Quantum Theory of Electronic Excitation and Sputtering by Transmission Electron Microscopy

Anthony Yoshimura[1,2*], Michael Lamparski[2], Joel Giedt[2],
David Lingerfelt[3], Jacek Jakowski[3], Panchapakesan Ganesh[3], Tao Yu[4],
Bobby Sumpter[3], 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]

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

where

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

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

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

while

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

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,

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

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

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

where we let p_1 denote the momentum of the beam electron so that $p_1^x = p_1^y = 0$. See the Mathematica[4] 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^4p is always multiplied by a delta function $\delta(p^2 - m^2)$, *i.e.*,

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

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

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

The last equality implies that

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

S3 Lab frame p_3^z

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

$$p_1 + p_2 = p_3 + p_4. \tag{S10}$$

This constrains four of the six components needed to specify the 3-momenta of the two outgoing particles, so that only two componenets 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

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

where

$$\epsilon_i = \epsilon_{\mathbf{p}_i} = \sqrt{\mathbf{p}_i^2 + m^2}, \qquad \Delta\epsilon = \epsilon_{n\mathbf{k}} - \epsilon_{n'\mathbf{k}'}, \tag{S12}$$

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

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

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

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

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

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

Asymptotic formula (S15) can then be solved for p_z^3, so that

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

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.
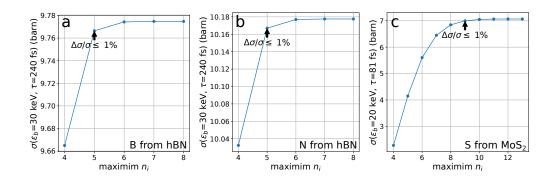
S4 Converging n_i^{max}

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%.

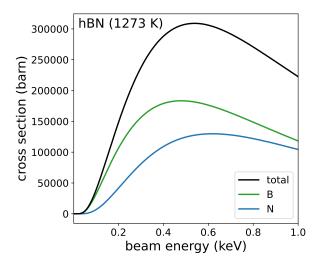# S5   Peaks in the sputtering cross section of hBN

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.
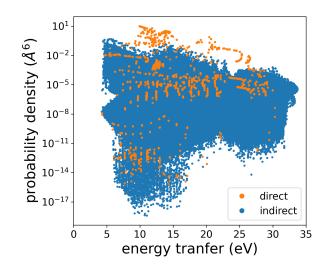
S6 Fitting and converging S

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 probabilites 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 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.
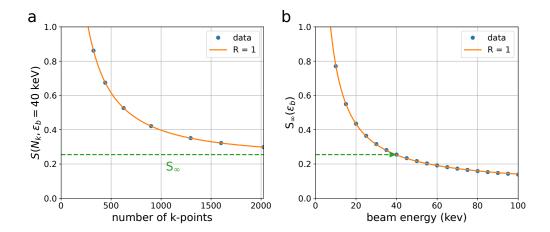
7

Figure S4: Finding the large crystal limit of $S(\epsilon_b)$ requires extrapolation. Panel (a) shows the dependence of S on the number k-points 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 MoS$_2$.

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

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

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,

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

where A, B, and C are fitting parameters. The parameters for hBN and MoS$_2$ 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

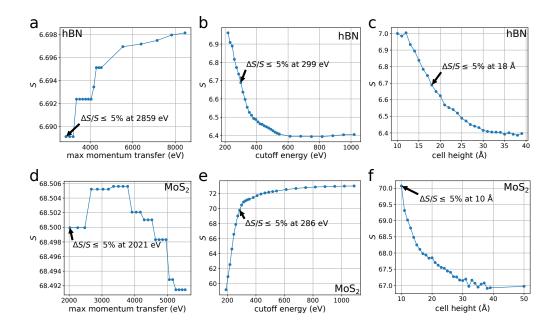Table S1: Fitting parameters of equation (S18) for hBN and MoS$_2$.

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 affect 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 MoS_2 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.
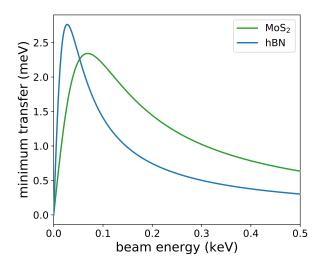
S8   Calculating E_{min}

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 $\sim$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 $\Omega_{\mathrm{hBN}}/2$, 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 $\mathrm{MoS_2}$.[8] This means that only the sputtering of sulfur needs to be considered. Additionally, of the two sulfur atoms in the $\mathrm{MoS_2}$ 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 $\mathrm{MoS_2}$ is $\Omega_{\mathrm{MoS_2}}$, the area of the $\mathrm{MoS_2}$ unit cell.

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

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

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

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

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 $E_d(n_f) < E_{\mathrm{min}}$ only for large n_f, and large n_i (and thus n_f) are much more likely for small beam energies (figure 3).

11

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 Fransisco, CA, 2nd edn, 2011.

[11] A. Yoshimura, M. Lamparski, N. Kharche and V. Meunier, *Nanoscale*, 2018, **10**, 2388–2397.