), each represented by a lattice vector $\mathbf{R} = (i, j, k)$ with $i = 0, \dots, n_1 - 1$, $j = 0, \dots, n_2 - 1$, $k = 0, \dots, n_3 - 1$, and by the band index n . The $w_{n\mathbf{R}}(\mathbf{r})$ functions are complex in general, periodic in the supercell, and all Wannier functions belonging to the same band are related by real-space lattice translations:

$$w_{n\mathbf{R}}(\mathbf{r}) = w_{n\mathbf{0}}(\mathbf{r} - \mathbf{R}) \quad (19)$$

(In the case of partially filled bands, the Wannier transformation is ill-defined, although a ‘‘disentanglement’’ method has been proposed by Marzari et al.[11] The rest of this discussion assumes filled bands, i.e. non-metallic solids.)

Wannier functions obey normalization relations:

$$\int_{\text{s.c.}} w_{n\mathbf{R}}(\mathbf{r})w_{n'\mathbf{R}'}(\mathbf{r})d\mathbf{r} = \delta_{nn'}\delta_{\mathbf{R}-\mathbf{R}'} \quad (20)$$

where the integral extends over the supercell (‘‘s.c.’’). In addition, Wannier functions can be used to write the density and related quantities in the usual way. For instance, the one-electron density matrix is:

$$\gamma(\mathbf{r}_1; \mathbf{r}_2) = \sum_{n\mathbf{R}} w_{n\mathbf{R}}^*(\mathbf{r}_1)w_{n\mathbf{R}}(\mathbf{r}_2) \quad (21)$$

The choice of which Bloch or Wannier functions to represent a system is arbitrary, in a sense more so than with molecular wavefunctions.[11] A particular brand of the Wannier transformation that has become popular in the solid-state community uses the maximum localization criterion proposed by Marzari and Vanderbilt.[11, 33, 34] These maximally localized Wannier functions (MLWF) are obtained from Eq. 18 but using rotated Bloch states:

$$\tilde{\psi}_{n\mathbf{k}}(\mathbf{r}) = \sum_m U_{nm}^{\mathbf{k}} \psi_{m\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} \sum_m U_{nm}^{\mathbf{k}} u_{m\mathbf{k}}(\mathbf{r}) \quad (22)$$

where $U_{mn}^{\mathbf{k}}$ is a set of unitary matrices such that the combined spread of all the resulting Wannier functions is minimized.[11] This criterion yields Wannier functions that are the solid-state equivalent of Foster and Boys' localized orbitals.[8, 35] MLWFs have found considerable use in the chemical interpretation of quantum mechanical results in periodic systems, due to their connection to the electron delocalization phenomenon.[11] MLWFs are also used[11] in the context of the modern theory of polarization and in the construction of model Hamiltonians, among other uses.[36, 37] In particular, MLWFs are useful in applications that require the calculation of the exact exchange density, such as the implementation of hybrid functional Kohn-Sham methods in periodic solids.[38] As we shall see, the adequacy of MLWFs for the latter application is intimately connected to the method proposed in the present work.

Implementation

Our method for the calculation of DIs in solids uses essentially the same equations originally proposed by Bader (Eq. 8 to 12), but calculates the atomic overlap matrices in terms of Wannier functions:

$$S_{n\mathbf{R},n'\mathbf{R}'}^{A+\mathbf{R}''\sigma} = \int_A w_{n\mathbf{R}}^{\sigma*}(\mathbf{r})w_{n'\mathbf{R}'}^{\sigma}(\mathbf{r})d\mathbf{r} \quad (23)$$

The calculation of DIs in solids presents a challenge. In addition to the usual complexity in the determination of atomic Bader basins, the number of overlaps given by Eq. 23 is very much larger than in a molecular calculation of similar size. For example, a small molecular crystal such as urea has 16 atoms in the unit cell and 24 occupied bands. A $4 \times 4 \times 4$ uniform \mathbf{k} -point grid is dense enough to attain reasonably converged Brillouin zone integrations. With this choice, there are 64 distinct values of \mathbf{R} , 24 values of n , and $16 \times 64 = 1024$ atomic basins, which puts the number of atomic overlaps required to calculate all the DIs in the system at around 2.4 billion. In contrast, the equivalent calculation in the gas-phase urea molecule necessitates only a few thousand atomic overlap integrals. It is clear that, in order to calculate DIs for even the simplest solids, it is necessary to circumvent the calculation of as many atomic overlaps as possible, and to compute the non-negligible overlap integrals very efficiently.

The simplest way to reduce the number of computed atomic overlap integrals is to use their equivalence properties. The overlap matrix is Hermitian:

$$S_{n\mathbf{R},n'\mathbf{R}'}^{A+\mathbf{R}''\sigma} = \left(S_{n'\mathbf{R}',n\mathbf{R}}^{A+\mathbf{R}''\sigma}\right)^* \quad (24)$$

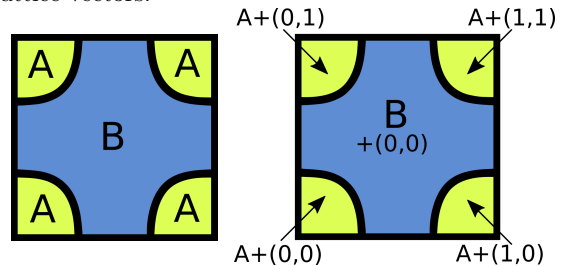
so only overlaps between unique $(n\mathbf{R},n'\mathbf{R}')$ pairs need to be calculated. In addition, the three lattice vectors in $S_{n\mathbf{R},n'\mathbf{R}'}^{A+\mathbf{R}''\sigma}$ are connected by:

$$S_{n\mathbf{R},n'\mathbf{R}'}^{A+\mathbf{R}''\sigma} = S_{n\mathbf{R}-\mathbf{R}'',n'\mathbf{R}'-\mathbf{R}''}^{A\sigma} = S_{n\mathbf{0},n'\mathbf{R}'-\mathbf{R}}^{A+\mathbf{R}''-\mathbf{R}\sigma} = S_{n\mathbf{R}-\mathbf{R}',n'\mathbf{0}}^{A+\mathbf{R}''-\mathbf{R}'\sigma} \quad (25)$$

Therefore, we only need to calculate overlaps over atomic basins in the main cell (i.e. the cell corresponding to $\mathbf{R} = \mathbf{0}$).

The calculation of atomic overlaps is greatly facilitated by the use of efficient atomic integration techniques specifically designed for scalar fields given on a uniform real-space grid, which is the natural expression for the electron density and related quantities in a plane-wave/pseudopotentials calculation. The two most popular of such approaches are the methods by Henkelman et al.[23–25] and by Yu and Trinkle (YT).[26, 27] For a given atom A , both methods determine a weight function $\omega_A(\mathbf{r})$ defined on the same grid as the electron density. For a given integrand $g(\mathbf{r})$ defined on the same grid, the corresponding atomic property (G_A) is determined by a

FIG. 1. Diagram illustrating the process of attractor remapping for the DI integration. The connected parts of the A basin that straddle the unit cell boundary are assigned different lattice vectors.



straightforward grid summation:

$$G_A = \int_A g(\mathbf{r})d\mathbf{r} = \int \omega_A(\mathbf{r})g(\mathbf{r})d\mathbf{r} \approx \frac{\left(\sum_{i=1}^N \omega_i^A g_i\right) V}{N} \quad (26)$$

where i runs over grid points (there is a total of N grid points), V is the cell volume, and ω_i^A and g_i are the values of $\omega_A(\mathbf{r})$ and $g(\mathbf{r})$ at grid point i . Unsystematic tests using the Laplacian of the electron density (which should integrate to zero inside each atomic basin) suggest that, in general, YT gives more accurate results than Henkelman et al.'s method. This is not surprising since, unlike the latter, YT weights can acquire non-zero values at the basin boundaries. The downside of the YT method is that it requires storing one grid per attractor, which in principle increases the amount of memory required to run a YT integration. However, in practice the YT atomic weights are generated on-the-fly from the sorted list of grid points and the grid point connectivity (for an in-depth description of the method, see Ref. 26). This increases the computational cost of the YT method only by a negligible amount, while drastically reducing the memory footprint. We therefore choose the YT method for all the DI integrations in this article.

As discussed above, only the atomic overlaps in the main cell need to be calculated. However, there is an important caveat. Because Wannier functions are periodic in the supercell but not in the unit cell, basins that straddle the cell boundary contribute to more than one atomic overlap integral. Before the overlaps are calculated, the connected parts of each atomic basin are determined. Each connected region is assigned a lattice vector that indicates the lattice-translated copy of the attractor to which that region belongs. This procedure remaps the list of attractors (A) into a list of attractor plus lattice vector pairs (A, \mathbf{L}) , as illustrated in Figure 1. Integration of the $(n\mathbf{R},n'\mathbf{R}')$ pair over a given connected region (A, \mathbf{L}) contributes to the atomic overlap integral $S_{n\mathbf{R}-\mathbf{L},n'\mathbf{R}'-\mathbf{L}}^{A\sigma}$ by virtue of Equation 25.

Another important aspect of the implementation concerns storage management. The calculation of the complete set of atomic overlap integrals involves a loop over

all $(n\mathbf{R}, n'\mathbf{R}')$ pairs. In a plane-wave/pseudopotential calculation, Bloch states are given by:

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{K}} c_{\mathbf{k}+\mathbf{K}}^n e^{i(\mathbf{k}+\mathbf{K})\cdot\mathbf{r}} \quad (27)$$

Combining this expression with the MLWF rotation (Eq. 22) and the Wannier transformation (Eq. 18) yields:

$$w_{n\mathbf{R}}(\mathbf{r}) = \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{R})} \sum_{\mathbf{K}} e^{i\mathbf{K}\cdot\mathbf{r}} \left(\sum_m U_{nm}^{\mathbf{k}} c_{\mathbf{k}+\mathbf{K}}^m \right) \quad (28)$$

where we have written the integral in Eq. 18 as a sum over the \mathbf{k} -points in the reciprocal-space sampling (and omitted the normalization factor for simplicity). In this equation, the factor involving the sum over reciprocal lattice vectors (\mathbf{K}) can be calculated easily by Fourier transform of the rotated plane-wave coefficients. Since all the $w_{n\mathbf{R}}(\mathbf{r})$ with the same band index are related by a phase factor, it is computationally advantageous to run a loop over pairs of bands (n, m) and generate the $w_{n\mathbf{R}}(\mathbf{r})$ and $w_{m\mathbf{R}}(\mathbf{r})$ for all \mathbf{R} simultaneously as two grids spanning the supercell. (Note that, since we only need the overlap integrals for the atoms in the main cell, a single $w_{n\mathbf{R}}(\mathbf{r})$ can be represented by a grid spanning only the main cell.) This procedure drastically reduces the computational cost, but increases the memory storage requirements. Hence, it is essential that at any given time only two supercell grids for two different bands are in memory at the same time (or one, if $n = m$). Luckily, as we shall see below, the DI calculation is not very sensitive to the number of grid points, and meaningful results can be obtained with the same density cutoffs and \mathbf{k} -point grid sizes used for a typical electronic structure calculation, or even lower.

The final implementation detail worth mentioning is that a significant number of atomic overlap integrals can be discarded *a priori* if MLWFs are used. In insulator crystals, MLWF decay exponentially.[11] Therefore, the integral in Equation 23 is assumed to give a negligible contribution if the distance between the MLWF centers is high enough. After testing, we found that a suitable criterion is to forego the calculation of all overlaps between MLWFs whose centers are at a distance higher than 4 times the sum of their spreads (defined as the square root of the quadratic spreads, Equation 18 in Ref. [11]). The adequacy of this criterion can be measured in practice by comparing the atomic electron populations (N_A) with the sum of localization and delocalization indices given in the right-hand side of Equation 4 to verify that only a negligible number of electrons has been “lost”. The DIs calculated using our method are independent of the choice of transformation used to obtain the set of Wannier functions, which is useful in cases when the iterative procedure that yields the MLWFs displays slow convergence. Using MLWFs, however, significantly reduces the computational cost in large systems.

COMPUTATIONAL DETAILS

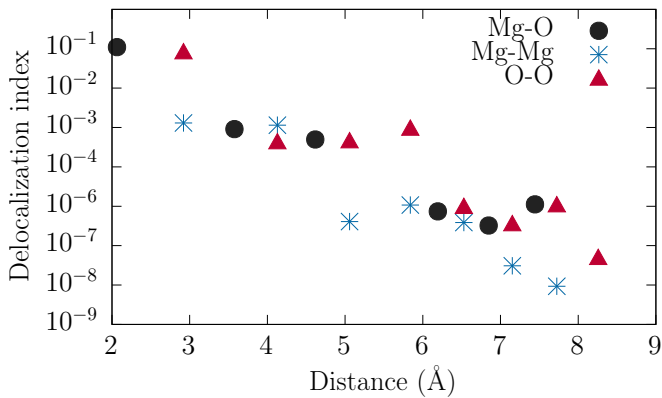
The proposed method for the calculation of the DIs in solids based on MLWFs has been implemented in the publicly available[39] critic2 code.[40, 41] All calculations were run using Quantum ESPRESSO[42] and the B86bPBE density functional[43, 44] plus the XDM dispersion correction,[45, 46] with the usual XDM parameters for this functional[47] ($a_1 = 0.6512$ and $a_2 = 1.4633\text{\AA}$). The XDM dispersion correction affects the calculated DIs indirectly through changes in the equilibrium geometry.

A single DI calculation requires the computation of the all-electron density, because pseudo-densities in plane-wave/pseudopotential calculations do not yield correct QTAIM atomic basins. In particular, the maxima associated with certain electron-depleted atoms may be too shallow or missing altogether. In this work, the all-electron densities were reconstructed from an additional projector-augmented-wave (PAW) calculation[48] at the same geometry and on the same real-space grid. These densities were used only for the purpose of computing the atomic basins but were not integrated, since they do not sum to the correct number of electrons. A norm-conserving pseudopotential calculation was used to obtain the electron density and the Bloch states from where the atomic populations and DIs were integrated. A 80 Ry cutoff for the wavefunctions and 320 Ry cutoff for the density was used for the simple and molecular solids and a 200 Ry cutoff for the wavefunctions and 800 Ry cutoff for the density in the other examples. An analysis of the dependence of the DIs on these cutoffs is performed in the next section. The interatomic distance up to which DIs are computed, and the overall cost of the DI calculation are controlled by the size of the \mathbf{k} -point grid. In all cases, we used a \mathbf{k} -point grid that was sufficient to converge the electronic energy of the system to a satisfactory degree.

Maximally-localized Wannier functions were obtained using the wannier90 code,[34] with random starting projections. A small modification to the Quantum ESPRESSO’s “pw2wannier.x” utility was necessary to write the Bloch state coefficients (an “evc” file). As discussed above, this is much more memory-efficient than handling the $u_{n\mathbf{k}}(\mathbf{r})$ or Wannier functions in real space, which would require enormous amount of disk space and memory for even the simplest systems. We verified that the calculated DIs are invariant to the choice of projection and to the convergence of the iterative procedure used in the calculation of the MLWFs. Indeed, we also verified that the calculated DIs do not change even if the wannier90 calculation is side-stepped altogether and the Bloch states are transformed directly into Wannier functions using Equation 18.

All crystal geometry plots were created using the mercury program.[49–51]

FIG. 2. Delocalization indices as a function of distance in MgO.



EXAMPLES

Simple solids

Table I shows the localization and delocalization indices calculated with our new method in some simple solids, as well as the integrated atomic populations. The picture that emerges from these results is consistent with the nature of bonding in these systems. Carbon atoms in diamond and graphite show strong covalent bonds with their nearest neighbors, with the bonds in graphite being slightly stronger than in diamond. The DIs decay with distance, and in graphite they do so more slowly inside a graphene layer. The other solids contain unlike atoms and thus feature varying degrees of interatomic charge transfer. SiC and ZnS only have partial ionic character, with relatively strong covalent bonds given by DIs in the range 0.5–0.6. In the hexagonal wurtzite-like phases (α -SiC and h-ZnS), there is a relatively large DI between the anions (C and S) and their second nearest neighbors, which are located on the z -axis opposite to the covalent bond. In the more ionic crystals (MgO and NaCl), the charge transfer between anions and cations is almost complete, and the DIs are consequently small, even between nearest neighbors. In both cases, there is a relatively large amount of delocalization between anions. The values in Table I are in quantitative agreement with the results previously reported by Baranov and Kohout[22] using completely different electronic structure and integration methods (diamond: $\delta = 0.91$; graphite: $\delta_{\text{intra}} = 1.20$, $\delta_{\text{inter}} = 0.02$; NaCl: $\delta_{\text{nn}} = 0.07$, $\delta_{\text{Cl-Cl}} = 0.05$).

Our new method calculates the DIs of all atoms in the supercell simultaneously. In order to calculate the DI between two given atoms in the crystal, the supercell must be large enough to contain both at the same time. Figure 2 shows the DIs calculated in MgO using a $8 \times 8 \times 8$ \mathbf{k} -point grid (with 1,024 atoms in the supercell and 2,048 unique DIs). The characteristic exponen-

TABLE I. Localization (λ) and delocalization indices between first (δ_1), second (δ_2), and third (δ_3) nearest neighbors in a few simple solids. The number of valence electrons (N_{elec}) and the electron population (N_A) of each attractor (“Attr.”) are given. Relevant details of the calculation are also indicated, such as the number of atoms in the unit cell (N_{cell}), the number of bands (N_{band}), the number of points in the real-space grid (“Grid”), and the \mathbf{k} -point grid (“ \mathbf{k} -point”). The two last columns are the wall time for the DI calculation[52] and the percentage of the atomic overlap integrals calculated.

System	N_{cell}	N_{band}	Grid	\mathbf{k} -point	Attr.	N_{elec}	N_A	λ	δ_1	δ_2	δ_3	Time	Overlaps
Diamond	2	4	$27 \times 27 \times 27$	$4 \times 4 \times 4$	C	4	4.000	1.809	0.920	0.040	0.008	10s	99.4%
Graph	4	8	$27 \times 27 \times 72$	$8 \times 8 \times 2$	C _{intra} C _{inter}	4	4.000	1.845	1.204	0.057	0.036	4m 22s	51.7%
α -SiC	4	8	$36 \times 36 \times 60$	$4 \times 4 \times 4$	Si	4	1.349	0.200	0.509	C	0.008	1m 50s	69.9%
c-ZnS	2	9	$45 \times 45 \times 45$	$4 \times 4 \times 4$	C	4	6.651	4.772	0.509	Si	0.124	C	80.4%
h-ZnS	4	18	$45 \times 45 \times 72$	$4 \times 4 \times 4$	Zn	12	11.118	9.797	0.599	S	0.013	Zn	2m 46s
					S	6	6.882	5.261	0.599	Zn	0.060	S	10m 56s
MgO	2	4	$32 \times 32 \times 32$	$4 \times 4 \times 4$	Mg	2	0.360	0.011	0.110	O	0.012	Zn	29.0%
					O	6	7.640	6.801	0.110	Mg	0.062	S	82.2%
NaCl	2	4	$45 \times 45 \times 45$	$4 \times 4 \times 4$	Na	1	0.186	0.004	0.059	Cl	0.000	Na	31s
					Cl	7	7.814	7.327	0.059	Na	0.051	Cl	76.6%

tial decay in insulators is apparent, as well as the fact that the delocalization index between anions (even distant ones) is almost always higher than between cations. Table I also shows some details regarding the computational performance of our method. All calculations were finished in a few minutes, with the longest being hexagonal ZnS due to the increased number of occupied bands. The last column shows the percentage of the overlaps required to calculate the DIs that were actually computed. It is clear that as the system becomes larger, the use of MLWF to discard negligible overlaps increases its effect on the computational cost of the calculation.

Finally, we examine the performance of our method as a function of the calculation parameters. Table II shows the DIs in MgO using different \mathbf{k} -point and density grids. The convergence behavior of the localization and delocalization indices mirrors that of the total energy, with minor variations in the order of 0.01 in λ and of 0.001 in δ for very large density grids or \mathbf{k} -point samplings. In the case of the \mathbf{k} -point convergence, small sampling sizes show the characteristic “folding”, whereby DIs between close atoms and localization indices are both overestimated. Convergence to chemically meaningful values of λ and δ is achieved for a $4 \times 4 \times 4$ or even a $3 \times 3 \times 3$ grid. As expected, the computational saving from using MLWFs to discard overlaps increases with the size of the \mathbf{k} -point grid. Surprisingly, the localization and delocalization indices are only mildly dependent on the density and wavefunction cutoffs, which control the size of the real-space density grid. In the whole range of grids examined, λ changes by at most 0.02, and the DIs are stable to within 0.003. In terms of its significance, this variation is much smaller than that of the total energy, which changes by about 0.2 Ry throughout the range. As expected, the computational cost increases with the size of the density grid.

Molecular crystals

In addition to simple covalent- and ionic solids, we also considered the five molecular solids whose crystal geometries are shown in Figure 3. Table III shows that the computational savings from using MLWFs are even more apparent in these cases. As expected, ice and urea show a significant amount of intermolecular delocalization between hydrogen-bonded molecules. The hydrogen bonds occur between nearest neighbors (NN) in ice, and with the first and second NN in urea. In the latter, the second NN corresponds to the chain of double hydrogen bonds along the c axis, and is correspondingly higher than the δ_1 . A similar effect is observed in the Br₂ crystal (Figure 3c), where it is the third NN who is involved in a halogen-bonding interaction with the parent molecule, forming chains along the c axis. The halogen-bonded DI (0.448) is substantially higher than the two hydrogen

bonds reported, indicative of a much stronger covalent character. Benzene and CO₂ do not feature strong directional interactions. However, while in CO₂ the first NN DIs are very small and the remaining DIs are negligible, there is a substantial degree of delocalization in benzene, probably caused by the overlap between π electron clouds.

Polymeric nitrogen

We now apply our new method to a few systems from recent studies in the literature. A particularly interesting field of application for the DIs is the analysis of chemical bonding under pressure because DIs provide a quantitative measure of bonding in situations where common chemical rules often do not apply. An example is the polymerization of nitrogen under extreme pressure.

Due to its extreme stability, the question of whether there is a high-pressure phase of nitrogen where its triple bond is absent has been debated extensively (see Ref. [53] and references therein). This problem also has potential practical implications, since a polymeric phase of nitrogen that is metastable at ambient pressure and temperature would make an excellent high-energy-density material.[54] At zero pressure, nitrogen exists in two similar molecular phases: a cubic (fcc) low-temperature α phase, and a high-temperature hexagonal (hcp) β phase.[55] For simplicity, we consider only β -N₂. [56] In 2004, Eremets et al. reported the first experimental observation of a stable polymeric phase of nitrogen at 110 GPa and over 2000 K.[54] After analyzing the powder diffraction patterns of the new phase, it was concluded that its structure coincided with the cubic gauche form of nitrogen (cg-N), previously studied computationally by Mailhot et al.[57]

The crystal geometries for the β -N₂ and cg-N phases are shown in Figure 4. Calculations were run using energy-converged \mathbf{k} -point grids ($2 \times 2 \times 2$ for β -N₂ and $4 \times 4 \times 4$ for cg-N) and wavefunction and density cutoffs of 200 and 800 Ry, respectively. Calculation of the DIs shows that in the β -N₂ phase, each nitrogen has valence delocalization $\Lambda = 1.58$, with the majority of the delocalized electrons employed in the triple bond with the adjacent nitrogen atom ($\delta = 3.00$). As expected from the relatively open structure, all the DIs between a single N atom and the atoms not directly bonded to it are very small, with the largest being the DIs to those atoms in closest proximity ($\delta = 0.033$). The intermolecular DIs in β -N₂ are similarly small ($\delta = 0.044$). These results are in sharp contrast with the cg-N phase. The increased pressure results in a higher number of delocalized valence electrons ($\Lambda = 2.13$ [58]), but in this phase each nitrogen atom is single-bonded to its immediate neighbors, with $\delta = 1.11$. In addition, there is significant delocalization to the second ($\delta = 0.072$), and particularly the three

TABLE II. Localization (λ) and delocalization indices between first (δ_1), second (δ_2), and third (δ_3) nearest neighbors in MgO as a function of \mathbf{k} -point grid and density and wavefunction cutoffs (in Ry). The calculations with variable \mathbf{k} -point grid use a $32 \times 32 \times 32$ density grid. The calculations with variable density grid (indicated in the second row for each entry) use a $4 \times 4 \times 4$ \mathbf{k} -point grid. Wall time for the calculation[52], the percentage of atomic overlaps calculated, and the SCF energy are indicated in the last three columns.

\mathbf{k} -point		λ	δ_1	δ_2	δ_3	Time	Overlaps	Energy (Ry)
$1 \times 1 \times 1$	Mg	0.005	0.355	—	—	0s	100%	-34.968 424 32
	O	7.640	0.355	0.000	0.000			
$2 \times 2 \times 2$	Mg	0.009	0.101	0.002	0.002	0s	100%	-34.973 131 49
	O	6.981	0.101	0.129	0.002			
$3 \times 3 \times 3$	Mg	0.012	0.108	0.001	0.001	3s	100%	-34.973 437 20
	O	6.830	0.108	0.080	0.001			
$4 \times 4 \times 4$	Mg	0.011	0.110	0.001	0.001	12s	82.2%	-34.973 579 78
	O	6.801	0.110	0.082	0.001			
$6 \times 6 \times 6$	Mg	0.010	0.110	0.001	0.001	1m 4s	27.4%	-34.973 636 67
	O	6.791	0.110	0.082	0.001			
$8 \times 8 \times 8$	Mg	0.011	0.110	0.001	0.001	5m 5s	12.1%	-34.973 662 09
	O	6.794	0.110	0.082	0.001			
$10 \times 10 \times 10$	Mg	0.009	0.110	0.001	0.001	16m 14s	6.2%	-34.973 673 21
	O	6.791	0.110	0.082	0.001			
$E_{\text{cut}}(\rho/\text{wfn})$		λ	δ_1	δ_2	δ_3	Time		Energy (Ry)
60/240	Mg	0.011	0.110	0.001	0.001	12s		-34.830 032 04
(27 \times 27 \times 27)	O	6.801	0.110	0.082	0.001			
80/320	Mg	0.012	0.110	0.001	0.001	13s		-34.973 579 78
(32 \times 32 \times 32)	O	6.802	0.110	0.081	0.001			
100/400	Mg	0.012	0.109	0.001	0.001	18s		-34.999 280 87
(36 \times 36 \times 36)	O	6.805	0.109	0.081	0.001			
120/480	Mg	0.011	0.109	0.001	0.001	26s		-35.002 234 76
(40 \times 40 \times 40)	O	6.808	0.109	0.082	0.001			
160/640	Mg	0.011	0.108	0.001	0.001	39s		-35.003 524 92
(45 \times 45 \times 45)	O	6.810	0.108	0.082	0.001			
200/800	Mg	0.011	0.108	0.001	0.001	54s		-35.004 278 17
(50 \times 50 \times 50)	O	6.812	0.108	0.082	0.001			
300/1200	Mg	0.011	0.107	0.001	0.001	2m 14s		-35.004 411 79
(64 \times 64 \times 64)	O	6.815	0.107	0.082	0.001			
400/1600	Mg	0.011	0.107	0.001	0.001	3m 18s		-35.004 424 42
(72 \times 72 \times 72)	O	6.816	0.107	0.082	0.001			
600/2400	Mg	0.011	0.107	0.001	0.001	7m 45s		-35.004 427 02
(90 \times 90 \times 90)	O	6.818	0.107	0.082	0.001			

TABLE III. Molecular valence delocalization (Λ) and intermolecular delocalization indices between first (δ_1), second (δ_2), and third (δ_3) nearest-neighbor molecules in a few simple molecular solids. Several details of the DI calculation are also indicated, such as the number of atoms in the unit cell (N_{cell}), the number of bands (N_{band}), the number of points in the real-space grid, and the \mathbf{k} -point grid. The two last columns are the wall time for the DI calculation[52] and the percentage of the atomic overlap integrals calculated. (Note: first, second, etc. nearest neighbors are defined using the distance between the molecular centers of mass.)

System	N_{cell}	N_{band}	Grid	\mathbf{k} -point	Λ	δ_1	δ_2	δ_3	Time	Overlaps
CO ₂	12	32	64 \times 64 \times 64	4 \times 4 \times 4	0.318	0.053	0.000	0.000	2h 31m 35s	19.0%
Br ₂ (c2m)	8	28	75 \times 45 \times 108	4 \times 4 \times 4	1.356	0.102	0.134	0.448	2h 2m 51s	22.5%
Ice (Ih)	48	64	96 \times 81 \times 80	2 \times 2 \times 2	0.566	0.273	0.003	0.000	3h 29m 45s	40.8%
Benzene	48	60	80 \times 108 \times 72	2 \times 2 \times 2	0.716	0.166	0.104	0.085	4h 8m 39s	46.9%
Urea	16	24	60 \times 60 \times 54	4 \times 4 \times 4	0.808	0.183	0.265	0.078	1h 20m 41s	19.0%

third nearest neighbors ($\delta = 0.110$). These results are consistent with previous bonding analyses based on the calculated band structure of the cg-N phase.[59]

Ice phases

Next we show how DIs calculated using our method can be used to examine intermolecular delocalization in solids by applying it to ten phases of water ice. The

FIG. 3. Geometries for the molecular crystals in Table III: (a) CO₂, (b) Br₂ (c2m phase), (c) Ice (Ih phase), (d) Benzene, and (e) urea. Hydrogen bonds are indicated by stippled lines in ice and urea.

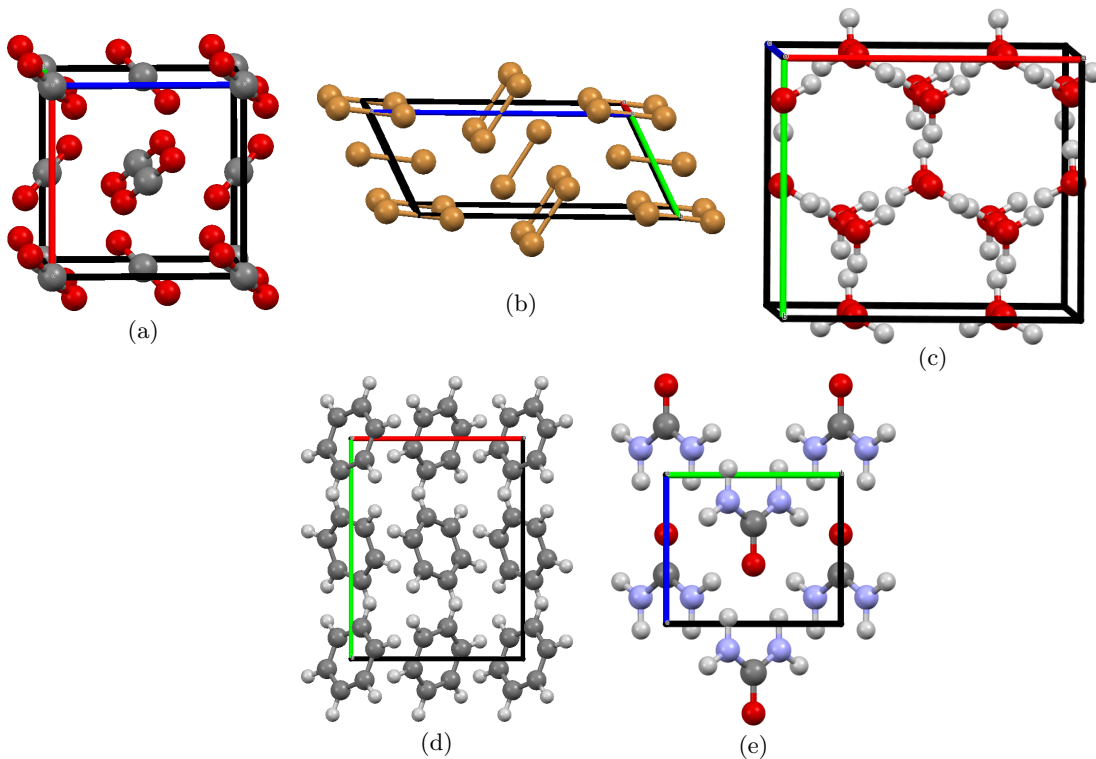
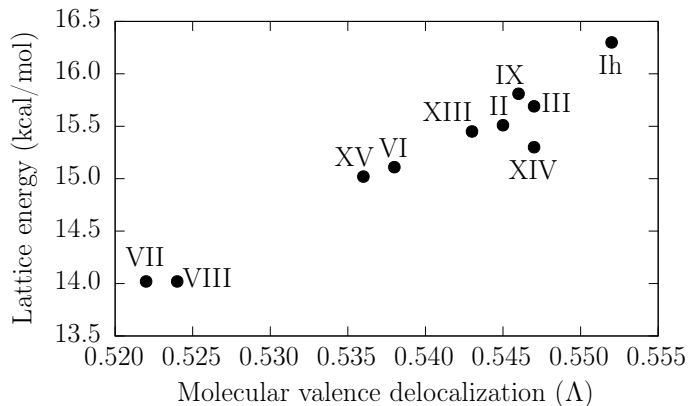


TABLE IV. Left: calculated lattice energies per molecule (E_{lat}) and nearest-neighbor intermolecular delocalization indices (δ_{nn}) for the ten ice polymorphs in the ICE10 set by Brandenburg et al.[60] The nearest neighbor distance (d_{nn}) and the molecular valence delocalization (Λ) are also given. Right: lattice energy per molecule plotted against molecular valence delocalization for the ten ice phases.

Poly	Λ	d_{nn} (Å)	δ_{nn}	E_{lat} (kcal/mol)
Ih	0.552	2.679	0.267	16.30
II	0.545	2.718	0.232	15.51
III	0.547	2.693	0.247	15.69
VI	0.538	2.742	0.210	15.11
VII	0.522	2.869	0.067	14.02
VIII	0.524	2.911	0.153	14.02
		2.912	0.154	
IX	0.546	2.692	0.246	15.81
XIII	0.543	2.699	0.238	15.45
XIV	0.547	2.747	0.218	15.30
XV	0.536	2.741	0.208	15.02

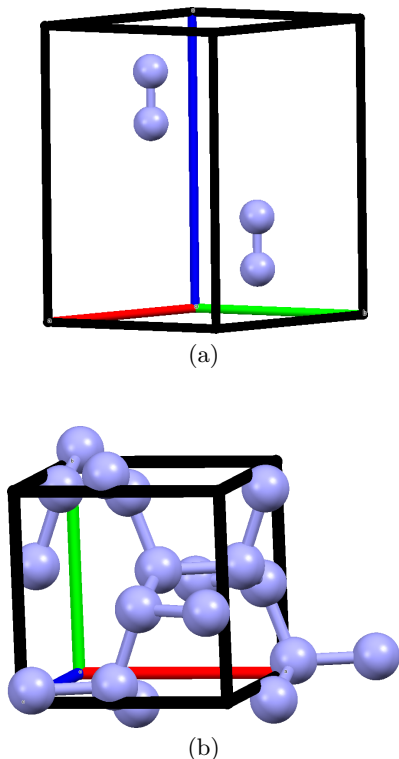


starting crystal geometries are from the ICE10 set by Brandenburg et al.[60] All ten crystals were relaxed using the PAW method and the B86bPBE-XDM functional, a $4 \times 4 \times 4$ \mathbf{k} -point grid, a 100 Ry wavefunction cutoff, and a 1000 Ry density cutoff. After an exploration of different \mathbf{k} -point grid sizes previous to the DI calculation, we decided to use a $2 \times 2 \times 2$ grid for all phases except VIII and XIV, which we ran using a $3 \times 3 \times 3$ grid. The wavefunction and density cutoffs were 200 and 800 Ry

respectively.

The results are shown in Table IV (see Figure 1 in Ref. [60] for a depiction of each phase). The lattice energies are very similar to the TPSS-D3 results reported in Brandenburg et al.[60], and in good agreement with the proposed reference values. Regarding bonding, ice is unlike water dimer in that none of its phases present any intermolecular charge transfer. However, there is a significant amount of intermolecular delocalization that,

FIG. 4. Crystal geometries of the two phases of nitrogen considered: (a) α -N₂ (zero pressure and 46 K[56]) and (b) cg-N (at 115 GPa and ≥ 2000 K[54]).



similarly to the results presented in Table III, extends only to the molecules in the immediate vicinity of a given water molecule. The plot in Table IV shows that there is a strong linear correlation between intermolecular delocalization and stability. To a lesser extent, this correlation also extends to the nearest-neighbor distances. The lowest-energy phase and also the phase with highest intermolecular delocalization is Ih, which is the thermodynamically stable phase of ice under ambient-conditions.

Phases VII and VIII are particular in that there is significant delocalization between molecules that are not directly hydrogen-bonded. Phase VII is approximately cubic, with water molecules at the corners of the unit cell, at the center of edges, faces, and body, and at the tetrahedral sites of the cube. A given molecule is therefore surrounded by other eight molecules, two of which give, and two receive, a hydrogen bond. The other four neighboring molecules are not hydrogen-bonded. Table IV shows that, while the DIs to the four hydrogen-bonded neighbors are significant (0.153), they are substantially smaller than the other phases, and there is very significant delocalization to the adjacent non-bonded molecules (≈ 0.07), who are also slightly closer to the central molecule. A similar effect can be observed in phase VIII, where each water molecule admits twelve other molecules in its vicinity, only four of which are directly hydrogen-bonded to it. While less stable than the rest at zero

pressure, phases VII and VIII present a more efficient crystal packing.[60] This explains why they are the thermodynamically stable phases under high pressure, from approximately 2 GPa up to tens of GPa.[61] Our DIs show that the effect of pressure on ice is to stabilize structures with significant delocalization between non-hydrogen-bonded water molecules, and to decrease overall per-molecule delocalization, as a means to obtain a more efficient packing.

Ammonia self-ionization under pressure

Another example of unusual chemical behavior under extreme pressure is the self-ionization of ammonia. Bonding in molecular materials under pressure is a topic of discussion in the literature. In particular, whether hydrogen bonds are symmetrized upon application of extreme pressure.[62] By performing a computational search for stable polymorphs in ammonia, Pickard and Needs[62] predicted a disproportionation to an ammonium amide phase with Pma2 space group in the 90–331 GPa range. In addition, this Pma2 phase was predicted to revert to a neutral hydrogen-bonded phase at pressures higher than 331 GPa. In a recent study, the transition to the ionic Pma2 phase was experimentally observed to occur at ≈ 120 GPa using a high-pressure diamond anvil cell technique.[63] Although rare, self-ionization under pressure is a phenomenon that had been observed previously in nitrogen oxides.[64–66] However, it is interesting that ammonia shows a radically different behavior under pressure than water, which symmetrizes its hydrogen bond at pressures higher than 60 GPa.[67] This observation could have relevant implications regarding our understanding of planetary geology in the Solar System.[62]

In addition to the self-ionized Pma2 phase, several other phases of ammonia are known.[62, 68] Phase I (P2₁3) is stable from 1 GPa (at room temperature) up to 4 GPa. Then, it transforms to phase IV (space group P2₁2₁2₁), which is stable up to ≈ 14 GPa, when it transforms into phase V, whose structure is not well known. Figure 5 shows the three phases examined. The geometries were taken from the work of Pickard and Needs[62] at 5 GPa (P2₁3 and P2₁2₁2₁) and at 100 GPa (Pma2). After exploration of the total energy as a function of \mathbf{k} -point grid size, the grids used were $2 \times 2 \times 2$ (P2₁3), $3 \times 3 \times 3$ (P2₁2₁2₁), and $2 \times 4 \times 4$ (Pma2). Wavefunction and density cutoffs of 200 Ry and 800 Ry, respectively, were employed in the calculation.

The calculated delocalization indices for the three phases are shown in Table V. Phase P2₁3 is reminiscent of the low-pressure phases of ice (Table IV) in that there is only a single hydrogen-bonded interaction with substantial delocalization (0.145, to each of the six neighboring molecules). However, this DI is smaller than in

FIG. 5. The three phases of ammonia studied in this work: a) Phase I ($P2_13$), stable between 1 and 4 GPa, b) phase IV ($P2_12_12_1$), observed between 4 GPa and 14 GPa, and c) self-ionized phase (Pma2), stable above approximately 120 GPa. The crystal geometries are from Pickard and Needs.[62]

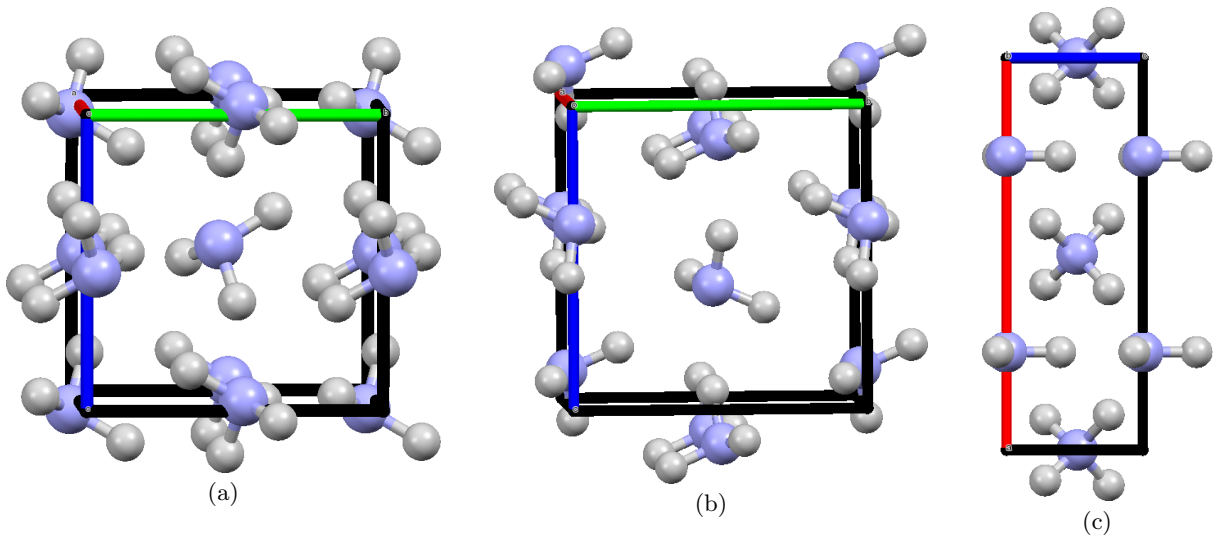


TABLE V. Molecular electron population (N), molecular localization index (λ), valence delocalization (Λ), and intermolecular delocalization indices as a function of increasing distance between centers of mass (δ). In the first two phases, the DIs followed by a * symbol are between hydrogen-bonded molecules. In Pma2, the + and - symbols indicate whether the delocalization is with a cation or an anion.

Phase	Mol.	N	λ	Λ	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6
$P2_13$	NH_3	8.000	7.412	0.588	0.145	*	0.047	0.002		
$P2_12_12_1$	NH_3	8.000	7.425	0.575	0.160	*	0.066	0.103	*	0.126
Pma2	NH_4^+	8.432	7.132	1.300	0.167	+	0.270	-	0.131	-
	NH_2^-	7.568	6.099	1.469	0.264	-	0.270	+	0.131	+

water, which is consistent with the fact that hydrogen bonds in ammonia are weaker.[62] Similarly to phases VII and VIII of ice (Table IV), high pressure stabilizes denser crystal packings at the expense of weaker hydrogen bonds and decreased molecular delocalization. This is the case in ammonia as well: Phase $P2_12_12_1$ is denser than $P2_13$ (unit cell volume = 97.2 \AA^3 at 5 GPa, c.f. 99.9 \AA^3 for $P2_13$ at the same pressure) and there is non-negligible delocalization between neighboring molecules that are not directly hydrogen-bonded. The six DIs for $P2_12_12_1$ in Table V involve 12 molecules in a relatively narrow distance range (3.1–3.4 \AA).

The self-ionized phase (Pma2) shows a transfer of almost half an electron between the ammonium and the amide moieties. The amide anion is more delocalized, which is similar to the behavior of other ionic solids (see Table I). As expected, the delocalization between unlike-charge moieties is strongest, despite being slightly farther away from each other. The DIs to the eight nearest unlike-charge neighbors range between 0.13 and 0.35. Interestingly, there is also significant delocalization between cations and, particularly, between anions, with all like-charge moieties appearing in layers perpendicular to the a axis. This relatively large intra-layer delocalization,

which occurs even between cations that are not hydrogen-bonded, can be used to justify the layered structure in this system.

CONCLUSIONS

In this work, we presented the development and implementation of an efficient method for the calculation of Bader's localization and delocalization indices in periodic solids. The new method is based on the pseudopotentials/plane-waves approach, and its efficiency rests on two key features: the rapid integration of atomic overlap integrals using grid-based integration techniques (the Yu-Trinkle algorithm), and the fact that, in large systems, the calculation of a significant fraction of the atomic overlaps required to build the DIs can be circumvented using maximally localized Wannier functions (MLWF). Specifically, the overlap between MLWFs whose centers are far enough from each other is assumed to give a negligible contribution to the delocalization indices. However, although using MLWFs is computationally more efficient, the calculated DIs are independent of the particular choice of Wannier transformation. Several

key features of the implementation, such as the management of disk storage and memory are also discussed.

Using the proposed method, all localization and delocalization indices in periodic solids with dozens of atoms in the unit cell can be calculated in minutes or hours on a desktop computer. Surprisingly, DIs are not very sensitive to the \mathbf{k} -point grid and, particularly, to the size of the real-space density grid, which allows reducing the computational cost even further. For now, our method is limited to systems with filled bands (i.e. non-metals) and to norm-conserving pseudopotential calculations.

A range of illustrative examples were presented, including some simple solids (diamond, graphite, SiC, MgO, NaCl, and ZnS) and five molecular solids (benzene, ice, urea, CO₂, and Br₂). The calculated DIs offer a picture that is consistent with common chemical knowledge, and numerically almost the same as those calculated using the method previously proposed by Baranov and Kohout in the context of the augmented-plane-wave (APW) method. We also showed the usefulness of our new method in practice by examining three recent cases from the literature: polymeric nitrogen, the relative stability of ten water ice phases, and the self-ionization of ammonia under pressure. This work opens the door to the quantitative description of bonding in periodic solids under pressure.

ACKNOWLEDGEMENTS

* aoterodelaroza@gmail.com

- [1] R. F. Bader, *Acc. Chem. Res.* **8**, 34 (1975).
- [2] R. F. Bader, *Acc. Chem. Res.* **18**, 9 (1985).
- [3] R. F. W. Bader, *Chem. Rev.* **91**, 893 (1991).
- [4] R. F. W. Bader, *Atoms in Molecules. A Quantum Theory* (Oxford University Press, Oxford, 1990).
- [5] R. Boyd and C. Matta, eds., *The Quantum Theory of Atoms in Molecules* (Weinheim, Germany, 2007) pp. 207–230.
- [6] A. D. Becke and K. E. Edgecombe, *J. Chem. Phys.* **92**, 5397 (1990).
- [7] B. Silvi and A. Savin, *Nature* **371**, 683 (1994).
- [8] J. Foster and S. Boys, *Rev. Mod. Phys.* **32**, 300 (1960).
- [9] C. Edmiston and K. Ruedenberg, *Rev. Mod. Phys.* **35**, 457 (1963).
- [10] J. Pipek and P. G. Mezey, *The Journal of Chemical Physics* **90**, 4916 (1989).
- [11] N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt, *Rev. Mod. Phys.* **84**, 1419 (2012).
- [12] R. F. Bader and M. E. Stephens, *J. Am. Chem. Soc.* **97**, 7391 (1975).
- [13] J. Cioslowski and S. T. Mixon, *J. Am. Chem. Soc.* **113**, 4142 (1991).
- [14] J. G. Ángyán, M. Loos, and I. Mayer, *J. Phys. Chem.* **98**, 5244 (1994).
- [15] J. G. Ángyán, E. Rosta, and P. R. Surján, *Chem. Phys. Lett.* **299**, 1 (1999).
- [16] X. Fradera, M. A. Austen, and R. F. Bader, *J. Phys. Chem. A* **103**, 304 (1999).
- [17] C. Matta and R. Boyd, in *The Quantum Theory of Atoms in Molecules*, edited by R. Boyd and C. Matta (Wiley-VCH, Weinheim, Germany, 2007) pp. 1–34.
- [18] E. Matito, J. Poater, M. Solà, M. Duran, P. Salvador, *et al.*, *J. Phys. Chem. A* **109**, 9904 (2005).
- [19] J. Hernández-Trujillo, F. Cortés-Guzmán, and G. Cuevas, in *The Quantum Theory of Atoms in Molecules*, edited by R. Boyd and C. Matta (Wiley-VCH, Weinheim, Germany, 2007) pp. 375–398.
- [20] E. Matito, J. Poater, and M. Solà, in *The Quantum Theory of Atoms in Molecules*, edited by R. Boyd and C. Matta (Wiley-VCH, Weinheim, Germany, 2007) pp. 399–424.
- [21] A. Otero-de-la-Roza, E. R. Johnson, and G. A. DiLabio, *J. Chem. Theory Comput.* **10**, 5436 (2014).
- [22] A. I. Baranov and M. Kohout, *J. Comput. Chem.* **32**, 2064 (2011).
- [23] G. Henkelman, A. Arnaldsson, and H. Jonsson, *Comput. Mater. Sci.* **36**, 354 (2006).
- [24] E. Sanville, S. D. Kenny, R. Smith, and G. Henkelman, *J. Comput. Chem.* **28**, 899 (2007).
- [25] W. Tang, E. Sanville, and G. Henkelman, *J. Phys. Condens. Matter* **21**, 084204 (2009).
- [26] M. Yu and D. R. Trinkle, *J. Chem. Phys.* **134**, 064111 (2011).
- [27] M. Yu, D. R. Trinkle, and R. M. Martin, *Phys. Rev. B* **83**, 115113 (2011).
- [28] T. Kar, J. G. Ángyán, and A. Sannigrahi, *J. Phys. Chem. A* **104**, 9953 (2000).
- [29] X. Fradera, J. Poater, S. Simon, M. Duran, and M. Solà, *Theor. Chem. Acc.* **108**, 214 (2002).
- [30] J. Poater, M. Sola, M. Duran, and X. Fradera, *Theor. Chem. Acc.* **107**, 362 (2002).
- [31] J. Slater and W. Shockley, *Phys. Rev.* **50**, 705 (1936).
- [32] G. H. Wannier, *Phys. Rev.* **52**, 191 (1937).
- [33] N. Marzari and D. Vanderbilt, *Phys. Rev. B* **56**, 12847 (1997).
- [34] A. A. Mostofi, J. R. Yates, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari, *Comput. Phys. Commun.* **178**, 685 (2008).
- [35] J. Foster and S. Boys, *Rev. Mod. Phys.* **32**, 303 (1960).
- [36] P. L. Silvestrelli, *Phys. Rev. Lett.* **100**, 053002 (2008).
- [37] L. Andrinopoulos, N. D. Hine, and A. A. Mostofi, *J. Chem. Phys.* **135**, 154105 (2011).
- [38] X. Wu, A. Selloni, and R. Car, *Physical Review B* **79**, 085102 (2009).
- [39] A. Otero-de-la-Roza, “The critic2 program,” <https://github.com/aoterodelaroza/critic2> (2018).
- [40] A. Otero-de-la-Roza, M. Blanco, A. M. Pendás, and V. Luaña, *Comput. Phys. Commun.* **180**, 157 (2009).
- [41] A. Otero-de-la-Roza, E. R. Johnson, and V. Luaña, *Comput. Phys. Commun.* **185**, 1007 (2014).
- [42] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, *et al.*, *J. Phys.: Condens. Matter* **29**, 465901 (2017).
- [43] A. Becke, *J. Chem. Phys.* **85**, 7184 (1986).
- [44] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).

- [45] A. D. Becke and E. R. Johnson, *J. Chem. Phys.* **127**, 154108 (2007).
- [46] A. Otero-de-la Roza and E. R. Johnson, *J. Chem. Phys.* **136**, 174109 (2012).
- [47] <http://schooner.chem.dal.ca/wiki/XDM> (2018).
- [48] P. Blöchl, *Phys. Rev. B* **50**, 17953 (1994).
- [49] C. F. Macrae, I. J. Bruno, J. A. Chisholm, P. R. Edgington, P. McCabe, E. Pidcock, L. Rodriguez-Monge, R. Taylor, J. v. Streek, and P. A. Wood, *J. Appl. Crystallogr.* **41**, 466 (2008).
- [50] C. F. Macrae, P. R. Edgington, P. McCabe, E. Pidcock, G. P. Shields, R. Taylor, M. Towler, and J. v. d. Streek, *J. Appl. Crystallogr.* **39**, 453 (2006).
- [51] I. J. Bruno, J. C. Cole, P. R. Edgington, M. Kessler, C. F. Macrae, P. McCabe, J. Pearson, and R. Taylor, *Acta Crystallogr. B* **58**, 389 (2002).
- [52] All timing results are using a PC desktop computer with an 8-core Intel i7-4790 processor.
- [53] L. Yakub, *Low Temp. Phys.* **42**, 1 (2016).
- [54] M. I. Eremets, A. G. Gavriliuk, I. A. Trojan, D. A. Dzivenko, and R. Boehler, *Nature Mater.* **3**, 558 (2004).
- [55] D. A. Young, *Phase diagrams of the elements* (Univ of California Press, 1991).
- [56] A. Schuch and R. Mills, *J. Chem. Phys.* **52**, 6000 (1970).
- [57] C. Mailhot, L. Yang, and A. McMahan, *Phys. Rev. B* **46**, 14419 (1992).
- [58] The calculated localization and delocalization indices, vary slightly (less than 0.1) depending on which atom is considered. The reason is that, even though all atoms in cg-N are equivalent, the density grid is not congruent with the crystal symmetry. This variability does not affect the analysis, and it decreases as the size of the density grid increases. Averages with the appropriate number of significant digits are reported where necessary.
- [59] X.-Q. Chen, C. L. Fu, and R. Podloucky, *Phys. Rev. B* **77**, 064103 (2008).
- [60] J. G. Brandenburg, T. Maas, and S. Grimme, *J. Chem. Phys.* **142**, 124104 (2015).
- [61] V. F. Petrenko and R. W. Whitworth, *Physics of ice* (OUP Oxford, 1999).
- [62] C. J. Pickard and R. Needs, *Nature Mater.* **7**, 775 (2008).
- [63] T. Palasyuk, I. Troyan, M. Eremets, V. Drozd, S. Medvedev, P. Zaleski-Ejgierd, E. Magos-Palasyuk, H. Wang, S. A. Bonev, D. Dudenko, *et al.*, *Nature Commun.* **5** (2014).
- [64] M. Somayazulu, P. Dera, A. F. Goncharov, S. A. Gramsch, P. Liermann, W. Yang, Z. Liu, H.-k. Mao, and R. J. Hemley, *Nature chemistry* **2**, 50 (2010).
- [65] D. Sihachakr and P. Loubeyre, *Phys. Rev. B* **74**, 064113 (2006).
- [66] Y. Meng, R. B. Von Dreele, B. H. Toby, P. Chow, M. Y. Hu, G. Shen, and H.-k. Mao, *Phys. Rev. B* **74**, 214107 (2006).
- [67] A. F. Goncharov, V. V. Struzhkin, H.-k. Mao, and R. J. Hemley, *Phys. Rev. Lett.* **83**, 1998 (1999).
- [68] F. Datchi, S. Ninet, M. Gauthier, A. Saitta, B. Canny, and F. Decremps, *Phys. Rev. B* **73**, 174111 (2006).