

S1: Dependence of NMF decomposition on the number of components

The dependence of the NMF decomposition on the number of components k were systematically examined. At smaller values of k , distinct spectral features associated with BL and TL graphene are merged into a single basis vector, obscuring physically meaningful distinctions. In contrast, at larger values of k , the decomposition yields multiple similar basis vectors that primarily capture minor spectral fluctuations without introducing new physical information. Based on these observations, $k = 7$ is adopted as a practical choice that provides a robust and physically interpretable decomposition. We note, however, that this selection is not unique, and the principal physical conclusions remain unchanged for nearby values of k (e.g., $k = 6$ and $k = 8$).

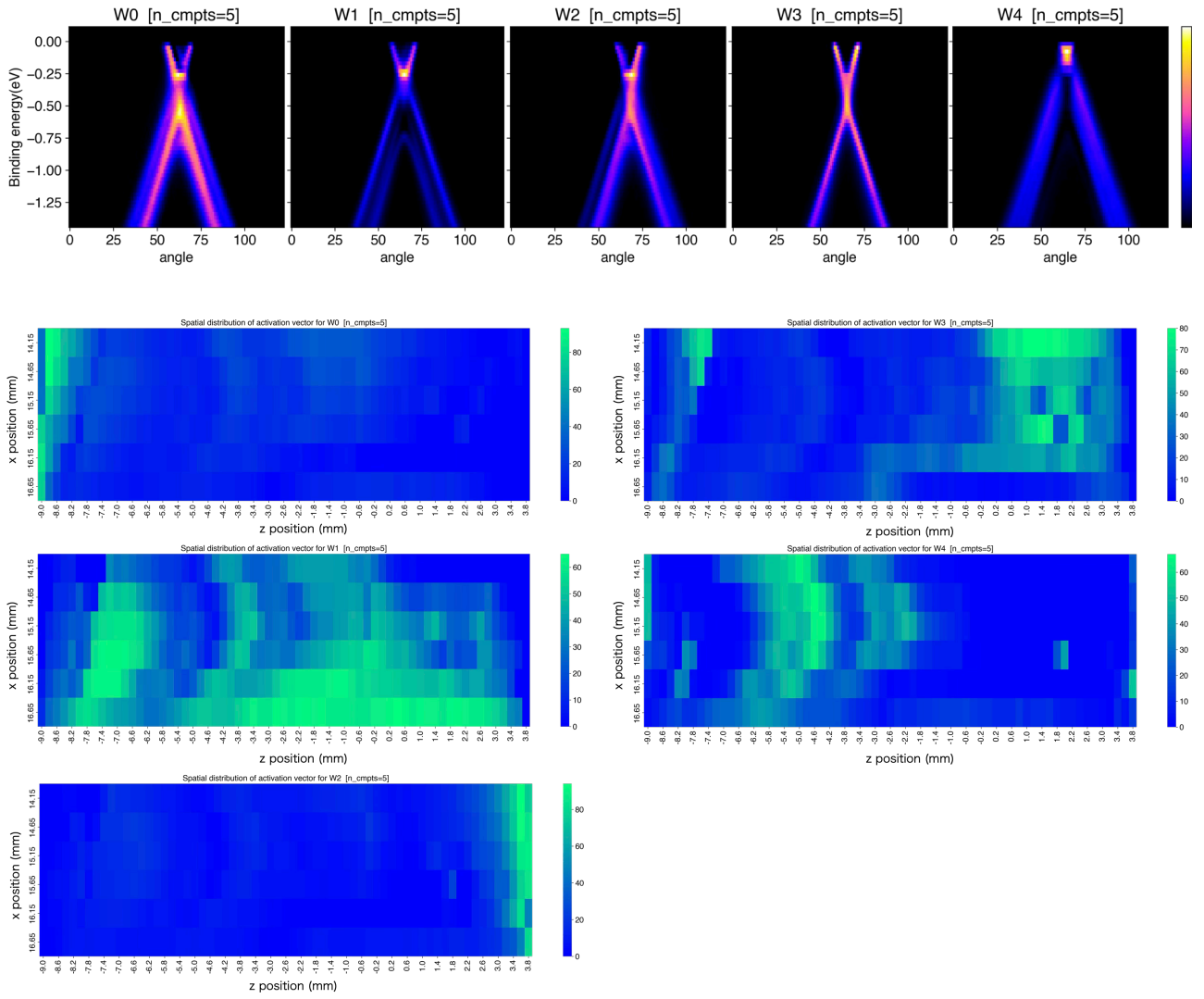


Figure. S1. (a) Basis vectors and heatmaps of activation vector obtained from NMF with $k=5$.

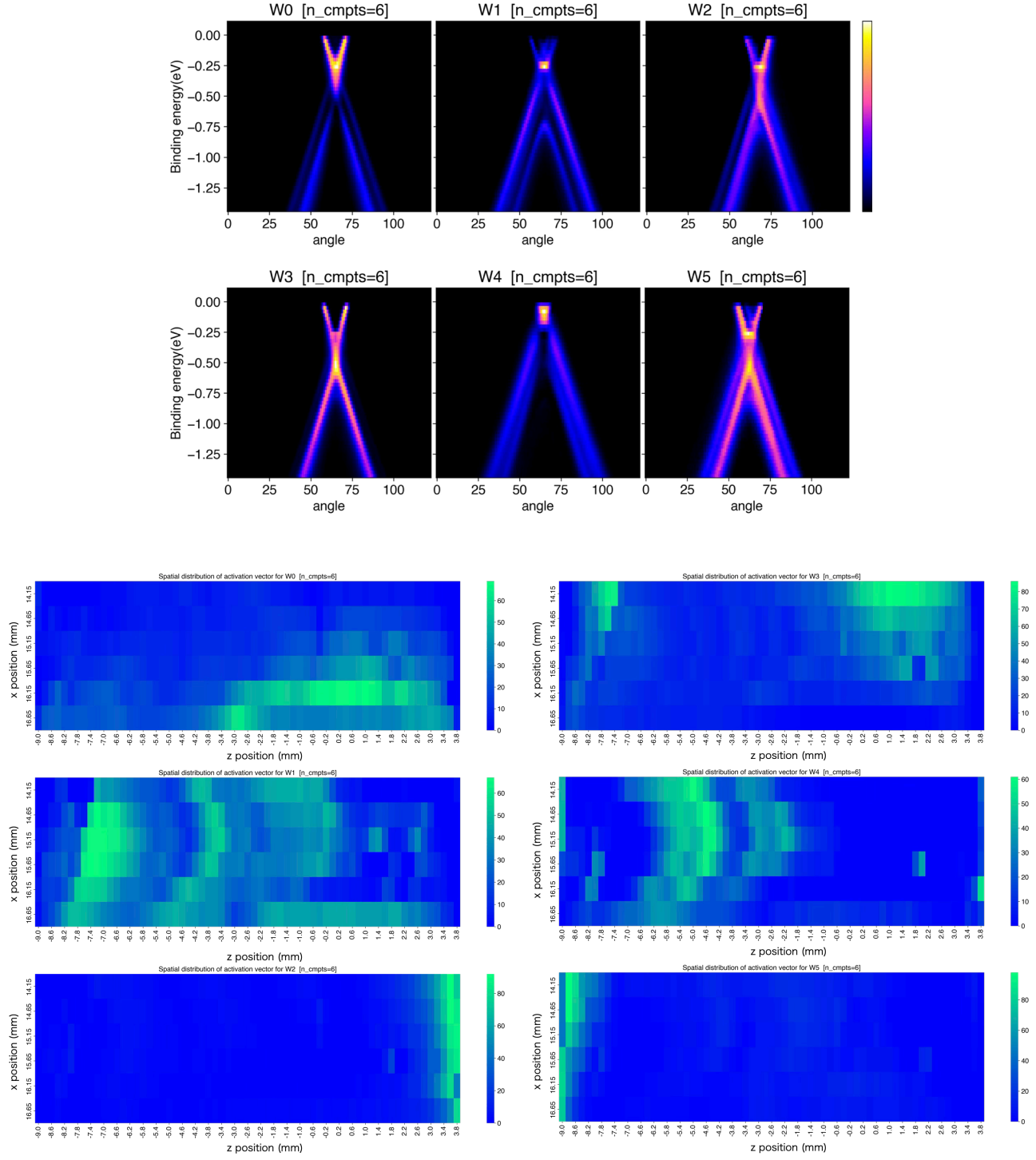


Figure. S1. (b) Basis vectors and heatmaps of activation vector obtained from NMF with $k=6$.

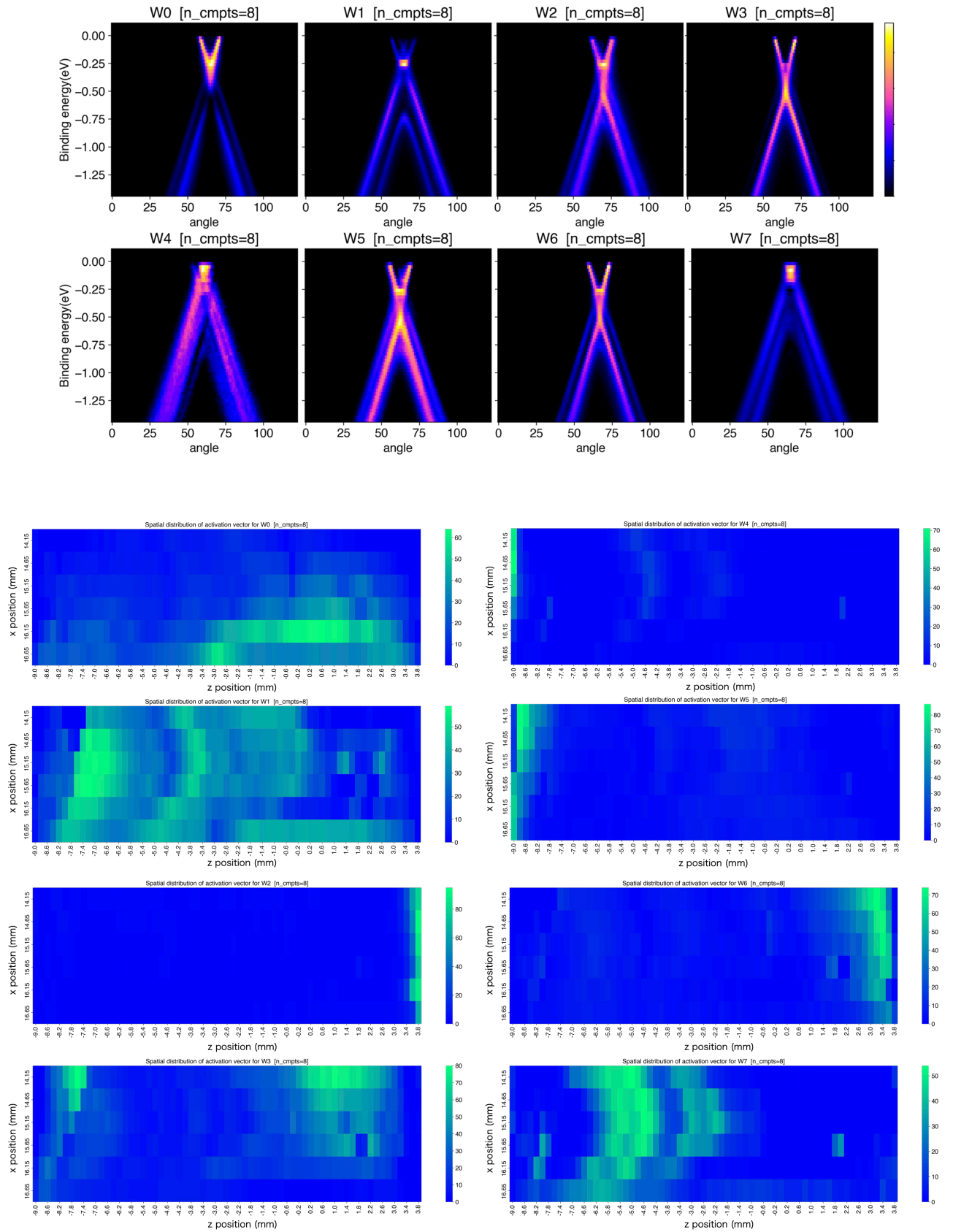
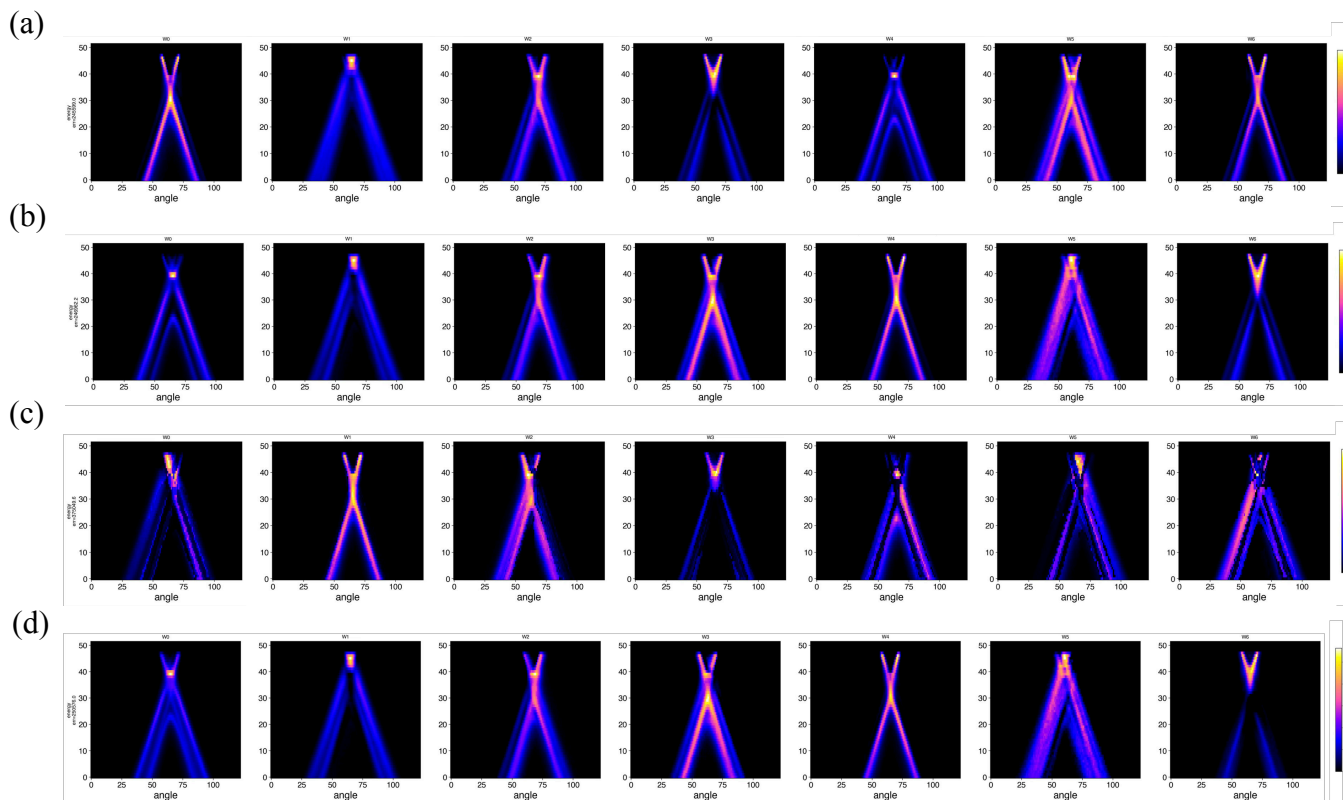


Figure. S1. (c) Basis vectors and heatmaps of activation vector obtained from NMF with $k=8$.

S2. Robustness for initialization method

To assess the dependence of the NMF decomposition on initialization, we performed the analysis at a fixed number of components ($k = 7$) using several initialization methods implemented in scikit-learn, namely 'random', 'nndsvd', 'nndsvda', and 'nndsvdar'. For the 'nndsvd' initialization, the solver was set to coordinate descent ('cd') instead of the default setting. The resulting basis matrices (W) are compared in Fig. S2 (a) - (d). The decompositions obtained with (a) 'random' and (b) 'nndsvd' initialization yield similar basis vectors and comparable reconstruction errors (SSE), indicating consistent and stable solutions. In contrast, (c) 'nndsvda' and (d) 'nndsvdar' result in larger SSE values and basis vectors that are less consistent with those obtained from other initializations, reflecting poorer reconstruction of the original spectra. The heatmaps of activation vector obtained with 'nndsvd' is qualitatively similar to that obtained with random initialization, further supporting the robustness of the results. (Fig. S2(e))

These results highlight the importance of initialization choice and demonstrate that both quantitative metrics (such as SSE) and qualitative inspection of the basis vectors are necessary to ensure physically meaningful decomposition.



(e)

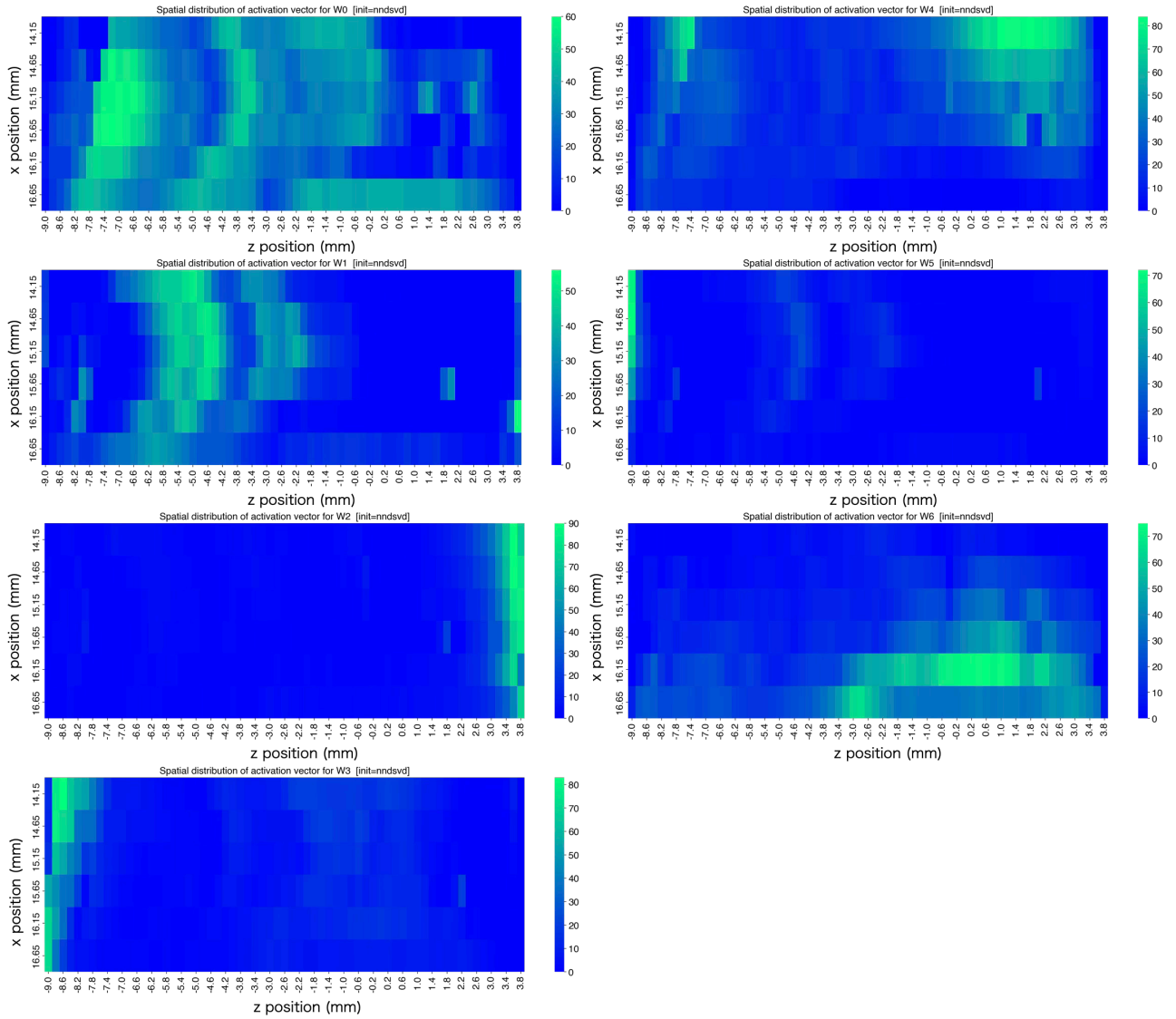


Figure S2. Comparison of NMF results obtained using different initialization methods at $k = 7$. Basis matrices (W) obtained with (a) ‘random’, (b) ‘nndsvd’, (c) ‘nndsvda’, and (d) ‘nndsvdar’ initialization. The results from ‘random’ and ‘nndsvd’ show similar basis vectors and comparable reconstruction errors (SSE), whereas ‘nndsvda’ and ‘nndsvdar’ yield larger SSE values and less consistent components. (e) Activation vector heatmaps corresponding to the ‘nndsvd’ initialization.

S3. Spatial distribution of clustering label with k-means method

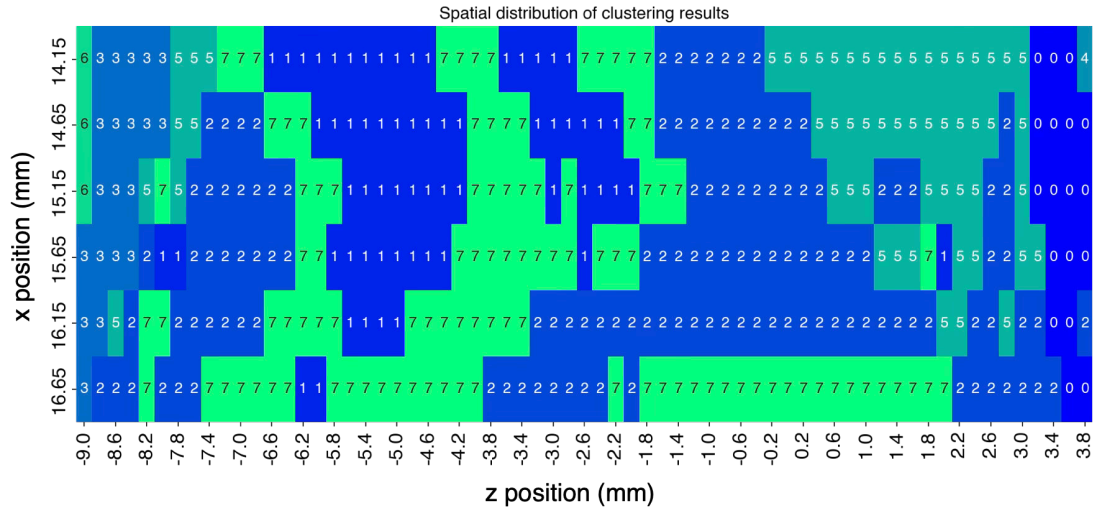


Figure. S3. Spatial distribution of cluster labels assigned to each ARPES spectrum based on the k-means method using the spectra as features. The vertical and horizontal axes correspond to the relative positions of the ARPES measurement stage (z-axis and x-axis, respectively), reflecting the spatial coordinates of the graphene on the SiC substrate. The numerical values in the map represent the cluster labels assigned to each measurement point.

S4. Deflection due to the electric field at the edge

Fig. S4(a), (b), and (c) show heatmaps of the spatial intensity distributions of the basis vectors W2, W5, and W6, respectively. The regions with high intensities of W2 and W6 are concentrated near the right edge of the spatial map. As shown in Fig. 2, both W2 and W6 exhibit spectral shifts toward higher emission angles. Among them, W2 shows a larger shift than W6, and its high-intensity region appears further to the right on the map. In contrast, the high-intensity region of W5, which exhibits a shift toward lower angles, is localized near the left edge of the spatial map.

The observed spectral shifts can be plausibly attributed to the presence of a localized electric field near the sample edges. During the ARPES measurements, the SiC substrate with grown graphene was fixed to the sample holder using Ta foil clamps placed outside the measurement region. Due to the difference in work function between the SiC substrate and the clamps, a local electrostatic field is expected to form near the contact regions. This electric field can slightly deflect the trajectories of photoelectrons, leading to deviations in the detected emission angles.

As a result, ARPES spectra collected from these regions exhibit systematic angular shifts relative to those from unaffected areas. Consistently, the shifted spectral components are localized near the substrate edges where the clamps are positioned, and W2, which exhibits a larger angular shift than W6, is distributed closer to the clamp region. These observations indicate that the local angular shifts observed in this study originate from electric fields induced at the sample edges. The spatial extent of this effect is estimated to be approximately 1 mm from both edges of the sample.

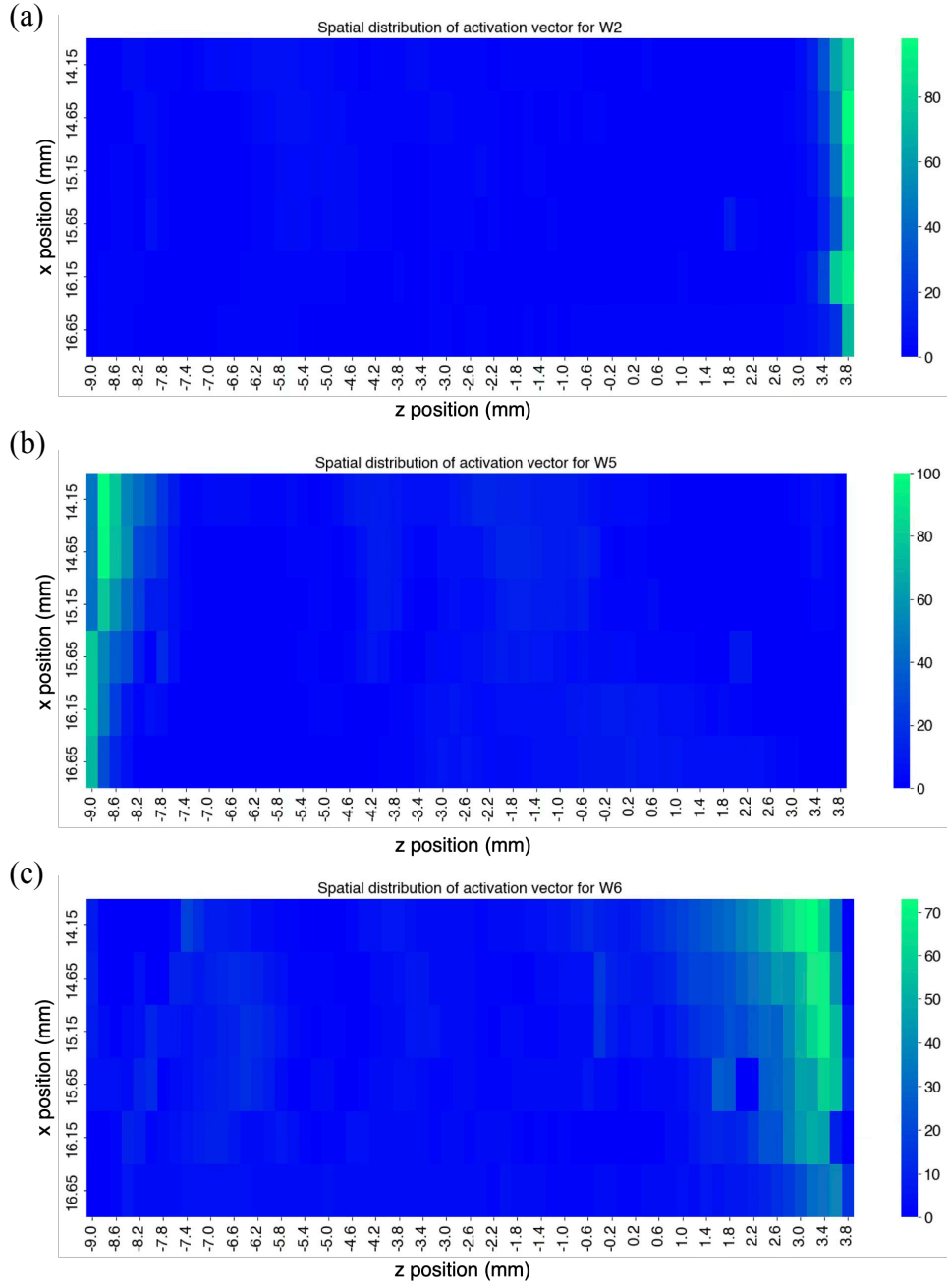


Figure. S4. Heatmaps of the activation vector intensities for basis vectors (a) W2, (b) W5, and (c) W6, respectively.

S.5 ARPES measurements deviated from K point

Fig. S5 shows ARPES spectra of TL graphene measured at a photon energy of 41 eV, with θ_y varied systematically in the vicinity of the K point. Fig. S5 (a) shows the constant-energy map at the Fermi level; the dashed line is drawn at the angular position corresponding to the K point ($\theta_y = 32.5^\circ$). Fig. S5(b) shows ARPES images acquired at different θ_y values, with the measurement angles indicated in the figure. As θ_y deviates from the K point, the spectral intensity changes systematically across the Dirac point: on the low-angle side, the spectral weight at higher binding energy increases and the apparent intensity maximum shifts, while on the high-angle side, the spectral weight at lower binding energy increases with a corresponding shift of the intensity maximum. This behavior indicates that the spectral intensity distribution evolves systematically with angle, while the energy position of the Dirac point itself remains unchanged. The apparent spectral variations therefore arise from intensity redistribution rather than from an intrinsic shift of the Dirac point energy. Accordingly, the spectral features observed in the main text, which may superficially resemble electronic-structure modifications, are more consistently attributed to geometric tilt effects associated with step bunching rather than to intrinsic electronic changes.

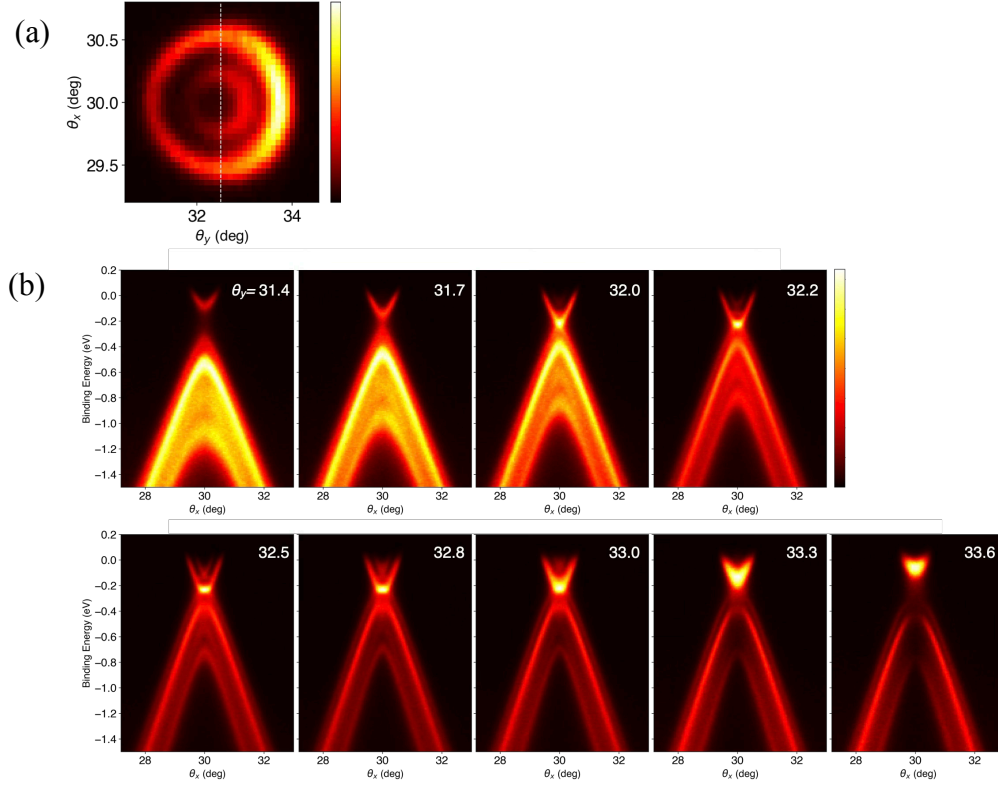


Figure S5. ARPES spectra of TL-graphene near the K point measured with the excitation energy of 41 eV. (a) Constant energy map at the Fermi level; the dashed line marks $\theta_y = 32.5^\circ$ (K point). (b) ARPES images measured at various θ_y . As θ_y deviates from the K point, the spectral intensity shifts across the Dirac point, accompanied by a change in the position of maximum intensity.