This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal’s standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.
Exciton Dynamics and Annihilation in WS$_2$ 2D Semiconductors

Long Yuan$^a$ and Libai Huang$^a$*

Department of Chemistry, Purdue University, West Lafayette, Indiana 47907

* Corresponding author, email: libai-huang@purdue.edu
Abstract

We systematically investigate the exciton dynamics in monolayer, bilayer, and trilayer WS$_2$ two-dimensional (2D) crystals by time-resolved photoluminescence (TRPL) spectroscopy. The exciton lifetime when free of exciton annihilation is determined to be $806 \pm 37$ ps, $401 \pm 25$ ps, and $332 \pm 19$ ps for WS$_2$ monolayer, bilayer, and trilayer, respectively. By measuring the fluorescence quantum yields, we also establish the radiative and nonradiative lifetimes of the direct and indirect excitons. The exciton decay in monolayer WS$_2$ exhibits a strong excitation density-dependence, which can be described using an exciton-exciton annihilation (two-particle Auger recombination) model. The exciton-exciton annihilation rate for monolayer, bilayer, and trilayer WS$_2$ is determined to be $0.41 \pm 0.02$, $(6.00 \pm 1.09) \times 10^{-3}$ and $(1.88 \pm 0.47) \times 10^{-3}$ cm$^2$/s, respectively. Notably, exciton-exciton annihilation rate is two orders of magnitude faster in monolayer than in bilayer and trilayer. We attribute the much slower exciton-exciton annihilation rate in bilayer and trilayer to reduced many-body interaction and phonon-assisted exciton-exciton annihilation of indirect excitons.

**Keyword:** Atomic layers, 2D exciton, many-body effect, exciton diffusion, phonon-assisted exciton-exciton annihilation
Introduction

Semiconducting atomically thin layers of transition metal dichalcogenides (TMDs) such as MoS$_2$, MoSe$_2$, WS$_2$, and WSe$_2$ have attracted much research interests due to their unique electronic structures and optical properties.$^{1, 2}$ These properties lead to potential applications in optoelectronics and electronics,$^{3, 4}$ including field-effect transistors,$^{5-9}$ atomically thin photovoltaics device$^{10-13}$ and ultrasensitive photodetectors.$^{14, 15}$

One of the unique properties in these atomically-thin 2D semiconductors is the indirect to direct bandgap transition and the extraordinarily large exciton binding energy at the monolayer limit.$^{16, 17}$ The atomically-thin nature of monolayer also leads to a strong enhancement of the Coulomb interaction between the electron and the hole. As a result, the bound electron-hole pairs, or known as excitons, dominate the optical and electrical properties of these materials. Recent theoretical calculations and experimental measurements showed that the exciton binding energy is on the range of 0.3-1.0 eV for the TMDs monolayers,$^{18-23}$ an order of magnitude larger than other previously investigated 2D excitonic structures, such as quantum well. Such a large exciton binding energy results in strongly bound exciton at room temperature, providing an ideal platform to study exciton behaviors in 2D systems. Despite exciton dynamics have been investigated in these 2D semiconductor,$^{24, 25}$ large variation in sample qualities due to different preparation methods have prevented intrinsic radiative and nonradiative lifetimes of exciton to be conclusively established.
Another hallmark of low-dimensional electronic systems is the enhanced many-body interaction due to reduced dimensionality. Upon the generation of a high density of electrons and holes, many-body scattering processes such as Auger recombination and exciton-exciton annihilation play an important role in nonradiative relaxation. These nonradiative recombination processes define the upper limit of excitation density and ultimately the efficiency for applications such as semiconductor lasers and light-emitting diodes. Exciton-exciton annihilation and Auger recombination has been intensively investigated in quantum dots,\textsuperscript{26, 27} carbon nanotube\textsuperscript{28, 29} and semiconductor nanowires.\textsuperscript{30-32} While recent works on MoS\textsubscript{2}, MoSe\textsubscript{2} and WSe\textsubscript{2} monolayers showed the existence of exciton-exciton annihilation at high excitation density,\textsuperscript{33-38} how quantum confinement of 2D exciton impacts many-body exciton interaction is still elusive.

In this paper, we investigate exciton dynamics and many-body exciton interaction in monolayer, bilayer, and trilayer exfoliated WS\textsubscript{2}. We choose WS\textsubscript{2} as a model system because the relatively low defect density in WS\textsubscript{2} as manifested by the higher photoluminescence (PL) quantum yield (QY) than other 2D semiconductors (~6% in WS\textsubscript{2}, compared to ~0.1% of MoS\textsubscript{2}). We measure exciton dynamics in both low and high exciton density regimes. Our results demonstrate that exciton-exciton annihilation dynamics in monolayer are drastically different from those in bilayer and trilayer as a result of enhanced many-body interaction and different annihilation mechanisms for direct and indirect excitons.
Results and Discussion

We identify the number of layers in exfoliated WS$_2$ flakes by optical microscopy and then verify using Raman and PL spectroscopy (Figures 1 and 2).\textsuperscript{39} To make sure the PL and Raman measurements are performed on the same flake, the substrate is marked with photolithography marker as described in our previous work.\textsuperscript{40} We have also performed PL and Raman measurements on different locations of the same flake and found no significant difference between different locations (Fig. S1 in the SI). We observe two characteristic Raman modes corresponding to the E$_{2g}^{1}$ (in-plane vibration) and A$_{1g}$ (out-of-plane vibration) consistent with previous reports. The frequency of these two modes is around 352 cm$^{-1}$ and 418 cm$^{-1}$, respectively. The frequency difference is thickness dependent and can be used to determine the number of layers.\textsuperscript{39} We obtain a frequency difference between E$_{2g}^{1}$ and A$_{1g}$ modes of 65.8 cm$^{-1}$, 67.1 cm$^{-1}$, and 68.1 cm$^{-1}$ for monolayer, bilayer, and trilayer, respectively.\textsuperscript{39}

The photoluminescence PL spectra of monolayer, bilayer, and trilayer WS$_2$ are shown in Figure 2(a). The monolayer shows a much higher PL intensity than that of the bilayer and trilayer, consistent with the indirect to direct exciton transition at the monolayer limit.\textsuperscript{16, 17, 41} The PL quantum yield (QY) was measured to be $\sim$ 6%, $1 \times 10^{-3}$ and $4 \times 10^{-4}$ for monolayer, bilayer and trilayer, respectively, and the details for the measurements of the QY are presented in the supporting information (SI). A single sharp peak corresponding to the A exciton resonance
at ~ 615 nm (2.02 eV) was observed for the monolayer. The PL in monolayer can be understood as originating from the direct exciton transition between the conduction band minimum (CBM) and the valance band maximum (VBM) both at K point. In contrast, the PL spectra of the bilayer and trilayer show additional broad and lower energy emission that corresponding to indirect exciton where VBM remains at K point but the CBM is between the K and Γ point (schematically shown in Figure 4).
Figure 1. (a) Optical image of WS\(_2\) monolayer, bilayer, and trilayer flake on Si wafer with 90 nm oxide thickness; (b) Raman spectra of WS\(_2\) monolayer, bilayer, and trilayer flake. The scale bar is 5 µm. The dashed lines mark the position of the \(E_{2g}^1\) and \(A_{1g}\) positions for the monolayer.

Figure 2. (a) PL spectra of monolayer, bilayer, and trilayer WS\(_2\); (b) PL decay without exciton-exciton annihilation for WS\(_2\) monolayer, bilayer, and trilayer at pump intensity of 5 nJ/cm\(^2\), 0.1 µJ/cm\(^2\), and 0.4 µJ/cm\(^2\), respectively. The photoluminescence lifetime measurements integrate the entire PL spectra from 500 nm to 850 nm. The black line is the instrument response function (IRF). The red lines are the fitting curves using a single exponential decay convoluted with a Gaussian response function.

Because photoluminescence observed was emitted from excitonic states and free carriers do not contribute to the emission, the dynamics discussed in this article reflect dominantly exciton behaviors. For WS\(_2\), trion binding energy is \(\sim 20 \text{ meV}^{43, 44}\) slightly smaller than \(k_BT\) at room temperature, which implies that trions can be dissociated by thermal energy. We do not observe significant trion contribution in the PL spectra and therefore conclude that the emission observed in our experiment is dominated by neutral excitons. Temperature dependent measurements will be very informative to elucidate exciton and trion dynamics, which are currently underway.
We first establish the PL decay of monolayer, bilayer, and trilayer at low excitation density where exciton-exciton annihilation effects are negligible, and the results shown in Figure 2(b) (also see Fig. S2). The photoluminescence lifetime measurements integrate the entire PL spectra from 500 nm to 850 nm. Interestingly, PL lifetime shows negligible wavelength dependence (Fig. S3). For all samples, the PL decays at low excitation density can be fitted satisfactorily with a single exponential decay convoluted with a Gaussian instrument response function (IRF) that is described by $I(t) = \int_{-\infty}^{t} IRF(t') e^{-\frac{t-t'}{\tau}} dt$. For the data presented in Figure 2, we obtain a PL lifetime of 856 ± 3 ps, 435 ± 1 ps, and 305 ± 2 ps for WS$_2$ monolayer, bilayer, and trilayer, respectively. Please note that uncertainties reported in the PL lifetimes are the uncertainties from the fitting of the data. We have also performed measurements at different locations of the same thickness (Fig. S4 in the SI). PL lifetimes vary slightly from position to position. Such variation is probably related to the difference in the local environments. The average PL lifetimes for monolayer, bilayer, and trilayer are found to be 806 ± 37 ps, 401 ± 25 ps, and 332 ± 19 ps.

The single exponential of hundreds of ps PL decay is in direct contrast with complex multi-exponential decay observed for MoS$_2$. Single exponential decay behavior implies that PL originates from a single state, indicative that other energy levels such as trap states induced by defects are not as significant in WS$_2$ in comparison to MoS$_2$. Lower defect density
is consistent with much higher PL quantum yield (~6%) of monolayer WS$_2$ than that of monolayer MoS$_2$ (~10$^{-3}$).

We can relate the observed PL lifetime, $\tau_{ob}$, to the radiative and nonradiative lifetimes, $\tau_r$ and $\tau_{nr}$, respectively by the equation:

$$\frac{1}{\tau_{ob}} = \frac{1}{\tau_r} + \frac{1}{\tau_{nr}}$$

and these lifetimes define the PL quantum yield,

$$\Phi_n = \frac{\tau_{ob}}{\tau_r}$$

Using a QY of ~6% and $\tau_{ob}$ of ~806 ps, the radiative and nonradiative lifetimes are determined to be ~13 ns and ~900 ps respectively for monolayer WS$_2$. In contrast, the radiative lifetime is determined to be ~400 ns and ~830 ns for the bilayer (QY of 1 × 10$^{-3}$, $\tau_{ob}$ of ~401 ps) and trilayer (QY of 4 × 10$^{-4}$, $\tau_{ob}$ of ~332 ps) respectively. The much longer radiative lifetime in bilayer and trilayer again is consistent with the indirect nature of the excitons. The nonradiative processes dominate the exciton decay and lifetimes for the bilayer and trilayer are estimated to be ~400 ps and ~300 ps respectively. Interestingly, while the nonradiative lifetime is on the same order for monolayer, bilayer and trilayer, it decreases as the number of layer increases. We speculate that the nonradiative pathways are similar in nature for monolayer, bilayer and trilayer. One likely candidate is electron-phonon scattering. The thickness dependence of the
nonradiative lifetime is likely to be related to change in dielectric environments. Further experiments will be needed to determine the exact nature of the nonradiative processes.

Next, we investigate exciton dynamics in the nonlinear regime. As shown in Figure 3, the PL dynamics of the monolayer is strongly pump intensity dependent. While for the bilayer and the trilayer, such excitation density dependence is observed but not nearly as significant (Figure 3). As the pump intensity increases, a faster decay component emerges. At such pump intensities, the integrated PL intensity also deviates from linear behavior (Fig. S5). The fast dynamics along with saturation of PL intensity at high exciton densities can be understood by Auger recombination or exciton-exciton annihilation, where two excitons interact with each other and fuse to form a higher energy exciton followed by a rapid internal conversion to the lowest energy excited state (Figure 4). It is possible that other mechanisms such as bi-exciton formation could occur at high pump intensities and contribute to fast dynamics. However, biexciton formation would lead to additional low energy emission peak. For the range of the pump intensity we used in the experiments, the spectral shape did not change as a function of pump intensity (Fig. S6). We only observed an obvious biexciton emission peak at about 10 times higher intensity than the highest pump intensity used in the dynamics measurements. Therefore, we concluded that biexciton formation is not responsible for the pump intensity dependent dynamics observed here.
Due to the strongly bound exciton in WS$_2$ with larger exciton binding energy of $\sim 0.3$ eV $\gg k_B T$ at room temperature, exciton-exciton annihilation (Auger recombination) is a bimolecular process that involves two excitons (Figure 4).$^{28,30}$ Note that biexciton formation could also lead bi-molecular dynamics, however, as discussed above, biexciton formation is negligible for the pump intensities used here. Including the exciton-exciton annihilation term, the rate equation of PL decay can be written as

$$\frac{dn}{dt} = -\frac{n}{\tau} - \gamma n^2 \quad (3)$$

where $n$ is the exciton population, $\tau$ is the exciton lifetime without annihilation, and $\gamma$ the annihilation rate constant. Assuming the $\gamma$ is time-independent, the solution to the above equation is$^{45}$:

$$n(t) = \frac{n(0)\exp(-\frac{t}{\tau})}{1 + \gamma n(0)[1 - \exp(-\frac{t}{\tau})]} \quad (4)$$

where $n(0)$ is the initial exciton density. The above equation can be rewritten as a linearized form:

$$\frac{1}{n(t)} = \left(\frac{1}{n(0)} + \gamma \tau\right) \exp \left(\frac{t}{\tau}\right) - \gamma \tau \quad (5)$$

The initial exciton density $n(0)$ is estimated by using absorption cross-section of $5 \times 10^5$ cm$^{-1}$ per layer of WS$_2$,$^{46}$ and assume every pump photon absorbed create one exciton.
Figure 3. PL decay at different excitation density for (a) monolayer, (c) bilayer, (e) trilayer; Linearized data using equation (5) for the PL decay of (b) monolayer, (d) bilayer, (f) trilayer. The red lines are linear fits to the data. $n(0)$ is the initial exciton density as described in the text.
For the monolayer, bilayer, and trilayer pump intensity dependent data from Figures 3(a), (c), and (e) is replotted in the linear form as shown in Figures 3(b), (d), and (f), respectively. \( \tau \) is given by the PL lifetime measurement when free of exciton-exciton annihilation. We applied a global fit to the whole data set to equation (5) to determine the exciton-exciton annihilation rate \( \gamma \). We obtain an exciton-exciton annihilation rate \( \gamma \) of \( 0.41 \pm 0.02 \) cm\(^2\)/s for the monolayer. The results for bilayer and trilayer are also shown in Figure 3. The exciton-exciton annihilation rate was determined to be \((6.0 \pm 1.1) \times 10^{-3}\) and \((1.88 \pm 0.47) \times 10^{-3}\) cm\(^2\)/s for bilayer and trilayer, respectively.

Most notably, the exciton-exciton annihilation rate is more than 100 times larger in monolayer than bilayer and trilayer. For monolayer, exciton-exciton annihilation occurs at exciton density as low as \(10^9\) excitons/cm\(^2\), or 1 exciton/\(10^5\) nm\(^2\), corresponding to an averaged inter-exciton distance of \(> 600\) nm. The spatial extent of exciton for the monolayer has been estimated to \(~ 3\) nm.\(^{22}\) This large inter-exciton distance implies that exciton diffusion has to precedes exciton annihilation.\(^{38}\) The following kinetic scheme describes the two rate-determining steps: (1) the diffusion of two excitons toward each other; (2) the annihilation of the two excitons when they are sufficiently close to each other with rate.

\[
E + E \xrightleftharpoons[k_{-1}]{k_1} (EE) \xrightarrow{k_a} E (\text{higher energy})
\]
Where $E$ is an isolated exciton, $(EE)$ denotes exciton pair sufficiently close that annihilation can take place, and $k_a$ is the annihilation rate proceeding from $(EE)$. $k_I$ is defined as the rate of change of the number of the close pair per unite area and $k_{-I}$ is rate for the reverse process. The overall exciton-exciton annihilation rate becomes

$$\gamma = \frac{k_I k_a}{k_{-I} + k_a} \tag{7}$$

Exciton diffusion rate is slightly higher in the bulk than the monolayer as demonstrated by recent transient absorption microscopy measurements of monolayer and bulk MoSe$_2$ with a diffusion constant $D$ measured to be 12 and 19 cm$^2$/s for monolayer and bulk, respectively.$^{47}$ Overall, exciton diffusion is an order of magnitude faster than the overall exciton-exciton annihilation rate if assuming similar exciton diffusion for WS$_2$ as MoSe$_2$, which implies $k_I, k_{-I} \gg k_a$. and annihilation $k_a$ is the rate-limited step(i.e. $\gamma \sim k_a$). For bilayer and trilayer, as thickness increases, exciton binding energy decrease, which also results in more delocalized excitons. The more delocalized nature of the indirect exciton could lead to even faster exciton diffusion. Therefore, much reduced $\gamma$ values measured in bilayer and trilayer is most likely due to smaller values of $k_a$ rather than $k_I$ and $k_{-I}$ in equation(7).

We explain thickness dependence of $k_a$ as follows. First, at the monolayer limit, stronger Coulomb interaction between the electrons and holes leads to stronger many-body interaction. Secondly, exciton-exciton annihilation requires conservation of both energy and momentum.$^{48}$
For bilayer and trilayer WS$_2$, which are indirect semiconductors, momentum conservation requires assistance from phonons.\textsuperscript{48,49} On the contrary, for the direct bandgap monolayer the involvement of phonon is not necessary. This is schematically illustrated in Figure 4.

Figure 4. Schematic summary of relaxation pathways for WS$_2$ monolayer and few-layer. (1) direct exciton recombination; (2) direct exciton-exciton annihilation; (3) indirect exciton recombination; and (4) indirect exciton-exciton annihilation; (5) other nonradiative pathways.

The additional requirement for phonon assistance makes exciton-exciton annihilation a much less probable event leading to at least two order of magnitude smaller $\gamma$ in bilayer and trilayer than in monolayer. In other words, in bilayer and trilayer WS$_2$, only a fraction of exciton encounters results in annihilation. The exciton-exciton annihilation rate is further reduced in trilayer in comparison to bilayer. Assuming the PL in bilayer and trilayer involve
similar phonon-assisted processes (Figure 4), the relative rate of $k_a$ should be similar to that of radiative rate. The radiative lifetimes for bilayer and trilayer are $\sim 400$ ns and $\sim 830$ ns, respectively. The relatively radiative rate is $\sim 2:1$ (bilayer: trilayer). The annihilation rate is $(6.0 \pm 1.1) \times 10^{-3}$ and $(1.88 \pm 0.47) \times 10^{-3}$ for bilayer, leading to a ration annihilation rate of 3:1 (bilayer: trilayer), similar to that of radiative rate.

The efficient exciton-exciton annihilation in monolayer WS$_2$ implies that the inverse process, impact ionization (multiple exciton generation)$^{50}$, could also be effective. By following the method used to obtain the Auger recombination time in quantum dots$^{23}$, we extract the exciton-exciton annihilation time to be $\sim 400$ ps in monolayer at an initial exciton density of $1.6 \times 10^9$ excitons/cm$^2$ (Fig. S7 in the supporting information). Such slow exciton-exciton annihilation on the hundreds of ps time scale makes it quite possible to extract the addition exciton generated.

For monolayer WS$_2$, extraordinary large exciton binding energy and significant correlation between electron and holes compared to bulk semiconductors can explain the enhanced many-body interaction. Here, we measured exciton-exciton annihilation rate of $0.41 \pm 0.02$ cm$^2$/s for monolayer, comparable to that of $0.33 \pm 0.06$ cm$^2$/s measured for monolayer MoSe$_2$ $^{37}$ and that of $\sim 0.35$ cm$^2$/s for monolayer WSe$_2$ $^{38}$. These values are about two order of magnitude larger than that in 2D semiconducting quantum wells ($10^{-3}$ cm$^2$/s)$^{51}$ consistent with increased correlation between electron and holes in 2D TMDs. Pervious theoretical calculation predicted
that Auger recombination rate could enhance by a factor of 50 for 2D systems when compared with bulk.\textsuperscript{52}

**Conclusion**

In summary, we have investigated exciton dynamics in monolayer, bilayer, and trilayer WS\textsubscript{2} using time-resolved PL under conditions with and without exciton-exciton annihilation. Exciton decays of monolayer, bilayer, and trilayer all exhibit mono-exponential decay behavior. PL lifetime is measured to be 806 ± 37 ps, 401 ± 25 ps, and 332 ± 19 ps for WS\textsubscript{2} monolayer, bilayer, and trilayer, respectively, when free of exciton annihilation. The radiative lifetime of exciton is determined to be ~13 ns, ~400 ns, and ~830 ns for monolayer, bilayer and trilayer, respectively. Further, two orders of magnitude enhancement of exciton-exciton annihilation rate has been observed in monolayer compared to that of the bilayer and trilayer. We attribute the strongly enhanced annihilation in monolayer WS\textsubscript{2} to enhanced electron hole interaction and to the transition to the direct semiconductor, which eliminates the need for phonon assistance in exciton-exciton annihilation.

**Experimental Methods**

WS\textsubscript{2} monolayers and few-layers samples were mechanically exfoliated from bulk crystals (2D Semiconductors) onto Si wafer with 90 nm oxide thickness (Graphene Supermarket). The number of layers of WS\textsubscript{2} flakes was identified by Raman microscope and further confirmed by
steady-state PL spectroscopy. Raman spectra were collected by a Renishaw Raman Microscope (RM1000) equipped with an argon-ion laser at 514 nm as an excitation source. The excitation beam was focused by a 50X (NA = 0.75) objective, and the Raman scattered light was collected with the same objective. The beam size in Raman measurement is ~ 1μm, which is much smaller than our sample size (about 5 μm). Steady-state PL and TRPL measurements were performed by employing a home-built confocal microphotoluminescence setup. A picosecond pulsed diode laser (PicoQuant, LDH-P-C-450B) with an excitation wavelength of 447 nm (FWHM = 50 ps) and repetition rate of 40 MHz was used to excite the sample, which was focused by a 100X (NA = 0.95) objective. The beam size in PL measurement is < 1μm, which is also much smaller than our sample size (about 5 μm). The PL emission was collected with the same objective, dispersed with a monochromator (Andor Technology) and detected by TE cooled charge coupled device (Andor Technology). The dynamics of PL were measured using a single photon avalanche diode (PicoQuant, PDM Series) and a single photon counting module (PicoQuant). The time resolution of this setup is ~ 100 ps. All optical measurements were conducted under ambient conditions and at room temperature.

Acknowledgement

The authors acknowledge the support from the National Science Foundation under award number 1433490. The authors thank Prof. Tony Heinz for useful discussions.
Notes and references

\(^a\) Department of Chemistry, Purdue University, West Lafayette, Indiana 47907

Electronic Supplementary Information (ESI) available. See DOI: 10.1039/b000000x/


47. N. Kumar, Q. Cui, F. Ceballos, D. He, Y. Wang and H. Zhao, Nanoscale, 2014, 6, 4915.