Peer Review File

Light-modulated van der Waals force microscopy

Corresponding Author: Professor Hong-Bo Sun

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The manuscript under review introduces a concept of light-enhanced van der Waals (vdW) force, potentially offering an innovative approach to atomic force microscopy (AFM). This concept could significantly contribute to the field. However, neither the theoretical principle nor the experimental evidence can clearly show this. The claimed 20 dB enhancement is the ratio between the measured signal with laser on and off, which could simply be given by the optical force related to the photothermal expansion. Furthermore, optical force microscopy with a similar setup has been reported in many papers, such as

• Lu, F., Jin, M., & Belkin, M. A. (2014). Tip-enhanced infrared nanospectroscopy via molecular expansion force detection. Nature photonics, 8(4), 307-312.

• Nowak, D., Morrison, W., Wickramasinghe, H. K., Jahng, J., Potma, E., Wan, L., ... & Park, S. (2016). Nanoscale chemical imaging by photoinduced force microscopy. Science advances, 2(3), e1501571.

• Wang, H., Meyer, S. M., Murphy, C. J., Chen, Y. S., & Zhao, Y. (2023). Visualizing ultrafast photothermal dynamics with decoupled optical force nanoscopy. Nature Communications, 14(1), 7267

While the manuscript proposes a novel approach in LvFM, it currently falls short of providing convincing evidence and theoretical support for the claimed light-enhanced vdW force. A more rigorous theoretical framework, comprehensive experimental validation, and a detailed comparison with existing techniques are required. In its current form, the manuscript is not suitable for publication and requires substantial revision.

Here are my detailed comments:

1. The authors report a light-enhanced vdW force; however, the principle is not clearly illustrated. The authors should show the theoretical model of how this enhancement is achieved. The enhancement should be intrinsic to the material. If the enhancement is achieved through a modulated photothermal expansion that changes the tip-sample distance, it does not count as an enhanced vdW force. Such properties can be directly reflected by the phase of the AC attractive mode (non-contact mode) of AFM, without using optical force microscopy.
2. The authors should provide the phase map of the topography to prove that light can change the intrinsic property of the material that is related to the vdW force.
3. vdW force varies dramatically with distance. The authors should justify why they use the tapping mode. At which distance is the vdW force measured?
4. The tapping mode of AFM involves a fixed engagement factor locked by the internal PID loop of the AFM system. Instead of showing the tip-sample force versus separation with a fixed free air amplitude (Fig. 1b), it may be more informative to show the force versus the engagement factor. How does the vdW force contribute to the curve?
5. How is the setup illustrated in Fig. 2 different from the one of PiFM? If it is the same, why does the measured signal lead to additional vdW force information?
6. What does the white curve represent in Fig. 2g?
7. The optical force could result from the photothermal expansion of not only the sample but also of the substrate and the AFM probe. To exclude this effect and show the contribution of the 2D material, the authors should compare the signal in Fig. 3c with the one without the 2D material (only with the substrate).

8. The histogram of the optical force during a single engagement in Fig. 3e cannot effectively represent the standard derivation of the vdW force due to the low-pass filtering of the lock-in amplifier. What is the integration time of the lock-in amplifier?

9. The signal measured at multiple engagements is recommended to eliminate the artifacts raised by mode hopping (between attractive mode and repulsive mode). How do the authors ensure the system operates in the attractive mode where the van der Waals interaction is dominant?

10. Why is the Raman microscopy and spectrum for the sample in Fig. 4g missing?

Reviewer #2

(Remarks to the Author)
The authors developed the Light-enhanced van der Waals force microscopy (LvFm), facilitating compositional identification of materials. This equipment can acquire information about the composition by analyzing the optothermal interaction between the tip and the sample. In contrast to AFM, which is unable to provide compositional data, the LvFm enables the determination of both the thickness and composition of the sample. The operational principles of this equipment are explained in details within the manuscript, highlighting its distinctive features compared to other AFM-IR, Raman and XRD. However, given that the primary driving force behind the LvFm lies in van der Waals interaction, a more compelling demonstration of its advantages would be achieved through the inclusion of diverse samples and experiments. The authors dedication to providing a thorough explanation of the LvFm's principles and distinctions is commendable. Expanding the various samples and experiments would undoubtedly enhance the credibility and appeal of this approach.

The utilization of IR and Raman, despite their comparatively lower resolution, is grounded in the certainty of vibrational modes associated with chemical bonding. A well- established vibration library, tailored to specific molecules, underscores the reliability of these methods.

In contrast, Van der Waals forces, being the weakest among intermolecular forces, introduce variability in force interactions, demanding careful consideration of sample quality. Even when measuring the same sample, the forces may exhibit differences. A more in-depth discussion should be needed. Therefore, this manuscript can be accepted after conducting additional experiments and explaining any unclear aspects.

1. The authors presented an example involving MoTe2 to demonstrate the potential of compositional indentification using LvFm. However, this aspect appears somewhat ambiquous. It aligns more closely with phase identification of TMDs rather than explicit compositional identification. To establish the capacity of LvFm for compositional identification, it is imperative to show additional examples such as organic compounds (polymer films), metal alloys, metal oxides, and other diverse materials.

2. The distinction between the 2H and 1T' phases of MoTe2 achieved through LvFm lacks a precise explanation of the underlying principle. Notably, the LvFm amplitude exhibits strength in the semiconducting 2H region and weakness in the metallic 1T' phase. The absence of a detailed explanation raises questions about the feasibility of compositional identification with LvFm, especially in the case of metallic materials.

3. What is the actual thickness of MoTe2 in Figure 4? How does the LvFm data change based on the thickness? In the case of 2D materials, Raman data typically shifts by about 1~3 cm-1, according to the thickness. How does LvFm compare to the sensitivity of Raman? Given that LvFm is significantly influenced by van der Waals forces, there appears to be substantial variation based on how the sample is prepared. How would the authors explain this?

Reviewer #3

(Remarks to the Author)
Dear editor,

The manuscript introduces and describes with important details a novel approach to easily discriminate 2D materials by combining AFM topography with light illumination. The authors present a very sound description of the phenomenon on the basis of enhanced van der Waals interaction between the AFM tip and the illuminated 2D material. They motivate well the importance of the technique by comparing it to other widely used technics. They also describe the way to implement it experimentally. They finally demonstrate its potential on 2 examples of 2D systems: 1T'/2H phases in MoTe2 and a hBN/WS2 stack, where the 2 materials can be discriminated thanks to the introduced technique. Overall the manuscript is of importance to the 2D materials community, and maybe to the larger surface science audience. It is well organized (to the most part, see issues below) and follows the standards for publication in your journal.

I would support publication given that the following issues are answered:

1. My main concern relates to the figures of merit introduced in the text. It is of high importance since they are also reported in the abstract. First, the 20dB enhancement factor is material dependent (here obtained on MoTe2). This is not stated in the abstract, and not clearly mentioned in the text. It is also not relevant to illustrate the discriminating power of the technique: if 2 materials have the same enhancement power, no contrast will be observed. I also have concerns on what is called "composition resolution". It appears to me simply as the signal contrast (or a derived way to express it). It is actually referred as is (i.e. U_H/U_T) in SI 10 and in figure 4f. It is also material dependent. Finally, the authors introduce a detection limit DL with the use of delta_1 and delta_2 which are no where defined. This DL is then used to normalize a contrast and define a

resolving power. This gives an impressive number that is to me suspicious, especially due to its mentioning in the abstract with less context and definition. Overall, this aspect on figures of merit requires to be revised with a much clearer definition of the extracted figures, and with their report in the abstract only if they are relevant and with additional context (material dependence character at least).

2. Another important concern relates to the last part introducing the experimental examples. First, MoTe2 phases are introduces with no description in the spatial/compositional resolution section and are only (quickly) described in the next section. Similarly, the hBN/WS2 stack is very vaguely introduced and the message related to this last paragraph is not clear. With the help of SI and thorough figure analysis, I am convinced the result is of importance, it is just poorly reported. I would advise reorganizing this last section.

3. I noted that, in page 5, the random amplitude variation "u" is defined later than its first introduction in equation 3. Also, its unit is not clear (unit of length given that we have "r+u" in the equation or unitless given that u = 0.01 later?), as well as its relation with sigma. This should be clearly defined in their first mention in page 5, and be coherent with their later use in page 10.

4. I would advise to reword the last sentences of page 6. "Astonishingly" appears too strong for a scientific assertion. "cone-like" and "copy of a Gaussian" may not be the most rational way to present the data.

Minor typos:
Space missing page 10 between "LvFM" and "requires"
I guess "Physical original" should be "Physical origin" in Fig 1 caption

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I appreciate the authors' great efforts in improving the quality of the manuscript. Most of my questions have been well-addressed. I have a few follow-up questions I hope the authors can address.
1. I agree that different physical effects may be dominated by different probe-tip working zones. I wonder if it is only a function of the tip-sample distances or is also related to other factors such as the choice of probe, type of samples, and method of modulation/demodulation. The author should demonstrate the specificity of the signal to the "probe tip working zone." How different would the signal be if it was not through the vdW force? How sensitive is it to the tip-sample distance?
2. It is interesting that the modulation of LvF does not require a photothermal expansion. For few-layer 2D material, I agree that there is no significant thermal expansion; therefore, the optical force mostly comes from the change of LvF. However, what if the sample absorbs the wavelength of light significantly? Will this technique still work? For example, in the Au nanoparticle and PS nanosphere experiment demonstrated in the SI, what specific experimental parameters, including the wavelength of light, type of probe, and modulation amplitude, were used?
3. Why do the tapping mode and sideband modulation/demodulation provide a better SNR? Shouldn't the non-contact mode with direct modulation provide a better SNR if the force is only decided by the tip-sample distance?
4. Since the measurement setup of LvFM and PiFM both use light-modulated tip-sample interaction, I still don't understand how one effect (light-induced dipole-dipole interaction, for example) will not appear in the other measurement techniques. I agree that the dominating effects may vary from sample to sample, and LvF may be dominant in the 2D material sample demonstrated in this manuscript. However, is this a general conclusion that LvFM can decouple LvF from the other light-induced tip-sample interactions (for example, photothermal expansion, dipole-dipole interactions, photoacoustic pressures, etc.)?
5. The authors mentioned that LvFM requires a closer tip-sample distance than PiFM; how is this achieved specifically? I don't think PiFM has specified the tip-sample distance in operation. The difference between LvFM and PiFM is unclear, and this needs to be addressed.

Reviewer #2

(Remarks to the Author)
The biggest weakness of this paper was the question of whether it could be applied to various samples, but the doubt was resolved by applying and proving it to various samples. The manuscript can be accepted in Nature Communications.

Reviewer #3

(Remarks to the Author)
I am satisfied by the answers given by the authors on the different comments I made previously. I confirm that the corresponding changes made to the manuscript are relevant and make the overall demonstration more compelling and scientifically sound. I do not have further comments.

Please note that, in the current given version, formatting errors are visible (line 138 of manuscript, and in several instances in SI text and equations, I guess they are missing delta or (number) signs).

Version 2:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I thank the authors for their efforts in addressing my questions. Most of the questions have been answered, but some parts were not addressed thoroughly.

1. It is interesting to see this force varying with distance z. What is the definition of z=0 and how is this calibrated? My question is specifically related to how this force is attributed to the light-enhanced van der Waals force rather than other light-induced forces discussed in prior works. Have the authors tried changing the modulation/demodulation methods or setpoint to see how these factors could affect the forces?

2. Based on Figure 4, the stronger the absorption, the stronger the force observed. How did the authors exclude the contribution of photothermal expansion of the sample? Additionally, gold absorbs more strongly than polystyrene in the visible spectrum, but Figure S10d shows no force for gold particles and significant force for polystyrene beads. This seems to contradict the finding in Figure 4.

===== **Detailed point-by point responses to all the review comments** ========
Reviewer #1

The manuscript under review introduces a concept of light-enhanced van der Waals (vdW) force, potentially offering an innovative approach to atomic force microscopy (AFM). This concept could significantly contribute to the field. However, neither the theoretical principle nor the experimental evidence can clearly show this. The claimed 20 dB enhancement is the ratio between the measured signal with laser on and off, which could simply be given by the optical force related to the photothermal expansion. Furthermore, optical force microscopy with a similar setup has been reported in many papers, such as

• Lu, F., Jin, M., & Belkin, M. A. (2014). Tip-enhanced infrared nanospectroscopy via molecular expansion force detection. Nature photonics, 8(4), 307-312.

• Nowak, D., Morrison, W., Wickramasinghe, H. K., Jahng, J., Potma, E., Wan, L., ... & Park, S. (2016). Nanoscale chemical imaging by photoinduced force microscopy. Science advances, 2(3), e1501571.

• Wang, H., Meyer, S. M., Murphy, C. J., Chen, Y. S., & Zhao, Y. (2023). Visualizing ultrafast photothermal dynamics with decoupled optical force nanoscopy. Nature Communications, 14(1), 7267

While the manuscript proposes a novel approach in LvFM, it currently falls short of providing convincing evidence and theoretical support for the claimed light-enhanced vdW force. A more rigorous theoretical framework, comprehensive experimental validation, and a detailed comparison with existing techniques are required. In its current form, the manuscript is not suitable for publication and requires substantial revision.

<u>**Response:**</u>
We are very grateful to the reviewer for the valuable comments and suggestions. We have followed the reviewer's suggestions to comprehensively improve the theoretical presentation and experimental evidences of LvFM, as detailed replied below.

    First of all, in the revision, we have put much emphasis on strengthening the clear explanation of the principle of LvFM with more derivation and numerical simulation, so as to help readers clearly understanding the intrinsic difference of LvFM with other near-field microscopic methods involving optical forces. Briefly speaking here, the optically induced force utilized in LvFM is firstly *a vdW force between the tip and sample*, which happens only in the very close vicinity of the probe tip (so-called "*vdW force interaction zone*" with tip-sample distance $h < 1$ nm, as shown in Fig. R1-1 and Fig. S1). In contrast, other near-field methods (including the three works mentioned by the reviewer) usually do not work in this vdW force interaction zone and do not utilize the tip-sample vdW force. Secondly, the tip-sample force utilized in LvFM is not a normal *static* vdW force, but a light-modulated *dynamic* vdW force (i.e., LvF), which

1

is generated by an excitation laser with modulation frequency f_m satisfying the condition $f_2 = |f_m \pm f_1|$ (where f_1 and f_2 are the fundamental and second-order resonance frequencies of the cantilever, respectively). Under the same excitation laser, different materials may have different LvF responses (as detailedly explained in the theoretical part of the paper and also responses below), just because the bonding forces between molecules or atoms in a material are different from those in another material. Hence, *only with this LvF, the different compositional regions in a sample can be directly distinguished with an AFM tip as a force sensor* (without the help of any spectroscopic technique), which cannot be realized with conventional AFM utilizing static tip-sample force.
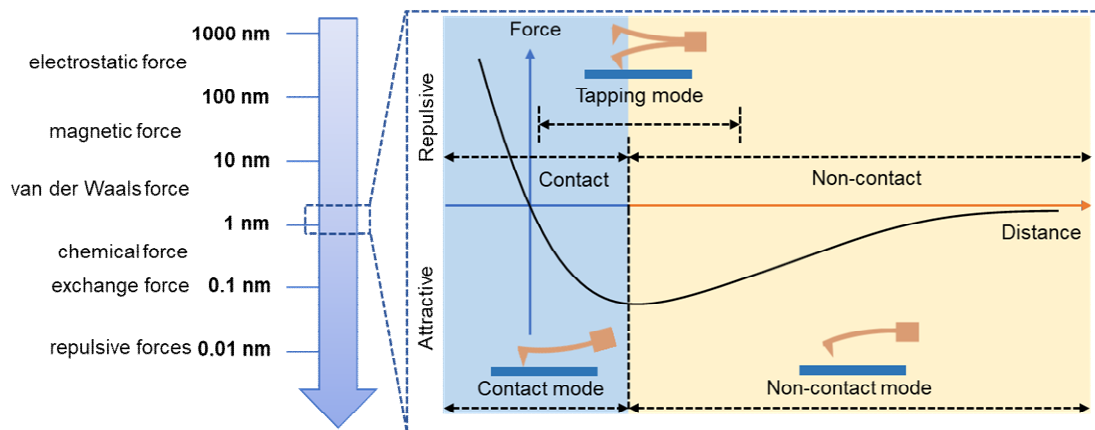

Fig. R1-1. Illustration of various interaction forces between a probe tip and a sample, where the dominant forces are dependent on the tip-sample distance. The tip-sample distance of $h < 1$ nm is the "vdW force interaction zone" where the vdW force is dominant.

To show the differences more clearly, we compare the utilized optical forces and probe working zones of our work and three works mentioned by the reviewer below:

Method	Utilized tip-sample optical force	Probe tip working zone (distance)	Realized functionality
LvFM (our work)	Light-modulated *dynamic* vdW force (thermal expansion is not necessary)	vdW force interaction zone (0~6 nm)	Nanoscale material composition discrimination
Lu, *et al.* Nature Photonics, 8(4), 307-312 (2014).	Photoinduced molecular expansion force	Non-contact mode zone (~30 nm)	Chemical constitution analysis by IR spectroscopy
Nowak, *et al.* Science Advances, 2(3), e1501571 (2016).	Photothermal expansion induced force	Contact mode zone (~ 0 nm)	Chemical constitution analysis by IR spectroscopy
Wang, *et al.* Nature Commun., 14(1), 7267 (2023).	Optical gradient force, photothermal force and photoacoustic force	Tapping mode with large amplitude (60 ~ 80 nm)	Ultrafast photothermal dynamics

According to the above discussion and thanks to the reviewer's question, we feel that "modulated" is a more appropriate term than "enhanced" to describe the principle of LvFM. Therefore, we have changed the full name of LvFM to "light-modulated van der Waals force microscopy" throughout the paper including the title.

Moreover, the LvF is caused by light-enhanced thermal motions of the molecules on the sample surface. Owing to the high sensitivity of vdW force with respect to the intermolecular distance change (see details in the theoretical derivation), the LvF is very sensitive to the small variation of molecular thermal motions. Therefore, even a small light excitation power without causing notable optothermal expansion of sample may lead to detectable LvF response, as can be seen in Fig. 2 and Fig. 4 (where the AFM topographies of the samples have no visible change under the excitation of a 633 nm laser of 1.2 mW, while the LvF responses are evident). In other words, *optothermal expansion of sample is not a prerequisite for LvF generation and detection*. This mechanism is different from other methods based on optothermal expansion. This is especially advantageous for the characterization of ultra-thin materials, such as single- or few-layer 2D materials, which have no obvious thermal expansion in the vertical direction.

As for the 20 dB enhancement, indeed it is the ratio between the measured signal with laser on and off under a specific excitation condition (here, 633 nm laser of 1.2 mW) on a specific sample (here, $MoTe_2$ film). As explained in the paper, since the LvF signal is weak and is obscured by other tip-sample forces and background noises, it is difficult to detect the LvF directly. To solve this problem, we propose to use a dual-modal cantilever in tapping mode and employed a sideband demodulation technique to extract the LvF signal. The 20 dB on-off contrast is therefore a manifestation of the signal-to-noise ratio (SNR) in this specific experiment. We agree with the reviewer that the 20 dB contrast cannot be a general FOM describing the performance of the LvFM method, which depends on sample material, excitation condition, instrument setup, etc. So, we have revised the statements in the manuscript.

The setups of combining optics with AFM have been adopted in many near-field methods, such as AFM-IR, TEPL, TERS, and PiFM, as mentioned in the paper. However, due to their different principles, the uses of light and the detected physical information are different. LvFM is a new method no matter in principle, instrumental realization, and functionality. More details are replied below to specific questions.

Revisions in the manuscript:

(1) We have changed the name of LvFM from "light-enhanced van der Waals force microscopy" to "light-modulated van der Waals force microscopy" throughout the paper including the title.

(2) To correct the "20 dB enhancement" statement, we have revised the text in the following places:

In Abstract: "Experiments of discriminating heterogeneous crystalline phases or compositions in typical materials with LvFM demonstrate not only a high compositional resolving capability represented by, for example, a 20 dB LvF

signal-to-noise ratio on a MoTe$_2$ film under the excitation of a 633 nm 1.2 mW laser, but also ...".

On page 9, "Moreover, the sensitivity of LvFM can be manifested by the signal-to-noise ratio (SNR) when the excitation laser is turn on or off. In the above experimental condition (a 633 nm laser excitation of 1.2 mW on the MoTe$_2$ film), the SNR of LvF is as high as 20 dB, as shown in Fig. 3c."

On page 11, in the caption of Fig. 3: "**c.** Detected LvF signals when the laser switches on or off, which shows a SNR of 20 dB."

On Page 13, in Summary: "This makes LvFM particularly attractive for at least three facts: (i) high compositional resolving capability represented by, for example, a 20 dB SNR on a MoTe$_2$ film under the excitation of a 633 nm laser of 1.2 mW ...".

(3) More rigorous theoretical descriptions of the LvFM principle have been added, including the following:

Pages 5-6, the presentation in the "**Light-modulated vdW force**" section is comprehensively rewritten to strengthen the completeness and clarity of the theoretical framework.

SI-5 is rewritten to show the theoretical derivation of LvF more rigorously and clearly. After revision, SI-2~SI-5 are sufficient supplementary materials to support the theoretical framework of the main text.

(4) Comprehensive experimental validation: we have conducted additional experiments to characterize five samples composed of various materials including organics, metals, metal oxides, inorganic compounds, etc. to show the wide material applicability and performance of LvFM. The experimental details and results were added to Supporting Information, SI-12~SI-16.

(5) A detailed comparison of LvFM with existing techniques has been given in the revised SI-1.

Here are my detailed comments:

1-1. The authors report a light-enhanced vdW force; however, the principle is not clearly illustrated. The authors should show the theoretical model of how this enhancement is achieved. The enhancement should be intrinsic to the material. If the enhancement is achieved through a modulated photothermal expansion that changes the tip-sample distance, it does not count as an enhanced vdW force. Such properties can be directly reflected by the phase of the AC attractive mode (non-contact mode) of AFM, without using optical force microscopy.

Response:
We thank the reviewer for the valuable comments. We are sorry for not explaining the principle of LvFM clearly. We have made our best effort to improve it in this revision.

As presented in the paper, a theoretical model has been established to describe and analyze the interaction force between an AFM tip and a sample surface when they are close to each other. To start with, we consider the interaction force between two closely

spaced molecules

$$f_{LJ} = 12\frac{\varepsilon}{r_0}(\frac{r_0^7}{r^7} - \frac{r_0^{13}}{r^{13}}) \,, \tag{1}$$

where the first term is an attractive vdW force and the second term is a repulsive force. Then the total interaction force between an AFM probe tip and a sample can be calculated by integration

$$F = \int n_1 dV_1(\int n_2 dV_2 \cdot f_{LJ}) = \int n_1 dV_1 \cdot f \,. \tag{2}$$

The detailed descriptions about Eq. (1) and Eq. (2) can be found in the manuscript.

Then we analyze the tip-sample force under light illumination, i.e., the LvF. Generally, the probe's equilibrium working position h is much higher than the repulsive force zone so that the total force F is an attractive force. In this case, to facilitate analysis, we can neglect the second repulsive force term in Eq. (1) and substitute it into Eq. (2), which yields

$$\begin{aligned} f &= n_2 \int_0^{+\infty} 2\pi\rho d\rho \int_h^{+\infty} \frac{z}{\sqrt{\rho^2 + z^2}} \frac{6\varepsilon r_0^6}{\left(\sqrt{\rho^2 + z^2}\right)^7} dz \\ &= \frac{\pi n_2 \varepsilon r_0^6}{12h^4} \end{aligned} \tag{3}$$

where ρ and z are coordinates as defined in Fig. 1a. In Eqs. (2) and (3), f and F are the vdW forces between a molecule of the tip and the whole sample and the total tip-sample force, respectively.

When light illuminates the sample surface, the thermal motions of the sample molecules are intensified. To describe the irregular thermal motion of a molecule located at (ρ, z) and at moment t, we define a random displacement vector $\mathbf{u}(\rho, z, t)$, as shown in the inset of Fig. 1(a). Taking this random molecular displacement into consideration, we can rewrite Eq. (3) as

$$\begin{aligned} f' &= n_2 \int_0^{+\infty} 2\pi\rho d\rho \int_h^{+\infty} \frac{z}{\sqrt{\rho^2 + z^2} + u(\rho,z,t)} \frac{6\varepsilon r_0^6}{\left(\sqrt{\rho^2 + z^2} + u(\rho,z,t)\right)^7} dz \\ &\Box \frac{\pi n_2 \varepsilon r_0^6}{12(h - \Box h)^4} \\ &\approx \frac{\pi n_2 \varepsilon r_0^6}{12h^4} + \frac{\pi n_2 \varepsilon r_0^6}{3h^5}\Box h \end{aligned} \tag{4}$$

Clearly, according to Eq. (4), when a sample molecule moves closer to the probe tip under light excitation, the vdW force f' will increase, due to the decrease of the tip-molecule distance to $h-\Delta h$. On the contrary, when the molecule moves farther away from the tip in the other half of its oscillation cycle, f' will decrease due to the increase of $h+\Delta h$. However, since f' has a high-order nonlinear dependence ($\propto h^{-4}$) on h, the integration of f' in one cycle of the molecule oscillation still leads to the increase of the overall tip-molecule vdW force, which can be validated by our numerical simulation later. Having obtained f', we can substitute Eq. (4) into Eq. (2) and calculate the total tip-sample vdW force F' by numerical integration. Since the probe tip in LvFM works

in tapping mode, the tip-sample forces vary periodically in accordance with the cantilever oscillation. So, we can calculate F or F' in a period and take their average values as the total forces.

Then the LvF enhancement is validated by Monte Carlo simulations, as detailedly described in the manuscript and shown in Figs. 1c and 1d. It is seen that the time-averaging F' is increased when a light excitation is applied and the enhancement factor is exponentially dependent on the incident light power (Fig. 1d). Since different materials may have different intrinsic intermolecular binding forces, they would have different LvF responses under the same light excitation. This is the key mechanism of LvFM for compositional discrimination of materials, which can be depicted more intuitively in Fig. R1-2 where two materials have different binding forces (sample 1> sample 2). When light illuminates the samples, the molecular thermal motion is enhanced so that $A_2>A_1$. Consequently, the tip-sample vdW forces $f_2>f_1$. For two samples with the same light excitation, $f_3>f_2$ because of $A_3>A_2$. This shows why the LvF is intrinsic to material and can be used to identify different materials, which are also validated by experimental results in Fig. 3 and Fig. 4.
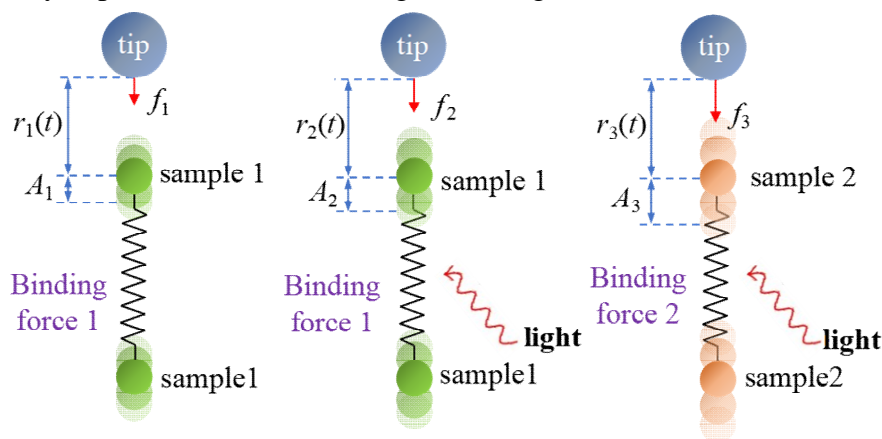

Fig. R1-2. Illustration of the intermolecular binding forces of two different materials with or without light excitation.

As regards the AFM phase imaging mentioned by the reviewer, it is based on the detection of viscous/adhesion force between a tip and a sample, which is totally different in principle from LvFM. Therefore, AFM phase imaging is only applicable to discriminate materials with evident difference in their tip-sample viscous/adhesion forces. In contrast, the vdW force is a more general interaction force between tip and sample and is more specific to materials because different materials have different intermolecular binding forces (see Fig. R1-2). This difference can be seen more clearly through the following two experiments. In Fig. R1-3, $MoTe_2$ films with 2H/1T' heterogeneous phases are characterized with both AFM phase imaging and LvFM. It is seen that AFM phase imaging can distinguish 2H phase and 1T' phase in the 5 nm-thick $MoTe_2$ film, but fails to do so for a thicker 10 nm $MoTe_2$ film. This means that the tip-sample viscous/adhesion force may be affected by thickness in ultra-thin materials such as 2D materials. In contrast, LvFM can still characterize the boundary between 2H phase and 1T' phase in the 10 nm thick film, because LvFM utilizes the vdW force between the tip and sample surface.
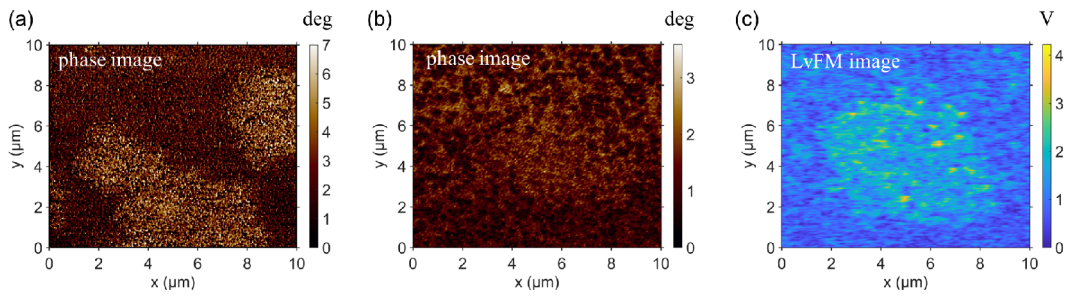
Fig. R1-3. Heterogeneous phase characterization of MoTe$_2$ films with 2H/1T' phases. (a) AFM phase image of a sample with 5 nm thickness; (b) AFM phase image of a sample with 10 nm thickness; (c) LvFM image of a sample with 10 nm thickness.

Figure R1-4 shows the experimental results of discriminating PS nanospheres and Au nanoparticles in a sample using both AFM phase imaging and LvFM. It is seen that when the PS NPs and Au NPs have the similar sizes, their AFM phase difference is too small (~4% difference) to distinguish them; while their LvF response difference is large enough (~42% difference) for discrimination. This shows again that LvFM has higher material compositional sensitivity than AFM phase imaging.
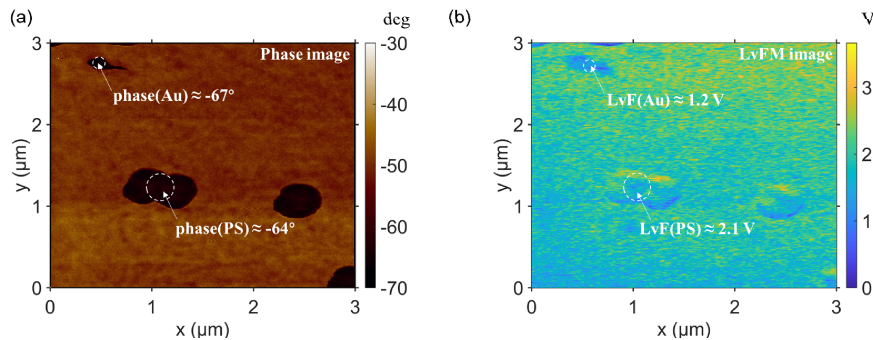

Fig. R1-4. Characterization of PS nanospheres and Au nanoparticles using (a) AFM phase imaging and (b) LvFM.

Moreover, LvFM has no special wavelength selectivity of the excitation light, only requiring that the light can intensify the thermal motions of sample molecules. This means that one can use the excitation light of various wavelengths in a wide band to characterize the LvF responses of a material but without the help of any optical spectrometer, as shown in the example in Fig. 4, SI-10, and SI-11. In this way, one may possibly get a LvF spectrum of the material with characteristic peaks related to vibrational and rotational absorption etc., just like the spectroscopic methods such as TERS, TEPL and AFM-IR can do, but without the use of spectrometer. Then, if we can build a LvF spectral library of various materials with characteristic peaks, we can identify the material directly with LvFM spectral sweeping. In this way, the LvFM can provide richer information of the material than AFM phase imaging. This would be an important task of our future work.

Revisions in the manuscript:

(1) More rigorous and detailed theoretical descriptions of the LvFM principle have

been added, including the following:

Pages 5-6, the presentation in the "**Light-modulated vdW force**" section is comprehensively rewritten to strengthen the completeness and clarity of the theoretical framework.

SI-5 is rewritten to show the theoretical derivation of LvF more rigorously and clearly. Especially, the mechanism of material discrimination based on LvF is explained in more detail. After revision, SI-2~SI-5 are sufficient supplementary materials to support the theoretical framework of the main text.

(2) The comparisons of characterization results of LvFM and AFM phase imaging have been provided in SI-12, SI-14, and SI-15.

1-2．The authors should provide the phase map of the topography to prove that light can change the intrinsic property of the material that is related to the vdW force.
<u>Response:</u>
Here may exist a misunderstanding. In LvFM, the excitation laser is not strong (~1 mW) so that it does not change the sample material property. As mentioned above, the use of excitation laser only intensifies the thermal motion of sample molecules, which is correlated to the light absorption in materials as measured in Figs. 4e and 4h.

As requested by the reviewer, we have also measured the AFM phase images of a $MoTe_2$ film before and after a 633 nm laser excitation, as shown in Fig. R1-5. Clearly, they are exactly the same, meaning that no viscous/adhesion force change is incurred under light illumination.
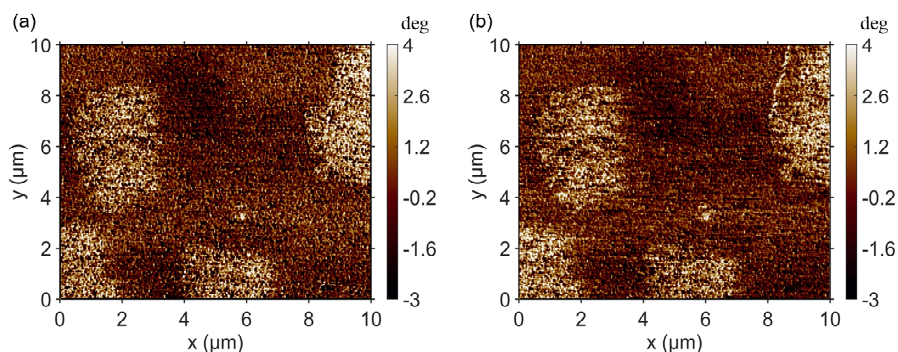

Fig. R1-5. AFM phase images of the $MoTe_2$ film (a) before and (b) after light illumination with a 633 nm laser of 1.2 mW.

1-3．vdW force varies dramatically with distance. The authors should justify why they use the tapping mode. At which distance is the vdW force measured?
<u>Response:</u>
We thank the reviewer for this important question. As explained above, the tapping mode is employed to extract weak LvF signal from the other tip-sample forces and background noises by sideband demodulation. To do this, the probe is installed on a dual-modal (with 1st- and 2nd-order resonance frequencies f_1 and f_2) cantilever working in tapping mode. The probe is driven by a dither piezo at frequency f_1, the excitation laser intensity is modulated at a frequency f_m, which satisfies $f_2 = |f_m \pm f_1|$, and then the

LvF signal can be demodulated at frequency f_2 with the best SNR.

As for the tip-sample distance, since the LvF is measured in tapping mode, the detected LvF signal is actually a time-averaged tip-sample interaction force, as mentioned on Page 4 of the paper. The probe must be launched in the close vicinity of the surface to ensure that it oscillates across the vdW force interaction zone during each vibration cycle, as explained in Fig. 1b. Therefore, in operation, we set the setpoint of AFM to 60%. As the free oscillation amplitude of the probe is ~10 nm, the actual amplitude is ~6 nm. It means that when the probe is launched on the sample surface, the tip-sample distance is in the range of 0~6 nm. Reasonable setpoint and oscillation amplitude of probe are the key factors to ensure that the LvFM can be well realized.

Revisions in the manuscript:
Revisions have been made to explain why the tapping mode is used:

On page 6, we added "Since the weak LvF signal is obscured by other tip-sample forces and background noises, it is difficult to detect the LvF directly. To solve this problem, we propose to use a dual-modal cantilever (with fundamental and second-order resonance frequencies f_1 and f_2) working in tapping mode and employed a sideband demodulation technique to extract the LvF signal, along with the topography mapping."

In SI-2, the tip-sample interaction forces in tapping mode are analyzed in detail.

In SI-5, thorough revisions have been made to analyze and calculate the LvF in tapping mode theoretically.

As for the working distance of the probe, the LvF is detected when the probe works in the vdW force interaction zone, as explained on page 5 and page 7.

1-4. The tapping mode of AFM involves a fixed engagement factor locked by the internal PID loop of the AFM system. Instead of showing the tip-sample force versus separation with a fixed free air amplitude (Fig. 1b), it may be more informative to show the force versus the engagement factor. How does the vdW force contribute to the curve?
Response:
We thank the reviewer for this comment. The engagement factor in AFM is defined as the ratio of the actual amplitude of probe after launching divided by its free amplitude. As can be seen from Fig. R1-6, the free amplitude of the probe is ~8.4 nm (meanwhile, h is close to 10 nm). The actual amplitude of the probe decreases when the probe is launched on the surface (and h decreases). So, the engagement factor can be calculated from the orange dashed line in Fig. R1-6(a) by taking the ratio of the actual amplitude to the free amplitude (8.4 nm) at different h, as shown in Fig. R1-6(b). The importance of Fig. R1-6 is that it correlates the engagement factor in the PID loop with the tip-sample force in the physical world. The above argument is under the precondition that the lowest position of the probe tip is on the sample surface during each vibration cycle of the probe. The tip-sample force in Fig. R1-6 consists of both the attractive vdW force and repulsive force. When h is large, the repulsive force is ineligibly weak, so that the tip-sample force is mainly vdW force. Therefore, the vdW force contributes to the right half of the curve.
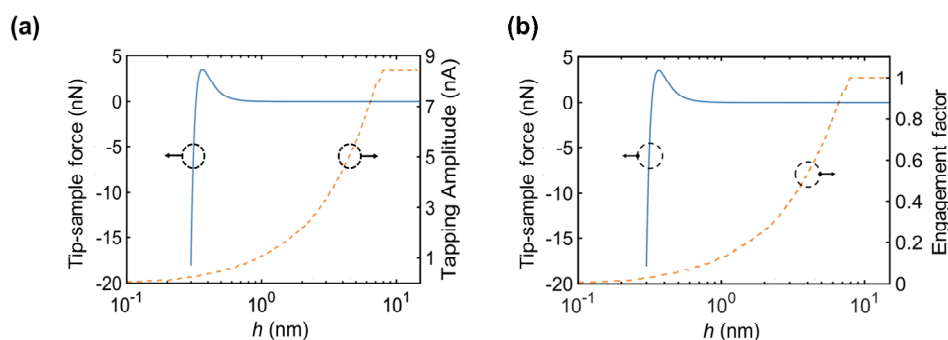
Fig. R1-6. Calculated total tip-sample force versus tip-sample distance h (blue). The orange lines show the h dependent (a) tapping amplitude and (b) engagement factor of the probe.

Revisions in the manuscript:
(1) We have accepted the reviewer's suggestion to change Fig. 1b to show the force versus the engagement factor.
(2) The contribution of the vdW force to the force curve have been discussed in more detail in the revision:

On page 3: "The repulsive bonding force F_B affects the engagement factor (oscillation amplitude) of an AFM cantilever in tapping mode (Fig. 1b) when it approaches the sample surface, which is utilized as a measure in LvF detection. More discussions about the tip-sample interaction forces are given in Supporting Information, SI-2."

On page 5: "Figure 1b shows the numerically calculated total tip-sample force F versus the tip-sample distance h as well as the engagement factor (tapping amplitude) of a probe with respect to h, by taking $\varepsilon = 1 \times 10^{-19}$ J, $r_0 = 0.3$ nm, and tip radius $R = 10$ nm. The h sensitive feature of the tip-sample force and its impact on the engagement factor of the probe in the vdW force interaction zone are clearly observed."

SI-3 and SI-5 are revised to give more details about the theoretical modeling of the tip-sample interaction force including the vdW force.

1-5．How is the setup illustrated in Fig. 2 different from the one of PiFM? If it is the same, why does the measured signal lead to additional vdW force information?
Response:
We thank the reviewer for this question. As stated in our paper, LvFM and PiFM are intrinsically different in principle, instrumental realization, and applications.

PiFM utilizes the *light induced dipole-dipole interaction force* between the tip and sample, which needs the polarization of the sample or the AFM tip. Therefore, PiFM has special requirement on sample material and excitation conditions, including the wavelength of excitation light, the polarizability of material (such as metals), etc. In setup, PiFM can use direct demodulation and sideband demodulation. The setpoint of PiFM is usually set at 85%[1], which corresponds to the *non-contact mode*.

In comparison, LvFM utilizes the *light-modulated vdW force* between tip and

sample, which is a universal force existing between any sample and the probe tip. Therefore, in principle, LvFM has no special requirement on sample material, provided that the thermal motion of molecules of the sample can be enhanced by light excitation. For this reason, LvFM has little requirement on the excitation light source. For example, a cw laser with moderate power is acceptable (such as the 633 nm laser of 1.2 mW used in our experiment) and the light wavelength can be arbitrarily chosen in a wide band, as demonstrated in Fig. 4, SI-10, and SI-11. Therefore, LvFM can be applied to a wide range of materials, including metals, dielectrics, semiconductors, and even ultra-thin materials such as 2D materials samples investigated in our experiments. According to these comparison, it is seen that LvFM can work in a wider spectral range, with better sample universality and simpler light excitation. As for the setup, since LvFM works in the vdW force interaction zone, it requires closer tip-sample distance than PiFM, so that a *tapping mode* detection scheme is adopted (the setpoint is usually set at ~60%) and a sideband demodulation technique based on a dual-modal cantilever must be used to extract the LvF signal.

See also the following paper [1] for the discussion of dipole and thermal effects in PiFM:

[1] Jahng, J. et al. Nanoscale spectroscopic studies of two different physical origins of the tip-enhanced force: dipole and thermal. arXiv preprint arXiv:1711.02479 (2017).

Revisions in the manuscript:
A detailed comparison of LvFM with other techniques including PiFM has been given in the revised SI-1. Especially, the differences in instrumental setup and application of LvFM and PiFM have been emphasized: "PiFM can use direct demodulation or sideband demodulation, where the setpoint is usually set at ~85%. In comparison, LvFM makes use of LvF, which is a universal force existing between any sample and the probe tip. Therefore, in principle, LvFM has no special requirement on sample material and excitation light, provided that the thermal motion of molecules of the sample can be enhanced by light excitation. In this sense, LvFM can be applied to a wider range of materials, including metals, dielectrics, semiconductors, and ultra-thin nanomaterials such as 2D materials, as will be shown in the following sections. Moreover, since LvFM works across the vdW force interaction zone, it requires a closer tip-sample distance than PiFM, so that the tapping mode is adopted with the setpoint usually set at ~60% and a sideband demodulation technique based on dual-modal cantilever is used to extract the LvF signal."

1-6．What does the white curve represent in Fig. 2g?
Response:
The white line shows the LvF signal magnitude along the red dashed line in Fig. 2g. We have revised Fig. 2g by adding the red indicative line to show it more clearly. It is also shown here as Fig. R1-7.
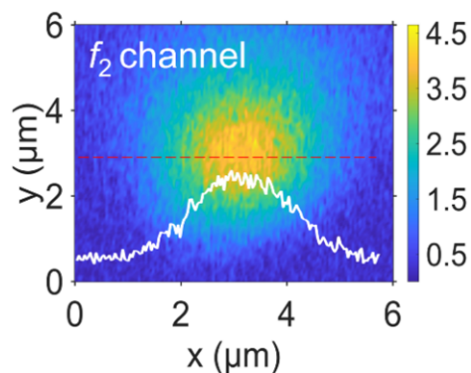
Fig. R1-7. LvFM image of an excitation laser of 1.2 mW at 633 nm focused on a smooth borosilicate glass substrate. The white curve is the line profile of LvF signal amplitude along the red dashed line.

Revisions in the manuscript:
We have revised Fig. 2g by adding the red indicative line to show it more clearly.

1-7．The optical force could result from the photothermal expansion of not only the sample but also of the substrate and the AFM probe. To exclude this effect and show the contribution of the 2D material, the authors should compare the signal in Fig. 3c with the one without the 2D material (only with the substrate).

Response:
We thank the reviewer for this question. First of all, as explained in the principle of LvFM, LvFM does not rely on the real photothermal expansion of material, but just needs the light intensified thermal vibration of molecules on sample surface. Owing to this, even a small or moderate light excitation power without causing optothermal expansion of sample may lead to notable LvF response, as can be seen in Fig. 2e and Fig. 4a (where the AFM topographies of the samples have no change under the excitation of a 633 nm laser of 1.2 mW, while the LvF responses are evident). This is quite different from other methods relying on thermal expansion of sample, such as AFM-IR.

Secondly, it is true that the thermal motions of molecules in the substrate and the tip are also intensified under light illumination. But even so, it does not prevent LvFM from discriminating materials. As required by the reviewer, we have measured the LvF responses on a bare sapphire substrate and on a $MoTe_2$ film on the sapphire substrate and compare them, as shown in Fig. R1-8. It is seen that the LvF responses from $MoTe_2$ film is 3.7 times of that from the bare substrate. This first shows that the LvF response of $MoTe_2$ is stronger than that of sapphire substrate, just because their molecular binding forces are different. This is a good example to show that LvFM can distinguish these two materials quite well. Consider the scanning process of LvFM, we know that when the probe scans over the sample, the influence of the planar uniform substrate and the probe to the LvF at different locations are always the same. Then the difference only comes from the sample (here, $MoTe_2$) itself. Since LvFM distinguishes the compositions of a sample based on the LvF difference or contrast, but not the absolute

value of LvF, the substrate and probe do not have any influence on the LvF imaging results. Furthermore, since the vdW force is very sensitive to the tip-sample separation, it mainly originates from the very top layer of sample surface. Then when the sample is placed on a substrate, the influence of substrate to the LvF results is also small, which at most just changes the baseline of the LvF image, while the contrast of the LvF image can still reflect the material composition difference.
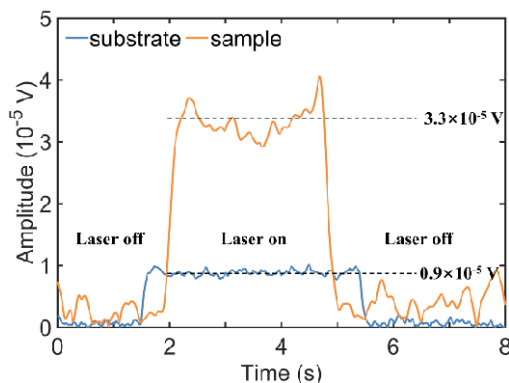

Fig. R1-8. Comparison of measured LvF amplitudes on a bare sapphire substrate (blue line) and a MoTe$_2$ film on the sapphire substrate (orange line).

Revisions in the manuscript:
We have made substantial revisions to clarify the principle of LvFM for composition discrimination. It is also emphasized that no thermal expansion of material is needed for the implementation of LvFM.

On page 7: "When a 633 nm laser of 1.2 mW illuminated the glass substrate surface, the topography was obtained via the f_1 channel (Fig. 2f), which was almost the same as Fig. 2e, meaning that light excitation did not cause visible expansion or deformation of the sample surface. However, evident LvF signal was probed via the f_2 channel (Fig. 2g) simultaneously. The distribution of the LvF mapping results highly resembled the Gaussian profile of the incident laser spot, showing its optothermal force nature instead of any physical protrusion."

On pages 9-10: the "**Spatial resolution and compositional resolving capability of LvFM**" section is thoroughly rewritten to describe and discuss the principle of LvFM for compositional discrimination in detail.

In the discussions of Fig. 2 and Fig. 4, the experimental evidences are provided to show that there is no thermal expansion of the samples.

1-8．The histogram of the optical force during a single engagement in Fig. 3e cannot effectively represent the standard derivation of the vdW force due to the low-pass filtering of the lock-in amplifier. What is the integration time of the lock-in amplifier?

Response:
We thank the reviewer for this question. Yes, the histograms of the measured LvF is associated with the time constant of the low-pass filter in the lock-in amplifier. Figure R1-9 shows the histograms for different time constants, ignoring the slight amplitude

drift. It can be observed that the time constant does not affect the mean value of the LvF amplitude. When the time constant is small (< 20 ms), its influence on the histogram is also small, because smaller time constant corresponds to higher filter cutoff frequency so that the LvF signal contains more high-frequency information. Thus, the output LvF is closer to the signal in the real environment. Moreover, the data of histograms in Fig. 3e is not the LvF amplitude in one vibration cycle, but thousands of times (depending on the sampling rate) of data collected in 10 seconds. The histogram is a representation of the quality of the LvF signal obtained during a certain time, which can reflect the stability of the LvFM system. The time constant is set to different values in different experiments, generally in the range of 3~21 ms, and the time constant of the LvF plot in Fig. 3e is 20.33 ms. But considering that these histograms may cause confusion and the time-dependent LvF signals in Fig. 3e are sufficient to show the stability of the detection, we have removed the histograms.
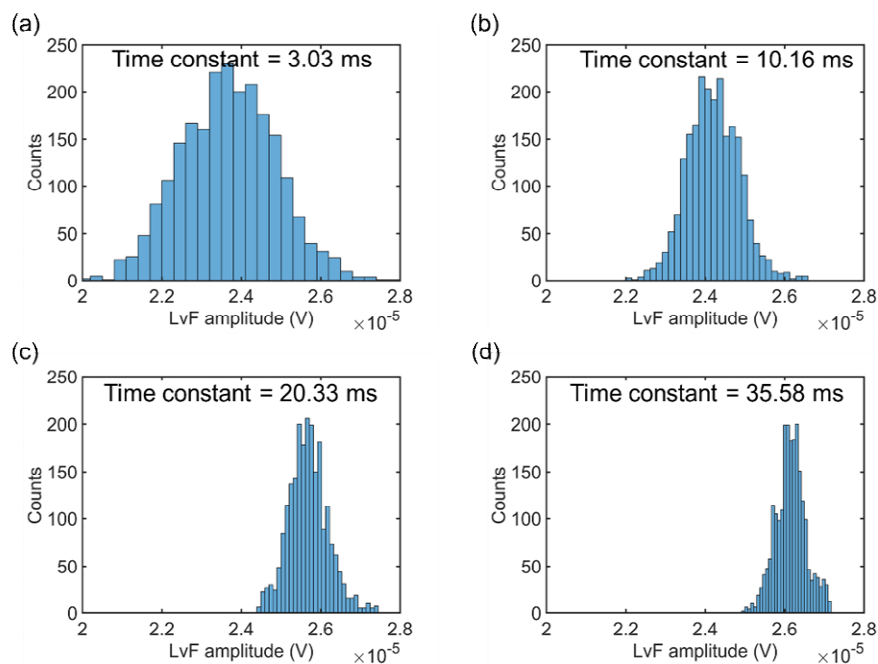

Fig. R1-9. Histograms of measured LvF amplitudes with time constants of (a) 3.03 ms, (b) 10.16 ms, (c) 20.33 ms, and (d) 35.58 ms in the low-pass filter of lock-in amplifier.

Revisions in the manuscript:
We have removed the histograms from Fig. 3e.

1-9．The signal measured at multiple engagements is recommended to eliminate the artifacts raised by mode hopping (between attractive mode and repulsive mode). How do the authors ensure the system operates in the attractive mode where the van der Waals interaction is dominant?
Response:
We thank the reviewer for this question. As mentioned above, the measured LvF response is the time-averaged LvF amplitude over many tapping cycles. So there is no artifact raised by mode hopping. According to the principle of LvFM, the probe is not

always in the vdW force interaction zone. In one tapping cycle, it is necessary that the tip goes through the vdW force interaction zone, which is ensured by loading the probe to the surface and controlling the setpoint.

In operation, there is a standard process to determine the vdW force interaction zone. According to the force-displacement curve of AFM in Fig. R1-10(a), the non-contact zone where the attractive force prevails, the bistable zone, and the contact zone where the repulsive force prevails are classified. The resonance peak of the cantilever in the free vibrational spectrum is noted as f_{res0}, and the corresponding amplitude is noted as A_0. When the probe approaches the sample surface, there are two cases. (1) At small tip-sample distance, the repulsive force affects the equivalent resonance frequency to a higher frequency f_{eff1}. If the driving force frequency is set to $f_{drive} < f_{res0}$, amplitude $A_1 < A_0$ can be realized after the tip approaches and reaches stabilization according to the setpoint, as shown in Fig. R1-10(b). This scenario corresponds to the tapping mode. (2) At larger tip-sample distance, the attractive force (namely, the vdW force) is dominant, which affects the equivalent resonance frequency to a lower frequency f_{eff2}. If the driving force frequency is set to $f_{drive} > f_{res0}$, amplitude $A_2 < A_0$ can be realized after the tip approaches and reaches stabilization according to the setpoint, as seen in Fig. R1-10(c). This scenario corresponds to the non-contact mode. Therefore, in practice, the AFM tip can be controlled in the right zone where the repulsive force or attractive vdW force is dominant by setting appropriate driving frequency f_{drive} and setpiont.
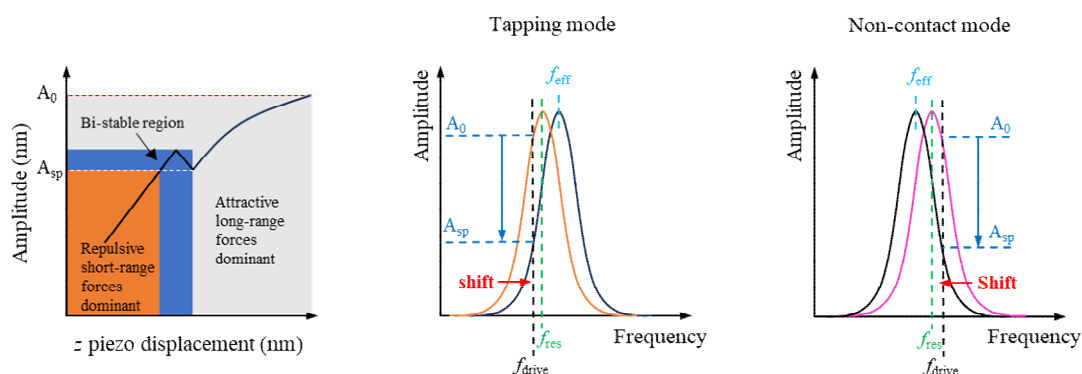

Fig. R1-10. (a) Schematic representation of the cantilever oscillation damping as a function of z piezo displacement (i.e., the tip-sample separation); (b) and (c) are schematic representations of the change of effective frequency f_{eff} of oscillation in tapping mode and non-contact mode, respectively.

Revisions in the manuscript:
In SI-1, we clarified: "Moreover, since LvFM works across the vdW force interaction zone, it requires a closer tip-sample distance than PiFM, so that the tapping mode is adopted with the setpoint usually set at ~60% and a sideband demodulation technique based on dual-modal cantilever is used to extract the LvF signal."

1-10．Why is the Raman microscopy and spectrum for the sample in Fig. 4g missing?
Response:

We thank the reviewer for this question. The Raman spectra of the h-BN/WS$_2$ sample have been obtained, as shown in Fig. R1-11, but were not added in Fig. 4 because of the limited space of figure. Now we have added the Raman spectra as well as the confocal Raman imaging results of the h-BN/WS$_2$ mixture in SI-11.
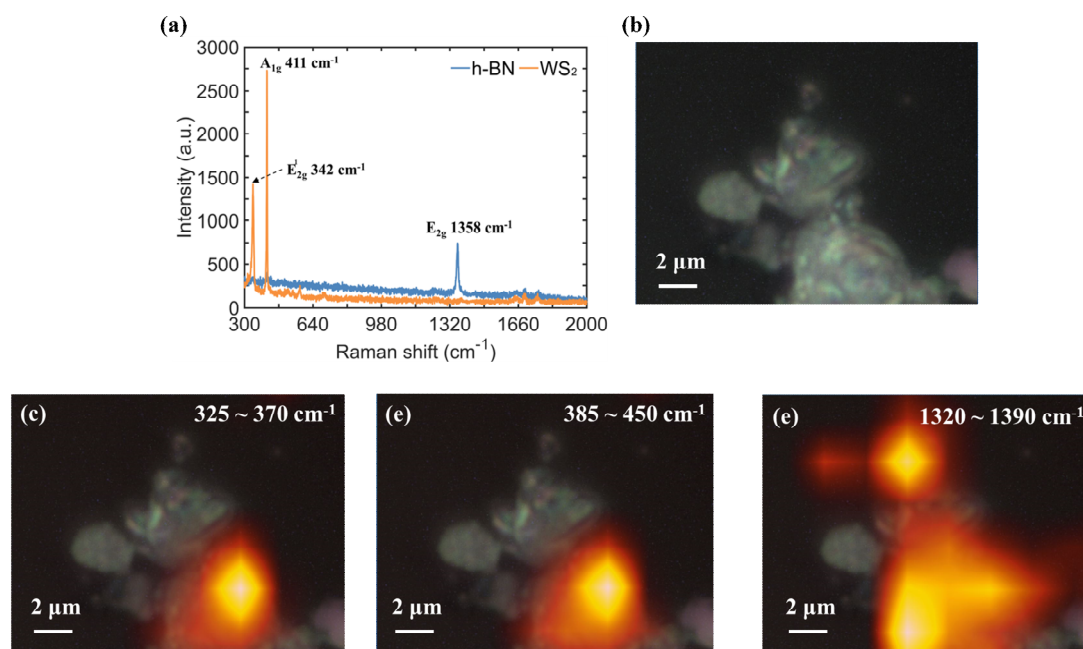

Fig. R1-11. (a) Raman spectra of the h-BN/WS$_2$ heteromaterial sample investigated in Fig. 4. (b) Optical microscopic image of the sample. (c), (d) and (e) are the overlaid images of the optical image of the sample and its confocal Raman mapping results obtained in the 325~370 cm^{-1}, 385~450 cm^{-1}, and 1320~1390 cm^{-1} integration intervals, respectively.

Revisions in the manuscript:
We have added the Raman spectra as well as the confocal Raman imaging results of the h-BN/WS$_2$ mixture in SI-11.

Reviewer #2
The authors developed the Light-enhanced van der Waals force microscopy (LvFm), facilitating compositional identification of materials. This equipment can acquire information about the composition by analyzing the optothermal interaction between the tip and the sample. In contrast to AFM, which is unable to provide compositional data, the LvFm enables the determination of both the thickness and composition of the sample. The operational principles of this equipment are explained in details within the manuscript, highlighting its distinctive features compared to other AFM-IR, Raman and XRD.

However, given that the primary driving force behind the LvFm lies in van der Waals interaction, a more compelling demonstration of its advantages would be achieved through the inclusion of diverse samples and experiments. The authors dedication to providing a thorough explanation of the LvFm's principles and distinctions is commendable. Expanding the various samples and experiments would

16

undoubtedly enhance the credibility and appeal of this approach.

The utilization of IR and Raman, despite their comparatively lower resolution, is grounded in the certainty of vibrational modes associated with chemical bonding. A well-established vibration library, tailored to specific molecules, underscores the reliability of these methods.

In contrast, Van der Waals forces, being the weakest among intermolecular forces, introduce variability in force interactions, demanding careful consideration of sample quality. Even when measuring the same sample, the forces may exhibit differences. A more in-depth discussion should be needed. Therefore, this manuscript can be accepted after conducting additional experiments and explaining any unclear aspects.

Response:

We are very grateful to the reviewer for the very valuable comments and suggestions. We have followed the reviewer's suggestion to conduct additional experiments to characterize various types of samples (including organics, metals, metal oxides, inorganic non-metallic compounds, etc.) using LvFM, to show the wide applicability and good performance of LvFM. We have also improved the explanation of all the unclear aspects that the reviewer has pointed out, as detailed below.

Moreover, as mentioned above, LvFM has no special wavelength selectivity of the excitation light, only requiring that the light can intensify the thermal motions of sample molecules. This means that one can use the excitation light of various wavelengths in a wide band to characterize the LvF responses of a material but without the help of any optical spectrometer, as shown in the example in Fig. 4, SI-10, and SI-11. In this way, one may possibly get a LvF spectrum of the material with characteristic peaks related to vibrational and rotational absorption etc., just like the spectroscopic methods such as TERS, TEPL and AFM-IR can do, but without the use of spectrometer. Then, if we can built a LvF spectral library of various materials with characteristic peaks, we can identify the material directly with LvFM spectral sweeping. In this way, the LvFM can provide richer information of the material than AFM phase imaging. This would be an important task of our future work.

Revisions in the manuscript:

(1) We have carried out additional experiments on various types of samples including organics, metals, metal oxides, inorganic non-metallic compounds, etc. to show the effectiveness and good performance of LvFM for a wide variety of materials. The details and results of these experiments are added to SI-12~SI-16.

(2) More theoretical descriptions of the LvFM principle have been added, including the following:

Pages 5-6, the presentation in the "**Light-modulated vdW force**" section is comprehensively rewritten to strengthen the completeness and clarity of the theoretical framework.

SI-5 is rewritten to show the theoretical derivation of LvF more rigorously and clearly. After revision, SI-2~SI-5 are sufficient supplementary materials to support the theoretical framework of the main text.

(3) A detailed comparison of LvFM with other techniques has been given in the revised SI-1.

2-1. The authors presented an example involving MoTe₂ to demonstrate the potential of compositional indentification using LvFm. However, this aspect appears somewhat ambiquous. It aligns more closely with phase identification of TMDs rather than explicit compositional identification. To establish the capacity of LvFm for compositional identification, it is imperative to show additional examples such as organic compounds (polymer films), metal alloys, metal oxides, and other diverse materials.

Response:

We thank the reviewer for this important suggestion. We have followed the reviewer's suggestion to perform more LvFM characterization experiments on various types of materials, including organics, metals, metal oxides, inorganic compounds, etc. to show the applicability of LvFM and the diversity of applicable materials. The additional experiments are detailed below. Due to the limited space of the manuscript, these new details and results have been added to SI-12~SI-16.

Experiment 1: Characterization of Au nanoparticles and PS nanospheres.

We deal with a mixture of Au NPs and polystyrene NPs with similar diameters ~100 nm, which are randomly dispersed onto the surface of a glass substrate. Both AFM and LvFM were employed to characterize the NP ensembles. The AFM topography image, LvF image, and AFM phase image are shown in Figs. R2-1(a), R2-1(b), and R2-1(c), respectively. Clearly, the two types of NPs cannot be distinguished by AFM topography. Surprisingly, they cannot be distinguished by AFM phase imaging either, whose phase difference is only ~4%, meaning that the viscous/ adhesion force between the tip and the two kinds of NPs are quite similar. In contrast, in LvFM imaging, the two types of NPs can be clearly distinguished, whose LvF response difference is as large as 42%. To cross-check that the NPs at positions ① and ② are indeed different kinds of NPs, we have used micro-photoluminescence (PL) spectroscopy to get their PL spectra, as shown in Fig. R2-1(d), from which Au NP and PS NP can be identified.

Moreover, to check the repeatability of the measurement, we have characterized multiple NPs in the sample with LvFM and make statistics of the LvF versus particle sizes and types, as shown in Fig. R2-1(e). Clearly, the data points are divided into two distinct groups, one for Au NPs with lower LvF response and the other for PS NPs with stronger LvF response. The LvF dependence on the particle size is not evident. This experiment is a good example of using LvFM to distinguish metal and polymer.
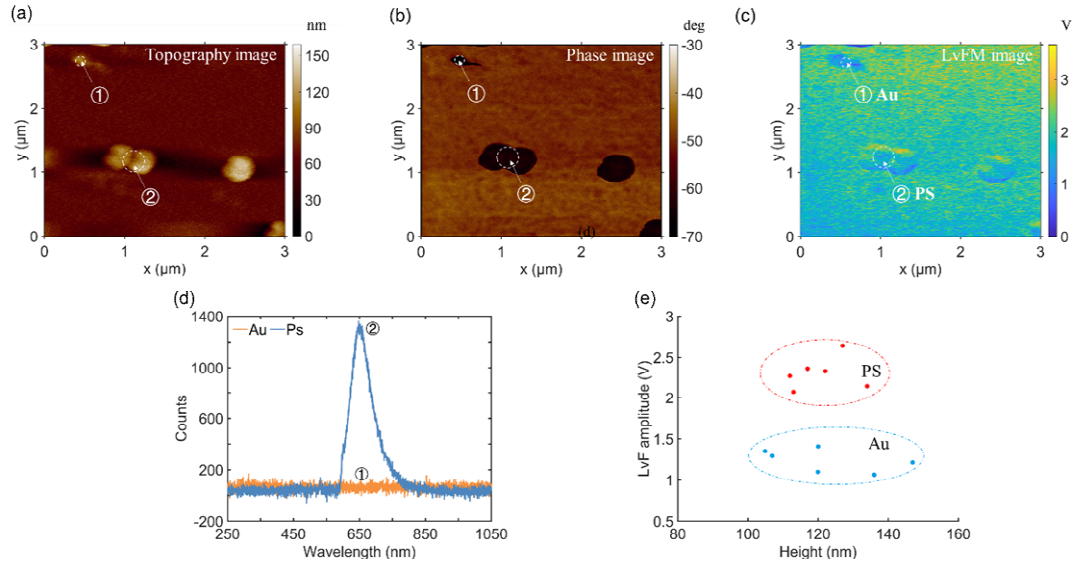
Fig. R2-1. (a) AFM topography image, (b) LvFM image, and (c) AFM phase image of the Au NPs and PS NPs randomly dispersed on the surface of a glass substrate. (d) PL spectra of the NPs taken from positions ① and ②. (e) Comparison of the measured LvF signal strengths of the two types of NPs.

Experiment 2: Characterization of organic compounds.

We characterize photoresists, a class of light-sensitive organic compounds. For this purpose, a heterogeneous structure containing two types of photoresists, AZ4620 (a positive photoresist) and SU8 (a negative photoresist) was prepared, as shown in Fig. R2-2(a). In this structure, SU8 is filled in a small volume of $12×12×7$ μm³ in the AZ4620 layer, as seen in the optical microscopic image in Fig. R2-2(b). AFM topography imaging and LvFM image results of the sample are shown in Figs. R2-2(c) and R2-2(d), respectively. Clearly, while the topography is rather flat, one can still distinguish the two photoresist areas according to their evident LvF response contrast. It is worth noting that in this experiment, a 532 nm laser of 1.1 mW was employed. Although the vibrational absorption energy levels of the organics are mainly in infrared band, the LvF response under visible light excitation is strong enough for material discrimination.
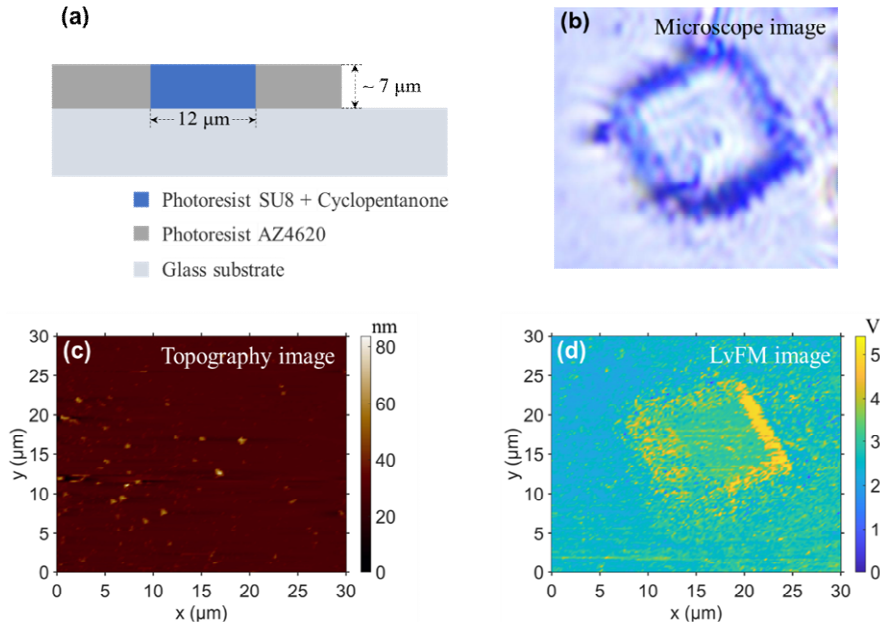
Fig. R2-2. (a) Schematic of the heterogeneous structure containing AZ4620 and SU8 photoresists. (b) Optical microscopic image, (c) AFM topography image, and (d) LvFM image of the sample.

Experiment 3: Characterization of metal oxides.

Ho_2O_3 and Al_2O_3 films were prepared and characterized by LvFM using a 633 nm excitation light. The scanning areas are across the edges of the films. The AFM topography, AFM phase, and LvFM image of the Ho_2O_3 film are shown in Figs. R2-3(a) ~ R2-3(c), respectively. According to the AFM topography, the Ho_2O_3 film is ~140 nm thick and there is a stacking area of ~2 μm wide in the Ho_2O_3 film. In the AFM phase image, the contrast of different areas is rather low. But in LvFM image, the LvF difference/contrast between Ho_2O_3 and the substrate is very clear. Moreover, the LvF strength is dependent on the thickness of the film: in thicker film area, the LvF amplitude is smaller, due to the strong absorption of Ho_2O_3 to visible light when it passes through the film from the bottom side (note that ~640 nm is one of the absorption bands of holmium ions). This shows that LvFM can not only discriminate different materials, but also has the potential to resolve the thickness difference of the sample material.

The same characterization procedure has been conducted on a Al_2O_3 film, as shown in R2-3(d) ~ R2-3(f). The major difference from the Ho_2O_3 film results is that the LvFM image contrast is lower, because the LvF response difference between Al_2O_3 film and the quartz substrate are not so large as that between Ho_2O_3 and quartz.
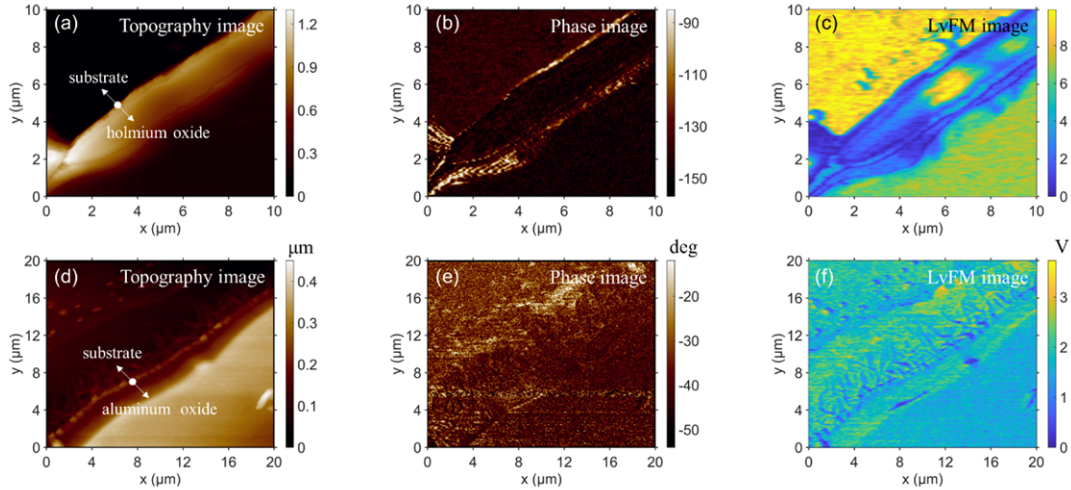
Fig. R2-3. (a) AFM topography, (b) AFM phase, and (c) LvFM images of a Ho_2O_3 film. (d) AFM topography, (e) AFM phase, and (f) LvFM images of a Al_2O_3 film.

Experiment 4: Characterization of single walled carbon nanotubes (SWCNTs).

The electronic properties of SWCNTs strongly depend on their chiral structures, which could be semiconducting or metallic with quite different electronic state densities and optical properties. Therefore, we have tried to use LvFM to identify these two types of SWCNTs. By carrying out a large number of experiments, we have found that the LvF responses of monodispersed SWCNTs with similar tube diameters may differ from each other significantly. Fig. R2-4 shows an example of the LvFM image of two SWCNTs under the same light excitation. Clearly, the LvF responses of SWCNTs ① and ② are quite different, whose difference is larger than both AFM topography and AFM phase. But due to limited time, we have not finished the thorough investigation on SWCNTs yet. We have not characterized the chirality of these SWCNTs by other means such as TEM, STM, and Raman. The connection of LvF response of specific SWCNT to its chirality property has not been established. But the obtained results already show that LvFM has the potential to distinguish different SWCNTs, which will be an important research task of us in the future.
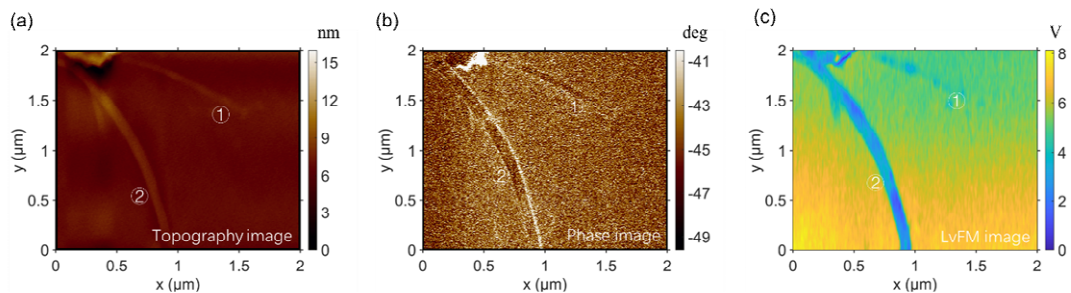

Fig. R2-4. (a) AFM topography, (b) LvFM image, and (c) AFM phase image of two SWCNTs.

Experiment 5: Characterization of nanodefects in van der Waals materials.

The mixture of two types of multilayered van der Waals materials prepared by

mechanical exfoliation, GaS and NbS$_2$, is characterized by LvFM. The AFM topography and LvF image of the sample are shown in Fig. R2-5. It is seen that nanodefects can be probed by LvFM, some of which can also be observed in AFM topography while some others cannot if they exist in flat area. Even for humps, some have no LvF contrast with the surrounding material (meaning their compositions the same), while some others have large LvF contrast (indicating the heterogenous materials). The detailed study of the types and forms of nanodefects with LvFM needs an in-depth investigation, which will be our work in the future. Nevertheless, these preliminary results already show that LvFM is a promising tool for finding and identifying nanodefects such as bubbles, wrinkles, or other impurities in van der Waals materials.
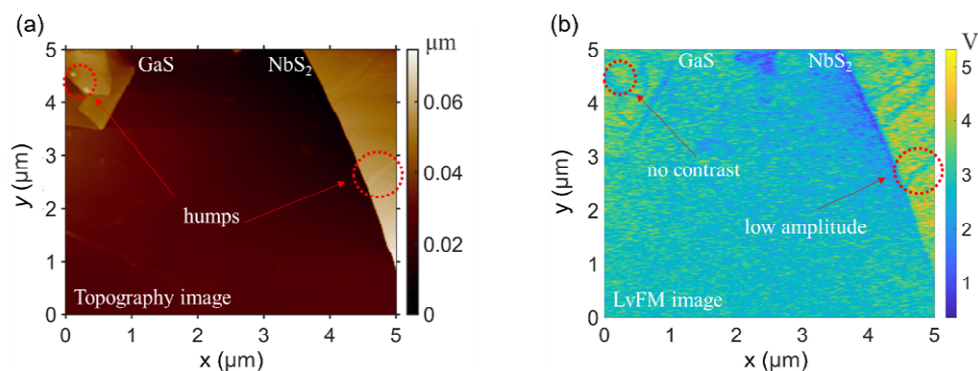

Fig. R2-5. (a) Topography and the (b) LvFM image of GaS and NbS$_2$ van der Waals multilayered materials

Finally, we should emphasize that LvFM is different from methods like EDS or XRD for atomic scale material compositional analysis. EDS is for elemental analysis with high spatial resolution, while XRD is for determining the crystallographic structure of materials. In contrast, LvFM works on a bulkier scale, which can distinguish the composition of different materials, but not the elemental composition or crystallographic structures of materials. For this reason, it cannot characterize the composition of, e.g., metal alloy whose inhomogeneity exists only on atomic scale.

Revisions in the manuscript:
We have added the details and results of the additional experiments to SI-12~SI-16 and mentioned these additions in the manuscript.

2-2.   The distinction between the 2H and 1T' phases of MoTe$_2$ achieved through LvFm lacks a precise explanation of the underlying principle. Notably, the LvFm amplitude exhibits strength in the semiconducting 2H region and weakness in the metallic 1T' phase. The absence of a detailed explanation raises questions about the feasibility of compositional identification with LvFm, especially in the case of metallic materials.

Response:
We thank the reviewer for this question. As explained in the revised principle part as

well as the above replies related to the principle of LvFM, the LvF response is positively correlated with the light intensified thermal motions of molecules on the sample surface. In our experiment, the LvFM works in transmission mode (see Fig. 2a and Methods), i.e., the samples were illuminated from the bottom side. In this case, the laser beam has to pass through the sample to reach the upper surface that has interaction with the probe tip. For visible light excitation, the absorption coefficient of 1T' phase is larger than that of 2H phase (as evidenced by the measured absorbance in Fig. 4e). This means that more light energy is lost in the process of passing through the 1T' phase area than the 2H phase area, so that the excitation light intensity on the 2H phase surface is larger than that on the 1T' phase surface. This causes the higher LvF response in 2H phase area than that in 1T' phase area. Anyway, provided that the LvF responses of two materials are distinguishably large, they can be discriminated.

<u>Revisions in the manuscript:</u>

Detailed discussions have been added to the analysis of the MoTe$_2$ characterization results, including:

On page 9: "Then, let us experimentally determine the spatial resolution and compositional resolving capability of LvFM by dealing with, e.g., a MoTe$_2$ film containing two crystalline phases 2H and 1T'. Even though the spatial resolution of LvFM is very high, the two phases of MoTe$_2$ cannot be distinguished without light excitation, since their static vdW force responses are almost the same (Fig. S4). Nevertheless, when an excitation light is applied, the magnitudes of LvF responses of the two phases are both enhanced and exhibit remarkable difference (Fig. S4), thanks to the distinct binding forces of molecules in the two phases. This is schematically illustrated in Fig. 3a when the tip scans across the boundary of the two phases. The measured LvF response during this scan is shown in Fig. 3b, from which the practical spatial resolution of LvFM can be estimated to be smaller than 10 nm, which is close to the theoretical estimate above. Note that the spatial resolution of AFM is restrained by deconvolution of tip shape and surface profile, and the spatial resolution of AFM-IR is exaggerated by material thermal diffusion length, while LvFM demonstrates for the first time a lateral resolution much smaller than the tip size. Moreover, the sensitivity of LvFM can be manifested by the signal-to-noise ratio (SNR) when the excitation laser is turn on or off. In the above experimental condition (a 633 nm laser excitation of 1.2 mW on the MoTe$_2$ film), the SNR of LvF is as high as 20 dB, as shown in Fig. 3c.

Such a high sensitivity at low excitation power of light is a special advantage of LvFM, which are related to material bonding property, amount of boosted bonds, light intensity, and material absorbance. To investigate these dependences, we have measured the LvF responses versus incident light power for six samples with different combinations of MoTe$_2$ crystalline phases (2H or 1T'), film thicknesses (10 nm or 5 nm), and probe tip materials (Au or Si), as shown in Fig. 3d. Clearly, when the laser power is above a threshold (~0.6 mW in our experiment), the LvF signals can be detected and increase monotonically with the increase of the incident power, which are in good agreement with the theoretical prediction in Fig. 1d. In addition, under the same light power and with the same film thickness, the 2H phase has stronger LvF response than 1T' phase, because the absorption coefficient of the metallic 1T' phase is larger so that the light transmitting through it from the bottom and arriving at the top surface is weaker. Hence, the light-enhanced thermal motions of the 1T' phase molecules are weaker. For the same phase of MoTe$_2$, thicker film produces stronger LvF, because

more light energy is absorbed to boost more bonds. Furthermore, when an Au-coated probe is used instead of the Si probe, larger LvF signal can be collected, which can be explained by the fact that the Hamaker constant of Au is larger than that of Si so that the tip-sample LvF interaction is stronger. All these observations validate the LvFM principle and demonstrate its excellent sensitivity and performance."

In SI-10: LvFM characterization of $MoTe_2$ under different excitation wavelengths is reported, which demonstrates the 2H and 1T' phase discrimination capability of LvFM in a wide band of 500 nm ~ 900 nm.

2-3.   What is the actual thickness of $MoTe_2$ in Figure 4? How does the LvFm data change based on the thickness? In the case of 2D materials, Raman data typically shifts by about 1~3 cm^{-1}, according to the thickness. How does LvFm compare to the sensitivity of Raman? Given that LvFm is significantly influenced by van der Waals forces, there appears to be substantial variation based on how the sample is prepared. How would the authors explain this?

Response:

We thank the reviewer for these questions and answer them one by one below.

(1) The thickness of $MoTe_2$ film in Fig. 4 was measured to be 12.3 nm by using an ellipsometer (V-VASE, J.A.Woollam Co.,Inc).

(2) The LvF response is indeed affected by the sample thickness in transmission mode, as can be observed in our experimental results in Fig. 3d, because the excitation light should go through the sample first and then arrive at the upper surface to affect the tip-sample interaction. This has been explained further in the revised manuscript on pages 9-10 when discussing the six samples with different configurations. But the dependence is complicated, which is specific to material. For example, for metals, their strong absorption makes the LvF negatively correlated with the thickness of sample, i.e., the thicker the film, the smaller the LvF response detected. To show this, we have prepared different thicknesses of Au and Pt films by e-beam evaporation and experimentally measured their LvF responses, as shown in Fig. R2-6 (a). Clearly, we can see the monotonic decline of LvF response with the increase of film thickness for both metals. This also means that it is possible to measure the thickness of the film thickness by LvFM. However, for some other materials, such as multilayered WS_2, the thickness dependence of LvF is not so simple. As shown in Fig. R2-6 (b), with the increase of WS_2 film thickness, the LvF response goes up first and then drops down. The mechanism of this behavior is unclear so far. But it implies that there is an optimum thickness of the WS_2 film for LvF characterization in transmission mode. Moreover, if the LvFM is operated in reflection mode, the sample thickness dependence should be loosened since the light does not need to go through the sample anymore. But this needs to change our instrumental system and needs a lot of work, which will be a task of our future work.
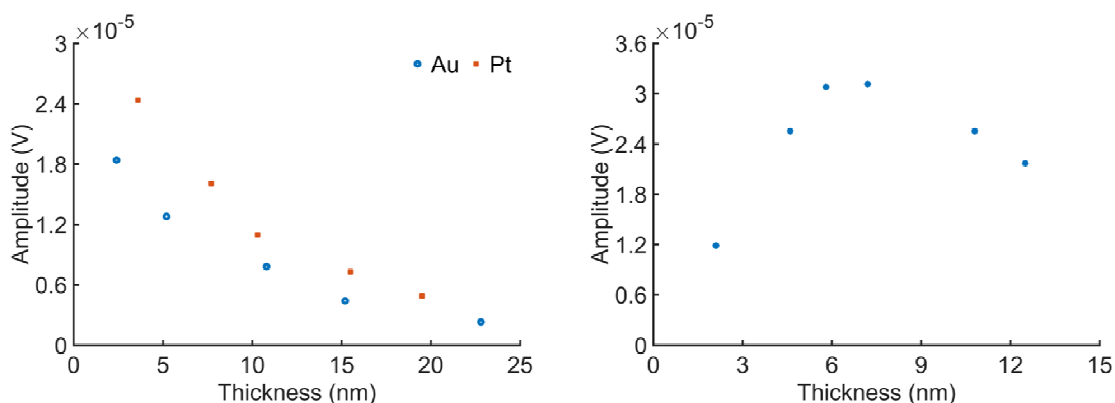
Fig. R2-6 (a) LvF amplitude versus the thicknesses of Au and Pt films. (b) LvF amplitude versus the thickness of multilayered WS_2 film.

(3) Our primary goal of proposing LvFM is for discriminating material composition. Although it has shown the potential for film thickness determination, we have not carried out a thorough investigation on it. But as discussed above, according to the preliminary experimental results on metal films, the LvF response bears a monotonic relation with the thickness, according to which the LvF sensitivity on thicknesses of Au and Pt films can be estimated to be around $-0.78\,\mu V/nm$ and $-1.22\,\mu V/nm$, respectively, according to Fig. R2-6 (a). It is hard to compare directly with Raman, since Raman uses peak shift to express the sensitivity.

(4) As explained in the principle part, LvFM works based on the tip-sample dynamic LvF, not the static vdW force. Therefore, under the same light excitation condition, the major factor influences the contrast of LvF responses of different materials is their intrinsic properties, i.e., the binding forces of their composing molecules. It is based on this principle that one can realize compositional identification with LvFM. In other words, LvFM characterizes the LvF variation over a sample surface, but not detects the absolute value of LvF. Moreover, the measured LvF mapping results are analyzed together with topography to discriminate the material composition and properties in a sample, as we did in all the experimental examples in the manuscript and SI.

Revisions in the manuscript:
(1) On page 11: "The first experiment is to distinguish the 2H and 1T' crystalline phases in a $MoTe_2$ film of 12 nm thickness, the same sample as that characterized in Fig. 3, under the excitation of a 633 nm laser of 1.2 mW."
(2) On pages 9-10: the LvF response dependence on sample thickness is discussed regarding the six samples with different configurations. "To investigate these dependences, we have measured the LvF responses versus incident light power for six samples with different combinations of $MoTe_2$ crystalline phases (2H or 1T'), film thicknesses (10 nm or 5 nm), and probe tip materials (Au or Si), as shown in Fig. 3d. Clearly, when the laser power is above a threshold (~0.6 mW in our experiment), the LvF signals can be detected and increase monotonically with the increase of the incident power, which are in good agreement with the theoretical prediction in Fig. 1d. In addition, under the same light power and with the same

25

film thickness, the 2H phase has stronger LvF response than 1T' phase, because the absorption coefficient of the metallic 1T' phase is larger so that the light transmitting through it from the bottom and arriving at the top surface is weaker. Hence, the light-enhanced thermal motions of the 1T' phase molecules are weaker. For the same phase of $MoTe_2$, thicker film produces stronger LvF, because more light energy is absorbed to boost more bonds. Furthermore, when an Au-coated probe is used instead of the Si probe, larger LvF signal can be collected, which can be explained by the fact that the Hamaker constant of Au is larger than that of Si so that the tip-sample LvF interaction is stronger."

Reviewer #3

The manuscript introduces and describes with important details a novel approach to easily discriminate 2D materials by combining AFM topography with light illumination. The authors present a very sound description of the phenomenon on the basis of enhanced van der Waals interaction between the AFM tip and the illuminated 2D material. They motivate well the importance of the technique by comparing it to other widely used technics. They also describe the way to implement it experimentally. They finally demonstrate its potential on 2 examples of 2D systems: 1T'/2H phases in MoTe2 and a hBN/WS2 stack, where the 2 materials can be discriminated thanks to the introduced technique. Overall the manuscript is of importance to the 2D materials community, and maybe to the larger surface science audience. It is well organized (to the most part, see issues below) and follows the standards for publication in your journal. I would support publication given that the following issues are answered:

Response:

We are very grateful to the reviewer for the positive comments and recommendations. We have taken this opportunity to improve our manuscript carefully by taking into account all the review comments and suggestions, as detailed below.

3-1.   My main concern relates to the figures of merit introduced in the text. It is of high importance since they are also reported in the abstract. First, the 20dB enhancement factor is material dependent (here obtained on MoTe2). This is not stated in the abstract, and not clearly mentioned in the text. It is also not relevant to illustrate the discriminating power of the technique: if 2 materials have the same enhancement power, no contrast will be observed. I also have concerns on what is called "composition resolution". It appears to me simply as the signal contrast (or a derived way to express it). It is actually referred as is (i.e. U_H/U_T) in SI 10 and in figure 4f. It is also material dependent. Finally, the authors introduce a detection limit DL with the use of delta_1 and delta_2 which are nowhere defined. This DL is then used to normalize a contrast and define a resolving power. This gives an impressive number that is to me suspicious, especially due to its mentioning in the abstract with less context and definition. Overall, this aspect on figures of merit requires to be revised with a much clearer definition of the extracted figures, and with their report in the abstract only if they are relevant and with additional context (material dependence character at

Response:

We thank the reviewer for giving these important comments, which we all accept. We are sorry for not presenting these figures of merit (FOMs) rigorously. Now we have modified the related expressions.

First of all, as we replied to Reviewer #1, the 20 dB enhancement is indeed a ratio between the measured LvF signals with laser on and off under a specific incident condition one a specific sample (here $MoTe_2$). This 20 dB on-off contrast is just an example showing the good SNR in a specific characterization, but cannot be a general description of the resolving power. We have clarified this in the revised manuscript.

The term "composition resolution" is also not rigorous. We have removed it and, alternatively, express the ability of compositional discrimination directly by the measured LvF contrast.

As for the detection limit, we initially want to use it as a parameter to describe the minimum detectable LvF difference in experiment. However, indeed as the reviewer commented, this DL is also dependent on instrument setup, sample property, noise level, etc. So, it also cannot be a general FOM to describe the LvFM method. In the revised manuscript, we have also removed this DL.

Revisions in the manuscript:

(1) To correct the "20 dB enhancement" statement, we have revised the text in the following places:

In Abstract: "Experiments of discriminating heterogeneous crystalline phases or compositions in typical materials with LvFM demonstrate not only a high compositional resolving capability represented by, for example, a 20 dB LvF signal-to-noise ratio on a $MoTe_2$ film under the excitation of a 633 nm 1.2 mW laser, but also ...".

On page 9, "Moreover, the sensitivity of LvFM can be manifested by the signal-to-noise ratio (SNR) when the excitation laser is turn on or off. In the above experimental condition (a 633 nm laser excitation of 1.2 mW on the $MoTe_2$ film), the SNR of LvF is as high as 20 dB, as shown in Fig. 3c."

On page 11, in the caption of Fig. 3: "**c.** Detected LvF signals when the laser switches on or off, which shows a SNR of 20 dB."

On Page 13, in Summary: "This makes LvFM particularly attractive for at least three facts: (i) high compositional resolving capability represented by, for example, a 20 dB SNR on a $MoTe_2$ film under the excitation of a 633 nm laser of 1.2 mW ...".

(2) We have deleted the term "composition resolution" throughout the paper and, instead, express the ability of compositional discrimination directly by the measured LvF contrast.

(3) we have also removed the statements and discussions about the detection limit DL throughout the paper.

Response:

We thank the reviewer for this important suggestion. In **Methods** section of the paper, the preparation of the MoTe$_2$ sample is described in detail. But these information is not enough. We have revised the paper by adding more detailed information.

Revisions in the manuscript:

(1) In the "**Spatial resolution and compositional resolving capability of LvFM**" section, we added more description of the MoTe$_2$ sample on page 10: "Then, let us experimentally determine the spatial resolution and compositional resolving capability of LvFM by dealing with, e.g., a MoTe$_2$ film containing two crystalline phases 2H and 1T'. Even though the spatial resolution of LvFM is very high, the two phases of MoTe$_2$ cannot be distinguished without light excitation, since their static vdW force responses are almost the same (Fig. S4). Nevertheless, when an excitation light is applied, the magnitudes of LvF responses of the two phases are both enhanced and exhibit remarkable difference (Fig. S4), thanks to the distinct binding forces of molecules in the two phases. This is schematically illustrated in Fig. 3a when the tip scans across the boundary of the two phases. The measured LvF response during this scan is shown in Fig. 3b, from which the practical spatial resolution of LvFM can be estimated to be smaller than 10 nm, which is close to the theoretical estimate above. Note that the spatial resolution of AFM is restrained by deconvolution of tip shape and surface profile, and the spatial resolution of AFM-IR is exaggerated by material thermal diffusion length, while LvFM demonstrates for the first time a lateral resolution much smaller than the tip size. Moreover, the sensitivity of LvFM can be manifested by the signal-to-noise ratio (SNR) when the excitation laser is turn on or off. In the above experimental condition (a 633 nm laser excitation of 1.2 mW on the MoTe$_2$ film), the SNR of LvF is as high as 20 dB, as shown in Fig. 3c."

(2) In the last experimental section, the details of the samples are described in more detail. Specifically, for the MoTe$_2$ sample, we added on page 11: "The first experiment is to distinguish the 2H and 1T' crystalline phases in a MoTe$_2$ film of 12 nm thickness, the same sample as that characterized in Fig. 3, under the excitation of a 633 nm laser of 1.2 mW. The topography and LvF distribution simultaneously obtained by 300 × 300 pixels in an area of 20 μm × 20 μm within 14 min are given in Figs. 4a and 4b, respectively. The topography shows that the sample is an atomically flat film with surface fluctuation smaller than 0.2 nm, but cannot distinguish the two phases. But the LvF mapping results clearly discriminate the different regions of 2H and 1T' phases. The characterization results are further confirmed by Raman peaks at 235 cm^{-1}, attributed to in-plane

E_{2g} mode of 2H phase, and 124 cm^{-1} (A_g) and 163 cm^{-1} (B_g) peaks associated with 1T' phase (Fig. 4c), as well as by its confocal Raman mapping (Fig. 4d). XRD characterizations (Supporting Information, SI-9) have also been conducted to cross check the results."

For the hBN/WS$_2$ sample, we added on page 13: "Then we use LvFM to characterize a heteromaterial sample composed of two types of 2D materials, h-BN and multilayer WS$_2$ nanosheets. By raster scan, the topography of the sample is obtained (Fig. 4g), from which some coarse structures can be observed but the composing materials are unknown. Similarly, the absorbance of h-BN and WS$_2$ can be measured (Fig. 4h), whose difference determines their LvF contrast. The LvF mapping image under excitation wavelength of 633 nm is given in Fig. 4i, and more mapping results under excitation wavelengths of 600 nm and 680 nm are given in Supporting Information, SI-11. In all these images, the h-BN and WS$_2$ regions can be clearly distinguished with a high spatial resolution of ~10 nm."

(3) In SI-8, SI-9, and SI-10, more details are given regarding the MoTe$_2$ sample. In SI-11, more details are given regarding the hBN/WS$_2$ sample.

3-3.   I noted that, in page 5, the random amplitude variation "u" is defined later than its first introduction in equation 3. Also, its unit is not clear (unit of length given that we have "r+u" in the equation or unitless given that u = 0.01 later?), as well as its relation with sigma. This should be clearly defined in their first mention in page 5, and be coherent with their later use in page 10.

Response:
We thank the reviewer for the valuable comments. Since **u** is a displacement vector, it has the unit of length, nm. σ is a variance of u, which is also in unit of nm. We have made all necessary changes according to the reviewer comments to eliminate the confusion.

Revisions in the manuscript:
(1) On page 5, the random displacement vector **u**(ρ, z, t) is defined at its first appearance, before Eq. (3). "To describe the irregular thermal motion of a molecule located at (ρ, z) and at moment t, we can define a random displacement vector **u**(ρ, z, t), as shown in the inset of Fig. 1a, whose direction is always along **r**, i.e., toward the tip apex. "
(2) On page 6, we give the description about σ: "A variance σ of $u(\rho, z, t)$ is introduced to describe the extent of random motion of the molecule, which is correlated to the power of incident light." All the values of σ in the manuscript have been added unit nm.
(3) Since we have corrected the FOMs of LvFM, the statements like "u = 0.01" have been removed.

3-4.   I would advise to reword the last sentences of page 6. "Astonishingly" appears too strong for a scientific assertion. "cone-like" and "copy of a Gaussian" may

not be the most rational way to present the data.

Response:

We thank the reviewer for the comments and have made the changes accordingly.

Revisions in the manuscript:

The sentence on page 6 has been modified as "However, evident LvF signal was probed via the f_2 channel (Fig. 2g) simultaneously. The distribution of the LvF mapping results highly resembled the Gaussian profile of the incident laser spot, showing its optothermal force nature instead of any physical protrusion."

Minor typos:

Space missing page 10 between "LvFM" and "requires"

I guess "Physical original" should be "Physical origin" in Fig 1 caption

Responses and revisions:

We thank the reviewer for pointing out these typos. We are sorry for the appearance of typos and have tried our best to have a thorough check of the whole manuscript and SI to correct all the typos.

Moreover, we have taken this opportunity to improve the presentation of the whole manuscript and SI.

Reference：

1   Murdick, R. A. et al. Photoinduced force microscopy: A technique for hyperspectral nanochemical mapping. Japanese Journal of Applied Physics **56**, 08la04 (2017).

===== **Detailed point-by point responses to all the review comments** ========
Reviewer #1

I appreciate the authors' great efforts in improving the quality of the manuscript. Most of my questions have been well-addressed. I have a few follow-up questions I hope the authors can address.

1. I agree that different physical effects may be dominated by different probe-tip working zones. I wonder if it is only a function of the tip-sample distances or is also related to other factors such as the choice of probe, type of samples, and method of modulation/demodulation. The author should demonstrate the specificity of the signal to the "probe tip working zone." How different would the signal be if it was not through the vdW force? How sensitive is it to the tip-sample distance?

Response:

We are very grateful to the reviewer for the valuable comments. As discussed in detail in the manuscript and in SI-1 and SI-2, the LvF can be detected only when the probe works across the vdW force interaction zone, which is ensured by controlling the setpoint in tapping mode. In our implementation, the setpoint is usually set at ~60%. We have performed more experiments on the LvF response of the $MoTe_2$ film with the same experimental conditions as those adopted in Fig. 3 of the manuscript. Firstly, we tested the LvF response difference when the probe approaches the sample surface to work in the LvFM mode (tapping mode with setpoint ~60%) and is lifted off from the sample, as shown in Fig. R1(a). Clearly, by repeatedly switching the probe between the two states many times, we can see that the LvF signal appears only when the probe works in the LvFM mode across the vdW force interaction zone. In Fig. R1(b), we also tested the variation of the LvF response during the process when the probe (in tapping mode) is launched onto or lifted off from the sample surface. Note that the horizontal axis of Fig. R1(b) is the relative displacement of the probe, but not the calibrated exact tip-sample distance. The negative value is caused by the elastic deformation of the cantilever when the probe contacts the sample surface. It is seen that LvF signal only appears in the close vicinity of the sample surface and reaches maximum when $z \sim 0$ nm, which validates, from another point of view, the LvFM mechanism that the probe must work in the vdW force interaction zone. The sensitivity to the tip-sample distance can also be evaluated from the slope of the curve. These facts show that the tip-sample distance and the working mode of probe are essential to realize successful detection of the LvF signal.

The other factors such as the choice of probe and type of samples only affect the sensitivity of the LvF detection, as can be seen from the detailed discussions in the manuscript (see Fig. 3d for various types of probes, sample thicknesses, and sample materials) and SI (see SI-10 ~ SI-11 for various sample materials). The method of modulation/demodulation affects the SNR as discussed in detail in the manuscript (see Fig. 3c) and SI-7.
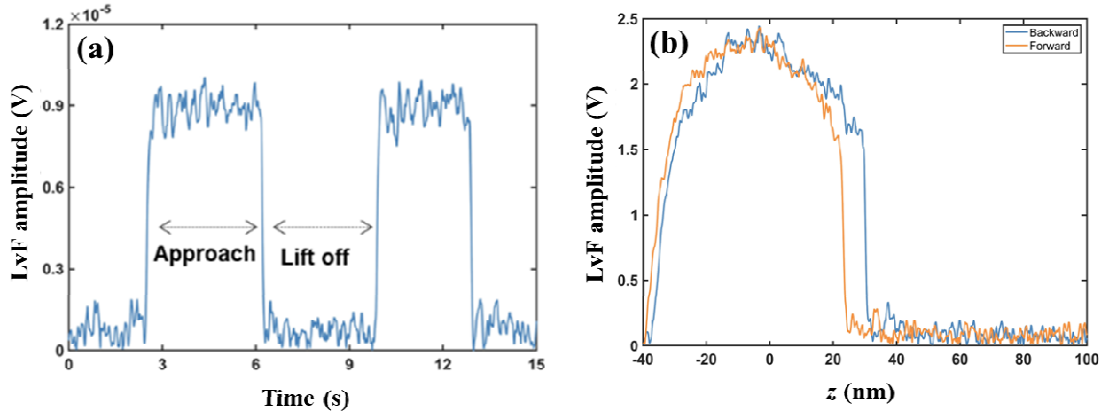
Fig. R1. Measured LvF signals from the MoTe$_2$ film samples with the same experimental conditions as those adopted in Fig. 3. (a) Detected LvF signals when the probe approaches or is lifted off from a sample surface. (b) Variation of the LvF signal when the probe is launched onto or lifted off from a sample surface.

Revisions in the manuscript:

We added on page 9: "We have further tested the dependence of the LvF signal on the tip-sample distance, see Supporting Information, SI-17."

We have added a new section in SI:

SI-17. Dependence of the LvF signal on the tip-sample distance

To show the dependence of the LvF detection on the tip-sample distance, we have performed more experiments on the LvF response of the MoTe$_2$ film with the same experimental conditions as those adopted in Fig. 3. Firstly, we tested the LvF response difference when the probe approaches the sample surface to work in the LvFM mode (tapping mode with setpoint ~60%) and is lifted off from the sample, as shown in Fig. R1(a). Clearly, by repeatedly switching the probe between the two states many times, we can see that the LvF signal appears only when the probe works in the LvFM mode in the vdW force interaction zone. In Fig. R1(b), we also tested the variation of the LvF response during the process when the probe (in tapping mode) is launched onto or lifted off from the sample surface. Note that the horizontal axis of Fig. R1(b) is the relative displacement of the probe, but not the calibrated tip-sample distance. The negative value is caused by the elastic deformation of the cantilever when the probe contacts the sample surface. It is seen that LvF signal only appears in the close vicinity of the sample surface and reaches maximum when z ~ 0 nm, which validates, from another point of view, the LvFM mechanism that the probe must work in the vdW force interaction zone. The sensitivity to the tip-sample distance can be evaluated from the slope of the curve. According to the above tests, we can see that the tip-sample distance and the working mode of probe are essential to realize successful detection of the LvF signal.
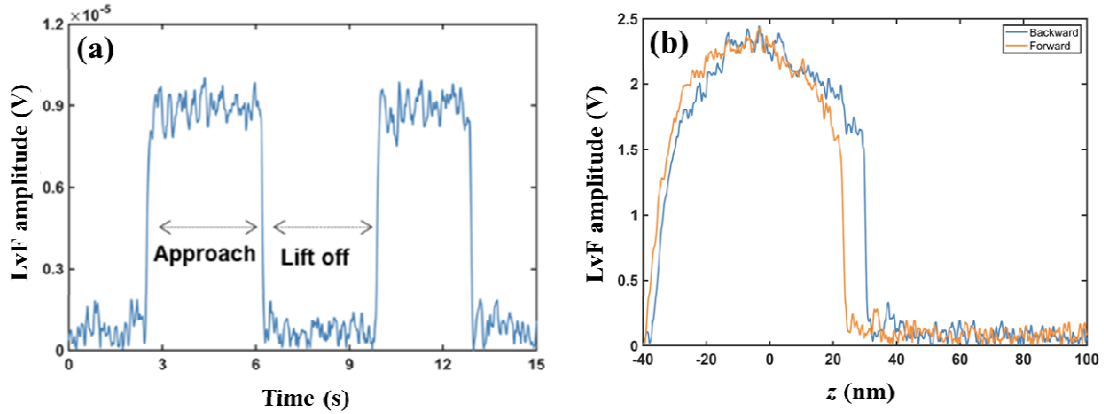
2. It is interesting that the modulation of LvF does not require a photothermal expansion. For few-layer 2D material, I agree that there is no significant thermal expansion; therefore, the optical force mostly comes from the change of LvF. However, what if the sample absorbs the wavelength of light significantly? Will this technique still work? For example, in the Au nanoparticle and PS nanosphere experiment demonstrated in the SI, what specific experimental parameters, including the wavelength of light, type of probe, and modulation amplitude, were used?

Response:

We thank the reviewer for this question. As shown in our paper, the LvFM still works well when the light absorption is strong. As seen from the detailed discussions in the manuscript regarding Fig. 1d, Fig. 3d, Fig. 4e, Fig. 4f, and Fig. 4h, when the material absorption gets stronger (which may be caused by the material itself, the selection of excitation wavelength, or the increase of the incident light power), the LvF response also gets stronger. The upper limit for various materials has not been studied yet. But LvFM generally works in the condition with small or moderate light excitation power, so that the sample is not damaged by the excitation light. Moreover, the LvF strength is also dependent on the thickness of the sample, as discussed in SI-14, which is also related to the absorption of light in the material.

In SI-12 (LvFM characterization of Au nanoparticles and PS nanospheres), the experimental parameters are: a 532 nm excitation laser of 1.0 mW, an Au probe from NANOSENSORS™ with f_1=231 kHz, Q_1= 438, k_1 = 40 N/m, f_2=1475 kHz, Q_2=642, and tip curvature radius ~ 25 nm, the setpoint of 60%, and the modulation frequency f_m = $f_2 - f_1$ = 1244 kHz.

Revisions in the manuscript:

In SI-12, we added the experimental parameters:

"The LvFM experimental parameters are: a 532 nm excitation laser of 1.0 mW, an Au probe from NANOSENSORS™ with f_1=231 kHz, Q_1= 438, k_1 = 40 N/m, f_2=1475 kHz, Q_2=642, and tip

curvature radius ~ 25 nm, the setpoint of 60%, and the modulation frequency $f_m = f_2 - f_1 = 1244$ kHz."

3. Why do the tapping mode and sideband modulation/demodulation provide a better SNR? Shouldn't the non-contact mode with direct modulation provide a better SNR if the force is only decided by the tip-sample distance?

Response:

We thank the reviewer for this question. The tapping mode is used to modulate the cantilever with its 1st resonance frequency f_1, so that it can beat with the modulation frequency of light f_m to generate f_2 that is the 2nd order resonance frequency of the cantilever. It is the use of this technique that one can exclude other forces in the detection. Specifically, as discussed in SI-7, the volume of sample material affected by the excitation light can be divided into three regions: the central region V_p right below the probe tip where the sample has the largest optothermal effect and the strongest LvF interaction with the probe, the intermediate region V_a where the material directly absorbs light but with negligible LvF interaction with the probe, and the thermal diffusion region V_d where the material receives heat not directly from the illumination but through heat diffusion from the nearby regions with higher temperature. When the sample is illuminated by a modulated laser beam, all the three regions have thermal responses with the same modulation frequency f_m. However, since the probe tip vibrates at tapping frequency f_1, only the response of the central region V_p contains a frequency component f_1. f_m and f_1 couple with each other, resulting in new vibration frequencies $f_2=|f_m \pm f_1|$. In this way, both the forces outside the central region V_p (which are not modulated by f_1) and the forces not influenced by the optothermal effect (which are not modulated by f_m) cannot respond and are therefore excluded in LvFM detection. Then the LvF signal can be demodulated from the resonance amplitude of the cantilever at f_2 with the best SNR. These are also discussed in detail in SI-7.

Moreover, by controlling the setpoint in tapping mode, one can ensure the probe working across the vdW force interaction zone to detect the LvF signal, as discussed in above reply to Comment #1.

Using non-contact mode cannot meet these goals of LvFM. On one hand, one cannot ensure that the probe enters the vdW force interaction zone in non-contact mode, which may cause the failure of the detection of LvF signal. Actually, we have tried to use the non-contact mode and indeed failed. On the other hand, the direct modulation (with only f_1 and without its coupling with f_m) can only measure the topography of the sample, as AFM does, but not the LvF signal. The mechanism lies in the difference of detecting the static vdW force and dynamic LvF, which has been discussed in detail in the manuscript (see sections "**Two types of atomic forces**" and "**Light-modulated vdW force**").

Furthermore, we want to emphasize that although LvF can only be detected in the vdW force interaction zone, it is not "only decided by the tip-sample distance". The optothermal effect of the sample material contributes to the strength of the LvF signal. In a word, the tip-sample distance determines where the LvF exists and is detectable, while the (material specific) tip-sample interaction determines the strength of LvF.

<u>**Revisions in the manuscript:**</u>

In SI-7, we added the sentence:

"In this way, both the forces outside the central region V_p (which are not modulated by f_1) and the forces not influenced by the optothermal effect (which are not modulated by f_m) cannot respond and are excluded in the LvFM detection."

4. Since the measurement setup of LvFM and PiFM both use light-modulated tip-sample interaction, I still don't understand how one effect (light-induced dipole-dipole interaction, for example) will not appear in the other measurement techniques. I agree that the dominating effects may vary from sample to sample, and LvF may be dominant in the 2D material sample demonstrated in this manuscript. However, is this a general conclusion that LvFM can decouple LvF from the other light-induced tip-sample interactions (for example, photothermal expansion, dipole-dipole interactions, photoacoustic pressures, etc.)?

<u>**Response:**</u>

We thank the reviewer for this question. Although LvFM and PiFM both use light modulation techniques, they work in quite different principles and probe working modes. Detailed discussions have been given in SI-1. More discussions are below.

As mentioned above, in PiFM the probe works in non-contact mode and therefore cannot be ensured to enter the vdW force interaction zone to probe the LvF. The force detected in PiFM is a strong dipole-dipole interaction force originating from materials with large polarizability under light excitation and is sensitive to the wavelength of light (*Appl. Phys. Lett.* **97,** 2010). Obviously, it is much stronger and has much larger tip-sample interaction distance than LvF that we are studying. For the same reason, the excitation light power is usually also much stronger than what we use in LvFM.

In comparison, LvFM detects LvF only existing in the vdW force interaction zone (no more than several nanometers above the sample). Hence, the probe has to work in tapping mode with small setpoint. Since LvF is a universal force existing between any sample and the probe tip, it in principle can be detected in any material no matter its polarizability is strong or not.

But, from another point of view, LvFM cannot exclude other forces caused by optothermal effects if they simultaneously exist in the vdW force interaction zone and can also be modulated by the same technique. For example, let us consider a scenario of characterizing a particular material with strong polarizability with LvFM, both the LvF and dipole-dipole interaction force may exist in the vdW force interaction zone under modulated light excitation. Then, the detected signal may inevitably contain the contributions from both LvF and the dipole-dipole interaction force. Nevertheless, this is not harmful but is advantageous for the discrimination of materials, since both forces originate from the intrinsic optothermal effect of the material itself. This can be understood by recalling the sideband demodulation mechanism mentioned in the reply to Comment #3. In a word, the LvFM works based on the detection of LvF, but may also receive contributions from other optothermal forces in the vdW force interaction zone (provided that they exist and bear the same sideband modulation). Further study

on the difference of LvFM and PiFM is interesting and will be our future work.

5. The authors mentioned that LvFM requires a closer tip-sample distance than PiFM; how is this achieved specifically? I don't think PiFM has specified the tip-sample distance in operation. The difference between LvFM and PiFM is unclear, and this needs to be addressed.

<u>Response:</u>

We thank the reviewer for this question. This can be anticipated from two aspects: (1) PiFM works in non-contact mode, where the probe is farther from the sample surface than tapping mode used in LvFM; (2) PiFM usually adopts a larger setpoint (say, ~85% as reported in literatures), which is larger than the setpoint of ~60% in LvFM. These facts show that LvFM probe works in a closer distance to the sample surface than that in PiFM. Since we do not have a PiFM setup, we cannot compare them directly. In the future, we will seek opportunity to collaborate with other research groups working with PiFM and do more comparative studies on the two methods.

Reviewer #2

The biggest weakness of this paper was the question of whether it could be applied to various samples, but the doubt was resolved by applying and proving it to various samples. The manuscript can be accepted in Nature Communications.

<u>Response:</u>

We are very grateful to the reviewer for giving the valuable comments and positive recommendations.

Reviewer #3

I am satisfied by the answers given by the authors on the different comments I made previously. I confirm that the corresponding changes made to the manuscript are relevant and make the overall demonstration more compelling and scientifically sound. I do not have further comments.

Please note that, in the current given version, formatting errors are visible (line 138 of manuscript, and in several instances in SI text and equations, I guess they are missing delta or (number) signs).

<u>Response:</u>

We are very grateful to the reviewer for giving the valuable comments and positive recommendations.

As for the formatting errors mentioned by the reviewer, we have checked again but did not find them. Probably, these are some displaying problems of some special symbols on different computers (e.g., "$\propto h^{-4}$" in line 138). We have replaced these symbols with font "Symbol" instead of "Times New Roman" in Word software, and hopefully this may help to solve the problems. If these problems still exist in the typesetting procedure, we will work with the copy editor to solve them.

Revisions in the manuscript:

We have tried our best to carefully check again the writing of the revised manuscript and SI thoroughly to ensure that there is no typesetting error remained. For example, in the captions of Fig. S10 and Fig. S13 in SI, we have modified the text to indicate the "AFM phase image" and "LvFM image" accurately.

Other revisions in the manuscript:

We have added a "**Code availability**" section in the manuscript according to the editorial request.

===== **Detailed point-by point responses to all the review comments** ========
Reviewer #1

I thank the authors for their efforts in addressing my questions. Most of the questions have been answered, but some parts were not addressed thoroughly.

1. It is interesting to see this force varying with distance z. What is the definition of z=0 and how is this calibrated? My question is specifically related to how this force is attributed to the light-enhanced van der Waals force rather than other light-induced forces discussed in prior works. Have the authors tried changing the modulation/demodulation methods or setpoint to see how these factors could affect the forces?

<u>**Response:**</u>

We are very grateful to the reviewer for the valuable comments. In Fig. S15 copied below, the horizontal axis z of Fig. S15(b) is the displacement of the dither piezo driving the probe. Therefore, it indicates the relative displacement of the probe, but not the absolute tip-sample distance. Actually, the absolute tip-sample distance cannot be measured or calibrated anyway. The only way to monitor the tip-sample distance is to control the setpoint, as discussed in the manuscript and SI. During the measurement of Fig. S15(b), with the decrease of z, the setpoint also decreases. When the setpoint approaches 60% as we adopt in LvFM, the displacement of the dither piezo is now denoted as $z = 0$. When $z < 0$, the probe is closer to the surface, and vice versa.
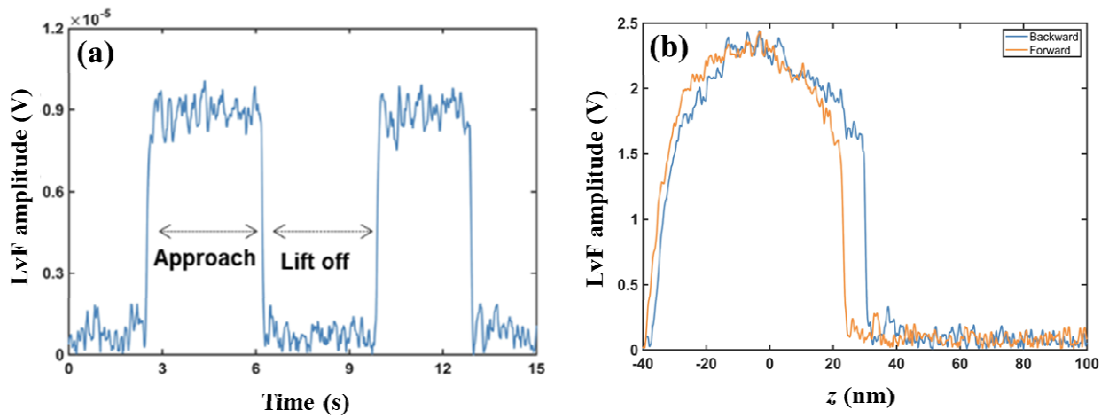

Fig. S15. Measured LvF signals from the MoTe$_2$ film with the same experimental conditions as those adopted in Fig. 3. (a) Detected LvF signals when the probe approaches or is lifted off from a sample surface. (b) Variation of the LvF signal when the probe is launched onto or lifted off from a sample surface.

As for the intrinsic difference of LvF from other light-induced forces, we have discussed in detail in SI-1. Taking PiFM as an example, the dipole-dipole interaction force in PiFM is dominant only in materials with large polarizability and is also sensitive to the wavelength of illumination light. PiFM can use direct demodulation or sideband demodulation, where the setpoint is usually set at ~85%.[1] In comparison, LvFM makes use of LvF, which is a universal force existing between any sample and the probe tip. Therefore, in principle, LvFM has no special requirement on sample

material and excitation light, provided that the thermal motion of molecules of the sample can be enhanced by light excitation. As shown in SI, LvFM can be applied to a wider range of materials, including metals, dielectrics, semiconductors, and ultra-thin nanomaterials such as 2D materials. Moreover, since LvFM works across the vdW force interaction zone, it requires a closer tip-sample distance than PiFM, so that the tapping mode is adopted with setpoint ~60%. As seen in Fig. S15(b), the largest force signal appears at $z = 0$ and setpoint ~60%, which verifies again that this is not the dipole-dipole interaction force in PiFM. Moreover, the light excitation conditions (including the light power density, wavelength, NA, etc.) in LvFM are quite different from those in PiFM. All these show that LvF is quite different in nature from other light-induced forces such as that in PiFM. It is named as light-modulated van der Waals force (LvF), because the probe must enter the vdW force interaction zone during tapping so as to detect the dynamic light-modulated vdW force signal, which is significantly different from that in PiFM.

Moreover, as explained in SI-7, by using the sideband demodulation, the other forces including the forces outside the central region V_p (which are not modulated by f_1) and the forces not influenced by the optothermal effect (which are not modulated by f_m) are excluded in LvFM detection.

As required by the reviewer, we have also tried different setpoints in the modulation/demodulation, as shown in Fig. R1 below. Specifically, at each setpoint, we experimentally scanned the laser modulation frequency f_m to get the dependence of LvF amplitude on f_m. It is seen that with the increase of the setpoint, the peak f_m moves to lower frequency. According to AFM theory, this shift represents the increase of the ratio of attractive force to repulsive force.[2] Since the attractive force in this tip-sample interaction region is mainly the vdW force, these results in Fig. R1 verify again, from another point of view, that LvF is in nature the vdW force.
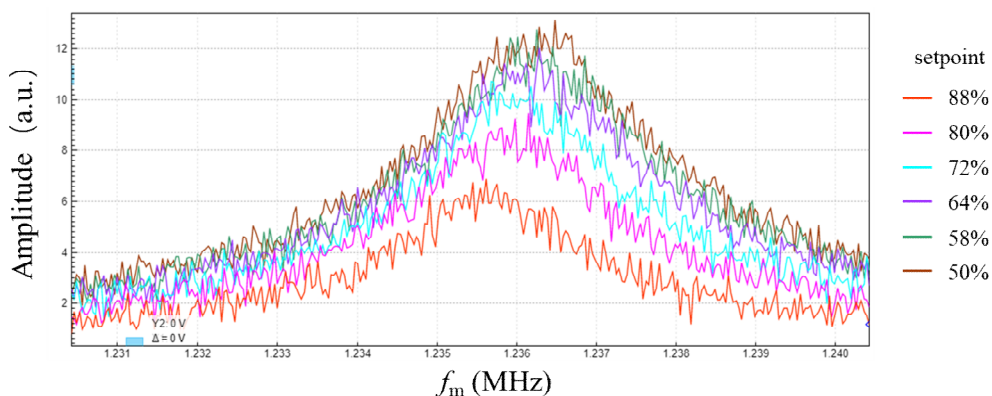

Fig. R1. Measured dependence of the LvF amplitude on the laser modulation frequency f_m at different setpoints.

Revisions in the paper:

In SI-17, we modified the text:

"Note that the horizontal axis z of Fig. S15(b) is the displacement of the dither piezo driving

the probe. Therefore, it indicates the relative displacement of the probe, but not the absolute tip-sample distance. Actually, the absolute tip-sample distance cannot be measured or calibrated anyway. The only way to monitor the tip-sample distance is to control the setpoint, as discussed before. During the measurement of Fig. S15(b), with the decrease of z, the setpoint also decreases. When the setpoint approaches 60% as we adopt in LvFM, the displacement of the dither piezo is now denoted as $z = 0$. When z < 0, the probe is closer to the surface, and vice versa. It is seen that LvF signal only appears in the close vicinity of the sample surface and reaches maximum when z ~ 0 nm (corresponding to setpoint ~60%) ..."

We have also added a paragraph of discussion and a figure:

Moreover, we have tried different setpoints in the modulation/demodulation, as shown in Fig. S16. At each setpoint, we scanned the laser modulation frequency f_m to get the dependence of LvF amplitude on f_m. It is seen that with the increase of the setpoint, the peak f_m moves to lower frequency. According to AFM theory, this shift represents the increase of the ratio of attractive force to repulsive force. Since the attractive force in this tip-sample interaction region is mainly the vdW force, it verifies again, from another point of view, that LvF is in nature the vdW force.
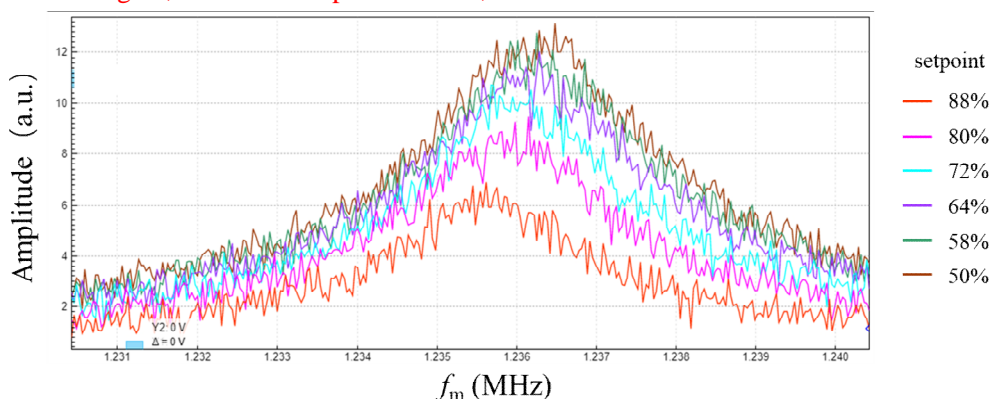

Fig. S16. Measured dependence of the LvF amplitude on the laser modulation frequency f_m at different setpoints.

2. Based on Figure 4, the stronger the absorption, the stronger the force observed. How did the authors exclude the contribution of photothermal expansion of the sample? Additionally, gold absorbs more strongly than polystyrene in the visible spectrum, but Figure S10d shows no force for gold particles and significant force for polystyrene beads. This seems to contradict the finding in Figure 4.

Response:
We are very grateful to the reviewer for the valuable comments. For a given material, the LvF response is stronger when the optothermal absorption is larger (which may be caused by many factors such as stronger light excitation and thicker material). For different materials, however, the relationships between LvF and absorption may be quite different and complicated, which are subject to the intermolecular binding forces of the materials. In other words, the optothermally intensified thermal motions of molecules are related to many factors mainly including the intermolecular binding forces, but not only the absorption. Therefore, it is difficult to directly compare the LvF responses of two different materials just according to their light absorption.

As explained in the theory part of the paper, LvFM works based on the intensified thermal motions of molecules on the sample surface under light excitation, which may exhibit as thermal expansion (e.g., bulk material under strong light excitation) or not (e.g, few layer 2D material[3] or bulk material under weak excitation), as can be seen from the many characterization examples in the manuscript and SI. In other words, photothermal expansion is not a prerequisite for LvFM and LvFM characterization also does not exclude thermal expansion (although in most cases, no thermal expansion is generated in LvFM characterization with low excitation power).

As for Fig. S10d, it shows the micro-photoluminescence (PL) spectra of the NPs, but not the LvF spectra. The PL characterization is to cross-check that the NPs at positions ① and ② are indeed different kinds of NPs. Since Au NP does not have PL radiation and PS NP have evident PL peak, we can confirm the two types of NPs according to the PL spectra in Fig. S10d. To avoid confusion, we have further clarified the caption of the vertical axis of Fig. S10d.

Revisions in the paper:
We have updated Fig. S10 with revised Fig. S10d with new vertical axis caption.

References:
[1] Murdick, R. A. et al. Photoinduced force microscopy: A technique for hyperspectral nanochemical mapping. *Japanese Journal of Applied Physics* **56**, 08la04 (2017).
[2] Korayem, M. H. et al. Comparison of frequency response of atomic force microscopy cantilevers under tip-sample interaction in air and liquids. *Scientia Iranica* **19**(1), 106-112 (2012).
[3] El-Mahalawy, S. H. et al. The thermal expansion of $2H-MoS_2$, $2H-MoSe_2$ and $2H-WSe_2$ between 20 and 800 °C. *Journal of Applied Crystallography* **9**(5), 403-406 (1976).