Describing Migdal effects in diamond crystal with atom-centered localized Wannier functions

Zheng-Liang Liang,1,∗ Lin Zhang,2,† Fawei Zheng,3,‡ and Ping Zhang3,§

1College of Science, Beijing University of Chemical Technology
Beijing 100029, China
2School of Nuclear Science and Technology, University of Chinese Academy of Sciences,
Beijing, 100049, China
3Institute of Applied Physics and Computational Mathematics
Beijing, 100088, China

Abstract
Recent studies have theoretically investigated the atomic excitation and ionization induced by the dark matter (DM)-nucleus scattering, and it is found that the suddenly recoiled atom is much more likely to excite or lose its electrons than expected. Such phenomenon is called the “Migdal effect”. In this paper, we extend the established strategy to describe the Migdal effect in isolated atoms to the case in semiconductors under the framework of tight-binding (TB) approximation. Since the localized aspects of electrons are respected in form of the Wannier functions (WFs), the extension of the existing Migdal approach for isolated atoms is much more natural, while the extensive nature of electrons in solids is reflected in the hopping integrals. We take diamond target as a concrete proof of principle for the methodology, and calculate relevant energy spectra and projected sensitivity of such diamond detector. It turns out that our method as a preliminary attempt is practically effective.

∗ liangzl@mail.buct.edu.cn
† zhanglin@ucas.ac.cn
‡ zheng_fawei@iapcm.ac.cn
§ zhang_ping@iapcm.ac.cn
1. Introduction

The identity of the dark matter (DM) is one of the most puzzling problems in modern physics. Although there has been overwhelming evidence for its existence from astrophysics and cosmology, its nature still remains a mystery from a particle physical perspective. In decades, tremendous efforts are invested into the search for the weakly interacting massive particles (WIMPs), which not only take root in mere theoretical motivations (e.g., supersymmetry), but also naturally explain the observed relic abundance in the context of thermal freeze-out. Owing to the spectacular improvements in sensitivity over recent years, the frontier of the detection has been pushed to the DM mass range around the sub-GeV scale, where traditional experiments (e.g., XENON1T [1, 2], LUX [3], PandaX [4], etc. [5–9]) are expected to turn insensitive. This has motivated various proposals based on novel detection channels with novel targets [10–43].

However, a recent research [44] clarified that the sensitivity of these conventional strategies have been significantly underestimated. Against previous impression that electrons are so tightly bound to the atom that the sudden boost of a recoiled atom cannot “shake off” the outer electrons, the authors of Ref. [44] pointed out that in realistic case it takes some time for the electrons to catch up with the struck atom, and excitation and ionization are found to be more frequent than anticipated. Such phenomenon is termed “Migdal effect” in the DM literature.

The purpose of this study is to extend the formalism developed for isolated atoms in Ref. [44] to the case of semiconductors. Unlike the electrons exclusively bound to an individual atom, the delocalized electrons in solids are free to hop between neighboring ions, which makes a direct application of the method developed in Ref. [44] to the crystalline environment rather dubious. On the one hand, if one follows the suddenly recoiled nucleus in solids, the rest of the ions in solids will no longer be stationary either. On the other hand, if one follows the whole recoiled crystal, the highly local impulsive effect caused by the incident DM particle cannot be appropriately accounted for. To pursue a reasonable extension, we resort to the tight-binding (TB) approximation in which both the local and extensive characteristics of the Migdal effect in crystal are taken into consideration simultaneously. To be specific, we first re-express the Bloch wavefunctions in terms of the Wannier functions (WFs) that reflect the localized aspects of the itinerant electrons, and then we impose the Galilean transformation operator exclusively onto those WFs associated with the struck atom, and as a consequence the extensive aspects of electrons are effectively encoded in the hopping integrals. As a natural representation of localized orbital for extended systems, WFs play a key in various applications related to local phenomena, such as defects, excitons, electronic polarization and magnetization, as well as in formal discussions of the Hubbard models of strongly correlated systems [45, 46]. Dealing with the Migdal effect in crystalline solids is another interesting application of the WFs.

In discussion we take diamond crystal as a concrete example to demonstrate the feasibility of the TB approach to describe the Migdal effect in solid detectors. For one thing, diamond itself is a promising material for DM detection [47], possessing numerous advantages over traditional silicon and germanium semiconductor detectors, such as the lighter mass of carbon nucleus that brings about a lower DM mass threshold, the long-lived and hard phonon modes that facilitate the phonon collection, and the ability to withstand strong electric fields that drive the ionized electrons across the bulk material, etc. For another thing, there is a technical reason to choose crystalline diamond rather than silicon and germanium for the proof of concept of our method, that is, it is easier for diamond to generate localized atom-centered WFs across an energy range up to $O(10\,\text{eV})$ in practical use of the Wannier 90 package [48], owing to its wider band width. We will return to this in Sec. 4.

To deal with the electronic excitation rate via recoiled ions, the quantization of vibration seems to be an alternative approach, since related studies and applicable tools have existed for long in areas such as neutron scattering in solid state (for a review, see [49]), and recently in DM detection [19, 27, 33, 50, 51]. However, taking a look at the Feynman rules of the phonon excitation process, one finds that each phonon external leg contributes a term proportional to $\epsilon_{q,\alpha} \cdot q/\sqrt{m_N \omega_{q,\alpha}}$, with $\epsilon_{q,\alpha}$ and $\omega_{q,\alpha}$ the phonon eigenvector and the eigen-frequency of branch $\alpha$ at momentum transfer $q$, respectively, and $m_N$ the mass of nucleus. Therefore, in the DM mass range above a few MeV, multi-phonon effects are no longer negligible, and if a large momentum transfer is involved, all kinetically possible processes have to be taken into account, making the problem even more intractable. However, it is found through an isotropic
harmonic oscillator toy model that in the limit \( q = |q| \to \infty \), the effects of all the multi-phonon terms can be well summarized with the impulse approximation [49], where all the vibrational effects are encoded in a free recoiled atom in a very short timescale. It is during this period of time, excitation occurs. So, in the DM mass range of sub-GeV, the impulse approximation, or the nuclear recoil interpretation, is still the appropriate approach to depict the Migdal effect in solids.

This paper is organized as follows. In Sec. 2, we first take a brief review of the Migdal effect in atoms, and then outline the TB framework to describe the Migdal effect in crystalline solids. In Sec. 3, we put into practice the TB approach by use of \textit{ab initio} density functional theory (DFT) code \textit{Quantum Espresso} [52] and WF-generation tool \textit{Wannier 90}, concretely calculating the Migdal excitation event rate and relevant energy spectrum for crystalline diamond. Conclusion and open discussions are arranged in Sec. 4.

Throughout the paper the natural units \( \hbar = c = 1 \) is adopted, while velocities are expressed with units of km/s in text for convenience.

2. Electronic excitation in the tight-binding description

We begin this section with a short review of the treatment of the Migdal effect in isolated atoms, and then generalize its application to the electronic bands in crystalline solids.

2.1. Migdal effect in isolated atoms

In an atom, the excitation/ionization of electrons can be reasonably estimated by using the Migdal’s approach [44], in which the excitation/ionization is a dynamical consequence of suddenly moving electrons in the rest frame of a recoiled nucleus. To account for the Migdal effect in the rest frame of the struck nucleus we invoke the Galilean boost operator \( e^{im_e \cdot r} \). For the given velocity \( \mathbf{v} \), the operator \( e^{im_e \cdot r} \) boosts the electron state at rest \( (m_e \text{ being the electron mass and } \hat{r} \text{ the electron position operator}) \) to the inertial frame moving with velocity \( \mathbf{v} \). To see this, assuming \( |\mathbf{p}\rangle \) is the eigenstate of the momentum operator \( \hat{p} \) with eigen momentum \( \mathbf{p} \), it is straightforward to verify that

\[
\hat{p} \left( e^{im_e \cdot r} |\mathbf{p}\rangle \right) = (\mathbf{p} + m_e \mathbf{v}) \left( e^{im_e \cdot r} |\mathbf{p}\rangle \right) . \tag{2.1}
\]

Thus, keeping pace with the struck nucleus, one can schematically express the excitation/ionization probability as

\[
P \propto |\langle \psi_2 | e^{i\mathbf{q} \cdot \hat{r}} |\psi_1 \rangle|^2 , \tag{2.2}
\]

where \( |\psi_1\rangle \) and \( |\psi_2\rangle \) represent the initial and final states, respectively, sandwiching the Galilean transformation operator \( e^{i\mathbf{q} \cdot \hat{r}} \) that boosts the bound electron in the opposite direction to the recoiled nucleus with Galilean momentum \( \mathbf{q} \equiv (m_e/m_N) \mathbf{q} \), with the nucleus mass \( m_N \), and the DM transferred momentum \( \mathbf{q} = \mathbf{p}_{\chi,f} - \mathbf{p}_{\chi,i} \), with \( \mathbf{p}_{\chi,i} \) and \( \mathbf{p}_{\chi,f} \) being the DM momenta before and after the scattering respectively. To compute the excitation/ionization rate, one also needs to integrate over momentum transfer \( \mathbf{q} \) and DM velocity \( \mathbf{w} \). As a result, the transition event rate for a DM particle with incident velocity \( \mathbf{w} \) to excite a bound electron via the Migdal process from level 1 to level 2 can be expressed in the following form*:

\[
\mathcal{R}_{1 \to 2} = \frac{\rho_\chi}{m_\chi} \langle \sigma_{1 \to 2} w \rangle = \frac{\rho_\chi}{m_\chi} \left( \frac{A^2 \sigma_{\chi n}}{4 \pi \mu_{\chi n}} \right) \int dq \int dw \ f_\chi (\mathbf{w}; \hat{q}) \ \delta \left( \frac{q^2}{2 \mu_{\chi N}} + q \cdot \mathbf{w} + \Delta E_{1 \to 2} \right) \left| \langle \psi_2 | e^{i\mathbf{q} \cdot \hat{r}} |\psi_1 \rangle \right|^2 , \tag{2.3}
\]

where \( \rho_\chi \) and \( m_\chi \) represent the DM local density and the DM mass, respectively, \( A \) is the atomic number of the target nucleus, \( \sigma_{\chi n} \) is the DM-nucleon cross section, the bracket \( \langle \cdots \rangle \) denotes the average over the DM velocity distribution.

* Here we omit the nuclear form factor since the transfer momentum is so small for sub-GeV DM that the structure of nucleus is irrelevant for a coherent DM-nucleus scattering.
$f_\chi(w, \hat{q})$ is the DM velocity distribution with unit vector $\hat{q}$ as its zenith direction, $g_\chi(w; \hat{q}) \equiv w^2 f_\chi(w; \hat{q})$, and $\Theta$ is the Heaviside step function. While $\phi_{q\mathbf{w}}$ is the azimuthal angle of the spherical coordinate system $(\hat{q}; w)$, the polar angle $d\cos \theta_{q\mathbf{w}}$ has integrated out the delta function in above derivation. $\mu_{\chi N} = m_\chi m_n / (m_\chi + m_n)$ ($\mu_{\chi N} = m_\chi m_N / (m_\chi + m_N)$) is the reduced mass of the DM-nucleon (DM-nucleus) pair, and $\Delta E_{1\rightarrow 2}$ denotes the relevant energy difference. For the given $q$ and $\Delta E_{1\rightarrow 2}$, function $w_{\text{min}}$ determines the minimum kinetically possible velocity for the transition:

$$w_{\text{min}}(q, \Delta E_{1\rightarrow 2}) = \frac{q}{2\mu_{\chi N}} + \frac{\Delta E_{1\rightarrow 2}}{q}.$$  \hfill (2.4)

In practice, we take $\rho_\chi = 0.3 \text{ GeV/cm}^3$, and the velocity distribution can be approximated as a truncated Maxwellian form in the galactic rest frame, i.e., $f_\chi(w, \hat{q}) \propto \exp \left[-\frac{|w + v_e|^2}{v_0^2}\right] \Theta(v_{\text{esc}} - |w + v_e|)$, with the earth’s velocity $v_e = 230 \text{ km/s}$, the dispersion velocity $v_0 = 220 \text{ km/s}$ and the galactic escape velocity $v_{\text{esc}} = 544 \text{ km/s}$.

### 2.2. Migdal effect with tight-binding approximation

It is a natural idea to extend the above Migdal approach to the electronic bands in the crystalline solids. However, due to the non-local nature of the itinerant electrons in solids such extension does not seem so straightforward, especially for the low energy excitation processes. In order to apply the Migdal approach for the localized electron system to the non-local electrons in crystal, we manage to describe the electrons with the Wannier functions (WFs) in the context of the tight-binding (TB) approximation, in which the extensive nature of itinerant electrons is encoded in the hopping integral.

Our strategy is outlined as follows. First, for simplicity it is assumed that there is only one atom in each primitive unit cell (PUC), and we express the Bloch wavefunction of an isolated electronic band $\{|i k\rangle\}$ (with band index $i$ and crystal momentum $k$) in terms of a complete set of localized WF $\{|\mathbf{R} i\rangle\}$ (with band index $i$ and cell index $\mathbf{R}$) as

$$|i k\rangle = \sum_{\mathbf{R}} \frac{e^{i k \mathbf{R}}}{\sqrt{N}} |\mathbf{R} i\rangle,$$  \hfill (2.5)

where $N$ is the number of unit cells, or equivalently, the number of mesh points in the first Brillouin Zone (1BZ). The orthonormality of WFs $\{|\mathbf{R} i\rangle\}$, i.e., $\langle \mathbf{R'} i'| \mathbf{R} i \rangle = \delta_{i'i} \delta_{\mathbf{R}'\mathbf{R}}$ corresponds to the normalization convention over the whole crystal such that $\langle i' k' | i' k \rangle = \delta_{i'i} \delta_{k'k}$. Accordingly it is straightforward to obtain the inverse relation

$$|\mathbf{R} i\rangle = \sum_{k} \frac{e^{-ik \mathbf{R}}}{\sqrt{N}} |i k\rangle.$$  \hfill (2.6)

Next, we impose the Galilean boost operator $\hat{G}_{\mathbf{R}}(\mathbf{q}_e) \equiv e^{i \mathbf{q}_e \cdot \mathbf{r}}$ exclusively on the recoiled atom located at $\mathbf{R}$, with velocity $-\mathbf{q}_e / m_N$. For convenience, we choose the site $\mathbf{R} = \mathbf{0}$ for illustration. Thus, as a direct extension of the Migdal effect in atoms, the transition amplitude between an initial valence state and final conducting state can be written as

$$\langle i' k' | \hat{G}_{\mathbf{R}}(\mathbf{q}_e) | i k \rangle = \frac{1}{N} \sum_{\mathbf{R'}} \langle \mathbf{R}' i' | \hat{G}_{\mathbf{R}}(\mathbf{q}_e) - 1 | 0 i \rangle e^{-i k' \mathbf{R}'}$$

$$= \frac{1}{N} \sum_{\mathbf{R'}} \langle \mathbf{R}' i' | e^{i \mathbf{q}_e \cdot \mathbf{r}} - 1 | 0 i \rangle e^{-i k' \mathbf{R}'}.$$  \hfill (2.7)

where the hopping integrals between neighboring atoms $\langle \mathbf{R}' i' | e^{i \mathbf{q}_e \cdot \mathbf{r}} | 0 i \rangle$ reflect the delocalized nature of the electrons in crystalline solids.

However, above extensive use of operator $e^{i \mathbf{q}_e \cdot \mathbf{r}}$ in crystalline environment needs to be carefully examined, considering that the recoiled nucleus no longer amounts to an evident reference to describe the electronic excitation process while ambient nuclei remain at rest. To this point we make some detailed explanation. First, a momentum transfer larger than $q \sim \mathcal{O}(\text{keV})$ is sufficient to resolve the diamond structure, and in this case the Galilean transformation
can be imposed onto a specific atom. Then imagine the atom residing at \( \mathbf{R} = 0 \) is struck, from Eq. (2.5) we assume that the instantaneous eigenstate responding to this recoiled atom takes the form
\[
|i\ k\rangle_{\text{struck}} = \frac{1}{\sqrt{N}} \left( \hat{G}_0 (-q e) |0\ i\rangle + \sum_{\mathbf{R} \neq 0} e^{-i\mathbf{k} \cdot \mathbf{R}} |\mathbf{R}\ i\rangle \right),
\]
while the actual electronic state after the collision remains intact (i.e., \(|i\ k\rangle\)) under the sudden approximation. Second, since the Debye frequency of the system \( \omega_D \sim O (10^{-1} \text{ eV}) \) is much lower than the energy gap \( E_g \sim O (\text{eV}) \) between valence and conduction bands, the evolution of the eigenstate \(|i\ k\rangle_{\text{struck}}\) can be approximated as adiabatic and is hence equivalent to the eigenstate of original Hamiltonian. Because these eigenstates eventually evolve adiabatically back to the original ones after the relaxation of the struck nucleus. In this sense, the state \(|i\ k\rangle\) is regarded as perturbed with respect to the adiabatically evolved Hamiltonian and can be projected to eigenstates \(|i'\ k'\rangle_{\text{struck}}\) to derive the transition amplitude \(\langle i'\ k' | i\ k \rangle\) given in Eq. (2.7). In addition, since the timescale of the excitation is roughly \(E_g^{-1}\), during which the recoiled nucleus at most travels a distance around \((q/m)E_g^{-1} \sim O (10^{-2} \text{ Å})\) for a momentum transfer \(q \sim 1 \text{ MeV}\), the excitation can be regarded as instantaneous, and thus the configuration effect of the displaced nucleus can be ignored.

3. Practical calculation of excitation event rates with WFs

3.1. Formalism

In this section we will derive the formalism of the excitation event rate induced by a recoiled nucleus in the bulk diamond. Here we consider a more realistic multiband case where a separate group of \(J\) bands cross with each other, and a Bloch orbital can be expressed in terms of WF states \(\{|\mathbf{R}\ m\rangle\}\) in following way:
\[
|i\ k\rangle = \sum_{\mathbf{R}} \frac{e^{i\mathbf{k} \cdot \mathbf{R}}}{\sqrt{N}} \left( \sum_{m=1}^{J} U_{mi}^{(k)\dagger} |\mathbf{R}\ m\rangle \right),
\]
and its inverse transformation
\[
|\mathbf{R}\ m\rangle = \sum_{\mathbf{k}} \frac{e^{-i\mathbf{k} \cdot \mathbf{R}}}{\sqrt{N}} \left( \sum_{i=1}^{J} U_{im}^{(k)} |i\ k\rangle \right),
\]
with the unitary matrix \(U^{(k)}\) that mixes different bands for each \(\mathbf{k}\)-vector. In practice, we use the code Wannier 90 [48] to realize this scheme, which specializes in constructing the maximally localized Wannier functions (MLWFs) from a set of Bloch states. To avoid distraction from present discussion, we arrange a short review of the MLWF in the
Appendix A. In order to transplant the treatment on the Migdal effect applied in the isolated atom to the crystalline environment, here we introduce the Galilean boost operator $\hat{G}_{01}(q_i)$ that accounts for recoil effect of the first diamond atom at the site $(\mathbf{R} = 0, \tau_1 = 0)$, so one has the following transition amplitude ($i' \neq i$):

$$
\langle i' k' | \hat{G}_{01}(q_i) | i k \rangle = \frac{i}{N} \sum_{m, m'} \sum_{\mathbf{R}'} (U^{(k')}_{m m'} \hat{c})^* \langle R' m' | \hat{G}_{01}(q_i) - 1 | 0 m \rangle U^{(k)}_{m' m} e^{-i k' \cdot \mathbf{R}'}
$$

$$
\simeq \frac{i}{N} q_i \cdot \left( \sum_{m_1, m'} \sum_{\mathbf{R}'} U^{(k')}_{m_1 m'} (0 m_1 \hat{r} | R' m')^* U^{(k')}_{m' m'} e^{-i k' \cdot \mathbf{R}'} \right)
$$

$$
= \frac{i}{N} q_i \cdot J_{1(i'k'; ik)},
$$

(3.3)

where the origin of coordinate operator $\hat{r}$ is placed at the atom 1 in the PUC, and correspondingly the partial summation is only over the WFs centered at atom 1. We use $J_{1(i'k'; ik)}$ to denote the summation in parenthesis in the second line. Since momentum transfer $q$ is highly suppressed by $m_e/m_N$, in derivation we assume $|q_i| \cdot |r| \ll 1$, which is a good approximation for a sub-GeV DM. Similar discussion can be easily applied to the second atom located at $\tau_2 = (1/4, 1/4, 1/4)$, only keep in mind that the origin of operator $\hat{r}$ is also placed at atom 1 in practical use of Wannier 90. After a translation, the corresponding transition amplitude for the atom 2 at the site $\mathbf{R} = 0$ is modified as

$$
\langle i' k' | \hat{G}_{02}(q_i) | i k \rangle \simeq \frac{i}{N} q_i \cdot \left( \sum_{m_2, m'} \sum_{\mathbf{R}'} U^{(k')}_{m_2 m'} (0 m_2 \hat{r} | R' m')^* U^{(k')}_{m' m'} e^{-i k' \cdot \mathbf{R}'} - d \sum_{m_2} U^{(k')}_{m_2 m} U^{(k')}_{m} \right)
$$

$$
= \frac{i}{N} q_i \cdot J_{2(i'k'; ik)},
$$

(3.4)

with $d$ being the position vector of atom 2 relative to atom 1, and $J_{2(i'k'; ik)}$ encodes the terms in parenthesis. Thus, after taking into account the two degenerate spin states, the total excitation event rate can be expressed as

$$
\mathcal{R} = \frac{\rho_e}{m_e} \left( \frac{A^2 \sigma_{\chi n}}{2 \pi \mu_{\chi n}^2} \right) \left( \frac{m_e}{m_N} \right)^2 \frac{V^2}{N} \int d^3 q \sum_{i} \sum_{i'} \int_{1BZ} \frac{d^3 k'}{(2\pi)^3} \frac{d^3 k}{(2\pi)^3} \left\{ \int \frac{g_{\chi}(w, q)}{q w} dw \right\} d\phi_{\chi w} \times \Theta [w - w_{\text{min}}(q, E_{i'k'} - E_{ik})] \left( |q \cdot J_{1(i'k'; ik)}|^2 + |q \cdot J_{2(i'k'; ik)}|^2 \right),
$$

(3.5)

where the sums are over the valence bands for initial states and the conducting bands for final states, respectively. For simplicity we approximate the velocity distribution as an isotropic one, and as a result the angular correlation between the laboratory velocity with respect to the galaxy and the orientation of the crystal is eliminated. Besides, in order to make the scan of parameters computationally more efficient, Eq. (3.5) can be further equivalently expressed as

$$
\mathcal{R} = \frac{\rho_e}{m_e} \left( \frac{A^2 \sigma_{\chi n} q_{\text{rel}}^2}{3 \pi \mu_{\chi n}^2} \right) \left( \frac{m_e}{m_N} \right)^2 N \int \frac{4 \pi g_{\chi}(w)}{w} dw \int d\ln E_e d\ln q \Theta [w - w_{\text{min}}(q, E_e)] \mathcal{F}(q, E_e),
$$

(3.6)

where a momentum reference value is constructed as $q_{\text{rel}} = 2\pi/a$, and a non-dimensional crystal form factor is introduced as

$$
\mathcal{F}(q, E_e) = \left( \frac{q}{q_{\text{rel}}} \right)^2 \sum_{i'} \sum_{i} \int_{1BZ} \frac{\Omega d^3 k'}{(2\pi)^3} \frac{\Omega d^3 k}{(2\pi)^3} \left[ E_e \delta(E_{i'k'} - E_{ik} - E_e) q^2 \left( |J_{1(i'k'; ik)}|^2 + |J_{2(i'k'; ik)}|^2 \right) \right],
$$

(3.7)

Specifically, after inserting the monochromatic velocity distribution $4 \pi g_{\chi}(v) = \delta(v - w)$ the total transition rate for parameter pair $(m_\chi, w)$ can be recast as

$$
\mathcal{R}(m_\chi, w) = \frac{\rho_e}{m_\chi} \left( \frac{A^2 \sigma_{\chi n} q_{\text{rel}}^2}{3 w \mu_{\chi n}^2} \right) \left( \frac{m_e}{m_N} \right)^2 N \int d\ln E_e d\ln q \Theta [w - w_{\text{min}}(q, E_e)] \mathcal{F}(q, E_e).
$$

(3.8)
3.2. Computational details and results

Now we put into practice of the estimate of the Migdal excitation event rate. With Quantum Espresso code [52], we first perform the DFT calculation to obtain the Bloch eigenfunctions and eigenvalues using the plane-wave basis set and Troullier-Martins norm-conserving (NC) pseudopotentials in the Kleinman-Bylander representation, on a \(20 \times 20 \times 20\) Monkhorst-Pack mesh of \(k\)-points. The exchange-correlation functional is treated within the generalized gradient approximation (GGA) parametrized by Perdew, Burke, and Ernzerhof (PBE) [53]. The energy cut is set to 90 Ry and a lattice constant \(a = 3.560\,\text{Å}\) obtained from relaxation is adopted.

Then we invoke the software package Wannier 90 [48] to compute the matrix element \(\langle 0 m | \hat{r} | R m' \rangle\) and the unitary matrix \(U^{(k)}\) using a smaller homogeneous set of \(16 \times 16 \times 16\) \(k\)-points. We generate \(J = 32\) WFs out of \(J_k = 72\) Bloch wavefunctions from the DFT calculation by beginning with a set of 32 localized trial orbitals \(\{g_n(\mathbf{r})\}\) that correspond to \(s, p, d, f\) orbitals as some rough initial guess for these WFs\(^\dagger\). In order to fix the WFs at the atom while containing their spreads, the gauge selection step is expediently spared because otherwise some ML WFs will be found located at interstitial sites, which makes the picture of recoiled atom ambiguous. The widest spread of these generated WFs is around \(1.4\,\text{Å}\), so hopping terms within up to the third neighbor PUCs are sufficient to guarantee convergence in calculation of \(\langle 0 m | \hat{r} | R m' \rangle\).

On the other hand, once the Hamiltonian matrices \(\langle R' m' | \hat{H} | 0 m \rangle\) have been determined, the band structures become an eigenvalue problem

\[
\sum_{m'=1}^{J} \sum_{\mathbf{R}} \langle 0 m | \hat{H} | R m' \rangle e^{-i\mathbf{k} \cdot \mathbf{R}} U_{m'i}^{(k)} = \bar{\epsilon}_{i,\mathbf{k}} U_{m'i}^{(k)},
\]

where the eigenvalue \(\bar{\epsilon}_{i,\mathbf{k}}\) corresponds to the \(i\)-th band energy at \(\mathbf{k}\), and the eigenvectors \(U_{m'i}^{(k)}\) that form a \(J \times J\) unitary matrix are deliberately written in accordance with Eq. (3.1). Such procedure is called the Wannier interpolation. The band structures of crystalline diamond from the DFT calculation and interpolation are presented in the left panel of Fig. 3.1, where the blue solid line (red dotted line) represents the DFT calculation (Wannier interpolation). To reproduce exactly the original DFT band structures in a specific energy range, Bloch states spanning relevant range are faithfully retained in the subspace selection procedure, and such energy range is referred to as “frozen energy

\(^\dagger\) See Appendix A for a brief review on the ML WFs.
window", or "inner window". In this work, we choose a frozen energy window ranging from the bottom of the valence band to 40 eV above the valence band maximums (VBM), which is also reflected in Fig. 3.1.

In practical computation of the form factor Eq. (3.7) we use a bin width $\Delta E = 0.059$ eV to smear the delta function, and the integrand is evaluated at the central value in each energy bin. Besides, the integrals of the continuous $k$-points in the 1BZ are replaced by corresponding summations over the uniform $16 \times 16 \times 16$ $k$-points. In the right panel of Fig. 3.1 shown is the crystal form factor $F(q, E_e)$ introduced in Eq. (3.7), where momentum transfer $q$ is expressed in terms of the diamond reciprocal lattice $2\pi/a \approx 3.48$ keV. For demonstration, we choose two benchmark pairs of parameters $(m_\chi, w)$ to calculate the transition event rate in Eq. (3.8). Due to the step function that embodies the kinetic constraint, only the areas enclosed in the contours contribute to the excitation event rate. For example, when a cross section $\sigma_{\chi n} = 10^{-38}$ cm$^2$ is assumed, parameters $m_\chi = 500$ MeV, $w = 100$ km/s presented in yellow corresponds to an excitation event rate of $0.01836 /$kg$/yr$, while $m_\chi = 10$ MeV, $w = 800$ km/s in blue corresponds to an excitation event rate of $0.1876 /$kg$/yr$. It is noted that although a heavier DM tends to prompt a larger momentum transfer and hence results in a larger transition probability, a suppression factor inversely proportional to $m_\chi^3$ may alleviate or even offset such mass effect in the Sub-GeV mass range, depending on velocity.

In left panel of Fig. 3.2 we plot the velocity-averaged energy spectra of the Migdal excitation for DM mass $m_\chi = 10$ MeV (blue), 100 MeV (orange) and 1 GeV (green), respectively. It is observed that these spectra do not fall off as quickly as the case in direct DM-electron excitation process. Hence a wider energy range is usually required to fully describe the Migdal effect in crystalline environment. From the energy spectra we also estimate the sensitivity of a 1 kg-yr diamond detector in the right panel of Fig. 3.2, assuming an average energy of 13 eV for producing one electron-hole pair [47]. The 90% C.L. exclusion contour for DM-nucleon cross section for both a single-electron (blue) and a two-electron (orange) threshold are presented with no background event assumed.

4. Summary and prospects

In this paper we presented a tight-binding approach to describe the Migdal effect in the diamond crystal. This localized description is a natural choice to generalize the well established treatment for isolated atoms to the case in crystalline solids. To achieve such TB description we generate a set of atom-centered WFs by use of software packages Quantum Espresso and Wannier 90. While the localization effect of the recoiled atom is preserved in forms of WFs, the delocalized nature of the Bloch states in solids is encoded in the hopping terms at the same time. Based on these hopping integrals, the electronic excitation rates induced by the recoiled ion were computed straightforwardly.
Here we make some comments on our methodology. As has been noted in previous section, our method deviates a little from the standard procedure to derive the MLWFs implemented in Wannier 90, which usually includes two steps, namely, subspace selection and gauge selection\(^\dagger\). The latter step is skipped in our computation so as to restrict the centers of WFs to the lattice sites. As a consequence, the spreads of those WFs have not been optimally minimized. From a conceptual point of view, this does not cause a severe problem because although WFs indicate some atomic properties and provide intuitive pictures of the chemical bonds, they do not necessarily bear definite physical meanings such as real atomic wavefunction in solids. Actually, to determine the WFs is experimentally infeasible, even in principle [45]. In this sense, even a set of relatively “fat” WFs suffice in calculation of the excitation event rate, as long as sufficiently large number of neighbors are included in hopping integrals. However, WFs with too large spreads indeed cause inconvenience in practice. This partly explains why we choose diamond crystal rather than other tetrahedral semiconductors as example to describe the Migdal effect in semiconductors: for the cases of silicon and germanium, Wannier 90 is found to generate atom-centered WFs with much larger spreads in the same energy window unless a gauge optimization is performed, and hence more computational efforts are required to reach convergence.

Another concern is how the calculation of Migdal excitation rates will rely on the selection of the WFs, considering there existing a gauge freedom. In fact, since it is required that the centers of the WFs locate at the lattice sites to account for the struck atoms, such invariance of arbitrary WF selection is unnecessary in formulating the transition probability. To see this, recall the transition amplitude in Eq. (3.3), while the final state \(\sum_{m'} R' e^{-i\mathbf{k'} \cdot \mathbf{R'}} \langle \mathbf{R}' m' | \cdots = \langle i' k' | \cdots \) as a whole is independent of the WFs, the valence initial part \(\sum_{m_1} |0 m_1 \rangle U^{(k)}_{m_1} \) is invariant only within the WF subspace attached to atom 1. This can be understood intuitively with the example of tetrahedral \(sp^3\) hybridization, where the valence wavefunctions around an atom can be expressed either as linear combinations of \(s\) and \(p\) atomic orbitals, or combinations of four identical \(sp\)-hybrid orbitals pointing along the directions from the center to the corners of a tetrahedron. All these different choices should keep the sum \(\sum_{m_1} |0 m_1 \rangle U^{(k)}_{m_1} \) invariant. To verify this, we used \(s\), \(p\) atomic orbitals and \(sp^3\) hybrids as trial orbitals \(\{g_n(\mathbf{r})\}\) to wannierize and calculate the transition event rate respectively, and the difference between the two results turns out to be well within a few percent, underlining a good agreement. Thus our approach has proved both effective and self-consistent.

As a last remark, we note that our method in dealing with the Migdal effect in semiconductor is rather ad hoc and preliminary, mainly due to the limitation that the WFs should be atom-centered, which prevents further minimizing the spreads of the WFs in implementation of established codes such as Wannier 90. Nevertheless, from above discussions, this hurdle can be circumvented if one wannierizes the initial valence states and the final conduction states in separate steps and in complementary energy windows. This brings two benefits: (1) the initial valence states alone can be readily wannierized as atom-centered WFs, e.g., either in the forms of \(sp^3\) hybrids or \(s\), \(p\) atomic orbitals; (2) the final conduction states alone can be optimally wannierized without the need to to atom-centered. Moreover, since the valence and conduction states are orthogonal, the corresponding sets of WFs are automatically orthogonal to each other, at least in theory. Of course, some modifications need to be made in underlying codes. We hope that our work will stimulate further optimization on these problems from a wider community.

\section{Maximally localized Wannier functions}

Since our practical realization of the atom-centered WFs are closely related to the generation of the maximally localized Wannier functions via implementing software Wannier 90 [48], here we take a brief review. From Eq. (3.2) it is straightforward to see that the WFs are non-unique due to the arbitrariness of the unitary matrix \(U^{(k)}\). In order to overcome such indeterminacy, Mazari and Vanderbilt [54] developed a procedure to minimize the second moment of the WFs around their centers so as to generate a set of well-defined and localized WFs, namely, the maximally localized Wannier functions (MLWFs). Given an isolated group of \(J\) Bloch bands, the procedure begins with a set of

\(^\dagger\) See Appendix A for more details.
localized trial orbitals \( \{ g_n(r) \} \) as some rough initial guess for corresponding WFs, which are then projected onto those Bloch wavefunctions as the following,

\[
    |\phi_{n\mathbf{k}}\rangle = \sum_{i=1}^{J} |i\mathbf{k}\rangle \langle i\mathbf{k}|g_n\rangle ,
\]

which are typically smooth in \( \mathbf{k} \) and hence are well-localized in space. By further Löwdin-orthonormalizing these functions one obtains \( J \) Bloch-like states \( |\tilde{i}\mathbf{k}\rangle \) that are related to the original \( |i\mathbf{k}\rangle \) via a \( J \times J \) unitary matrix \( U_{\mathbf{k}} \) in the following manner,

\[
    |\tilde{i}\mathbf{k}\rangle = \sum_{j=1}^{J} |j\mathbf{k}\rangle U_{\mathbf{k},ij} .
\]

Substituting \( |i\mathbf{k}\rangle \) in Eq. (2.6) with \( |\tilde{i}\mathbf{k}\rangle \) we then have \( J \) localized WFs. The minimization criterion proposed by Mazari and Vanderbilt mentioned above is to minimize the functional

\[
    \Omega = \sum_{n} \left[ \langle 0n|\hat{r}^2|0n\rangle - |\langle 0n|\hat{r}|0n\rangle|^2 \right]
    = \Omega_1 + \tilde{\Omega} ,
\]

with

\[
    \Omega_1 = \sum_{n} \left[ \langle 0n|\hat{r}^2|0n\rangle - \sum_{R,m} |\langle Rm|\hat{r}|0n\rangle|^2 \right]
\]

and

\[
    \tilde{\Omega} = \sum_{n} \sum_{Rm \neq 0n} |\langle Rm|\hat{r}|0n\rangle|^2 .
\]

The purpose of such separation is that \( \Omega_1 \) is invariant under arbitrary unitary transformation \( \mathbf{U}^{(\mathbf{k})} \) in Eq. (3.2), and hence the minimization of \( \Omega \) depends only on the variation effects of gauge \( \mathbf{U}^{(\mathbf{k})} \) on the term \( \tilde{\Omega} \). Since the matrix \( U_{\mathbf{k}} \) in Eq. (A.2) has already provided \( J \) sufficiently localized WFs, we use them as the starting point for the iterative steepest-descent method to reach the optimal unitary transformation \( \mathbf{U}^{(\mathbf{k})} \) that minimizes \( \tilde{\Omega} \). This procedure is called gauge selection.

In more general cases, additional modifications are required to pick the Bloch orbitals of interest from those unwanted ones when the bands are crossing with each other. The Souza-Marzari-Vanderbilt method proposed in Ref. [55] successfully realized such disentanglement. In this approach, one first identifies a set of \( J_\mathbf{k} \geq J \) Bloch orbitals that form a \( J_\mathbf{k} \)-dimensional Hilbert space at each point \( \mathbf{k} \) in the 1BZ in a sufficiently large energy range, which is dubbed as “disentanglement window”, or “outer window”, and then chooses a \( J_\mathbf{k} \)-dimensional subspace that gives the smallest possible value of \( \Omega_1 \) via iterative procedures. Similarly to above discussion, the minimization of \( \Omega_1 \) also begins with a rough initial guess that are usually obtained by first projecting \( J \) localized trial orbitals \( \{ g_n(r) \} \) onto \( J_\mathbf{k} \) Bloch states such that

\[
    |\phi_{n\mathbf{k}}\rangle = \sum_{j=1}^{J_\mathbf{k}} |j\mathbf{k}\rangle \langle j\mathbf{k}|g_n\rangle ,
\]

and then constructing the following \( J \) orthonormalized Bloch-like states \( |\tilde{i}\mathbf{k}\rangle \) that are related to the original \( |i\mathbf{k}\rangle \) via a \( J_\mathbf{k} \times J \) matrix \( V_{\mathbf{k}} \) such as

\[
    |\tilde{i}\mathbf{k}\rangle = \sum_{j=1}^{J_\mathbf{k}} |j\mathbf{k}\rangle V_{\mathbf{k},ij} .
\]

An algebraic algorithm that updates the minimization iteratively is then used to obtained \( J \)-dimensional optimal subspace in a self-consistent manner. This process is called subspace selection. Once the optimal subspace is determined,
i.e., the minimization of $\Omega_I$ is achieved, the gauge-selection step to further minimize the noninvariant part $\tilde{\Omega}$ are followed, until a set of self-consistent MLWFs are obtained.

[1] XENON collaboration, *Dark Matter Search Results from a One Ton-Year Exposure of XENON1T*, Phys. Rev. Lett. **121** (2018) 111302 [1805.12562].
[2] XENON collaboration, *Light Dark Matter Search with Ionization Signals in XENON1T*, 1907.11485.
[3] LUX collaboration, *Results of a Search for Sub-GeV Dark Matter Using 2013 LUX Data*, Phys. Rev. Lett. **122** (2019) 131301 [1811.11241].
[4] PandaX-II collaboration, *Constraining Dark Matter Models with a Light Mediator at the PandaX-II Experiment*, Phys. Rev. Lett. **121** (2018) 021304 [1802.06912].
[5] SUPERCDMS collaboration, *Low-mass dark matter search with CDMSlite*, Phys. Rev. **D97** (2018) 022002 [1707.01632].
[6] CDEX collaboration, *Constraints on Spin-Independent Nucleus Scattering with sub-GeV Weakly Interacting Massive Particle Dark Matter from the CDEX-1B Experiment at the China Jinping Underground Laboratory*, Phys. Rev. Lett. **123** (2019) 161301 [1905.00354].
[7] CRESST collaboration, *First results from the CRESST-III low-mass dark matter program*, 1904.00498.
[8] DarkSide collaboration, *Low-Mass Dark Matter Search with the DarkSide-50 Experiment*, Phys. Rev. Lett. **121** (2018) 081307 [1802.06994].
[9] EDELWEISS collaboration, *Searching for low-mass dark matter particles with a massive Ge bolometer operated above-ground*, Phys. Rev. **D99** (2019) 082003 [1901.03588].
[10] R. Essig, J. Mardon and T. Volansky, *Direct Detection of Sub-GeV Dark Matter*, Phys. Rev. **D85** (2012) 076007 [1108.5383].
[11] P. W. Graham, D. E. Kaplan, S. Rajendran and M. T. Walters, *Semiconductor Probes of Light Dark Matter*, Phys. Dark Univ. 1 (2012) 32 [1203.2531].
[12] R. Essig, A. Manalaysay, J. Mardon, P. Sorensen and T. Volansky, *First Direct Detection Limits on sub-GeV Dark Matter from XENON10*, Phys. Rev. Lett. **109** (2012) 021301 [1206.2644].
[13] R. Essig, M. Fernandez-Serra, J. Mardon, A. Soto, T. Volansky and T.-T. Yu, *Direct Detection of sub-GeV Dark Matter with Semiconductor Targets*, JHEP **05** (2016) 046 [1509.01598].
[14] Y. Hochberg, Y. Zhao and K. M. Zurek, *Superconducting Detectors for Superlight Dark Matter*, Phys. Rev. Lett. **116** (2016) 011301 [1504.07237].
[15] S. K. Lee, M. Lisanti, S. Mishra-Sharma and B. R. Safdi, *Modulation Effects in Dark Matter-Electron Scattering Experiments*, Phys. Rev. **D92** (2015) 083517 [1508.07361].
[16] Y. Hochberg, M. Pyle, Y. Zhao and K. M. Zurek, *Detecting Superlight Dark Matter with Fermi-Degenerate Materials*, JHEP **08** (2016) 057 [1512.04533].
[17] Y. Hochberg, T. Lin and K. M. Zurek, *Detecting Ultralight Bosonic Dark Matter via Absorption in Superconductors*, Phys. Rev. **D94** (2016) 015019 [1604.06800].
[18] I. M. Bloch, R. Essig, K. Tobioka, T. Volansky and T.-T. Yu, *Searching for Dark Absorption with Direct Detection Experiments*, JHEP **06** (2017) 087 [1608.02123].
[19] S. Knapen, T. Lin and K. M. Zurek, *Light Dark Matter in Superfluid Helium: Detection with Multi-excitation Production*, Phys. Rev. **D95** (2017) 056019 [1611.06228].
[20] S. Derenzo, R. Essig, A. Massari, A. Soto and T.-T. Yu, *Direct Detection of sub-GeV Dark Matter with Scintillating Targets*, Phys. Rev. **D96** (2017) 016026 [1607.01009].
[21] Y. Hochberg, Y. Kahn, M. Lisanti, C. G. Tully and K. M. Zurek, *Directional detection of dark matter with two-dimensional targets*, Phys. Lett. **B772** (2017) 239 [1606.08849].
[22] R. Essig, J. Mardon, O. Slone and T. Volansky, *Detection of sub-GeV Dark Matter and Solar Neutrinos via
[23] Y. Hochberg, T. Lin and K. M. Zurek, *Absorption of light dark matter in semiconductors*, Phys. Rev. **D95** (2017) 023013 [1608.01994].

[24] R. Essig, T. Volansky and T.-T. Yu, *New Constraints and Prospects for sub-GeV Dark Matter Scattering off Electrons in Xenon*, Phys. Rev. **D96** (2017) 043017 [1703.00910].

[25] S. Knapen, T. Lin and K. M. Zurek, *Directional Detection of Light Dark Matter with Optical Phonons in Polar Materials*, Phys. Lett. **B785** (2018) 386 [1712.06598].

[26] R. M. Geilhufe, F. Kahlhoefer and M. W. Winkler, *Dirac Materials for Sub-MeV Dark Matter Detection: New Targets and Improved Formalism*, 2019.02091.

[27] M. Heikinheimo, K. Nordlund, K. Tuominen and N. Mirabolfathi, *Velocity Dependent Dark Matter Interactions in Single-Electron Resolution Semiconductor Detectors with Directional Sensitivity*, Phys. Rev. **D99** (2019) 103018 [1903.08654].

[28] T. Trickle, Z. Zhang, K. M. Zurek, K. Inzani and S. Griffin, *Multi-Channel Direct Detection of Light Dark Matter: Theoretical Framework*, 1910.08092.

[29] T. Emken, R. Essig, C. Kouvaris and M. Sholapurkar, *Direct Detection of Strongly Interacting Sub-GeV Dark Matter via Electron Recoils*, JCAP **1909** (2019) 070 [1905.06348].

[30] R. M. Geilhufe, F. Kahlhoefer and M. W. Winkler, *Dirac Materials for Sub-MeV Dark Matter Detection: New Targets and Improved Formalism*, 2019.02091.

[31] M. Heikinheimo, K. Nordlund, K. Tuominen and N. Mirabolfathi, *Velocity Dependent Dark Matter Interactions in Single-Electron Resolution Semiconductor Detectors with Directional Sensitivity*, Phys. Rev. **D99** (2019) 103018 [1903.08654].

[32] A. Caputo, A. Esposito and A. D. Polosa, *Sub-MeV Dark Matter and the Goldstone Modes of Superfluid Helium*, 1901.07569.

[33] N. A. Kurinsky, T. C. Yu, Y. Hochberg and B. Cabrera, *Diamond Detectors for Direct Detection of Sub-GeV Dark Matter*, Phys. Rev. **D99** (2019) 123005 [1901.07569].
[48] A. A. Mostofi, J. R. Yates, G. Pizzi, Y.-S. Lee, I. Souza, D. Vanderbilt et al., *An updated version of wannier90: A tool for obtaining maximally-localised wannier functions*, *Computer Physics Communications* **185** (2014) 2309.

[49] H. Schober, *An introduction to the theory of nuclear neutron scattering in condensed matter*, *Journal of Neutron Research* (2014) 109.

[50] K. Schutz and K. M. Zurek, *Detectability of Light Dark Matter with Superfluid Helium*, *Phys. Rev. Lett.* **117** (2016) 121302 [1604.08206].

[51] F. Acanfora, A. Esposito and A. D. Polosa, *Sub-GeV Dark Matter in Superfluid He-4: an Effective Theory Approach*, *Eur. Phys. J. C79* (2019) 549 [1902.02361].

[52] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni et al., *QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials*, *Journal of Physics: Condensed Matter* **21** (2009) 395502.

[53] J. P. Perdew, K. Burke and M. Ernzerhof, *Generalized gradient approximation made simple*, *Phys. Rev. Lett.* **77** (1996) 3865.

[54] N. Marzari and D. Vanderbilt, *Maximally localized generalized wannier functions for composite energy bands*, *Phys. Rev. B* **56** (1997) 12847.

[55] I. Souza, N. Marzari and D. Vanderbilt, *Maximally localized wannier functions for entangled energy bands*, *Phys. Rev. B* **65** (2001) 035109.