Layer-by-layer thinning of MoS$_2$ by thermal annealing

Xin Lu,$^a$ Muhammad Iqbal Bakti Utama,$^a$ Jun Zhang,$^a$ Yanyuan Zhao$^a$ and Qihua Xiong$^{ab}$

By thermal annealing, few-layer MoS$_2$ flakes can be thinned down. In one hour, the upper layer is peeled off due to sublimation. Eventually, monolayer MoS$_2$ is achieved. We have characterized the process by optical contrast, Raman spectroscopy and atomic force microscopy (AFM), and observed a mixture of surfaces of $N$ and $N-1$ layers.

Atomically thin layered transition metal dichalcogenides (TMDs) have aroused lots of interest recently, especially molybdenum disulfide (MoS$_2$). As a typical TMD, MoS$_2$ possesses special layer-related electronic and optical properties.$^1$ Bulk MoS$_2$ has an indirect gap, which increases gradually as we move towards atomic scale. The indirect gap evolves into a direct one in a monolayer.$^2$ Both single-layer and multilayer MoS$_2$ have been shown to be suitable materials for electronic devices, such as field effect transistors (FETs), FET-based small signal amplifiers, FET-based sensing devices and transistor-based integrated circuits.$^3-6$ Monolayer MoS$_2$ is considered a suitable material for valleytronics,$^7$ since valley polarization has already been demonstrated at low temperature.$^8-9$

Other features have been observed at low temperature, such as an increasing photoluminescence (PL) intensity in few-layer MoS$_2$$^{10}$ and tightly bound trions in monolayer MoS$_2$.$^{11}$ Due to the direct gap, single-layer MoS$_2$ has a stronger PL than the few-layer MoS$_2$, even at room temperature.$^{12,13}$ All the evidence shows monolayer MoS$_2$ to be a suitable candidate for optoelectronics devices,$^{1,12,13}$ and the optoelectronic application of multilayer MoS$_2$ is also worth investigating. Meanwhile, layer-related properties have also been observed by Raman spectroscopy. Two fingerprint peaks, the $A_{1g}$ and $E_{2g}$ modes, shift according to the number of layers.$^{14-18}$ $A_{1g}$ shifts to higher frequencies as the number of layers increases, due to stronger van der Waals forces. In contrast, $E_{2g}^1$ down shifts because of an enhancement in the dielectric screening of the long-range Coulomb interactions, which is much stronger than the short range van der Waals forces.$^{15}$ Hence, it is proposed, that from the frequency difference between the $A_{1g}$ and $E_{2g}$ modes, it would be possible to determine the number of layers under five. Recently, a few groups including ours have demonstrated that low frequency (~50 cm$^{-1}$) vibrational breathing and shear modes show highly sensitive but distinct dependence on thickness.$^{17-19}$ The distinct shift of the low-frequency interlayer breathing mode (called B1 for convenience) can facilitate distinguishing between single- to nine-layer MoS$_2$ flakes very accurately. Beyond nine layers, the breathing mode energy is so small that it is beyond the detection limit of the spectrometer. We have mainly used this method to determine the number of layers in our thinning experiment, in addition to AFM.

Several approaches have been reported to produce few-layer MoS$_2$ flakes, including scotch tape-based mechanical exfoliation, liquid-based exfoliation$^{20,21}$ and chemical vapor deposition (CVD) growth.$^{22,23}$ Mechanical exfoliation is a random process and has a very low yield. The liquid method and CVD growth will introduce impurities unavoidably. Other new methods have been developed recently, such as laser thinning and anodic binding.$^{24,25}$ While the latter one is actually an application of the method reported in graphene, laser thinning is relatively new. A green laser has been adopted to attenuate multilayer MoS$_2$ flakes into monolayers. Laser-induced heat causes the sublimation of the upper layers when the power exceeds 10 mW, while the bottom layer would not disappear until the power reaches another threshold, around 17 mW. But laser-thinning can only produce monolayer samples. No bilayers or trilayers can be achieved in this way. However, layer-by-layer thinning is urgently needed for the study of atomically thin MoS$_2$. For example, we could conduct in situ measurements to further investigate layer-dependent electronic and optical properties from the same original sample. Even with the same thickness, some properties vary, e.g. the PL intensity. Recently, layer-by-layer thinning of MoS$_2$ has been achieved by argon plasma,

$^a$Division of Physics and Applied Physics, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371. E-mail: Qihua@ntu.edu.sg

$^b$NOVITAS, Nanoelectronics Centre of Excellence, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798.
which is more likely to cause damage than other thinning methods due to the energetic plasma. Here, we demonstrate that simple thermal annealing in a controlled atmosphere can also lead to layer-by-layer thinning.

We used Scotch tape (3M) to exfoliate few-layer flakes from bulk MoS2 and deposited the thin samples on a Si substrate with 285 nm of SiO2. Monolayer to eight-layer samples could be observed and located under an optical microscope (Olympus BX51). We could even estimate the layer thickness by optical contrast. The contrast spectra of pristine few-layer MoS2 and thinned MoS2 were measured in a microspectrophotometer (Cricad 20) under a 36 × objective, with a numerical aperture of 0.5. A contrast spectrum calculated from Fresnel equations was also obtained, and agreed with the experimental results. In addition, AFM (Park NX-10) and Raman spectroscopy (Horiba-JY T64000) were used to determine the number of layers in a much faster way. MoS2 flakes were measured under a back scattering configuration and excited with a 532 nm laser. The low-frequency interlayer vibrational modes can be used to verify the number of layers from one to eight.

Thermal annealing was conducted in a home built CVD. By placing samples with the silicon substrate at the centre of a quartz tube, we made sure the temperature on the samples is the same as the set point. In this case, 650 °C and 10 Torr argon gas were set, with a gas flow rate of 5 standard cubic centimeters per minute (scm). We set one hour as one cycle, during which one layer was lost. In fact, these parameters gave us the optimized results, considering the thinning rate, the size and surface quality of the thinned samples. Fig. 1a (eight-layer) and 1e (seven-layer) are pristine MoS2 flakes before annealing thinning. Fig. 1b–d and Fig. 1f–h display the gradual changes of these few-layer samples.

Since thicker MoS2 flakes can reflect more light and show a brighter colour than thinner ones, the change of colour, from bright to dark, indicates that the flakes are getting thinner and thinner. Meanwhile, a shrinkage of the surface area is noticeable. This means that sublimation not only occurs perpendicularly on the surface, but also along the surface, thus resulting in smaller areas. To make a detailed comparison, we measured the contrast spectra of pristine single-layer to eight-layer MoS2 (Fig. 2b) and thinned MoS2 by thermal annealing (Fig. 2a), whose original number of layers was eight (shown in Fig. 1a).

From Fig. 2a and b, the vertical axis is the optical contrast C(λ), which is defined as:

\[
C(\lambda) = \frac{R_0(\lambda) - R(\lambda)}{R_0(\lambda)}
\]

where \(R_0(\lambda)\) is the reflectance from the substrate, and \(R(\lambda)\) is the reflectance from the MoS2 sample. If \(C(\lambda) > 0\), such as the spectra between 600 and 700 nm, it means that the SiO2-Si substrate reflects more light than the MoS2 sample. If \(C(\lambda) < 0\), e.g. the region of 450–500 nm, it means that MoS2 reflects more light than SiO2–Si. From Fig. 2b, we can see that thicker MoS2 flakes have deeper valleys in the negative contrast region, which indicates that as the thickness increases, the reflected light from the samples within the range of 450–500 nm also increases since the few-layer MoS2 on 285 nm SiO2–Si substrate shows a blue colour (see Fig. 1). Comparing Fig. 2a and b, we can observe a similar change of contrast between thinned and pristine MoS2. Even though there is a slight difference (e.g. the change from seven-layer to six-layer), we believe it is mostly due to systematic errors in the spectrum measurements. From Fig. 2a, as the thinning cycle increases, the contrast decreases correspondingly. This is the first evidence that our thermal annealing method effectively thins down few-layer MoS2 samples layer-by-layer. The contrast spectra calculated using the Fresnel equations are shown in Fig. 2c. In our simulation, we considered Si as semi-infinite and its extinction coefficient \(\kappa = 0\). The refractive index \(n\) of air is 1 and is wavelength-independent. The refractive index and extinction coefficient of MoS2 in our calculations came from a fitting, and they were about 5% smaller than the values from a previous report. and \(\kappa\) are plotted in Fig. 2d. Apparently, the simulation shows a larger negative value compared to the experimental results. We attribute this to two possible reasons. First of all, we did not consider the influence of the numerical aperture (NA) in the calculations. We simply hypothesized that the incident angle of light was 0°, while the NA is equal to 0.5 in reality. Therefore, we overestimated the absolute contrast in the simulations. Another reason, we believe, is the efficiency of collection. The aperture size we used is a relatively small square, corresponding to a
3.6 × 3.6 μm² area. If we used a larger size, we could observe higher peaks and lower valleys, while the position of the peaks and valleys would remain unchanged. As a matter of fact, the raw data from the microphotometer mean:

\[
\frac{R_{\text{sample}}}{R_{\text{substrate}}} = \frac{R_{\text{dark}}}{R_{\text{dark}}}
\]

\(R_{\text{dark}}\) is the dark scan of the detector, and it can be treated as constant. If \(R_{\text{sample}}\) and \(R_{\text{substrate}}\) are much larger than \(R_{\text{dark}}\), the above ratio would be much closer to \(R_{\text{sample}}/R_{\text{substrate}}\), the relative reflectance that we are looking for. Therefore, it is reasonable that a smaller aperture size results in a smaller contrast than the true value. However, due to the limited size and irregular shape of the samples, we cannot use a larger aperture size, or it will be difficult to guarantee that the area within the aperture is of uniform thickness. Despite this, the good agreement between calculations and experiments (both pristine and thinned MoS₂) shows that our results are robust.

To characterize the thinning process of the original seven-layer MoS₂ sample (shown in Fig. 1d), we conducted a Raman scattering experiment. Compared to the high frequency regime, low frequency modes are much more sensitive to thickness.19 For instance, the B1 mode shifts from 10 cm⁻¹ in a nine-layer sample to 40 cm⁻¹ in a bilayer sample. However, the frequency difference between \(A_{1g}\) and \(E_{2g}\) increased from about 18 cm⁻¹ in a monolayer sample to 24.5 cm⁻¹ in a five-layer sample, which is already close to that from bulk. The original and thinned MoS₂ flakes were measured after each cycle (one hour) under both parallel \(z(xz)z\) and cross \(z(xy)z\) polarization configurations.19 From Fig. 3a, we can see that the low frequency breathing mode B1 peak shifts to a higher frequency and the shear mode \(E_{2g}\) (called S1 for convenience) down shifts as the annealing cycle increases, indicating that the layer thickness is decreasing. The mixture of surfaces is also reflected in the spectra. The two B1 peaks after the second and third cycle of annealing confirm that the surface is a mixture of \(N\) and \(N - 1\) layers (\(N = 6, 5\)) even in a small area, within a laser spot. After the seventh cycle, the interlayer breathing mode B1 and shear mode S1, which are mainly caused by van der Waals forces, disappear. The presence of the fingerprint modes \(A_{1g}\) and \(E_{2g}\) in Fig. 3b suggests that the sample is in monolayer regime. Meanwhile, the shifts in \(A_{1g}\) and \(E_{2g}\), especially from the fourth cycle to the seventh cycle, are distinct. It can be regarded as another sign of thinning, which is in agreement with previous reports.14,16–19 Fig. 3c and d are the spectra from the cross-polarized configuration. Only the shear mode vibration can be observed, since breathing modes, \(A_{1g}\) and B1, are greatly suppressed. The down-shifted S1 in Fig. 3c is evidence of thinning. Two S1 peaks after the fourth and fifth cycle were observed, since the S1 peak shifts more in thinner MoS₂ flake. This phenomenon indicates a mixed surface of \(N\) and \(N - 1\) layers (\(N = 4, 3\)).

We summarize the shifts of the S1, B1, \(A_{1g}\) and \(E_{2g}\) modes as a function of the annealing cycles in Fig. 3e and f. The general trend of shifting confirms that thermal annealing can thin down few-layer MoS₂. From Fig. 3d, we can see that after the sixth cycle, \(A_{1g}\) was not as suppressed as in other cycles. We believe there may be some heat-induced disorders after this cycle, which mainly affect out-of-plane vibrations. Although the full width at half maximum (FWHM) of the pristine bilayer is comparably larger than that of the other multilayers, it almost reaches 8.5 cm⁻¹ after the sixth thinning cycle. For the rest of the thinning cycles, FWHM of the S1, B1, \(A_{1g}\) and \(E_{2g}\) peaks is below 2.0 cm⁻¹, mostly around 1.5 cm⁻¹ (data not shown here), which indicates that thermal annealing does not compromise the crystalline quality of few-layer MoS₂ in terms of its optical properties. The frequency difference of \(A_{1g}\) and \(E_{2g}\) is about 21.5 cm⁻¹ in a monolayer, and close to 23 cm⁻¹ in a bilayer, which is larger than that of the pristine MoS₂ produced by the mechanical exfoliated method.14 Similar effects were observed in monolayer MoS₂ flakes fabricated by the chemical method12 and laser thinning.24 On the other hand, the smaller frequency difference in the monolayer (16 cm⁻¹) and the bilayer (21 cm⁻¹) was also noticed in mechanically exfoliated MoS₂ in a previous report.23 Instead of the other likely reasons such as local thickness variation12 and a rougher surface caused by the laser,24 we believe this small difference is related to the lifetime of phonons. Similarly to laser thinning, we obtain a large frequency difference in thinned monolayers, as both thinning methods share a similar mechanism, a heat-induced sublimation. From our results, a larger frequency difference is mainly caused by the shear mode \(E_{2g}\), which did not shift too much after the fifth cycle. Therefore, we conclude that too many thermal thinning cycles will influence the force in shear mode vibration, while that of the breathing mode can hardly be affected.21
We further used AFM to check the uniformity and surface roughness of a thinned MoS$_2$ flake. Fig. 4a shows a flat surface of an exfoliated MoS$_2$ flake (the corresponding optical image is Fig. 1a), which is about 5.5 nm in thickness, corresponding to an eight-layer sample. After one annealing cycle, an inhomogeneous surface appears, indicating that some parts are getting thinned to six-layer flakes, while some parts are seven-layer flakes, as shown in Fig. 4b. In repeated experiments, although we have succeeded in thinning a small portion of samples to homogeneous thickness, for most atomically thin MoS$_2$ flakes an inhomogeneous surface appears. We therefore considered vacancy defects as a possible reason to cause the uneven surface. The vacancy defects could come naturally from the samples themselves, as well as during the thinning process. In a “perfect” and “defect-free” MoS$_2$ sample, sublimation can happen randomly in any area, leaving vacancies (thickness is $N-1$ layers after sublimation), which make their surrounding parts ($N$ layers) easier to sublime due to the broken covalent bonds. As to the surface roughness, since there is inhomogeneity, we just selected some parts with uniform thickness and find that the roughness hardly changed during the thinning process, even when it came to the thinned bilayer and monolayer samples.

**Conclusion**

To conclude, via thermal annealing, few-layer MoS$_2$ flakes can be peeled off layer-by-layer. The thinning process was characterized by microspectrophotometry, Raman spectroscopy and AFM. During thinning, we can see that the crystalline quality does not change much, and the surface roughness mostly stays stable. Although the thinning process is not perfectly layer-by-layer yet, we believe that thinned MoS$_2$ flakes with homogenous surfaces can be achieved by thermal annealing. Thus, we consider this easy and low-cost approach a promising top-down method for fabricating single- and few-layer MoS$_2$ as well as facilitating future in situ measurements on electronic and optical properties. This thinning process may apply to other transition metal dichalcogenides as well.

**Acknowledgements**

Q. X. acknowledges the strong support from the Singapore National Research Foundation through a Fellowship grant (NRF-RF2009-06), the Singapore Ministry of Education via two tier2 grants (MOE2011-T2-2-051 and MOE2012-T2-2-086) and the Nanyang Technological University via a start-up grant support (M58110061).

**Notes and references**


27 M. M. Benamer, B. Radisavljevic, J. S. Heron, S. Sahoo, H. Berger and A. Kis, Nanotechnology, 2011, 22, 125706.


