UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title

Localization of Excitons due to Inhomogeneous Nanoscopic Strain in Monolayer Transition Metal Dichalcogenides investigated by nano-PL and nano-Raman Microscopies

Permalink

https://escholarship.org/uc/item/91d0x9c4

Author

Darlington, Thomas Phillip

Publication Date

2020

Peer reviewed|Thesis/dissertation

Localization of Excitons due to Inhomogeneous Nanoscopic Strain in Monolayer Transition Metal Dichalcogenides investigated by nano-PL and nano-Raman Microscopies

By

Thomas P. Darlington

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Feng Wang, Co-Chair
Professor P. James Schuck, Co-Chair
Professor Mary C. Scott
Professor Jeffrey B. Neaton

Spring 2020

Abstract:

Localization of Excitons due to Inhomogeneous Nanoscopic Strain in Monolayer Transition Metal Dichalcogenides investigated by nano-PL and nano-Raman Microscopies

By

Thomas Darlington

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Feng Wang, Co-Chair

Professor P. James Schuck, Co-Chair

The monolayer semiconducting transition metal dichalcogenides are a family of 2D crystals that, beginning with MoS_2 and soon expanding to include tungsten and other chalcogen variants, has become a focus of much research in 2D materials. Their semiconductor character, with wide direct bandgaps in the visible and near-infrared energy spectrum, make them particular amenable to optoelectronic applications. The low-dimensional character leads the direct gaps to host tightly bound exciton states that are stable at room temperature, allowing for the experiments on exciton physics that were only accessible at low temperatures in classical 2D excitonic system such as GaAs 2D quantum wells. The crystal structure of TMDs further gives the excitons and other quasi-particles valley contrasting physics, which can be discriminated optically with circularly polarized light, and leads to phenomena such as the valley Hall effect. The combination of strong light-matter interactions in TMDs with novel physics makes optical spectroscopy and microscopy powerful tools in the characterization of these materials. However, the native length scales of the excitons, while larger than the lattice constant, are still far below the resolution of optical probes. This resolution is set by the diffraction limit of light, which limits the focusing of propagating electromagnetic fields to roughly half the wavelength. Optical probes will effectively average over large areas (200 x 200 nm^2) of the crystal, obscuring the effects of highly localized and heterogeneous perturbations.

The need for local experimental probes in the study of TMDs has been highlighted by the recent discovery of single-photon emission from monolayer WSe_2. Many far-field experimental works have shown a strong association of the emergence of these emitters and regions inhomogeneous tensile strain, leading to the hypothesis of quantum confinement of the excitons by the strain-created potential wells. However, experimental investigation of this hypothesis is challenging due to the scale mismatch of the localized excitons, which must be on order the exciton Bohr radius of $\sim$ 1 nm, and the $\sim$ 250 nm size of the optical probes.

The focus of this work is to bridge the resolution gap of optical microscopy and spectroscopy in the study of excitons in the 2D TMDs using scanning near-field optical microscopy (SNOM), which allows optical resolutions down to a few nanometers. Using this technique to measure spectra of photoluminescence (PL) and scattering from lattice vibrations (Raman scattering), I investigate the nanoscale interplay of strain and the excitons in several of the commonly studied semiconducting TMDs, including WS_2, WSe_2, and $MoSe_2$. In this thesis, I first discuss methods of performing near-field optical

studies on 2D materials, including aperture and apertureless SNOM methods, and demonstrate them by collecting real nanoscale PL (nano-PL) and nano-Raman on WS_2 and WSe_2. Following that, I show how to estimate the strain in nanobubbles from classical models of plate bending, and show by comparison with shifts in the lattice vibrations that such models on average correctly estimate the strain. Further, by combining high-resolution nano-PL of WSe_2 with advanced microscopic theories, I show evidence that strain alone can efficiently localized excitons at room temperature due to formation of atomic scale wrinkles of the crystal. Using gap-mode nano-PL with nanometer resolution, I demonstrate that the spatial location of localized exciton emission in WSe_2 occurs along a common axis over several nanobubbles, and that it is highly localized in the wrinkle-like rings in asymmetric nanobubbles. Lastly, investigation of nanobubbles in $MoSe_2/WSe_2$ heterostructures shows that similar wrinkle-like ring structures occur commonly in large nanobubbles, and that such structures host strong exciton complexes in energetic regions associated with the interlayer exciton. This emission is unusually strong at room temperature, and has emission peaks that extend passed 1000 nm, which has not been previous described in such heterostructures. The results presented here show the strong and rich interplay of excitons with strain, demonstrating the utility of strain-engineering in future TMD devices, but also represent a frontier in studying the physics of confined excitons both spectrally and spatially.

Introduction:

A fundamental insight of materials science and condensed matter physics is the dependence of the macroscopic measured physical properties on the underlying atomic structure. Perturbations of this structure thus permit manipulation and control of these properties. For example, introduction of defects in insulators like diamond or boron nitride create deep sub-bandgap color centers that are being developed for quantum computing applications. Strain applied to semiconductors has long been known to change the bandgap, but also can create polarization fields through piezo-electricity and modify carriers dynamics by changing the electron-phonon coupling and by breaking degeneracies. In 2D materials, inhomogeneous strain has been shown to lead to novel physical phenomena, such as pseudo-magnetic fields in graphene[1] and quantum dot-like bound exciton states in the transition metal dichalcogenides (TMDs)[2-5]. Classically, crystal structure has been investigated in momentum space with diffraction techniques to great success. However, because these techniques average over large areas of the crystal, they are ill-suited for probing structural perturbations that are highly local and inhomogeneous. Real-space techniques must be employed.

Techniques such as scanning tunneling microscopy and transmission electron microscopy routinely measure single atomic defects and localized electronic states resulting for instance from the above-mentioned pseudo-magnetic fields in graphene. While these techniques rule in terms of resolution, they are limited in the study of excited state phenomena like excitons. Excitons in 2D materials have extraordinarily high binding energies (~500 meV)[6] and boast strong interaction with light. Optical microscopy and spectroscopy are thereby powerful probes of these excited states. However, conventional optical microscopy has a lower bound on the achievable resolution, set by the diffraction of light, typically >250 nm. The scale of many excitonic properties is unfortunately much smaller than this limit, on the order of ~1-100 nm. For instance, the quantum-dot-like states mentioned above are localized on scales ~10 nm or lower. This scale mismatch of light and excitons leads to a "resolution gap" in the study of the fundamental properties of these states.

Bridging the resolution gap is the topic of super-resolution microscopy. Super-resolution typically refers to far-field techniques to enhance resolution using photon localization methods and/or manipulation of the excited states of fluorescent molecules. Higher resolution can also be achieved by collection of higher wavevectors corresponding to the evanescent fields. These fields are non-propagating and typically decay exponentially away from the source, be they scattering or luminescent in origin. Inclusion of the evanescent fields can be achieved in the far-field by mixing the fields with lower wavevector fields that do propagate using structured light. However, because the structured light is itself diffraction limited, the resolution enhancement is capped at approximately double for linear structured illumination. Higher wavevectors can be included by going to the near-field, defined as a distance from the source where the evanescent fields are still significant. Though these techniques now routine achieve resolution in the tens to even single nanometer scale, the requirement that the probe be extremely close to the sample limits these techniques to surface investigations

2D materials are atomically thin are thus an ideal system for characterization with near-field methods, as the entire volume can be probed. This work is focused on the application of near-field optical methods to study excitons in 2D materials, including the effect of nanoscopic strain due to nanobubble formation on the exciton properties. In the following sections, I outline the classical origin of the diffraction limit from Maxwell's equations, and the principal techniques of scanning near-field optical microscopy, followed by its application to nanobubbles in WS_2, WSe_2, and heterostructures of

MoSe$_2$/WSe$_2$. A broad theme of these sections is the localization of these excitons and the spatial structure of the localized exciton (LX) distributions, and I show that LX positioning is not random as might be expected from crystalline defects, but reflects complex strain distributions at the nanoscale.

Section 1 – the transition metal dichalcogenides

There are example materials for each level of dimensionality below the usual "3D" (excluding here the extra dimensions in some theoretical speculations). Quantum dots are perhaps the most commonly studied 0-dimensional systems today. For 1-dimensional systems, carbon nanotubes kicked off a wave of research in their properties in the 1990's and early 2000's as well as analogous materials such as boron nitride tubes and wires. For 2-dimensional systems, it was carbon again, in the form of graphene, that launched a massive effort in the physics of 2D semi-metals, semiconductors, superconductors, and insulators (of which one of the most ubiquitous examples is again boron nitride in its 2D form).

In this section we introduce the sub-family of 2D materials that constitutes the subject of this thesis: the molybdenum and tungsten family of 2D semiconductors collectively referred to as the group VI transition metal dichalcogenides. In introducing these materials, I will first describe how the 2D versions of the transition metal dichalcogenides (TMDs) are different from their 3-dimensional counterparts. This will begin with a short discussion of graphene, which has a very similar structure to the TMDs. Next, I will discuss of how band structures changes in the monolayer form of TMDs. In the last section I introduce excitons and their general properties in 2D TMDs.

1.1: The van der Waals structure of graphene and TMDs
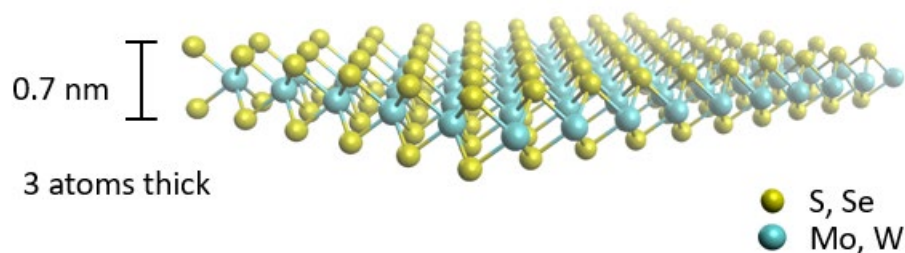
Figure 1-1: Ball and stick model of the common 2H-phase of group-VI semiconducting transition metal chalcogenides.

Before delving into the 2D world, it is worth discussing the structure of the 3D versions of graphene and TMDs. Bulk graphene is graphite, commonly found in pencil tips among its many uses. The atomic structure of graphite is hexagonal. The common 2H-phase of the TMDs like MoS_2 has a very similar structure, but instead of a single atomic layer of carbon, it is a tri-atomic layer of 2 chalcogen planes sandwiching the molybdenum or tungsten atoms, shown in Figure 1-1. Bulk graphene (graphite) and the TMDs are made up of many atomic layers[7,8]. An important physics feature to mention at this point is the nature of the bonds layered materials such as graphite and TMDs, both within each layer and between each layer. For graphite and TMDs, the intralayer bonds are covalent bonds are very strong. However, between the layers, the material is held together by comparatively much weaker van der Waals forces. For this reason, graphite, bulk TMDs, and similar materials like hexagonal boron nitride are often referred to as van der Waals materials.

The van der Waals structure is dramatically important. It is this comparative strength mismatch of the chemical bonds that allows the individual layers to be separated without complete destruction of the

hexagonal structures of the layers, often through physical exfoliation by means of cheap scotch tape. The comparative ease of producing atomically thin materials with no dangling surface bonds and without expensive epitaxial growth methods is a major advantage of these materials.

1.2: Band structure of TMDs

When a single layer of van der Waals crystal is isolated from the bulk, the resulting structure can have dramatically different properties from the bulk. Graphene for instance transitions from metallic to semi-metallic. For TMDs, both the bulk and single layers are semiconducting, but the band structure transitions from an indirect to direct bandgap semiconductor[8,9].

Why this happens for TMDs can be seen from the relative contributions to the band structures from the metal and chalcogen atoms. In the common trigonal prismatic (2H-) phase of TMDs, the conduction band minimum is located between the gamma and K-point, sometimes called the sigma point, of the Brillouine Zone. Meanwhile, the valence band maximum is located at the gamma-point. The band states at the K-point are primarily composed of the d-orbitals from the metal atom, while the p-orbitals of the chalcogenides make greater contributions to the valence and conduction band states as one moves toward the gamma point from K-point[8,9] . Given the layered structured of TMDs, with the metal atom sandwiched between two chalcogens, we might expect that the bands would be more affected by the layer number the closer we are to the gamma point. This has been shown in ab initio density functional theory, where the valence band maximum at the gamma-point moves down in energy and the sigma point conduction band minimum moves up as layer number is reduced, till in the monolayer both the conduction band minimum and the valence band maximum are at the K- and the K'-point[8]. This can be understood as increased quantum confinement in the direction perpendicular to the layer.

1.3: Excitons in monolayer TMDs

Excitons are elementary excitations in semiconductors that occur when an electron in the conduction band binds to a hole in valence band, via the Coulomb interaction[10]. Excitons are thus analogous to a hydrogen atom with the proton replaced by the hole in the valence band. However, the ground state of the exciton is an excited state of the material. While this description may seem exotic, excitons are a well-known and studied phenomenon in 2D and 3D semiconductors, and are important for explaining their optical interactions. To illustrate this, Figure 1-2 shows a cartoon of an absorption spectrum for a 2D and 3D semiconductor. Using a simple particle in-a-box model it easy to show that without excitons the spectrum has step-function character in 2D and a square root character in 3D. If excitons are included these functional forms are still present. However, below their onset, discrete resonances appear. The first, and typically strongest is the 1s state of the exciton, in further analogy to the hydrogen atom, with the higher resonances identified with the 2s, 2p, etc.

(a) (b)
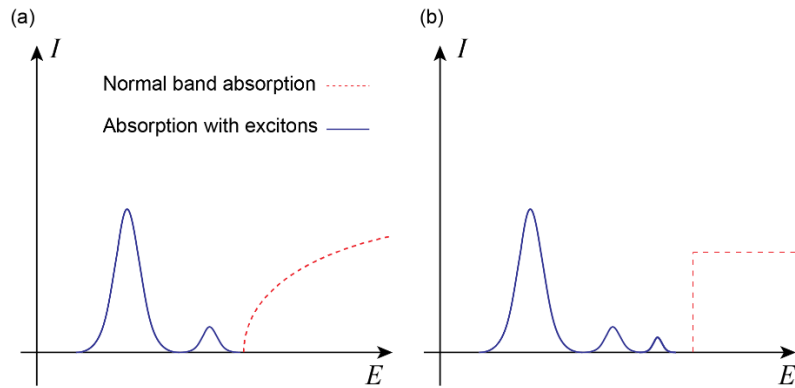
Figure 1-2: diagram of the semiconductor absorption in 3D and 2D, (a) and (b) respectively. Red dashed lines are the normal band edge absorption while the blue curves are the exciton states.

Spectra like those in Figure 1-2 are observed in semiconductors like bulk GaAs and in 2D quantum wells at temperatures ~<10 K. For most semiconductors and insulators, when the temperature is increased the excitonic features in the absorption are suppressed. This occurs because the thermal energy of the lattice quickly overcomes the binding energy of the electron and hole, and the two disassociate. For most bulk semiconductors, the binding energy is on ~1-10 meV[10] , typical for most Wannier-Mott excitons, and thus excitons are negligible in optical interactions at room temperature where kT ~ 26 meV.

The relatively weak binding energy in Wannier-Mott excitons is a consequence of dielectric screening of the Coulomb interaction[8,9] within most materials. An important consequence of the reduced dimensionality in 2D materials, particularly the reduction from multi- to single layer, is the change in the dielectric screening of the charge carriers, which results in very significant excitonic effects. This can be seen in the diagram of Figure 1-3, which considers an electron and hole in a 2D material.
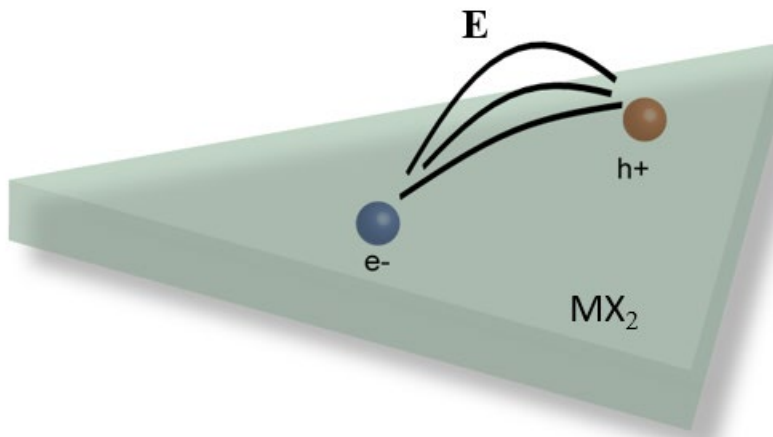
Figure 1-3: Illustration of lowered screening in 2D materials.

Since many of the electric field lines can pass outside the material, the attraction felt between the electron and hole is much greater. In the bulk, all of the field lines must pass through the material and are thus screened by dielectric response. However, in a monolayer, the lack of material on either side reduces this screening significantly, leading to stronger interactions of the electron and hole and a correspondingly

larger binding. Because monolayer TMDs are only 3 atoms thick, this increase in screening is dramatic, leading to a binding energy on the order of $300 - 500$ meV[6,11]. Because of this high binding energy, excitons in TMDs are observable in optical studies even at room temperature, and they lead to strong photoluminescence.

The strong photoluminescence of TMDs was first observed in monolayer MoS_2[8,9]. Since then, there has been a explosion of scientific interest in the properties of the TMDs. In addition to the strong PL and room temperature excitons, monolayer TMDs exhibit a number of other interesting physical properties. Unlike graphene, TMDs have significant spin-orbit coupling. The effect of this is to break the spin-degeneracy at the K and K' points in both the valence and conduction band[12]. The splitting has different strengths in the bands, being stronger in the valence than the conduction bands. In terms of energy, the spitting of the valence band is on order ~100 or 400 meV for the molybdenum and tungsten species, respectively, compared to the splitting in the conduction band on order 10 or 30 meV[12] . In addition, the lack of in-plane inversion leads to the so-called "spin-valley" locking, where valley here refers to weather the quasi-particle is located at K or K' points of the Brillouin zone. This entails that the direction of the spin polarization flips in switching from the K to K' valley.
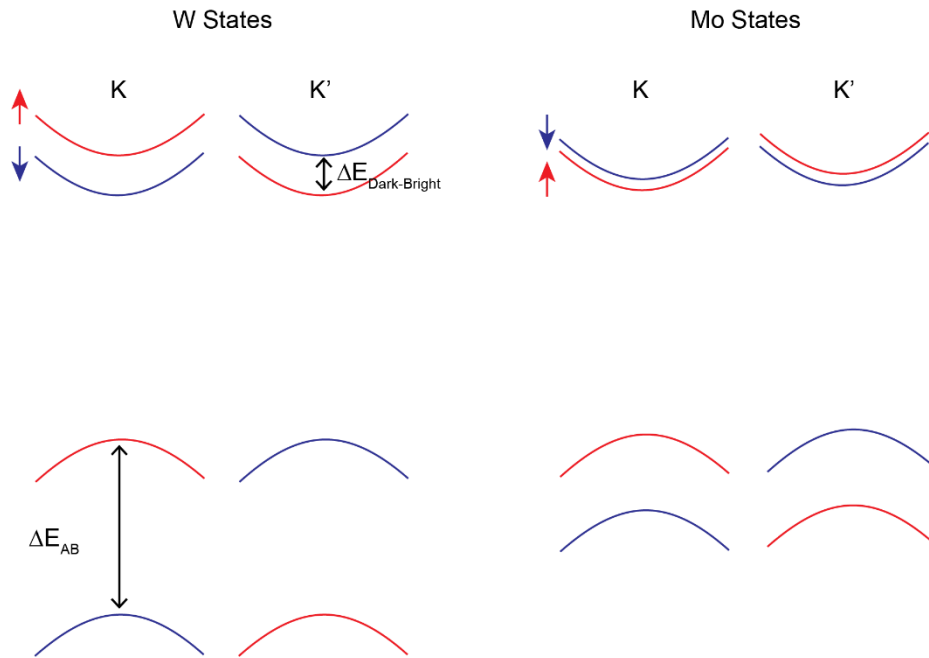
Figure 1-4: Diagram showing the band edge states at the K and (K') points in W- and Mo- transition metal dichalcogenides

A diagram of this splitting is shown in Figure 1-4 for the molybdenum and tungsten K and K' states, with the different spins shown in red and blue (spin up and spin down). For each valley we see there are 4 pairs of valence and conduction band states, implying four different excitons. Two of the excitons are formed of valence and conduction bands of parallel spins, with the other two of valence band of opposite spin from the conduction band. In the latter, an electron transition from one band to another

requires a spin flip, making the transition from these states are optically forbidden, and are generally referred to as dark excitons.with the other pairings referred to a bright excitons[13]The splitting of the two energies is shown in Figure 1-4 at $\Delta E_{\text{Dark-Bright}}$. One important difference between the Mo and W species is the opposite ordering of the spin splitting in the conduction band, sometimes referred to as the spin texture[12]As a result, the lowest lying exciton in molybdenum TMDs is the bright exciton, where it is the opposite for the tungsten sulfide and selenide.

Because the spin splitting in the conduction band is on the order of $k_{\text{B}}T$ at room temperature, its effects are most prominently observed at lower temperatures. However, by use of apertureless SNOM (a-SNOM) methods, the dark exciton in WSe_2 can also be observed at room temperature[14]. The excitons formed from the splitting of the valence band are a different story, with the energy difference between the split states far above the mean thermal energy, and are often seen in absorption spectroscopy [6,11,15]. Excitons formed with the top valence band are referred to as the A-excitons, and the excitons formed with the lower valence band, thereby at higher transition energy, are the B-excitons.
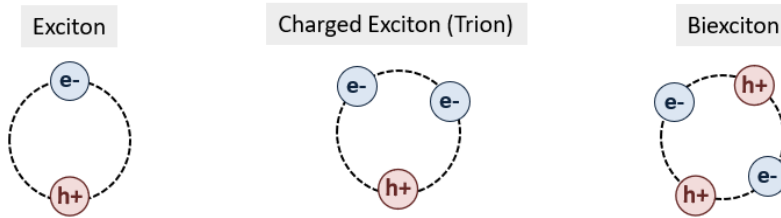
Figure 1-5: Excitons states, from left, normal exciton, charged exciton (trion), biexciton**.**

Beyond the standard hydrogen-like excitons, with a single electron bound to a single hole, there exists also excitons formed with additional electrons. The two most important types that we will discuss are the charged exciton (trion) and the biexciton. A cartoon of these states is shown in Figure 1-5. Trions may be positive or negative depending on the sign of the free carriersThe biexciton is a four particle bound state of two electrons and two holes. Naturally, the biexciton is seen as the result of two normal excitons interacting to form a bound stateand not unexpectedly formation of biexcitons depends strongly on the density of normal excitons[16]. Biexcitons themselves may also become charged, forming a five-particle state, and also can be formed of bright and dark excitons, leading to a rich excitonic landscape, which was recently studied in WSe_2 at low temperatures[16].

1.4: Interlayer excitons in van der Waals heterostructures

Beyond the fascinating physics and promising electronic properties seen in monolayer TMDs, individual layers of vdW materials can be stacked onto one another to form a heterostructure. This idea extends beyond the semiconductor transition metal dichalcogenides of course, originating in the study of graphene and progressing rapidly after discovery that graphene on boron nitride substrates greatly increased the mobility[17]. Like assembling a structure with Lego bricks, a great variety of heterostructures can be created by stacking individual van der Waals onto one another[18]. In this section we will briefly discuss the properties of heterostructures formed with the group VI TMDs.

When two monolayers are brought into contact, the electronic states of the two layers will mix, called hybridization. This hybridization happens strongly at the gamma and sigma points of the Brillouin zone, but is forbidden by symmetry at the K points, leaving the electronic states at these mostly confined

to the individual layers of the respective monolayers. The different work functions and bandgaps of each of the TMDs results in type-II band alignment at the K-point[19,20] , with the minimum energy states of the electron and holes located in different layers of the heterostructure. An interlayer exciton (ILE) can be formed between electron and holes at the K and K' points in different layers.

The ILE is then the lowest energy exciton in the TMD heterostructure. For instance, the ILE is ~300 meV below the WSe_2 exciton in the $MoSe_2/WSe_2$ heterostructures. Compared to the intralayer exciton, the ILE has significantly less overlap of the electron and hole wavefunctions and is an *indirect* exciton owing to the lattice mismatch and/or twist of the TMDs monolayers in the heterostructure[21-24] . These effects lead to significantly weaker oscillator strength and longer lifetimes[19,20]. Further, the spatial separation of the charge leads to formation of static dipoles, which leads to stronger interactions between excitons. The combination of long lifetimes and strong exciton-exciton interactions makes the ILE physics very rich, and they are among the most actively studied excitations in TMD heterostructures.

Section 2 – limitations of a classical optical microscope

Optical microscopy constitutes of the most ubiquitous as well as one of the most powerful experimental tools in use by modern researches. Though the word "optical" in principal includes all wavelengths of light, from x-ray to radio frequencies. However, most commonly it refers any of the wavelengths of light that are perceivable by humans. In terms of the wavelength this corresponds to light waves between 400 – 700 nm in length. These values are not hard cut offs, for instance some optical microscopes will include light sources ~ 394 nm or ~785 nm. Upright visible light microscopes can be found in nearly every experimental laboratory in the life and physical sciences, as well as being among the most common tools used for science education and outreach. It is arguable that the image of the upright microscope itself is the most recognized symbol of experimental science bar none.

The purpose of any microscope, by its name, is to observe and measure small things, often (but not necessarily) imperceptible to normal human vision. With this in mind, the preeminence of visible light microscopes in modern research may seem surprising since higher resolutions cousins, such as electron microscopes are well developed. Indeed, the development of electron microscopy, as well as the many scanning probe microscopes, is chiefly motivated by the limited resolution of microscopes using visible light. So why then the ubiquity?

There are of course practical considerations to this question. One may assume, correctly, that more advanced microscopes are more expensive, both in monetary cost and difficulty of operation. First developed in the 17th century by English scientist Robert Hooke and Dutch scientist Antoni van Leeuwenhoek (Nikon MicoU), technology of crafting visible light microscopes is by a wide margin the most developed. We therefore would expect more microbes using visible light from a purely economic perspective.

However visible light microscopy has several advantages over its higher resolution cousins. Firstly, it worthwhile to discuss light detection. The most common detector used is, perhaps obviously, human eyes, and, by definition, visible light would be the best choice. This is true also for current generation of electronic light detectors, including point as well as array detectors such as cameras. Si based CCDs can have quantum efficiencies approach unity in the visible and near-infrared wavelengths of light. This high sensitivity has allowed studies of single molecules, which being only few nanometers in dimensions where previously thought impossible to observe. Secondly, many things of interest in the life and physical sciences interact with visible light strongly, through absorption, light emission scattering, in the visible spectrum. For example, in fluorescence microscopy many fluorescence molecules, such as the revolutionary green fluorescent protein, absorb and emit in the visible spectrum. But this true for many inorganic materials as well. The bandgaps of the many of the most widely used semiconductors are in the mid or near-infrared and the visible. Examples include the aforementioned Si (1.14 eV bandgap), but also Ge (0.67 eV), GaN (3.4 eV), GaAs (1.43 eV), GaP (2.26 eV)[25]. We can compare this to the energy values of visible light photons 1.77 eV (700 nm) to 3.1 eV (400 nm). Even very insulating materials such as

diamond (5.5 eV bandgap) and BN (6 eV) sometimes have defect states which lie squarely in the range of visible light. If we take the values of the bandgap as a measure of the energy separation between the electronic states of many common materials, it appears that photon energies in the energies ~1 eV correspond well, making visible and near-infrared light natural probes. Thirdly, and this applies to many optical investigation methods, visible light microscopy is *non-destructive*. This is often not the case electronic microscopies, especially TEM. Though this technique can achieve extraordinary resolution, even resolving atoms, the high energies of the electrons are extremely damaging, even on hard metallic samples.

Clearly there is great utility to visible light microscopes in experimental science. However, when the size of the system to be interrogated shrinks significantly below the size of a small cell (~1 um), resolution becomes an issue, namely the image will blur to the point where no spatial features can be distinguished. The source of this blurring is the wave nature of light, which as discussed above is inversely proportional the energy of the photon. Resolution of the microscope scales linearly with the wavelength, $\Delta x \sim \lambda$. Scaling down the wavelength by increasing the energy of each photon will improve the resolution. Unfortunately, like electrons, higher-energy photons such as ultra-violet and x-rays very quickly begin to damage many samples. Focusing light also becomes difficult as the photon energy increases. Normal refractive glass, which =commonly used for visible light, becomes opaque in the UV. Specialized focusing mirrors are required. The sources themselves can also pose challenging. For example, while lasers are readily available throughout the visible and near infrared, they are dramatically less so for higher energy light.

A natural question is can we improve the resolution of a microscope without decreasing the wavelength, wherein we retain the advantages of the visible and near infrared? In this section we will explain the origin of the resolution limit of in optical microscopy (the diffraction limit), outlining the key results from electromagnetic theory. Following we will explore methods of circumventing the diffraction limit, briefly mentioning far-field techniques. In the final part we will explain the principals of near-field microscopy, which will be the topic of following section.

1.1: Mathematical representations of optical fields

Like all wave phenomenon, electromagnetic waves are characterized by two fundamental parameters: the wavelength (λ), the distance between two peaks of the wavelength, and the wave frequency (v), which is the inverse of the period. Multiplied together these two quantities define the speed of the wave, (or formerly the phase velocity):

$$\lambda v = \frac{c}{n} \tag{1}$$

The speed is a fundamental constant of physics, and in vacuum all light travels at this speed. However inside matter the speed in equation (n) is reduced by a quantity called the refractive index. We will see bellow that this modification of the phase velocity partly determines the ultimate resolution of a microscope, and allow provides a means to increase the resolution.

In the theory below it's is convenient to use the angular version of the frequency, which are found by simple multiplication by 2π. And instead of the length is advantageous to a related quantity called the wavevector

$$\tag{2}$$

$$k = \frac{2\pi}{\lambda}$$

When expressed as a vector, the wavevector also denotes the direction of propagation of the light wave: $\mathbf{k} = (k_x, k_y, k_z)$ with that condition that $k_x^2 + k_y^2 + k_z^2 = k^2 = \left(\frac{2\pi}{\lambda}\right)^2$. With the definitions of the angular variables, the phase velocity is then:

$$\frac{\omega}{k} = \frac{c}{n} \tag{3}$$

The angular versions are no more fundamental, but simply allow the writing of a general light field in terms of harmonic light fields simpler. Harmonic fields here refer plane waves, or waves of the form:

$$E \sim \exp\big(i(\mathbf{k} \cdot \mathbf{r} - \omega t)\big) \tag{4}$$

1.2: The wave equation for light in free space

To understand the limitations of an optical microscope we will first begin with the fundamental equations of a classical light field, to derive the wave equation for light in free space. The wave equation combined with the angular spectrum representation we can established the wavelength dependence of the cutoff between evanescent and propagating optical fields which provide an ultimate limit on the resolution of a standard microscope.

We start with Maxwell's equations in vacuum. Commonly Maxwell's equations are presented in differential form. In vacuum they are a set of four equations, the divergence and curl for the electric and the magnetic fields:

$$\nabla \cdot \mathbf{E} = 0 \qquad \nabla \cdot \mathbf{B} = 0$$
$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad \nabla \times \mathbf{B} = \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \tag{5}$$

The divergence equation for the magnetic field is identically zero in all situations since no net source of magnetic field is yet known. Without any source of charge the divergence is likewise zero.

We proceed then by examining the curl for both fields, which shows that the two fields are coupled, with the curl of one being proportional to the time rate of change of the other. We can very straight forwardly show that these equations directly imply that light is wave for standard vector calculus operations. What is typically done now in most textbooks is to take the curl of equation (xx) for the curl of the electric field again. We may then switch the order of differentiation on the right hand side showing that the double curl of the electric field is equally to the partial time derivative of the magnetic field. Substituting the other Maxwell's equation for the curl of magnetic field, we derive the following:

$$\nabla \times \nabla \times \mathbf{E} = -\mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} \tag{6}$$

Looking up vector identities for double-curl of the gradient operator ($\nabla \times \nabla \times$) and exploiting the zero divergence of the electric field in absence of net charge we arrive at a wave equation for the electric field.

$$\nabla^2 \mathbf{E} = \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} \quad \text{with the definition of the wave speed } c^2 = \frac{1}{\mu_0 \varepsilon_0}. \tag{7}$$

11

A very similar equation is obtained for the magnetic field by the exact same procedure beginning with the curl for magnetic field.

We can remove the time derivative by assuming our electric field has a harmonic time-dependence, which more complex time dependencies obtained a superposition of harmonic components. Applying the double tip derivative to equation (4) we find:

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\mathbf{E} = -\frac{\omega^2}{c^2}\mathbf{E} = -k^2\mathbf{E} \text{ recalling dispersion relationship of equation (3)}.$$

We then arrive at the homogeneous Helmholtz equation.

$$(\nabla^2 + k^2)\mathbf{E} = \mathbf{0} \tag{8}$$

Specifically, this is the vector version of the Helmholtz equation, meaning that each vector component of the electric field separately satisfies the scalar Helmholtz equation. The parameter k here is the wave vector for the optical field, and is defined as (equation for wavevector). Further in this section we will show that the requirement that light satisfy this equation this will provide a fundamental limit of the resolution of an imaging system. Before doing so is useful to represent that optical field in what is known as the angular spectrum representation.

1.3: The angular spectrum representation:

If the medium in which the electric field is propagating is homogeneous we can represent the spatial dependence of the electric in the same manner as the time dependence where in place of omega we use the components of the k-vector[26]. Essentially this allows to write the spatial dependence as a 2D Fourier transform.

$$\mathbf{E}(x, y, z) = \iint\limits_{-\infty}^{\infty} \hat{\mathbf{E}}\left(k_x, k_y; z\right)e^{i[k_x x + k_y y]}dk_x dk_y \tag{9}$$

Physically what this equation signifies is that we can represent an arbitrary electric field in a homogeneous medium as the sum of propagating and evanescent plain waves.

The essential point of equation (9) is that we can present an arbitrary optical field as a 2D Fourier transform and it provides a useful framework for understanding many properties of optical fields, including strong focusing of a laser spot. For the purposes of this section we can use this equation the understand the fundamental limitation of a far-field imaging system.

Let's consider an optical field as it propagates along the z-axis that we can represent that is described in the angular spectrum representation. This field still must satisfy the Helmholtz equation. Inserting equation (9) into the Helmholtz equation:

$$(\nabla^2 + k^2)\iint\limits_{-\infty}^{\infty} \hat{\mathbf{E}}\left(k_x, k_y; z\right)e^{i[k_x x + k_y y]}dk_x dk_y$$

$$= \left(k^2 - k_x^2 - k_y^2 + \frac{\partial^2}{\partial z^2}\right)\iint\limits_{-\infty}^{\infty} \hat{\mathbf{E}}\left(k_x, k_y; z\right)e^{i[k_x x + k_y y]}dk_x dk_y = 0$$

12

With the common exp() terms we can see that this is just a 2D inverse Fourier Transform for of differential equation for $\hat{\mathbf{E}}$. Applying the forward transform we arrive at a harmonic oscillator equation for the z-dependence of the Fourier coefficients of our optical field.

$$\frac{\partial^2}{\partial z^2}\hat{\mathbf{E}} = -k_z^2\hat{\mathbf{E}}$$

The solution to which is of course: $\hat{\mathbf{E}}(k_z, k_y; z) = \hat{\mathbf{E}}(k_z, k_y; 0)\exp(\pm ik_z z)$

While the mathematics here were relatively simple, the significance of this equation is worth pondering. For a field propagating along the z-axis this states that the Fourier components at some distance along the axis are related by the multiplication of $\exp(\pm ik_z z)$. We can then think of $\exp(\pm ik_z z)$ as a propagator, moving our optical field along the z-axis.

To see how this impacts an imaging system we can first rewrite k_z in terms of the total wavevector minus the transverse components

$$k_z = \sqrt{k^2 - k_x^2 - k_y^2} \qquad\qquad\qquad (10)$$

We see that k_z is real provided that the square of k_x and k_y is less than the square of the total wavevector and imaginary otherwise. For k_z real the propagator oscillates in z, whereas for k_z imaginary it decays exponentially. The former thus describes a propagating field whereas the latter describes an evanescent field. The cutoff between the two is given by the total wavevector.

This cutoff allows us to see how the resolution limit of an optical microscope depends on the wavelength of light observed. Light with in-plane k-vectors whose sum of squares is greater than the square of the total wavevector decay exponentially along the propagation direction. This implies for a detector that is far from the source emitter, all the evanescent fields will have decayed to zero. Considering a field where $k_y = 0$, the maximum propagating in-plane wave vector is $k_x = k$. The smallest feature our detect can observed is then

$$\Delta x \sim \frac{1}{k} = \frac{\lambda}{2\pi n} \qquad\qquad\qquad (11)$$

In this derivation, we made no assumptions about the imaging system we are using, aside from the fact that evanescent waves are not detected. What equation (11) shows is that no matter the sophistication of the imaging system   the smallest features we can observe are on order the inverse of the detected wavelength absent any special light-matter interactions that break our assumptions of light invacuum. We also see the connection between the index of refraction and resolution. With a higher index, the greater in-plane k-vectors can propagate. For this reason, it is common for standard microscopes to use water ($n \sim 1.33$) or special oils ($n \sim 1.5$) between the sample and objective to increase resolution.

While equation (11) shows the correct scaling of resolution with wavelength and refractive index, it is not a formal statement about resolution. The angular spectrum representation can also be used to derive a more formal expression by considering the focusing of a propagating light beam.Thinking backwars imaging is easily seen to be focusing in reverse, i.e. the light source imaged by our focusing element (a lens or microscope objective). The relationship between the focused spot on the incoming light packet $\hat{\mathbf{E}}_\infty$ can be shown to be:

$$\qquad\qquad\qquad\qquad\qquad (12)$$

13

$$E(x, y, z) = \frac{ire^{-ikr}}{2\pi} \iint\limits_{(k_x^2 + k_y^2) \leq k^2} \hat{E}_\infty(\frac{k_x}{k}, \frac{k_y}{k}) \frac{1}{k_z} e^{i[k_x x + k_y y \pm k_z z]} dk_x dk_y$$

It is seen that mathematically focusing is approximately that a of a Fourier transform, the difference arising from the phase factor before the integral and the factor of $\frac{1}{k_z}$. In case of light angles close to the propagating axis (the paraxial approximation) this factor can be neglected and $k_z \approx k$ and the Fourier transform relationship becomes exact[26].

Calculation of even a highly symmetric incoming beam via equation (12) is nottrivial. We can obtain an approximate value by considering the 2D Fourier transformed image of a circular aperture or radius a in the paraxial approximation in terms of angle[27]:

$$|E(\theta)|^2 \sim \left[2\frac{J_1(ka\sin(\theta))}{ka\sin(\theta)}\right]^2$$

This is the famous Airy disk, and represents the intensity image of a point source. For this reason, the Airy disk is referred to as the point spread function (PSF) for the given imaging system in question. A plot of this function is shown in Figure 2-1. In general the curve is bell shaped with oscillation maxima and minima as you proceed outward from the center of the disk. The main feature to observe about pattern is not sharp, infinitesimal point, a but an extended structure that slowly decays. To derive the resolution limit we consider how close two such patterns can be before they become indistinguishable. Commonly distinguishability is defined as the maximum of the first pattern overlapping with the first minimum of the second. Defining $x = a\sin(\theta)$, the result is:

$$\Delta x = 1.22\frac{\lambda}{2n} \tag{13}$$

Very close to our earlier result considering the cutoff of the propagator.
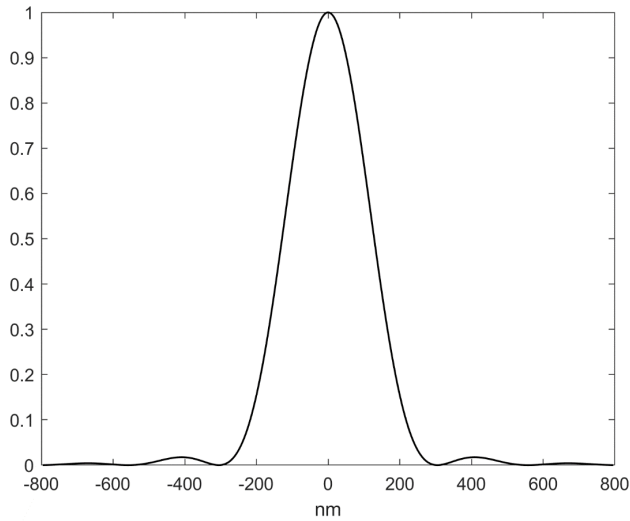

Figure 2-1: the Airy disk corresponding 500 nm laser light.

The full vector theory calculation of a strongly focused Gaussian beam is found in ref[26][26]. There important differences from the scalar approximation given by equation (13), for instance the dependence of the focused spot size on the incident polarization of the electric field, however the result is very close to that obtained by the scalar theory.

The optical transfer function (OTF)

From the discussing above we can understand the resolution limitation of an optical microscope in two ways, first by considering conditions the cutoff on the propagating planes in the angular spectrum representation, and the second by computing the finite extent of a focused laser spot, which when interpreted as the image of an ideal point source represents the point spread function (PSF) for the microscope. Both considerations yield a resolution that scales inversely with the measured wavevectors of light.

The connection between the image of a point source and the recorded spatial frequencies can be more explicitly seen if we consider the 2D Fourier transform of the PSF, or the image of a PSF in wavevector space, commonly called k-space (or frequency space).
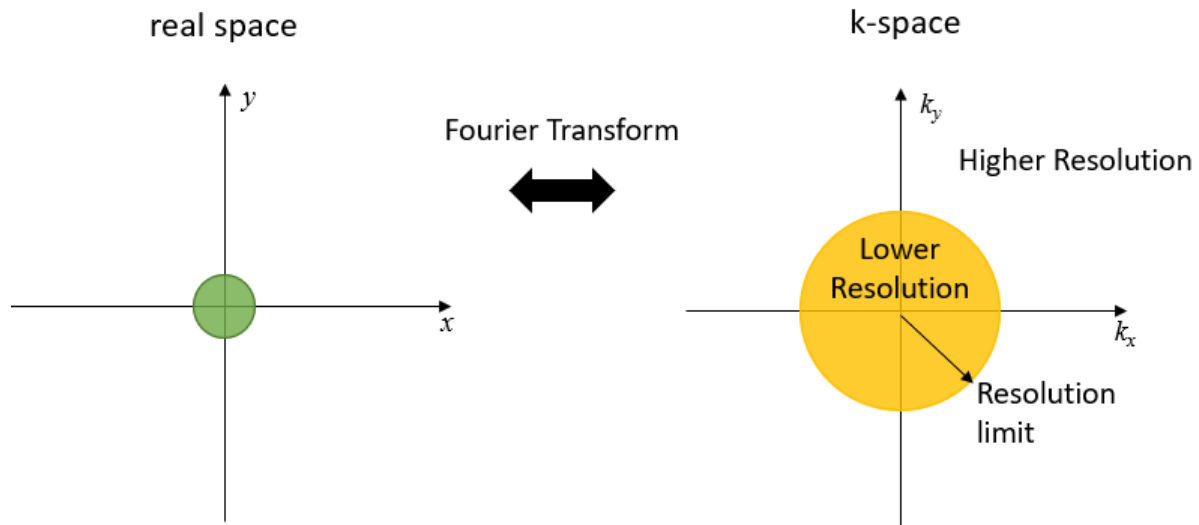
Figure 2-2: Illustration of the inverse scale relationship between the real space and k-space image of the point spread function. Figure 2-2 shows the Airy disk and its corresponding image in k-space. We that the k-space image has the same rotational symmetry, however now the distance between the central maximum and the first minimum is given by Abbe criterion.

To see how this limits resolution of an arbitrary object of interest, we positive that any general extended source can be represented as a sum of point sources. Then every point of the extended source is replaced by the PSF or

$$I_m[j] = \sum_{i<N} \text{PSF}[j - i]I_0[i] \qquad (14)$$

Mathematically, the above expression is a discrete convolution, which is in the continuous case is:

$$I_m(x,y) = \iint\limits_{-\infty}^{+\infty} \text{PSF}(x-x', y-y') \times I_0(x,y) dx' dy'$$

If we now apply the 2D Fourier transform to this expression it can be see shown that

$$\hat{I}_m(k_x, k_y) = \widehat{\text{PSF}}(k_x, k_y) \times \hat{I}_0(k_x, k_y) \qquad (15)$$

This result is more generally known as the Fourier-convolution theorem and generally applies to the convolution of functions. In our context we see that the k-space image of our object of interest is equal to the k-space image of the PSF multiplied by the "true" k-space image of our original object of interest. If we imagine that our original object is composed of all possible in-plane wavevectors then we see that the convolution of the PSF acts as a low-pass filter in k-space, setting all wavevectors higher that edge of the Fourier image of the PSF to zero.

Because of this filtering acting, k-space image of the PSF is thus referred to thus as the optical transfer function, OTF. We've seen that the resolution limitation is a result of the value of the OTF going to zero, or having finite support in k-space. However, the picture of the OTF also shows us that to increase the resolution we simply have to increase the support of the OTF in k-space, or equivalently shrinking the PSF in the image space.

Optical resolution beyond the diffraction limit: super-resolution

Equation (11) shows that this finite extent is due to the finite extent of included k-vectors in the calculation of the focused image our beam. In general, the spreading out of wavevectors due to the finite extent of the beam is called diffraction, and for that reason the resolution limit defined above is called the diffraction limit. The limitation of wavevectors has two sources. The first is due to the finite extent of the focused laser beam. This connection between the diameter of the focused beam is more clear if we write equation in terms of the smallest resolved angle or.

$$\theta = 1.22 \frac{\lambda}{\text{D}} \qquad (16)$$

This equation is commonly used in astronomy where Drepresent the diameter the light collection aperture. If we examine the above equation a lone, we might expect that we could simply increase D to obtain arbitrarily high resolution. However, the paraxial approximation when the angle formed by D and the distance to the sample becomes large, and equation (16) ceases to be a good approximation of resolution. Furtherrecalling the condition from the propagator, all propagating plane waves are subject to the condition $k_x^2 + k_y^2 < k^2$, ultimately keeping the resolution on order $\lambda/2$

In order to beat this limit, we must break one or more of the assumptions behind equation (12), which considered a uniform circular light field focused onto a plane. This can be done by using "structured light" or non-uniform illumination, leveraging light-matter interactions, or by including the evanescent fields. The first two approaches, used alone or in combination, are generally referred to as "far-field" techniques as light fields measured a still those with wavevectors with a positive k_z. Microscopes based on these techniques thus have very similar considerations to standard diffraction limited microscopes.

The first demonstration of far-field super-resolution was done by Betzig[28] using photo-activated localization microscopy (PALM). PALM works by using densely distributed photo-activated dyes. By controlling the activation such that no two neighboring dyes are activated simultaneously centroid fitting

can be used to estimate the position of the dye molecule to high precision, provided adequate signal-to-noise. A very similar technique uses the natural random switching on/off of dye molecules, called stochastic optical reconstruction microscopy (STORM).

Structured illumination microscopy (SIM), and leverages the Fourier transform properties of equations to increase the measured by creating difference, for beat, frequencies with the spatial frequencies of the structured light. Recalling the Fourier-convolution theorem, just as the multiplication of the k-pace object image by the OTF corresponds to a convolution by the PSF in real space, multiplication of the structure light in real space corresponds to a convolution in k-space. Structured illumination thus 'moves' higher k-vectors to within the support of the OTF by computing their difference with the structured light. Because the spatial frequencies of the spatially structured light is the known, the higher wavevectors can be extracted and a higher resolution image can be reconstructed[29].

A third approach that in a sense combines structured illumination with the fluorescence dye manipulation of PALM and STORM, is known as stimulated emission depletion, or STED. STED is employs of confocal scanning approach with two laser sources. The first is the fundamental gaussian mode that excites a fluorophore near saturation, and a second "depletion" laser, that has a spatial a "doughnut pattern" with the intensity of the laser spot going to zero at the center.For a sample that is densely labeled, the through the process of stimulated emission the depletion beam will quickly deexcite the doughnut area, resulting in only the smaller center to emit light through spontaneous emission of the fluorophore[30], thereby increasing the resolution.

The above four techniques constitute the principal "far-field" super-resolution approaches, as the higher wavevectors are extracted by engineering the interaction of propagating optical fields. That is the exciting light fields of PALM/STORM, SIM, and STED are all themselves subjected to the diffraction limit. Higher wavevectors are recorded by using these fields to carefully manipulate the light matter interaction as in PALM/STORM and STED to localized the emitter to smaller spatial extent, or by "folding in" normally lost wavevectors to within the support of the OTF by introducing wavevectors into the illumination pattern itself. Because the first three rely on very specific interactions the fluorescent dyes, the photophysical requirements are very strict and dense labeling is crucial to obtain a sample that corresponds well to true structure. Nonetheless well-developed biological tagging techniques have allowed their extensive application in biological imaging. An advantage of SIM is that it does not rely any specific light matter interaction can be generally applied to imaging any sample. However, because the structured light is itself diffraction limited, there is a constraint the smallest size, or highest wavevector, of the structured light, limiting the resolution to ~ 2x the classical limit.

Including the evanescent waves – near field optics

A different approach to beating the diffraction limit relies on including the evanescent components of the optical fields in the excitation and/or the emission or scattering of the object of interest. Whereas with previous techniques were called the far-field because they only collection propagating optical fields, here we can define "near-field" as distances from the source where the evanescent waves are still significant. By considering an evanescent wave with wavevector k, it is easily seen that this corresponds to distances $\Delta z \ll \lambda$.

For visible light, this corresponds to distances ~1 – 100 nm. Comparing this to working distance of high numerical aperture objective, WD ~ 0.2 mm, we see this distance is far below the usual length scales of optical microscopes. For this reason near-field microscopies make use of precision distance control techniques used in scanning probe microscopies, especially atomic force microscopy (AFM) and scanning tunneling microscopy (STM), where in addition to the tunnel current or the atomic force

deflection, the optical interaction is also recorded. Techniques that collect near-field signal using a scanning probe microscope fall under the name scanning near-field optical microscopy (SNOM or NSOM).

The concept of for SNOM was first approach was first proposed in 1928 by Synge[31] however the first experiment demonstration was not done until 1984[32,33]. Since then there have several different approaches to SNOM that have been implemented. In the remainder of this section we will briefly review the early approaches to SNOM. Modern SNOM approaches will be discussed at length in the following section.

Classical SNOM – small aperture probes

Early SNOM techniques were used to luminesce, fluorescence, from light emitting molecules and semiconductor. The nano-optical probes used in these early measurements were aperture probes, a diagram of which is shown in Figure 2-3. Aperture probes can be thought of as waveguides, with a small hole at the end that provides. Because the sample is brought very close to the aperture the image of the hole on the sample is approximately the same size as the aperture. In principle the desired resolution can be achieved by using an aperture of the appropriate size.

Typically apertures probes were constructed out of tapered optical fibers coated with metal[26,34], with the aperture cut by a focused ion beam or even by purposely crashing the tapered tip into the sample. While this design is very simple and in principle provides nanoscopic imaging without any background, it suffers from low throughput. This occurs because the boundary conditions of the type of waveguide shown are such that no propagating mode can make it to aperture and only evanescent waves transmit across the aperture[34]. This transmission further is incredibly sensitive to the aperture size, and decrease with the fourth power of the aperture radius[34]. This creates between signal and resolution, and practically limited choices of the aperture size to 50 – 100 nm. While below the diffraction limit by 2 – 3x, this is relatively modest gain of resolution. When relative ease of operation and instrumentation, significantly higher signal, and speed of conventional microscope is considered, the advantage this form SNOM is limited.
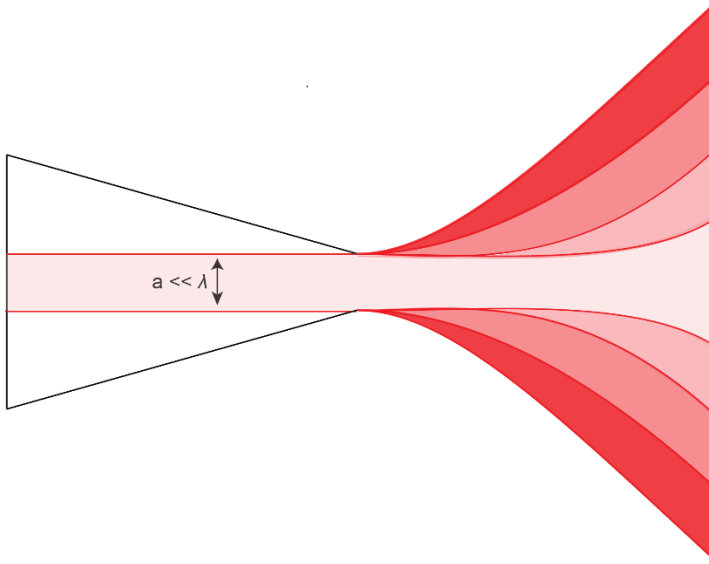
Figure 2-3: Diagram of aperture probe used in classical SNOM, adapted from[35]. The smaller the aperture the better the corresponding resolution, however the transmission also extremely diminishes.

The challenges of classical aperture SNOM led to a general collapse in its use, with even Eric Betzig, and early pioneer, abandoning the technique (He would later earn the 2014 Nobel prize in chemistry for his co-development of PALM). However, significant advances have occurred in design and fabrication of nano-optics, which has greatly improved the observed signal and resolution, making the above trade-off moot. The next section we will describe the main approaches in to SNOM and the consideration for the probes for each.

Section 3 – SNOM techniques

As mentioned in the previous section, there are several different techniques scanning near-field optical microscopy. The principle difference relates to the photon interaction with the sample that is being detected. We can divide the light-matter interaction into in-elastic and elastic, the former including phenomena such as fluorescence and Raman scattering, and the latter representing energy conserving scattering and absorption. The focus of this dissertation is one the former cases, particularly their application to the study of excitons in TMDs, however the techniques overlap in the design the nano-optical probe in some cases.

In this section we will discuss the two primary paradigms of probe for achieving nanoscale resolution in the visible and NIR. For the first paradigm, we will discuss an example of the next generation of aperture-based probes, the Campanile probe, and showing the resolution obtained when applying this probe to imaging luminescence from excitons in monolayer tungsten diselenide. In the second we will describe apertureless SNOM techniques, which are perhaps the mostly widely used near-field optical technique currently. Diagrams of the two types of probes are shown in Figure 3-1, a-SNOM probes shown on the left and a Campanile probe on the right. Following we will show an example of applying this technique to monolayer WS_2 and WSe_2, showing both hyperspectral Raman and photoluminescence imaging, how these techniques can be combined with complimentary scanning probe techniques, namely atomic force microscopy and Kelvin-probe microscopy.
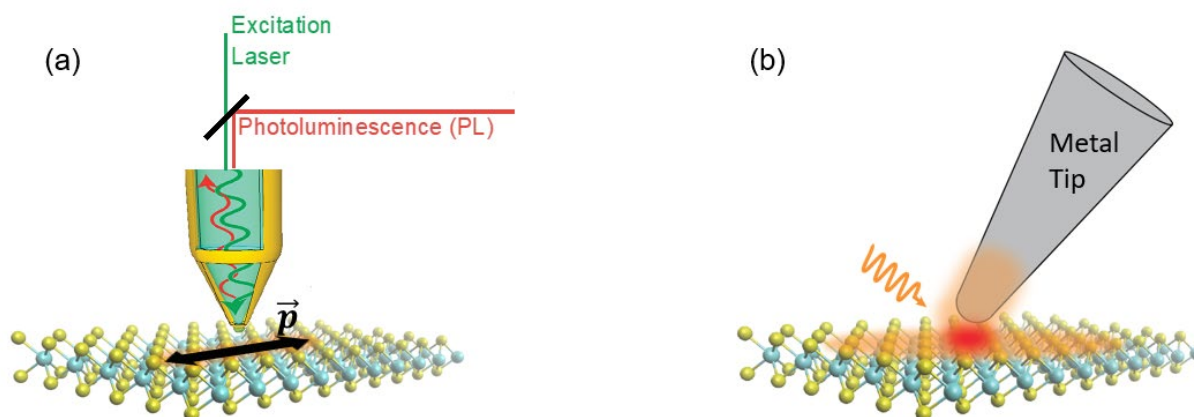
Figure 3-1: Diagrams of the common SNOM techniques: (a) aperture type SNOM using the Campanile probe, (a) apertureless-SNOM (a-SNOM), where an metalized scanning probe tip acts as an optical antenna to convert the evanescent near-field to the far-field.

Aperture probes – the near-field to far-field transduction problem

As was mentioned at the end of Section 2 discussing ways to circumvent the diffraction limit of conventional microscopes, a fundamental problem of classical SNOM using aperture tips was the low signal throughput. This is directly attributable to the electromagnetic boundary conditions that prevent any propagating modes from escaping the waveguide, allowing only the evanescent waves to pass[26,34].. When coupling a laser through the aperture probes the transmission of the light could be as small as one in a million photons are "tunneled" through the small aperture[26], and the remainder of the light is reflected back into the waveguide, in this case an optical fiber. This abysmal coupling efficiently obvious limits the

signal and practically limits the use of the probe use to either collection or excitation. Ideally both collection and excitation would be done at the desired scale.

The limiting problem here is transmission of the far-field light to the near-field, which is sometimes referred to and the optical transduction. Ideally one would want to "convert" significant fraction the propagating far-fields to the near-field. In the context of a waveguide, this requires finding a way to couple more light to the probe aperture in spite of the fact that there is no propagating electromagnetic mode at the apex of a tapered waveguide. The strategy to do this it to coupling light plasmons which can be compressed to sub diffraction volumes efficiently.

A plasmon is wave-like excitation of the free electrons in a metal. They have be most often studied in the "noble metals" with many free electrons, that is Au, Ag, and Cu, but are general phenomenon of free electrons, existing even in wide bandgap oxides[36]. Plasmons can be excited both the metallic bulk and of the surface, where, as one would expect, they are referred to as surface plasmons. It is latter that will prove to be key to efficient coupling of propagating light to the nanoscale.

When a wavelike excitation in a material couples strongly to light, the resulting coupled-wave excitation has modified properties from the material excitation or light alone, and is called a polariton[37]. A surface-plasmon coupled to light would then be a surface-plasmon-polariton (SPP). To get some insight into how these SPPs are utilized in near-field optics, consider an electromagnetic wave that is confined to propagate along the surface of a metal and an insulator, for example an Au/SiO$_2$ interface. Figure 3-2 shows a schematic of such SPP at metal-insulator interface.
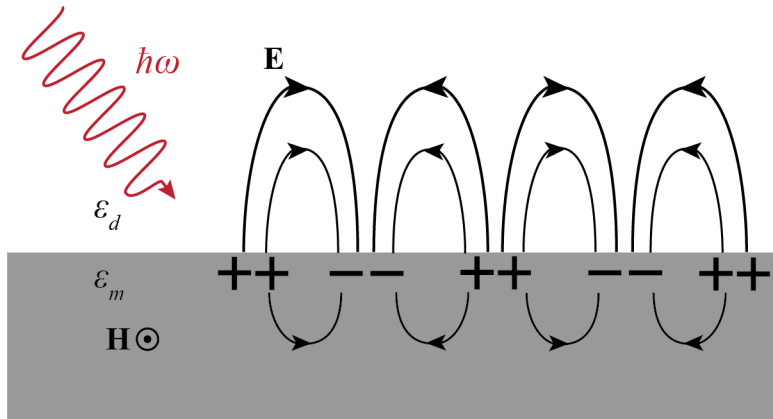
Figure 3-2: diagram of a photon exciting a propagating surface plasmon polariton (SPP) at a metal-dielectric (metal-insulator) interface.

For such a surface wave to satisfy the electromagnetic boundary conditions of Maxwell's equations, it can be shown that the k-vector for the surface bound wave must be:

$$k_{\text{spp}} = k_0 \sqrt{\frac{\varepsilon_r^m \varepsilon_r^i}{\varepsilon_r^m + \varepsilon_r^i}}$$

Where ε_r^m and ε_r^i are the real parts of the metal and insulator dielectric functions respectively. We see that the surface k-vector is related the k-vector in vacuum by a ratio the dielectric functions of both the insulator and the metal. Plugging in the values for an electromagnetic wave in the visible (say 600 nm wavelength in vacuum) and corresponding values of the dielectric functions for SiO$_2$ and Au we see that

k-vector for the SPP is larger. Recalling that higher k-vectors have smaller wavelengths, and thus can be focused to smaller dimensions.

Utilizing SPPs – the Campanile probe

The basic structure of the campanile probe is shown in Figure 3-3. It consists of a metal insulator metal waveguide that leads to a taper. The angle of the taper is chosen to excite the surface plasmon polaritons, which then propagate along metal-insulator boundary to tip apex. The focusing of the plasmons is adiabatic[38], leading to a high conversion efficiency of the photon to plasmon, approaching 50%. The apex of the tip then acts as the local source of optical field that can then be scanned over the sample. The first demonstration of the campanile concept was performed in 2012 on-chip[39], followed shortly by its integration into a shear-force AFM which was used to image nanoscale photoluminescence from indium phosphate nanowires[40] and later monolayer MoS_2[41]. The high transduction of light from the near-filed to far-field allows it to be used both for both excitation and collection. This excitation of is ideal for extended samples such as monolayer transition metal dichalcogenides.

While the Campanile probe has significant advantages in nano-imaging of photoluminescence, the complex structure makes fabrication of the probe challenging. In the previous examples the Campanile probes were fabrication at the ends of optical fibers. The fiber is first chemical etched into a taper structure. The structure of the waveguide and taper are carved using focused ion beam milling with electron-beam metal evaporation of Au with a Ti adhesion layer. The final step is the carving of the nanoscale gap with a focused ion beam. This approach produces high quality tips which very reliable and importantly reproducible confinement of the optical field. However, the heavy use focused ion beam milling in the fabrication imposes a sever bottleneck on production of tips. In the next sub-section we will present an alternative structure of the Campanile probe that allows batch fabrication of several tips at once with need for focuses ion beam milling, and show the application of these tips to the imaging of nano-PL of monolayer WSe_2.

22
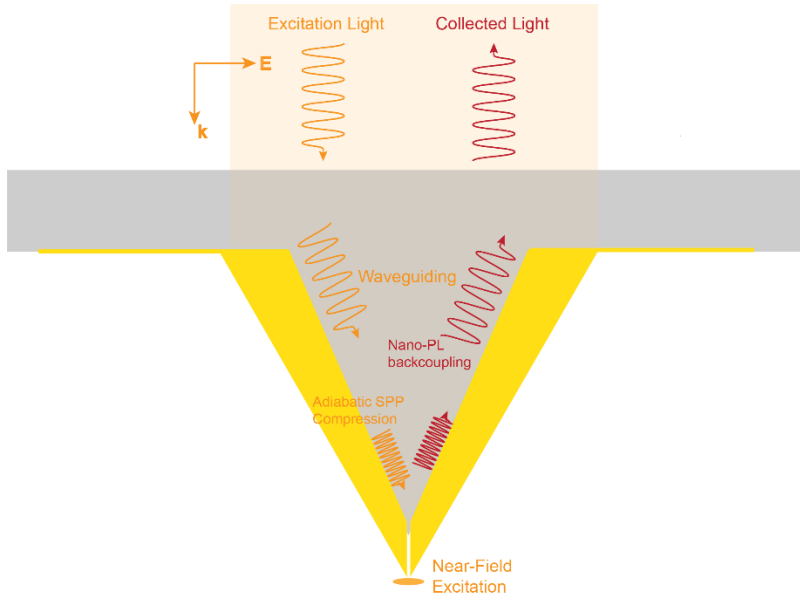
Figure 3-3: Basic structure of the campanile probe.
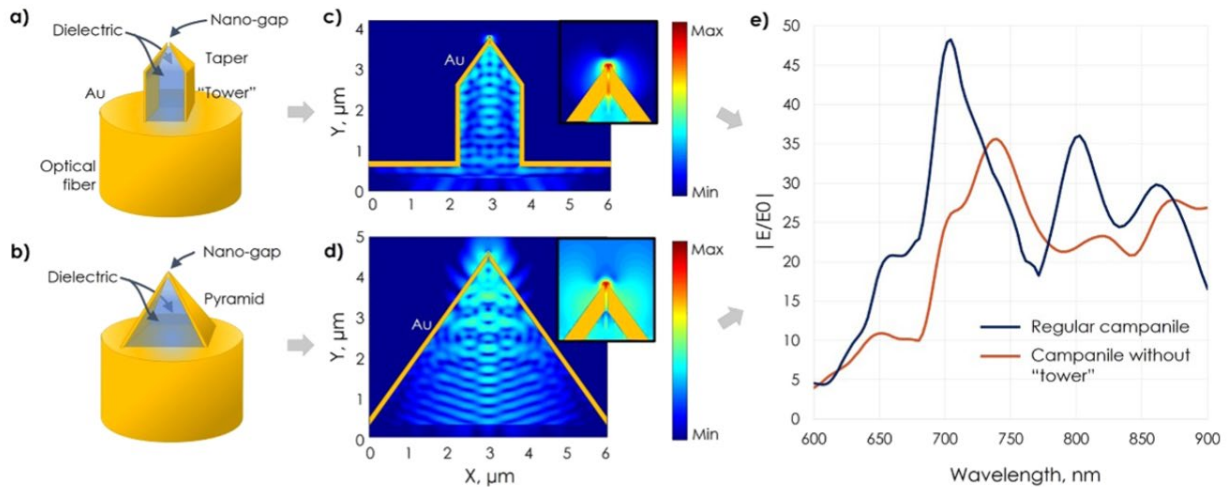
Figure 3-4: comparison of the original campanile structure (a) and the modified structured where the initial "tower" section of the waveguide is excluded (b). (c), (d) electric field intensity maps simulated with FTDT Lumerical software. (e) comparison of the wavelength/trasnsmission response of the probe designs.

Imprinted Campanile probes

A comparison of the two campanile structures in shown in Figure 3-4. The top of panels (a) and (c) are the "traditional structure" consisting of a MIM waveguide that is initial parallel, the "tower" in the diagram, before the tapered section that couples the propagating SPPs. The alternative design is shown in the bottom panels of (b), and (d). Here the parallel section of the waveguide is removed and replaced by a larger base of the taper, in 3D forming a purely pyramid structure. To test if this structure performed in comparable to the Campanile fabrication previously, we performed finite-different-time-domain simulations (FTDT) of the electric field intensity, compared to the normal campanile structure. The results of the simulations are shown the panels (c) and (d) referring the to the normal and alternative

campanile structures respectively. There are some differences in the mode character, however both structures have the greatest intensity confined to the nanoscale volume near the apex of the tip, seen more clearly in the insets panes (c) and (d). Figure 3-4(e) shows a comparison of electric field enhancement at the tip apex for the traditional and alternative campanile structures, showing a similar enhancement of the electric over a broad range wavelengths though notably has smaller peaks in the spectrum providing a more uniform enhancement in the near-infrared.

The simplification of the campanile structure allows it to be fabricated without the time-consuming cutting of large sections of optical fiber with the focused ion beam. The method to do this is use UV-cured nano-imprint lithography. The central idea is to "stamp", or imprint, the dielectric center of the campanile structure onto the core of a cleaved optical fiber then evaporate metal layers in such manner to the leave two sides and gap uncoated. The process consists of first etching the master mold into a double-polished Si/Si_3N_4 wafer using electron-beam lithography followed by reactive ion etching of silicon nitride. The pyramid is then created in the mastermold by KOH etching. An intermediate mold is then created in Orthostamp on top a quartz substrate. Using a focused ion beam (Zeiss Orion Nanofab) a slit is carved into the apex, with will serve to stop the nanogap from being closed during metal evaporation. The mold is then inverted again in to Orthocomp to secondary, which can then be used to imprint the campanile structure on the base of the optical fibers. Full details of the simulations and fabrication can be found in G. Califiore, A. Koshelev, and T. Darlington et al. [42].
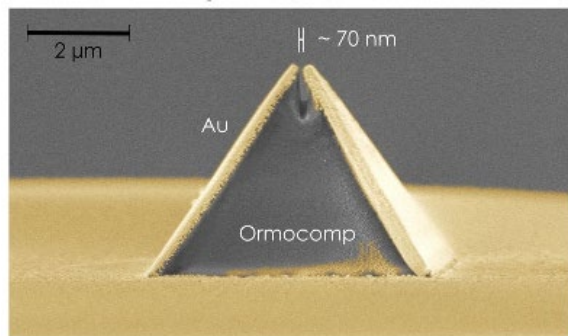

Figure 3-5: SEM micrograph of finished imprinter campanile. Gold has been false colored for added clarity.

Figure 3-5 shows a SEM micrography of a finished tip, with the evaporated gold false colored. To test the performance of the tips we imaged a field a polystyrene beads fields with fluorescent dyes (AlecFluor 680 nm). The diameter of the beads is 40 nm seen in, near half the gap size is seen in Figure 3-5, and provides bright sample with consistent intensity and topographic profile. The microscope to image the same was an atomic force microscope with an NTMDT NTEGRA spectrometer and an Andor CCD camera.

To load into AFM, optical fibers with imprinted campaniles need to be attached to oscillator while will be used to sense the surface when the tip approaches. This can be done by locking into the normal or shear force depending on how the oscillator is oriented relative to the surface. In this study we use the shear force modality, with the oscillation direction parallel to the surface. The oscillator frequency of the quartz tuning fork is 150 kHz. Tuning forks in this frequency range are fairly small compared common tuning forks around 25 kHz. Because of this smaller size attachment of the optical fiber of the campanile probe represents a non-trivial perturbation, and careful attached must be paid to the attachment of the fiber to the tuning fork arms, resulting in a low quality-factor (Q-factor). Since the sensitivity of

shear force feedback is directly related to the Q-factor[26], if it is to low the feedback will be poor, risking the tip. For our system we found that minimum Q-factor to be ~100.

Figure 3-6(a) shows atomic force micrography for the AlecFlour 680 beads drop cast onto a fused silica glass slide. The beads are approximately spherical, however the micrograph displays a double-imaging character separated by ~ 100 nm, is indicative of the "double-tip" seen in the campanile diagram of Figure 3-3. This distance is typically larger than the true size of the gap.

Figure 3-7 shows a diagram of the light coupling setup for a hybrid excitation/collect nano-PL experiment with the campanile probe. The HeNe laser light is coupled into the back of the fiber with a 10x microscope objective from Olympus. To get efficient coupling into the core of the optical fiber the incident focused laser needs to be within the admittance angle of the core of optical, which is practice means matching as close as possible the respective numerical apertures of the core and focusing lens. The numerical aperture of the common PlanApo Olympus 10x objective is $NA_{obj} \sim= 0.21$, which is slightly larger the numerical aperture of the fiber ~ 0.1-1.4 of single mode optical fibers in the visible and NIR. To make them better matched the back aperture of the objective can be underfilled, with lowers the effective NA of the focused laser light.
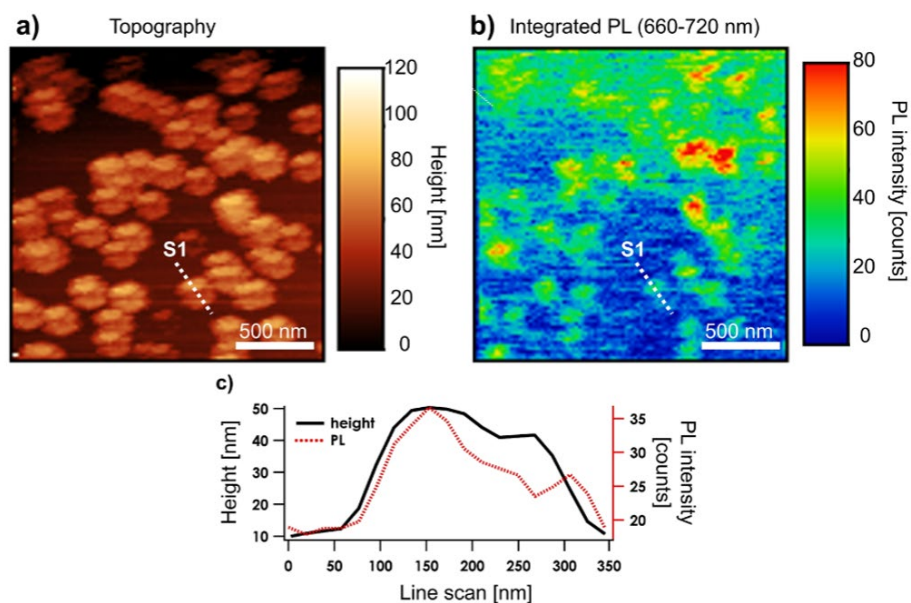
Figure 3-6: Demonstration of the imprinted campanile on polystyrene beads containing fluorescence molecules.

Figure 3-6b shows the integrated nanoscale photoluminescence (nano-PL map that was recorded simultaneously with the AFM topography in Fig. 3-6a. The fluorescence molecules inside the polystyrene beads emit at a mean wavelength of 680 nm, and were excited by a HeNe gas laser at 632.7 nm. Comparing the results of Figure 3-6a with 3-6b the integrated nano-PL has approximately the same resolution. This is more clearly seen in Figure 3-6c which shows a comparison of line-scans from the same location in Figures 3-6a and 3-6b, which should the optical resolution is slightly larger than AFM micrograph, but still clearly below the diffraction limit.
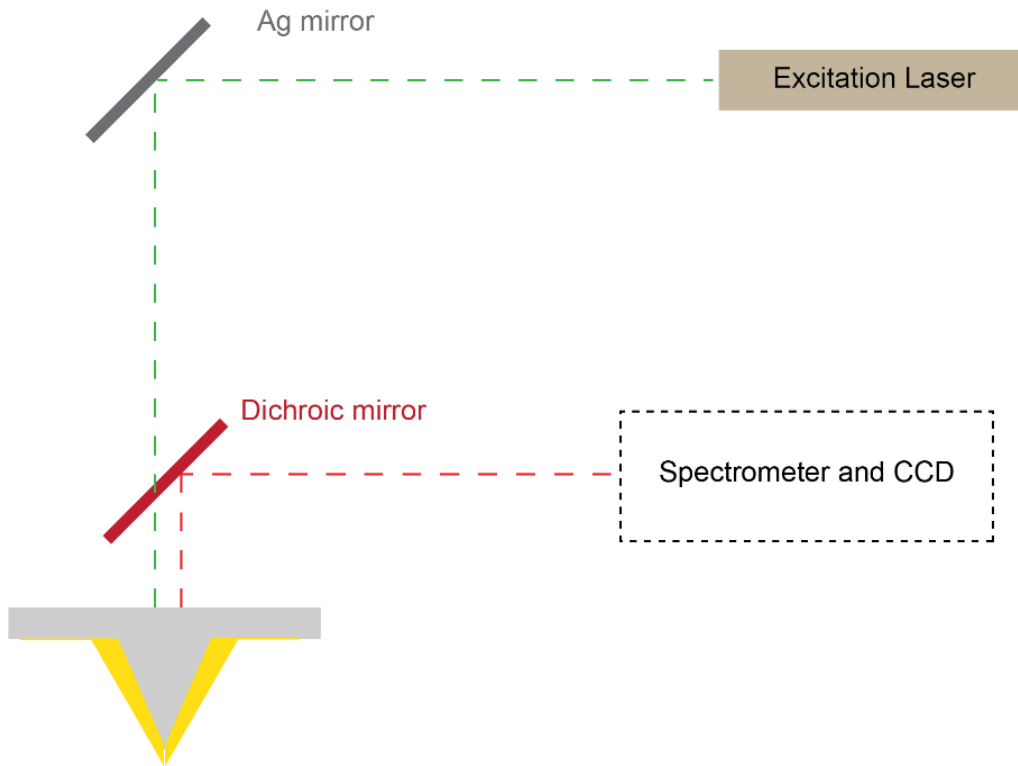
Figure 3-7: Diagram of light coupling in hybrid excitation/collection mode using the campanile probe.

While the integrate nano-PL in Figure 3-6 is proof of the concept that the alternative campanile probe design functions a near-field tip. To give a more direction comparison we apply the imprinted campanile tip to imaging a monolayer TMDs. In Figure 3-8 we show a comparison integrated PL of the exciton emission from monolayer WSe_2 taking with the imprinted campanile probe and scanning confocal microscope. The sample is a heterostructure of WSe_2 and h-BN were the 1L-WSe_2 is sandwiched between a larger bottom flake of h-BN (~200 nm thick) and a thing (~3 nm thick) top layer that leaves part of the WSe_2 exposed. A diagram of the sample structure is shown in Figure 3-8(a). Fig. 3-8(b) shows confocal scan of the entire 1L-WSe_2 flake. Figure 3-8c shows a zoom of the confocal indicated by the white dashed box. The same area was reconstructed using multiple overlapping near-field scans in Figure 3-8(d), which clearly show considerable improvement compared the far-field map.

Figure 3-8: Applying the imprinted campanile to measurements of photoluminescence on monolayer WSe$_2$ on hexagonal boron nitride. (a) heterostructure diagram of the sample. (b) Far-field confocal photoluminescence map of whole sample, and (c) zoom in on the same area measured by the campanile probe in (d). Excitation wavelength was a HeNe 632.75 nm gas laser for both near-field and far-field maps. The microscope for each was the NTMDT NTEGRA Spectra.

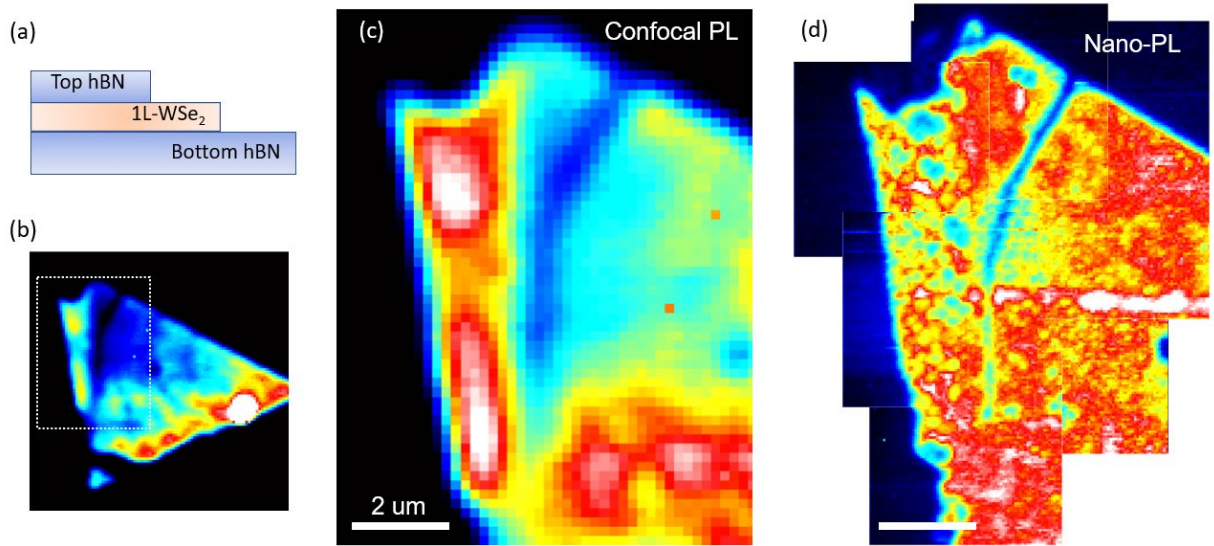## General principals and operations of apertureless-SNOM

Apertureless SNOM (a-SNOM) is perhaps the most common form scanning near-field measurements in use by researchers. The difference from the classical SNOM with small aperture tips on optical fibers, and the recently developed advanced aperture tips like the Campanile probe, is obvious from the name. Unlike the Campanile probe were the tip effectively replaces the microscope objective for excitation and collection, a-SNOM the near-field probe is used to locally enhance the optical field. All probes with a-SNOM are variations of the design shown in Fig 3-1b: a metallized AFM or STM probe that acts as a local source which provides the sub-diffraction resolution.

The near-field probe enhances the optical field primarily through two mechanisms. The first of these is a geometric effect, the sharpness itself will locally focus charge and polarization[34]. The second, and most important in context for in-elastic light scattering, is the enhancement by excitation of local surface plasmon resonances. Recalling that the campanile "transduces" the far-field to the near-field, we see the operation of the a-SNOM probes has commonalities. Unlike the campanile waveguide, where the momentum mismatch of the photon and plasmon is provided by the slope and index of the insulator, the tip of the probe itself bridges the missing momentum[34]. Again, similar to the campanile, the SPPs may be excited over a broad range around the tip, and then adiabatically propagate to the tip apex, resulting in a large local field enhancement.

While the picture above is true in principal and is well reflected in the FEM and FTDT simulations of sharp metal probes on various substates[43-45], in practice it is unclear if this reflects the enhancement of a-SNOM measurements under experimental conditions. A primary reason for this the

quality of the metal film is very difficult to control through common deposition techniques at the length scales of the radius of curvature of common AFM probes (typically reported ~ 10 nm), especially considering its complex 3D shape. Realistic metal films will thus have significant surface roughness, which is difficult to model in simulation. Recent reports have demonstrated that control of this roughness, in particular generation of some particular roughness, is critical for common a-SNOM techniques such as tip enhanced Rama scattering (TERS) and tip-enhanced photoluminescence (TEPL)[46]. We will comment more on this subject in Probe's sub-section below.

Coupling light to the a-SNOM probe

A primary difference between aperture based near-field methods and a-SNOM is coupling of light to the probe. As we discussed because all the action of converting the far-field propagating light the evanescent near-field and the reversal happens at the metal probe apex (in other words the tip), external optics are needed to initially focus the light and collect the emitted or scattered light. In Figure 3-9 we show an expanded diagram of the tip shown before in Figure 3-1b, which also the optical setup used for much of the SNOM data presented here. An a-SNOM probe is illuminated from the side onto the tip apex and the back scattered light is collected along the same optical path. An advantage of this geometry is that it allows non-transparent samples to be addressed. If a sample is transparent the light can be incident from the bottom. This latter approach was the one employed in the first demonstrations of TERS in 2000 by Stockle et al. and Hayazawa et al.[47,48]. While the later illumination method is less flexible concerning samples, in practice the geometry allows higher numerical aperture for the focusing optics to be employed as the angle of side illumination forces the optics to often be several centimeters away from the sample.
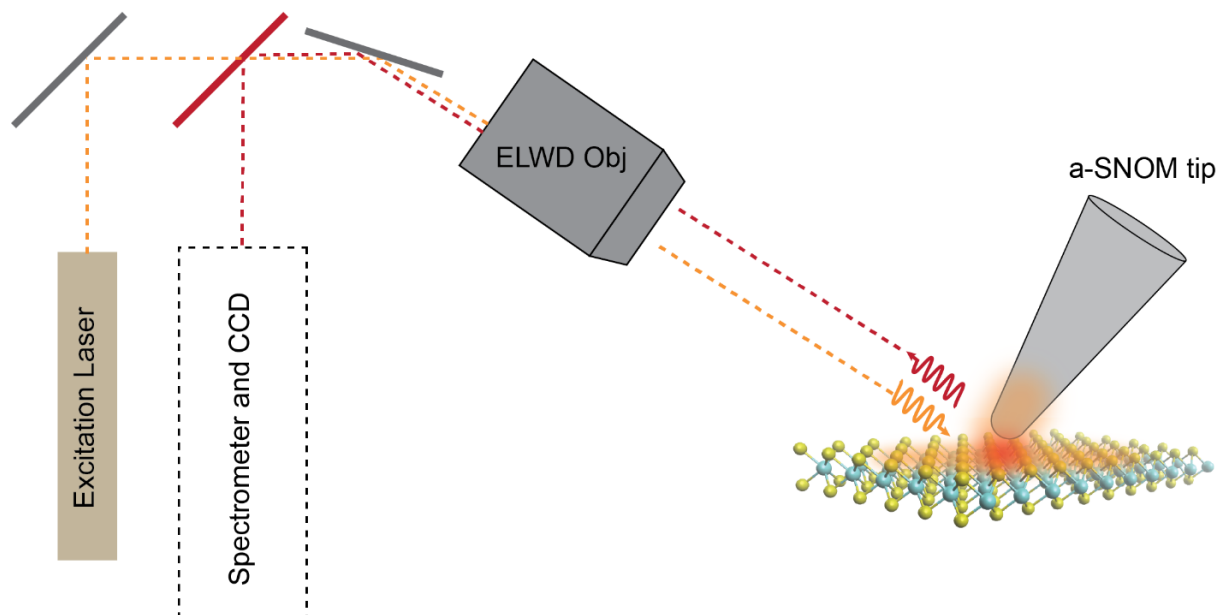

Figure 3-9: Diagram of coupling light an a-SNOM probe for PL or Raman spectroscopy.

A distinction in the application of the two illumination schemes is the polarization of the incident light. With the geometry a metallized SPM probe, strongest resonance by the surface plasmon polaritons

is achieved with the electric field is polarized along the long axis of the probe[43]. In the side illumination scheme this corresponds to the *p*-polarization of the incident laser light, and achievement of this polarization can be done using the typically already linearly polarized laser light and use of the wave-plates and polarizers. In the cases of bottom illumination *p*-polarized light will not create a strong electric field at the point, since the probe lies along the z-direction. In fact, for ideal Gaussian beam the z-polarization is exactly zero at the focus[26]. In this case, it is necessarily to for the light to be radially polarized prior to the entering the focusing objective, which will produce a significant z-component of the electric field at the focal point.

In principal the above description is all that is necessary for coupling the light to a-SNOM probe. In practice the small radius of the prove makes alignment of the light critical for observation of significant enhanced, which is especially important for weak processes such as Raman scattering. Precise control of the focus as well as the angle of the objective (or other focusing element) is required, both for the initial alignment as well as stability of far-field to near-field coupling as the probe is scanned over the sample.

Probe considerations of a-SNOM for elastic (s-SNOM) and inelastic near-field techniques

In any a-SNOM experiment the enhancement of the probe is typically the critical factor that determines if observation the near-field light from the sample is possible. This is especially true for the extended 2D samples like the monolayer TMDs, where the whole illuminated area of the focused light in Figure 3-9 interacts strongly with light.

A rough estimation of the enhancement requirements can be done by consideration of the far-field interaction area with the probe interaction areas with the sample. Assuming both areas to be circular, a 10 nm radius of the near-field compared to a 250 nm radius of the far-field, and a linear light matter interaction, the ratio of the near-field to far-field is $\sim 10^{-3}$ emission intensities. Thus, for the near-field to on the same order of the far-field, the enhancement of the electric field from probe must be $\sim 10^{1.5}$. This enhancement requirement is the chief motivation for exploiting the strong light confinement offered via SPP resonances discussed previously.

The enhancement requirements are relaxed somewhat in the case of the scattering SNOM measurements. Because the light-matter interaction with the tip and the sample is elastic, the back scattered light from the tip and sample can be mixed with the excitation source in an interferometer which allows for homodyne and pseudo-heterodyne amplification of the near-field light. By use of lock-amplifiers at the probe oscillation frequency, these methods can efficiently discriminate near-field and far-field contributions of the scattered light[34]. In s-SNOM experiments strong and consistent scattering of probe is desired, which can be provided by metals that do not have strong surface plasmon resonances. While the use of lock-in detection and interferometry does require significant instrumentation above that of that of nanoscale Raman or PL, the ability to reliability suppress the far-field background has made s-SNOM one of the most successful a-SNOM techniques to date.

However, these methods of near-field/far-field discrimination are not amenable to inelastic light-matter interactions such spontaneous Raman scattering and photoluminescence. Instead the enhancement of the probe alone must be sufficient to observe the near-field light above the far-field light. Strongly enhancing probes are thus critical in these near-field experiments, surface Raman scattering or photoluminescence, where the far-field is always present. This entails near-field probes with strong surface plasmon resonances and probes used in a-SNOM Raman and PL measurements typically use silver and gold.

In the earlier section discussion coupling light to a-SNOM probes, we mentioned that surface roughness of the metal film has proved to be an important parameter for high-enhancing probes[46]. Because the surface roughness of the deposited metal film is incredible sensitive to many parameters not readily in the experimenter's control, a-SNOM production is rampant with "black magic," in that procedures that produce reliable a-SNOM probes in one laboratory, fail when replicated in another for unclear reasons. This variation itself is part of the motivation behind development of the advance alternative SNOM techniques, including alternative probes like the Campanile probe, but also measurement approaches to amplify the near-field to far-field contrast such as pseudo-heterodyne amplification in scattering-SNOM. Another such measurement approach is gap-mode TERS/TEPL which can be used to study thin samples deposited on a metal substrate, and can provide significantly higher enhancements compared to a probe alone[44]. Gap-mode type measurements will comprise a significant fraction of this manuscript and we will discuss in a following sub-section.

As an aside, the focus on controlling the surface roughness of metal film for TERS is ironic given that development of TERS itself was motivation by similar challenges in controlling the surface conditions in surface enhanced Raman spectroscopy (SERS)[48]. That a problem that plagued SERS should similarly impact its offspring, TERS, is perhaps not surprising in hindsight, given the similarities in the techniques. With comparatively complex geometry of an AFM or STM compared to a flat substrate, it is likely that roughness challenge for inelastic a-SNOM studies may remain for some time.

Example: TERS/TEPL of WS$_2$ monolayers

As an illustration for the characterization power of a-SNOM, in this section we will show characterization of monolayer tungsten disulfide grown by chemical vapor deposition using high purity tungsten oxide precursors. Chemical vapor deposition is one of the most widely used methods for controlled growth of high quality TMD monolayers. However, the growth kinetics are perniciously complicated, and even small changes in the growth conditions can lead to significant differences in the material quality even within the same monolayer film. Example of these effects could be differences in the concentrations of point defects in the crystal, or 1-D defects such as grain boundaries, both of which affect the local electronic environment a free charge carrier, or an exciton, interact with. Because the length scales are typically in the 1 nm – 100 nm, discriminating their corresponding effects is challenging with conventional optical microscopy

In order to elucidate some of these effects in the growth of WS$_2$, we used a-SNOM to perform TEPL and TERS of a single monolayer. The experimental set up is diagramed in Fig. 3-9. The microscope used here is a Horiba Omegascope-R, which is a-SNOM optimized for side illumination, with the incident excitation laser striking the a-SNOM probe at angle of 65 degrees from the vertical. The probe used was a metallized, angled Si tip (ACCESS tip from Applied Nanostructures). The plasmonic active coating for the tip is evaporated silver, with a thin metallic Au/Ag layer to protect the silver from oxidation. Excitation wavelength for TEPL was 532 nm, which is will above the ground exciton wavelength in WS$_2$ of ~ 620 nm. Since the expected TERS signal is expected to be low, owing to the non-gap mode measurement geometry, a wavelength of 637 nm was used for Raman measures due to its closer resonance with the exciton. AFM topographies were first acquired in semi-contact mode with the same probe prior to optical measurements.

Figure 3-10 breaks down the results for the AFM, TERS, and TEPL imaging of a single flake. Fig. 3-10a shows the integrated tip-enhanced photoluminescence of the single CVD grown WS$_2$ flake on 300 nm SiO$_2$/Si chip. The outer dashed line denotes the boundary of the flake defined from image registration with the topographic image shown in the inset. As can been, the TEPL intensity is notably

less at the edge than in the interior of the flake, and decays to zero detected intensity before the flake terminates. The second dashed line is a user defined boundary between the edge and interior, which will be discussed shortly below. In addition, there is clearly major inhomogeneity within the interior of the flake on the length scales of 10 – 50 nm, limited here only by the sampling of the image.
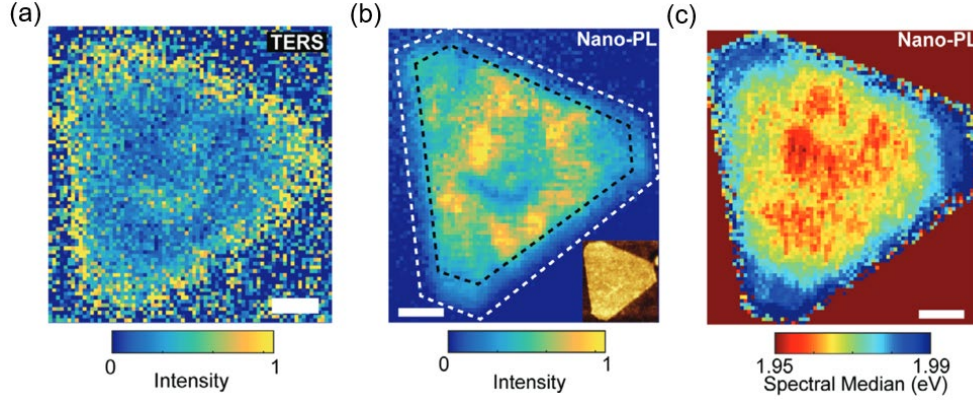

Figure 3-10: Hyperspectral TEPL/TERS imaging of monolayer WS$_2$ on SiO$_2$ grown by chemical vapor deposition. (a) integrated TERS scattering intensity between 100 – 1900 cm^{-1}. (b) Integrated TEPL of the exciton luminescence on the same flake, inset is the semi-contact (tapping mode) AFM topography. (c) Spectral median of TEPL emission. All scale bars are 250 nm.

Figure 3-10b shows the spectral median of the TEPL spectra for each measured pixel in the image of Fig 3-10a. The spectral median is defined by the following formula,

$$\text{Specral Median} = \varepsilon \quad \text{such that} \quad \frac{\int_{-\infty}^{\varepsilon} \text{PL}(\varepsilon\prime)d\varepsilon\prime}{\int_{-\infty}^{+\infty} \text{PL}(\varepsilon\prime)d\varepsilon\prime} = \frac{1}{2} \qquad (1)$$

the point where the integrated intensity defined is equal to have the total intensity. The spectral median provides a computationally quick method to detect shifts in the emission peaks, and emergence of stronger or weaker emission tails due wither to new peak appearance or asymmetric changes in linewidth. Because the spectra median integrates over many spectral bins, it has a sensitivity advantage over standard curve fitting, which can become ill-defined when the signal-to-noise is poor. An example of this seen can be seen by closely examining the spectral median map with the integrated TEPL intensity map. Even as the detected integrated intensity approaches zero, the spectral median still clearly shows the blue shifting spectral of the exciton emission. Comparing Figs. 3-10a with 3-10b clearly shows that the dim edge clearly corresponds with the blue shift in the spectral median. Even considering the variation of in the interior, the differences clearly mark the edge as having different excitonic environment.

Figure 3-10c shows the integrated TERS intensity of the WS$_2$ flake between 100 – 1900 cm^{-1}, after the PL background from the main and charged excitons has been removed by single exponential fit, excluding the spectral bins which also correspond the Raman scattering bins for 1L-WS$_2$, ~ 355 and 407 cm-1 for the E' and A' modes respectively[49]. The first order Raman modes in 1L-TMDs, the E' is 2-fold degenerate mode that corresponds to the vibration in the plane of the material, while the A' corresponds to the non-degenerate out-of-plane vibration, from the D$_{3h}$ symmetry. The A' should be preferentially excited since the dominate polarization of the plasmon-polariton of the probe is in the normal direction to the sample surface. However, it has been shown for resonant TERS that both parallel and orthogonal modes can be significantly enhanced, as was demonstrate on STM TERS of molecular copper naphthalocyanine[50]. (This is reflected in the TERS spectra shown in the following figure). The TERS

intensity map shows that there is greater scattering intensity at the edges compared to the interior, in opposite fashion to that of the integrated TEPL map discussed previously. Closer inspection shows areas of increased TERS intensity in the interior, which corresponds to raised areas in the AFM topography of less than 1 nm, possibly indicating a grain boundary.
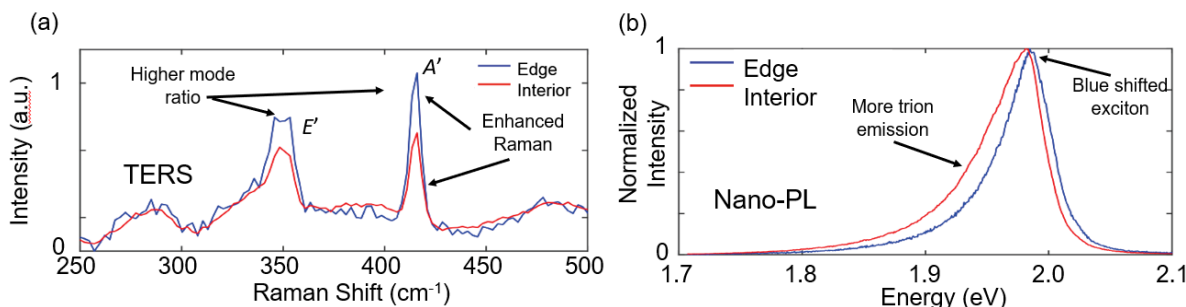

Figure 3-11: Spatial averaged spectra of the edge and interior region for the TERS (a) and TEPL (b) hyperspectral maps shown in Fig. 3-10.

The clear distinction of the edge versus the interior of Figure 3-10 of both the integrated TERS and TEPL maps invites an ensemble analysis, whereby we average the spectra over all the edge pixels and interior pixels defined by the dashed regions in Fig 3-10a. This provides the strong single-to-noise as one would have in far-field spectroscopy, but retains the spatial discrimination of the near-field. The spatially averaged spectra are shown in Figure 3-11, (a) the normalized photoluminescence and (b) Raman scattering spectra. Blue curves in both correspond to the edge, while the red correspond the interior Regions. Considering the PL first, it is straightforward to see that spectra of the edge if marked different from the interior, with the interior pixels in addition to having greater intensity the peak positions show a clear redshift to lower emission energies. The emission tail is also greater in interior compared to the edge. Two phenomenon could explain this: that is greater charged exciton (trion) emission, or changes in the exciton-phonon coupling. Trions are well known and commonly observed excitation in all TMDs[51], and inhomogeneity in trion emission at the nanoscale has been clearly observed in MoS_2[41]. Changes in the exciton-phonon can also change the PL linewidth, however this effect was shown to be small in WS_2 due to already comparatively weak exciton-phonon coupling[52]. For this reason, we attribute this effect to a greater formation and emission of trions.

The spatially averaged TERS spectra shown in Fig. 3-11a allows a way to discriminate between these two effects through observing A' mode relative to the E'. Because the former represents an out of plane vibration normal to the plane it has stronger coupling to charges via electron-phonon coupling[53]. Greater carrier concentration thus leads to stronger softening of A' mode, thereby changing the relative intensity of the two modes. Examination of the Fig. 3-11a show that in addition to greater scattering intensity for both modes, there is a higher ratio of the A' to E' modes, compared to the interior, consistent with the observation of the greater trion emission.

The difference between edge and interior Raman and photoluminescence properties can possibly be attributed to changes in growth conditions during conversion of the oxide precursor to the monolayer TMD. This has been observed in monolayer CVD grown MoS_2, using aperture based nano-PL using a plasmonic Campanile tip[41].

Example: Combined TERS and other SPM imaging and gap-mode a-SNOM

Finally, in this subsection we will show that near-field optical imaging techniques can be readily combined with other scanning probe modalities other than standard AFM. Here we will so characterization of the CVD grown WSe_2 transferred to Au and Ag substrates, with Kelvin-prove force microscopy (KPFM) and tip-enhances Raman scattering. KPFM is electronic scanning probe technique uses a controlled AC-bias between a conducting AFM tip and conductive substrate, which forms a capacitor. If one considers the energy of the capacitor, it can be show that when there is potential difference between the tip and sample there will be a force acting on the cantilever which can be detect using the same feedback controls as AFM. The recorded signal is referred as the contact potential difference (CPD) which is related to the work function of the sample. Because the probes used in TERS are also metallic, the same probe can be used for both measurements, allowing convenient and comparable characterization of the CPD and TERS signals.
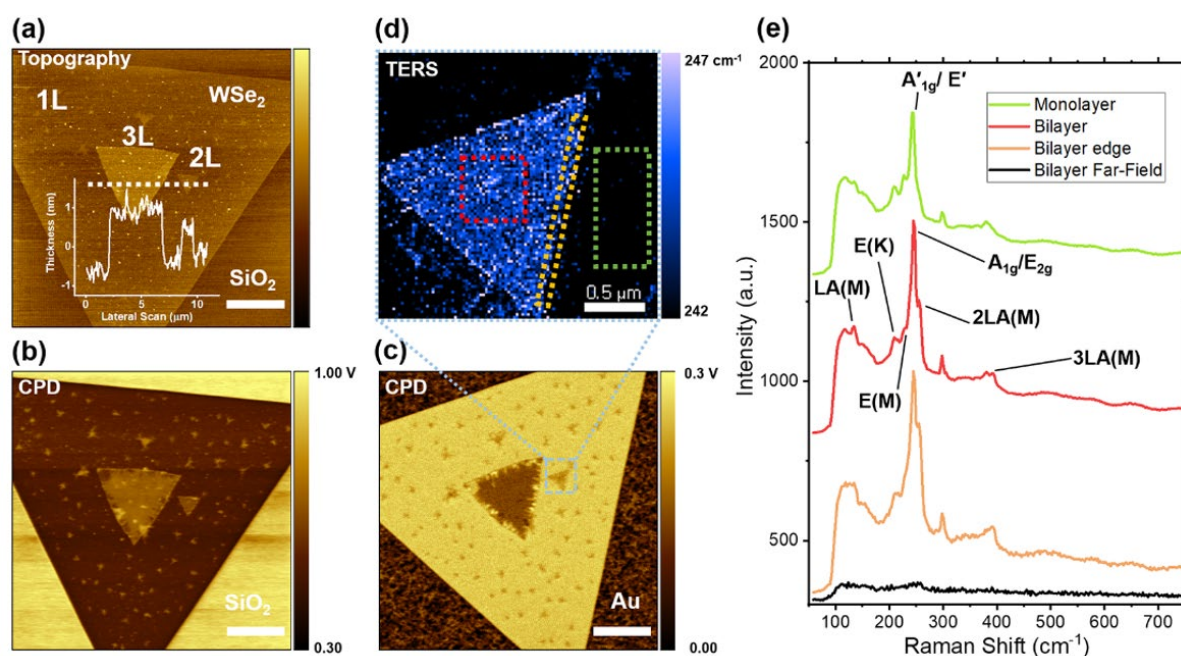

Figure 3-12: AFM, KPFM, and TERS imaging of WSe_2 CVD flakes embedded in Au by metal evaporation. (a) tapping AFM topography identifying the layer number of the WSe_2 edges prior to transfer to metal. Contact potential difference measured by KPFM before (a) after (b) transfer to the Au. Gap-mode TERS (d) peak wavenumber of the mode complex between 200-260 cm^{-1}. (e) spatially averaged spectra of the dashed colored boxes in (d).

Figure 3-12 and 3-13 shows an example of TERS and KPFM imaging for a WSe_2 on gold and silver substrates. The samples of the WSe_2 were first grown via chemical vapor deposition on the standard SiO_2/Si wafers. Instead of using dry-stamp methods to transfer the flakes potentially introducing contamination of the TMD-metal interface, ~50 nm of Au and Ag where directly deposited onto the WSe_2 flakes using metal evaporation. The TMD flakes are then directly embedded in the metal. Because the TMD adherence to Ag and Au is stronger than SiO_2, the whole metal film with the WSe_2 flakes can be stripped and inverted, allowing the bottom surface to be accessed. Full details of the sample preparation are discussed in A. Krayev et al.[54]

Figure 3-12a – 3-12b shows an example of WSe$_2$ before transfer. Figure 3-12a shows the atomic force micrograph for the as grown flake on the SiO$_2$/Si growth substrate. The flake hosts a large monolayer triangle with regions of multilayer growth scattering over the monolayer. The inset cure shows line scan indicated by the dashed black line shows step sizes consistent with tri-layer growth corresponding to center triangle, with the smaller triangles likely bi-layer. Figure 3-12b shows the CPD map taken concurrently with the topography of (a). The WSe$_2$ shows a much lower work function than underly SiO$_2$/Si substrate, not unexpected given the change from a wide bandgap insulator to semiconductor and within the flake the multilayer regions showed higher work function than the monolayer.

Because the flake is in intimate contract with the Au substrate, a-SNOM becomes a very powerful technique to characterize monolayer samples. This occurs because when both the probe and the substrate are metallic, the so-called gap-mode plasmon can be excited[34,44]. Compared to the normal plasmon of the tip alone the gap mode plasmon has much higher field intensity and tighter mode volume. The principals by which this works are similar to the Campanile probe, where the gap of the waveguide determines the resolution[41,42]. The gap here is instead formed between the tip and sample substrate. Unlike the Campanile probe where it is very challenging to fabricate gaps far below ~50 nm, the fine sample distance control that scanning probe methods such as AFM and STM, combined with the atomically thin character of TMDs, enable a truly nanoscale gaps to be formed. For atomically thin materials, this is easily below 1 nm.
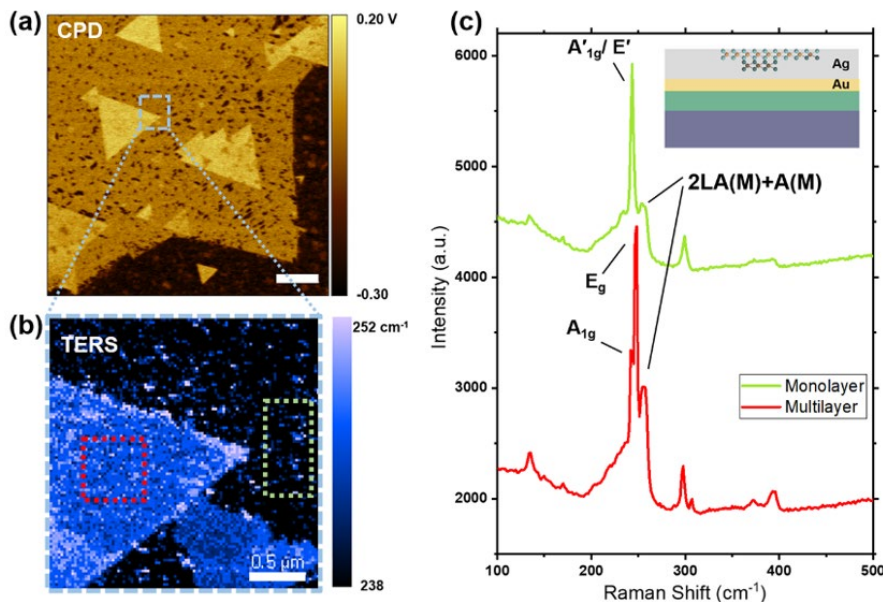
Figure 3-13: TERS and KPFM of WSe2 embedded in Ag. (a) CPD measured by frequency modulated KPRM. (b) TERS maximum wavenumber of Raman modes between 200-260 cm^{-1}. (c) averaged TERS spectra of the colored boxes shown in (b).

An example of gap mode TERS is shown in Figures 3-12d and 12e. Figure 3-12d shows the position of maximum TERS intensity between 200 and 260 cm^{-1}. Figure 3-12e shows the average Raman spectrum for the colored boxes. The position of the maximum Raman scattering is clearly different between the layers which is consist in the change from monolayer to multilayer[55]. Note however the sharpness of the bi-layer monolayer interface, which is limited to on order to 1 pixel (~ 20 nm).

We can gain insight into the KPFM and TERS maps shown Figure 3-12 by repeated the same experiment replacing the metal Ag. Compared with Au, silver has a much lower work function and Fermi level, ergo we may expect to see these differences reflected in the contact potential difference. Figure 3-13 shows that this is in fact the case. Figure 3-13a shows the CPD map for similar flake, monolayer containing islands of multilayers. Like Au the monolayer region has a higher CPD contrast compared to the metal, but the difference is notably lower. Unlike Au however, the multilayer areas increase in CPD as opposed to the decrease seen in Figure 3-12c. The peak position in the hyperspectral TERS data shows a similar character as that for gold, however some of the second order Raman peaks are shifted compared the spectra from Au. This may reflect differences in the screening between the Au and Ag which can affect shift in the resonance of the 632.7 nm laser.

Advantages and limitations of gap-mode

Figures 3-12 and 3-13 nicely demonstrate the high resolution that gap-mode TERS provides. We related this back to the very small gaps that can be formed and analogized the gap of the probe/sample to the gap of the Campanile structure. The comparison goes further, as the gap of the probe/sample has a similar broadband spectral response as the campanile tip, allowing for wide spectrum measurements[44]. Because metallization of standard AFM probes and sample substrates is straightforward and only requires common instrumentation such as metal evaporators, gap-mode are among the most widely used near-field optical measurement schemes, particularly for tip-enhanced Raman scattering.

The disadvantage gap-mode measurements is the smallness of the gap itself, i.e. the enhancement from gap is greatest when the gaps are very small. This has been shown theoretically in both finite element (FEM) and FTDT modeling[44,45]. We can also show this experimentally. Figure 3-14 shows an approach of simultaneous tip-enhanced Raman and tip-enhanced photoluminescence measurements on nanobubbles in a heterostructure of WSe_2 and $MoSe_2$ on the Au substrate. Nanobubbles in TMDs will form the primary subject of study of this dissertation and the details of nanobubbles, and this sample in particular will be described at length in the subsequent sections. For the moment it suffices to describe this nanobubble as small (~100 nm across) light emitters in the monolayer TMDs.
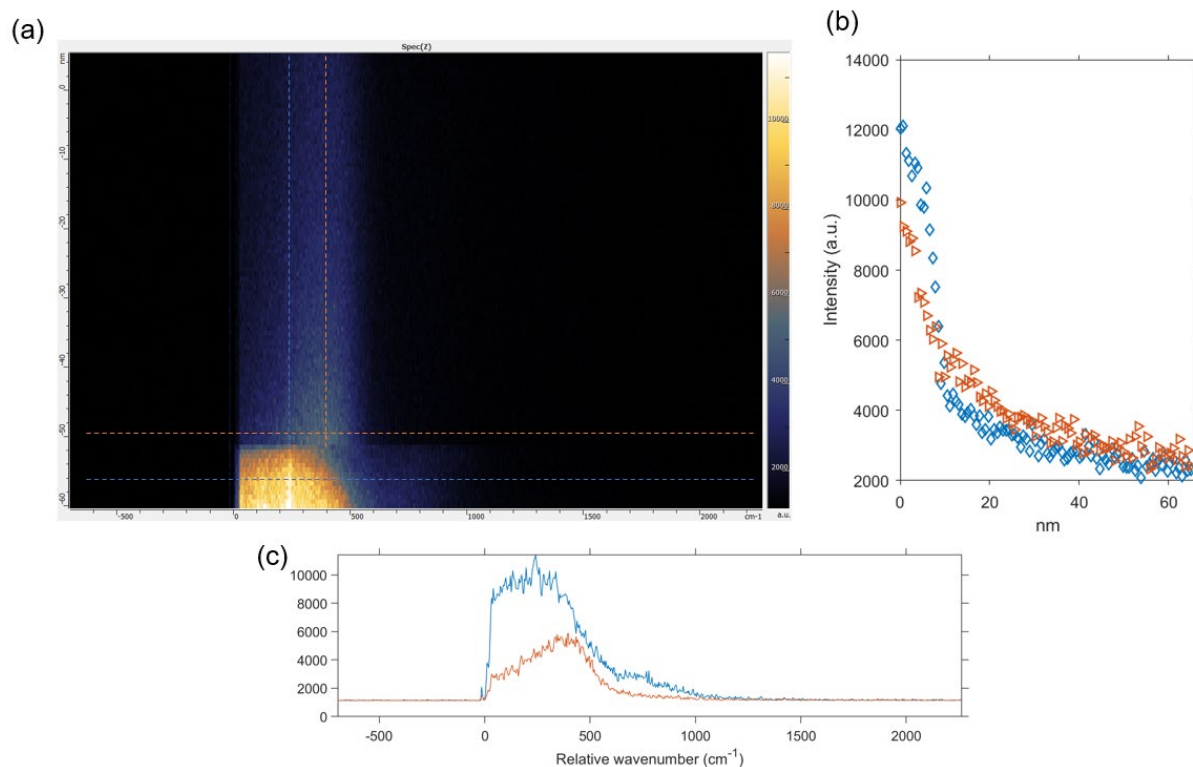
Figure 3-14: approach curve of TERS/TEPL (a), heat map showing spectral intensity as a function of sample probe separation, (b) and (c) vertical and horizontal linecuts for the dashed lines in (a).

Figure 3-14 shows a TERS/TEPL approach curve on such a nanobubble. Approach curves in the context of near-field optics are distance dependent optical measures, when a spectrum of the emitter (or Raman scatterer) is taken as a function of the separation of the emitter and the probe. Since the goal of SNOM and all near-field techniques is to record the typically lost evanescent fields, if the recorded spectrum is truly near-field an exponential dependence with distance should be observed. This is seen readily in Fig. 3-14b which shows vertical linecut the full spectrum color/heat map in Fig 3-14a where the y-axis is the distance from the sample in nanometers. On close inspection of the linecut one will notice that that are clearly two regimes, or a double exponential behavior with one decaying very quickly in about 2 – 10 nm while the other decaying much more slowly or 50 – 100 nm. This is perhaps more obvious in the full colormap of (a) with the sudden step-change that occurs around -52 nm. The fast dying regime is the gap-mode enhancement while the latter is the enhancement of the probe alone. Because of this fast die-off, gap mode is most effective for studies of very thin samples, like bubbles in 2D-semicondcutors and single molecules.

Another feature of gap-mode, which may be an asset or a complication, is that the extremely high fields of the small gap, in addition to other modification such as charge transfer from the tip to sample, can substantial modify the emission and the Raman scattering properties of the sample. Charge transfer in particular is understood to be both part of the answer behind the huge enhancement of the Raman, and the fine spatial resolution[45]. An extreme example of this was published recently in low temperature STM TERS on single molecules, where the combination of the ultra-small gap and the tunneling charge allowed contrast to be observed on scale < 0.5 nm[50]. Modification of the spectrum also appears to be the case of the photoluminescence and Raman in Figure 3-14. We can more easily see this by plotting example spectra given by the horizontal lines plotted in Figure 3 In the additional to the $WSe_2/MoSe_2$

36

Raman being visible at ~240-250 cm^{-1}, there are additional spectral lower and higher wavenumbers compared the non-gap-mode spectra where only a PL peak centered 400 cm^{-1} is seen.

The modification of the emission properties of the emitter is an example of a challenge for all scanning probe microscopies. Because is in intimate contact with the target in question, the recorded signal is always of probe sample system, and careful consideration is always warranted in the interpretation of scanning probe maps. For optical measurements we have already noted that charge transfer is known to enhance the Raman emission, and for excitons in TMDs we already seen that higher carrier density is associated with increased trion emission. We might speculate the apparent new PL at ~ 700 cm^{-1} in the blue spectrum is charged state of the enhance PL peak in the red-orange spectrum. However, caution is warranted before peaks from nano-PL recorded in gap-mode are assigned, as other phenomena are possible.

For instance, the interaction in the gap-mode geometry has been shown to be strong enough for strong coupling to occur. In the same manner that the surface plasmon at the metal/insulator interface forms a surface-plasmon-polariton in the Campanile probe, the gap-mode plasmon polariton can form hybrid state with the emitter. This has been observed in SNOM experiments measuring the excitons in quantum dots, forming what are called plexciton (or plasmon-exciton-polariton) states[56,57]. For the spectra in Fig 3-14c it is possible strong coupling occur with the huge plasmon enhanced fields of the gap-mode plasmon.

Another possibility for a-SNOM measurement performed in contact mode is modification of the emission due to strain. This possibility is especially important for stiffer cantilevers as the force even at small deflections can range into tens of nano-Newtons[58], which is significant for soft materials. To show that strain can have a significant effect on samples such as nanobubbles we can perform a similar measurement to that of the approach curve, where instead we start near in contact and progressively apply more force. Figure 3-15 shows the color/heat map of the result on a flat area of MoSe$_2$/WSe$_2$/Au (same sample as the approach shown in previous figure). As the tip first makes contact the strong gap-mode TERS signal from the Raman modes of WSe$_2$ and MoSe$_2$ are observed, along with significant PL, which perhaps indicates relatively poor contact of the Au and the semiconductor heterostructure. As more force is applied the PL is quenched, but later reemerges, which we may interpret as the heterostructure coming into good contact of the metal, followed some buckling as excess force is applied. We should stress that tip used in this experiment was attached to stiff non-contact cantilever (resonance frequency ~ 400 kHz) which is very high compared to the common near-field probes used in the AFM TERS/TEPL measurements, where resonant frequencies are usually between 50 – 150 kHz.
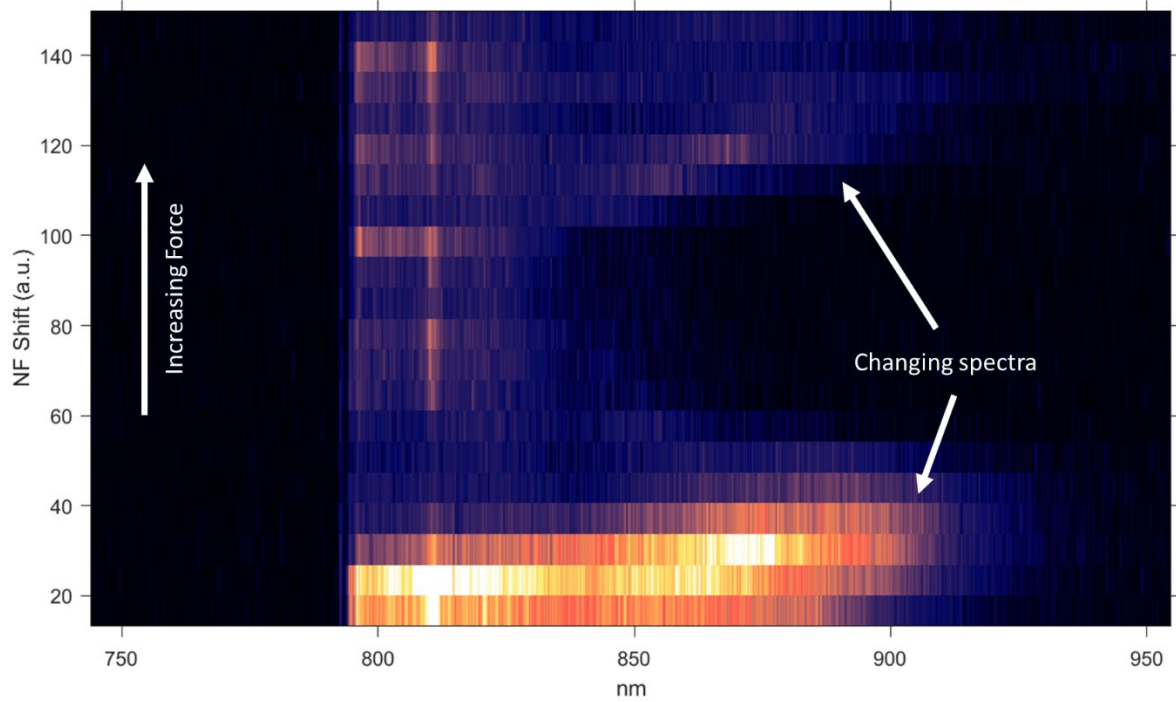
Figure 3-15: Gap-mode TEPL and TERS spectroscopy of MoSe$_2$/WSe$_2$, as a function of the normal force.

Section Conclusion

In this section we have overviewed aperture and apertureless SNOM measurements and their general working principles showing examples of hyperspectral nano-photoluminescence and nano-Raman characterization of monolayer WS$_2$ and WSe$_2$. In the following section we will focus on the application of these techniques to study interplay of nanoscale strain with the excitons in TMDs.

Section 4 – Nanoscale strain in TMDS: WS₂ nanobubbles

When atomically thin materials are deposited on a substrate by exfoliation, delamination of the material from the host substrate is a common occurrence, leading to formation of structural imperfections such as nanoscale wrinkles, folds, and bubbles. Nanobubbles, which form when trapped materials between the atomic layer and the substrate coalesce to form a local pocket of pressure, are among the most ubiquitous of these. Though often undesired for device applications, the ability of nanobubbles to modify the electronic structure of the capping atomic layer has spurred significant interest in the physics of nanobubbles themselves. In the case of the molybdenum and tungsten family of the transition metal dichalcogenides (TMDs), nanobubbles have been shown to significantly alter the optical bandgap and exciton photoluminescence up to several hundred meV[59,60]. The inhomogeneity of the strain has also been shown to induce complex phenomena such as exciton funneling[61], and single photon emission at cryogenic temperatures[2]. These exotic and technologically relevant properties have spurred interest in developing devices based on the 2D TMDs for solar light harvesting and optical quantum information processing respectively.

Critical to understanding the fundamental properties of nanobubbles is the determination of the magnitude and distribution of strain within them. This often done with a combination of atomic force microscopy (AFM), to measure topographical parameters such as the nanobubble maximum height and radius, and elasticity theory, where the monolayer TMD is treated as a thin plate subjected to a transverse load[60,62,63]. This approach, however, presents a challenge as the governing equations for a thin plate, the Föppl-von Kármán equations (FvK), are non-trivial to solve owing to their coupled non-linearity[63,64]. Many studies have made use of Hencky's solutions for thin circular plates with clamped edges under uniform transverse load[65]. This solution was used to estimate the Grüneisen parameters of graphene, which agreed well with values derived from ab initio density functional theory[66]. Hencky's solution has also been employed to estimate the van der Waals adhesion energy of both graphene[67] and MoS₂[68] using a mechanical blister test, as well the shift of the optical bandgap of MoS₂ under biaxial strain[60]. However, in nanobubbles formed by trapped material, the edges are not clamped but adapt according the balance of the attractive forces of the monolayer to the substrate, and the pressure of the trapped material. Nanobubbles, for instance, can even be moved relative to the top layer and substrate by means of contact AFM[58]. For nanobubbles of this type, Khestanova *et al*. solved the FvK for circular symmetric nanobubbles in MoS₂, graphene, and hBN. Remarkably, the authors found a universal bubble strain distribution and self-consistent scaling of nanobubbles[62,63], where an ideal nanobubble of a van der Waals material on a given substrate will have a constant shape and strain distribution for nanobubbles of different sizes. The magnitude of the strain depended only the square of the nanobubble aspect ratio, $\frac{h_{\max}}{r}$, which is determined by the attractive force of the monolayer to the substrate and material stiffness. This allows the nanobubble strain to be tuned by changing the substrate, which was utilized to control the approximate strain in circular MoS₂ nanobubbles[61].

While the application of elasticity theory to circular nanobubbles has yielded great insight into adhesion energies and strain dependence of free exciton emission of TMDs, the restriction to nanobubbles of circular symmetry has so far limited the use of these models in the quantitative study of localized exciton (LX) states that are also known to inhabit the bubbles, particularly LXs that become single-photon emitting at low temperatures. It has been observed that typical nanobubbles exhibited three to four energetically distinct single-photon-emitting LX states per bubble[2], suggesting strong asymmetry in the strain distribution. Moreover, nanobubbles in TMDs are not generally circularly symmetric, and can take

on arbitrary and amorphous shapes[58], which will have significantly different strain distributions compared to the relatively simple case of circular nanobubbles.

Recently, advanced atomistic theory has shown that the effect of strain from nanobubbles in monolayer TMDs can be dramatically different than the results of classical elasticity theory[69]. According to classical plate theory, the maximum strain of a circularly symmetric nanobubble is at its center[70,71], which would then correspond to the maximum detuning of band-edge electronic states[60]. By contrast, calculations of MoS_2 nanobubbles using valence force field theory combined with tight-binding models, including effects both of strain and changes in dielectric screening when the material separates from the substrate, showed that the greatest strains and energy level shifts are concentrated at the nanobubble edges[69]. Further, the atomistic theory predicted that for nanobubbles of sufficiently high aspect ratio, the atomic lattice will form wrinkles around the nanobubble boundary, a phenomenon not captured in the FvK equations. A key advantage of this approach is that assumptions of circular symmetry are unnecessary in the valence force field calculations, and arbitrary nanobubble morphologies can be considered. We recently showed using high-resolution scanning near-field optical microscopy studies (SNOM) that the energies and spatial distributions of photoluminescence in WSe_2 nanobubbles indeed correspond to the nanoscale signatures predicted by the atomistic theory, where experimentally measured AFM topographical profiles were explicitly used as the input for the valence force field calculations[72].

While the atomistic approach offers an accurate description of the nanoscopic strain and its effect on the electronic band structure, such calculations are computationally expensive, limiting its use for evaluating complex strain profiles in TMDs and other atomically thin van der Waals materials. Further, the universal scaling observed in ref[62] shows that the FvK equations do describe the average mechanical properties of nanobubbles to a reasonable degree[62]. And estimations of the overall energy shift of exciton emissions under biaxial strain by the mechanical blister tests in refs[60][60] agree at least qualitatively with predictions from ab initio density functional theory[73]. This leads us to ask: how reasonable are the FvK equations in the description of excitons TMD nanobubbles, and can they be utilized in a manner that enables the analysis of more complex geometries?

Here, to better understand the applicability of the FvK equations to TMD nanobubbles, we develop a numerical method to estimate the strain in nanobubbles of arbitrary symmetry. This method relies only on the experimentally measured AFM topography and the material's Poisson ratio to calculate the full strain tensor for nanobubbles of arbitrary shape. We verify that this method reproduces the strain for circularly symmetric bubbles different height profiles where the analytical solutions are available. We then apply this method to nanobubbles formed in monolayer WS_2 deposited on template stripped Au. We use nanoscale photoluminescence measurements to show that the predicated strain tracks with the PL intensity and energy. Lastly, we show that our method can be applied to large fields of interacting nanobubbles imaged with AFM, simultaneously providing nanoscopic strain distributions of all imaged nanobubbles. Using this approach, we rederive the Grüneisen parameter for the E' mode of 1L-WS_2, which agrees well with previous experiments[74,75]. Our method is thus fully consistent with established plate theory models, and by back solving the measured topography is able to capture structural details, such as kinks and wrinkles provided the measured topography is good estimate the true surface, are not captured in the known analytic solutions. This procedure thus provides a facial way to map strain distributions in heterostructures of TMDs and van der Waals materials generally.

Determination of the strain profile of a nanobubble from its AFM topography

We outline here a method to calculate the strain of a nanobubble in monolayer WSe_2 or other van der Waals materials based on its measured height profile. The general theory is described in detail in many

classic references[64,76]. Here we briefly summarize the important aspects that are relevant for the calculation of the strain profile of a bubble with an arbitrary height profile, h. For an elastic deformation of a continuous body the internal stress must balance with applied forces on the boundary. In terms of the stress tensor $\mathbf{\Sigma}$ this condition is written as:

$$\sigma_{ik}n_k = P_i \qquad (1)$$

Where P_i is applied traction (pressure in this case), σ_{ik} is the stress tensor, and n_k is the unit normal vector. Repeated indices imply summation. For a thin plate, only small forces are necessary to bend it relative to the magnitude of the internal stresses caused by the bending. Thus, the pressure term may be neglected in the normal direction to the surface of the plate. Assuming the unit-normal lies in the z-direction we have the condition of plane stress with $\sigma_{zx} = \sigma_{zy} = \sigma_{zz} = 0$. The generalized Hooke's law relations for a plate then take the form[64]:

$$\varepsilon_{xx} = \frac{(\sigma_{xx} - v\sigma_{yy})}{E} \qquad\qquad \varepsilon_{xy} = \frac{(1+v)\sigma_{xy}}{E} \qquad (2)$$

$$\varepsilon_{yy} = \frac{(\sigma_{yy} - v\sigma_{xx})}{E}$$

where ε_{ij} is the strain tensor, v is the Poisson ratio, and E is the Young's modulus. The components of the strain tensor are defined in terms of the displacement of the system from the reference configuration. Following the standard formulation, we define the displacement in terms of an in-plane displacement vector $\mathbf{u}$ and the out-of-plane displacement, h, which we later identify as the nanobubble height. Considering an infinitesimal element of length of the deformed body, we adopt the Green-Lagrange strain measure that has non-linear strain-displacement relationships given by:

$$\varepsilon_{xx} = \frac{\partial u_x}{\partial x} + \frac{1}{2}\left(\frac{\partial h}{\partial x}\right)^2 \qquad \varepsilon_{xy} = \frac{1}{2}\left(\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x} + \frac{\partial h}{\partial x}\frac{\partial h}{\partial y}\right) \qquad (3)$$

$$\varepsilon_{yy} = \frac{\partial u_y}{\partial y} + \frac{1}{2}\left(\frac{\partial h}{\partial y}\right)^2.$$

The equations of equilibrium of a plate are obtained by minimizing the total energy, defined as the sum of the mechanical free energy, F, and the potential due to external forces, U. The mechanical free energy can be divided into the contributions due to the out of plane (bending energy) and the in-plane (stretching energy) displacements. Per unit volume, the bending energy is then given explicitly by:

$$f_b = \frac{Ed^3}{24(1-v^2)}\left[\left(\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2}\right)^2 + 2(1-v)\left\{\left(\frac{\partial^2 h}{\partial y\partial x}\right)^2 - \frac{\partial^2 h}{\partial x^2}\frac{\partial^2 h}{\partial y^2}\right\}\right]$$

The stretching energy is given by the formula where repeated indexes are summed:

$$f_s = \frac{1}{2}\varepsilon_{ij}\sigma_{ij}$$

The minimization of these quantities is described in reference[64]. The resulting equations are:

$$\frac{Ed^3}{12(1-v^2)}\nabla^2\nabla^2 h - d\frac{\partial}{\partial x_j}\left(\sigma_{ij}\frac{\partial h}{\partial x_i}\right) = P \qquad (4a)$$

$$\partial\sigma_{ij}/\partial x_j = 0 \qquad (4b)$$

where the stress tensor can be written in terms of the unknown displacements by (2) and (3). Equation (4b) derives from the condition that the stress tensor has zero divergence $\nabla \cdot \mathbf{\Sigma} = 0$, which is satisfied when the

body forces can be neglected. Under these conditions and because the stress tensor is symmetric, the stress tensor can be expressed as derivatives of a scalar potential called the Airy stress function χ, defined by[64,77]:

$$\sigma_{xx} = \frac{\partial^2 \chi}{\partial y^2} \qquad \sigma_{xy} = -\frac{\partial^2 \chi}{\partial y \partial x} \qquad (5)$$

$$\sigma_{yy} = \frac{\partial^2 \chi}{\partial x^2}$$

Equation (4b) is then satisfied automatically. A second equation is obtained by combining equations (2), (3), and (5) for the stress function in terms of second order partials derivatives of the height, h. One then obtains a complete system of equations, referred to as the von Karman equations for a thin plate[64,76,77]:

$$\frac{K_b}{d}\nabla^2\nabla^2 h - \frac{\partial^2 \chi}{\partial y^2}\frac{\partial^2 h}{\partial x^2} - \frac{\partial^2 \chi}{\partial x^2}\frac{\partial^2 h}{\partial y^2} + 2\frac{\partial^2 \chi}{\partial y \partial x}\frac{\partial^2 h}{\partial y \partial x} = P/d \qquad (6)$$

$$\nabla^2\nabla^2 \chi + E\left\{\frac{\partial^2 h}{\partial x^2}\frac{\partial^2 h}{\partial y^2} - \left(\frac{\partial^2 h}{\partial y \partial x}\right)^2\right\} = 0 \qquad (7)$$

where we have introduced the bending rigidity $K_b = \frac{Ed^3}{12(1-\nu^2)}$, and d is the thickness of the plate. Note that the term in the curly brackets of equation (7) is the Gaussian curvature. The von Karman equations are valid for small deformations and moderate rotations.

These equations form a coupled non-linear system for the Airy stress function and out-of-plane displacement height, h, and are difficult to solve in general. Many past studies have applied these equations in the membrane limit[59,60]. In this limit, the first term in equation (6) is neglected, yielding a linear equation for the height profile in terms of the applied stress at the nanobubble boundaries and transverse pressure, P. In this regime, Khestanova et al.[62] showed that circularly symmetric nanobubbles in van der Waals solids have a characteristic universal height profile. In the opposite limit where bending energies are significantly larger than the stretching energy, the second and later terms on the left side of equation (6) are neglected, yielding a linear equation for the height profile. This regime is sometimes called the "pure-bending" limit[71].

For the case of nanobubbles in this work, the out-of-plane displacement is known. Then from (3), to determine the strains, we need only solve for the in-plane displacement vector, $\mathbf{u}$, making the assumptions that Hooke's law is valid and the system is in equilibrium, i.e. (4b) is satisfied. There are several approaches to solve this system. For instance, one can use equation (4b) with the stress and strain given by (2) and (3) to solve two second order linear equations for the in-plane displacements, or equivalently solve the fourth order equation (7) for the Airy stress function. In the former case, Rostami et al. used a Green's function approach to solve for the in-plane displacement given an arbitrary height profile and assuming circular symmetry[71]. Since the nanobubbles in this work are not universally circularly symmetric, we chose the latter approach, numerically solving equation (7) for the Airy stress function normalized to the Young's modulus, E.

Many efficient numerical techniques exist for solving such linear partial differential equations. In this manuscript we used a spectral interpolation method[78]. Spectral methods solve for the target function as linear sum of orthogonal basis functions, i.e., an interpolation. This approach allows the solution to be approximated from a relatively small number of points (termed collocation points) that have been solved in the basis function space. The choice of the collocation points depends strongly on the chosen basis functions. For example, in periodic problems a Fourier basis is commonly used, and the collocation points are an equally spaced grid defined by the frequencies of the discrete Fourier transform. For the nanobubble

in Fig. 3 of the main text we choose a basis of Chebyshev polynomials of the first kind for the x and y direction. Our collocation points are the Chebyshev points defined by:

$$x_i = \cos(i\pi/N) \qquad i = 0,1,\dots,N$$

$$y_j = \cos(j\pi/N) \qquad j = 0,1,\dots,N.$$

With a given choice of basis functions, the differential operator in (7) can be replaced by a matrix, in this case an $(N+1)^2 \times (N+1)^2$ square matrix. For the Chebyshev polynomials, the derivative matrix in each dimension is given by the formula:

$$D_{ij} = \frac{c_i}{c_j}\frac{(-1)^{i+j}}{x_i-x_j}, \; 0 \le i,j \le N, i \ne j$$

$$D_{ii} = -\frac{x_i}{2(1-x_i^2)}, \qquad 1 \le i \le N-1$$

$$D_{00} = -D_{NN} = \frac{2N^2+1}{6}$$

where the constants c are $c_0 = c_N = 2$, $c_i = 1$, for $i = 1,2,\dots,N-1$. Higher order derivatives are found by computing appropriate powers of the matrix[78,79].

For boundary conditions on the stress function, we assume that the ends of the plate are traction free and not pinned to the substrate. Physically this implies that the component of the stress tensor normal to the boundary should be zero from equation (1)[77]. Explicitly in polar coordinates:

$$\sigma_{rr} = \sigma_{r\theta} = 0 \qquad\qquad\qquad (8)$$

The boundary here is the actual boundary of the plate and not the boundary of the nanobubble, where the radial stress is not in general zero. However, for nanobubbles with circular symmetry it can be shown that sum of the diagonal components of the stress should vanish. This condition is equivalent to the Laplacian of the Airy stress function vanishing at the edge of the nanobubble:

$$\nabla^2\chi = 0 \quad \text{when} \quad |\nabla h| = h = 0. \qquad\qquad (9)$$

In the section below we derive these conditions for circularly symmetric bubbles and then generalize them for nanobubbles of arbitrary geometry.

With these boundary conditions there are still many functions that satisfy equation (7) and are thus compatible stress functions. For the additional degrees of freedom, we set $\chi = 0$ at the boundaries. The boundary conditions are imposed on the stress function itself and its 2nd order derivative which allows one to separate equation (7) into two weakly coupled Poisson equations that can be solved sequentially[79]:

$$\nabla^2\phi = -E\left\{\frac{\partial^2 h}{\partial x^2}\frac{\partial^2 h}{\partial y^2} - \left(\frac{\partial^2 h}{\partial y\partial x}\right)^2\right\} \qquad (10a)$$

$$\nabla^2\chi = \phi. \qquad\qquad\qquad (10b)$$

Practically solving equation (7) numerically by spectral interpolation using a Chebyshev basis amounts to the following steps:

1. Fit a smooth function to the AFM topography of the nanobubble.
2. Differentiate the smooth surface in x, y to determine the Gaussian curvature at each of the given collocation points.

3. Define the derivative operators in terms of the Chebyshev polynomials at each of the collocation points. The Laplacian operator can then be constructed using tensor products of the derivative operator and the identity matrix[78].
4. Apply the boundary conditions by replacing the corresponding columns and rows of the Laplacian with the boundary conditions defined by equation (8) or equation (9).
5. Invert equation (7) or equations (10a) and (10b) to find the Airy stress function.
6. Apply interpolation to find the function throughout the real space.

For an excellent reference on Chebyshev based techniques for solving boundary value problems see reference[78].

For verification that the outlined numerical recipe correctly solves for the strain, we solved for the radial strain of the three circularly symmetric models that were solved by Rostami *et al.*[71] In SI Fig. 1a, we show the three profiles normalized such that the radius and maximum height are unity. The corresponding radial components of the strain are graphed in SI Fig. 1b for a Poisson ratio of 0.19 . The solid lines denote the exact solution taken from appendix D of Rostami *et al.*[71], and the symbols are the numerical solution obtained by solving equations (9a) and (9b). As can be seen there is excellent agreement between the exact and numerical solutions.

Derivation of the boundary condition of equation (9)

Here we derive the boundary conditions for the Laplacian of the Airy stress function. To begin we apply the Laplacian to the Airy stress function. Invoking the definition of the function and Hooke's Law we find that the Laplacian is proportional to the trace of the strain tensor:

$$\nabla^2 \chi = \sigma_{xx} + \sigma_{yy} = \frac{E}{1-\nu}(\varepsilon_{xx} + \varepsilon_{yy}). \qquad (11)$$

Recalling that the strain gives the relative extension of the plate on each axis, since strain is typically small (~1 %), the trace of the tensor is then equal to the relative increase in area of the plate to first-order (because strain values are typically small, the second-order terms become negligible). In our case the plate is deformed by an outward pressure (i.e., the load applied to the plate) in the normal direction of the surface to give a deflection in z (given by the height profile, h), and we might expect the relative increase of area to go to zero at the boundary where both the deflection and load go to zero. Thus, under these conditions at the boundary, equation (11) implies that the Laplacian of the Airy stress function also goes to zero. Below we show that this is exactly the case for circularly symmetric nanobubbles. We then discuss the case for bubbles of arbitrary shape.

For circularly symmetric nanobubbles we assume the in-plane displacement has only a radial component

$$\mathbf{u} = u(r)\hat{\mathbf{r}}.$$

It can be shown then the shear components of the strain tensor vanish, $\varepsilon_{r\theta} = 0$. The radial and azimuthal strain are then given by[71],

$$\varepsilon_{rr} = \frac{\partial u}{\partial r} + \frac{1}{2}\left(\frac{\partial h}{\partial r}\right)^2 \qquad\qquad \varepsilon_{\theta\theta} = \frac{u}{r}. \qquad (12)$$

The stresses associated with these strains are then given by Hooke's law:

$$\sigma_{rr} = \frac{E}{1-\nu^2}(\varepsilon_{rr} + \nu\varepsilon_{\theta\theta}) \qquad \sigma_{\theta\theta} = \frac{E}{1-\nu^2}(\varepsilon_{\theta\theta} + \nu\varepsilon_{rr}). \qquad (13)$$

The stress tensor can be written as $\boldsymbol{\Sigma} = \hat{\mathbf{r}}\sigma_{rr}\hat{\mathbf{r}} + \hat{\boldsymbol{\theta}}\sigma_{\theta\theta}\hat{\boldsymbol{\theta}}$.

To find boundary condition of the Laplacian of the Airy stress function we need to solve for the radial and azimuthal strains given by equation (12). Since the height is known, we thus only need to solve for the value of the in-plane radial displacement. To solve this system, we assume again that nanobubble is in equilibrium, thereby the divergence of the strain tensor must be zero. Inserting (12) into (13) and taking the divergence we arrive then at the equation for the in-plane displacement,

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} - \frac{u}{r^2} + \frac{\partial h}{\partial r}\frac{\partial^2 h}{\partial r^2} + \frac{1-\nu}{2r}\left(\frac{\partial h}{\partial r}\right)^2 = 0. \qquad (14)$$

At the nanobubble edge we expect the height of the nanobubble and the normal of the gradient of the height to vanish. Because h is only a function of r, this condition entails that $\partial h/\partial r|_{r\to R} = 0$. Therefore at the nanobubble edge the displacement must satisfy

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} - \frac{u}{r^2} = 0. \qquad (15)$$

One will recognize this as equivalent to the vector Laplacian applied to the displacement vector, $\nabla^2\mathbf{u} = \mathbf{0}$. The solution to equation (15) can be shown to be

$$u(r) = \begin{cases} Ar & \text{for } r \leq R \\ Br^{-1} & \text{for } r > R \end{cases}, \qquad (16)$$

where the constants A and B are set by requiring the strain to be continuous at $r = R$.

With equation (16) we can calculate the strains for the nanobubble boundary by equation (12). We find then for $r > R$ and $\partial h/\partial r = 0$ that

$$\varepsilon_{rr} = -\varepsilon_{\theta\theta} = -\frac{B}{r^2}. \qquad (17)$$

Since the trace is invariant of coordinates, we conclude that

$$\nabla^2\chi = \frac{E}{1-\nu}(\varepsilon_{rr} + \varepsilon_{\theta\theta}) = 0 \text{ for } r > R. \qquad (18)$$

Thus for nanobubbles of circular symmetry the Laplacian is zero when $r = R$, where R is the radius of the nanobubble. This condition was applied for the numerical solutions of the parabolic-like and pure-bending height profiles in SI Fig. 1, which show excellent agreement with the exact solutions that were derived by Rostami et al.[71] For the Gaussian nanobubble, R is not the boundary of the bubble but the standard deviation of the Gaussian profile. Thus for equation (15) to be valid, it suffices to set the boundary condition at $r \gg R$, such that $\partial h/\partial r \approx 0$.

For a nanobubble of arbitrary shape we expect that far away from the nanobubble, the strain should be approximately circularly symmetric. We can thus invoke (18) for the boundary condition on the Laplacian. To verify this assertion, we first rewrite the Laplacian of the Airy stress function with the explicit expressions for the x and y components of the strain given by the set of equations in (3). One finds that:

$$\nabla^2\chi = \frac{E}{1-\nu}(\varepsilon_{xx} + \varepsilon_{yy}) = \frac{E}{1-\nu}\left(\nabla\cdot\mathbf{u} + \frac{1}{2}|\nabla h|^2\right). \qquad (19)$$

At the boundary of a nanobubble, the component of the gradient of the height that is normal to the boundary vanishes:

$$\nabla h \cdot \mathbf{n} = 0,$$

where $\mathbf{n}$ is the unit normal to the nanobubble boundary. For a boundary defined by $h = \text{constant} = 0$, this is equivalent to the gradient vanishing $|\nabla h| = \partial h / \partial r = 0$. If we choose the nanobubble boundary as a contour defined by the gradient of the height being zero, equation (19) becomes

$$\nabla^2 \chi = \frac{E}{1-\nu}(\nabla \cdot \mathbf{u}). \qquad (20)$$

We see that assuming the Laplacian of the Airy stress function is zero is equivalent to assuming that the divergence of the in-plane displacement is zero. Further, along this contour where the gradient of the height is zero, equation (20) also indicates that the in-plane displacement is given by gradient vector field.

$$\mathbf{u} = \nabla \frac{1-\nu}{E} \chi. \qquad (21)$$

Since $h = |\nabla h| = 0$, the equations of equilibrium for the displacement are the same as for a general isotropic medium with zero body forces[64]:

$$2(1-\nu)\nabla(\nabla \cdot \mathbf{u}) - (1-2\nu)\nabla \times \nabla \times \mathbf{u} = 0. \qquad (22)$$

Because $\mathbf{u}$ is a gradient vector field the second term vanishes. We find that the divergence of $\mathbf{u}$ must be a constant:

$$\nabla(\nabla \cdot \mathbf{u}) = 0 \Rightarrow \nabla \cdot \mathbf{u} = \text{const.}$$

Since as r tends to infinity we require the strain, and thus the trace of strain tensor to vanish, this constant must be zero.

From this argument we see that if we define the solution boundary such that $h = |\nabla h| = 0$, the Laplacian of the Airy stress function should vanish in absence any extrinsic in-plane stress applied at the boundaries.

Verifying the method

Before applying the technique to the topography of real nanobubbles, we verified that the method correctly reproduces the strain for nanobubbles with circular symmetry, where analytic solutions have been derived[71]. Figure 1b shows three radial cross sections of the symmetric nanobubbles, solved originally in ref [71]. Figure 1c shows the corresponding radial strain for the exact analytic (solid lines) and numerical solutions to equation (5). As can be seen, the numerical solutions reproduce the analytic solutions for each nanobubble case.
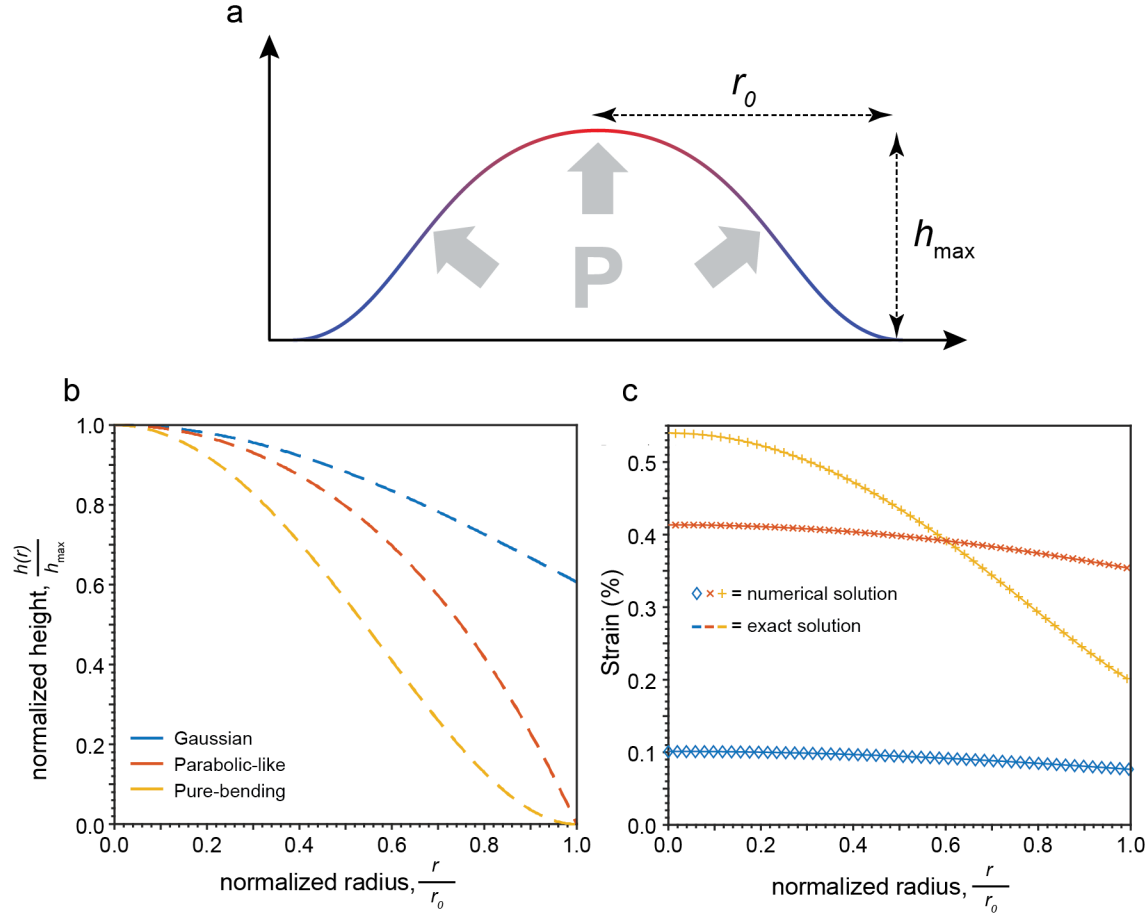
Figure 4-1: Comparison of the numerically calculated strain of circularly symmetric bubbles of different profiles to the exact strain as derived in Rostami *et al.*[8] **(a)** Diagram of a nanobubble in 1L-WS$_2$; **(b)** Normalized height profiles of the three nanobubble cases: Gaussian (blue), parabolic-like (red), pure bending (yellow-orange). **(c)** The corresponding strain values from each case. The solid is the exact radial strain as derived in Rostami *et al.*[8], and the symbols are the numerical values calculated by solving for the Airy stress function in equations (7) and (9). For the complete formula of the height profile and the strain, see Appendix D of Rostami *et al.*[8].

We next aim to verify that our computational method correctly solves the FvK. To do this, it is necessary to have a measure of strain that is independent of the estimate obtained from the AFM topography. For the atomically thin TMDs, both the photoluminescence and Raman scattering signals are well known to be redshifted by tensile strain[60,61,69,75,80,81]. This allows us to use hyperspectral mapping to resolve the energetic shifts of the PL and Raman signals, which will correlate with the tensile strain. However, because of the nanometric size of many nanobubbles, with radii of order ~100 nm or less, this approach is confounded by the resolution limit of conventional optical microscopes (~250 nm for visible light). As detailed in the previous section, a-SNOM is technique well suited to studying spatial heterogeneities and serves to bridge the length scale between optical wavelength and the native length scale of excitons in these materials. In the following we can use the hyperspectral maps in both Raman and PL to verifying the strain values calculated from the AFM topography.

We have chosen to study nanobubbles in monolayer WS_2 as the benchmark for our approach. WS_2 possesses excellent optical properties, with its PL and Raman dependencies on strain well-characterized both experimentally[75] and theoretically[73], making it an ideal material for experimental verification of the strain estimation presented here. Our sample was prepared by exfoliation of a bulk WS_2 crystal onto gold. A template-stripped gold substrate is used, which produces metallic films with very low surface roughness rms values < 1 nm[82], allowing for excellent adhesion of the exfoliated WS_2 onto the substrate. More details of the sample preparation are provided in ref[82][82]. The use of the Au substrate also allows both TERS and TEPL to operate in "gap-mode", producing significantly higher field enhancements compared to dielectric or insulating substrates[34].

An excitation wavelength of 637.25 nm was used for TERS and TEPL at incident excitation power of ~70 uW. This wavelength is near resonant with the A exciton of 1L-WS_2 , emitting near ~ 620 nm[49]. While this choice of the excitation does limit the observable photoluminescence, the resonance condition greatly enhances the Raman scattering response [83], and as will be shown the effect of strain is substantial enough that over the nanobubble significant TEPL is observed. Additionally, this choice of excitation allows simultaneous measurement of the TEPL and TERS signals, eliminating the risk of sample modification between measurements. All AFM and TERS/TEPL measurements where conducted with an Horiba OmegaScope-R coupled to an Horiba Xplora micro-Raman spectrometer.
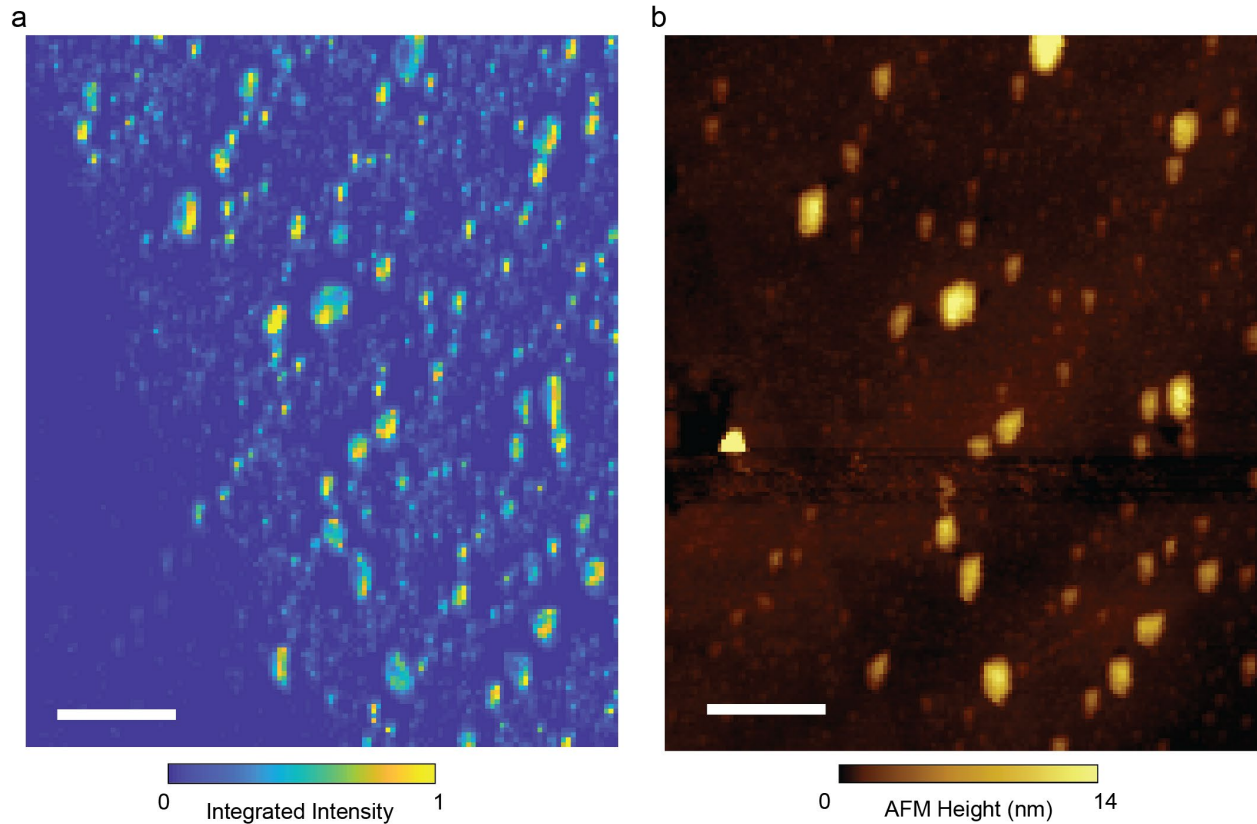

Figure 4-2: (a) atomic force micrograph and (b) nano-PL and nano-Raman map of nanobubbles in WS_2 on Au substrate.

Figure 4-2a shows a semi-contact AFM micrograph of a field of nanobubbles in the sample. The field contains dozens of the nanobubbles, the majority of which are irregular in shape. Additionally, several nanobubbles are separated by distances where the strain fields could potentially overlap, further complicating study by solutions to the FvK equations solved for a single nanobubble. Figure 4-2a shows the integrated TEPL and TERS intensity for the same field of nanobubbles in Figure 4-2b. As can clearly be seen, each resolvable nanobubble in the AFM micrograph has a substantial optical response. Notably, the emission intensity is non-uniform over the nanobubble surface with certain nanobubble points exhibiting greater intensity than neighboring points. A possible explanation for this is funneling of excitons to points of greatest strain, which was theoretic predicted[59] and recently observed[61] in monolayer MoS$_2$. We comment more on this possibility below.

On inspecting the distribution of nanobubbles in Fig 4-2a, one observes that there are many small nanobubbles dispersed throughout the map. A closer look reveals that the larger nanobubbles have relatively few of these smaller nanobubbles immediately around them. We hypothesize that this is result of the attractive force between neighboring nanobubbles which was predicted theoretically by Khestonva et al.[62]. The corresponding larger strain field of a larger nanobubble may then pull in and merge with surrounding smaller bubbles, effectively clearing the area around it. While the data in Figure 4-2 is not sufficient to prove this, the consistency of this feature does imply an underlying physical mechanism rather than vagaries of material deposition.

We next solve equation (7) using the AFM topography of a single large nanobubble in Figure 4-2b and compare this to the corresponding TEPL/TERS data from the Fig. 4-2a. For this test, we choose the largest nanobubble located in the top center of the field, which shows a relative lack of close neighbors. In Figure 3a we show the resulting map of the strain tensor trace. This value is proportion to the total relative change in volume. We see that maximum strain of ~0.9% is located at the center of the nanobubble or at the point of maximum deflection when compared to the AFM topography (inset of Figure 4-3a), as expected from circularly symmetric plates in Figure 4-1. However, there is also a secondary maximum located near the edge of the nanobubble.

The corresponding nano-optical data for this nanobubble are shown in Figures 4-3b and 4-3c. The integrated TEPL/TERS intensity is shown in Figure 4-3b and select point spectra are shown in Figure 3c, corresponding to the colored crosses in the integrated intensity map in Fig. 4-3b. We see that the reddest emission correlates to the area of greatest strain, as expected.

However, when considering the intensity, both the max intensity of the point spectra as well as the integrated intensity show the greatest signal near the secondary maximum. Exciton funneling would suggest that maximum strain would exhibit the greatest PL emission intensity since it is energetically favorable for exciton diffusion[59,61]. This discrepancy can be fully accounted for when the excitation geometry of the gap-mode TEPL/TERS measurement is considered. It is well known that the enhancement of the electric field due to the plasmonic probe has a strong dependence on the probe-sample distance. Thus, for points of a nanobubble farther away from the substrate, the increased emission intensity due to exciton funneling is offset and eventually overcome by the large fall-off of excitation intensity and reduction in emission enhancement. This effect is seen by diminished TERS intensity relative the TEPL intensity at points towards the nanobubble peak in Fig. 3. The dramatic reduction in Raman to PL ratio of the blue spectra implies a corresponding reduction in the excitation intensity.
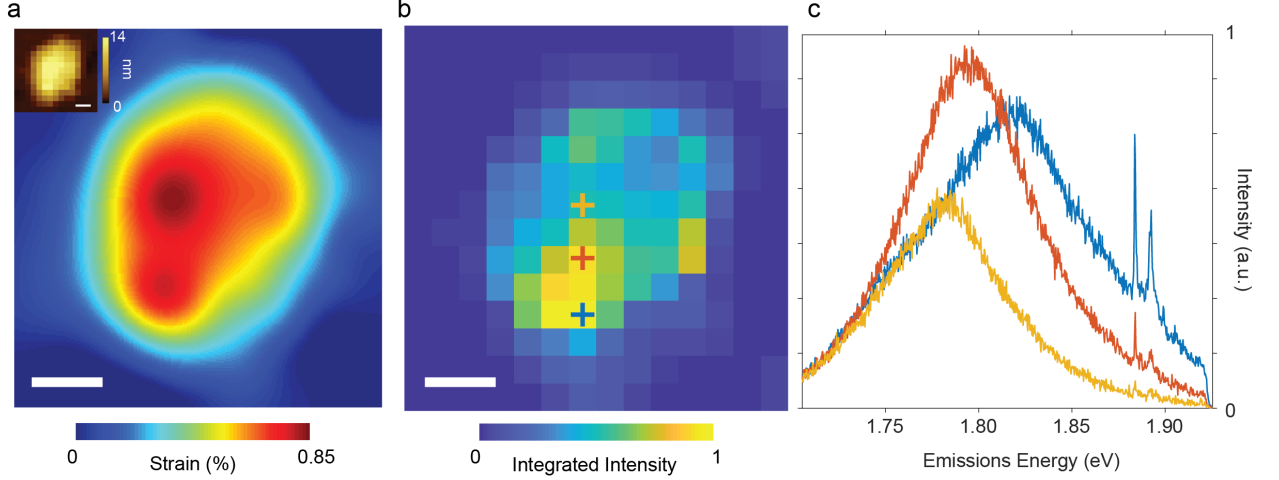
Figure 4-3: (a) trace of the strain tensor calculated from the second Föppl-von Kármán equation for the nanobubbles given in Fig 1c, inset the corresponding AFM micrograph; (b) Nano-PL map of the same nanobubble showing the strain shifted exciton luminescence as well as localized exciton (LX) states.

Finally, we show that the strain-analysis method can be applied over the large field of nanobubbles shown in Fig. 4-2. Figures 4-4a – 4-4c show three square fields of nanobubbles with the corresponding TEPL/TERS intensity of Figure 4-2 inset for ease of reference. Figure 4-4a includes the nanobubble analyzed in detail in Figure 4-3a. As can be seen by inspection of Figures 4-4a and 4-3a, the two solutions have nearly the same strain morphology, the discrepancies between which we attribute to the solution of Fig. 4-3a implicitly assuming the nanobubble to be completely isolated (the larger field solution of Fig. 4-4a relaxes this assumption).

To show the quantitative accuracy of the strain solution, we leverage the high number of nanobubbles to calculate the Grüneisen parameter for E' mode. The Grüneisen parameter relates the change in frequency of a crystal vibration to a hydrostatic stretching (without any shear) for a single vibrational mode, *i,* and is defined by[74]:

$$\gamma_i = \frac{V}{\omega_i}\frac{\partial \omega_i}{\partial V} \tag{23}$$

Where V is volume of the unit cell, and omega is the angular frequency of the vibrational mode. Since the E' mode in 1L-WS$_2$ is confined to the plane of the material, we can express the Grüneisen parameter as the change in the frequency due to pure in-plane biaxial strain:

$$\gamma_i = \frac{1}{\omega_i}\frac{\partial \omega_i}{\partial \epsilon} \tag{24}$$

where ϵ refers to the change in area due to stretching. This case would apply at the center of circular nanobubble, and this approach has been used to estimate the Grüneisen parameter for MoS$_2$ nanobubbles[60].
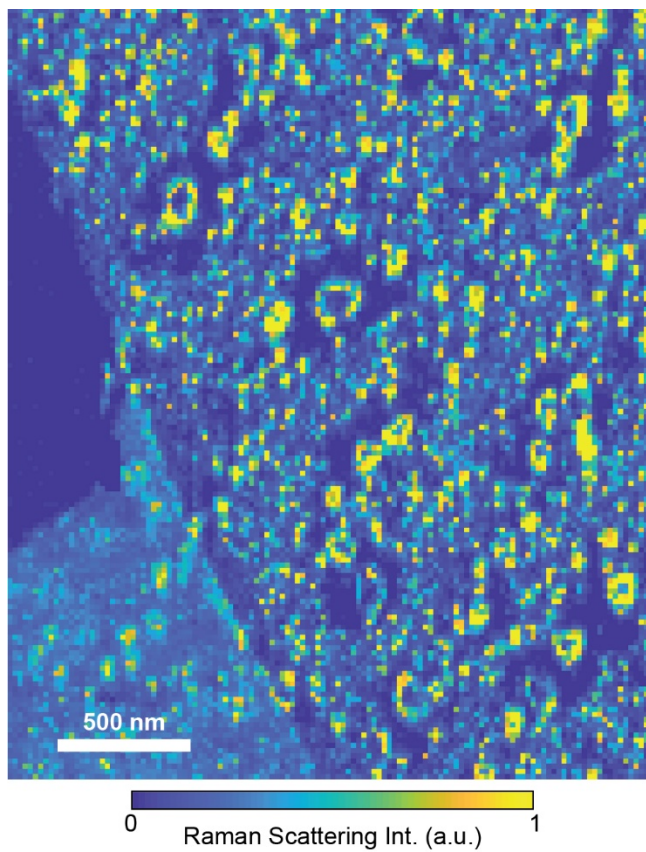
Fig 4-4: Integrated tip-enhanced Raman scattering intensity. Tip-enhanced photoluminescence was removed via a three Gaussian fit.
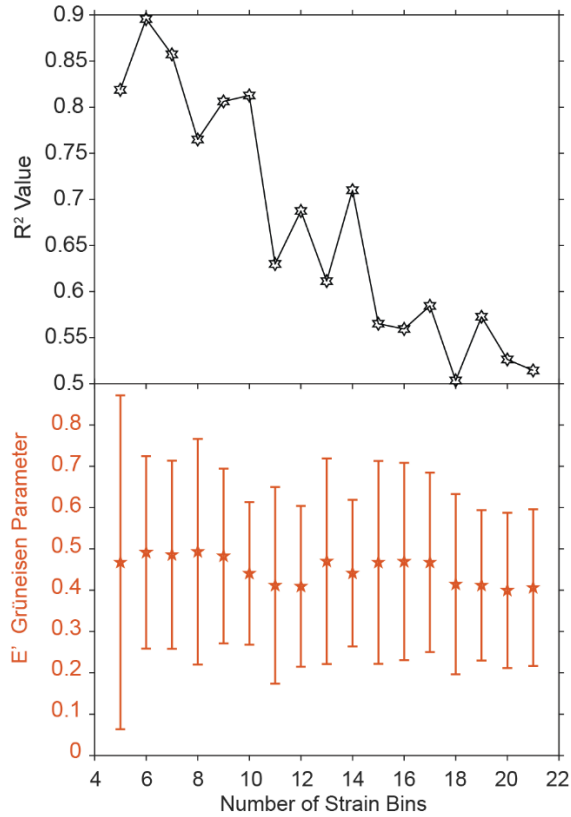
Figure 4-5: Variation of the number of strain bins for estimation of the E' mode Grünesein parameter. (a) the r-squared value for the linear fit to the average strain shifted E' mode for each. (b) the corresponding Grüneisen parameter calculated via equation 7 in the main. Error bars denote the 95% confidence interval.
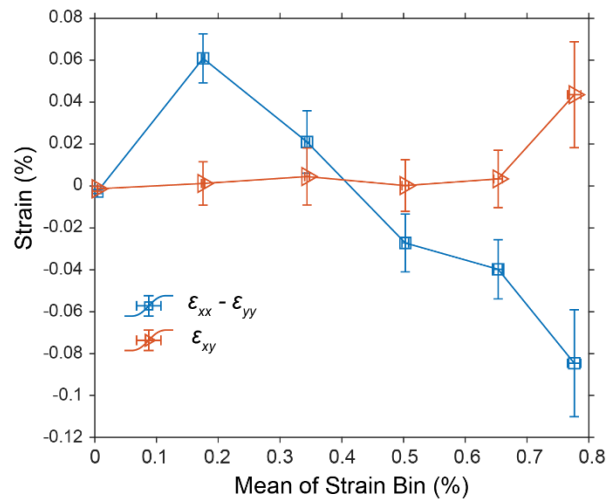

Figure 4-6: Mean values of the shear components of the strain tensor for each value of the trace, for a bin size of 0.16 % Blue squares are the difference of the x- and y-diagonal components of the strain tenors, while orange triangles show the mean of the off-diagonal component. Error bars denote the standard error of the mean.

For the non-circular nanobubbles studied here, this case does not apply. This can be seen especially in the TERS signal, which is greatest at the edges of nanobubbles as seen in Fig 4-4. However, unlike the stretching, which is net tensile, a shear can be both positive and negative. Therefore, when averaging over a large number of nanobubbles, we expect the result to correspond approximately to the hydrostatic case. We show that this is indeed the case in Fig 4-5 for the strain maps in Fig. 4-7. As can be seen, the value is <0.1 % for all strain bins for the shear components of the strain tensor. The averaging of the strain tensor trace over the nanobubbles can thus be treated as approximately hydrostatic or pure biaxial strain, and the corresponding slope should reproduce the Grüneisen parameter for the E' mode.

Figure 4d shows the mean Raman shift of the E' mode versus the strain tensor trace for each of the strain bins, where the error bars are the standard error of the mean. A linear fit to the data is plotted in red, which yields a corresponding slope of -1.75 cm^{-1}/%. Plugging this value into equation (7), we obtain a Grüneisen parameter of 0.49 +- 0.23, which agrees well with the experimental value of 0.5 measured via diamond anvil cell[74] and 0.54 using uniaxial plate-bending techniques[75] respectively. In Fig 4-6b we show that this value is robust against variations of the strain bin size, which indicates that our nanobubble sample size is sufficiently large for the strain to be considered approximately hydrostatic.
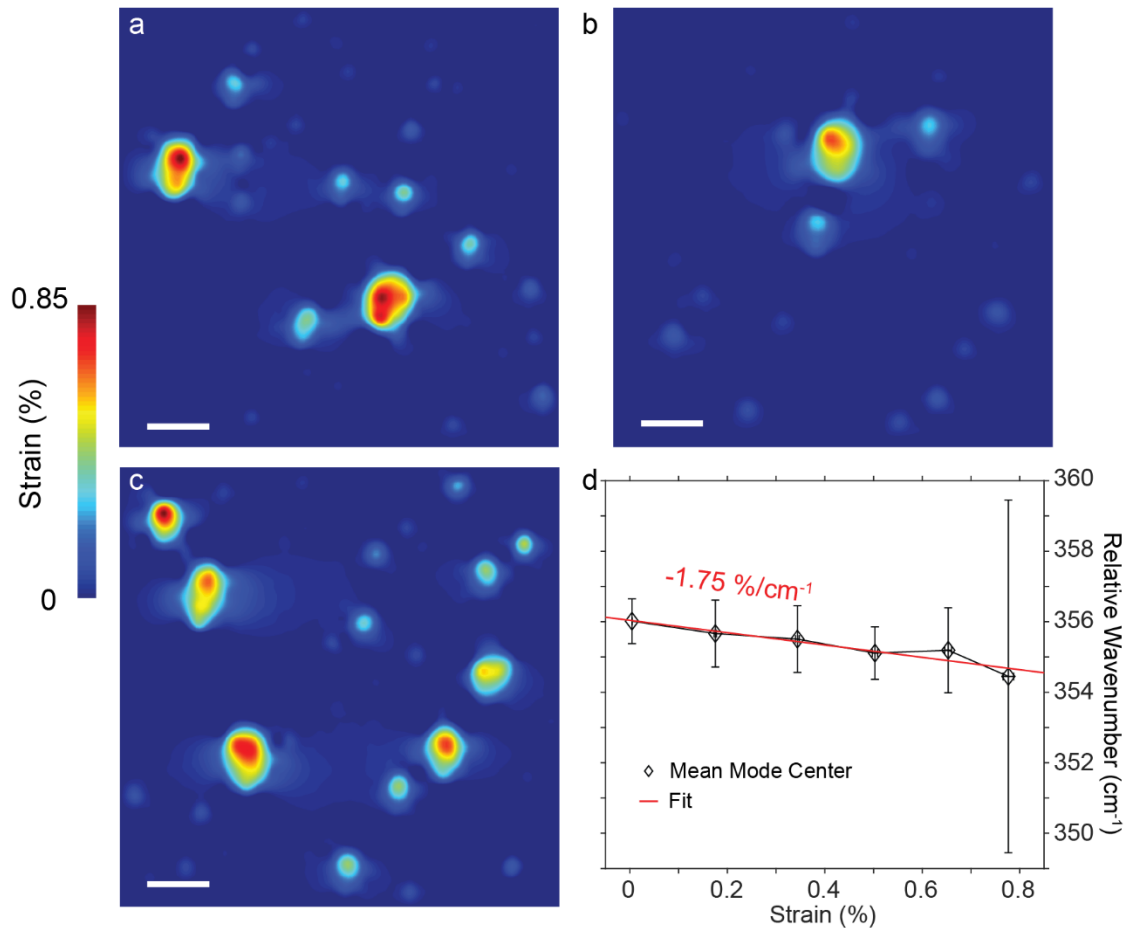

Fig 4-7: (a)-(c) trace of the strain tensor calculated from the second Foppl-von Karman equation for the nanobubbles given in Fig 1c, inset the corresponding AFM micrograph; (d) plot of the center of the WS$_2$ E' Raman mode averaged over multiple nanobubbles for a given value of strain. Error bars are estimated

using the standard error of the mean. A linear fit is denoted in red, yielding a slope of -1.75 cm^{-1}/% corresponding to a Grüneisen parameter 0.49+/- 0.23, agreeing well with other experimental values.

Intensity dependence of TERS/TEPL on probe/sample to substrate

Before concluding the section, the data in the above figures provides an opportunity to investigation the effect of the substrate sample on the Raman and PL intensity. A disadvantage of gap-mode is that the metal will significantly quench the luminescence, which is demonstrated clearly in Figure 4-2, where only the bubbled monolayer emits significantly. Experimentally it is desirable to the separate the target sample and the metal substrate from substrate, which will decrease however at the expense of the field enhancement of the gap-mode. By analyzing the height intensity relationship, we can extract scales for the quenching of the PL and the effects of strain on enhancing both Raman and PL signals. This analysis is shown in Figure 4-8. Both the PL (left) and Raman (right) are exponentially enhanced for the first few nanobubble. The Raman reaches its maximum first around 5 nm away, while the PL has a slower increase after the initial fast phase. The solid lines are fits based on the expressions below. For the PL the emission k_{em} is written as the quantum yield time the exponentially decaying excitation ration, where the quantum yield is written as the ratio of the radiative recombination rate and the total recombination. We assume that the non-radiative recombination consists solely of metal induced quenching that exponentially depends on the metal emitter separation. Further we allow the radiative to change non-linearly to account for effects such as the Purcell enhancement[34].

For the Raman we applied the perturbation theory expression to account for the resonance enhancement[84], times the exponentially decaying excitation field as in the PL expression that has omitted for clarity. Both expressions roughly capture the intensity dependences and provide a quenching distance of about ~0.5 nm. Importantly these curves show that optical nanobubble height to study in gap-mode in terms of signal is around 5 nm.
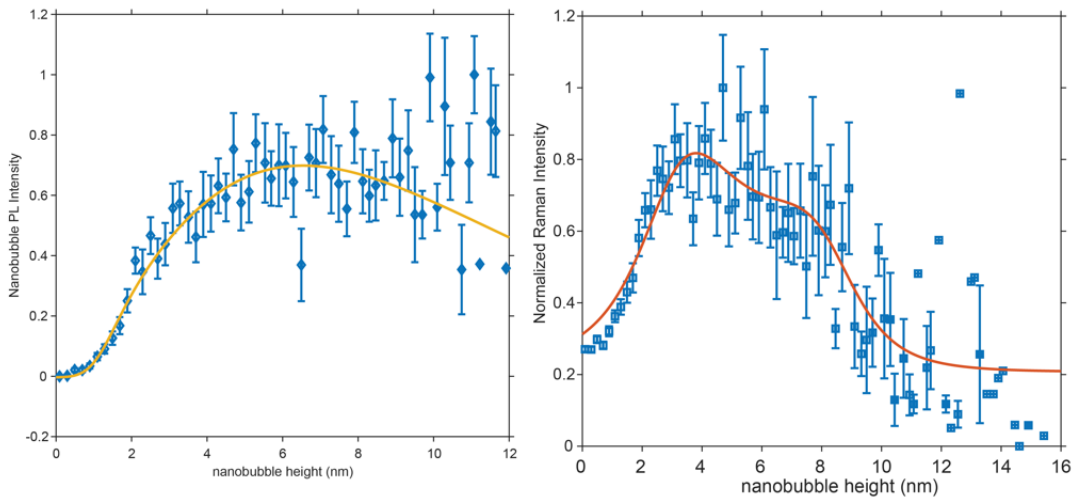
$$k_{em} = Ae^{-\alpha d} \times \frac{k_r^0 d^n}{k_r^0 d^n + k_{nr}^0 + Be^{-\beta d}}$$

$$I \sim \left| \sum_{a,b} \frac{\langle f|H_{el-pt}|a\rangle\langle a|H_{el-ph}|b\rangle\langle b|H_{el-pt}|i\rangle}{(E_{laser} - E_g - i\gamma)(E_{laser} - E_g + E_{ph} - i\gamma)} \right|^2$$

Figure 4-8: Intensity-height dependence of the WS$_2$ nanobubble PL (left) and Raman (right), averaged over the full field in Fig. 4-7. Error bars denote the standard error of mean. Solid curves are fits using the expression below each figure.

Section Conclusion

In conclusion, we have demonstrated a facile numerical recipe using the AFM topography to estimate the strain of nanobubbles in 2D materials. This method is generally applicable to nanobubbles of arbitrary shape and can be easily applied to complex fields of interacting nanobubbles. Using state of the art nano-optical characterization tools, we show that the photoluminescence energy shift agrees with the estimated strain of single non-circular nanobubble. By combining the simultaneously solved strain maps of dozens of irregular nanobubbles, we reproduce the Grüneisen parameter of the in-plane Raman active mode for the 1L-WS$_2$, showing the quantitative accuracy of our method for truly "nano" nanobubbles with radii on order 10-100 nm. Though we have applied this method to nanobubbles, we note that equation (5) makes no assumption of a "bubble" of trapped material, and the recipe may be applied to other deflections of 2D materials where the deformation is induced by a normal load to the material. The resulting accuracy of our method rests on the quality of the atomic force micrograph and the ultimate applicability of modeling the 2D material as thin plate. Because of the wide availability of powerful numerical solvers for partial differential equations and the ubiquity of AFM microscopes for topographic characterization, our approach provides a practical and accessible recipe for estimation of strain in low dimensional materials.

Section 5 – WSe₂ nanobubbles on hBN

The intense light-matter interactions of two-dimensional (2D) monolayer transition metal dichalcogenides (1L-TMDs) are mediated by a diverse suite of excitonic phenomena that present a wealth of opportunities for novel optoelectronic functionalities in areas spanning from high-performance sensing and non-traditional photovoltaics to the quantum information sciences. Many of these opportunities emerge from—and heavily rely on—the unique ways in which the 2D TMD semiconductors enable the manipulation of excitonic phenomena on the nanoscale. Within this quiver of capabilities[6,21-24,85-87], strain engineering is preeminent, offering unprecedented flexibility and precision, and opening new routes for highly tailored optoelectronic materials. Extrinsic strains of up to several percent can be endured without fracture[2,60,62] and have been shown to continuously reduce the optical bandgap[60,81] as well as modify the exciton-phonon coupling, thus narrowing the photoluminescence (PL) linewidth[52]. Further, such strain can be localized, embedding nanoscale potential energy wells for the funneling and localization of excitons[59,61,88]. Notably, in 1L-WSe₂, it has become evident that local strain is a key ingredient for the formation of low-temperature quantum emitters[2,4,89], and that strain-engineering can potentially extend their operating range to room-temperature[5,90]. The resulting highly integrable solid-state non-classical light sources[91] underscore the remarkable technological potential of harnessing nanoscale strain engineering of exciton localization in 2D TMD semiconductors.

Despite the observations of exciton funneling and single-photon emission in localized strained regions of 2D TMDs, an understanding of the induced exciton localization is notably lacking, especially on the nanoscale. This absence of a fundamental picture has led to critical ambiguities, particularly for the formation of quantum emitters in 1L-WSe₂ where the roles and interplay between strain, excitons, and crystallographic defects remain largely a mystery. For instance, it has been shown that the quantum emission can be co-localized with a nanobubble[2] where strain models derived from continuum elastic plate theory predict that a single region of maximum strain (and thus exciton localization) occurs at the apex of the nanobubble[60,62]. However, multiple emitters per nanobubble are typically observed, which is at odds with a single localization site and suggests a possible role of crystallographic defect states or some other sub-nanobubble inhomogeneity[2,89,92-95]. Recently, improved microscopic theoretical models predict that quantum-dot-like electronic states in TMDs can form within nanobubbles in the absence of defects[69,96]. In particular, a first-principles approach that carefully considers the strain-induced atomic structure[69] (and which we employ here) predicts that strain maxima and multiple low-energy states form in a doughnut-like distribution near the nanobubble periphery. This surprising strain distribution is an apparent result of atomic scale wrinkling around the edges of highly-strained nanobubbles, leading to more localized lower energy states relative to the case of a nanobubble with a smooth topography. Such predicted states provide a critical missing puzzle piece for understanding and controlling strain-localization of excitons in 2D TMD semiconductors. However, due to their nanoscale size and distribution within a nanobubble, these localized energy states cannot be directly resolved with far-field optical characterization, necessitating interrogation by more sophisticated higher-resolution techniques.

In this section, we employ apertureless scanning near-field optical microscopy to image the localized exciton (LX) states within nanobubbles in 1L-WSe₂. Using hyperspectral nano-photoluminescence (nano-PL) mapping, we achieve a sub-20 nm spatial resolution and resolve individual localized low-energy states that are separated by distances less than 50 nm within single nanobubbles at room temperature. Furthermore, we observe that these localized states form in doughnut-like patterns that are consistent with the first-principles predictions[69] that refine those of continuum plate and membrane models[60,62]. These LXs are also identified in nanobubbles of "flux-grown" 1L-WSe₂, which has a much lower defect density as compared to monolayers exfoliated from commercially available crystals (ca. 100× lower defect density)[97]. Our results are consistent and reproducible across multiple samples and numerous bubbles (N>50; see, for example, Fig. 5-5). The combination of nano-optical characterization with

atomistic modeling and materials engineering indicates that (i) exciton localization by inhomogeneous strain occurs via the formation of wrinkles near the edges and bending near the base of nanobubbles; (ii) that these excitons remain localized at elevated (room) temperature; and (iii) the optical emission of LXs at room temperature can be enhanced by nano-optical/plasmonic techniques by at least 100×. Together, these findings provide key experimental evidence of highly-confined localized states in nanobubbles, which constitute a new paradigm of nanoscale strain engineering in 2D TMD semiconductors that is particularly relevant for developing non-classical light sources for practical quantum optical devices.

Nano-PL characterization of the localized excitons

Numerous processes that are critical for developing advanced functionality and novel devices with 1L-TMD semiconductors have been studied with nano-optical methods, revealing a rich suite of highly localized optoelectronic phenomena[34,41,49,98-100]. Figure 1a illustrates the experimental configuration of our nano-PL investigations of individual nanobubbles of 1L-WSe$_2$ where a sharp silver tip (Horiba Scientific) has been positioned within 2 nm of the 1L-WSe$_2$ (Fig. 1a). The tip-sample junction was illuminated from the side with p-polarized, continuous wave laser excitation (637.27 nm; $P = 100$ µW) at an oblique angle of incidence such that the induced polarization of the tip can couple to the in-plane and out-of-plane transition dipoles[101]. As the sample is raster scanned, nano-PL from the tip-sample junction and local topography of the 1L-WSe$_2$ are simultaneously recorded.
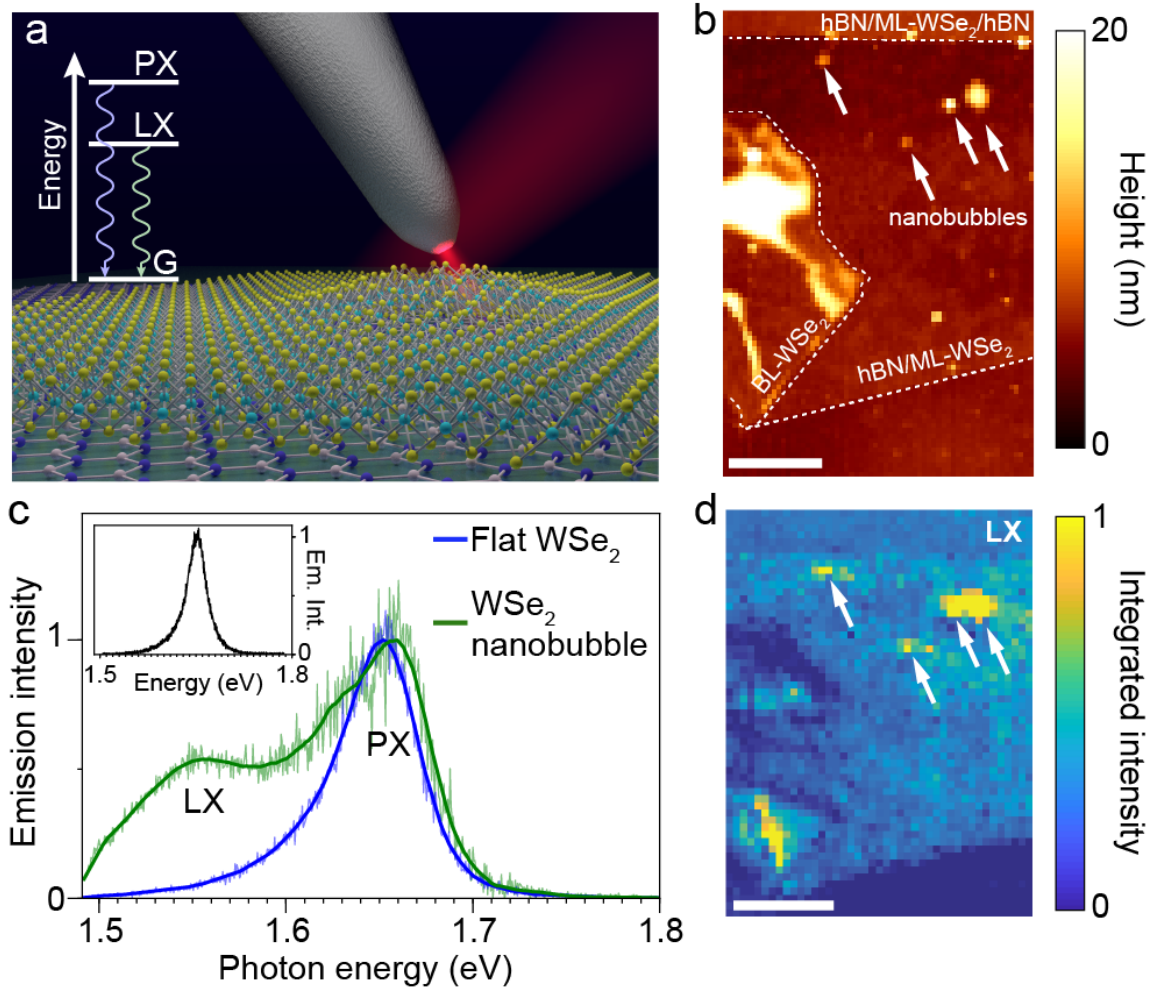

Figure 5-1: nano-optical detection of room-temperature photoluminescence from localized\ exciton (LX) states localized to nanobubbles in 1L-WSe$_2$. (a) Schematic of the room-temperature nano-PL imaging and

spectroscopy of 1L-WSe$_2$ on top of a hBN. Inset: simple level diagram illustrating the energetic ordering between the primary exciton (PX), LX and ground (G) states of 1L-WSe$_2$. (b) AFM topography of the 1L-WSe$_2$ flake exfoliated on top of the hBN substrate. The white arrows mark several nanobubbles in the 1L-WSe$_2$. (c) Comparison of nano-PL emission spectra collected from the 1L-WSe$_2$ on and off a nanobubble with a resolution of 20 nm. Inset: typical far field spectrum of the 1L-WSe$_2$. (d) Spatial map of the LX emission (integrated from 1.5-1.6 eV) showing that the LX emission in the 1L-WSe$_2$ is localized to spatially discrete regions that correspond to nanobubbles in the topography. Inset: similar map of the PX state showing diminished intensity in the bilayer region. All scale bars are 500 nm.

As indicated in the atomic force microscopy (AFM) micrograph (Fig. 5-1b), the entirety of the 1L-WSe$_2$ in Figs 5-1 – 5-3 is supported by an underlying crystal of hexagonal boron nitride (hBN; ~200 nm thick) forming a 2D 1L-WSe$_2$/hBN heterostructure. We note that a portion of this 1L-WSe$_2$ flake is fully encapsulated and covered by a second, thin hBN crystal (~3 nm). However, all of the results presented here were acquired from the exposed 1L-WSe$_2$ that is directly accessible to the nano-optical antenna (as illustrated in Fig. 5-1a). The details of the "dry-stamping" sample fabrication are provided in earlier work[2]. This process is known to produce structural "imperfections" such as nanobubbles, tears, folds and wrinkles. In our case, a portion of the 1L-WSe$_2$ is folded on top of itself to form a bilayer and several pronounced creases (Fig. 5-1b). In addition, small topographical nanobubbles, with characteristic heights of ~2-10 nm, are observed in the 1L-WSe$_2$ and are attributed to trapped substances between the layers[62]. Such nanobubbles have been repeatedly observed in heterostructures of layered materials[2,61,102], and as noted above are known to be correlated with quantum emitters in 1L-WSe$_2$ at cryogenic temperatures[2].

Figure 5-1c contrasts nano-PL spectra collected from a flat region and a nanobubble of the 1L-WSe$_2$ acquired at room temperature. On the flat region, the spectrum is composed of the standard excitonic PL of 1L-WSe$_2$ at ~1.65 eV,[98,103] which we will refer to as the primary exciton (PX). When the nano-optical antenna is positioned over a nanobubble, the emission spectrum shows a dramatic change, exhibiting red-shifted excitonic emission and an additional intense low-energy emission band that is centered at ~1.56 eV, as determined by Gaussian peak fitting . In the 1L-WSe$_2$, every region that was found to exhibit strong low-energy emission was correlated with a nanobubble. A selection of such emission spectra from four individual nanobubbles and their associated topographies is shown in Figure 5-2 The lateral extents of the nanobubbles range from ~30–140 nm with heights from 1.5-20 nm, and depending on size, exhibit a diverse set of broad emission spectra, some of which show signatures of multiple low-energy states. In general, low-energy emission occurs at emission energies that span a range of about 70 meV, from ~80 meV to ~150 meV below the PX, and exhibits linewidths of ~150 meV that are attributed to thermal broadening[90]rand plasmonic coupling to the tip[104].

Close inspection of the nano-PL spectra in Fig. 5-1c reveals several characteristics of the low-energy PL that are different from the well-studied PX state. First, we concurrently observe both the low-energy emission and strong emission at ~1.65 eV, corresponding to the PX spectra measured in the far-field (inset Fig. 5-1c), within our nano-optical mode volume. Second, the low-energy band here is well below the energies of other native excitonic complexes in this material (e.g. trions, the dark exciton[13,105], and biexcitons[106] and in an energetic region often associated with defect states[41]. Third, the linewidth is substantially larger than that of the PX, which contrasts with what has been observed for strain-tuned PX states[52]. And finally, the energy separation of this band with respect to the PX state (see also Fig. 5-2) and its localization to nanobubble regions are in-line with corresponding observations for single emitter states at cryogenic temperatures in 1L-WSe$_2$[2,89,93-95,107,108] Based on these considerations, we deduce that this low-energy band, which is amplified significantly by the nano-optical antenna (consistent with the observation in ref[90][90] of plasmon-enhanced localized emission at elevated temperatures), originates from LX states in the nanobubbles.
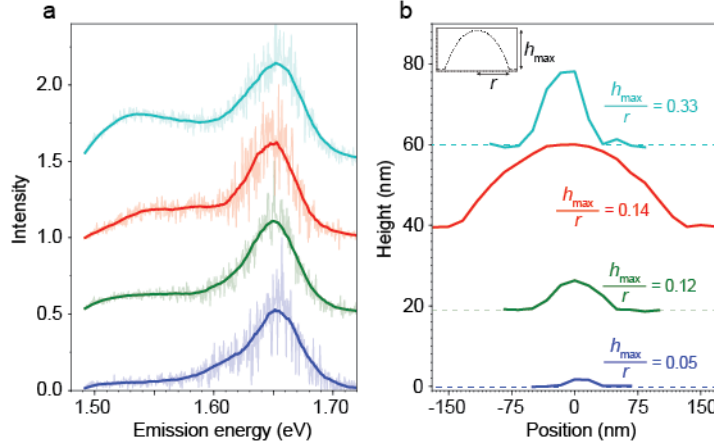
Figure 5-2: average localized LX emission and topography from four individual nanobubbles of ML-WSe$_2$ at room temperature. (a) Emission spectrum for each nanobubble. (b) Corresponding topography for each nanobubble. Each profile is offset for clarity and the dashed lines mark the height of the surrounding ML-WSe$_2$. Inset is a schematic of the "universal bubble" shape with an aspect ratio (h_{max}/r) of 0.14. The spectra and topography are sorted from bottom to top with increasing aspect ratio where the bottommost (topmost) spectrum-topography combination is the nanobubble with the smallest (largest) aspect ratio. We note that the lateral extent of the smallest nanobubble might be resolution-limited by our AFM topography measurement and that the actual aspect ratio could be larger.

The nano-PL directly probes the LX emission on length scales that are commensurate with – and even smaller than – the nanobubbles themselves, allowing us to explore the origins of the diverse emission spectra and nanoscale structure-property relationships in more detail. Here, we estimate that our spatial resolution is ~15 nm. In Figure 5-3, the spatial distribution of LX emission is shown for a larger nanobubble (approximate radius of 75 nm; peak height of 12 nm; aspect ratio of 0.16). With nano-PL, we resolve nanoscale variations of the LX emission within the nanobubble itself, finding that the integrated emission intensity is not uniform (integrated over the LX spectral region of 1.5-1.6 eV; Fig. 5-3a): it is concentrated to specific locations of the nanobubble that do not clearly correspond to its apex (cf. inset of Fig. 5-3a). Furthermore, different points of the nanobubble separated by distances as small as 30 nm exhibit clearly distinct spectra (Fig. 5-3b). The dashed curve is drawn as guide to the eye between the maxima of LX spectral peaks, clearly showing that the redder LX states are positioned to the edges of the nanobubble, where the maximum confinement potential and possible wrinkle formation are expected to occur. Clearly, larger nanobubbles host multiple emissive localization centers for LXs, an effect that to date has only been inferred spectroscopically[2,107].
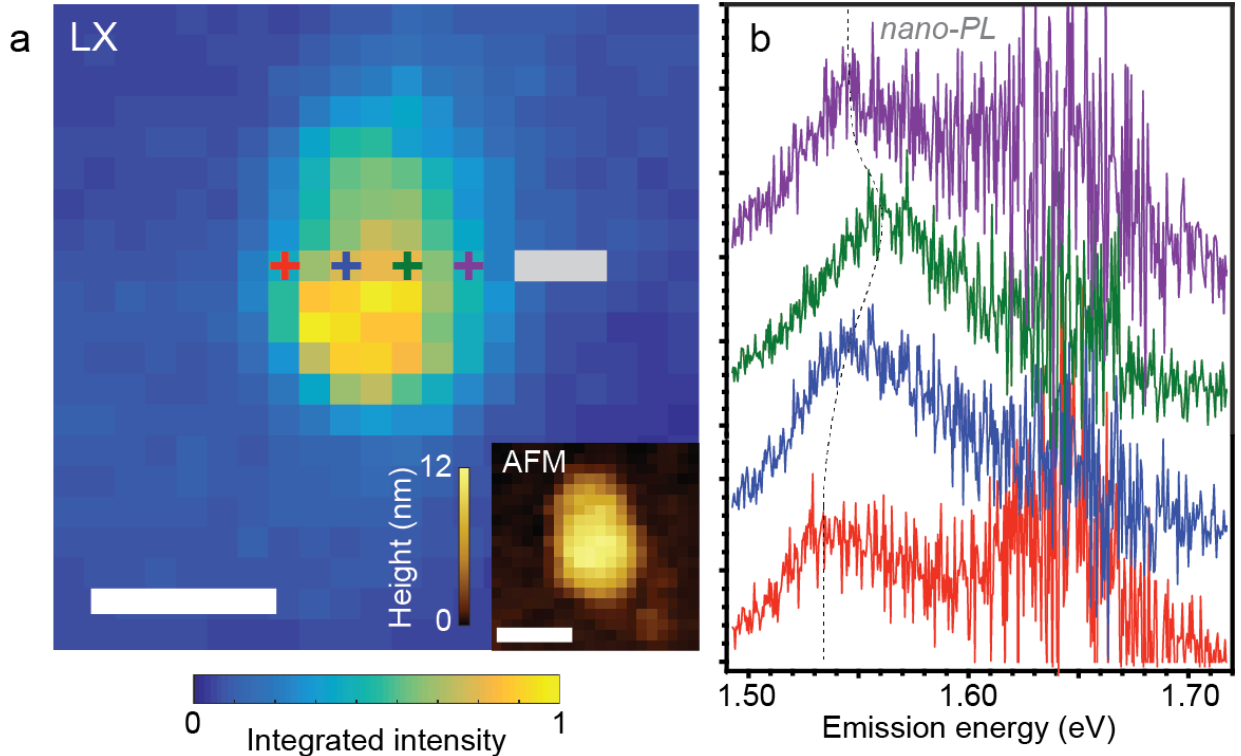
Figure 5-3: high-resolution nano-PL imaging and spectroscopy of distinct LX localization centers within a single nanobubble. (a) Nano-PL image of the spatial distribution of LX emission (integrated intensity from 1.5-1.6 eV) of a single nanobubble. The three grey pixels are a scan artifact where nano-optical signal was not acquired. Scale bar: 100 nm. Inset is the topography of the nanobubble (scale bar: 100 nm). Its overall height is 12 nm and its average radius is 74 nm. (b) Sample emission spectra from the corresponding points labeled in panel (a). The distance between the red and blue marks is 45 nm.

Determining the strain field

Our hyperspectral mapping of the nanobubble in Fig 5-3 shows that the lowest-energy states (and thus deepest confinement potentials) are concentrated around the edge. We showed previously that the AFM topography can be combine with theoretical modeling to approximate the strain field within the 2D monolayer. With this approach we developed a continuum plate approximately that agreed well with the general shift in the exciton energy and allowed us to rederive the Grüneisen parameter for the 1L-WS$_2$ E' mode. To gain more insight into the physical mechanism, we applied a more advanced atomistic the theory of ref[69][69], which considers the influence of local strain within the nanobubbles on electronic and optical properties. Our method is based on atomistic calculations, in which the relevant part of the nanostructure is modeled using a lattice of atomic sites. For the present calculations, we use the nanobubble geometry obtained from AFM measurements and determine the corresponding relaxed positions of the individual atoms in a valence force field simulation. The result provides the strain field in the structure and confirms the existence of atomic-scale wrinkling. In a second step, the information about the displaced atomic positions is used within a tight-binding calculation in order to quantify how the strain field of the individual atoms within the nanobubble translates into a confinement potential and local electronic states.

Our calculations are performed in four subsequent steps, starting with the experimentally determined morphologies, followed by a computation of the structural relaxation of the atomic lattice, electronic-state calculation, and the determination of corresponding optical spectra.

The starting point of all calculations is a rectangular supercell of a MoS$_2$ bilayer with periodic boundary conditions and an interlayer equilibrium distance of 3.0 Å in accordance with ref[109][109] and comparable to recent experimental investigations[62]. Since REAX parameters are presently available only

for MoS₂, our calculations are performed for this material. Similar results are expected for WSe₂ due to the very similar elastic properties that determine strain-induced localization effects.

In the first step, the experimental AFM topography (see Fig. 5 Row I) is interpolated onto the triangular TMD lattice of the upper layer. While the size is scaled down by a factor of five to facilitate the highly demanding computation, the size of nanobubbles has only a minor influence on the localization pattern[69]. The upper layer is shorter in both lateral directions than the lower layer so that the number of atoms forming the nanobubble geometry can change during the subsequent relaxation of the upper TMD layer, while keeping the atoms in the lower layer fixed. For this part, we utilize a REAX potential[110] with the parametrization from ref[109] within a valence force field calculation. The relaxation is performed until a force tolerance of 10^{-3} eV/Å or a relative energy tolerance of 10^{-16} is reached. In each iteration step, the charge equilibration energy is minimized until it reaches a value below 10^{-3} eV. During the 0 K energy minimization the bubble geometry is optimized.

The new equilibrium positions of the individual atoms are used in subsequent electronic-state calculations. For this part, we solve a 6-band tight-binding Hamiltonian that accounts for both strain-induced band gap changes[69,110] and changes of the dielectric environment[69]. In our tight-binding model, the local change of interatomic distances translates into a confinement landscape (see Fig. 4 Row III). From a diagonalization of the Hamiltonian, we obtain the electron and hole wave functions, as well as the corresponding overlap (see Fig. 4 Row IV).

In the last step, the PL energy shift (see Fig. 5 Row V) is calculated. For this part, we use a spatial grid and determine free particle PL spectra for each spatial area based on the corresponding wave function contributions from this area. By extracting the shift of the highest PL intensity from each local area, we obtain a PL map that is directly comparable to the experimentally measured LX emission.

Comparison of experiment and theory

For the nanobubble topography given by the AFM data of Fig. 5-3a (inset), the calculated confinement potential (Fig. 5-4a) exhibits a doughnut pattern, with deeper potentials located on the nanobubble periphery that traces the wrinkling effect (Fig. 5-4b) and corresponds remarkably well to the nanoscale spatial distribution of measured LX energies (Fig. 5-4c). This correspondence is further in direct accordance with the experimental spectra in Figure 5-3b, which shows that the low-energy emission is isolated to the nanobubble periphery. To demonstrate the correlation between the theoretical predictions and experimental data, Fig. 5-4d shows the measured LX energy plotted against the predicted confinement potential at the corresponding region of the nanobubble. Around the periphery of the nanobubble, a strong correlation between the LX energy and the depth of the confinement potential is observed: regions of stronger confinement correspond to lower LX energies. In contrast, in the center region, very little correlation of energy with confinement potential is observed, which suggests a lack of strongly confined excitons in that area. This different behavior may indicate a more-general exciton funneling effect in the central region[59], whereas the periphery is dominated by highly confined states.

The correspondence between the theoretical analysis of the strain and resulting electronic states within the nanobubble and our nano-optical dataset presented in Fig. 4 strongly suggests that the experimentally localized states observed are the predicted quantum-dot-like states[69]. In principle, point defects in the 1L-WSe₂ lattice can also localize excitons in a similarly inhomogeneous way. To better elucidate the potential role of defects, we repeated our nanobubble PL mapping and theoretical calculations in high-quality flux-grown 1L-WSe₂, which has been shown to have defect densities approximately two orders of magnitude lower than commercially grown crystals–providing for the first time an average spacing between defects on the order of nano-optical resolution[97]. In fact, every nanobubble identified in this higher-quality material exhibits similar LX emission to that of the commercial material More

specifically, the occurrence of the LX emission band in the nanobubbles does not depend on the defect density in the ranges probed here.
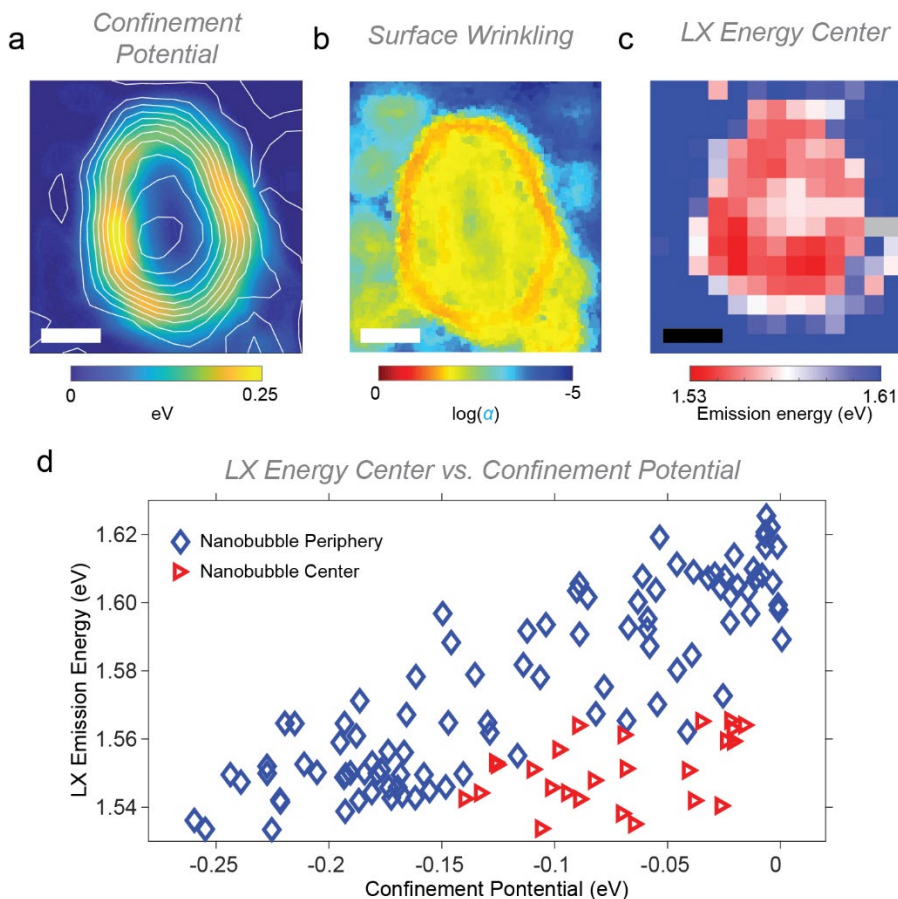

Figure 5-4: (a) Calculated carrier confinement potential for the nanobubble topography shown in Fig. 5-3a inset. The solid lines represent topographic contours based on the AFM measurement. (b) Corresponding values of the surface normal deviation, defined as the average angle α between normal vectors at neighboring unit cells, obtained from relaxed atomic positions, showing wrinkling of the unit cell. (c) Map of experimentally measured LX emission energy from the hyperspectral nano-PL shown in Fig 2. (d) Correlation between the local confinement potentials in (a) and LX emission energy shown in (c). The datapoints are separated into points from the center of the nanobubble (red triangles) and points from the periphery of the nanobubble (blue diamonds). The scale bar of the experiment in (c) is 50 nm, while theory calculations with atomic resolution are scaled up by 5x.

Below bandgap excitation to highlight localized exciton states

Because the primary exciton (PX) of WSe_2 is strongly emissive in all regions of the monolayer (including regions surrounding nanobubbles), when the monolayer is excited with 633 nm, which is well above the optical bandgap, it presents a considerable amount of background signal that can overwhelm the study of highly localized excitons (LXs) at nanometer resolution. To decrease background from the normal excitons of the monolayer and bilayer areas, we scanned the same area but instead excited with a 785 nm laser, which is below the optical bandgap of the monolayer and bilayer, but still above the emission energies of the LXs states. The resulting map is shown in Figure 5-5. As can be seen, the monolayer and bilayer free excitons are dramatically suppressed, allowing observation of LX emission from nearly every nanobubble

in the field. Notably, smaller nanobubbles in bilayer region show a "doughnut" intensity pattern. This feature arises from the effect of the interference filter used to block the incident laser intensity. For the smaller nanobubbles, the emission in their centers, which is at somewhat higher energy than the emission near the periphery, is significantly attenuated by the interference filters. The edges of the nanobubbles emitting at redder energies thus experience less attenuation, appearing brighter in emission intensity. The consistency of the nano-PL image is striking, and may indicate that these nanobubbles are all of comparable sizes, or of sizes below our resolution.
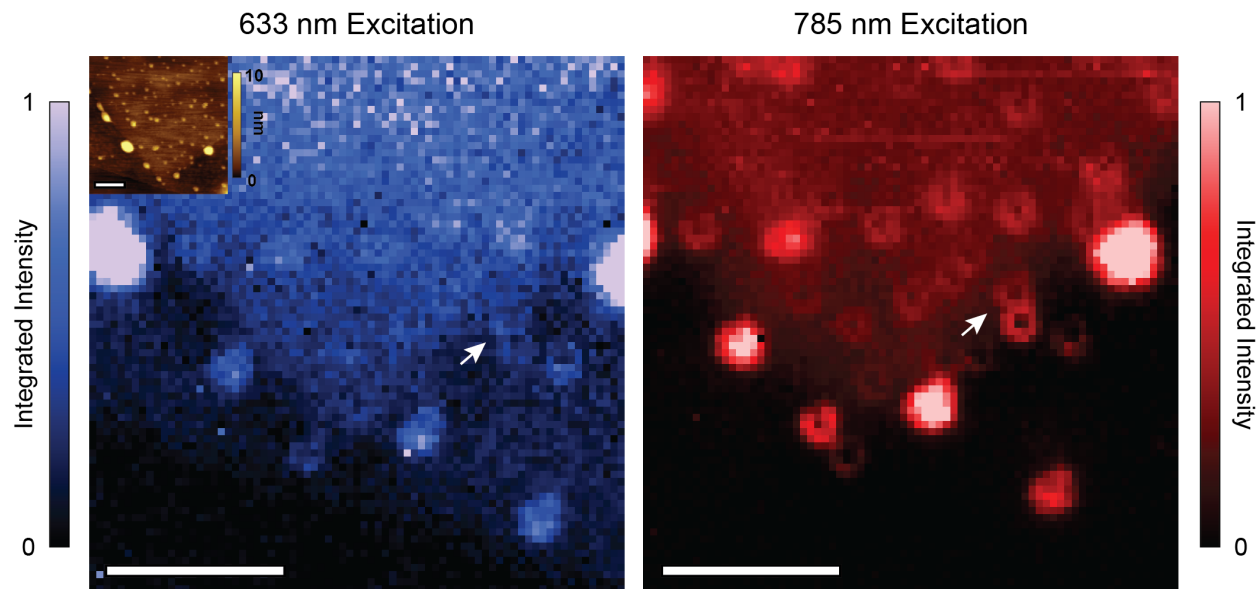

Figure 5-5: Nano-PL integrated intensity maps (integrated emission between 800-820 nm) of nanobubbles presented in Figure 4 using two different excitation wavelengths, 633 nm and 785 nm, which are above and below the optical bandgap of 1L-WSe$_2$, respectively. The white arrow highlights a bubble that is difficult to discern using 633 nm excitation, but is clearly seen with 785 nm excitation.

Imaging similar LX emission in low defect 1L-WSe$_2$

In Figure 5-6, we provide a collage of four nanobubbles in the monolayer area of the high-quality (i.e., flux-grown) 1L-WSe$_2$. Row I of Figure 5-6 contains the measured AFM topography for the four nanobubbles with their corresponding spatial distributions of LX emission energies using 785 nm excitation shown in Row II. Exciting with 785 nm light, which is below the energy of the PX, enhances the contrast between the LX and PX states, allowing the LX states to be observed with a greater signal-to-noise ratio. Rows III, IV, and V summarize the theoretical results of the confinement potentials, overlap of the electron-hole wavefunctions, and the predicted shifts in PL emission energies, respectively. As with the nanobubble shown in Fig. 5-4b, each nanobubble in the flux-grown 1L-WSe$_2$ exhibits lower energy emission concentrated at the nanobubble periphery. This peripheral spatial distribution of low-energy states is replicated in the predicted confinement potentials and electron-hole wavefunction overlap shown Rows III and IV of Fig. 5-6, respectively. Finally, the predicted PL energy shifts and corresponding spatial distributions shown in Row V agree well with the measured LX emission energy (Row II). Considering the challenges of room-temperature hyperspectral imaging at the 10-20 nm length scale, limitations of the topographical maps in Row I to a resolution of ~20 nm, as well as the complexities to directly incorporate these measurements into a feasible atomistic model, the agreement between experiment and theory is particularly remarkable.
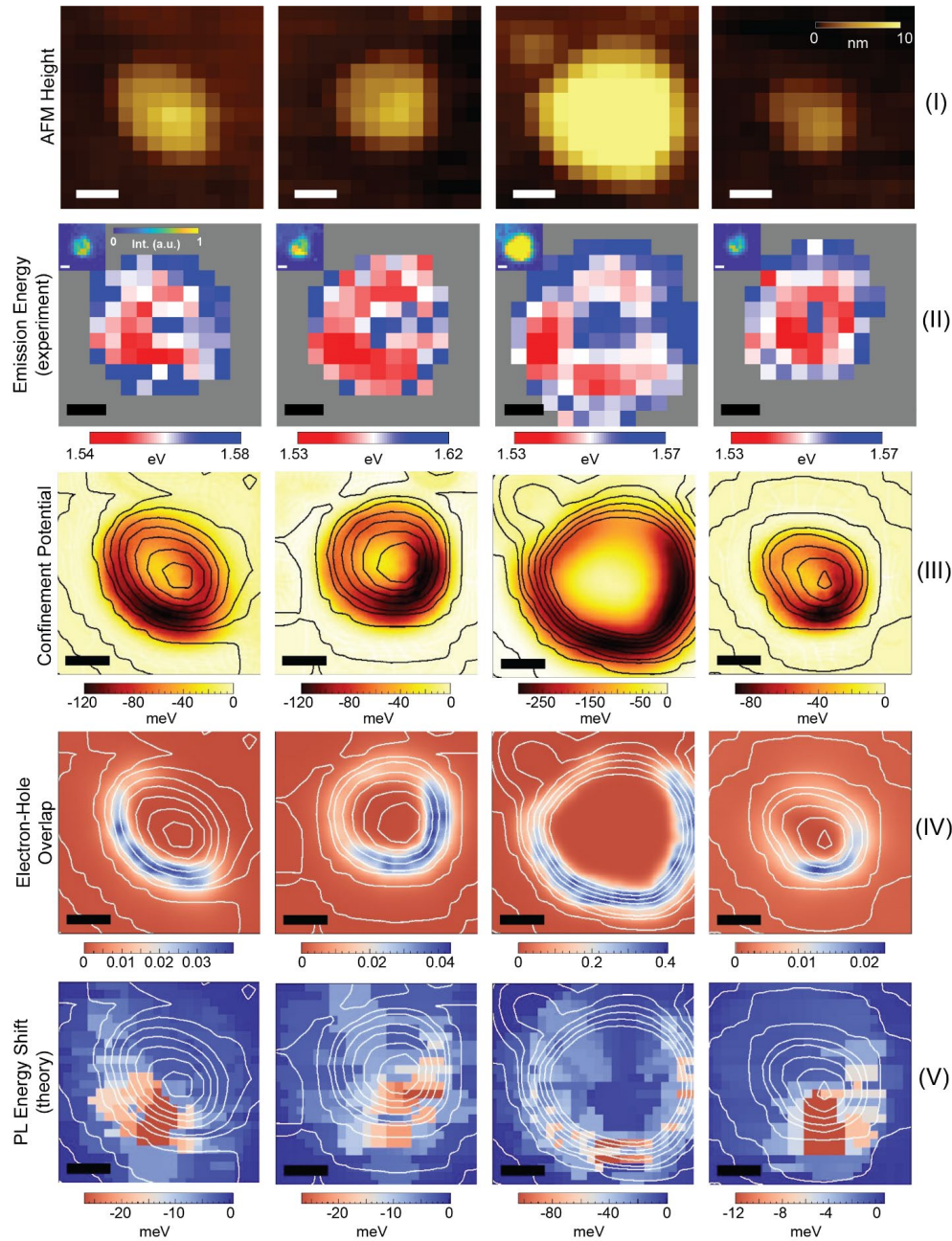
Figure 5-6: Comparison of experiment and theory for 4 nanobubbles in low-defect monolayer WSe_2. Rows I and II show the measured AFM topography and nano-PL emission energy, respectively. The inset of row II shows the integrated PL intensity after a low pass energy filter to isolate the LX emission. Row III provides theoretical results of the confinement potential obtained from displaced atomic positions using nanobubble strain calculations and the Harrison rule[111]. Rows IV and V contain electron and hole wavefunction overlap, and predicted PL energy shift, respectively. The solid lines in rows III-V represent topographic contours using the nanobubble topographies shown in Row 1. Scale bars are 50 nm in Row I and II. Theory calculations with atomic resolution are scaled up by 5x.

The similarities between Figures 5-4 and 5-6—despite dramatically different defect densities—suggest that LX emission distribution in nanobubbles is not significantly determined by atomic defects in the lattice. As noted above, since there is approximately one defect every 15 nm × 15 nm on average in flux-grown material – near the spatial resolution of the nano-optical probe – then major defect-associated optical changes would be resolvable. While defects may still play a role in the emission, the implication here is that the primary effect in terms of the nanoscale spatial distribution and creation of quantum-dot-like states is the local strain profile, with localized maximum strain occurring at locations around the edge of nanobubbles leading to LX emission observable at room temperature with a nano-optical probe. Still, it is certainly possible that defects play a role in e.g. further localization or the exact distribution of wrinkles at the atomic scale. Such angstrom-level effects are beyond the resolution of our current nano-optical probes, which utilize AFM tips with radii of curvature ranging from ~15 to ~40 nm. Future correlations with e.g. atomic-resolution AFM[112] should be able to map these effects.

Section Conclusion

In conclusion, using nano-optical techniques, strain localized exciton states within single nanobubbles of 1L-WSe$_2$ are observed at room temperature with localization lengths scales of less than 50 nm and significantly smaller than the nanobubble itself. In general, the spatial distribution of these low-energy states forms a doughnut-like pattern around the periphery of the nanobubble that surrounds a region of higher-energy states at its center, which is consistent with previous theoretical predictions of atomic-scale strain effects. By integrating the experimentally determined topography with this atomistic model, the experimental results are qualitatively replicated by the theoretical model, establishing the first robust experiment-theory connection of the formation of strain-induced quantum-dot states in monolayer TMD semiconductors. Furthermore, we reproduce these findings in 1L-WSe$_2$ with orders-of-magnitude lower defect densities, implying the nanoscale localization is primarily influenced by strain and not crystalline defects. Our results show that LX states accessed with nano-optical techniques remain localized at room temperature, leading to compelling potential applications in practical strain-engineered optoelectronic and quantum-optical architectures.

Section 6 – Gap-mode nano-PL and nano-Raman on WSe₂/Au

To follow up on the localized exciton measurements of WSe$_2$/hBN substrates, I utilizing gap-mode tip-enhanced PL and Raman spectroscopy to probe nanobubbles of WSe$_2$ exfoliated onto Au. Here we can take advantage of the high field localization shown in Section 3 and allow us to investigate the localized exciton emission, with the same high resolution and contrast seen in the nanobubbles of WS$_2$ in Section 4. Further we can also use the same excitation wavelength 785 nm as the previous section. We will show that the combination of gap-mode and near-IR excitation wavelength allows high sampling of nanobubbles with pixel sizes that approach the scanning system limits, and use this to gain more insight into the localization observed for WSe$_2$ nanobubbles in Section 5.

High resolution characterization of a single nanobubble

Figure 6-1 shows a large area AFM of such a sample. The sample was prepared with the same method presented earlier for WS$_2$ nanobubbles, and similarly the field contains multiple nanobubbles, including several dozen larger nanobubbles, height ~> 10 nm, in addition to numerous smaller nanobubbles.
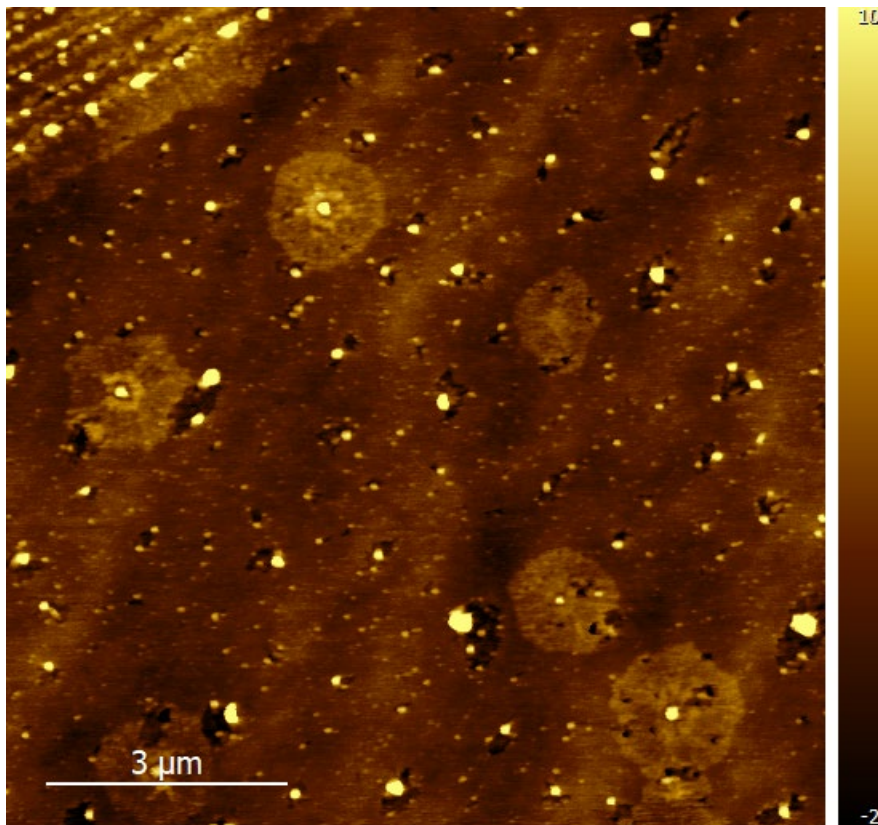

Figure 6-1: Semi-contact AFM topography of nanobubbles in monolayer-WSe2 on template-stripped Au.

An example integrated TEPL map is shown in 6-2, showing the integrated nano-PL in two spectral bins excited with HeNe laser (632.8 nm), left 750 – 820 nm, plotted in a blue colormap (b), and

820 – 950 nm plotted in red (c). The red map has some clear similarities the nano-doughnuts in Fig. 5-5 using 785 nm, however with integrated intensity asymmetric weighted in the horizontal direction. Interestingly the higher energy blue integrated-TEPL map, the intensity concentrated in the center, compared to the AFM topography shown in Figure 6-2a, the bright center distribution is apparently smaller than physical extent of nanobubble while there is also observable emission outside the nanobubble boundary. The red emission on the other hand seems to correspond roughly to the AFM spatial extent, though does appear to extend slightly beyond it in the top left.
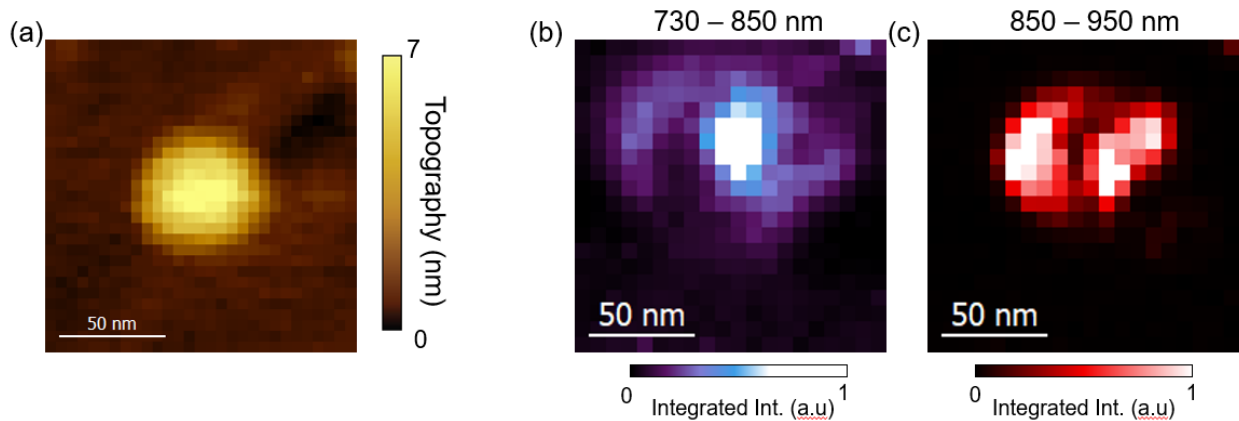

Figure 6-2: Integrated TEPL of WSe$_2$/Au nanobubble. (a) Tapping mode topography, and (b) and (c) integrated PL in higher (blue) and lower (red) spectral bins.
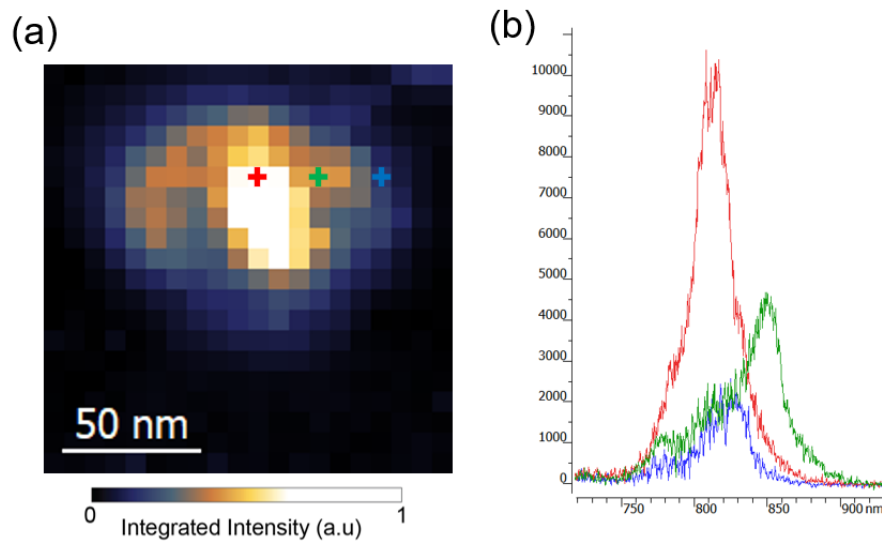

Figure 6-3: point spectra from the nanobubble in Figure 6-2.

Figure 6-3 shows representative spectra in panel (b), indicated by the same-color crosses in the total integrated TEPL intensity map shown in (a). All the emission is notably redshifted from the normal exciton energy of WSe$_2$ at room temperature, emitting at around 750 nm (1.65 eV), which is expected from the effects of strain and increased screening from the Au substrate[69]. PL from the outmost edge, the blue spectra consists overlapping peaks, with local maxima found around 775, 810, and 820 nm (1.60,

1.53, and 1.51 eV respectively). The green spectra has similar peaks, with the addition of a strongly redshifted peak at ~855 nm (1.45 eV). Emission at the center of the nanobubble, the red spectrum by contrast is highly localized at near 810 nm, and is much more intense than previous spectra.

Like the WS_2 nanobubbles on Au, we can interpret the spectra observed in Fig. 6-2 as a combination of the formation of localized exciton (LX) states and exciton funneling. The later effect explains the bright center spectra straightforward, as it is energetically favorable for free or weakly localized excitons to diffuse toward the nanobubble center where there is significant strain, as showed in WS_2 and was observed recently in larger MoS_2 nanobubbles[61]. Funneling may also explain the larger optical radius of the nanobubble compared to the AFM. Since the strain of the nanobubble extends outside the nanobubble boundary, the funneling effect on excitons should still be present despite the flatness of the monolayer. The predominance of the LX emission at the nanobubble edge is expected from the previously WSe_2 nanobubble on hBN. We can verify this by computing the spectral median of the emission, shown in Fig. 6-5a, which as expected displays the doughnut like pattern the median emission energy.
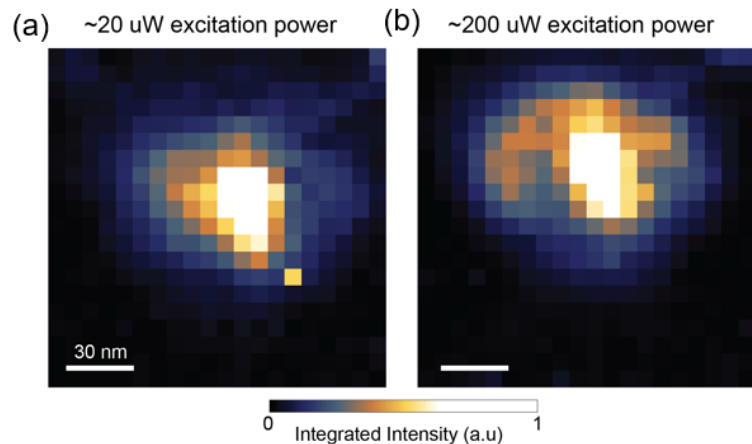

Figure 6-4: comparison of integrated TEPL for nanobubble at different excitation powers.
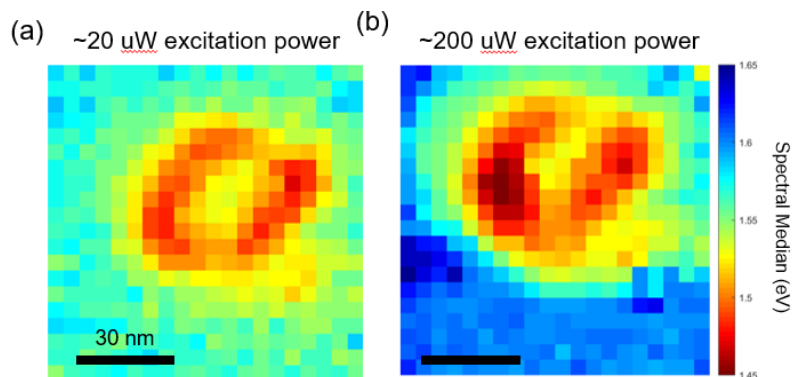

Figure 6-5: spectral median maps (eV) for the TEPL maps in Figure 6-4. (a) 20 uW incident power and (b) 200 uW incident power.

The asymmetry of the LX emission in 6-2c has several possible origins. The first that is a real effect, possibly reflecting a corresponding asymmetry of the monolayer, either extrinsic strain, i.e. strain not induced by the pressure of trapped material, or other inhomogeneity not captured in the atomic force micrograph. The second is that it is measurement artifact, reflecting possibly asymmetry of the probe's

enhancement of the electric field, or drift of the scanning piezo over the scan duration. Since the asymmetry is also reflected in the emission energy the former is unlikely presuming a linear response of the LX emission. However, drift in the scanning piezo however always possible, and usually expected, in all scanning probe measurements. It is straightforward to the check this by rescanning the nanobubble in question, changing the total scan time. Figure 6-4 shows the a comparison of the total integrated nano-PL intensity of the previous scan (a), where each pixels was scanned with an integration time of 80 miliseconds (ms) with ~200 uW incident power, against the longer scan at 150 ms integration time and ~20 uW incident power. As can be seen, the morphology of the nanobubble PL emission is same in both intensity maps, suggesting the asymmterty reflects the true emission intensity distribution.

However, there are slight differences in the distribution of the LX emission between the incident powers, for instance comparing the top and bottom of the nanobubble. One might guess that some such differences may be due saturation effects, which we would have a greater effect on the more localized excitons. If saturation where present would expect then a higher incident powers that the LX may shift to blue, as the lower lying excitons saturate, force more of the excited excitons to emit a higher energy, which would be reflected in spectral median. Unexpectedly when the map of the spectral medians of the two powers, shown in Figure 6-5, we observed a red shift with increasing laser power ~ 10 meV, especially localized to a few pixels in the bottom left. This surprising results in not consistent with saturation.

The redshift with increased laser power as two possible explanations. The first is that the scan was under-sampling the nanobubble, and the that small shifts in the x,y-positioning led to the probe measuring as different LX emitter between the higher and lower power scans. This would imply that resolution of the probe is better than one pixel (7.5 x 7.5 nm^2). Such fine resolutions are possible in gap mode. This possibility can be checked by comparing neighboring spectra in slow and fast scan directions (vertical and horizontal for maps). Such a comparison is shown in Figure 6-6, showing spectra indicated by the colored crosses in same manner as Figure 6-4. While there is certainly discernable difference in the spectral weight implying that the maps are somewhat under sampled, there are also significant commonalities.
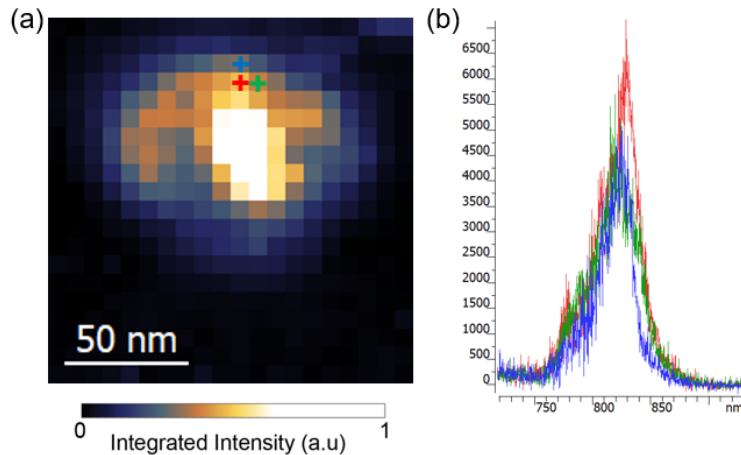

Figure 6-6: Neighboring spectra comparison of nanobubble scan in Figure 6-2.

While it is still possible that the lower intensity scan missed the state seen in the higher power nano-PL map, we feel this is unlikely. Direct comparison of the two spectra from the sample area of the

69

nanobubble supports this argument, plotted in Figure 6-7. The higher power spectra (plotted in orange-red) has clearly defined peak approximately 30 meV below the lowest emission peak observed in the lower power spectra. Moreover, the emission intensity out-competes all other emitters in terms of peak intensity. Because of the strength and spectral separation, we argue of that this peak is an emergent feature in the higher power scan.

The observed energy separation of the blue and red-orange spectra of 30 meV is similar to the observed binding energies of the biexciton in WSe_2[16]. Because the biexciton is formed by binding of two excitons into a four-particle state, it appears only at sufficient excitation densities where exciton-exciton interaction becomes significant. Ascribing the observed peak in Figure 6-7 to biexciton emission would explain it emergence at higher powers as well as the coincidence of its energetic shift from A localized biexciton state at the scale of 7.5 nm has not been observed experimentally in strain-confined excitons in $1L-WSe_2$ to our knowledge, but is well known in the quantum dot literature[113]. More experimental observations are necessarily before such assignment can be conclusively made, but it is nonetheless and intriguing possibility.
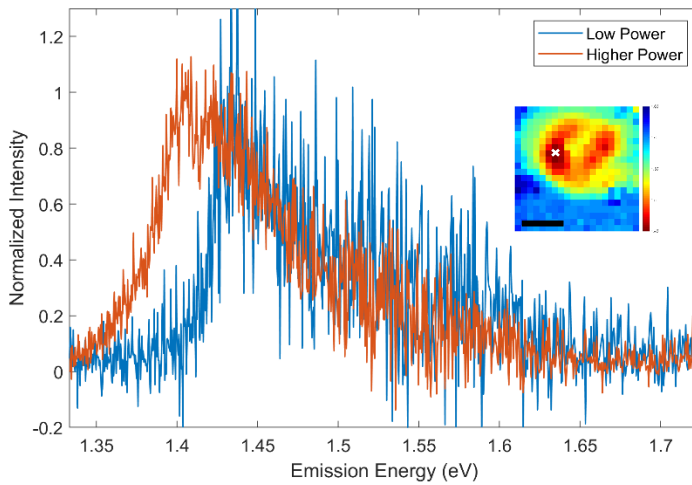

Figure 6-7: Comparison of point spectra of the two power maps in Figures 6-5 and 6-6, location indicatied by the white "x" in spectral median inset, scale bar 30 nm.

633 versus 785 nm excitation of WSe$_2$ nanobubbles in gap-mode

In our earlier study of WSe$_2$ nanobubble we showed how below band gap using a 785 nm laser can enhance the contrast of the LX states, with the added benefit of a stronger plasmonic response from the a-SNOM near-field probe[44]. We can apply the same technique to WSe$_2$ nanobubble on Au, which with the benefit the much higher electromagnetic enhancement in gap-mode allows simultaneous measurement of the Raman modes. In this way, gap-mode excitation with the 785 nm light on WSe$_2$ is an analogue of the experiment we reported on WS$_2$ (Section 4).

Figure 6-8 shows a map of the same nanobubble in Figs. 6-2 – 6-7, excited with ~50 uW of 785 nm laser light. When total spectral window (about 790 nm – 1000 nm) is mapped as in Fig. 6-8a the integrated intensity map appears much the same as was seen in the maps using HeNe excitation, which is to be expect since the majority of the exciton emission is seen to be at wavelengths > 800nm (energies < 1.55 eV). When the spectral window is adjusted to only integrated the LX emission, we again see the

intensity map shifts to edges of the nanobubble, however now the emission is preferentially localized to the two lobes. In addition, some LX is observed outside the apparent nanobubble optical boundary in the bottom right. Fig 6-8c plot point-spectra from each of these points and multiple LX distinct LX emitters are seen, with the lobes peaks around 860 nm (1.44 eV), and outside point (blue spectrum) showing two LX peaks at ~ 840 and 880 nm (1.48 and 1.41 eV). These are not observable in the earlier map with 633 nm excitation.
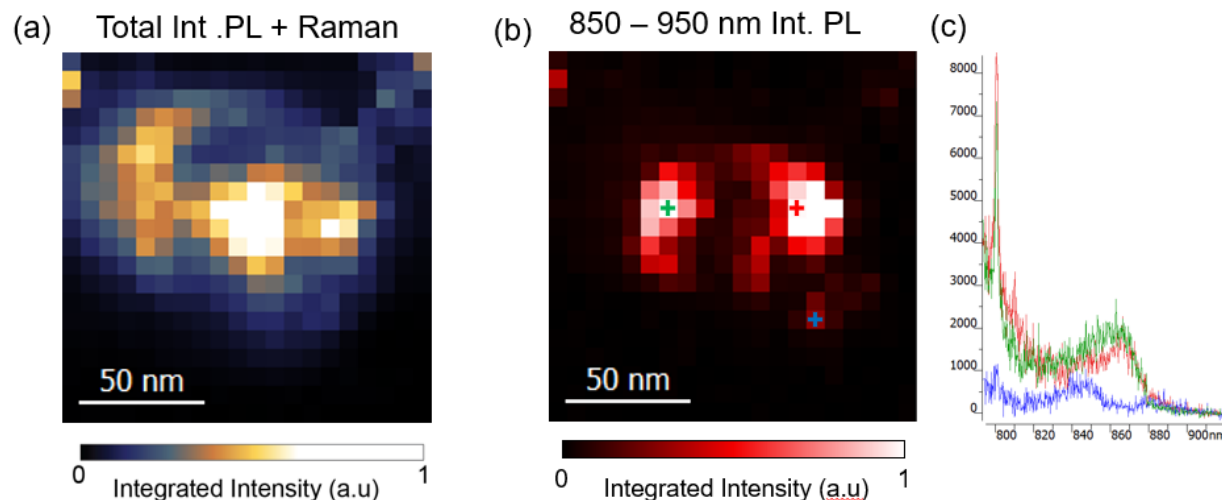

Figure 6-8: TEPL of nanobubble with 785 nm. (a) the total integrated TEPL intensity, (b) integrated PL of the LX between 850-950 nm, and (c) point spectra showing the LX peaks between 830 nm and 890 nm.

LX emission mapping in multiple WSe$_2$/Au nanobubbles

The data presented in Fig 6-8 shows that the same strategies of probing the LX emitters for WSe$_2$ nanobubbles on dielectric substrates can yield similar insight for nanobubble on Au, and the small mode volumes possible in gap mode allow measurements at scales difficult to access otherwise. In addition, the plethora nanobubble present in same flake of WSe$_2$ will allow us to perform high resolution scans on multiple nanobubbles.

Figure 6-9 shows a collage of the TEPL hyperspectral map integrated between 850 – 950 nm for four similarly sized nanobubbles in the region. Strikingly the lobes of each of the four nanobubbles are also aligned the horizontal axis. Example spectra from the first nanobubble of the collage is shown in Figure 6-10 which displays similar spectra to the those in Figure 6-8. The commonality of these axes suggests strongly that the strain is distributed approximately in the same way for each, which itself implies a common morphology of these nanobubbles in WSe$_2$ with diameters ~ 50 nm. This is consistent with the universal strain profile of relaxed nanobubbles in references[62,63]. The fact the LX emission is further localized to particular axis may also indicate the presence of wrinkles, which localize to specific points partly determined by the crystal orientation[69,72].
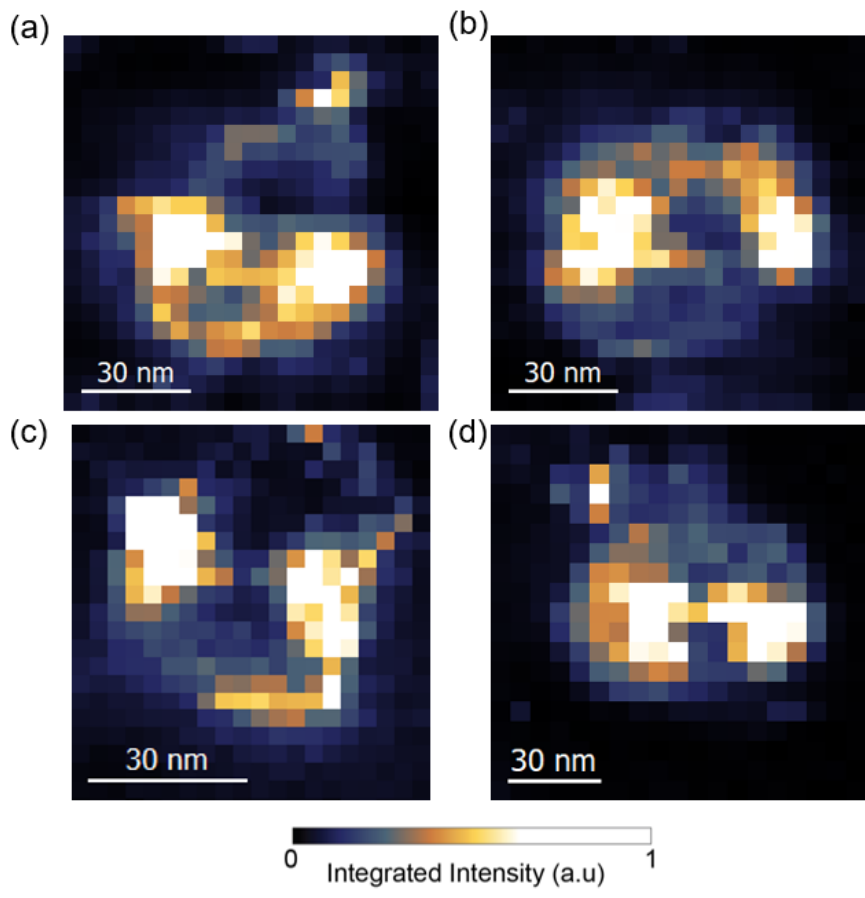
Figure 6-9: Collage of integrated TEPL between 850 – 950 nm showing the LX emission profiles in multiple additional nanobubbles in the vicinity as the nanobubble shown in Figures 6-2 – 6-8.
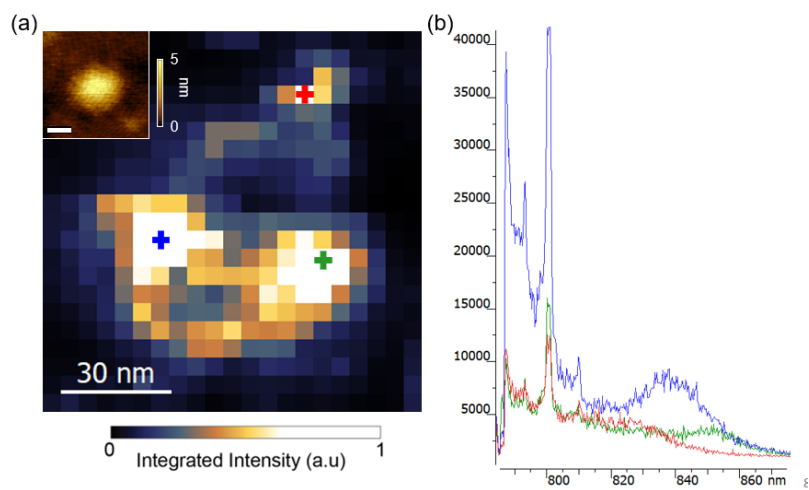

Figure 6-10: Point spectra from nanobubble in Figure 6-9(a).

Gap-mode TEPL of asymmetric nanobubble in WSe$_2$

For nanobubbles of different symmetry we naturally expect to see deviations from a uniform strain profile, in addition to differences in the formation of any wrinkles. The high resolution of gap-mode likewise allows us to investigate these deviations. An example of such is shown in Figure 6-11. The total integrated intensity of the TERS and TEPL as well and the recorded atomic force micrograph are shown as insets of Figure 6-11b. The consist of a pair of nanobubbles in close proximity, a larger one with a highly elliptical shape and max height of ~ 6 nm, while the smaller is circular with a diameter of ~ 50 nm and a maximum height of ~2 nm. Note that the separation of the nanobubbles is less than respective radii of either of the pair, which suggest that their strain fields likely overlap[62].

Figure 6-11a shows the integrated TEPL from 810 – 850 nm, in contrast to the single smaller nanoubbbles shown in Figs. 6-2 – 6-9, the emission in this range is heavily concentrated on the peripheries of long ais of the larger elliptical nanobubble. Interestingly this asymmetry also is true is also observed in the smaller nanobubble, possibly reflecting influence of the overlapping strain fields, or an underlying asymmetry that is not capture in the AFM topography. Figure 6-11(b) shows the TEPL of the 850 – 950 nm spectral bin, the same choice as the collage of the nanobubble discussed above. The emission the larger elliptical nanobubble is similar pattern as the map the (a), emission is predominately weighted toward the periphery, but like the single nanobubble in Fig. 6-2, the lower energy emission is less extended than the higher energy map. Looking at several point spectra in Fig. 6-11d several distinct LX spectra are observed, at similar energies to those reported for the smaller nanobubbles.
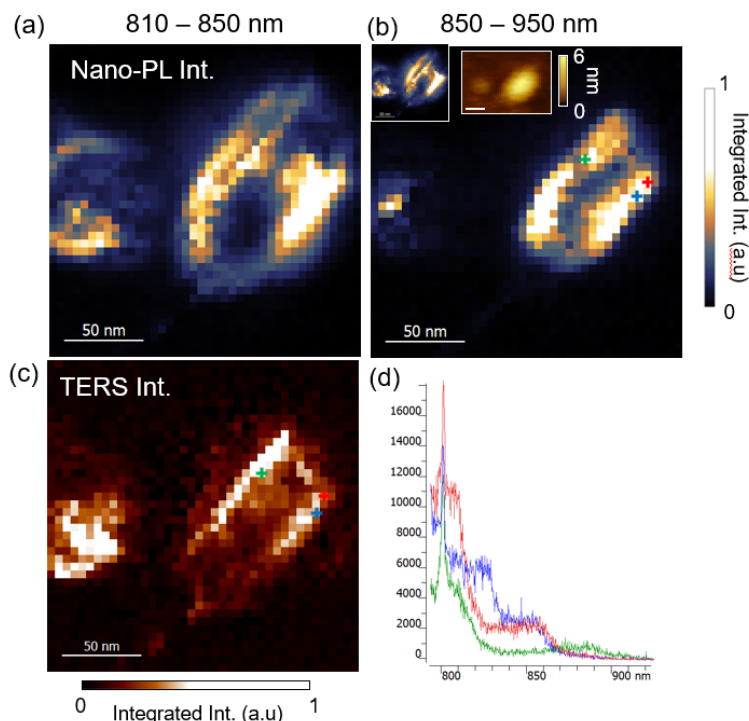

Figure 6-11: TEPL with 785 nm excitation of the an asymmetric nanobubble of WSe$_2$ on Au. (a) integrated TEPL between 810 – 850 nm, and (b) between TEPL 850 – 950 nm, inset shows the total integrated TERS/TEPL and semi-contact AFM topography. (c) shows the TERS intensity between 240 – 260 cm^{-1}. (d) shows point spectra colored crosses in (b) and (c). The scale bars in the insets are 50 nm.

Because the Raman is also observed it allows to compare the strain effects on the lattice vibrations. While the spectral resolution of this TERS/TEPL data is coarse, making observations of the mode shifts difficult, the intensity of TERS emission relate to the underlying strain through changes in the resonance condition. This approach follows the analysis of the TERS intensity from WS_2 nanobubbles (section 4). Fig. 6-11(c) in the integrated intensity of the peak at 800 nm, which corresponds to the complex of first and second order Raman modes between $240 – 260$ cm^{-1}[1114]. For the elliptical bubble the TERS signal is stronger along the edges, with the strongest Raman features appear along the longer edges.

While a comparison of Fig 6-11c and 6-11d appear to imply an association between the strong LX emission and TERS integrated intensity. However, the correspondence is clearly not one-to-one as strong LX emission is also seen areas where the TERS is relatively diminished. To get a better sense of how the TERS intensity compares to the LX emission, we compute the spectral median, again the excluding the Raman modes around 250 cm^{-1}. Note that because the spectra are artificially cutoff by the long-pass filter, the spectral median calculated in this can only be used as a relative measure of the emission energy center. The calculated spectral median is shown in Figure 6-12. As can be seen, the greatest shift of the spectral median is located toward the bottom left edge of the ellipse, which shows weaker TERS.
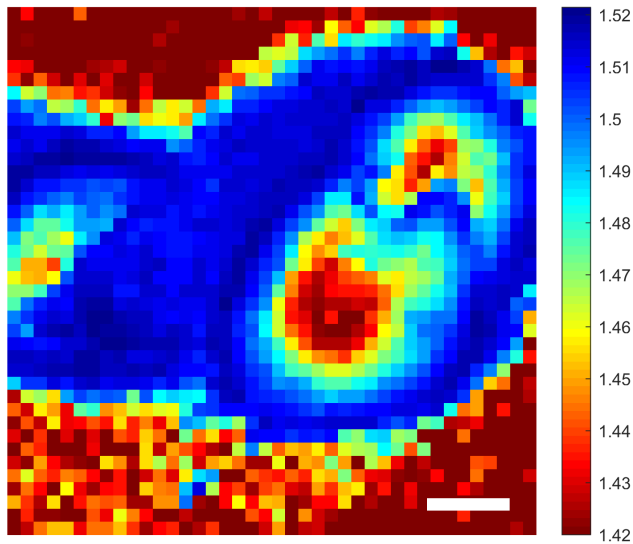

Figure 6-12: Spectral median for the TEPL data shown in Figure 6-11. Scale bar = 30 nm.

The spectral median Figure 6-12 additional shows a complex structure in the energy-spatial LX emission, and clearly points to complex environment of asymmetric nanobubbles at scales on order 1 nm or less. Despite the high resolution of this map (5 x 5 nm^2 pixels), it does appear to be under samples. Pushing sampling to the single nanometer, or less, level however is difficult to achieve in SPM measurements at ambient conditions. Ultra-high vacuum and low temperature TEPL is likely necessarily to full capture the dynamics of these maps.

LX emission forming nanoscale wrinkle-like rings

The TEPL maps in Figures 6-10 – 6-12 host extremely heterogenous LX emission reflecting a complex nanoscale strain. For larger nanobubbles that have clear asymmetry we may expect that some would have geometry such that wrinkling would be greater to large enough features that would be resolvable in nano-PL. In Figure 6-13 we show an example of LX emission localized around a wrinkle-like ring. This feature is already visible in the integrated intensity in 810 – 850 spectral window, shown in (a), but becomes prominent in the LX emission range between 850 – 950 nm (b).
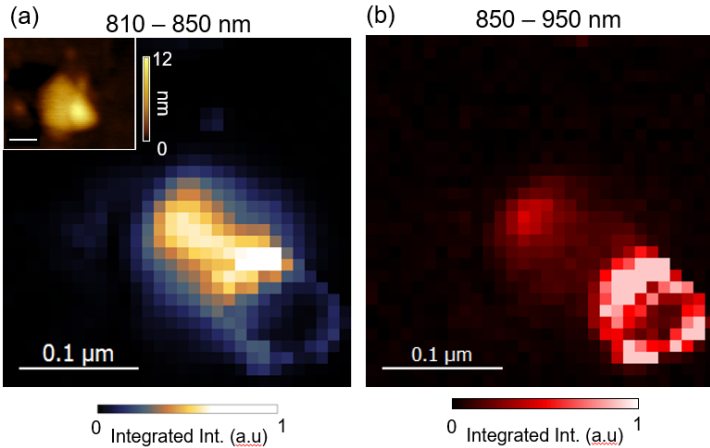

Figure 6-13: TEPL of an asymmetric nanobubble showing a wrinkle like feature that emits strongly in low energy. (a) integrated TEPL between 810 – 850 nm, inset AFM topography. (b) integrate TEPL between 850 – 950 nm.

Figure 6-14 shows a high resolution scan of the high resolution (3.75 x 3.75 nm^2). Despite some visible instability, the wrinkle optical thickness can be seen to be < 30 nm, with a minimum thickness of ~ 10 nm. Example spectra are shown in in Figure 6-14b. Similar to the LX observed in previous nanobubbles this is prominent emission 870 nm (1.43 eV), seen in the red spectrum. Surprising we find measurable emission past 950 nm (1.30 eV). Compared to the free exciton energy ~1.65 eV which would correspond to potential well of ~300 meV, which is certainly extreme, but within the depths of the potential wells predicted theoretically for wrinkles in TMDs[69].
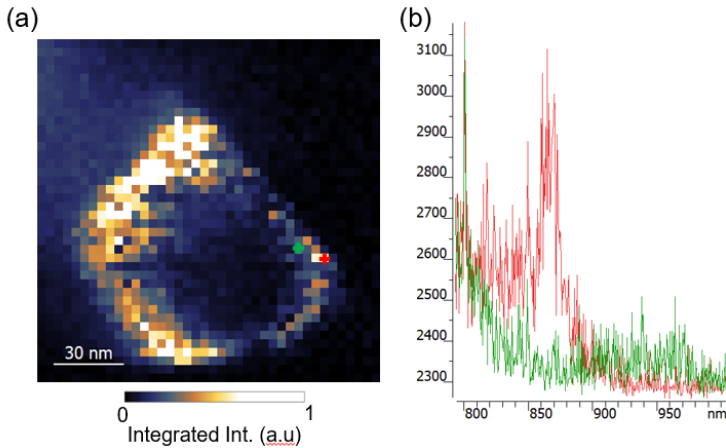
Figure 6-15: high resolution TEPL data from the wrinkle-like "ring" shown in Figure 6-13. (a) integrated PL between 850 – 950 nm, and (b) point spectra from the colored crosses in (a).

Section Conclusion

In this section we present in depth comparison of WSe$_2$ nanobubbles on template stripped Au, using both above and below bandgap excitation. Taking advantage of the <10 nm in gap-mode, we reproduced the observations of LX emission at the periphery. Further we observed a common spatial pattern of the LX emission in small nanobubbles ~50 nm in diameter, implying a common nanoscale strain profile. Further in large nanobubbles we observed wrinkle-like features in the nano-PL, with extremely redshifted PL peaks, with some emission shifted more 300 meV below the free exciton in WSe$_2$. In the following section we will show such features are also observed and are enhanced in MoSe$_2$/WSe$_2$ heterostructures.

Section 7 – Interlayer Localized Exciton States in MoSe$_2$/WSe$_2$/Au

The previous sections detailing the TEPL response in WSe$_2$ and WS$_2$ on dielectric and metal substrates showed the strong localization of excitons due to the potential well provided by the inhomogeneous strain profile of the nanobubble. We can also apply these same experimental techniques to interlayer excitons in "hetero-bubbles."

As described in the earlier section on excitons in TMDs, the interlayer exciton (ILE) is hybrid exciton between 2 non-equivalent layers, where the bound electron and hole reside in different layers. For instance, a commonly studies interlayer sample, in vertical heterostructure of MoSe$_2$ and WSe$_2$, diagrammed in Fig. Figure 7-1 (inset) below. With the advances in producing high quality monolayers of all 4 of the commonly excitonic 2D semiconductors, such vertical structures allow high quality interfaces between the materials. In turn this has led to a surge in recent studies of theses states[19,21-24,115], with particular interplay of these states and underlying effect of the Moire potential. Given the novelty of these excitations, the effect of strain on the properties of the ILE have not yet been thoroughly investigated, in particular strain due to structural imperfections such as nanobubbles.

In section we will show that nanobubbles formed in a heterostructures of MoSe$_2$/WSe$_2$/Au can indeed strongly localized ILE states. Using high resolution gap-mode TEPL and TERS we are able to optically map "wrinkle" structures on multiple nanobubbles that host strong interlayer localized exciton (IL-LX) emitters. The width of some wrinkles (or rings) is resolved to be as small as 7 nm, with some spectrally distinct emitters separated by distances <2.25 nm. In the following we will outline the imaging of several large and small hetero-bubbles.

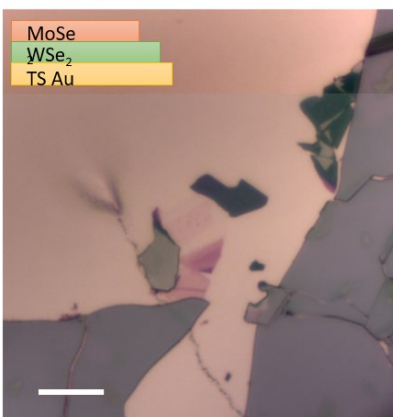**Overall sample characterization and far-field PL mapping of hetero-layer**

Figure 7-1: Optical micrograph MoSe2/WSe2/Au heterostructure. Scale bar = 10 μm.

A cartoon diagram of the sample studies is shown in Figure 7-1 (inset). The gold substrate was prepared by template stripping a 40 nm layer of evaporated goal on standard commercial Si wafers with a 300 nm SiO$_2$ layer. The Au later was stripped from the growth chip using epoxy. Monolayers of WSe$_2$ and MoSe$_2$ were first mechanical exfoliated onto separate equivalent Si substrates and monolayers were

identified via a calibrated RGB color contrast. The heterostructure was assembled using PCL polymer to first assemble the stack and deposit the heterostructure on the Au substrate.

The nanobubbles were created during the final step of sample preparation where the PCL is removed via submersion in a hot-acetone bath. While the bath successfully removed the PCL polymer, the acetone also attacked the epoxy beneath the gold film. The effect of this was to cause in-plane stresses in the Au, inducing widespread buckling and some cracking. An optical image of the finished sample is shown in Figure 7-1. A long crack moving from the bottom right toward the top left is seen in the image. The large dark colored flakes are bulk WSe_2 and $MoSe_2$ that are commonly also transferred during sample preparation. The actual heterostructure is situated between two smaller bulk flakes in the center. Multilayerd WSe_2 (darker purple-pink in color) can be seen just below the heterostructure which is a faint pink in color-tone. The fainter pink to the top-left of the heterostructure is monolayer $MoSe_2$.
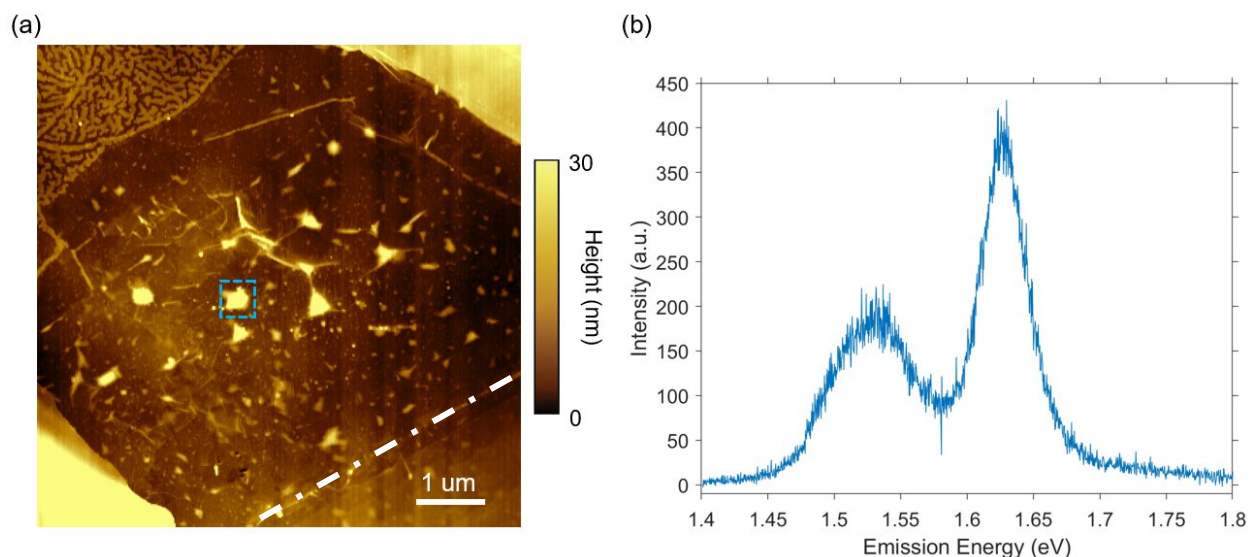

Figure 7-2: (a) tapping mode AFM micrograph heterostructure. (b) photoluminescence spectrum of nanobubble in dashed box.

While the crack does clearly pass beneath some of the bulk TMD, it did not strike the heterostructure or the monolayers, and the Au gold beneath the TMDs appears to be continuous and relatively smooth. The contrast in the TMDs mono- and heterolayer is mostly continuous, aside from a few larger bubbles. Fig 7-2a shows a tapping AFM micrography of the heterostructure region. Dashed lines demark the edges of the heterostructure from the monolayer WSe_2 (bottom right) and $MoSe_2$ (top left). And a dashed-dotted line denotes the edge of the bare Au. Despite appearing smooth optically, the AFM shows the Au topography to be extremely buckled, with a surface roughness of 5-10 nm, similar to that of evaporated Au films. In addition to the larger nanobubbles visible in Fig 7-2a (~40 nm in height), the heterostructure and monolayers display numerous nanobubbles of various sizes throughout as well as several extended wrinkles and folds. While the later also can have an extreme and heterogeneous response as was discussed previous for nanobubbles in WS_2, the remainder of this section will be devoted to characterization of the heteronanobubbles.

Though there are clearly nanobubbles in the heterostructure, the AFM micrograph cannot distinguish between bubbles where both materials are delaminated from the substrate and alternatively where just the top $MoSe_2$ layer is separated from the Au. In order to verify the "double" character of the bubbles, as well as verify the monolayer-heterolayer demarcations in Fig 7-2b, we performed standard

78

far-field confocal PL mapping of the hetero- monolayer nanobubbles. If the nanobubbles are solely in top layer, then the PL should show only a single exciton of $MoSe_2$, shifted in energy and broadened due to strain inhomogeneity of the nanobubbles, with the luminescence from the bottom WSe2 quenched by metal substrate. By contrast a hetero-bubble, both layers delaminated from the Au, and the PL spectrum will contain emission from the both of the intralayer excitons. Fig. 7-2b shows the PL spectrum corresponding the average of the blue box in the AFM micrograph. Two strong peaks are seen, the first centered ~ 1.63 eV and the second ~ 1.54, slightly below the normal free exciton energy in WSe_2 and $MoSe_2$ respectively.

The spectra in Figure 7-2b provides definite evidence of the double-layer character of the nanobubbles. However, the relative strengths of the $MoSe_2$ or WSe_2 excitons show great heterogeneity depending on the nanobubbles studied. This can be seen more readily in Figure 7-3, which shows average spectra of 2x2 pixel (400 x 400 nm^2) of three select areas given by the colored boxes in the intensity map. The red-orange and blue boxed regions both so dual peaks, but very disparate peak positions, with low energy peak of the blue centered at ~1.54 eV while the red-orange is centered at ~ 1.47 eV. Emission from the green box was approximately 10x more intense. The center of this emission is around 1.62 eV suggestion that it primarily from intralayer WSe_2 exciton.
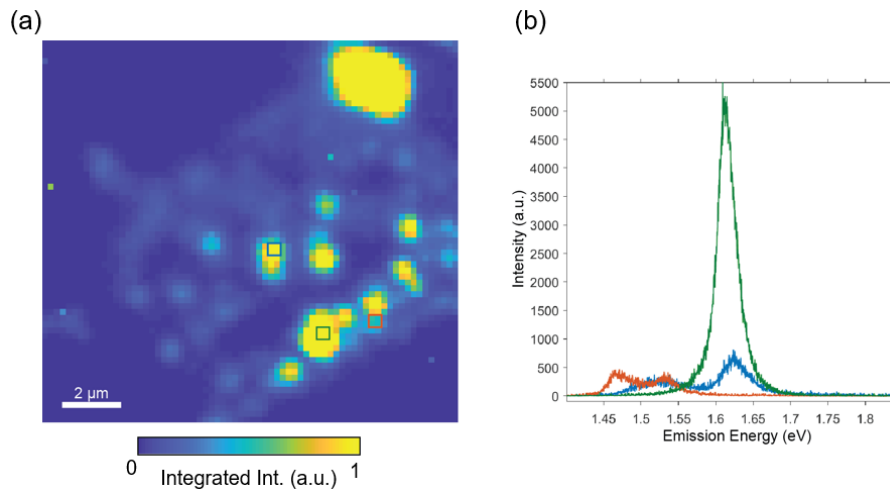

Figure 7-3: Comparison of point spectra (b) in corresponding to the color-coded boxes in integrated PL intensity shown in (a).

To better capture this emission energy and intensity variation Figure 7-4 shows integrated photoluminescence maps in three selected energy ranges: 1.45 – 1.6 eV, 1.6 – 1.75 eV, and 1.70 – 1.80 eV. The first two bins correspond roughly to the $MoSe_2$ and WSe_2 excitons emission respectively while the third represents hot-luminescence from higher order exciton states, such as the B-exciton in $MoSe_2$, and higher excited states of the WSe_2 A-exciton[116]. Examination of the first two maps shows strong variation, with some bubbles having strong emission of both intralayer exciton species, with emission from other nanobubbles dominated by a single exciton species. For the last bin the majority of nanobubbles show no detectable emission, which is expected with the predominately tensile strain of each nanobubble. Surprisingly however, significant emission was observed between 1.7 – 1.8 eV in the larger nanobubbles. As far as we are aware such enhanced emission of higher order exciton states due to

inhomogeneous strain has not been described experimentally and may arise from the increased exciton density due to the funneling effect[61] combined with enhanced radiative recombination due to the gold substrate[117].
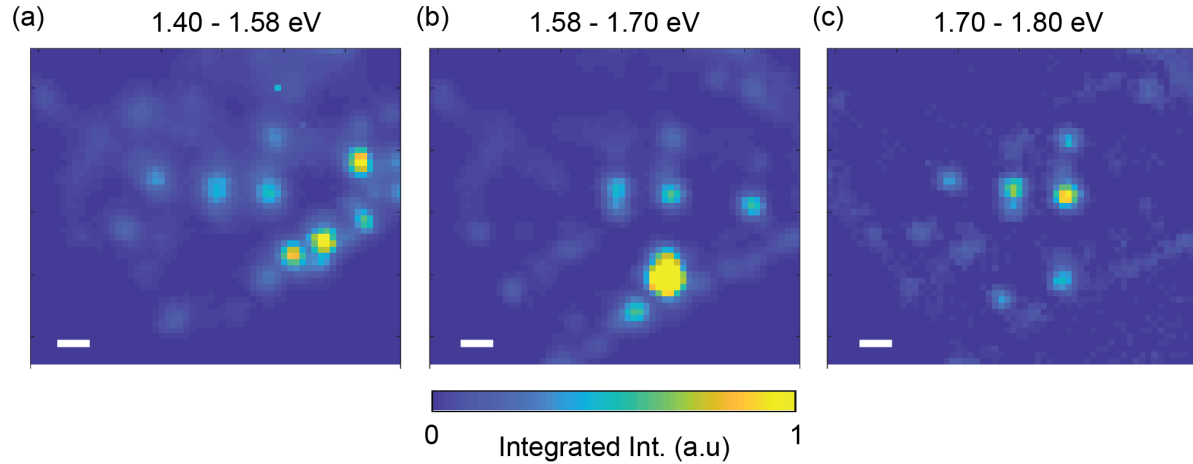

Figure 7-4: integrated PL maps of different spectral windows. (a) 1.4 – 1.58 eV, (b) 1.58 – 1.7 eV, and (c) 1.7 – 1.8 eV) from left. Scale bars are 1 μm.

To summarize this subsection, far-field confocal photoluminescence mapping confirms that nanobubbles shown in the heterolayer region in Fig 7-2a are double-bubbles of both stack layers. Further the nanobubbles show strong and very heterogeneous emission of the two intralayer exciton specifies, implying a complex strain environment at the nanoscale. In the following subsections we will examine the luminescent structure of these nanobubbles at the few nanometer length scale.

Gap-mode TEPL scans of localized excitons

The region of interest for the nano-PL, scans to follow is indicated by the red arrow in the inset of Fig 7-5a. This region is roughly that of green box of Fig 7-3, showing the strong emission seemingly comprised mostly of strained intralayer WSe$_2$ exciton. We find that this is the case at nanoscale. Fig 7-5a shows the largest nanobubble in the chosen ROI, which has a "witch's hat" topographic profile in the AFM micrograph. Fig. 7-5b shows the corresponding nano-PL integrated map, exciting the sample with 633 nm (1.96 eV) HeNe laser, well above the bandgaps for both TMDs. The total emission intensity is mostly constant over the structure of the nanobubble, with two edge regions showing slightly bright emission. Figure 7-5c shows the corresponding average spectra of the colored boxes in Fig. 7-5b. With some variation the dominant emission is at 770-790 nm, slightly redshifted but near the far-field emission shown in Figure 7-3, with the brightest emission in the bottom right corner at ~ 775 nm. However, some localized emission ~820 nm is observed at the top-left edge in the blue spectrum. Integrated the hyperspectral maps in wavelength ranges 750 – 800 nm and 790 – 850 nm, denoted by the dashed boxes in the spectra, confirm this observation showing the ~820 nm emission seems to be localized to the top-left edge region.
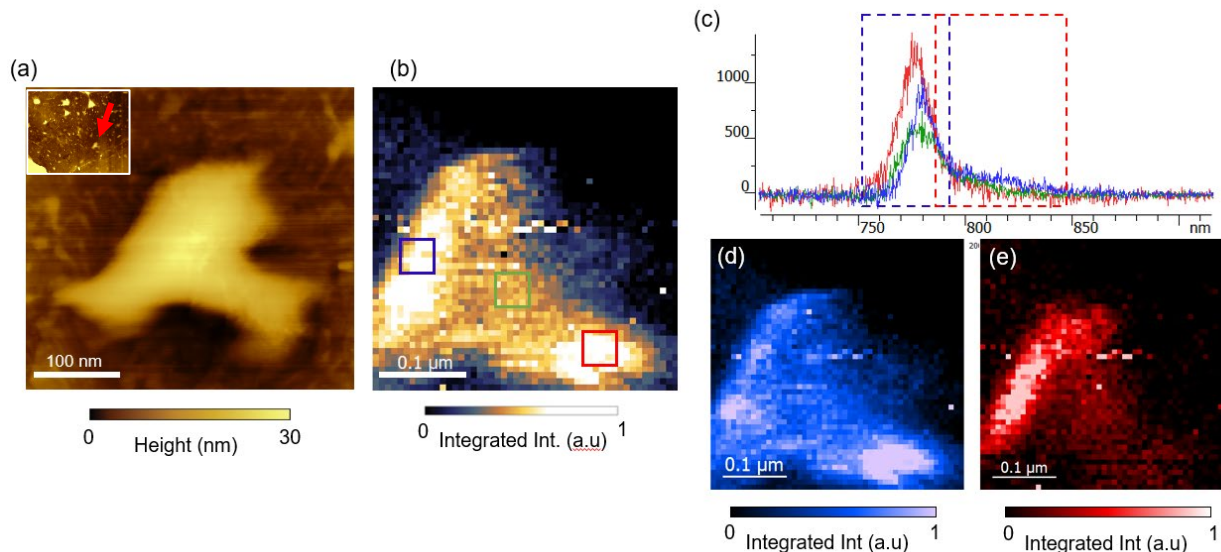
Figure 7-5: TEPL of large MoSe$_2$/WSe$_2$ nanobubble. (a) topography, (b) total integrated TEPL, (c) box-averaged spectra displayed in (b). (d) and (e) are integrate sub-spectral ranges corresponding to the dashed boxes in (c), ranges 750 – 790 nm and 785 – 850 nm respectively.

While the nano-PL scans in Figure 7-5 show some LX emission, the relatively poor signal-to-noise combined with overwhelming light emission from the delocalized excitons may be swamping additional LX states. As we showed for WSe$_2$ nanobubbles on Au and hBN, greater contrast of LX states can be experimentally obtained by using below gap excitation of WSe$_2$, using a 785 nm (1.57 eV) laser, below the strong emission observed at 1.62 eV (766 nm) in Fig 7-3. Figure 7-6 shows the same nanobubble using this excitation with same near-field probe. As can be seen, additional LX emission emerges at the bottom and right edges of the nanobubble. Representative point spectra are shown in Figure 7-6b, of the top and left LX regions, blue and red, as well as the bright in green. The center of the green spectra is partially cutoff by the filter, suggesting the peak is near the laser line at 785 nm. By contrast the localized emission of the blue and red is significantly redshifted, with peaks in range 820-830 nm.
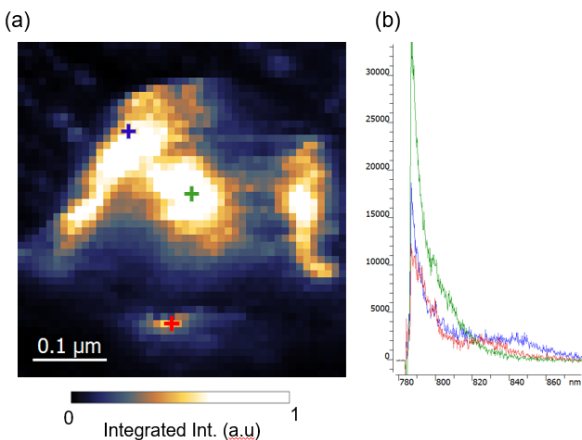

Figure 7-6: TEPL of the nanobubble in Figure 7-5 excited 785 nm. (a) total integrated PL intensity and (b) point spectra of colored crosses in (a).

In total the nano-PL emission characteristics of this nanobubble are very similar to measurements of the pure WSe_2 nanobubbles on hBN and Au substrates, judging by the peak positions of the LX states and the strong emission from the more weakly localized emitters at the center. This is consistent with the far-field measurements of the previous subsection. This behavior extends to neighboring nanobubbles, and is not a result of the unique nanobubble morphology. Figure 7-7a shows the integrate LX luminescence for a nanobubble ~55 nm in radius. The LX emissions is clearly seen localized to three locations around the periphery, with emission peaks in the 810 – 850 nm range. By comparison with Figure 6-9 of section 6 one observes a similar spatial profile for the LX emission intensity, however the redshift are somewhat less than solo WSe_2 nanobubbles on Au.
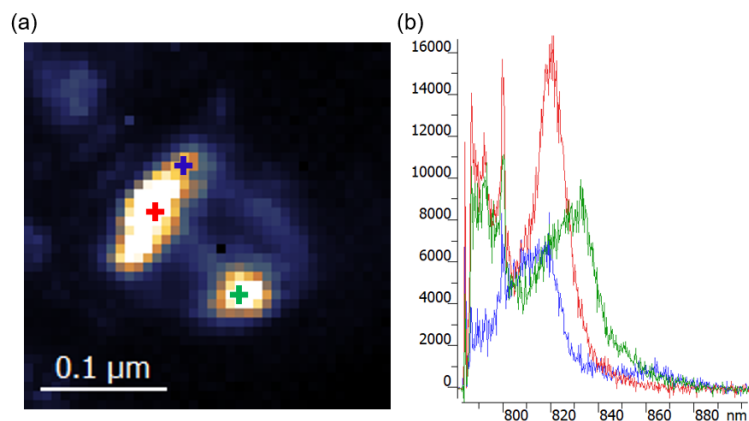
Figure 7-7: TEPL of smaller nanobubble than in Figure 7-6, (a) integrated between 810 – 950 nm, and (b) point spectra given by colored crosses.

Given that these are nanobubbles of heterolayer of $MoSe_2/WSe_2$, it is curious that the emission has such similarities to nanobubbles in WSe_2 alone. The emission properties of these nanobubbles is also strikingly different from other hetero-nanobubbles in the vicinity, as will be shown in the following section. This contrast can reasonably be attributed to vagaries in the nanobubble formation, a more concrete explanation requires further investigation.

Interlayer localized exciton emission from nanoscopic wrinkles

Figure 7-8: overlay of integrated nano-PL of neighboring nanobubbles of MoSe$_2$/WSe$_2$.
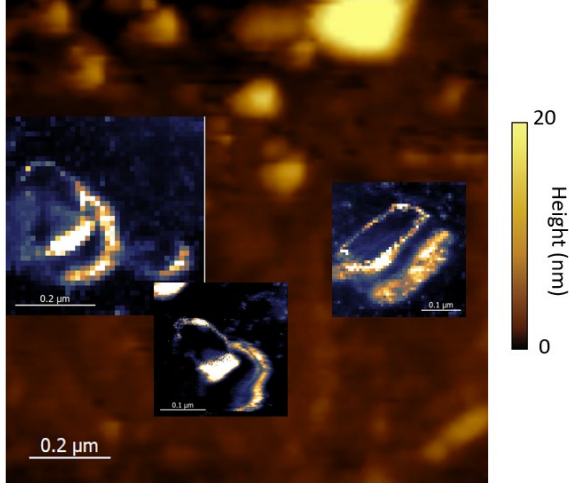
Figure 7-8 shows integrated TEPL scans for several different nanobubbles overlaid on the AFM micrograph. Like the previous nano-PL data, there is significant LX exciton emission located around the periphery of the nanobubble. In addition, however, strong emission seen from the ring- and wrinkle-like features which extends across the body to the periphery of each nanobubble. These unexpected features of the nanobubbles host rich complex of LX states. Figure 7-9 shows an example of such rich behavior from two neighboring pixels (area 7x7 nm^2) on one such wrinkle like features. In blue spectrum four clearly defined LX peaks are observed, the centered at approximately 820, 865, 892, and 915 nm, with the longer two wavelength peaks at greater emission intensity. The red spectrum contains these, plus an additional peak centered at 935 nm, at yet greater emission intensity.

The position of the redder peaks ~ 890 – 940 nm are in the energy range of the interlayer exciton, which has been reported to be between 885 – 920 nm (1.35 – 1.40 eV) for MoSe$_2$/WSe$_2$ heterostructures at room temperature[115]. This close correspondence suggests the observed peaks arise from the localizations of the interlayer exciton. For this reason we will refer to excitonic peaks in this energy range as the interlayer localized exciton, or IL-LX. However, we note that given the possible extreme strain that possible according to the wrinkling model proposed in ref[69][69], as was discussed for LX emission in the pure WSe$_2$ nanobubbles of the previous section. However, the emission intensities of Figure 7-9 compared to Figure 6-15 is much more intense, suggesting these states are of different character. Further analysis of the differences of LX emission in monolayer and heterolayers will be discussed below.
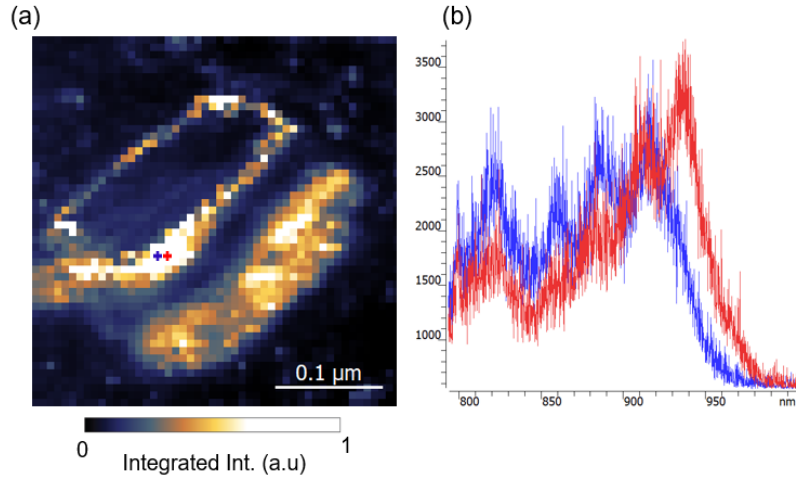
Figure 7-9: point spectra (b) of the integrated nano-PL map in the top right of 7-8, indicated by the colored crosses in (a).
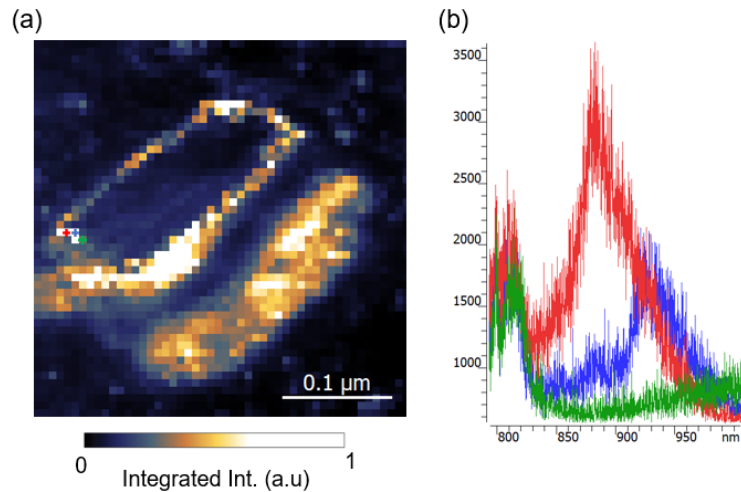

Figure 7-10: same map as 7-9 with additional point spectra in the left edge of the nanobubble.

As with the far-field observations the morphology and relative strengths of these emission peaks shows great variation over the spatial extent of these nanobubbles. For instance, Figure 7-10 shows three neighboring spectra on the left edge of the nanobubble in Fig. 7-9, which have entirely different character from the previous two spectra. The green spectra in Figure 7-10 shows far red shifted emission with a broad peak that is centered near 1000 nm. To our knowledge such red-shifted localized exciton emission has not been described in $MoSe_2/WSe_2$ heterostructures at room temperature. Similar emission characteristics are observed along wrinkle structures in the neighboring heterobubbles. Figures 7-11 and 7-12 show such spectra from two other nanobubbles in Figure 7-8. Notably while the spectra are similar, the IL-LX emission is substantially different, no two pixels so the peak ratios which speaks to the likely complex nanoscopic variations of the confinement potential.

While the composition of IL-LX emission energies is very heterogeneous, there is a common trend between the three nanobubbles shown in Figures 7-8 – 7-12. Namely the diameter of the ring varies depending on the emitter energy, with redder IL-LX states predominately located toward the interior of the ring. To show this we integrated the spectra in three separate spectral ranges of the IL-LX emission, in

a similar way to the far-field hyperspectral maps of Fig. 7-4, and overlaid the resulting maps using different colormaps. The spectral bins were in rising order of the wavelength: 860 – 910, 900 – 950, and 940 – 1000 nm, with the corresponding colormaps blue, green, and red. These bins are also given by the dashed boxes over the emission shown in Figures 7-13b – 7-15b. For each hetero-bubble the lower energy IL-LX emission is toward the center of the ring, with bluer emission toward the periphery. Note that Figures 7-13 and 7-15 is the same nanobubble as in Figure 78-8, scanned at higher pixel density.
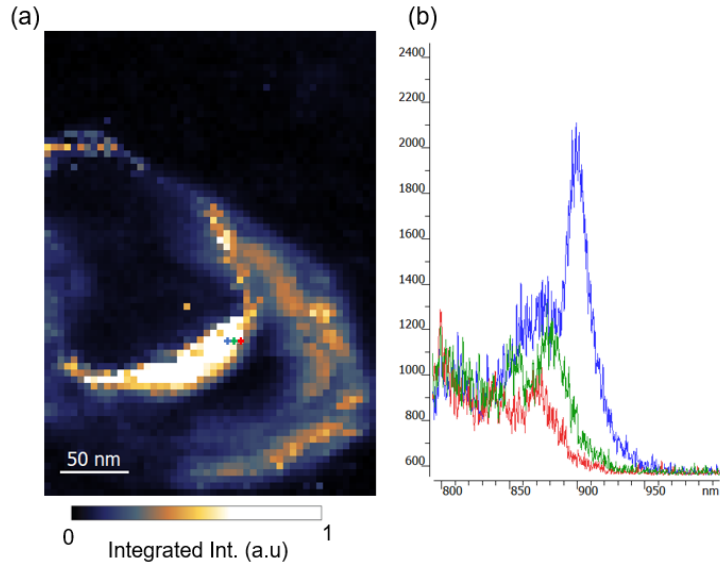
Figure 7-11: neighboring pixel point spectra of the left nanobubble in Figure 7-8.
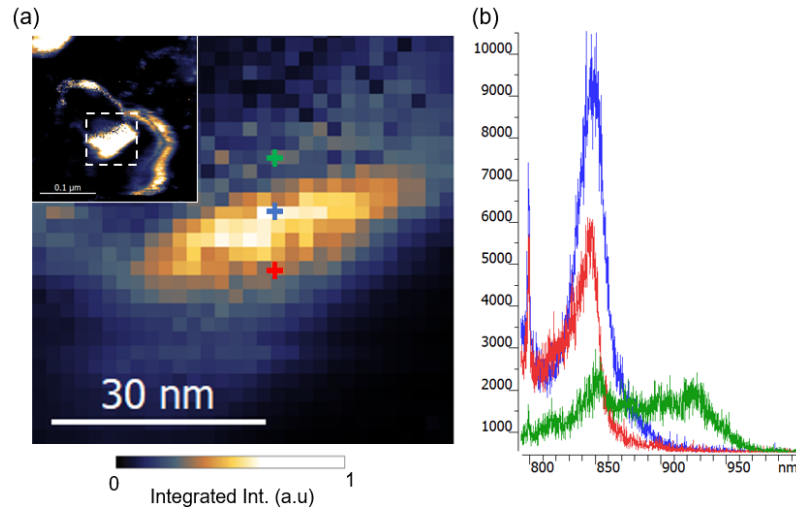
Figure 7-12: neighboring pixel point spectra of the bottom nanobubble in Figure 7-8. (a) zoom in of dashed white box in the inset. (b) point spectra of the colored crosses in (a).

To show the extreme localization of the emission, to the right of each map, three spectra are shown taken from directly neighboring pixels. The location of each spectra is shown by the same colored crosses in the inset of the 3-color integrated TEPL map. As can be seen, each neighboring spectra is distinct, and the total width of the ring (or wrinkle) appears to be < 7 nm by the observed contrast.
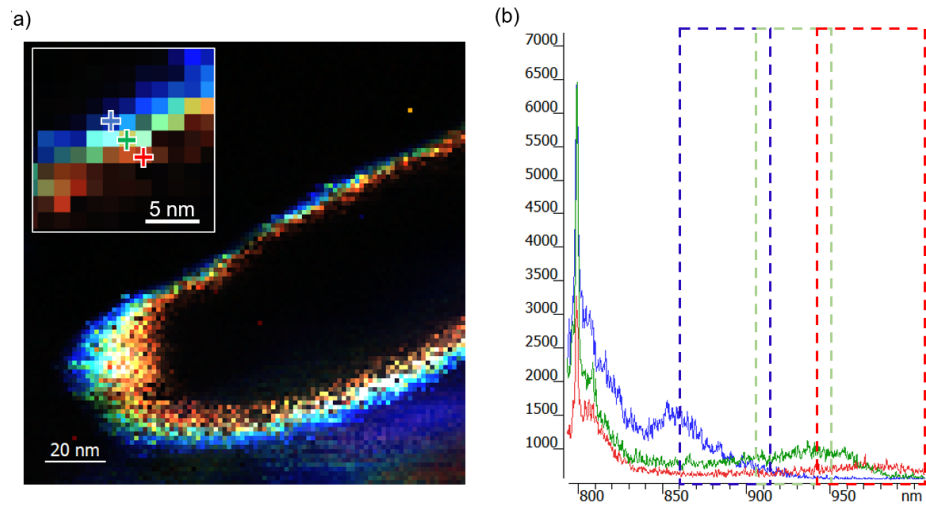
Figure 7-13: (a) Overlay of 3 integrated TEPL maps color-coded to the dashed boxes in the point spectra shown in (b). Spectra are from the colored crosses in the inset of (a).

Figure 7-14: (a) same overlaid TEPL plot as 7-13 for the bottom nanobubble shown in Figure 7-12.

Figure 7-15: (a) same overlaid 3-color integrated TEPL map of the dashed boxes in (b). Nanobubble is higher resolution of the left most nanobubble shown in Figure 7-8.

It is worth commenting that 1000 nm is near the cutoff for Si based detectors, which rapidly lose efficiency as the photon approaches bandgap near ~1100 nm. We may imagine the emission pattern observed in Figures 7-13 – 7-15 continues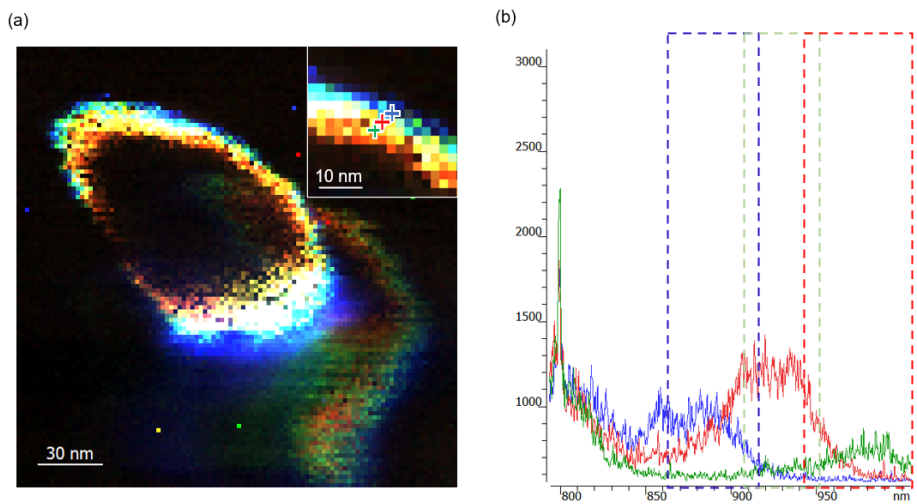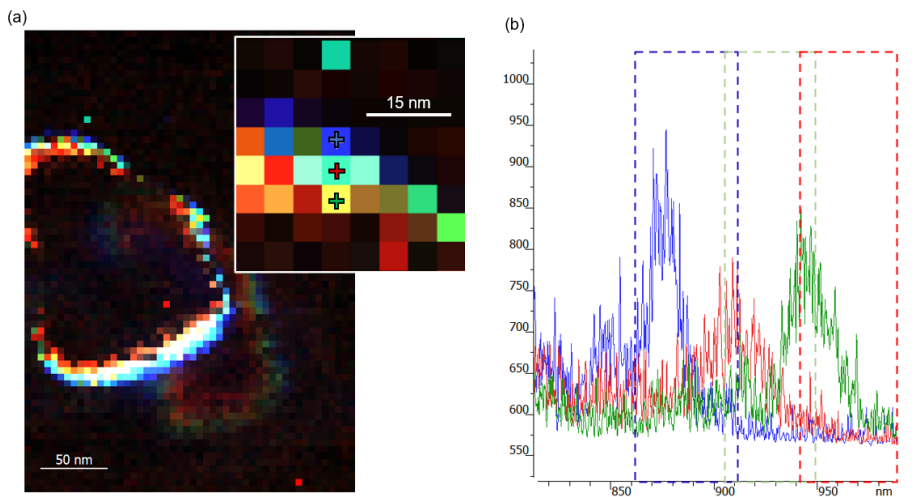, with the shrinking with longer emission wavelength, till the radius of ring shrinks to zero. Such a possibility is fascinating, and invites further work with NIR a-SNOM methods.

Discussion of interlayer nano-PL of MoSe₂/WSe₂/Au versus the monolayers alone

It mentioned earlier the primary evidence for the interlayer character of the observed emission in the hetero-nanobubbles is the observed emission energies of the peaks, grouped around 920 nm, with some emission extending even passed 1000 um, well in the near-infrared. Another factor in favor of this assignment is the comparison with previous sections of 1L-WSe₂/Au nanobubbles. The TEPL maps of the LX states in pure WSe₂ shown in Figure 6-8 show have emission in the range 790 – 900 nm (1.37 – 1.56 eV). This does overlap with the spectra shown in Figure 7-9 – 7-11 but the strongest LX emission observed in 1L-WSe₂ was concentrated in the 800 – 850 nm, which is sharp contrast to some of spectra observed for example in Figure 7-9.

While the emission energies are supported of the interlayer exciton assignment, given the deep potential wells of the electron and hole wavefunctions, predicted by theory to be as much as 500 meV below the monolayer quasiparticle bandgap[69]. The general heterogeneity of the exciton of the nanobubbles also strongly demonstrates the importance of the local strain on the emission characteristics. The lack of wrinkle-like features in Figures 7-5 has a likely origins in deviations of the nanoscopic strain from neighboring nanobubbles shown in the subsequent figures.

We may then hypothesis the strong emission observed are in fact intralayer excitons in the monolayer WSe₂ or MoSe₂. For the former we can compare to the TEPL data in section 6 of nanobubbles in monolayer 1L-WSe₂ on Au. While similar wrinkles are seen in 1L-WSe₂ (Fig 6-15), they did not typically show very strong passed 900 nm, it might be expected that such a wrinkle in 1L-MoSe₂ may have such states in these energy ranges.

Further experiments combined with detailed theoretical modeling will be needed to conclusively identify the layer composition of these excitons.

Section Conclusion:

In this section we have overviewed TEPL emission from double-layer nanobubbles in a MoSe₂/WSe₂ heterostructure. Several of these nanobubbles exhibit strong NIR emission in the energy range corresponding to the interlayer exciton, localized to the nanoscopic wrinkles. These quasi-1D features formed rings over the nanobubble. Comparing these maps to theoretical models from ref[69][69] suggest that the emission originates from the localized intra-layer emission in strain induced potential wells. Further emission peaks were observed over wide range of wavelengths, from ~900 nm to past 1000 nm, with redder emission localized toward the interior of the "ring" of the wrinkle. This data shows the strong tunability of the ILE emission with strain, and shows that the ILE can be strongly enhanced in gap-mode plasmonic cavities.

Our observations of localized ILE in wrinkles as well as more generalized bright weakly localized ILE offers some intriguing possible utilizing strain combine with plasmonic cavities to enhanced control the ILE emission. The fact that all measurements presented in this section further motivates this fact, as the ILE is has been consistently observed to be weak at room temperature[115]. The high degree of spatial localization also suggests a possible mechanism for the strong emission the IL-LX. Since the bound electron and hole of the interlayer exciton are momentum mismatched owing to the indirect band alignment of the $MoSe_2$ and WSe_2 layers[19]. Direct ILE are therefore optically dark, and only ILE with non-zero kinematic momentum, which can be supplied by phonons or the Moire lattice, the latter accounting for the dependence of the ILE transition dipole with twist angle[19]. The highly localized strain of a lattice wrinkle will dramatically break the symmetry and also serve as a scattering center providing the missing momentum between the electron and hole, increasing the light-cone of optically bright ILE emitters. Localized strain appears then to be a mechanism of the controlling the ILE intensity and energy, in a manner very similar to intralayer excitons where strain interactions have been studied extensively.

8. Concluding Remarks:

In this work we have shown how scanning near-field optical microscopy can be applied to the study of excitons in the semiconducting transition metal dichalcogenides. We began by detailing how the diffraction limit arises for standard microscopes and how near-field methods circumvent this limit by including the evanescent fields of the source. We proceeded to demonstrate how the two dominant flavors of SNOM are applied to imaging nano-photoluminescence from 2D transition metal dichalcogenides of monolayer WS_2 and WSe_2.

The main subject of this work was to use nano-PL and nano-Raman, investigated by TEPL and TERS, to study the localization of excitons in TMDs by the inhomogeneous strain produced by naturally forming nanobubbles. In the four sections that detailed this work (sections $4 - 7$) we showed that:

- **The AFM topography combined with classical plate theory can estimate the strain tensor of the monolayer from nanobubbles. With TERS we showed that the method reproduces the Grüneisen parameter from the in-plane E' vibration of monolayer WS_2.**
- **Studying the hyperspectral TEPL of nanobubble in WSe_2 we showed how the greatest localization of excitons, judged by the spectral shift of the emission, occurs at the edges of the bubbles, and that this localization is insensitive to the defect concentrations by observing the same phenomena in flux-grown low defect WSe_2. We found are results are in good agreement with atomistic theory which predicts the greatest localization at nanobubble edges due bending of the monolayer and formation of atomic scale wrinkles.**
- **With few nanometer resolution gap-mode TEPL of WSe_2 nanobubbles we showed the localized exciton states in single crystal domain form preferentially along a common axis, as well localized exciton emission in wrinkle-like ring, with nano-PL emission detectable up to ~ 300 meV below the main exciton in regions ~ 5 nm in extent.**
- **Observing double-layer nanobubbles formed in heterostructures of $MoSe_2/WSe_2$ we observed convincing evidence for localized interlayer excitons similar wrinkle-like ring structures, with unusually intense room temperature emission. The strength of this emission we attribute with expansion of the light-cone due to scattering or symmetry breaking from the high strains in the nanoscale wrinkle.**

These studies are a first-of-their-kind investigation of strain of excitons in the monolayer TMDs and many of the conclusions will undoubted be modified or expanded upon by future works. However, the work presented here clearly shows the strong control of exciton emission provided by non-uniform strain in 2D materials.

References:

1   Levy, N. *et al.* Strain-Induced Pseudo–Magnetic Fields Greater Than 300 Tesla in Graphene Nanobubbles. *Science* **329**, 544, doi:10.1126/science.1191700 (2010).

2   Shepard, G. D. *et al.* Nanobubble induced formation of quantum emitters in monolayer semiconductors. *2D Materials* **4**, doi:10.1088/2053-1583/aa629d (2017).

3   Kumar, S., Kaczmarczyk, A. & Gerardot, B. D. Strain-Induced Spatial and Spectral Isolation of Quantum Emitters in Mono- and Bilayer WSe_2. *Nano Letters* **15**, 7567-7573, doi:10.1021/acs.nanolett.5b03312 (2015).

4   Branny, A., Kumar, S., Proux, R. & Gerardot, B. D. Deterministic strain-induced arrays of quantum emitters in a two-dimensional semiconductor. *Nature Communications* **8**, 15053, doi:10.1038/ncomms15053 (2017).

5   Rosenberger, M. R. *et al.* Quantum Calligraphy: Writing Single-Photon Emitters in a Two-Dimensional Materials Platform. *ACS Nano* **13**, 904-912, doi:10.1021/acsnano.8b08730 (2019).

6   Yao, K. *et al.* Optically Discriminating Carrier-Induced Quasiparticle Band Gap and Exciton Energy Renormalization in Monolayer MoS_2. *Physical Review Letters* **119**, 087401, doi:10.1103/PhysRevLett.119.087401 (2017).

7   Novoselov, K. S. *et al.* Two-dimensional gas of massless Dirac fermions in graphene. *Nature* **438**, 197-200, doi:10.1038/nature04233 (2005).

8   Splendiani, A. *et al.* Emerging Photoluminescence in Monolayer MoS2. *Nano Letters* **10**, 1271-1275, doi:10.1021/nl903868w (2010).

9   Mak, K. F., Lee, C., Hone, J., Shan, J. & Heinz, T. F. Atomically Thin MoS_2: A New Direct-Gap Semiconductor. *Physical Review Letters* **105**, 136805, doi:10.1103/PhysRevLett.105.136805 (2010).

10   Miller, R. C., Kleinman, D. A., Tsang, W. T. & Gossard, A. C. Observation of the excited level of excitons in GaAs quantum wells. *Phys. Rev. B: Condens. Matter Mater. Phys.* **24**, 1134-1136, doi:10.1103/PhysRevB.24.1134 (1981).

11   Hill, H. M. *et al.* Observation of Excitonic Rydberg States in Monolayer MoS2 and WS2 by Photoluminescence Excitation Spectroscopy. *Nano Letters* **15**, 2992-2997, doi:10.1021/nl504868p (2015).

12   Xu, X., Yao, W., Xiao, D. & Heinz, T. F. Spin and pseudospins in layered transition metal dichalcogenides. *Nature Physics* **10**, 343-350, doi:10.1038/nphys2942 (2014).

13   Zhang, X. X., You, Y., Zhao, S. Y. & Heinz, T. F. Experimental Evidence for Dark Excitons in Monolayer WSe_2. *Physical Review Letters* **115**, 257403, doi:10.1103/PhysRevLett.115.257403 (2015).

14   Park, K.-D., Jiang, T., Clark, G., Xu, X. & Raschke, M. B. Radiative control of dark excitons at room temperature by nano-optical antenna-tip Purcell effect. *Nature Nanotechnology*, doi:10.1038/s41565-017-0003-0 (2017).

15   Borys, N. J. *et al.* Anomalous Above-Gap Photoexcitations and Optical Signatures of Localized Charge Puddles in Monolayer Molybdenum Disulfide. *ACS Nano* **11**, 2115-2123, doi:10.1021/acsnano.6b08278 (2017).

16   Ye, Z. *et al.* Efficient generation of neutral and charged biexcitons in encapsulated WSe_2 monolayers. *Nat Commun* **9**, 3718, doi:10.1038/s41467-018-05917-8 (2018).

17   Dean, C. R. *et al.* Boron nitride substrates for high-quality graphene electronics. *Nature Nanotechnology* **5**, 722-726, doi:10.1038/nnano.2010.172 (2010).

18   Geim, A. K. & Grigorieva, I. V. Van der Waals heterostructures. *Nature* **499**, 419-425, doi:10.1038/nature12385 (2013).

19   Rivera, P. *et al.* Interlayer valley excitons in heterobilayers of transition metal dichalcogenides. *Nature Nanotechnology* **13**, 1004-1015, doi:10.1038/s41565-018-0193-0 (2018).

20   Jin, C. *et al.* Ultrafast dynamics in van der Waals heterostructures. *Nature Nanotechnology* **13**, 994-1003, doi:10.1038/s41565-018-0298-5 (2018).

21   Seyler, K. L. *et al.* Signatures of moire-trapped valley excitons in MoSe$_2$/WSe$_2$ heterobilayers. *Nature* **567**, 66, doi:10.1038/s41586-019-0957-1 (2019).

22   Tran, K. *et al.* Evidence for moire excitons in van der Waals heterostructures. *Nature* **567**, 71, doi:10.1038/s41586-019-0975-z (2019).

23   Alexeev, E. M. *et al.* Resonantly hybridized excitons in moire superlattices in van der Waals heterostructures. *Nature* **567**, 81, doi:10.1038/s41586-019-0986-9 (2019).

24   Jin, C. H. *et al.* Observation of moire excitons in WSe$_2$/WS$_2$ heterostructure superlattices. *Nature* **567**, 76, doi:10.1038/s41586-019-0976-y (2019).

25   Streetman, B. G., Banerjee, Sanjay. *Solid State Electronic Devices* 5th edn, (Prentence Hall, 2000).

26   Novotny, L. & Hecht, B. *Principles of Nano-Optics*. (Cambridge University Press, 2012).

27   Goodman, J. W. *Introduction to Fourier Optics*. 3rd edn, (Roberts and Company Publishers 2005).

28   Betzig, E. *et al.* Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. *Science* **313**, 1642, doi:10.1126/science.1127344 (2006).

29   Müller, C. B. & Enderlein, J. Image Scanning Microscopy. *Physical Review Letters* **104**, 198101, doi:10.1103/PhysRevLett.104.198101 (2010).

30   Hell, S., Wichmann, Jan. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy *Optics Letters Vol.* **1994**, 780-782, doi:https://doi.org/10.1364/OL.19.000780 (1994).

31   Synge, E. H. XXXVIII. A suggested method for extending microscopic resolution into the ultra-microscopic region. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **6**, 356-362, doi:10.1080/14786440808564615 (1928).

32   Lewis, A., Isaacson, M., Harootunian, A. & Muray, A. Development of a 500 Å spatial resolution light microscope: I. light is efficiently transmitted through λ/16 diameter apertures. *Ultramicroscopy* **13**, 227-231, doi:https://doi.org/10.1016/0304-3991(84)90201-8 (1984).

33   Pohl, D. W., Denk, W. & Lanz, M. Optical stethoscopy: Image recording with resolution λ/20. *Appl. Phys. Lett.* **44**, 651-653, doi:10.1063/1.94865 (1984).

34   Ogletree, D. F. *et al.* Near-Field Imaging: Revealing Optical Properties of Reduced-Dimensionality Materials at Relevant Length Scales. *Advanced Materials* **27**, 5692-5692, doi:10.1002/adma.201570255 (2015).

35   Herzog, J. B. *Optical spectroscopy of colloidal CdSe semiconductor nanostructures* Doctor of Philosophy thesis, University of Notre Dame.

36   Kriegel, I. *et al.* Light-Driven Permanent Charge Separation across a Hybrid Zero-Dimensional/Two-Dimensional Interface. *The Journal of Physical Chemistry C*, doi:10.1021/acs.jpcc.0c01147 (2020).

37   Klingshirn, C. *Semiconductor Optics*. (Springer Berlin Heidelberg, 2004).

38   Stockman, M. I. Nanofocusing of Optical Energy in Tapered Plasmonic Waveguides. *Physical Review Letters* **93**, 137404, doi:10.1103/PhysRevLett.93.137404 (2004).

39   Choo, H. *et al.* Nanofocusing in a metal–insulator–metal gap plasmon waveguide with a three-dimensional linear taper. *Nature Photonics* **6**, 838-844, doi:10.1038/nphoton.2012.277 (2012).

40   Bao, W. *et al.* Mapping Local Charge Recombination Heterogeneity by Multidimensional Nanospectroscopic Imaging. *Science* **338**, 1317, doi:10.1126/science.1227977 (2012).

41   Bao, W. *et al.* Visualizing nanoscale excitonic relaxation properties of disordered edges and grain boundaries in monolayer molybdenum disulfide. *Nature Communications* **6**, 7993, doi:10.1038/ncomms8993 (2015).

42   Calafiore, G. *et al.* Campanile Near-Field Probes Fabricated by Nanoimprint Lithography on the Facet of an Optical Fiber. *Scientific Reports* **7**, 1651, doi:10.1038/s41598-017-01871-5 (2017).

43   Park, K.-D. & Raschke, M. B. Polarization Control with Plasmonic Antenna Tips: A Universal Approach to Optical Nanocrystallography and Vector-Field Imaging. *Nano Letters* **18**, 2912-2917, doi:10.1021/acs.nanolett.8b00108 (2018).

44   Bao, W. *et al.* Plasmonic near-field probes: a comparison of the campanile geometry with other sharp tips. *Opt. Express* **21**, 8166-8176, doi:10.1364/OE.21.008166 (2013).

45   Yang, Z., Aizpurua, J. & Xu, H. Electromagnetic field enhancement in TERS configurations. *Journal of Raman Spectroscopy* **40**, 1343-1348, doi:10.1002/jrs.2429 (2009).

46   Huang, T.-X. *et al.* Tip-enhanced Raman spectroscopy: tip-related issues. *Analytical and Bioanalytical Chemistry* **407**, 8177-8195, doi:10.1007/s00216-015-8968-8 (2015).

47   Hayazawa, N., Inouye, Y., Sekkat, Z. & Kawata, S. Metallized tip amplification of near-field Raman scattering. *Optics Communications* **183**, 333-336, doi:https://doi.org/10.1016/S0030-4018(00)00894-4 (2000).

48   Stöckle, R. M., Suh, Y. D., Deckert, V. & Zenobi, R. Nanoscale chemical analysis by tip-enhanced Raman spectroscopy. *Chem. Phys. Lett.* **318**, 131-136, doi:https://doi.org/10.1016/S0009-2614(99)01451-7 (2000).

49   Kastl, C. *et al.* The important role of water in growth of monolayer transition metal dichalcogenides. *2D Materials* **4**, doi:10.1088/2053-1583/aa5f4d (2017).

50   Jaculbia, R. B. *et al.* Single-molecule resonance Raman effect in a plasmonic nanocavity. *Nature Nanotechnology*, doi:10.1038/s41565-019-0614-8 (2020).

51   Mak, K. F. *et al.* Tightly bound trions in monolayer MoS_2. *Nature Materials* **12**, 207, doi:10.1038/nmat3505

https://www.nature.com/articles/nmat3505#supplementary-information (2012).

52   Niehues, I. *et al.* Strain Control of Exciton–Phonon Coupling in Atomically Thin Semiconductors. *Nano Letters* **18**, 1751-1757, doi:10.1021/acs.nanolett.7b04868 (2018).

53   Carvalho, B. R., Malard, L. M., Alves, J. M., Fantini, C. & Pimenta, M. A. Symmetry-dependent exciton-phonon coupling in 2D and bulk MoS2 observed by resonance Raman scattering. *Phys. Rev. Lett.* **114**, 136403, doi:10.1103/PhysRevLett.114.136403 (2015).

54   Krayev, A. *et al.* Dry Transfer of van der Waals Crystals to Noble Metal Surfaces To Enable Characterization of Buried Interfaces. *ACS Applied Materials & Interfaces* **11**, 38218-38225, doi:10.1021/acsami.9b09798 (2019).

55   Li, H. *et al.* From Bulk to Monolayer MoS2: Evolution of Raman Scattering. *Adv. Funct. Mater.* **22**, 1385-1390, doi:10.1002/adfm.201102111 (2012).

56   Groß, H., Hamm, J. M., Tufarelli, T., Hess, O. & Hecht, B. Near-field strong coupling of single quantum dots. *Science Advances* **4**, eaar4906, doi:10.1126/sciadv.aar4906 (2018).

57   Park, K.-D. *et al.* Tip-enhanced strong coupling spectroscopy, imaging, and control of a single quantum emitter. *Science Advances* **5**, eaav5931, doi:10.1126/sciadv.aav5931 (2019).

58   Rosenberger, M. R. *et al.* Nano-"Squeegee" for the Creation of Clean 2D Material Interfaces. *ACS Applied Materials & Interfaces* **10**, 10379-10387, doi:10.1021/acsami.8b01224 (2018).

59   Feng, J., Qian, X., Huang, C.-W. & Li, J. Strain-engineered artificial atom as a broad-spectrum solar energy funnel. *Nature Photonics* **6**, 866-872, doi:10.1038/nphoton.2012.285 (2012).

60   Lloyd, D. *et al.* Band Gap Engineering with Ultralarge Biaxial Strains in Suspended Monolayer MoS_2. *Nano Letters* **16**, 5836-5841, doi:10.1021/acs.nanolett.6b02615 (2016).

61   Tyurnina, A. V. *et al.* Strained Bubbles in van der Waals Heterostructures as Local Emitters of Photoluminescence with Adjustable Wavelength. *ACS Photonics* **6**, 516-524, doi:10.1021/acsphotonics.8b01497 (2019).

62   Khestanova, E., Guinea, F., Fumagalli, L., Geim, A. K. & Grigorieva, I. V. Universal shape and pressure inside bubbles appearing in van der Waals heterostructures. *Nature Communications* **7**, 12587, doi:10.1038/ncomms12587 (2016).

63   Dai, Z. *et al.* Interface-Governed Deformation of Nanobubbles and Nanotents Formed by Two-Dimensional Materials. *Physical Review Letters* **121**, 266101, doi:10.1103/PhysRevLett.121.266101 (2018).

64   Landau, L. D., Lifshiftz, E. M. *Theory of Elasticity* (Pergamon Press, 1959).

65   Hencky, H. On the Stress State in Circular Plates With Vanishing Bending Stiffness. *Zeitchrift fur Mathematik und Physik* **63**, 311-317 (1915).

66   Zabel, J. *et al.* Raman Spectroscopy of Graphene and Bilayer under Biaxial Strain: Bubbles and Balloons. *Nano Letters* **12**, 617-621, doi:10.1021/nl203359n (2012).

67   Koenig, S. P., Boddeti, N. G., Dunn, M. L. & Bunch, J. S. Ultrastrong adhesion of graphene membranes. *Nature Nanotechnology* **6**, 543-546, doi:10.1038/nnano.2011.123 (2011).

68   Lloyd, D. *et al.* Adhesion, Stiffness, and Instability in Atomically Thin MoS2 Bubbles. *Nano Letters* **17**, 5329-5334, doi:10.1021/acs.nanolett.7b01735 (2017).

69   Carmesin, C. *et al.* Quantum-Dot-Like States in Molybdenum Disulfide Nanostructures Due to the Interplay of Local Surface Wrinkling, Strain, and Dielectric Confinement. *Nano Lett.* **19**, 3182-3186, doi:10.1021/acs.nanolett.9b00641 (2019).

70   Ma, Y. *et al.* Extended Hencky solution for the blister test of nanomembrane. *Extreme Mechanics Letters* **22**, 69-78, doi:https://doi.org/10.1016/j.eml.2018.05.006 (2018).

71   Rostami, H., Guinea, F., Polini, M. & Roldán, R. Piezoelectricity and valley chern number in inhomogeneous hexagonal 2D crystals. *npj 2D Materials and Applications* **2**, 15, doi:10.1038/s41699-018-0061-7 (2018).

72   Thomas P. Darlington, C. C., Matthias Florian, Emanuil Yanev, Obafunso Ajayi, Jenny Ardelean, Daniel A. Rhodes, Augusto Ghiotto, Andrey Krayev, K. Watanabe, T. Taniguchi, Jeffrey W. Kysar, Abhay N. Pasupathy, James C. Hone, Frank Jahnke, Nicholas J. Borys, P. James Schuck. Imaging strain-localized exciton states in nanoscale bubbles in monolayer WSe2 at room temperature. *arXiv:2003.01789 [cond-mat.mes-hall]* (2020).

73   Frisenda, R. *et al.* Biaxial strain tuning of the optical properties of single-layer transition metal dichalcogenides. *npj 2D Materials and Applications* **1**, doi:10.1038/s41699-017-0013-7 (2017).

74   Fan, W. *et al.* Vibrational spectrum renormalization by enforced coupling across the van der Waals gap between MoS_2 and WS_2 monolayers. *Phys. Rev. B: Condens. Matter Mater. Phys.* **92**, 241408, doi:10.1103/PhysRevB.92.241408 (2015).

75   Dadgar, A. M. *et al.* Strain Engineering and Raman Spectroscopy of Monolayer Transition Metal Dichalcogenides. *Chem. Mater.* **30**, 5148-5155, doi:10.1021/acs.chemmater.8b01672 (2018).

76   Timoshenko, S., Woinowsky-Krieger, S. *Theory of Plates and Shells*. (McGraw-Hill, 1959).

77   Seung, H. S. & Nelson, D. R. Defects in flexible membranes with crystalline order. *Physical Review A* **38**, 1005-1018, doi:10.1103/PhysRevA.38.1005 (1988).

78   Trefethen, L. N. *Spectral Methods in MATLAB*. (Society for Industrial and Applied Mathematics, 2000).

79   Mai-Duy, N. & Tanner, R. I. A spectral collocation method based on integrated Chebyshev polynomials for two-dimensional biharmonic boundary-value problems. *Journal of Computational and Applied Mathematics* **201**, 30-47, doi:https://doi.org/10.1016/j.cam.2006.01.030 (2007).

80   Steinhoff, A., Rösner, M., Jahnke, F., Wehling, T. O. & Gies, C. Influence of Excited Carriers on the Optical and Electronic Properties of MoS2. *Nano Letters* **14**, 3743-3748, doi:10.1021/nl500595u (2014).

81   Desai, S. B. *et al.* Strain-induced indirect to direct bandgap transition in multilayer WSe_2. *Nano Letters* **14**, 4592-4597, doi:10.1021/nl501638a (2014).

82   Jariwala D, K., A, Wong, J, Robinson, A E, Sherrott, M C, Wang, S, Liu, G-Y, Terrones, M, Atwater, H A. Nanoscale doping heterogeneity in few-layer WSe2 exfoliated onto noble metals revealed by correlated SPM and TERS imaging. *2D Materials* **5**, 035003 (2018).

83   del Corro, E. *et al.* Atypical Exciton–Phonon Interactions in WS_2 and WSe_2 Monolayers Revealed by Resonance Raman Spectroscopy. *Nano Letters* **16**, 2363-2368, doi:10.1021/acs.nanolett.5b05096 (2016).

84   del Corro, E. *et al.* Excited Excitonic States in 1L, 2L, 3L, and Bulk WSe2 Observed by Resonant Raman Spectroscopy. *ACS Nano* **8**, 9629-9635, doi:10.1021/nn504088g (2014).

85   Mak, K. F. *et al.* Tightly bound trions in monolayer MoS_2. *Nature Materials* **12**, 207, doi:10.1038/nmat3505 (2013).

86   Raja, A. *et al.* Coulomb engineering of the bandgap and excitons in two-dimensional materials. *Nature Communications* **8**, 15251, doi:10.1038/ncomms15251 (2017).

87   Unuchek, D. *et al.* Room-temperature electrical control of exciton flux in a van der Waals heterostructure. *Nature* **560**, 340, doi:10.1038/s41586-018-0357-y (2018).

88   Li, H. *et al.* Optoelectronic crystal of artificial atoms in strain-textured molybdenum disulphide. *Nature Communications* **6**, 7381, doi:10.1038/ncomms8381 (2015).

89   Tonndorf, P. *et al.* Single-photon emission from localized excitons in an atomically thin semiconductor. *Optica* **2**, doi:10.1364/optica.2.000347 (2015).

90   Luo, Y., Liu, N., Li, X., Hone, J. C. & Strauf, S. Single photon emission in WSe_2 up 160 K by quantum yield control. *2D Materials* **6**, doi:10.1088/2053-1583/ab15fe (2019).

91   Aharonovich, I., Englund, D. & Toth, M. Solid-state single-photon emitters. *Nature Photonics* **10**, 631-641, doi:10.1038/nphoton.2016.186 (2016).

92   He, Y. M. *et al.* Single quantum emitters in monolayer semiconductors. *Nature Nanotechnology* **10**, 497-502, doi:10.1038/nnano.2015.75 (2015).

93   Koperski, M. *et al.* Single photon emitters in exfoliated WSe_2 structures. *Nature Nanotechnology* **10**, 503-506, doi:10.1038/nnano.2015.67 (2015).

94   Srivastava, A. *et al.* Optically active quantum dots in monolayer WSe_2. *Nature Nanotechnology* **10**, 491-496, doi:10.1038/nnano.2015.60 (2015).

95   Chakraborty, C., Kinnischtzke, L., Goodfellow, K. M., Beams, R. & Vamivakas, A. N. Voltage-controlled quantum light from an atomically thin semiconductor. *Nature Nanotechnology* **10**, 507-511, doi:10.1038/nnano.2015.79 (2015).

96   Chirolli, L., Prada, E., Guinea, F., Roldán, R. & San-Jose, P. Strain-induced bound states in transition-metal dichalcogenide bubbles. *2D Materials* **6**, 025010, doi:10.1088/2053-1583/ab0113 (2019).

97   Edelberg, D. *et al.* Approaching the Intrinsic Limit in Transition Metal Diselenides via Point Defect Control. *Nano Letters* **19**, 4371-4379, doi:10.1021/acs.nanolett.9b00985 (2019).

98   Park, K. D. *et al.* Hybrid Tip-Enhanced Nanospectroscopy and Nanoimaging of Monolayer WSe_2 with Local Strain Control. *Nano Letters* **16**, 2621-2627, doi:10.1021/acs.nanolett.6b00238 (2016).

99   Lin, Z. *et al.* 2D materials advances: from large scale synthesis and controlled heterostructures to improved characterization techniques, defects and applications. *2D Materials* **3**, doi:10.1088/2053-1583/3/4/042001 (2016).

100   Lee, Y. *et al.* Near-field spectral mapping of individual exciton complexes of monolayer WS_2 correlated with local defects and charge population. *Nanoscale* **9**, 2272-2278, doi:10.1039/c6nr08813a (2017).

101   Schuck, P. J., Bao, W. & Borys, N. J. A polarizing situation: Taking an in-plane perspective for next-generation near-field studies. *Frontiers of Physics* **11**, doi:10.1007/s11467-015-0526-5 (2016).

102   Haigh, S. J. *et al.* Cross-sectional imaging of individual layers and buried interfaces of graphene-based heterostructures and superlattices. *Nature Materials* **11**, 764, doi:10.1038/nmat3386 (2012).

103   Shi, W. *et al.* Raman and photoluminescence spectra of two-dimensional nanocrystallites of monolayer WS_2 and WSe_2. *2D Materials* **3**, doi:10.1088/2053-1583/3/2/025016 (2016).

104   Eggleston, M. S., Messer, K., Zhang, L., Yablonovitch, E. & Wu, M. C. Optical antenna enhanced spontaneous emission. *Proceedings of the National Academy of Sciences* **112**, 1704-1709 (2015).

105   Park, K. D., Jiang, T., Clark, G., Xu, X. D. & Raschke, M. B. Radiative control of dark excitons at room temperature by nano-optical antenna-tip Purcell effect. *Nature Nanotechnology* **13**, 59-+, doi:10.1038/s41565-017-0003-0 (2018).

106   Chow, P. K. *et al.* Defect-Induced Photoluminescence in Monolayer Semiconducting Transition Metal Dichalcogenides. *ACS Nano* **9**, 8 (2015).

107   Kern, J. *et al.* Nanoscale Positioning of Single-Photon Emitters in Atomically Thin WSe_2. *Advanced Materials* **28**, 7101-7105, doi:10.1002/adma.201600560 (2016).

108   Luo, Y. *et al.* Deterministic coupling of site-controlled quantum emitters in monolayer WSe_2 to plasmonic nanocavities. *Nat Nanotechnol* **13**, 1137-1142, doi:10.1038/s41565-018-0275-z (2018).

109   Ostadhossein, A. *et al.* ReaxFF Reactive Force-Field Study of Molybdenum Disulfide (MoS2). *The Journal of Physical Chemistry Letters* **8**, 631-640, doi:10.1021/acs.jpclett.6b02902 (2017).

110   van Duin, A. C. T., Dasgupta, S., Lorant, F. & Goddard, W. A. ReaxFF: A Reactive Force Field for Hydrocarbons. *The Journal of Physical Chemistry A* **105**, 9396-9409, doi:10.1021/jp004368u (2001).

111   Froyen, S. & Harrison, W. A. Elementary prediction of linear combination of atomic orbitals matrix elements. *Phys. Rev. B: Condens. Matter Mater. Phys.* **20**, 2420-2422, doi:10.1103/PhysRevB.20.2420 (1979).

112   Chiang, C. L., Xu, C., Han, Z. M. & Ho, W. Real-space imaging of molecular structure and chemical bonding by single-molecule inelastic tunneling probe. *Science* **344**, 885-888, doi:10.1126/science.1253405 (2014).

113   Brunner, K., Abstreiter, G., Böhm, G., Tränkle, G. & Weimann, G. Sharp-Line Photoluminescence and Two-Photon Absorption of Zero-Dimensional Biexcitons in a GaAs/AlGaAs Structure. *Physical Review Letters* **73**, 1138-1141, doi:10.1103/PhysRevLett.73.1138 (1994).

114   De Luca, M. *et al.* New insights in the lattice dynamics of monolayers, bilayers, and trilayers of WSe2 and unambiguous determination of few-layer-flakes' thickness. *2D Materials* **7**, 025004, doi:10.1088/2053-1583/ab5dec (2020).

115   Hanbicki, A. T. *et al.* Double Indirect Interlayer Exciton in a $MoSe_2$/WSe_2 van der Waals Heterostructure. *ACS Nano* **12**, 4719-4726, doi:10.1021/acsnano.8b01369 (2018).

116   He, K. *et al.* Tightly Bound Excitons in Monolayer ${\mathrm{WSe}}_{2}$. *Physical Review Letters* **113**, 026803, doi:10.1103/PhysRevLett.113.026803 (2014).

117   Karanikolas, V., Thanopulos, I. & Paspalakis, E. Strong interaction of quantum emitters with a WS2 layer enhanced by a gold substrate. *Opt. Lett.* **44**, 2049-2052, doi:10.1364/OL.44.002049 (2019).