

**Quantum theory of electronic excitation and sputtering by transmission electron microscopy**

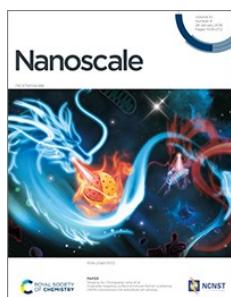
Journal:	<i>Nanoscale</i>
Manuscript ID	NR-ART-02-2022-001018.R1
Article Type:	Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Yoshimura, Anthony; Lawrence Livermore National Laboratory, Lamparski, Michael; Rensselaer Polytechnic Institute School of Science Giedt, Joel; Rensselaer Polytechnic Institute, Physics, Applied Physics, and Astronomy Lingerfelt, David; Oak Ridge National Laboratory, Center for Nanophase Materials Sciences Jakowski, Jacek; Oak Ridge National Laboratory, Center for Nanophase Materials Sciences Ganesh, Panchapakesan; Oak Ridge National Laboratory, Center for Nanophase Materials Sciences Yu, Tao; University of North Dakota, Chemistry Sumpter, Bobby; Oak Ridge National Laboratory, Computer Science and Mathematics Division and Center for Nanophase Materials Sciences Meunier, Vincent; Rensselaer Polytechnic Institute, Department of Physics, Applied Physics, and Astronomy
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
main.tar.gz	

SCHOLARONE™
Manuscripts

Nanoscale

Guidelines for Referees

Thank you very much for agreeing to review this manuscript for *Nanoscale*.



Nanoscale is a high impact international journal, publishing high quality research across nanoscience and nanotechnology. It publishes a full mix of research articles on experimental and theoretical work, including reviews, communications and full papers.

Nanoscale Associate Editors stress very high standards for acceptance in the journal. Articles must report extremely novel, very high quality, reproducible new work of broad general interest.

Nanoscale's Impact Factor is **7.790** (2020 Journal Citation Reports®)

The following manuscript has been submitted for consideration as a
PAPER

For acceptance, papers must report original scientific work that has not been published previously. Full papers do not have a page limit and should be appropriate in length for scientific content.

When preparing your report, please:

- Focus on the **originality, importance, impact** and **reproducibility** of the science.
- Refer to the [journal scope and expectations](#).
- **State clearly** whether you think the article should be accepted or rejected and give detailed comments (with references) both to help the Editor to make a decision on the paper and the authors to improve it
- **Inform the Editor** if there is a conflict of interest, a significant part of the work you cannot review with confidence or if parts of the work have previously been published.
- **Provide your report rapidly** or inform the Editor if you are unable to do so.

Best regards,

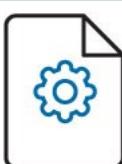
Professor Chunli Bai
Editor-in-Chief, *Nanoscale*

Professor Dirk Guldin
Editor-in-Chief, *Nanoscale*

Dr Heather Montgomery
Managing Editor, *Nanoscale*

Contact us

Please visit our [reviewer hub](#) for further details of our processes, policies and reviewer responsibilities as well as guidance on how to review, or click the links below.



What to do
when you
review



Reviewer
responsibilities



Process &
policies

Dear Editor,

Please find our uploaded manuscript “Quantum theory of electronic excitation and sputtering by transmission electron microscopy” that we are hereby submitting for consideration as an Paper in *Nanoscale*.

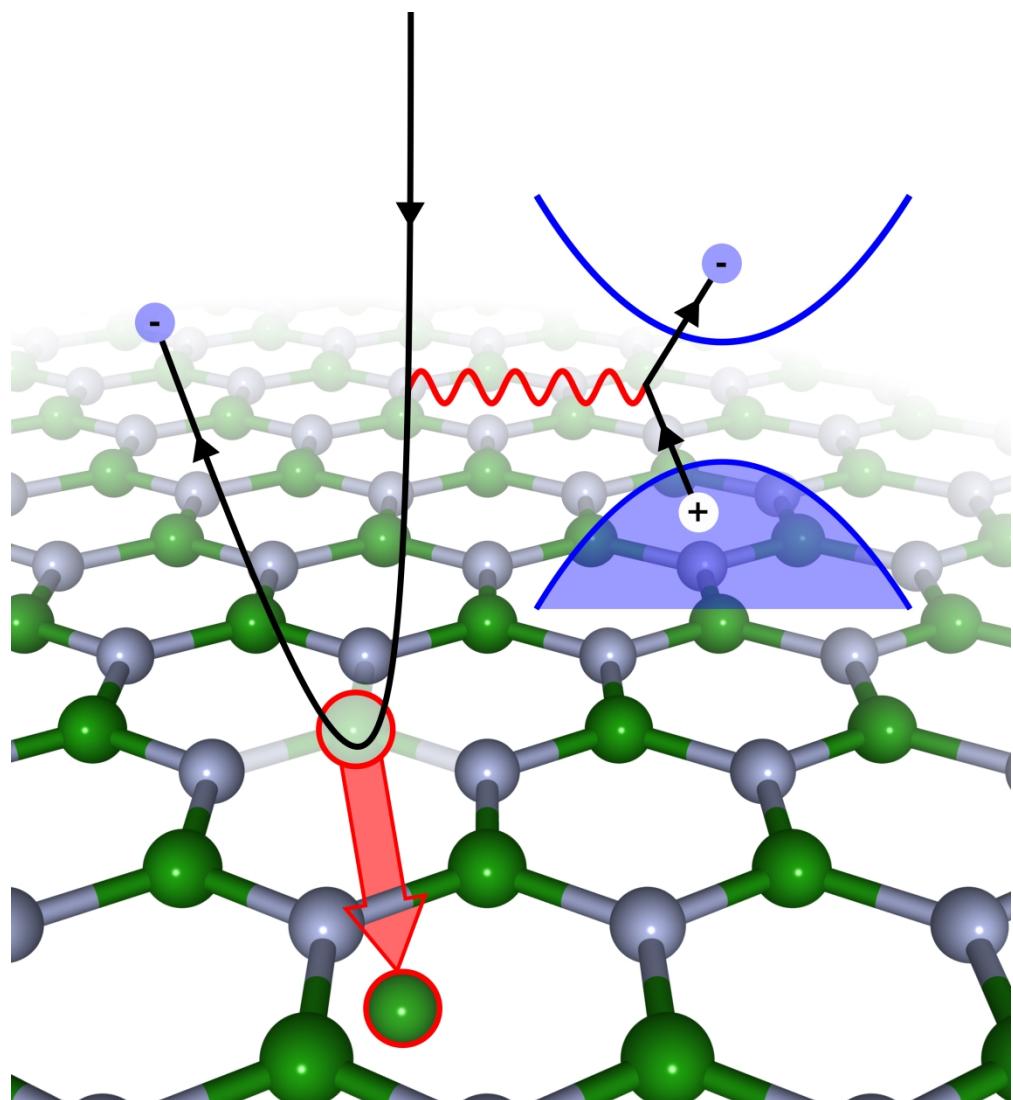
Electron beam irradiation by transmission electron microscopy (TEM) is an effective tool for defect engineering in 2-dimensional (2D) materials, potentially with atomic-scale precision. This degree of spatial control can provide powerful solutions to many of today’s most pressing challenges in nanotechnology, including quantum device fabrication, hydrogen evolution catalysis, and electronic device miniaturization. However, despite the broad application of TEM-driven defect engineering in materials science, current theoretical models fail to quantitatively describe defect formation rates, as the non-equilibrium physics taking place during irradiation remains poorly understood. Specifically, existing models drastically underestimate sputtering rates in insulators, in which beam-induced electronic excitations may weaken the bonding between the irradiated atoms. To address this anomalous behavior, this work develops the first quantum electrodynamics (QED)-based method to accurately describe electron beam-induced sputtering cross sections in 2D crystals by explicitly calculating the probabilities of beam-induced electronic excitations and their effects on sputtering kinetics. The consideration of excitations yields cross sections that quantitatively match experiment and correctly predict appreciable sputtering rates at beam energies previously predicted to leave the crystal intact.

We feel that this work is appropriate for *Nanoscale*, as it significantly advances the state of the art in modeling the response of nanomaterials to TEM irradiation. The QED-based prediction of excitation rates could also push efforts to simulate both TEM images and electron energy loss spectra towards a more analytical, quantum field theory-based direction. Lastly, we believe this new quantitative theory can help arm materials scientists and engineers with precise nanoscale control of any 2D material.

Finally, a preprint for this work is available on arXiv under identifier 2112.10002.

Thank you for your consideration.

Anthony Yoshimura, PhD
Lawrence Livermore National Laboratory
(805) 886-4810
yoshimura4@llnl.gov



887x968mm (118 x 118 DPI)

Combining quantum electrodynamics with density function theory, we model electronic excitation and sputtering by high-energy beam electrons in two-dimensional materials such as hexagonal boron nitride. Electronic excitations can drastically increase the sputtering rates in these materials.

Journal Name

ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

Quantum theory of electronic excitation and sputtering by transmission electron microscopy[†]

Anthony Yoshimura,^{a,b} Michael Lamparski,^b Joel Giedt,^b David Lingerfelt,^c Jacek Jakowski,^c Panchapakesan Ganesh,^c Tao Yu,^d Bobby Sumpter,^c and Vincent Meunier^{b,e}

Received Date

Accepted Date

DOI: 00.0000/xxxxxxxxxx

Many computational models have been developed to predict the rates of atomic displacements in two-dimensional (2D) materials under electron beam irradiation. However, these models often drastically underestimate the displacement rates in 2D insulators, in which beam-induced electronic excitations can reduce the binding energies of the irradiated atoms. This bond softening leads to a qualitative disagreement between theory and experiment, in that substantial sputtering is experimentally observed at beam energies deemed far too small to drive atomic dislocation by many current models. To address these theoretical shortcomings, this paper develops a first-principles method to calculate the probability of beam-induced electronic excitations by coupling quantum electrodynamics (QED) scattering amplitudes to density functional theory (DFT) single-particle orbitals. The presented theory then explicitly considers the effect of these electronic excitations on the sputtering cross section. Applying this method to 2D hexagonal BN and MoS₂ significantly increases their calculated sputtering cross sections and correctly yields appreciable sputtering rates at beam energies previously predicted to leave the crystals intact. The proposed QED-DFT approach can be easily extended to describe a rich variety of beam-driven phenomena in any crystalline material.

1 Introduction

The holy grail of materials engineering is atomic scale control of the material structure. Towards this aim, electron irradiation by transmission electron microscopy (TEM) can be an effective means of structural manipulation with spatial control.^{1–4} These structural changes can arise *via* atomic displacement in which an incident electron nudges a material atom from its initial site. We call this atom the primary knock-on atom (PKA). Two-dimensional (2D) crystals provide an excellent platform to measure the rates of these PKA displacements. When electron ir-

radiation is normal to a 2D crystal's surface, a displacement likely propels the PKA away from the crystal. As such, these PKA are often ejected from the crystal in a process called *sputtering*. Sputtering events leave behind vacancies, which can then be counted using TEM. Counting the number of vacancies for a given dose and beam energy allows one to experimentally determine the sputtering cross section of that crystal.

Sputtering occurs when the energy transferred to the PKA is greater than the PKA's displacement threshold E_d . This means that a displacement is possible only if the kinetic energy of the beam electron exceeds some critical energy ϵ_c . Many computational models have been proposed to predict both E_d and ϵ_c to calculate electron beam-induced sputtering rates in 2D crystals.^{4–7} However, the vast majority of current methods focus solely on interactions between the beam electrons and material nuclei, neglecting any coupling with the material's electrons. Thus, while present-day models give reasonable predictions for conductors,⁵ where electronic relaxation is rapid, they often vastly underestimate the atomic displacement rates in insulators. For example, the critical energy for sputtering boron or nitrogen from hexagonal boron nitride (hBN) is predicted to be about 80 keV.⁸ However, sputtering has been observed in hBN under 30 keV irradiation.⁹ Furthermore, selenium sputters from WSe₂ and MoSe₂ under irradiation energies of 60 and 80 keV, respectively. These energies are almost 150 keV below their predicted critical ener-

^aLawrence Livermore National Laboratory, Livermore, CA 94550, USA E-mail: yoshimura4@llnl.gov

^bDepartment of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

^cCenter for Nanophase Material Sciences, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

^dCenter for Nanophase Material Sciences, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

^eDepartment of Materials Science and Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

[†]Electronic Supplementary Information (ESI) available: Derivation of the invariant matrix element \mathcal{M} introduced in equation (4), description of the normalization scheme used for the integral in equation (6), derivation of p_3^z that conserves energy and momentum, convergence of σ with respect to n_i^{\max} , peaks in the sputtering cross section of hBN for low beam energies, fitting and convergence of S with respect to various parameters, derivation of the minimum energy transfer E_{\min} mentioned after equation (24), and details on the derivation of \mathcal{M} . See DOI: 00.0000/00000000.

gies.^{10,11} Lastly, while the calculated critical energy for sulfur sputtering in MoS₂ is about 90 keV, sulfur has been shown to sputter under 20 keV beams.¹² Discrepancies like these suggest that the displacement thresholds in insulating crystals are much smaller than what is predicted by ground-state theory. Lehnert *et al.* have proposed that the consideration of inelastic scattering, *i.e.*, beam-induced electronic excitation, can lead to such a reduction in the displacement threshold.¹¹ This would increase the sputtering cross section for all beam energies and enable sputtering for energies well-below the ground state ε_c .

To account for these effects, this paper combines quantum electrodynamics (QED) and density functional theory (DFT) to derive the probability of beam-induced electronic excitation in 2D insulating crystals. The basic idea is as follows: DFT can provide effective single-particle states that can be decomposed into a plane-wave basis,^{13–15} while QED is well-equipped to describe how each plane-wave evolves in time through interactions with an electromagnetic field.^{16–18} Thus, a plane-wave decomposition of the Kohn-Sham orbitals can allow for a component-by-component treatment of the interactions between the beam and material electrons. This generalized QED-DFT approach enables, for the first time, a first-principles description of any beam-matter interaction process. The only limitations of this method are the order to which the time-evolution operator is expanded and the sophistication of the theory used to determine the material's electronic structure. Additionally, while DFT is used here, our method is compatible with any first-principles formalism that can produce single-particle eigenstates and eigenvalues for a given material.

This paper is organized as follows. First, we derive the theoretical model to determine beam-induced sputtering cross sections in 2D insulating crystals. We then use this model to calculate sputtering cross sections in hBN and MoS₂ that quantitatively agree with experiment.

2 Theoretical Model

The majority of present-day beam-damage models focus solely on the interaction between the beam electron and target nucleus. These models are thus centered on one process: energy transfer from the beam electron to the nucleus. For this, one defines a differential cross section $d\sigma/dE(\varepsilon_b, E)$ providing the distribution of energy transfers E for a given beam energy ε_b . The McKinley and Feshbach differential cross section can adequately describe $d\sigma/dE$ for light nuclei ($Z < 20$).^{19–22} Setting $\hbar = c = 1$,

$$\frac{d\sigma}{dE} = \pi \left(\frac{Z\alpha}{|\mathbf{p}_b|\beta} \right)^2 \left[\frac{E_{\max}}{E^2} - \beta(\pi Z\alpha + \beta) \frac{1}{E} + \pi Z\alpha\beta \sqrt{\frac{E_{\max}}{E^3}} \right], \quad (1)$$

where \mathbf{p}_b is the momentum of the beam electron, β is its velocity, Z is the atomic number of the target nucleus, and α is the fine structure constant.

One can then calculate the displacement cross section by integrating $d\sigma/dE$ over all E large enough to cause a displacement, so that

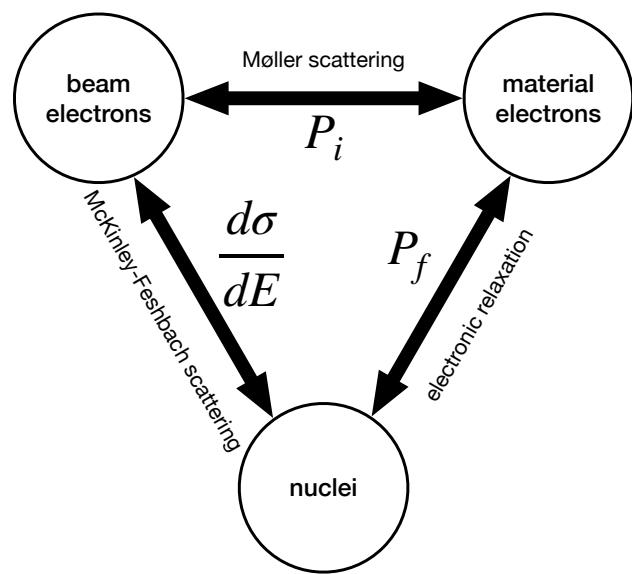


Fig. 1 Three interactions between three pairs of particles involved in electron beam-induced sputtering. In insulators, the probability of material electronic excitation (P_i) and the probability that those excitations are substantially long-lived (P_f) can be large. Therefore, the effect of exciting the material electrons on the sputtering rate should not be ignored in these materials.

$$\sigma_0(\varepsilon_b, E_d) = \theta(E_{\max}(\varepsilon_b) - E_d) \int_{E_d}^{E_{\max}(\varepsilon_b)} \frac{d\sigma}{dE}(\varepsilon_b, E) dE, \quad (2)$$

where $E_{\max}(\varepsilon_b)$ is the maximum possible energy transfer for a given beam energy, *i.e.*, the energy transfer resulting from a direct collision. The step function θ (which we will leave implicit going forward for the sake of compactness) enforces that the cross section is zero when $E_{\max} < E_d$. The critical energy ε_c is then the beam energy for which $E_{\max} = E_d$. Therefore, the observation of sputtering at energies well below ε_c is completely at odds with equation (2). With this in mind, a fair amount of work has been done to treat deviations from equation (2). Notably, several studies have explored the effects of temperature on displacement cross sections.^{5–7} This consideration involves calculating the degree to which the pre-collision thermal motion of the nucleus increases the cross section. However, these techniques essentially amount to smearing the beam energy dependence of the cross section, so that the cross section only strays significantly from equation (2) for beam energies very close to ε_c . Thus, temperature-induced increases in the cross section cannot account for the disparities between equation (2) and experiment. This necessitates the consideration of additional phenomena that can reduce E_d .

To address the limitations of equation (2), this work introduces a third party: the material's electrons. Doing so brings two new interactions into play: one between the beam and material electrons and another between material electrons and nuclei. This yields a total of three interactions between the three pairs of particles (figure 1). Therefore, the rate of beam-induced sputtering hinges on the rates of three processes mediated by these interac-

tions.

1. Beam and material electrons: a beam electron can excite some number n_i ground state electrons to the conduction band (i denotes the *initial* interaction with the beam). The probability of this event for a given beam energy ϵ_b is $P_i(\epsilon_b, n_i)$.
2. Material electrons and nuclei: Some number n_f beam-induced excitations can survive long enough for the target atom to leave its original site (f denotes the *final* system at the completion of sputtering). This depends on the nuclear kinetic energy E and the excitation lifetime τ . The probability that n_f of the n_i excitations survive is $P_f(E, \tau, n_i, n_f)$.
3. Beam electrons and material nuclei: the energy transferred to a material nucleus by the beam electron can exceed the PKA's displacement threshold $E_d(n_f)$, which depends on the number of surviving excitations n_f . We define $\{E_d\}$ as the set of all displacement thresholds for all possible n_f . Sputtering occurs when $E > E_d(n_f)$. The differential cross section for an energy transfer E from the beam electron to material nucleus is $d\sigma/dE(\epsilon_b, E)$.

The sputtering cross section can then be calculated by coupling $d\sigma/dE$ to P_i and P_f for all possible n_i and n_f . With the terms defined above, this excitation-sensitive sputtering cross section can be written as

$$\begin{aligned} \sigma(\epsilon_b, \{E_d\}, \tau) = & \sum_{n_i=0}^{\infty} P_i(\epsilon_b, n_i) \\ & \times \sum_{n_f=0}^{n_i} \int_{E_d(n_f)}^{E_{\max}(\epsilon_b)} P_f(E, \tau, n_i, n_f) \frac{d\sigma}{dE}(\epsilon_b, E) dE. \end{aligned} \quad (3)$$

If P_i and P_f are non-negligible when n_i and n_f are nonzero, and E_d depends strongly on n_f , then interactions with the material electrons must be considered. We will later show that this makes σ in equation (3) larger than σ_0 in equation (2) for all beam energies, most prominently when $\epsilon_b < \epsilon_c$.

The remainder of this section focuses on the derivations of P_i , P_f , and $\{E_d\}$. Section 2.1 describes how to combine QED with DFT to obtain P_i . Section 2.2 then considers the evolution of the excited states during the sputtering process to derive $\{E_d\}$ and P_f .

2.1 Probability of beam-induced excitation

For a crystal in its ground state, an occupied electron energy eigenstate has zero overlap with any unoccupied state. However, the collision of a beam electron can give an occupied state a momentum boost that breaks this orthogonality. Thus, the boosted ground state has a nonzero probability of being measured in an excited state. We can use this idea to derive $P_i(\epsilon_b, n_i)$, the probability that a beam electron with kinetic energy ϵ_b excites exactly n_i material electrons. The derivation can be broken down into four steps: (i) determine the amplitude for a free electron to scatter from one momentum eigenstate into another after collision with

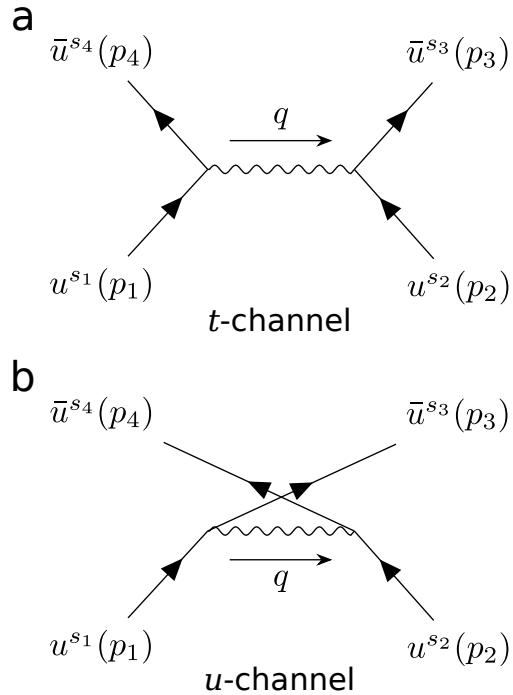


Fig. 2 The lowest order electron-electron scattering perturbation includes two Feynman diagrams called the (a) *t*-channel and (b) *u*-channel. The incoming and outgoing electron states are represented by Dirac spinors $u^s(p)$ and $\bar{u}^s(p)$ respectively, where p and s are the electron's 4-momentum and spin index respectively. Subscripts 1 and 2 label components of the initial beam and material states respectively. The virtual photon 4-momentum q is the momentum transfer between the electrons.

another free electron; (ii) generalize the formalism to obtain the amplitude for scattering from one wave packet into another by summing over the amplitudes for each momentum component of one wave packet to scatter into each momentum component of the other; (iii) decompose a pair of occupied and unoccupied crystal states into a momentum basis and plug them in as incoming and outgoing wave packets respectively, then square the amplitude to obtain the corresponding excitation probability for a particular transition; (iv) Compute the sum of all transition probabilities and use combinatorics to determine $P_i(\epsilon_b, n_i)$. The following subsections address each step (i-iv) in detail.

2.1.1 Scattering of free electrons

We first derive the scattering amplitude for momentum transfer between two free electrons via Møller scattering.^{23–25} Going forward, we label the 4-momenta of the incoming electrons as p_1 and p_2 , while the outgoing electrons have momenta p_3 and p_4 . We also choose to make p_1 and p_2 components of the initial beam and material states respectively. The 4-momentum of the n th electron can be written as $p_n = (\epsilon_n, p_n^x, p_n^y, p_n^z) = (\epsilon_n, \mathbf{p}_n)$, where ϵ_n is the particle's energy and \mathbf{p}_n is its 3-momentum. Dot products between 4-vectors are then taken over Minkowski space, so that $p_n \cdot p_m = g_{\mu\nu} p_n^\mu p_m^\nu = \epsilon_n \epsilon_m - \mathbf{p}_n \cdot \mathbf{p}_m$, where $\mathbf{p}_n \cdot \mathbf{p}_m = p_n^x p_m^x + p_n^y p_m^y + p_n^z p_m^z$.

To lowest order, the amplitude for free electron scattering can be represented by two tree-level diagrams, which we call the *t*- and *u*-channels (figure 2). Using Feynman's rules,^{17,18} we can

write these diagrams in terms of Dirac spinors, yielding the invariant matrix element

$$\mathcal{M}(p_4 p_3 \leftarrow p_2 p_1) = \frac{e^2}{2} \sum_{s_1} \sum_{s_2} \sum_{s_3} \sum_{s_4} \left[\bar{u}^{s_4}(p_4) \gamma^\mu u^{s_1}(p_1) \left(\frac{1}{p_3 - p_2} \right)^2 \bar{u}^{s_3}(p_3) \gamma_\mu u^{s_2}(p_2) + \bar{u}^{s_3}(p_3) \gamma^\mu u^{s_1}(p_1) \left(\frac{1}{p_4 - p_2} \right)^2 \bar{u}^{s_4}(p_4) \gamma_\mu u^{s_2}(p_2) \right], \quad (4)$$

where $s_n = 1$ or 2 denotes the spin of the n th electron, $u^s(p)$ is a Dirac spinor, and $\bar{u}^s(p)$ is its conjugate (section S1). The factor of $1/2$ before the summation arises from the assumption that the incoming states are spin unpolarized.

The first term in brackets is the t -channel describing momentum transfer $p_3 - p_2$ and the second is the u -channel describing momentum transfer $p_4 - p_2$. Because the DFT cutoff energy (hundreds of eV) is much smaller than the beam energy (tens of keV), it is always the case that $|\mathbf{p}_2|$ is much smaller than $|\mathbf{p}_1|$. Furthermore, we need only consider outgoing momenta for which the kinetic energy associated with either $|\mathbf{p}_3|$ or $|\mathbf{p}_4|$ falls within the DFT cutoff energy. In these cases, the magnitude of one outgoing momentum is similar to $|\mathbf{p}_2|$, while that of the other is much greater. This means that one channel's momentum transfer is always much larger than the other's. As the momentum transfers reside in the denominators of either channel in equation (4), it follows that one channel always contributes much more to \mathcal{M} than the other. Thus, when the t -channel is significant, the u -channel is negligible, and vice versa. Additionally, when integrating over all possible outgoing momenta, the contribution of the t -channel is equal to that of the u -channel. Taking advantage of this along with the indistinguishability of the electrons, we calculate only the t -channel and multiply the resulting amplitude by 2 instead of calculating both channels and adding them. We can then define the 4-momentum transfer between the electrons as that of the t -channel: $q \equiv p_3 - p_2$. Because the t -channel has a q^2 in the denominator, the resulting scattering probability is proportional to q^{-4} . This makes large momentum transfers statistically irrelevant, allowing us to only consider momentum transfers inside the first Brillouin zone (BZ). The evaluation of the t -channel in terms of the components of the electrons' 4-momenta is straightforward, though cumbersome, and is described in section S1.

We can use the resulting \mathcal{M} to obtain the free electron scattering amplitude

$$\langle p_4 p_3 | \hat{T} | p_2 p_1 \rangle = (2\pi)^4 \delta(p_1 + p_2 - p_3 - p_4) \mathcal{M}(p_4 p_3 \leftarrow p_2 p_1), \quad (5)$$

where \hat{T} is the scattering operator.^{17,18} This gives the amplitude for two free electrons with momenta p_1 and p_2 to scatter into p_3 and p_4 . Equation (5) is used to derive the scattering amplitude between two arbitrary wave packets in the next subsection.

2.1.2 Scattering of wave packets

The free particle scattering amplitude in equation (5) can be used to determine the amplitude for the scattering of two arbitrary electron states $|\phi_1\rangle$ and $|\phi_2\rangle$ into $|\phi_3\rangle$ and $|\phi_4\rangle$. This is obtained by sandwiching the scattering operator between the initial and final 2-particle states, *i.e.*,

$$\langle \phi_4 \phi_3 | \hat{T} | \phi_2 \phi_1 \rangle = \int \frac{d^3 p_4 d^3 p_3 d^3 p_2 d^3 p_1}{(2\pi)^{12} 16 \epsilon_4 \epsilon_3 \epsilon_2 \epsilon_1} \times \langle \phi_4 \phi_3 | p_4 p_3 \rangle \langle p_4 p_3 | \hat{T} | p_2 p_1 \rangle \langle p_2 p_1 | \phi_2 \phi_1 \rangle. \quad (6)$$

On the right side, we have inserted two resolutions of the identity given in equation (S8). Inserting equation (5) into the integrand, the amplitude can be written in terms of the invariant matrix element \mathcal{M} , becoming

$$\int \frac{d^3 p_4 d^3 p_3 d^3 p_2 d^3 p_1}{(2\pi)^{12} 16 \epsilon_4 \epsilon_3 \epsilon_2 \epsilon_1} \langle \phi_4 | p_4 \rangle \langle \phi_3 | p_3 \rangle \langle p_2 | \phi_2 \rangle \langle p_1 | \phi_1 \rangle \times \mathcal{M}(p_4 p_3 \leftarrow p_2 p_1) (2\pi)^4 \delta(p_1 + p_2 - p_3 - p_4). \quad (7)$$

Using the delta function to integrate over \mathbf{p}_4 and \mathbf{p}_3^z yields

$$\int \frac{d^2 p_3^\perp d^3 p_2 d^3 p_1}{(2\pi)^8 16 \epsilon_2 \epsilon_1} \frac{\mathcal{M}(p_4 p_3 \leftarrow p_2 p_1)}{|p_3^z \epsilon_4 - p_4^z \epsilon_3|} \times \langle \phi_4 | p_4 \rangle \langle \phi_3 | p_3 \rangle \langle p_2 | \phi_2 \rangle \langle p_1 | \phi_1 \rangle, \quad (8)$$

where it is understood that p_4 and p_3 satisfy $p_1 + p_2 = p_3 + p_4$ (section S3). The normalization of 4-momentum states $|p_n\rangle$ in terms of 3-momentum states $|\mathbf{p}_n\rangle$ as defined in equation (S9) allows us to rewrite the expression as

$$\int \frac{d^2 p_3^\perp d^3 p_2 d^3 p_1}{(2\pi)^8} \sqrt{\frac{\epsilon_4 \epsilon_3}{\epsilon_2 \epsilon_1}} \frac{\mathcal{M}(p_4 p_3 \leftarrow p_2 p_1)}{4 |p_3^z \epsilon_4 - p_4^z \epsilon_3|} \times \langle \phi_4 | \mathbf{p}_4 \rangle \langle \phi_3 | \mathbf{p}_3 \rangle \langle \mathbf{p}_2 | \phi_2 \rangle \langle \mathbf{p}_1 | \phi_1 \rangle. \quad (9)$$

We can then discretize the momenta by replacing $d^3 p_i / (2\pi)^3$ with V^{-1} and $d^2 p_i^\perp / (2\pi)^2$ with A^{-1} , where V and A are the volume and cross sectional area of the simulated crystal respectively, *i.e.*, the volume and cross sectional area of the unit cell times the number of k-points used to sample the BZ. With this, the amplitude for electron states $|\phi_1\rangle$ and $|\phi_2\rangle$ to scatter into $|\phi_3\rangle$ and $|\phi_4\rangle$ takes the form

$$\frac{1}{AV^2} \sum_{\mathbf{p}_3} \sum_{\mathbf{p}_2} \sum_{\mathbf{p}_1} \sqrt{\frac{\epsilon_4 \epsilon_3}{\epsilon_2 \epsilon_1}} \frac{\mathcal{M}(p_4 p_3 \leftarrow p_2 p_1)}{4 |p_3^z \epsilon_4 - p_4^z \epsilon_3|} \times \langle \phi_4 | \mathbf{p}_4 \rangle \langle \phi_3 | \mathbf{p}_3 \rangle \langle \mathbf{p}_2 | \phi_2 \rangle \langle \mathbf{p}_1 | \phi_1 \rangle. \quad (10)$$

In the next subsection, we replace $|\phi_{1\dots 4}\rangle$ with states relevant to electron beam-induced excitation.

2.1.3 Probability of a crystal excitation

We now consider the specific case of beam-induced excitations to determine the form of the four electron states in equation (10). We assign $|\phi_1\rangle$ and $|\phi_4\rangle$ to the initial and final beam states $|\mathbf{p}\rangle$

and $|\mathbf{p}'_b\rangle$ respectively. States $|\phi_2\rangle$ and $|\phi_3\rangle$ are then the ground and excited crystal states $|n\mathbf{k}\rangle$ and $|n'\mathbf{k}'\rangle$ respectively, where n and n' are band indices and \mathbf{k} and \mathbf{k}' are k-points. Substituting these specific states into expression (10), the amplitude for exciting $|n\mathbf{k}\rangle$ to $|n'\mathbf{k}'\rangle$ becomes

$$\begin{aligned} \langle \mathbf{p}'_b, n'\mathbf{k}' | \hat{T} | n\mathbf{k}, \mathbf{p}_b \rangle &= \frac{1}{AV^2} \sum_{\mathbf{p}_3} \sum_{\mathbf{p}_2} \sum_{\mathbf{p}_1} \sqrt{\frac{\epsilon_4 \epsilon_3}{\epsilon_2 \epsilon_1}} \frac{\mathcal{M}(p_4 p_3 \leftarrow p_2 p_1)}{4 |p_3^\perp \epsilon_4 - p_4^\perp \epsilon_3|} \\ &\quad \times \langle \mathbf{p}'_b | \mathbf{p}_4 \rangle \langle n'\mathbf{k}' | \mathbf{p}_3 \rangle \langle \mathbf{p}_2 | n\mathbf{k} \rangle \langle \mathbf{p}_1 | \mathbf{p}_b \rangle. \end{aligned} \quad (11)$$

The values of $\epsilon_{1\dots 4}$ need to be clarified before moving forward. The zeroth components of the initial and final beam momenta obey the free particle dispersion relations, so $\epsilon_1 = \sqrt{|\mathbf{p}_1|^2 + m^2}$ and $\epsilon_4 = \sqrt{|\mathbf{p}_4|^2 + m^2}$. The beam energy that appears in equation (3) is then defined as the beam electron's total energy minus its rest mass: $\epsilon_b \equiv \epsilon_1 - m$. Meanwhile, the momentum of the crystal states can be treated nonrelativistically. Thus, the zeroth components of the crystal state momenta are the energy eigenvalues of the crystal state plus the electron rest mass, *i.e.*, $\epsilon_2 = \epsilon_{n\mathbf{k}} + m$ and $\epsilon_3 = \epsilon_{n'\mathbf{k}'} + m$. For the remainder of this derivation, we continue to leave our expressions in terms of $\epsilon_{1\dots 4}$ for compactness.

Sputtering from a 2D crystal often requires that the beam electron is backscattered or nearly backscattered, in which case, its final trajectory after collision with the nucleus is nearly antiparallel to its initial trajectory and perpendicular to the crystal surface. Given that many 2D materials (including hBN and MoS₂) possess inversion and/or reflection symmetry about the crystal plane, we assume that the likelihood of excitation before and after the collision are about equal. In light of this, we calculate the excitation probability during a sputtering event assuming the beam electron's trajectory is not altered by its collision with the nucleus. That is, we impose that $\mathbf{p}_1 = |\mathbf{p}_1| \hat{\mathbf{z}}$ until an electronic excitation is induced.

We can now evaluate the bra-ket products in equation (11). The initial beam state is highly localized on \mathbf{p}_b , meaning that

$$|\langle \mathbf{p}_1 | \mathbf{p}_b \rangle|^2 = V \delta_{\mathbf{p}_1, \mathbf{p}_b} \Rightarrow \langle \mathbf{p}_1 | \mathbf{p}_b \rangle = \sqrt{V} \delta_{\mathbf{p}_1, \mathbf{p}_b}. \quad (12)$$

Meanwhile, the ground and excited crystal states can be expanded into a plane-wave basis, so that

$$\begin{aligned} |n\mathbf{k}\rangle &= \frac{1}{\sqrt{V}} \sum_{\mathbf{G}} C_{\mathbf{G}+\mathbf{k}}^n |\mathbf{G} + \mathbf{k}\rangle \\ |n'\mathbf{k}'\rangle &= \frac{1}{\sqrt{V}} \sum_{\mathbf{G}} C_{\mathbf{G}+\mathbf{k}'}^{n'} |\mathbf{G} + \mathbf{k}'\rangle, \end{aligned} \quad (13)$$

where each \mathbf{G} is a reciprocal lattice vector. By re-expressing \mathbf{p}_2 and \mathbf{p}_3 as $\mathbf{G}_2 + \mathbf{k}_2$ and $\mathbf{G}_3 + \mathbf{k}_3$ respectively, we find

$$\begin{aligned} \langle \mathbf{p}_2 | n\mathbf{k} \rangle &= \langle \mathbf{G}_2 + \mathbf{k}_2 | \frac{1}{\sqrt{V}} \sum_{\mathbf{G}} C_{\mathbf{G}+\mathbf{k}}^n |\mathbf{G} + \mathbf{k}\rangle = \sqrt{V} C_{\mathbf{G}_2 + \mathbf{k}_2}^n \delta_{\mathbf{k}_2, \mathbf{k}}, \\ \langle \mathbf{p}_3 | n'\mathbf{k}' \rangle &= \langle \mathbf{G}_3 + \mathbf{k}_3 | \frac{1}{\sqrt{V}} \sum_{\mathbf{G}} C_{\mathbf{G}+\mathbf{k}'}^{n'} |\mathbf{G} + \mathbf{k}'\rangle = \sqrt{V} C_{\mathbf{G}_3 + \mathbf{k}_3}^{n'} \delta_{\mathbf{k}_3, \mathbf{k}'}. \end{aligned} \quad (14)$$

Lastly, we do not care where the outgoing scattered electron ends up, so we wish for $|\mathbf{p}'_b\rangle$ to satisfy

$$|\langle \mathbf{p}_4 | \mathbf{p}'_b \rangle|^2 = V \Rightarrow \langle \mathbf{p}_4 | \mathbf{p}'_b \rangle = \sqrt{V}. \quad (15)$$

The excitation amplitude is then obtained by plugging in the bracket products from equations (12), (14), and (15) into equation (11). This gives us the excitation amplitude

$$\begin{aligned} \langle \mathbf{p}'_b, n'\mathbf{k}' | \hat{T} | n\mathbf{k}, \mathbf{p}_b \rangle &= \frac{1}{A} \sum_{\mathbf{G}_3^\perp} \sum_{\mathbf{G}_2} \sqrt{\frac{\epsilon_4 \epsilon_3}{\epsilon_1 \epsilon_2}} \\ &\quad \times \frac{\mathcal{M}(p_4 p_3 \leftarrow p_2 p_1)}{4 |p_3^\perp \epsilon_4 - p_4^\perp \epsilon_3|} C_{\mathbf{G}_3 + \mathbf{k}'}^{n'*} C_{\mathbf{G}_2 + \mathbf{k}}^n, \end{aligned} \quad (16)$$

where it is understood that $p_2 = (m + \epsilon_{n\mathbf{k}}, \mathbf{G}_2 + \mathbf{k})$ and $p_3 = (m + \epsilon_{n'\mathbf{k}'}, \mathbf{G}_3 + \mathbf{k}')$, and p_4 and p_3^\perp satisfy $p_1 + p_2 = p_4 + p_3$. Squaring this amplitude yields the probability of a single electronic excitation from the valence band state $|n\mathbf{k}\rangle$ to the conduction band state $|n'\mathbf{k}'\rangle$ for a given beam energy ϵ_b , that is,

$$P(n'\mathbf{k}' \leftarrow n\mathbf{k} | \epsilon_b) = |\langle \mathbf{p}'_b, n'\mathbf{k}' | \hat{T} | n\mathbf{k}, \mathbf{p}_b \rangle|^2. \quad (17)$$

2.1.4 Probability of n_i excitations

We are finally ready to derive $P_i(\epsilon_b, n_i)$, the probability that a beam electron excites a particular number of electrons n_i . First, we define the sum of all transition probabilities

$$S(\epsilon_b) \equiv \sum_{\mathbf{k}} \sum_{\mathbf{k}'} \sum_n \sum_{n'} P(n'\mathbf{k}' \leftarrow n\mathbf{k} | \epsilon_b) \equiv \sum_j^N P_j(\epsilon_b), \quad (18)$$

where \mathbf{k} and \mathbf{k}' run over all k-points, n over the valence bands, and n' over the conduction bands (those possessing states with energy between the Fermi level and the work function). The index j in the right-most expression labels the possible single-particle excitations (*e.g.*, $j = n'\mathbf{k}' \leftarrow n\mathbf{k}$).

Combinatorics tells us that the probability of exciting exactly one excitation is

$$\begin{aligned} P_i(n_i = 1) &= \sum_{j_1} P_{j_1} \prod_{j_2 \neq j_1} (1 - P_{j_2}) \\ &= \sum_{j_1} P_{j_1} \left(1 - \sum_{j_2 \neq j_1} P_{j_2} + \frac{1}{2} \sum_{j_3 \neq j_1, j_2} \sum_{j_2 \neq j_1} P_{j_2} P_{j_3} - \dots \right). \end{aligned} \quad (19)$$

In the large-crystal limit, the number of states, and thus the number of transitions, is large so that the summations over j_i in equation (19) are approximately equal to one another. That is,

$$\sum_{j_2 \neq j_1} P_{j_2} \sim \sum_{j_2} P_{j_2} = S. \quad (20)$$

In this limit, the probability of exactly one beam-induced excitation can be written as

$$P_i(n_i = 1) \sim \sum_j P_j \left(1 - S + \frac{1}{2} S^2 - \dots \right) = e^{-S} \sum_j P_j = S e^{-S}. \quad (21)$$

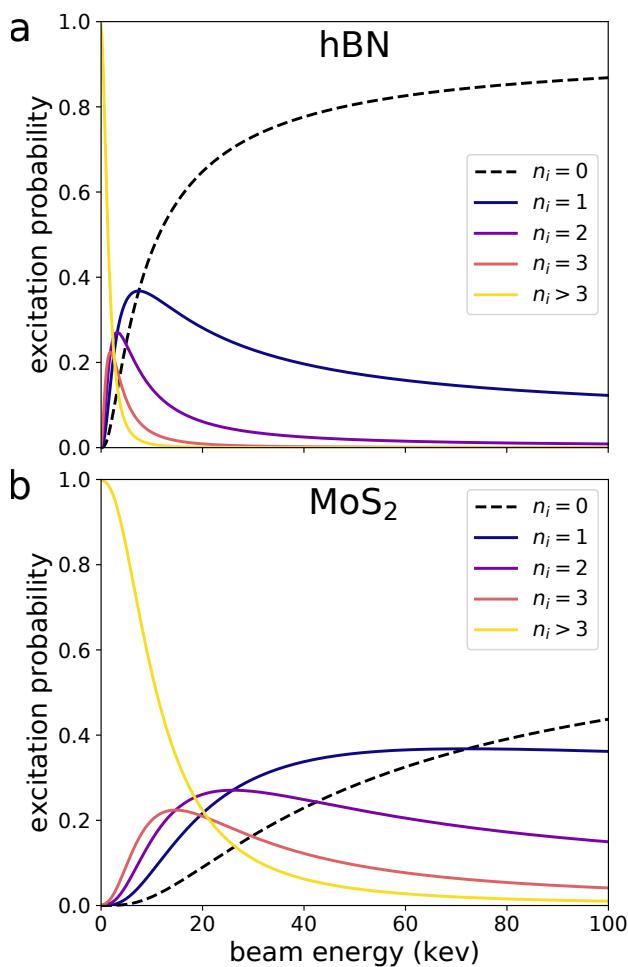


Fig. 3 Probabilities of initially exciting a certain number of electrons n_i in (a) hBN and (b) MoS₂ with a beam electron. The excitation probabilities tend to decrease with increasing beam energy. The excitation probabilities in MoS₂ are larger than those in hBN because MoS₂ has a significantly smaller band gap.

In the same way, the probability of two excitations is

$$P_i(n_i = 2) = \frac{1}{2} \sum_{j_1} \sum_{j_2 \neq j_1} P_{j_1} P_{j_2} \prod_{j_3 \neq j_1, j_2} (1 - P_{j_3}) \sim \frac{S^2}{2} e^{-S}. \quad (22)$$

In general, the probability of exactly n_i beam-induced excitations is approximately

$$P_i(\epsilon_b, n_i) \sim \frac{S(\epsilon_b)^{n_i}}{n_i!} e^{-S(\epsilon_b)}. \quad (23)$$

Thus, we see that the probability P_i can be written purely in terms of S .

We can now use formula (23) to calculate excitation probabilities in hBN and MoS₂ (figure 3). DFT is used to obtain the plane-wave coefficients $C_{\mathbf{G}_2 + \mathbf{k}}$ and $C_{\mathbf{G}_3 + \mathbf{k}'}$ and eigenvalues $\epsilon_{n\mathbf{k}}$ and $\epsilon_{n'\mathbf{k}'}$ for the pristine unit cell of each material. These are plugged into equation (16) to obtain the amplitude for each transition. We sum over the squares of all resulting amplitudes to obtain S , which is then plugged into formula (23) to obtain P_i for both materials.

We emphasize that the only DFT calculation needed to determine P_i is the electronic structure relaxation of a pristine unit cell, a very inexpensive calculation.

The probabilities plotted in figure 3 reveal some notable trends. First, for sufficiently large beam energies, $P_i(\epsilon_b, n_i)$ decreases with increasing ϵ_b for all $n_i > 0$. This is because a faster beam electron has less time to interact with the material and cause an excitation. In this regime, $P_i(\epsilon_b)$ is proportional to ϵ_b^{-1} , a relationship originally predicted by Bethe.^{12,26,27} This means that multiple excitations are more likely at low beam energies. Furthermore, the probability of remaining in the ground state $P(\epsilon_b, n_i = 0)$ vanishes as ϵ_b goes to zero. This implies that a stationary electron in the vicinity of a material is guaranteed to interact with the material's electrons and affect its electronic structure. However, the validity of formula (23) diminishes as ϵ_b approaches zero. In the case of a slow beam electron, the interaction between the beam and material electrons can no longer be approximated by the single virtual photon transfer processes depicted in figure 2, as the amplitudes for higher-order processes become more significant. The effects of processes beyond the tree-level should be the subject of future work. Lastly, the excitation probability is inversely proportional to the material's band gap. This is because the zeroth component of the momentum transfer q depicted in figure 2 is the difference in energy eigenvalues between the occupied and unoccupied states (section 2.1.3). Thus, the smallest possible denominator of the t -channel in equation (4) is proportional to the difference in eigenvalues squared. The experimentally measured band gap of MoS₂ is about 1.9 eV²⁸ while that of hBN is about 6.1 eV.²⁹ This means MoS₂ hosts transitions with smaller eigenvalue differences, making the summands in equation (16) larger.

With $P_i(\epsilon_b, n_i)$ derived in formula (23) and $d\sigma/dE(\epsilon_b, E)$ defined in equation (1), we now have two of the three functions depicted in figure 1 needed to calculate the sputtering cross section in equation (3). The final ingredients are $P_f(E, \tau, n_i, n_f)$, the probability that n_f excitations survive the displacement event given n_i initial beam-induced excitations, and $\{E_d\}$, the set of all displacement thresholds for all n_f . The derivations of these objects are described in the next section.

2.2 Sputtering cross section

We have demonstrated how the interaction between the beam and material electrons can induce electronic excitations in the material. In this section, we show how these excitations bring about much larger sputtering cross sections than those predicted by a ground state theory. We start by showing how electronic excitations can reduce the displacement threshold E_d . We then show that longer excitation lifetimes increase P_f for nonzero n_f , giving the beam-induced excitations more opportunities to lessen E_d . This motivates us to write the sputtering cross section in equation (3) analytically in terms of the excitation lifetime τ . Finally, in the next section, the resulting equation is used to predict the sputtering rates of boron and nitrogen in hBN and sulfur in MoS₂, which can be made to agree well with experiment for appropriate values of τ for each material.

2.2.1 Effect of excitations on the displacement threshold

We begin by describing how beam-induced excitations can reduce the displacement threshold E_d in a process called *bond softening*. As E_d is the lower bound of integration over E in equation (3), and the differential cross section in equation (1) behaves like E^{-2} for small E , reductions in E_d can greatly increase the sputtering cross section. Exactly how excitations change E_d is an ambitious study on its own, requiring a careful consideration of the excited electrons' evolution and various relaxation pathways.³⁰ Here instead, we make three simplifying assumptions that allow us to calculate E_d with only ground state DFT (see section S7 for a comment on the validity of these assumptions).

Before describing these assumptions, we first define some important terms. Consider the moment immediately after the beam electron collides with a material nucleus. The nucleus now has a velocity corresponding to the kinetic energy E transferred from the beam electron. The resulting nuclear motion away from its equilibrium position causes the energy of the system to increase. In this sense, the system climbs an energy *surface* from the bottom of its equilibrium well. Far away from the well's bottom, the energy surface eventually plateaus. If the system reaches this plateau, the displaced PKA moves freely away from its initial site without deceleration. At this point, we consider the PKA to have sputtered. We call the energy at the well's bottom E_0 and that at the plateau E_s . If the energy surface is static throughout the entire process, then the displacement threshold is simply $E_d = E_s - E_0$. When $E > E_d$, the system has enough energy to climb out of the well, and the PKA sputters. Our task now is to determine how beam-induced excitations in the material electrons and their subsequent relaxation affect E_0 and E_s . To facilitate this, we make the following assumptions.

Assumption 1: the excited electrons and holes occupy the band edges. To justify this, we shift our focus to the material electrons immediately after the collision. There are now n_i electrons in the conduction band and n_i holes in the valence bands. Kretschmer *et al.* simulated the time-evolution of these excitations using Ehrenfest dynamics.^{12,31} They found that the excited electrons relax nonradiatively to the conduction band minimum (CBM) in a few femtoseconds, while the holes in the valence band take a similar amount of time to relax to the valence band maximum (VBM). In contrast, the PKA takes several hundreds of femtoseconds to fully sputter.⁷ Thus, the nonradiative relaxation of the electronic structure is essentially instantaneous, and we can assume that all excited electrons and holes occupy the CBM and VBM respectively before the PKA has been displaced by an appreciable amount.

These findings greatly simplify the calculation of E_0 . Given n_i excitations, E_0 is just the energy of the pristine system with n_i electrons and holes in the CBM and VBM respectively. In the large crystal limit, this amounts to adding the pristine band gap E_g to the system's ground state energy n_i times. In other words, $E_0(n_i) = E_0(0) + n_i E_g$. This approximation of course ignores any binding energy between the electron and hole. However, we believe that this formalism allows for an efficient treatment of the lowest order effects of excitation on E_0 .

Assumption 2: the fully sputtered system is in its ground state.

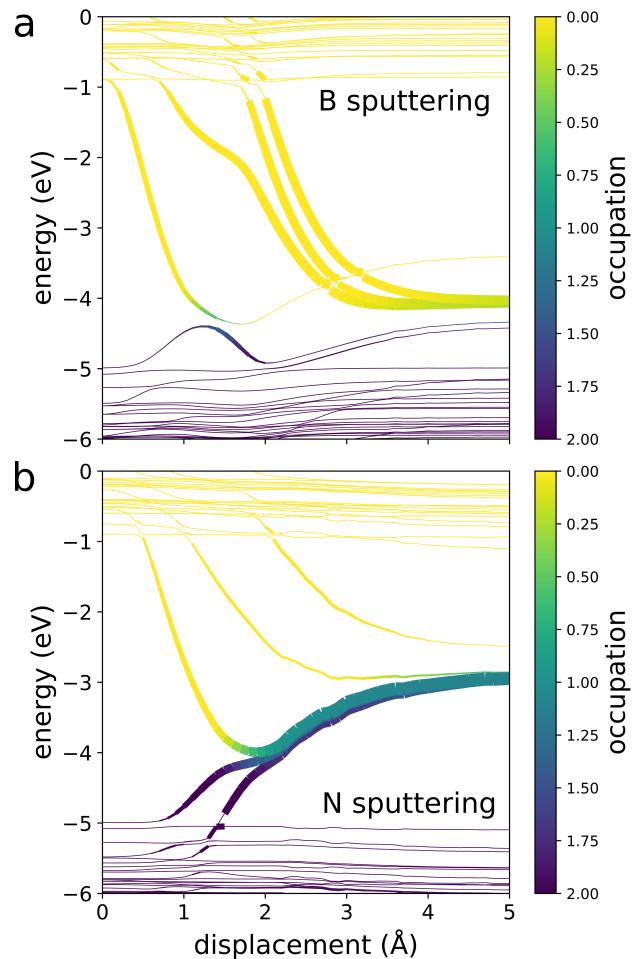


Fig. 4 Eigenvalues and occupations of the eigenstates of armchair hBN throughout the sputtering process with respect to the distance between sputtered (a) boron and (b) nitrogen and their pristine sites. The thickness of each curve is proportional to the localization of the corresponding state on the sputtered atom. In hBN, the electronegative nitrogen atoms borrow charge from the neighboring boron atoms. Thus, the p-orbital states that would be occupied on an isolated boron atom reside in the conduction band of hBN and descend into the gap as boron sputters. On the other hand, most of the p-orbital states of nitrogen lie in the valence band and rise into the gap during nitrogen sputtering.

Finding the plateau energy E_s can be a bit more involved. To calculate it properly, one must track how the excitations in the CBM and VBM evolve as the PKA moves away from the crystal. This would require Ehrenfest dynamics of a supercell over timescales of hundreds of femtoseconds for each excitation considered, which is prohibitively expensive. Here we seek a much less costly set of DFT calculations that can still provide a reasonable approximation for E_s . To this end, we draw upon another finding of Kretschmer *et al.*¹² After excitation, the displacement of the PKA causes the occupied and unoccupied CBM and VBM states to converge into the band gap and localize on the resulting defect. These converging states are the bonding and antibonding states that connect the PKA to the host crystal. Thus, we should consider how the beam-induced excitations affect the electronic structure on both the PKA and the remaining vacancy.

We start by considering how the ground state eigenvalues evolve as the PKA moves away from the crystal. For hBN, the evolution of eigenvalues differ depending on whether boron or nitrogen is sputtered (figure 4). Boron is electropositive while nitrogen is electronegative. Thus, in an hBN crystal, the nitrogen atoms borrow negative charge from the neighboring boron atoms. This means that the p-orbital states that would be occupied on an isolated boron atom are vacant, hovering in the conduction band of hBN. Conversely, the p-orbital states of nitrogen lie occupied in the valence band. For both atoms, separation from the crystal causes those states to converge to degenerate p-orbitals on their respective atoms. When boron is sputtered, these states come from the conduction band. This means that excited electrons residing in the CBM tend to transfer negative charge to the sputtered boron. However, this charge transfer is quite energetically unfavorable since boron is electropositive. The energy cost of this is much larger than the band gap of hBN, compelling any excess electrons on boron to relax to the host crystal. On the other hand, the states that localize on sputtered nitrogen must rise from the valence band. This means that the holes at the VBM tend to transfer positive charge to the sputtered nitrogen, which is again energetically unfavorable for the electronegative atom. Thus, because charge transfer to the sputtered PKA has such a large energy cost, we presume that in most cases, the PKA becomes charge neutral by the time sputtering has occurred.

Furthermore, because the states that localize on the PKA converge to the same degenerate p-orbital, the PKA assumes its ground state after it has sputtered, regardless of how many electrons the beam excites. We call the energy of the isolated ground state atom E_a . Meanwhile, the convergence of the bonding and antibonding states localized on the remaining vacancy during the displacement causes the excited electron and hole energy levels to cross multiples times. This means that there are plenty of chances for an excited electron to relax nonradiatively as the PKA separates from the material. Thus, we assume that the host crystal also relaxes to its ground state for most sputtering events.

Lastly, the sputtering threshold is always larger than the vacancy formation energy, because some of the energy transferred to the PKA disperses to the neighboring atoms. For this reason, we calculate the energy of the vacant system without relaxing the structure, as it has been shown that the free energy gained by leaving the structure unrelaxed roughly matches the energy dissipated to the surrounding material.³² We call the energy of the unrelaxed ground state vacancy E_v . Thus, we calculate the plateau energy using $E_s = E_v + E_a$, the sum of the ground state vacancy and isolated atom free energies.

Assumption 3: the PKA velocity is initially constant. Each number of excitations creates its own energy surface. In this way, electronic excitation and relaxation cause the system to hop from surface to surface. In our problem, the initial beam-induced excitation perches the system on an elevated surface. Once there, electronic relaxation via spontaneous emission (SE) enables downward surface hopping at the cost of the emitted photon energy. In this framework, E_d is the energy gained along all energy surfaces that the system traverses during sputtering. For example, suppose

the system emits one photon as it sputters. E_d is then the energy gained along the portion of the excited energy surface traversed before emission plus that gained along the relaxed surface after emission. This is how SE can change E_d . However, the effect of SE on E_d must diminish as the PKA moves further away from the crystal. This means there must be some distance d beyond which SE no longer affects the forces on the sputtered atom, leaving the energy surface unchanged. We take $d = 4.5 \text{ \AA}$ in line with Kretschmer *et al.*'s study, though we acknowledge that the exact meaning of this distance is different in their work.¹²

We wish to capture this idea while making the calculation of P_f as intuitive as possible. We do this by assuming that each energy surface is constant except for a step when the PKA reaches a distance d away from its equilibrium site. This means that immediately after the collision that excites n_i electrons, the surface assumes a constant energy $E_0(n_i)$. Each relaxation via SE causes the system to drop to the surface beneath it, decreasing its energy by E_g . We define n_f as the number of electronic excitations that survive the PKA's traversal of distance d to the energy step. Thus, the energy of the system just before the PKA reaches the step is $E_0(n_f)$. Beyond the step, all energy surfaces have energy E_s , and SE no longer changes the energy surface. With this formalism, E_d is simply the height of the step,

$$E_d(n_f) = E_s - E_0(n_f) = E_v + E_a - E_0(0) - n_f E_g. \quad (24)$$

Thus we have derived a simple relationship for how electronic excitations affect E_d . However, before we show how this relationship affects the sputtering cross section, we must first point out an important caveat. Equation (24) suggests that E_d can be negative for large n_f . In these cases, the affected atom accelerates away from its pristine site even without energy transfer from the beam electron. This would seem to suggest that the sputtering cross section is infinite, since the PKA would sputter for any energy transfer, no matter how small. This is corroborated by the fact that the integral in equation (3) diverges as E_d approaches zero, meaning that the sputtering cross section σ would approach infinity. However, this is nonsensical, as it would imply an infinite sputtering rate for a finite beam current. In reality, even though the beam electron does nudge all of the material nuclei to some extent, one nucleus always receives a bigger nudge than the rest. This is the atom onto which the beam-induced excitations localize. Thus, for each beam electron, only one atom's displacement threshold is reduced by electronic excitations. For this reason, the displacement cross section should never exceed the cross sectional area occupied by a single target atom. This means that E_d , the lower integration bound in equation (3), must have a lower bound itself, which we call E_{\min} . When E_d falls below E_{\min} , we replace it with E_{\min} . We can approximate E_{\min} by setting the displacement cross section in equation (2) equal to the area occupied by the PKA and solving for the displacement threshold (section S8). With this, we are finally ready to use equation (24) to determine the set of displacement thresholds $\{E_d\}$ to insert into equation (3).

All positive displacement thresholds computed for this work are listed in table 1. Explanations for the calculations of E_v , E_a ,

n_f	displacement threshold (eV)			
	0	1	2	3
B from hBN	12.85	8.78	4.71	0.64
N from hBN	12.71	8.64	4.57	0.50
S from MoS ₂	6.92	5.04	3.16	1.28

Table 1 All computed positive displacement thresholds for sputtering from the hBN armchair edge and MoS₂ surface. The thresholds decrease as the number of surviving excitations n_f increases

$E_0(0)$, and E_g are given in subsections 3.1 and 3.2. The three largest displacement thresholds for MoS₂ with $n_f = 0, 1$, and 2 excitations are similar to those calculated with DFT-based molecular dynamics simulations.¹² This provides some assurance that our simplified approach to calculating E_d yields reasonable results. Excitation numbers n_f greater than those listed in table 1 make E_d negative, in which case the exact value of E_d is not important since it is replaced with E_{\min} in the calculation of σ . With that said, it is critical to consider large n_f , even if the resulting E_d is less than E_{\min} . These large n_f can make appreciable contributions to the total cross section, especially at small beam energies for which $P_i(\epsilon_b, n_i)$ is significant for large n_i (figure 3). As n_f cannot exceed n_i , this means that we must consider sufficiently large n_i to acknowledge these contributions. Eventually, these contributions diminish as we increment n_i , because the $n_i!$ in the denominator of P_i in formula (23) eventually outgrows the S^{n_i} in the numerator. We can therefore truncate the summation over n_i in equation (3) at some adequately large n_i^{\max} . This means that n_i^{\max} must be converged for each material, as materials with greater S require greater n_i^{\max} . In this work, we choose n_i^{\max} large enough so that including more excitations increases the sputtering cross section by less than 1% for the smallest experimental beam energy considered for each material (figure S1).

Thus, we have shown how E_d depends only on the number of surviving excitations n_f . We must now determine the likelihood that n_f excitations survive given n_i beam-induced excitations. This is done by comparing the excitation lifetime τ to t_s , the time it takes for the PKA to travel a distance d . We handle this task in the next subsection, where we write the sputtering cross section in equation (3) in terms of τ .

2.2.2 Sputtering cross section in terms of the excitation lifetime

In hBN and MoS₂, occupation of the antibonding state localized on the PKA does not affect the free energy of the sputtered system, since the antibonding and bonding states converge to the same degenerate p-orbital in the limit that the PKA is isolated from the host material. Therefore, only the excitation lifetime of the host material needs to be considered in calculating $P_f(n_i, n_f)$, the probability that n_f of the n_i beam-induced excitations survive long enough to reduce the displacement threshold.

We define the ratio of surviving excitations as

$$R(E, \tau) = e^{-t_s(E)/\tau}, \quad (25)$$

where t_s is the time it takes for the sputtered atom to travel a

distance d , and the excitation lifetime τ is determined for each material by fitting the cross section to experimental data. Given n_i beam-induced excitations, the probability that n_f excitations survive is

$$\begin{aligned} P_f(n_i, n_f) &= \binom{n_i}{n_f} R^{n_f} (1-R)^{n_i-n_f} \\ &= \binom{n_i}{n_f} \sum_{n=0}^{n_i-n_f} \binom{n_i-n_f}{n} (-1)^n R^{n_f+n}. \end{aligned} \quad (26)$$

Equation (26) allows us to rewrite the integral in equation (3) as

$$\int_{E_d}^{E_{\max}} P_f \frac{d\sigma}{dE} dE = \binom{n_i}{n_f} \sum_{n=0}^{n_i-n_f} \binom{n_i-n_f}{n} (-1)^n \int_{E_d}^{E_{\max}} R^{n_f+n} \frac{d\sigma}{dE} dE. \quad (27)$$

Assumption 3 from the previous subsection again aids us here. Because we posit that the PKA travels at a constant velocity until it reaches the sputtering distance d , we can write t_s from equation (25) quite simply as $t_s = d\sqrt{M/2E}$. This allows for the straightforward analytical integration of equation (27). Using $d\sigma/dE$ defined in equation (1), the integral on the right can be written as

$$\begin{aligned} \int R^{n_f+n} \frac{d\sigma}{dE} dE &= \pi \left(\frac{Z\alpha}{|\mathbf{p}|\beta} \right)^2 \left[\frac{4\tau^2 E_{\max}}{d^2 M} I_1(\xi, \mu) \right. \\ &\quad \left. + 2\beta(\pi Z\alpha + \beta) I_2(\xi, \mu) \right] \\ &\quad + \frac{\pi Z\alpha \beta \tau}{d} \sqrt{\frac{8E_{\max}}{M}} I_3(\xi, \mu) + \text{const.}, \end{aligned} \quad (28)$$

where we define

$$\xi \equiv \frac{d}{\tau} \sqrt{\frac{M}{2E}} \quad \text{and} \quad \mu \equiv n_f + n, \quad (29)$$

and the functions I_1 , I_2 , and I_3 are

$$I_1(\xi, \mu) = \begin{cases} -\frac{\xi^2}{2} & \text{if } \mu = 0 \\ \frac{\mu\xi+1}{\mu^2} e^{-\mu\xi} & \text{if } \mu > 0, \end{cases} \quad (30)$$

$$I_2(\xi, \mu) = \begin{cases} \ln(\xi) & \text{if } \mu = 0 \\ \text{Ei}(-\mu\xi) & \text{if } \mu > 0, \end{cases} \quad (31)$$

and

$$I_3(\xi, \mu) = \begin{cases} -\xi & \text{if } \mu = 0 \\ \frac{e^{-\mu\xi}}{\mu} & \text{if } \mu > 0. \end{cases} \quad (32)$$

Lastly, the function Ei in equation (31) is the exponential integral,

$$\text{Ei}(x) = \int_{-\infty}^x \frac{e^t}{t} dt. \quad (33)$$

Looking back at equation (3), we now have everything we need to evaluate the sputtering cross section. We derived $P_i(\epsilon_b, n_i)$ in formula (23), and we performed the integration over E analytically in equations (27) and (28), setting $\{E_d\}$ in equation (24) as

the lower integration bounds. We also have a criterion to truncate the summation over n_i at n_i^{\max} for a given material, as described at the end of section 2.2.1. In the following section, we use these results to calculate the sputtering cross sections of hBN and MoS₂.

3 Results and Discussion

We now apply our theoretical model to calculate the sputtering cross sections of boron and nitrogen from hBN and sulfur from MoS₂. In doing so, we show that the consideration of electronic excitations is necessary to make quantitative predictions about beam-induced sputtering rates in 2D insulators.

3.1 Boron and nitrogen sputtering from hexagonal BN

Electron beam irradiation has been shown to bore nanopores in monolayer hBN at beam energies far beneath the calculated ground state critical energy of $\varepsilon_c \sim 80$ keV.^{8,9,33,34} In a pristine hBN layer, these beam-induced pores can be initialized from isolated boron and nitrogen vacancies. The atoms surrounding these vacancies have reduced coordination numbers, and thus, smaller displacement thresholds than those in the pristine material. Therefore, the atoms lining the defect are more likely to sputter than their surface counterparts for a given beam energy. As the edge atoms continue to sputter away at a high rate, the nanopore grows, eventually extending up to a few nm in diameter.^{9,34,35}

Cretu *et al.* measured the radial growth of these nanopores under electron irradiation for several temperatures ranging from 673 to 1473 K.⁹ By dividing the radial growth rate by the beam current, they estimated the sputtering cross section to be around 25 barn under both 30 and 60 keV beams at temperatures of 1273 K and below (figure 5c). The cross sections were fairly temperature independent at these relatively low temperatures. Cretu *et al.* also found that the edges of these pores most often assume an armchair structure. Therefore, in an attempt to reproduce these measurements, we calculate the cross section of boron and nitrogen sputtering from an armchair edge at 1273 K (figure 5). To evaluate the displacement thresholds defined in equation (24) and listed in table 1, E_v is the free energy of the armchair edge supercell with a single boron or nitrogen vacancy, while E_a is that of an isolated boron or nitrogen atom. $E_0(0)$ and E_g are then the calculated free energy and band gap of the pristine armchair edge supercell. We plug the resulting set of displacement thresholds $\{E_d\}$ into equation (3) to calculate the sum of the boron and nitrogen sputtering cross sections. The excitation probabilities of hBN plotted in figure 3a are then used for P_i . Strictly speaking, the calculation of P_i should consider the effects of the edge state orbitals. However, in our formalism, the beam electron is in a momentum eigenstate that is highly delocalized in real space. We therefore assume that the radius of the beam is much larger than that of the nanopore. This means that the majority of the beam-matter interactions occur in regions of pristine material, validating the use of the P_i calculated for pristine hBN. Future work should consider the effects of localized beam electron states to simulate beams with smaller focal points. Lastly, temperature effects on the cross section are considered in the manner described

in our previous work.⁷

Fitting our cross section curves to the data of Cretu *et al.* yields an excellent agreement if the predicted excitation lifetime is set to $\tau \sim 240$ fs. This predicted lifetime is much shorter than the reported excitation lifetime of ~ 0.75 ns in pristine hBN,³⁶ indicating that the sputtering process can significantly reduce the excitation lifetimes of hBN. We suspect that the atomic motion gives rise to non-radiative relaxation pathways that are not explicitly accounted for here. This motivates a closer investigation into the full electronic evolution of the hBN system post-collision. However, such a study is beyond the scope of this work. With that said, the electronic structure in the vicinity of a sputtering PKA differs significantly from that of the pristine room-temperature systems in which excitation lifetimes are experimentally measured. There is therefore no reason to expect the predicted lifetimes in this work to match those obtained by experiments on systems not undergoing sputtering.

We also see that the sputtering cross section, and thus the growth rate of the nanopore, is minimized for beam energies between 30 and 60 keV. Below these energies, the sputtering cross section begins to grow as the beam energy decreases. This is due to a non-negligible probability of final excitation numbers $n_f > 3$, for which $E_d(n_f)$ falls below E_{\min} (table 1). In these cases, the sputtering cross section is $\Omega \sim 5 \times 10^8$ barn, seven orders of magnitude larger than the measured cross section at 30 keV. This suggests that one can expect the beam-induced nanopores in hBN to grow under beam energies as low as 1 keV. However, one must again be cautious of the predicted cross section at low beam energies in which the tree-level theory starts to break down. Nevertheless, figure 5c demonstrates a strong beam-energy dependence in the sputtering cross section across a wide range of experimentally relevant beam-energies. These findings could facilitate precise control of nanopore growth rates in hBN under electron irradiation.

3.2 Sulfur sputtering from MoS₂

Kretschmer *et al.* measured the sputtering cross section of sulfur from MoS₂ for several beam energies ranging from 20 to 80 keV.¹² They found a peak in the cross section at 30 keV, much less than the predicted ground state $\varepsilon_c \sim 90$ keV.¹² To help explain this unexpected peak, we calculate the cross section for sulfur sputtering from pristine MoS₂ (figure 6). The vacant system free energy E_v for the calculation of $\{E_d\}$ is that of a MoS₂ supercell with a single sulfur vacancy. E_a is then the free energy of an isolated sulfur atom, and $E_0(0)$ is that of the pristine MoS₂ supercell. We then set E_g equal to the experimental band gap of 1.88 eV.²⁸ Using the resulting $\{E_d\}$, we find that summing over the contributions of final excitation numbers $n_f > 2$ to the sputtering cross section produces a peak just below 30 keV, matching Kretschmer *et al.*'s findings remarkably well. In fitting to this peak, we predict an excitation lifetime of $\tau \sim 81$ fs. This is again much shorter than the excitation lifetime of pristine MoS₂, which is on the order of a few picoseconds.^{37–39} However, this is also much shorter than the fitted lifetime of hBN found in the previous subsection, consistent with the fact that the excitation lifetime of pristine hBN is

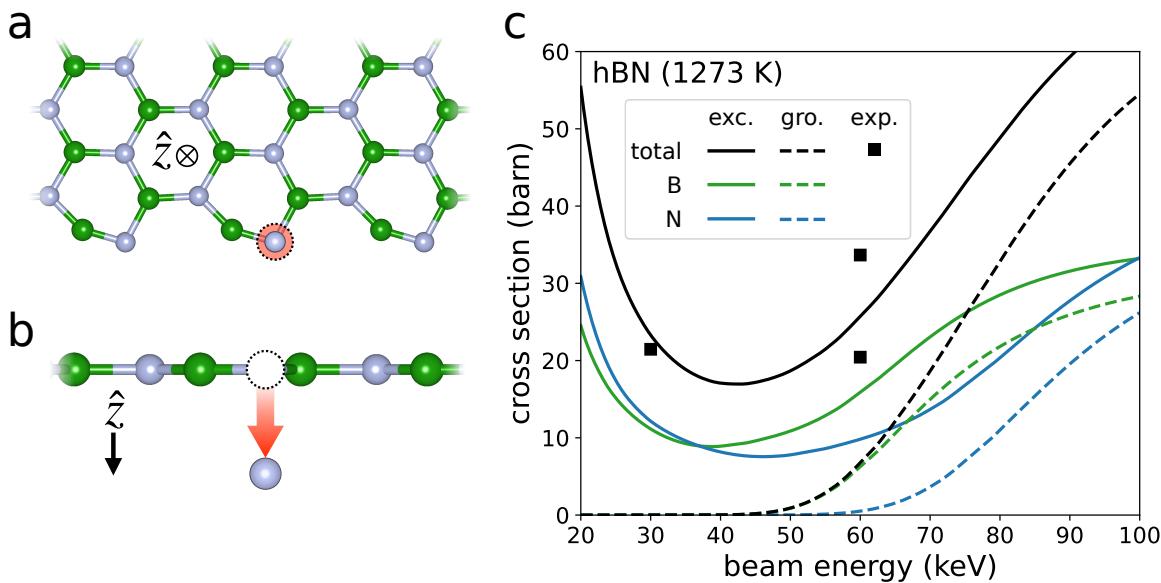


Fig. 5 Sputtering from the armchair edge of hBN. The schematics in panels (a) and (b) depict the sputtering of nitrogen from out-of-plane and in-plane perspectives respectively. Panel (c) plots the sputtering cross sections of both boron and nitrogen in green and blue respectively. The black curves are the total sputtering cross section, *i.e.*, the sum of their corresponding blue and green curves. The solid curves account for beam-induced excitations assuming an excitation lifetime of $\tau = 240$ fs (labelled “exc.” for excited). The dashed curves ignore the possibility of beam-induced excitation (labelled “gro.” for ground). The black squares are experimentally observed sputtering cross sections for hBN at 1273 K.⁹ The consideration of beam-induced excitations reduces the disagreement between theory and experiment substantially.

much longer than that of MoS₂.

The difference between the two materials’ lifetimes leads to markedly dissimilar cross section behaviors at low beam energies. Below beam energies of 30 keV, the cross section of MoS₂ gradually drops to zero with decreasing ε_b . In contrast, hBN’s total cross section has a minimum at around 40 keV and begins to increase as ε_b decreases. Eventually, hBN’s sputtering cross section peaks before dropping quickly to zero as the beam energy goes to zero (figure S2). However, this peak occurs at around 0.5 keV, far below the lower energy bound of figure 5c. These distinct cross section behaviors can be explained by the amplified sensitivity of n_f to τ at low beam energies. Equations (25) and (26) tell us that larger τ makes large n_f more likely. At the same time, n_f cannot exceed n_i . Thus, n_f is more sensitive to τ at low beam energies for which n_i is large. Accordingly, because τ is greater in hBN than in MoS₂, the expected values of n_f in hBN are much larger than those of MoS₂ at low beam energies. This explains why the effect of considering excitations is much more pronounced in hBN than in MoS₂. Furthermore, the difference in the cross section behavior is exacerbated by the fact that sulfur is heavier than both boron and nitrogen, so that its post-collision velocity is always smaller for a given energy transfer E . It follows that t_s is larger, and thus, P_f is smaller for a given E in MoS₂. This again increases the likelihood that hBN has more final excitations than MoS₂, meaning that the effects of beam-induced excitation on hBN’s sputtering cross section are greater.

We also plot the contributions of $n_f = 0, 1$, and 2 excitations to MoS₂’s sputtering cross section separately. In doing so, we see that the contributions of $n_f = 1$ or 2 would conceal the peak at 30 keV. Thus, it seems that the likelihoods of $n_f = 1$ or 2 are some-

how suppressed. This suggests that the individual beam-induced excitations and subsequent relaxation are in some cases correlated. That is, the excitation and/or relaxation rates of a given electronic transition are affected by the coinciding distribution of electronic excitations. Thus, a proper treatment of these excitations and their effect on the sputtering cross section requires that the excitation probabilities of every possible transition are calculated for every possible excitation configuration. Such a nonlinear calculation is beyond the scope of this work, but should certainly be pursued in a future study. Nonetheless, the methods laid out here demonstrate that the consideration of beam-induced excitation can provide a quantitative justification for the sulfur sputtering rate to peak at a beam energy well-below the expected ground state critical energy.

4 Conclusions

In this paper, we developed a first-principles method to more accurately describe electron beam-induced sputtering cross sections in 2D insulating crystals by accounting for beam-induced electronic excitations and their subsequent relaxations. The method combines QED scattering theory with DFT electronic structure calculations to determine the likelihood of beam-induced excitation. The results show that the excitation probability is inversely proportional to both the material’s band gap and beam electron’s kinetic energy. We then show how these nonzero excitation probabilities increase the predicted sputtering cross sections of both hBN and MoS₂. These cross sections can be made quantitatively similar to those obtained experimentally by treating the excitation lifetime as a fitting parameter. The methods laid out are computationally efficient, requiring only a few ground state electronic

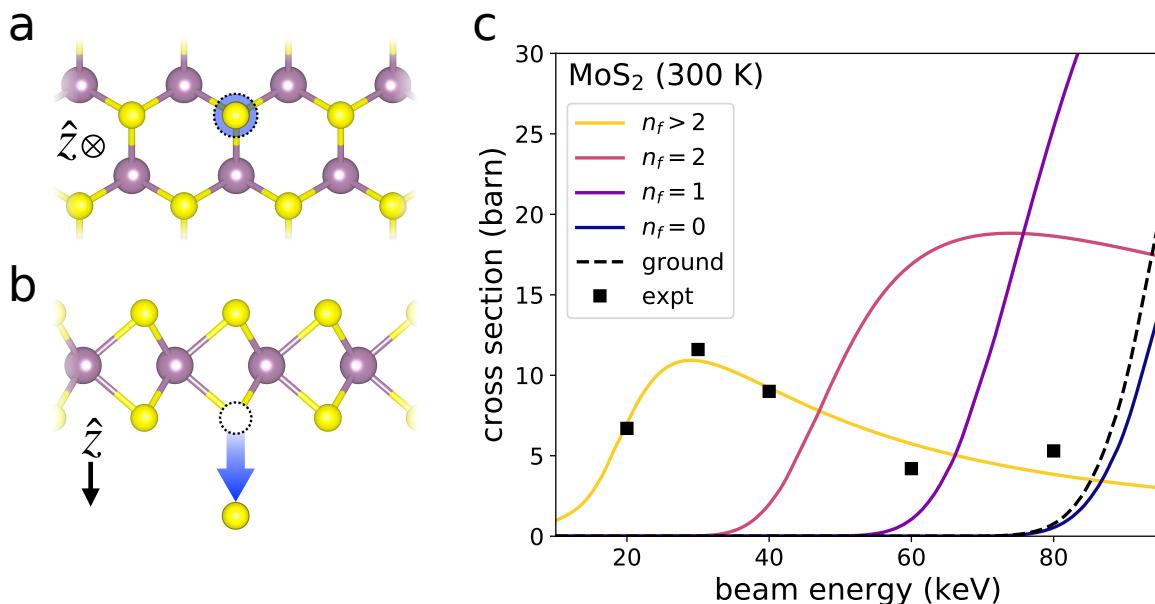


Fig. 6 Sputtering of sulfur from the outgoing MoS₂ surface. The sputtering of molybdenum from MoS₂ can be ignored because its displacement threshold is significantly larger than that of sulfur. The schematics in panels (a) and (b) depict the sputtering process from out-of-plane and in-plane perspectives respectively. Panel (c) plots the contributions of various numbers of final excitations n_f to the sputtering cross section assuming an excitation lifetime of $\tau = 81$ fs. The black dashed curve plots the predicted cross section ignoring the possibility of beam-induced excitation. The black squares are experimentally observed cross sections at 300 K.¹² The contribution from the sum of all $n_f > 2$ final excitations matches the experimental cross section remarkably well.

optimization calculations for each cross section curve. Thus, the formalism that we have developed can be easily applied to any 2D crystalline material to simulate the rates of atomic displacement under electron irradiation.

With that said, several questions naturally arise from our study. For example, why is the excitation lifetime reduced so drastically during sputtering? How might excitation and relaxation rates of different transitions be correlated? How might preexisting defects affect those rates? These questions urge follow up work to address the full breadth of physical processes involved in beam-matter interactions. Future studies should also consider additional electronic relaxation pathways to determine their effect on P_f . Correspondingly, other electronic responses such as ionization, core excitations, and second order electronic excitation effects such as Auger scattering can be incorporated into the calculation of P_i .^{38,40–42} Moreover, spin polarization effects can also be examined by making \mathcal{M} spin-dependent, *i.e.*, not averaging over spins as is done in equation (4). The beam electron path can also change significantly after collision with the nucleus. Ensuing research should investigate how these altered trajectories generate new excitation probabilities P_i , which must certainly play a role in 3D bulk materials as well as layered and confined 2D systems.^{43–46} Furthermore, the methods here can be modified to accommodate more advanced DFT techniques. For example, the calculation of P_i should be made compatible with ultrasoft and projector augmented wave pseudopotentials⁴⁷ in a manner similar to that employed in modern GW codes.^{48–50} In addition, the plane-wave coefficients of excited electronic states can be calculated self-consistently using constrained DFT.⁵¹ Perhaps most im-

portantly, progress in this field requires much more experimental data. We therefore hope this paper encourages new experimental investigation into beam-induced sputtering for beam energies below the predicted ground state critical energy.

Clearly, the combined QED-DFT approach to modeling beam-induced excitations and their effect on the sputtering cross section opens up a rich and diverse field of physics for both theoretical and experimental exploration. We hope that this work and the work it may stimulate can eventually enable the use of electron beams for precise atomic-scale engineering of crystalline materials.

5 Computational Details

DFT^{13,14} was used to calculate (i) the plane wave coefficients C_{G+k}^n and eigenvalues ε_{nk} of each material's electronic orbitals to determine the excitation probabilities (section 2.1.3) and (ii) the free energies of the pristine and sputtered structures to determine the displacement thresholds for sputtering (section 2.2.1).

Plane-wave coefficients and eigenvalues were calculated using Quantum ESPRESSO.⁵² Because equation (16) relies on the orthogonality of the Kohn-Sham orbitals, the optimized norm-conserving Vanderbilt pseudopotentials^{53,54} were employed. The sum of all transition probabilities S defined in equation (18) was used to gage the convergence of all parameters, which were deemed converged when any increase in precision changed S by less than 5% (figure S5). The cutoff energies for hBN and MoS₂ were set to 490 and 408 eV respectively. While these cutoffs are lower than what is typically used for norm-conserving pseudopotentials, both cutoffs yield well-converged values of S (figures S5b

and e). Meanwhile, the hBN and MoS₂ unit cells were given heights of 18 and 12 Å respectively. For both materials, the maximum virtual photon momentum $|\mathbf{q}_{\max}|$ required for convergence fell well within their first BZs. We therefore chose $|\mathbf{q}_{\max}|$ to be the magnitude of high-symmetry point K in each respective BZ. The maximum number of initial excitations considered for hBN and MoS₂ were $n_i^{\max} = 5$ and 9 respectively (figure S1). Finally, convergence of S requires extremely dense k-point sampling of the BZ. This necessitates fitting a curve to S calculated for various k-point mesh densities and extrapolating to an infinitely fine mesh to estimate the converged value of S (section S6). The most dense k-point meshes used to fit these curves had dimensions of $45 \times 45 \times 1$ and $36 \times 36 \times 1$ for hBN and MoS₂ respectively, corresponding to very similar linear k-point spacings of 0.056 and 0.055 Å⁻¹ respectively.

Next, free energy calculations were carried out with the Vienna ab initio Simulation Package (VASP)^{15,55} implementing the projector augmented wave (PAW) method⁴⁷ along with the Perdew-Burke-Ernserhof (PBE) generalized gradient approximation (GGA) to the exchange correlation functional.⁵⁶ Van der Waals interactions were accounted for using the optB88-vdW density functional methods.^{57,58} All parameters were converged so that any increase in precision would change the total free energy by less than 1 meV per atom. The parameter values that satisfy this criteria generally differed from those listed in the previous paragraph. The cutoff energies for hBN and MoS₂ were set to 800 and 550 eV respectively, while the BZs of both materials' pristine unit cells were sampled with Γ -centered $6 \times 6 \times 1$ Monkhorst-Pack meshes,⁵⁹ corresponding to linear k-point spacings of 0.417 and 0.328 Å⁻¹ for hBN and MoS₂ respectively. To achieve the same k-point densities, surface vacancies in MoS₂ and edge vacancies in hBN were placed in respective $6 \times 6 \times 1$ and $4 \times 1 \times 1$ supercells whose BZs were sampled with a single k-point on Γ . The heights of the hBN and MoS₂ cells were 12 Å and 20 Å respectively to provide sufficient separation from periodic images. Nanoribbon structures were used to simulate isolated armchair edges in hBN. These ribbons were more than 16 Å across and placed in cells 28 Å wide to avoid interactions between opposing edges and periodic images. Lastly, relaxation iterations of ionic positions and lattice constants persisted until the all Hellmann-Feynman forces settled below 1 meV/Å.

Author Contributions

AY and ML designed the code used to calculate the excitation probabilities and their effect on the sputtering cross section. AY and JG developed the theoretical model to describe beam-induced electronic excitation. AY, DL, JJ, PG, TY, BGS, and VM, developed the theoretical model to describe how excitations affect the sputtering cross section. AY carried out all calculations.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Calculations were performed at the Center for Computational Innovations at Rensselaer Polytechnic Institute, Livermore Comput-

ing Center at Lawrence Livermore National Laboratory (LLNL), and Compute and Data Environment for Science at Oak Ridge National Laboratory (ORNL), which is supported by the Office of Science of the U.S. Department of Energy (DOE) under Contract No. DE-AC05-00OR22725. This work was performed under the auspices of the U.S. Department of Energy by LLNL under Contract DE-AC52-07NA27344. This material is also based upon work supported by the U.S. DOE, Office of Science, Office of Workforce Development for Teachers and Scientists, Office of Science Graduate Student Research (SCGSR) program. The SCGSR program is administered by the Oak Ridge Institute for Science and Education for the DOE under contract number DE-SC0014664. Funding was provided by the National Science Foundation (Award 1608171). Work (BGS, PG, DL, JJ) was also supported by ORNL's Center for Nanophase Materials Sciences, a U.S. DOE Office of Science User Facility.

Notes and references

- 1 F. Banhart, *Rep. Prog. Phys.*, 1999, **62**, 1181–1221.
- 2 R. F. Egerton, *Ultramicroscopy*, 2013, **127**, 100–108.
- 3 X. Zhao, J. Kotakoski, J. C. Meyer, E. Sutter, P. Sutter, A. V. Krasheninnikov, U. Kaiser and W. Zhou, *MRS Bull.*, 2017, **42**, 667–676.
- 4 T. Susi, J. C. Meyer and J. Kotakoski, *Nat. Rev. Phys.*, 2019, **1**, 397–405.
- 5 J. C. Meyer, F. Eder, S. Kurasch, V. Skakalova, J. Kotakoski, H. J. Park, S. Roth, A. Chuvilin, S. Eyhusen, G. Benner, A. V. Krasheninnikov and U. Kaiser, *Phys. Rev. Lett.*, 2012, **108**, 196102–196107.
- 6 T. Susi, C. Hofer, G. Argentero, G. T. Leuthner, T. J. Pennycook, C. Mangler, J. C. Meyer and J. Kotakoski, *Nat. Commun.*, 2016, **7**, 13040.
- 7 A. Yoshimura, M. Lamparski, N. Kharche and V. Meunier, *Nanoscale*, 2018, **10**, 2388–2397.
- 8 J. Kotakoski, C. H. Jin, O. Lehtinen, K. Suenaga and A. V. Krasheninnikov, *Phys. Rev. B*, 2010, **82**, 113404.
- 9 O. Cretu, Y. C. Lin and K. Suenaga, *Micron*, 2015, **72**, 21–27.
- 10 Y.-C. Lin, T. Björkman, H.-P. Komsa, P.-Y. Teng, C.-H. Yeh, F.-S. Huang, K.-H. Lin, J. Jadczak, Y.-S. Huang, P.-W. Chiu, A. V. Krasheninnikov and K. Suenaga, *Nat. Commun.*, 2015, **6**, 6736–6741.
- 11 T. Lehnert, O. Lehtinen, G. A. Siller, U. Kaiser, G. Algara-Siller and U. Kaiser, *Appl. Phys. Lett.*, 2017, **110**, 033106.
- 12 S. Kretschmer, T. Lehnert, U. Kaiser, A. V. Krasheninnikov and A. V. Krasheninnikov, *Nano Lett.*, 2020, **20**, 2865–2870.
- 13 P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864–871.
- 14 W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133–1138.
- 15 G. Kresse and J. Furthmüller, *Phys. Rev. B*, 1996, **54**, 11169–11186.
- 16 H. Lourenço-Martins, A. Lubk and M. Kociak, *SciPost Phys.*, 2021, **10**, 31.
- 17 M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Perseus Books Publishing, L.L.C., Reading, MA, 1st edn, 1995.

- 18 T. Lancaster and S. J. Blundell, *Quantum Field Theory for the Gifted Amateur*, Oxford University Press, New York, NY, United States of America, 1st edn, 2014, p. 350.
- 19 N. F. Mott, *Proc. R. Soc. A*, 1929, **124**, 425–442.
- 20 W. A. McKinley and H. Feshbach, *Phys. Rev.*, 1948, **74**, 1759–1763.
- 21 O. S. Oen, *Cross sections for atomic displacements in solids by fast electrons*, Oak ridge national laboratory technical report, 1973.
- 22 R. F. Egerton, R. McLeod, F. Wang and M. Malac, *Ultramicroscopy*, 2010, **110**, 991–997.
- 23 C. Møller, *Ann. Phys.*, 1932, **406**, 531–585.
- 24 H. Kragh, *Arch. Hist. Exact Sci.*, 1992, **43**, 299–328.
- 25 X. Roqué, *Arch. Hist. Exact Sci.*, 1992, **44**, 197–264.
- 26 H. Bethe, *Ann. Phys.*, 1930, **397**, 325–400.
- 27 M. Inokuti, Y.-K. Kim and R. L. Platzman, *Phys. Rev.*, 1967, **164**, 55–61.
- 28 J. Gusakova, X. Wang, L. L. Shiau, A. Krivosheeva, V. Shaposhnikov, V. Borisenko, V. Gusakov and B. K. Tay, *Phys. status solidi*, 2017, **214**, 1700218.
- 29 C. Elias, P. Valvin, T. Pelini, A. Summerfield, C. J. Mellor, T. S. Cheng, L. Eaves, C. T. Foxon, P. H. Beton, S. V. Novikov, B. Gil and G. Cassabois, *Nat. Commun.*, 2019, **10**, 2639.
- 30 D. B. Lingerfelt, P. Ganesh, J. Jakowski and B. G. Sumpter, *Adv. Funct. Mater.*, 2019, **1901901**, 1901901.
- 31 X. Li, J. C. Tully, H. B. Schlegel and M. J. Frisch, *J. Chem. Phys.*, 2005, **123**, 84106.
- 32 H. P. Komsa, J. Kotakoski, S. Kurasch, O. Lehtinen, U. Kaiser and A. V. Krasheninnikov, *Phys. Rev. Lett.*, 2012, **109**, 035503–1–5.
- 33 C. Jin, F. Lin, K. Suenaga and S. Iijima, *Phys. Rev. Lett.*, 2009, **102**, 195505.
- 34 J. C. Meyer, A. Chuvilin, G. Algara-Siller, J. Biskupek and U. Kaiser, *Nano Lett.*, 2009, **9**, 2683–2689.
- 35 M. Dogan, S. M. Gilbert, T. Pham, B. Shevitski, P. Ercius, S. Aloni, A. Zettl and M. L. Cohen, *Appl. Phys. Lett.*, 2020, **117**, 9–13.
- 36 J. Li, X. K. Cao, T. B. Hoffman, J. H. Edgar, J. Y. Lin and H. X. Jiang, *Appl. Phys. Lett.*, 2016, **108**, 122101–1–4.
- 37 T. Korn, S. Heydrich, M. Hirmer, J. Schmutzler, C. Schüller and C. Schller, *Appl. Phys. Lett.*, 2011, **99**, 102109.
- 38 D. Lagarde, L. Bouet, X. Marie, C. R. Zhu, B. L. Liu, T. Amand, P. H. Tan and B. Urbaszek, *Phys. Rev. Lett.*, 2014, **112**, 47401.
- 39 M. Palummo, M. Bernardi and J. C. Grossman, *Nano Lett.*, 2015, **15**, 2794–2800.
- 40 D. Kozawa, R. Kumar, A. Carvalho, K. Kumar Amara, W. Zhao, S. Wang, M. Toh, R. M. Ribeiro, A. H. Castro Neto, K. Matsuda and G. Eda, *Nat. Commun.*, 2014, **5**, 4543.
- 41 Z. Nie, R. Long, J. S. Teguh, C.-C. Huang, D. W. Hewak, E. K. L. Yeow, Z. Shen, O. V. Prezhdo and Z.-H. Loh, *J. Phys. Chem. C*, 2015, **119**, 20698–20708.
- 42 H. Shi, R. Yan, S. Bertolazzi, J. Brivio, B. Gao, A. Kis, D. Jena, H. G. Xing and L. Huang, *ACS Nano*, 2013, **7**, 1072–1080.
- 43 W. Wang, Y. Zheng, X. Li, Y. Li, H. Zhao, L. Huang, Z. Yang, X. Zhang and G. Li, *Adv. Mater.*, 2019, **31**, e1803448.
- 44 A. Kakanakova-Georgieva, G. K. Gueorguiev, D. G. Sangiovanni, N. Suwannaharn, I. G. Ivanov, I. Cora, B. Pécz, G. Nicotra and F. Giannazzo, *Nanoscale*, 2020, **12**, 19470–19476.
- 45 A. Kakanakova-Georgieva, F. Giannazzo, G. Nicotra, I. Cora, G. K. Gueorguiev, P. O. Å. Persson and B. Pécz, *Appl. Surf. Sci.*, 2021, **548**, 149275.
- 46 Z. Wang, G. Wang, X. Liu, S. Wang, T. Wang, S. Zhang, J. Yu, G. Zhao and L. Zhang, *J. Mater. Chem. C*, 2021, **9**, 17201–17232.
- 47 P. E. Blöchl, *Phys. Rev. B*, 1994, **50**, 17953–17979.
- 48 M. Shishkin and G. Kresse, *Phys. Rev. B*, 2006, **74**, 035101.
- 49 M. Gajdoš, K. Hummer, G. Kresse, J. Furthmüller and F. Bechstedt, *Phys. Rev. B*, 2006, **73**, 045112.
- 50 J. Paier, R. Hirschl, M. Marsman and G. Kresse, *J. Chem. Phys.*, 2005, **122**, 234102.
- 51 B. Kaduk, T. Kowalczyk and T. Van Voorhis, *Chem. Rev.*, 2012, **112**, 321–370.
- 52 P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougaudis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari and R. M. Wentzcovitch, *J. Phys. Condens. Matter*, 2009, **21**, 395502.
- 53 D. R. Hamann, *Phys. Rev. B*, 2013, **88**, 85117.
- 54 M. Schlipf and F. Gygi, *Comput. Phys. Commun.*, 2015, **196**, 36–44.
- 55 G. Kresse and J. Furthmüller, *Comput. Mater. Sci.*, 1996, **6**, 15–50.
- 56 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.
- 57 J. Klimeš, D. Bowler and A. Michaelides, *J. Phys. Condens. Matter*, 2009, **22**, 022201–022205.
- 58 J. Klimeš, D. R. Bowler and A. Michaelides, *Phys. Rev. B*, 2011, **83**, 195131–1–13.
- 59 H. Monkhorst and J. Pack, *Phys. Rev. B*, 1976, **13**, 5188–5192.

Supporting Information

Quantum Theory of Electronic Excitation and Sputtering by Transmission Electron Microscopy

Anthony Yoshimura^{1,2*}, Michael Lamparski², Joel Giedt²,
David Lingerfelt³, Jacek Jakowski³, Panchapakesan Ganesh³, Tao Yu⁴,
Bobby Sumpter³, and Vincent Meunier^{2,5}

¹*Lawrence Livermore National Laboratory, Livermore, CA 94550, USA*

²*Department of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute,
Troy, New York 12180, USA*

³*Center for Nanophase Material Sciences, Oak Ridge National Laboratory, Oak Ridge, TN
37831, USA*

⁴*Department of Chemistry, University of North Dakota, Grand Forks, ND 58202, USA*

⁵*Department of Materials Science and Engineering, Rensselaer Polytechnic Institute, Troy,
NY 12180, USA*

**Correspondence to be addressed to yoshimura4@llnl.gov*

S1 Invariant matrix element \mathcal{M}

As the excitation amplitude in equation (16) contains the invariant matrix element \mathcal{M} and not $|\mathcal{M}|^2$, we are unable to use the spin sum identities typically used to derive cross sections in quantum electrodynamics.^{1,2} Nonetheless, the evaluation of \mathcal{M} in equation (4) is straightforward, though a bit cumbersome. In the Dirac basis,³

$$\gamma^0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \quad \text{and} \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \quad (\text{S1})$$

where

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \vec{\sigma} = \left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\}. \quad (\text{S2})$$

The electron spinors in equation (4) can be written as

$$u^1(p) = \sqrt{\epsilon + m} \begin{pmatrix} 1 \\ 0 \\ \frac{p^z}{\epsilon + m} \\ \frac{p^x + ip^y}{\epsilon + m} \end{pmatrix} \quad \text{and} \quad u^2(p) = \sqrt{\epsilon + m} \begin{pmatrix} 0 \\ 1 \\ \frac{p^x - ip^y}{\epsilon + m} \\ \frac{-p^z}{\epsilon + m} \end{pmatrix}, \quad (\text{S3})$$

while

$$\bar{u}^s(p) = u^{s\dagger}(p)\gamma^0. \quad (\text{S4})$$

As justified in section 2.1.1, we need only evaluate the t-channel contribution to \mathcal{M} and multiply the result by 2. That is,

$$\mathcal{M}(p_4 p_3 \leftarrow p_2 p_1) \sim e^2 \sum_{s_1, s_2, s_3, s_4} \bar{u}^{s_4}(p_4) \gamma^\mu u^{s_1}(p_1) \left(\frac{1}{p_3 - p_2} \right)^2 \bar{u}^{s_3}(p_3) \gamma_\mu u^{s_2}(p_2). \quad (\text{S5})$$

Substituting equations (S1) through (S4) into (S5) yields \mathcal{M} in terms of the components of the electrons' 4-momenta. That is,

$$\begin{aligned} \mathcal{M}(p_4 p_3 \leftarrow p_2 p_1) &\sim -\frac{2e^2}{(p_3 - p_2)^2} [(\epsilon_1 + m)(\epsilon_2 + m)(\epsilon_3 + m)(\epsilon_4 + m)]^{-1/2} \\ &\times \left\{ (\epsilon_1 + m)p_4^x [(\epsilon_2 + m)p_3^x + (\epsilon_3 + m)p_2^x] \right. \\ &+ 2(\epsilon_1 + m)(\epsilon_3 + m)(p_2^y + ip_2^z)(p_4^y - ip_4^z) \\ &+ 2(\epsilon_2 + m)(\epsilon_4 + m)ip_1^z(p_3^y - ip_3^z) \\ &- [(\epsilon_2 + m)(\epsilon_3 + m) + p_2^x p_3^x + (p_2^y + ip_2^z)(p_3^y - ip_3^z)] \\ &\left. \times [(\epsilon_1 + m)(\epsilon_4 + m) + ip_1^z(p_4^y - ip_4^z)] \right\}, \end{aligned} \quad (\text{S6})$$

where we let p_1 denote the momentum of the beam electron so that $p_1^x = p_1^y = 0$. See the Mathematica⁴ notebook in the Supporting Information for more details.

S2 Normalization

When integrating over 4-momentum space as is done in section 2.1.2, Lorentz invariance constrains a particle's 4-momentum to obey $p^2 = m^2$. It follows that the 4-momentum integration measure $d^4 p$ is always multiplied by a delta function $\delta(p^2 - m^2)$, *i.e.*,

$$d^4 p \delta(p^2 - m^2) \theta(p^0) = \frac{d^4 p}{2p^0} \delta(p_0 - \epsilon_{\mathbf{p}}), \quad (\text{S7})$$

where the Heaviside step function restricts our consideration to particles of positive mass (we note that antiparticles are interpreted as positive mass particles that propagate backwards in time). With this integration measure, the identity operator can be written as

$$\begin{aligned} \hat{I} &= \int \frac{d^4 p}{(2\pi)^4} (2\pi) \delta(p^2 - m^2) \theta(p^0) |p\rangle \langle p| \\ &= \int \frac{d^3 p}{(2\pi)^3 2\epsilon_{\mathbf{p}}} |p\rangle \langle p| \\ &= \int \frac{d^3 p}{(2\pi)^3} |\mathbf{p}\rangle \langle \mathbf{p}|. \end{aligned} \quad (\text{S8})$$

The last equality implies that

$$|p\rangle = (2\epsilon_{\mathbf{p}})^{1/2} |\mathbf{p}\rangle. \quad (\text{S9})$$

S3 Lab frame p_3^z

For the scattering of two free electrons, conservation of 4-momentum allows us to write

$$p_1 + p_2 = p_3 + p_4. \quad (\text{S10})$$

This constrains four of the six components needed to specify the 3-momenta of the two outgoing particles, so that only two components are independent. We choose the independent components to be p_3^x and p_3^y . Thus, we wish to find p_3^z as a function of p_1 , p_2 , p_3^x , and p_3^y . From this, $p_4 = p_1 + p_2 - p_3$ is easily obtained, and we have all six components of outgoing 3-momenta needed to calculate the scattering amplitude from equation (16). We start by setting the z -direction parallel to \mathbf{p}_1 . This means that

$$\begin{aligned} p_4^x &= p_2^x - p_3^x \\ p_4^y &= p_2^y - p_3^y \\ p_4^z &= p_1^z + p_2^z - p_3^z \\ \epsilon_4 &= \epsilon_1 + \Delta\epsilon \end{aligned} \quad (\text{S11})$$

where

$$\epsilon_i = \epsilon_{\mathbf{p}_i} = \sqrt{\mathbf{p}_i^2 + m^2}, \quad \Delta\epsilon = \epsilon_{n\mathbf{k}} - \epsilon_{n'\mathbf{k}'}, \quad (\text{S12})$$

and $\epsilon_{n\mathbf{k}}$ and $\epsilon_{n'\mathbf{k}'}$ are the eigenvalues of the excited hole and electron states, respectively. We can also write $\epsilon_1 = \gamma m$ where $\gamma = 1/\sqrt{1 - \beta^2}$ is the Lorentz factor corresponding to the beam electron's velocity β . Squaring the last line in (S11) and subtracting m^2 then gives

$$\mathbf{p}_4^2 = \mathbf{p}_1^2 + 2\gamma m\Delta\epsilon + \mathcal{O}(\Delta\epsilon^2). \quad (\text{S13})$$

We can ignore terms of order $\mathcal{O}(\Delta\epsilon^2)$ since $\Delta\epsilon \ll m$. Meanwhile, squaring and summing the first three equations in (S11) tells us that

$$\mathbf{p}_4^2 = \mathbf{p}_1^2 + \mathbf{p}_2^2 + \mathbf{p}_3^2 + 2 [p_1^z p_2^z - \mathbf{p}_2^\perp \cdot \mathbf{p}_3^\perp - (p_1^z + p_2^z)p_3^z], \quad (\text{S14})$$

where the \perp superscript denotes the projection perpendicular to \hat{z} . Subtracting equation (S14) from equation (S13) then yields

$$0 \sim \mathbf{p}_2^2 + \mathbf{p}_3^2 + 2 [p_1^z p_2^z - \mathbf{p}_2^\perp \cdot \mathbf{p}_3^\perp - (p_1^z + p_2^z)p_3^z] - 2\gamma m\Delta\epsilon. \quad (\text{S15})$$

Asymptotic formula (S15) can then be solved for p_3^z , so that

$$p_3^z \sim p_1^z + p_2^z \pm \sqrt{(p_1^z)^2 - (\mathbf{p}_2^\perp - \mathbf{p}_3^\perp)^2 + 2\gamma m\Delta\epsilon}. \quad (\text{S16})$$

We choose the $-$ from \pm as we impose that \mathbf{p}_3 is a component of the outgoing crystal electron state, whose z -momentum is much smaller than that of the beam electron.

S4 Converging n_i^{\max}

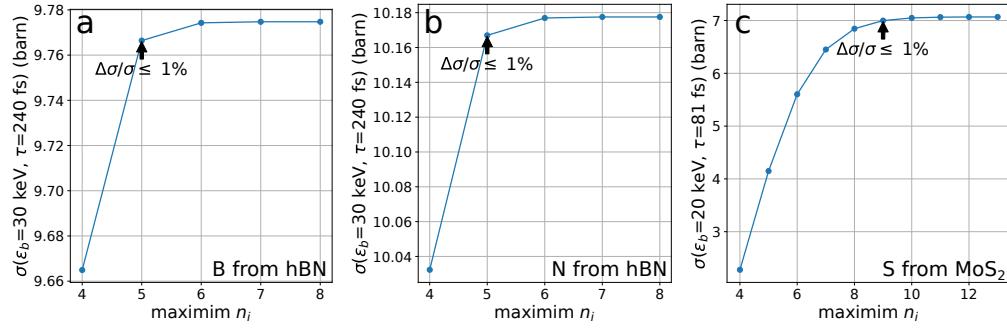


Figure S1: Convergence of the sputtering cross section with respect to the maximum number of beam-induced excitations n_i^{\max} considered. The simulated beam energies ϵ_b are the lowest experimental beam energies used for each material. The excitation lifetimes τ are those used to fit the experimental data in figures 5 and 6. Each cross section is deemed converged when any increase in n_i^{\max} increases the cross section by less than 1%.

S5 Peaks in the sputtering cross section of hBN

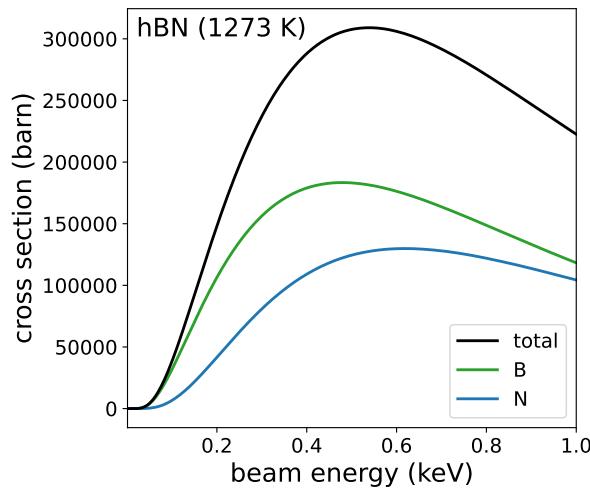


Figure S2: The sputtering cross sections of boron and nitrogen from the hBN armchair edge peak at beam energies much lower than those typically used for microscopy and defect engineering. However, as mentioned in the main text, the validity of our perturbative approximation to the scattering operator breaks down at low beam energies (section 2.1.4). Thus, the values and positions of these peaks may change if higher order perturbation terms are considered.

S6 Fitting and converging S

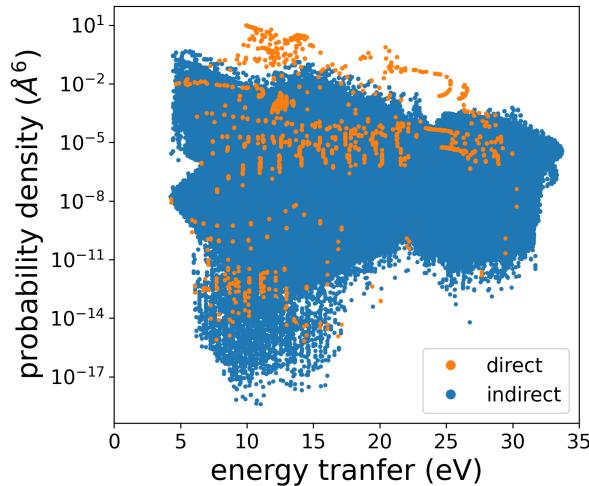


Figure S3: The most probable beam-induced electronic excitations are direct transitions. Each circle denotes the probability density centered on a particular excitation instigated by a 40 keV beam electron in hexagonal boron nitride, whose Brillouin zone was sampled by a $18 \times 18 \times 1$ Monkhorst-Pack mesh.

As shown in formula (23), the probability $P_i(\epsilon_b, n_i)$ that a beam electron of energy ϵ_b excites a specific number n_i of material electrons depends solely $S(\epsilon_b)$, the sum of all transition probabilities defined in equation (18). As such, convergence of S with respect to all simulation parameters is crucial to calculating the effect of electronic excitation on sputtering. In this section, we describe the general scheme that we use to converge S with respect to our simulation parameters. In particular, the convergence with respect to k-point density must be treated with care, and will be the primary focus of the remainder of this section.

Because the square of the momentum transfer appears in the denominator of the t -channel in equation (4), the probability density for electronic excitation peaks sharply when the momentum transfer is minimized. This minimization occurs when the beam electron induces a direct transition, in which the resulting electron and hole lie on the same k-point. Thus, direct excitations that conserve crystal momentum tend to be much more likely than indirect excitations (figure S3).

As a result, a sparse k-point mesh leads to an overestimation of S . This is because a coarser k-point mesh yields a higher proportion of direct transitions, as there are fewer indirect transitions available for each k-point. For a given number of k-points N_k , the number of direct transitions is proportional to N_k , while the number of indirect transitions is proportional to N_k^2 . Thus, increasing N_k increases the proportion of possible indirect transitions, lowering S . For a sufficiently dense mesh, the distinction between direct and nearly direct transitions is inconsequential, in which case, S is converged with respect to N_k .

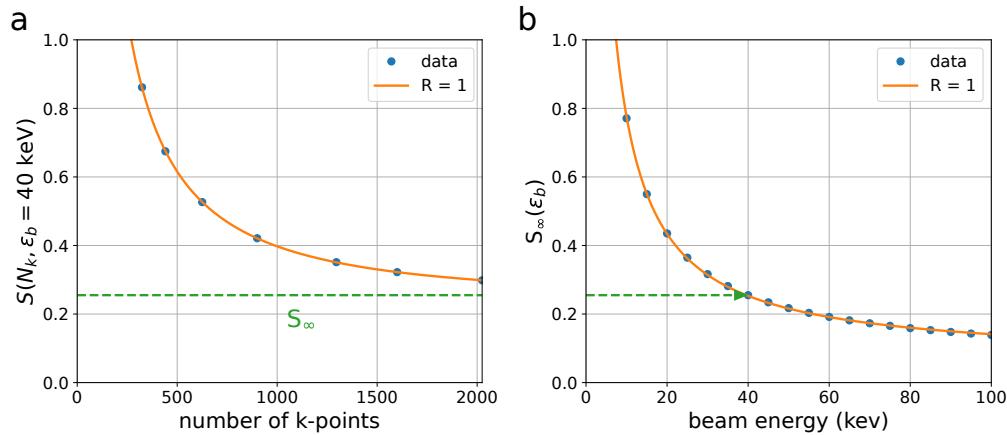


Figure S4: Finding the large crystal limit of $S(\epsilon_b)$ requires extrapolation. Panel (a) shows the dependence of S on the number N_k in the Brillouin zone of hBN under 40 keV irradiation. The simulated points are fitted well by equation (S17). The green dashed line denotes S_∞ , the asymptotic limit of $S(N_k)$ for large N_k . Panel (b) then shows the dependence of S_∞ on the beam energy ϵ_b , which is fitted well by equation (S18). The green arrow in panel (b) illustrates how $S_\infty(\epsilon_b = 40 \text{ keV})$ is determined by the fit in panel (a). The fitted curve in panel (b) is then used as $S(\epsilon_b)$ in formula (23) to calculate P_i for hBN. The same fitting and extrapolation method is used for MoS₂.

Satisfactory convergence of S requires an extremely dense k-point mesh. We therefore use extrapolation to determine the converged value of S . For a given beam energy ϵ_b , we calculate S for various N_k . We then fit the points to a curve of the form

$$S(N_k, \epsilon_b) = \frac{a}{N_k - b} e^{-cN_k} + S_\infty, \quad (\text{S17})$$

where a , b , c , and S_∞ are all fitting parameters that depend on ϵ_b (figure S4a). The last parameter S_∞ is the limit of S for large N_k , and is taken as the converged value of S for that beam energy. We repeat this for multiple values of ϵ_b ranging from 5 to 100 keV and record the best fit S_∞ for each ϵ_b . Based on the work of Bethe,^{5–7} we fit the resulting values of $S_\infty(\epsilon_b)$ to an inverse function,

$$S_\infty(\epsilon_b) = \frac{A}{\epsilon_b - B} + C. \quad (\text{S18})$$

where A , B , and C are fitting parameters. The parameters for hBN and MoS₂ are given in table S1. The fitted curve can then be substituted for $S(\epsilon_b)$ in formula (23) to obtain $P_i(\epsilon_b, n_i)$.

Finally, S is also converged with respect to the maximum virtual photon momentum, DFT cutoff energy, and height of the pristine unit cell (figure S5). In these cases, S is considered converged with respect to a parameter when any increase in the parameter's precision changes S by less than 5%.

material	A (keV)	B (keV)	C
hBN	7.655	-0.770	0.06526
MoS ₂	49.05	-3.867	0.3547

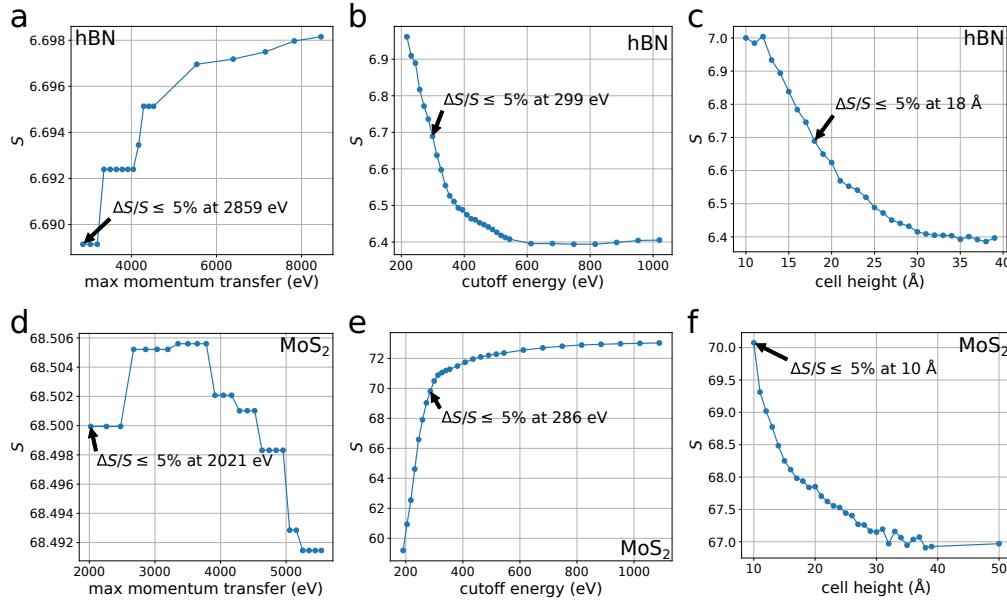
Table S1: Fitting parameters of equation (S18) for hBN and MoS₂.

Figure S5: Convergence of S with respect to the (a and d) maximum magnitude of virtual photon momentum considered, (b and e) plane-wave DFT cutoff energy, and (c and f) height of the unit cell. A $6 \times 6 \times 1$ k-point mesh and a beam energy of 60 keV is used to generate all six plots. The converged parameters were found to be insensitive to changes in k-point density and beam energy. S is deemed converged when any increase in precision changes S by less than 5%.

S7 Validity of the three simplifying assumptions

While the three assumptions laid out in section 2.2.1 may seem drastic, their effect on the sputtering cross section is small due to a cancellation of errors. To see this, consider the connection between the potential energy surfaces (ESs) and the PKA velocity. By assumption 1, the ESs at the PKA's initial position are equally spaced, separated by the material's band gap. By assumption 2, the ESs converge to the same energy as the PKA approaches a distance d . We shall ignore assumption 3 for now. If we consider that the ESs are not constant as the PKA moves towards d , so that the PKA slows down in transit, then the energy required to get to d depends on when the system relaxes to a lower ES. This is because higher ESs have shallower slopes than do the lower ones, so the PKA's deceleration will be smaller while on a higher ES. Thus, as the PKA moves, the ESs converge by assumption 2, and the energy lost during relaxation tends to shrink.

By introducing assumption 3, the energy lost during relaxation does not shrink but remains constant while the PKA sputters, and is therefore generally greater than it would be with gradually changing ESs. Our method therefore overestimates the energy lost during relaxation. On the other hand, because the PKA does not slow down until it reaches d , assumption 3 also underestimates the sputtering time t_s , introduced in equation (25). A smaller t_s means that fewer excitations relax during sputtering, increasing the cross section. However, each relaxation comes at a higher energy cost, which makes the PKA less likely to overcome the displacement threshold, decreasing the cross section. Thus, the underestimation t_s counteracts the overestimation of the energy lost during relaxation, so that final effect on the sputtering cross section is small.

Considering the ESs also sheds light on the computational cost saved by assumption 3. Without assumption 3, t_s would depend on the exact times that each relaxation occurs. In this case, calculating the cross section would require integration over time for each excited ES considered. We have found that the consideration of 5 and 9 excited ESs are necessary to converge the cross sections of hBN and MoS₂ respectively (figure S1). For an N-point Gaussian quadrature, this would multiply the computational costs by N^5 and N^9 respectively. Assumption 3 removes this complication by eliminating the need to consider exactly when each relaxation happens. Instead, only the total number of relaxations before the PKA reaches d need be considered.

S8 Calculating E_{\min}

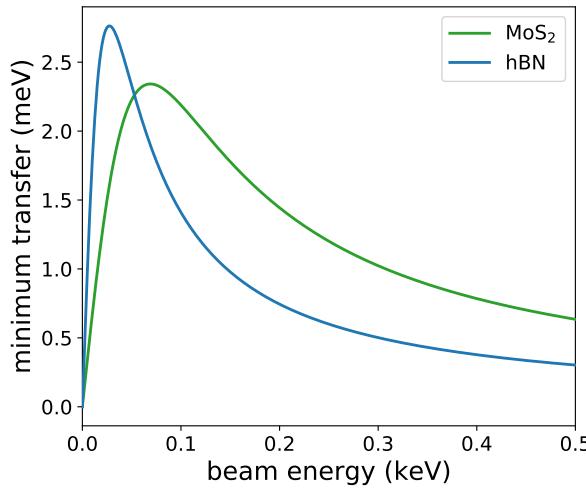


Figure S6: The minimum energy transfer from the beam electron to the target nucleus peaks at very low beam energies and is always much smaller than the nuclei's average thermal kinetic energy of ~ 39 meV at room temperature.

In hBN, a unit cell contains a boron and nitrogen atom. As these atoms have similar masses and displacement thresholds (table 1), the sputtering of both atoms should be considered for a given beam energy. Thus, we approximate the maximum cross sectional area σ_{\max} of these atoms to be $\Omega_{\text{hBN}}/2$, where Ω_{hBN} is the area of the hBN unit cell. On the other hand, molybdenum is much heavier than sulfur and has a much larger displacement threshold in MoS₂.⁸ This means that only the sputtering of sulfur needs to be considered. Additionally, of the two sulfur atoms in the MoS₂ unit cell, only the atom on the outgoing surface is eligible to sputter from a pristine system.⁸ Therefore, σ_{\max} for sulfur sputtering from MoS₂ is Ω_{MoS_2} , the area of the MoS₂ unit cell.

To approximate E_{\min} , we use the Rutherford displacement cross section^{9–11} as an approximation to equation (2),

$$\sigma_R = \pi \left(\frac{Z\alpha}{|\mathbf{p}| \beta} \right)^2 \left(\frac{E_{\max}}{E_d} - 1 \right). \quad (\text{S19})$$

Setting $\sigma_R = \sigma_{\max}$ and $E_d = E_{\min}$ and solving for E_{\min} yields

$$E_{\min}(\epsilon_b) = E_{\max} \left[\frac{\Omega}{\pi} \left(\frac{|\mathbf{p}_b| \beta}{Z\alpha} \right)^2 + 1 \right]^{-1}. \quad (\text{S20})$$

Using the Rutherford cross section instead of the McKinley-Feshbach cross section should be accurate for small beam energies for which $\beta \ll 1$. This makes it well-suited for finding E_{\min} , since $E_d(n_f) < E_{\min}$ only for large n_f , and large n_i (and thus n_f) are much more likely for small beam energies (figure 3).

With that said, the sputtering cross section is fairly insensitive to the exact value of E_{\min} when $E_{\min} \ll E_{\max}$. This is because the post-collision velocity of the PKA diminishes when the energy transfer E shrinks. The beam-induced excitations therefore have more time to relax before the PKA reaches the step in the energy surface (section 2.2.1). This makes P_f small for small E , so that contributions to the integral in equation (3) are extremely tiny for E near E_{\min} . As a result, changes in E_{\min} are essentially immeasurable for beam energies greater than 1 keV. Nonetheless, the use of E_{\min} is necessary for the calculation of a finite sputtering cross section.

References

- [1] Peskin, M. E.; Schroeder, D. V. *An Introduction to Quantum Field Theory*, 1st ed.; Perseus Books Publishing, L.L.C.: Reading, MA, 1995.
- [2] Lancaster, T.; Blundell, S. J. *Quantum Field Theory for the Gifted Amateur*, 1st ed.; Oxford University Press: New York, NY, United States of America, 2014; p 350.
- [3] Bjorken, J. D.; Drell, S. D. *Relativistic Quantum Mechanics*; International series in pure and applied physics; McGraw-Hill, 1964.
- [4] Inc., W. R. Mathematica, Version 12.3. <https://www.wolfram.com/mathematica>.
- [5] Bethe, H. Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie. *Ann. Phys.* **1930**, *397*, 325–400.
- [6] Susi, T.; Meyer, J. C.; Kotakoski, J. Quantifying Transmission Electron Microscopy Irradiation Effects Using Two-Dimensional Materials. *Nat. Rev. Phys.* **2019**, *1*, 397–405.
- [7] Kretschmer, S.; Lehnert, T.; Kaiser, U.; V. Krasheninnikov, A.; Krasheninnikov, A. V. Formation of Defects in Two-Dimensional MoS₂ in Transmission Electron Microscope at Electron Energies Below the Knock-On Threshold: The Role of Electronic Excitations. *Nano Lett.* **2020**, *20*, 2865–2870.
- [8] Komsa, H. P.; Kotakoski, J.; Kurasch, S.; Lehtinen, O.; Kaiser, U.; Krasheninnikov, A. V. Two-Dimensional Transition Metal Dichalcogenides Under Electron Irradiation: Defect Production and Doping. *Phys. Rev. Lett.* **2012**, *109*, 035503–1–5.
- [9] Thornton, S. T.; Marion, J. B. *Classical Dynamics of Particles and Systems*, 5th ed.; Brooks/Cole: Boston, MA, 2004.
- [10] Sakurai, J. J.; Napolitano, J. *Modern Quantum Mechanics*, 2nd ed.; Pearson Education: San Francisco, CA, 2011.
- [11] Yoshimura, A.; Lamparski, M.; Kharche, N.; Meunier, V. First-Principles Simulation of Local Response in Transition Metal Dichalcogenides under Electron Irradiation. *Nanoscale* **2018**, *10*, 2388–2397.

Dear Editor,

Please find our responses to the referees' comments (bold text) below.

Changes to the manuscript and SI are highlighted in blue text in hl_main.pdf and hl_sup.pdf.

On behalf of all the authors,

Anthony Yoshimura

Lawrence Livermore National Laboratory

Referee 1

The authors could enlarge the sampling and mention applicability to also other classes of 2D materials of relevance whereby their methodology may be very useful, not least some confined materials containing hexagonal group III nitrides (beyond h-BN) and oxides, namely, e.g., Nanoscale12 (2020) Pages 19470 - 19476; Applied Surface Science 548 (2021) Article number 149275. Such works are supportive to the transferability of the scheme reported in the present manuscript.

Thank you for this suggestion. We now mention the prospective application to layered and confined 2D systems in the conclusion, citing both articles you mention.

The authors discuss thermal kinetic energy at room temperature which is part of the scheme assessment. However, isn't that also important to explicitly compare these energies to the thermal stability of the studied 2D material(s)? Many materials would delaminate/disintegrate at similar energies...

This is a great point to keep in mind in general. However, to the best of our knowledge, the temperatures considered in our manuscript are those employed in the experiments we reference. Specifically, Cretu et al. [10.1016/j.micron.2015.02.002] irradiated their hBN at temperatures up to 1473 K. Kretchmer et al. [10.1021/acs.nanolett.0c00670] did not mention any specific temperatures in their study, so we assumed that their MoS₂ was irradiated at 300 K. Neither articles mention any delamination nor disintegration, so we do not believe such stability issues are a major concern in our study.

Section S6: Fitting and converging S. How specific is the convergence routine considered each particular 2D material? The conclusion that Sec. S6 leaves the reader with is that it is very specific and needs lots of tests and simulations. Or am I missing something? In any case it would be worth to revised the final sentences of this section.

Thank you for pointing this out. While it may not have been clear in Section S6, the method for converging S is material independent. The goal of the k-point fitting method is to determine the limit of S for an infinitely dense k-point mesh. This requires calculating S for sufficiently dense k-point meshes, i.e., S(Nk) plotted in figure S4a should appear to approach an asymptote for the densest k-point mesh

considered. We found that meshes of 45x45x1 and 36x36x1 were sufficiently dense hBN and MoS₂ respectively. These dense k-point meshes have very similar linear k-point spacings of 0.056 and 0.055 inverse angstrom for hBN and MoS₂ respectively. The hBN mesh contains more k-points because hBN has a larger Brillouin zone.

Admittedly, this method relies on “eyeballing” the $S(N_k)$ curve, relying on the user’s discretion as to whether the curve is approaching an asymptote. We are currently working to establishing a more formal convergence criteria for future work.

Convergence of S with respect to other parameters is also material independent, and a bit more straightforward. For a given parameter, we chose a value such that any increase in the parameter’s precision changes S by less than 5%.

We have added a paragraph at the beginning of section S6 that hopefully contextualizes the section to better orient the reader. We have also added two sentences to the end of the caption of figure S4 to clarify that the convergence methods for hBN and MoS₂ are identical.

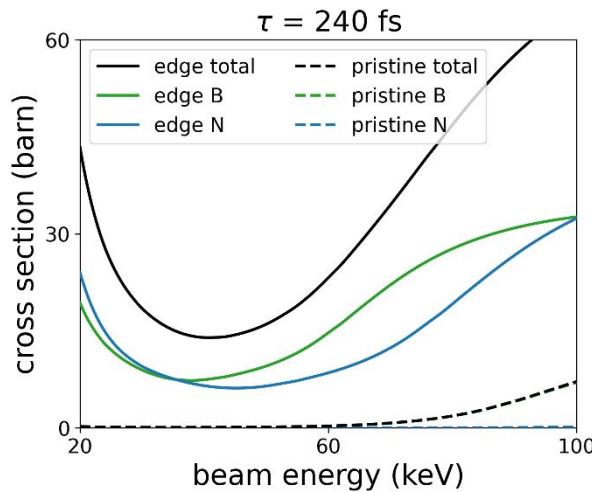
Spell-check and stylistic revision of the paper are still necessary. Some long sentences, misspellings, etc., still are noticeable throughout the text.

We have proofread the manuscript again and have done our best to fix any remaining errors. Please let us know if there are still any specific errors you would like us to change.

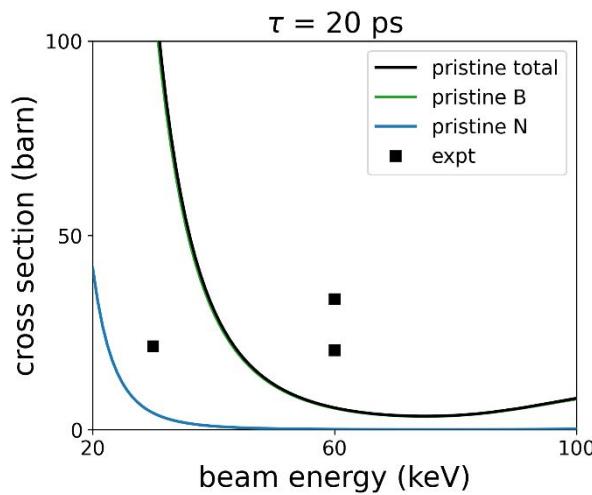
Referee: 2

Have you calculated the displacement thresholds of B and N in a pristine hBN, instead of on the edge of hBN? How does it compare to the experiment?

We have not calculated displacement thresholds for B and N sputtering from pristine hBN. However, Kotakoski et al. [10.1103/PhysRevB.82.113404] calculate thresholds of 19.36 eV and 23.06 eV for B and N respectively. As expected, these thresholds are several eV larger than those from the hBN edge. Assuming the excitation lifetime of 240 fs that we used to plot figure 5, these larger thresholds reduce the sputtering cross sections considerably for all beam energies. This is shown in the plot below, where the dashed “pristine” curves sit well below the solid “edge” curves.



There also does not exist an excitation lifetime that aligns the simulated pristine curve to the experimental edge sputtering data of Cretu et al. [10.1016/j.micron.2015.02.002]. Below is our best fit to the data with a lifetime of 20 ps.



Finally, experimental evaluation of the pristine displacement thresholds and corresponding cross sections would require a careful TEM study in which only single B and N vacancies (not multiatom pores) are counted for various TEM doses. To our knowledge, such a study does not exist.

Is your code open-source and accessible by the public?

We intend to provide a public GitHub repository soon after publication.

Journal Name

ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

Quantum theory of electronic excitation and sputtering by transmission electron microscopy[†]

Anthony Yoshimura,^{a,b} Michael Lamparski,^b Joel Giedt,^b David Lingerfelt,^c Jacek Jakowski,^c Panchapakesan Ganesh,^c Tao Yu,^d Bobby Sumpter,^c and Vincent Meunier^{b,e}

Received Date

Accepted Date

DOI: 00.0000/xxxxxxxxxx

Many computational models have been developed to predict the rates of atomic displacements in two-dimensional (2D) materials under electron beam irradiation. However, these models often drastically underestimate the displacement rates in 2D insulators, in which beam-induced electronic excitations can reduce the binding energies of the irradiated atoms. This bond softening leads to a qualitative disagreement between theory and experiment, in that substantial sputtering is experimentally observed at beam energies deemed far too small to drive atomic dislocation by many current models. To address these theoretical shortcomings, this paper develops a first-principles method to calculate the probability of beam-induced electronic excitations by coupling quantum electrodynamics (QED) scattering amplitudes to density functional theory (DFT) single-particle orbitals. The presented theory then explicitly considers the effect of these electronic excitations on the sputtering cross section. Applying this method to 2D hexagonal BN and MoS₂ significantly increases their calculated sputtering cross sections and correctly yields appreciable sputtering rates at beam energies previously predicted to leave the crystals intact. The proposed QED-DFT approach can be easily extended to describe a rich variety of beam-driven phenomena in any crystalline material.

1 Introduction

The holy grail of materials engineering is atomic scale control of the material structure. Towards this aim, electron irradiation by transmission electron microscopy (TEM) can be an effective means of structural manipulation with spatial control.^{1–4} These structural changes can arise *via* atomic displacement in which an incident electron nudges a material atom from its initial site. We call this atom the primary knock-on atom (PKA). Two-dimensional (2D) crystals provide an excellent platform to measure the rates of these PKA displacements. When electron ir-

radiation is normal to a 2D crystal's surface, a displacement likely propels the PKA away from the crystal. As such, these PKA are often ejected from the crystal in a process called *sputtering*. Sputtering events leave behind vacancies, which can then be counted using TEM. Counting the number of vacancies for a given *dose* and beam energy allows one to experimentally determine the sputtering cross section of that crystal.

Sputtering occurs when the energy transferred to the PKA is greater than the PKA's displacement threshold E_d . This means that a displacement is possible only if the kinetic energy of the beam electron exceeds some critical energy ϵ_c . Many computational models have been proposed to predict both E_d and ϵ_c to calculate electron beam-induced sputtering rates in 2D crystals.^{4–7} However, the vast majority of current methods focus solely on interactions between the beam electrons and material nuclei, neglecting any coupling with the material's electrons. Thus, while present-day models give reasonable predictions for conductors,⁵ where electronic relaxation is rapid, they often vastly underestimate the atomic displacement rates in insulators. For example, the critical energy for sputtering boron or nitrogen from hexagonal boron nitride (hBN) is predicted to be about 80 keV.⁸ However, sputtering has been observed in hBN under 30 keV irradiation.⁹ Furthermore, selenium sputters from WSe₂ and MoSe₂ under irradiation energies of 60 and 80 keV, respectively. These energies are almost 150 keV below their predicted critical ener-

^aLawrence Livermore National Laboratory, Livermore, CA 94550, USA E-mail: yoshimura4@llnl.gov

^bDepartment of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

^cCenter for Nanophase Material Sciences, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

^dCenter for Nanophase Material Sciences, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

^eDepartment of Materials Science and Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

[†]Electronic Supplementary Information (ESI) available: Derivation of the invariant matrix element \mathcal{M} introduced in equation (4), description of the normalization scheme used for the integral in equation (6), derivation of p_3^z that conserves energy and momentum, convergence of σ with respect to n_i^{\max} , peaks in the sputtering cross section of hBN for low beam energies, fitting and convergence of S with respect to various parameters, derivation of the minimum energy transfer E_{\min} mentioned after equation (24), and details on the derivation of \mathcal{M} . See DOI: 00.0000/00000000.

gies.^{10,11} Lastly, while the calculated critical energy for sulfur sputtering in MoS₂ is about 90 keV, sulfur has been shown to sputter under 20 keV beams.¹² Discrepancies like these suggest that the displacement thresholds in insulating crystals are much smaller than what is predicted by ground-state theory. Lehnert *et al.* have proposed that the consideration of inelastic scattering, *i.e.*, beam-induced electronic excitation, can lead to such a reduction in the displacement threshold.¹¹ This would increase the sputtering cross section for all beam energies and enable sputtering for energies well-below the ground state ε_c .

To account for these effects, this paper combines quantum electrodynamics (QED) and density functional theory (DFT) to derive the probability of beam-induced electronic excitation in 2D insulating crystals. The basic idea is as follows: DFT can provide effective single-particle states that can be decomposed into a plane-wave basis,^{13–15} while QED is well-equipped to describe how each plane-wave evolves in time through interactions with an electromagnetic field.^{16–18} Thus, a plane-wave decomposition of the Kohn-Sham orbitals can allow for a component-by-component treatment of the interactions between the beam and material electrons. This generalized QED-DFT approach enables, for the first time, a first-principles description of any beam-matter interaction process. The only limitations of this method are the order to which the time-evolution operator is expanded and the sophistication of the theory used to determine the material's electronic structure. Additionally, while DFT is used here, our method is compatible with any first-principles formalism that can produce single-particle eigenstates and eigenvalues for a given material.

This paper is organized as follows. First, we derive the theoretical model to determine beam-induced sputtering cross sections in 2D insulating crystals. We then use this model to calculate sputtering cross sections in hBN and MoS₂ that quantitatively agree with experiment.

2 Theoretical Model

The majority of present-day beam-damage models focus solely on the interaction between the beam electron and target nucleus. These models are thus centered on one process: energy transfer from the beam electron to the nucleus. For this, one defines a differential cross section $d\sigma/dE(\varepsilon_b, E)$ providing the distribution of energy transfers E for a given beam energy ε_b . The McKinley and Feshbach differential cross section can adequately describe $d\sigma/dE$ for light nuclei ($Z < 20$).^{19–22} Setting $\hbar = c = 1$,

$$\frac{d\sigma}{dE} = \pi \left(\frac{Z\alpha}{|\mathbf{p}_b|\beta} \right)^2 \left[\frac{E_{\max}}{E^2} - \beta(\pi Z\alpha + \beta) \frac{1}{E} + \pi Z\alpha\beta \sqrt{\frac{E_{\max}}{E^3}} \right], \quad (1)$$

where \mathbf{p}_b is the momentum of the beam electron, β is its velocity, Z is the atomic number of the target nucleus, and α is the fine structure constant.

One can then calculate the displacement cross section by integrating $d\sigma/dE$ over all E large enough to cause a displacement, so that

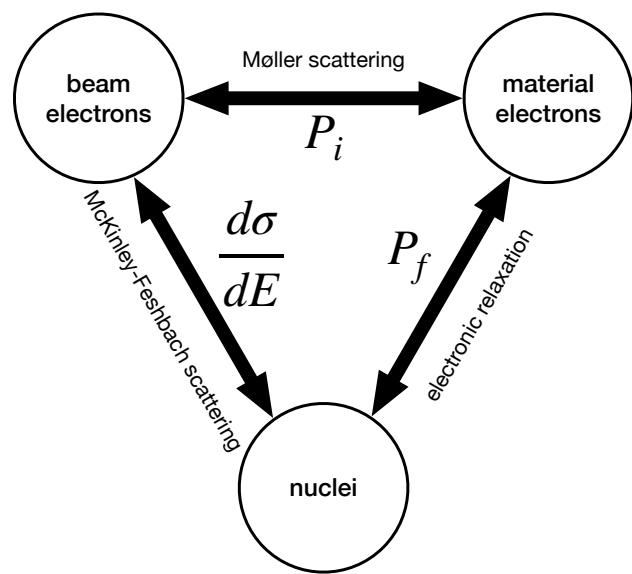


Fig. 1 Three interactions between three pairs of particles involved in electron beam-induced sputtering. In insulators, the probability of material electronic excitation (P_i) and the probability that those excitations are substantially long-lived (P_f) can be large. Therefore, the effect of exciting the material electrons on the sputtering rate should not be ignored in these materials.

$$\sigma_0(\varepsilon_b, E_d) = \theta(E_{\max}(\varepsilon_b) - E_d) \int_{E_d}^{E_{\max}(\varepsilon_b)} \frac{d\sigma}{dE}(\varepsilon_b, E) dE, \quad (2)$$

where $E_{\max}(\varepsilon_b)$ is the maximum possible energy transfer for a given beam energy, *i.e.*, the energy transfer resulting from a direct collision. The step function θ (which we will leave implicit going forward for the sake of compactness) enforces that the cross section is zero when $E_{\max} < E_d$. The critical energy ε_c is then the beam energy for which $E_{\max} = E_d$. Therefore, the observation of sputtering at energies well below ε_c is completely at odds with equation (2). With this in mind, a fair amount of work has been done to treat deviations from equation (2). Notably, several studies have explored the effects of temperature on displacement cross sections.^{5–7} This consideration involves calculating the degree to which the pre-collision thermal motion of the nucleus increases the cross section. However, these techniques essentially amount to smearing the beam energy dependence of the cross section, so that the cross section only strays significantly from equation (2) for beam energies very close to ε_c . Thus, temperature-induced increases in the cross section cannot account for the disparities between equation (2) and experiment. This necessitates the consideration of additional phenomena that can reduce E_d .

To address the limitations of equation (2), this work introduces a third party: the material's electrons. Doing so brings two new interactions into play: one between the beam and material electrons and another between material electrons and nuclei. This yields a total of three interactions between the three pairs of particles (figure 1). Therefore, the rate of beam-induced sputtering hinges on the rates of three processes mediated by these interac-

tions.

1. Beam and material electrons: a beam electron can excite some number n_i ground state electrons to the conduction band (i denotes the *initial* interaction with the beam). The probability of this event for a given beam energy ϵ_b is $P_i(\epsilon_b, n_i)$.
2. Material electrons and nuclei: Some number n_f beam-induced excitations can survive long enough for the target atom to leave its original site (f denotes the *final* system at the completion of sputtering). This depends on the nuclear kinetic energy E and the excitation lifetime τ . The probability that n_f of the n_i excitations survive is $P_f(E, \tau, n_i, n_f)$.
3. Beam electrons and material nuclei: the energy transferred to a material nucleus by the beam electron can exceed the PKA's displacement threshold $E_d(n_f)$, which depends on the number of surviving excitations n_f . We define $\{E_d\}$ as the set of all displacement thresholds for all possible n_f . Sputtering occurs when $E > E_d(n_f)$. The differential cross section for an energy transfer E from the beam electron to material nucleus is $d\sigma/dE(\epsilon_b, E)$.

The sputtering cross section can then be calculated by coupling $d\sigma/dE$ to P_i and P_f for all possible n_i and n_f . With the terms defined above, this excitation-sensitive sputtering cross section can be written as

$$\begin{aligned} \sigma(\epsilon_b, \{E_d\}, \tau) = & \sum_{n_i=0}^{\infty} P_i(\epsilon_b, n_i) \\ & \times \sum_{n_f=0}^{n_i} \int_{E_d(n_f)}^{E_{\max}(\epsilon_b)} P_f(E, \tau, n_i, n_f) \frac{d\sigma}{dE}(\epsilon_b, E) dE. \end{aligned} \quad (3)$$

If P_i and P_f are non-negligible when n_i and n_f are nonzero, and E_d depends strongly on n_f , then interactions with the material electrons must be considered. We will later show that this makes σ in equation (3) larger than σ_0 in equation (2) for all beam energies, most prominently when $\epsilon_b < \epsilon_c$.

The remainder of this section focuses on the derivations of P_i , P_f , and $\{E_d\}$. Section 2.1 describes how to combine QED with DFT to obtain P_i . Section 2.2 then considers the evolution of the excited states during the sputtering process to derive $\{E_d\}$ and P_f .

2.1 Probability of beam-induced excitation

For a crystal in its ground state, an occupied electron energy eigenstate has zero overlap with any unoccupied state. However, the collision of a beam electron can give an occupied state a momentum boost that breaks this orthogonality. Thus, the boosted ground state has a nonzero probability of being measured in an excited state. We can use this idea to derive $P_i(\epsilon_b, n_i)$, the probability that a beam electron with kinetic energy ϵ_b excites exactly n_i material electrons. The derivation can be broken down into four steps: (i) determine the amplitude for a free electron to scatter from one momentum eigenstate into another after collision with

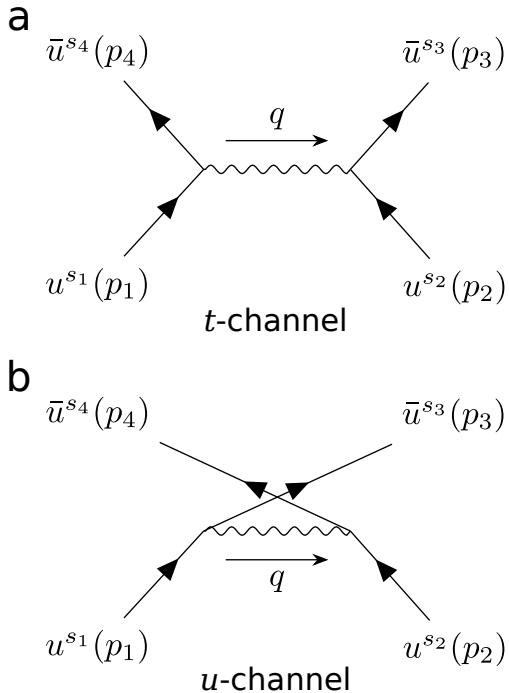


Fig. 2 The lowest order electron-electron scattering perturbation includes two Feynman diagrams called the (a) *t*-channel and (b) *u*-channel. The incoming and outgoing electron states are represented by Dirac spinors $u^s(p)$ and $\bar{u}^s(p)$ respectively, where p and s are the electron's 4-momentum and spin index respectively. Subscripts 1 and 2 label components of the initial beam and material states respectively. The virtual photon 4-momentum q is the momentum transfer between the electrons.

another free electron; (ii) generalize the formalism to obtain the amplitude for scattering from one wave packet into another by summing over the amplitudes for each momentum component of one wave packet to scatter into each momentum component of the other; (iii) decompose a pair of occupied and unoccupied crystal states into a momentum basis and plug them in as incoming and outgoing wave packets respectively, then square the amplitude to obtain the corresponding excitation probability for a particular transition; (iv) Compute the sum of all transition probabilities and use combinatorics to determine $P_i(\epsilon_b, n_i)$. The following subsections address each step (i-iv) in detail.

2.1.1 Scattering of free electrons

We first derive the scattering amplitude for momentum transfer between two free electrons via Møller scattering.^{23–25} Going forward, we label the 4-momenta of the incoming electrons as p_1 and p_2 , while the outgoing electrons have momenta p_3 and p_4 . We also choose to make p_1 and p_2 components of the initial beam and material states respectively. The 4-momentum of the n th electron can be written as $p_n = (\epsilon_n, p_n^x, p_n^y, p_n^z) = (\epsilon_n, \mathbf{p}_n)$, where ϵ_n is the particle's energy and \mathbf{p}_n is its 3-momentum. Dot products between 4-vectors are then taken over Minkowski space, so that $p_n \cdot p_m = g_{\mu\nu} p_n^\mu p_m^\nu = \epsilon_n \epsilon_m - \mathbf{p}_n \cdot \mathbf{p}_m$, where $\mathbf{p}_n \cdot \mathbf{p}_m = p_n^x p_m^x + p_n^y p_m^y + p_n^z p_m^z$.

To lowest order, the amplitude for free electron scattering can be represented by two tree-level diagrams, which we call the *t*- and *u*-channels (figure 2). Using Feynman's rules,^{17,18} we can

write these diagrams in terms of Dirac spinors, yielding the invariant matrix element

$$\mathcal{M}(p_4 p_3 \leftarrow p_2 p_1) = \frac{e^2}{2} \sum_{s_1} \sum_{s_2} \sum_{s_3} \sum_{s_4} \left[\bar{u}^{s_4}(p_4) \gamma^\mu u^{s_1}(p_1) \left(\frac{1}{p_3 - p_2} \right)^2 \bar{u}^{s_3}(p_3) \gamma_\mu u^{s_2}(p_2) + \bar{u}^{s_3}(p_3) \gamma^\mu u^{s_1}(p_1) \left(\frac{1}{p_4 - p_2} \right)^2 \bar{u}^{s_4}(p_4) \gamma_\mu u^{s_2}(p_2) \right], \quad (4)$$

where $s_n = 1$ or 2 denotes the spin of the n th electron, $u^s(p)$ is a Dirac spinor, and $\bar{u}^s(p)$ is its conjugate (section S1). The factor of $1/2$ before the summation arises from the assumption that the incoming states are spin unpolarized.

The first term in brackets is the t -channel describing momentum transfer $p_3 - p_2$ and the second is the u -channel describing momentum transfer $p_4 - p_2$. Because the DFT cutoff energy (hundreds of eV) is much smaller than the beam energy (tens of keV), it is always the case that $|\mathbf{p}_2|$ is much smaller than $|\mathbf{p}_1|$. Furthermore, we need only consider outgoing momenta for which the kinetic energy associated with either $|\mathbf{p}_3|$ or $|\mathbf{p}_4|$ falls within the DFT cutoff energy. In these cases, the magnitude of one outgoing momentum is similar to $|\mathbf{p}_2|$, while that of the other is much greater. This means that one channel's momentum transfer is always much larger than the other's. As the momentum transfers reside in the denominators of either channel in equation (4), it follows that one channel always contributes much more to \mathcal{M} than the other. Thus, when the t -channel is significant, the u -channel is negligible, and vice versa. Additionally, when integrating over all possible outgoing momenta, the contribution of the t -channel is equal to that of the u -channel. Taking advantage of this along with the indistinguishability of the electrons, we calculate only the t -channel and multiply the resulting amplitude by 2 instead of calculating both channels and adding them. We can then define the 4-momentum transfer between the electrons as that of the t -channel: $q \equiv p_3 - p_2$. Because the t -channel has a q^2 in the denominator, the resulting scattering probability is proportional to q^{-4} . This makes large momentum transfers statistically irrelevant, allowing us to only consider momentum transfers inside the first Brillouin zone (BZ). The evaluation of the t -channel in terms of the components of the electrons' 4-momenta is straightforward, though cumbersome, and is described in section S1.

We can use the resulting \mathcal{M} to obtain the free electron scattering amplitude

$$\langle p_4 p_3 | \hat{T} | p_2 p_1 \rangle = (2\pi)^4 \delta(p_1 + p_2 - p_3 - p_4) \mathcal{M}(p_4 p_3 \leftarrow p_2 p_1), \quad (5)$$

where \hat{T} is the scattering operator.^{17,18} This gives the amplitude for two free electrons with momenta p_1 and p_2 to scatter into p_3 and p_4 . Equation (5) is used to derive the scattering amplitude between two arbitrary wave packets in the next subsection.

2.1.2 Scattering of wave packets

The free particle scattering amplitude in equation (5) can be used to determine the amplitude for the scattering of two arbitrary electron states ϕ_1 and ϕ_2 into ϕ_3 and ϕ_4 . This is obtained by sandwiching the scattering operator between the initial and final 2-particle states, *i.e.*,

$$\langle \phi_4 \phi_3 | \hat{T} | \phi_2 \phi_1 \rangle = \int \frac{d^3 p_4 d^3 p_3 d^3 p_2 d^3 p_1}{(2\pi)^{12} 16 \epsilon_4 \epsilon_3 \epsilon_2 \epsilon_1} \times \langle \phi_4 \phi_3 | p_4 p_3 \rangle \langle p_4 p_3 | \hat{T} | p_2 p_1 \rangle \langle p_2 p_1 | \phi_2 \phi_1 \rangle. \quad (6)$$

On the right side, we have inserted two resolutions of the identity given in equation (S8). Inserting equation (5) into the integrand, the amplitude can be written in terms of the invariant matrix element \mathcal{M} , becoming

$$\int \frac{d^3 p_4 d^3 p_3 d^3 p_2 d^3 p_1}{(2\pi)^{12} 16 \epsilon_4 \epsilon_3 \epsilon_2 \epsilon_1} \langle \phi_4 | p_4 \rangle \langle \phi_3 | p_3 \rangle \langle p_2 | \phi_2 \rangle \langle p_1 | \phi_1 \rangle \times \mathcal{M}(p_4 p_3 \leftarrow p_2 p_1) (2\pi)^4 \delta(p_1 + p_2 - p_3 - p_4). \quad (7)$$

Using the delta function to integrate over \mathbf{p}_4 and \mathbf{p}_3^z yields

$$\int \frac{d^2 p_3^\perp d^3 p_2 d^3 p_1}{(2\pi)^8 16 \epsilon_2 \epsilon_1} \frac{\mathcal{M}(p_4 p_3 \leftarrow p_2 p_1)}{|p_3^z \epsilon_4 - p_4^z \epsilon_3|} \times \langle \phi_4 | p_4 \rangle \langle \phi_3 | p_3 \rangle \langle p_2 | \phi_2 \rangle \langle p_1 | \phi_1 \rangle, \quad (8)$$

where it is understood that p_4 and p_3 satisfy $p_1 + p_2 = p_3 + p_4$ (section S3). The normalization of 4-momentum states $|p_n\rangle$ in terms of 3-momentum states $|\mathbf{p}_n\rangle$ as defined in equation (S9) allows us to rewrite the expression as

$$\int \frac{d^2 p_3^\perp d^3 p_2 d^3 p_1}{(2\pi)^8} \sqrt{\frac{\epsilon_4 \epsilon_3}{\epsilon_2 \epsilon_1}} \frac{\mathcal{M}(p_4 p_3 \leftarrow p_2 p_1)}{4 |p_3^z \epsilon_4 - p_4^z \epsilon_3|} \times \langle \phi_4 | \mathbf{p}_4 \rangle \langle \phi_3 | \mathbf{p}_3 \rangle \langle \mathbf{p}_2 | \phi_2 \rangle \langle \mathbf{p}_1 | \phi_1 \rangle. \quad (9)$$

We can then discretize the momenta by replacing $d^3 p_i / (2\pi)^3$ with V^{-1} and $d^2 p_i^\perp / (2\pi)^2$ with A^{-1} , where V and A are the volume and cross sectional area of the simulated crystal respectively, *i.e.*, the volume and cross sectional area of the unit cell times the number of k-points used to sample the BZ. With this, the amplitude for electron states $|\phi_1\rangle$ and $|\phi_2\rangle$ to scatter into $|\phi_3\rangle$ and $|\phi_4\rangle$ takes the form

$$\frac{1}{AV^2} \sum_{\mathbf{p}_3} \sum_{\mathbf{p}_2} \sum_{\mathbf{p}_1} \sqrt{\frac{\epsilon_4 \epsilon_3}{\epsilon_2 \epsilon_1}} \frac{\mathcal{M}(p_4 p_3 \leftarrow p_2 p_1)}{4 |p_3^z \epsilon_4 - p_4^z \epsilon_3|} \times \langle \phi_4 | \mathbf{p}_4 \rangle \langle \phi_3 | \mathbf{p}_3 \rangle \langle \mathbf{p}_2 | \phi_2 \rangle \langle \mathbf{p}_1 | \phi_1 \rangle. \quad (10)$$

In the next subsection, we replace $|\phi_{1\dots 4}\rangle$ with states relevant to electron beam-induced excitation.

2.1.3 Probability of a crystal excitation

We now consider the specific case of beam-induced excitations to determine the form of the four electron states in equation (10). We assign $|\phi_1\rangle$ and $|\phi_4\rangle$ to the initial and final beam states $|\mathbf{p}_b\rangle$

and $|\mathbf{p}'_b\rangle$ respectively. States $|\phi_2\rangle$ and $|\phi_3\rangle$ are then the ground and excited crystal states $|n\mathbf{k}\rangle$ and $|n'\mathbf{k}'\rangle$ respectively, where n and n' are band indices and \mathbf{k} and \mathbf{k}' are k-points. Substituting these specific states into expression (10), the amplitude for exciting $|n\mathbf{k}\rangle$ to $|n'\mathbf{k}'\rangle$ becomes

$$\begin{aligned} \langle \mathbf{p}'_b, n'\mathbf{k}' | \hat{T} | n\mathbf{k}, \mathbf{p}_b \rangle &= \frac{1}{AV^2} \sum_{\mathbf{p}_3} \sum_{\mathbf{p}_2} \sum_{\mathbf{p}_1} \sqrt{\frac{\epsilon_4 \epsilon_3}{\epsilon_2 \epsilon_1}} \frac{\mathcal{M}(p_4 p_3 \leftarrow p_2 p_1)}{4 |p_3^\perp \epsilon_4 - p_4^\perp \epsilon_3|} \\ &\quad \times \langle \mathbf{p}'_b | \mathbf{p}_4 \rangle \langle n'\mathbf{k}' | \mathbf{p}_3 \rangle \langle \mathbf{p}_2 | n\mathbf{k} \rangle \langle \mathbf{p}_1 | \mathbf{p}_b \rangle. \end{aligned} \quad (11)$$

The values of $\epsilon_{1\dots 4}$ need to be clarified before moving forward. The zeroth components of the initial and final beam momenta obey the free particle dispersion relations, so $\epsilon_1 = \sqrt{|\mathbf{p}_1|^2 + m^2}$ and $\epsilon_4 = \sqrt{|\mathbf{p}_4|^2 + m^2}$. The beam energy that appears in equation (3) is then defined as the beam electron's total energy minus its rest mass: $\epsilon_b \equiv \epsilon_1 - m$. Meanwhile, the momentum of the crystal states can be treated nonrelativistically. Thus, the zeroth components of the crystal state momenta are the energy eigenvalues of the crystal state plus the electron rest mass, i.e., $\epsilon_2 = \epsilon_{n\mathbf{k}} + m$ and $\epsilon_3 = \epsilon_{n'\mathbf{k}'} + m$. For the remainder of this derivation, we continue to leave our expressions in terms of $\epsilon_{1\dots 4}$ for compactness.

Sputtering from a 2D crystal often requires that the beam electron is backscattered or nearly backscattered, in which case, its final trajectory after collision with the nucleus is nearly antiparallel to its initial trajectory and perpendicular to the crystal surface. Given that many 2D materials (including hBN and MoS₂) possess inversion and/or reflection symmetry about the crystal plane, we assume that the likelihood of excitation before and after the collision are about equal. In light of this, we calculate the excitation probability during a sputtering event assuming the beam electron's trajectory is not altered by its collision with the nucleus. That is, we impose that $\mathbf{p}_1 = |\mathbf{p}_1| \hat{\mathbf{z}}$ until an electronic excitation is induced.

We can now evaluate the bra-ket products in equation (11). The initial beam state is highly localized on \mathbf{p}_b , meaning that

$$|\langle \mathbf{p}_1 | \mathbf{p}_b \rangle|^2 = V \delta_{\mathbf{p}_1, \mathbf{p}_b} \Rightarrow \langle \mathbf{p}_1 | \mathbf{p}_b \rangle = \sqrt{V} \delta_{\mathbf{p}_1, \mathbf{p}_b}. \quad (12)$$

Meanwhile, the ground and excited crystal states can be expanded into a plane-wave basis, so that

$$\begin{aligned} |n\mathbf{k}\rangle &= \frac{1}{\sqrt{V}} \sum_{\mathbf{G}} C_{\mathbf{G}+\mathbf{k}}^n |\mathbf{G} + \mathbf{k}\rangle \\ |n'\mathbf{k}'\rangle &= \frac{1}{\sqrt{V}} \sum_{\mathbf{G}} C_{\mathbf{G}+\mathbf{k}'}^{n'} |\mathbf{G} + \mathbf{k}'\rangle, \end{aligned} \quad (13)$$

where each \mathbf{G} is a reciprocal lattice vector. By re-expressing \mathbf{p}_2 and \mathbf{p}_3 as $\mathbf{G}_2 + \mathbf{k}_2$ and $\mathbf{G}_3 + \mathbf{k}_3$ respectively, we find

$$\begin{aligned} \langle \mathbf{p}_2 | n\mathbf{k} \rangle &= \langle \mathbf{G}_2 + \mathbf{k}_2 | \frac{1}{\sqrt{V}} \sum_{\mathbf{G}} C_{\mathbf{G}+\mathbf{k}}^n |\mathbf{G} + \mathbf{k}\rangle = \sqrt{V} C_{\mathbf{G}_2 + \mathbf{k}_2}^n \delta_{\mathbf{k}_2, \mathbf{k}}, \\ \langle \mathbf{p}_3 | n'\mathbf{k}' \rangle &= \langle \mathbf{G}_3 + \mathbf{k}_3 | \frac{1}{\sqrt{V}} \sum_{\mathbf{G}} C_{\mathbf{G}+\mathbf{k}'}^{n'} |\mathbf{G} + \mathbf{k}'\rangle = \sqrt{V} C_{\mathbf{G}_3 + \mathbf{k}_3}^{n'} \delta_{\mathbf{k}_3, \mathbf{k}'}. \end{aligned} \quad (14)$$

Lastly, we do not care where the outgoing scattered electron ends up, so we wish for $|\mathbf{p}'_b\rangle$ to satisfy

$$|\langle \mathbf{p}_4 | \mathbf{p}'_b \rangle|^2 = V \Rightarrow \langle \mathbf{p}_4 | \mathbf{p}'_b \rangle = \sqrt{V}. \quad (15)$$

The excitation amplitude is then obtained by plugging in the bracket products from equations (12), (14), and (15) into equation (11). This gives us the excitation amplitude

$$\begin{aligned} \langle \mathbf{p}'_b, n'\mathbf{k}' | \hat{T} | n\mathbf{k}, \mathbf{p}_b \rangle &= \frac{1}{A} \sum_{\mathbf{G}_3^\perp} \sum_{\mathbf{G}_2} \sqrt{\frac{\epsilon_4 \epsilon_3}{\epsilon_1 \epsilon_2}} \\ &\quad \times \frac{\mathcal{M}(p_4 p_3 \leftarrow p_2 p_1)}{4 |p_3^\perp \epsilon_4 - p_4^\perp \epsilon_3|} C_{\mathbf{G}_3 + \mathbf{k}'}^{n'*} C_{\mathbf{G}_2 + \mathbf{k}}^n, \end{aligned} \quad (16)$$

where it is understood that $p_2 = (m + \epsilon_{n\mathbf{k}}, \mathbf{G}_2 + \mathbf{k})$ and $p_3 = (m + \epsilon_{n'\mathbf{k}'}, \mathbf{G}_3 + \mathbf{k}')$, and p_4 and p_3^\perp satisfy $p_1 + p_2 = p_4 + p_3$. Squaring this amplitude yields the probability of a single electronic excitation from the valence band state $|n\mathbf{k}\rangle$ to the conduction band state $|n'\mathbf{k}'\rangle$ for a given beam energy ϵ_b , that is,

$$P(n'\mathbf{k}' \leftarrow n\mathbf{k} | \epsilon_b) = |\langle \mathbf{p}'_b, n'\mathbf{k}' | \hat{T} | n\mathbf{k}, \mathbf{p}_b \rangle|^2. \quad (17)$$

2.1.4 Probability of n_i excitations

We are finally ready to derive $P_i(\epsilon_b, n_i)$, the probability that a beam electron excites a particular number of electrons n_i . First, we define the sum of all transition probabilities

$$S(\epsilon_b) \equiv \sum_{\mathbf{k}} \sum_{\mathbf{k}'} \sum_n \sum_{n'} P(n'\mathbf{k}' \leftarrow n\mathbf{k} | \epsilon_b) \equiv \sum_j^N P_j(\epsilon_b), \quad (18)$$

where \mathbf{k} and \mathbf{k}' run over all k-points, n over the valence bands, and n' over the conduction bands (those possessing states with energy between the Fermi level and the work function). The index j in the right-most expression labels the possible single-particle excitations (e.g., $j = n'\mathbf{k}' \leftarrow n\mathbf{k}$).

Combinatorics tells us that the probability of exciting exactly one excitation is

$$\begin{aligned} P_i(n_i = 1) &= \sum_{j_1} P_{j_1} \prod_{j_2 \neq j_1} (1 - P_{j_2}) \\ &= \sum_{j_1} P_{j_1} \left(1 - \sum_{j_2 \neq j_1} P_{j_2} + \frac{1}{2} \sum_{j_3 \neq j_1, j_2} \sum_{j_2 \neq j_1} P_{j_2} P_{j_3} - \dots \right). \end{aligned} \quad (19)$$

In the large-crystal limit, the number of states, and thus the number of transitions, is large so that the summations over j_i in equation (19) are approximately equal to one another. That is,

$$\sum_{j_2 \neq j_1} P_{j_2} \sim \sum_{j_2} P_{j_2} = S. \quad (20)$$

In this limit, the probability of exactly one beam-induced excitation can be written as

$$P_i(n_i = 1) \sim \sum_j P_j \left(1 - S + \frac{1}{2} S^2 - \dots \right) = e^{-S} \sum_j P_j = S e^{-S}. \quad (21)$$

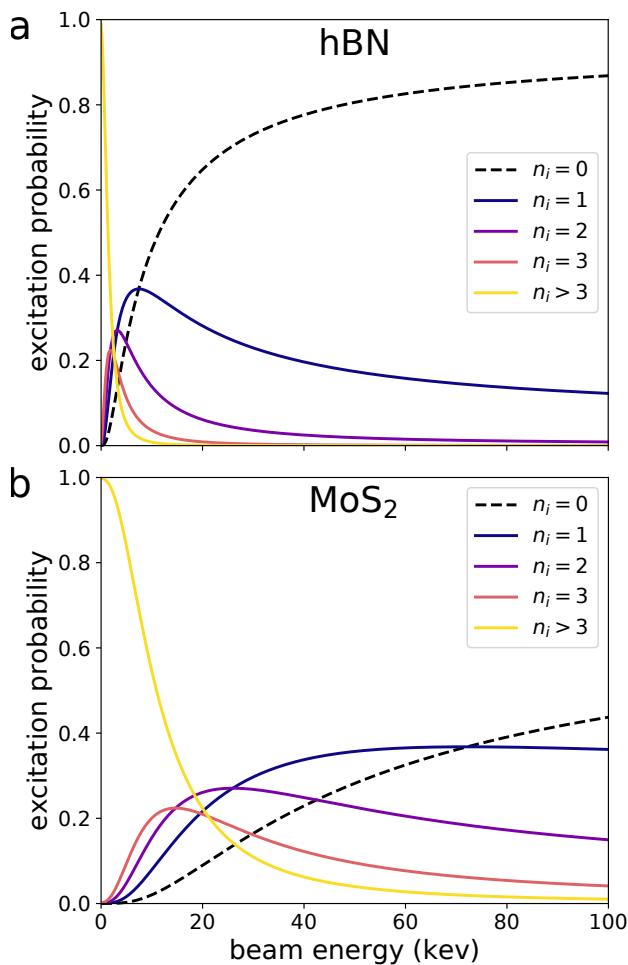


Fig. 3 Probabilities of initially exciting a certain number of electrons n_i in (a) hBN and (b) MoS₂ with a beam electron. The excitation probabilities tend to decrease with increasing beam energy. The excitation probabilities in MoS₂ are larger than those in hBN because MoS₂ has a significantly smaller band gap.

In the same way, the probability of two excitations is

$$P_i(n_i = 2) = \frac{1}{2} \sum_{j_1} \sum_{j_2 \neq j_1} P_{j_1} P_{j_2} \prod_{j_3 \neq j_1, j_2} (1 - P_{j_3}) \sim \frac{S^2}{2} e^{-S}. \quad (22)$$

In general, the probability of exactly n_i beam-induced excitations is approximately

$$P_i(\epsilon_b, n_i) \sim \frac{S(\epsilon_b)^{n_i}}{n_i!} e^{-S(\epsilon_b)}. \quad (23)$$

Thus, we see that the probability P_i can be written purely in terms of S .

We can now use formula (23) to calculate excitation probabilities in hBN and MoS₂ (figure 3). DFT is used to obtain the plane-wave coefficients $C_{\mathbf{G}_2 + \mathbf{k}}^n$ and $C_{\mathbf{G}_3 + \mathbf{k}'}^{n'}$ and eigenvalues $\epsilon_{n\mathbf{k}}$ and $\epsilon_{n'\mathbf{k}'}$ for the pristine unit cell of each material. These are plugged into equation (16) to obtain the amplitude for each transition. We sum over the squares of all resulting amplitudes to obtain S , which is then plugged into formula (23) to obtain P_i for both materials.

We emphasize that the only DFT calculation needed to determine P_i is the electronic structure relaxation of a pristine unit cell, a very inexpensive calculation.

The probabilities plotted in figure 3 reveal some notable trends. First, for sufficiently large beam energies, $P_i(\epsilon_b, n_i)$ decreases with increasing ϵ_b for all $n_i > 0$. This is because a faster beam electron has less time to interact with the material and cause an excitation. In this regime, $P_i(\epsilon_b)$ is proportional to ϵ_b^{-1} , a relationship originally predicted by Bethe.^{12,26,27} This means that multiple excitations are more likely at low beam energies. Furthermore, the probability of remaining in the ground state $P(\epsilon_b, n_i = 0)$ vanishes as ϵ_b goes to zero. This implies that a stationary electron in the vicinity of a material is guaranteed to interact with the material's electrons and affect its electronic structure. However, the validity of formula (23) diminishes as ϵ_b approaches zero. In the case of a slow beam electron, the interaction between the beam and material electrons can no longer be approximated by the single virtual photon transfer processes depicted in figure 2, as the amplitudes for higher-order processes become more significant. The effects of processes beyond the tree-level should be the subject of future work. Lastly, the excitation probability is inversely proportional to the material's band gap. This is because the zeroth component of the momentum transfer q depicted in figure 2 is the difference in energy eigenvalues between the occupied and unoccupied states (section 2.1.3). Thus, the smallest possible denominator of the t -channel in equation (4) is proportional to the difference in eigenvalues squared. The experimentally measured band gap of MoS₂ is about 1.9 eV²⁸ while that of hBN is about 6.1 eV.²⁹ This means MoS₂ hosts transitions with smaller eigenvalue differences, making the summands in equation (16) larger.

With $P_i(\epsilon_b, n_i)$ derived in formula (23) and $d\sigma/dE(\epsilon_b, E)$ defined in equation (1), we now have two of the three functions depicted in figure 1 needed to calculate the sputtering cross section in equation (3). The final ingredients are $P_f(E, \tau, n_i, n_f)$, the probability that n_f excitations survive the displacement event given n_i initial beam-induced excitations, and $\{E_d\}$, the set of all displacement thresholds for all n_f . The derivations of these objects are described in the next section.

2.2 Sputtering cross section

We have demonstrated how the interaction between the beam and material electrons can induce electronic excitations in the material. In this section, we show how these excitations bring about much larger sputtering cross sections than those predicted by a ground state theory. We start by showing how electronic excitations can reduce the displacement threshold E_d . We then show that longer excitation lifetimes increase P_f for nonzero n_f , giving the beam-induced excitations more opportunities to lessen E_d . This motivates us to write the sputtering cross section in equation (3) analytically in terms of the excitation lifetime τ . Finally, in the next section, the resulting equation is used to predict the sputtering rates of boron and nitrogen in hBN and sulfur in MoS₂, which can be made to agree well with experiment for appropriate values of τ for each material.

2.2.1 Effect of excitations on the displacement threshold

We begin by describing how beam-induced excitations can reduce the displacement threshold E_d in a process called *bond softening*. As E_d is the lower bound of integration over E in equation (3), and the differential cross section in equation (1) behaves like E^{-2} for small E , reductions in E_d can greatly increase the sputtering cross section. Exactly how excitations change E_d is an ambitious study on its own, requiring a careful consideration of the excited electrons' evolution and various relaxation pathways.³⁰ Here instead, we make three simplifying assumptions that allow us to calculate E_d with only ground state DFT (see section S7 for a comment on the validity of these assumptions).

Before describing these assumptions, we first define some important terms. Consider the moment immediately after the beam electron collides with a material nucleus. The nucleus now has a velocity corresponding to the kinetic energy E transferred from the beam electron. The resulting nuclear motion away from its equilibrium position causes the energy of the system to increase. In this sense, the system climbs an energy *surface* from the bottom of its equilibrium well. Far away from the well's bottom, the energy surface eventually plateaus. If the system reaches this plateau, the displaced PKA moves freely away from its initial site without deceleration. At this point, we consider the PKA to have sputtered. We call the energy at the well's bottom E_0 and that at the plateau E_s . If the energy surface is static throughout the entire process, then the displacement threshold is simply $E_d = E_s - E_0$. When $E > E_d$, the system has enough energy to climb out of the well, and the PKA sputters. Our task now is to determine how beam-induced excitations in the material electrons and their subsequent relaxation affect E_0 and E_s . To facilitate this, we make the following assumptions.

Assumption 1: the excited electrons and holes occupy the band edges. To justify this, we shift our focus to the material electrons immediately after the collision. There are now n_i electrons in the conduction band and n_i holes in the valence bands. Kretschmer *et al.* simulated the time-evolution of these excitations using Ehrenfest dynamics.^{12,31} They found that the excited electrons relax nonradiatively to the conduction band minimum (CBM) in a few femtoseconds, while the holes in the valence band take a similar amount of time to relax to the valence band maximum (VBM). In contrast, the PKA takes several hundreds of femtoseconds to fully sputter.⁷ Thus, the nonradiative relaxation of the electronic structure is essentially instantaneous, and we can assume that all excited electrons and holes occupy the CBM and VBM respectively before the PKA has been displaced by an appreciable amount.

These findings greatly simplify the calculation of E_0 . Given n_i excitations, E_0 is just the energy of the pristine system with n_i electrons and holes in the CBM and VBM respectively. In the large crystal limit, this amounts to adding the pristine band gap E_g to the system's ground state energy n_i times. In other words, $E_0(n_i) = E_0(0) + n_i E_g$. This approximation of course ignores any binding energy between the electron and hole. However, we believe that this formalism allows for an efficient treatment of the lowest order effects of excitation on E_0 .

Assumption 2: the fully sputtered system is in its ground state.

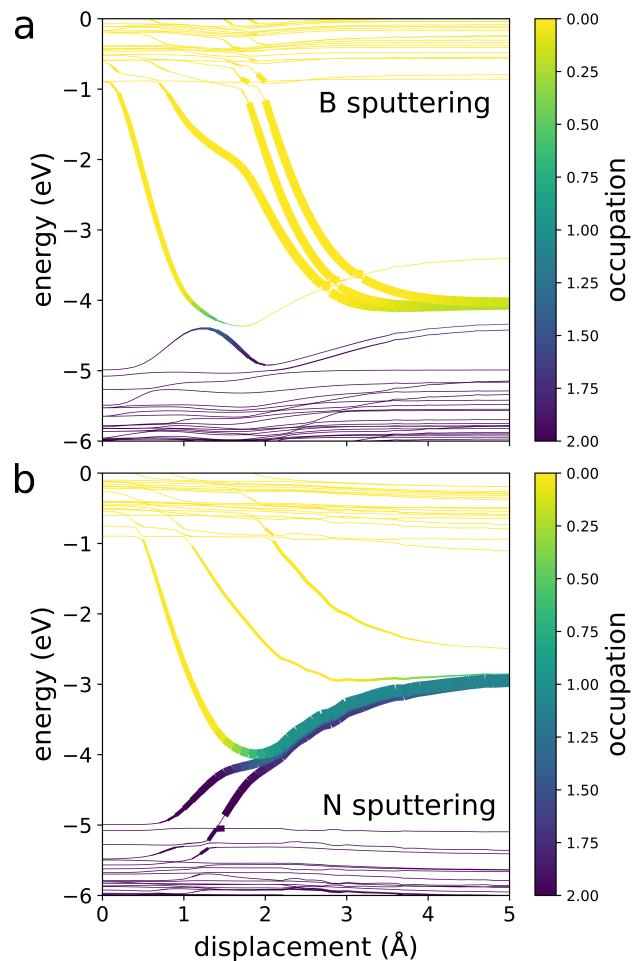


Fig. 4 Eigenvalues and occupations of the eigenstates of armchair hBN throughout the sputtering process with respect to the distance between sputtered (a) boron and (b) nitrogen and their pristine sites. The thickness of each curve is proportional to the localization of the corresponding state on the sputtered atom. In hBN, the electronegative nitrogen atoms borrow charge from the neighboring boron atoms. Thus, the p-orbital states that would be occupied on an isolated boron atom reside in the conduction band of hBN and descend into the gap as boron sputters. On the other hand, most of the p-orbital states of nitrogen lie in the valence band and rise into the gap during nitrogen sputtering.

Finding the plateau energy E_s can be a bit more involved. To calculate it properly, one must track how the excitations in the CBM and VBM evolve as the PKA moves away from the crystal. This would require Ehrenfest dynamics of a supercell over timescales of hundreds of femtoseconds for each excitation considered, which is prohibitively expensive. Here we seek a much less costly set of DFT calculations that can still provide a reasonable approximation for E_s . To this end, we draw upon another finding of Kretschmer *et al.*¹² After excitation, the displacement of the PKA causes the occupied and unoccupied CBM and VBM states to converge into the band gap and localize on the resulting defect. These converging states are the bonding and antibonding states that connect the PKA to the host crystal. Thus, we should consider how the beam-induced excitations affect the electronic structure on both the PKA and the remaining vacancy.

We start by considering how the ground state eigenvalues evolve as the PKA moves away from the crystal. For hBN, the evolution of eigenvalues differ depending on whether boron or nitrogen is sputtered (figure 4). Boron is electropositive while nitrogen is electronegative. Thus, in an hBN crystal, the nitrogen atoms borrow negative charge from the neighboring boron atoms. This means that the p-orbital states that would be occupied on an isolated boron atom are vacant, hovering in the conduction band of hBN. Conversely, the p-orbital states of nitrogen lie occupied in the valence band. For both atoms, separation from the crystal causes those states to converge to degenerate p-orbitals on their respective atoms. When boron is sputtered, these states come from the conduction band. This means that excited electrons residing in the CBM tend to transfer negative charge to the sputtered boron. However, this charge transfer is quite energetically unfavorable since boron is electropositive. The energy cost of this is much larger than the band gap of hBN, compelling any excess electrons on boron to relax to the host crystal. On the other hand, the states that localize on sputtered nitrogen must rise from the valence band. This means that the holes at the VBM tend to transfer positive charge to the sputtered nitrogen, which is again energetically unfavorable for the electronegative atom. Thus, because charge transfer to the sputtered PKA has such a large energy cost, we presume that in most cases, the PKA becomes charge neutral by the time sputtering has occurred.

Furthermore, because the states that localize on the PKA converge to the same degenerate p-orbital, the PKA assumes its ground state after it has sputtered, regardless of how many electrons the beam excites. We call the energy of the isolated ground state atom E_a . Meanwhile, the convergence of the bonding and antibonding states localized on the remaining vacancy during the displacement causes the excited electron and hole energy levels to cross multiples times. This means that there are plenty of chances for an excited electron to relax nonradiatively as the PKA separates from the material. Thus, we assume that the host crystal also relaxes to its ground state for most sputtering events.

Lastly, the sputtering threshold is always larger than the vacancy formation energy, because some of the energy transferred to the PKA disperses to the neighboring atoms. For this reason, we calculate the energy of the vacant system without relaxing the structure, as it has been shown that the free energy gained by leaving the structure unrelaxed roughly matches the energy dissipated to the surrounding material.³² We call the energy of the unrelaxed ground state vacancy E_v . Thus, we calculate the plateau energy using $E_s = E_v + E_a$, the sum of the ground state vacancy and isolated atom free energies.

Assumption 3: the PKA velocity is initially constant. Each number of excitations creates its own energy surface. In this way, electronic excitation and relaxation cause the system to hop from surface to surface. In our problem, the initial beam-induced excitation perches the system on an elevated surface. Once there, electronic relaxation via spontaneous emission (SE) enables downward surface hopping at the cost of the emitted photon energy. In this framework, E_d is the energy gained along all energy surfaces that the system traverses during sputtering. For example, suppose

the system emits one photon as it sputters. E_d is then the energy gained along the portion of the excited energy surface traversed before emission plus that gained along the relaxed surface after emission. This is how SE can change E_d . However, the effect of SE on E_d must diminish as the PKA moves further away from the crystal. This means there must be some distance d beyond which SE no longer affects the forces on the sputtered atom, leaving the energy surface unchanged. We take $d = 4.5 \text{ \AA}$ in line with Kretschmer *et al.*'s study, though we acknowledge that the exact meaning of this distance is different in their work.¹²

We wish to capture this idea while making the calculation of P_f as intuitive as possible. We do this by assuming that each energy surface is constant except for a step when the PKA reaches a distance d away from its equilibrium site. This means that immediately after the collision that excites n_i electrons, the surface assumes a constant energy $E_0(n_i)$. Each relaxation via SE causes the system to drop to the surface beneath it, decreasing its energy by E_g . We define n_f as the number of electronic excitations that survive the PKA's traversal of distance d to the energy step. Thus, the energy of the system just before the PKA reaches the step is $E_0(n_f)$. Beyond the step, all energy surfaces have energy E_s , and SE no longer changes the energy surface. With this formalism, E_d is simply the height of the step,

$$E_d(n_f) = E_s - E_0(n_f) = E_v + E_a - E_0(0) - n_f E_g. \quad (24)$$

Thus we have derived a simple relationship for how electronic excitations affect E_d . However, before we show how this relationship affects the sputtering cross section, we must first point out an important caveat. Equation (24) suggests that E_d can be negative for large n_f . In these cases, the affected atom accelerates away from its pristine site even without energy transfer from the beam electron. This would seem to suggest that the sputtering cross section is infinite, since the PKA would sputter for any energy transfer, no matter how small. This is corroborated by the fact that the integral in equation (3) diverges as E_d approaches zero, meaning that the sputtering cross section σ would approach infinity. However, this is nonsensical, as it would imply an infinite sputtering rate for a finite beam current. In reality, even though the beam electron does nudge all of the material nuclei to some extent, one nucleus always receives a bigger nudge than the rest. This is the atom onto which the beam-induced excitations localize. Thus, for each beam electron, only one atom's displacement threshold is reduced by electronic excitations. For this reason, the displacement cross section should never exceed the cross sectional area occupied by a single target atom. This means that E_d , the lower integration bound in equation (3), must have a lower bound itself, which we call E_{\min} . When E_d falls below E_{\min} , we replace it with E_{\min} . We can approximate E_{\min} by setting the displacement cross section in equation (2) equal to the area occupied by the PKA and solving for the displacement threshold (section S8). With this, we are finally ready to use equation (24) to determine the set of displacement thresholds $\{E_d\}$ to insert into equation (3).

All positive displacement thresholds computed for this work are listed in table 1. Explanations for the calculations of E_v , E_a ,

n_f	displacement threshold (eV)			
	0	1	2	3
B from hBN	12.85	8.78	4.71	0.64
N from hBN	12.71	8.64	4.57	0.50
S from MoS ₂	6.92	5.04	3.16	1.28

Table 1 All computed positive displacement thresholds for sputtering from the hBN armchair edge and MoS₂ surface. The thresholds decrease as the number of surviving excitations n_f increases

$E_0(0)$, and E_g are given in subsections 3.1 and 3.2. The three largest displacement thresholds for MoS₂ with $n_f = 0, 1$, and 2 excitations are similar to those calculated with DFT-based molecular dynamics simulations.¹² This provides some assurance that our simplified approach to calculating E_d yields reasonable results. Excitation numbers n_f greater than those listed in table 1 make E_d negative, in which case the exact value of E_d is not important since it is replaced with E_{\min} in the calculation of σ . With that said, it is critical to consider large n_f , even if the resulting E_d is less than E_{\min} . These large n_f can make appreciable contributions to the total cross section, especially at small beam energies for which $P_i(\epsilon_b, n_i)$ is significant for large n_i (figure 3). As n_f cannot exceed n_i , this means that we must consider sufficiently large n_i to acknowledge these contributions. Eventually, these contributions diminish as we increment n_i , because the $n_i!$ in the denominator of P_i in formula (23) eventually outgrows the S^{n_i} in the numerator. We can therefore truncate the summation over n_i in equation (3) at some adequately large n_i^{\max} . This means that n_i^{\max} must be converged for each material, as materials with greater S require greater n_i^{\max} . In this work, we choose n_i^{\max} large enough so that including more excitations increases the sputtering cross section by less than 1% for the smallest experimental beam energy considered for each material (figure S1).

Thus, we have shown how E_d depends only on the number of surviving excitations n_f . We must now determine the likelihood that n_f excitations survive given n_i beam-induced excitations. This is done by comparing the excitation lifetime τ to t_s , the time it takes for the PKA to travel a distance d . We handle this task in the next subsection, where we write the sputtering cross section in equation (3) in terms of τ .

2.2.2 Sputtering cross section in terms of the excitation lifetime

In hBN and MoS₂, occupation of the antibonding state localized on the PKA does not affect the free energy of the sputtered system, since the antibonding and bonding states converge to the same degenerate p-orbital in the limit that the PKA is isolated from the host material. Therefore, only the excitation lifetime of the host material needs to be considered in calculating $P_f(n_i, n_f)$, the probability that n_f of the n_i beam-induced excitations survive long enough to reduce the displacement threshold.

We define the ratio of surviving excitations as

$$R(E, \tau) = e^{-t_s(E)/\tau}, \quad (25)$$

where t_s is the time it takes for the sputtered atom to travel a

distance d , and the excitation lifetime τ is determined for each material by fitting the cross section to experimental data. Given n_i beam-induced excitations, the probability that n_f excitations survive is

$$\begin{aligned} P_f(n_i, n_f) &= \binom{n_i}{n_f} R^{n_f} (1-R)^{n_i-n_f} \\ &= \binom{n_i}{n_f} \sum_{n=0}^{n_i-n_f} \binom{n_i-n_f}{n} (-1)^n R^{n_f+n}. \end{aligned} \quad (26)$$

Equation (26) allows us to rewrite the integral in equation (3) as

$$\int_{E_d}^{E_{\max}} P_f \frac{d\sigma}{dE} dE = \binom{n_i}{n_f} \sum_{n=0}^{n_i-n_f} \binom{n_i-n_f}{n} (-1)^n \int_{E_d}^{E_{\max}} R^{n_f+n} \frac{d\sigma}{dE} dE. \quad (27)$$

Assumption 3 from the previous subsection again aids us here. Because we posit that the PKA travels at a constant velocity until it reaches the sputtering distance d , we can write t_s from equation (25) quite simply as $t_s = d\sqrt{M/2E}$. This allows for the straightforward analytical integration of equation (27). Using $d\sigma/dE$ defined in equation (1), the integral on the right can be written as

$$\begin{aligned} \int R^{n_f+n} \frac{d\sigma}{dE} dE &= \pi \left(\frac{Z\alpha}{|\mathbf{p}|\beta} \right)^2 \left[\frac{4\tau^2 E_{\max}}{d^2 M} I_1(\xi, \mu) \right. \\ &\quad \left. + 2\beta(\pi Z\alpha + \beta) I_2(\xi, \mu) \right] \\ &\quad + \frac{\pi Z\alpha \beta \tau}{d} \sqrt{\frac{8E_{\max}}{M}} I_3(\xi, \mu) + \text{const.}, \end{aligned} \quad (28)$$

where we define

$$\xi \equiv \frac{d}{\tau} \sqrt{\frac{M}{2E}} \quad \text{and} \quad \mu \equiv n_f + n, \quad (29)$$

and the functions I_1 , I_2 , and I_3 are

$$I_1(\xi, \mu) = \begin{cases} -\frac{\xi^2}{2} & \text{if } \mu = 0 \\ \frac{\mu\xi+1}{\mu^2} e^{-\mu\xi} & \text{if } \mu > 0, \end{cases} \quad (30)$$

$$I_2(\xi, \mu) = \begin{cases} \ln(\xi) & \text{if } \mu = 0 \\ \text{Ei}(-\mu\xi) & \text{if } \mu > 0, \end{cases} \quad (31)$$

and

$$I_3(\xi, \mu) = \begin{cases} -\xi & \text{if } \mu = 0 \\ \frac{e^{-\mu\xi}}{\mu} & \text{if } \mu > 0. \end{cases} \quad (32)$$

Lastly, the function Ei in equation (31) is the exponential integral,

$$\text{Ei}(x) = \int_{-\infty}^x \frac{e^t}{t} dt. \quad (33)$$

Looking back at equation (3), we now have everything we need to evaluate the sputtering cross section. We derived $P_i(\epsilon_b, n_i)$ in formula (23), and we performed the integration over E analyti-

cally in equations (27) and (28), setting $\{E_d\}$ in equation (24) as the lower integration bounds. We also have a criterion to truncate the summation over n_i at n_i^{\max} for a given material, as described at the end of section 2.2.1. In the following section, we use these results to calculate the sputtering cross sections of hBN and MoS₂.

3 Results and Discussion

We now apply our theoretical model to calculate the sputtering cross sections of boron and nitrogen from hBN and sulfur from MoS₂. In doing so, we show that the consideration of electronic excitations is necessary to make quantitative predictions about beam-induced sputtering rates in 2D insulators.

3.1 Boron and nitrogen sputtering from hexagonal BN

Electron beam irradiation has been shown to bore nanopores in monolayer hBN at beam energies far beneath the calculated ground state critical energy of $\epsilon_c \sim 80$ keV.^{8,9,33,34} In a pristine hBN layer, these beam-induced pores can be initialized from isolated boron and nitrogen vacancies. The atoms surrounding these vacancies have reduced coordination numbers, and thus, smaller displacement thresholds than those in the pristine material. Therefore, the atoms lining the defect are more likely to sputter than their surface counterparts for a given beam energy. As the edge atoms continue to sputter away at a high rate, the nanopore grows, eventually extending up to a few nm in diameter.^{9,34,35}

Cretu *et al.* measured the radial growth of these nanopores under electron irradiation for several temperatures ranging from 673 to 1473 K.⁹ By dividing the radial growth rate by the beam current, they estimated the sputtering cross section to be around 25 barn under both 30 and 60 keV beams at temperatures of 1273 K and below (figure 5c). The cross sections were fairly temperature independent at these relatively low temperatures. Cretu *et al.* also found that the edges of these pores most often assume an armchair structure. Therefore, in an attempt to reproduce these measurements, we calculate the cross section of boron and nitrogen sputtering from an armchair edge at 1273 K (figure 5). To evaluate the displacement thresholds defined in equation (24) and listed in table 1, E_v is the free energy of the armchair edge supercell with a single boron or nitrogen vacancy, while E_a is that of an isolated boron or nitrogen atom. $E_0(0)$ and E_g are then the calculated free energy and band gap of the pristine armchair edge supercell. We plug the resulting set of displacement thresholds $\{E_d\}$ into equation (3) to calculate the sum of the boron and nitrogen sputtering cross sections. The excitation probabilities of hBN plotted in figure 3a are then used for P_i . Strictly speaking, the calculation of P_i should consider the effects of the edge state orbitals. However, in our formalism, the beam electron is in a momentum eigenstate that is highly delocalized in real space. We therefore assume that the radius of the beam is much larger than that of the nanopore. This means that the majority of the beam-matter interactions occur in regions of pristine material, validating the use of the P_i calculated for pristine hBN. Future work should consider the effects of localized beam electron states to simulate beams with smaller focal points. Lastly, temperature ef-

fects on the cross section are considered in the manner described in our previous work.⁷

Fitting our cross section curves to the data of Cretu *et al.* yields an excellent agreement if the predicted excitation lifetime is set to $\tau \sim 240$ fs. This predicted lifetime is much shorter than the reported excitation lifetime of ~ 0.75 ns in pristine hBN,³⁶ indicating that the sputtering process can significantly reduce the excitation lifetimes of hBN. We suspect that the atomic motion gives rise to non-radiative relaxation pathways that are not explicitly accounted for here. This motivates a closer investigation into the full electronic evolution of the hBN system post-collision. However, such a study is beyond the scope of this work. With that said, the electronic structure in the vicinity of a sputtering PKA differs significantly from that of the pristine room-temperature systems in which excitation lifetimes are experimentally measured. There is therefore no reason to expect the predicted lifetimes in this work to match those obtained by experiments on systems not undergoing sputtering.

We also see that the sputtering cross section, and thus the growth rate of the nanopore, is minimized for beam energies between 30 and 60 keV. Below these energies, the sputtering cross section begins to grow as the beam energy decreases. This is due to a non-negligible probability of final excitation numbers $n_f > 3$, for which $E_d(n_f)$ falls below E_{\min} (table 1). In these cases, the sputtering cross section is $\Omega \sim 5 \times 10^8$ barn, seven orders of magnitude larger than the measured cross section at 30 keV. This suggests that one can expect the beam-induced nanopores in hBN to grow under beam energies as low as 1 keV. However, one must again be cautious of the predicted cross section at low beam energies in which the tree-level theory starts to break down. Nevertheless, figure 5c demonstrates a strong beam-energy dependence in the sputtering cross section across a wide range of experimentally relevant beam-energies. These findings could facilitate precise control of nanopore growth rates in hBN under electron irradiation.

3.2 Sulfur sputtering from MoS₂

Kretschmer *et al.* measured the sputtering cross section of sulfur from MoS₂ for several beam energies ranging from 20 to 80 keV.¹² They found a peak in the cross section at 30 keV, much less than the predicted ground state $\epsilon_c \sim 90$ keV.¹² To help explain this unexpected peak, we calculate the cross section for sulfur sputtering from pristine MoS₂ (figure 6). The vacant system free energy E_v for the calculation of $\{E_d\}$ is that of a MoS₂ supercell with a single sulfur vacancy. E_a is then the free energy of an isolated sulfur atom, and $E_0(0)$ is that of the pristine MoS₂ supercell. We then set E_g equal to the experimental band gap of 1.88 eV.²⁸ Using the resulting $\{E_d\}$, we find that summing over the contributions of final excitation numbers $n_f > 2$ to the sputtering cross section produces a peak just below 30 keV, matching Kretschmer *et al.*'s findings remarkably well. In fitting to this peak, we predict an excitation lifetime of $\tau \sim 81$ fs. This is again much shorter than the excitation lifetime of pristine MoS₂, which is on the order of a few picoseconds.^{37–39} However, this is also much shorter than the fitted lifetime of hBN found in the previous subsection, con-

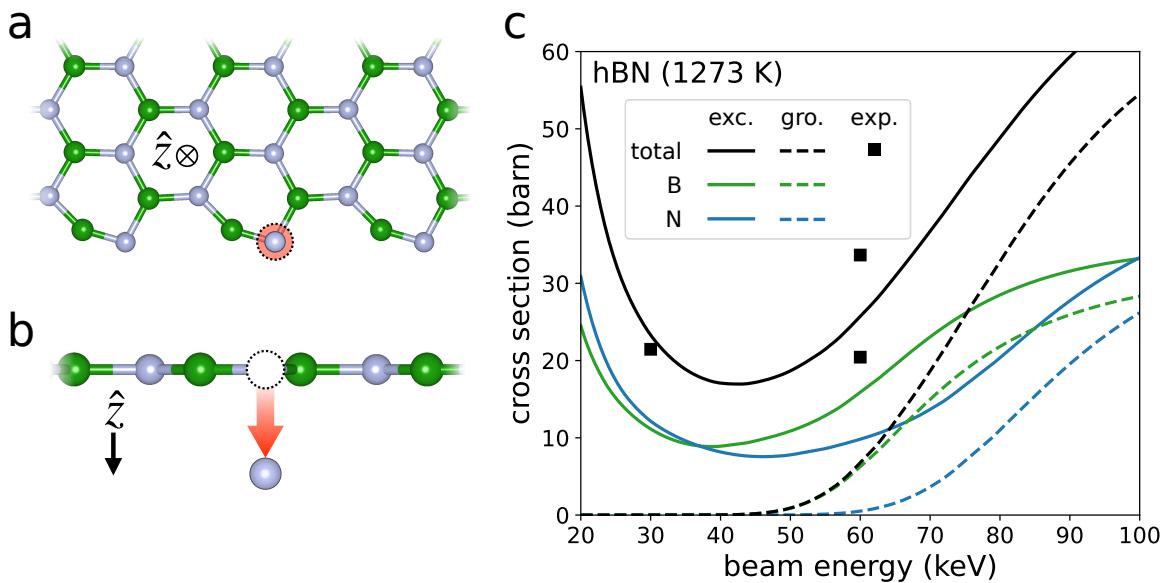


Fig. 5 Sputtering from the armchair edge of hBN. The schematics in panels (a) and (b) depict the sputtering of nitrogen from out-of-plane and in-plane perspectives respectively. Panel (c) plots the sputtering cross sections of both boron and nitrogen in green and blue respectively. The black curves are the total sputtering cross section, *i.e.*, the sum of their corresponding blue and green curves. The solid curves account for beam-induced excitations assuming an excitation lifetime of $\tau = 240$ fs (labelled “exc.” for excited). The dashed curves ignore the possibility of beam-induced excitation (labelled “gro.” for ground). The black squares are experimentally observed sputtering cross sections for hBN at 1273 K.⁹ The consideration of beam-induced excitations reduces the disagreement between theory and experiment substantially.

sistent with the fact that the excitation lifetime of pristine hBN is much longer than that of MoS₂.

The difference between the two materials’ lifetimes leads to markedly dissimilar cross section behaviors at low beam energies. Below beam energies of 30 keV, the cross section of MoS₂ gradually drops to zero with decreasing ε_b . In contrast, hBN’s total cross section has a minimum at around 40 keV and begins to increase as ε_b decreases. Eventually, hBN’s sputtering cross section peaks before dropping quickly to zero as the beam energy goes to zero (figure S2). However, this peak occurs at around 0.5 keV, far below the lower energy bound of figure 5c. These distinct cross section behaviors can be explained by the amplified sensitivity of n_f to τ at low beam energies. Equations (25) and (26) tell us that larger τ makes large n_f more likely. At the same time, n_f cannot exceed n_i . Thus, n_f is more sensitive to τ at low beam energies for which n_i is large. Accordingly, because τ is greater in hBN than in MoS₂, the expected values of n_f in hBN are much larger than those of MoS₂ at low beam energies. This explains why the effect of considering excitations is much more pronounced in hBN than in MoS₂. Furthermore, the difference in the cross section behavior is exacerbated by the fact that sulfur is heavier than both boron and nitrogen, so that its post-collision velocity is always smaller for a given energy transfer E . It follows that t_s is larger, and thus, P_f is smaller for a given E in MoS₂. This again increases the likelihood that hBN has more final excitations than MoS₂, meaning that the effects of beam-induced excitation on hBN’s sputtering cross section are greater.

We also plot the contributions of $n_f = 0, 1$, and 2 excitations to MoS₂’s sputtering cross section separately. In doing so, we see that the contributions of $n_f = 1$ or 2 would conceal the peak at 30

keV. Thus, it seems that the likelihoods of $n_f = 1$ or 2 are somehow suppressed. This suggests that the individual beam-induced excitations and subsequent relaxation are in some cases correlated. That is, the excitation and/or relaxation rates of a given electronic transition are affected by the coinciding distribution of electronic excitations. Thus, a proper treatment of these excitations and their effect on the sputtering cross section requires that the excitation probabilities of every possible transition are calculated for every possible excitation configuration. Such a nonlinear calculation is beyond the scope of this work, but should certainly be pursued in a future study. Nonetheless, the methods laid out here demonstrate that the consideration of beam-induced excitation can provide a quantitative justification for the sulfur sputtering rate to peak at a beam energy well-below the expected ground state critical energy.

4 Conclusions

In this paper, we developed a first-principles method to more accurately describe electron beam-induced sputtering cross sections in 2D insulating crystals by accounting for beam-induced electronic excitations and their subsequent relaxations. The method combines QED scattering theory with DFT electronic structure calculations to determine the likelihood of beam-induced excitation. The results show that the excitation probability is inversely proportional to both the material’s band gap and beam electron’s kinetic energy. We then show how these nonzero excitation probabilities increase the predicted sputtering cross sections of both hBN and MoS₂. These cross sections can be made quantitatively similar to those obtained experimentally by treating the excitation lifetime as a fitting parameter. The methods laid out are com-

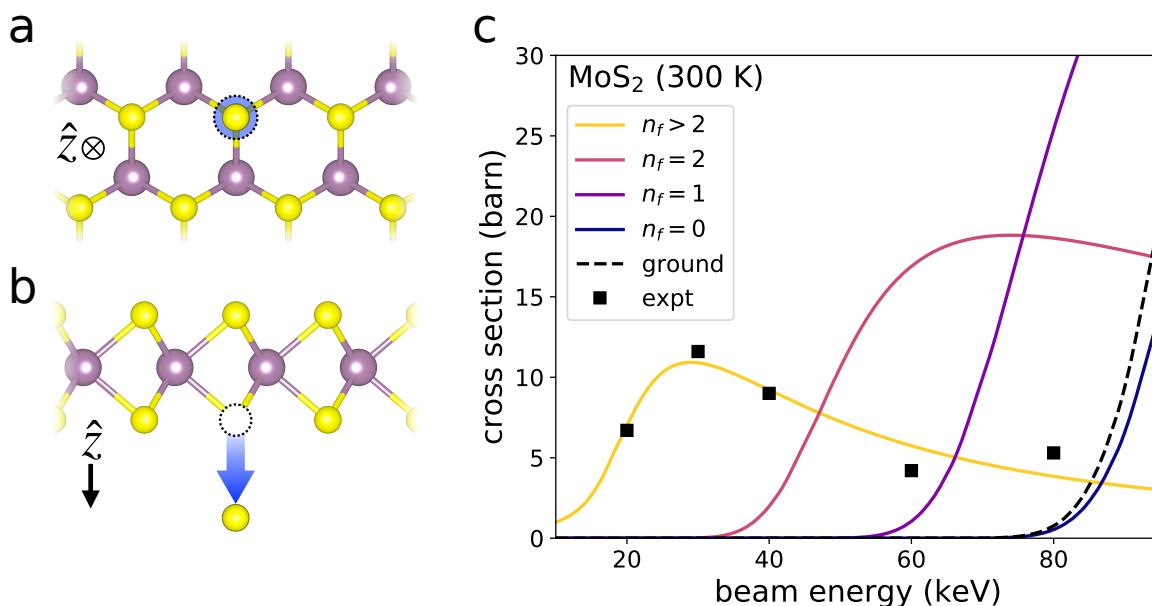


Fig. 6 Sputtering of sulfur from the outgoing MoS₂ surface. The sputtering of molybdenum from MoS₂ can be ignored because its displacement threshold is significantly larger than that of sulfur. The schematics in panels (a) and (b) depict the sputtering process from out-of-plane and in-plane perspectives respectively. Panel (c) plots the contributions of various numbers of final excitations n_f to the sputtering cross section assuming an excitation lifetime of $\tau = 81$ fs. The black dashed curve plots the predicted cross section ignoring the possibility of beam-induced excitation. The black squares are experimentally observed cross sections at 300 K.¹² The contribution from the sum of all $n_f > 2$ final excitations matches the experimental cross section remarkably well.

tationally efficient, requiring only a few ground state electronic optimization calculations for each cross section curve. Thus, the formalism that we have developed can be easily applied to any 2D crystalline material to simulate the rates of atomic displacement under electron irradiation.

With that said, several questions naturally arise from our study. For example, why is the excitation lifetime reduced so drastically during sputtering? How might excitation and relaxation rates of different transitions be correlated? How might preexisting defects affect those rates? These questions urge follow up work to address the full breadth of physical processes involved in beam-matter interactions. Future studies should also consider additional electronic relaxation pathways to determine their effect on P_f . Correspondingly, other electronic responses such as ionization, core excitations, and second order electronic excitation effects such as Auger scattering can be incorporated into the calculation of P_f .^{38,40–42} Moreover, spin polarization effects can also be examined by making \mathcal{M} spin-dependent, *i.e.*, not averaging over spins as is done in equation (4). The beam electron path can also change significantly after collision with the nucleus. Ensuing research should investigate how these altered trajectories generate new excitation probabilities P_i , which must certainly play a role in 3D bulk materials **as well as layered and confined 2D systems**.^{43–46} Furthermore, the methods here can be modified to accommodate more advanced DFT techniques. For example, the calculation of P_i should be made compatible with ultrasoft and projector augmented wave pseudopotentials⁴⁷ in a manner similar to that employed in modern GW codes.^{48–50} In addition, the plane-wave coefficients of excited electronic states can be calcu-

lated self-consistently using constrained DFT.⁵¹ Perhaps most importantly, progress in this field requires much more experimental data. We therefore hope this paper encourages new experimental investigation into beam-induced sputtering for beam energies below the predicted ground state critical energy.

Clearly, the combined QED-DFT approach to modeling beam-induced excitations and their effect on the sputtering cross section opens up a rich and diverse field of physics for both theoretical and experimental exploration. We hope that this work and the work it may stimulate can eventually enable the use of electron beams for precise atomic-scale engineering of crystalline materials.

5 Computational Details

DFT^{13,14} was used to calculate (i) the plane wave coefficients C_{G+k}^n and eigenvalues ϵ_{nk} of each material's electronic orbitals to determine the excitation probabilities (section 2.1.3) and (ii) the free energies of the pristine and sputtered structures to determine the displacement thresholds for sputtering (section 2.2.1).

Plane-wave coefficients and eigenvalues were calculated using Quantum ESPRESSO.⁵² Because equation (16) relies on the orthogonality of the Kohn-Sham orbitals, the optimized norm-conserving Vanderbilt pseudopotentials^{53,54} were employed. The sum of all transition probabilities S defined in equation (18) was used to gage the convergence of all parameters, which were deemed converged when any increase in precision changed S by less than 5% (figure S5). The cutoff energies for hBN and MoS₂ were set to 490 and 408 eV respectively. While these cutoffs are lower than what is typically used for norm-conserving psuedopo-

tentials, both cutoffs yield well-converged values of S (figures S5b and e). Meanwhile, the hBN and MoS₂ unit cells were given heights of 18 and 12 Å respectively. For both materials, the maximum virtual photon momentum $|\mathbf{q}_{\max}|$ required for convergence fell well within their first BZs. We therefore chose $|\mathbf{q}_{\max}|$ to be the magnitude of high-symmetry point K in each respective BZ. The maximum number of initial excitations considered for hBN and MoS₂ were $n_i^{\max} = 5$ and 9 respectively (figure S1). Finally, convergence of S requires extremely dense k-point sampling of the BZ. This necessitates fitting a curve to S calculated for various k-point mesh densities and extrapolating to an infinitely fine mesh to estimate the converged value of S (section S6). The most dense k-point meshes used to fit these curves had dimensions of $45 \times 45 \times 1$ and $36 \times 36 \times 1$ for hBN and MoS₂ respectively, corresponding to very similar linear k-point spacings of 0.056 and 0.055 Å⁻¹ respectively.

Next, free energy calculations were carried out with the Vienna ab initio Simulation Package (VASP)^{15,55} implementing the projector augmented wave (PAW) method⁴⁷ along with the Perdew-Burke-Ernserhof (PBE) generalized gradient approximation (GGA) to the exchange correlation functional.⁵⁶ Van der Waals interactions were accounted for using the optB88-vdW density functional methods.^{57,58} All parameters were converged so that any increase in precision would change the total free energy by less than 1 meV per atom. The parameter values that satisfy this criteria generally differed from those listed in the previous paragraph. The cutoff energies for hBN and MoS₂ were set to 800 and 550 eV respectively, while the BZs of both materials' pristine unit cells were sampled with Γ -centered $6 \times 6 \times 1$ Monkhorst-Pack meshes,⁵⁹ corresponding to linear k-point spacings of 0.417 and 0.328 Å⁻¹ for hBN and MoS₂ respectively. To achieve the same k-point densities, surface vacancies in MoS₂ and edge vacancies in hBN were placed in respective $6 \times 6 \times 1$ and $4 \times 1 \times 1$ supercells whose BZs were sampled with a single k-point on Γ . The heights of the hBN and MoS₂ cells were 12 Å and 20 Å respectively to provide sufficient separation from periodic images. Nanoribbon structures were used to simulate isolated armchair edges in hBN. These ribbons were more than 16 Å across and placed in cells 28 Å wide to avoid interactions between opposing edges and periodic images. Lastly, relaxation iterations of ionic positions and lattice constants persisted until the all Hellmann-Feynman forces settled below 1 meV/Å.

Author Contributions

AY and ML designed the code used to calculate the excitation probabilities and their effect on the sputtering cross section. AY and JG developed the theoretical model to describe beam-induced electronic excitation. AY, DL, JJ, PG, TY, BGS, and VM, developed the theoretical model to describe how excitations affect the sputtering cross section. AY carried out all calculations.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Calculations were performed at the Center for Computational Innovations at Rensselaer Polytechnic Institute, Livermore Computing Center at Lawrence Livermore National Laboratory (LLNL), and Compute and Data Environment for Science at Oak Ridge National Laboratory (ORNL), which is supported by the Office of Science of the U.S. Department of Energy (DOE) under Contract No. DE-AC05-00OR22725. This work was performed under the auspices of the U.S. Department of Energy by LLNL under Contract DE-AC52-07NA27344. This material is also based upon work supported by the U.S. DOE, Office of Science, Office of Workforce Development for Teachers and Scientists, Office of Science Graduate Student Research (SCGSR) program. The SCGSR program is administered by the Oak Ridge Institute for Science and Education for the DOE under contract number DE-SC0014664. Funding was provided by the National Science Foundation (Award 1608171). Work (BGS, PG, DL, JJ) was also supported by ORNL's Center for Nanophase Materials Sciences, a U.S. DOE Office of Science User Facility.

Notes and references

- 1 F. Banhart, *Rep. Prog. Phys.*, 1999, **62**, 1181–1221.
- 2 R. F. Egerton, *Ultramicroscopy*, 2013, **127**, 100–108.
- 3 X. Zhao, J. Kotakoski, J. C. Meyer, E. Sutter, P. Sutter, A. V. Krasheninnikov, U. Kaiser and W. Zhou, *MRS Bull.*, 2017, **42**, 667–676.
- 4 T. Susi, J. C. Meyer and J. Kotakoski, *Nat. Rev. Phys.*, 2019, **1**, 397–405.
- 5 J. C. Meyer, F. Eder, S. Kurasch, V. Skakalova, J. Kotakoski, H. J. Park, S. Roth, A. Chuvalin, S. Eyhusen, G. Benner, A. V. Krasheninnikov and U. Kaiser, *Phys. Rev. Lett.*, 2012, **108**, 196102–196107.
- 6 T. Susi, C. Hofer, G. Argentero, G. T. Leuthner, T. J. Pennycook, C. Mangler, J. C. Meyer and J. Kotakoski, *Nat. Commun.*, 2016, **7**, 13040.
- 7 A. Yoshimura, M. Lamparski, N. Kharche and V. Meunier, *Nanoscale*, 2018, **10**, 2388–2397.
- 8 J. Kotakoski, C. H. Jin, O. Lehtinen, K. Suenaga and A. V. Krasheninnikov, *Phys. Rev. B*, 2010, **82**, 113404.
- 9 O. Cretu, Y. C. Lin and K. Suenaga, *Micron*, 2015, **72**, 21–27.
- 10 Y.-C. Lin, T. Björkman, H.-P. Komsa, P.-Y. Teng, C.-H. Yeh, F.-S. Huang, K.-H. Lin, J. Jadczak, Y.-S. Huang, P.-W. Chiu, A. V. Krasheninnikov and K. Suenaga, *Nat. Commun.*, 2015, **6**, 6736–6741.
- 11 T. Lehnert, O. Lehtinen, G. A. Siller, U. Kaiser, G. Algara-Siller and U. Kaiser, *Appl. Phys. Lett.*, 2017, **110**, 033106.
- 12 S. Kretschmer, T. Lehnert, U. Kaiser, A. V. Krasheninnikov and A. V. Krasheninnikov, *Nano Lett.*, 2020, **20**, 2865–2870.
- 13 P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864–871.
- 14 W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133–1138.
- 15 G. Kresse and J. Furthmüller, *Phys. Rev. B*, 1996, **54**, 11169–11186.
- 16 H. Lourenço-Martins, A. Lubk and M. Kociak, *SciPost Phys.*, 2021, **10**, 31.

- 17 M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Perseus Books Publishing, L.L.C., Reading, MA, 1st edn, 1995.
- 18 T. Lancaster and S. J. Blundell, *Quantum Field Theory for the Gifted Amateur*, Oxford University Press, New York, NY, United States of America, 1st edn, 2014, p. 350.
- 19 N. F. Mott, *Proc. R. Soc. A*, 1929, **124**, 425–442.
- 20 W. A. McKinley and H. Feshbach, *Phys. Rev.*, 1948, **74**, 1759–1763.
- 21 O. S. Oen, *Cross sections for atomic displacements in solids by fast electrons*, Oak ridge national laboratory technical report, 1973.
- 22 R. F. Egerton, R. McLeod, F. Wang and M. Malac, *Ultramicroscopy*, 2010, **110**, 991–997.
- 23 C. Møller, *Ann. Phys.*, 1932, **406**, 531–585.
- 24 H. Kragh, *Arch. Hist. Exact Sci.*, 1992, **43**, 299–328.
- 25 X. Roqué, *Arch. Hist. Exact Sci.*, 1992, **44**, 197–264.
- 26 H. Bethe, *Ann. Phys.*, 1930, **397**, 325–400.
- 27 M. Inokuti, Y.-K. Kim and R. L. Platzman, *Phys. Rev.*, 1967, **164**, 55–61.
- 28 J. Gusakova, X. Wang, L. L. Shiau, A. Krivosheeva, V. Shaposhnikov, V. Borisenko, V. Gusakov and B. K. Tay, *Phys. status solidi*, 2017, **214**, 1700218.
- 29 C. Elias, P. Valvin, T. Pelini, A. Summerfield, C. J. Mellor, T. S. Cheng, L. Eaves, C. T. Foxon, P. H. Beton, S. V. Novikov, B. Gil and G. Cassabois, *Nat. Commun.*, 2019, **10**, 2639.
- 30 D. B. Lingerfelt, P. Ganesh, J. Jakowski and B. G. Sumpter, *Adv. Funct. Mater.*, 2019, **1901901**, 1901901.
- 31 X. Li, J. C. Tully, H. B. Schlegel and M. J. Frisch, *J. Chem. Phys.*, 2005, **123**, 84106.
- 32 H. P. Komsa, J. Kotakoski, S. Kurasch, O. Lehtinen, U. Kaiser and A. V. Krasheninnikov, *Phys. Rev. Lett.*, 2012, **109**, 035503–1–5.
- 33 C. Jin, F. Lin, K. Suenaga and S. Iijima, *Phys. Rev. Lett.*, 2009, **102**, 195505.
- 34 J. C. Meyer, A. Chuvalin, G. Algara-Siller, J. Biskupek and U. Kaiser, *Nano Lett.*, 2009, **9**, 2683–2689.
- 35 M. Dogan, S. M. Gilbert, T. Pham, B. Shevitski, P. Ercius, S. Aloni, A. Zettl and M. L. Cohen, *Appl. Phys. Lett.*, 2020, **117**, 9–13.
- 36 J. Li, X. K. Cao, T. B. Hoffman, J. H. Edgar, J. Y. Lin and H. X. Jiang, *Appl. Phys. Lett.*, 2016, **108**, 122101–1–4.
- 37 T. Korn, S. Heydrich, M. Hirmer, J. Schmutzler, C. Schüller and C. Schller, *Appl. Phys. Lett.*, 2011, **99**, 102109.
- 38 D. Lagarde, L. Bouet, X. Marie, C. R. Zhu, B. L. Liu, T. Amand, P. H. Tan and B. Urbaszek, *Phys. Rev. Lett.*, 2014, **112**, 47401.
- 39 M. Palummo, M. Bernardi and J. C. Grossman, *Nano Lett.*, 2015, **15**, 2794–2800.
- 40 D. Kozawa, R. Kumar, A. Carvalho, K. Kumar Amara, W. Zhao, S. Wang, M. Toh, R. M. Ribeiro, A. H. Castro Neto, K. Matsuda and G. Eda, *Nat. Commun.*, 2014, **5**, 4543.
- 41 Z. Nie, R. Long, J. S. Teguh, C.-C. Huang, D. W. Hewak, E. K. L. Yeow, Z. Shen, O. V. Prezhdo and Z.-H. Loh, *J. Phys. Chem. C*, 2015, **119**, 20698–20708.
- 42 H. Shi, R. Yan, S. Bertolazzi, J. Brivio, B. Gao, A. Kis, D. Jena, H. G. Xing and L. Huang, *ACS Nano*, 2013, **7**, 1072–1080.
- 43 W. Wang, Y. Zheng, X. Li, Y. Li, H. Zhao, L. Huang, Z. Yang, X. Zhang and G. Li, *Adv. Mater.*, 2019, **31**, e1803448.
- 44 A. Kakanakova-Georgieva, G. K. Gueorguiev, D. G. Sangiovanni, N. Suwannaharn, I. G. Ivanov, I. Cora, B. Pécz, G. Nicotra and F. Giannazzo, *Nanoscale*, 2020, **12**, 19470–19476.
- 45 A. Kakanakova-Georgieva, F. Giannazzo, G. Nicotra, I. Cora, G. K. Gueorguiev, P. O. Å. Persson and B. Pécz, *Appl. Surf. Sci.*, 2021, **548**, 149275.
- 46 Z. Wang, G. Wang, X. Liu, S. Wang, T. Wang, S. Zhang, J. Yu, G. Zhao and L. Zhang, *J. Mater. Chem. C*, 2021, **9**, 17201–17232.
- 47 P. E. Blöchl, *Phys. Rev. B*, 1994, **50**, 17953–17979.
- 48 M. Shishkin and G. Kresse, *Phys. Rev. B*, 2006, **74**, 035101.
- 49 M. Gajdoš, K. Hummer, G. Kresse, J. Furthmüller and F. Bechstedt, *Phys. Rev. B*, 2006, **73**, 045112.
- 50 J. Paier, R. Hirschl, M. Marsman and G. Kresse, *J. Chem. Phys.*, 2005, **122**, 234102.
- 51 B. Kaduk, T. Kowalczyk and T. Van Voorhis, *Chem. Rev.*, 2012, **112**, 321–370.
- 52 P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougaudis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari and R. M. Wentzcovitch, *J. Phys. Condens. Matter*, 2009, **21**, 395502.
- 53 D. R. Hamann, *Phys. Rev. B*, 2013, **88**, 85117.
- 54 M. Schlipf and F. Gygi, *Comput. Phys. Commun.*, 2015, **196**, 36–44.
- 55 G. Kresse and J. Furthmüller, *Comput. Mater. Sci.*, 1996, **6**, 15–50.
- 56 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.
- 57 J. Klimeš, D. Bowler and A. Michaelides, *J. Phys. Condens. Matter*, 2009, **22**, 022201–022205.
- 58 J. Klimeš, D. R. Bowler and A. Michaelides, *Phys. Rev. B*, 2011, **83**, 195131–1–13.
- 59 H. Monkhorst and J. Pack, *Phys. Rev. B*, 1976, **13**, 5188–5192.

Supporting Information

Quantum Theory of Electronic Excitation and Sputtering by Transmission Electron Microscopy

Anthony Yoshimura^{1,2*}, Michael Lamparski², Joel Giedt²,
David Lingerfelt³, Jacek Jakowski³, Panchapakesan Ganesh³, Tao Yu⁴,
Bobby Sumpter³, and Vincent Meunier^{2,5}

¹*Lawrence Livermore National Laboratory, Livermore, CA 94550, USA*

²*Department of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute,
Troy, New York 12180, USA*

³*Center for Nanophase Material Sciences, Oak Ridge National Laboratory, Oak Ridge, TN
37831, USA*

⁴*Department of Chemistry, University of North Dakota, Grand Forks, ND 58202, USA*

⁵*Department of Materials Science and Engineering, Rensselaer Polytechnic Institute, Troy,
NY 12180, USA*

**Correspondence to be addressed to yoshimura4@llnl.gov*

S1 Invariant matrix element \mathcal{M}

As the excitation amplitude in equation (16) contains the invariant matrix element \mathcal{M} and not $|\mathcal{M}|^2$, we are unable to use the spin sum identities typically used to derive cross sections in quantum electrodynamics.^{1,2} Nonetheless, the evaluation of \mathcal{M} in equation (4) is straightforward, though a bit cumbersome. In the Dirac basis,³

$$\gamma^0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \quad \text{and} \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \quad (\text{S1})$$

where

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \vec{\sigma} = \left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\}. \quad (\text{S2})$$

The electron spinors in equation (4) can be written as

$$u^1(p) = \sqrt{\epsilon + m} \begin{pmatrix} 1 \\ 0 \\ \frac{p^z}{\epsilon + m} \\ \frac{p^x + ip^y}{\epsilon + m} \end{pmatrix} \quad \text{and} \quad u^2(p) = \sqrt{\epsilon + m} \begin{pmatrix} 0 \\ 1 \\ \frac{p^x - ip^y}{\epsilon + m} \\ \frac{-p^z}{\epsilon + m} \end{pmatrix}, \quad (\text{S3})$$

while

$$\bar{u}^s(p) = u^{s\dagger}(p)\gamma^0. \quad (\text{S4})$$

As justified in section 2.1.1, we need only evaluate the t-channel contribution to \mathcal{M} and multiply the result by 2. That is,

$$\mathcal{M}(p_4 p_3 \leftarrow p_2 p_1) \sim e^2 \sum_{s_1, s_2, s_3, s_4} \bar{u}^{s_4}(p_4) \gamma^\mu u^{s_1}(p_1) \left(\frac{1}{p_3 - p_2} \right)^2 \bar{u}^{s_3}(p_3) \gamma_\mu u^{s_2}(p_2). \quad (\text{S5})$$

Substituting equations (S1) through (S4) into (S5) yields \mathcal{M} in terms of the components of the electrons' 4-momenta. That is,

$$\begin{aligned} \mathcal{M}(p_4 p_3 \leftarrow p_2 p_1) &\sim -\frac{2e^2}{(p_3 - p_2)^2} [(\epsilon_1 + m)(\epsilon_2 + m)(\epsilon_3 + m)(\epsilon_4 + m)]^{-1/2} \\ &\times \left\{ (\epsilon_1 + m)p_4^x [(\epsilon_2 + m)p_3^x + (\epsilon_3 + m)p_2^x] \right. \\ &+ 2(\epsilon_1 + m)(\epsilon_3 + m)(p_2^y + ip_2^z)(p_4^y - ip_4^z) \\ &+ 2(\epsilon_2 + m)(\epsilon_4 + m)ip_1^z(p_3^y - ip_3^z) \\ &- [(\epsilon_2 + m)(\epsilon_3 + m) + p_2^x p_3^x + (p_2^y + ip_2^z)(p_3^y - ip_3^z)] \\ &\left. \times [(\epsilon_1 + m)(\epsilon_4 + m) + ip_1^z(p_4^y - ip_4^z)] \right\}, \end{aligned} \quad (\text{S6})$$

where we let p_1 denote the momentum of the beam electron so that $p_1^x = p_1^y = 0$. See the Mathematica⁴ notebook in the Supporting Information for more details.

S2 Normalization

When integrating over 4-momentum space as is done in section 2.1.2, Lorentz invariance constrains a particle's 4-momentum to obey $p^2 = m^2$. It follows that the 4-momentum integration measure $d^4 p$ is always multiplied by a delta function $\delta(p^2 - m^2)$, *i.e.*,

$$d^4 p \delta(p^2 - m^2) \theta(p^0) = \frac{d^4 p}{2p^0} \delta(p_0 - \epsilon_{\mathbf{p}}), \quad (\text{S7})$$

where the Heaviside step function restricts our consideration to particles of positive mass (we note that antiparticles are interpreted as positive mass particles that propagate backwards in time). With this integration measure, the identity operator can be written as

$$\begin{aligned} \hat{I} &= \int \frac{d^4 p}{(2\pi)^4} (2\pi) \delta(p^2 - m^2) \theta(p^0) |p\rangle \langle p| \\ &= \int \frac{d^3 p}{(2\pi)^3 2\epsilon_{\mathbf{p}}} |p\rangle \langle p| \\ &= \int \frac{d^3 p}{(2\pi)^3} |\mathbf{p}\rangle \langle \mathbf{p}|. \end{aligned} \quad (\text{S8})$$

The last equality implies that

$$|p\rangle = (2\epsilon_{\mathbf{p}})^{1/2} |\mathbf{p}\rangle. \quad (\text{S9})$$

S3 Lab frame p_3^z

For the scattering of two free electrons, conservation of 4-momentum allows us to write

$$p_1 + p_2 = p_3 + p_4. \quad (\text{S10})$$

This constrains four of the six components needed to specify the 3-momenta of the two outgoing particles, so that only two components are independent. We choose the independent components to be p_3^x and p_3^y . Thus, we wish to find p_3^z as a function of p_1 , p_2 , p_3^x , and p_3^y . From this, $p_4 = p_1 + p_2 - p_3$ is easily obtained, and we have all six components of outgoing 3-momenta needed to calculate the scattering amplitude from equation (16). We start by setting the z -direction parallel to \mathbf{p}_1 . This means that

$$\begin{aligned} p_4^x &= p_2^x - p_3^x \\ p_4^y &= p_2^y - p_3^y \\ p_4^z &= p_1^z + p_2^z - p_3^z \\ \epsilon_4 &= \epsilon_1 + \Delta\epsilon \end{aligned} \quad (\text{S11})$$

where

$$\epsilon_i = \epsilon_{\mathbf{p}_i} = \sqrt{\mathbf{p}_i^2 + m^2}, \quad \Delta\epsilon = \epsilon_{n\mathbf{k}} - \epsilon_{n'\mathbf{k}'}, \quad (\text{S12})$$

and $\epsilon_{n\mathbf{k}}$ and $\epsilon_{n'\mathbf{k}'}$ are the eigenvalues of the excited hole and electron states, respectively. We can also write $\epsilon_1 = \gamma m$ where $\gamma = 1/\sqrt{1 - \beta^2}$ is the Lorentz factor corresponding to the beam electron's velocity β . Squaring the last line in (S11) and subtracting m^2 then gives

$$\mathbf{p}_4^2 = \mathbf{p}_1^2 + 2\gamma m\Delta\epsilon + \mathcal{O}(\Delta\epsilon^2). \quad (\text{S13})$$

We can ignore terms of order $\mathcal{O}(\Delta\epsilon^2)$ since $\Delta\epsilon \ll m$. Meanwhile, squaring and summing the first three equations in (S11) tells us that

$$\mathbf{p}_4^2 = \mathbf{p}_1^2 + \mathbf{p}_2^2 + \mathbf{p}_3^2 + 2 [p_1^z p_2^z - \mathbf{p}_2^\perp \cdot \mathbf{p}_3^\perp - (p_1^z + p_2^z)p_3^z], \quad (\text{S14})$$

where the \perp superscript denotes the projection perpendicular to \hat{z} . Subtracting equation (S14) from equation (S13) then yields

$$0 \sim \mathbf{p}_2^2 + \mathbf{p}_3^2 + 2 [p_1^z p_2^z - \mathbf{p}_2^\perp \cdot \mathbf{p}_3^\perp - (p_1^z + p_2^z)p_3^z] - 2\gamma m\Delta\epsilon. \quad (\text{S15})$$

Asymptotic formula (S15) can then be solved for p_3^z , so that

$$p_3^z \sim p_1^z + p_2^z \pm \sqrt{(p_1^z)^2 - (\mathbf{p}_2^\perp - \mathbf{p}_3^\perp)^2 + 2\gamma m\Delta\epsilon}. \quad (\text{S16})$$

We choose the $-$ from \pm as we impose that \mathbf{p}_3 is a component of the outgoing crystal electron state, whose z -momentum is much smaller than that of the beam electron.

S4 Converging n_i^{\max}

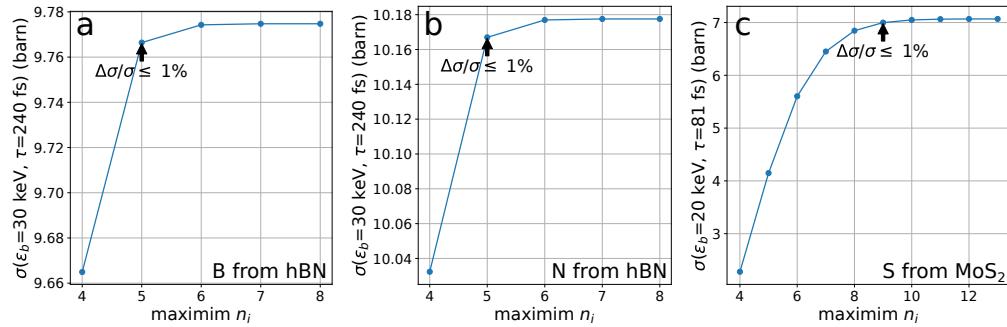


Figure S1: Convergence of the sputtering cross section with respect to the maximum number of beam-induced excitations n_i^{\max} considered. The simulated beam energies ϵ_b are the lowest experimental beam energies used for each material. The excitation lifetimes τ are those used to fit the experimental data in figures 5 and 6. Each cross section is deemed converged when any increase in n_i^{\max} increases the cross section by less than 1%.

S5 Peaks in the sputtering cross section of hBN

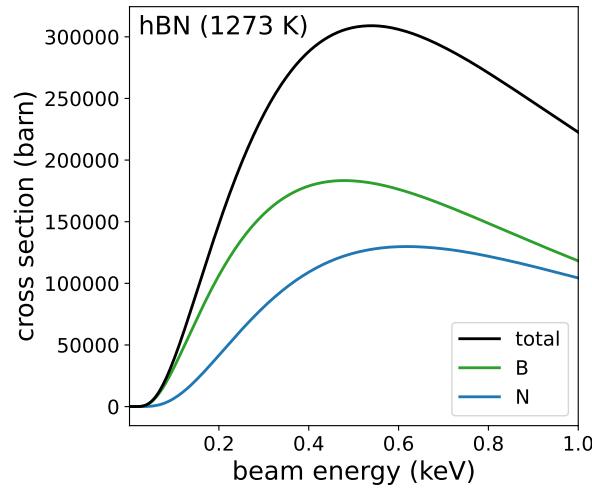


Figure S2: The sputtering cross sections of boron and nitrogen from the hBN armchair edge peak at beam energies much lower than those typically used for microscopy and defect engineering. However, as mentioned in the main text, the validity of our perturbative approximation to the scattering operator breaks down at low beam energies (section 2.1.4). Thus, the values and positions of these peaks may change if higher order perturbation terms are considered.

S6 Fitting and converging S

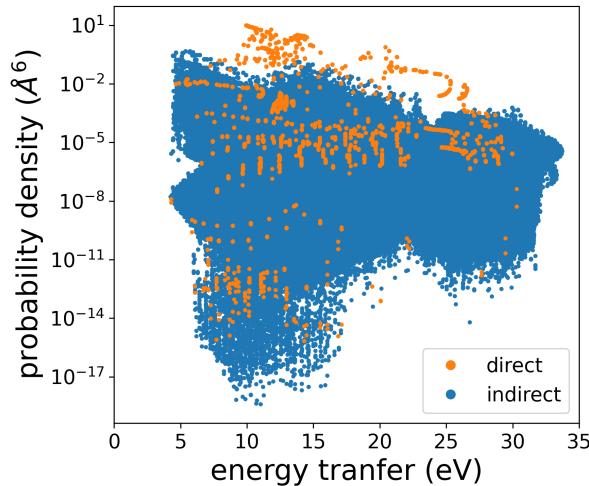


Figure S3: The most probable beam-induced electronic excitations are direct transitions. Each circle denotes the probability density centered on a particular excitation instigated by a 40 keV beam electron in hexagonal boron nitride, whose Brillouin zone was sampled by a $18 \times 18 \times 1$ Monkhorst-Pack mesh.

As shown in formula (23), the probability $P_i(\epsilon_b, n_i)$ that a beam electron of energy ϵ_b excites a specific number n_i of material electrons depends solely $S(\epsilon_b)$, the sum of all transition probabilities defined in equation (18). As such, convergence of S with respect to all simulation parameters is crucial to calculating the effect of electronic excitation on sputtering. In this section, we describe the general scheme that we use to converge S with respect to our simulation parameters. In particular, the convergence with respect to k-point density must be treated with care, and will be the primary focus of the remainder of this section.

Because the square of the momentum transfer appears in the denominator of the t -channel in equation (4), the probability density for electronic excitation peaks sharply when the momentum transfer is minimized. This minimization occurs when the beam electron induces a direct transition, in which the resulting electron and hole lie on the same k-point. Thus, direct excitations that conserve crystal momentum tend to be much more likely than indirect excitations (figure S3).

As a result, a sparse k-point mesh leads to an overestimation of S . This is because a coarser k-point mesh yields a higher proportion of direct transitions, as there are fewer indirect transitions available for each k-point. For a given number of k-points N_k , the number of direct transitions is proportional to N_k , while the number of indirect transitions is proportional to N_k^2 . Thus, increasing N_k increases the proportion of possible indirect transitions, lowering S . For a sufficiently dense mesh, the distinction between direct and nearly direct transitions is inconsequential, in which case, S is converged with respect to N_k .

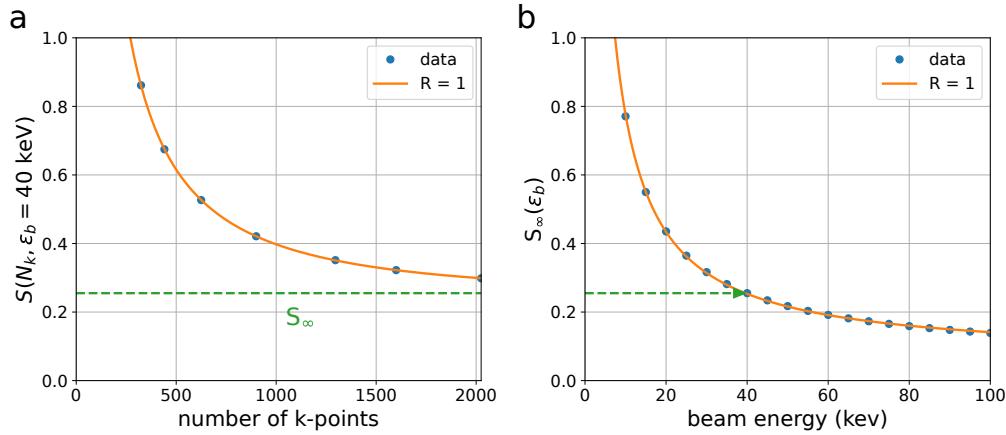


Figure S4: Finding the large crystal limit of $S(\epsilon_b)$ requires extrapolation. Panel (a) shows the dependence of S on the number N_k in the Brillouin zone of hBN under 40 keV irradiation. The simulated points are fitted well by equation (S17). The green dashed line denotes S_∞ , the asymptotic limit of $S(N_k)$ for large N_k . Panel (b) then shows the dependence of S_∞ on the beam energy ϵ_b , which is fitted well by equation (S18). The green arrow in panel (b) illustrates how $S_\infty(\epsilon_b = 40 \text{ keV})$ is determined by the fit in panel (a). [The fitted curve in panel \(b\) is then used as \$S\(\epsilon_b\)\$ in formula \(23\) to calculate \$P_i\$ for hBN.](#) The same fitting and extrapolation method is used for MoS₂.

Satisfactory convergence of S requires an extremely dense k-point mesh. We therefore use extrapolation to determine the converged value of S . For a given beam energy ϵ_b , we calculate S for various N_k . We then fit the points to a curve of the form

$$S(N_k, \epsilon_b) = \frac{a}{N_k - b} e^{-cN_k} + S_\infty, \quad (\text{S17})$$

where a , b , c , and S_∞ are all fitting parameters that depend on ϵ_b (figure S4a). The last parameter S_∞ is the limit of S for large N_k , and is taken as the converged value of S for that beam energy. We repeat this for multiple values of ϵ_b ranging from 5 to 100 keV and record the best fit S_∞ for each ϵ_b . Based on the work of Bethe,^{5–7} we fit the resulting values of $S_\infty(\epsilon_b)$ to an inverse function,

$$S_\infty(\epsilon_b) = \frac{A}{\epsilon_b - B} + C. \quad (\text{S18})$$

where A , B , and C are fitting parameters. The parameters for hBN and MoS₂ are given in table S1. The fitted curve can then be substituted for $S(\epsilon_b)$ in formula (23) to obtain $P_i(\epsilon_b, n_i)$.

Finally, S is also converged with respect to the maximum virtual photon momentum, DFT cutoff energy, and height of the pristine unit cell (figure S5). In these cases, S is considered converged with respect to a parameter when any increase in the parameter's precision changes S by less than 5%.

material	A (keV)	B (keV)	C
hBN	7.655	-0.770	0.06526
MoS ₂	49.05	-3.867	0.3547

Table S1: Fitting parameters of equation (S18) for hBN and MoS₂.

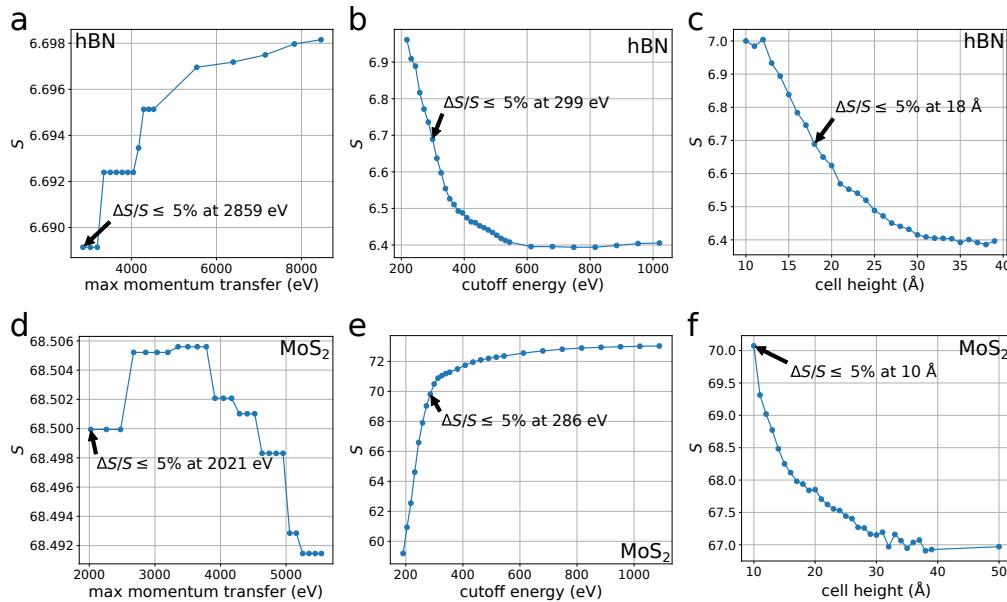


Figure S5: Convergence of S with respect to the (a and d) maximum magnitude of virtual photon momentum considered, (b and e) plane-wave DFT cutoff energy, and (c and f) height of the unit cell. A $6 \times 6 \times 1$ k-point mesh and a beam energy of 60 keV is used to generate all six plots. The converged parameters were found to be insensitive to changes in k-point density and beam energy. S is deemed converged when any increase in precision changes S by less than 5%.

S7 Validity of the three simplifying assumptions

While the three assumptions laid out in section 2.2.1 may seem drastic, their effect on the sputtering cross section is small due to a cancellation of errors. To see this, consider the connection between the potential energy surfaces (ESs) and the PKA velocity. By assumption 1, the ESs at the PKA's initial position are equally spaced, separated by the material's band gap. By assumption 2, the ESs converge to the same energy as the PKA approaches a distance d . We shall ignore assumption 3 for now. If we consider that the ESs are not constant as the PKA moves towards d , so that the PKA slows down in transit, then the energy required to get to d depends on when the system relaxes to a lower ES. This is because higher ESs have shallower slopes than do the lower ones, so the PKA's deceleration will be smaller while on a higher ES. Thus, as the PKA moves, the ESs converge by assumption 2, and the energy lost during relaxation tends to shrink.

By introducing assumption 3, the energy lost during relaxation does not shrink but remains constant while the PKA sputters, and is therefore generally greater than it would be with gradually changing ESs. Our method therefore overestimates the energy lost during relaxation. On the other hand, because the PKA does not slow down until it reaches d , assumption 3 also underestimates the sputtering time t_s , introduced in equation (25). A smaller t_s means that fewer excitations relax during sputtering, increasing the cross section. However, each relaxation comes at a higher energy cost, which makes the PKA less likely to overcome the displacement threshold, decreasing the cross section. Thus, the underestimation t_s counteracts the overestimation of the energy lost during relaxation, so that final effect on the sputtering cross section is small.

Considering the ESs also sheds light on the computational cost saved by assumption 3. Without assumption 3, t_s would depend on the exact times that each relaxation occurs. In this case, calculating the cross section would require integration over time for each excited ES considered. We have found that the consideration of 5 and 9 excited ESs are necessary to converge the cross sections of hBN and MoS₂ respectively (figure S1). For an N-point Gaussian quadrature, this would multiply the computational costs by N^5 and N^9 respectively. Assumption 3 removes this complication by eliminating the need to consider exactly when each relaxation happens. Instead, only the total number of relaxations before the PKA reaches d need be considered.

S8 Calculating E_{\min}

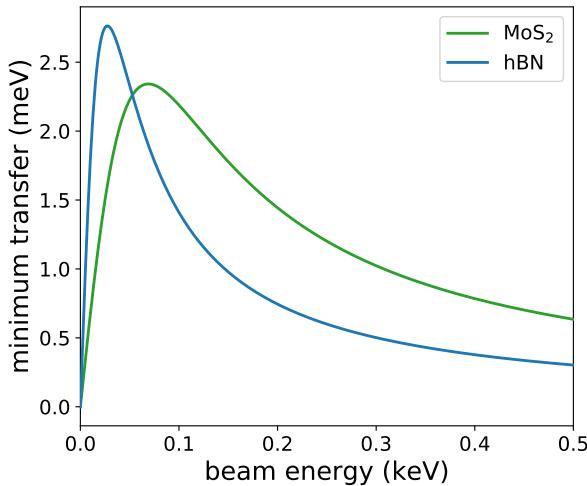


Figure S6: The minimum energy transfer from the beam electron to the target nucleus peaks at very low beam energies and is always much smaller than the nuclei's average thermal kinetic energy of ~ 39 meV at room temperature.

In hBN, a unit cell contains a boron and nitrogen atom. As these atoms have similar masses and displacement thresholds (table 1), the sputtering of both atoms should be considered for a given beam energy. Thus, we approximate the maximum cross sectional area σ_{\max} of these atoms to be $\Omega_{\text{hBN}}/2$, where Ω_{hBN} is the area of the hBN unit cell. On the other hand, molybdenum is much heavier than sulfur and has a much larger displacement threshold in MoS₂.⁸ This means that only the sputtering of sulfur needs to be considered. Additionally, of the two sulfur atoms in the MoS₂ unit cell, only the atom on the outgoing surface is eligible to sputter from a pristine system.⁸ Therefore, σ_{\max} for sulfur sputtering from MoS₂ is Ω_{MoS_2} , the area of the MoS₂ unit cell.

To approximate E_{\min} , we use the Rutherford displacement cross section^{9–11} as an approximation to equation (2),

$$\sigma_R = \pi \left(\frac{Z\alpha}{|\mathbf{p}|\beta} \right)^2 \left(\frac{E_{\max}}{E_d} - 1 \right). \quad (\text{S19})$$

Setting $\sigma_R = \sigma_{\max}$ and $E_d = E_{\min}$ and solving for E_{\min} yields

$$E_{\min}(\epsilon_b) = E_{\max} \left[\frac{\Omega}{\pi} \left(\frac{|\mathbf{p}_b|\beta}{Z\alpha} \right)^2 + 1 \right]^{-1}. \quad (\text{S20})$$

Using the Rutherford cross section instead of the McKinley-Feshbach cross section should be accurate for small beam energies for which $\beta \ll 1$. This makes it well-suited for finding E_{\min} , since $E_d(n_f) < E_{\min}$ only for large n_f , and large n_i (and thus n_f) are much more likely for small beam energies (figure 3).

With that said, the sputtering cross section is fairly insensitive to the exact value of E_{\min} when $E_{\min} \ll E_{\max}$. This is because the post-collision velocity of the PKA diminishes when the energy transfer E shrinks. The beam-induced excitations therefore have more time to relax before the PKA reaches the step in the energy surface (section 2.2.1). This makes P_f small for small E , so that contributions to the integral in equation (3) are extremely tiny for E near E_{\min} . As a result, changes in E_{\min} are essentially immeasurable for beam energies greater than 1 keV. Nonetheless, the use of E_{\min} is necessary for the calculation of a finite sputtering cross section.

References

- [1] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Perseus Books Publishing, L.L.C., Reading, MA, 1st edn, 1995.
- [2] T. Lancaster and S. J. Blundell, *Quantum Field Theory for the Gifted Amateur*, Oxford University Press, New York, NY, United States of America, 1st edn, 2014, p. 350.
- [3] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics*, McGraw-Hill, 1964.
- [4] W. R. Inc., *Mathematica*, Version 12.3, <https://www.wolfram.com/mathematica>.
- [5] H. Bethe, *Ann. Phys.*, 1930, **397**, 325–400.
- [6] T. Susi, J. C. Meyer and J. Kotakoski, *Nat. Rev. Phys.*, 2019, **1**, 397–405.
- [7] S. Kretschmer, T. Lehnert, U. Kaiser, A. V. Krasheninnikov and A. V. Krasheninnikov, *Nano Lett.*, 2020, **20**, 2865–2870.
- [8] H. P. Komsa, J. Kotakoski, S. Kurasch, O. Lehtinen, U. Kaiser and A. V. Krasheninnikov, *Phys. Rev. Lett.*, 2012, **109**, 035503–1–5.
- [9] S. T. Thornton and J. B. Marion, *Classical Dynamics of Particles and Systems*, Brooks/Cole, Boston, MA, 5th edn, 2004.
- [10] J. J. Sakurai and J. Napolitano, *Modern Quantum Mechanics*, Pearson Education, San Francisco, CA, 2nd edn, 2011.
- [11] A. Yoshimura, M. Lamparski, N. Kharche and V. Meunier, *Nanoscale*, 2018, **10**, 2388–2397.