PCCP

ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

Electronic Supplementary Information (ESI) – Surface phase diagrams from nested sampling†

Mingrui Yang,[a] Livia B. Pártay,[b] and Robert B. Wexler*[a]

S1 Finite-size effects

We first examine how the calculated thermodynamic properties change on a smaller 2×2 per monolayer (ML) unit cell, where the finite-size effects should be more pronounced. Additionally, a smaller unit cell allows us to increase the number of free particles in the system to form more than one ML. The results are presented in Figures S1a and S1b.

One can see that the finite-size effect is significant. The lower-temperature, entropy-induced peaks in the heat capacity curves are not found (besides the one- and two-free-particle cases, $\theta = 1/4$ and $2/4$ ML), which is intuitive, as the number of configurations on a 2×2 surface is significantly lower than the one with 4×4 surface particles. Nevertheless, the smaller system allows us to examine the phase transitions with surface coverages beyond one ML, where the transition temperatures exhibit weak coverage dependency.

To verify the accuracy of our 4×4 surface results and to ensure minimal finite-size effects, we performed nested sampling (NS) calculations using a larger 6×6 particles-per-ML surface unit cell. These calculations are computationally demanding due to the inclusion of more free particles for equivalent fractional coverages and a significantly larger sampling volume. For efficient phase-space sampling, we utilized 256 walkers per free particle. Our calculations converged to a fractional coverage of $\theta = 31/36$ ML, necessitating roughly three million NS iterations to reach the lowest potential energy state.

The heat capacity curves derived from these calculations can be found in Figure S1c, with the corresponding phase diagram in Figure S1d. The results for the 4×4 surface align closely with those for the 6×6 surface, both qualitatively and quantitatively. They produce similar phase diagrams with consistent transition temperatures and a triple point. Additionally, the surface structures observed on the 4×4 and 6×6 surfaces at equivalent fractional coverages verify the adequacy of the 4×4 surface size. In conclusion, the more computationally modest 4×4 surface effectively captures the relevant physics. This computational efficiency enables us to explore a broader spectrum of surfaces beyond Lennard-Jones(111) [LJ(111)], making the 4×4 surface our system size of choice.

S2 Additional details on surface system setup

In this section, we compare the computational outputs from simulations with different system setups, illustrating the importance of having the appropriate configuration for the sampled system.

To adequately capture the interplay between free particles and the host surface, we must ensure that the LJ interactions from the surface cover the entire volume of the available space. Otherwise, particles may move into regions with no interactions during sampling, leading to either no potential energy contribution or inaccurate contributions from particle-particle interactions rather than particle-surface interactions. Figure S2 depicts several scenarios of system setups that lead to different final results.

We have conducted tests on how the presence of "voids" and the placement of initial walkers influence the NS process and the calculated heat capacities. Three tests were carried out on a 2×2 LJ(111) surface with three free particles. The results are summarized in Figure S3. The first set of tests used the "double volume" setup, as shown in Figure S2(a). The second set of tests utilized the "reduced volume" setup, as displayed in Figure S2(b). The third set of tests used the setup shown in Figure S2(a) but placed all initial walkers within the LJ cutoff range (labeled as "low starting positions" in Figure S3), where the overall volume is the same and all free particles can initially interact with the surface. However, this setup does not have a uniform distribution of initial walkers in the configuration space. We observed that the overall volume determines the height of the peak in the C_V curves and has minimal influences on the transition temperature. This volume dependence is unsurprising as the difference in volume leads to changes in other conditions, *e.g.* pressure, thus altering the C_V. Interestingly, the "low starting positions" setup recovers the same C_V, albeit with significant variance, even though a non-uniform distribution of initial walkers is used. We consider the setup used to produce the orange curves in Figure 1 appropriate for this work.

Furthermore, we encountered a technical challenge when running NS for any surfaces with a single free particle using the "double volume" setup, as shown in Figure S2(a). The sampling can stall when the free particle is placed in the "void" region, either

[a] *Department of Chemistry and Institute of Materials Science and Engineering, Washington University in St. Louis, St. Louis, MO 63130, USA. E-mail: wexler@wustl.edu*
[b] *Department of Chemistry, University of Warwick, Coventry, CV4 7AL, UK.*

† Electronic Supplementary Information (ESI) available: finite-size analysis, system setup procedures, specifications of surface structure, calculations of order parameters, and illustrations of maximum-probability surface structures. See DOI: https://doi.org/10.1039/cXCP00000x

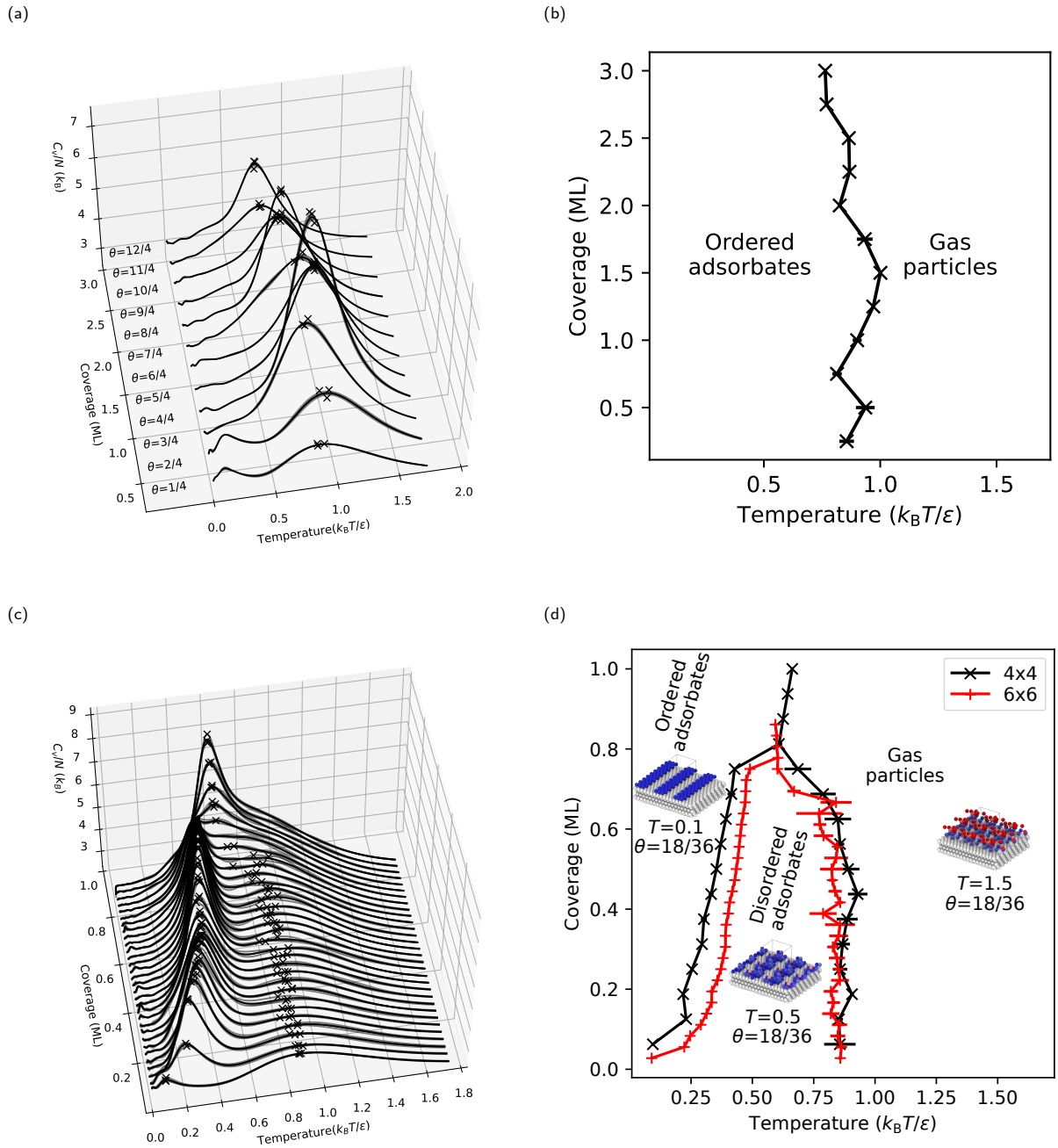

Fig. S1 Coverage-temperature properties for the flat Lennard-Jones(111) surface are presented as follows: (a) heat capacity per free particle on a 2×2 surface with fractional coverages ranging from $\theta = 1/4$ ML to three MLs; (b) phase diagram for the same 2×2 surface; (c) heat capacity per free particle on an expanded 6×6 surface with fractional coverages ranging from $\theta = 1/36$ ML to $\theta = 31/36$ ML; and (d) the phase diagram for this 6×6 surface (in red), with insets showing the maximum-probability structures of each phase at selected temperatures for $\theta = 18/36$ ML, juxtaposed with the phase diagram for the 4×4 surface as shown in Figure 4b (in black). The 6×6 unit cell is repeated three times in x- and y-directions for better visibility of the surface structures. The lines between the crosses are only guides for the eye.

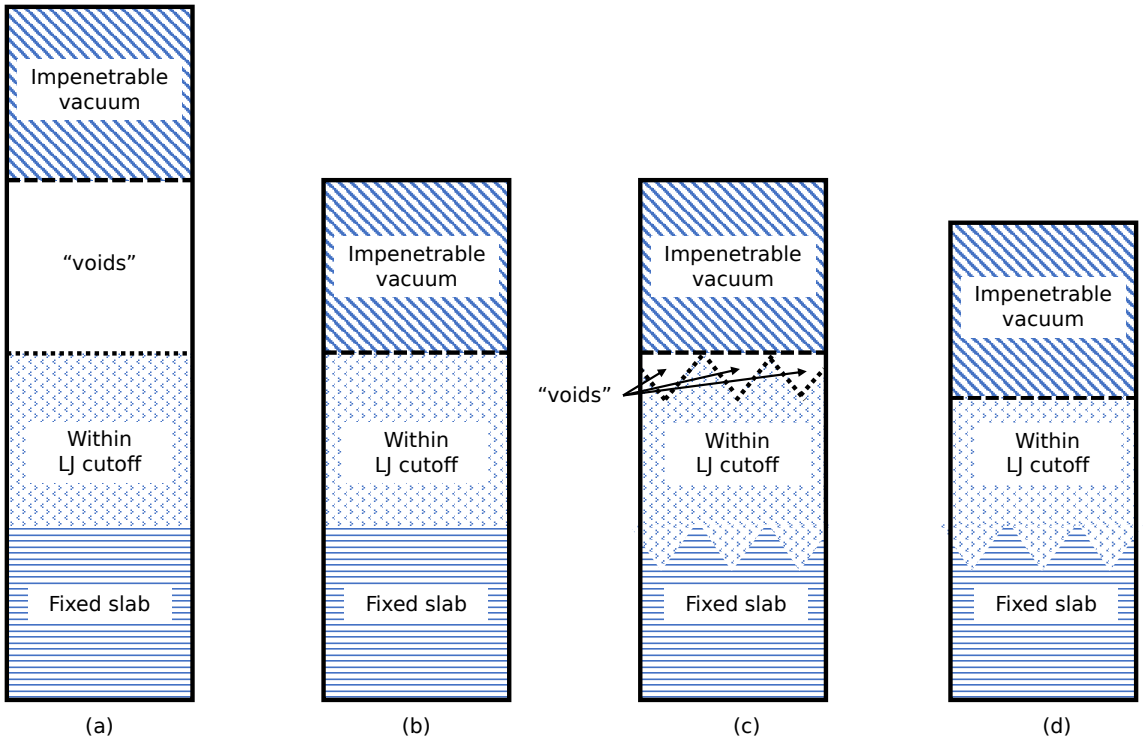

Fig. S2 Different ways to construct the fixed surface-free particles model: The key is to ensure that all free particles can interact with the surface. In setup (a), there is a large volume outside the LJ cutoff, labeled as "voids", where the free particles cannot interact with the fixed surface. Such a setup is considered problematic in this work, as we may accidentally sample the clustering of particles in the "voids". Setup (b) is appropriate for flat surfaces. The "voids" are removed by reducing the overall height of the cell. Such a setup is used for LJ(111) and LJ(100) surfaces. However, even with reduced cell height, significant "voids" can still appear, as shown in (c), with stepped surfaces such as LJ(311) and LJ(110). We further reduced the cell's height to correct it, as shown in (d).

by a Monte Carlo (MC) move or generated as an initial walker. As the "void" region is rather large and there are no other particles with which to interact, generating a new configuration with lower potential energy using random MC moves is challenging due to the small step size relative to the cell's dimensions. Currently, `pymatnest` updates the MC step size to maintain an acceptance ratio of 0.25-0.75. However, when particles enter these voids during MC steps, the potential energy does not change. This condition decreases the acceptance ratio because we only accept the MC step if the potential energy remains below $E_i^{\max}$. As a result, adjusting the MC step size to infinitesimal values becomes necessary to restore the acceptance ratio to 0.25 and 0.75. NS with molecular dynamics can proceed with the "double volume" setup (green curves in Figure S4) via NVE propagation. As discussed above, the phase transitions, especially the higher-temperature condensation processes, change significantly with additional volume. To produce consistent C_V with only MC propagation, we prefer eliminating any additional vacuum beyond the LJ cutoff range.

S3 Surface structures

In Table S1, we report the parameters of the four facets of the LJ solid. The cell dimensions refer to the dimensions of the simulation cell used in the NS calculations. The vacuum thickness is the distance between the topmost fixed layers and the reflective wall (at 4σ below the top of the simulation cell), and the thickness is less than or equal to the LJ potential cutoff. The slab thickness is the distance between the topmost fixed layers and the bottom of the cell. For each facet, there are different numbers of layers depending on the interlayer spacing. The trough spacing is the distance between the centers of two neighboring troughs, applicable for the (110) and (311) facets.

S4 Phase diagrams for flat LJ(100) and stepped LJ(311) surfaces

S4.1 Flat LJ(100) surface

As the LJ(100) and LJ(111) surfaces share similarities in terms of their surface features and phase behaviors, we only present the LJ(111) results in the main text. Most of the phase behaviors of the LJ(100) surface can be understood following the discussions of the LJ(111) surface. Here, we only highlight specific differences between the two surfaces. Compared to LJ(111), LJ(100) is also considered a flat surface but with reduced surface symmetry (four-fold *vs.* six-fold). This reduction leads to broader and lower low-temperature peaks in the C_V curves (Figure 5a) for coverages $\theta \leq 10/16$ ML. For $\theta > 10/16$ ML, the low-T peaks disappear. We manually determine a "shoulder" peak for each C_V curve with $\theta > 10/16$ ML and mark them with ♦ in Figure 5d. One

Facet	Cell dimensions	Vacuum thickness	Slab thickness	Number of layers	Interlayer spacing	Trough spacing
(111)	$4.49 \times 3.89 \times 11.66$	4.00	3.66	5	0.92	–
(110)	$6.34 \times 4.49 \times 11.37$	3.44	3.93	8	0.56	1.60
(100)	$4.49 \times 4.49 \times 11.97$	4.00	3.97	6	0.79	–
(311)	$4.49 \times 7.44 \times 11.35$	3.52	3.83	9	0.48	1.86

Table S1 The parameters of the four facets of the LJ solid are all in units of σ. They include cell dimensions, represented by $x \times y \times z$; the vacuum thickness, which is the distance between the topmost fixed layers and the reflective wall; the slab thickness, defined as the distance between the topmost fixed layers and the bottom of the cell; the number of layers within the fixed slab; the interlayer spacing, which is the distance between two layers in the fixed slab; and the trough spacing, the distance between the centers of two neighboring troughs.

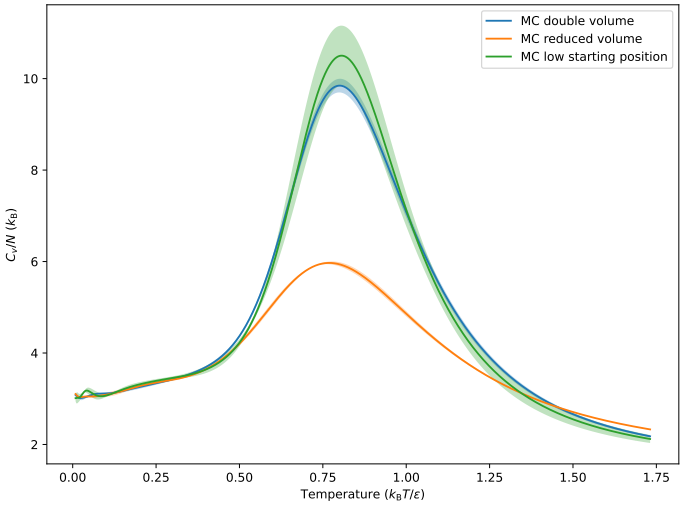

Fig. S3 Heat capacity of a 2×2 LJ(111) surface with three free particles. The "double volume" (blue) curves were produced from NS using the setup shown in Figure S2(a). The "reduced volume" curves were produced using the setup shown in Figure S2(b). The "low starting positions" (green) curves were generated by utilizing the setup in Figure S2(a), but all initial walkers were placed within the LJ cutoff range. Although the overall volume remains the same in this setup, all free particles can interact initially with the surface. Only the "reduced volume" (orange) curves are considered for this study.

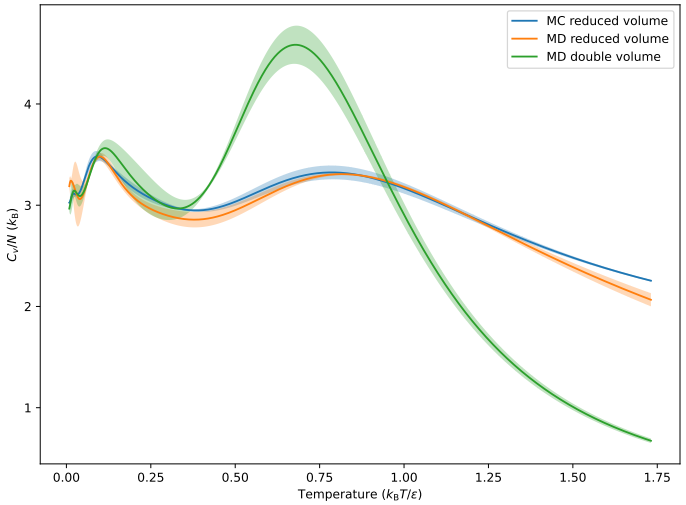

Fig. S4 Heat capacity of a 2×2 LJ(111) surface with a single free particle. "MD" in the legend indicates that the free particle is propagated using molecular dynamics (MD) in the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS).[?] The NS runs with the "double volume" setup, as shown in Figure S2(a), get stuck whenever the single particle is in a "void." The "reduced volume" curves, shown in blue, are from the setup as shown in Figure S2(b). These are the correct results, as verified by the MD (represented in orange) results, using LAMMPS with an identical system setup.

can see that heat capacity peaks for the entropy-driven ordering phase transition are being diminished as coverage increases beyond the half-ML ($\theta = 8/16$ ML), where the coordination number is being maximized. Note that the maximum coordination number for surface particles in a (100) ML is eight, with four from the slab and four from neighboring particles. Significantly, the triple point is not present in the LJ(100) phase diagram (Figure 5d), which is present in the LJ(111) phase diagram. This absence indicates that, on LJ(111), for high surface coverages, the entropic and enthalpic effects compete, whereas, on LJ(100), the entropy-driven (lower temperature) ordering phase transition disappears while the enthalpy-driven (higher temperature) condensation phase transition remains unaffected.

S4.2 Stepped LJ(311) surface

The stepped surface with a higher Miller index has surface features such as "troughs" along one of the lateral directions, offering several different binding sites compared to the flat LJ(111) facet. The potential energy landscape of LJ(311) is more complex than LJ(111), and its adsorbate phase diagram is more difficult to determine. We compute the $C_V(T)$ for the LJ(311) surface with the same coverages as those for the flat LJ(111) surface. The $C_V(T)$ curves are shown in Figure 5b.

Interpreting the $C_V(T)$ curves for the LJ(311) surface is more challenging than those for the LJ(111) surface. First, the LJ(311) $C_V(T)$ curves from the NS runs show larger deviations near the peaks than those for LJ(111). Moreover, the peaks are generally broader and significantly overlap, making them less precise. For example, at $\theta = 6/16$ and $7/16$, the C_V curves have a plateau-like feature from $T = 0.5 - 0.8\ k_BT/\varepsilon$. Despite these difficulties, we utilized the same automated procedure to find the peaks. We noted that for each of the lower coverages ($\theta \leq 10$), there are two separate peaks: one at a higher temperature between 0.7 and $0.9\ k_BT/\varepsilon$ (see $\times$s), and another peak at a lower temperature between 0.2 and $0.6\ k_BT/\varepsilon$. Compared to the C_V curves of LJ(111), LJ(311)'s low-coverage, high-temperature peaks are more dominant.

For higher coverages ($\theta \geq 11/16$), the peaks on the C_V curves have merged into one, similar to the behavior observed on LJ(111) at higher coverages. The merged peaks are typically lo-

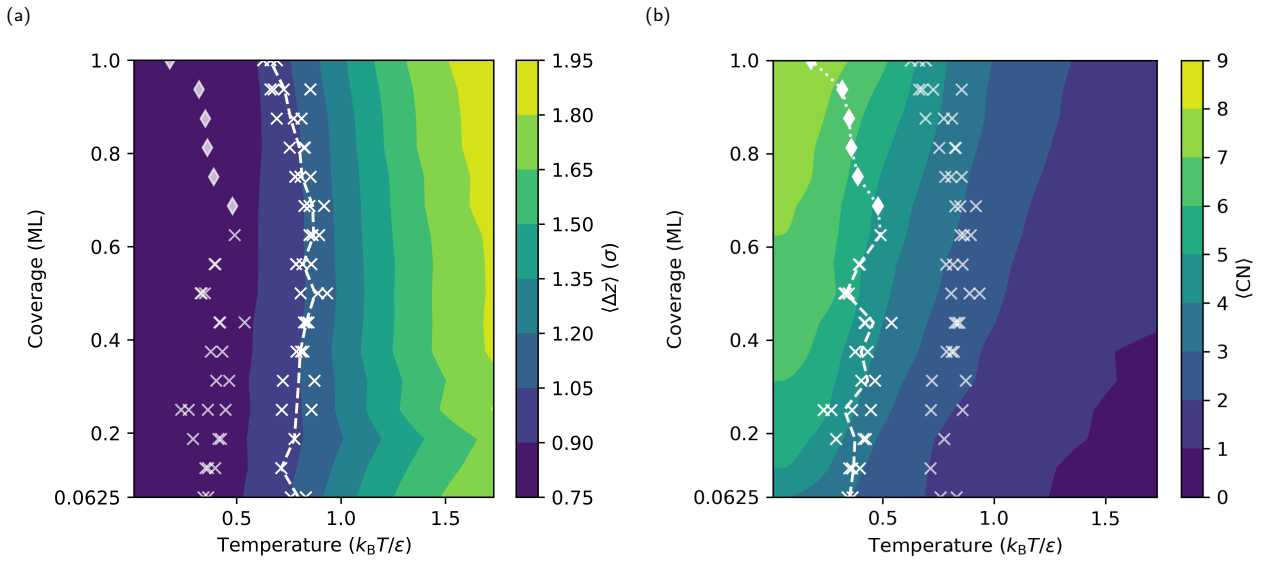

Fig. S5 Calculated coverage-temperature properties of the flat LJ(100) surface with fractional coverages from $\theta = 1/16$ ML up to one ML. (a) The average z-coordinates of the free particles are shown relative to the topmost layer in the fixed slab. Note that the bulk (100) interlayer spacing is 0.79σ. (b) The free particles' average coordination numbers, including particle-particle and particle-surface bonding, are indicated. The lines between the crosses are just guides for the eye. The diamond markers (♦) point to the disappearing peaks not found by the automated procedure.

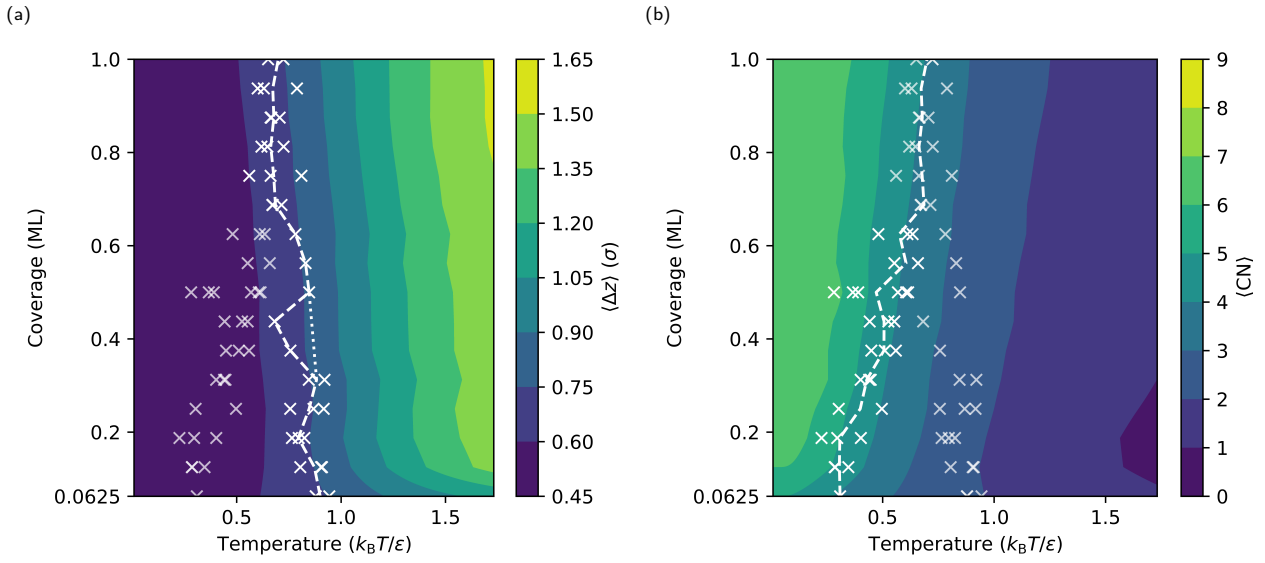

Fig. S6 Calculated coverage-temperature properties of the stepped LJ(311) surface with fractional coverages from $\theta = 1/16$ ML up to one ML. (a) The average z-coordinates of the free particles are shown relative to the topmost layer in the fixed slab. Note that the bulk (311) interlayer spacing is 0.48σ. (b) The free particles' average coordination numbers, including particle-particle and particle-surface bonding, are indicated. The lines between the crosses are just guides for the eye.

cated at a temperature between 0.6 and 0.8 $k_B T/\varepsilon$ and are often broader and lower than those for LJ(111).

We further project the temperatures of all peaks onto the coverage-temperature plot to construct the phase diagram for LJ(311), as shown in Figure 5e. Due to the more significant deviations of the peak temperatures from the C_V curves, the phase boundaries are less clearly defined than those for LJ(111). That said, the overall features of the LJ(311) phase diagram are still qualitatively similar to LJ(111): two phase transitions for coverages $\theta \leq 10$, a triple point located at $T \approx 0.7\ k_B T/\varepsilon$ and $\theta = 11/16$, and a single phase transition for coverages $\theta \geq 12/16$. Note that, due to the flatness of the C_V curves at $\theta = 6/16$ and $7/16$, the peaks found by the `scipy.signal.find_peaks()` function may not be reliable. Therefore, we join high-temperature peaks at $\theta = 5/16$ and $8/16$, marked by the dotted line in Figure 5e, to show a more likely phase boundary.

In order to further understand the phase transitions on the LJ(311) surface, we also computed surface order parameters, including $\langle \Delta z \rangle$ (Figure S6a) and $\langle \mathrm{CN} \rangle$ (Figure S6b). We can utilize the results from LJ(111) and LJ(110) to infer the phase transitions on the LJ(311) surface. LJ(311) can be understood as an intermediate surface between the completely flat LJ(111) and the stepped LJ(110) surfaces. The troughs on LJ(311) are shallower than LJ(110), and the inter-trough spacing is larger (see Table S1), hence it is more "flat."

S5 Stepped LJ(110) surface

We compute $C_V\,(T^*)$ for the stepped LJ(110) surface with the same coverages as those for the flat LJ(111) surface. We refer to the LJ(110) surface as stepped because it has "troughs" along one of the lateral directions. The C_V curves in Figure 5c show a dominant and almost coverage-independent peak (using the same automated procedure as that used in Section 3.1) at temperatures between $0.7\ k_B T/\varepsilon$ and $0.9\ k_B T/\varepsilon$ (see $\times$s). Note that a side peak can be observed at certain lower coverages (*e.g.*, $\theta = 2/16$ ML, $\theta = 4/16$ ML, and $\theta = 8/16$ ML), but no clear trend emerges. For the coverages without a visible side peak, such peaks may be masked by the dominant peaks. We do not yet have a reliable tool to resolve them, and we wish to develop a technique for masked peak detection in the future. The dominant and almost coverage-independent peak corresponds to an adsorbate phase transition, with a coverage-averaged phase transition temperature, $\overline{T}_\theta$, of $0.80(3)\ k_B T/\varepsilon$, as indicated by the dashed line in Figure 5f.

We calculated $\langle \Delta z \rangle$ to characterize the adsorbate phases and their transitions. Figure S7a shows that, at the highest temperature considered, *i.e.*, $T \approx 1.73\ k_B T/\varepsilon$, $\langle \Delta z \rangle = 1.65\sigma$ is nearly equal to the vertical center of **Region 2** in the simulation cell, *i.e.*, 1.72σ (See Table S1 in the ESI). At $\overline{T}_\theta$, $0.75\sigma < \langle \Delta z \rangle < 0.90\sigma$. Note that in a perfect LJ(110) bulk, the $\Delta z_{\mathrm{bulk}} = 0.84\sigma$ (the average of the trough site at 0.56σ and the atop site at 1.12σ). This correspondence between $\langle \Delta z \rangle$ and Δz_{bulk} suggests that the free particles "condense" on the stepped LJ(110) about one interlayer spacing above the surface, which we also observed on the LJ(111) in the previous section.

However, the nature of surface condensation on LJ(110) differs from that on LJ(111), as seen from the widths of their C_V peaks. To further characterize surface condensation, we calculate $\langle \mathrm{CN} \rangle$ (see Figure S7b), which is approximately three at $\overline{T}_\theta$. The average is derived from two scenarios: two free-fixed bonds when an adsorbed particle touches only one side of the trough and four when the particle touches both sides but has not fully settled into position. The coordination is at least fivefold when the particles are completely situated within the trough sites because they bind the fixed surface particles from two distinct layers, forming a square pyramidal structure. At near-zero temperature, $\langle \mathrm{CN} \rangle$ is five for $\theta = 1/16$ ML (*i.e.*, five free-fixed bonds), six for $\theta = 2/16$ ML (*i.e.*, one free-free and five free-fixed bonds), and seven for $\theta \geq 3/16$ ML (*i.e.*, two free-free and five free-fixed bonds). Figure S7b also indicates that the free particles prefer to occupy the same trough if it has at least one unoccupied site because the maximum $\langle \mathrm{CN} \rangle$ for $\theta \geq 3/16$ ML is seven, which is not possible if the free particles occupy different troughs.

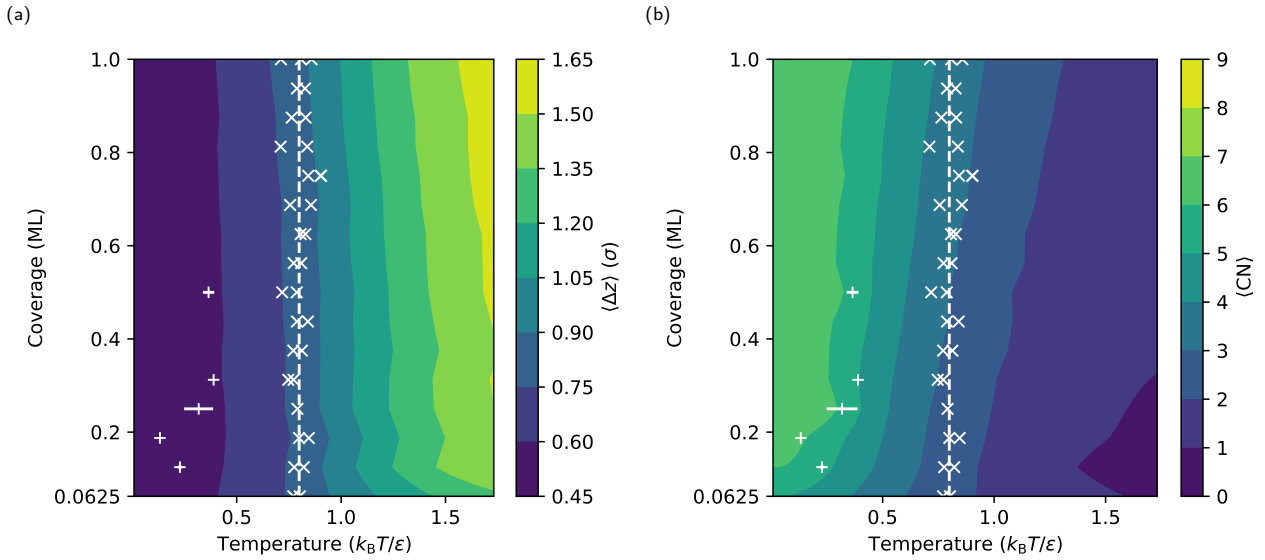

Fig. S7 Calculated coverage-temperature properties of the stepped LJ(110) surface with fractional coverages from $\theta = 1/16$ ML up to one ML. (a) The average z-coordinates of the free particles are shown relative to the topmost layer in the fixed slab. Note that the bulk (110) interlayer spacing is 0.56σ. (b) The free particles' average coordination numbers, including particle-particle and particle-surface bonding, are indicated. The dashed line at $\overline{T}_\theta = 0.80\ k_B T/\varepsilon$ shows the coverage-averaged phase transition temperature.

S6 Maximum-probability structures

We include top and side views of all four different facets on 4×4 surfaces at selected temperatures at all coverages as separate files:

- LJ(111) surface: `111-4x4-grid-top.png` and `111-4x4-grid-side.png`
- LJ(311) surface: `311-4x4-grid-top.png` and `311-4x4-grid-side.png`
- LJ(100) surface: `100-4x4-grid-top.png` and `100-4x4-grid-side.png`
- LJ(110) surface: `110-4x4-grid-top.png` and `110-4x4-grid-side.png`

We also include the top and side views of the (111) facet on the 6×6 at selected temperatures at coverages up to $\theta = 31/36$ ML as separate files:

- LJ(111) surface: `111-6x6-selected-grid-top.png` and `111-6x6-selected-grid-side.png`