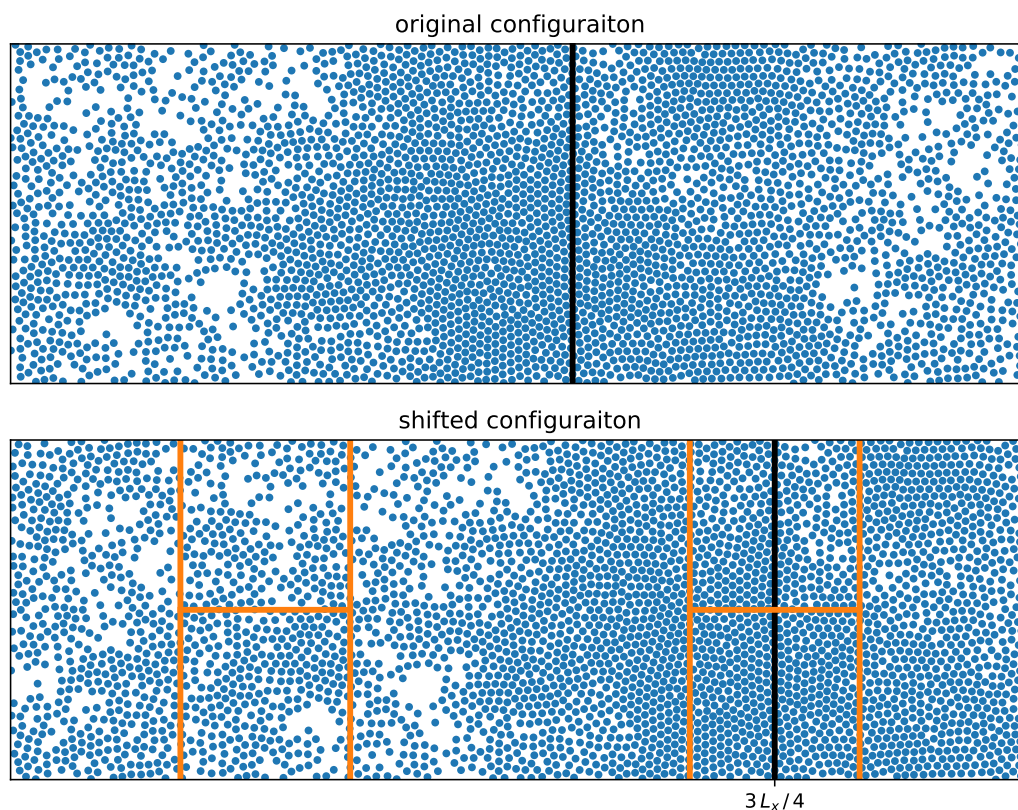


## Electronic Supplemental Information

Claudio Maggi,<sup>a,b†</sup> Matteo Paoluzzi,<sup>c</sup> Andrea Crisanti,<sup>b,d</sup> Emanuela Zaccarelli,<sup>d,b</sup> and Nicoletta Gnan<sup>d,b‡</sup>

### Center of mass shifting



**Fig. 1** (Top-panel) One original configuration for  $N = 3750$ ,  $\rho = 0.95$  and  $\tau = 16.5$ . The  $x$ -coordinate of the center of mass found is indicated by the black vertical line. (Bottom-panel) Shifter veriosn of configuration in top panel. Particles positions are shifted so that the  $x$ -coordinate of the center of mass coincides with  $3L_x/4$ . The analysis sub-boxes (orange lines) are centered onto  $3L_x/4$  (dense phase) and onto  $L_x/4$  (dilute phase).

We analyze the data following the same procedure of Ref.<sup>1</sup> and used also in Ref.<sup>2</sup>. For each configuration we first find the  $x$ -coordinate of the center of mass with periodic boundary conditions as shown in Fig. 1(top). We then shift particles coordinate so that the center of mass  $x$ -coordinate coincides with  $3L_x/4$  (see Fig. 1(bottom)). Finally we perform the analysis on the sub-boxes (orange lines) that are centered onto  $3L_x/4$  (dense phase) and onto  $L_x/4$  (dilute phase) as shown in Fig. 1(bottom).

### Length of the simulation runs

We show here that the length of the simulation runs is long enough to allow the full relaxation of the density correlation function  $C(t) = \langle \Delta N_b(0) \Delta N_b(t) \rangle / \langle \Delta N_b^2 \rangle$ , where  $\Delta N_b = N_b - \langle N_b \rangle$  is the fluctuation of the number of particles in the sub-box. More specifically we analyze separately the relaxation in the dense and dilute phases considering the two sub-boxes on the left and two on the right as shown

<sup>a</sup> NANOTEC-CNR, Institute of Nanotechnology, Soft and Living Matter Laboratory - Piazzale A. Moro 2, I-00185, Roma, Italy

<sup>b</sup> Dipartimento di Fisica, Università di Roma "Sapienza", I-00185, Roma, Italy

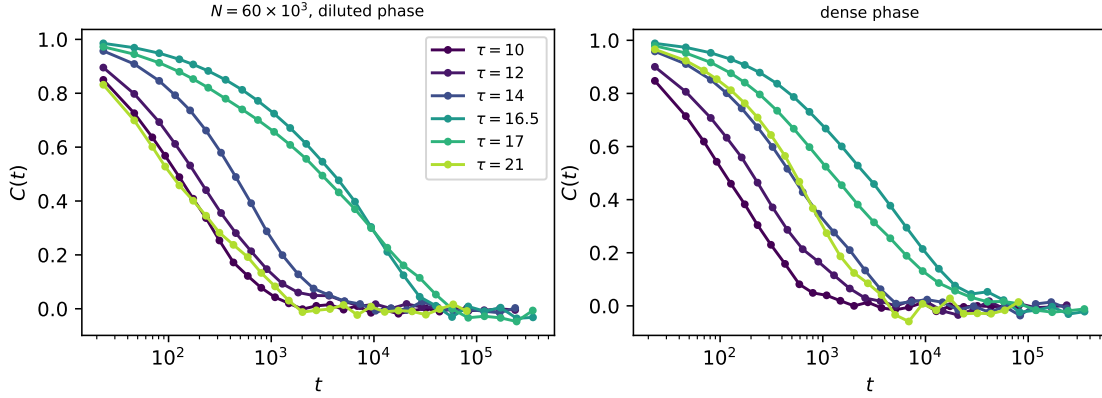
<sup>c</sup> Departament de Física de la Matèria Condensada, Universitat de Barcelona, C. Martí Franquès 1, 08028 Barcelona, Spain

<sup>d</sup> CNR-ISC, Institute of Complex Systems, Roma, Italy

<sup>†</sup> E-mail: claudio.maggi@cnr.it

<sup>‡</sup> E-mail: nicoletta.gnan@cnr.it

in Fig 1 of the main text. Figure 2(a) and (b) display  $C(t)$  for the largest system investigate (i.e  $N = 60 \times 10^3$ ) at different  $\tau$  values both in the dense and dilute phase: in both cases  $C(t)$  decay to zero at all  $\tau$ .



**Fig. 2** Density auto-correlation function for the system with  $N = 60 \times 10^3$  particles at  $\rho = 0.95$  and different values of  $\tau$ . Left panel: dilute phase. Right panel: dense phase.

## Averages and error estimation

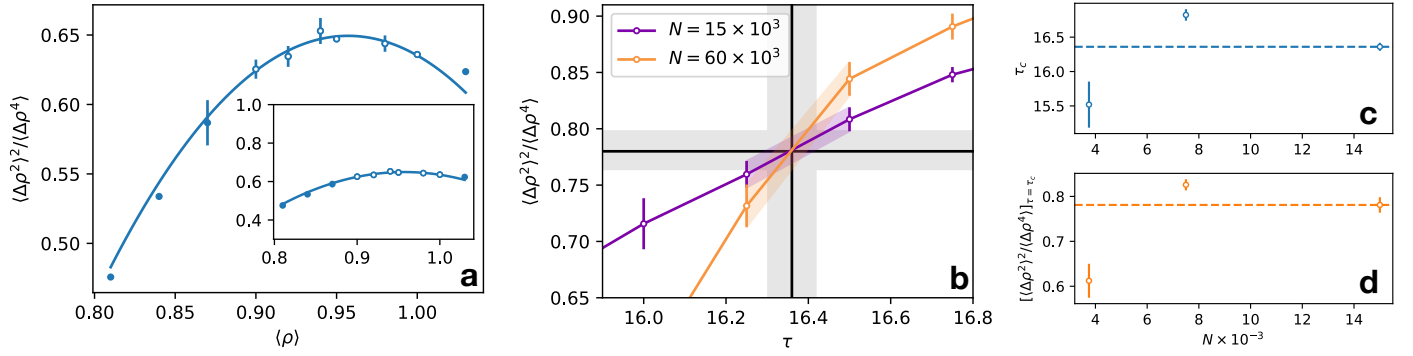
All quantities appearing in the main text (i.e. the Binder cumulant, the susceptibility, the order parameter and the kinetic temperature difference between the dense and dilute phase) are averaged over 16000 to 36000 configurations and over all sub-boxes. Individual configurations are taken at time intervals of duration  $\approx 23$  in reduced units, which is approximately equal the largest  $\tau$  explored. Moreover for the values of  $\tau$  close to  $\tau_c$  these quantities of interest are averaged over multiple initial random configurations (up to 12 for the largest system sizes). To estimate the error on these quantities we proceed as follows: we first divide each run in time windows larger than the relaxation time of the density correlation function  $C(t)$  (Fig. 2) we then compute the observable in each time-window and compute the standard error of the mean over all time windows. In Fig. 2 of the main text we report the error as twice the standard error of the mean. We have also checked that by computing the average of the Binder cumulant over these time windows (instead that on all configurations) we get almost identical results than those reported in Fig. 2 of the main text.

## Estimate of the critical $\rho$ and $\tau$

We roughly estimate the critical density by performing a density scan, at fixed  $\tau$ , in the proximity of the critical point for the smallest system investigated (i.e.  $N = 3750$ ). According to Ref.<sup>3</sup> the cumulant should exhibit a maximum at  $\rho = \rho_c$  when plotted as a function of  $\rho$  and at fixed  $\tau = \tau_c$ . Fig. 3(a) shows that the Binder parameter indeed displays a maximum if we fix  $\tau = 16$ . This value has been chosen based on preliminary simulations and it is close to the critical value  $\tau_c = 16.36$  estimated in the following. Note also that the cumulant varies much less upon changing density in this small interval than upon changing  $\tau$  on a large interval as in Fig. 2(a) of the main text. This is shown in the inset of Fig. 3(a) plotting the same data of the main panel in the y-range (0.3; 1), i.e. the range of the Binder cumulant as a function of  $\tau$ . To extract  $\rho_c$  we have fitted the six closest points to the maximum in Fig. 3(a) with a 2nd order polynomial finding  $\rho_c = 0.953(0.037)$  where the fit error is reported in brackets. We set the value  $\rho_c = 0.95$  for all the sizes discussed in the main text neglecting the dependence of  $\rho_c$  on the size of the system.

The critical value  $\tau$ , indicated by  $\tau_c$ , has been determined by finding the crossing of the cumulants for two sizes  $N = 15 \times 10^3$  and  $N = 60 \times 10^3$ . These have been chosen because  $N = 60 \times 10^3$  is the largest size simulated and  $N = 15 \times 10^3$  has a linear size which is two times smaller than the largest one. The procedure followed to find the intersection is illustrated in Fig. 3(b). We linearly interpolate the cumulant curves and find  $\tau_c$  as the x-intersection of the two lines, while the critical cumulant value  $\mathcal{B} = [\langle \Delta \rho^2 \rangle^2 / \langle \Delta \rho^4 \rangle]_{\tau=\tau_c}$  is found from the intersection on the y-axis. We obtain  $\tau_c = 16.361(0.058)$  and  $\mathcal{B} = 0.781(0.017)$ . This  $\mathcal{B}$  value is lower than the one found for the triangular lattice gas ( $\mathcal{B} = 0.8321(0.0023)$ , see section below) and for the square lattice gas ( $\mathcal{B} \approx 0.83$  extracted from Ref.<sup>1</sup>). However it is close to  $\mathcal{B} \approx 0.75$  which is the critical cumulant value of the active lattice model found in Ref.<sup>2</sup>. Note that previous studies on the Lennard-Jones fluid have also reported a lower value of the critical Binder parameter with respect to the Ising model<sup>4</sup>.

To further check the size dependence of  $\tau_c$  and  $\mathcal{B}$ , we have repeated this procedure also considering other system sizes, that are separated by a factor of 2 in linear size, i.e.: ( $N = 3750, N = 15 \times 10^3$ ) and ( $N = 7500, N = 30 \times 10^3$ ). The resulting  $\tau_c(N)$  and  $\mathcal{B}(N)$  are shown in Fig. 3(c) and (d) respectively where each value of  $\tau_c(N)$  and  $\mathcal{B}(N)$  is associated with the smallest size in the pair. In Fig. 3(c) and (d) we see that the values of both quantities at  $N = 7500$  are closer to those of the largest system size (dashed lines) than to those corresponding to  $N = 3750$ . This suggests that, upon increasing the size,  $\tau_c(N)$  and  $\mathcal{B}(N)$  progressively converge to the infinite-system



**Fig. 3** (a) Fourth order cumulant as a function of density. Data points represent the Binder parameter varying the density at fixed  $\tau = 16$  and  $N = 3750$ . The open symbols are the closest six points to the maximum used in the parabolic fit (full line) for determining  $\rho_c$ . The inset shows the same data of the main panel plotted on the same  $y$ -range of Fig. 2(a) of the main text. (b) Intersection of the cumulants (colored points) for  $N = 60 \times 10^3$  and  $N = 15 \times 10^3$  used for locating  $\tau_c$ . The intersection point  $\tau_c = 16.361(0.058)$  and  $B = 0.781(0.017)$  (black lines) is found by a piece-wise interpolation of the cumulant curves. The error on  $\tau_c$  and  $B$  (gray areas) is obtained by propagating the  $y$ -error on the points nearby the intersection (colored areas). (c) and (d) (same  $x$ -axis) show the obtained values of  $\tau_c$  and  $B$  (open symbols) as a function of the system size. The values of the reference point for the largest size is also reported as a dashed line.

critical values.

## Direct estimates of the critical exponents

Here we show how we directly estimate the critical exponents. Following Ref.<sup>1</sup> we first focus on the dependence of the slope of the critical Binder cumulant on  $L$  whose scaling with size is controlled by the exponent  $\nu$ :

$$\left[ \frac{\partial}{\partial \tau} \left( \frac{\langle \Delta \rho^2 \rangle^2}{\langle \Delta \rho^4 \rangle} \right) \right]_{\tau=\tau_c} \sim L^{1/\nu} \quad (1)$$

To use Eq. (1) we evaluate the derivative of the cumulant by fitting the numerical data with a generalized logistic function of the form

$$y(\tau) = A_1 + \frac{A_2}{[A_3 + A_4 e^{-(\tau-\tau_0)/w}]^{1/\theta}} \quad (2)$$

Where  $A_1, A_2, A_3, A_4, w, \tau_0$  and  $\theta$  are fitting parameters. As shown in Fig. 4(a) this function fits well the data especially around  $\tau_c$  and allows us to estimate the derivative (1). This derivative is reported in Fig. 4(b) as a function of the system size and it is indeed well fitted by  $\sim L^{1/\nu}$  with  $\nu = 1$ . In fact a direct fit with a power law gives  $1/\nu = 0.968(0.096)$  (the fit error is indicated in brackets) and, by linear error propagation,  $\nu = 1.03(0.10)$ . Contrarily the best fit with the exponent  $\nu = 1.5$  proposed in Ref.<sup>1</sup> deviates considerably from the data.

Next we consider the size dependence of the susceptibility  $\chi$  at  $\tau_c$  (i.e.  $\chi_{\tau=\tau_c}$ ), that should scale as  $\chi_{\tau=\tau_c} \sim L^{\gamma/\nu}$ . To do this we simply linearly interpolate the  $\chi$  at  $\tau_c$  at all sizes and report the results in Fig. 4(c). This quantity is also well fitted by the Ising exponent  $\gamma/\nu = 7/4 = 1.75$ . A direct fit with a power law yields  $\gamma/\nu = 1.787(0.090)$ . Using  $\nu$  found above and propagating also its error we get  $\gamma = 1.84(0.20)$ . Also in this case by fixing the values of  $\gamma = 2.2$  and  $\tau = 1.5$  (i.e.  $\gamma/\nu \approx 1.47$ ) from Ref.<sup>1</sup> we obtain a worse fit of our data.

To directly estimate  $\beta$  we interpolate the points of the order parameter  $(\rho_h - \rho_l)$  at  $\tau_c$  for all sizes and we plot them as a function of  $L$  in Fig. 5(a). It is evident that these points are quite noisy and also that  $\beta/\nu$  is small. However we are still able to obtain an estimate of  $\beta$  compatible with the Ising value ( $\beta = 0.125$ ) when we fit these points with a power law  $(\rho_h - \rho_l) \sim L^{-\beta/\nu}$ . This yields  $\beta/\nu = 0.110(0.053)$  (full line in Fig. 5(a)) and  $\beta = 0.113(0.055)$  if the error on  $\nu$  is propagated linearly. Note that this is close to the Ising value as shown by the orange line in Fig. 5(a) and appreciably smaller than the  $\beta = 0.45$  of Ref.<sup>1</sup> shown by the green line Fig. 5(a).

To obtain a more accurate estimate of  $\beta$  we also implement a finer method which finds the exponent by minimizing the deviation between collapsed data. This type of technique has been applied in the past to extract the critical exponents of various spin models<sup>5,6</sup>. To practically apply this method we fix the values of  $\tau_c$  and  $\nu$  to the values determined above ( $\tau_c = 16.36$  and  $\nu = 1.03$ ). We then consider the following error function to be minimized:

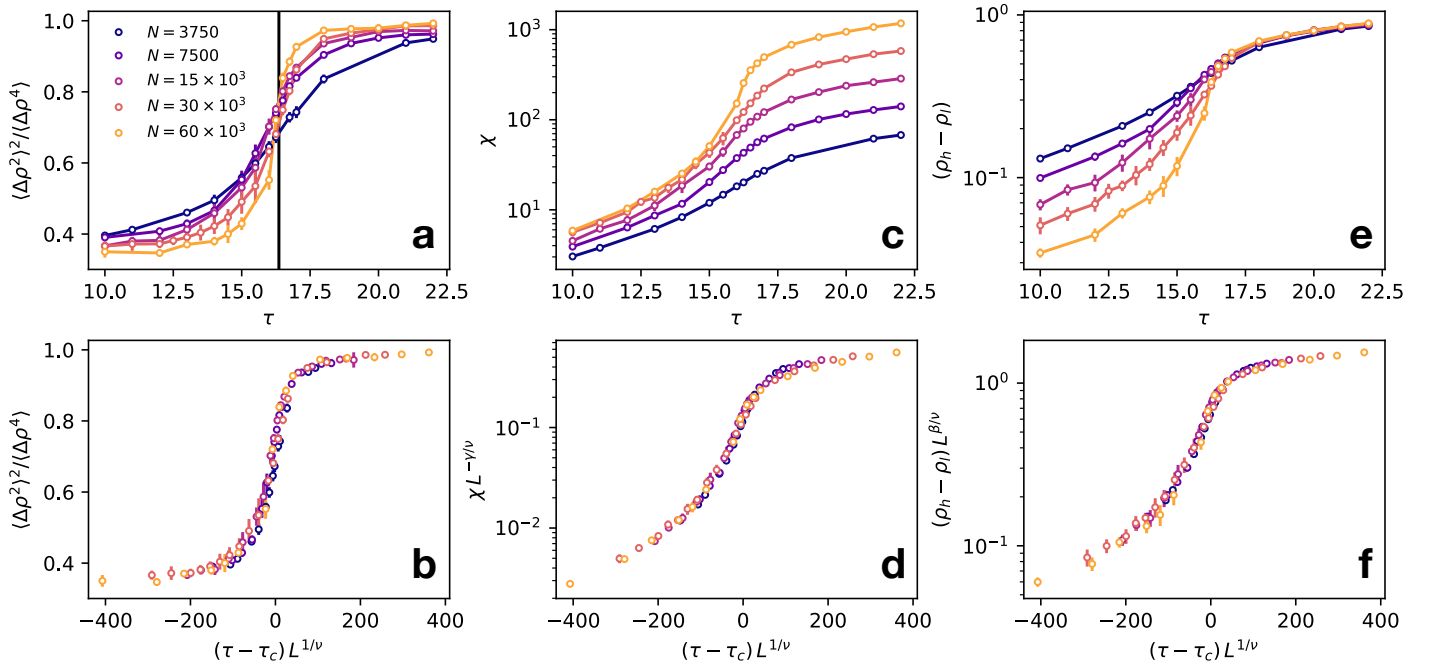
$$\mathcal{E} = \sum_i \left[ L_i^{\beta/\nu} m(\tilde{\tau}_i, L_i) - G(\tilde{\tau}_i) \right]^2 \quad (3)$$

where  $L_i$  and  $\tilde{\tau}_i = L_i^{1/\nu}(\tau_i - \tau_c)$  are respectively the system size and the scaled control parameter of the  $i$ -th data-point, while  $m(\tilde{\tau}_i, L_i) = [\rho_h - \rho_l]_{(\tau=\tilde{\tau}_i, L=L_i)}$  is the order parameter value at  $L_i$  and  $\tilde{\tau}_i$  (the sum runs over all available data). The function  $G$  in Eq. (3) is the scaling









**Fig. 8** Data collapse of the analyzed quantities with the exponents estimated directly  $\nu = 1.03$ ,  $\gamma = 1.84$  and  $\beta = 0.133$  (including the smallest system with size  $N = 3750$ ).

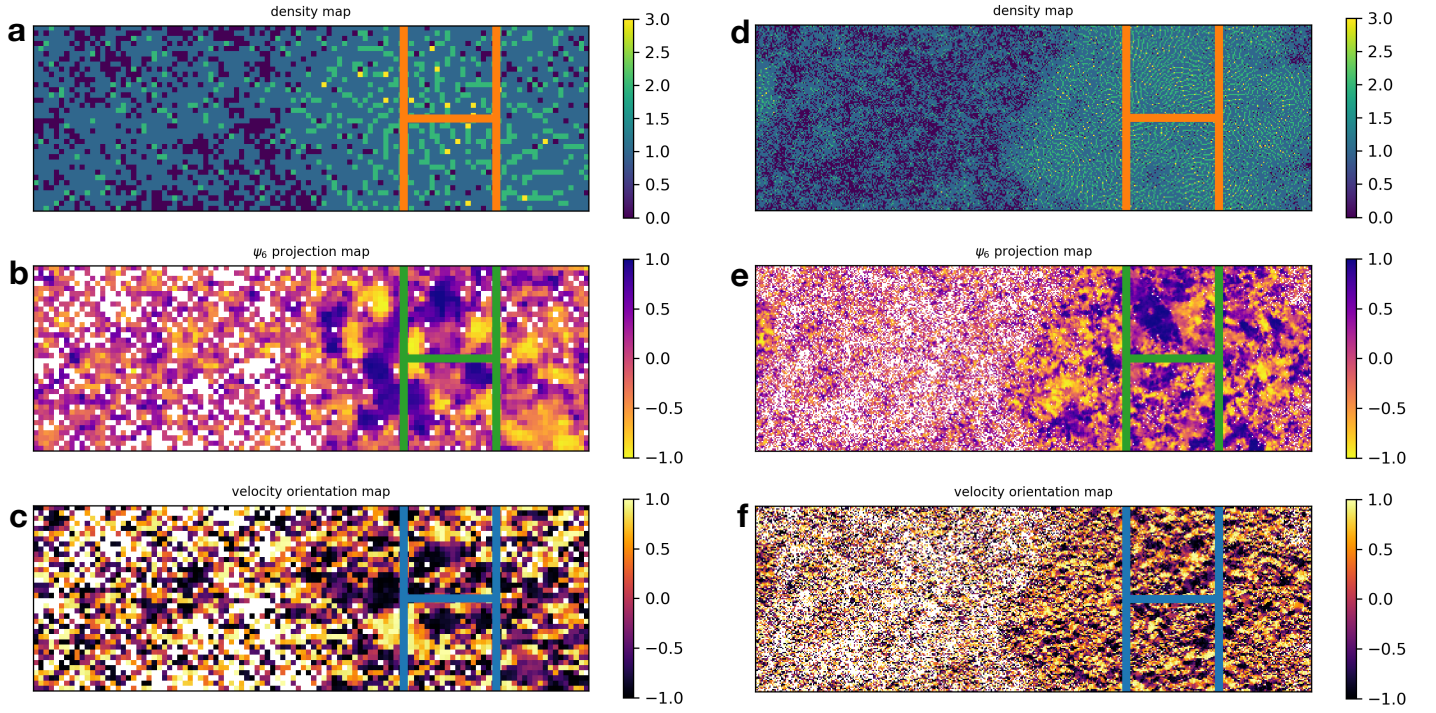
be mapped onto the Ising model with spin  $\sigma_i = \pm 1$  on the triangular lattice having critical temperature  $T_c = 4/\ln 3 \approx 3.641$  (for  $J = 1$  and  $k_B = 1$ )<sup>8</sup>. As a consequence, the  $T_c$  of the lattice gas model turns out to be  $T_c = 1/(\ln 3) \approx 0.91$ , i.e. an inverse critical temperature  $T_c^{-1} \approx 1.099$  while the critical average occupancy is  $n_c = 0.5$ .

The configurations of the lattice gas are analyzed as described in the main text for the active system. We start by shifting each configuration so that its center of mass is positioned at  $x = 3L_x/4$  as also shown in Fig. 7(a) and (b). Subsequently the quantities of interest are averaged over all four  $L \times L$  sub-boxes, where  $L = L_y/2$ . The density  $\rho$  in one sub-box is computed as  $\rho = \sum_i' n_i / L^2$ , where the prime indicates the sum runs only on those sites within the sub-box. Using this method we further check the correctness of the critical temperature by showing the Binder parameter and its good scaling with  $\nu = 1$  in Figs 7(c) and (d). By interpolating and averaging the values of the cumulants at the known value of  $T_c$  for all sizes we get  $\mathcal{B} = 0.8321(0.0023)$  which is close to the value of  $\mathcal{B}$  of the square lattice gas found in Ref.<sup>1</sup>. In the main text we have mentioned that the  $\chi$ , computed by averaging over all sub-boxes, does not show the typical peaked shape but rather forms a s-shaped curve when plotted as a function of the control parameter. This is the case also for the equilibrium lattice gas as shown in Fig. 7(e). In Fig. 7(f) we also show that this  $\chi$  scales well with  $\nu = 1$  and  $\gamma = 7/4$ . In the main text we have also used the average difference of the density in the high-density phase  $\rho_h$  and of the low-density phase  $\rho_l$  as an order parameter. To check if this quantity behaves as expected at criticality also in the equilibrium case we compute  $\rho_h$  and  $\rho_l$  as the average density of the two sub-boxes on the right and on the left respectively. The resulting  $(\rho_h - \rho_l)$  is shown in Fig. 7(b) as a function of  $T^{-1}$ . In Fig. 7(c) we show that we obtain a good data collapse by using the Ising exponents  $\beta = 1/8$  and  $\nu = 1$ . These data are clearly compatible with those presented for the off-lattice active system discussed in the main text, thus reinforcing the robustness of the analysis bringing to the Ising universality class in the case of the active system.

## System sizes, hexatic order and velocity correlation length

We discuss here the data collapse for the smallest system simulated, i.e.  $N = 3750$  (not included in the main text), which is comparable with the sizes used in a previous investigation on the critical behaviour of an off-lattice active system<sup>1</sup>. For this size, in Fig. 8 (a)-(c), we report the Binder cumulant, the susceptibility and the order parameter. It is evident that a reasonable data collapse with the exponents calculated above is found also for  $N = 3750$  (see Fig. 8(b),(d) and (f)). However the crossing point of the Binder cumulant for this size seems significantly lower in  $\mathcal{B}$  and  $\tau$  than the larger sizes (see also Fig. 3(c) and (d)). As mentioned in the main text we speculate that this could be due to the presence of another growing (but not diverging) correlation length. In the following we identify and compare two of them: the first related to the hexatic order and the second associated to velocity correlations.

A very recent work<sup>9</sup> has shown that the dense phase formed by active particles undergoing MIPS is made of a mosaic of hexatic micro-domains whose size does not diverge. To compare the size of these regions with our smallest system size, near the critical point, we consider the state point  $\tau = 16.5$ ,  $\rho = 0.95$  for  $N = 3750$ . In Fig. 9(a) we show a high-resolution density map of one configuration of this system (near  $\tau_c$ , already showing phase separation). This  $\rho$ -map is obtained by counting the number of particles in small squared



**Fig. 9** (a), (b) and (c) represent, respectively the maps of the density field, the  $\psi_6$  projection and the velocity direction projection for a typical configuration of sytem with  $N = 3750$ ,  $\tau = 16.5$  and  $\rho = 0.95$ . The map is calculated for a single configuration choosing bins of the order of the particles size. Different colors represent different values of the fields (see color-bars on the right). White pixels in (b) and (c) correspond to bins where no particles are found. The dense-phase sub-boxes (employed for the FSS) are drawn on (a), (b) and (c) to compare its size with the size of hexaitc and velocity-oriented domains. (d), (e) and (f) are the same of (a),(b) and (c) respectively but for a configuration of a large system ( $N = 60 \times 10^3$ ,  $\tau = 16.5$  and  $\rho = 0.95$ ).

bins of linear size  $s = 1$ . To characterize the hexatic order we calculate the parameter  $\psi_{6j} = N_j^{-1} \sum_k e^{i\theta_{jk}}$  for each particle. Here  $\theta_{jk}$  is orientation angle of the segment connecting the position of the  $j$ -th particle with its  $k$ -th (out of  $N_j$ ) nearest neighbors found with a Voronoi tessellation. To visualize the regions with the same orientation we project  $\psi_{6j}$  onto the direction of the mean orientation  $N^{-1} \sum_i \psi_{6i}$  where the sum runs over all particles in the system. In Fig. 9(b) we show the  $\psi_6$ -projection map obtained by averaging the  $\psi_6$ -projection of the particles found in each small bin (white pixels correspond to empty bins). In Fig. 9(b) it is evident that, in the dense phase, hexatic domains (i.e. regions with the same color) have an extent comparable to the size  $L$  of the FSS analysis boxes (we have  $L \approx 18$  for  $N = 3750$ ).

Recent works<sup>10,11</sup> have also shown that in active systems the colored noise induces an effective coupling between particles velocities. This effect gives rise to regions of densely packed particles with correlated speed and velocity orientation. We show here that, close to  $\tau_c$ , these regions have a size similar to the one of the hexatic regions. To visualize the extent of these velocity correlations we show in Fig. 9(c) the orientation map of particle velocities. This map is obtained by averaging the projected particle velocity vector on the  $x$ -axis, i.e.  $\cos(\vartheta_j)$  (where  $\vartheta_j$  is the orientation angle of the  $j$ -th particle velocity). Fig. 9(c) shows that the “islands” of velocity-correlated particles have a size comparable with the size of hexatic regions. Note however that when we consider a larger system ( $N = 60 \times 10^3$  and  $L \approx 72$ ) at the same  $\tau$  and  $\rho$  that is phase separating (Fig. 9(d)) the extension of these correlated hexatic and velocity regions does not scale up but remains approximately of the same size (see Fig. 9(e) and (f)). To quantify this more precisely we compute the correlation function of the hexatic order parameter  $g_6(r) = \langle \psi_{6j}^* \psi_{6k} \rangle_{|\mathbf{r}_k - \mathbf{r}_j| = r} / \langle |\psi_{6j}|^2 \rangle$  and the correlation function of the velocity orientation vector  $g_v(r) = \langle \hat{\mathbf{v}}_j \cdot \hat{\mathbf{v}}_k \rangle_{|\mathbf{r}_k - \mathbf{r}_j| = r}$ . These functions are computed and reported in Fig. 10 considering only particles in the dense phase of the largest system. We find that both  $g_6$  and  $g_v$  decay to zero in an exponential-like fashion as shown in the double-log inset Fig. 10. We assume that both correlators are well described by a Ornstein-Zernike form in  $q$ -space (i.e.  $g(q) \sim (\xi^{-2} + q^2)^{-1}$ ) and therefore we fit both data-sets with a function of the form  $g(r) = A K_0(r/\xi) + B$  where  $K_0$  is the modified Bessel function of the second kind  $\xi$  is the correlation length and  $A$  and  $B$  are amplitude and shift factors. The fit is quite good and reveals (in agreement with the qualitative map analysis discussed above) that the typical correlation lengths of the hexatic domains and velocity-oriented domains are, respectively,  $\xi_6 = 9.9(1.2)$  and  $\xi_v = 4.89(0.15)$ .

