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

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

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

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal’s standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.
**In situ** imaging and control of layer-by-layer femtosecond laser thinning of graphene

D.W. Li, a Y.S. Zhou, a X. Huang, a L. Jiang, b J.-F. Silvain, c and Y.F. Lu* a

Although existing methods (chemical vapor deposition, mechanical exfoliation, etc.) are available to produce graphene, the lack of thickness control limits further graphene applications. In this study, we demonstrated an approach to precisely thinning graphene films to a specific thickness using a femtosecond (fs) laser raster scanning. By using appropriate laser fluence and scanning times, graphene thinning with an atomic layer precision, namely layer-by-layer graphene removal, has been realized. The fs laser used was configured in a four-wave mixing (FWM) system which can be used to distinguish graphene layer thickness and count the number of layers using the linear relationship between the FWM signal intensity and the graphene thickness. Furthermore, the FWM imaging has been successfully applied to achieve **in situ**, real-time monitoring of the fs laser graphene thinning process. This method can not only realize large-scale thinning of graphene with atomic layer precision, but also provide **in situ**, rapid imaging capability of graphene for accurate assessment of the number of layers.

**Introduction**

In the field of nanotechnology, layered nanomaterials represent a diverse and largely untapped source of two-dimensional (2D) systems with unusual electronic properties and high specific surface areas that are critical for sensing devices, catalysis, and energy storage applications. Graphene as a archetypical 2D nanomaterial, which was first discovered experimentally in 2004 by Geim and Novoselov, exhibits unique and fascinating physical properties, such as high carrier mobility and optical transparency, remarkable magnetotransport, and fascinating mechanical property. These excellent properties make graphene suitable for a wide variety of applications, such as transistors, transparent conducting electrodes, optoelectronics, Hall effect sensors. To realize the use of graphene in these applications, the most important challenge is controlled fabrication.

It is known that different layer thicknesses of graphene have different physical properties. For example: 1) thermal conductivity increases with the number of graphene layers; 2) hydrogen coverage investigation demonstrates that the hydrogenation of bilayer- and multilayer- graphene is much more feasible than that of single-layer graphene; 3) the electrical transport property of graphene shows a metallic- to semiconductor transition with respect to the graphene layer thickness; and 4) the work function of graphene increases with the number of graphene layers. Despite recent advances in the synthesis of graphene using various methods, such as mechanical exfoliation of large crystals using adhesive tape, chemical exfoliation by dispersing in a solvent, solution-phase growth, chemical vapor deposition (CVD) growth, the precise control of graphene layer thickness is still a challenge. Therefore, new procedures for fabricating “on-demand” graphene are urgently needed for future applications.

Laser-thinning is a relatively new technique for modifying graphene and other related 2D layered nanomaterials. Previously, Zhou et al. reported laser thinning of graphene oxide (GO) film from a multi-layer (> five layers) to a tri-layer film. The laser thinning arose from the oxidative burning of the GO films in air. Recently, Han et al. utilized a laser with Raman spectroscopy to attenuate multi-layer to monolayer graphene. The accumulation of heat induced by a laser leads to the oxidative burning of upper graphene. It was found that the substrate plays a crucial role as a heat sink for the bottom monolayer of graphene, resulting in no burning or etching. The laser used in Raman spectroscopy has also been adopted to attenuate multi-layer molybdenum disulphide (MoS2) into a monolayer. Similar to graphene thinning, laser-induced heat causes the sublimation of the upper MoS2 layers at laser powers higher than 10 mW, while the bottom layer remains until the laser power is about 17 mW. In general, laser thinning can only produce trilayer or monolayer graphene samples. No other layers can be controlled in this way. This is mainly because the laser sources described above are all continuous-wave (CW) lasers. Due to their heat transfer and dissipation mechanism, they are not suitable for layer-by-layer thinning of graphene. Recently, layer-by-layer thinning of graphene was achieved by sputter coating graphene with zinc and dissolving the latter with dilute acid; however, this is a complex, time consuming process and is likely to cause damage or contamination due to acid corrosion. Unlike CW laser irradiation, femtosecond (fs) laser pulse excitation induces a very different response in graphene. By optimizing the pulse energy, duration, and number of exposures, we believe that fs laser thinning of graphene layer by layer can be realized.

In this study, we demonstrated a new approach to obtaining graphene with a controlled number of layers based on fs laser thinning of few-layer graphene with single atomic layer precision, as shown in Fig. 1. The fs laser used is configured in a four-wave mixing (FWM) system, the use of which enables **in situ**, real-time monitoring of the laser thinning process based on the large optical nonlinearity of graphene. It was determined that there is a linear relationship between the FWM signal intensity and the graphene layer thickness, which can be used as a new, efficient method for identifying the number of graphene layers. We also found that fs laser thinning of few-layer graphene with single atomic layer
precision is a novel, effective, and fast strategy for fabricating graphene with a specific layer thickness. This method not only provides the ability to quickly realize and scale up the fabrication of graphene with accurate thickness control with atomic layer precision, but also can be optimized to perform as a kind of fast fs-laser-based lithography technique.

**Experimental section**

**Material preparation**

In this study, three kinds of graphene samples were used: a commercial CVD graphene product, rapid thermal processing (RTP) graphene, and micromechanically cleaved, few-layer graphene.

(a) **Commercial CVD graphene:** The commercial few-layer CVD graphene on a quartz substrate with six to eight layers was purchased from ACS MATERIAL.

(b) **RTP graphene:** Few-layer graphene was directly grown on a fused silica substrate via a single-step RTP of substrates coated with amorphous carbon (C) and nickel (Ni) thin films. The growth process can be briefly described as follows. The cleaned fused silica substrate was first deposited with an amorphous C film of 8 nm and Ni film of 65 nm using a magnetron sputtering system at room temperature. The Ni/C coated substrate was then loaded into an RTP system. The RTP tube was pumped down and maintained at about 15 mTorr. The temperature was then increased to 1100 °C with an average rate of 500 °C min⁻¹ and kept unchanged at 1100 °C for 120 s. After the RTP tube was cooled to room temperature, the graphene sample was taken out for further characterization.

(c) **Micromechanically cleaved few-layer graphene:** Few-layer graphene samples were fabricated by tapping and shearing highly oriented pyrolytic graphite (HOPG) onto the surface of a transparent fused silica substrate. To make it easier to find the few-layer (less than ten layers) graphene, we located the prepared graphene substrate under an optical microscope and initially estimated the number of layers through their different optical contrast.

**Characterization of graphene samples**

Atomic force microscopy (AFM) and Raman spectroscopy were used to determine the number of graphene layers. An AFM (Agilent 5500, CA, U.S.A) operated in the contact mode was used to study the morphology and to determine the number of graphene layers. A Raman spectrometer (Renishaw InVia plus, Renishaw, Gloucetershire, U.K.) with an excitation wavelength of 514 nm and a lateral resolution of approximately 1 μm was used to evaluate the quality of graphene and to check the number of graphene layers. The Raman spectra and Raman mapping were collected through a 100× objective with an accumulation time of 10 s at each position.

In addition, an FWM system was used to investigate the nonlinear optical properties of graphene and to determine the number of graphene layers. A schematic setup of the FWM system used in this study is shown in Fig. 2a. A Ti:Sapphire fs laser (MaiTai DeepSee HP, SpectraPhysics), in conjunction with a supercontinuum generator (SCG-800, Newport), provides two incident pump laser beams. The
Ti:Sapphire fs laser provides a laser beam with a wavelength of 800 nm, whose power, pulse duration, and repetition rate were 2.95 W, 100 fs, and 80 MHz, respectively. The laser power of the two pump laser beams divided by a Faraday isolator was controlled independently. One pump laser beam was formed using a 500 mW laser beam to generate the supercontinuum using the supercontinuum generator, and then the laser beam was filtered through a long-pass filter (10CGA-830, Newport); the other pump laser beam was formed by introducing an 800 nm laser beam through an attenuator and a delay line. Then the two pump lasers were focused collinearly onto the sample surfaces using a water-immersion objective with a numerical aperture of 1.05 and a working distance of 2 mm. The signal was collected in the forward direction by a sensitive photomultiplier (PMT) tube. Imaging was obtained by raster scanning of the excitation laser beam. A detailed description of our FWM system can be seen in references.28, 29 The inset in Fig. 2a shows a diagram of energy conservation in the FWM process, which involves the generation of mixed optical frequency harmonics $2\omega_p - \omega_s$ under irradiation by two monochromatic waves with two frequencies of $\omega_p$ and $\omega_s$.30-32

**Fs laser thinning of few-layer graphene**

For thinning, few-layer graphene samples were raster scanned and thinned by the same fs laser in the FWM system in the process of imaging, which means that FWM imaging of thinned graphene can be immediately obtained after each cycle of the laser scanning process. The fs laser fluence and scanning rate were adjusted to realize laser thinning of graphene layer by layer. Our FWM system is superior in scan speed, which can realize the maximum areas of 500 μm x 500 μm scanning in 1 second. Therefore, FWM imaging and graphene thinning processes are expected to be completed in few seconds, depending on the scanning area and rate.
Results and discussion

Characterization and imaging of graphene samples

Optical microscopy, AFM, and Raman spectroscopy are currently three most important tools to identify and characterize graphene samples. While it is possible to observe even monolayer graphene using optical microscopy, it is practically difficult due to very low contrasts. Although AFM is powerful to determine the layer thickness, its imaging speed is slow and discrepancies arise from differences in the interactions of the tip with the sample and substrate. With Raman spectroscopy, the Raman spectra of graphene are only sensitive to monolayer, bilayer, and trilayer. To address these issues, we introduced a new approach to distinguishing the number of graphene layers based on the large optical nonlinearity of graphene.

First, we demonstrated and characterized the nonlinear optical property of few-layer graphene using this nonlinear, coherent FWM technique. FWM, as a kind of coherent anti-Stokes form, is enhanced by on-photon resonances and is very sensitive to the nonlinear electronic response. Similar to carbon nanotubes, strong FWM signals are expected when the excitation energies overlap with the electronic excitation energies of graphene. Figs. 2b-d show the FWM imaging for different types of few-layer graphene samples prepared by (b) CVD method (six to eight layers) after transfer onto a quartz substrate, (c) RTP after etching Ni catalysts, and (d) mechanical exfoliation. From these images, although the physical appearances of CVD, RTP, and exfoliated graphene are different, the FWM images are all clearly observed, indicating their similar nonlinear optical properties. The successful detection of nonlinear optical signals from graphene opens up opportunities for optical investigations on an atomic level resolution. In addition, the optical nonlinearity of graphene can be used for high-contrast imaging to distinguish different types of graphene.

For comparison, a few-layer graphene flake with different number of layers prepared by mechanical exfoliation was characterized by optical microscopy, Raman spectroscopy, AFM, and FWM, respectively. Fig. 3a shows a typical optical micrograph of a few-layer graphene transferred onto a fused silica substrate. It is possible, but difficult, for us to observe the monolayer graphene and distinguish different layer thicknesses due to the low contrast. Distinguishing the number of graphene layers was further carried out...
using Raman spectroscopy by mapping the $I_{2D}/I_G$ ratio (Fig. 3b). The $I_{2D}/I_G$ ratio decreases as the number of graphene layers increases (Fig. 2c) (Raman mappings of 2D-band and G-band are shown in Figs. S1a and b). Through Raman characterization, it is easy to identify monolayer graphene from bilayer and multilayer graphene; however, multilayer graphene with different layers of thickness is difficult to distinguish. An AFM was used to measure the morphology and phase of the same sample (Figs. 3d and S1c). The number of graphene layers at different positions can easily be measured, as shown in Figs. 3d and S1d. Fig. 3e shows the FWM imaging for the few-layer graphene as shown in Fig. 3a. Similar to the morphological information shown in Fig. 3d, it is extremely easy to distinguish and quantify graphene with different thicknesses from one to five layers, due to the high contrast in the FWM images, which allow us to count the number of graphene layers in each region.

To more quantitatively investigate the relationship between Raman scattering/FWM signal and graphene layer thickness (or numbers), the dependencies of the Raman $I_{2D}/I_G$ ratio and FWM signal intensity on graphene thickness are plotted in Fig. 4. In Fig. 4a, we can observe the dependence of the Raman $I_{2D}/I_G$ ratio on graphene layer thickness. Monolayer/bilayer graphene is sensitive to the Raman $I_{2D}/I_G$ ratio; however, there is not much difference in the Raman $I_{2D}/I_G$ ratio for few-layer or multi-layer graphene. The ratio decreases slightly as the graphene layer thickness increases. In general, the $I_{2D}/I_G$ ratio of graphene decreases exponentially as the graphene layer thickness increases, which fits well the equation of $I_{2D}/I_G = a \cdot (1 - \exp(-b \cdot H))^c$, where $H$ is graphene layer thickness, $a$, $b$, and $c$ are 0.016, 0.275 and -0.746, respectively. Although Raman spectroscopy has been proved to be a valuable tool for distinguishing monolayer graphene from the graphene of bilayers and few-layers, it is not an efficient, accurate, quantitative method for characterizing, or distinguishing few-layer graphene and accurately judging the layer number of few-layer graphene. Fig. 4b depicts the dependence of the FWM signal on the graphene layer thickness. Similar to Fig. 3e, the FWM signal intensity varied among graphene of different thicknesses. In fact, the FWM signal scales linearly with the number of graphene layers, which is much different from Raman scattering, as shown in Fig. 4a. Fig. 4c compares the graphene layer thickness and Raman $I_{2D}/I_G$ ratio dependences of FWM signal intensity. A linearly increasing relationship between the graphene layer thickness and the FWM signal as well as an exponentially decreasing relationship between the Raman $I_{2D}/I_G$ ratio of graphene and the FWM signal were observed. The linear relationship between the few-layer graphene thickness and the FWM signal can be briefly explained as follows. First, few-layer graphene exhibits a very strong nonlinear optical response in the near-infrared region used in the FWM technique, which enables high-contrast imaging of few-layer graphene compared to dielectric substrate with weak nonlinearity (fused silica). Second, the linear relation is caused by the different constructive interferences of the radiated fields from different layers, as the thickness of few-layer graphene is significantly smaller than the wavelength of the light (800 nm for pump pulses). Based on the results of Fig. 4c, we can conclude that after calibration, an FWM signal can be used to distinguish graphene layer thickness and count the number of layers, which is more accurate than Raman spectroscopy, especially for few-layer graphene. In addition, FWM measurement is much faster (in seconds) compared with Raman and AFM measurements. Therefore, FWM imaging can be used to rapidly realize real time to monitor graphene layer control by fs laser thinning.

**FWM imaging to monitor fs laser thinning of few-layer graphene**

Few-layer graphene flakes, prepared by mechanical exfoliation from HOPG, were used for fs laser thinning in this study. A fs laser configured in the FWM system with a wavelength of 800 nm was used to scan few-layer graphene to a specific number of layers. The FWM system (Fig. 2a) was used to perform both laser-thinning and real-time monitoring through in situ imaging.

First, we investigated the laser thinning threshold of the exfoliated graphene. The thinning threshold of graphene was defined as the laser fluence at which ten times of fs laser scanning cause the FWM signal intensity to weaken. The thinning threshold for the 800 nm fs laser with a pulse duration of 100 fs was obtained, as shown in Fig. S2, around 0.125 J/cm².

We investigated the fs laser thinning of graphene with different layer thicknesses. Fig. 5 shows an AFM image of a typical few-layer graphene with different layer thicknesses before laser thinning. The
average layer thicknesses for Regions 1 to 6 were approximately 2.0, 4.1, 4.4, 8.5, 10.0, and 10.0 nm, respectively (Fig. 5b). The FWM image of the same sample before laser thinning is shown in Fig. 5c. To distinguish different thickness of the graphene, dotted lines were drawn on the FWM images to indicate four different regions with the numbers. The fs laser with a fluence of about 0.475 J/cm² was used to scan and thin the same area with different scanning times, as shown in Figs. 5d-f. Fig. 5d shows the FWM image of the sample for ten times of laser scanning. The FWM signal in the whole graphene area decreased, and the signal from Region 1 almost disappeared, which indicates that the graphene layer thickness has been reduced by approximately 2 nm with ten times of the fs laser scanning. Continuing with 20 and 40 more times of laser scanning, the FWM signal from Regions 2 and 3 disappeared successively, as shown in Figs. 5e and 5f, indicating that the graphene was thinned for about 4 and 8 nm, respectively. The thickness of graphene can be effectively thinned by fs laser scanning with precise control of thickness. The graphene lattice survived without much modification up to a certain laser fluence. Beyond that value, it began to be thinned. By choosing fs laser fluence and scanning times appropriately, graphene thinning with single atomic layer precision, namely layer-by-layer graphene thinning can be realized.

Fig. 6a shows an optical micrograph of a few-layer graphene flake deposited on a fused silica substrate. The Raman mapping of the $I_{2D}/I_G$ ratio is shown in Fig. 6b. There is no obvious contrast in the whole graphene area except for the right margin, indicating a uniform graphene layer thickness. The FWM image is shown in Fig. 6d. The nonlinear signal intensities for the whole graphene area are almost the same due to the uniform layer thickness, which is consistent with the Raman results. Fig. 6e shows the FWM image of the graphene flake after scanning by a fs laser with a fluence around 0.263 J/cm² for ten times. The FWM imaging of the thinned graphene area is uniform with weaker signal intensity as compared with the imaging of the sample before laser thinning. Fig. 6f compares the FWM signal intensity profiles of the graphene before (d) and after (e) the laser thinning, where the FWM signal intensity of the graphene is reduced by approximately 50%. According to the linear relationship between the FWM signal intensity and the graphene layer thickness as described above, we can quantitatively judge how many graphene layers removed or remained if we know the initial graphene layer thickness. From FWM imaging, we can obtain the expression for the number of graphene layers after laser thinning as follows:

$$N_{\text{thinned}} = \frac{I_{\text{initial}} - I_{\text{background}}}{I_{\text{background}}} \times N_{\text{initial}}$$

(1)

where $N_{\text{thinned}}$ is the number of graphene layers that have been removed, $N_{\text{initial}}$ is the number of graphene layers before laser thinning. $I_{\text{initial}}$, $I_{\text{thinned}}$ and $I_{\text{background}}$ are FWM intensities of initial graphene, thinned graphene, and background signal, respectively. To determine the initial graphene layer thickness (number of layers), an AFM image of the same graphene sample before laser thinning was measured (Figs. 7a and c). The initial thickness of the graphene was approximately 1.4 nm. Therefore, using equation (1), the thickness of the thinned graphene was calculated to be approximately 0.7 nm, suggesting that approximately one atomic layer of the graphene was removed during this laser thinning process. To further confirm this conclusion, AFM imaging of the same graphene...
Fig. 6. (a) Optical micrograph of few-layer graphene deposited on a fused silica substrate; (b) Raman mapping of the I\textsubscript{2D}/I\textsubscript{G} ratio of the graphene in the area marked by a dashed rectangle in (a); (c) Raman spectra of the graphene shown in (a) before (black line) and after (red line) fs laser scanning with a fluence around 0.263 J/cm\textsuperscript{2}; (d) FWM image of the graphene in (a); (e) FWM image of the graphene after fs laser scanning with a fluence around 0.263 J/cm\textsuperscript{2}; (f) FWM intensity profiles along the lines drawn in (d) and (e).

Fig. 7. AFM images of the graphene sample shown in Fig. 6(a) before (a) and after (b) fs laser scanning with a fluence around 0.263 J/cm\textsuperscript{2}; (c) and (d) show the thickness profiles of the graphene along the lines in (a) and (b), respectively.

After laser thinning was performed, as shown in Figs. 7b and 7d. It can be seen from Fig. 7d that approximately 0.8 nm thick graphene remained after laser thinning. Namely, approximately 0.6 nm thick of the graphene was removed, which is consistent with the calculation results. Therefore, fs laser thinning of graphene with precision of a single atomic layer has been successfully realized.

To investigate the effect of fs laser irradiation on the graphene’s quality, we use Raman probing. Figure 6c compares the Raman spectra of the graphene before (black line) and after (red line) fs laser scanning with a fluence around 0.263 J/cm\textsuperscript{2}. Raman spectra consist of D-band, G-band, and 2D-band, where D-band is a measure of defect in the sp\textsuperscript{2} graphitic structure. D-band peak is not observable in the Raman spectrum before (black line) laser thinning, indicating very good crystalline quality for this graphene. After laser thinning, a weak D-band peak at around 1336 cm\textsuperscript{-1} has emerged. It indicates an increase in the number of defects due to the laser irradiation induced breaking of sp\textsuperscript{2} carbon-carbon bonds on the sample surfaces. The defect information is considered to mainly come from the nanocrystallites formed on the upper graphene layer during the laser thinning process, but not from the bottom graphene layer on the substrate.

Finally, we carried out an experiment to realize in situ, real-time optical monitoring of a uniform graphene during the laser thinning process using the FWM system. The sample we used was an
exfoliated graphene with a layer thickness of approximately 4.2 nm, corresponding to a six-layer sample (see Fig. S3). To ensure each time of the fs laser scanning corresponded to the reduction of one single layer of graphene, the laser fluence and the scanning speed were adjusted. We found that the laser fluence around 2.5 J/cm² is ideal to realize this thinning requirement. Fig. 8a shows the in situ, real-time monitoring of fs laser thinning of the six-layer graphene using FWM imaging. Numbers in (a) indicate cumulative laser scanning times. It took only a few seconds (~ 2-3 s) to complete one time of FWM imaging and laser scanning, with a scanning rate of approximately 25 μm²/s, which is much faster than Raman spectroscopy. After each time of laser scanning, a homogenous reduction of FWM intensity was observed on the graphene surface, indicating that the whole graphene area was uniformly getting thinned. Fig. 8b shows the changes in FWM intensity of the graphene, along the dotted yellow line shown in (a), as the laser scanning time increased. It was clearly observed that the graphene was completely removed by six to seven fs laser scans. To further prove the layer-by-layer thinning of graphene, the dependence of the FWM intensity from the graphene on the laser scanning time was plotted in Fig. 8c. A linear relationship between the FWM intensity and the scanning time (< 7 times) was observed, which can be expressed as:

$$I_{FWM} = a \times N_{scan} + b,$$

(2)

where $I_{FWM}$ and $N_{scan}$ represent the FWM intensity arising from the graphene and laser scanning times, respectively. Equation (2) includes two constants of $a$ (< 0) and $b$. With more than six scans, the FWM intensity did not change further due to the complete removal of graphene. The linear relationship between the FWM intensity and the graphene layer number can be simply described by:

$$I_{FWM} = a' \times H_{graphene} + b',$$

(3)
where $I_{\text{FWM}}$ and $H_{\text{graphene}}$ represent the FWM intensity and graphene layer number with two constants of $a' (> 0)$ and $b'$. Substituting equation (3) into equation (2), we obtain

$$H_{\text{graphene}} = \frac{a}{b} \times N_{\text{scan}} + \frac{b-b'}{b}$$

(4)

According to Fig. 8c, when $H_{\text{graphene}} = 6$, $N_{\text{scan}} = 0$ and when $H_{\text{graphene}} = 0$, $N_{\text{scan}} \approx 6$. Then equation (4) can be transformed to

$$H_{\text{graphene}} = -N_{\text{scan}} \text{ or } \Delta H_{\text{graphene}} = -\Delta N_{\text{scan}}.$$ 

(5)

This means that layer-by-layer thinning of graphene has been realized by fs laser scanning. In addition, the successful realization of fs laser thinning of graphene with such a fast scanning rate (25 $\mu$m/s) provides the capability to scale up the modification of graphene.

Laser thinning of graphene involves either thermal or non-thermal effects. CW laser thinning of graphene, in general, arises from a heat transfer by absorption of photons and subsequent energy dissipation through phonons (thermal effect). Unlike CW laser irradiation, fs laser produces a different response in graphene during the thinning process. The energy from the fs laser pulse is transferred at rates much faster than the phonon relaxation time. Thus, the hot electrons are created and cool by transferring their energy to phonons on a time much shorter than thermal diffusion. This ultrafast absorption will create unique energy transfer mechanism within the graphene, which depends on the amount of energy absorbed. In this study, the fs laser irradiation energy levels are much lower than those of aforementioned CW laser thinning processes. Therefore, thermal effect is minimized, and non-thermal effects play an important role in the layer-by-layer thinning of graphene via fs laser scanning. This laser thinning process can be used as a new laser-based lithography for modifying graphene with atomic layer precision.

**Conclusions**

In summary, we have developed a rapid layer-by-layer graphene controlled thinning and imaging method via fs laser raster scanning. In situ and real time monitoring of the fs laser thinning process was achieved using the FWM imaging technique. The laser fluence and scanning repetition time play crucial roles in the controlled laser thinning process. Smooth surfaces of graphene after laser thinning can be achieved. In addition, the in situ FWM imaging can be utilized to quantify the number of graphene layers, which is more accurate and much faster than the Raman microscopy. Due to the high speed (in seconds) and simplicity of this method, it is a promising method to achieve large-scale fabrication of graphene with accurate thickness control with atomic layer precision.

**Acknowledgements**

This research work was financially supported by Nebraska Center of Energy Sciences Research (NCSER) and National Science Foundation (ILA 1430519, CMMI 1129613, and CMMI 1265122).

**Notes and references**


