Supplemental Material for:

First-principles ionized impurity scattering and charge transport in doped materials

I-Te Lu, Jin-Jian Zhou, Jinsoo Park, and Marco Bernardi*

* Department of Applied Physics and Materials Science, California Institute of Technology, Pasadena, California 91125
I. ADDITIONAL COMPUTATIONAL DETAILS

We carry out plane-wave DFT calculations, using the Quantum ESPRESSO (QE) code [1], to obtain the ground state electronic structure of a primitive unit cell of Si as well as pristine and defect-containing 512-atom Si supercells. The calculations employ the PBE exchange-correlation functional [2] together with norm-conserving pseudopotentials [3]; spin-orbit coupling is not included. We use plane-wave kinetic energy cutoffs of 40 Ry and 70 Ry, respectively, for calculations on P-doped and B-doped Si. The self-consistent step uses a uniform $12 \times 12 \times 12$ $k$-point grid, which converges the total energy to within 10 meV/atom. To obtain maximally localized WFs for the primitive unit cell, we use the Wannier90 code [4, 5], starting with four $sp^3$ orbitals on each Si atom as the initial projections, together with KS wave functions and eigenvalues obtained from a non-self-consistent DFT calculation using an $8 \times 8 \times 8$ $k$-point grid (referred to as the “coarse” grid). The resulting electronic Hamiltonian in the Wannier basis can accurately interpolate the DFT electronic band structure. Lattice dynamical properties and the $e$-$ph$ perturbation potentials are computed on the primitive unit cell with an $8 \times 8 \times 8$ $q$-point grid, using density functional perturbation theory [6] as implemented in the QE code [1].

The supercells employed in this study are constructed by repeating the primitive cell in the directions of the primitive lattice vectors $\mathbf{a}_i$ (where $i = 1, 2, 3$). The supercell sizes are denoted as $N_1 \times N_2 \times N_3$, where $N_i$ is the number of primitive cells along the direction of the lattice vector $\mathbf{a}_i$. For P-doped Si, we use a $6 \times 6 \times 6$ supercell and replace one Si atom at the center of the supercell with one P atom. In the self-consistent DFT calculation for the ionized P impurity, namely the charged defect $P^+$, we remove an electron from the supercell and set the $G_{\text{sup}} = 0$ component of the Hartree potential to zero (using a feature implemented in QE). The BZ is sampled only at the $\Gamma$ point to obtain the KS potential. The atomic forces are relaxed to within 50 meV/Å to account for the structural changes induced by the charged defect. To save computational cost, we relax only the atoms within the 4th nearest-neighbor shell of the charged defect, and fix the position of the atoms outside that region. We apply a similar approach for the the ionized B impurity, i.e. the charged defect $B^-$, in the B-doped Si supercell. The $e$-$d$ perturbation potentials for the $P^+$ and $B^-$ charged defects are obtained using the method described in the main text, which accounts for potential alignment and removal of spurious image-supercell contributions.

In the screened Coulomb potential, we use the model dielectric function $\epsilon(q)$ from Ref. 7,
which can reproduce the *ab initio* dielectric function of Si and other semiconductors within the random-phase approximation [7]:

\[
\epsilon(q) = 1 + \left\{ [\epsilon(0) - 1]^{-1} + 1.536 \times \frac{(q/q_{\text{TF}})^2}{(q/q_{\text{TF}})^2 + \hbar^2 q^4/(4m_e^2 \omega_p^2)} \right\}^{-1},
\]

where \(\epsilon(0)\) is the dielectric constant, taken to be 11.3 for Si, while \(q_{\text{TF}}\) is the Thomas-Fermi (TF) wave vector and \(\omega_p\) is the plasma frequency. The TF wave vector and plasma frequency are computed using the valence electron density [7].

The potential alignment, \(V_{\text{align}}\) in Eq. (4) of the main text, is determined by averaging the potential difference \(\Delta V_{\text{KS}} - V_{\text{scr}}^{\text{np}}\) in a spherical shell of radius 14–16 Å around the defect. We use a smaller sphere with an 8 Å radius for the neutral part of the defect perturbation potential, \(\Delta V_{\text{c-d}}^{\text{np}}\), which is short-ranged and vanishes outside that sphere. The \(e\)-\(d\) matrix elements on the coarse grid (size 8 × 8 × 8 \(k_c\)-points) are computed using the KS Bloch wave functions from the primitive unit cell calculation [8] and the defect perturbation potentials of the charged defects. These coarse-grid \(e\)-\(d\) matrix elements in the Bloch basis are used in the WF-interpolation procedure described below in Sec. II.

The \(e\)-\(ph\) matrix elements, first computed on coarse 8 × 8 × 8 \(k_c\)- and \(q_c\)-point grids, are interpolated to fine BZ grids using the PERTURBO code [9]. The RTs and MFPs shown in Figs. 1–2 of the main text are computed using very fine random grids: for the \(e\)-\(d\) interactions, we use 16 M (M: million) \(k_f\)-points drawn from a Cauchy distribution, and for the \(e\)-\(ph\) interactions we use a uniformly distributed random \(q_f\)-grid with 8 M points. The carrier mobilities are computed in PERTURBO using energy windows spanning 200 meV above the CBM for electrons, and below the VBM for holes. For the iterative approach (ITA), we use a uniform 200 × 200 × 200 \(k_f\)-grid for both \(e\)-\(ph\) and \(e\)-\(d\) interaction. For the relaxation times approximation (RTA), we use a random \(k_f\)-grid with 4 M points generated from a Cauchy distribution for the \(e\)-\(d\) interactions, and a uniform random \(q_f\)-grid with 8M points for the \(e\)-\(ph\) interactions. The energy broadening of the Dirac delta functions is 2 meV. The Fermi energy in each calculation is determined from the doping concentration [9]. The dopants are assumed to be completely ionized except for the temperature-dependent mobility calculation on P-doped Si in Fig. 4. For the partially ionized case, the inverse screening length, ionized dopant concentration, and DOS of the impurity band are computed in a self-consistent way as discussed below in Sec. III.
II. WANNIER INTERPOLATION OF THE E-D MATRIX ELEMENTS

The WF interpolation of the charge-neutral part of the e-d matrix elements uses the approach developed in our previous work [10]. Briefly, the maximally localized WFs $|j \vec{R}\rangle$, with index $j$ at lattice point $\vec{R}$, can be constructed from the Bloch states on a coarse BZ $\vec{k}_c$-grid using the Wannier unitary matrices $U(\vec{k}_c)$ [11]:

$$|j \vec{R}\rangle = \frac{1}{N_{k_c}} \sum_{n_{k_c}} e^{-i\vec{k} \cdot \vec{R}} U_{nj}(\vec{k}_c) |n_{k_c}\rangle,$$  \hspace{2cm} (S2)

where $N_{k_c}$ is the number of $\vec{k}$-points in the coarse $\vec{k}_c$-grid. Conversely, a Bloch state $|n \vec{k}\rangle$ can be constructed using the inverse transformation:

$$|n \vec{k}\rangle = \sum_{j \vec{R}} e^{i\vec{k} \cdot \vec{R}} U_{jn}(\vec{k}) |j \vec{R}\rangle.$$  \hspace{2cm} (S3)

The e-d interaction matrix elements in the Wannier basis are defined as

$$M_{ij}(\vec{R}', \vec{R}) = \langle i \vec{R}'|\Delta V_{e-d}|j \vec{R}\rangle.$$  \hspace{2cm} (S4)

Using Eqs. (S2) – (S4), we can write them in terms of the coarse grid e-d matrix elements in the Bloch basis, $M(\vec{k}_c', \vec{k}_c)$, which are computed directly as explained in the main text. The resulting relation between the Bloch- and Wannier-basis matrix elements is a generalized double Fourier transform:

$$M(\vec{R}', \vec{R}) = \left(\frac{1}{N_{k_c}}\right)^2 \sum_{\vec{k}_c'} e^{i(\vec{k}_c' \cdot \vec{R}' - \vec{k}_c \cdot \vec{R})}$$

$$\times U^\dagger(\vec{k}_c') M(\vec{k}_c', \vec{k}_c) U(\vec{k}_c).$$  \hspace{2cm} (S5)

Here and below, we omit the band indices. By performing the inverse transformation, we can obtain the e-d matrix elements in the Bloch representation for any arbitrary pair ($\vec{k}_f', \vec{k}_f$) of fine-grid BZ crystal momenta:

$$M(\vec{k}_f', \vec{k}_f) = \sum_{\vec{R}', \vec{R}} e^{-i(\vec{k}_f' \cdot \vec{R}' - \vec{k}_f \cdot \vec{R})} U(\vec{k}_f') M(\vec{R}', \vec{R}) U^\dagger(\vec{k}_f).$$  \hspace{2cm} (S6)

The unitary matrix $U(\vec{k}_f)$ on a fine $\vec{k}_f$-grid is obtained by diagonalizing the electronic Hamiltonian [12]

$$H(\vec{k}_f) = \sum_{\vec{R}} H(\vec{R}) e^{i\vec{k}_f \cdot \vec{R}},$$  \hspace{2cm} (S7)
where $H(R)$ is the electronic Hamiltonian in the Wannier basis. For charge-neutral defects, the success of the above interpolation scheme relies on the spatial decay of the $e$-$d$ matrix elements in the Wannier basis [10]. For charged defects, this interpolation scheme can be applied directly only to the short-ranged, charge-neutral part of the $e$-$d$ matrix elements, $M^{\text{sp}}(k_c', k_c)$ in Eq. (8) of the main text. However, without a special treatment of the long-range part of the $e$-$d$ matrix elements, $M^{\text{src}}(k_c', k_c)$ in Eq. (8) of the main text, the interpolation method would fail for charged defects. To address this issue, we define the long-range contribution of the screened Coulomb potential to the $e$-$d$ matrix elements as

$$M^{\text{src,LR}}(k', k) = \frac{1}{\Omega_{\text{uc}}} \sum_{G} \tilde{V}_{\text{ex}}^{\text{src}}(k' - k + G) U(k') U^\dagger(k).$$

(S8)

Similar to the approach used in the literature to interpolate long-range $e$-$ph$ interactions [9], this long-range contribution is first removed from the $e$-$d$ matrix elements on the coarse $k_c$-grid; the resulting, short-range part of the $e$-$d$ matrix elements,

$$M^{\text{src,SR}}(k_c', k_c) = M^{\text{src}}(k_c', k_c) - M^{\text{src,LR}}(k_c', k_c),$$

(S9)

is then interpolated using Eq. (S5), and finally the long-range part $M^{\text{src,LR}}(k_f', k_f)$ is computed and added back on the fine-grid.

### III. PARTIAL DOPANT IONIZATION

Following Ref. 13, we write the Coulomb potential fluctuations due to the ionized dopants as

$$\langle U^2 \rangle = \left[ e^2 / \epsilon(0) \right]^2 \left[ (N^+_D + N^-_A) / 8\pi \right] q_0^{-1},$$

(S10)

where $N^+_D$ and $N^-_A$ are the concentrations of the ionized donors and acceptors, respectively, $e$ is the electron charge, and $\epsilon(0)$ is the dielectric constant. We assume that no compensation occurs in the doped semiconductor, so $N^+_D = n$, where $n$ is the electron concentration in the conduction band, or $N^-_A = p$, where $p$ is the hole concentration in the valence band. For electrons, the inverse screening length $q_0$, which accounts for screening from both the valence electrons around the impurity and the free carriers, reads:

$$q_0 = \left[ \frac{e^2}{\epsilon_0 \epsilon(0)} \int dE \left[ \rho(E) + g(E) \right] \left( -\frac{\partial f^0}{\partial E} \right) \right]^{1/2},$$

(S11)
where $g(E)$ is the DOS of the conduction band at energy $E$. The DOS of the donor level, $\rho(E)$, which is broadened into a band due to the Coulomb potential fluctuation, can be written as

$$\rho(E) = \left[ N_D / (2\pi\langle U^2 \rangle)^{1/2} \right] \exp \left[ -\frac{(E - E_D)^2}{2\langle U^2 \rangle} \right],$$

(S12)

where $N_D$ is the donor concentration and $E_D$ is the energy level of the donor. The potential fluctuation $\langle U^2 \rangle$, the inverse screening length $q_0$, and the DOS of the donor $\rho(E)$ are obtained by solving Eqs. (S10)–(S12) self-consistently for a given value of the donor energy level $E_D$. Here, the donor energy level is set to 45 meV below the CBM for the P impurity in Si [14].

The Fermi level $E_F$ used in the Fermi-Dirac distribution $f^0(E, E_F)$ for the partially ionized calculations is also obtained self-consistently by solving

$$N_D = \int dE \left[ \rho(E) + g(E) \right] f^0(E, E_F).$$

(S13)

The computed Fermi energy is then employed in our mobility calculations. The inverse screening length $q_0$ is used as the value of $q_{scr}$ in the screened Coulomb potential due to the isolated point charge (see Eq. (6) in the main text). The concentration of the ionized donors, $N_D$, is converted to a dimensionless concentration $C_{imp}$ and used in the $e$-$d$ scattering rate and transport calculations. This framework was employed to obtain the P-doped mobility with partial dopant ionization shown in Fig. 4 of the main text.


