Supplementary Information for:

Human-AI collaborative autonomous synthesis with pulsed laser deposition for remote epitaxy

Asraful Haque[1], Daniel T. Yimam[1], Jawad Chowdhury[1], Ralph Bulanadi[1], Ivan Vlassiouk[1], John Lasseter[1], Sujoy Ghosh[1], Christopher M. Rouleau[1], Kai Xiao[1], Yongtao Liu[1], Eva Zarkadoula[1], Rama K. Vasudevan[1], Sumner B. Harris[1]*

1. Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States.

*Correspondence should be addressed to: harrissb@ornl.gov

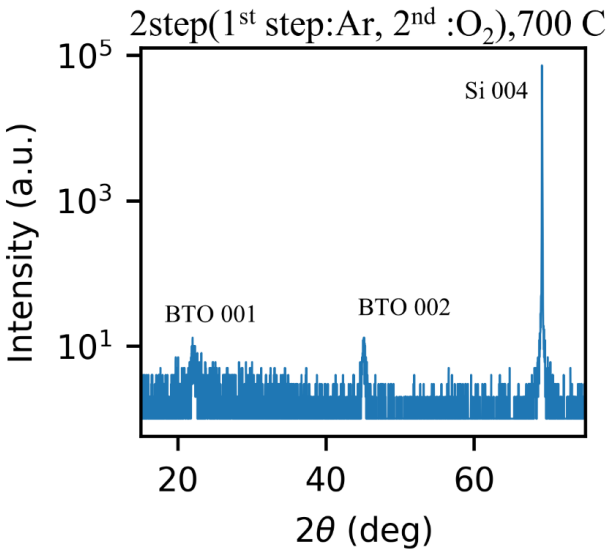


Figure S1. XRD θ-2θ scan of BTO grown on Gr/SiO_2/Si using two step deposition shows (00l) orientation.

Note S1 – Verbatim LLM Hypothesis

Using the RAG pipeline with our knowledge base, we "chatted" with the LLM to gain background on remote epitaxy, PLD of BTO, Raman analysis of graphene, providing details of our experimental capabilities. After these chats, we prompted the LLM to generate a hypothesis and propose and experiment, which is the hypothesis/experiment that we chose as a basis for the HAIC autonomous workflow presented in the main text.

User prompt:

> Based on the previous conversation, given below, state a testable hypothesis for filling the knowledge gaps in remote epitaxy of PLD using autonomous synthesis with in situ Raman spectroscopy. Suggest an experiment and provide the parameter ranges. Use only 2 to 4 PLD parameters. Target-to-substrate distance is not a parameter that we can tune so do not suggest that one.
>
> --- Previous Conversation ---
>
> [All previous prompts and responses were appended here]

ChatGPT said:

> There exists a well‑defined "soft‑landing" window of PLD growth parameters—specifically O_2 background pressure (pO_2), laser fluence (F), and substrate temperature

(T_s)—in which (a) the kinetic energy of both plume species and sputtering‐gas neutrals falls below the displacement threshold of graphene (<∼22 eV/C‐atom) and (b) $BaTiO_3$ still nucleates in exact registry to the underlying polar substrate through the graphene interlayer. In that window, the Raman defect metric I(D)/I(G) will remain below 0.1 (indicating negligible graphene damage), and ex situ structural probes (RHEED, XRD pole figures) will confirm remote epitaxy (single‐crystal, substrate‐matched $BaTiO_3$). Outside this window, either graphene damage rises sharply (I(D)/I(G)≫0.1) or film registry is lost (transition to van der Waals epitaxy).

Proposed Autonomous‐PLD Experiment

• Knobs (3):

- pO_2: 50 to 250 mTorr (steps of ∼50 mTorr)
- F (KrF excimer, 248 nm): 1.2 to 2.2 J/cm² (steps of ∼0.2 J/cm²)
- T_s: 650 to 800 °C (steps of ∼50 °C)

• Fixed parameters:

- Repetition rate: 5 Hz
- On-axis geometry, monolayer graphene on TiO_2-terminated $SrTiO_3$(001)
- Deposition thickness per run: ∼10 nm of $BaTiO_3$

• Workflow:

AI (Bayesian‐optimization) proposes (pO_2, F, T_s).

Deposit ∼10 nm $BaTiO_3$ by PLD.

Immediately acquire in situ Raman on graphene to extract I(D)/I(G).

Feed I(D)/I(G) into the optimizer (target: minimize I(D)/I(G)).

Iterate for ∼20–30 runs until convergence of parameters.

Grow a full‐thickness film at the optimized conditions and confirm remote epitaxy by RHEED oscillations and XRD pole‐figure analysis.

Expected Outcome

- A well‐localized optimum, e.g. pO_2 ≈ 120–160 mTorr, F ≈ 1.6–1.8 J/cm², T_s ≈ 700 °C, yielding I(D)/I(G) < 0.1 (minimal graphene damage) and single‐crystal $BaTiO_3$ registry.
- Outside this window: at lower pO_2 (<100 mTorr) or higher F (>2.0 J/cm²), fast neutrals and high‐energy species induce graphene defects (I(D)/I(G)≫0.1) [Chen et al. APL 103,033109 (2013) | chunk 14]; at higher T_s (>800 °C) graphene oxidizes [Wohlgemuth et al. APL Mater.12,021113 (2024) | chunk 5], and at too-low T_s (<650 °C) $BaTiO_3$ fails to crystallize [Choi et al. Science 306,1005 (2004) | chunk 8].

Filling the Gap

This experiment directly correlates PLD variables to graphene integrity (via in situ Raman) and epitaxial registry, establishing quantitative growth rules for damage‐free remote epitaxy by PLD—a gap in the current mechanistic understanding.

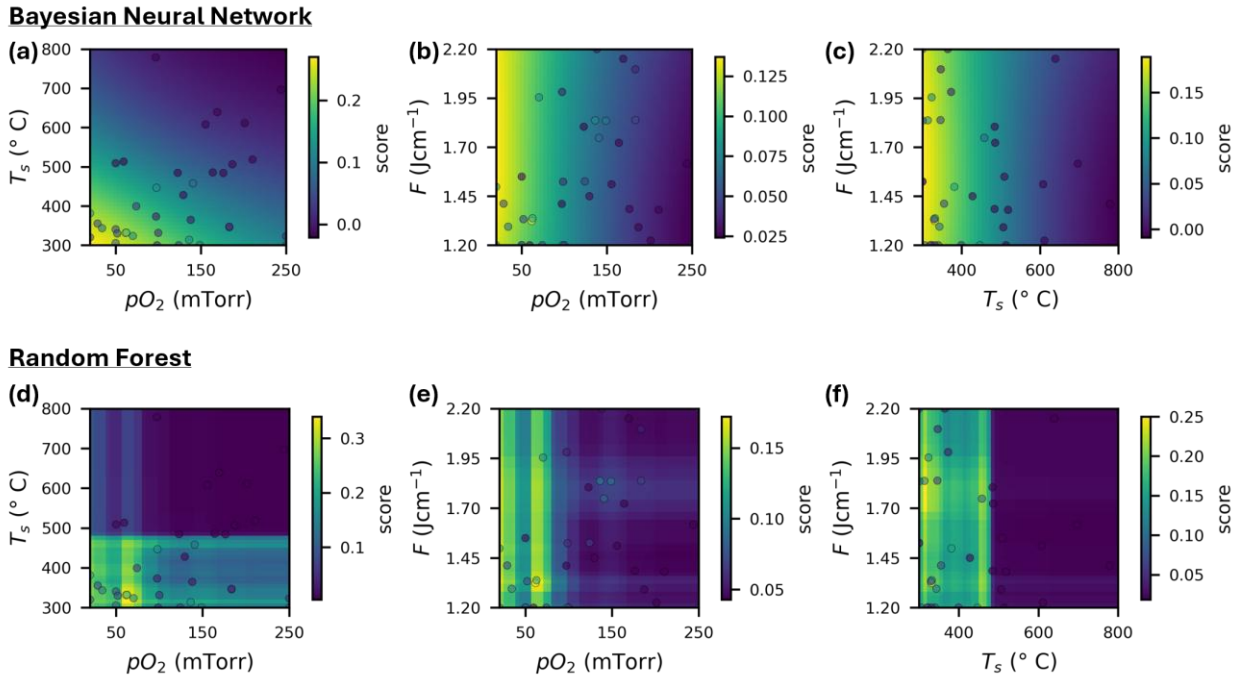


Figure S2. Kernel-free surrogate models fit with the same data from **Figure 2a-c** in the main text. (a-c) Bayesian neural network (BNN) and (d-f) Random forest (RF) surrogate models yield similar response surfaces to the Gaussian process (GP) surrogate which aid in validating the GP results. The BNN predictions are nearly identical to the GP predictions while the RF displays the usual "blocky" lookup-table-like surface. All 3 surrogate models have similar maxima regions and qualitative shapes.

Note S2 – Notable aspects of offline human-AI conference after phase 1

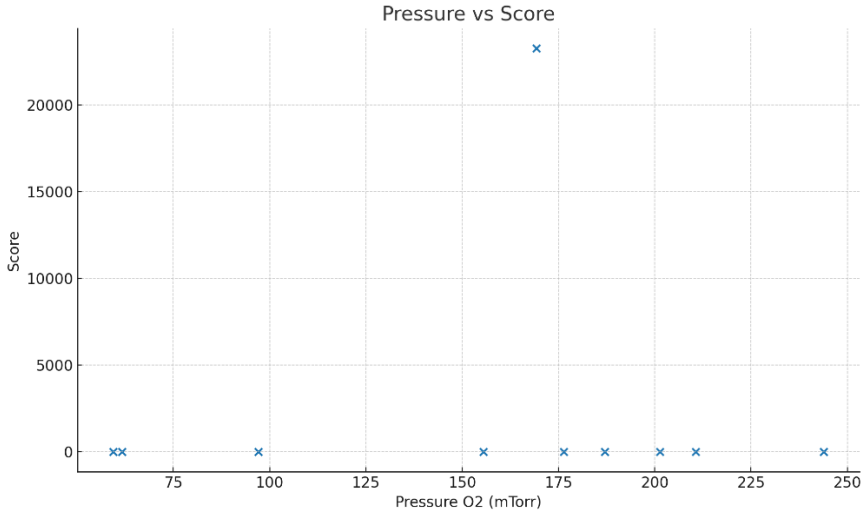
After phase 1 of the autonomous BTO growth (10 samples grown), the offline human AI conference resulted in the first of several changes to the Raman analysis and found an outlier point (caused by bad curve fitting) which skewed the GP. Below, we show an excerpt from this human-AI conference to highlight the automatic data analysis the LLM conducted to identify the outlier and the recommended fixes. Parts of the prompts and responses are abbreviated in brackets for brevity.

User prompt:

> *[PNG images of all 10 Raman spectra fits were uploaded to the LLM]*
>
> I am fitting Raman spectra of graphene before and after depositing BaTiO3 on it to search for conditions which lead to the maximum of our score metric. My score should be bounded between 0-1 but the fitting function is not working robustly, leading to bad fits and anomalously large Y values, which may mislead my gaussian process surrogate model during Bayesian optimization. Please analyze the current graphene fitting function code and plot results so far as well as the X,Y values used in the BO. X values are [Pressure O2, Temperature, Laser energy] and Y is the score returned from the graphene_objective function. Several of the samples show that the graphene is completely destroyed after deposition which is expected. But it is clear that the fitting function and score calculation is returning unexpected values. Plots of the graphene fitting results are attached.
>
> --- Current X, Y Data for GP model ---
>
> X=(tensor([[61.6309, 331.4551, 29.1189], [176.3493, 484.1429, 30.4635], [201.3447, 611.2147, 26.9216], [210.6028, 518.4784, 30.3651], [97.0653, 778.7671, 31.0331], [169.2077, 639.2903, 47.3119], [187.0194, 506.3742, 28.4088], [155.4756, 607.8483, 33.2359], [243.8909, 696.5065, 35.5431], [59.3433, 513.2381, 26.4000]], dtype=torch.float64), Y=tensor([[5.4303e-01], [4.4023e-08], [7.7948e-01], [1.6241e-22], [4.0972e-01], [2.3269e+04], [5.7596e-02], [5.3415e-08], [1.1697e-14], [1.8283e-02]], dtype=torch.float64))
>
> --- Code for graphene fitting ---
>
> [*The full code for the graphene fitting functions is pasted here*]

ChatGPT said:

> [*ChatGPT created scatter plots of the score vs growth parameter, shown below is just 1 of the 3 plots to display the outlier point]*



We can see straight away from the table and scatter - plots that, while most of your scores lie nicely in [0,1], the point at (P≈169 mTorr, T≈639 °C, E≈47 mJ) shoots up to ≈2.3 × 10^4. That single outlier will dominate your GP and lead BO astray.

What's going wrong in your fitting / scoring

1. Normalization by a tiny "global max".

When the graphene is effectively destroyed, your filtered intensity has almost no G-peak and the global maximum (used to normalize) collapses to some small noise level. Dividing by this tiny number wildly amplifies both D and G areas and produces huge ratios.

2. Dividing by near-zero denominators.

You compute

$rg=I_{G,treated}/I_{G,ref}$, $\Delta=I_{D,treated}/I_{G,treated}-I_{D,ref}/I_{G,ref}$, $score=rg/(1+\Delta)$.

If $I_{G,ref}$ or $1+\Delta$ is nearly zero you get a numerical blow-up.

3. Poor backgrounds → bad peak - area estimates.

A simple quadratic background fit over the entire 1100–1700 cm^{-1} range is not robust when only a tiny residual "bump" remains.

Recommended mitigations

1. Reject or clamp destroyed samples up front.

If the total G-peak amplitude from the reference is below some threshold (e.g. $I_{G,ref}<\epsilon$), or if the treated spectrum shows no discernible G, assign a safe score (e.g. zero) rather than trying to fit.

2. Normalize to a more stable reference.

Instead of global_max of the treated spectrum, normalize both reference and treated to the same baseline. For example, fit the G-peak height on the reference and divide both spectra by that. That way a missing peak on the treated side gives rg≈0 rather than ∞.

3. Bound your score.

After computing, clip: score = max(0.0, min(score, 1.0))

This prevents any outliers from wrecking your surrogate's kernel fit.

4. Improve background fitting locally.

Fit the baseline only in flanking regions around each peak (e.g. [1500–1550] and [1650–1700]) rather than across the entire Raman window. This gives a more accurate area under the peak and avoids large negative residuals.

5. Add fit quality checks.

Inspect result_trt, redchi or the ratio of residual RMS to peak height. If the fit quality is poor (e.g. reduced $\chi^2 \gg 1$), again fallback to a default score.

With those changes you'll keep all returned scores safely in [0,1], and the BO surrogate will see a consistent, meaningful response surface rather than be dominated by a single numerical artifact.

Following this initial suggestion, we continued a lengthy exchange with the LLM to update and test the fitting functions, refining the LLM's initial suggestions.

Note S3 - LLM assisted iterative refinement of Raman analysis

The LLM assisted in 3 revisions of the graphene fitting function after phase 1, phase 2, and phase 3 of the human-AI cooperative autonomous synthesis experiment. Revision 1 introduced a shared normalization, peak flank-based local baseline, and a 3-peak model which conditionally added polymer contaminate peaks. Including contaminate peaks remedied a "known unknown" in the spectra that we did not realize we needed to account for before experiments. Revision 2 added the D' peak to the fit and a G peak amplitude-based retention ratio (changing from an area ratio) to improve fitting in the G region. Finally, revision 3 added a global gain correction to re-scale the "after deposition" spectra to account for the increased background scattering caused by the BTO deposition (an example of dealing with an "unknown unknown" effect that was only found after conducting experiments.) Below, the changes are detailed in **Table S1**.

Table S1. Summary of Raman curve fitting function updates resulting from offline human-AI conferences

Aspect	Original	rev1	rev2	rev3
Pre-baseline step	none	Subtract last point	same	same
Shared normalization	No (each by own max)	Yes (norm_ref from ref)	Yes	Yes
Local baseline (flanks)	In-model poly bkg on full region	Yes (1100–1300 & 1650–1700)	Yes	Yes (with data window effectively 1200–1700)
Data window for fit (cm^{-1})	1100–1700	1100–1700	1100–1700	1200–1700 (code)
Default peaks	D, G	D, 1450, G	D, 1450, G, D′	D, 1450, G, D′ (+ initial guesses)
D′ > G constraint	—	—	Yes	No
Contaminants (1480, 1520)	—	Conditional (only if fit poor)	Always (if-check disabled)	Always
Gain correction (treated↔ref)	—	—	—	Yes (s_est)
r_g definition	G area ratio	G area ratio	G amplitude ratio	G amplitude ratio (after gain correction)
Reference gate (I_g_ref, redchi)	—	Yes	Yes	Removed (only norm_ref too small aborts)

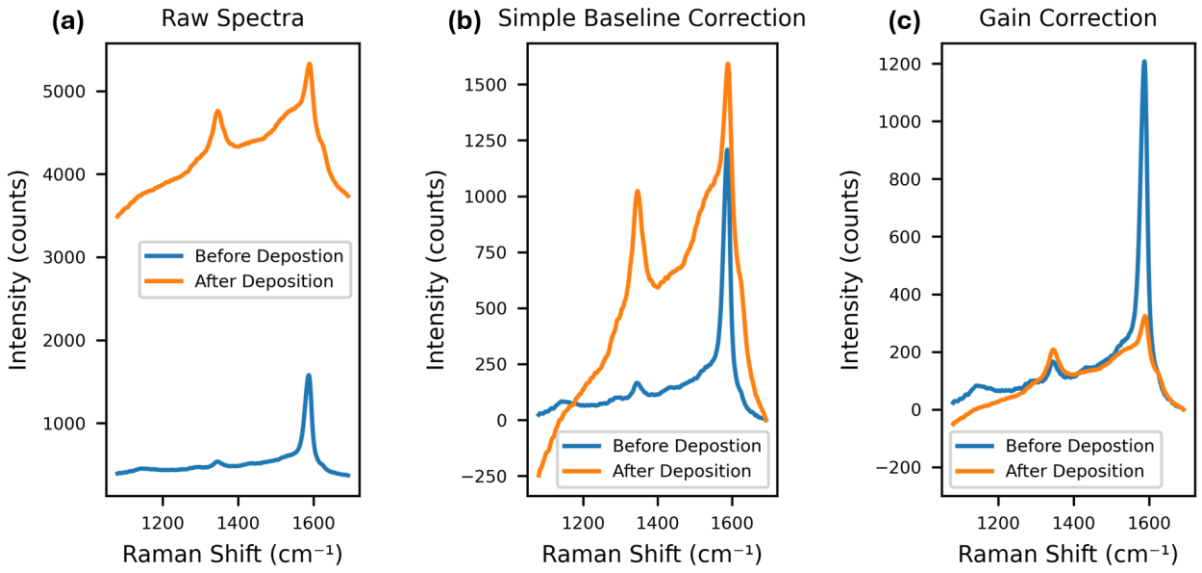


Figure S3. Representative example of the final preprocessing procedure for Raman spectra before curve fitting for the sample grown at pO_2 = 50 mTorr, T_s = 305 °C, F = 1.2 J cm^{-2}. (a) Raw Raman spectra after deposition have increased background scattering due to the presence of the $BaTiO_3$ thin film, which complicates the relative analysis of Raman peak areas. (b) A simple baseline correction is applied by subtracting the intensity of the last data point. This is not enough because it is physically impossible for the G peak area to increase after deposition (potential for graphene to etch/oxidize implies $G/G_0 \leq 1$) (c) Finally, the global scale is estimated and the "after deposition" spectrum is corrected to yield the final preprocessed spectra for curve fitting. The global scale (s) is estimated as the median of the ratio of the D and G peak region median intensities i.e. $r_D = m_{D,after}/m_{D,before}$, $r_G = m_{G,after}/m_{G,before}$, give the global scale factor $s = median(r_D, r_G)$. The "after deposition" spectrum is divided by s to correct the intensity. Notably, this generally maintains the intensity of polymer contaminate peaks, which should not change significantly (see **Figure S4** below). The global scale factor was $s = 4.912$ in this case.

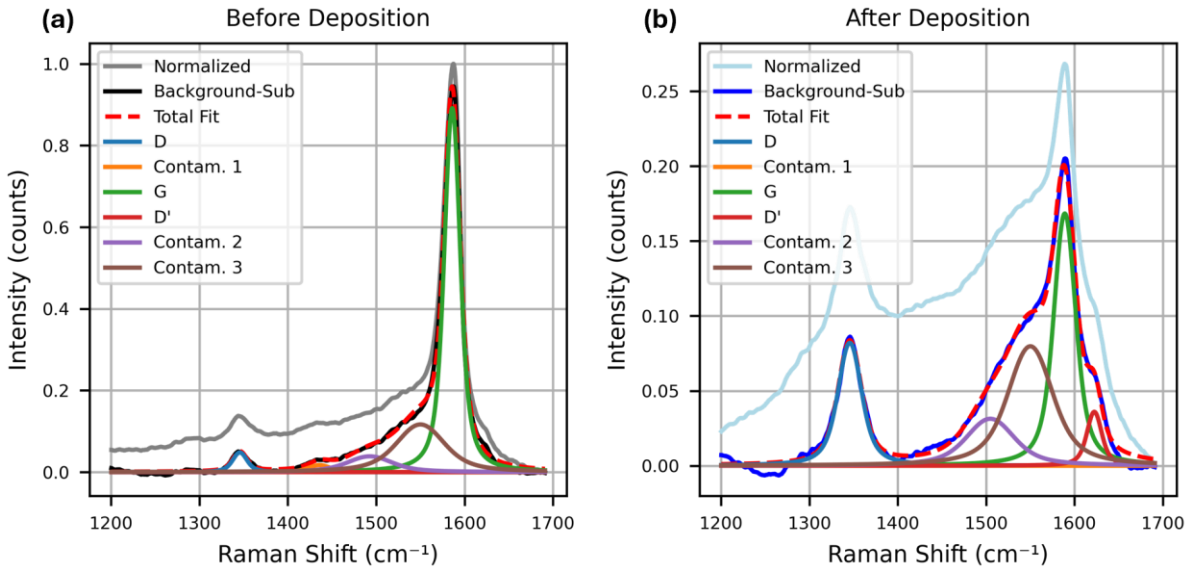


Figure S4. Representative example of graphene Raman spectra curve fitting procedure, following the preprocessing for intensity correction shown in Figure S3, for the sample grown at pO_2 = 50 mTorr, T_s = 305 °C, F = 1.2 J cm^{-2}. Each spectrum is fit to extract I_D and I_G accounting for polymer contaminate peaks and a polynomial background. (a) Before $BaTiO_3$ deposition and (b) after $BaTiO_3$ deposition. Note that the polymer contaminates peaks (Contam. 2, Contam. 3) approximately maintain their absolute intensity, which helps verify that the global intensity correction applied to the "after deposition" spectrum is appropriate. For this sample, the change in I_D/I_G was Δ = (0.05-0.529) = 0.480 and the retention ratio was $r_G = 0.259$ to give a score of 0.175.

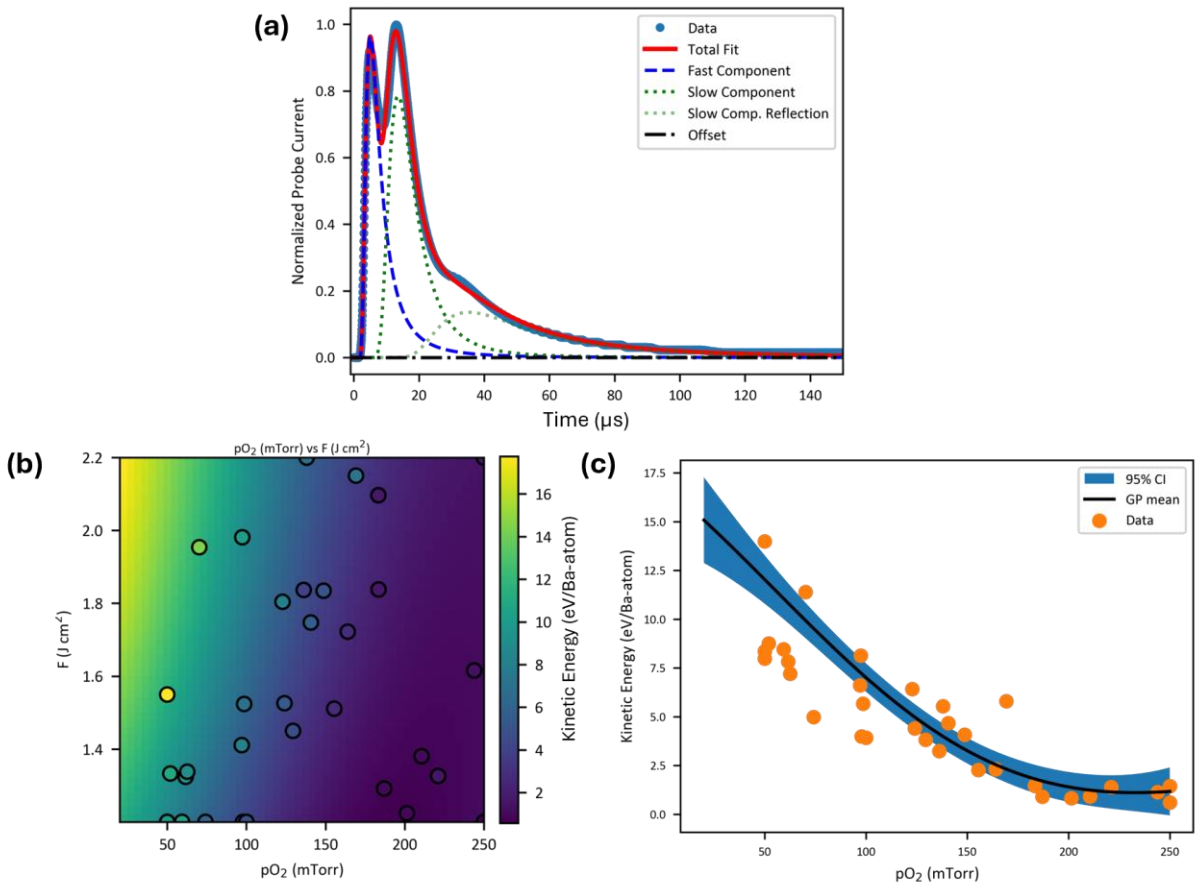


Figure S5. Most probable kinetic energy (MPKE) of the slow component of the $BaTiO_3$ plume was fit from the ion probe collected during the autonomous experiment. (a) Example ion probe trace taken at pO_2 = 61.6 mTorr and F = 1.32 J cm^{-2} showing the "fast" component, the scattered "slow" component, and an artifact peak (slow comp. reflection) arising from the plume "bouncing off" the substrate holder and interacting with the ion probe again. MPKE is taken from the time-of-flight (TOF) of a component peak position and calculated as $MPKE = 0.5\, m\, (D/TOF)^2$ where D is the target-probe distance which is 4.5 cm in this case and m is the atomic mass of Ba (137.33 u). (b) Gaussian process model mean predictions of the slow component in the F, pO_2 parameter space. (c) GP predictions averaged over fluence displays the usual exponential decay behavior of

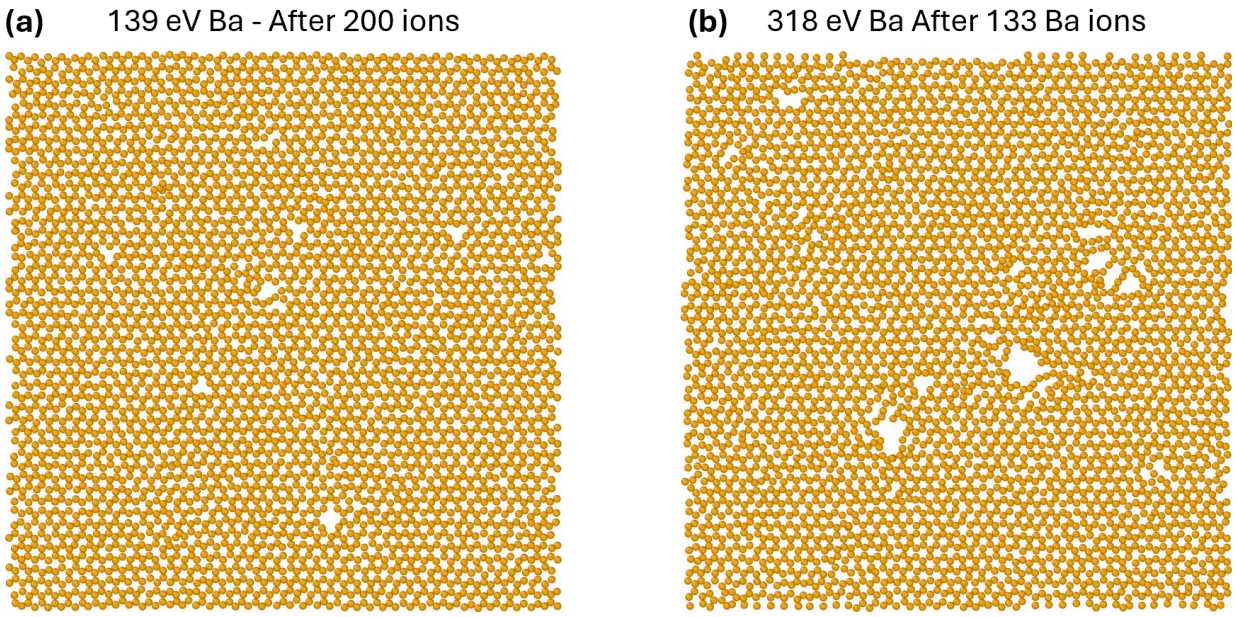


Figure S6. Final results of Ba ion bombardment of graphene on an SiO_2 substrate. (a) After 200 ions with kinetic energy of 138 eV/Ba-atom. (b) After 133 ions with kinetic energy of 318 eV/Ba-atom.

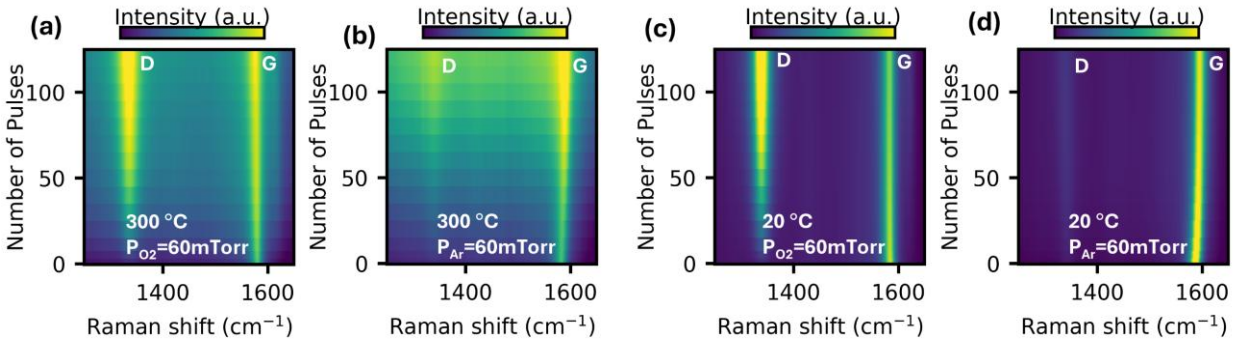


Figure S7. Raman heat maps at 300 °C versus laser pulses in (a) Ar and (b) O_2. Raman heat maps at 20 °C versus laser pulses in (c) O_2 and (d) Ar.

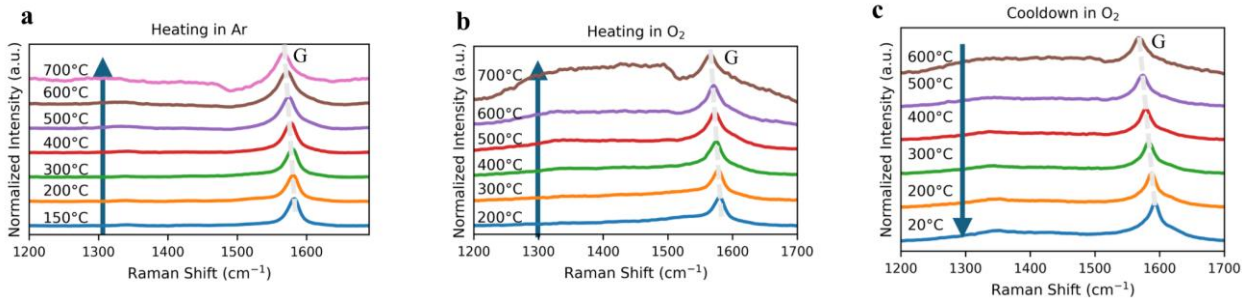


Figure S8. Raman spectra of Graphene on SiO_2/Si collected during annealing: (a) Heating in Ar (150 °C to 700 °C): no prominent D band; (b) Heating in O_2 (200 °C to 700 °C): similar absence of D activation; (c) Cooldown in O_2 (600 °C to 20 °C): G shifts without D-band activation. We observe G shifts after cooldown indicating charge transfer rather than disorder, marking the usual p-type doping when transferred graphene is annealed on SiO_2.

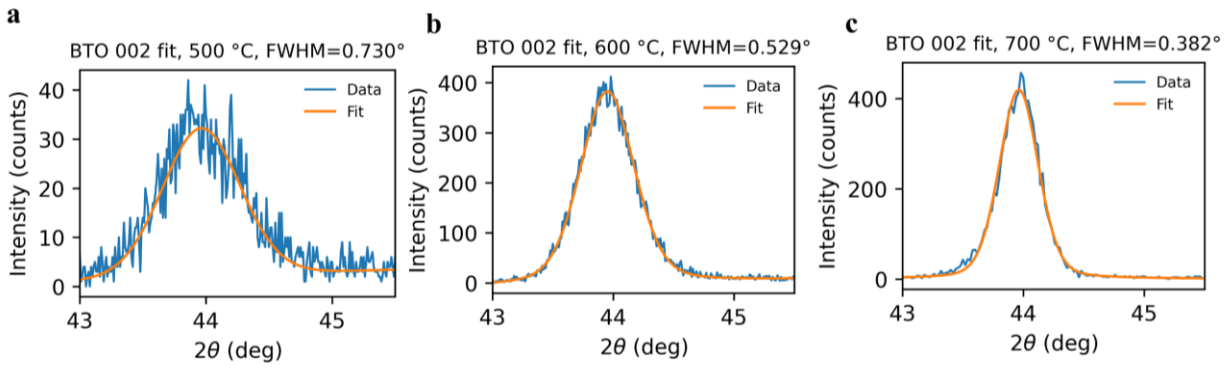


Figure S9. (a-c) XRD θ-2θ data and pseudo-Voigt fits of the BTO 002 reflection for films grown at 500, 600, and 700 °C, showing a systematic reduction in FWHM with increasing growth temperature.

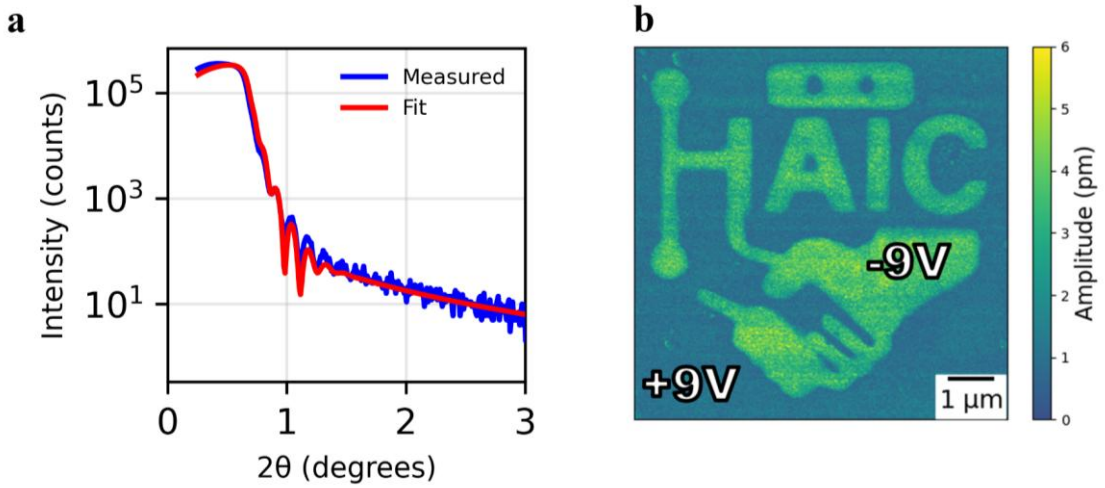


Figure S10. (a) X-ray reflectivity (XRR) of BTO (50.7 nm thick) grown on Nb-STO substrate following two step growth method. (b) PFM amplitude of 2-step grown BTO on Nb-STO substrate after patterning.

GRAPHENE CHARACTERIZATION

Monolayer appearance

Generally, a Raman intensity ratio of the 2D to G band (I_D/I_G) greater than 1, together with a single Lorentzian line shape of the 2D band, is considered a fingerprint of monolayer graphene[1,2]. Undoped graphene typically exhibits an I_D/I_G ratio of ~3, which decreases to ~1.5 at doping levels of $5x10^{12}$ cm^{-2} [3] accompanied by other spectral changes such as band position shifts and variations in FWHM. Our synthesized graphene shows an I_D/I_G ratio exceeding 3.4 in the undoped, suspended state and decreasing to ~1.5 under electrostatic doping induced by substrate surface termination groups (hydroxyls) when supported on a silica substrate. Placement on the substrate also, as expected, increases the 2D-band FWHM from ~30 cm^{-1} to ~42 cm^{-1} and blueshifts the band position by ~15 cm^{-1} (**Figure S11 a, b**).

Uniformity

Our graphene samples are >99% monolayer, as assessed by optical microscopy and scanning electron microscopy. We provide Raman mapping acquired with a 50x objective, confirming the uniformity of the samples. (**Figure S11 c, d**).

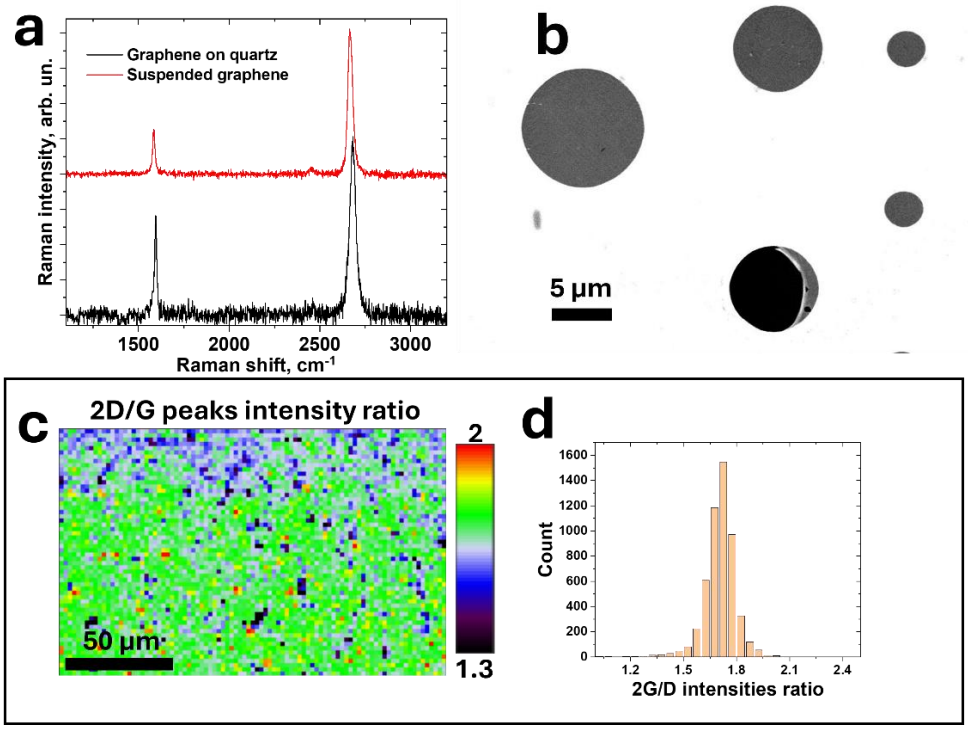


Figure S11. *Raman characterization of synthesized monolayer graphene (532nm).* (**a**) Comparison of Raman spectra for suspended graphene and graphene supported on a substrate. (**b**) SEM images of suspended graphene; a ruptured hole is also shown for contrast. (**c**) I_{2D}/I_G mapping confirming the high uniformity of the synthesized graphene. (**d**) Histogram corresponding to panel (c).

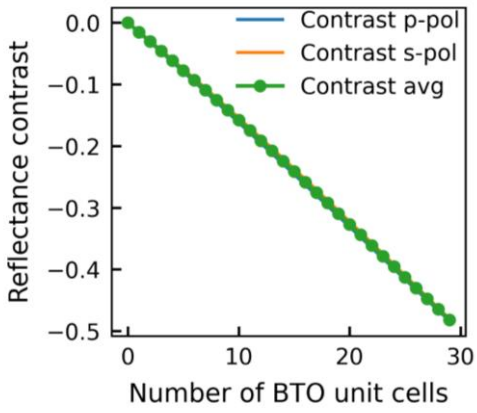


Figure S12. Reflectance contrast vs number of BTO unit cell for a heterostructure of BTO on monolayer graphene coated SiO_2/Si substrate.

Table S2: PLD grown oxides using remote epitaxy and exfoliation status: data extracted using RAG tool.

S.No	**Heterostructure**	**Synthesis method**	**Graphene (synthesis method and no. of layers used)**	**Exfoliation demonstrated**	**Ref**
1	BTO/Gr//STO	PLD	SiC, Bilayer	Yes	[4]
2	BTO/Gr//STO	MBE	SiC, Bilayer	Yes	[5]
3	BTO/Gr//STO	PLD	CVD, Bilayer	Yes	[6]
4	BTO/Gr/SVO//STO	PLD	CVD, Mono	Transferred utilizing graphene enhanced dissolution of SVO	[7]
4	VO_2/Gr//Sapphire	PLD	CVD, Monolayer	< 10 µm	[8]
5	STO/Gr//STO	MBE	CVD, Monolayer	Yes	[9]
	BTO/Gr//Ge	PLD	Graphene grown on Ge substrate CVD	yes	[10]
6	STO/Gr//STO	PLD	CVD, Monolayer	No	[11]
7	HfO_2/Gr//STO	PLD	CVD, Monolayer	No	[12]
8	BTO/Gr//STO	PLD	CVD, Monolayer	Yes	This work

References:

1 Ferrari, A. C.; Meyer, J. C.; Scardaci, V.; Casiraghi, C.; Lazzeri, M.; Mauri, F.; Piscanec, S.; Jiang, D.; Novoselov, K. S.; Roth, S.; Geim, A. K. Raman Spectrum of Graphene and Graphene Layers. *Physical Review Letters* **97**, 187401 (2006). https://doi.org/10.1103/PhysRevLett.97.187401

2 Graf, D.; Molitor, F.; Ensslin, K.; Stampfer, C.; Jungen, A.; Hierold, C.; Wirtz, L. Spatially Resolved Raman Spectroscopy of Single- and Few-Layer Graphene. *Nano Letters* **7**, 238-242 (2007). https://doi.org/10.1021/nl061702a

3 Das, A.; Pisana, S.; Chakraborty, B.; Piscanec, S.; Saha, S. K.; Waghmare, U. V.; Novoselov, K. S.; Krishnamurthy, H. R.; Geim, A. K.; Ferrari, A. C.; Sood, A. K. Monitoring dopants by Raman scattering in an electrochemically top-gated graphene transistor. *Nature Nanotechnology* **3**, 210-215 (2008). https://doi.org/10.1038/nnano.2008.67

4 Kum, H. S.; Lee, H.; Kim, S.; Lindemann, S.; Kong, W.; Qiao, K.; Chen, P.; Irwin, J.; Lee, J. H.; Xie, S.; Subramanian, S.; Shim, J.; Bae, S. H.; Choi, C.; Ranno, L.; Seo, S.; Lee, S.; Bauer, J.; Li, H.; Lee, K. *et al.* Heterogeneous integration of single-crystalline complex-oxide membranes. *Nature* **578**, 75-81 (2020). https://doi.org/10.1038/s41586-020-1939-z

5 Lee, S.; Zhang, X.; Abdollahi, P.; Barone, M. R.; Dong, C.; Yoo, Y. J.; Song, M. K.; Lee, D.; Ryu, J. E.; Choi, J. H.; Lee, J. H.; Robinson, J. A.; Schlom, D. G.; Kum, H. S.; Chang, C. S.; Seo, A.; Kim, J. Route to Enhancing Remote Epitaxy of Perovskite Complex Oxide Thin Films. *ACS Nano* **18**, 31225-31233 (2024). https://doi.org/10.1021/acsnano.4c09445

6 Haque, A.; Mandal, S. K.; Parate, S. K.; HJ, D. S.; Nukala, P.; Raghavan, S. Free standing epitaxial oxides through remote epitaxy: the role of the evolving graphene microstructure. *Nanoscale* **17**, 2020-2031 (2025). https://doi.org/10.1039/d4nr03356f

7 Haque, A.; Mandal, S. K.; Jeyaseelan, A.; Vura, S.; Nukala, P.; Raghavan, S. Remote epitaxy-based atmospherically stable hybrid graphene template for fast and versatile transfer of complex ferroelectric oxides onto Si. *Materials Today Electronics* **8** (2024). https://doi.org/10.1016/j.mtelec.2024.100091

8 Guo, Y.; Sun, X.; Jiang, J.; Wang, B.; Chen, X.; Yin, X.; Qi, W.; Gao, L.; Zhang, L.; Lu, Z.; Jia, R.; Pendse, S.; Hu, Y.; Chen, Z.; Wertz, E.; Gall, D.; Feng, J.; Lu, T. M.; Shi, J. A Reconfigurable Remotely Epitaxial VO(2) Electrical Heterostructure. *Nano Lett* **20**, 33-42 (2020). https://doi.org/10.1021/acs.nanolett.9b02696

9 Yoon, H.; Truttmann, T. K.; Liu, F.; Matthews, B. E.; Choo, S.; Su, Q.; Saraswat, V.; Manzo, S.; Arnold, M. S.; Bowden, M. E.; Kawasaki, J. K.; Koester, S. J.; Spurgeon, S. R.; Chambers, S. A.; Jalan, B. Freestanding epitaxial SrTiO(3) nanomembranes via remote epitaxy using hybrid molecular beam epitaxy. *Sci Adv* **8**, eadd5328 (2022). https://doi.org/10.1126/sciadv.add5328

10 Dai, L.; Zhao, J.; Li, J.; Chen, B.; Zhai, S.; Xue, Z.; Di, Z.; Feng, B.; Sun, Y.; Luo, Y.; Ma, M.; Zhang, J.; Ding, S.; Zhao, L.; Jiang, Z.; Luo, W.; Quan, Y.; Schwarzkopf, J.; Schroeder, T.; Ye, Z. G. *et al.* Highly heterogeneous epitaxy of flexoelectric BaTiO(3-delta) membrane on Ge. *Nat Commun* **13**, 2990 (2022). https://doi.org/10.1038/s41467-022-30724-7

11 Wohlgemuth, M. A.; Trstenjak, U.; Sarantopoulos, A.; Gunkel, F.; Dittmann, R. Control of growth kinetics during remote epitaxy of complex oxides on graphene by pulsed laser deposition. *APL Materials* **12** (2024). https://doi.org/10.1063/5.0180001

12 Trstenjak, U.; Goß, K.; Gutsche, A.; Jo, J.; Wohlgemuth, M.; Dunin-Borkowski, R. E.; Gunkel, F.; Dittmann, R. Heterogeneous Integration of Graphene and HfO2 Memristors. *Advanced Functional Materials* **34**, 2309558 (2024). https://doi.org/https://doi.org/10.1002/adfm.202309558