Phonon-assisted exciton dissociation in transition metal dichalcogenides†

Raül Perea-Causín, a Samuel Brem b and Ermin Malic a,b

Monolayers of transition metal dichalcogenides (TMDs) have been established in the last years as promising materials for novel optoelectronic devices. However, the performance of such devices is often limited by the dissociation of tightly bound excitons into free electrons and holes. While previous studies have investigated tunneling at large electric fields, we focus in this work on phonon-assisted exciton dissociation that is expected to be the dominant mechanism at small fields. We present a microscopic model based on the density matrix formalism providing access to time- and momentum-resolved exciton dynamics including phonon-assisted dissociation. We track the pathway of excitons from optical excitation via thermalization to dissociation, identifying the main transitions and dissociation channels. Furthermore, we find intrinsic limits for the quantum efficiency and response time of a TMD-based photodetector and investigate their tunability with externally accessible knobs, such as excitation energy, substrate screening, temperature and strain. Our work provides microscopic insights in fundamental mechanisms behind exciton dissociation and can serve as a guide for the optimization of TMD-based optoelectronic devices.

1. Introduction

The extensive research on two-dimensional materials during the last decade has put monolayers of transition metal dichalcogenides (TMDs) in the spotlight for next-generation optoelectronic applications.1–4 The strong light–matter interaction and the ultrafast non-equilibrium dynamics5–9 makes them excellent candidates for active materials in photodetectors and solar cells. The understanding of their properties has significantly advanced in the last years with experimental and theoretical studies shining light on the optical response,10–14 exciton relaxation dynamics15–19 and exciton propagation20–24 in different TMD materials. While these are important properties for the operation of a photodetector, exciton dissociation plays a central role as the bridge between optical excitation and photocurrent generation.

During the last years the characteristics of TMD-based optoelectronic devices have been investigated, reporting promising features, such as fast photosresponse,25,26 large responsivities27–30 and high tunability.31,32 To reach optimal operation of these devices, exciton dissociation mechanisms and their tunability need to be understood. Previous studies have investigated exciton dissociation via tunneling to the continuum in presence of strong electric fields,33,34 showing that field-induced dissociation dominates the response at large fields needed to dissociate excitons with large binding energies.35 However, the dissociation of excitons via scattering with phonons, which is expected to play a major role at low electric fields, has been overlooked so far—despite its potential implications on the fundamental efficiency limits of TMD-based photodetectors and solar cells. After an optical excitation, the generated excitons scatter with phonons to reach a thermal equilibrium leading to a finite population of unbound electrons and holes as described by the Saha equation.36 Under the presence of an electric field, these free carriers are dragged away and excitons dissociate to preserve the thermal equilibrium (cf. Fig. 1). Exciton dissociation hence poses an upper fundamental limit on the efficiency of TMD-based photodetectors and solar cells.

The aim of this work is to provide microscopic insights in fundamental processes governing exciton dissociation in TMD materials. Based on a fully quantum mechanical approach, we resolve the dissociation dynamics of excitons in energy, momentum and time. We take into account the Rydberg-like series of bright and dark excitonic states giving rise to a multi-excitonic landscape, which plays an essential role for the dissociation process. We track the pathway of excitons from optical excitation via thermalization to dissociation into free carriers and reveal the underlying microscopic channels (cf. Fig. 1). In particular, we find that intervalley scattering dominates the dissociation...
Nanoscale

Fig. 1 Exciton dissociation. Schematic representation of exciton thermalization and dissociation in WSe₂. Excitons are generated by optical excitation in the KK valley (represented by G in the figure). Scattering with phonons leads to a thermalization of excitons and a redistribution across all available states ( Kum). Excitons close to the continuum can dissociate into free electrons and holes (γnr) preserving the thermal equilibrium. The dissociation of bound excitons (γrad) results in a finite population of unbound carriers. Finally, free carriers are dragged away by an electric field and produce a photocurrent (jph). Excitons in the light cone at the KK valley can recombine radiatively (γrad), while excitons with non-zero momentum recombine non-radiatively (γnr).

dynamics due to its strong exciton–phonon coupling, scattering excitons to higher energies and dissipating them into free electrons and holes. Furthermore, we investigate the tunability of the dissociation efficiency and response time through externally accessible knobs and find the best conditions for optimal operation of a TMD-based photodetector. We find a dissociation-limited response time for WSe₂ monolayers that is in good agreement to reported experimental values, supporting the accuracy of our model and the crucial role of exciton–phonon scattering at weak fields.

2. Microscopic model

The photocurrent created through dissociation of excitons is the key quantity in TMD-based optoelectronic devices, such as photodetectors and solar cells. In excitonic basis, the current can be expressed as

\[ I_j = \sum_{jk} f'_k n_j, \]

where \( f'_k \) is the single-particle occupation, \( \phi'_k \) the exciton wavefunction, and \( n_j \) the exciton density in the state \( j \). Since the squared wavefunction \( |\phi'_k|^2 \) for bound states is an even function and \( k \) is odd, the momentum integration yields zero current. This is not the case for unbound electron–hole pairs, and hence the photocurrent is generated only by free carriers, cf. the ESL. When TMD-based devices are optically excited, excitons are formed and scatter with phonons to reach a thermal distribution with a finite population of unbound carriers. Under the presence of an electric field, free carriers are dragged away generating a current and excitons close to the continuum continuously dissociate preserving the thermal equilibrium. The dissociation process can be slow compared to transport and therefore poses a fundamental limit to the current generation and the response time of photodetectors. In the following, we present a quantum-mechanical model to describe the dynamics of excitons including optical excitation, intra- and intervalley thermalization and, in particular, exciton dissociation.

Based on the density matrix formalism in second quantization, we make use of the many-particle Hamilton operator and Heisenberg’s equation of motion to describe the dynamics of incoherent excitons \( N^e_Q \) accounting for the annihilation (creation) of an exciton in the state \( \nu \) with momentum \( Q \). The Hamilton operator including electron–electron, electron–phonon, and electron–photon interactions is transformed into an excitonic basis resulting in an excitonic Hamiltonian. The excitonic Hamilton operator describes the system well for low electron–hole pair densities in an undoped system. Note that doping can result in the appearance of trion features in optical spectra, and large doping levels can cause energy renormalization due to many-particle dielectric screening. We focus on low-density conditions, where exciton–exciton interaction is negligible. The central part of our work lies in the exciton–phonon interaction that drives exciton thermalization and dissociation. In general, the exciton–phonon Hamiltonian describing scattering from \( \nu, Q \) to \( \nu', Q' + q \) has the form

\[ H_{e-ph} = \sum_{\nu, Q} G^{\nu}_{Q} X^\dagger_{Q' + q} \left( b^\dagger_{q} + b^\dagger_{q'} \right), \]

(1)

where we have introduced the annihilation (creation) operator \( b^\dagger_{q} \) for phonons with the mode \( j \) and the momentum \( q \). The exciton–phonon matrix element \( G^{\nu}_{Q} = \mathcal{F}^{\nu}_{Q} - \mathcal{F}^{\nu}_{Q + q} \) contains the excitonic form factor \( \mathcal{F}^{\nu}_{Q} = \sum_k \phi_k^\dagger \phi_{k+q} \), the electron–phonon coupling strength \( \mathcal{F}_{Q} \), and the factor \( \alpha_\nu = m_\nu / (m_e + m_h) \).

The excitonic wavefunctions \( \phi_k^\dagger \) are obtained by solving the Wannier equation with a thin-film Coulomb potential. We describe excitonic states with positive binding energy (free states) by orthogonal plane waves (OPW). Free states form a continuum in which the quantum number \( \nu \) becomes a continuous momentum \( p \) with the energy \( E^{\nu}_{Q} = E^{\nu}_{Q} + \hbar^2 q^2 / 2M + \hbar q^2 / 2m_e \) (in an effective mass approximation with \( m_e \) being the reduced mass and \( M \) the total mass). The OPWs describing the states in the continuum can be written as \( \psi^\dagger_{k,\nu} = \delta_{pk} - \sum_{\nu'} \phi^\dagger_{k,\nu'} \phi^\dagger_{k,\nu} \), where \( N^e_{\nu} \) is the number of bound states. While the first term corresponds to the plane wave description of continuum states, the second term accounts for the orthogonalization of these states with respect to the bound states in order to avoid unphysical wavefunction overlaps, cf. the ESL.

Using Heisenberg’s equation of motion we obtain an expression for momentum- and time-resolved exciton occupation:

\[ N^e_Q(t) = \sum_{\nu} \int_0^t \mathrm{d}t' |\mathcal{F}^{\nu}_{Q}|^2 \delta_{Q,\nu} \mathcal{R}^{\text{in}}_{\nu} N^e_Q(t') + \mathcal{R}^{\text{out}}_{\nu} N^e_Q(t') \]

(2)

The first line in eqn (2) describes the interaction with light. Here, the first term accounts for the absorption of incident
light expressed with the microscopic polarization $p'(t)$, while the second term describes the loss of exciton occupation via radiative recombination.\textsuperscript{16} A more thorough description of exciton–light interaction can be found in the ESI.\textsuperscript{†} The second line in eqn (2) accounts for exciton–phonon scattering within the second-order Born–Markov approximation.\textsuperscript{16,37–39} The first two terms describe bound-to-bound transitions with the in- and out-scattering rates $\Gamma_{i,O}^\text{in} = \sum \frac{N}{\hbar q} G_q^\text{in} N_q^\text{out}$ and $\Gamma_{i,O}^\text{out} = \sum \frac{N}{\hbar q} G_q^\text{out} N_q^\text{in}$. The appearing scattering matrix reads $G_q^\text{in} = \frac{2\pi}{\hbar} \sum_j \langle n_j^+ \mid \eta_q^j \delta(E_q - E_j + \hbar Q_q) \rangle$. Here, we have introduced the abbreviation $\eta_q^j = \eta_q + \frac{1}{2}$ with $\eta_q = \langle \frac{1}{2} \rangle$ being the phonon number. Furthermore, $q = Q - Q$ is the momentum exchange, $\pm$ accounts for the emission (+) and absorption (−) of phonons, and $E_q$ and $\hbar Q_q$ are the exciton and phonon energies, respectively. While the first two terms in the second line of eqn (2) account for scattering within bound exciton states, the last two terms account for the formation/dissociation of excitons from/to free electron–hole pairs with the rates $\Gamma_{i,O}^\text{form} = \sum \frac{N}{\hbar q} G_q^\text{in} N_q^\text{out}$ and $\Gamma_{i,O}^\text{dis} = \sum \frac{N}{\hbar q} G_q^\text{out} N_q^\text{in}$ including a sum over all scattering possibilities.

We now make the assumption that free carriers are immediately extracted from the system after dissociation due to a finite electric field resulting in a zero contribution from the exciton formation term. The electric field is assumed to be weak enough, so that field-induced dissociation is negligibly small. Previous experiments suggest that this is the case for electric fields below 10–13 V µm$^{-1}$ in hBN-encapsulated WS$_2$ monolayers.\textsuperscript{35} Here, we focus on the microscopic origin of dissociation and its implications on the fundamental limits of real devices. Therefore, we assume a ballistic transport and a complete collection at the leads, so that the current is determined by the dissociation. The latter poses an upper fundamental limit for the maximum photocurrent one can obtain in a real device. Furthermore, in addition to the discussed terms appearing in eqn (2), we also include a phenomenological decay $N_q^\text{dis} = -\tau_{\text{dis}} N_q$ to account for non-radiative recombination, which can be significant in real samples due to defects\textsuperscript{24} and is influenced by doping.\textsuperscript{46} The decay time $\tau_{\text{dis}}$ has been experimentally and theoretically determined to range from few to hundreds of picoseconds.\textsuperscript{20,24,35,49}

In this study, we discuss exciton dissociation in the four most studied semiconducting TMDs including WS$_2$, WS$_3$, MoSe$_2$, and MoS$_2$. However, for now, we focus on hBN-encapsulated WS$_2$ monolayers due to their well established excitonic landscape containing bright (KK) and momentum-dark (KA and KK) excitons.\textsuperscript{15,19,23,50,51} We consider up to 7 bound s-states for each valley, assigning states with higher main quantum number to the continuum and disregarding states with non-zero angular momentum due to their lower exciton–phonon cross section.\textsuperscript{16} Moreover, we exclude the influence of spin-dark states, since the timescale of spin-flip scattering processes is considerably slow compared to the spin-conserving ones.\textsuperscript{19,52} We use input parameters for the bandstructure and electron–phonon coupling from \textit{ab initio} studies,\textsuperscript{53–56} cf. the ESI.\textsuperscript{†}

$\frac{\hbar q}{\pi}$

3. Results and discussion

3.1. Dissociation dynamics

Now, we resolve the exciton dissociation dynamics in hBN-encapsulated WSe$_2$ monolayers. We solve eqn (2) considering a continuous wave optical excitation resonant to the 1s exciton with a rise time of 1 ps and a power density of 1 µW µm$^{-2}$. In order to account for a realistic sample with defects, we set the non-radiative decay time $\tau_{\text{nr}} = 100$ ps, similar to experimentally reported values.\textsuperscript{20} We define the dissociation current density $i_d = e_0 A^\text{in} \sum_{\hbar q} \Gamma_{\text{dis}} N_q^\text{in}$ and disentangle the contributions from different channels, \textit{i.e.} scattering from the state $\nu$ to the continuum of a given valley (cf. Fig. 2(a)). We identify the most important dissociation channel to be KA$\rightarrow$KK$\text{cont.}$ followed by KK$\rightarrow$KK$\text{cont.}$. There is also a number of other dissociation channels with minor contributions that however sum up to account for a significant portion of the total current. Furthermore, we find that the dissociation current increases on a timescale of 50–100 ps, corresponding to $\tau^{-1} = \tau_{\text{dis}}^{-1} + \tau_{\text{nr}}^{-1}$. This is in agreement with the fact that the response time is dominated by the shortest decay time, \textit{cf.} the ESI.\textsuperscript{†} In the considered case, both dissociation and non-radiative decay show characteristic lifetimes of 100 ps, and hence both dominate the response. Although radiative recombination is very fast (below 1 ps), radiative decay of excitons is rather slow (1–100 ns) because only excitons in the light cone ($Q \approx 0$) can recombine\textsuperscript{13} and hence radiative decay does not influence the response time.

The efficiency of the dissociation channel KA$\rightarrow$KK$\text{cont.}$ is a result of the strong coupling of excitons with M phonons,\textsuperscript{55} leading to a strong intervalley scattering from KA to KK$'$ excitons. This is manifested in high rates $\Gamma_{\text{dis}}$ for dissociation from KA$\rightarrow$KK$\text{cont.}$ compared to other exciton states with a similar energy (\textit{cf.} Fig. 2(b)). Moreover, the importance of KA$\rightarrow$KK$\text{cont.}$ results from the interplay between occupation and dissociation rates: while lower states are largely occupied (\textit{cf.} Fig. 2(e)), higher states exhibit faster dissociation rates due to their proximity to the continuum (\textit{cf.} Fig. 2(b)). Such faster dissociation for higher excitonic states has been already observed in carbon nanotubes.\textsuperscript{57,58} We also determine the main exciton pathway, starting from the optical excitation at KK$\text{cont.}$ and ending with the main dissociation channel KA$\rightarrow$KK$\text{cont.}$, identified above (\textit{cf.} arrows in Fig. 1). For this purpose, we calculate the net scattering rates from $\nu$ to $\mu$ with $\Gamma_{\nu} = A^{-1} \sum_{\hbar q} (\Gamma_{\nu} N_q^\text{in} - \Gamma_{\nu} N_q^\text{out}\nu_{\text{dis}})$, as illustrated in Fig. 2(c) and tabulated in the ESI.\textsuperscript{†} We show that most excitons follow the...
following path: after optical excitation of KK1s excitons, they scatter to KΛ1s states followed by KΛ1s → KK′2s and KK′2s → KΛ2s transitions, and finally they dissociate through KΛ2s → KK_cont. Thus, not only is the dissociation dominated by scattering with M phonons, but also the transitions within bound states, which occur mostly between KK′ and KΛ states.

Exploiting the microscopic character of our theoretical approach, we can resolve the dynamics of excitons in energy, momentum and time. At short times during the start of the optical excitation (t = 0.1 ps, Fig. 2(d)), a small occupation has already been generated at the 1s level of all valleys as a result of ultrafast polarization-to-population transfer and inter-valley scattering, with sharp features appearing due to scattering with optical Γ phonons and intervalley Λ and M phonons.15,16 Higher KK states (2s, 3s) show minor occupation due to off-resonant excitation. At later times, when the stationary state has nearly been reached (t = 200 ps, Fig. 2(e)), the occupation shows a more thermalized distribution, with larger occupation in higher exciton states compared to earlier times, and a strong occupation in the light cone at Q = 0 due to the continuous optical excitation. The equilibration between bound and unbound states described here is opposite to the relaxation cascade16 but occurs on the same timescale of a few picoseconds. This timescale would govern the photoresponse in systems pumped by an ultra-short femtosecond pulse.

Finally, we compare the exciton occupation at t = 200 ps with the case where dissociation has not been considered (NQnd). In order to be able to compare the occupation in these two cases, we normalize them by the total density n, i.e. NQ̃ = NQ / n. We compute the relative difference (NQ − NQnd)/NQnd and find that higher states are clearly less occupied when the effect of dissociation is taken into account (cf. Fig. 2(f)). The lower occupation of higher states is a direct consequence of their efficient dissociation (cf. Fig. 2(b)).

3.2. Quantum efficiency and response time
The rise time of the dissociation current is limited by the fastest decay mechanism. An efficient dissociation can hence pose a fundamental limit on how fast the response time can be in a real device. Moreover, the magnitude of the dissociation current represents as well a fundamental limit for the internal (IQE) and external (EQE) quantum efficiencies. In real devices, the response time will be longer and the quantum efficiency will be lower because of imperfect charge transport and collection at the leads. In this work, we show the optimal
efficiency and response time that can be reached in 2D material based photodetectors.

In the last section, we have resolved and understood the microscopic channels responsible for exciton dissociation. Here, we investigate their impact on key quantities for a photodetector: EQE, IQE and the response time. EQE and IQE describe the efficiency of dissociation with respect to the rate of incident photons \( \Phi = \frac{I}{\hbar \omega} \) and of optically generated excitons \( G \), respectively, and are directly related through the optical absorption \( a(\omega) \), i.e. EQE = \( a(\omega) \)IQE. Moreover, the response time \( \tau_{\text{rp}} \) is determined by the dominant decay time, which can be non-radiative decay (\( \tau_{\text{nr}} \)), dissociation (\( \tau_{\text{d}} \)), or radiative decay (\( \tau_{\text{r}} \)). Thus, we can write

\[
\text{IQE} = \frac{\dot{n}_d}{G}, \quad \text{EQE} = \frac{\dot{n}_d \hbar \omega / I, \quad \tau_{\text{rp}} = \sum_i \tau_i^{-1}}.
\]

where \( \omega \) and \( I \) are the frequency and power density of the incident light, respectively, and \( i = \text{nr, d, r} \). From our microscopic model, we obtain the exciton generation rate \( G = A^{-1} \sum_{i=Q} \Gamma_{iQ} |p_i(t)|^2 = \frac{I}{\hbar \omega} a(\omega) \) and the dissociation rate

\[
\dot{n}_d = A^{-1} \sum_{i=Q} \Gamma_{iQ} N_{iQ}.
\]

In the following, we reveal how quantum efficiency and response time can be controlled and optimized through externally accessible knobs, such as excitation energy, non-radiative decay time, substrate screening, temperature and strain. We model the effect of defects by investigating two cases of non-radiative decay time: \( \tau_{\text{nr}} \to \infty \) and \( \tau_{\text{nr}} = 100 \) ps, accounting for a pristine defect-free sample and a more realistic sample with defects,\(^{20}\) respectively. Unless otherwise stated, the excitation energy is centered at the 1s exciton resonance, the dielectric environment corresponds to hBN encapsulation, the temperature is 300 K and the sample is unstrained. We solve eqn (2) in the stationary state, setting \( N_i = 0 \), and use the calculated exciton occupations to determine the EQE, IQE, and the response time \( \tau_{\text{rp}} \) in eqn (3).

In a pristine defect-free sample, dissociation is the fastest decay mechanism and hence all optically generated excitons dissociate, yielding an IQE of 100%. The EQE follows the absorption spectrum for varying excitation energies (cf. blue solid line in Fig. 3(a)), with maxima at the exciton resonance energies (1s, 2s, 3s). When including a realistic non-radiative recombination time of 100 ps, the IQE drops down to 50%, and the EQE decreases by the same factor (cf. red lines in Fig. 3(a)). The IQE is constant for a wide range of energies where the 1s exciton is excited. As the excitation energy approaches the 2s resonance, excitons in this state are excited as well. Since dissociation from higher states is more efficient (cf. Fig. 2(b)), the IQE increases with excitation energy, though showing again a plateau around 2s and 3s resonances (cf. the red dashed line in Fig. 3(a)). The most optimal performance is thus obtained by exciting at excitonic resonances, where the optical absorption is maximized, with a trade-off between lower absorption and faster dissociation at higher states. Note that above-bandgap excitation would result in a large fraction of photoexcited free electron–hole pairs,\(^{59,60}\) giving rise to an efficient IQE. Moreover, we expect the large absorption of the B and C exciton absorption peaks above the bandgap to boost the EQE as well.

The effect of defects can be studied by varying the non-radiative decay time \( \tau_{\text{nr}} \) (cf. Fig. 3(b)). When non-radiative recombination is much faster than dissociation, most excitons are lost via this decay channel and the IQE is 0%. As \( \tau_{\text{nr}} \) progressively approaches the dissociation time \( \tau_{\text{dis}} \approx 100 \) ps, dissociation starts to take over and the IQE increases. At even longer \( \tau_{\text{nr}} > 10^5 \) ps dissociation dominates yielding an IQE of 100%. Nevertheless, there is a trade-off between efficiency and response time. For short \( \tau_{\text{nr}} \), the response time will be dominated by this value and can be tuned, while the quantum efficiency becomes low (cf. inset in Fig. 4(a)). Remarkably, similar response times on the order of 100 ps have been also reported in experimental studies for the same system at low electric fields in ref. 35, where \( \tau_{\text{nr}} \sim 300 \) ps. It was shown that the measured response times deviate from the field-assisted tunneling model at low electric fields, showing saturation at values around 100 ps. This deviation is consistent with our phonon-induced dissociation model and indicates that exciton–phonon scattering indeed limits the photoresponse at...
weak fields. Our theoretical predictions thus have important implications on the limitations of real TMD-based optoelectronic devices.

Another experimentally accessible knob is dielectric engineering via substrate screening that influences the material’s characteristics in many different ways. First, stronger screening results in lower exciton binding energies, i.e., the excitonic levels are closer to each other and to the continuum, leading to a more efficient exciton dissociation. This is manifested in a significant increase of IQE and EQE (cf. Fig. 3(c)) and a decrease of the response time (Fig. 4(a)). On the other hand, screening also affects the absorption and thus the EQE. The optical absorption is directly proportional to the inverse dielectric function, $\alpha \propto \varepsilon_0^{-1}$. Moreover, the oscillator strength is proportional to the exciton probability $|\psi(r=0)|^2$, which also decreases for stronger screening due to more delocalized excitonic wavefunctions. Both facts contribute to a reduction of the absorption, resulting in the decrease of the EQE at large dielectric screening constants, cf. Fig. 3(c). The reduction of the oscillator strength also weakens radiative recombination, causing the response time to slightly increase for low $\varepsilon_s$ in a pristine sample (cf. blue line in Fig. 4(a)). For weak screening, dissociation is so inefficient that the main decay channel is radiative recombination in a defect-free sample. Including a finite $\tau_{nr}$ of 100 ps results in an overall reduction of the efficiency (cf. red lines in Fig. 3(c)). However, while dissociation is very inefficient on a SiO$_2$ substrate ($\tau_d \sim 4$ ns), the IQE increases to approximately 50% on an hBN-encapsulated sample and the dissociation can compete with non-radiative recombination. The most optimal quantum efficiency is thus obtained for substrates with a stronger screening (such as hBN) giving rise to a more efficient exciton dissociation.

A similar behaviour is observed when exploring temperature as a potential knob to tune the performance of a photo-detector. As intra- and intervalley scattering of excitons is governed by phonons in the considered low-excitation regime, the strength of these scattering channels strongly depends on temperature.11,13,61,62 Here, the boost in the IQE (Fig. 3(d)) and the drop in response time (Fig. 4(b)) are caused by an increase in the phonon number $n_q^f$ with temperature. As more phonons are available, exciton–phonon scattering becomes stronger, populating higher excitonic states and increasing the dissociation efficiency. Note that at very low temperatures there is a 0% plateau in the IQE and EQE at which radiative recombination is more efficient than dissociation. While the IQE displays a monotonic increase with temperature, the EQE reaches a maximum and then decreases due to lower absorption at higher temperatures. The increase in the phonon-induced linewidth of excitonic resonances11 results in a broader absorption. Since the oscillator strength is conserved, the absorption peak decreases causing a clear reduction of the EQE at the exciton resonance (cf. solid lines in Fig. 3(d)). The inclusion of a finite $\tau_{nr}$ shifts the IQE plateau to higher temperatures, at which exciton dissociation can compete with non-radiative decay, giving rise to an overall decrease of the EQE at low temperatures (cf. red lines in Fig. 3(d)). Overall, the optimal quantum efficiency is found for temperatures slightly above room temperature for a realistic finite non-radiative decay.

Finally, we study the effect of strain, which is known to shift the electronic K and \( \Lambda \) valleys in opposite directions with respect to the valence band.56 These shifts have a direct impact on the linewidth and thus on the EQE of a pristine sample. For increasing tensile strain at approximately +1%, the $K_{1s}$ state approaches the $K\bar{K}_{1s}$ state, suppressing the scattering channel $K\bar{K}_{1s} \rightarrow K\bar{\Lambda}_{1s}$ induced by phonon emission. Similarly, further increasing strain towards +2% results in a shift of $K\Lambda_{1s}$ above $K\bar{K}_{1s}$ and the inhibition of the intervalley scattering induced by phonon absorption. As a result, the exciton linewidth decreases and the absorption peak increases with tensile strain resulting in a step-wise increase of the EQE whenever a scattering channel is blocked (cf. solid blue line in Fig. 4(e)). The slight increase in the EQE at negative (compressive) strain can be traced back to a larger spectral distance between the KK and KA excitons, which shifts the final state $|\mu,q\rangle$ to larger momenta, resulting in a smaller wavefunction overlap $F_q^{\mu}$ for the intervalley scattering. In order to understand the effect of strain in a more realistic sample with defects ($\tau_{nr} = 100$ ps), we study the changes in the IQE, which reflects the efficiency of dissociation with respect to the non-radiative decay. For tensile strain, the $\Lambda$ valley is lifted up so that the $K\Lambda_{1s}$ state lies closer to the continuum of the KK valley. This optimal situation favours the dissociation of excitons from $K\bar{A}_{1s}$ to $K\bar{K}_{cont}$, yielding an enhancement of the IQE (cf. dashed red line in Fig. 4(e)). In contrast, when compressive strain is applied, the $K$ valley shifts down. As a result, $K\Lambda_{1s}$ has by far the largest occupation. Since this state moves further away from $K\bar{K}_{cont}$ as compressive strain increases, the dissociation becomes less efficient and the IQE decreases.

Considering realistic values for non-radiative decay time, the quantum efficiency for exciton dissociation is the highest and the response time the fastest at the following conditions: temperatures slightly above 300 K, substrates with strong dielectric screening, excitation energies resonant to exciton states, and tensile strain values around 1%. For short non-radiative decay times, there is a trade-off between a fast response time and a low quantum efficiency.
3.3. Comparison of TMD materials

So far, we have focused on exciton dissociation in a WSe$_2$ monolayer encapsulated in hBN. Now, we compare EQE, IQE and the response time for the four most studied semiconducting TMDs including WSe$_2$, WS$_2$, MoSe$_2$, and MoS$_2$ placed on the two most common dielectric environments (SiO$_2$ substrate and hBN encapsulation), cf. Fig. 5. The highest EQE in a defect-free monolayer (i.e. $\tau_{nr} \to \infty$, cf. light colors in Fig. 5) on a SiO$_2$ substrate is found for WS$_2$ (23%) due to its large optical absorption,$^{63}$ followed by WSe$_2$ (8%), MoSe$_2$ (5%), and MoS$_2$ (2%). Since here the dissociation is faster than the radiative decay, the IQE is close to 100% in the absence of non-radiative recombination. Despite the relatively large EQE and IQE values, exciton dissociation in TMDs on a SiO$_2$ substrate is very inefficient as a result of the large binding energies and leads to very long response times ranging from few to tens of nanoseconds. Due to such inefficiency, including a non-radiative decay time of 100 ps completely suppresses dissociation, resulting in quantum efficiencies much below 1%, which are barely visible in Fig. 5(a) and (b). Under hBN encapsulation, the excitonic oscillator strength decreases resulting in a weaker optical absorption and hence in lower EQE values. In MoS$_2$, however, this effect is countered by the suppression of the intervalley $\Gamma K \to \Gamma K'$ dephasing channel via phonon emission. This is due to a lower energetic separation between the two valleys, resulting in a much smaller linewidth and hence a larger absorption and EQE. In hBN encapsulated samples, exciton dissociation is much more efficient because of the smaller binding energies and yields response times on the order of 100 ps. Including a realistic non-radiative decay of 100 ps causes the EQE and IQE to decrease to moderate values around 2–6% and 30–60% respectively, including a modest improvement of the response time to values below 100 ps. We note that sulfur-based TMDs typically show larger values of intrinsic doping than selenium-based TMDs, hence WS$_2$ and MoS$_2$ samples usually exhibit faster non-radiative recombination channels than WSe$_2$ and MoSe$_2$ ones.$^{48}$

In order to understand why dissociation is more efficient in some TMDs, we need to consider their complex excitonic landscape. The long response time in pristine MoS$_2$ on any substrate compared to the other TMDs is a consequence of the slow exciton dissociation in this material (cf. Fig. 5(c)). This can be traced back to the fact that the largest occupation is found in $\Gamma K$ and $\Gamma K'$ excitons exhibiting a large effective mass and hence large exciton binding energies. Thus, it is difficult to scatter into higher exciton states, which are energetically far away. Note however that the excitonic landscape and in particular the question which states is the lowest is still under debate in the case of MoS$_2$.$^{12,64-66}$

The TMD material with the fastest dissociation on a SiO$_2$ substrate is WSe$_2$ (cf. light blue boxes in Fig. 5(c)) due to its optimal band structure favouring the dissociation channel $K\Lambda_{2s} \to K\Theta$ cont.. Under hBN encapsulation, however, it is MoSe$_2$ that shows the fastest dissociation due to the proximity of $K\Lambda_{1s}$ excitons to the $K\Gamma$ continuum. In both dielectric environments, the dissociation in WS$_2$ is not as fast as in WSe$_2$ and MoSe$_2$. While the band structure of WS$_2$ is similar to the one of WSe$_2$, the main state for dissociation $K\Lambda_{2s}$ is too high in energy exhibiting a very low occupation and thus giving rise to a longer dissociation time. Note that the dielectric environment and in particular hBN encapsulation will also modify the transport properties affecting the performance of a real device. Our calculations provide an upper fundamental limit of the performance limited by dissociation.

4. Conclusions

Using a microscopic approach, we have resolved the complex many-particle processes behind phonon-assisted exciton dissociation in TMD monolayers exhibiting a multi-valley exciton landscape. Focusing on WSe$_2$, we have discerned the distinct transition channels dominating exciton dissociation, which involve phonon-driven intervalley scattering from $K\Lambda$ to $K\Gamma$ excitons. Due to the trade-off between lower occupation and larger dissociation rates at higher exciton states, most of the dissociation arises from the $2s$ exciton of the $K\Lambda$ valley. Furthermore, we have determined fundamental limits for the efficiency and response time of TMD-based optoelectronic devices. In particular, we find a dissociation-limited response time of 100 ps for hBN-encapsulated WSe$_2$ monolayer, which is similar to experimental findings. Moreover, we have studied the tunability of key quantities with externally accessible knobs such as excitation energy, dielectric engineering, temperature and strain on a microscopic footing. We find a trade-

![Fig. 5 Comparison of TMD materials. (a) External quantum efficiency (EQE), (b) internal quantum efficiency (IQE) and (c) response time for four TMDs on SiO$_2$ substrate (blue) and in hBN encapsulation (red) studied for two non-radiative decay times $\tau_{nr} = 100$ ps (dark color) and $\tau_{nr} \to \infty$ (light color).](image-url)
off between faster response time and lower EQE for increasing substrate screening, temperature, non-radiative decay rate, and excitation energy.

**Note after first publication**

This article replaces the version published on 13 January 2021, which contained an incorrect version of Equation 1. This error was only present in the article PDF, and therefore the HTML version of the article remains unchanged. Equation 1 has been updated in the current article PDF to address any confusion the discrepancy may have caused.

**Conflicts of interest**

There are no conflicts to declare.

**Acknowledgements**

This project has received funding from the Swedish Research Council (VR, project number 2018-00734), the European Union’s Horizon 2020 research and innovation programme under grant agreement no. 881603 (Graphene Flagship), and Chalmers’ Excellence Initiative Nano under its excellence PhD programme. The computations were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at C3SE partially funded by the Swedish Research Council through grant agreement no. 2016-07213. We thank Thomas Müller (TU Wien) and Roberto Rosati (Chalmers) for fruitful discussions.

**References**

64 T. Deilmann and K. S. Thygesen, 2D Mater., 2019, 6, 035003.
66 M. Feierabend, S. Brem, A. Ekman and E. Malic, 2D Mater., 2021, 8, 015113.