

Ultra-strong nonlinear optical processes and trigonal warping in MoS₂ layers

A. Säynätjoki^{1,2,*}, L. Karvonen^{1,*}, H. Rostami^{3,*}, A. Autere¹, S. Mehravar⁴, A. Lombardo⁵, R. A. Norwood⁴, T. Hasan⁵, N. Peyghambarian^{1,2,4}, H. Lipsanen¹, K. Kieu⁴, A. C. Ferrari⁵, M. Polini³, Z. Sun¹

¹*Aalto University, Department of Micro and Nanosciences, Tietotie 3, FI-02150 Espoo, Finland*

²*University of Eastern Finland, Institute of Photonics, Yliopistokatu 7, FI-80100 Joensuu, Finland*

³*Istituto Italiano di Tecnologia, Graphene Labs, Via Morego 30, I-16163 Genova, Italy*

⁴*University of Arizona, College of Optical Sciences,
1630 EUniversity Blvd, Tucson, AZ 85721, USA and*

⁵*Cambridge Graphene Centre, University of Cambridge, Cambridge CB3 0FA, UK*

We report ultra-strong high-order nonlinear multiphoton processes in monolayer MoS₂ (1L-MoS₂): the third harmonic is thirty times stronger than the second harmonic, and the fourth harmonic is comparable to the second harmonic. We find that second and third harmonic processes are strongly dependent on elliptical polarization, which can be used to selectively tune harmonic generation with different orders. We explain this by calculating the nonlinear response functions of 1L-MoS₂ with a continuum-model Hamiltonian and quantum-mechanical diagrammatic perturbation theory, highlighting the crucial role of trigonal warping. A similar effect is expected for all other transition-metal dichalcogenides. Our results pave the way for efficient and tunable harmonic generation based on layered materials for various applications, including microscopy and imaging.

[73] Nonlinear optical processes, such as harmonic generation[1], are of great interest for various applications, e.g. microscopy[2, 3], therapy[2, 3], frequency conversion[1, 4] and data storage[3]. Nonlinear optical phenomena can generate high-energy photons by converting $n = 2, 3, 4, \dots$ low-energy photons into one high-energy photon. These are usually referred to as second-, third- and fourth-harmonic generation (SHG, THG and FHG)[1–4]. Due to different selection rules[1, 5], various harmonic processes are distinct from optically-pumped laser phenomena (e.g. optically-pumped amplification[6]), and other typical single-photon processes (e.g. single-photon excited photoluminescence[1]), in which the energy of the generated photons is smaller than the pump photons. Therefore, multiphoton harmonic processes have been widely exploited for various applications (e.g. all-optical signal processing in telecommunications[1, 7], medicine[2, 3], and data storage[3]), as well as to study various transitions forbidden under low-energy single-photon excitation[2, 3]. The physical origin of these processes is the nonlinear polarization induced by an electromagnetic field. This gives rise to higher harmonic components, the n -th harmonic component amplitude being proportional to $|E|^n$ [1]. Quantum mechanically, higher-harmonic generation consists in the annihilation of n pump photons and generation of a photon with n times the pump energy. Because an n -th order nonlinear optical process requires n photons to be present simultaneously, the probability for higher-order processes is lower than for lower order[1]. Thus, higher-order processes are typically weaker and require higher intensities[8, 9].

Graphene and related materials (GRMs) are at the center of an ever increasing research effort due to their unique and complementary properties, making them appealing for a wide range of photonic and optoelec-

tronic applications[10–16]. Amongst these, semiconducting transition-metal dichalcogenides (TMDs) are of particular interest due to their direct bandgap when in monolayer form[17, 18], leading to an increase in luminescence efficient by a few orders of magnitude compared with the bulk material[17–22]. 1L-MoS₂ has a single layer of Mo atoms sandwiched between two layers of S atoms in a trigonal prismatic lattice. Therefore, in contrast to graphene, it is non-centrosymmetric and belongs to the space group D_{3h}^1 [23]. The lack of spatial inversion symmetry makes 1L-MoS₂ an interesting material for nonlinear optics, since second-order nonlinear processes are present only in non-centrosymmetric materials[1]. However, when stacked, MoS₂ layers are arranged mirrored with respect to one another[23], therefore MoS₂ with an even number of layers (EN) is centrosymmetric and belongs to the D_{3d}^3 space group[23], producing no second harmonic (SH) signal. On the other hand, MoS₂ with any odd number of layers (ON) is non-centrosymmetric. SHG from 1L-MoS₂ has already been experimentally demonstrated by several groups[23–31].

Here, we present experimental and theoretical work on nonlinear harmonic generation in 1L and few-layer (FL) MoS₂ flakes. We report ultra-strong THG and FHG from 1L-MoS₂. In comparison to SHG, the THG is more than one order of magnitude larger and FHG has the same magnitude as SHG. This is surprising, since one normally expects the intensity of non-linear optical processes to decrease with n [1, 5]. One therefore expects the SHG intensity to be much larger than THG and FHG, although even-order processes only exist in non-centrosymmetric materials. Our results show that this expectation is wrong in the case of 1L-MoS₂. The point is that, at sufficiently low photon-frequencies (in our experiments the photon energy of the pump is 0.8eV), SHG only probes the low-energy band structure of 1L-

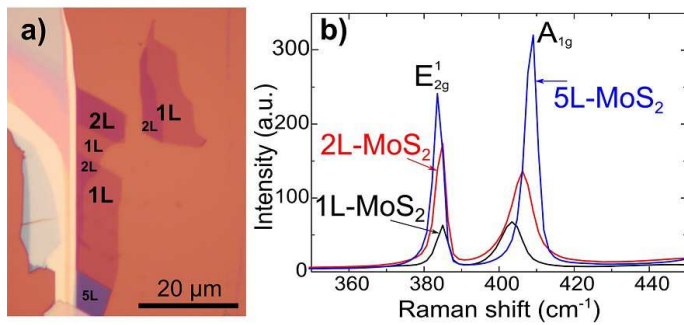


FIG. 1: a) Optical micrograph with single-layer, bilayer, and five-layer areas marked by 1L, 2L and 5L, respectively. b) Raman spectra of the same sample.

MoS₂. This is nearly rotationally invariant[14, 32–38], but with corrections due to trigonal warping. It is because of these corrections[34–36], fully compatible with the D_{3h}^1 space group[1], but reducing the full rotational symmetry of the low-energy bands to a three-fold rotational symmetry[1], that a finite amplitude of non-linear harmonic processes with even n can exist at low photon energies. Thus, lack of spatial inversion symmetry is only a necessary but not sufficient condition for the occurrence of SHG. We demonstrate that the observed THG/SHG intensity ratio can be explained by quantum mechanical calculations based on finite-temperature many-body diagrammatic perturbation theory[39] and low-energy continuum-model Hamiltonians that include trigonal warping[40]. We show that these nonlinear processes are sensitive to the number of layers, their symmetry, relative orientation, and the elliptical polarization of the excitation light. Similar effects are expected for all other TMDs. This paves the way for the assembly of heterostructures with tailored nonlinear properties.

MoS₂ flakes are produced by micromechanical cleavage (MC) of bulk MoS₂[41, 42] onto Si+285nm SiO₂ substrate. 1L-MoS₂ and bilayer (2L-MoS₂) samples are identified by a combination of optical contrast[43, 44] and Raman spectroscopy[45, 46]. Raman spectra are acquired by a Renishaw micro-Raman spectrometer equipped with a 600 line/mm grating and coupled with an Ar⁺ ion laser at 514.5nm. Fig.1 shows the MoS₂ flakes studied in this work and their Raman signatures. A reference MC graphene sample is also prepared on a similar substrate.

Nonlinear optical measurements are carried out with the setup of Fig.2[47, 48]. As excitation source, we use an erbium doped mode-locked fiber laser with a \sim 50MHz repetition rate, maximum average power \sim 60mW and pulse duration \sim 150fs, which yields an estimated pulse peak power \sim 8kW[49]. The laser beam is scanned with a galvo mirror and focused on the sample using a microscope objective. The back-scattered second and third harmonic signals are split into different branches using a

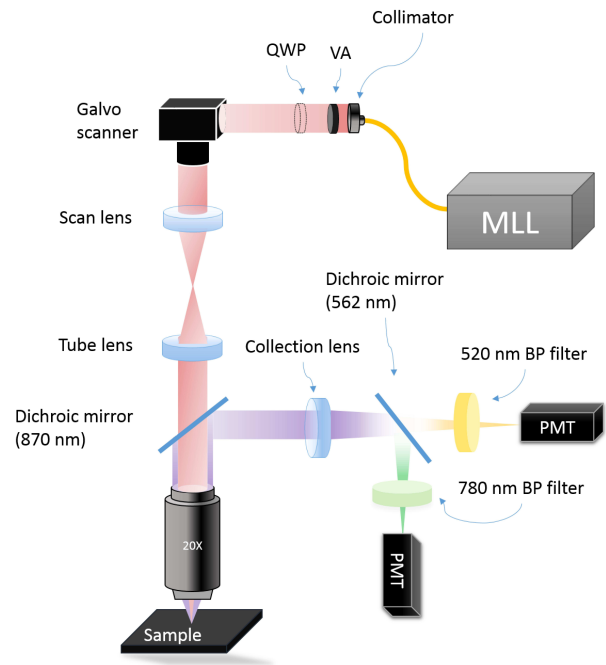


FIG. 2: Schematic diagram of the multiphoton microscope. MLL: linearly polarized mode-locked fiber laser. VA: variable attenuator. QWP: quarter-wave plate. QWP is inserted only when we study the dependence of SHG and THG on the elliptical polarization of the pump light. BP filter: Bandpass filter. PMT: Photomultiplier tube.

dichroic mirror and then detected using photomultiplier tubes (PMTs). For two-channel detection, the light is split into two PMTs using a dichroic mirror with 560nm cut-off. After the dichroic mirror, the detected wavelength range can be further refined using bandpass filters. The light can also be directed to a spectrometer (OceanOptics QE Pro-FL) to analyze the spectral properties of the generated light. The average power on sample is kept between 10 and 28mW with a typical measurement time \sim 5 μ s, which prevents sample damage and enables high signal-to-noise-ratio, even with acquisition time per pixel in the μ s range.

SHG and THG images of the MoS₂ sample are shown in Figs.3a,b). The SHG signal is generated in 1L-MoS₂, while 2L-MoS₂ appears dark. As discussed above, the second-order nonlinear response is present in 1L-MoS₂, which is non-centrosymmetric. However, when stacked to form 2L-MoS₂, MoS₂ layers are mirrored one with respect to another[23, 24]. Therefore, EN-MoS₂ is centrosymmetric [23, 24], and belongs to the D_{3d}^3 space group[23, 24], producing no SHG signal. On the other hand, ON-MoS₂ is non-centrosymmetric [23, 24].

We note that strong THG is detected compared with SHG, even for 1L-MoS₂, as shown in Fig.3b). THG was previously reported for a thick ($N \geq 10$) MoS₂ flake[27], but here we see it down to 1L-MoS₂. However, THG is not observed from the thickest areas of our flake, with

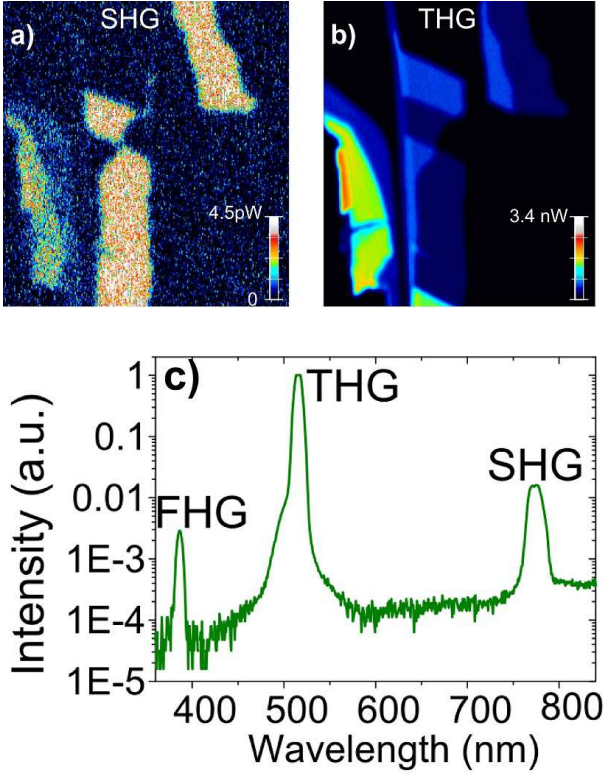


FIG. 3: a) SHG and b) THG map of the MoS₂ flake in Fig.1a). c) Optical spectrum of the nonlinear signal from 1L-MoS₂ with a peak irradiance ~ 30 GW/cm².

N 30, as in Ref.[27]. The output spectrum in Fig.3c) further confirms that we are observing SHG and THG together. Peaks for THG and SHG at ~ 520 and ~ 780 nm can be seen, as well as a peak at ~ 390 nm, corresponding to a four-photon process. This is detected only in 1L-MoS₂. Its intensity is ~ 5.5 times lower than SHG, and two orders of magnitude smaller than THG.

SHG signals on areas with $N=3, 5, 7$ have nearly the same intensity as 1L-MoS₂, Fig.4a). This contrasts Ref.[23], where a pump laser at 810 nm was used. We attribute this difference to the fact that photons generated in the second-order nonlinear process in our setup with a 1560 nm pump wavelength have an energy ~ 1.6 eV (780 nm), below the band gap of 1L-MoS₂ [17, 18], therefore not adsorbed, unlike the SHG signal in Ref.[23].

Based on the measured SHG and THG intensities, we can estimate the nonlinear susceptibilities $\chi^{(2)}$ and $\chi^{(3)}$. $\chi^{(2)}$ can be calculated from the measured average powers of the fundamental and SH signals as follows[50]:

$$\chi_s^{(2)} = \sqrt{\frac{\epsilon_0 c \lambda_2^4 P_{2\omega} R \tau^2 (n_2 + 1)^2 (n_1 + 1)^2}{32 \text{NA}^2 \tau_2 P_{\text{pump}} \phi}}, \quad (1)$$

where τ is the pulse width, P_{pump} is the average power of the incident fundamental (pump) beam and $P_{2\omega}$ stands for the generated SH beam power, R is the repetition

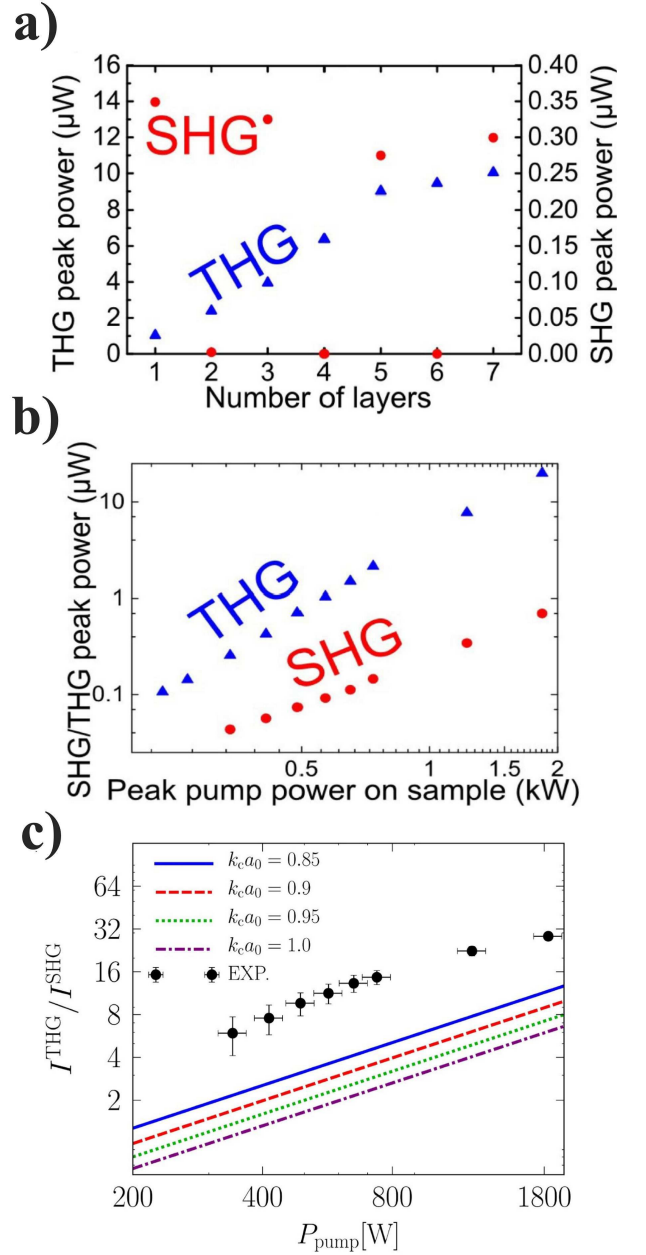


FIG. 4: a) SHG and THG intensities as functions of N . b) Power dependence of SHG and THG in 1L-MoS₂. c) Experimental and theoretical THG/SHG irradiance ratio as a function of P_{pump} . Different theoretical curves refer to different values of the ultra-violet cut-off k_c (measured in units of $1/a_0 = \sqrt{3}/a$ where $a \approx 3.16$ Å is the lattice constant of 1L-MoS₂ [17, 18]).

rate, $\text{NA}=0.5$ is the numerical aperture, $\lambda_2=780$ nm is the SH wavelength, $\tau = \tau_2 = 150$ fs are the pulse durations at fundamental and SH wavelengths, $\phi = 8\pi \int_0^1 |\cos^{-1} \rho - \rho \sqrt{1 - \rho^2}| \rho d\rho = 3.56$ from Ref.[50], and $n_1 = n_2 = \sim 1.45$ are the refractive indices of the substrate at the wavelengths of the fundamental and SHG, respectively. The effective bulk-like second order suscep-

tibility of MoS₂ ($\chi_{\text{eff}}^{(2)}$) can be obtained from Eq.1 with $\chi_{\text{eff}}^{(2)} = \frac{\chi_s^{(2)}}{t_{\text{MoS}_2}}$, where $t_{\text{MoS}_2} = 0.75\text{nm}$ is the 1L-MoS₂ thickness[13, 14]. We obtain the effective second order susceptibility $\chi_{\text{eff}}^{(2)} \sim 2.2 \text{ pm/V}$ for 1L-MoS₂. The third-order susceptibility $\chi^{(3)}$ of MoS₂ is estimated by comparing the measured THG signal from MoS₂ to that of 1L-graphene (SLG):

$$\chi^{(3)} \approx \frac{t_{\text{gr}}}{t_{\text{MoS}_2}} \sqrt{\frac{\text{THG}_{\text{MoS}_2}}{\text{THG}_{\text{gr}}}} \chi_{\text{gr}}^{(3)}. \quad (2)$$

With $t_{\text{gr}} \sim 0.33\text{nm}$ the SLG thickness, and THG_{gr} and $\text{THG}_{\text{MoS}_2}$ the measured signals from SLG and MoS₂, respectively. Using $\chi_{\text{gr}}^{(3)} \sim 3 \times 10^{-7} \text{ esu} \sim 4.2 \times 10^{-15} \text{ m}^2/\text{V}^2$ [47], we find $\chi^{(3)} \sim 2.8 \times 10^{-7} \text{ esu} \sim 3.9 \times 10^{-15} \text{ m}^2/\text{V}^2$, comparable to that of SLG in the same frequency range that we used in our experiment. This is remarkable, as SLG is known to have a large $\chi^{(3)}$ [51, 52], 2 orders of magnitude larger than that of bulk glass[53] and ~ 5 times larger than gold[53]. Furthermore, MoS₂ is transparent at this telecommunication wavelength due to its $\sim 1.9\text{eV}$ gap[17, 18, 20], while SLG absorbs 2.3% of the light[13, 54]. Therefore, MoS₂ and possibly other TMDs are promising for integration with optical waveguides or fibers for all-optical nonlinear devices, where materials with nonlinear properties are essential, such as all-optical modulators and signal processing devices[16].

The SHG and THG power dependence follows quadratic and cubic trends, respectively, Fig.4b). At the power levels of our measurements, THG is up to 30 times stronger than SHG. We attribute such a large THG/SHG ratio to the approximate rotational invariance of the MoS₂ band structure at low energies, which is broken by trigonal warping. Fig.4c) plots the THG/SHG ratio obtained from the experiments and microscopic calculations based on the $k \cdot p$ theory[40] and finite-temperature diagrammatic perturbation theory[39] (details in Methods). The calculations are factor of two smaller than the experimental data. Considering the complexity of the investigated non-linear optical processes and the fact that our calculations ignore high-energy band structure effects[38] and many-body renormalizations[55], we believe this to be a satisfactory agreement, indicating the importance of trigonal warping in harmonic generation.

FHG generally derives from cascades of lower-order nonlinear multi-photon processes[8, 56]. With an excitation wavelength of 1560nm, this could be, e.g., a cascade of two SHG processes, where 780nm photons are first generated through SHG ($\omega_{1560\text{nm}} + \omega_{1560\text{nm}} \Rightarrow \omega_{780\text{nm}}$) and then undergo another SHG process ($\omega_{780\text{nm}} + \omega_{780\text{nm}} \Rightarrow \omega_{390\text{nm}}$). To yield a FHG at 390 nm of the same intensity as SHG at 780nm in this cascaded process, one would need a conversion efficiency (defined as $P_{2\omega}/P_{\text{pump}}$ [1]) for the second SHG process (i.e., $\omega_{780\text{nm}} + \omega_{780\text{nm}} \Rightarrow$

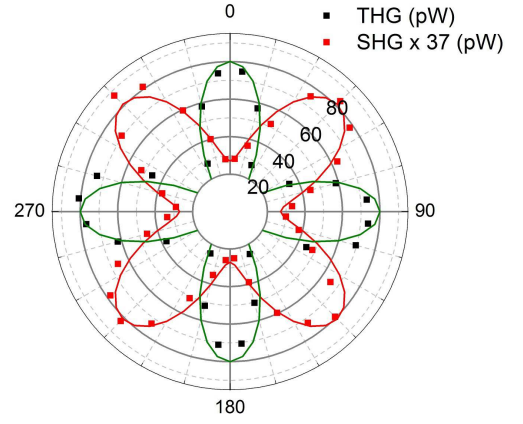


FIG. 5: Dependence of SHG and THG intensities on the elliptical polarization of the pump light in 1L-MoS₂. The polar plot angle corresponds to linearly polarized light when $\theta = 0^\circ + m \cdot 90^\circ$, and gives circularly polarized pump light when $\theta = 45^\circ + m \cdot 90^\circ$. The SHG power is multiplied by a factor of 37 to fit in the same scale as THG.

$\omega_{390\text{nm}}$) to be close to unity. However, we observe a conversion efficiency $\sim 10^{-10}$ for SHG. Therefore, we conclude that our FHG does not arise from cascaded SHGs. Another possible cascade process is based on THG ($\omega_{1560\text{nm}} + \omega_{1560\text{nm}} \Rightarrow \omega_{520\text{nm}}$) and sum-frequency generation ($\omega_{520\text{nm}} + \omega_{1560\text{nm}} \Rightarrow \omega_{390\text{nm}}$). We find that THG strongly increases up to $N=5$, as for Fig.4a). Therefore, we expect this cascaded process to have a similar trend with N . However, we only observe FHG in 1L-MoS₂. Thus, we can also exclude this cascade process, and conclude that we observe a direct $\chi^{(4)}$ process in 1L-MoS₂. To the best of our knowledge, this is the first observation of FHG in a monolayer GRM.

We now discuss the dependence of our results on the elliptical polarization of the incident light. We consider an incident laser beam with arbitrary polarization, i.e. $E = |E|\hat{\epsilon}_\pm$ with $\hat{\epsilon}_\pm = \hat{x} \cos(\theta) \pm i\hat{y} \sin(\theta)$. Using the crystal symmetries of 1L-MoS₂, we derive (see Methods) the following expressions for the second- and third-order polarizations $P^{(2)}$ and $P^{(3)}$:

$$P^{(2)} = \epsilon_0 \chi_{yyy}^{(2)} |E|^2 [\mp i \sin(2\theta) \hat{x} - \hat{y}] \quad (3)$$

and

$$P^{(3)} = \epsilon_0 \chi_{yyyy}^{(3)} |E|^3 \hat{\epsilon}_\pm \cos(2\theta). \quad (4)$$

Note that $\theta = 0^\circ$ corresponds to a linearly polarized laser along the \hat{x} direction which is perpendicular to the mirror symmetry plane of D_{3h}^1 symmetry group, while $\theta = 45^\circ$ corresponds to a circularly polarized laser. From Eq.3 we expect the intensity of SHG in response to a circularly polarized pump laser to be twice that of a linearly polarized laser. Eq.4 implies vanishing THG in response to a circularly polarized pump laser.

We thus measure the SHG and THG dependence on elliptical polarization using a linearly polarized laser and

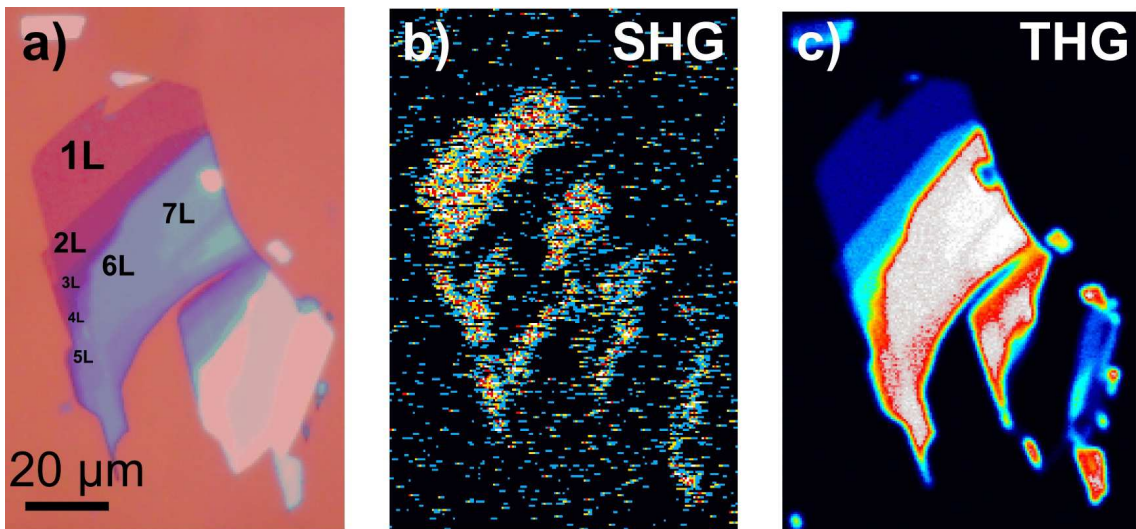


FIG. 6: a) Optical micrograph, b) SHG and c) THG images of flake with few-layer areas under 1560nm excitation.

a rotating quarter-wave plate (QWP). Depending on the angle θ between the QWP axes and the polarization, the excitation light will be linearly ($\theta = 0^\circ + m \cdot 90^\circ$) or circularly ($\theta = 45^\circ + m \cdot 90^\circ$) polarized. Fig.5 shows that the experiments are in excellent agreement with Eqs.3,4. The THG signal is maximum for a linearly polarized excitation laser, while it vanishes for circularly polarized excitation. Note that SHG is always visible, but its intensity is maximum for circularly polarized light.

Given that harmonic generation is strongly dependent on the symmetry and stacking of layers and different monolayer TMDs (e.g. WSe₂, MoSe₂), all have similar nonlinear response[16, 23, 24, 31], one could use heterostructures (e.g. MoS₂/WSe₂) to engineer SHG and other nonlinear processes for high photon-conversion efficiency for a wide range of applications requiring the generation of higher frequencies. This may lead to the use of layered materials and heterostructures for applications utilizing optical nonlinearities (e.g. all-optical devices, frequency combs, high-order harmonic generation, multiphoton microscopy and therapy etc.).

METHODS

Determination of MoS₂ thickness from SHG and THG signals

SHG and THG for few-layer MoS₂ ($N=1\dots7$) are studied on the flake in Fig.6a. SHG and THG images are shown in Figs.6b,c. At 1560 nm, the contrast between 1 and 3L areas is small, as well as the contrast between 3-, 5- and 7L regions (Fig.6b).

The THG signal increases up to $N=7$, Figs.4a,6c. On the other hand, the SHG signal (Fig.6b) is only generated in ON areas, due to symmetry[23]. Therefore, the areas

that have intensity between the 3-, 5- and 7L areas in Fig.6c but appear dark in SHG, can be identified as 4 and 6L. The dependence of the intensities of THG and SHG on N is plotted in Fig.4a). Thus, the combination of SHG and THG can be used to accurately identify N at least up to 7. The THG signal develops as a function of N . Using Maxwell's equations for a non-linear medium with thickness t and considering the slowly varying amplitude approximation[1, 57], we obtain:

$$\frac{I_{3\omega}}{I_{\text{in}}} \approx \frac{(3\omega)^2 I_{\text{in}}^2}{16n_1^3 n_3 \epsilon_0^2 c^4} \left| \chi^{(3)}(-3\omega; \omega, \omega, \omega) \right|^2 t^2 \text{sinc}^2 \left(\frac{\Delta kt}{2} \right), \quad (5)$$

where I_{in} and $I_{3\omega}$ are the intensity of the incident and THG light, respectively and $\chi^{(3)}(-3\omega; \omega, \omega, \omega)$ is the third order optical susceptibility. Note that $n_{j=1,3} = \sqrt{\epsilon^{(1)}(j\omega)}$ in which $\epsilon^{(1)}$ is the linear dielectric function of multilayer TMD. Δkt is the phase mismatch between the fundamental and third harmonic generated waves. For $\Delta kt \approx 0$, THG adds up quadratically with light propagation length (i.e. $t \propto N$). The signal starts to saturate for $N=6$. The possible reasons for sub-quadratic signal build-up can be either phase mismatch, or absorption[19]. For THG, $\Delta k = 3k_{\text{in}} \pm k_{3\omega}$, where k_{in} and $k_{3\omega}$ are the wavevectors of the incident and THG signals, respectively, where the plus sign indicates THG generated in the backward direction, while minus identifies forward generated THG. Even for backward generated THG, $\Delta kt \approx 0$ for 6L-MoS₂ (~ 4.3 nm[58]). This rules out phase mismatch as the origin of the signal saturation when $N \leq 6$. Therefore we assume that the signal saturation is due to absorption of the third harmonic light.

Continuum-model Hamiltonian and current matrix elements for 1L-MoS₂

For 1L-MoS₂ we use the low-energy $k \cdot p$ continuum-model Hamiltonian described in Ref.[40]. Around the K and K' points the model Hamiltonian contains isotropic \mathcal{H}_i and trigonal warping \mathcal{H}_{tw} contributions, i.e. $\mathcal{H} = \mathcal{H}_i + \mathcal{H}_{tw}$, with:

$$\begin{aligned} \mathcal{H}_i(k, \tau, s) &= \frac{\lambda_0 \tau s}{2} + \frac{\Delta + \lambda \tau s}{2} \sigma_z + t_0 a_0 k \cdot \sigma_\tau \\ &+ \frac{\hbar^2 |k|^2}{4m_0} (\alpha + \beta \sigma_z), \end{aligned} \quad (6)$$

and

$$\begin{aligned} \mathcal{H}_{tw}(k, \tau, s) &= t_1 a_0^2 (k \cdot \sigma_\tau^*) \sigma_x (k \cdot \sigma_\tau^*) \\ &+ t_2 a_0^3 \tau (k_x^3 - 3k_x k_y^2) (\alpha' + \beta' \sigma_z). \end{aligned} \quad (7)$$

Here, $s = \pm$ is a spin index, $\tau = \pm$ is a valley index, and $\sigma_\tau = (\tau \sigma_x, \tau \sigma_y)$, with σ_x and σ_y ordinary 2×2 Pauli matrices operating on a suitable conduction/valence band basis[40]. We note that the terms in the Hamiltonian that contain the parameters Δ , β , β' and λ_0 are related to broken spatial inversion symmetry in 1L-MoS₂. The trigonal warping term contains three parameters, α' , β' , and t_1 . The contribution to the band dispersion due to trigonal warping has the characteristic form $z_\pm \cos(3\phi)$, where $z_\pm = t_2(\alpha' \pm \beta') \pm 4t_0 t_1 / [2\Delta - (\lambda_0 - \lambda)\tau s]$, and z_+ (z_-) stands for conduction (valence) band[59]. According to *ab-initio* calculations[35, 36], symmetry considerations[36, 60], and experimental evidence[61], the valence band of 1L-MoS₂ is strongly warped, while the conduction band is nearly isotropic.

The Hamiltonian \mathcal{H} can be diagonalized. Eigenvalues $\epsilon_{k,\tau,s}^{c(v)}$ and eigenvectors $|u_{k,\tau,s}^{c(v)}\rangle$ are:

$$\epsilon_{k,\tau,s}^{c(v)} = h_0(k, \tau, s) \pm \sqrt{[h_z(k, \tau, s)]^2 + |h_{12}(k, \tau, s)|^2} \quad (8)$$

and

$$|u_{k,\tau,s}^{c(v)}\rangle = \frac{1}{\sqrt{[D^{c(v)}(k, \tau, s)]^2 + |h_{12}(k, \tau, s)|^2}} \begin{bmatrix} -h_{12}(k, \tau, s) \\ D^{c(v)}(k, \tau, s) \end{bmatrix}, \quad (9)$$

where

$$h_0(k, \tau, s) = \frac{\lambda_0}{2} \tau s + \frac{\hbar^2 k^2}{4m_0} \alpha + t_2 a_0^3 \tau (k_x^3 - 3k_x k_y^2) \alpha', \quad (10)$$

$$h_z(k, \tau, s) = \frac{\Delta + \lambda \tau s}{2} + \frac{\hbar^2 k^2}{4m_0} \beta + t_2 a_0^3 \tau (k_x^3 - 3k_x k_y^2) \beta', \quad (11)$$

$$h_{12}(k, \tau, s) = t_0 a_0 (\tau k_x - i k_y) + t_1 a_0^2 (\tau k_x + i k_y)^2, \quad (12)$$

and

$$\begin{aligned} D^{c(v)}(k, \tau, s) &= h_z(k, \tau, s) \\ &\mp \sqrt{[h_z(k, \tau, s)]^2 + |h_{12}(k, \tau, s)|^2}. \end{aligned} \quad (13)$$

We need the matrix elements of the current operator for the evaluation of the nonlinear response functions. We start by introducing the so-called paramagnetic current operator[62] ($c = 1$, where c is the speed of light, $-e < 0$ is the electron charge):

$$j_\ell(k) \equiv - \left. \frac{\delta \mathcal{H}(k + eA/\hbar)}{\delta A_\ell} \right|_{A=0} = - \frac{e}{\hbar} \frac{\partial \mathcal{H}}{\partial k_\ell}, \quad (14)$$

where $\ell = x, y$ is a Cartesian index. The diamagnetic contributions to the current operator can be written as follows[63]:

$$\kappa_{\ell_1 \ell_2}(k) \equiv - \left. \frac{\delta^2 \mathcal{H}(k + eA/\hbar)}{\delta A_{\ell_1} \delta A_{\ell_2}} \right|_{A=0} = - \left(\frac{e}{\hbar} \right)^2 \frac{\partial^2 \mathcal{H}}{\partial k_{\ell_1} \partial k_{\ell_2}} \quad (15)$$

and

$$\xi_{\ell_1 \ell_2 \ell_3}(k) \equiv - \left. \frac{\delta^3 \mathcal{H}(k + eA/\hbar)}{\delta A_{\ell_1} \delta A_{\ell_2} \delta A_{\ell_3}} \right|_{A=0} \quad (16)$$

$$= - \left(\frac{e}{\hbar} \right)^3 \frac{\partial^3 \mathcal{H}}{\partial k_{\ell_1} \partial k_{\ell_2} \partial k_{\ell_3}} \quad (17)$$

Using the continuum-model Hamiltonian introduced in Eqs. (6) and (7), we find:

$$j_\ell = - \frac{e}{\hbar} \left\{ \frac{\partial h_0}{\partial k_\ell} + \frac{\partial h_z}{\partial k_\ell} \sigma_z + \text{Re} \left[\frac{\partial h_{12}}{\partial k_\ell} \right] \sigma_x - \text{Im} \left[\frac{\partial h_{12}}{\partial k_\ell} \right] \sigma_y \right\} \quad (18)$$

and

$$\begin{aligned} \kappa_{\ell\ell} &= - \left(\frac{e}{\hbar} \right)^2 \left\{ \frac{\partial^2 h_0}{\partial k_\ell^2} + \frac{\partial^2 h_z}{\partial k_\ell^2} \sigma_z + \text{Re} \left[\frac{\partial^2 h_{12}}{\partial k_\ell^2} \right] \sigma_x \right. \\ &\quad \left. - \text{Im} \left[\frac{\partial^2 h_{12}}{\partial k_\ell^2} \right] \sigma_y \right\}. \end{aligned} \quad (19)$$

Similarly, one can derive an explicit expression for $\xi_{\ell\ell\ell}$.

The required matrix elements of j_ℓ and $\kappa_{\ell\ell}$ between the eigenspinors (9) are given by:

$$j_\ell^{\text{cv}}(k, \tau, s) \equiv \langle u_{k,\tau,s}^{\text{c}} | j_\ell | u_{k,\tau,s}^{\text{v}} \rangle = \frac{e}{\hbar} \left\{ \frac{h_z(k, \tau, s) \text{Re} [h_{12}(k, \tau, s) \partial h_{12}^*(k, \tau, s) / \partial k_\ell]}{|h_{12}(k, \tau, s)| \sqrt{[h_z(k, \tau, s)]^2 + |h_{12}(k, \tau, s)|^2}} + i \frac{\text{Im} [h_{12}(k, \tau, s) \partial h_{12}^*(k, \tau, s) / \partial k_\ell]}{|h_{12}(k, \tau, s)|} - \frac{|h_{12}(k, \tau, s)| \partial h_z(k, \tau, s) / \partial k_\ell}{\sqrt{[h_z(k, \tau, s)]^2 + |h_{12}(k, \tau, s)|^2}} \right\}, \quad (20)$$

$$j_\ell^{\text{cc(vv)}}(k, \tau, s) \equiv \langle u_{k,\tau,s}^{\text{c(v)}} | j_\ell | u_{k,\tau,s}^{\text{c(v)}} \rangle = -\frac{e}{\hbar} \left\{ \frac{\partial h_0(k, \tau, s)}{\partial k_\ell} \pm \frac{h_z(k, \tau, s) \partial h_z(k, \tau, s) / \partial k_\ell + \text{Re} [h_{12}(k, \tau, s) \partial h_{12}^*(k, \tau, s) / \partial k_\ell]}{\sqrt{[h_z(k, \tau, s)]^2 + |h_{12}(k, \tau, s)|^2}} \right\}, \quad (21)$$

$$\kappa_{\ell\ell}^{\text{cv}}(k, \tau, s) \equiv \langle u_{k,\tau,s}^{\text{c}} | \kappa_{\ell\ell} | u_{k,\tau,s}^{\text{v}} \rangle = \left(\frac{e}{\hbar} \right)^2 \left\{ \frac{h_z(k, \tau, s) \text{Re} [h_{12}(k, \tau, s) \partial^2 h_{12}^*(k, \tau, s) / \partial k_\ell^2]}{|h_{12}(k, \tau, s)| \sqrt{[h_z(k, \tau, s)]^2 + |h_{12}(k, \tau, s)|^2}} + i \frac{\text{Im} [h_{12}(k, \tau, s) \partial^2 h_{12}^*(k, \tau, s) / \partial k_\ell^2]}{|h_{12}(k, \tau, s)|} - \frac{|h_{12}(k, \tau, s)| \partial^2 h_z(k, \tau, s) / \partial k_\ell^2}{\sqrt{[h_z(k, \tau, s)]^2 + |h_{12}(k, \tau, s)|^2}} \right\}, \quad (22)$$

and

$$\kappa_{\ell\ell}^{\text{cc(vv)}}(k, \tau, s) \equiv \langle u_{k,\tau,s}^{\text{c(v)}} | \kappa_{\ell\ell} | u_{k,\tau,s}^{\text{c(v)}} \rangle = -\left(\frac{e}{\hbar} \right)^2 \left\{ \frac{\partial^2 h_0(k, \tau, s)}{\partial k_\ell^2} \pm \frac{h_z(k, \tau, s) \partial^2 h_z(k, \tau, s) / \partial k_\ell^2 + \text{Re} [h_{12}(k, \tau, s) \partial^2 h_{12}^*(k, \tau, s) / \partial k_\ell^2]}{\sqrt{[h_z(k, \tau, s)]^2 + |h_{12}(k, \tau, s)|^2}} \right\}. \quad (23)$$

We note that intra-band matrix elements (e.g. j_y^{cc} and κ_{yy}^{cc}) have a definite parity while inter-band ones (e.g. j_y^{cv} and κ_{yy}^{cv}) do not. This is at the origin of the vanishing of the paramagnetic contribution to even harmonic-generation response functions. Therefore, as we will see later, only diamagnetic terms yield a finite contribution to even harmonic-generation responses.

General symmetry considerations

Our continuum-model Hamiltonian is derived from a tight-binding Hamiltonian in which the zigzag direction of the lattice coincides with the \hat{x} direction. The zigzag direction is perpendicular to the reflection (mirror) symmetry plane of the 1L-MoS₂ lattice (see Fig.7).

The n -th order optical susceptibilities $\chi_{\ell i_1 i_2 \dots i_n}^{(n)}$ are defined as:

$$P_\ell^{(n)}(\omega_\Sigma) = \epsilon_0 \sum_{i_1 i_2 \dots i_n} \chi_{\ell i_1 i_2 \dots i_n}^{(n)}(-\omega_\Sigma; \omega_1, \omega_2, \dots, \omega_n) \times E_{i_1}(\omega_1) E_{i_2}(\omega_2) \dots E_{i_n}(\omega_n), \quad (24)$$

where E_i and $P_\ell^{(n)}$ are the Cartesian components of the

electric field E and the n -th order macroscopic polarization $P^{(n)}$, respectively, and ϵ_0 is the vacuum permittivity. Note that i_1, i_2, \dots, i_n are Cartesian indices and $\omega_\Sigma \equiv \sum_i \omega_i$.

Since 1L-MoS₂ belongs to the D_{3h}^1 symmetry group, the only non-vanishing elements of the second-order susceptibility are[1]:

$$\chi_{yyy}^{(2)} = -\chi_{yxx}^{(2)} = -\chi_{xxy}^{(2)} = -\chi_{xyx}^{(2)}, \quad (25)$$

while for the case of the third-order response we have[1]:

$$\chi_{yyyy}^{(3)} = \chi_{xxxx}^{(3)} = \chi_{yyxx}^{(3)} + \chi_{xyxy}^{(3)} + \chi_{yxxy}^{(3)}, \quad (26)$$

and

$$\begin{aligned} \chi_{xxyy}^{(3)} &= \chi_{yyxx}^{(3)}, \\ \chi_{xyyy}^{(3)} &= \chi_{yxyx}^{(3)}, \\ \chi_{xyxy}^{(3)} &= \chi_{yxxy}^{(3)}. \end{aligned} \quad (27)$$

In the case of a linearly-polarized pump laser, we expect a SHG maximum when the laser is polarized along the \hat{y} direction, i.e. perpendicular to the zigzag direction. On the contrary, if the incident light is polarized

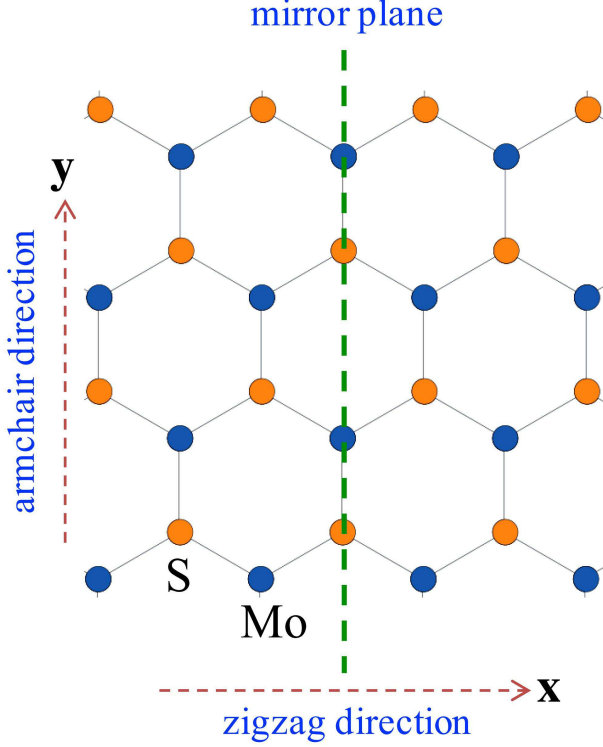


FIG. 7: Top view of the 1L-MoS₂ lattice.

along the \hat{x} direction, i.e. the zigzag direction, we expect a vanishing SHG signal due to the reflection symmetry (i.e. $\sigma_v : x \rightarrow -x$) along this axis. Our continuum-model Hamiltonian is consistent with these general expectations based on symmetry. We find $\chi_{xxx}^{(2)} = 0$, even in the presence of trigonal warping, because the contribution in the two valleys identically cancel each other.

Using Eqs.24,25,26,27 we obtain Eqs.3,4 of the main text, which describe the dependence between induced charge polarization, P , and the polarization of the incident laser. In the case of a circularly-polarized pump laser, we have $E = |E|\hat{\varepsilon}_{\pm}$ with $\hat{\varepsilon}_{\pm} = (\hat{x} \pm i\hat{y})/\sqrt{2}$. Using Eqs.3,4 of the main text we arrive at the following results for the circularly-polarized pump laser:

$$P^{(2)} = \mp i\sqrt{2}\epsilon_0\chi_{yyy}^{(2)}|E|^2\hat{\varepsilon}_{\mp} \quad (28)$$

and

$$P^{(3)} = 0. \quad (29)$$

Eq.28 implies an *opposite* polarization of the SHG signal with respect to the laser, while Eq.29 implies no THG signal in response to a circularly-polarized pump laser.

For quantitative results, only the three tensor elements $\chi_{yyy}^{(2)}$, $\chi_{yyy}^{(3)}$ and $\chi_{yyy}^{(4)}$ are required for second-, third- and fourth-order nonlinear response functions.

Nonlinear response functions

The response of an electron system to light can be calculated by adopting different gauges for describing the electric field of incident light. The gauge in which a uniform electric field $E(t)$ is described in terms of a uniform time-dependent vector potential, $E(t) = -\partial A(t)/\partial t$, is convenient in solids as it does not break Bloch translational invariance. The vector potential couples to matter degrees of freedom through the minimal coupling, i.e. $k \rightarrow k + eA/\hbar$. The external vector potential induces a current $J(t)$, which can be expanded in a power series of $A(t)$. For each Cartesian component, $J_{\ell} = \sum_n J_{\ell}^{(n)}$ where n denotes the n -th order in powers of $A(t)$. In Fourier transform with respect to time we get:

$$J_{\ell}^{(n)}(\omega_{\Sigma}) \equiv \sum_{i_1, i_2, \dots, i_n} \Pi_{\ell i_1 i_2 \dots i_n}^{(n)}(-\omega_{\Sigma}; \omega_1, \omega_2, \dots, \omega_n) \times A_{i_1}(\omega_1)A_{i_2}(\omega_2) \dots A_{i_n}(\omega_n), \quad (30)$$

where $A(\omega_i) = -iE(\omega_i)/(\omega_i + i\eta/\hbar)$ and η is an infinitesimal positive real number, needed to ensure that the external field is absent in the remote past ($t \rightarrow -\infty$).

Since the macroscopic current is related to the macroscopic polarization by $J(t) = \partial P/\partial t$ [64], we get $J^{(n)}(\omega_{\Sigma}) = -i(\omega_{\Sigma} + i\eta/\hbar)P^{(n)}(\omega_{\Sigma})$, for each order in perturbation theory.

We finally find the following relation between nonlinear response functions and optical susceptibilities:

$$\epsilon_0\chi_{\ell i_1 i_2 \dots i_n}^{(n)}(-\omega_{\Sigma}; \omega_1, \dots, \omega_n) = i(-i)^n \times \frac{\Pi_{\ell i_1 i_2 \dots i_n}^{(n)}(-\omega_{\Sigma}; \omega_1, \dots, \omega_n)}{(\omega_{\Sigma} + i\eta/\hbar)(\omega_n + i\eta/\hbar) \dots (\omega_1 + i\eta/\hbar)}. \quad (31)$$

The n -th order nonlinear response $\Pi_{\ell i_1 i_2 \dots i_n}^{(n)}$ contains both paramagnetic and diamagnetic current contributions, which will be denoted by $\Pi_{\ell i_1 i_2 \dots i_n}^{(n),P}$ and $\Pi_{\ell i_1 i_2 \dots i_n}^{(n),D}$, respectively. The paramagnetic current correlators, which are diagrammatically illustrated in Fig. 8, read:

$$\Pi_{\ell i_1}^{(1),P}(i\nu) \equiv \langle \hat{j}_{i_1}(-i\nu)\hat{j}_{\ell}(i\nu) \rangle, \quad (32)$$

$$\Pi_{\ell i_1 i_2}^{(2),P}(-i\nu_{\Sigma}; i\nu_1, i\nu_2) \equiv \sum_{\mathcal{P}} \langle \hat{j}_{i_1}(-i\nu_1)\hat{j}_{i_2}(-i\nu_2)\hat{j}_{\ell}(i\nu_{\Sigma}) \rangle, \quad (33)$$

$$\begin{aligned} & \Pi_{\ell i_1 i_2 i_3}^{(3),P}(-i\nu_{\Sigma}; i\nu_1, i\nu_2, i\nu_3) \\ & \equiv \sum_{\mathcal{P}} \langle \hat{j}_{i_1}(-i\nu_1)\hat{j}_{i_2}(-i\nu_2)\hat{j}_{i_3}(-i\nu_3)\hat{j}_{\ell}(i\nu_{\Sigma}) \rangle, \end{aligned} \quad (34)$$

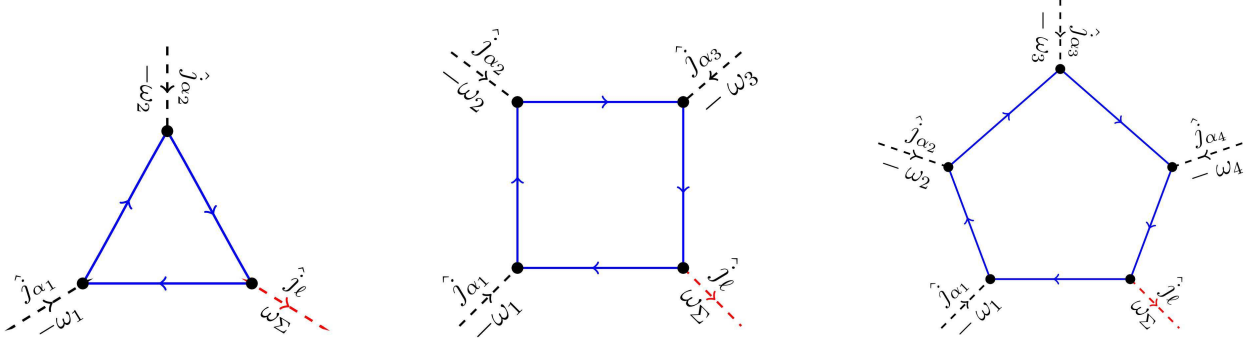


FIG. 8: Three-, four-, and five-leg Feynman diagrams for the second-, third-, and fourth-order nonlinear *paramagnetic* response functions. Solid lines denote electron propagators while dashed lines denote photons. The quantities $\omega_1 = \dots = \omega_n = \omega$ indicate the incoming photon frequencies, while \hat{j}_α denotes the α -th Cartesian component of the paramagnetic current operator.

and

$$\begin{aligned} & \Pi_{i_1 i_2 i_3 i_4}^{(4),P}(-i\nu_\Sigma; i\nu_1, i\nu_2, i\nu_3, i\nu_4) \\ & \equiv \sum_{\mathcal{P}} \left\langle \hat{j}_{i_1}(-i\nu_1) \hat{j}_{i_2}(-i\nu_2) \hat{j}_{i_3}(-i\nu_3) \hat{j}_{i_4}(-i\nu_4) \hat{j}_\ell(i\nu_\Sigma) \right\rangle. \end{aligned} \quad (35)$$

Here, $\langle \dots \rangle$ denotes the thermal averaging[62, 65], \hat{j}_i indicates the second-quantized form of i -th Cartesian component of the paramagnetic current operator, $\sum_{\mathcal{P}}$ enforces the so-called “*intrinsic permutation symmetry*” among all dummy variables (i_n, ν_n) [57], and $\nu_\Sigma = \sum_i \nu_i$, where $\nu_i = 2\pi n/\beta$ ’s are bosonic Matsubara energies corresponding to the incident photon energies. n is a relative integer. $\beta = 1/(k_B T)$, with T the electron temperature.

The paramagnetic current correlators in Eqs.32-35 can be calculated by using the many-body diagrammatic perturbation theory[39, 66]. Following Ref.[39], we first sum over the Fermionic Matsubara energies and then carry out the analytical continuation $\nu_i = \nu \rightarrow \hbar\omega + i\eta$ where $\eta \rightarrow 0^+$. We find the following relations for the case of $\ell = i_n = y$:

$$\Pi_{yy}^{(1),P}(\omega) = \sum_{k,\tau,s} \sum_{\{\lambda_i\}} U_{\lambda_1 \lambda_2} j_y^{\lambda_2 \lambda_1} j_y^{\lambda_1 \lambda_2}, \quad (36)$$

$$\begin{aligned} \Pi_{yyy}^{(2),P}(-2\omega; \omega, \omega) &= \sum_{k,\tau,s} \sum_{\{\lambda_i\}} \frac{j_y^{\lambda_3 \lambda_2} j_y^{\lambda_2 \lambda_1} j_y^{\lambda_1 \lambda_3}}{2(\hbar\omega + i\eta) + \epsilon_{k,\tau,s}^{\lambda_1} - \epsilon_{k,\tau,s}^{\lambda_3}} \\ &\times (U_{\lambda_1 \lambda_2} - U_{\lambda_2 \lambda_3}), \end{aligned} \quad (37)$$

$$\Pi_{yyyy}^{(3),P}(-3\omega; \omega, \omega, \omega) = \sum_{k,\tau,s} \sum_{\{\lambda_i\}} \frac{j_y^{\lambda_4 \lambda_3} j_y^{\lambda_3 \lambda_2} j_y^{\lambda_2 \lambda_1} j_y^{\lambda_1 \lambda_4}}{3(\hbar\omega + i\eta) + \epsilon_k^{\lambda_1} - \epsilon_k^{\lambda_4}} \left\{ \frac{U_{\lambda_1 \lambda_2} - U_{\lambda_2 \lambda_3}}{2(\hbar\omega + i\eta) + \epsilon_{k,\tau,s}^{\lambda_1} - \epsilon_{k,\tau,s}^{\lambda_3}} - \frac{U_{\lambda_2 \lambda_3} - U_{\lambda_3 \lambda_4}}{2(\hbar\omega + i\eta) + \epsilon_{k,\tau,s}^{\lambda_2} - \epsilon_{k,\tau,s}^{\lambda_4}} \right\}, \quad (38)$$

and

$$\begin{aligned} \Pi_{yyyyy}^{(4),P}(-4\omega; \omega, \omega, \omega, \omega) &= \sum_{k,\tau,s} \sum_{\{\lambda_i\}} \frac{j_y^{\lambda_5 \lambda_4} j_y^{\lambda_4 \lambda_3} j_y^{\lambda_3 \lambda_2} j_y^{\lambda_2 \lambda_1} j_y^{\lambda_1 \lambda_5}}{4(\hbar\omega + i\eta) + \epsilon_{k,\tau,s}^{\lambda_1} - \epsilon_{k,\tau,s}^{\lambda_5}} \\ &\left\{ \frac{1}{3(\hbar\omega + i\eta) + \epsilon_{k,\tau,s}^{\lambda_1} - \epsilon_{k,\tau,s}^{\lambda_4}} \left[\frac{U_{\lambda_1 \lambda_2} - U_{\lambda_2 \lambda_3}}{2(\hbar\omega + i\eta) + \epsilon_{k,\tau,s}^{\lambda_1} - \epsilon_{k,\tau,s}^{\lambda_3}} - \frac{U_{\lambda_2 \lambda_3} - U_{\lambda_3 \lambda_4}}{2(\hbar\omega + i\eta) + \epsilon_{k,\tau,s}^{\lambda_2} - \epsilon_{k,\tau,s}^{\lambda_4}} \right] \right. \\ &\left. - \frac{1}{3(\hbar\omega + i\eta) + \epsilon_{k,\tau,s}^{\lambda_2} - \epsilon_{k,\tau,s}^{\lambda_5}} \left[\frac{U_{\lambda_2 \lambda_3} - U_{\lambda_3 \lambda_4}}{2(\hbar\omega + i\eta) + \epsilon_{k,\tau,s}^{\lambda_2} - \epsilon_{k,\tau,s}^{\lambda_4}} - \frac{U_{\lambda_3 \lambda_4} - U_{\lambda_4 \lambda_5}}{2(\hbar\omega + i\eta) + \epsilon_{k,\tau,s}^{\lambda_3} - \epsilon_{k,\tau,s}^{\lambda_5}} \right] \right\}. \end{aligned} \quad (39)$$

For simplicity, we introduce the quantity $U_{\lambda\lambda'}$ as follows:

$$U_{\lambda\lambda'}(k, \omega, \tau, s) \equiv \frac{1}{\mathcal{S}} \frac{n_F(\epsilon_{k,\tau,s}^\lambda) - n_F(\epsilon_{k,\tau,s}^{\lambda'})}{\hbar\omega + \epsilon_{k,\tau,s}^\lambda - \epsilon_{k,\tau,s}^{\lambda'} + i\eta}, \quad (40)$$

where \mathcal{S} is the sample area, $\lambda, \lambda' = c, v$, and

$$n_F(E) = \left\{ \exp\left(\frac{E - \mu}{k_B T}\right) + 1 \right\}^{-1} \quad (41)$$

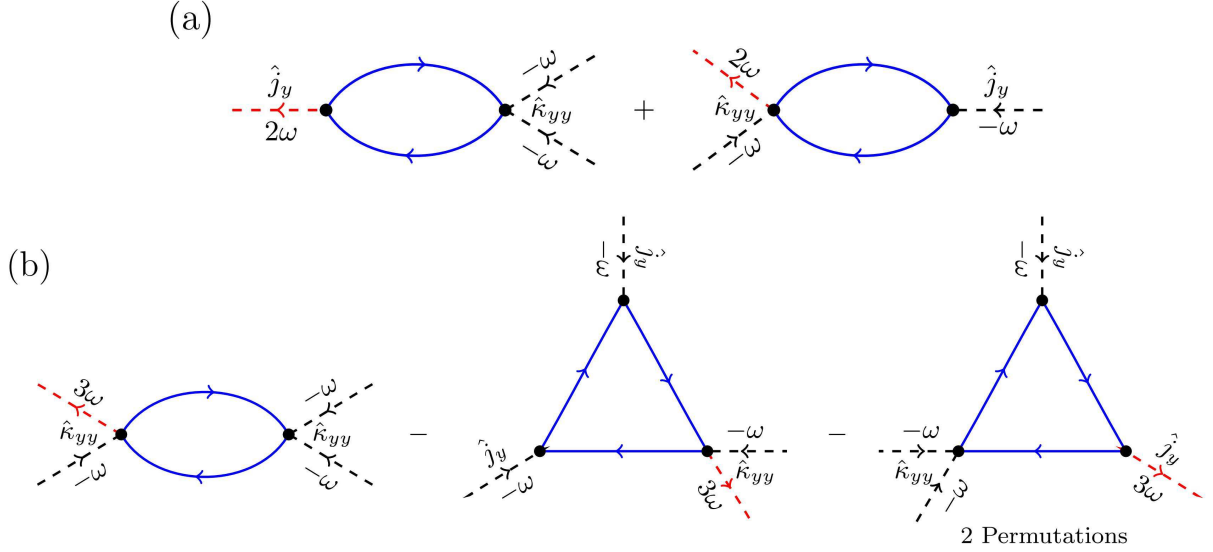


FIG. 9: Feynman diagrams for the *diamagnetic* contributions to the second- and third-order response functions. a) second-order response. b) third-order response.

is the Fermi-Dirac distribution function at finite temperature T and chemical potential μ . In Eqs.36-39 we dropped the explicit functional dependence on k, τ, s , e.g. $j_y^{mn} = j_y^{mn}(k, \tau, s)$ and $U_{mn} = U_{mn}(k, \omega, \tau, s)$. We find most convenient to first carry out the sum over the band indices λ_i and then carry out numerically the integral over the wave vector k .

The paramagnetic contributions to the even-order response functions, $\Pi_{yyy}^{(2),P}$ and $\Pi_{yyyy}^{(4),P}$, vanish because $\epsilon_{k,\tau,s}^{c(v)}$ is an even function of k_y . This property of the energy dispersion is protected by symmetry, and stems from time-reversal (\mathcal{T}) and reflection (σ_v) symmetries.

A microscopic calculation of even-order response functions requires the knowledge of diamagnetic contributions. These can be included with the aid of correlation functions involving the $\hat{\kappa}_{yy}$ operator. In fact, $\hat{\xi}_{yyy}$ could also contribute to diamagnetic responses. However, in our low-energy model $\hat{\xi}_{yyy}$ is identically zero. Similar to the paramagnetic case, $\hat{\kappa}_{yy}$ and $\hat{\xi}_{yyy}$ indicate the second-quantized form of the diamagnetic current operators (i.e. κ_{yy} and ξ_{yyy}). Diamagnetic contributions to the second- and third-order response functions are reported in Fig.9, in terms of Feynman diagrams. For the sake of simplicity, we have not calculated diamagnetic contributions to the fourth-order response.

According to Fig.9a, the diamagnetic contribution to the second-order response is given by:

$$\begin{aligned} \Pi_{yyy}^{(2),D}(-2\omega; \omega, \omega) = & - \sum_{k,\tau,s} \sum_{\{\lambda_i\}} \left[U_{\lambda_1\lambda_2} j_y^{\lambda_1\lambda_2} \kappa_{yy}^{\lambda_2\lambda_1} \right. \\ & \left. + \tilde{U}_{\lambda_1\lambda_2} \kappa_{yy}^{\lambda_1\lambda_2} j_y^{\lambda_2\lambda_1} \right]. \end{aligned} \quad (42)$$

Similarly, the diamagnetic contribution to the third-order response, Fig.9b, is given by:

$$\begin{aligned} \Pi_{yyy}^{(3),D}(-3\omega; \omega, \omega, \omega) = & \sum_{k,\tau,s} \sum_{\{\lambda_i\}} \left\{ \tilde{U}_{\lambda_1\lambda_2} \kappa_{yy}^{\lambda_1\lambda_2} \kappa_{yy}^{\lambda_2\lambda_1} \right. \\ & - \frac{j_y^{\lambda_3\lambda_2} j_y^{\lambda_2\lambda_1} \kappa_{yy}^{\lambda_1\lambda_3}}{2(\hbar\omega + i\eta) + \epsilon_{k,\tau,s}^{\lambda_1} - \epsilon_{k,\tau,s}^{\lambda_3}} (U_{\lambda_1\lambda_2} - U_{\lambda_2\lambda_3}) \\ & \left. - \sum_{\mathcal{P}} \frac{\kappa_{yy}^{\lambda_3\lambda_2} j_y^{\lambda_2\lambda_1} j_y^{\lambda_1\lambda_3}}{3(\hbar\omega + i\eta) + \epsilon_{k,\tau,s}^{\lambda_1} - \epsilon_{k,\tau,s}^{\lambda_3}} (\tilde{U}_{\lambda_1\lambda_2} - U_{\lambda_2\lambda_3}) \right\}. \end{aligned} \quad (43)$$

Here, $\tilde{U}_{\lambda_1\lambda_2} = U_{\lambda_1\lambda_2}(k, 2\omega, \tau, s)$ with $\kappa_{yy}^{mn} = \kappa_{yy}^{mn}(k, \tau, s)$ is the matrix element of κ_{yy} .

Since our low-energy model is valid for a limited range of values of the wave vector k , we must introduce an ultra-violet cut-off, which breaks gauge invariance[67]. We therefore need to regularize our final results to avoid unphysical response function. This can be accomplished[67] by considering the following gauge-regularized response tensors: $\Pi_{\ell_1 i_2 \dots i_n}^{(n)} \equiv \Pi_{\ell_1 i_2 \dots i_n}^{(n)} - \Pi_{\ell_1 i_2 \dots i_n}^{(n)}|_{\{\omega_i\} \rightarrow 0}$.

We note that the summands in Eqs.37,39,42 contain an *odd* number of matrix elements of the paramagnetic (j_y) and diamagnetic (κ_{yy}) current operators. In the absence of trigonal warping, the overall form-factor, which is proportional to these matrix elements, is an odd function of k_y : we therefore conclude that, *in the absence of trigonal warping*, $\Pi_{yyy}^{(2)}(-2\omega; \omega, \omega) = \Pi_{yyyy}^{(4)}(-4\omega; \omega, \omega, \omega, \omega) = 0$. An identical conclusion was reached for other isotropic low-energy continuum

model Hamiltonians, such as those describing gapped graphene[68] and biased bilayer graphene[69, 70]. We expect the second-order nonlinear response function $\Pi_{yyy}^{(2)}$ to be small compared to the third-order one, since it is controlled by a small trigonal warping correction (\mathcal{H}_{tw}) in comparison with the fully isotropic leading term (\mathcal{H}_i) in the low-energy model Hamiltonian. Of course, this conclusion is valid within the single-particle picture and in the low-energy limit, which we have relied on so far.

Relative magnitude of nonlinear responses: ratios of irradiances

To quantify the *relative* magnitude of nonlinear harmonic signals, we calculate ratios between induced polarizations $P_y^{(n)}$ at different orders n in perturbation theory. For a linearly-polarized laser (e.g. $E = |E|\hat{y}$):

$$\begin{aligned} \left| \frac{P_y^{(n+1)}}{P_y^{(n)}} \right| &= \left| \frac{\chi_{y\dots y}^{(n+1)} |E|}{\chi_{y\dots y}^{(n)}} \right| = \left| \frac{\Pi_{y\dots y}^{(n+1)} / \Pi_0^{(n+1)}}{(\hbar\omega + i\eta)/(eV) \times \Pi_{y\dots y}^{(n)} / \Pi_0^{(n)}} \right| \times \left(\frac{n\Pi_0^{(n+1)}\hbar}{(n+1)\Pi_0^{(n)}(eV)} \right) \times |E| \\ &= \frac{n}{n+1} \times \frac{t_0}{eV} \times \frac{a_0}{m} \times \frac{|E|}{V/m} \times X_{n+1,n}(\omega), \end{aligned} \quad (44)$$

where

$$\begin{aligned} \Pi_0^{(n)} &\equiv \frac{(et_0a_0/\hbar)^{n+1}}{8\pi a_0^2(eV)^n} = \frac{(eV)m^{n-1}}{8\pi} \left(\frac{t_0}{eV} \right)^{n+1} \\ &\quad \times \left(\frac{a_0}{m} \right)^{n-1} \left(\frac{e}{\hbar} \right)^{n+1} \end{aligned} \quad (45)$$

and the quantities t_0 and a_0 have been introduced in the Hamiltonian \mathcal{H} . Leaving aside pre-factors, Π_0 represents the physical dimensions of the nonlinear current correlator $\Pi_{i_1 i_2 \dots i_n}^{(n)}(-\omega_\Sigma; \omega_1, \omega_2, \dots, \omega_n)$. In the SI system, the unit of $\Pi_0^{(n)}$ is $\text{C m}^{n-1} \text{V}^{-n} \text{s}^{-(n+1)}$. The dimensionless quantities $X_{n+1,n}$ are given by:

$$X_{n+1,n}(\omega) = \left| \frac{\Pi_{y\dots y}^{(n+1)} / \Pi_0^{(n+1)}}{(\hbar\omega + i\eta)/(eV) \times \Pi_{y\dots y}^{(n)} / \Pi_0^{(n)}} \right|. \quad (46)$$

The amplitude of the electric field ($|E|$) in Eq.44 can be replaced by the power of the pump laser (P_{pump}) by using the following relation:

$$\frac{P_{\text{pump}}}{\pi(D/2)^2} = \frac{1}{2} n_r c \epsilon_0 |E|^2, \quad (47)$$

where $D \approx 1.85 \mu\text{m}$ is the experimental spot size diameter, $n_r \approx 1$ is the refractive index of air, $c \approx 3 \times 10^8 \text{ m/s}$ is the speed of light in vacuum, and $\epsilon_0 \approx 8.85 \times 10^{-12} \text{ C/(Vm)}$ is the vacuum electrical permittivity. Using Maxwell's equations, we can obtain the following wave equation in a nonlinear medium[1]:

$$\nabla^2 E^{(n)} + \left(\frac{\omega_n}{c} \right)^2 \epsilon^{(1)}(\omega_n) \cdot E^{(n)} = -\frac{1}{\epsilon_0} \left(\frac{\omega_n}{c} \right)^2 P^{(n)}. \quad (48)$$

where $n = 2, 3, \dots$ indicates the order of nonlinearity, $\epsilon^{(1)}$ is the linear dielectric tensor and $P^{(n)}$ is the n -th order polarization vector. The intensity $I^{(n)}$ of the n -th order nonlinear signal is proportional to the square of the induced electric field amplitude $E^{(n)} \propto \omega_n^2 P_y^{(n)}$ where $\omega_n = n\omega$ for the harmonic generation case. Replacing Eq.47 in Eq.44 we find:

$$\frac{I^{(n+1)}}{I^{(n)}} = \left(\frac{n+1}{n} \right)^2 \left| \frac{P_y^{(n+1)}}{P_y^{(n)}} \right|^2 = R_{n+1,n}(\omega) P_{\text{pump}}, \quad (49)$$

where $R_{n+1,n}(\omega)$ in units of $1/W$ is given by:

$$\begin{aligned} R_{n+1,n}(\omega) &= \frac{8[\text{m/s}][\text{C}/(\text{Vm})]}{\pi n_r c \epsilon_0} \left[\frac{t_0/(eV) \times a_0/m}{D/m} \right]^2 \\ &\quad \times [X_{n+1,n}(\omega)]^2. \end{aligned} \quad (50)$$

If we assume that the spot size of different harmonic-generated signals on the detector are equal to each other, we can write the following relation between power and intensity ratios:

$$\frac{I^{(n+1)}}{I^{(n)}} \approx \frac{P_{(n+1)\omega}}{P_{n\omega}}, \quad (51)$$

where $P_{n\omega}$ denotes the signal power of the n -th harmonic-generated signal.

Our main results for nonlinear response functions of 1L-MoS₂ are summarized in Figs.10-12. We use the following values for the parameters of the model: $\Delta = 1.82 \text{ eV}$, $\lambda_0 = 69 \text{ meV}$, $\lambda = -80 \text{ meV}$, $t_0 = 2.34 \text{ eV}$, $\alpha = -0.01$, $\beta = -1.54$, $t_1 = -0.14 \text{ eV}$, $t_2 = 1 \text{ eV}$, $\alpha' = 0.44$, and $\beta' = -0.53$. These parameters are obtained from a tight-binding fitting[40] of LDA-DFT band

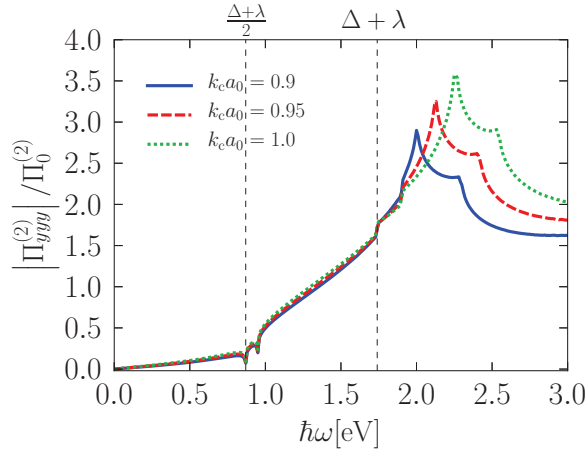


FIG. 10: Frequency dependence of the second-order response function $\Pi_{yyy}^{(2)}$ (in units of $\Pi_0^{(2)}$). Different curves refer to different values of the parameter k_c .

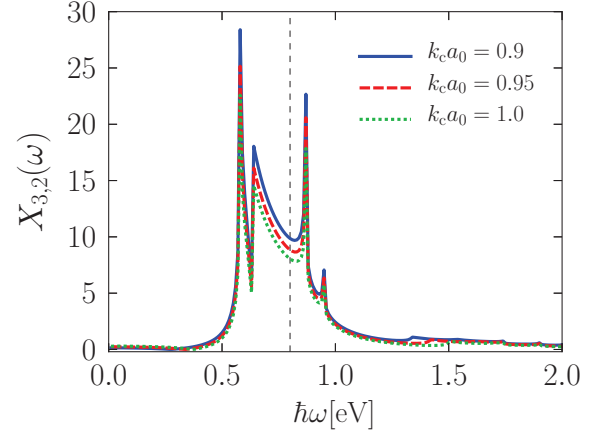


FIG. 12: Results for the $X_{3,2}$ as function of the pump laser frequency. Vertical dashed lines is positioned at $\hbar\omega = 0.8$ eV.

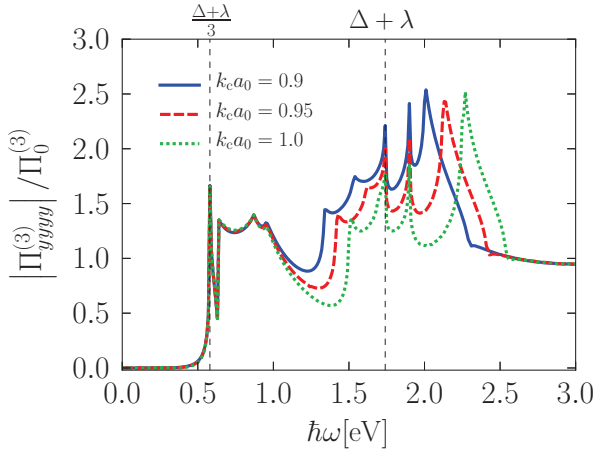


FIG. 11: Same as in Fig.10, but for the case of the third-order response function.

structure calculations[71, 72]. In all our numerical results, we use $T = 300$ K and $\mu = 0$. In Figs.10-12, we check the dependence of our results on the value of the ultra-violet cut-off, $k_c \propto 1/a_0$. Note that $a_0 = a/\sqrt{3}$ with $a \approx 3.16$ Å is the lattice constant of 1L-MoS₂.

According to Figs.10,11, the nonlinear response functions start to grow when $\hbar\omega$ is larger than $(\Delta + \lambda)/2$ and $(\Delta + \lambda)/3$ for the SHG and THG cases, respectively. $\Delta + \lambda$ is the optical band gap of MoS₂. Moreover, in the frequency range of our interest (< 1 eV) the spectra of the second and third order response functions are not very sensitive to the value of k_c . The theoretical result shown in Fig.4c) of the main text is obtained by using Eqs.50,49 for $\hbar\omega = 0.8$ eV.

ACKNOWLEDGMENTS

We thank M. J. Huttunen and R. W. Boyd for useful discussions. We acknowledge funding from the Academy of Finland (No.:276376, 284548), TEKES (NP-Nano,OPEC), Fondazione Istituto Italiano di Tecnologia, the Graphene Flagship, ERC grant Hetero2D, Nokia Foundation, EPSRC grants EP/K01711X/1, EP/K017144/1, EP/L016087/1, the AFOSR COMAS MURI (FA9550-10-1-0558), the ONR NECom MURI, the CIAN NSF ERC under grant EEC-0812072, and TRIF Photonics funding from the state of Arizona and the Micronova, Nanofabrication Centre of Aalto University.

-
- [1] Boyd, R. W. *Nonlinear Optics*, (Academic Press, 2003).
 - [2] Zipfel, W. R., Williams, R. M. & Webb, W. W. Nonlinear magic: multiphoton microscopy in the biosciences. *Nat. Biotechnol.* **21**, 1369 (2003).
 - [3] Bhawalkar, J. D., He, G. S. & Prasad, P. N. Nonlinear multiphoton processes in organic and polymeric materials. *Rep. Prog. Phys.* **59**, 1041 (1996).
 - [4] Broderick, N. G. R., Bratfalean, R. T., Monro, T. M. Richardson, D. J. & de Sterke, C. M. Temperature and wavelength tuning of second-, third-, and fourth-harmonic generation in a two-dimensional hexagonally poled nonlinear crystal. *J. Opt. Soc. Am. B* **19**, 2263 (2002).
 - [5] Pavone, F. S. & Campagnola, P.J. *Second Harmonic Generation Imaging*. (CRC Press, 2013).
 - [6] Saleh, B. E. A. & Teich, M. C. *Fundamentals of Photonics*. (Wiley, 2007).
 - [7] Willner, A. E., Khaleghi, S., Chitgarha, M. R. & Yilmaz, O. F. All-Optical Signal Processing. *J. Lightwave Tech.* **32**, 660 (2014).
 - [8] Zhu, S. -N., Zhu, Y. -Y. & Ming, N. -B., Quasi-phase-matched third-harmonic generation in a quasi-periodic optical superlattice. *Science* **278**, 843 (1997).

- [9] Tsang, T. Y. F., Optical third-harmonic generation at interfaces. *Phys. Rev. A* **52**, 4116 (1995).
- [10] Bonaccorso, F., Sun, Z., Hasan, T. & Ferrari, A. C. Graphene photonics and optoelectronics. *Nat. Photon.* **4**, 611 (2010).
- [11] Butler, S. Z. et al. Progress, challenges, and opportunities in two-dimensional materials beyond graphene. *ACS Nano* **7**, 2898-2926 (2013).
- [12] Koppens, F. H. L. et al. Photodetectors based on graphene, other two-dimensional materials and hybrid systems. *Nat. Nanotechnol.* **9**, 780-793 (2014).
- [13] Ferrari, A. C. et al. Science and technology roadmap for graphene, related two-dimensional crystals, and hybrid systems. *Nanoscale* **7**, 4598 (2015).
- [14] Wang, Q. H., Kalantar-Zadeh, K., Kis, A., Coleman, J. N. & Strano, M. S. Electronics and optoelectronics of two-dimensional transition metal dichalcogenides. *Nat. Nanotechnol.* **7**, 699-712 (2012).
- [15] Xu, X., Yao, W., Xiao, D. & Heinz, T. F. Spin and pseudospins in layered transition metal dichalcogenides. *Nat. Phys.* **10**, 343 (2014).
- [16] Sun, Z., Martinez, A., & Wang, F. Optical modulators with two-dimensional layered materials. *Nat. Photon.* **10**, 227-238 (2016).
- [17] Mak, F. K., Lee, C., Hone, J., Shan, J. & Heinz, T. F. Atomically thin MoS₂: a new direct-gap semiconductor. *Phys. Rev. Lett.* **105**, 136805 (2010).
- [18] Splendiani, A. et al. Emerging photoluminescence in monolayer MoS₂. *Nano Lett.* **10**, 1271 (2010).
- [19] Eda, G., Yamaguchi, H., Voiry, D., Fujita, T., Chen, M., & Chhowalla, M. Photoluminescence from chemically exfoliated MoS₂. *Nano Lett.* **11**, 5111 (2011).
- [20] G. Eda & Maier, S. A. Two-dimensional crystals: managing light for optoelectronics. *ACS Nano* **7**, 5660 (2013).
- [21] Berraquero, C. P. et al. Atomically thin quantum light emitting diodes. *arXiv:1603.08795* (2016).
- [22] Amani, M. et al. Near-unity photoluminescence quantum yield in MoS₂. *Science* **350**, 1065 (2015).
- [23] Li, Y. et al. Probing symmetry properties of few-layer MoS₂ and h-BN by optical second-harmonic generation. *Nano Letters* **13**, 3329 (2013).
- [24] Kumar, N. et al. Second harmonic microscopy of monolayer MoS₂. *Phys. Rev. B* **87**, 161403 (2013).
- [25] Wang, K. et al. Ultrafast saturable absorption of two-dimensional MoS₂ nanosheets. *ACS Nano* **7**, 9260 (2013).
- [26] Malard, L. M., Alencar, T.V., Barboza, A. P. M., Mak, K. F. & de Paula, A. M. Observation of intense second harmonic generation from MoS₂ atomic crystals. *Phys. Rev. B* **87**, 201401 (2013).
- [27] Wang, R. et al. Third-harmonic generation in ultrathin films of MoS₂. *ACS Appl. Mater. Interfaces* **6** 314 (2014).
- [28] Trolle, M. L., Seifert, G. & Pedersen, T. G. Theory of excitonic second-harmonic generation in monolayer MoS₂. *Phys. Rev. B* **89**, 235410 (2014).
- [29] Clark, D. et al. Strong optical nonlinearity of CVD-grown MoS₂ monolayer as probed by wavelength-dependent second-harmonic generation. *Phys. Rev. B* **90**, 121409 (2014).
- [30] Bonaccorso, F. & Sun, Z. Solution processing of graphene, topological insulators and other 2d crystals for ultrafast photonics. *Opt. Mater. Express* **4**, 63 (2014).
- [31] Seyler, K. L. et al. Electrical control of second-harmonic generation in a WSe₂ monolayer transistor. *Nat. Nanotechnol.* **10**, 407-411 (2015).
- [32] Kuc, A., Zibouche, N. & Heine, T. Influence of quantum confinement on the electronic structure of the transition metal sulfide TS₂. *Phys. Rev. B* **83**, 245213 (2011).
- [33] Shi, H., Pan, H., Zhang, Y. -W. & Yakobson, B. I. Quasiparticle band structures and optical properties of strained monolayer MoS₂ and WS₂. *Phys. Rev. B* **87**, 155304 (2013).
- [34] Kadantsev, E. S. & Hawrylak, P. Electronic structure of a single MoS₂ monolayer. *Solid State Commun.* **152**, 909 (2012).
- [35] Zahid, F., Liu, L., Zhu, Y., Wang, J. & Guo, H. A generic tight-binding model for monolayer, bilayer and bulk MoS₂. *AIP Advances* **3**, 052111 (2013).
- [36] Kormányos, A. et al. Monolayer MoS₂: Trigonal warping, the Γ valley, and spin-orbit coupling effects. *Phys. Rev. B* **88**, 045416 (2013).
- [37] Y. Qiu, D., da Jornada, F.H. & Louie, S. G. Optical Spectrum of MoS₂: Many-Body Effects and Diversity of Exciton States *Phys. Rev. Lett.* **111**, 216805 (2013).
- [38] Gibertini, M., Pellegrino, F. M. D., Marzari, N. & Polini, M. Spin-resolved optical conductivity of two-dimensional group-VIB transition-metal dichalcogenides. *Phys. Rev. B* **90**, 245411 (2014).
- [39] Rostami, H. & Polini, M. Theory of third harmonic generation in graphene: a diagrammatic approach. *Phys. Rev. B* **93**, 161411 (2016).
- [40] Rostami, H., Roldán, R., Cappelluti, E., Asgari, R. & Guinea, F. Theory of strain in single-layer transition metal dichalcogenides. *Phys. Rev. B* **92**, 195402 (2015).
- [41] Novoselov, K. S. et al. Two-dimensional atomic crystals. *Proc. Natl. Acad. Sci. USA* **102**, 10451 (2005).
- [42] Bonaccorso, F. et al. Production and processing of graphene and 2d crystals. *Mater. Today* **15**, 564 (2012).
- [43] Sundaram, R. S. et al. Electroluminescence in single layer MoS₂. *Nano Lett.* **13**, 1416 (2013).
- [44] Casiraghi, C. et al. Rayleigh imaging of graphene and graphene layers. *Nano Lett.* **7**, 2711 (2007).
- [45] Lee, J., Novoselov, K. S. & Shin, H. S. Interaction between metal and graphene: dependence on the layer number of graphene. *ACS Nano* **4**, 2695 (2010).
- [46] Zhang, X. et al. Raman spectroscopy of shear and layer breathing modes in multilayer MoS₂. *Phys. Rev. B* **87**, 115413 (2013).
- [47] Säynätjoki, A. et al. Rapid large-area multiphoton microscopy for characterization of graphene. *ACS Nano* **7**, 8441 (2013).
- [48] Kieu, K, Jones, R. & Peyghambarian, N. Generation of few-cycle pulses from an amplified carbon nanotube mode-locked fiber laser system. *IEEE Photon. Technol. Lett.* **22**, 1521 (2010).
- [49] Kieu, K, Jones, R. & Peyghambarian, N. High power femtosecond source near 1 micron based on an all-fiber Er-doped mode-locked laser. *Opt. Express* **18**, 21350 (2010).
- [50] Janisch, C. et al. Extraordinary second harmonic generation in tungsten disulfide monolayers. *Sci. Rep.* **4**, 5530 (2014).
- [51] Hendry, E., Hale, P. J., Moger, J., Savchenko, A. K. & Mikhailov, S. A. Coherent nonlinear optical response of graphene. *Phys. Rev. Lett.* **105**, 097401 (2010).
- [52] Kumar, N. et al. Third harmonic generation in graphene and few-layer graphite films. *Phys. Rev. B* **87**, 121406(R) (2013).
- [53] Hong, S. -Y. et al. Optical third-harmonic generation in graphene. *Phys. Rev. X* **3**, 021014 (2013).

- [54] Nair, R. R.; Blake, P.; Grigorenko, A. N.; Novoselov, K. S.; Booth, T. J.; Stauber, T.; Peres, N. M. R.; Geim, A. K. Fine Structure Constant Defines Visual Transparency of Graphene. *Science* **320**, 1308-1308 (2008).
- [55] Grüning, M. & Attacalite, C. Second harmonic generation in h-BN and MoS₂ monolayers: Role of electron-hole interaction. *Phys. Rev. B* **89**, 081102 (2014).
- [56] Pfister, O. et al. Continuous-wave frequency tripling and quadrupling by simultaneous three-wave mixings in periodically poled crystals: application to a two-step 1.19-10.71- μ m frequency bridge. *Opt. Lett.* **22**, 1211 (1997).
- [57] Butcher, P. N. & Cotter, D. *The elements of nonlinear optics* (Cambridge University Press, 1990).
- [58] Radisavljevic, B., Radenovic, A., Brivio, J., Giacometti, V. & Kis, A. Single-layer MoS₂ transistors. *Nat. Nanotechnol.* **6**, 147-150 (2011).
- [59] Rostami, H., Asgari, R. & Guinea, F. Edge modes in zigzag and armchair ribbons of monolayer MoS₂. *arXiv:1511.07003* (2015).
- [60] Rostami, H., Moghaddam, A. G. & Asgari, R. Effective lattice Hamiltonian for monolayer MoS₂: Tailoring electronic structure with perpendicular electric and magnetic fields. *Phys. Rev. B* **88**, 085440 (2013).
- [61] Alidoust, N. et al. Observation of monolayer valence band spin-orbit effect and induced quantum well states in MoX₂. *Nature Commun.* **5** 4673 (2014).
- [62] Giuliani, G. F. & Vignale, G. *Quantum Theory of the Electron Liquid* (Cambridge University Press, 2005).
- [63] Rostami, H., Katsnelson, M.K. & Polini, M. Gauge invariance in second-order nonlinear optics. (in preparation).
- [64] Griffiths, D. J., *Introduction to Electrodynamics* (Pearson, 2012).
- [65] Mahan, G. D. *Many-Particle Physics: Physics of Solids and Liquids* (Springer US, 1981).
- [66] Jafari, S. A. Nonlinear optical response in gapped graphene. *J. Phys.: Condens. Matter* **24**, 205802 (2012).
- [67] Chirolli, L., Polini, M., Giovannetti, V. & MacDonald, A. H. Drude weight, cyclotron resonance, and the Dicke model of graphene cavity QED. *Phys. Rev. Lett.* **109**, 267404 (2012).
- [68] Margulis, V.I. A., Muryumin, E. E. & Gaiduk, E. A. Optical second-harmonic generation from two-dimensional hexagonal crystals with broken space inversion symmetry. *J. Phys.: Condens. Matter* **25**, 195302 (2013).
- [69] Wu, S., Mao, L., Jones, A. M., Yao, W., Zhang, C. & Xu, X. Quantum-enhanced tunable second-order optical nonlinearity in bilayer graphene. *Nano Lett.* **12**, 2032 (2012).
- [70] Brun, S. J., & Pedersen, T. G. Intense and tunable second-harmonic generation in biased bilayer graphene. *Phys. Rev. B* **91**, 205405 (2015).
- [71] Cappelluti, E., Roldán, R., Silva-Guillén, J. A., Ordejón, P. & Guinea, F. Tight-binding model and direct-gap/indirect-gap transition in single-layer and multilayer MoS₂. *Phys. Rev. B* **88**, 075409 (2013).
- [72] Roldán, R. et al. Momentum dependence of spin-orbit interaction effects in single-layer and multi-layer transition metal dichalcogenides. *2D Mater.* **1**, 034003 (2014).
- [73] *These authors contributed equally to this work