PAPER

Simple graphene film production by chemical vapour deposition and the use of AI for quality analysis

Selene Muñoz-Vargas, [ID] ab Marcos Fernando Perez-Pucheta [ID] ab and Karl S. Coleman [ID] *b

The reproducible production of graphene films at low cost, while minimizing inter-batch variability, is essential for the development of affordable and reliable devices. By using a simple and reproducible methodology along with readily available low-cost equipment, graphene films can easily become accessible to non-graphene specialist research laboratories. This will help accelerate and encourage the development of graphene application research. Here, we describe in detail a simple hot-walled chemical vapour deposition (CVD) procedure to grow graphene with low-cost equipment, using non-processed commercially available copper foil. The quality of the film on copper foil and on Si/SiO$_2$ substrates, after film transfer, was studied by scanning electron microscopy (SEM) and transmission electron microscopy (TEM) to confirm the predominantly monolayer nature of graphene. The quality of graphene was analysed by comparing against standardisation Raman metrics, full width at half maximum of the 2D peak and the ratio of the 2D : G peaks, proposed by the UK National Physical Laboratory showing the films produced were similar or better quality than commercially available monolayer graphene. The random forest machine learning method was used as a predictive tool for the classification of the quality of samples according to the batch and position on the copper substrate relative to the direction of gas flow, revealing that the position and full width half maximum of the G and 2D peaks are the important features in Raman for classification, supporting the idea that unintentional strain and doping are crucial to understand the inherent variations found in the production of monolayer graphene in hot-wall CVD setups.

Introduction

Since the work of Geim and Novoselov[1] in 2004, graphene has attracted the attention of many scientists due to its unique properties. Graphene is a 2D material, in which the carbon atoms are oriented in a hexagonal arrangement and one atom thick. The production of high-quality graphene has been a topic with particular interest. Many methods to produce it have been proposed, which range from chemical, mechanical, and electrochemical exfoliation to chemical vapour deposition (CVD).[2]

CVD can be conducted at atmospheric pressure (APCVD) or low pressure (LPCVD),[3] with LPCVD producing continuous and high-quality graphene film on copper foil substrates.[4] Graphene films are transparent, conductive, and highly flexible, so are considered excellent candidates for transparent conductive electrodes in photovoltaic cells, functional touch-screen panel devices and transparent antennas.[5–7] Other applications include

graphene-based electrochemical biosensors, liquid gate field effect transistors, and physiological sensors.[8,9] To reliably use graphene in these applications, defect-free and high-quality graphene is essential. CVD can be used to grow graphene on different metal substrates, but copper remains the most common.[10]

Wafer scale production of graphene can be achieved using commercial industry-ready cold-wall reactors where only the substrate is heated.[11] Although, the cost can be prohibitive for non-graphene specialist research laboratories and the parameter space complex for growth conditions.[11] In contrast, hot-wall CVD reactors using a tube furnace are less expensive, have precise temperature control, and can be fitted with simple gas control (such as needle valves or mass flow controllers) and vacuum systems. Although straightforward setups have been reported, they are only helpful for demonstration purposes.[12] The pretreatment of copper substrates, through electropolishing and the use of chemicals such as acetic acid, ammonium persulfate, among others,[13–16] is routinely thought to be essential but adds complexity and increases processing time.[17–20]

aDepartment of Chemistry, Durham University, South Road, Durham DH1 3LE, UK
bDepartment of Chemistry, School of Physical Sciences, University of Liverpool, Peach Street, Liverpool L69 7ZE, UK. E-mail: karl.coleman@liverpool.ac.uk

To characterise graphene films, confocal Raman spectroscopy offers rapid, non-destructive analysis of CVD graphene samples across a statistically significant area, thanks to its spatial resolution. However, as individual graphene domains are joined at their boundaries to create a continuous film, and Raman of graphene films is very sensitive to inhomogeneities, taking a single point spectrum to draw conclusions for the whole sample can be misleading.[21,22] It is therefore more meaningful to use a statistical technique, for example bootstrapping, to establish the minimum number of data points necessary for statistically robust conclusions.[22,24] Moreover, interpretation of Raman spectra must be carefully considered as it is sensitive to strain,[24–27] doping[28,29] and chemical functionalisation.[30,31] The main features extracted from the graphene spectrum are typically the 2D, G, and D band intensities (I), positions (ω), full width at half maximum (Γ) and ratios (I_{2D}/I_G and I_D/I_G).

The UK National Physical Laboratory (NPL), an international leader in graphene standardisation, has proposed $\Gamma_{2D} \leq$ 35 cm^{-1} and $I_{2D}/I_G \geq 2$ as parameters for identifying monolayer graphene,[23,32] however these two metrics are sensitive to external factors.[23] I_{2D}/I_G is decreased by doping level[28] and defects,[33] while Γ_{2D} increases with the presence of adlayers and sub-spot size strain variations from the substrate.[34,35] Therefore, nanometre strain, unintentional doping from the Si/SiO$_2$ substrate,[36,37] residual PMMA[38] and FeCl$_3$,[39] used for graphene film transfer, and annealing could all be a source of sample variation. By considering additional metrics such as ω_G and ω_{2D} we can get additional information about the strain and doping state in the sample.[40] As a result, monolayer graphene with varying levels of strain and doping can be characterised.

As a case study, we synthesized large area monolayer graphene in a hot wall LPCVD setup using as provided untreated copper foil to simplify the graphene growth process. The predominantly monolayer nature of graphene was confirmed by transmission electron microscopy (TEM), scanning electron microscopy (SEM), atomic force microscopy (AFM) and Raman spectroscopy, confirming that the synthesised graphene meets these two criteria ($\Gamma_{2D} \leq 35$ cm^{-1} and $I_{2D}/I_G \geq 2$), set out by NPL, just as commercially available monolayer graphene samples.

Given that the reactive species are in a non-equilibrium steady state along the hot tube, the gas composition will vary,[41] using machine learning we anticipate these subtle variations can be detected using Raman mapping of graphene collected along the gas flow direction. Machine learning has proven to be a powerful tool in materials science, with recent work demonstrating its utility in optimizing graphene fabrication processes. For instance, Beckham et al.[42] used machine learning to guide the synthesis of flash graphene by exploring the effects of process parameters on crystallinity, and Jain et al.[43] employed Bayesian optimization in laser induced graphene patterning. However, a significant gap remains in applying these techniques to quality control and post-synthesis analysis for scalable and continuous production methods. Supervised methods such as random forests (RF) outperform unsupervised methods, demonstrating higher accuracy, robustness, and computational efficiency, while also providing implicit feature importance estimations.[44,45] For example, RF has demonstrated to be particularly useful in the 2D materials field for predicting magnetic ordering,[46] discriminating between monolayer and multilayer regions in MoS$_2$ by optical microscopy[44] and estimation of the twist angle in bilayer graphene.[47] In this work, we have utilized random forests for the classification of graphene samples from different batches and positions on the copper surface substrate. Our approach specifically highlights the potential of using machine learning techniques and Raman spectroscopy for quality control in monolayer graphene production, enabling the differentiation between batches and substrate position relative to the gas flow by finding generalizable predictive patterns based on Raman mapping. This is a new direction that moves beyond process optimization to product verification, addressing a critical need for industrial applications.

Results and discussion

Fig. 1a shows the schematic setup for graphene growth, where a simple flow meter is used to control the argon gas flow and two mass flow controllers to feed hydrogen and methane gases. To grow graphene, a piece of as supplied untreated copper (6.5 × 4.0 cm^2) foil was rolled against the quartz tube wall (Fig. 1b and c) and placed in the centre of the quartz tube and within the uniform temperature zone of the furnace. The temperature profile from the centre of the tube was determined in an argon flow to ensure any temperature gradients were minimized (see Fig. S1). Once the gas lines (argon, hydrogen and methane) were purged and the system pressure lowered to 43 m torr, the temperature was increased under 10 sccm hydrogen gas flow in two steps: from room temperature to 900 °C over 20 minutes and from 900 °C to 1000 °C at a rate of 10 °C min^{-1}, followed by an annealing step at 1000 °C for 30 min. Then, 7 sccm methane was introduced, and the hydrogen flow was increased to 15 sccm. These conditions were maintained for 10 min to allow graphene to grow. Upon completion, the chamber was rapidly cooled down by opening the furnace whilst maintaining the methane and hydrogen gas flows until the temperature reached 140 °C, at which point they were stopped, while maintaining the continuous flow of argon. The cooling rate is shown in Fig. 1d, which shows cooling rate of c.38 °C min^{-1} between 1000–300 °C, and 13 °C min^{-1} between 300–140 °C. Fig. 1d illustrates this procedure, and full details are given in the Methods Section. To conduct the Raman analysis, five distinct areas were selected from various substrate positions in each batch of graphene. These areas were then transferred onto silicon/silicon oxide (Si/SiO$_2$) substrates, as depicted in Fig. 1e.

Fig. 2a shows the optical image of as grown graphene on copper exposed to air and the absence of any copper oxide areas. Fig. 2b shows an SEM image of the 'as-grown' graphene on copper foil where the copper grain boundary and wrinkles can be readily observed due to the difference in the thermal expansion coefficients between graphene and copper.[48] Graphene adlayers can also be observed. The average copper grain size is 167.5 ± 15.9 μm^2, while the average graphene grain size is 3.4 ± 0.3 μm^2 (see Fig. S2 and S3).
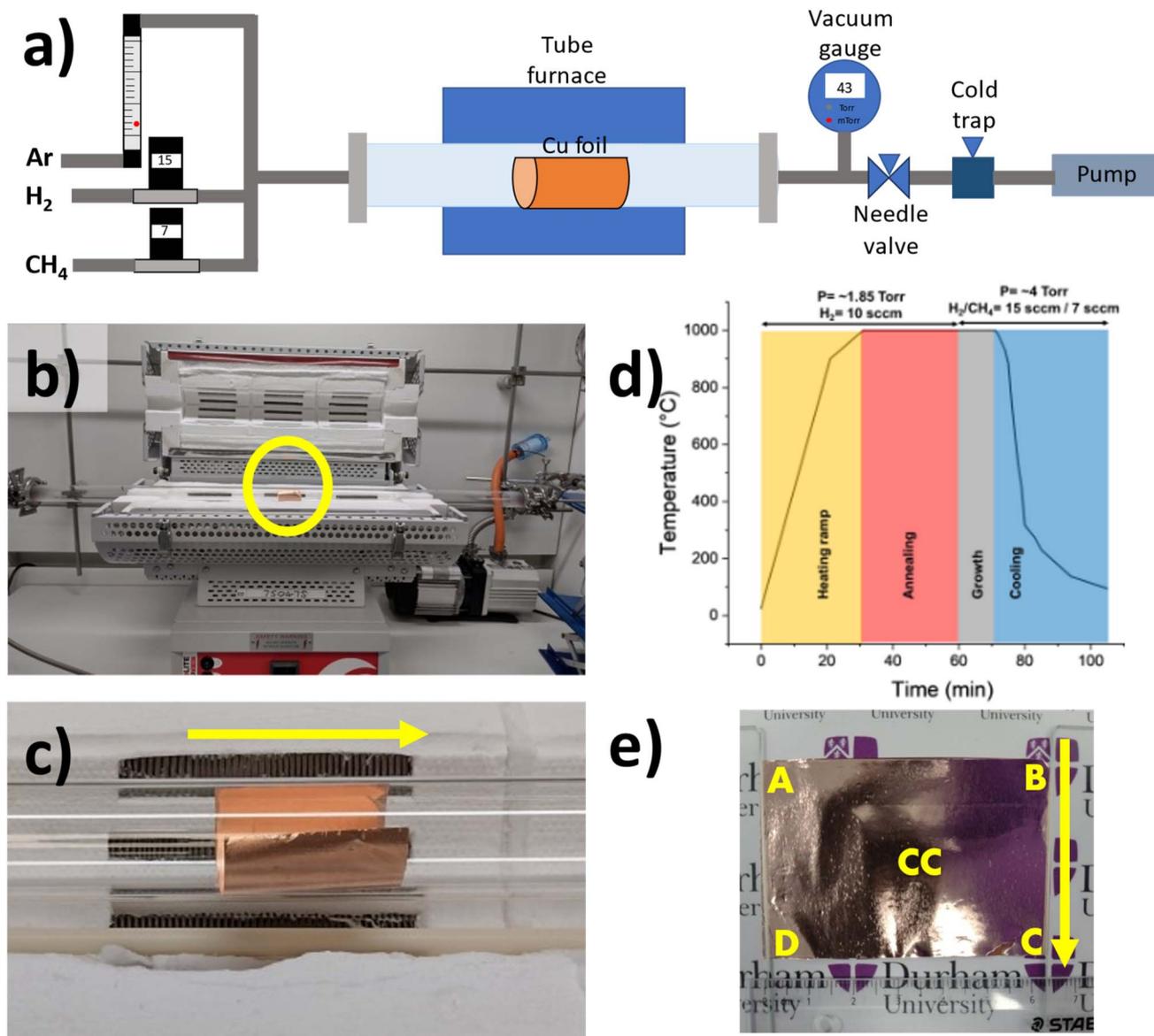
Fig. 1 (a) Schematic diagram for graphene synthesis by LPCVD. (b) Photograph of setup, yellow circle shows copper foil placed at the centre of the furnace. (c) Image of the graphene foil wrapped against tube wall. The yellow arrow indicates the flow direction. (d) Time dependence of experimental parameters. (e) Large area graphene synthesised (6.5×4.0 cm^2) on copper foil showing where the samples were taken. The yellow arrow indicates the flow direction of the reactive gas in the growth process.

The graphene was transferred onto a Si/SiO$_2$, using the established PMMA transfer method,[4] for further characterisation. In brief, graphene on copper was coated with a thin layer of polymethyl methacrylate (PMMA) as protective layer and the copper was dissolved in a 1 M FeCl$_3$ solution. The film was washed in deionised water and transferred onto the Si/SiO$_2$. Finally, the PMMA layer was dissolved in acetone, leaving the graphene on Si/SiO$_2$. A full description of the transfer process can be found in the Methods Section. Fig. 2c shows graphene transferred on Si/SiO$_2$ where a continuous film, with some adlayers and polymer residues, can be observed. Fig. 3a–c shows a set of TEM image at different magnifications of the graphene free-standing films on a holey carbon matrix with some polymer

residues. Fig. 3d shows the corresponding selected area electron diffraction (SAED) pattern. The inset figure illustrates the intensity profile through the dashed line, showing higher intensities of the inner reflections ($0\bar{1}10$ and $\bar{1}010$) than those of outer reflections ($1\bar{2}10$ and $\bar{2}110$), which is consistent with, and confirming, monolayer graphene.[49]

Monolayer graphene shows a characteristic Raman spectrum in which three peaks can be identified and used for quality assessment. The D peak at $\sim$1350 cm^{-1} is caused by the breathing mode of the six-atoms ring and is activated by defects such as grain boundaries, imperfect stitching and sp^3 hybridization.[50,51] The G peak at $\sim$1580 cm^{-1} corresponds to the in-plane bond-stretching motion of pairs of sp^2 carbon and does
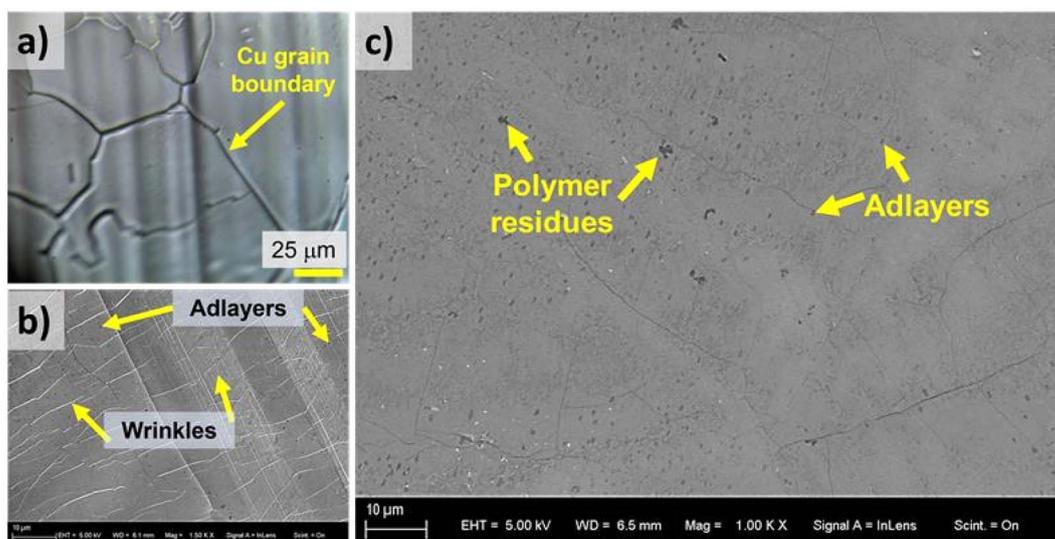
Fig. 2 (a) Optical image of as grown graphene on copper foil. (b) SEM image of graphene as grown on copper showing the wrinkles resulting from the cooling step and the difference in thermal expansion coefficients from copper and graphene. (c) SEM image of graphene transferred on Si/SiO$_2$ illustrating some polymer residues, adlayers and wrinkles.
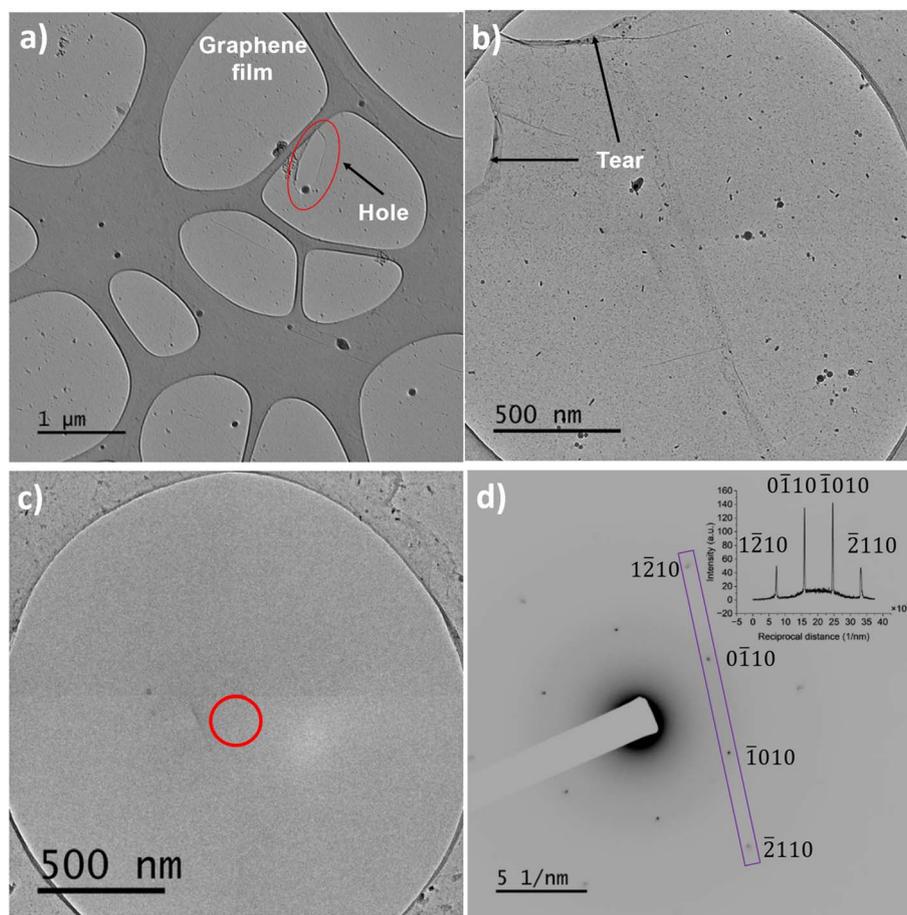

Fig. 3 (a)–(c) Low magnification TEM images of monolayer graphene on holey carbon TEM grid showing tears and a hole in the continuous film. (d) SAED pattern from red circled area in c, inset shows the profile along the rectangle, demonstrating monolayer graphene.
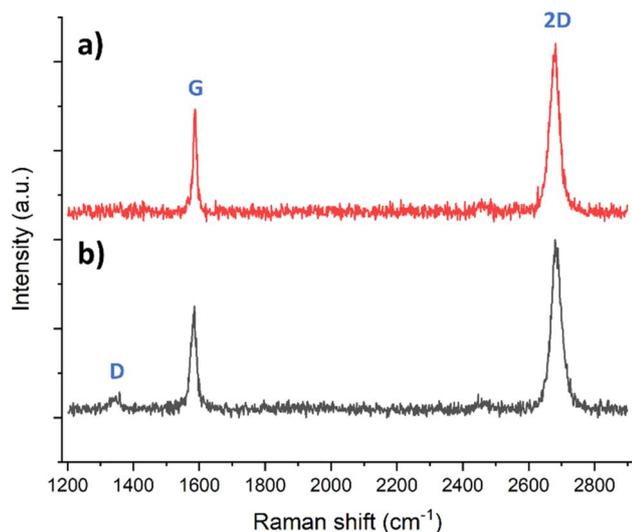
Fig. 4 Raman point spectra from this work of (a) defect-free graphene and (b) minimally defective graphene using a 532 nm laser excitation wavelength showing the main peaks of graphene: D, G and 2D.

not require the presence of a six-folded ring.[52] The peak at ~ 2700 cm^{-1} is caused by a second-order overtone of the D band. Since the 2D band for graphene can be fitted to a single Lorentzian peak, it can be used as a distinctive fingerprint to distinguish between single and multilayer graphene.[53] Fig. 4a shows a typical monolayer graphene Raman point spectrum exhibiting the characteristic bands for graphene: G (~ 1583 cm^{-1}) and 2D (~ 2678 cm^{-1}) bands, with no apparent D peak. Fig. 4b shows the case for minimally defective graphene, exhibiting a small D band (~ 1340 cm^{-1}).

As suggested from the NPL guidelines, $\Gamma_{2D} \leq 35$ cm^{-1}, where Γ is fullwidth at half maximum of the peak, is typical for monolayer graphene on Si/SiO$_2$, based on experimental observations.[4,29,54,55] However, it is worth noting that a turbostratic stacking of graphene layers can also be represented by a single Lorentzian peak, with $\Gamma_{2D} \leq 35$ cm^{-1}, similar to that of single layer graphene.[56] Similarly, the $I_{2D}/I_G > 2.0$ ratio, where I is intensity, is characteristic of single-layer graphene and an indicator of low level of doping.[57] In this study, TEM results, however, suggest a predominantly monolayer graphene structure; therefore, these two metrics are suitable for identifying monolayer graphene via Raman spectroscopy.

Five specific graphene growth areas on the copper substrate, labelled A, B, C, D, and CC (centre), as illustrated in Fig. 1e, were examined. This enables variations caused by changes along the gas flow direction to be determined. Given that the average grain size is approximately 3.4 μm^2, the objective was to determine the minimum sampling area required to obtain statistically meaningful results from the entire sample. Consequently, we acquired data from a large map consisting of 1024 spectra (32 μm $\times$ 32 μm in 1 μm steps), so on average around 301 graphene grains were sampled. Here, the methodology described by Turner et al.[23] was used, where the Γ_{2D} and the I_{2D}/I_G ratio are the parameters of interest. For calculations, D, G, and 2D peaks were fitted in OriginPro 2024[58] using a Lorentzian function with the fit bounds suggested by Turner et al.[23] Cosmic rays on the CCD detector were manually removed and spectra were normalized to 1 with respect to the intensity of the 2D peak.

The large map was divided into smaller sub-maps ranging from 1 $\times$ 1 to 32 $\times$ 32 μm^2 with 1 μm increments. Each sub-map was randomly positioned, and a square of the specified size was traced from that point. One hundred sub-maps of each size were created throughout the large area map, and the average values of Γ_{2D} and I_{2D}/I_G ratio were calculated. The difference between
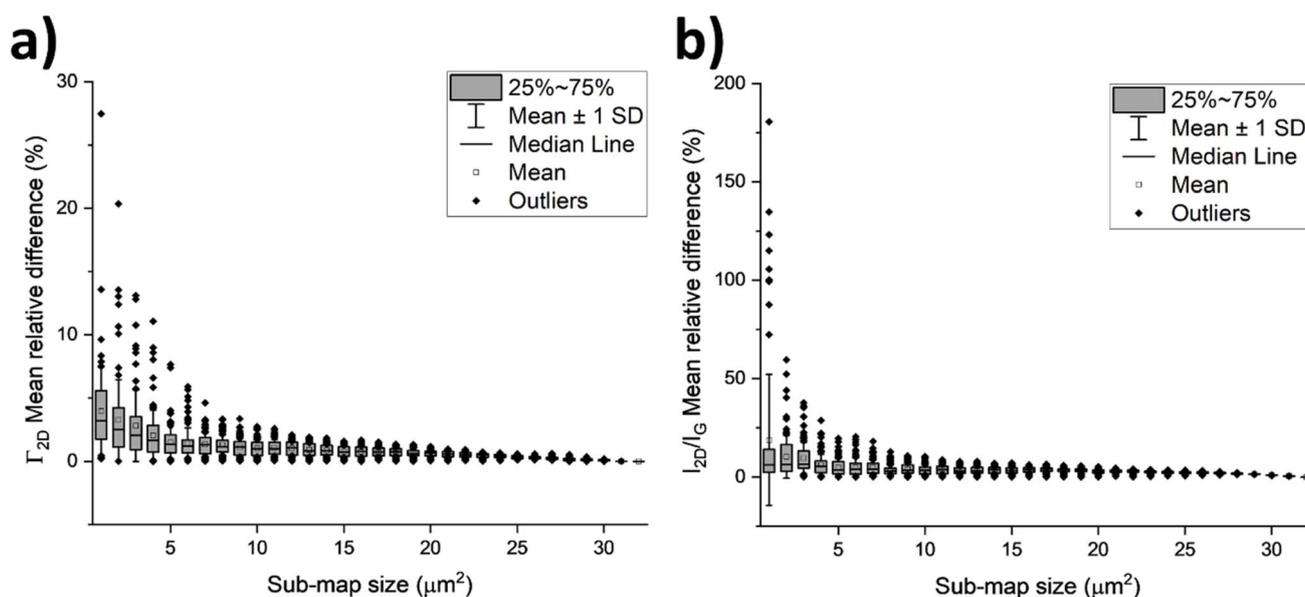

Fig. 5 The mean relative difference between the mean from each sub-map and the mean from the large area map (32 $\times$ 32 μm^2) for (a) Γ_{2D} and (b) I_{2D}/I_G ratio, as a function of sub-map size.

the average values of each sub-map and the large area map was then calculated for each sub-map size (Fig. 5a and b).

The mean of the 1024 spectra analysed was 26.2 ± 1.7 cm^{-1} for Γ_{2D} and 2.4 ± 0.7 for the I_{2D}/I_G ratio. It was observed that for maps 13×13 μm^2, the relative difference with respect to the large map (32 μm × 32 μm in 1 μm steps) is 0.9%, which is an acceptable error given the fewer spectra collected. Statistical analysis was conducted on a 13×13 μm^2 map recorded for each position A, B, C, D, and CC (Fig. 1e) in 5 independent and separate batches. The G and 2D peaks are observed for all recorded spectra, and only a minimal number of spectra showed the presence of a small D peak, suggesting a graphene film with very few defects.

By measuring the mean and standard deviation from the background, $\mu_{background}$ and $\sigma_{background}$ respectively, in all normalized spectra in the region 1800–2200 cm^{-1} (where no graphene peaks are found), it is observed that in all batches the D band has only a minimal intensity, and of the same order of magnitude as the background signal $(\mu + \sigma)_{background}$ (see Table S1). Moreover, D intensities from the average spectra are comparable with those of commercial brands, suggesting that graphene films have a low defect density. For comparison, Fig. S4–S9 show all Raman spectra collected for each sample from each batch and two separate commercial brands.

Fig. 6 shows the Raman maps for the spatial distribution of Γ_{2D} for each position and batch. According to the Γ_{2D}, most of the scanned areas can be considered monolayer graphene. The

Fig. 6 Raman mapping of Γ_{2D} for each 5 positions (A, B, C, D and centre CC, see Fig. 1e) from each batch. The last two maps correspond to two commercially available samples of monolayer graphene. For every map an area of 13×13 μm^2 with 1 μm step was sampled and all maps were graphed using the same scale colour bar.
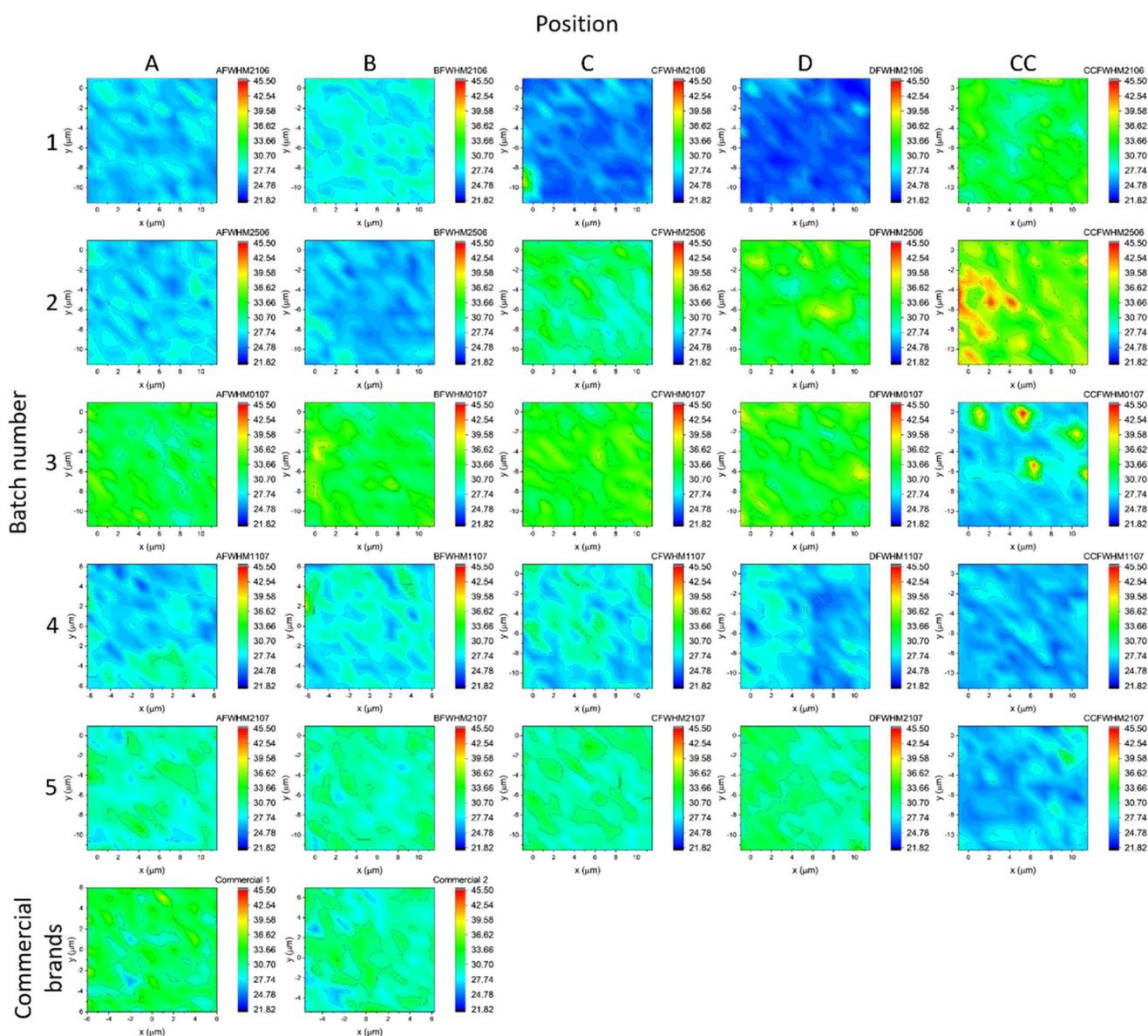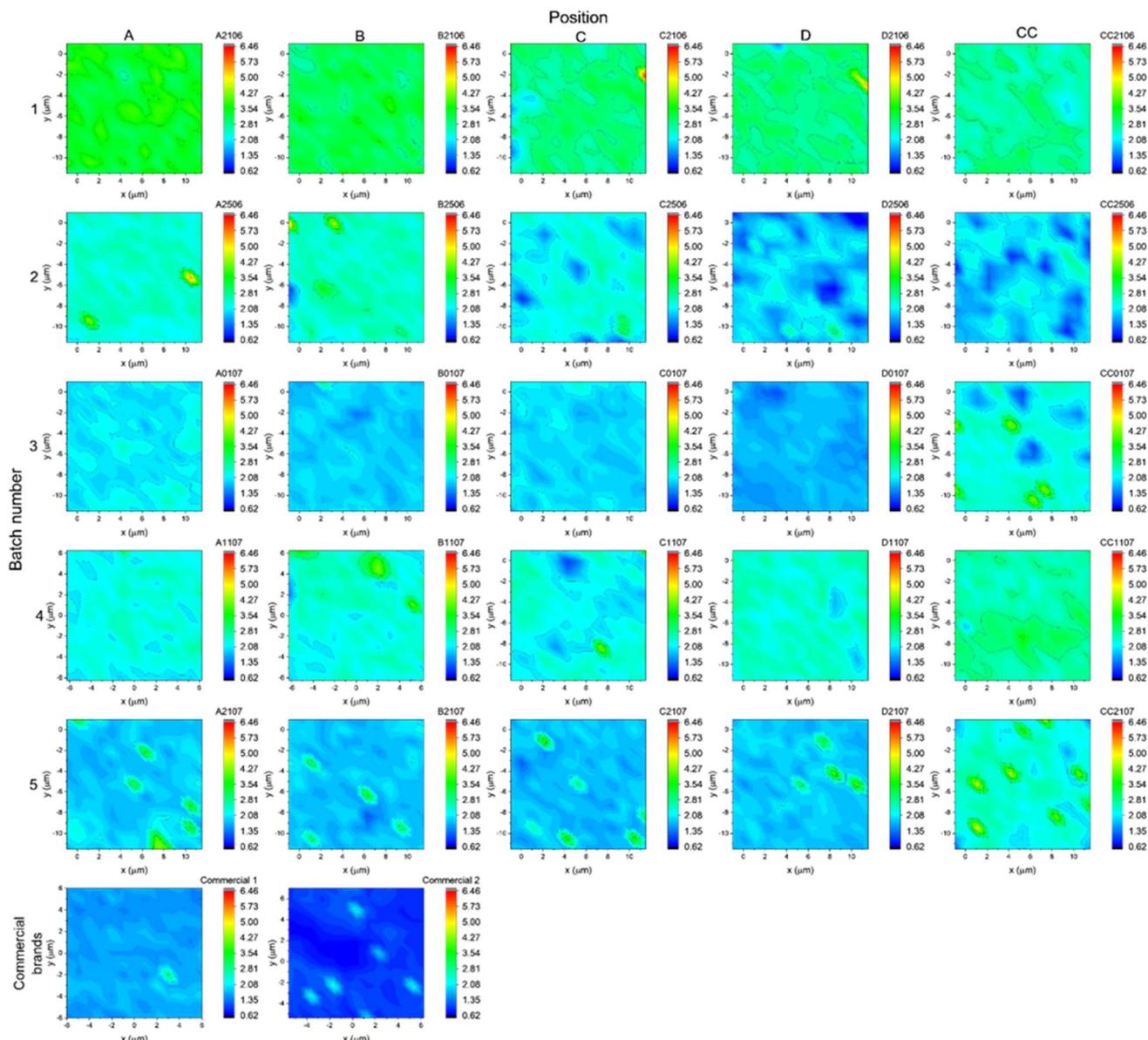
Fig. 7 Raman mapping of I_{2D}/I_G for each 5 positions (A, B, C, D and centre CC, see Fig. 1e) from each batch. The last two maps correspond to two commercially available samples of monolayer graphene. For every map an area of $13 \times 13~\mu m^2$ with 1 μm step was sampled and all maps were graphed using the same scale colour bar.

largest Γ_{2D} obtained is 45.5 cm^{-1}, present in only a small number of areas in two samples (CC position batches 2 and 3), while the lowest value is 21.8 cm^{-1}. The varied values observed may be attributed to the p-doping effects caused by ferric chloride used during the copper etching process and residual PMMA present after film transfer.[38,39] Interestingly, all samples from all batches show higher I_{2D}/I_G ratios than the commercial product (see Fig. 7); the 2D peak intensity is almost twice the intensity of the G peak, with the highest ratio of 6.4 and the lowest 0.62.

The data presented in Fig. 6 and 7 have been summarised in box plots in Fig. 8a–d. By establishing a threshold of 35 cm^{-1} for the Γ_{2D} peak, it can be observed that, with the exception of sample CC250622 from batch 2, the average Γ_{2D} value of all the

samples falls below this threshold. The mean value of 36 cm^{-1} for the sample CC250622 slightly exceeds the predetermined threshold. Moreover, comparative analysis of the samples against two separate commercially available monolayer graphene samples, under the same mapping conditions, revealed that although Γ_{2D} was less than 35 cm^{-1}, the I_{2D}/I_G ratio was less than 2, which indicates that graphene films in this work are of higher quality and support the conclusion that the synthesized graphene was primarily monolayer.

Interestingly, although the analysis of the I_{2D}/I_G ratio and Γ_{2D} reveals the high quality of the synthesized graphene, Fig. 8 shows possible parameter fluctuations, indicating the likelihood of variations within a batch (intra-batch) and among different batches (inter-batch). To analyse this further, machine
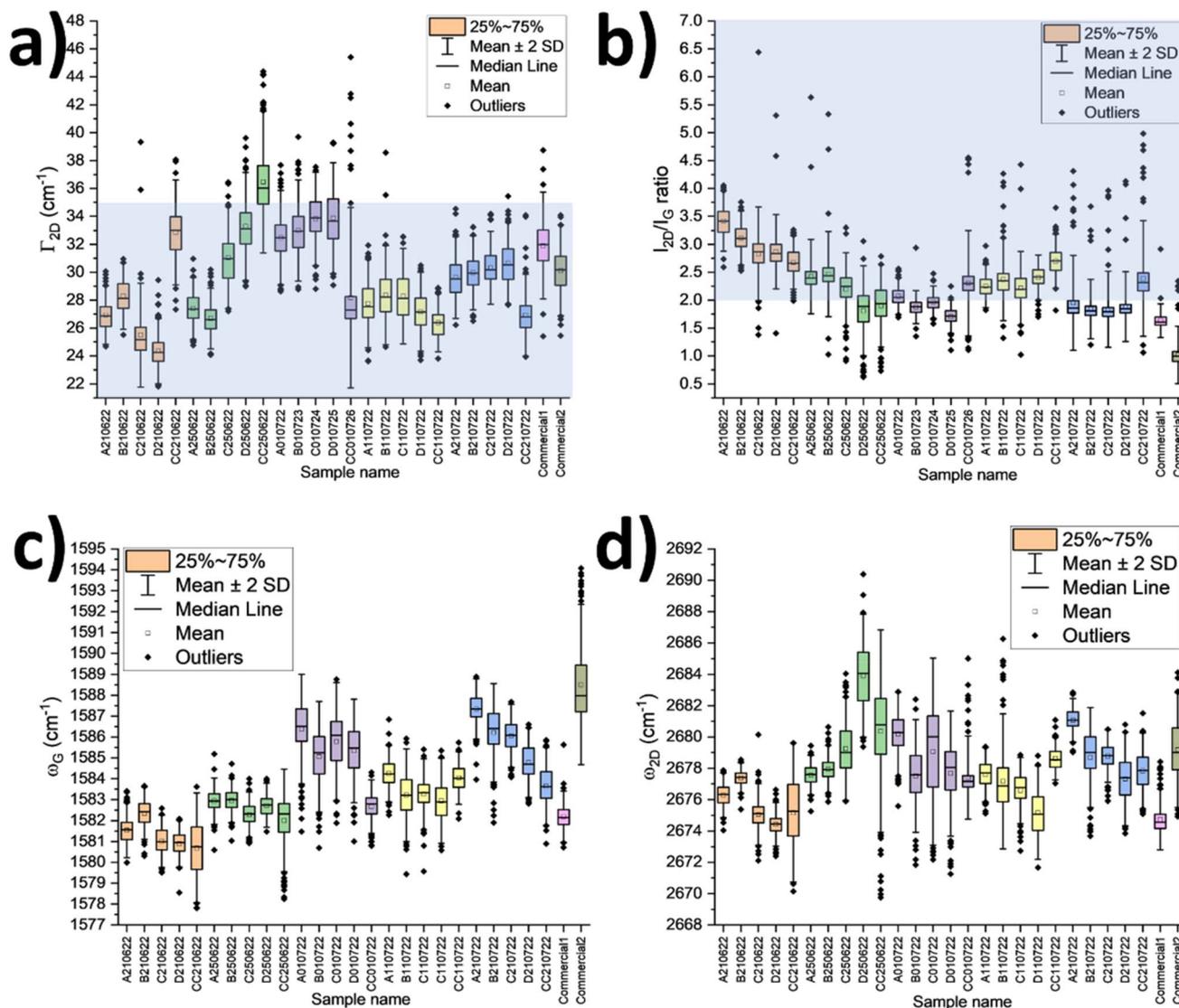
Fig. 8 Box plots showing the data variations for each position (A, B, C, C and centre CC) from each batch and for the two commercially available brands, (a) Γ_{2D}, (b) I_{2D}/I_G peak ratio, (c) ω_G and (d) ω_{2D}. Blue shaded regions in (a) and (b) illustrates the thresholds for which $\Gamma_{2D} < 35$ cm^{-1} and $I_{2D}/I_G > 2$ associated with monolayer graphene according to the National Physical Laboratory (UK).[23]

learning (ML) tools were used to help identify the most statistically informative Raman features for data classification. To predict sample batch assignment and rank the importance of predictor variables, a random forest algorithm was employed.[44] The algorithm was coded in Python by using the open-source Pandas and Scikit-learn modules.[59] To train the model 7 characteristics from the Raman fittings were used: ω_G, ω_{2D}, Γ_G, Γ_{2D}, I_G, I_{2D}, and I_{2D}/I_G ratio (where ω is the peak position, Γ fullwidth at half height and I intensity). Overall 4225 data points were used, with 70% allocated for model training and 30% for evaluation, using the batch number or sample position as the output variable.

The model was able to correctly classify the batch number with an accuracy of 83% and F1-score of 83%, which indicates inter-batch variation. This high level of performance is particularly significant as it demonstrates that the subtle fluctuations in Raman spectra are not random but contain distinct, quantifiable information that can be used as a "fingerprint" for each batch. The confusion matrix in Fig. 9a further validates this finding, showing a minimum accuracy of 76% and confirming the model's robustness across all batches. By utilizing the model to identify the key parameters responsible for these variations, we can further analyse them, which allows the parameters to be linked to their respective underlying physical influences.

From the feature importance, Fig. 9b, is clear that ω_G, ω_{2D}, Γ_G, Γ_{2D} are the main features to consider for data classification, suggesting that strain and unintentional doping are essential to understand the sources of variation between batches. This could be expected as it has been widely demonstrated that shifts in G and 2D peak positions are correlated to strain and doping effects.[25,40] Pristine graphene sheets can display a random
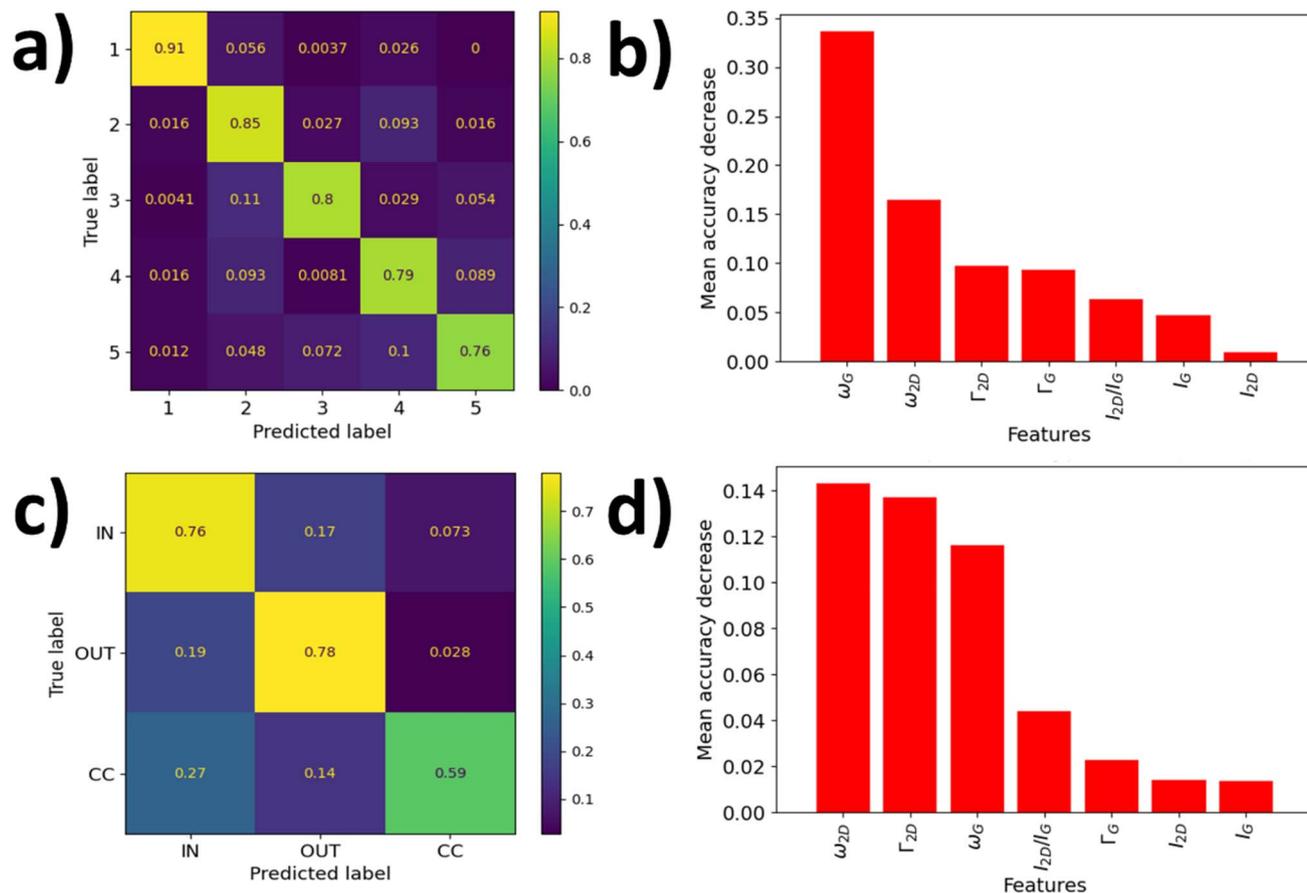
Fig. 9 Permutation feature importance and normalised confusion matrix for batch (a and b) and in–out-centre (c and d) prediction. The higher the diagonal values in the confusion matrix the better, with the incorrect predictions out of the diagonal. The sum of each row equals one. A perfect score would have a 1 across the diagonal. The remaining cells contain inaccurate predictions.

distribution of compressive and tensile strains ranging between −0.2% and 0.4%.[40] The difference in thermal expansion coefficients between graphene and copper is likely to subject the whole sample to this level of strain during the rapid cooling step as part of graphene film production, which in turn could be used by the model as a fingerprint to distinguish between

batches.[48,60] Experimental and theoretical calculations of the Gruneisen parameter report shifts of the G and 2D bands with respect to the strain applied (% ε) as $\Delta\omega_G/\Delta\varepsilon = -57.8$ cm^{-1}/% and $\Delta\omega_{2D}/\Delta\varepsilon = -140$ cm^{-1}/%, so under biaxial strain $\frac{\Delta\omega_{2D}}{\Delta\omega_G} \sim 2.2$.[25,26] Similar findings have been observed for

Fig. 10 CVD set up; 1 and 2 hydrogen and methane mass flow controllers, respectively, 3 rotameter for argon, and 4 the vent valve.

the Γ, showing a $\Delta\Gamma_{2D}/\Delta\Gamma_G$ ratio of approximately $\sim$2.2.[34] Hence, even small strains of approximately $\pm$0.1% can lead to a broad range of ω_G, ω_{2D} and Γ_{2D} values. These variations can be seen in graphene films produced both in this work and in the commercial samples. Unintentional p-doping from the Si/SiO$_2$ substrate, residual ferric chloride, or polymer residue from the transfer process could also play a part.[36,37,39,54] It has been previously shown that the average shift produced by p-doping is $\frac{\Delta\omega_{2D}}{\Delta\omega_G} \sim 0.7$.[40] The I_{2D}/I_G ratio should similarly be influenced by doping, but interestingly it was not identified as a key feature by the random forest algorithm, indicating that the strain is likely the primary factor contributing to the shift in ω_G and ω_{2D}.

As for the data classification by position respect to the flow inlet, the model was able to classify the position of the sample with a training set accuracy of 73%, and F1 score of 73% and a minimal accuracy of 59% to predict centre position in the batch (see Fig. 9c). The significantly reduced accuracy of the model suggests lower intra-batch variation. Again, the most useful Raman features to distinguish between front (IN), centre (CC) and back (OUT) positions respect to flow direction were ω_{2D}, Γ_{2D}, and ω_G, Fig. 9d. Reduced accuracy suggests the model's difficulty in determining sample position relative to flow direction, indicating an acceptable batch homogeneity.

Conclusion

A simple and readily accessible low-cost LPCVD setup that uses off-the-shelf copper foil, with no pretreatment, has been shown to produce consistently high-quality large area graphene films (6.5 $\times$ 4 cm^2) that are comparable, or better, than commercially available material. SEM observations show a continuous graphene film with very few adlayers and wrinkles. Moreover, AFM measurements and TEM observations confirmed the predominantly monolayer nature of the films synthesized. Raman mapping was used to obtain statistically meaningful estimates of the mean of two key parameters Γ_{2D} and I_{2D}/I_G to characterise the quality of monolayer graphene. It was found that in most of the samples, $\Gamma_{2D} < 35$ cm^{-1}, and the I_{2D}/I_G ratio was larger than that of commercial material, suggesting a predominantly monolayer nature of the graphene films with lower doping or defect levels. Remarkably, although minimal variations between batches was observed, machine learning using the Random Forest algorithm could be used to determine the main features for batch classification and position of the copper growth substrate with respect to the gas flow. Interestingly, these features were the peak positions and full width at half maximum of the G and 2D peaks and not the I_{2D}/I_G ratio, as one might expect given the vast amount of literature, as they are related primarily to residual strain which is present due to differing thermal expansion coefficients of graphene and the copper substrate. Although here we highlight the inherent variability anticipated in hot wall systems, we anticipate that this methodology could be readily applicable to the spatial characterization of graphene films from cold wall systems. Ultimately, this would enable a full comparative analysis of inter-batch differences and spatial homogeneity of CVD graphene films, which would be highly relevant for future quality control in either a laboratory or commercial setting.

Methods

Graphene growth process

Graphene was grown using a homemade set up which consists of a mixer gas chamber with two mass flow controllers for CH$_4$ and H$_2$ (Brooks 5850 TR) and one rotameter for Ar (Brooks 1355/D2C2D1C00000), a commercial furnace (Carbolite, EST12450B-230SN), pressure gauge (KJL275806), needle valve, cold trap, and a pump (TRIVAC D 2,5 E). A 32 mm diameter and 1 m-long quartz tube was used as the reaction chamber (with a 100 mm uniform heating zone), (Fig. 10).

Copper foil (Alfa Aesar 46986.RF, 25 μm thickness, 99% purity) was cut to desired size (6.5 cm × 4.0 cm), rinsed with acetone followed by isopropyl alcohol rinse and dried with nitrogen gas. The substrate was then wrapped up in the quartz tube at the middle of the furnace, the lines were purged with 27 lpm: 15 sccm: 15 sccm Ar:CH$_4$:H$_2$ for 10 min and another 5 min just under Ar flow. At this point the vent valve was closed and the argon flow was stopped and the pump was turn on, the system was then pumped until 43 mTorr was reached. Then 10 sccm of hydrogen was introduced before beginning the heating cycle, the first heating ramp was 43.9 °C min^{-1} until 900 °C was reached and reduced to 10 °C min^{-1} until 1000 °C was reached. The substrate was annealed for 30 min, after which the hydrogen flow was increased to 15 sccm and 7 sccm methane was introduced into the chamber to start the graphene growth process. After 10 min, enough for full graphene coverage on the copper substrate, heating was stopped and the furnace was opened when the temperature reached 900 °C, when the temperature reached 140 °C, argon was introduced at 27 lpm and the hydrogen and methane feeds were stopped. Argon flow was maintained for 5 min after which the pump was turned off and the vent valve was opened. Finally, the quartz tube was opened to remove the graphene on copper foil. The fresh graphene was placed in a Petri dish and sealed with parafilm for future use.

Graphene transfer process

Fresh graphene was transferred on to glass and silicon/silicon oxide (Si/SiO$_2$) using a wet transfer method.[61] A 1 cm^2 graphene/copper foil (Gr/Cu) was cut and attached to thermal release tape, spin-coated with 50 μl of PMMA (8% w/w in anisole, M$_w$ $\sim$120 000) at 2000 rpm for 1 min, cured at 115 °C for 30 s and the thermal release tape was released. The Gr/Cu was placed for 2 min in a oxidizing solution (2 : 2 : 21 v/v HCl (37%) : H$_2$O$_2$(30% w v^{-1}) : H$_2$O), placed in deionized water (DI) bath for 13 min, placed in 1 M FeCl$_3$ for 15 min at 55 °C, followed by washing in 3 DI water baths, 10 min for each one. The graphene film was then carefully removed by transferring on to a glass or Si/SiO$_2$ substrate (Gr/S). The Gr/S was placed on hot plate at 40 °C for 30 min, then placed in a drying oven at 130 °C for 30 min with $\sim$45° inclination. Finally, it was placed in an acetone bath

at 55 °C for 2 h, rinsed in acetone, rinsed with isopropanol, and dried under nitrogen (N_2).

Characterization methods

SEM images to analyse the surface morphology of graphene on copper, and after the transfer process on Si/SiO$_2$ substrate, were recorded using Zeiss Sigma 360 VP microscope at 5.0 kV. TEM characterization was carried out using JEOL 2100F FEG microscope at 200 kV, the samples were prepared using the polymer free method[62] to transfer graphene from copper on to holey carbon TEM grids.

Raman maps were collected using a Horiba LabRam Evolution confocal Raman microscope with $50\times$ long-working-distance and a 532 nm diode laser operating at $\sim$1.1 mW. The instrument was calibrated using the Si line at 520.7 cm^{-1}. Five samples (one at each corner and one at the centre) were taken from each analysed batch and Raman maps were taken over a 13 μm × 13 μm area with 169 point spectra. All the fittings were done using Lorentzian fitting in Origin Pro 2024 software.

For machine learning training, an Excel file was imported, with each feature (ω_G, I_G, Γ_G, ω_{2D}, I_{2D}, Γ_{2D}, I_{2D}/I_G) and output label separated columns. For batch prediction, output labels were 1, 2, 3, 4, 5 and for sample position respect to the gas flow were IN, OUT and CC (centre). A total of 4225 Raman spectra were used to implement this algorithm. For the random forest classifier 100 decision trees were chosen and bootstrap sampling is used to generate randomly new data sets, for each tree, extracted from the original data set. Bootstrap sampling ensures the resulting training sets are independent and preserve the original data distribution. The original data (4225 points) set was randomly split across all spectra, ensuring all positions or batches contributed equally. The model was trained on 70% of the data, with the remaining 30% used for validation. Accuracy and the F1 score served as our performance metrics. The top permutation feature importance was considered as the indicator of which Raman features are most useful to the model for distinguishing between batches or sample position in the batch. The confusion matrix was calculated from the validation data set.

Author contributions

SMV, KSC and MFPP conceptualized the research, completed formal analysis, methodology, figure visualization and wrote the original draft. KSC was responsible for supervision, and review of the manuscript. All authors have given approval to the final version of the manuscript.

Conflicts of interest

There are no conflicts to declare.

Data availability

The data supporting this article have been included as part of the main manuscript and the Supplementary Information. All raw data are available upon request from the corresponding author.

Supplementary information: Fig. S1: temperature profile in the tube furnace measured from the centre. Fig. S2: (a) SEM image of graphene on copper showing the grain boundaries of copper, (b) AFM image of graphene on SiO$_2$/Si and (c) graphene domains exposed after growing graphene for 2 minutes and oxidation in air at 200 °C. Fig. S3: (a) histograms of the copper grains distribution and (b) graphene domains distributions from Fig. S2(c). Table S1: mean and standard deviations from the background (1800–2200 cm^{-1}) of every map and the average spectrum. Fig. S4–S8: all spectra measured for each position of each batch, after cosmic ray contribution removal, baseline correction and normalisation. The red spectrum is the averaged spectra for each position. Fig. S9: all spectra measured for each commercial brand, after cosmic ray contribution removal, baseline correction and normalisation. The red spectrum is the averaged spectra for each commercial brand. Fig. S10: G peak FWHM for every sample from each batch and the two commercial brands analysed. See DOI: https://doi.org/10.1039/d5na00668f.

Acknowledgements

References

1 K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva and A. A. Firsov, *Science*, 2004, **306**, 666–669.

2 C. Backes, A. M. Abdelkader, C. Alonso, A. Andrieux-Ledier, R. Arenal, J. Azpeitia, N. Balakrishnan, L. Banszerus, J. Barjon, R. Bartali, S. Bellani, C. Berger, R. Berger, M. M. B. Ortega, C. Bernard, P. H. Beton, A. Beyer, A. Bianco, P. Bøggild, F. Bonaccorso, G. B. Barin, C. Botas, R. A. Bueno, D. Carriazo, A. Castellanos-Gomez, M. Christian, A. Ciesielski, T. Ciuk, M. T. Cole, J. Coleman, C. Coletti, L. Crema, H. Cun, D. Dasler, D. De Fazio, N. Díez, S. Drieschner, G. S. Duesberg, R. Fasel, X. Feng, A. Fina, S. Forti, C. Galiotis, G. Garberoglio, J. M. García, J. A. Garrido, M. Gibertini, A. Gölzhäuser, J. Gómez, T. Greber, F. Hauke, A. Hemmi, I. Hernandez-Rodriguez, A. Hirsch, S. A. Hodge, Y. Huttel, P. U. Jepsen, I. Jimenez, U. Kaiser, T. Kaplas, H. Kim, A. Kis, K. Papagelis, K. Kostarelos, A. Krajewska, K. Lee, C. Li, H. Lipsanen, A. Liscio, M. R. Lohe, A. Loiseau, L. Lombardi, M. Francisca López, O. Martin, C. Martín, L. Martínez, J. A. Martin-Gago, J. Ignacio Martínez, N. Marzari, Á. Mayoral, J. McManus, M. Melucci, J. Méndez, C. Merino, P. Merino, A. P. Meyer, E. Miniussi, V. Miseikis, N. Mishra, V. Morandi, C. Munuera, R. Muñoz, H. Nolan, L. Ortolani, A. K. Ott, I. Palacio, V. Palermo, J. Parthenios, I. Pasternak, A. Patane, M. Prato, H. Prevost, V. Prudkovskiy, N. Pugno, T. Rojo, A. Rossi, P. Ruffieux, P. Samorì, L. Schué,

E. Setijadi, T. Seyller, G. Speranza, C. Stampfer, I. Stenger, W. Strupinski, Y. Svirko, S. Taioli, K. B. K. Teo, M. Testi, F. Tomarchio, M. Tortello, E. Treossi, A. Turchanin, E. Vazquez, E. Villaro, P. R. Whelan, Z. Xia, R. Yakimova, S. Yang, G. R. Yazdi, C. Yim, D. Yoon, X. Zhang, X. Zhuang, L. Colombo, A. C. Ferrari and M. Garcia-Hernandez, *2D Mater.*, 2020, **7**, 022001.

3 S. Bhaviripudi, X. Jia, M. S. Dresselhaus and J. Kong, *Nano Lett.*, 2010, **10**, 4128–4133.

4 X. Li, W. Cai, J. An, S. Kim, J. Nah, D. Yang, R. Piner, A. Velamakanni, I. Jung, E. Tutuc, S. K. Banerjee, L. Colombo and R. S. Ruoff, *Science*, 2009, **324**, 1312–1314.

5 Y. Zhang, L. Zhang and C. Zhou, *Acc. Chem. Res.*, 2013, **46**, 2329–2339.

6 S. Koh, S. Kosuga, R. Suga, S. Nagata, S. Kuromatsu, T. Watanabe and O. Hashimoto, *Carbon Reports*, 2023, **2**, 020104.

7 S. Bae, H. Kim, Y. Lee, X. Xu, J.-S. Park, Y. Zheng, J. Balakrishnan, T. Lei, H. Ri Kim, Y. Il Song, Y.-J. Kim, K. S. Kim, B. Özyilmaz, J.-H. Ahn, B. H. Hong and S. Iijima, *Nat. Nanotechnol.*, 2010, **5**, 574–578.

8 X. Hu, Y. Zhang, X. Chen, M. He and Z. Jin, *Electroanalysis*, 2023, **35**, 26.

9 D. Kireev, S. K. Ameri, A. Nederveld, J. Kampfe, H. Jang, N. Lu and D. Akinwande, *Nat. Protoc.*, 2021, **16**, 2395–2417.

10 Q. Shi, K. Tokarska, H. Q. Ta, X. Yang, Y. Liu, S. Ullah, L. Liu, B. Trzebicka, A. Bachmatiuk, J. Sun, L. Fu, Z. Liu and M. H. Rümmeli, *Adv. Mater. Interfaces*, 2020, **7**, 1902024.

11 A. Boscá, A. Ladrón-de-Guevara, J. Pedros, J. Martínez, R. Fandan and F. Calle, *Cryst. Growth Des.*, 2023, **23**, 6349–6358.

12 R. M. Jacobberger, R. Machhi, J. Wroblewski, B. Taylor, A. L. Gillian-Daniel and M. S. Arnold, *J. Chem. Educ.*, 2015, **92**, 1903–1907.

13 N. Li, R.-J. Zhang, Z. Zhen, Z.-H. Xu, R.-D. Mu and L.-M. He, *Nanotechnology*, 2021, **32**, 095607.

14 T. J. Gnanaprakasa, Y. Gu, S. K. Eddy, Z. Han, W. J. Beck, K. Muralidharan and S. Raghavan, *Microelectron. Eng.*, 2015, **131**, 1–7.

15 K. Lee and J. Ye, *Carbon*, 2016, **100**, 441–449.

16 Z. Luo, Y. Lu, D. W. Singer, M. E. Berck, L. A. Somers, B. R. Goldsmith and A. T. C. Johnson, *Chem. Mater.*, 2011, **23**, 1441–1447.

17 A. Alnuaimi, I. Almansouri, I. Saadat and A. Nayfeh, *RSC Adv.*, 2017, **7**, 51951–51957.

18 G. Deokar, J. Avila, I. Razado-Colombo, J.-L. Codron, C. Boyaval, E. Galopin, M.-C. Asensio and D. Vignaud, *Carbon*, 2015, **89**, 82–92.

19 F. Kiani, Z. Razzaghi, B. Ghadiani, M. Tamizifar, M. H. Mohmmadi and A. Simchi, *Ceram. Int.*, 2017, **43**, 15010–15017.

20 V. Miseikis, D. Convertino, N. Mishra, M. Gemmi, T. Mashoff, S. Heun, N. Haghighian, F. Bisio, M. Canepa, V. Piazza and C. Coletti, *2D Mater.*, 2015, **2**, 014006.

21 P. K. Kashyap, I. Sharma and B. K. Gupta, *ACS Omega*, 2019, **4**, 2893–2901.

22 S. J. Goldie, S. Bush, J. A. Cumming and K. S. Coleman, *ACS Appl. Nano Mater.*, 2020, **3**, 11229–11239.

23 P. Turner, K. R. Paton, E. J. Legge, A. de Luna Bugallo, A. K. S. Rocha-Robledo, A.-A. Zahab, A. Centeno, A. Sacco, A. Pesquera, A. Zurutuza, A. M. Rossi, D. N. H. Tran, D. L Silva, D. Losic, F. Farivar, H. Kerdoncuff, H. Kwon, J. Pirart, J. L. E. Campos, K. M. Subhedar, L.-L. Tay, L. Ren, L. G. Cançado, M. Paillet, P. Finnie, P. L. Yap, R. Arenal, S. R. Dhakate, S. Wood, S. Jiménez-Sandoval, T. Batten, V. Nagyte, Y. Yao, A. R. Hight Walker, E. H. Martins Ferreira, C. Casiraghi and A. J. Pollard, *2D Mater.*, 2022, **9**, 035010.

24 J. E. Proctor, E. Gregoryanz, K. S. Novoselov, M. Lotya, J. N. Coleman and M. P. Halsall, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2009, **80**, 073408.

25 T. M. G. Mohiuddin, A. Lombardo, R. R. Nair, A. Bonetti, G. Savini, R. Jalil, N. Bonini, D. M. Basko, C. Galiotis, N. Marzari, K. S. Novoselov, A. K. Geim and A. C. Ferrari, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2009, **79**, 205433.

26 J. Zabel, R. R. Nair, A. Ott, T. Georgiou, A. K. Geim, K. S. Novoselov and C. Casiraghi, *Nano Lett.*, 2012, **12**, 617–621.

27 F. Ding, H. Ji, Y. Chen, A. Herklotz, K. Dörr, Y. Mei, A. Rastelli and O. G. Schmidt, *Nano Lett.*, 2010, **10**, 3453–3458.

28 C. Casiraghi, S. Pisana, K. S. Novoselov, A. K. Geim and A. C. Ferrari, *Appl. Phys. Lett.*, 2007, **91**, 233108.

29 M. Kalbac, A. Reina-Cecco, H. Farhat, J. Kong, L. Kavan and M. S. Dresselhaus, *ACS Nano*, 2010, **4**, 6055–6063.

30 D. C. Elias, R. R. Nair, T. M. G. Mohiuddin, S. V. Morozov, P. Blake, M. P. Halsall, A. C. Ferrari, D. W. Boukhvalov, M. I. Katsnelson, A. K. Geim and K. S. Novoselov, *Science*, 2009, **323**, 610–613.

31 J. E. Johns and M. C. Hersam, *Acc. Chem. Res.*, 2013, **46**, 77–86.

32 A. J. Pollard, K. R. Paton, C. A. Clifford and E. Legge, *Characterisation of the Structure of Graphene*, NPL Management Limited, 2017.

33 Y. Xu, A. Ali, K. Shehzad, N. Meng, M. Xu, Y. Zhang, X. Wang, C. Jin, H. Wang, Y. Guo, Z. Yang, B. Yu, Y. Liu, Q. He, X. Duan, X. Wang, P. H. Tan, W. Hu, H. Lu and T. Hasan, *Adv. Mater. Technol.*, 2017, **2**, 1600241.

34 C. Neumann, S. Reichardt, P. Venezuela, M. Drögeler, L. Banszerus, M. Schmitz, K. Watanabe, T. Taniguchi, F. Mauri, B. Beschoten, S. V Rotkin and C. Stampfer, *Nat. Commun.*, 2015, **6**, 8429.

35 Y. Wang, S. W. Tong, X. F. Xu, B. Özyilmaz and K. P. Loh, *Adv. Mater.*, 2011, **23**, 1514–1518.

36 Y. Shi, X. Dong, P. Chen, J. Wang and L.-J. Li, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 2009, **79**, 115402.

37 E. Ji, M. J. Kim, J.-Y. Lee, D. Sung, N. Kim, J.-W. Park, S. Hong and G.-H. Lee, *Carbon*, 2021, **184**, 651–658.

38 O. Sul, K. Kim, E. Choi, J. Kil, W. Park and S.-B. Lee, *Nanotechnology*, 2016, **27**, 505205.

39 Z. T. Wu, W. W. Zhao, W. Y. Chen, J. Jiang, H. Y. Nan, X. T. Guo, Z. Liang, Y. M. Chen, Y. F. Chen and Z. H. Ni, *J. Raman Spectrosc.*, 2015, **46**, 21–24.

40 J. E. Lee, G. Ahn, J. Shim, Y. S. Lee and S. Ryu, *Nat. Commun.*, 2012, **3**, 1024.

41 Z. Li, W. Zhang, X. Fan, P. Wu, C. Zeng, Z. Li, X. Zhai, J. Yang and J. Hou, *J. Phys. Chem. C*, 2012, **116**, 10557–10562.

42 J. L. Beckham, K. M. Wyss, Y. Xie, E. A. McHugh, J. T. Li, P. A. Advincula, W. Chen, J. Lin and J. M. Tour, *Adv. Mater.*, 2022, **34**, 2106506.

43 H. Wahab, V. Jain, A. S. Tyrrell, M. A. Seas, L. Kotthoff and P. A. Johnson, *Carbon*, 2020, **167**, 609–619.

44 Y. Mao, N. Dong, L. Wang, X. Chen, H. Wang, Z. Wang, I. M. Kislyakov and J. Wang, *Nanomaterials*, 2020, **10**, 2223.

45 Z. Hu, J. Liu, X. Li, Z. Huang, X. Qi and W. Du, *Chemistry*, 2025, **7**, 80.

46 C. M. Acosta, E. Ogoshi, J. A. Souza and G. M. Dalpian, *ACS Appl. Mater. Interfaces*, 2022, **14**, 9418–9432.

47 P. Solís-Fernández and H. Ago, *ACS Appl. Nano Mater.*, 2022, **5**, 1356–1366.

48 L. Meng, Y. Su, D. Geng, G. Yu, Y. Liu, R.-F. Dou, J.-C. Nie and L. He, *Appl. Phys. Lett.*, 2013, **103**, 251610.

49 J. C. Meyer, A. K. Geim, M. I. Katsnelson, K. S. Novoselov, D. Obergfell, S. Roth, C. Girit and A. Zettl, *Solid State Commun.*, 2007, **143**, 101–109.

50 A. C. Ferrari and J. Robertson, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2000, **61**, 14095–14107.

51 C. Thomsen and S. Reich, *Phys. Rev. Lett.*, 2000, **85**, 5214–5217.

52 F. Tuinstra and J. L. Koenig, *J. Chem. Phys.*, 1970, **53**, 1126–1130.

53 A. C. Ferrari, J. C. Meyer, V. Scardaci, C. Casiraghi, M. Lazzeri, F. Mauri, S. Piscanec, D. Jiang, K. S. Novoselov, S. Roth and A. K. Geim, *Phys. Rev. Lett.*, 2006, **97**, 187401.

54 Y. ying Wang, Z. hua Ni, T. Yu, Z. X. Shen, H. min Wang, Y. hong Wu, W. Chen and A. T. Shen Wee, *J. Phys. Chem. C*, 2008, **112**, 10637–10640.

55 A. V Babichev, S. A. Rykov, M. Tchernycheva, A. N. Smirnov, V. Y. Davydov, Y. A. Kumzerov and V. Y. Butko, *ACS Appl. Mater. Interfaces*, 2016, **8**, 240–246.

56 R. Okuda, K. Niwano, K. Hatada, K. Kokubu, R. Suga, T. Watanabe and S. Koh, *Sci. Rep.*, 2023, **13**, 13878.

57 A. Das, B. Chakraborty, S. Piscanec, S. Pisana, A. K. Sood and A. C. Ferrari, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2009, **79**, 155417.

58 *Origin Pro, Version 2024*, OriginLab Corporation, Northampton, MA, USA.

59 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss and V. Dubourg, *J. Mach. Learn. Res.*, 2011, **12**, 2825–2830.

60 R. Kato, Y. Koga, K. Matsuishi and M. Hasegawa, *Jpn. J. Appl. Phys.*, 2017, **56**, 30307.

61 K. S. Kim, Y. Zhao, H. Jang, S. Y. Lee, J. M. Kim, K. S. Kim, J.-H. Ahn, P. Kim, J.-Y. Choi and B. H. Hong, *Nature*, 2009, **457**, 706–710.

62 G. Zhang, A. G. Güell, P. M. Kirkman, R. A. Lazenby, T. S. Miller and P. R. Unwin, *ACS Appl. Mater. Interfaces*, 2016, **8**, 8008–8016.