Localized orbital scaling correction for periodic systems

Aaron Mahler
Duke University, Department of Physics, Durham, NC 27708

Jacob Williams
Duke University, Department of Chemistry, Durham, NC 27708

Neil Qiang Su
Department of Chemistry, Key Laboratory of Advanced Energy Materials Chemistry (Ministry of Education) and Renewable Energy Conversion and Storage Center (RECAST), Nankai University, Tianjin 300071, China and Duke University, Department of Chemistry, Durham, NC 27708

Weitao Yang*
Duke University, Department of Chemistry, Durham, NC 27708 and Duke University, Department of Physics, Durham, NC 27708

(Dated: June 2, 2022)

Density functional theory offers accurate structure prediction at acceptable computational cost, but commonly used approximations suffer from delocalization error; this results in inaccurate predictions of quantities such as energy band gaps of finite and bulk systems, energy level alignments, and electron distributions at interfaces. The localized orbital scaling correction (LOSC) was developed to correct delocalization error by using orbitals localized in space and energy. These localized orbitals span both the occupied and unoccupied spaces and can have fractional occupations in order to correct both the total energy and the one-electron energy eigenvalues. We extend the LOSC method to periodic systems, in which the localized orbitals employed are dually localized Wannier functions. In light of the effect of the bulk environment on the electrostatic interaction between localized orbitals, we modify the LOSC energy correction to include a screened Coulomb kernel. For a test set of semiconductors and large-gap insulators, we show that the screened LOSC (sLOSC) method consistently improves the band gap compared to the parent density functional approximation.

I. INTRODUCTION

The cost of solving the electronic Schrödinger equation scales exponentially with the size of the system, exceeding the computational resources available on the planet for any system larger than a few tens of electrons.[1] Density functional theory (DFT) sidesteps this exponential cost by treating the electron density as the fundamental variable instead of computing the wavefunction directly and by constructing an auxiliary noninteracting reference system sharing the density of the physical system.[2,3] Due to the accuracy attainable at a cost only cubic in the number of electrons N, DFT has become a mainstay of computational chemistry and materials science.[4–7] While DFT is exact in theory, the form of the universal exchange-correlation functional is unknown, and density functional approximations (DFAs) must be used in practice. Commonly used DFAs suffer from systematic delocalization and static correlation errors.[8,9] The delocalization error underlies the failure of DFAs to describe energy band gaps of finite and bulk systems, energy level alignments, and electron distributions at interfaces.[10] Overcoming delocalization error remains an active and challenging research effort.

Connecting single-particle orbital energies ϵ to observable quantities was another longstanding question in Kohn–Sham DFT. As a contrast, Koopmans[11] showed in 1934 that the Hartree–Fock ionization potential (IP) and electron affinity (EA) are given under the frozen orbital aproximation by the negative of the highest occupied and lowest unoccupied molecular orbital eigenvalues respectively. A series of three results established a rigorous connection for DFT.

First, Janak[12] derived a link between the Kohn–Sham orbital energies ϵ_m and the total energy E, viewed as a function of the orbital occupation numbers n_m:

$$\epsilon_m = \frac{\partial E}{\partial n_m}. \tag{1}$$

However, $\partial E/\partial n_m$ was not yet linked to a physical observable.

A few years later, Perdew, Parr, Levy, and Balduz[13] showed that E is piecewise linear in the number of electrons N when computed with the exact functional; that is, for all $|\delta| \leq 1$, we have

$$E(N+\delta) = \begin{cases} (1+\delta)E(N) - \delta E(N-1) & \delta < 0, \\ (1-\delta)E(N) + \delta E(N+1) & \delta \geq 0. \end{cases} \tag{2}$$

This relationship, called the PPLB condition, connects the chemical potential $\mu(N) = \partial E/\partial N$ to the IP and EA; observe that

$$\mu(N) = \begin{cases} -I(N) = E(N) - E(N-1) & \partial N < 0, \\ -A(N) = E(N+1) - E(N) & \partial N > 0. \end{cases} \tag{3}$$

Finally, Cohen *et al.*[14] proved that the chemical potential is given by the partial derivative of E with respect

to the frontier orbital eigenvalues,

$$\mu(N) = \frac{\partial E}{\partial n_f}. \tag{4}$$

Crucially, f labels not only the highest unoccupied molecular orbital (HOMO) if $\partial N < 0$, but also the lowest unoccupied molecular orbital (LUMO) if $\partial N > 0$; this was the first time a physical meaning for the energy of the Kohn–Sham LUMO was established. This result holds for any local functional continuous in the electron density, as well as any nonlocal functional continuous in the Kohn–Sham density matrix; in the latter case, the work also extends Janak's theorem to the eigenvalues from the generalized Kohn–Sham equations.

Combining these three results, we see that

$$\mu(N) = \begin{cases} -I(N) = \epsilon_{\text{HOMO}} & \partial N < 0, \\ -A(N) = \epsilon_{\text{LUMO}} & \partial N \geq 0. \end{cases} \tag{5}$$

Thus, the frontier eigenvalues obtained from an N-electron DFT calculation correspond rigorously to physically relevant quantities;[14] if the PPLB condition is obeyed and the functional predicts the exact energies for $N-1$, N, and $N+1$ electrons, the correspondence is exact.

A feature derivable from these quantities is the fundamental or integer gap, defined as the difference between the IP and the EA:

$$\begin{aligned} E_{\text{gap}}^{\text{integer}} &= I - A \\ &= E(N-1) - 2E(N) + E(N+1). \end{aligned} \tag{6}$$

$E_{\text{gap}}^{\text{integer}}$ quantifies the difference between positively and negatively ionizing the system and is a crucial part of the accurate modeling of semiconductor electronic structure. If the PPLB condition is obeyed, Eqs. (3) and (5) also allow the gap to be computed from a single N-electron calculation as the discontinuity in the chemical potential; in this form, it is called the derivative gap, defined as

$$\begin{aligned} E_{\text{gap}}^{\text{deriv}} &= \left.\frac{\partial E}{\partial N}\right|_{+} - \left.\frac{\partial E}{\partial N}\right|_{-} \\ &= \epsilon_{\text{LUMO}} - \epsilon_{\text{HOMO}}. \end{aligned} \tag{7}$$

If the PPLB condition is obeyed, the derivative gap and the integer gap are equal.[14] In bulk systems with periodic boundary conditions, the PPLB condition is satisfied by any DFA continuous in the Kohn–Sham density or density matrix, regardless of systematic errors in its definition;[10] thus, the fundamental gap of bulk systems can be predicted by the (generalized) Kohn–Sham orbital gap, for functionals continuous in the density (density matrix), as in Eq. (7). In finite systems, however, the PPLB condition is not in general obeyed, and the gap computed from Eq. (7) may differ from that computed by calculating the $(N \pm 1)$-electron energies to obtain the integer gap as in Eq. (6), the ΔSCF method.

A. Delocalization error

The delocalization error has a dramatic size-dependent manifestation. In finite systems, standard DFAs fail to obey the PPLB linearity condition, so the derivative gap is not equal to the integer gap. This is due to the error in the approximate exchange-correlation functional, which nearly always yields E convex in N, underestimating the piecewise linearity prescribed by the PPLB condition. This convex deviation has been identified as the cause for an unphysical smearing of the electron density in space, as well as underestimation of the total energy in a delocalized electron density; thus, we may identify it with delocalization error, as exhibited in small systems. In bulk systems, the delocalized nature of the orbitals produces a total energy linear with respect to fractional charge, yielding no deviation from the PPLB condition; however, delocalization error manifests as an incorrect slope of the $E(N)$ line at integer N.[10]

The effects of delocalization error include the underestimation of band gaps and reaction barriers,[15] undervaluation of dissociation curves,[16–18] overestimation of conductance and polarizability,[19] and incorrect energy level alignment and charge transfer across interfaces.[20,21] To capture the full derivative discontinuity and hence the band gap, it has been shown that the exact functional, whether local or nonlocal, cannot be a differentiable functional of the electron density or of the Kohn-Sham density matrix.[22,23] To reduce the systematic delocalization error, many approaches have been developed, including range-separated functionals,[24–30] the screened range-separated hybrid functional,[31] self-interaction error corrected functionals,[17,32–38] Koopmans-compliant functionals,[39,40] and generalized transition state methods,[41] along with related developments using localized Wannier functions.[42]

The localized orbital scaling correction (LOSC) method was developed to eliminate delocalization error systematically.[43,44] Previous incarnations of LOSC were implemented for molecular systems with real orbitals and the boundary condition $\lim_{|\mathbf{r}|\to\infty} \rho(\mathbf{r}) = 0$. They accurately model IP, EA, photoemission spectra, dissociation curves, and polarizabilities, as well as restore size-consistency.[43–46] In this work, we extend LOSC to periodic boundary conditions (PBCs) and complex orbitals. Additionally, we introduce a screened Coulomb interaction to the LOSC energy correction to enable the accurate computation of bulk system band structures.

B. Periodic boundary conditions

In PBCs, the eigenfunctions of the single-particle Hamiltonian are known as Bloch orbitals; they satisfy $h_s|\psi_n^{\mathbf{k}}\rangle = \epsilon_n^{\mathbf{k}}|\psi_n^{\mathbf{k}}\rangle$. The Bloch orbitals are also eigenfunctions of the unit cell translation operator, so they take the form $|\psi_n^{\mathbf{k}}\rangle = e^{i\mathbf{k}\cdot\mathbf{r}}|u_n^{\mathbf{k}}\rangle$, where $|u_n^{\mathbf{k}}\rangle$ has the periodicity of the unit cell and $\mathbf{k}$ is a point in the Brillouin zone (the reciprocal-space unit cell).[47,48] The Bloch orbitals obey

the normalization convention $\langle\psi_m^{\mathbf{q}}|\psi_n^{\mathbf{k}}\rangle = \delta(\mathbf{k}-\mathbf{q})\delta_{mn}$, where $\langle f|g\rangle = \int_{\mathcal{D}} d\mathbf{r}\, \overline{f}(\mathbf{r})g(\mathbf{r})$. Here, $\delta(\mathbf{k})$ is the Dirac delta distribution, δ_{mn} is the Kronecker delta, and $\overline{f}$ is the complex conjugate of f. The domain of integration $\mathcal{D}$ is the periodic unit for the functions being integrated; for Bloch orbitals, $\mathcal{D} = \mathbb{R}^3$. The $|u_n^{\mathbf{k}}\rangle$ are orthonormal in the band index n at a fixed $\mathbf{k}$-point in reciprocal space: that is, $\langle u_m^{\mathbf{k}}|u_n^{\mathbf{k}}\rangle = \delta_{mn}$, where the inner product integrates over one unit cell. Note that we assume closed-shell systems in this work.

The single-particle density can be represented in real space by the occupied Bloch orbitals as

$$\rho_s(\mathbf{r}) = \sum_n^{\text{occ}} \frac{V}{(2\pi^3)} \int_{\text{BZ}} d\mathbf{k}\, |\psi_n^{\mathbf{k}}(\mathbf{r})|^2, \tag{8}$$

where V is the volume of the unit cell and the integral is over the first Brillouin zone. Since the Hamiltonian is diagonal in $\mathbf{k}$, we can solve for the Bloch orbitals in reciprocal space, requiring diagonalization only in one unit cell. In practice, the Brillouin zone is sampled with a finite number of points; in this work we use a Monkhorst–Pack mesh centered at the origin Γ of the Brillouin zone.[49] Thus, integrals over the Brillouin zone become equally weighted sums over the $\mathbf{k}$-mesh

$$\frac{V}{(2\pi)^3} \int d\mathbf{k}\, f(\mathbf{k}) \mapsto \frac{1}{N_k} \sum_{\mathbf{k}} f(\mathbf{k}), \tag{9}$$

where N_k is the number of $\mathbf{k}$-points in the mesh. Using a Monkhorst–Pack mesh centered at Γ yields Bloch orbitals having the periodicity of an unfolded supercell comprised of N_k primitive unit cells; this supercell is referred to as the Born–von Karman cell.[48] The Bloch orbitals then obey the normalization convention $\langle\psi_m^{\mathbf{q}}|\psi_n^{\mathbf{k}}\rangle = N_k\delta_{\mathbf{kq}}\delta_{mn}$, where the integral is over the Born–von Karman cell.

II. METHODS

The LOSC method consists of two steps. First, we find orbitals that are spatially localized while remaining associated with specific energy ranges. Next, we compute a curvature matrix modeling the magnitude of the deviation from linearity. This is combined with the fractionally occupied localized orbitals to correct the convex deviation of $E(N)$ from linearity at non-integer N, as well as incorrect total energies at integer N. Both steps are implemented as post-processing after a converged self-consistent field calculation.

A. Localization

The wave-like nature of the Bloch orbitals prohibits them from being spatially localized. In order to obtain a state that is localized in space, the discrete Fourier transform of the Bloch orbitals is used to produce Wannier functions[50]

$$|w_n^{\mathbf{R}}\rangle = \frac{1}{N_k} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}} |\psi_n^{\mathbf{k}}\rangle. \tag{10}$$

Wannier functions inherit the periodicity of the Born–von Karman cell, and are indexed by electron bands n and unit cells $\mathbf{R}$ in the supercell. They are symmetric under translation by unit cell vectors, so that $w_n^{\mathbf{R}}(\mathbf{r}) = w_n^{\mathbf{0}}(\mathbf{r}-\mathbf{R})$; $\mathbf{R} = \mathbf{0}$ is referred to as the home unit cell. There is a unitary, or gauge, freedom at each $\mathbf{k}$-point in the choice of Bloch orbitals that comprise a Wannier function, so we can define *generalized* Wannier functions[51]

$$\begin{aligned} |w_i^{\mathbf{R}}\rangle &= \frac{1}{N_k} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}} \sum_n U_{ni}^{\mathbf{k}} |\psi_n^{\mathbf{k}}\rangle \\ &= \frac{1}{N_k} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}} |\phi_i^{\mathbf{k}}\rangle, \end{aligned} \tag{11}$$

where we refer to $|\phi_i^{\mathbf{k}}\rangle$ as a transformed Bloch orbital (TBO). From now on, we will refer to generalized Wannier functions as Wannier functions.

The gauge freedom $U^{\mathbf{k}}$ in the TBOs can be chosen such that the resulting set of Wannier functions have advantageous properties. In order to obtain localization in space, Marzari and Vanderbilt suggested choosing $U^{\mathbf{k}}$ that minimize the Wannier functions' spatial variance $\langle\Delta r^2\rangle_i = \langle r^2\rangle_i - \langle\mathbf{r}\rangle_i^2$, where $\langle x\rangle_i = \langle w_i^{\mathbf{0}}|x|w_i^{\mathbf{0}}\rangle$; the resulting orbitals are called maximally localized Wannier functions.[52] In molecules, the scheme of minimizing spatial variance is referred to as Foster–Boys localization.[53] However, constructing maximally localized Wannier functions from both valence and conduction bands is physically ill-motivated. Because bands far apart in energy can mix freely and the Bloch bands form a complete basis, adding more virtual bands will result in increasing spatial localization of the maximally localized Wannier functions, with a corresponding loss of information about the energy dispersion of the bands.

In order to preserve locality in energy while maintaining spatial localization, enabling simultaneous treatment of the occupied and unoccupied spaces, we choose the Wannier gauge that minimizes a cost function considering both energy and spatial variance:

$$F = (1-\gamma)\sum_i \langle\Delta r^2\rangle_i + \gamma C \sum_i \langle\Delta h_{\text{s}}^2\rangle_i, \tag{12}$$

where $0 \le \gamma \le 1$ and in the units used here $C = 1\,\text{a}_0^2/\text{eV}^2$, where a_0 is the Bohr radius. This cost function was first proposed by Gygi *et al.*[54] for computations sampling the Brillouin zone only at Γ and implemented for such systems by Giustino and Pasquarello[55]. It was used for LOSC in molecules[44] to treat system symmetries and degeneracies more robustly than the original localization, which used soft energy windows;[43] in molecular LOSC, the localized orbitals are called *orbitalets*. We recently extended F to systems with $N_k \ge 1$;[56] we refer to such orbitals as dually localized Wannier functions (DLWFs).

These formulations show how the combination of occupied and unoccupied spaces can be localized simultaneously to produce Wannier functions that are localized in both space and energy. This construction is critical for addressing delocalization error in finite systems because it allows for dynamic localization in the resulting orbitals; the orbitals can qualitatively and quantitatively differ depending on the geometry of the system.[43,44] In keeping with the principle of universality in functional development, we use the same mixing parameter in Eq. (12), setting $\gamma = 0.47714$. The value of γ has important implications for the LOSC method; see Section VI of the Supplementary Information of Su *et al.*[44]. Setting $\gamma = 0$, for instance, yields maximally localized Wannier functions,[52] while $\gamma = 1$ yields DLWFs that are pure Fourier transforms (up to $\mathbf{k}$-dependent phases) of the Kohn–Sham bands.

Note that we have not proven that a unique global minimum of F exists; in practice, F has a fairly rugged landscape of solutions, and we have observed multiple local minima. We choose the DLWFs yielding the smallest total cost. Different DLWFs can produce somewhat different sLOSC corrections, with eigenvalues varying by up to a few tenths of an eV. The problem of multiple minima of F has also been observed in molecular LOSC,[57] but was not found to be the dominant source of error. An additional question worth exploring is the effect of symmetry breaking, such as that due to perturbations of the crystal lattice, on the localization procedure.

The compromise between spatial and energy localization and the inclusion of unoccupied orbitals are key to producing localized orbitals that can address delocalization error while retaining size-consistency. For example, the DFA HOMO and LUMO of H_2^+ at (or near) the dissociation limit are delocalized over the whole molecule; since they are (nearly) degenerate, there exists a unitary freedom in the subspace spanned by both. Due to the symmetry of the system, we expect to obtain two separate $H^{0.5+}$ fragments; the (small or) vanishing gap means that any choice of $\gamma < 1$ in Eq. (12) will result in half-occupied orbitals localized on each H atom. This is the physical motivation for a localization scheme that minimizes the spatial variance of occupied and unoccupied orbitals while allowing only orbitals that are close in energy to mix.[43]

B. Energy corrections

The deviation from energy linearity with respect to fractional charges is characteristically quadratic in most exchange-correlation functionals.[43,58,59] To restore compliance with the PPLB condition for small finite systems, the global scaling correction (GSC) was developed. GSC corrects the total energy by an amount quadratic in the occupation numbers of the canonical molecular orbitals.[58,60] This method is effective at correcting the systematic deviation from the PPLB condition for systems with fractional charges and leads to accurate prediction of quasiparticle energies as the eigenvalues from the resulting one-electron Hamiltonian. However, GSC is applicable only for systems of small and moderate size; the convex deviation of conventional DFAs from the piecewise linearity prescribed by the PPLB condition decreases with increasing system size, and the delocalization error manifests instead as underestimated ground-state energies for integer systems and incorrect linear $E_{\mathrm{gs}}(N)$ curves with wrong slopes at the bulk limit.[10] The localized orbital scaling correction (LOSC) applies its energy correction adaptively by the construction of localized orbitals, allowing systematic and size-consistent correction of delocalization error.[43,44] In this section, we discuss the extension of LOSC to periodic systems.

A basic quantity in LOSC is the density matrix in the basis of DLWFs; its elements are occupations

$$\lambda_{ij}^{\mathbf{TR}} = \langle w_i^{\mathbf{T}} | \rho_s | w_j^{\mathbf{R}} \rangle. \tag{13}$$

The occupations between all pairs of DLWFs are used to remove quadratic deviations, while the diagonal terms are used to restore linearity. The energy correction defined by LOSC for each unit cell is given by

$$\Delta E^{\mathrm{LOSC}} = \frac{1}{2N_k} \sum_{\mathbf{TR}} \sum_{ij} \widetilde{\kappa}_{ij}^{\mathbf{TR}} \lambda_{ij}^{\mathbf{TR}} (\delta_{ij}^{\mathbf{TR}} - \overline{\lambda}_{ij}^{\mathbf{TR}}), \tag{14}$$

where $\delta_{ij}^{\mathbf{0R}} = \delta_{ij}\delta_{\mathbf{0R}}$; $\widetilde{\kappa}$ models the curvature of the deviation from linearity.

The diagonal terms in the energy correction are proportional to $\lambda_{ii}^{\mathbf{TR}} - |\lambda_{ii}^{\mathbf{TR}}|^2$; thus, if a DLWF has integer occupancy (implying $\lambda_{ij}^{\mathbf{TR}} = 0$ whenever $i \neq j$ or $\mathbf{T} \neq \mathbf{R}$), then the energy correction due to that DLWF will also be zero.

The matrix $[\lambda_{ij}^{\mathbf{TR}}]$ of occupations between the DLWFs is the discrete Fourier transform of the occupation matrix between the TBOs. As such, it is positive semidefinite and Hermitian, and

$$\mathrm{tr}_c[\lambda_{ij}^{\mathbf{TR}}] = \frac{1}{N_k} \sum_{\mathbf{k}} N_f^{\mathbf{k}}, \tag{15}$$

where tr_c denotes the trace per unit cell and $N_f^{\mathbf{k}}$ is the number of electrons below the Fermi energy at $\mathbf{k}$.

Following Su *et al.*[44], the elements of the curvature matrix are given by

$$\begin{aligned} \widetilde{\kappa}_{ij}^{\mathbf{TR}} = & \,\mathrm{erf}(8S_{ij}^{\mathbf{TR}}) \sqrt{\kappa_{ii}^{\mathbf{TT}} \kappa_{jj}^{\mathbf{RR}}} \\ & + \mathrm{erfc}(8S_{ij}^{\mathbf{TR}}) \kappa_{ij}^{\mathbf{TR}}. \end{aligned} \tag{16}$$

Here, $\mathrm{erfc}(r) = 1 - \mathrm{erf}(r)$ is the complementary error function. $S_{ij}^{\mathbf{TR}}$ is the absolute overlap between DLWFs,

$$S_{ij}^{\mathbf{TR}} = \int d\mathbf{r} \sqrt{\rho_i^{\mathbf{T}}(\mathbf{r}) \rho_j^{\mathbf{R}}(\mathbf{r})}, \tag{17}$$

where $\rho_i^{\mathbf{T}}(\mathbf{r}) = |w_i^{\mathbf{T}}(\mathbf{r})|^2$ is a DLWF's charge density. The matrix elements $\kappa_{ij}^{\mathbf{TR}}$ in Eq. (16) are given by

$$\kappa_{ij}^{\mathbf{TR}} = J[\rho_i^{\mathbf{T}}, \rho_j^{\mathbf{R}}] - X[\rho_i^{\mathbf{T}}, \rho_j^{\mathbf{R}}], \tag{18}$$

with

$$J[\rho_i^{\mathbf{T}}, \rho_j^{\mathbf{R}}] = \iint d\mathbf{r}\, d\mathbf{r}'\, \rho_i^{\mathbf{T}}(\mathbf{r}) \rho_j^{\mathbf{R}}(\mathbf{r}') K(|\mathbf{r} - \mathbf{r}'|), \quad (19a)$$

$$X[\rho_i^{\mathbf{T}}, \rho_j^{\mathbf{R}}] = \tau \frac{2C_{\mathrm{X}}}{3} \int d\mathbf{r} \left[\rho_i^{\mathbf{T}}(\mathbf{r}) \rho_j^{\mathbf{R}}(\mathbf{r})\right]^{2/3}. \quad (19b)$$

In the above, $K(r) = 1/r$ is the Coulomb kernel, $C_X = \frac{3}{4}(\frac{6}{\pi})^{1/3}$ is the Dirac exchange constant,[61] and $\tau = 6(1 - 2^{-1/3}) \approx 1.2378$ is a nonempirical parameter.[43] The derivation of how this correction restores the PPLB condition can be found in the supplementary data of Li *et al.*[43]. The use of $\widetilde{\kappa}$ instead of κ was introduced because the cost function in Eq. (12) can induce discontinuous jumps between localization characters during molecular dissociation.[44] The diagonal elements of $\widetilde{\kappa}$ and κ are equal, so when $\lambda_{ii} \in \{0, 1\}$ the corrections from $\widetilde{\kappa}$ and κ are the same. In practice, $X[\rho_i^{\mathbf{T}}, \rho_j^{\mathbf{R}}]$ term is evaluated using numerical integration on a grid of real-space points. The Coulomb term $J[\rho_i^{\mathbf{T}}, \rho_j^{\mathbf{R}}]$ is evaluated in a plane wave basis, detailed in Sec. II C.

Applying the extension of Janak's theorem[12] to the generalized Kohn-Sham theory,[14] the LOSC energy correction of Eq. (14) yields corrections to the Bloch orbital energy eigenvalues $\epsilon_n^{\mathbf{k}}$ given by

$$\Delta\epsilon_n^{\mathbf{k}} = \sum_i \widetilde{\kappa}_{ii}^{\mathbf{00}} \left(\frac{1}{2} - \lambda_{ii}^{\mathbf{00}}\right) |U_{ni}^{\mathbf{k}}|^2 - \sum_{\mathbf{R}i \neq \mathbf{0}j} \widetilde{\kappa}_{ij}^{\mathbf{0R}} \mathrm{Re}\left\{\lambda_{ij}^{\mathbf{0R}} e^{i\mathbf{k}\cdot\mathbf{R}} \overline{U}_{ni}^{\mathbf{k}} U_{nj}^{\mathbf{k}}\right\}. \quad (20)$$

Consider the diagonal corrections given by the first summand in Eq. (20). There is no correction to the eigenvalue when a DLWF is half-occupied ($\lambda_{ii} = \frac{1}{2}$); on the other hand, the correction is maximal when it is completely occupied or unoccupied. Li *et al.*[43] observed that the slopes of the quadratic DFA and the correct linear $E(N)$ curves agree at half-integer N. Since the frontier orbital energy corresponds to this slope, we see that accurate frontier orbital energies are given by half-occupied frontier orbitals. The LOSC correction to the orbital energies arrives naturally at this conclusion, additionally agreeing with Slater transition state theory.[62,63]

We may also view LOSC as a correction to the Kohn–Sham Hamiltonian. It is given by the functional derivative of the energy correction with respect to the density operator under the frozen orbital approximation,

$$\Delta v = \left.\frac{\delta \Delta E^{\mathrm{LOSC}}}{\delta \rho_s}\right|_{\{w_i^{\mathbf{R}}\}}, \quad (21)$$

and can be written in operator form as

$$\Delta v = \frac{1}{2} \sum_{ij,\mathbf{TR}} \widetilde{\kappa}_{ij}^{\mathbf{TR}} \left(\frac{\delta_{ij}^{\mathbf{TR}}}{2} - \lambda_{ij}^{\mathbf{TR}}\right) |w_i^{\mathbf{T}}\rangle\langle w_j^{\mathbf{R}}| + \mathrm{h.c.} \quad (22)$$

(See the Supplemental Material for details on this derivation.) The correction to the n^{th} Bloch orbital eigenvalue is then given by $\Delta\epsilon_n^{\mathbf{k}} = \langle\psi_n^{\mathbf{k}}|\Delta v|\psi_n^{\mathbf{k}}\rangle$.

In practice, the energy corrections are applied to disentangled Bloch orbitals. The conduction bands of most systems cannot be formed into sets of bands that do not cross anywhere in the Brillouin zone, a condition referred to as band entanglement. In order to obtain a finite set of bands for localization and energy correction, we use the disentanglement procedure outlined by Souza *et al.*[64] This procedure obtains N_w bands from a set of $N_b \geq N_w$ Bloch orbitals at each $\mathbf{k}$-point, chosen such that the subspace spanned by the disentangled bands is as smooth as possible in $\mathbf{k}$. To correct the band gap of semiconductors and insulators, we include sufficiently many virtual bands in the construction of the Wannier functions to converge the localization of the frontier bands (that is, the valence band maximum and conduction band minimum).[56] We find that $N_b = N_{\mathrm{occ}} + 3N_{\mathrm{coord}}$ and $N_w = N_{\mathrm{occ}} + 2N_{\mathrm{coord}}$, where N_{occ} is the number of occupied bands per unit cell and N_{coord} is the coordination number of the lattice, are sufficient; see the Supplemental Material for details. The N_w disentangled Bloch bands yields N_w DLWFs per unit cell, so there are $N_w N_k$ DLWFs in the Born–von Karman supercell on which they are periodic. The energy corrections for the N_w disentangled Bloch orbitals at each $\mathbf{k}$-point are implemented using Eq. (20).

In this work, we restrict our attentions to closed-shell systems. Extending the LOSC method to spin-polarized materials is accomplished by finding the corrections from the spin-up and spin-down DLWFs independently and summing them to obtain ΔE^{LOSC}; this functionality is planned for the next version of LOSC. However, treating the strong correlation common to open-shell materials brings its own set of challenges beyond the scope of this work. We discuss them briefly in Sec. IV below.

C. Coulomb integrals

Accurate calculation of the Coulomb interaction $J[\rho_i^{\mathbf{T}}, \rho_j^{\mathbf{R}}]$ is needed for LOSC to restore the PPLB condition. In PBCs, a plane-wave basis is typically employed. The Coulomb energy is diagonal in this basis, and the double integral required in real space collapses to a single sum over basis vectors $\mathbf{G}$:

$$J[\rho_i^{\mathbf{T}}, \rho_j^{\mathbf{R}}] = \sum_{\mathbf{G}} \frac{4\pi}{G^2} \overline{\rho}_i^{\mathbf{T}}(\mathbf{G}) \rho_j^{\mathbf{R}}(\mathbf{G}), \quad (23)$$

where $G = |\mathbf{G}|$. However, this sum converges only for neutral charge distributions, for which the $\mathbf{G} = \mathbf{0}$ term vanishes. The DLWF densities are individually charged, so ignoring the divergent term coming from the net charge will significantly underestimate the Coulomb energy. There are many methods to evaluate the Coulomb energy for charged densities in the plane-wave basis accurately, including those of Makov and Payne[65], Kantorovich[66], Dabo *et al.*[67], and Li and Dabo[68]. We choose the spherical cutoff method,[69–71] truncating the Coulomb kernel in Eq. (19a) at a cutoff radius R_c; this is taken to be half the length of the shortest Born–von Karman supercell lattice vector, ensuring that the Coulomb

interactions between the a DLWF density and its images in neighboring supercells are zero. Thus, the spherical cutoff Coulomb kernel is

$$K_c(r;R_c) = \begin{cases} 1/r & r < R_c \\ 0 & r \geq R_c, \end{cases} \tag{24}$$

which has Fourier coefficients

$$K_c(G;R_c) = \begin{cases} \frac{4\pi}{G^2}[1-\cos(GR_c)] & G \neq 0 \\ 2\pi R_c^2 & G = 0. \end{cases} \tag{25}$$

Observe that $K_c(G;R_c)$ does not diverge for any $\mathbf{G}$. As long as the pair of DLWF densities in Eq. (19a) lie in a sphere of radius R_c, the spherical cutoff method is also accurate in highly anisotropic unit cells, unlike schemes such as that of Makov and Payne.[66] We enforce this containment condition in practice by checking that each DLWF density is well contained in a volume spanned by half of each Born–von Karman supercell lattice vector, and only compute curvature elements between pairs of DLWF densities that have centers closer together than R_c. We evaluate the Coulomb integrals on the unfolded supercell in the plane-wave basis, which requires a fast Fourier transform (FFT) of the DLWF densities on the supercell.

D. Screening

Applying LOSC with a bare Coulomb interaction leads to severe overcorrection of semiconductors' band gaps in PBCs. However, this is not surprising; we anticipate an effect of the other electrons in the lattice on $J[\rho_i^{\mathbf{T}}, \rho_j^{\mathbf{R}}]$. As shown by highly accurate methods such as GW, a screened Coulomb interaction is required to model the interaction between electrons in a periodic system accurately.[72,73] Recently, Mei and coworkers also found that the deviation from linearity of the total energy as a function of canonical orbital occupations is given to second order by a screened interaction.[60] We model the screening phenomenologically, attenuating the long-range $1/r$ behavior of the spherical cutoff Coulomb interaction by a complementary error function. This modifies the Coulomb kernel to read

$$K_s(r;R_c,\alpha) = \begin{cases} \operatorname{erfc}(\alpha r)/r & r < R_c \\ 0 & r \geq R_c, \end{cases} \tag{26}$$

where α is a screening parameter. We choose the α that best reproduces the experimental band gaps of a test set of semiconductors and insulators. For r larger than the screening radius α^{-1}, $K_s(r;R_c,\alpha)$ decays exponentially instead of as $1/r$. The Fourier coefficients of K_s are

$$K_s(G;R_c,\alpha) = \begin{cases} \frac{4\pi}{G^2}\left[1-\cos(GR_c)\operatorname{erfc}(\alpha R_c) - e^{-(G/2\alpha)^2}\operatorname{Re}\left\{\operatorname{erf}\left(\alpha R_c + \frac{iG}{2\alpha}\right)\right\}\right] & G \neq 0 \\ 2\pi R_c^2 + \pi \operatorname{erf}(\alpha R_c)\left(\alpha^{-2} - 2R_c^2\right) - 2\sqrt{\pi}e^{-(\alpha R_c)^2}R_c/\alpha & G = 0. \end{cases} \tag{27}$$

The error function is unbounded for complex arguments, overflowing double-precision floating-point numbers even for relatively small G. Thus, we evaluate $K_s(G;R_c,\alpha)$ with a scaled form of erf z called the Faddeeva function, implemented in the numerically stable ACM Algorithm 916.[74,75] For details, see the Supplemental Material.

In principle, the screening is system-dependent. Improved accuracy would be attainable by setting its value to best reproduce each material's band gap. However, the phenomenological screening model of sLOSC does not enable doing so while retaining predictive ability. This would require a rigorously screened Coulomb (or Hartree-exchange-correlation) interaction based on the linear response function $\chi(\mathbf{r},\mathbf{r}') = \delta\rho_s(\mathbf{r})/\delta v(\mathbf{r}')$. *Ab initio* screening of this kind appears in the extensions of the GSC method to hybrid functionals[76] and in the following exploration of orbital relaxation on GSC,[77], the GSC2 method,[60] as well as in recent work on Koopmans-compliant functionals.[78–80] For small, finite systems, the delocalization error is quantified by $\partial^2 E/\partial n_i^2$, where n_i is the occupation number of the Kohn–Sham orbital $|\psi_i\rangle$; analytical expressions for $\partial^2 E/\partial n_i^2$ were derived in Yang *et al.*[81] In sLOSC, linear-response screening would very likely increase the accuracy, but at substantial computational cost to compute the nonlocal $\chi(\mathbf{r},\mathbf{r}')$.

III. RESULTS

We use the PBE functional[6] for the parent DFA calculations, with optimized norm-conserving Vanderbilt pseudopotentials[82] generated by PseudoDojo.[83] Both self-consistent field (SCF) and non-SCF calculations are carried out in the `Quantum ESPRESSO` code suite.[84,85]

The energy cutoff for the fast Fourier transform is set to 100 Ry for wavefunctions and 400 Ry for densities. The Brillouin zone is sampled with Monkhorst–Pack meshes centered at Γ, which is necessary for the Wannier functions to be periodic on the Born–von Karman supercell. For SCF calculations, we use a 16×16×16 **k**-mesh, while the other calculations are performed on 6×6×6 grids. The localization step of LOSC is implemented in a modified fork of the `wannier90` code,[86,87] and the energy correction as module to a fork of `Quantum ESPRESSO`.

To determine an optimal screening parameter α, we minimize the mean absolute percent error (MAPE) on

the SC/40 set of semiconductors with experimentally available band gaps[88] together with six additional large-gap insulators. The experimental band gaps studied range from 0.23 eV to 21.7 eV. We find that $\alpha = 0.15\,a_0^{-1}$ achieves the lowest MAPE; coincidentally, this value is numerically equal to the screening parameter used in the HSE density functional.[88] As shown in Fig. 1, LOSC with Coulomb screening (sLOSC) yields marked improvement of the band gap for the test set in comparison with the parent functional. It is also apparent that unscreened LOSC overcorrects the band gaps; indeed, it is less accurate than the parent functional. The performance of sLOSC in molecules is better than the parent functional, but unscreened LOSC achieves the best performance in molecular systems (see Table I). The Supplemental Material details the variation in performance of screened LOSC for both bulk systems and molecules with the creening parameter α.

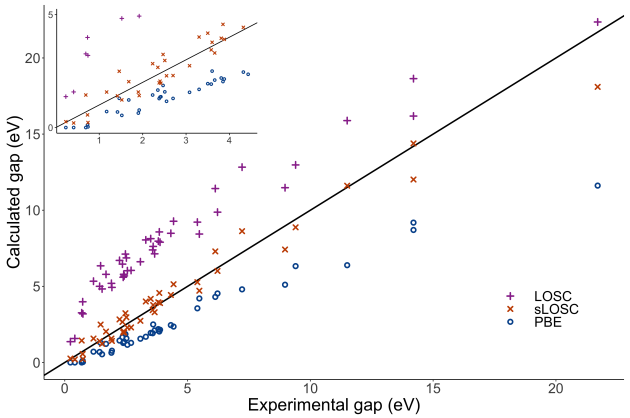


Figure 1. Comparison of experimental band gaps with those calculated by PBE (o), sLOSC (×), and unscreened LOSC (+). The inset shows systems with an experimental band gap less than 5 eV.

Table I. Mean absolute percent error of the band gap for PBC and molecular test sets. For details on the systems tested, see the Supplemental Material.

Method	PBE	LOSC	sLOSC
PBC	47.5%	158.6%	19.7%
Molecule	79.8%	10.1%	43.6%

The band structures of sLOSC and of the parent functional are shown for the small-gapped semiconductor silicon in Fig. 2 and the larger-gapped insulator lithium fluoride in Fig. 3. They use the disentangled band structures, which are numerically indistinguishable from the true band structure at and below the conduction band minimum for the parent functional. Wannier interpolation in the same DLWF basis as that used in sLOSC is used to find the energy at the points in the Brillouin zone not explicitly treated by the localization and energy correction. The sLOSC correction to the band structure comes largely from the more localized DLWFs, for which there is a larger Coulomb self-energy $J[\rho_i^{\mathbf{R}}, \rho_i^{\mathbf{R}}]$. Because of this, sLOSC mostly corrects the energy of the occupied bands (which we observe to correspond closely to the occupied DLWFs in semiconductors); it affects the virtual bands much less. More work on molecule-surface and surface-surface interactions is required to determine whether LOSC yields correct energy level alignment and whether the larger correction to the valence bands is physically meaningful.

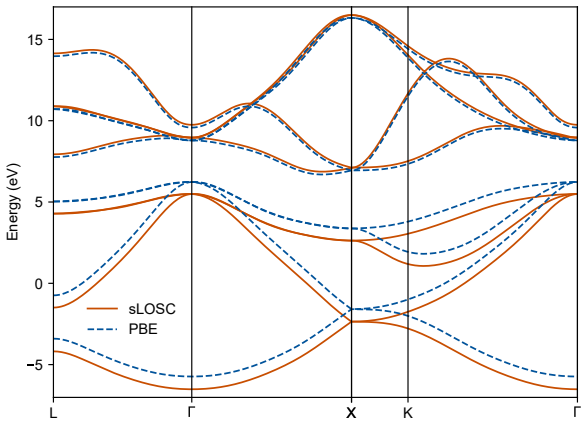


Figure 2. Band structure of silicon under the PBE functional (dashes) and sLOSC (solid). The Fermi energy of the PBE calculation was 6.23 eV, while the PBE with LOSC Fermi energy was 5.49 eV.

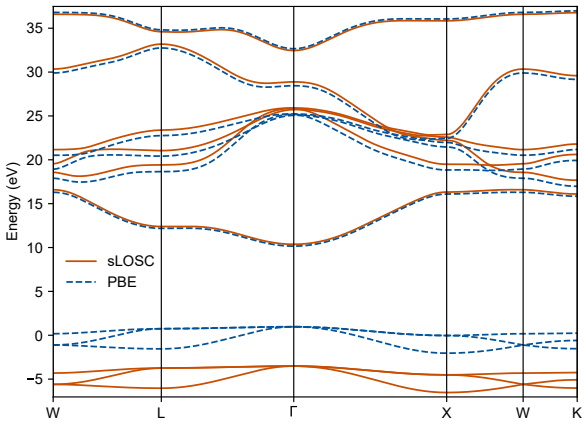


Figure 3. Band structure of lithium fluoride under the PBE functional (dashes) and sLOSC (solid). The Fermi energy of the PBE calculation was 0.97 eV, while the PBE with LOSC Fermi energy was −3.52 eV. The core states are not included in the figure.

IV. DISCUSSION

We have shown that despite the simple form of its phenomenological Coulomb screening, sLOSC systematically

corrects the band gap error associated with the parent functional for materials spanning a large range of band gaps. Screening improves the correction of delocalization error in bulk systems, but degrades the accuracy of molecular systems' band gaps relative to unscreened LOSC; however, sLOSC still offers better band gaps than those computed by the parent functional. One key remaining challenge is to model the curvature more accurately for all systems; we expect that linear response of the electron density, used by Mei *et al.*[60] for accurate screening of the Kohn–Sham orbitals, could be used to find the exact expression for $\partial^2 E/\partial\lambda_{ij}$. This would alleviate the error imposed by modeling κ as a difference between Coulomb repulsion and Dirac exchange.

We implement the energy correction as a post-processing step to a self-consistent calculation; such corrections are accurate when the change in electron density is small and hence the total energy correction is small. For every system considered in this work, ΔE^{LOSC} does not exceed three parts in 10^5. The corresponding change to the density for such systems is also expected to be minimal. LOSC can also be implemented self-consistently;[45] this can correct the delocalization error of the total density, improving the accuracy of LOSC for systems with large total energy corrections. A self-consistent implementation of sLOSC could be necessary for the accurate computation of heterogeneous and interfacial systems. Since delocalization error leads to incorrect charge distributions, the sLOSC energy correction is likely to be larger, and self-consistently correcting the delocalized density is expected to yield better orbital energies.

Work is ongoing to implement sLOSC for spin-polarized materials and to investigate its treatment of metals. The DLWFs of gapless systems constructed from Bloch bands near the Fermi energy are expected to have occupations λ_{ii} close to $\frac{1}{2}$, which means that the sLOSC correction to those eigenvalues will be small. While there may be changes to the overall band structure, it is likely that such systems will remain gapless.

This may not hold in semimetals, whose valence and conduction bands cross only in a small volume (or a single point) of the Brillouin zone. Metals, on the other hand, have one band that crosses the Fermi energy. sLOSC can open a gap in systems the DFA predicts to be semimetals; this occurs with the smallest-gapped system in our test set, InSb. Thus, it is not certain that true semimetals would remain so after the sLOSC correction. In addition, the treatment of strong correlation due to (near) degeneracy of spin states and the inclusion of topological or spin-orbit effects are beyond the scope of the current work. A modification of molecular LOSC to include fractional spins was developed in Su *et al.*[89]; it could possibly be extended to bulk materials as well.

A. Comparison with other methods

1. DFT+U(+V)

(s)LOSC is related to the DFT+U[90,91] method for correcting delocalization error. The kinship can be seen in the similarity of the sLOSC energy correction, Eq. (14), to the rotationally invariant DFT+U correction[92]

$$\Delta E_{\mathrm{DFT+U}} = \frac{1}{2}\sum_{\ell,\sigma} \mathrm{tr}\left[\boldsymbol{n}_{\ell\sigma}(1-\boldsymbol{n}_{\ell\sigma})\right] U_\ell, \tag{28}$$

where U_ℓ is the effective Hubbard parameter for orthonormal local orbital (LO) ℓ, combining atom and orbital indices, and $\boldsymbol{n}_{\ell\sigma}$ is the LO occupation matrix. Both offer an adjustment to the total energy quadratic in the occupation of the LOs. The energy correction of DFT+U comes only from interactions between LOs on the same atom, although DFT+U+V[93] extends this to interactions between atoms, analogous to the off-diagonal curvature elements $\widetilde{\kappa}_{ij}$ of (s)LOSC. However, the LOs of DFT+U(+V) are static (usually being d and f orbitals on transition metal centers), while the DLWFs of sLOSC dynamically localize based on the gauge set by the cost function F. Thus, where DFT+U(+V) must recompute the effective Hubbard parameter for every perturbation of the crystal structure or molecular geometry, the size of the correction in (s)LOSC follows from the DLWFs.

It is worth noting that, while DFT+U(+V) does not explicitly include energy localization in its construction, the Hubbard correction applies primarily to the Kohn–Sham orbitals that have the most overlap with the LOs; viewed another way, the (spatially) localized orbitals that have the most energy-local character (via their large overlap with energy eigenstates).[94] In particular, the d and f atomic orbitals of transition metals correspond closely to flat bands in reciprocal space, which carry some energy information implicitly. However, they are independent of the of the system's geometry. In contrast, the LOs of LOSC are dynamic: the orbitals can change with the geometric structure of the system. This is key to their utility in finite systems. In compact structures near equilibrium, the LOSC LOs can replicate the Kohn–Sham canonical orbitals, while becoming localized as chemical bonds are stretched. This allows the LOSC total energy correction to change with the geometry, as seen in Li *et al.*[43] and Su *et al.*[44]

Neither sLOSC nor DFT+U are suitable for solving the analogue of delocalization error for systems with fractional spin.[14,22] For molecules, fractional-spin LOSC (FSLOSC)[89] extends the original LOSC method to this case; the judiciously modified DFT (jmDFT) method[95,96] does the same for DFT+U.

2. Koopmans-compliant functionals

Koopmans-compliant functionals[97,98] mitigate delocalization error by enforcing the PPLB linearity condition

directly: in the Koopmans integral (KI) formulation,[39]

$$\Delta E_{\mathrm{KI}} = \sum_i \alpha_i \left[f_i \eta_i - \int_0^{f_i} ds_i \, \langle \phi_i | h_s(s_i) | \phi_i \rangle \right]. \quad (29)$$

Here α_i is an orbital-dependent screening function based on the relaxation of the LOs ϕ_i; f_i is the (fractional) occupation of ϕ_i; $\eta_i = \int_0^1 ds_i \, \langle \phi_i | h_{\mathrm{PZ}}(s_i) | \phi_i \rangle$, integrating the Perdew–Zunger self-interaction corrected Kohn–Sham Hamiltonian,[32] gives the linearized slope of the energy with respect to f_i; and the last term computes the nonlinearity in E that is replaced by $f_i \eta_i$. Like (s)LOSC, Koopmans-compliant functionals are dependent on the choice of localized orbitals.[99] In extended systems, localized orbitals are necessary for a Koopmans-compliant correction to have any effect,[79] and screening has also been found to effect improvements in band gap calculation.[78,80]

An advantage of sLOSC over the Koopmans-compliant functionals for extended systems is that the DLWFs treat the valence and conduction bands together; on the one hand, DLWFs are empirically robust to increasing the number of conduction bands from which they are constructed, and on the other, sLOSC can in principle be applied to metals without additional modification. For gapped systems, the energy localization inherent in the DLWF cost function enforces separation between the occupied and virtual electronic manifolds without manual input. The system with the smallest gap in our analysis, indium antimonide (InSb, experimental gap 0.23 eV), which is predicted to be gapless by the DFA (and whose sLOSC gap is 0.260 eV) has DLWF occupations $\lambda_{ii}^{\mathbf{RR}} \geq 0.98$ in the valence manifold and ≤ 0.0074 in the conduction manifold. It is conceivable that DLWFs could also serve as effective LOs for Koopmans-compliant methods, even if their subspaces corresponding to the valence and conduction bands are not variational for total Koopmans-compliant energy.

3. *Additional methods*

The Fermi–Löwdin orbital (FLO) self-interaction correction (SIC),[36,100,101] which has its roots in the Perdew–Zunger (PZ) self-interaction correction method,[32] also uses localized orbitals for an energy correction. However, the self-interaction error treated by both FLOSIC and PZ-SIC is well-defined only for one-electron systems; LOSC and its derivatives account for the many-electron nature of delocalization error explicitly.[17]

As mentioned in Su *et al.*[89], the generalized transition state method[41] and the Wannier-function method of Ma and Wang[42] are effective at improving band gap predictions. However, since they do not mix valence and conduction bands to create fractionally occupied orbitals, they cannot change the total energy of the DFA calculation; thus, they cannot restore size-consistency to DFAs and will not be able to capture (for instance) molecular dissociation at the same time as improving band gap predictions. This problem is shared by early Koopmans-compliant methods, which used the Kohn–Sham orbitals as the ϕ_i; it underlies the observation of Nguyen *et al.*[79] that localized orbitals such as Wannier functions are required for Koopmans compliance in extended systems.

B. Computational efficiency

The sLOSC method as implemented in this work scales as $O(N_w^3 N_k)$ for the localization step, $O(N_w^2 N_G)$ for the computation of curvature elements, and $O(N_w N_G \log N_G)$ for the FFT of the DLWF densities. Here, N_G is the number of plane waves in the unfolded supercell, which is N_k times the number of plane waves in the unit cell. Calculating the curvature and energy corrections is the computational bottleneck for the systems evaluated in this work, with wall times for each system reaching a few hours using 16 threads on an Intel Xeon E5-2630v3 processor. The systems that took the longest time were those with the largest number of core states, which have no effect on frontier state corrections; these could be neglected if only a correction to the band gap is desired. The running time was divided fairly evenly between the computation of the matrix elements defined in Eqs. (17, 19a, 19b). We note that the size of the integration domain for these quantities could be reduced from the full Born–von Karman supercell if the relevant DLWF densities are contained in a smaller region. This is supported by the fact that systems that had similar localizations for a 4×4×4 and 6×6×6 **k**-mesh yielded very similar energy corrections.

Along the same lines, we find that for the systems in the test set the correction results are converged with a 6×6×6 **k**-mesh, but this is only necessary to achieve a converged localization. Certain systems exhibit a qualitatively different localization with a smaller 4×4×4 **k**-mesh; however, by decreasing the value of γ in the localization cost function F, a set of DLWFs qualitatively similar to the 6×6×6 case can be obtained.

Some other methods that attempt to address delocalization error in bulk calculations, such as the approach of Ma and Wang[42] and the screened range-separated hybrid functional,[31] rely on supercell self-consistent calculations. These have cubic scaling in the number of electrons, so an unfolded Born–von Karman supercell arising from N_k **k**-points sampling a unit cell with N_w Wannier functions scales as $O(N_k^3 N_w^3)$. Both of the aforementioned methods use Wannier functions as a localized charge representation and rely on manually choosing the Bloch orbitals to comprise the Wannier functions representing the frontier of the occupied space. The sLOSC method uses DLWFs, which naturally supply Wannier functions representing the frontier of the occupied and unoccupied spaces without the need for manual energy windowing.

ACKNOWLEDGMENTS

A.M., J.Z.W., N.Q.S., and W.Y. acknowledge support from the National Science Foundation (Grant No. CHE-1900338); A.M. was additionally supported by the Molecular Sciences Software Institute Phase-II Software Fellowship, and J.Z.W. by the National Institutes of Health (Grant No. 5R01GM061870).

* weitao.yang@duke.edu

[1] W. Kohn, Rev. Mod. Phys. **71**, 1253 (1999), publisher: American Physical Society.

[2] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964), publisher: American Physical Society.

[3] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965), publisher: American Physical Society.

[4] A. D. Becke, The Journal of Chemical Physics **98**, 5648 (1993).

[5] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988), publisher: American Physical Society.

[6] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[7] R. Van Noorden, B. Maher, and R. Nuzzo, Nature News **514**, 550 (2014), cg_type: Nature News Section: News Feature.

[8] A. J. Cohen, P. Mori-Sánchez, and W. Yang, Science **321**, 792 (2008), publisher: American Association for the Advancement of Science Section: Special Perspectives.

[9] A. J. Cohen, P. Mori-Sánchez, and W. Yang, Chem. Rev. **112**, 289 (2012), publisher: American Chemical Society.

[10] P. Mori-Sánchez, A. J. Cohen, and W. Yang, Phys. Rev. Lett. **100**, 146401 (2008), publisher: American Physical Society.

[11] T. Koopmans, Physica **1**, 104 (1934).

[12] J. F. Janak, Phys. Rev. B **18**, 7165 (1978), publisher: American Physical Society.

[13] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, PHYSICAL REVIEW LETTERS **49**, 4 (1982).

[14] A. J. Cohen, P. Mori-Sánchez, and W. Yang, Phys. Rev. B **77**, 115123 (2008).

[15] R. Merkle, A. Savin, and H. Preuss, J. Chem. Phys. **97**, 9216 (1992), publisher: American Institute of Physics.

[16] Y. Zhang and W. Yang, J. Chem. Phys. **109**, 2604 (1998), publisher: American Institute of Physics.

[17] P. Mori-Sánchez, A. J. Cohen, and W. Yang, J. Chem. Phys. **125**, 201102 (2006), publisher: American Institute of Physics.

[18] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, V. N. Staroverov, and J. Tao, Phys. Rev. A **76**, 040501 (2007), publisher: American Physical Society.

[19] T. Körzdörfer, S. Kümmel, and M. Mundt, J. Chem. Phys. **129**, 014110 (2008), publisher: American Institute of Physics.

[20] R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989), publisher: American Physical Society.

[21] A. D. Becke, J. Chem. Phys. **140**, 18A301 (2014), publisher: American Institute of Physics.

[22] P. Mori-Sánchez, A. J. Cohen, and W. Yang, Phys. Rev. Lett. **102**, 066403 (2009), publisher: American Physical Society.

[23] W. Yang, A. J. Cohen, and P. Mori-Sánchez, J. Chem. Phys. **136**, 204111 (2012), publisher: American Institute of Physics.

[24] A. Savin and H.-J. Flad, International Journal of Quantum Chemistry **56**, 327 (1995), _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qua.560560417.

[25] A. Savin, in *Theoretical and Computational Chemistry*, Recent Developments and Applications of Modern Density Functional Theory, Vol. 4, edited by J. M. Seminario (Elsevier, 1996) pp. 327–357.

[26] H. Iikura, T. Tsuneda, T. Yanai, and K. Hirao, J. Chem. Phys. **115**, 3540 (2001), publisher: American Institute of Physics.

[27] T. Yanai, D. P. Tew, and N. C. Handy, Chemical Physics Letters **393**, 51 (2004).

[28] O. A. Vydrov and G. E. Scuseria, J. Chem. Phys. **125**, 234109 (2006), publisher: American Institute of Physics.

[29] J.-D. Chai and M. Head-Gordon, Phys. Chem. Chem. Phys. **10**, 6615 (2008), publisher: The Royal Society of Chemistry.

[30] R. Baer, E. Livshits, and U. Salzner, Annual Review of Physical Chemistry **61**, 85 (2010), _eprint: https://doi.org/10.1146/annurev.physchem.012809.103321.

[31] D. Wing, G. Ohad, J. B. Haber, M. R. Filip, S. E. Gant, J. B. Neaton, and L. Kronik, PNAS **118**, 10.1073/pnas.2104556118 (2021), publisher: National Academy of Sciences Section: Physical Sciences.

[32] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981), publisher: American Physical Society.

[33] P. Mori-Sánchez, A. J. Cohen, and W. Yang, J. Chem. Phys. **124**, 091102 (2006), publisher: American Institute of Physics.

[34] J. P. Perdew, V. N. Staroverov, J. Tao, and G. E. Scuseria, Phys. Rev. A **78**, 052513 (2008), publisher: American Physical Society.

[35] T. Schmidt, E. Kraisler, L. Kronik, and S. Kümmel, Physical Chemistry Chemical Physics **16**, 14357 (2014), publisher: Royal Society of Chemistry.

[36] M. R. Pederson, A. Ruzsinszky, and J. P. Perdew, J. Chem. Phys. **140**, 121103 (2014), publisher: American Institute of Physics.

[37] T. Schmidt and S. Kümmel, Phys. Rev. B **93**, 165120 (2016), publisher: American Physical Society.

[38] Z.-h. Yang, M. R. Pederson, and J. P. Perdew, Phys. Rev. A **95**, 052505 (2017), publisher: American Physical Society.

[39] G. Borghi, A. Ferretti, N. L. Nguyen, I. Dabo, and N. Marzari, Phys. Rev. B **90**, 075135 (2014).

[40] N. Colonna, N. L. Nguyen, A. Ferretti, and N. Marzari, J. Chem. Theory Comput. **15**, 1905 (2019), publisher: American Chemical Society.

[41] V. I. Anisimov and A. V. Kozhevnikov, Phys. Rev. B **72**, 075125 (2005), publisher: American Physical Society.

[42] J. Ma and L.-W. Wang, Sci Rep **6**, 24924 (2016).

[43] C. Li, X. Zheng, N. Q. Su, and W. Yang, National Science Review **5**, 203 (2018).

[44] N. Q. Su, A. Mahler, and W. Yang, J. Phys. Chem. Lett. **11**, 1528 (2020).

[45] Y. Mei, Z. Chen, and W. Yang, The Journal of Physical Chemistry Letters 10.1021/acs.jpclett.0c03133 (2020), publisher: American Chemical Society.

[46] Y. Mei, N. Yang, and W. Yang, J. Chem. Phys. **154**, 054302 (2021), publisher: American Institute of Physics.
[47] F. Bloch, Z. Physik **52**, 555 (1929).
[48] N. Ashcroft and D. Mermin, *Solid State Physics* (Thomson Press, 2003).
[49] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).
[50] G. Wannier, Physical Review **52**, 191 (1937).
[51] J. D. Cloizeaux, Phys. Rev. **129**, 554 (1963), publisher: American Physical Society.
[52] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997), publisher: American Physical Society.
[53] J. M. Foster and S. F. Boys, Rev. Mod. Phys. **32**, 300 (1960), publisher: American Physical Society.
[54] F. Gygi, J.-L. Fattebert, and E. Schwegler, Computer Physics Communications **155**, 1 (2003).
[55] F. Giustino and A. Pasquarello, Phys. Rev. Lett. **96**, 216403 (2006).
[56] A. Mahler, J. Z. Williams, N. Q. Su, and W. Yang, arXiv:2201.07751 [cond-mat] (2022), arXiv: 2201.07751.
[57] Y. Mei, J. Yu, Z. Chen, N. Q. Su, and W. Yang, J. Chem. Theory Comput. 10.1021/acs.jctc.1c01058 (2022).
[58] X. Zheng, A. J. Cohen, P. Mori-Sánchez, X. Hu, and W. Yang, Phys. Rev. Lett. **107**, 026403 (2011), publisher: American Physical Society.
[59] D. Hait and M. Head-Gordon, J. Phys. Chem. Lett. **9**, 6280 (2018), publisher: American Chemical Society.
[60] Y. Mei, Z. Chen, and W. Yang, arXiv:2106.10358 [physics, physics:quant-ph] (2021), arXiv: 2106.10358.
[61] P. a. M. Dirac, Mathematical Proceedings of the Cambridge Philosophical Society **26**, 376 (1930), publisher: Cambridge University Press.
[62] J. C. Slater and J. C. Phillips, Physics Today **27**, 49 (1974), publisher: American Institute of Physics.
[63] J. C. Slater, *Quantum Theory of Molecules and Solids*, Vol. IV (McGrow-Hill, 1974).
[64] I. Souza, N. Marzari, and D. Vanderbilt, Phys. Rev. B **65**, 035109 (2001), publisher: American Physical Society.
[65] G. Makov and M. C. Payne, Phys. Rev. B **51**, 4014 (1995).
[66] L. N. Kantorovich, Phys. Rev. B **60**, 15476 (1999), publisher: American Physical Society.
[67] I. Dabo, B. Kozinsky, N. E. Singh-Miller, and N. Marzari, Phys. Rev. B **77**, 115139 (2008), arXiv: 0709.4647.
[68] Y. Li and I. Dabo, Phys. Rev. B **84**, 155127 (2011), publisher: American Physical Society.
[69] G. Onida, L. Reining, R. W. Godby, R. Del Sole, and W. Andreoni, Phys. Rev. Lett. **75**, 818 (1995), publisher: American Physical Society.
[70] M. R. Jarvis, I. D. White, R. W. Godby, and M. C. Payne, Phys. Rev. B **56**, 14972 (1997), publisher: American Physical Society.
[71] C. A. Rozzi, D. Varsano, A. Marini, E. K. U. Gross, and A. Rubio, Phys. Rev. B **73**, 205119 (2006).
[72] L. Hedin, Physical Review series I **139**, 796 (1965), number: 3A Publisher: American Institute of Physics (AIP).
[73] R. Martin, L. Reining, and D. Ceperley, *Interacting Electrons: Theory and Computational Approaches* (Cambridge University Press, 2016).
[74] M. R. Zaghloul and A. N. Ali, ACM Trans. Math. Softw. **38**, 15:1 (2012).
[75] M. R. Zaghloul, ACM Trans. Math. Softw. **42**, 26:1 (2016).
[76] X. Zheng, T. Zhou, and W. Yang, J. Chem. Phys. **138**, 174105 (2013).
[77] D. Zhang, X. Zheng, C. Li, and W. Yang, J. Chem. Phys. **142**, 154113 (2015).
[78] N. Colonna, N. L. Nguyen, A. Ferretti, and N. Marzari, J. Chem. Theory Comput. **14**, 2549 (2018).
[79] N. L. Nguyen, N. Colonna, A. Ferretti, and N. Marzari, Phys. Rev. X **8**, 021051 (2018).
[80] N. Colonna, R. De Gennaro, E. Linscott, and N. Marzari, arXiv (2022), arXiv:2202.08155 [cond-mat, physics:physics].
[81] W. Yang, A. J. Cohen, F. De Proft, and P. Geerlings, J. Chem. Phys. **136**, 144110 (2012).
[82] D. R. Hamann, Phys. Rev. B **88**, 085117 (2013).
[83] M. J. van Setten, M. Giantomassi, E. Bousquet, M. J. Verstraete, D. R. Hamann, X. Gonze, and G. M. Rignanese, Computer Physics Communications **226**, 39 (2018).
[84] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. d. Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, J. Phys.: Condens. Matter **21**, 395502 (2009), publisher: IOP Publishing.
[85] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. D. Corso, S. d. Gironcoli, P. Delugas, R. A. DiStasio, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. Otero-de-la Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, and S. Baroni, J. Phys.: Condens. Matter **29**, 465901 (2017), publisher: IOP Publishing.
[86] A. A. Mostofi, J. R. Yates, G. Pizzi, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari, Computer Physics Communications **185**, 2309 (2014).
[87] G. Pizzi, V. Vitale, R. Arita, S. Blügel, F. Freimuth, G. Géranton, M. Gibertini, D. Gresch, C. Johnson, T. Koretsune, J. Ibañez-Azpiroz, H. Lee, J.-M. Lihm, D. Marchand, A. Marrazzo, Y. Mokrousov, J. I. Mustafa, Y. Nohara, Y. Nomura, L. Paulatto, S. Poncé, T. Ponweiser, J. Qiao, F. Thöle, S. S. Tsirkin, M. Wierzbowska, N. Marzari, D. Vanderbilt, I. Souza, A. A. Mostofi, and J. R. Yates, J. Phys.: Condens. Matter **32**, 165902 (2020).
[88] J. Heyd, J. E. Peralta, G. E. Scuseria, and R. L. Martin, J. Chem. Phys. **123**, 174101 (2005), publisher: American Institute of Physics.
[89] N. Q. Su, C. Li, and W. Yang, PNAS **115**, 9678 (2018).
[90] V. I. Anisimov, J. Zaanen, and O. K. Andersen, Phys. Rev. B **44**, 943 (1991), publisher: American Physical Society.
[91] M. Cococcioni and S. de Gironcoli, Phys. Rev. B **71**, 035105 (2005).
[92] D. D. O'Regan, in *Optimised Projections for the Ab Initio Simulation of Large and Strongly Correlated Systems*, edited by D. D. O'Regan (Springer, Berlin, Heidelberg, 2012) pp. 89–123.
[93] V. L. J. Campo and M. Cococcioni, J. Phys.: Condens. Matter **22**, 055602 (2010).
[94] B. Himmetoglu, A. Floris, S. de Gironcoli, and M. Cococcioni, Int. J. Quantum Chem. **114**, 14 (2014).

[95] A. Bajaj, J. P. Janet, and H. J. Kulik, J. Chem. Phys. **147**, 191101 (2017).
[96] A. Bajaj, F. Liu, and H. J. Kulik, J. Chem. Phys. **150**, 154115 (2019).
[97] I. Dabo, M. Cococcioni, and N. Marzari, arXiv:0901.2637 [cond-mat] (2009), arXiv:0901.2637 [cond-mat].
[98] I. Dabo, A. Ferretti, N. Poilvert, Y. Li, N. Marzari, and M. Cococcioni, Phys. Rev. B **82**, 115121 (2010).
[99] G. Borghi, C.-H. Park, N. L. Nguyen, A. Ferretti, and N. Marzari, Phys. Rev. B **91**, 155112 (2015).
[100] D.-y. Kao, K. Withanage, T. Hahn, J. Batool, J. Kortus, and K. Jackson, J. Chem. Phys. **147**, 164107 (2017), publisher: American Institute of Physics.
[101] K. A. Jackson, J. E. Peralta, R. P. Joshi, K. P. Withanage, K. Trepte, K. Sharkas, and A. I. Johnson, J. Phys.: Conf. Ser. **1290**, 012002 (2019).

Supplemental Material for: Localized orbital scaling correction for periodic systems

Aaron Mahler

Duke University, Department of Physics, Durham, NC 27708

Jacob Williams

Duke University, Department of Chemistry, Durham, NC 27708

Neil Qiang Su

Department of Chemistry, Key Laboratory of Advanced Energy Materials Chemistry (Ministry of Education) and Renewable Energy Conversion and Storage Center (RECAST), Nankai University, Tianjin 300071, China

Weitao Yang*

Duke University, Department of Chemistry, Durham, NC 27708 and

Duke University, Department of Physics, Durham, NC 27708

(Dated: June 2, 2022)

S1. THE LOSC HAMILTONIAN CORRECTION

Recall from the main text that the energy correction due to LOSC is

$$\Delta E = \frac{1}{2}\sum_{ij} \overline{\lambda}_{ij}(\delta_{ij} - \lambda_{ij})\widetilde{\kappa}_{ij}, \tag{S1}$$

where $\overline{\lambda}$ is the complex conjugate of λ, δ_{ij} is the Kronecker delta function, and we have combined the band and unit cell indices of the dually localized Wannier functions (DLWFs) into a single label i. Then, under the frozen orbital approximation, the LOSC correction to the Hamiltonian is

$$\Delta v = \left.\frac{\delta \Delta E}{\delta \rho_s(\mathbf{x}, \mathbf{x}')}\right|_{\{w_i\}}, \tag{S2}$$

where

$$\rho_s = \sum_{ij} \lambda_{ij} |w_i\rangle\langle w_j| \tag{S3}$$

is the Kohn–Sham density matrix. Note that ρ_s is not diagonal in the basis of DLWFs; furthermore, unlike the canonical orbital occupations, which are either 0 or 1, DLWFs can include virtual orbital character, and their occupations are given by $\lambda_{ij} = \langle w_i|\rho_s|w_j\rangle \in \mathbb{C}$. In the position basis, we write

$$\rho_s(\mathbf{x}, \mathbf{x}') = \sum_{ij} \lambda_{ij} w_i(\mathbf{x})\overline{w}_j(\mathbf{x}'), \tag{S4}$$

where $\mathbf{x} = (\mathbf{r}, \sigma)$ combines space and spin variables. Then we may write the Hamiltonian correction as

$$\Delta v = \sum_{ij} \left[\frac{\partial \Delta E}{\partial \lambda_{ij}} \frac{\delta \lambda_{ij}}{\delta \rho_s(\mathbf{x}, \mathbf{x}')} + \frac{\partial \Delta E}{\partial \overline{\lambda}_{ij}} \frac{\delta \overline{\lambda}_{ij}}{\delta \rho_s(\mathbf{x}, \mathbf{x}')}\right]. \tag{S5}$$

Let us consider the terms in this expression individually. First, consider the (Wirtinger [1]) partial derivative of ΔE with respect to λ_{ij} and its complex conjugate. The diagonal elements ($i = j$) give

$$\frac{\partial \Delta E}{\partial \lambda_{ii}} = -\frac{1}{2}\overline{\lambda}_{ij}\widetilde{\kappa}_{ii}; \quad \frac{\partial \Delta E}{\partial \overline{\lambda}_{ii}} = \frac{1}{2}(1 - \lambda_{ii})\widetilde{\kappa}_{ii}. \tag{S6}$$

If $i \neq j$, on the other hand,

$$\frac{\partial \Delta E}{\partial \lambda_{ij}} = -\frac{1}{2}\overline{\lambda}_{ij}\widetilde{\kappa}_{ij}; \quad \frac{\partial \Delta E}{\partial \overline{\lambda}_{ij}} = -\frac{1}{2}\lambda_{ij}\widetilde{\kappa}_{ij}. \tag{S7}$$

* weitao.yang@duke.edu

Finally, consider the variation in the local occupations with respect to the density matrix. From the relationship of ρ_s and λ_{ij}, we obtain

$$\frac{\delta\lambda_{ij}}{\delta\rho_s(\mathbf{x},\mathbf{x}')} = w_j(\mathbf{x})\overline{w}_i(\mathbf{x}'); \quad \frac{\delta\overline{\lambda}_{ij}}{\delta\rho_s(\mathbf{x},\mathbf{x}')} = w_i(\mathbf{x})\overline{w}_j(\mathbf{x}'). \tag{S8}$$

Combining yields

$$\begin{aligned}\Delta v = \frac{1}{2}\Bigg[&\sum_i (1-2\lambda_{ii})\,\widetilde{\kappa}_{ii} w_i(\mathbf{x})\overline{w}_i(\mathbf{x}') \\ &- \sum_{i\neq j}\left(\lambda_{ij}\widetilde{\kappa}_{ij} w_j(\mathbf{x})\overline{w}_i(\mathbf{x}') + \overline{\lambda}_{ji}\widetilde{\kappa}_{ij} w_i(\mathbf{x})\overline{w}_j(\mathbf{x}')\right)\Bigg].\end{aligned} \tag{S9}$$

Noting that the matrix (λ_{ij}) of local occupations is Hermitian and that the curvature matrix $(\widetilde{\kappa}_{ij})$ is real symmetric, we can rewrite our expression in operator form as

$$\begin{aligned}\Delta v &= \sum_{ij}\left(\frac{1}{2}\delta_{ij} - \operatorname{Re}\lambda_{ij}\right)\widetilde{\kappa}_{ij}|w_i\rangle\langle w_j| \\ &= \frac{1}{2}\sum_{ij}\left(\frac{1}{2}\delta_{ij} - \lambda_{ij}\right)\widetilde{\kappa}_{ij}|w_i\rangle\langle w_j| + \text{h.c.},\end{aligned} \tag{S10}$$

as seen in the main text.

S2. ENERGY CORRECTION WITH DISENTANGLEMENT

The construction of dually localized Wannier functions (DLWFs) requires a set of composite Bloch bands: that is, one which is separated by an energy gap from all other bands at each point in the Brillouin zone. The valence bands of semiconductors and insulators form a composite set by definition; however, including virtual bands of semiconductors means that a composite set may not be attainable, and metals may not have any composite set of bands at all. We refer to a set of bands which cannot be separated by an energy gap from the rest as *entangled*. Souza et al. showed in [2] that from N_b entangled bands a set of $N_w < N_b$ of Bloch orbitals at each $\mathbf{k}$-point can be obtained; this procedure is called *disentanglement*. Note that these disentangled orbitals are no longer eigenfunctions of the single-particle Hamiltonian, but they may be treated as a composite set. When LOSC is applied to the DLWFs found from disentangled Bloch orbitals, the orbital energy correction is applied to the disentangled Bloch orbitals to obtain the correction to the frontier orbital energies and the band gap. If the disentangled set of eigenvalues is a subset of the true eigenvalues in an area of interest (close to the Fermi energy and the conduction band minimum), then correcting the disentangled eigenvalues is equivalent to correcting the original eigenvalues. We found in [3] that

including sufficiently many virtual orbitals in the LOSC procedure satisfies this condition. For completeness, we demonstrate the form of the eigenvalue correction when applied to the original Bloch eigenvalues. This can be found by finding the expectation of the LOSC Hamiltonian correction, Δv, in the Bloch orbital basis.

To begin, let us represent the disentangled Bloch orbitals $|\varphi_n^{\mathbf{k}}\rangle$ in the original Bloch orbital basis $|\psi_b^{\mathbf{k}}\rangle$,

$$|\varphi_n^{\mathbf{k}}\rangle = \sum_b^{N_b} V_{nb}^{\mathbf{k}} |\psi_b^{\mathbf{k}}\rangle. \tag{S11}$$

Here, $V^{\mathbf{k}}$ is an $N_w \times N_b$ matrix that obeys the relation $V^{\mathbf{k}}(V^{\mathbf{k}})^\dagger = I_{N_w}$, where I_{N_w} is the $N_w \times N_w$ identity operator. We may then write a transformed Bloch orbital for generalized Wannier function construction in the disentangled Bloch orbital basis as

$$|\phi_n^{\mathbf{k}}\rangle = \sum_a^{N_w} U_{na}^{\mathbf{k}} |\varphi_a^{\mathbf{k}}\rangle. \tag{S12}$$

A generalized Wannier function constructed from a set of disentangled Bloch orbitals then takes the form

$$\begin{aligned} |w_m^{\mathbf{R}}\rangle &= \frac{1}{N_k} \sum_{\mathbf{k}}^{N_k} e^{-i\mathbf{k}\cdot\mathbf{R}} \sum_a^{N_w} U_{ma}^{\mathbf{k}} \sum_b^{N_b} V_{ab}^{\mathbf{k}} |\psi_b^{\mathbf{k}}\rangle \\ &= \frac{1}{N_k} \sum_{\mathbf{k}}^{N_k} e^{-i\mathbf{k}\cdot\mathbf{R}} \sum_b^{N_b} L_{mb}^{\mathbf{k}} |\psi_b^{\mathbf{k}}\rangle, \end{aligned} \tag{S13}$$

where $L_{mb}^{\mathbf{k}} = [U^{\mathbf{k}} V^{\mathbf{k}}]_{mb}$. Recall that the LOSC method corrects the Kohn–Sham Hamiltonian by

$$\begin{aligned} \Delta v &= \frac{\delta \Delta E^{\mathrm{LOSC}}}{\delta \rho_s} \\ &= \sum_{\mathbf{R}mn} \widetilde{\kappa}_{mn}^{\mathbf{0R}} \mathrm{Re} \left\{ \left(\frac{\delta_{mn}^{\mathbf{0R}}}{2} - \lambda_{mn}^{\mathbf{0R}} \right) |w_m^{\mathbf{0}}\rangle\langle w_n^{\mathbf{R}}| \right\}. \end{aligned} \tag{S14}$$

To compute the matrix elements of Δv in the basis of Bloch orbitals, we require the overlap of a Wannier function and Bloch orbital:

$$\begin{aligned} \langle w_n^{\mathbf{R}} | \psi_i^{\mathbf{k}} \rangle &= \frac{1}{N_k} \sum_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{R}} \langle \phi_n^{\mathbf{q}} | \psi_i^{\mathbf{k}} \rangle \\ &= \frac{1}{N_k} \sum_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{R}} \sum_b^{N_b} \overline{L}_{nb}^{\mathbf{q}} \langle \psi_b^{\mathbf{q}} | \psi_i^{\mathbf{k}} \rangle \\ &= e^{i\mathbf{k}\cdot\mathbf{R}} \overline{L}_{ni}^{\mathbf{k}}, \end{aligned} \tag{S15}$$

where we have used the normalization convention $\langle \psi_b^{\mathbf{q}} | \psi_i^{\mathbf{k}} \rangle = N \delta_{\mathbf{qk}} \delta_{bi}$. We write $\overline{z}$ for the complex conjugate of z. We can now easily find the expectation of the LOSC Hamiltonian correction in the Bloch orbital basis,

$$\begin{aligned} \langle \psi_i^{\mathbf{k}} | \Delta v | \psi_i^{\mathbf{k}} \rangle &= \sum_{\mathbf{R}mn} \widetilde{\kappa}_{mn}^{\mathbf{0R}} \mathrm{Re} \left\{ \left(\frac{\delta_{mn}^{\mathbf{0R}}}{2} - \lambda_{mn}^{\mathbf{0R}} \right) \langle \psi_i^{\mathbf{k}} | w_m^{\mathbf{0}} \rangle \langle w_n^{\mathbf{R}} | \psi_i^{\mathbf{k}} \rangle \right\} \\ &= \sum_{\mathbf{R}mn} \widetilde{\kappa}_{mn}^{\mathbf{0R}} \mathrm{Re} \left\{ \left(\frac{\delta_{mn}^{\mathbf{0R}}}{2} - \lambda_{mn}^{\mathbf{0R}} \right) L_{mi}^{\mathbf{k}} \overline{L}_{ni}^{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}}. \right\} \end{aligned} \tag{S16}$$

Functionally, the only difference between this form and using the disentangled basis is that instead of $L^{\mathbf{k}}$ the correction would have $U^{\mathbf{k}}$, the transformation from the (possibly) disentangled basis to the transformed basis.

S3. THE COULOMB INTEGRAL

In this section, we discuss details of the Coulomb interaction J between two dually localized Wannier functions (DLWFs) $|w_i\rangle$ and $|w_j\rangle$; note that we combine the DLWFs' band-like and unit cell indices into a single subscript, so $|w_i\rangle = |w_m^{\mathbf{R}}\rangle$. J depends on the charge density ρ associated with each DLWF, $\rho_i(\mathbf{r}) = |w_i(\mathbf{r})|^2$, and is given by

$$J[\rho_i, \rho_j] = \iint d\mathbf{r}\, d\mathbf{r}' \, \frac{\rho_i(\mathbf{r}) \rho_j(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} = \iint d\mathbf{r}\, d\mathbf{r}' \, \rho_i(\mathbf{r}) \rho_j(\mathbf{r}') \, K(|\mathbf{r} - \mathbf{r}'|), \tag{S17}$$

where we term $K(r) = 1/r$ the *Coulomb kernel*.

Naïvely, the cost of computing $J[\rho_i, \rho_j]$ scales as $O(N_{\mathbf{r}})^2$, where $N_{\mathbf{r}}$ is the size of the densities on the real grid. However, the DLWF densities are periodic on the Born–von Karman supercell (see the main text). Recall that we may write the Fourier series for sufficiently well-behaved (that is, obeying Dirichlet conditions) periodic functions $f(\mathbf{r})$ as

$$f(\mathbf{r}) = \frac{1}{\Omega_{\mathrm{BvK}}} \sum_{\mathbf{G}} f(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}, \tag{S18}$$

where the basis vectors $\mathbf{G}$ belong to the reciprocal lattice and Ω_{BvK} is the volume of the Born–von Karman supercell. Each DLWF density can be so represented, yielding $N_{\mathbf{G}}$ Fourier coefficients $\rho(\mathbf{G})$. We can then exploit the fast Fourier transform to compute $J[\rho_i, \rho_j]$ in $O(N_{\mathbf{G}} \log N_{\mathbf{G}})$ time.

The Fourier-space representation of $J[\rho_i, \rho_j]$ is a classical result in electrodynamics, but we present it briefly here. First, observe that we can write

$$J[\rho_i, \rho_j] = \iint d\mathbf{r}\, d\mathbf{r}' \, \frac{\rho_i(\mathbf{r}) \rho_j(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} = \int d\mathbf{r} \, V_i(\mathbf{r}) \rho_j(\mathbf{r}), \tag{S19}$$

where V_i is the potential due to ρ_i experienced by ρ_j,

$$V_i(\mathbf{r}) = \int d\mathbf{r}' \, \frac{\rho_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \tag{S20}$$

V_i inherits the periodicity of ρ_i and can be written as a Fourier series; it also satisfies the Poisson equation

$$\nabla^2 V_i(\mathbf{r}) = -4\pi\rho_i(\mathbf{r}). \tag{S21}$$

This equation is algebraic in the plane-wave basis since $\nabla^2 e^{i\mathbf{G}\cdot\mathbf{r}} = -G^2$, where $G = |\mathbf{G}|$; thus

$$\nabla^2 \frac{1}{\Omega_{\mathrm{BvK}}} \left(\sum_{\mathbf{G}} V_i(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}} \right) = -\frac{1}{\Omega_{\mathrm{BvK}}} \sum_{\mathbf{G}} G^2 V_i(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}} = -\frac{4\pi}{\Omega_{\mathrm{BvK}}} \sum_{\mathbf{G}} \rho_i(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}. \tag{S22}$$

Since the plane waves $e^{i\mathbf{G}\cdot\mathbf{r}}$ are orthonormal, we can solve this equation term by term, obtaining the Fourier coefficients

$$V_i(\mathbf{G}) = 4\pi \frac{\rho_i(\mathbf{G})}{G^2}. \tag{S23}$$

Substituting V_i into the original expression and replacing ρ_j by its Fourier series gives

$$\begin{aligned} J[\rho_i, \rho_j] &= \frac{4\pi}{\Omega_{\mathrm{BvK}}^2} \int d\mathbf{r} \, \left(\sum_{\mathbf{G}} \frac{\rho_i(\mathbf{G})}{G^2} e^{i\mathbf{G}\cdot\mathbf{r}} \right) \left(\sum_{\mathbf{G}'} \rho_j(\mathbf{G}') e^{i\mathbf{G}'\cdot\mathbf{r}} \right) \\ &= \frac{4\pi}{\Omega_{\mathrm{BvK}}^2} \int d\mathbf{r} \, \left(\sum_{\mathbf{G}} \frac{\overline{\rho}_i(\mathbf{G})}{G^2} e^{-i\mathbf{G}\cdot\mathbf{r}} \right) \left(\sum_{\mathbf{G}'} \rho_j(\mathbf{G}') e^{i\mathbf{G}'\cdot\mathbf{r}} \right) \\ &= \frac{4\pi}{\Omega_{\mathrm{BvK}}^2} \int d\mathbf{r} \sum_{\mathbf{G}\mathbf{G}'} \frac{\overline{\rho}_i(\mathbf{G}) \rho_j(\mathbf{G}')}{G^2} e^{i(\mathbf{G}'-\mathbf{G})\cdot\mathbf{r}} = \frac{4\pi}{\Omega_{\mathrm{BvK}}} \sum_{\mathbf{G}} \frac{\overline{\rho}_i(\mathbf{G}) \rho_j(\mathbf{G})}{G^2}. \end{aligned} \tag{S24}$$

We have that $\rho_i(-\mathbf{G}) = \overline{\rho}_i(\mathbf{G})$, where $\overline{\rho}$ is the complex conjugate of ρ, since $\rho_i(\mathbf{r})$ is real. The integral and double sum over $\mathbf{G}, \mathbf{G}'$ collapse into a single sum by the orthonormality of the plane waves $e^{i\mathbf{G}\cdot\mathbf{r}}$:

$$\int d\mathbf{r} \, e^{i(\mathbf{G}-\mathbf{G}')\cdot\mathbf{r}} = \Omega_{\mathrm{BvK}} \delta_{\mathbf{G},\mathbf{G}'}, \tag{S25}$$

where $\delta_{\mathbf{G},\mathbf{G}'}$ is the Kronecker delta function.

In perfect analogy with the real-space description, we can express the Coulomb integral in reciprocal space as a sum over Fourier coefficients with a reciprocal-space Coulomb kernel. That is,

$$J[\rho_i, \rho_j] = \frac{1}{\Omega_{\mathrm{BvK}}} \sum_{\mathbf{G}} \overline{\rho_i}(\mathbf{G}) \rho_j(\mathbf{G}) K(G), \tag{S26}$$

where $K(G) = 4\pi/G^2$. This kernel has a singularity at $G = 0$, as does $K(r)$; however, unlike its real-space counterpart, the singularity in $K(G)$ is not integrable in three dimensions.

For the density of the entire unit cell, the singularity is circumvented by requiring the cell to be charge neutral, which is equivalent to setting $K(0) = 0$. However, the DLWFs are individually charged, so an alternative approach must be chosen. We address two problems in this work: the spurious interaction of a DLWF density with its periodically repeating images, and the suppression by the bulk environment of the long-range behavior of the Coulomb kernel. Both of these can be addressed by modifying the Coulomb kernel $K(r)$.

S3.1. The spherical cutoff kernel

The Fourier coefficients $K(G)$ of the Coulomb kernel can be derived by taking the Fourier transform of the Yukawa kernel $K_y(r) = e^{-\lambda r}/r$ in the limit $\lambda \to 0$. Likewise, the Fourier transforms of other modified kernels $K_{\text{mod}}(r)$ yield the modified Coulomb interaction in the plane-wave basis as

$$J_{\text{mod}}[\rho_i, \rho_j] = \iint d\mathbf{r}\, d\mathbf{r}'\, \rho_i(\mathbf{r})\rho_j(\mathbf{r}')K_{\text{mod}}(|\mathbf{r} - \mathbf{r}'|) = \frac{1}{\Omega_{\text{BvK}}} \sum_{\mathbf{G}} \overline{\rho}_i(\mathbf{G})\rho_j(\mathbf{G})K_{\text{mod}}(|\mathbf{G}|). \quad \text{(S27)}$$

In order to prevent spurious interaction with the densities' periodic images, we use the spherical cutoff kernel K_c, used first by Onida and coworkers in [4] for GW calculations and applied to DFT by Jarvis and coworkers in [5]. It is defined as

$$K_c(r; R_c) = \frac{1 - \Theta(r - R_c)}{r} = \begin{cases} \dfrac{1}{r} & r < R_c \\ 0 & r \geq R_c, \end{cases} \quad \text{(S28)}$$

where $\Theta(r)$ is the Heaviside step function. R_c is the cutoff radius, taken here to be half the length of the shortest Born–von Karman supercell lattice vector. Performing the Fourier transform yields

$$K_c(G; R_c) = \begin{cases} \dfrac{4\pi}{G^2}\left[1 - \cos(GR_c)\right] & G \neq 0 \\ 2\pi R_c^2 & G = 0. \end{cases} \quad \text{(S29)}$$

Note that there is no singularity in $K_c(G; R_c)$. The sinusoidal oscillations are expected; they exemplify the Gibbs phenomenon due to the jump discontinuity of $K_c(r; R_c)$ at R_c.

S3.2. The sLOSC kernel

Even after introducing the spherical cutoff, the effect of the lattice means that the long-range $1/r$ behavior of $J[\rho_i, \rho_j]$ must be attenuated to give an accurate correction to the band

gap of bulk systems. To complete the sLOSC kernel, we therefore introduce an additional screening term proportional to the complementary error function $\mathrm{erfc}(r) = 1 - \mathrm{erf}(r)$, yielding

$$K_s(r; R_c, \alpha) = \frac{\mathrm{erfc}(\alpha r)\,[1 - \Theta(r - R_c)]}{r} = \begin{cases} \dfrac{\mathrm{erfc}(\alpha r)}{r} & r < R_c \\ 0 & r \geq R_c. \end{cases} \tag{S30}$$

Let us derive the reciprocal-space kernel $K_s(G)$ by Fourier transforming $K_s(r)$ directly, utilizing the rotational symmetry of the real-space kernel to convert into spherical coordinates. Thus

$$\begin{aligned} K_s(G; R_c, \alpha) &= \int d\mathbf{r}\, K_s(r; R_c, \alpha) e^{-i\mathbf{G}\cdot\mathbf{r}} \\ &= \int_0^{2\pi} d\theta \int_0^{\pi} d\phi \int_0^{\infty} dr\, r^2 \sin\phi \frac{\mathrm{erfc}(\alpha r)\,[1 - \Theta(r - R_c)]}{r} e^{-iGr\cos\phi} \\ &= 2\pi \int_0^{\pi} d\phi\, \sin\phi \int_0^{R_c} r\,\mathrm{erfc}(\alpha r) e^{-iGr\cos\phi}. \end{aligned} \tag{S31}$$

Setting $u = \cos\phi$, whence $du = -d\phi\,\sin\phi$, we next obtain

$$\begin{aligned} K_s(G; R_c, \alpha) &= 2\pi \int_0^{R_c} d\mathbf{r}\, r\,\mathrm{erfc}(\alpha r) \int_{-1}^{1} du\, e^{-iGru} = 2\pi \int_0^{R_c} d\mathbf{r}\, r\,\mathrm{erfc}(\alpha r) \left[\frac{e^{iGr} - e^{-iGr}}{iGr}\right] \\ &= \frac{4\pi}{G} \int_0^{R_c} d\mathbf{r}\ \mathrm{erfc}(\alpha r)\sin(Gr). \end{aligned} \tag{S32}$$

Performing the final integral with the Mathematica software package [6] yields

$$\begin{aligned} K_s(G; R_c, \alpha) = \frac{4\pi}{G^2}\Bigg\{ & 1 - \cos(GR_c)\,\mathrm{erfc}(\alpha R_c) + \\ & \frac{e^{-G^2/4\alpha^2}}{2}\left[\mathrm{erf}\left(-\alpha R_c + \frac{iG}{2\alpha}\right) - \mathrm{erf}\left(\alpha R_c + \frac{iG}{2\alpha}\right)\right]\Bigg\}. \end{aligned} \tag{S33}$$

Writing $z = \alpha R_c + iG/2\alpha$ and noting that $\mathrm{erf}(\overline{z}) = \overline{\mathrm{erf}\, z}$ and $\mathrm{erf}(-z) = -\,\mathrm{erf}\, z$, we may simplify the relevant part of $K_s(G; R_c, \alpha)$:

$$\mathrm{erf}(-\overline{z}) - \mathrm{erf}(z) = -(\mathrm{erf}\,\overline{z} + \mathrm{erf}\, z) = -(\overline{\mathrm{erf}\, z} + \mathrm{erf}\, z) = -2\,\mathrm{Re}\{\mathrm{erf}\, z\}. \tag{S34}$$

Thus, we obtain in accordance with the main text that when $G \neq 0$

$$K_s(G; R_c, \alpha) = \frac{4\pi}{G^2}\left[1 - \cos(GR_c)\,\mathrm{erfc}(\alpha R_c) - e^{-G^2/4\alpha^2}\,\mathrm{Re}\left\{\mathrm{erf}\left(\alpha R_c + \frac{iG}{2\alpha}\right)\right\}\right]. \tag{S35}$$

Note that $K_s(G)$ is not singular at $G = 0$; we have

$$K_s(G = 0; \alpha) = 2\pi R_c^2 + \pi\,\mathrm{erf}(\alpha R_c)\left(\frac{1}{\alpha^2} - 2R_c^2\right) - \frac{2\sqrt{\pi} R_c e^{-(\alpha R_c)^2}}{\alpha}. \tag{S36}$$

S3.2.1. The Faddeeva function

By Liouville's theorem, $\operatorname{erf} z$ is unbounded for $z \in \mathbb{C}$; in practice, it overflows floating-point arithmetic even for **G**-vectors of relatively small modulus. Therefore, to use the sLOSC kernel in practice, we require a scaled implementation of the complex error function, namely the Faddeeva function

$$w(z) = e^{-z^2} \operatorname{erfc}(-iz). \tag{S37}$$

In particular, we seek to replace

$$e^{-G^2/4\alpha^2} \operatorname{Re}\left\{\operatorname{erf}\left(\alpha R_c + \frac{iG}{2\alpha}\right)\right\}$$

by a term containing $w(z)$. As mentioned in the main text, we implement the Faddeeva function using ACM Algorithm 916 [7, 8].

Note that whenever $\operatorname{Re} z > 0$ we may write

$$\operatorname{erf} z = 1 - e^{-z^2} w(iz); \tag{S38}$$

in our case, $\operatorname{Re} z = \alpha R_c > 0$. Thus, we may write

$$\begin{aligned}
&e^{-G^2/4\alpha^2} \operatorname{Re}\left\{\operatorname{erf}\left(\alpha R_c + \frac{iG}{2\alpha}\right)\right\} \\
&\quad = e^{-G^2/4\alpha^2} \operatorname{Re}\left\{1 - e^{-(\alpha R_c + iG/2\alpha)^2}\, w\left(i\alpha R_c - \frac{G}{2\alpha}\right)\right\} \\
&\quad = e^{-G^2/4\alpha^2} - \operatorname{Re}\left\{e^{-(\alpha R_c + iG/2\alpha)^2 - (G/2\alpha)^2}\, w\left(i\alpha R_c - \frac{G}{2\alpha}\right)\right\} \\
&\quad = e^{-G^2/4\alpha^2} - \operatorname{Re}\left\{e^{-(\alpha R_c)^2} e^{-iGR_c}\, w\left(i\alpha R_c - \frac{G}{2\alpha}\right)\right\} \\
&\quad = e^{-G^2/4\alpha^2} - e^{-(\alpha R_c)^2} \operatorname{Re}\left\{e^{-iGR_c}\, w\left(i\alpha R_c - \frac{G}{2\alpha}\right)\right\}.
\end{aligned} \tag{S39}$$

Breaking the Faddeeva function into its real and imaginary parts as

$$w\left(i\alpha R_c - \frac{G}{2\alpha}\right) = V + iL, \tag{S40}$$

we obtain

$$\begin{aligned}
\operatorname{Re}\left\{e^{-iGR_c}\, w\left(i\alpha R_c - \frac{G}{2\alpha}\right)\right\} &= \operatorname{Re}\left\{[\cos(GR_c) - i\sin(GR_c)]\; w\left(i\alpha R_c - \frac{G}{2\alpha}\right)\right\} \\
&= \operatorname{Re}\left\{[\cos(GR_c) - i\sin(GR_c)]\,(V + iL)\right\} \\
&= V\,\cos(GR_c) + L\,\sin(GR_c).
\end{aligned} \tag{S41}$$

The Fourier coefficients of the sLOSC kernel as implemented in our code are thus

$$K_s(G;R_c,\alpha)=\frac{4\pi}{G^2}\Bigg\{1-\cos(GR_c)\,\mathrm{erfc}(\alpha R_c)-e^{-G^2/4\alpha^2}+e^{-(\alpha R_c)^2}\times \left[\mathrm{Re}\left\{w\left(i\alpha R_c-\frac{G}{2\alpha}\right)\right\}\cos(GR_c)+\mathrm{Im}\left\{w\left(i\alpha R_c-\frac{G}{2\alpha}\right)\right\}\sin(GR_c)\right]\Bigg\}; \quad \text{(S42)}$$

$K_s(G=0)$, having only a real argument to the error function, is implemented as written in Eq. (S36).

S4. COMPUTATIONAL DETAILS

In any DFT calculation in periodic boundary conditions, several computational parameters must be selected. In this work, we use the PBE density functional [9]; set the kinetic energy cutoff of the plane-wave basis at 100 Ry; sample the Brillouin zone uniformly by 16×16×16 **k**-points for self-consistent calculations (yielding the Kohn–Sham density matrix), and 6×6×6 **k**-points for the other calculations (conduction Bloch bands, localization, and sLOSC); and use the norm-conserving optimized Vanderbilt pseudopotentials with scalar relativistic correction [10], sourced from the PseudoDojo [11].

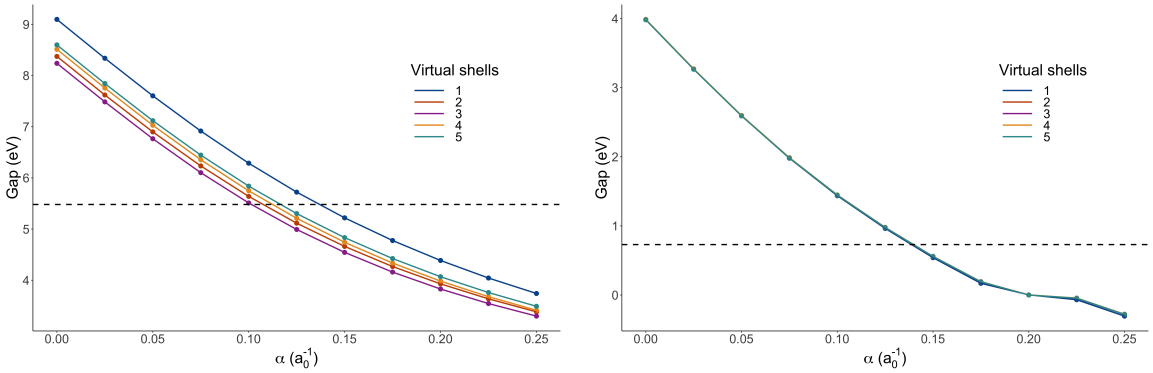
Localization is performed using a fork of the `wannier90` code [12–14] maintained by one of us (A.M.) [3]. We set the frozen (inner) window for disentanglement at 0.5 eV above the Fermi energy. For an initial guess, we use the selected columns of the density matrix (SCDM)[15, 16], with midpoint μ set at the top of the frozen window and spread $\sigma = 4.0$ eV. Worth noting is that some systems presented difficulty in converging to the minimal value of the cost function; we are not guaranteed a unique global minimum of our cost function, and the DLWFs can display imaginary components, especially in the virtual space [17]. We searched manually over a range of conjugate gradient step sizes to seek convergence, choosing the one that yielded the smallest total cost.

Several additional parameters must be set for sLOSC computations. First, the mixing parameter γ in the cost function, which controls the relative importance of spatial and energy localization of the DLWFs, must be chosen. In the main text, we use $\gamma = 0.47714$ to match the molecular LOSC results; see below for a discussion of the effects of varying γ. The exchange factor τ in the sLOSC curvature can also be modified; we choose the nonempirical value $\tau = 6(1-2^{-1/3})$, which enforces the condition $X[\rho_1] = 2X[\frac{1}{2}\rho_1]$ for any one-electron density ρ_1 (for a derivation of this fact, see the Supporting Information of [18]).

S4.1. The number of virtual orbitals

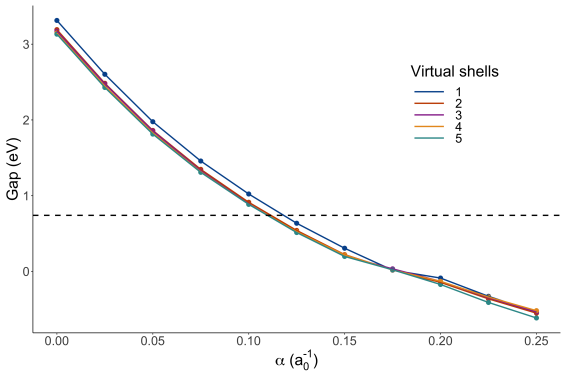
Because the DLWFs mix both occupied (valence) and unoccupied (conduction) states, it is necessary to include enough virtual bands in their construction that the Fermi level is reproduced accurately. To ensure the frontier orbitals are converged in the number of virtual states, we chose ten representative systems and perform sLOSC on them with $n_v = 2, 3, 4, 5, 6$ coordination shells of virtual Bloch bands. By "coordination shell", we mean the coordination number of the lattice; thus, for diamond, zincblende, and wurtzite structures we include $4n_v$ virtual , while for cubic structures we include $6n_v$ virtual bands. In all cases, we disentangle to one fewer coordination shell of transformed Bloch orbitals (and, hence, DLWFs).

In S1, we compare the fundamental gap calculated with sLOSC as a function of the Coulombic screening parameter α (for details on α, see the next subsection), for different numbers of virtual shells. We find that disentangling to two virtual shells is sufficient to converge the gap; the notable differences between different numbers of virtual shells come from qualitative differences in the localization procedure.

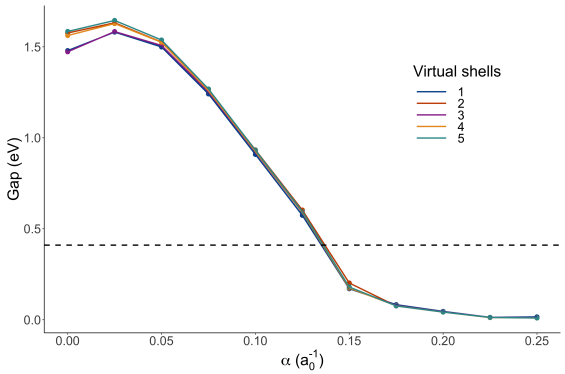


(a) C (diamond)

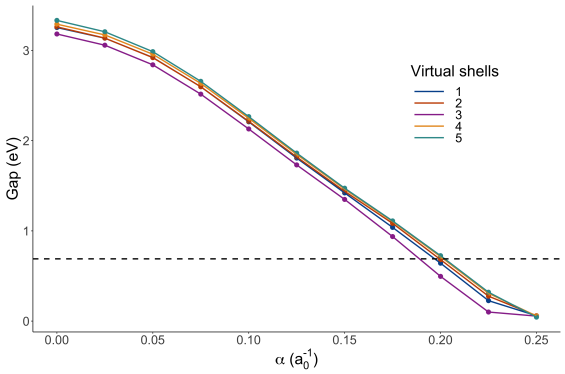
(b) GaSb (zincblende)



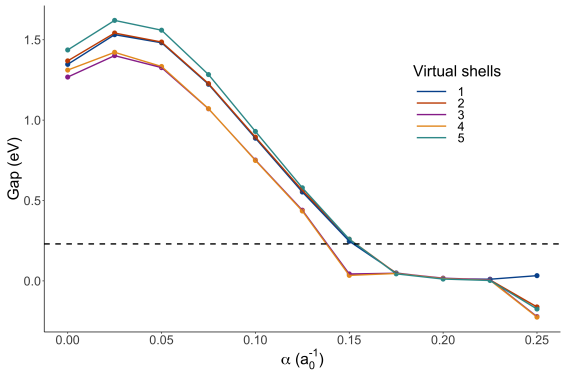
(c) Ge (diamond)



(d) InAs (zincblende)

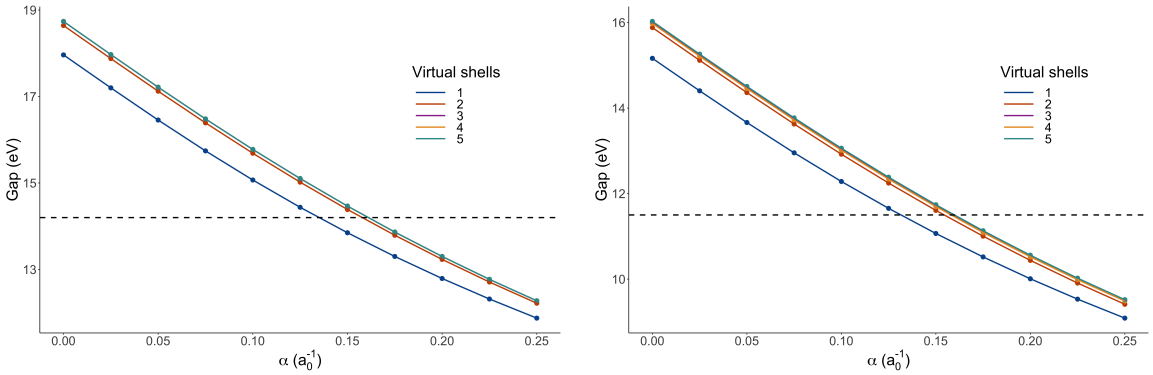


(e) InN (wurtzite)



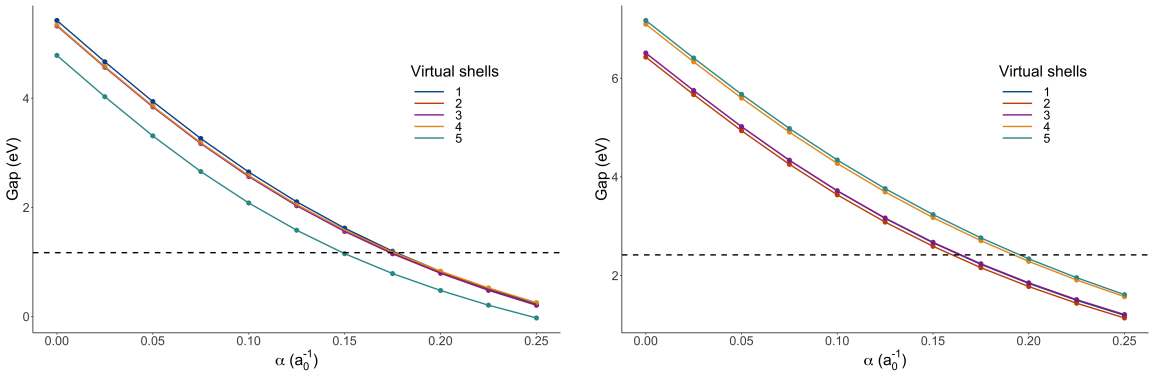
(f) InSb (zincblende)

Fig. S1. sLOSC band gaps as a function of screening parameter α for $n = 1, 2, 3, 4, 5$ virtual shells of DLWFs. The dashed line is the experimental gap. The subcaption title gives the chemical formula, with the lattice type in parentheses.



(g) LiF (cubic)

(h) NaF (cubic)



(i) Si (diamond)

(j) SiC (zincblende)

Fig. S1. sLOSC band gaps as a function of screening parameter α for $n = 1, 2, 3, 4, 5$ virtual shells of DLWFs. The dashed line is the experimental gap. The subcaption title gives the chemical formula, with the lattice type in parentheses.

S4.2. Choosing the screening parameter

The sLOSC method's utility in calculating the band gap of bulk materials comes at the cost of one additional empirical parameter α, which measures the strength of the screening. To choose α, we computed the band gaps of the full dataset below, varying α from 0 (unscreened) to $0.25\,\mathrm{a}_0^{-1}$ in steps of 0.025, where a_0 denotes the Bohr radius. We selected the value minimizing the mean absolute percentage error (MAPE) of the bulk systems' band gaps relative to experimentally determined values. As can be seen from Fig. S2, the minimal MAPE for bulk systems is attained at $\alpha = 0.15\,\mathrm{a}_0^{-1}$, so we choose this value for the screening in sLOSC.

We also show the MAPE for a test set of small molecules' fundamental gaps (ionization potential minus electron affinity), relative to CCSD(T) calculations. The smallest MAPE is given by LOSC2 [19], with $\alpha = 0$ (no screening); however, even with $\alpha = 0.15\,\mathrm{a}_0^{-1}$ as in sLOSC, the MAPE is substantially improved compared to the uncorrected PBE computation.

For the results of individual uncorrected and sLOSC computations, see Tables S2 and S3 in the section below.

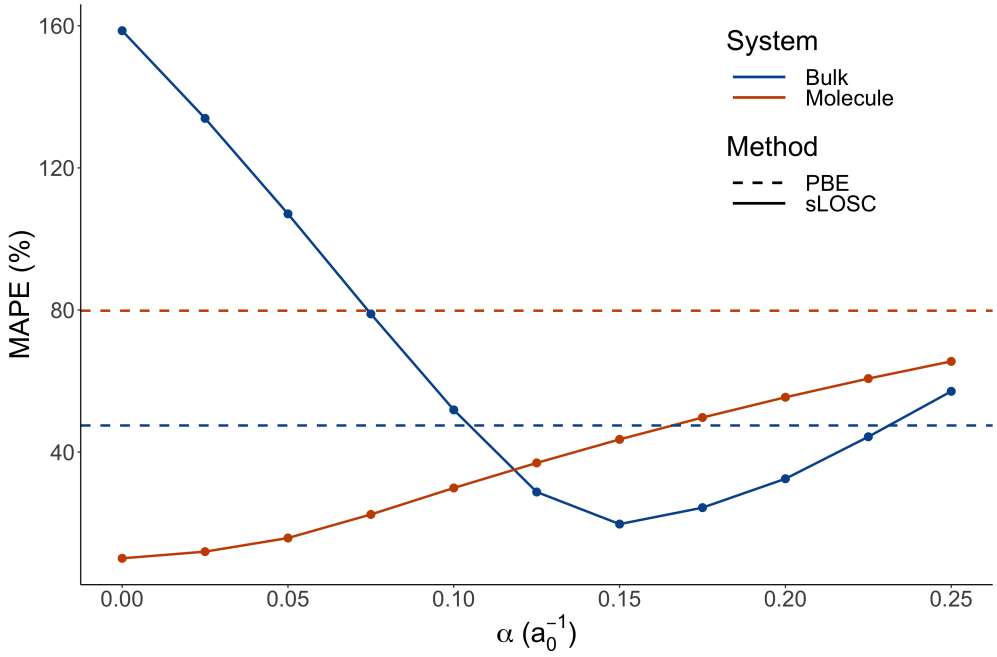


Fig. S2. MAPE of the band gap of bulk (orange) and molecular (blue) systems, as α is varied. The dashed horizontal lines give the MAPE of the uncorrected DFA (PBE).

S4.3. The DLWF space/energy mixing parameter

Recall that the DLWFs are obtained by minimizing the cost function

$$F = (1-\gamma)\sum_i \langle \Delta r^2 \rangle_i + \gamma C \sum_i \langle \Delta h_s^2 \rangle_i, \tag{S43}$$

where $\langle \Delta r^2 \rangle_i$ ($\langle \Delta h_s^2 \rangle_i$) is the spatial (energy) variance of DLWF i, $0 \leq \gamma \leq 1$, $C = 1\,\mathrm{a}_0^2/\mathrm{eV}^2$, and a_0 is the Bohr radius.

As recommended by one of the referees, we investigated the effect of varying γ, which controls the relative importance of spatial and energy localization, on a subset of materials in the dataset: GaSb, Ge, InAs, InN, InSb, NaF, Si, SiC. γ was varied between 0, the maximally localized Wannier functions (MLWFs) of Marzari and Vanderbilt [20], and 1, yielding pure Fourier transforms of the disentangled Bloch orbitals [2] (up to phases at each **k**-point). At each value of γ, the screening parameter α was scanned from $0.000\,\mathrm{a}_0^{-1}$ to $0.250\,\mathrm{a}_0^{-1}$. Observe from Fig. S3 that the optimal α required increases as γ decreases; in other words, more screening is required for accurate band gaps when the sLOSC correction is produced by more spatially localized orbitals. Note that setting $\gamma > 0.50$ resulted in DLWFs that were not well localized within the $6 \times 6 \times 6$ Born–von Karman supercell; they would require an even larger supercell to give accurate answers, which is computationally expensive.

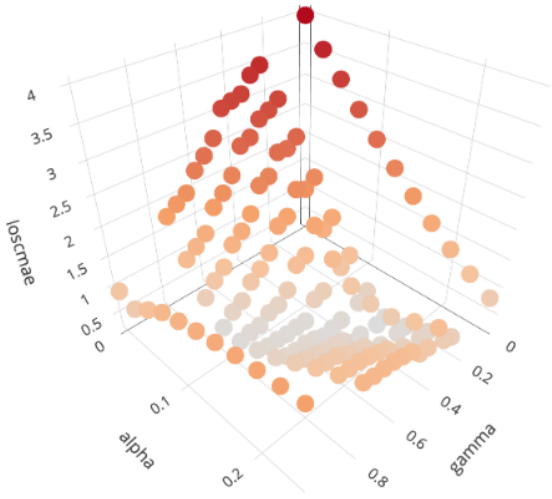


Fig. S3. The mean absolute error (MAE, eV) in the band gap as a function of γ and α. Orange points have larger MAE, while grey points have smaller MAE.

Therefore, we chose six combinations of $\gamma \leq 0.50$ and α with small MAE in the bulk subset and obtained the mean absolute percentage error (MAPE) of the molecular systems (those in Table S3 below) at those values. There is not an appreciable change compared to the γ and α in the main text, which is approximated by the $(0.50, 0.150)$ data point. Thus, we are satisfied that the value of γ in the main text is good for the sLOSC method.

Table S1: The MAPE in the fundamental gap of the molecular systems as γ and α are varied.

γ	α (a_0^{-1})	MAPE (%)
0.25	0.175	46.08
0.30	0.175	47.59
0.40	0.175	48.58
0.45	0.150	43.41
0.45	0.175	49.58
0.50	0.150	43.68

S5. SLOSC DATA

sLOSC is evaluated on a modified version of the SC/40 test set found in [21], with modifications as follows: Systems for which no experimental lattice constant could be found (BSb, CaTe, SrTe) were excluded; and the large-gap insulators LiF, LiCl, NaF, NaCl, Ar, Ne were added, for a total of 43 systems. Computations were carried out in the `Quantum ESPRESSO` code suite [22, 23], to which sLOSC is implemented as an add-on package, and a version of the `wannier90` code [12–14] modified for the computation of DLWFs [3].

For every system in which [21] lists an experimental lattice constant, we use it; the other experimental lattice constants are sourced from [24]. Experimental band gaps are also sourced from [21] whenever they are available; for the rest of the systems, the experimental gaps are sourced from [25] (LiF, LiCl, Ne, Ar); [26] (NaF); [27] (NaCl); and [28] (CaS, CaSe, SrS, SrSe).

Table S2: The data presented in Fig. 1 of the main text. The second column gives the Strukturbericht designation for each system; note that B4 (wurtzite) structures have two lattice constants.

System parameters			Band gaps (eV)			Energy (eV)	
Formula	Lattice	Constant (Å)	DFA	sLOSC	Exp.	E_{DFA}	$\Delta E_{\mathrm{sLOSC}}$
AlAs	B3	5.661	1.440	2.828	2.23	−2950.23	6.550×10^{-3}
AlN	B4	(a) 3.111	4.306	7.296	6.13	−696.28	4.290×10^{-3}
		(c) 4.981					
AlP	B3	5.463	1.580	2.968	2.51	−252.76	5.886×10^{-3}
AlSb	B3	6.136	1.221	2.032	1.68	−2538.75	1.074×10^{-2}
Ar	A1	5.256	8.703	12.017	14.20	−614.70	2.626×10^{-4}
BaS	B1	6.389	2.148	3.903	3.88	−1068.90	1.382×10^{-2}
BAs	B3	4.777	1.294	2.490	1.46	−2966.27	2.450×10^{-3}
BaSe	B1	6.595	1.922	3.437	3.58	−4067.49	1.465×10^{-2}
BaTe	B1	7.007	1.561	2.736	3.08	−3621.33	1.290×10^{-2}
BN	B3	3.616	4.533	6.012	6.22	−364.71	1.429×10^{-4}
BP	B3	4.538	1.348	2.053	2.40	−269.31	8.857×10^{-4}
C	A4	3.567	4.206	4.712	5.48	−327.57	8.299×10^{-5}
CaS	B1	5.689	2.358	5.139	4.43	−1358.08	7.687×10^{-3}
CaSe	B1	5.916	2.035	4.567	3.85	−4356.48	1.220×10^{-2}
CdS	B3	5.818	1.163	2.291	2.55	−5072.67	1.416×10^{-2}
CdSe	B3	6.052	0.639	1.571	1.90	−8071.28	1.732×10^{-2}
CdTe	B3	6.480	0.773	1.414	1.92	−7625.31	1.835×10^{-2}
GaAs	B3	5.648	0.538	1.224	1.52	−4967.79	8.015×10^{-3}
GaN	B4	(a) 3.189	1.941	4.177	3.50	−4728.22	7.120×10^{-3}
		(c) 5.185					
β-GaN	B3	4.523	1.689	4.007	3.30	−2364.11	3.380×10^{-3}
GaP	B3	5.451	1.658	2.664	2.35	−2270.19	9.008×10^{-3}
GaSb	B3	6.096	0.116	0.560	0.73	−4556.57	7.704×10^{-3}

System parameters			Band gaps (eV)			Energy (eV)	
Formula	Lattice	Constant (Å)	DFA	sLOSC	Exp.	E_{DFA}	$\Delta E_{\mathrm{sLOSC}}$
Ge	A4	5.658	0.031	0.222	0.74	−4861.09	3.093×10^{-3}
InAs	B3	6.058	0.000	0.202	0.41	−4694.91	1.463×10^{-2}
InN	B4	(a) 3.537	0.000	1.440	0.69	−4180.63	5.138×10^{-2}
		(c) 5.704					
InP	B3	5.869	0.689	1.402	1.42	−1997.15	9.919×10^{-3}
InSb	B3	6.479	0.000	0.260	0.23	−4283.83	1.943×10^{-2}
LiCl	B1	5.130	6.328	8.877	9.40	−645.92	9.402×10^{-4}
LiF	B1	4.017	9.186	14.387	14.20	−870.52	1.918×10^{-4}
MgO	B1	4.207	4.803	8.623	7.22	−2069.45	6.925×10^{-4}
MgS	B3	5.622	3.558	5.273	5.40	−1919.41	4.867×10^{-3}
MgSe	B1	5.400	1.863	3.235	2.47	−4917.70	5.720×10^{-3}
MgTe	B3	6.420	2.499	3.773	3.60	−4471.36	6.453×10^{-3}
NaCl	B1	5.641	5.107	7.419	8.97	−1693.81	1.595×10^{-3}
NaF	B1	4.620	6.389	11.610	11.50	−1917.95	5.157×10^{-4}
Ne	A1	4.429	11.617	18.101	21.70	−907.70	6.803×10^{-5}
Si	A4	5.430	0.709	1.572	1.17	−230.28	6.914×10^{-3}
SiC	B3	4.358	1.363	1.962	2.42	−279.46	1.000×10^{-3}
SrS	B1	5.990	2.455	4.434	4.32	−1236.19	6.467×10^{-3}
SrSe	B1	6.234	2.192	3.953	3.81	−4234.69	7.301×10^{-3}
ZnS	B3	5.409	2.101	3.300	3.66	−5990.28	6.740×10^{-3}
ZnSe	B3	5.668	1.289	2.301	2.70	−8988.74	9.112×10^{-3}
ZnTe	B3	6.089	1.282	1.978	2.38	−8542.59	1.592×10^{-2}

We additionally compute sLOSC results on a small set of molecular systems using the in-house QM^{4}D code. We use the PBE exchange-correlation functional, as in the bulk case, and the 6-311++G(3df,3pd) basis set, with the Dunning augmented correlation-consistent triple-zeta (aug-cc-pVTZ) basis used for density fitting.

Instead of experimental values, we use CCSD(T) computations for the reference fundamental gaps. These are computed under both the aug-cc-pVTZ and aug-cc-pVQZ (Dunning augmented correlation-consistent triple- and quadruple-zeta) basis sets and extrapolated

using the formula from Eq. (44) of [29] as

$$\epsilon_\infty = \frac{\epsilon_{\mathrm{TZ}} \times 3^3 - \epsilon_{\mathrm{QZ}} \times 4^3}{3^3 - 4^3}, \tag{S44}$$

where ϵ_{nZ} is the aug-cc-pVnZ eigenvalue.

Table S3: The molecular sLOSC data.

	Band gaps (eV)			Energy (eV)	
Formula	DFA	sLOSC	Exp.	E_{DFA}	$\Delta E_{\mathrm{sLOSC}}$
SH	0.430	3.676	8.05	−10845.99	2.346×10^{-4}
O_2	2.306	6.959	12.51	−4088.33	2.072×10^{-5}
S_2	1.292	2.497	7.92	−21662.11	2.906×10^{-4}
Cl_2	2.740	4.395	10.61	−25035.22	1.035×10^{-4}
CF_2	3.626	7.544	12.68	−6464.52	5.676×10^{-4}
CH_2	2.413	6.845	9.57	−1064.09	9.100×10^{-4}
CH_3	2.472	6.531	9.84	−1082.70	1.029×10^{-3}
CN	1.709	6.125	10.29	−2520.88	2.064×10^{-4}
CHO	1.693	4.568	9.44	−3095.93	1.407×10^{-3}
CH_3O	0.723	6.102	9.51	−3128.24	2.093×10^{-3}
CH_2S	1.891	4.067	9.16	−11898.67	9.618×10^{-4}
CH_2SH	1.567	3.743	5.91	−11914.33	9.567×10^{-4}
OH	0.916	8.599	11.24	−2059.42	1.223×10^{-4}
NH	3.597	9.751	13.15	−1501.30	3.693×10^{-4}
NH_2	2.738	8.527	11.37	−1519.24	1.025×10^{-3}
SiH_2	1.813	3.671	8.48	−7903.86	4.788×10^{-4}
SiH_3	1.828	3.924	7.93	−7920.69	1.044×10^{-3}
PH	2.213	4.728	9.16	−9298.83	1.938×10^{-4}
PH_2	1.844	4.349	8.56	−9315.89	7.920×10^{-4}

We plot the data from Table S3 in Fig. S4, together with PBE and unscreened LOSC results, in analogy with Fig. 1 in the main text. We can see again that the best gaps are uniformly those computed without screening, in agrement with Fig. S2 and as expected for

small molecules with open boundary conditions; however, sLOSC still improves the accuracy of the gap compared to the PBE calculation.

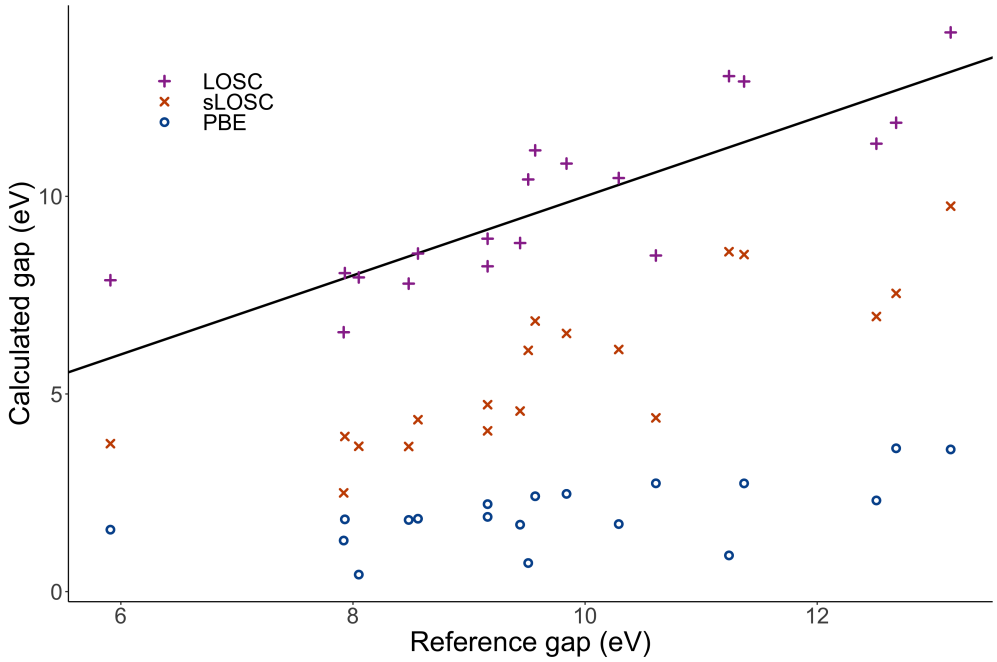


Fig. S4. Comparison of experimental fundamental gaps of molecular systems with those calculated by PBE (o), sLOSC (×), and unscreened LOSC (+).

[1] W. Wirtinger, Math. Ann. **97**, 357 (1927).

[2] I. Souza, N. Marzari, and D. Vanderbilt, Phys. Rev. B **65**, 035109 (2001), publisher: American Physical Society.

[3] A. Mahler, J. Z. Williams, N. Q. Su, and W. Yang, arXiv:2201.07751 [cond-mat] (2022), arXiv: 2201.07751.

[4] G. Onida, L. Reining, R. W. Godby, R. Del Sole, and W. Andreoni, Phys. Rev. Lett. **75**, 818 (1995), publisher: American Physical Society.

[5] M. R. Jarvis, I. D. White, R. W. Godby, and M. C. Payne, Phys. Rev. B **56**, 14972 (1997), publisher: American Physical Society.

[6] Wolfram Research, Inc., Mathematica, Version 12.3.1 (2021).

[7] M. R. Zaghloul and A. N. Ali, ACM Trans. Math. Softw. **38**, 15:1 (2012).

[8] M. R. Zaghloul, ACM Trans. Math. Softw. **42**, 26:1 (2016).

[9] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[10] D. R. Hamann, Phys. Rev. B **88**, 085117 (2013).

[11] M. J. van Setten, M. Giantomassi, E. Bousquet, M. J. Verstraete, D. R. Hamann, X. Gonze, and G. M. Rignanese, Computer Physics Communications **226**, 39 (2018).

[12] A. A. Mostofi, J. R. Yates, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari, Computer Physics Communications **178**, 685 (2008).

[13] A. A. Mostofi, J. R. Yates, G. Pizzi, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari, Computer Physics Communications **185**, 2309 (2014).

[14] G. Pizzi, V. Vitale, R. Arita, S. Blügel, F. Freimuth, G. Géranton, M. Gibertini, D. Gresch, C. Johnson, T. Koretsune, J. Ibañez-Azpiroz, H. Lee, J.-M. Lihm, D. Marchand, A. Marrazzo, Y. Mokrousov, J. I. Mustafa, Y. Nohara, Y. Nomura, L. Paulatto, S. Poncé, T. Ponweiser, J. Qiao, F. Thöle, S. S. Tsirkin, M. Wierzbowska, N. Marzari, D. Vanderbilt, I. Souza, A. A. Mostofi, and J. R. Yates, J. Phys.: Condens. Matter **32**, 165902 (2020).

[15] A. Damle, L. Lin, and L. Ying, J. Chem. Theory Comput. **11**, 1463 (2015).

[16] A. Damle, L. Lin, and L. Ying, Journal of Computational Physics **334**, 1 (2017).

[17] C. Brouder, G. Panati, M. Calandra, C. Mourougane, and N. Marzari, Phys. Rev. Lett. **98**, 046402 (2007), publisher: American Physical Society.

[18] C. Li, X. Zheng, N. Q. Su, and W. Yang, National Science Review **5**, 203 (2018).

[19] N. Q. Su, A. Mahler, and W. Yang, J. Phys. Chem. Lett. **11**, 1528 (2020).

[20] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997), publisher: American Physical Society.

[21] J. Heyd, J. E. Peralta, G. E. Scuseria, and R. L. Martin, J. Chem. Phys. **123**, 174101 (2005), publisher: American Institute of Physics.

[22] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. d. Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, J. Phys.: Condens. Matter **21**, 395502 (2009), publisher: IOP Publishing.

[23] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. D. Corso, S. d. Gironcoli, P. Delugas, R. A. DiStasio, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. Otero-de-la Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, and S. Baroni, J. Phys.: Condens. Matter **29**, 465901 (2017), publisher: IOP Publishing.

[24] R. W. G. Wyckoff, *Crystal Structures*, 2nd ed. (Wiley, New York, 1973).

[25] F. Tran and P. Blaha, Phys. Rev. Lett. **102**, 226401 (2009), publisher: American Physical Society.

[26] R. T. Poole, J. G. Jenkin, J. Liesegang, and R. C. G. Leckey, Phys. Rev. B **11**, 5179 (1975), publisher: American Physical Society.

[27] D. M. Roessler and W. C. Walker, Phys. Rev. **166**, 599 (1968), publisher: American Physical Society.

[28] Y. Kaneko and T. Koda, Journal of Crystal Growth **86**, 72 (1988).

[29] Y.-S. Lin, C.-W. Tsai, G.-D. Li, and J.-D. Chai, J. Chem. Phys. **136**, 154109 (2012), publisher: American Institute of Physics.