Fully *ab-initio* all-electron calculation of dark matter-electron scattering in crystals with evaluation of systematic uncertainties

Cyrus E. Dreyer,[1,2] Rouven Essig,[3] Marivi Fernandez-Serra,[1,4] Aman Singal,[1,3,4] and Cheng Zhen[1,3]

[1]*Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York 11794, USA*
[2]*Center for Computational Quantum Physics, Flatiron Institute, 162 Fifth Avenue, New York, New York 10010, USA*
[3]*C. N. Yang Institute for Theoretical Physics, Stony Brook University, Stony Brook, New York 11794, USA*
[4]*Institute for Advanced Computational Sciences, Stony Brook University, Stony Brook, New York 11794, USA*

(Received 7 August 2023; accepted 8 May 2024; published 7 June 2024)

We calculate target-material responses for dark matter–electron scattering at the *ab initio* all-electron level using atom-centered Gaussian basis sets. The all-electron effects enhance the material response at high momentum transfers from dark matter to electrons, $q \gtrsim \mathcal{O}(10\alpha m_e)$, compared to calculations using conventional plane wave methods, including those used in QEDARK; this enhances the expected event rates at energy transfers $E \gtrsim 10$ eV, especially when scattering through heavy mediators. We carefully test a range of systematic uncertainties in the theory calculation, including those arising from the choice of basis set, exchange-correlation functional, number of unit cells in the Bloch sum, **k**-mesh, and neglect of scatters with very high momentum transfers. We provide state-of-the-art crystal form factors, focusing on silicon and germanium. Our code and results are made publicly available as a new tool, called *quantum chemistry dark* ("QCDARK").

DOI: 10.1103/PhysRevD.109.115008

I. INTRODUCTION

There has been rapid progress in direct-detection searches of sub-GeV dark matter (DM) by looking for electron recoils from DM-electron scattering in noble liquids and crystals (see, e.g., [1–26]). A theoretical description of such DM–electron scattering processes requires a quantitative description of the electronic structure of the detector material. This can be achieved by utilizing methods from condensed-matter physics and quantum chemistry. In particular, density functional theory (DFT) has been demonstrated to be a powerful tool for determining the ground-state electronic structure in a wide variety of materials from first principles, as well their response to various perturbations [27]. Also, DFT is the necessary first step to performing calculations using more advanced methods to treat systems with, e.g., stronger electron-electron correlations [28].

Several choices need to be made when calculating the electronic structure and DM–electron scattering rate from DFT. This includes choosing the type of basis functions to describe the electronic wavefunctions; whether to separate the core electrons in the material from the valence electrons by, e.g., pseudopotentials [27] or the projector-augmented wave (PAW) method [29]; and the exchange-correlation (XC) functional, which incorporates the many body effects of the electron–electron interactions [27,30]. These choices often represent trade-offs between accuracy and computational efficiency [31], which differ for different types of materials. Moreover, once these choices are made, the relevant properties must be tested for convergence with respect to the numerical parameters of the calculation.

DFT-based techniques have been applied to calculating DM–electron scattering in a variety of materials relevant for detectors including semiconductors [5,32–35], semimetals [36], superconducting nanowire single-photon detectors [37,38], quantum dots [39], etc. For the most part, studies of electron recoils in semiconductors have taken the approach of a plane wave basis set, core electrons frozen in pseudopotentials of PAW potentials, and either local/semilocal XC functionals based on the local density approximation (LDA), the generalized gradient approximation (GGA), or hybrid functionals that include a fraction of exact electron-electron exchange interaction [40]. DFT using plane wave basis sets are the one most commonly used for studying solids in condensed-matter physics and materials science.

However, there are motivations for choosing an alternative approach. First, it was recently shown that explicit

treatment of the core electrons significantly affects the DM–electron scattering rates [34,35,41]. While sub-GeV DM does not typically excite an electron from a core orbital to the conduction band, the inclusion of the rapidly oscillating part of the valence all-electron wavefunction near the atomic cores (where it must be orthogonal to the core electron wave functions [42]) is necessary to capture DM–electron scattering events with high momentum transfer. For a calculation with a plane wave basis to be tractable, these oscillations must be smoothed through the use of pseudopotentials, though the all-electron wavefunction can be reconstructed if PAWs are used [34,35].

A basis set made up of localized functions, e.g., atom-centered Gaussians, can treat core and valence electrons on the same footing without significant increase in computational cost. In addition, such basis sets can be used either with periodic boundary conditions for solids [43] or for finite systems, such as molecules or nanostructures [43], increasing the flexibility to explore different DM detector materials. Finally, using localized basis sets allows the use of quantum chemistry methods, which allows many-body correlations to be included when calculating the wave functions for atoms, molecules, liquids, and solids.

In this work, we develop the computational methodology to perform all-electron calculations of DM–electron scattering based on localized Gaussian basis sets. The resulting code, which we call *quantum chemistry dark* (QCDARK), is based on the python-based simulations of chemistry framework (PySCF) [43–45] package, which allows for DFT and quantum chemistry methods to be used on both finite and extended systems.[1] Via benchmark calculations on silicon and germanium, we show that the basis sets can be converged. PySCF has previously been employed in the context of DM–electron scattering in isolated atoms and molecules in [46]. We compare our results for DM-electron scattering with previous work, and find good agreement with EXCEED-DM [34,35], which reconstructs all–electron effects with PAW reconstruction. We also explore uncertainties related to the choice of XC functional and numerical convergence parameters. Though a quantitative determination of such errors in general is not possible for DFT methods, we demonstrate the potential situations where different parameters may result in significant changes.

The rest of the paper is organized as follows. In Sec. II, we describe electronic structure calculations, include a comparison between plane-wave and atomic-centered bases, discuss how to include all-electron effects, and describe the DM–electron scattering rate calculations with quantum chemistry basis sets. In Sec. III, we describe the results for Si and Ge, including the systematic uncertainties and the effects of the secondary ionization modeling; we also calculate the annual modulation rates, and compare our results with those of previous works. We conclude in Sec. IV. Two appendices contain technical information, including the properties of Cartesian Gaussians (Appendix A) and a derivation of the scattering rate formulas (Appendix C).

[1]The code is available at https://github.com/asingal14/QCDark.

II. CALCULATING DARK MATTER-ELECTRON SCATTERING RATES

In this section, we introduce the computational approach that underlies QCDARK, comparing the all-electron localized basis set approach used in this work with previous implementations based on plane waves and pseudopotentials.

A. Electronic structure

The electronic structure of the material is described by Kohn-Sham (KS) [30] DFT, in which the equations that describe the system of many interacting electrons are mapped onto a set of single-particle equations in an effective potential constructed to reproduce the ground-state electron density and total energy. The energy functional of the density n is given by (Gaussian units are assumed throughout)

$$E[n] = T_S[n] + \int d\mathbf{r} v_{ext}(\mathbf{r}) n(\mathbf{r}) + \frac{1}{2}\int d\mathbf{r} \int d\mathbf{r}' \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + E_{xc}[n], \quad (1)$$

where T_S is the sum of the kinetic energies of the non-interacting orbitals; v_{ext} is the external potential given by, e.g., the atomic nuclei in an all-electron calculation or the ions if pseudopotentials are used; and E_{xc} is the so-called exchange-correlation energy, which accounts for the many-body and quantum effects that are neglected in T_S and the Hartree electron-electron interaction [third term in Eq. (1)] [27]. Writing the density as a sum over auxiliary single-particle orbitals, i.e., $n(\mathbf{r}) = \sum_i |\psi_i(\mathbf{r})|^2$, and minimizing Eq. (1) with respect to variations in ψ results in the KS equations,

$$\left[-\frac{1}{2}\nabla^2 + v_{eff}(\mathbf{r})\right]\psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}), \quad (2)$$

where $v_{eff} = v_{ext} + \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + \delta E_{xc}[n]/\delta n(\mathbf{r})$, with the last term being the functional derivative of the exchange-correlation energy with respect to the density, which results in the XC potential $v_{xc}(\mathbf{r})$. The KS equations in Eq. (2) must be solved self-consistently, since the effective potential depends on the density.

Several choices exist in the above calculation, which we discuss in subsequent sections. First, the KS wave functions ψ must be expressed in terms of some basis functions (Sec. II A 1); the choice of these basis functions has a significant impact on many other aspects of the calculation,

including the possible choices for the boundary conditions and the treatment of core electrons (Sec. II A 2). Also, the exact form for the exchange-correlation potential is not known, and the choice of approximate $v_{\rm xc}(\mathbf{r})$ can result in qualitatively different results for the electronic structure of the material (Sec. II A 3).

1. Basis sets

Previous works [5,32–35] used plane waves as basis function (PW-basis), together with periodic boundary conditions. Then the KS wave functions for a given wave vector $\mathbf{k}$ in the first Brillouin Zone can be written as

$$\psi_{i\mathbf{k}}(\mathbf{r}) = \sqrt{\frac{1}{V}}\sum_{\mathbf{K}} u_i(\mathbf{k}+\mathbf{K}) e^{i(\mathbf{k}+\mathbf{K})\cdot\mathbf{r}}, \tag{3}$$

where $\mathbf{K}$ is a reciprocal lattice vector, $e^{i(\mathbf{k}+\mathbf{K})\cdot\mathbf{r}}$ are the basis functions, $u_i(\mathbf{k}+\mathbf{K})$ are the coefficients, and V is the volume of the crystal. The accuracy of the PW-basis for describing ψ is governed by the number of reciprocal lattice vectors included in the basis, which is usually specified as a kinetic energy cutoff of the plane waves. The PW-basis is good at describing relatively delocalized states, e.g., the states near and above the Fermi level in most solids; however, as we will discuss in Sec. II A 2, it is computationally expensive to capture the core electrons with a PW basis, since their localized nature requires the inclusion of very high energy plane waves and hence a large PW-basis size.

In this work, we use atom-centered Cartesian Gaussian basis sets. These basis sets are efficient at treating localized states, including the core electrons, and may be used for periodic or finite systems. The building blocks of this basis are *primitive* Gaussians,

$$G_{ijk}(\mathbf{r},\xi_\mu,\mathbf{A}) = (x-A_x)^i (y-A_y)^j (z-A_z)^k \times \exp\left[-\xi_\mu(\mathbf{r}-\mathbf{A})^2\right], \tag{4}$$

where $\mathbf{A}$ is the atomic position, ξ_μ is an adjustable parameter, and the Cartesian exponents i, j, and k are all integers that satisfy the condition $i+j+k=l$, with l being the angular quantum number of the shell. Then, the basis functions are given by *contracted* Gaussians, i.e., weighted sums of $N_{\rm prim}$ primitive Gaussians:

$$\tilde{G}_\alpha(\mathbf{r}) = \sum_{\mu=1}^{N_{\rm prim}} N_\mu c_\mu G_{ijk}(\mathbf{r},\xi_\mu,\mathbf{A}), \tag{5}$$

where $\alpha = \{\kappa, ijk\}$ is a composite index that runs over all nuclei κ in the system (or unit cells for periodic boundary conditions) as well as the Cartesian exponents, N_μ is the normalization of the primitive Gaussian, and c_μ is the coefficient of the primitive Gaussian. Note that N_μ and c_μ are fixed throughout the DFT calculation, as they are defined by the basis set.

For periodic calculations, the atomic orbitals in the unit cell are then

$$\phi_{\alpha\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} \tilde{G}_\alpha(\mathbf{r}-\mathbf{R}), \tag{6}$$

where $\mathbf{R}$ are the real-space lattice vectors. Upon self-consistently solving the DFT Hamiltonian, we obtain a coefficient matrix for each $\mathbf{k}$, $C_{i\alpha}(\mathbf{k})$, so that the Kohn-Sham wavefunctions (often referred to as molecular orbitals) are

$$\psi_{i\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{N_{\rm cell}}}\sum_\alpha C_{i\alpha}(\mathbf{k})\phi_{\alpha\mathbf{k}}(\mathbf{r}), \tag{7}$$

where $N_{\rm cell}$ is the number of unit cells in the crystal.

Gaussian basis sets have been tested for use in crystalline solids [43–45]. There is also a broad use of numerical atomic orbital (NAO) basis sets, which follow a very similar size-dependent nomenclature as Gaussian basis sets. They are employed in, for example, the SIESTA package [47]. We compare the band structures obtained using Gaussian basis sets to the ones obtained using a plane wave code such as QUANTUM ESPRESSO [48], previously employed in QEDARK [5] in appendix B.

Compared to plane waves, Gaussian basis sets are significantly more complex. First of all, Gaussian basis sets are element-specific and the total basis set will be the sum of those from the individual atoms. For a given element, the construction of the basis set requires defining c_μ, N_μ, and ξ_μ for each atomic shell n and angular momentum $l = i+j+k$. The choice of basis presents a source of systematic uncertainty in our calculation, which we estimate by varying the basis sets for each system.

The size of the basis set determines the number of orbitals, with multiple important parameters to consider. Moreover, there are several naming conventions that are widely used. In this paper, we use several basis sets, all taken from [49]. The number of "zetas" in a basis set refers to the number of orbitals as a function of the number of occupied orbitals. An N-zeta (NZ) basis set would have N basis functions for each fully or partly occupied orbital in an atomic species. Si has an electronic configuration of $1s^2 2s^2 2p^6 3s^2 3p^2$ (where the $3s^2 3p^2$ are valence electrons), and so has 3 s-orbitals and 2×3 p-orbitals, and so a DZ (double zeta) basis set would have 6 s- and 4×3 p-orbitals. The treatment of $l \geq 2$ orbitals differs between spherical and cartesian Gaussians. This is because there exists a linear combination of, for example, three d-orbitals in Cartesian Gaussians $(ijk) = (200), (020)$, and (002) orbitals, which mimics (in this case) an s-orbital. In addition, certain basis sets contain N zeta only for the valence orbitals, with one atomic orbital for core states. This is called a NZ-valence or NZV basis set. One can further add polarization functions on the valence orbitals, which

adds new orbitals. For example, for Si, a DZP (double-zeta polarized) basis set would add a d-orbital to allow for the electrons to be polarized in the atoms. In general, addition of extra orbitals and polarization functions improves convergence, though diffuse components in basis sets may become pseudolinearly dependent in periodic calculations [43,45].

It should be noted that the aforementioned details represent a very brief overview of the very complex field of quantum chemical basis sets [50], and different basis sets, especially with different naming conventions can add complexities to these considerations.

We employ the TZP and the def2-TZVP basis set for Si and Ge respectively, and show the uncertainties associated with the choice of basis set in Sec. III C 1.

2. Treatment of core electrons in DFT

The treatment of core electrons, i.e., those tightly bound to the nuclei, is closely connected to the choice of basis set. For atom-centered Gaussians, it is straightforward to include all of the electrons in the DFT calculation, since localized basis functions can just as easily describe core electrons as those near and above the Fermi level. However, for plane waves, describing such localized wave functions would require a prohibitively large energy cutoff. In addition, including core orbitals would require the wave functions of the valence electrons to be orthogonalized to the wave functions of the core electrons; this would introduce rapid oscillations in the region around the atomic nuclei, which would also require too many plane waves to be computationally tractable. Therefore, plane-wave calculations usually freeze the core orbitals via pseudopotentials, effective core potentials (ECPs), or the projector-augmented wave (PAW) method. This has three main effects. First, the core orbitals do not participate in hybridization and bonding in the crystal; this is usually an excellent approximation, since such orbitals are so tightly localized around the atoms, so that there is negligible overlap between atoms. Second, since they are not explicitly included, DM–electron scattering transitions between core orbitals and the conduction band are neglected; this is also not usually an issue, since the energies of such transitions is often beyond the scope of light DM searches. Third, and most crucially, the "pseudowave functions" are smooth in the core region, since they no longer must be explicitly orthogonalized to the core orbitals.

As was shown in [34,41] (see also Sec. IV), all-electron effects are crucial for describing DM–electron scattering events with high momentum transfer, since such events couple high-frequency modes, which are only present in the all-electron valence and conduction bands due to the rapid oscillations near the core. In [34], the all-electron wave functions were recovered after a plane-wave calculation via the PAWs. In Sec. III, we benchmark this methodology against the full all-electron calculation allowed by our Gaussian basis set.

3. Exchange-correlation functional

In practice, the exact form of the exchange and correlation energy in Eq. (1) is not known; however, there are many well-motivated approximations (see Ref. [51]). The choice of exchange and correlation functionals presents a major source of systematic uncertainty in the calculation. This systematic uncertainty can be estimated by calculating the electronic structure with different functionals, and comparing the results. The broad categories of functionals are

(1) Local density approximation (LDA)—$E_{\rm xc}$ is a functional of only n, $E_{\rm xc} \to E_{\rm xc}[n]$.
(2) Generalized gradient approximation (GGA)—$E_{\rm xc}$ is a functional of n and $\boldsymbol{\nabla} n$, $E_{\rm xc} \to E_{\rm xc}[n, \boldsymbol{\nabla} n]$.
(3) Meta-GGAs (mGGA) contain higher order derivatives of n, including terms like $\partial \cdot \partial n$ and $\nabla^2 n$.
(4) Hybrid functional are GGA and mGGA functionals with added exact Hartree-Fock exchange.
(5) Double hybrid functionals add Møller-Plesset perturbation theory at second order ("MP2 level") to hybrid functionals in an effort to better model correlations.

We use the well tested PBE0 hybrid exchange-correlation functional. We further test several GGAs, mGGAs, and other hybrid functionals to show the dependence of DM–electron scattering rates on the choice of exchange-correlation functionals. Moreover, the choice of $E_{\rm xc}$ affects the calculated band gap $E_{\rm gap}$ of the material, and so we apply a scissor correction to match the experimental band gap.

B. Excitation rates in atom-centered bases

1. Theory

As described in Sec. II A, we self-consistently solve the Kohn-Sham equations to obtain the coefficients $C_{i\alpha}(\mathbf{k})$ in Eq. (7). The key quantity required from the DFT calculation for calculating DM–electron scattering is the crystal form factor (equivalent to Eq. (3.17) of [5]),

$$|f_{\rm crystal}(q, E_e)|^2 = \frac{2\pi^2}{E_e} \frac{1}{\alpha m_e^2 V_{\rm cell}} \sum_{ij} \int_{\rm BZ} \frac{V_{\rm cell} d^3k}{(2\pi)^3} \frac{V_{\rm cell} d^3k'}{(2\pi)^3} E_e \delta(E_e - (E_{j\mathbf{k}'} - E_{i\mathbf{k}})) \times \sum_{\mathbf{K}'} q \delta(q - |\mathbf{k}' + \mathbf{K}' - \mathbf{k}|) |f_{[j\mathbf{k}', i\mathbf{k}]}(\mathbf{q})|^2 \Big|_{\theta_q = \theta_U, \phi_q = \phi_U}, \tag{8}$$

where $E_{i\mathbf{k}}$ is the energy of the ith orbital from ground state at $\mathbf{k}$ in the reciprocal cell and $\mathbf{U} = \mathbf{k}' + \mathbf{K}' - \mathbf{k}$. Here, E_e and q are the energy and momentum transferred from the DM particle to the electron, and θ_v and ϕ_v refer to the polar and azimuthal angles of $\mathbf{v}$, respectively. Indices i and j run over occupied and unoccupied orbitals, respectively, and

$$f_{[j\mathbf{k}',i\mathbf{k}]}(\mathbf{q}) = \sum_{\mathbf{R}} e^{-i\mathbf{k}'\cdot\mathbf{R}} \int d^3r C^{\dagger}_{j\beta}(\mathbf{k}')\tilde{G}^{*}_{\beta}(\mathbf{r}-\mathbf{R}) \times e^{i\mathbf{q}\cdot\mathbf{r}}\tilde{G}_{\alpha}(\mathbf{r})C_{\alpha i}(\mathbf{k}). \tag{9}$$

The DM–electron scattering rate for a DM particle of mass m_χ and local density ρ_χ is

$$\frac{dR_{\text{crystal}}}{d\ln E_e} = \frac{\rho_\chi}{m_\chi} N_{\text{cell}} \bar{\sigma}_e \alpha \frac{m_e^2}{\mu_{\chi e}^2} \int d\ln q \frac{E_e}{q} \eta(v_{\min}(q, E_e)) \times |F_\chi(q)^2||f_{\text{crystal}}(q, E_e)|^2 |f_e/f_e^0|^2, \tag{10}$$

where

$$\sigma_e(q) = \bar{\sigma}_e |F_\chi(q)|^2 = \bar{\sigma}_e \left(\frac{(\alpha m_e)^2 + m_V^2}{q^2 + m_V^2}\right)^2$$

is the DM–electron interaction cross section assuming a bosonic mediator with mass m_V. In general, we assume two limits, $m_V \gg q$ (called the heavy-mediator limit) and $m_V \ll q$ (called the light-mediator limit). The factor $|f_e/f_e^0|^2$ is a screening factor discussed next (see also [34]), while $\eta(v_{\min}(q, E_e))$ is the average inverse speed of DM in the galaxy (for more details, see Ref. [5] or Appendix C) and $\mu_{\chi e}^{-1} = m_\chi^{-1} + m_e^{-1}$. We describe the calculation of the matrix elements, Eq. (9), in Appendix A.

If the DM–electron interaction is mediated by a dark photon or scalar, the interaction is screened due to the in-medium effects [36,52]. Recent works [32,53] have emphasized the importance of this electrostatic screening, especially for recoils at low energy transfer. In the results shown below, we follow the prescription of [34,35], and multiply the crystal form factor, $|f_{\text{crystal}}|^2$ by a factor of $|f_e/f_e^0|^2$, with $f_e/f_e^0 = (\hat{\mathbf{q}}\cdot\epsilon\cdot\hat{\mathbf{q}})^{-1}$. Here ϵ is the dielectric function, which is modeled as [54]

$$\epsilon(q, E_e) = 1 + \left[\frac{1}{\epsilon_0 - 1} + \tau\left(\frac{q}{q_{\text{TF}}}\right)^2 + \frac{q^4}{4m_e^2\omega_p^2} - \left(\frac{E_e}{\omega_p}\right)^2\right]^{-1}, \tag{11}$$

TABLE I. Parameters used in the dielectric function calculation in Eq. (11) for Si and Ge from [54].

Target	ϵ_0	τ	ω_p (eV)	q_{TF}(KeV)
Si	11.3	1.563	16.6	4.13
Ge	14.0	1.563	15.2	3.99

where $\epsilon_0 \equiv \epsilon(0,0)$ is the empirically measured static dielectric constant, τ is a fitting parameter (we use the results from [54], which fit their dielectric function to the results of [55]), ω_p is the plasma frequency, and q_{TF} is the Thomas-Fermi momentum, with values listed in Table I. This equation ignores the tensorial nature of ϵ, since the crystals we consider here are cubic. This equation only estimates the real part of the dielectric function, and we do not account for the imaginary part in our screening.

Ref. [35] compares the differences in the expected DM–electron scattering rates using the analytical model in Eq. (11) for the dielectric function with an RPA calculation of the dielectric function. We expect an $\mathcal{O}(10\%)$ correction for creating 3 or fewer electrons-hole pairs ($Q \leq 3e^-$) with a numerically calculated dielectric function using a preliminary RPA dielectric function calculation, and $\mathcal{O}(1\%)$ or less for creating more than 3 electron-hole pairs.

2. Numerical implementation

In practice, the numerical integration in $\mathbf{k}$-space requires the replacement

$$\int_{\text{BZ}} \frac{V_{\text{cell}} d^3k}{(2\pi)^3}(\ldots) \to \frac{1}{N_{\mathbf{k}}}\sum_{\mathbf{k}}(\ldots), \tag{12}$$

where $N_{\mathbf{k}}$ is the number of $\mathbf{k}$-points in the chosen $\mathbf{k}$-grid.

Moreover, to discretize $|f_{\text{crystal}}(q, E_e)|^2$ in q and E_e, we define bins with width Δq and ΔE_e and bin centers $\{q_n\}$ and $\{E_e^m\}$, respectively. Then, the discretization procedure follows as

$$|f_{\text{crystal}}(q_n, E_e^m)|^2 \equiv \int_{q_n-\frac{1}{2}\Delta q}^{q_n+\frac{1}{2}\Delta q} \frac{dq}{\Delta q} \times \int_{E_e^m-\frac{1}{2}\Delta E_e}^{E_e^m+\frac{1}{2}\Delta E_e} \frac{dE_e}{\Delta E_e} |f_{\text{crystal}}(q, E_e)|^2. \tag{13}$$

Consequently, the numerical crystal form factor is

$$|f_{\text{crystal}}(q_n, E_e^m)|^2 = \frac{2\pi^2}{E_e^m} \frac{1}{\alpha m_e^2 V_{\text{cell}}} \frac{1}{N_{\mathbf{k}}^2} \sum_{ij} \sum_{\mathbf{k},\mathbf{k}'} \sum_{\mathbf{K}'} \frac{E_e^m}{\Delta E_e} \frac{q_n}{\Delta q} |f_{[j\mathbf{k}',i\mathbf{k}]}(\mathbf{k}' + \mathbf{K}' - \mathbf{k})|^2 \times \Theta\left(1 - \frac{|E_{j\mathbf{k}'} - E_{i\mathbf{k}} - E_e^m|}{\frac{1}{2}\Delta E_e}\right) \Theta\left(1 - \frac{||\mathbf{k}' + \mathbf{K}' - \mathbf{k}| - q_n|}{\frac{1}{2}\Delta q}\right). \tag{14}$$

Here, as for Eq. (8), the recoil electron is assumed to be transferred from occupied molecular orbital $|i, \mathbf{k}\rangle$ to unoccupied orbital $|j, \mathbf{k}' + \mathbf{K}'\rangle$.

In principle, there should be $n_k = N_T$ points modeled in the calculation, where N_T is the number of unit cells in this crystal, with $N_T \gtrsim \mathcal{O}(10^{23})$ for a 10 g crystal. However, such a dense mesh is impractical. Since we need to calculate rates for transition from each $\mathbf{k}$ to each $\mathbf{k}'$, the computational complexity scales as n_k^2, and the calculation quickly becomes unfeasible.

We generally model the grid in reciprocal space as a Γ-centered Monkhorst-Pack grid, with $n_{k,i}$ points along the b_i reciprocal lattice vector. We use the shorthand $n_{k,1} \times n_{k,2} \times n_{k,3}$ to denote the mesh in the reciprocal space, which has $n_k = \prod_{i=1}^{3} n_{k,i}\mathbf{k}$-points. Because both Si and Ge crystallize in FCC cells, it is reasonable to set $n_{k,1} = n_{k,2} = n_{k,3}$. We choose $n_{k,i} = 4$ and 6 for Si and Ge, respectively, but check the dependence of DM–electron scattering rates on our choice of $\mathbf{k}$-grid in Sec. III C 3.

In addition, it becomes computationally expensive to include a large number of $\mathbf{K}$ vectors, and so in practice, we must limit the vectors to some $q_{\max}$. The choice of $q_{\max}$ is another source of systematic error. We choose $q_{\max} = 25\alpha m_e$ and $q_{\max} = 20\alpha m_e$ for Si and Ge, respectively, and further show the dependence of DM–electron scattering rates on $q_{\max}$ in Sec. III C 5.

III. RESULTS AND DISCUSSION

In this section, we describe the results from our calculation of the crystal form factors [Eq. (14)] and the DM–electron scattering rates [Eq. (10)] for Si and Ge performed with QCDARK. We evaluate the various systematic uncertainties, and compare our results to those from other available codes. Finally, we look at the annual modulation rate as a function of the DM mass, m_χ and the DM form factor, F_χ.

TABLE II. Parameters used for DFT calculation of electronic structure of Si and Ge crystals. E_{gap} refers to the scissor corrected band gap we employ for the materials.

Target	Crystal structure	Lattice constant (Å)	E_{gap} (eV)
Si	FCC	5.43	1.11
Ge	FCC	5.65	0.67

Table II lists the values of crystal parameters used for our calculation of DM–electron scattering rates in both Si and Ge, including the experimental band gap used for the scissor correction procedure.

A. Crystal form factor

The calculated crystal form factors using Eq. (14) are shown in Fig. 1, for silicon and germanium crystals in the left and right panels respectively. The region below the black lines are kinematically inaccessible for halo DM, i.e., the halo DM, irrespective of the mass of the fermion, is unable to transfer energy E_e with a momentum transfer $q < E_e/(v_{\text{Escape}} + v_{\text{Earth}})$.

Both panels show an enhancement of the crystal form factor at $q \gtrsim 4\alpha m_e$ compared to Fig. 5 of [5], which is a consequence of including all-electron effects in our calculation. The enhancement in the crystal form factor for germanium crystals at $E_e \gtrsim 30$ eV corresponds to the transitions from the semicore 3d-shell to the conduction bands, which has an energy of −28.6 eV relative to the top of the valence bands in our calculation.

B. Dark matter–electron scattering rates

Figure 2 shows the DM–electron scattering rates expected in silicon and germanium crystals for $m_\chi = 10$ MeV, 100 MeV, and 1 GeV for both heavy and light mediators, assuming the crystal form factors shown in Fig. 1. In this and

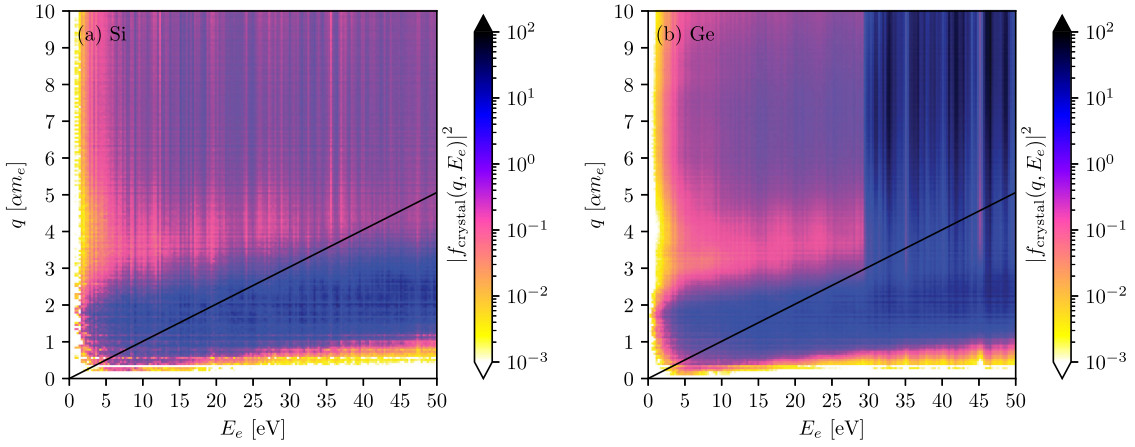

FIG. 1. Panels (a) and (b) show the calculated crystal form factor $|f_{\text{crystal}}(q, E_e)|^2$ [see Eq. (14)] for silicon and germanium, respectively. The details of the calculation are described in §III A. The region beneath the black line is kinematically inaccessible for halo DM, as it would require $v_{\min}(q, E_e) > v_{\text{Escape}} + v_{\text{Earth}}$ (see Appendix C for more details).

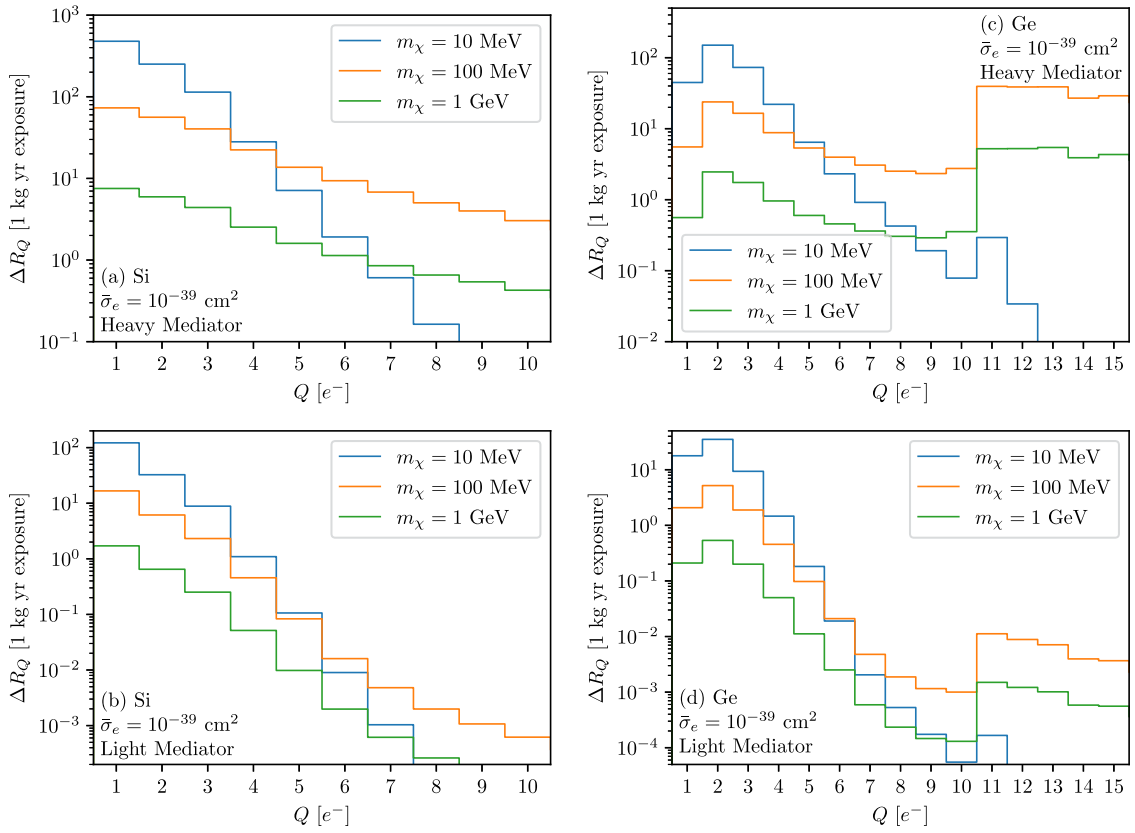

FIG. 2. Panels (a) and (b) show the DM–electron scattering rates in a silicon crystal for heavy and light mediators respectively. These values are calculated using a TZP basis set with PBE0 exchange and correlation functionals, $4 \times 4 \times 4$ **k**-grid, and $q_{\max} = 25\alpha m_e$. Panels (c) and (d) show the DM–electron scattering rates in a germanium crystal with interaction mediated by a heavy and a light mediator respectively calculated using a def2-TZVP basis set with PBE0 exchange–correlation functional, $6 \times 6 \times 6$ **k**-grid, and $q_{\max} = 20\alpha m_e$.

subsequent figures, we plot

$$\Delta R_Q = \int dE_e \frac{dR}{dE_e} p(Q, E_e), \tag{15}$$

where $p(Q, E_e)$ is the probability that a transition with recoil energy E_e excites Q electrons (for more details, see Sec. III D). We use the ionization model at 100 K from [56] for Si. For Ge, we use an electron-hole-pair creation model,

$$Q = \sum_{n=0}^{\infty} \Theta\left(E_e - n \times E_{\rm pp} - E_{\rm gap}\right), \tag{16}$$

where $\Theta(x)$ is the Heaviside step function, $E_{\rm pp}$ is the electron-hole-pair creation energy ($E_{\rm pp} = 2.9$ eV for Ge), and $E_{\rm gap}$ is the band gap of the material.

Panels (a) and (b) of Fig. 2 show DM–electron scattering rates in a silicon crystal mediated by a heavy and a light mediator, respectively. For Si, we use a TZP basis set with a PBE0 exchange–correlation functional, $4 \times 4 \times 4$ **k**-grid and $q_{\max} = 25\alpha m_e$.

Panels (c) and (d) of Fig. 2 show DM–electron scattering rates in a germanium crystal mediated by a heavy and a light mediator, respectively, calculated using a def2-TZVP basis set with a PBE0 exchange–correlation functional, $6 \times 6 \times 6$ **k**-grid and $q_{\max} = 20\alpha m_e$.

Because the 3d-dominated bands in Ge are flat in **k**-space (i.e., they are highly localized in real space), we need a denser **k**-grid to reduce the numerical noise in the (unbinned) rate spectra dR/dE_e. An ionization model for Ge akin to the model in [56] for Si (which includes a Fano factor) remains unavailable, which would smooth out the numerical noise while calculating ΔR_Q (see Sec. III D).

In principle, increasing the density of the **k**-grid would reduce the noise, at the expense of computation time. However the rates, barring ∼10% systematics coming from the numerical noise at high Q, are robust (see Sec. III C 3 and Fig. 6 for more details), and indicate that for $m_\chi \gtrsim \mathcal{O}(100\ \text{MeV})$ the rates are higher for $Q \geq 11e^-$ than for lower Q (for lower masses, interactions with high q transfers are kinematically suppressed). This would imply that germanium-based detectors with relatively high thresholds can still probe significant regions of DM parameter space, assuming $m_\chi \gtrsim 100$ MeV and a heavy mediator.

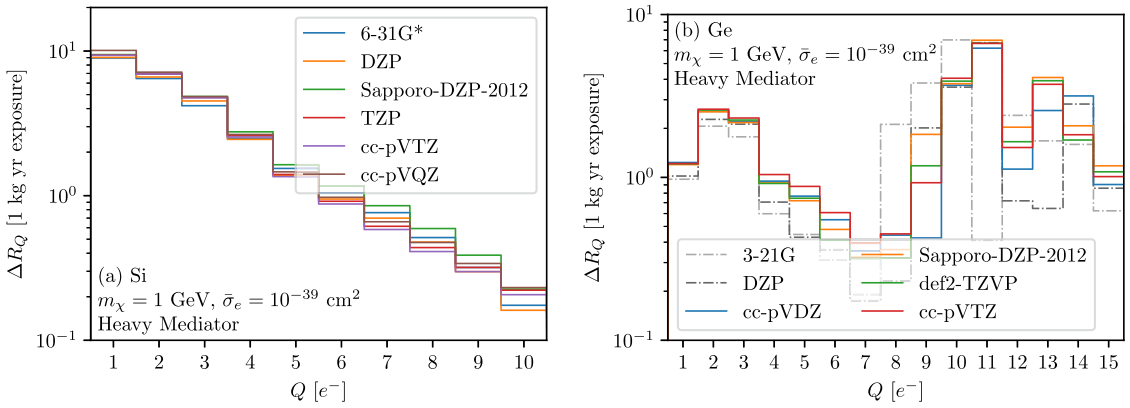

FIG. 3. Panels (a) and (b) show the DM–electron scattering rates for Si and Ge respectively, calculated with a $4 \times 4 \times 4$ **k**-grid, with DM parameters noted in the figures. The plots use $q_{\max} = 10\alpha m_e$ and PBE exchange-correlation. `cc-pVQZ` is the most accurate basis set we test for Si, while `cc-pVTZ` is the same for Ge. Moving further, we use TZP and def2-TZVP for Si and Ge respectively, which mimic the most accurate bases at lower computational costs.

The effects of all–electron modes are visible for heavy mediators, and not as much for light mediators. This is because of an effective $|F_\chi|^2 \propto 1/q^4$ suppression in the DM–electron scattering cross-section in the case of light mediators.

C. Evaluation of systematic uncertainties

There are multiple sources of theoretical uncertainties as well as several convergence parameters (i.e., parameters that can be improved with more computational time) in our calculation of DM–electron scattering rates. The choice of the exchange–correlation functional, $E_{\mathrm{xc}}[n]$, in Eq. (1) is a source of theoretical uncertainty. The real-space cut-off for constructing our Bloch atomic orbitals, the size of the **k**-grid, and the choice of $q_{\max}$ are convergence parameters. The choice of the atomic centered Gaussian basis set is both a convergence parameter (since increasing the number of basis functions allows us to model the conduction states better) and a theoretical uncertainty (since different basis sets are optimized for different types of calculations, be it molecular or periodic boundary conditions). In this section, we go through each of these choices and determine their effects on our DM–electron scattering rate calculation. Note that it is computationally intractable to perform all of the tests of numerical convergence/uncertainty at the fully converged $q_{\max}$ we will finally adopt. Therefore, we perform such tests at somewhat lower values, mentioned in each section. In appendix D we explore the effects of the lower $q_{\max}$ on our conclusions about systematic uncertainty.

1. Choice of basis set

There are many choices of atom-centered Gaussian basis sets available for use [49]. However, most of these basis sets are optimized for molecular calculations, and we have to choose among the few optimized for periodic boundary conditions. In addition, the size of the basis sets determines the number of conduction bands.

Figure 3 shows the DM–electron scattering rates calculated for various basis sets. While the DM–electron scattering rates are consistent across all the basis sets we test, the cc-pVQZ (correlation-consistent polarized valence quadruple zeta) and cc-pVTZ (correlation-consistent polarized valence triple zeta) are the best optimized basis sets that we test for Si and Ge, respectively (these are also computationally very expensive). For Si, DM–electron scattering rates calculated using TZP (shown in Fig. 2) differ by only ∼5% from those derived using cc-pVQZ. Similarly for Ge, DM–electron scattering rates calculated using def2-TZVP (shown in Fig. 2) differ by only ∼5% on average from those derived using cc-pVTZ. Both of the TZP and def2-TZVP basis sets provide a good balance of computational efficiency and accuracy, and we use these in further analyses.

Figure 3 also shows that DM–electron scattering rates in Ge are heavily dependent on the choice of basis set, especially for large $Q \gtrsim 9e^-$. This is because basis sets like 3-21G and DZP are unable to capture conduction bands well, while the energy of the semicore 3d electrons is highly dependent on accurate modeling of core shells.

Note that we use a $q_{\max} = 10\alpha m_e$ in Fig. 3, which is lower than the fully converged value. We show in appendix D that for low charge ionization $Q \lesssim 5e^-$, the basis sets converge at the same rate with as a function of $q_{\max}$ for $q_{\max} \gtrsim 9\alpha m_e$, and hence we can safely compare them at $q_{\max} = 10\alpha m_e$. For higher charge ionizations, the basis sets differ in their $q_{\max}$ convergence, since they differ in their treatment of high-lying conduction bands. Therefore, the comparison between basis sets at $Q \gtrsim 6e^-$ should be taken with caution.

2. Real space cutoff

Our atomic orbitals are Bloch sums in real space as in Eq. (6); in principle, one must sum over an infinite number of Gaussians displaced by real-space lattice vectors $\mathbf{R}$ to form each atomic orbital. In practice, however, Gaussians are rapidly decaying functions, and so it suffices to include a finite number of neighbors depending on the exponents in the contracted Gaussians. This is generally accomplished by setting a real space cutoff, which we call $R_{\rm cut}$.

While PySCF is capable of choosing a dynamic real space cut off for each orbital, which lowers the computational cost to calculating the matrix elements [see Eq. (9)], this complicates our analytical approach to calculating the matrix elements (see Appendix A for more details on an analytical calculation). Hence, we choose a constant $R_{\rm cut}$ for all orbitals, the value of which is chosen via the following procedure. We first calculate the electron loss function, i.e., the imaginary part of the inverse dielectric function $\mathrm{Im}[-\epsilon(\omega,\mathbf{q})^{-1}]$. For this, we assume that the real part of the dielectric function is modeled by Eq. (11), and calculate the imaginary part of the RPA dielectric function using Eq. (16) of [57]. We then integrate $\omega\mathrm{Im}[-\epsilon(\omega,\mathbf{q})^{-1}]$ over energy for a given magnitude of $\mathbf{q}$; by the f-sum rule,

$$\int_0^\infty dE_e E_e \mathrm{Im}\left\{-\frac{1}{\epsilon(E_e,q)}\right\} = \frac{\pi}{2}\omega_{\rm p}^2 \tag{17}$$

the result should equal $\pi/2\omega_{\rm p}^2$, where $\omega_{\rm p}$ is the plasma frequency of the material. The f-sum rule is only achieved in limit of a complete basis set, however we have found that the convergence of this quantity is a useful diagnostic as to whether a given $R_{\rm cut}$ is sufficient for the relevant range of q [53]. We show the electron loss function integrated over E_e up to 50 eV for Si (using the TZP basis set) and Ge (using the def2-TZVP basis set) in Fig. 4.

The high momentum transfer modes $q \gtrsim 3\alpha m_e$ are captured well even by low $R_{\rm cut}$, except $R_{\rm cut} = 1$ cell for Si. Hence, for Si, we disregard this result. Moreover, for Ge we observe that $R_{\rm cut} = 1$ and 2 cells match with $R_{\rm cut} = 4$ cells for $q \gtrsim 2\alpha m_e$ and $1\alpha m_e$ respectively.

One may understand this in the real space from the viewpoint of our KS wave functions—the high frequency modes of the molecular orbitals, resulting from orthogonalization with inner orbitals, are more localized near the nuclei, and hence have smaller overlaps in real space with counterparts from more distant atoms. The low q modes, on the other hand, correspond to long distance behavior of the matrix elements, and so necessitate the usage of larger $R_{\rm cut}$ in order to satisfy the f-sum rule. The curves going to zero at very low $q \lesssim 0.3\alpha m_e$ is an artifact of our choice of a sparse $\mathbf{k}$-grid, and is not the true behavior of f sum rule.

Motivated by these results, our final calculations use a hybrid real space cutoff. For Si (Ge), we use $R_{\rm cut} = 4(3)$ cells for $q \leq 3\alpha m_e$ and $R_{\rm cut} = 2(1)$ cells otherwise. Figure 5 shows that this is also cautious, as the low momenta deviation only occurs in a prohibited region of the parameter space $(v_{\rm min}(q,E_e) > v_{\rm Escape} + v_{\rm Earth})$ for DM–electron scattering (see Appendix C for more details).

3. Convergence of k-mesh in reciprocal space

A potential source of systematic error in the rate calculation comes from the density of k-points in the first Brillouin zone (1BZ). As discussed above, because the computational cost scales as the square of the number of k-points, it is infeasible to include N_T k-points, where $N_T \gtrsim \mathcal{O}(10^{23})$ is the number of unit cells in the crystal. In this section, we discuss the effects of modeling the 1BZ with an $N \times N \times N$ $\mathbf{k}$-grid, with a total of N^3 k-points.

Figure 6 shows the convergence of our calculations with $\mathbf{k}$-grid for both Si and Ge. Note that Si is already converged at a $4 \times 4 \times 4$ $\mathbf{k}$-grid. One reason for this is that the

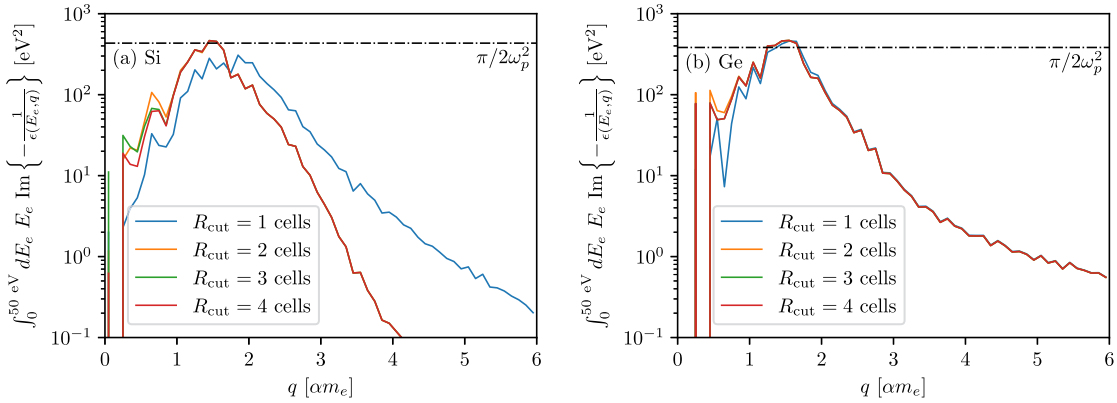

FIG. 4. Panels (a) and (b) show the electron loss functions integrated over electron recoil energy, E_e, for Si and Ge, respectively, with the dash-dotted line showing the theoretical upper bound from the f-sum rule. For Si, we use here a TZP basis set, at $4 \times 4 \times 4$ $\mathbf{k}$-grid and PBE functional to calculate these results. For Ge, we use here a def2-TZVP basis set, with the same $\mathbf{k}$-grid and PBE functional. Note that a good description of high momentum transfer $q \gtrsim 3\alpha m_e$ does not require a large $R_{\rm cut}$ for either element.

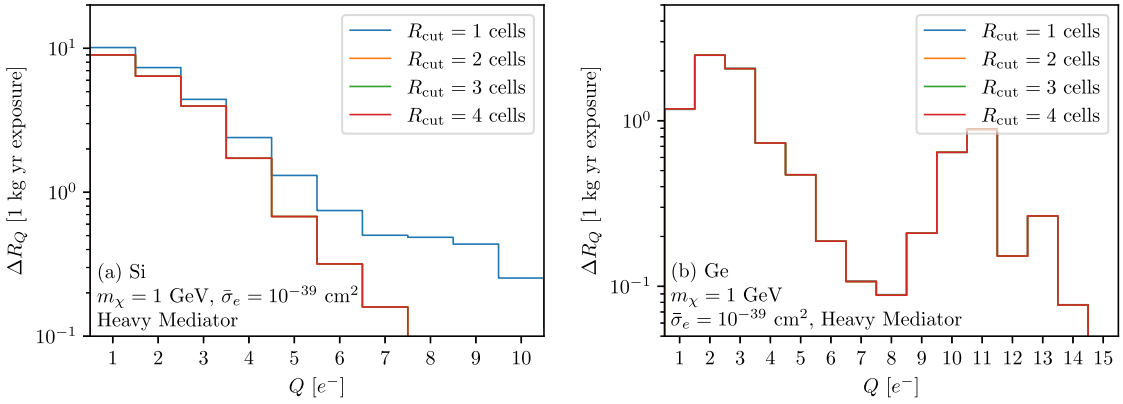

FIG. 5. Panels (a) and (b) show the DM–electron scattering rates in Si and Ge, respectively, for various values of the real space cut-off $R_{\rm cut}$, assuming an exposure of 1 kg − year. We use a $4 \times 4 \times 4$ **k**-grid, set $q_{\rm max} = 6\alpha m_e$, and use the PBE functional.

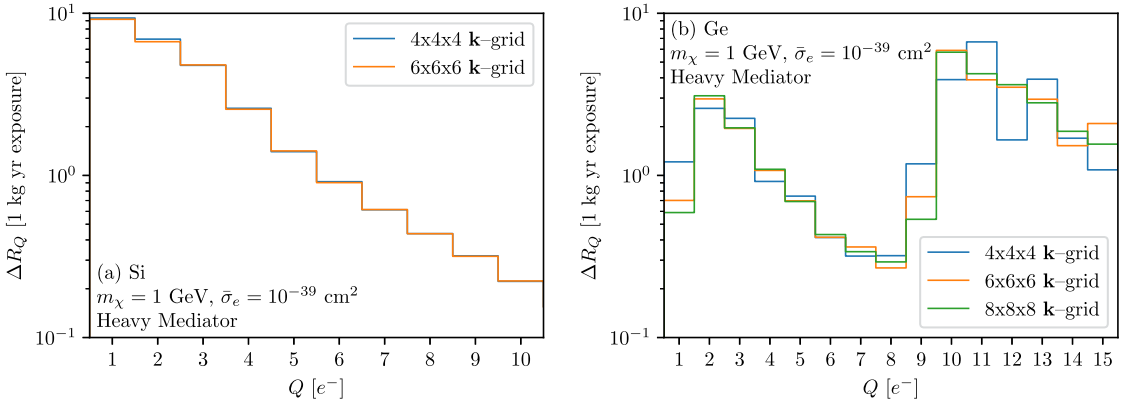

FIG. 6. Panels (a) and (b) show DM–electron scattering rates for Si and Ge, respectively, for various **k**-grid densities, assuming an exposure of 1 kg − year. We use the PBE exchange–correlation functional and set $q_{\rm max} = 10\alpha m_e$. It is evident that the Si calculation is converged, even at the sparse $4 \times 4 \times 4$ **k**-grid level, while Ge only converges at higher **k**-grid densities.

numerical uncertainties are smoothed out from applying the ionization model (see Sec. III D for more details). For Ge, the dispersionless $3d$-derived band at $E_e \sim 29$ eV below the Fermi level results in less distinct energies sampled, thus requiring a finer **k**-grid for convergence. We find that a $6 \times 6 \times 6$ grid performs adequately, with errors of $\lesssim$15% compared to $8 \times 8 \times 8$ in each bin. A probabilistic ionization modeling for Ge, akin to 56 for Si, will aid in smoothing out the recoil spectrum.

Given a basis set and an exchange–correlation functional, a choice of **k**-grid is a convergence parameter, and therefore is converged independently of $q_{\rm max}$.

4. Exchange-correlation functional

Figure 7 shows DM–electron scattering rates in Si and Ge crystals for various exchange-correlation functionals. We test the commonly used PBE GGA functional, along with SCAN and TPSS mGGAs. For Si, we test multiple hybrids—PBE0, SCAN0 and TPSS0, along with a hybrid semiempirical functional optimized for molecules rather than crystals (B3LYP). For Ge, we test PBE, SCAN, TPSS, and PBE0.

It is important to note that a scissor correction has been applied to the band gaps of Si and Ge, so DM–electron scattering calculations have the same gap regardless of functional. For materials where the experimental gap is not known, the differences in gaps predicted by different functionals is expected to lead to a significant source of variation in the scattering rates. A related issue observed for Ge (right panel of Fig. 7) is the dependence of the energy of the 3d shell. This results in significant differences in the DM–electron scattering rates in the 8–11 electron-hole-pair bins. It is apparent that PBE, TPSS, and SCAN underestimate the electron binding energy for the $3d$-shell electrons, with values ∼25 eV from the top of the valence band. The PBE0 functional results in values between 28.6 and 29.0 eV, which are much closer to the experimental values of ∼29.5 eV of 3d-shell electrons, respectively [58]. For Si, we scissor correct the band gap, and core orbitals do not get involved until energies of ∼99.2 eV [58].

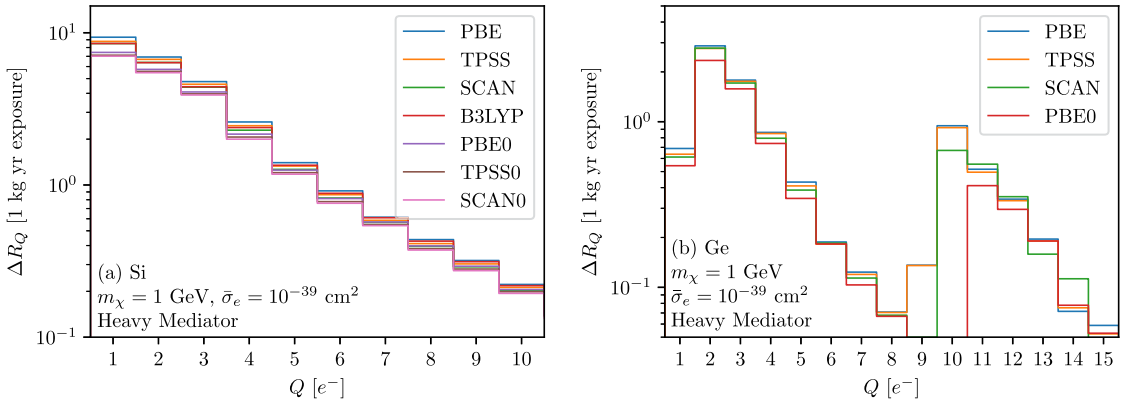

FIG. 7. Panels (a) and (b) show DM–electron scattering rates for Si and Ge, respectively, calculated with various exchange-correlation functionals, assuming an exposure of 1 kg − year. We use a $4 \times 4 \times 4$ **k**-grid for Si and $6 \times 6 \times 6$ **k**-grid for Ge, and set $q_{\max} = 6\alpha m_e$. We favor the well-tested PBE0 functional for our calculations. Note the dependence of the energy of the 3d-shell of Ge on the choice of $E_{\rm xc}$.

We use $q_{\max} = 10\alpha m_e$ in Fig. 7 rather than the higher reference value mentioned in Sec. III A. In appendix D we show that this value is sufficient such that all exchange-correlation functions converge at the same rate as a function of $q_{\max}$, and thus can be directly compared.

5. Maximum momentum transfer

The implementation of atom-centered basis sets without using an effective core potential has one direct effect—we are able to capture the high-momentum transfer regime of the crystal form factor. These high-q contributions come from orthogonalizing the valence and conduction bands against the core orbitals, which introduces high wave number modes to the valence and conduction wave functions. This allows the wave functions to be modeled to arbitrarily high wavenumbers, and allows us to fully capture the crystal form factor. Figure 8 shows the impact of adding high q modes on the rates of 1 GeV DM particle interacting with Si (left panels) or Ge (right panels) via a heavy (top row) or light (bottom row) mediator.

As expected, DM–electron scattering mediated by a light mediator is not significantly influenced by the high-q contributions in silicon. This is due to the $|F_\chi(q)|^2 \propto q^{-4}$ dependence of the rate in the integrand of Eq. (10). Ge, on the other hand, is sensitive to $q_{\max}$ even for scattering through a light mediator, since the $3d$-shell in germanium dominates the high q regime.

For interactions mediated by a heavy boson, however, there are important high-q contributions even for relatively small charge bins. Moreover, when including the high-q contributions, we see that DM with $m_\chi \gtrsim 50$ MeV and scattering through a heavy mediator ($F_\chi \approx 1$), the rates from $Q \geq 11e^-$ bins dominate over the $1e^- \leq Q \leq 10e^-$ bins. Similarly, for Si with a heavy mediator, high q contributions are important, with, e.g., a ∼75% increase in rates if we go from $q_{\max} = 8\alpha m_e$ to $q_{\max} = 25\alpha m_e$ for the $10e^-$ bin.

D. Effects of the secondary ionization model for silicon

Our results for the DM-electron scattering rates in silicon are shown using the ionization modeling from [56]. In Fig. 9, we compare these rates with those from a simple step function model from Eq. (16) for Si. For the latter, we use $E_{\rm gap} = 1.1$ eV and $E_{\rm pp} = 3.8$ eV (see, e.g., [59]). We see significant differences between the two ionization models for the 1 e^- and 2 e^--bin, although the rates are similar for the bins with $Q \geq 3$. As we observed for Ge, for which only a step-function model is available, the probabilistic model from [56] smoothes out the numerical fluctuations introduced by the sparse **k**-grid.

E. Annual modulation

The DM–electron scattering rates are dependent on the DM flux incident on the target material, which in turn depends on the velocity of the detector in the galactocentric frame. For table-top experiments, there are three major contributions to this velocity. First, there is the local circular velocity, which we take to be $v_0 = 230$ km s^{-1}. The second contribution comes from the Sun's peculiar velocity, $\mathbf{v}_\odot - \mathbf{v}_0$. We use the recommended value, $v_\odot = 250.2$ km s^{-1} [60]. Finally, the earth revolves around the sun with an average speed $\langle v_\oplus \rangle = 29.8$ km s^{-1}. This revolution causes an annual modulation in DM–electron scattering rates, as the total velocity, $v_{\rm Earth} = |\mathbf{v}_\odot + \mathbf{v}_\oplus|$, varies from 220.4 km s^{-1} on December 2 to 280 km s^{-1} on June 2.

We calculate the modulation amplitude following [5],

$$f^Q_{\rm mod} = \frac{\Delta R_{Q,\rm Jun\,2} - \Delta R_{Q,\rm Dec\,2}}{2\Delta R_{Q,0}}, \tag{18}$$

where $\Delta R_{Q,0} = \Delta R_{Q,\rm Sept\,2} = \Delta R_{Q,\rm Mar\,2}$. Even in the presence of backgrounds, a measurement of $f_{\rm mod}$ could allow

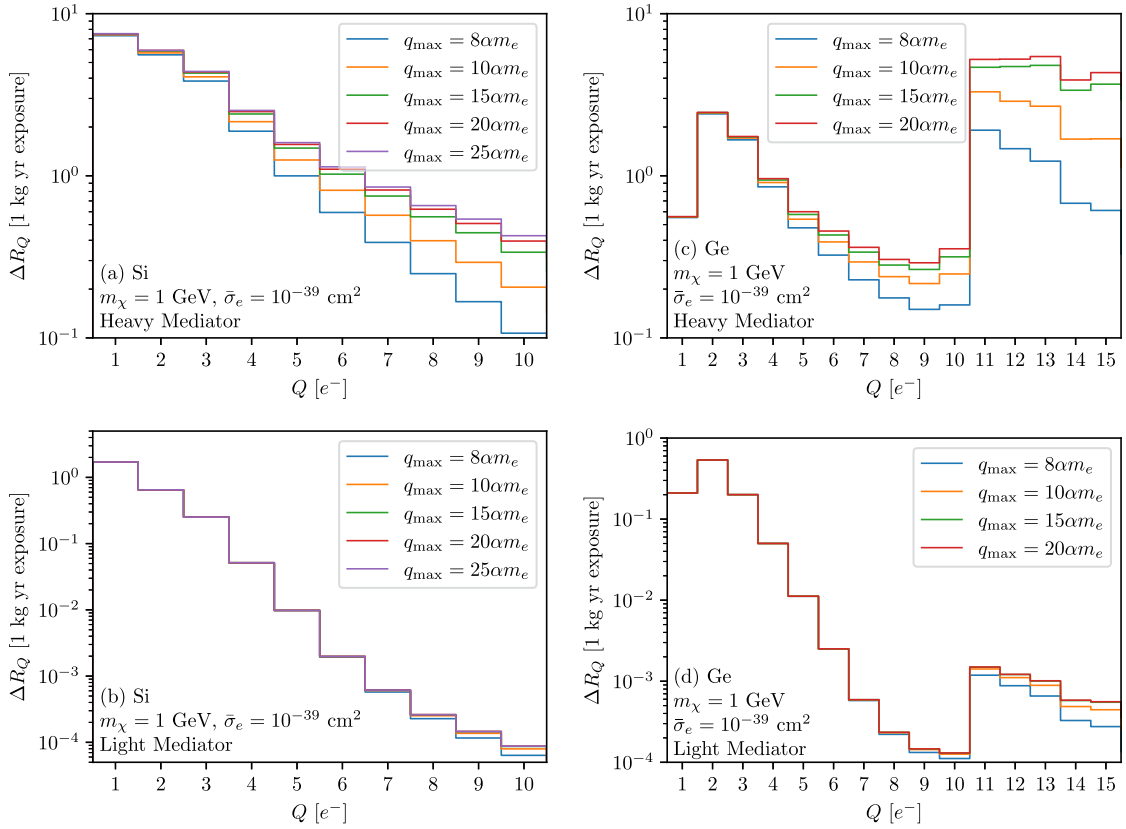

FIG. 8. Panels (a) and (b) shows DM–electron scattering rates for Si with DM–electron interaction mediated by a heavy and a light mediator, respectively, with different $q_{\max}$ cutoffs, assuming an exposure of 1 kg – year. Panels (c) and (d) show the same for a Ge crystal. We use a $4 \times 4 \times 4$ **k**-grid for Si and a $6 \times 6 \times 6$ **k**-grid for Ge, and the PBE0 exchange correlation funtional.

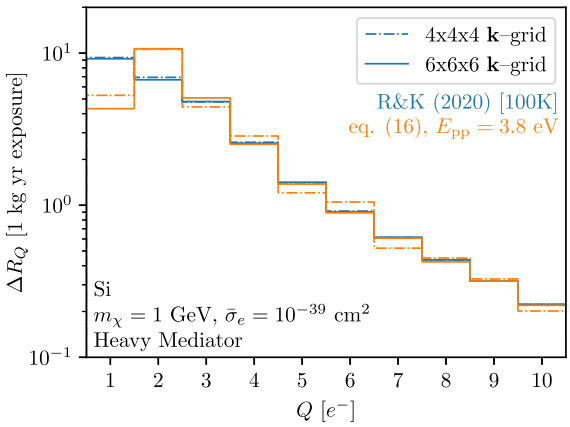

FIG. 9. The effects on the DM-electron scattering rates in silicon of using the secondary ionization modeling from [56] ("R&K") versus the step-function model from Eq. (16). We use the PBE exchange-correlation functional and $q_{\max} = 10\alpha m_e$.

for the detection of DM in such an experiment. We plot $f_{\rm mod}$ as a function of the DM mass, the mediator form factor, and the charge ionized in the target material Q for both Si and Ge in Fig. 10.

Comparing to Fig. 8 of [5], the difference in rates from the inclusion of high wave number modes in the crystal form factor allows the electron to scatter into a larger parameter space, which reduces $f_{\rm mod}$, especially for the $m_\chi = 1$ GeV case. The same is visible for Ge, and is in fact even more pronounced for the 3d-shells, which dominate the rate.

A measurement of the annual modulation signal will be an important step in confirming a potential DM signal. We calculate the 5σ-sensitivity by requiring

$$\frac{\Delta S}{\sqrt{S_{\rm tot} + B}} \geq 5, \tag{19}$$

where $\Delta S = f_{\rm mod} S_{\rm tot}$ is the modulation amplitude, $S_{\rm tot}$ is the total number of signal events, and B is the number of background events. Here $f_{\rm mod}$ is calculated using Eq. (18), except we sum the rates over $1e^- \leq Q \leq 10e^-$ for Si, and $1e^- \leq Q \leq 15e^-$ for Ge. Assuming no background events, we show the 5σ discovery reach in Fig. 11 with dash-dotted lines for Si (black) and Ge (blue). The left panel shows the reach for heavy mediators, while the right panel corresponds to light mediators.

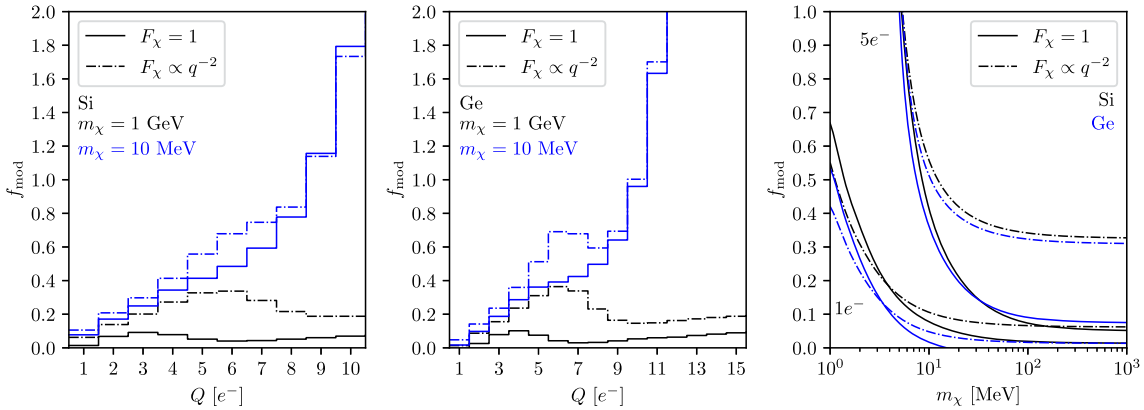

FIG. 10. The modulation amplitude, $f_{\rm mod}$ from Eq. (18), versus Q for Si (left) and Ge (middle), for $m_\chi = 10$ MeV and 1 GeV, for both heavy and light mediator-mediated scattering. The right panel shows $f_{\rm mod}$ versus m_χ, for $Q = 1e^-$ and $Q = 5e^-$, for both Si and Ge. We use the PBE0 exchange-correlation functional, along with $q_{\rm max} = 25\alpha m_e(20\alpha m_e)$ and a $4\times4\times4(6\times6\times6)$ **k**-grid for Si (Ge).

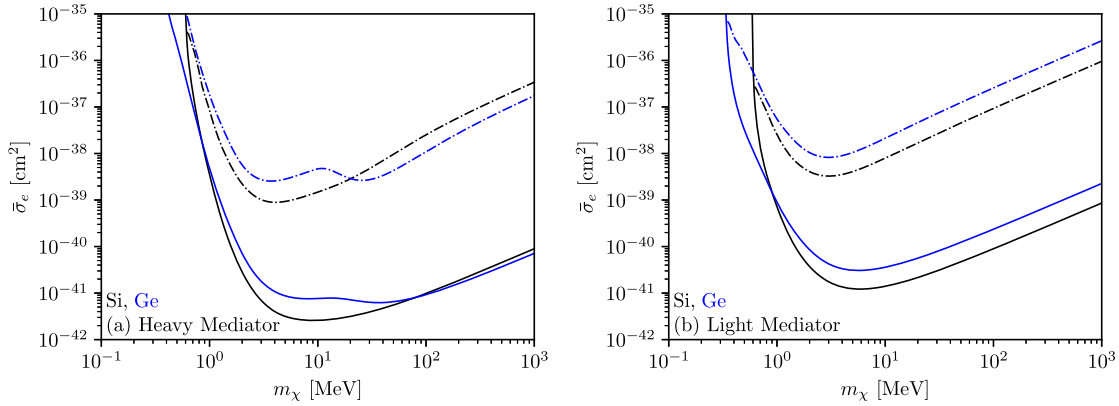

FIG. 11. This plot shows the reach for both, obtaining 2.3 events for 1 kg · yr exposure of our target material (solid line), as well as the threshold for a 5-σ discovery by annual modulation with the same exposure (dash-dotted line). Panel (a) shows the upper bounds on cross-section that can be placed for heavy mediators, while panel (b) shows the same for light mediators. We assume no background for either panel, and include $1 \le Q \le 10\ e^-$ for Si (black) and $1 \le Q \le 15e^-$ for Ge (blue).

F. Comparison with other codes

QUANTUM ESPRESSO was used in the first numerical calculation of the crystal form factor for Ge in [1]. It was also used in [5], which presented a detailed calculation of the crystal form factor for both Ge and Si, and made the resulting code, QEDARK, publicly available. In Fig. 12, the crystal form factor is recalculated with QEDARK with improved computational parameters, including a higher energy cutoff for plane wave calculations, a denser **k**-grid for both Si and Ge, and the analytical screening described in Sec. II B 1. We use the PBE functional, which, as shown above, underestimates the energy of the 3d-shell in Ge, and also excludes the effects of high frequency modes that are visible in both Si and Ge, especially for heavy mediators.

DarkELF [32] emphasized the need for better screening, especially for low energy excitations. Here we use GPAW RPA dielectric function with local field effects (LFE) for both Si and Ge. However, it also does not include the high frequency modes, which dominate the rates at high energy, and the Ge 3d-shell is frozen in the pseudopotential, which otherwise dominate the DM–electron scattering rates at $E_e \gtrsim 30$ eV.

EXCEED-DM [35] is able to reconstruct the high-frequency modes, and is also able to capture the dielectric screening with an RPA dielectric function. It also employs the well-tested and commonly employed HSE06 functional. However, it reconstructs the semi-core and core orbitals after a pseudopotential calculation. We find good agreement between the rates calculated with EXCEED-DM and `QCDark`, showing that the PAW method is accurate in these materials.

`QCDark` implements *ab-initio* calculation of the crystal form factor, along with an analytical approximation to the dielectric function. However, as [35] recently showed, the analytic screening only approximates the true screening, and does not capture all the effects completely. On the other

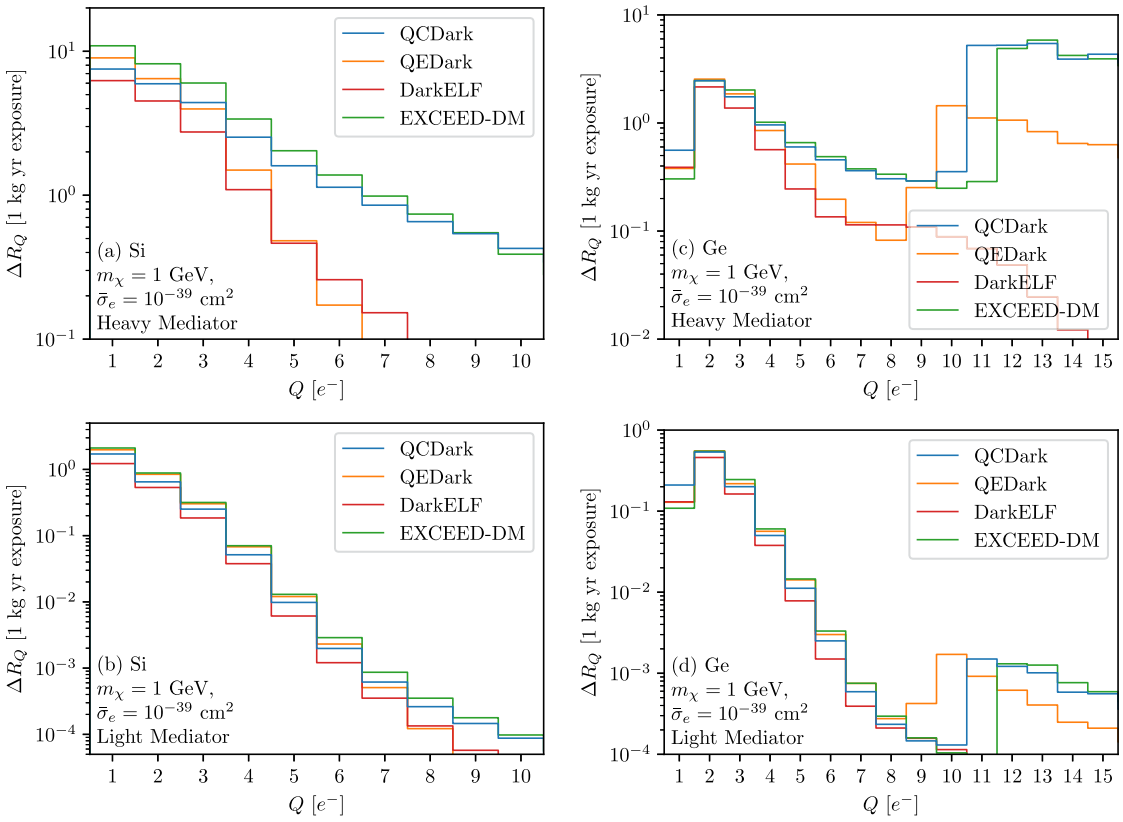

FIG. 12. Panels (a) and (b) show the comparison among DM–electron scattering rates in Si assuming a heavy and a light mediator, respectively, calculated using different codes. This work and QEDARK both implement the screening described in Sec. II B 1, while EXCEED-DM and DarkELF use their numerically calculated dielectric function. Panels (c) and (d) show the comparative plots for Ge assuming a heavy and a light mediator respectively.

hand, QCDARK allows for a much better handle on systematics by giving users control over the theory parameters, as discussed in Sec. III C.

IV. CONCLUSION

In this paper, we present DM–electron scattering rates in silicon and germanium crystals calculated using a new code, which we make public as QCDARK. We use a novel approach that naturally includes all core electrons, and treats them on the same level as valence electrons of the crystal. This implies that all-electron effects are automatically included from the beginning. Moreover, we present the theoretical uncertainties associated with the calculation, including those associated with DFT (basis set, exchange–correlation functional, and **k**-grid), along with uncertainties associated with the transition matrix elements (real space cutoff and the maximum momentum transfer modeled, $q_{\max}$).

The major sources of systematic error include the choice of basis set and exchange–correlation functional, even after we apply the scissor correction, and the choice of $q_{\max}$, though the rates converge quickly in the $E_e \in [0, 50\ \text{eV}]$ range for both Si and Ge. The rates also converge quickly as finer **k**-grids are chosen, assuming the secondary ionization model in [56] for Si. The rates also converge quickly for small values of the real-space cutoff, R_{cut}.

We find that modeling high momentum transfers by including all-electron effects is necessary for accurately modeling DM–electron scattering rates, especially at high recoil energies, in line with the findings in [34,35]. This is especially important for Ge crystals, in which the transition rates from the 3D-shell (when kinematically accessible) dominate the rates if the high momentum transfer modes are modeled accurately.

ACKNOWLEDGMENTS

We thank Daniel Baxter, Timothy Berkelbach, Yonit Hochberg, Simon Knapen, Yutaro Shoji, Greg Suczewski, Tanner Trickle, and Tien-Tien Yu for valuable discussions. C. E. D. acknowledges support from the National Science Foundation under Grant No. DMR-2237674. The Flatiron Institute is a division of the Simons Foundation. R. E. acknowledges support from DoE Grant No. DE-SC0009854, Simons Investigator in Physics Award No. 623940, and the US-Israel Binational Science Foundation Grant No. 2016153. A. S. and C. Z. were

supported in part by a Stony Brook IACS Seed Grant, from Fermilab subcontract 664693 for the DoE DMNI award for Oscura, from DoE Grant No. DE-SC0009854, and from the Simons Investigator in Physics Award No. 623940. We also thank Stony Brook Research Computing and Cyberinfrastructure, and the Institute for Advanced Computational Science at Stony Brook University for access to the high-performance SeaWulf computing system, which was made possible by a National Science Foundation Grant No. 1531492.

APPENDIX A: PROPERTIES OF CARTESIAN GAUSSIANS

In this section we discuss the properties of Cartesian Gaussians, including the calculation of the matrix elements in Eq. (9). Our Cartesian Gaussian basis sets contain primitive Gaussians as building blocks [see Eq. (4)],

$$G_{ijk}(\mathbf{r},\xi,\mathbf{A}) = (x-A_x)^i(y-A_y)^j(z-A_z)^k \times \exp\{-\xi(\mathbf{r}-\mathbf{A})^2\}, \tag{A1}$$

which we separate into three independent Gaussians,

$$G_{ijk}(\mathbf{r},\xi,\mathbf{A}) = G_i(x,\xi,A_x)G_j(y,\xi,A_y) \times G_k(z,\xi,A_z), \tag{A2}$$

where $G_i(x,\xi,A_x) = (x-A_x)^i \exp\{-\xi(x-A_x)^2\}$. This separation will be instrumental in obtaining an analytical form for calculation of atomic orbital overlaps.

It is useful to define Hermite Gaussian functions (see Ref. [61] for more details),

$$\Lambda_t(x,\xi,A_x) = \left(\frac{\partial}{\partial A_x}\right)^t \exp\{-\xi(x-A_x)^2\}. \tag{A3}$$

These Hermite Gaussians will appear below and are related to Hermite polynomials $H_t(x)$ as

$$\Lambda_t(x,\xi,A_x) = \sqrt{\xi}H_t\big(\sqrt{\xi}(x-A_x)\big)\exp\{-\xi(x-A_x)^2\}. \tag{A4}$$

We now discuss the overlap between two orbitals,

$$\Omega_{ij}(x,a,b,A_x,B_x) \equiv G_i(x,a,A_x)G_j(x,b,B_x) = x_A^i x_B^j \exp\{-ax_A^2\}\exp\{-bx_B^2\}, \tag{A5}$$

where $x_A \equiv x - A_x$. Now,

$$\exp\{-ax_A^2\}\exp\{-bx_B^2\} = \exp\{-qQ_x^2\}\exp\{-px_P^2\}, \tag{A6}$$

where

$$\begin{aligned} pP_x &= aA_x + bB_x, \\ Q_x &= A_x - B_x, \\ p &= a+b, \\ \text{and} \quad q &= \frac{ab}{a+b}. \end{aligned} \tag{A7}$$

Note that the x-dependence in Eq. (A6) only comes from $\exp\{-px_P^2\}$, and so we can define the constant $K_{AB} \equiv \{\exp -qQ_x^2\}$. Thus we can write (see Eqs. (49)–(53), (59), (60), and (70)–(75) in [61])

$$\Omega_{ij}(x,a,b,A_x,B_x) = \sum_{t=0}^{i+j} E_t^{ij}\Lambda_t(x,p,P_x), \tag{A8}$$

where the expansion coefficients have the recurrence relations,

$$\begin{aligned} E_0^{00} &= K_{AB}, \\ E_t^{i+1,j} &= \frac{1}{2p}E_{t-1}^{ij} - \frac{qQ_x}{a}E_t^{ij} + (t+1)E_{t+1}^{ij}, \\ E_t^{i+1,j} &= \frac{1}{2p}E_{t-1}^{ij} + \frac{qQ_x}{b}E_t^{ij} + (t+1)E_{t+1}^{ij}. \end{aligned} \tag{A9}$$

We can now finally calculate integrals of the form

$$\begin{aligned} &\langle G_i(x,a,A_x)|\exp\{ik_x x_C\}|G_j(x,b,B_x)\rangle \\ &= \int_{-\infty}^{\infty} dx\,\Omega_{ij}(x,a,b,A_x,B_x)\exp\{ik_x x_C\} \\ &= \sum_{t=0}^{i+j} E_t^{ij}\int_{-\infty}^{\infty} dx\,\Lambda_t(x,p,P_x)\exp\{ik_x x_C\} \\ &= \sum_{t=0}^{i+j} E_t^{ij}K_t^x, \end{aligned} \tag{A10}$$

where $K_t^x \equiv \int_{-\infty}^{\infty} dx\,\Lambda_t(x,p,P_x)\exp\{ik_x x_C\}$. Expanding the Hermite Gaussian, we get

$$\begin{aligned} K_t^x &= \left(\frac{\partial}{\partial P_x}\right)^t \int_{-\infty}^{\infty} dx \exp\{ik_x x_C - px_P^2\} \\ &= \left(\frac{\partial}{\partial P_x}\right)^t \exp\{ik_x X_{PC}\}\int_{-\infty}^{\infty} dx \exp\{ik_x x_P - px_P^2\}, \end{aligned} \tag{A11}$$

with $X_{PC} \equiv P_x - C_x$. The integral term is now independent of P_x, and so the differential only applies to $\exp\{ik_x X_{PC}\}$, giving

$$K_t^x = \sqrt{\frac{\pi}{p}}\exp\left\{ik_x X_{PC} - \frac{k_x^2}{4p}\right\}(ik_x)^t. \tag{A12}$$

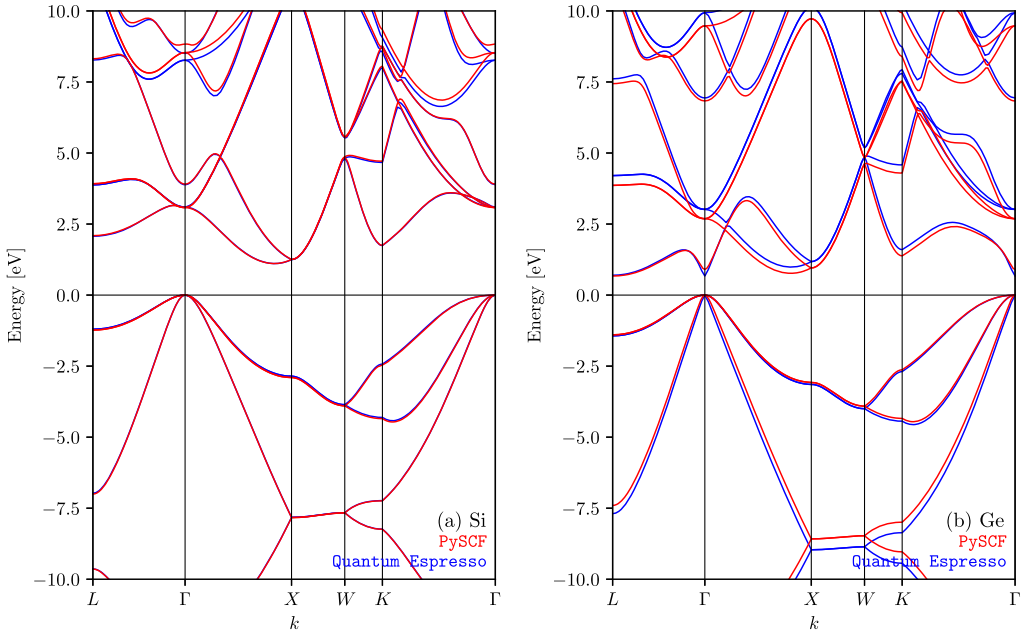

FIG. 13. Panels (a) and (b) show the band structure of Si and Ge respectively, obtained with PySCF [43–45] using a cc-pVTZ basis set and PBE functional (in red). These band structure are compared to those computed using plane wave basis sets with QUANTUM ESPRESSO [62,63] using pseudopotentials from Ref. [64] (in blue).

This gives us the analytical solution to the matrix element between two primitive Gaussians,

$$\langle G_i(x,a,A_x)\exp\{ik_x x_C\}|G_j(x,b,B_x)\rangle = \sqrt{\frac{\pi}{p}}\exp\left\{ik_x X_{PC}-\frac{k_x^2}{4p}\right\}\sum_{t=0}^{i+j}E_t^{ij}(ik_x)^t, \tag{A13}$$

which can be plugged into Eq. (9),

$$\begin{aligned} f_{[j\mathbf{k}',i\mathbf{k}]}(\mathbf{q}) &= \sum_{\mathbf{R}}e^{i\mathbf{k}'\cdot\mathbf{R}}\sum_{\alpha}\sum_{\beta}C^{\dagger}_{\beta i}(\mathbf{k})C_{j_\alpha}(\mathbf{k}')\sum_{\mu\in\beta}\sum_{\nu\in\alpha}N_\mu c_\mu N_\nu c_\nu \\ &\times\langle G^x_\mu(x,\xi_\mu,A_{\mu x})|\exp\{iq_x x\}|G^x_\nu(x,\xi_\nu,A_{\nu x})\rangle \\ &\times\langle G^y_\mu(y,\xi_\mu,A_{\mu y})|\exp\{iq_y y\}|G^y_\nu(y,\xi_\nu,A_{\nu y})\rangle \\ &\times\langle G^z_\mu(z,\xi_\mu,A_{\mu z})|\exp\{iq_z z\}|G^z_\nu(z,\xi_\nu,A_{\nu z})\rangle. \end{aligned} \tag{A14}$$

APPENDIX B: BAND STRUCTURE CALCULATION WITH PySCF

Gaussian basis sets have been widely used in DFT calculations involving crystalline solids [43–45]. In this section, we use Cartesian Gaussians in PySCF to calculate the band structure of Si and Ge. For both calculations, we use a `cc-pVTZ` basis set.

We compare the obtained band structure to that obtained using plane wave basis set, calculated using QUANTUM ESPRESSO [62,63] using pseudopotentials from Ref. [64], in Fig. 13. The band structures have been scissor corrected to experimental band gaps listed in table II. We use an $8\times8\times8$ **k**-grid and a PBE exchange-correlation functional for these calculations.

The band structure for Si in panel (a) of Fig. 13 shows that Cartesian Gaussians are consistent with established PW method in the valence and conduction band regimes, while incorporating *ab initio* all-electron effects. Indeed, the discrepancy in panel (b) around energies ∼5 eV are a result of plane wave methods underestimating the band gap for Ge and the conduction bands being raised to higher energies due to the scissor correction.

APPENDIX C: DERIVATION OF SCATTERING RATE FORMULA

In this section, we will briefly review derivation of the scattering rate formulas, mostly following Appendix A of [5]. If a DM particle scatters with an electron in a stationary bound state, such as in a crystal, it can excite the electron from some initial energy $E_{e,1}$ to some final energy $E_{e,2}$ by transferring four-momentum $(E_e,\vec{q})$. We describe the derivation in the context of field theory, treating the electron as being bound in a static background potential—in other words, treating it nonrelativistically during the interaction. This is a valid approximation because the momentum transfers are $q\sim\mathcal{O}(\text{several keV})\ll m_e$.

1. General formula for DM induced transitions

The cross section for free $2 \to 2$ scattering is given by

$$\sigma v_{\text{free}} = \frac{1}{4E'_\chi E'_e}\int \frac{d^3q}{(2\pi)^3}\frac{d^3k'}{(2\pi)^3}\frac{1}{4E_\chi E_e}|\overline{\mathcal{M}_{\text{free}}(\mathbf{q})}|^2 \times (2\pi)^4\delta(E_i - E_f)\delta^3(\mathbf{k}-\mathbf{q}-\mathbf{k}'), \tag{C1}$$

where $\mathcal{M}_{\text{free}}$ is the field-theory matrix element and $|\overline{\mathcal{M}_{\text{free}}}|^2$ is its absolute squared averaged over initial spins and summed over final spins.

For bound electron initial and final states, say ψ_1 and ψ_2, respectively, the cross section is modified as

$$V(2\pi)^3\delta^3(\mathbf{k}-\mathbf{q}-\mathbf{k}')|\mathcal{M}_{\text{free}}|^2 \to V^2|\mathcal{M}_{\text{free}}|^2|f_{1\to 2}(\mathbf{q})|^2, \tag{C2}$$

where

$$f_{1\to 2}(\mathbf{q}) \equiv \int d^3x\psi_2^*(\mathbf{x})e^{i\mathbf{q}\cdot\mathbf{x}}\psi_1(\mathbf{x}). \tag{C3}$$

Moreover, because there is only one electron final state being considered, we can make the replacement $V\int\frac{d^3k'}{(2\pi)^3} \to 1$.

Combining these observations, we obtain

$$\sigma v_{1\to 2} = \frac{1}{4E'_\chi E'_e}\int\frac{d^3q}{(2\pi)^3}\frac{1}{4E_\chi E_e}2\pi\delta(E_i - E_f) \times |\overline{\mathcal{M}_{\text{free}}(\mathbf{q})}|^2|f_{1\to 2}(\mathbf{q})|^2. \tag{C4}$$

For nonrelativistic scattering,

$$E_i = m_\chi + m_e + \frac{1}{2}m_\chi v^2 + E_{e,1}, \quad \text{and}$$
$$E_f = m_\chi + m_e + \frac{|m_\chi\mathbf{v}-\mathbf{q}^2}{2m_\chi} + E_{e,2}. \tag{C5}$$

Moreover, we can parametrize

$$|\overline{\mathcal{M}_{\text{free}}(\mathbf{q})}|^2 = |\overline{\mathcal{M}_{\text{free}}(\alpha m_e)}|^2 \times |F_\chi(\mathbf{q})|^2,$$
$$\text{and} \quad \bar{\sigma}_e \equiv \frac{\mu_{\chi e}^2|\overline{\mathcal{M}_{\text{free}}(\alpha m_e)}|^2}{16\pi m_\chi^2 m_e^2}, \tag{C6}$$

so the cross section simplifies to

$$\sigma v_{1\to 2} = \frac{\bar{\sigma}_e}{\mu_{\chi e}^2}\int\frac{d^3q}{4\pi}\delta\left(E_e + \frac{q^2}{2m_\chi} - \mathbf{q}\cdot\mathbf{v}\right) \times |F_\chi(\mathbf{q})|^2|f_{1\to 2}(\mathbf{q})|^2. \tag{C7}$$

2. Average rate in a DM halo

The rate of the specific transitions induced by DM hitting a target electron is then

$$R_{1\to 2} = n_\chi\int d^3v g_\chi(\mathbf{v})\sigma v_{1\to 2}, \tag{C8}$$

where n_χ and $g_\chi(\mathbf{v})$ are the DM number density and velocity distribution, respectively. In this work, we use the parameters recommended by [60].

Note that the velocity distribution of DM in the standard halo model implies that the speed of the DM wind we observe must follow $v_\chi < v_{\text{Escape}} + v_\odot + v_\oplus = v_{\text{Escape}} + v_{\text{Earth}}$, where v_{Escape} is the escape velocity at the Sun's location in the galactic gravitational potential well, $v_\odot$ is the sun's galactocentric speed, $v_\oplus$ is the Earth's heliocentric speed, and v_{Earth} is the Earth's galactocentric speed.

In this paper, we assume both DM velocity distribution and electron wave functions to be spherically symmetric, which is not true in general. We then use the integral over DM velocity to eliminate the δ–function in Eq. (C7), obtaining

$$R_{1\to 2} = \frac{n_\chi\bar{\sigma}_e}{\mu_{\chi e}^2}\int\frac{d^3q}{4\pi}\int\frac{v^2dvd\phi_v}{qv}g_\chi(\mathbf{v}) \times \Theta(v - v_{\min}(q,E_e))|F_\chi(\mathbf{q})|^2|f_{1\to 2}(\mathbf{q})|^2. \tag{C9}$$

Here $v_{\min}$ is the minimum velocity of the DM particle required for an energy-momentum transfer of $(E_e,\mathbf{q})$ to be feasible,

$$v_{\min}(q,E_e) = \frac{E_e}{q} + \frac{q}{2m_\chi}. \tag{C10}$$

We define

$$\eta(v_{\min}(q,E_e)) \equiv \int\frac{d^3v}{v}g_\chi(\mathbf{v})\Theta(v - v_{\min}(q,E_e)), \tag{C11}$$

and obtain

$$R_{1\to 2} = \frac{n_\chi\bar{\sigma}_e}{8\pi\mu_{\chi e}^2}\int d^3q\frac{1}{q}\eta(v_{\min}(q,E_e))|F_\chi(\mathbf{q})|^2|f_{1\to 2}(\mathbf{q})|^2. \tag{C12}$$

3. Excitation rates in crystals

So far, we have not discussed the initial and final states of the electron being acted upon. In a crystal system, an electron may transition from an occupied orbital (core or valence) to an unoccupied state (conduction or free). Since we treat both core and valence shells equivalently, we simply call these occupied orbitals. Using the setup

discussed in Sec. II A, we describe now the transition form factors, $f_{1\to 2}(\mathbf{q})$.

The electron is excited from an occupied state $|\psi_{i\mathbf{k}}\rangle$ to a conduction state $|\psi_{i'\mathbf{k}'}\rangle$, and so $f_{1\to 2}(\mathbf{q}) \to f_{i\mathbf{k}\to i'\mathbf{k}'}$, with

$$\begin{aligned} f_{i\mathbf{k}\to i'\mathbf{k}'} &= \langle \psi_{i'\mathbf{k}'} \exp\{i\mathbf{q}\cdot\mathbf{r}\}|\psi_{i\mathbf{k}}\rangle \\ &= \frac{1}{N_{\text{cell}}} C^{\dagger}_{i'\beta}(\mathbf{k}')\langle \phi_{\beta\mathbf{k}'}|e^{i\mathbf{q}\cdot\mathbf{r}}|\phi_{\alpha\mathbf{k}}\rangle C_{\alpha i}(\mathbf{k}), \end{aligned} \tag{C13}$$

where we have maintained the PySCF normalization, and Einstein summation over α and β indices is implied. We shall invoke the orthogonality relation

$$\frac{V_{\text{cell}}}{(2\pi)^3}\sum_{\mathbf{R}} e^{i\mathbf{q}\cdot\mathbf{R}} = \sum_{\mathbf{G}} \delta^3(\mathbf{q}-\mathbf{G}). \tag{C14}$$

The matrix element is then given by

$$\begin{aligned} f_{i\mathbf{k}\to i'\mathbf{k}'}(\mathbf{q}) &= \frac{1}{N_{\text{cell}}} C^{\dagger}_{i'\beta}(\mathbf{k}')\sum_{\mathbf{R}}\sum_{\mathbf{R}'} e^{-i\mathbf{k}'\cdot\mathbf{R}'}e^{i\mathbf{k}\cdot\mathbf{R}}\int d^3r \tilde{G}^*_{\beta}(\mathbf{r}-\mathbf{R}')e^{i\mathbf{q}\cdot\mathbf{r}}\tilde{G}_{\alpha}(\mathbf{r}-\mathbf{R})C_{\alpha i}(\mathbf{k}) \\ &= \frac{1}{N_{\text{cell}}} C^{\dagger}_{i'\beta}(\mathbf{k}')\sum_{\mathbf{R}} e^{i(\mathbf{k}+\mathbf{q}-\mathbf{k}')\cdot\mathbf{R}}\sum_{\mathbf{R}'} e^{-i\mathbf{k}'\cdot\mathbf{R}'}\int d^3r \tilde{G}^*_{\beta}(\mathbf{r}-\mathbf{R}')e^{i\mathbf{q}\cdot\mathbf{r}}\tilde{G}_{\alpha}(\mathbf{r})C_{\alpha i}(\mathbf{k}) \\ &= \frac{(2\pi)^3}{V}\sum_{\mathbf{G}}\delta^3(\mathbf{k}+\mathbf{q}-\mathbf{k}'-\mathbf{G})\sum_{\mathbf{R}} e^{-i\mathbf{k}'\cdot\mathbf{R}}\int d^3r C^{\dagger}_{i'\beta}(\mathbf{k}')\tilde{G}^*_{\beta}(\mathbf{r}-\mathbf{R})e^{i\mathbf{q}\cdot\mathbf{r}}\tilde{G}_{\alpha}(\mathbf{r})C_{\alpha i}(\mathbf{k}) \\ &= \frac{(2\pi)^3}{V} f_{[i'\mathbf{k}',i\mathbf{k}]}(\mathbf{q})\sum_{\mathbf{G}}\delta^3(\mathbf{k}+\mathbf{q}-\mathbf{k}'-\mathbf{G}), \end{aligned} \tag{C15}$$

where we have used the definition of $f_{[i'\mathbf{k}',i\mathbf{k}]}(\mathbf{q})$ from Eq. (9). We can now plug this into Eq. (C12) to obtain

$$\begin{aligned} R_{i\mathbf{k}\to i'\mathbf{k}'} = \frac{\pi^2 n_\chi \bar{\sigma}_e}{V\mu^2_{\chi e}}\sum_{\mathbf{G}}\frac{1}{q}\eta(v_{\min}(q, E_{i'\mathbf{k}'}-E_{i\mathbf{k}})) \\ \times |F_\chi(\mathbf{q})|^2 |f_{i\mathbf{k}\to i'\mathbf{k}'}(\mathbf{q})|^2|_{\mathbf{q}=\mathbf{k}'+\mathbf{G}-\mathbf{k}}. \end{aligned} \tag{C16}$$

To calculate the total event rate, we must sum over occupied orbitals i and unoccupied orbitals i', and integrate over both $\mathbf{k}$ and $\mathbf{k}'$. Moreover, we must also consider the spin of the electrons in the occupied bands, giving

$$R_{\text{crystal}} = 2\sum_{i}^{\text{occ}}\sum_{i'}^{\text{unocc}}\int_{\text{BZ}}\frac{Vd^3k}{(2\pi)^3}\int_{\text{BZ}}\frac{Vd^3k'}{(2\pi)^3}R_{i\mathbf{k}\to i'\mathbf{k}'}. \tag{C17}$$

Expanding this, and inserting the relevant δ-distributions in $\mathbf{q}$ and E_e,

$$\begin{aligned} R_{\text{crystal}} = \frac{2\pi^2 n_\chi \bar{\sigma}_e}{\mu_{\chi e}} V\int_{-\infty}^{\infty} d\ln E_e E_e \\ \times \int d^3q \frac{1}{q}\eta(v_{\min}(q,E_e))|F_\chi(\mathbf{q})|^2 \\ \times \sum_{ii'}\int_{\text{BZ}}\frac{d^3k d^3k'}{(2\pi)^6}\delta\big(E_e-(E_{i'\mathbf{k}'}-E_{i\mathbf{k}})\big) \\ \times f_{[i'\mathbf{k}',i\mathbf{k}]}(\mathbf{q})\sum_{\mathbf{G}}\delta^3(\mathbf{k}+\mathbf{q}-\mathbf{k}'-\mathbf{G}). \end{aligned} \tag{C18}$$

To simplify this form, we take two steps. First, we define $\mathbf{U}\equiv\mathbf{k}'+\mathbf{G}-\mathbf{k}$, and so $\delta^3(\mathbf{k}+\mathbf{q}-\mathbf{k}'-\mathbf{G}) = \delta^3(\mathbf{q}-\mathbf{U})$, which can be further expanded as

$$\delta^3(\mathbf{q}-\mathbf{U}) = \frac{1}{q^2\sin\theta_q}\delta(q-U)\delta(\theta_q-\theta_U)\delta(\phi_q-\phi_U). \tag{C19}$$

Here θ_U and ϕ_U have the usual definitions as the inclination and azimuthal angles, respectively. From here, we can integrate over Ω_q. Second, we differentiate the rate equation with respect to $\ln E_e$, and obtain Eqs. (8) and (10),

$$\frac{\mathrm{d}R_{\text{crystal}}}{\mathrm{d}\ln E_e} = n_\chi N_{\text{cell}}\bar{\sigma}_e\alpha\frac{m_e^2}{\mu^2_{\chi e}}\int d\ln q\frac{E_e}{q}\eta(v_{\min}(q,E_e))|F_\chi(q)|^2|f_{\text{crystal}}(q,E_e)|^2, \quad \text{where} \tag{C20}$$

$$\begin{aligned} |f_{\text{crystal}}(q,E_e)|^2 \equiv \frac{2\pi^2}{E_e}\frac{1}{\alpha m_e^2 V_{\text{cell}}}\sum_{ii'}\int_{\text{BZ}}\frac{V_{\text{cell}}d^3k}{(2\pi)^3}\frac{V_{\text{cell}}d^3k'}{(2\pi)^3}E_e\delta(E_e-(E_{j\mathbf{k}'}-E_{i\mathbf{k}})) \\ \times\sum_{\mathbf{G}'} q\delta(q-|\mathbf{k}'+\mathbf{G}'-\mathbf{k}|)|f_{[i'\mathbf{k}',i\mathbf{k}]}(\mathbf{q})|^2\Big|_{\theta_q=\theta_U,\phi_q=\phi_U}. \end{aligned} \tag{C21}$$

APPENDIX D: CROSS CORRELATION OF SYSTEMATIC UNCERTAINTIES AND $q_{\max}$ CONVERGENCE

In Sec. III C, we discussed multiple systematic uncertainties and convergence parameters in our calculation. In this section, we show the dependence of systematic uncertainties on the chosen $q_{\max}$ for a calculation. Note that the **k**-grid and R_{cut} are convergence parameters, where both a denser **k**-grid and a higher R_{cut} are objectively better. In fact, the **k**-grid only affects the energies sampled for a calculation and are hence orthogonal to $q_{\max}$, while, as discussed above, R_{cut} effects low-momentum transfers and has no effect at high q.

Therefore in this section we focus only on basis sets, which is both a systematic uncertainty and a convergence parameter, and exchange–correlation functionals. We find that the effect are converged for low charge ionization ($Q \lesssim 4e^-$) for the basis sets, while the effect of exchange-correlation functional remains consistent through $1e^- \leq Q \leq 10e^-$ for all $q_{\max} \gtrsim 8\alpha m_e$.

1. Basis sets

The choice of basis set presents a major source of systematic uncertainty in our calculations, which we previously discussed in Sec. III C 1. However, to reduce computational complexity, we limit the discussion to a maximum momentum transfer of $q_{\max} = 10\alpha m_e$ compared to the model $q_{\max} = 25\alpha m_e$ for Si or $q_{\max} = 20\alpha m_e$ for Ge (see Sec. III C 5 for more details). Hence, it is important to discuss the convergence of the rates with respect to $q_{\max}$ when choosing an alternate basis set.

To accomplish this, we calculate the fractional change in the expected rates for a DM particle with mass $m_\chi = 1$ GeV, where the DM–electron interaction is mediated by a heavy mediator. Note that this model is more sensitive to changes in the crystal form factor, $|f_{\text{crystal}}(q, E_e)|^2$, at high momentum transfers q than models with lower m_χ, or where the interactions are mediated by a low mass mediator [with mass $\lesssim \mathcal{O}(1\alpha m_e)$].

We define the fractional change in rates upon going from TZP to another basis set BS as

$$f^{(\text{BS}|\text{TZP})}_{Q,q_{\max}} \equiv \frac{\Delta R_{Q|q_{\max},\text{BS}} - \Delta R_{Q|q_{\max},\text{TZP}}}{\Delta R_{Q|q_{\max},\text{TZP}}}, \tag{D1}$$

where $\Delta R_{Q|q_{\max},\text{BS}}$ are the DM–electron scattering rates for the aforementioned DM model given a maximum momentum transfer $q_{\max}$ and a basis set BS. We plot the fractional change for a Si target in Fig. 14, where we limit ourselves to the more reliable cc-pvtz and cc-pvqz basis sets (see discussion in Sec. III C 1). Similarly, we plot the results for a Ge target in Fig. 16 where we use the more reliable Sapporo-DZP-2012 and cc-pvtz basis sets. The title of each subplot corresponds to the chosen basis set BS.

For a given Q, we would like to see that $f^{(\text{BS}|\text{TZP})}_{Q,q_{\max}}$ converges to a constant value with increasing $q_{\max}$. This means that BS (the chosen basis set) and TZP (the reference basis set) converge at the same rate with $q_{\max}$, and thus we can faithfully compare the results of the two basis sets at the given $q_{\max}$. In Fig. 14 (and in Figs. 15–17 which we discuss below), this is demonstrated by the color becoming constant with increasing $q_{\max}$.

In Si, for large charge ionization $Q \geq 6e^-$, we find that the uncertainties are not converged for $q_{\max} = 10\alpha m_e$, since the uncertainties stem from the basis sets providing different descriptions of the conduction bands at higher energies. However, at low charge ionizations $Q \lesssim 5e^-$, which in any case dominate the DM scattering rates (see Sec. III B and Fig. 2), we find that the systematic uncertainties are converged for each basis set BS, and are hence reliable to higher $q_{\max}$.

Similarly, in Ge, we find that the basis sets are converged for both low and high charge ionizations, $Q \lesssim 4e^-$ and $Q \geq 11e^-$, again due to differences in modeling of the higher energy conduction bands. However, we again find that the systematic uncertainties are converged in the region where the rates dominate.

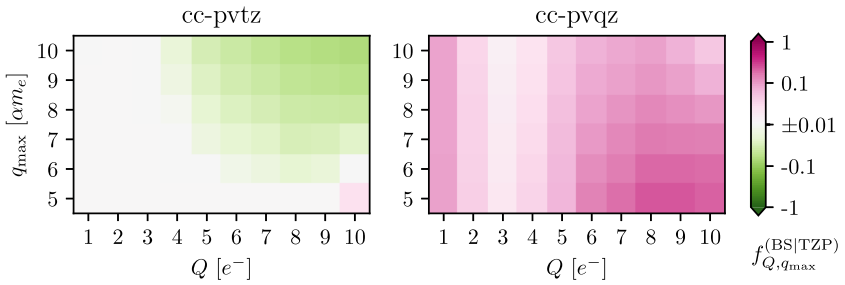

FIG. 14. The plot shows the convergence of the systematic uncertainty sourced from the choice of basis set as a function of charge ionization Q and maximum momentum transfer modeled $q_{\max}$ by showing the fractional change in rates observed for a DM–electron scattering ($m_\chi = 1$ GeV, $F_\chi = 1$ in Si target) upon going from a TZP functional to a new functional BS [see Eq. (D1)]. Note that for all BS tested, we observe that the fractional change becomes constant as a function of $q_{\max}$ by $q_{\max} \leq 10\alpha m_e$ at low charge ionization $Q \lesssim 5e^-$.

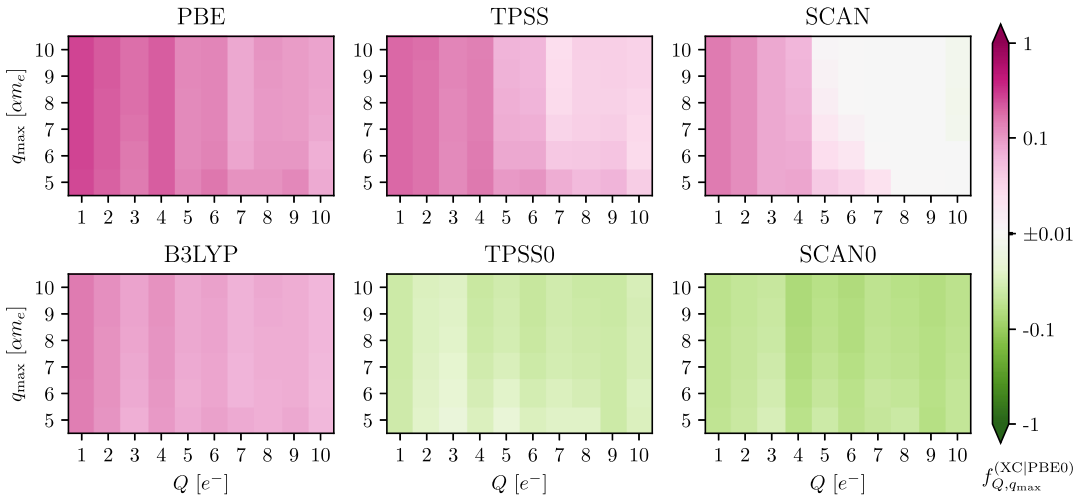

FIG. 15. The plot shows the convergence of the systematic uncertainty sourced from the choice of exchange–correlation functional as a function of charge ionization Q and maximum momentum transfer modeled $q_{\max}$ by showing the fractional change in rates observed for a DM–electron scattering ($m_\chi = 1$ GeV, $F_\chi = 1$ in Si target) upon going from PBE0 functional to a new functional XC [see Eq. (D2)]. Note that for all XC tested, we observe that the fractional change becomes constant as a function of $q_{\max}$ by $q_{\max} \leq 8\alpha m_e$.

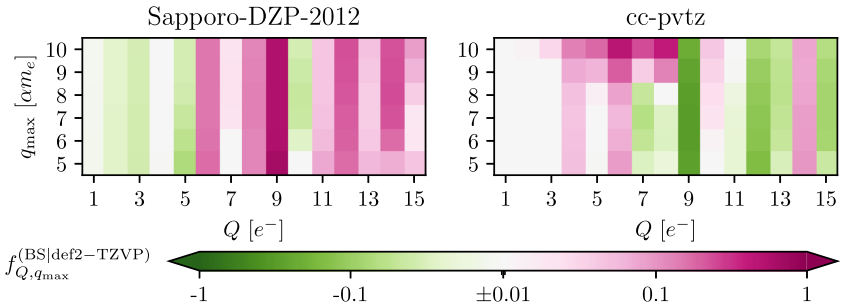

FIG. 16. The plot shows the same as Fig. 14, except for Ge instead of Si, calculated using a PBE functional. Note that the basis set is converged for large $Q \geq 11e^-$, and also for low $Q \lesssim 4e^-$.

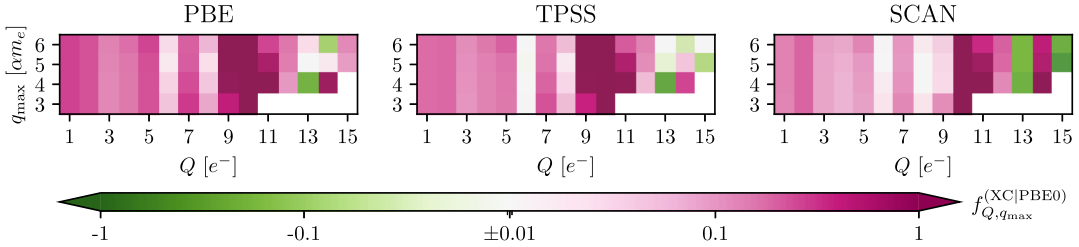

FIG. 17. The plot shows the equivalent of Fig. 15, except for Ge instead of Si. Note that we test to a much lower $q_{\max} = 6\alpha m_e$, as in Fig. 7, however we find that the we attain convergence for low $Q \lesssim 8e^-$.

2. Exchange–correlation functionals

In Sec. III C 4, we discuss the effect of exchange–correlation functional on systematic uncertainties. In a similar vein to the discussion above, we must ensure that the calculations are converged with respect to the other parameters (in particular, $q_{\max}$) while we vary the exchange–correlation functional.

We follow the procedure outlined above for the basis set uncertainty by calculating the fractional change upon going from a PBE0 functional to a new functional XC as

$$f^{(\mathrm{XC}|\mathrm{PBE0})}_{Q,q_{\max}} \equiv \frac{\Delta R_{Q|q_{\max},\mathrm{XC}} - \Delta R_{Q|q_{\max},\mathrm{PBE0}}}{\Delta R_{Q|q_{\max},\mathrm{PBE0}}}, \tag{D2}$$

and plot the same for a Si target in Fig. 15, and Ge target in Fig. 17. The title of each subplot corresponds to the functional XC chosen. Note that the fractional change remains consistent while we increase our $q_{\max}$, especially at $q_{\max} \gtrsim 8\alpha m_e$ in Si. In Ge, we find convergence for $Q \lesssim 8e^-$ even to $q_{\max} = 6\alpha m_e$.

[1] R. Essig, J. Mardon, and T. Volansky, Direct detection of sub-Gev dark matter, Phys. Rev. D **85**, 076007 (2012).

[2] 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**, 021301 (2012).

[3] R. Essig, T. Volansky, and T.-T. Yu, New constraints and prospects for sub-Gev dark matter scattering off electrons in Xenon, Phys. Rev. D **96**, 043017 (2017).

[4] J. Angle *et al.* (XENON10 Collaboration), A search for light dark matter in Xenon10 data, Phys. Rev. Lett. **107**, 051301 (2011); **110**, 249901(E) (2013).

[5] 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, J. High Energy Phys. 05 (2016) 046.

[6] S. K. Lee, M. Lisanti, S. Mishra-Sharma, and B. R. Safdi, Modulation effects in dark matter-electron scattering experiments, Phys. Rev. D **92**, 083517 (2015).

[7] M. Crisler, R. Essig, J. Estrada, G. Fernandez, J. Tiffenberg, M. Sofo haro, T. Volansky, and T.-T. Yu (SENSEI Collaboration), Sensei: First direct-detection constraints on sub-Gev dark matter from a surface run, Phys. Rev. Lett. **121**, 061803 (2018).

[8] O. Abramoff *et al.* (SENSEI Collaboration), SENSEI: Direct-detection constraints on sub-GeV dark matter from a shallow underground run using a prototype Skipper-CCD, Phys. Rev. Lett. **122**, 161801 (2019).

[9] L. Barak *et al.* (SENSEI Collaboration), SENSEI: Direct-detection results on sub-GeV dark matter from a new Skipper-CCD, Phys. Rev. Lett. **125**, 171802 (2020).

[10] A. Aguilar-Arevalo *et al.* (DAMIC Collaboration), Results on low-mass weakly interacting massive particles from a 11 kg − day target exposure of DAMIC at SNOLAB, Phys. Rev. Lett. **125**, 241803 (2020).

[11] R. Agnese *et al.* (SuperCDMS Collaboration), First dark matter constraints from a supercdms single-charge sensitive detector, Phys. Rev. Lett. **121**, 051301 (2018).

[12] A. Aguilar-Arevalo *et al.* (DAMIC Collaboration), Constraints on light dark matter particles interacting with electrons from damic at Snolab, Phys. Rev. Lett. **123**, 181802 (2019).

[13] Q. Arnaud *et al.* (EDELWEISS Collaboration), First germanium-based constraints on sub-MeV dark matter with the EDELWEISS experiment, Phys. Rev. Lett. **125**, 141301 (2020).

[14] P. Agnes *et al.* (DarkSide Collaboration), Constraints on sub-GeV dark-matter–electron scattering from the DarkSide-50 experiment, Phys. Rev. Lett. **121**, 111303 (2018).

[15] E. Aprile *et al.* (XENON Collaboration), Low-mass dark matter search using ionization signals in Xenon100, Phys. Rev. D **94**, 092001 (2016); **95**, 059901(E) (2017).

[16] E. Aprile *et al.* (XENON Collaboration), Light dark matter search with ionization signals in Xenon1T, Phys. Rev. Lett. **123**, 251801 (2019).

[17] E. Aprile *et al.* (XENON Collaboration), Emission of single and few electrons in XENON1T and limits on light dark matter, Phys. Rev. D **106**, 022001 (2022).

[18] C. Cheng *et al.* (PandaX-II Collaboration), Search for light dark matter-electron scatterings in the PandaX-II experiment, Phys. Rev. Lett. **126**, 211803 (2021).

[19] C. Blanco, J. I. Collar, Y. Kahn, and B. Lillard, Dark matter-electron scattering from aromatic organic targets, Phys. Rev. D **101**, 056001 (2020).

[20] I. Arnquist *et al.* (DAMIC-M Collaboration), The DAMIC-M experiment: Status and first results, SciPost Phys. Proc. **12**, 014 (2023).

[21] D. W. Amaral *et al.* (SuperCDMS Collaboration), Constraints on low-mass, relic dark matter candidates from a surface-operated SuperCDMS single-charge sensitive detector, Phys. Rev. D **102**, 091101 (2020).

[22] E. Aprile *et al.* (XENON Collaboration), Excess electronic recoil events in XENON1T, Phys. Rev. D **102**, 072004 (2020).

[23] J. Tiffenberg, M. Sofo-Haro, A. Drlica-Wagner, R. Essig, Y. Guardincerri, S. Holland, T. Volansky, and T.-T. Yu (SENSEI Collaboration), Single-electron and single-photon sensitivity with a silicon skipper CCD, Phys. Rev. Lett. **119**, 131802 (2017).

[24] M. Settimo (DAMIC Collaboration), The damic experiment at Snolab, in *Proceedings, 53Rd Rencontres De Moriond on Cosmology: La Thuile, Italy, 2018* (2018); arXiv:1805.10001.

[25] N. Castelló-Mor, Damic-m experiment: Thick, silicon CCDS to search for light dark matter, Nucl. Instrum. Methods Phys. Res., Sect. A **958**, 162933 (2020).

[26] A. Aguilar-Arevalo *et al.*, The oscura experiment, arXiv:2202.10518.

[27] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, England, 2020).

[28] R. M. Martin, L. Reining, and D. M. Ceperley, *Interacting Electrons* (Cambridge University Press, Cambridge, England, 2016).

[29] P. E. Blöchl, Projector augmented-wave method, Phys. Rev. B **50**, 17953 (1994).

[30] W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, Phys. Rev. **140**, A1133 (1965).

[31] S. Dick and M. Fernandez-Serra, Highly accurate and constrained density functional obtained with differentiable programming, Phys. Rev. B **104**, L161109 (2021).

[32] S. Knapen, J. Kozaczuk, and T. Lin, Dark matter-electron scattering in dielectrics, Phys. Rev. D **104**, 015031 (2021).

[33] S. Knapen, J. Kozaczuk, and T. Lin, PYTHON package for dark matter scattering in dielectric targets, Phys. Rev. D **105**, 015014 (2022).

[34] S. M. Griffin, K. Inzani, T. Trickle, Z. Zhang, and K. M. Zurek, Extended calculation of dark matter-electron scattering in crystal targets, Phys. Rev. D **104**, 095015 (2021).

[35] T. Trickle, Extended calculation of electronic excitations for direct detection of dark matter, Phys. Rev. D **107**, 035035 (2023).

[36] Y. Hochberg, M. Pyle, Y. Zhao, and K. M. Zurek, Detecting superlight dark matter with Fermi-degenerate materials, J. High Energy Phys. 08 (2016) 057.

[37] Y. Hochberg, I. Charaev, S.-W. Nam, V. Verma, M. Colangelo, and K. K. Berggren, Detecting sub-GeV dark matter with superconducting nanowires, Phys. Rev. Lett. **123**, 151802 (2019).

[38] Y. Hochberg, B. V. Lehmann, I. Charaev, J. Chiles, M. Colangelo, S. W. Nam, and K. K. Berggren, New constraints on dark matter from superconducting nanowires, Phys. Rev. D **106**, 112005 (2022).

[39] C. Blanco, R. Essig, M. Fernandez-Serra, H. Ramani, and O. Slone, Dark matter direct detection with quantum dots, Phys. Rev. D **107**, 095035 (2023).

[40] Y. Kahn and T. Lin, Searches for light dark matter using condensed matter systems, Rep. Prog. Phys. **85**, 066901 (2022).

[41] Z.-L. Liang, L. Zhang, P. Zhang, and F. Zheng, The wave function reconstruction effects in calculation of DM-induced electronic transition in semiconductor targets, J. High Energy Phys. 01 (2019) 149.

[42] V. Heine, *The Pseudopotential Concept* (Academic Press, New York, 1970), pp. 1–36.

[43] Q. Sun *et al.*, Recent developments in the PySCF program package, J. Chem. Phys. **153**, 024109 (2020).

[44] Q. Sun, Libcint: An efficient general integral library for Gaussian basis functions, J. Comput. Chem. **36**, 1664 (2015).

[45] Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters, and G. K.-L. Chan, PySCF: The PYTHON-based simulations of chemistry framework, WIREs Comput. Mol. Sci. **8**, e1340 (2018).

[46] L. Hamaide and C. McCabe, Fueling the search for light dark matter-electron scattering with spherical proportional counters, Phys. Rev. D **107**, 063002 (2023).

[47] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal, The siesta method for *ab initio* order-n materials simulation, J. Phys. Condens. Matter **14**, 2745 (2002).

[48] P. Giannozzi *et al.*, QUANTUM ESPRESSO: A modular and open-source software project for quantum simulations of materials, J. Phys. Condens. Matter **21**, 395502 (2009).

[49] B. P. Pritchard, D. Altarawy, B. Didier, T. D. Gibson, and T. L. Windus, New basis set exchange: An open, up-to-date resource for the molecular sciences community, J. Chem. Inf. Model. **59**, 4814 (2019).

[50] J. Dunning and H. Thom, Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen, J. Chem. Phys. **90**, 1007 (1989).

[51] S. Lehtola, C. Steigemann, M. J. Oliveira, and M. A. Marques, Recent developments in LIBXC—A comprehensive library of functionals for density functional theory, SoftwareX **7**, 1 (2018).

[52] G. B. Gelmini, V. Takhistov, and E. Vitagliano, Scalar direct detection: In-medium effects, Phys. Lett. B **809**, 135779 (2020).

[53] Y. Hochberg, Y. Kahn, N. Kurinsky, B. V. Lehmann, T. C. Yu, and K. K. Berggren, Determining dark matter-electron scattering rates from the dielectric function, Phys. Rev. Lett. **127**, 151802 (2021).

[54] G. Cappellini, R. Del Sole, L. Reining, and F. Bechstedt, Model dielectric function for semiconductors, Phys. Rev. B **47**, 9892 (1993).

[55] J. P. Walter and M. L. Cohen, Wave-vector-dependent dielectric function for Si, Ge, GaAs, and ZnSe, Phys. Rev. B **2**, 1821 (1970).

[56] K. Ramanathan and N. Kurinsky, Ionization yield in silicon for eV-scale electron-recoil processes, Phys. Rev. D **102**, 063026 (2020).

[57] S. Knapen, J. Kozaczuk, and T. Lin, Dark matter-electron scattering in dielectrics, Phys. Rev. D **104**, 015031 (2021).

[58] J. A. Bearden and A. F. Burr, Reevaluation of x-ray atomic energy levels, Rev. Mod. Phys. **39**, 125 (1967).

[59] C. A. Klein, Bandgap dependence and related features of radiation ionization energies in semiconductors, J. Appl. Phys. **39**, 2029 (1968).

[60] D. Baxter *et al.*, Recommended conventions for reporting results from direct dark matter searches, Eur. Phys. J. C **81**, 907 (2021).

[61] T. Helgaker and P. R. Taylor, Gaussian basis sets and molecular integrals, in *Modern Electronic Structure Theory Part II* (World Scientific Publishing Company, Singapore, 1995), pp. 725–856, 10.1142/9789812832115_0001.

[62] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo *et al.*, QUANTUM ESPRESSO: A modular and open-source software project for quantum simulations of materials, J. Phys. Condens. Matter **21**, 395502 (2009).

[63] P. Giannozzi *et al.*, Advanced capabilities for materials modelling with quantum espresso, J. Phys. Condens. Matter **29**, 465901 (2017).

[64] D. R. Hamann, Optimized norm-conserving vanderbilt pseudopotentials, Phys. Rev. B **88**, 085117 (2013).