In the Auxiliary Material we give details of the 2D fermion Hubbard model QMC calculations from which the energy and entropy scales reported in the main paper are obtained.

Our methodology is determinant QMC, a technique which solves interacting fermion Hamiltonians exactly on finite lattices, with error bars which can be systematically reduced to zero by increasing the run time and utilizing a finer mesh when discretizing the inverse temperature. For details see [1]. Most of our results are on lattices with $N=10^2$ sites, with some checks (see, e.g., inset to Fig. 1(c) of the main paper) on larger lattices of sizes up to $N=14^2$.

We have given all our results in terms of the parameters in the Hubbard Hamiltonian, and, in particular, the ratio $U/t$ of the interaction to kinetic energy scales. For optical lattices, the nearest-neighbor hopping amplitude $t$ is determined by the overlap of Wannier functions centered around neighboring lattice sites, which in turn can be written in terms of the lattice depth, wavelength and recoil energy. The Hubbard interaction $U$ depends both on the scattering length for fermions and on the Wannier functions. For a detailed discussion, see ref. [2, 3].

**Entropy:** The entropy $S$ is a quantity of central interest in our paper, and we calculate it in two different ways. The first method [4, 5] is to integrate the energy $E = \langle \mathcal{H} \rangle$ down from infinite temperatures to the temperature $T = \beta^{-1}$ of interest, using

$$ S(\beta)/k_B = N \ln 4 + \beta E(\beta) - \int_0^\beta d\beta' E(\beta'). \tag{1} $$

The second method [6] is to integrate up from $T = 0$. Here we fit the data for $E(T)$ using a suitable functional form (sum of exponentials), find the specific heat $C(T) = dE(T)/dT$ and calculate

$$ S(T) = \int_0^T dT' C(T')/T'. \tag{2} $$

We have checked that the answers obtained from the two methods agree to within a few percent.

**Repulsive Hubbard Model at half-filled:**

Compressibility and insulating gap: To quantify the opening of the insulating gap in the repulsive Hubbard model at half-filling, we calculate the compressibility $dp/d\mu$ where $\rho = \sum_i (n_i)/N$ is the density on an $N$-site lattice and $\mu$ is the chemical potential. In Fig. 1(a) we plot the density $\rho$ versus chemical potential $\mu$. Note that as we change $\mu$ away from zero (the chemical potential required to be at $\mu = 0$), we necessarily move away from half-filling. We thus encounter the fermion sign problem, resulting in large error bars. Nevertheless, we clearly see that for $|\mu| < \Delta_{ch}$, the insulating (or “charge”) gap, the compressibility $dp/d\mu$ is very small; it should vanish at $T = 0$. This allows us to estimate $\Delta_{ch}$, whose $U/t$-dependence is shown in the paper.

To better understand the $U$-dependence of the gap, we use a simple mean field theory (MFT) of spin density wave (SDW) ordering, that shows the system is an
insulator at \( \rho = 1 \) for any \( U > 0 \). The noninteracting metallic ground state is unstable to an arbitrarily small repulsion \( U \), due to perfect nesting of the Fermi surface: \( \epsilon(k + Q) = \epsilon(k) \) for \( Q = (\pi, \pi) \). The ground state is then an insulator with long-range antiferromagnetic order at ordering wavevector \( Q \) and gap \( \Delta_{\text{MF}} \) given by solving \( N^{-1} \sum_k 1/[2\sqrt{\epsilon_k^2 + \Delta_{\text{MF}}^2}] = 1/U \), shown in Fig. 1(b). We thus find that the MFT gap is a monotonically increasing function of \( U/t \) which starts off as \( \Delta_{\text{MF}} \sim t \exp[-2\pi \sqrt{t/U}] \) for \( U/t \ll 1 \) and smoothly crosses over to the Mott gap \( \Delta_{\text{MF}} \sim U/2 \) for \( U/t \gg 1 \). As already noted in the paper, this evolution from a SDW insulator to a Mott insulator is the \( U > 0 \) analog of the BCS to BEC crossover in attractive Fermi systems.

Although it correctly captures the qualitative behavior of the gap as a function of \( U/t \), MFT is quantitatively not accurate. \( \Delta_{\text{MF}} \) overestimates \( \Delta_{\text{SF}} \), obtained from the QMC data for \( \rho(\mu) \), by a surprisingly large factor (of approximately \( 3 \) – \( 4 \)) for large coupling, as shown in Fig. 1(b). MFT also makes the incorrect prediction of antiferromagnetic (AF) long range order at finite temperature, that is destroyed by thermal fluctuations in 2D. Spin correlations: To study the antiferromagnetic correlations we have looked at a variety of observables, including the uniform magnetic susceptibility \( \chi(q = 0, T) \), spin-spin correlation function \( \langle S_i \cdot S_{i+d} \rangle \), and its Fourier transform, the structure factor \( S(q) \). Henceforth we will suppress the momentum argument \( q = 0 \) of the susceptibility and just refer to \( \chi(T) \).

The uniform susceptibility \( \chi(T) \) for various \( U/t \) is shown in Fig. 2. We find a broad peak in \( \chi(T) \) at a temperature \( T_{\text{spin}} \) below which short range AFM spin correlations are significant. We have also analyzed the short range spin-spin correlation function \( \langle S_i \cdot S_{i+d} \rangle \) as a function of \( T \) for various \( U/t \) and found that it too shows a growth in correlations at the temperature scale \( T_{\text{spin}} \).

To get a better feel for \( T_{\text{spin}} \) we also plot in Fig. 2 the uniform susceptibility for the 2D nearest-neighbor \( S = 1/2 \) Heisenberg AF \( \chi(T) \), that describes the low energy spin physics of the \( U/t \gg 1 \) Mott-Hubbard insulator. Here too we see a peak in \( \chi \) at \( T_{\text{spin}} = J_{\text{AF}} \). To plot this data together with our Hubbard model results, we have chosen (somewhat arbitrarily) \( U/t = 12 \), so that the AF superexchange \( J_{\text{AF}} = 4t^2/U = t/3 \).

Finally, let us discuss the way in which \( T_{\text{spin}} \) is closely related to the crossover temperature below which the system appears to be ordered on finite lattices. At sufficiently low \( T \), the AFM correlation length grows as \( \xi \sim \exp(\alpha T_{\text{spin}}/T) \) where \( \alpha \) is of order unity. The temperature \( T_0 \) at which \( \xi \) reaches the linear size of the system \( L \) is then given by \( T_0 \sim \alpha T_{\text{spin}}/\ln L \). The scale of \( T_0 \) is set by \( T_{\text{spin}} \) and it is a very weak function of \( L \), even though it goes to zero for extremely large systems.

**Attractive Hubbard model at arbitrary filling:** At \( \rho = 1 \), the \( U < 0 \) model can be mapped to the \( U > 0 \) model using a particle-hole transformation. Thus, at half-filling the \( U < 0 \) model has \( SU(2) \) symmetry, corresponding to a degeneracy between charge-density-wave (CDW) ordering and s-wave superfluidity, which also means an absence of a finite temperature phase transition in 2D. However we observe the \( U < 0 \)
model away from half-filling ($\rho \neq 1$) which changes the symmetry to $U(1)$. This leads to a ground state with only superfluid order, and a finite temperature Berezinskii-Kosterlitz-Thouless (BKT) phase transition in 2D.

Pairing Pseudogap scale: The susceptibility $\chi(T)$, plotted in Fig. 3(a) for a sequence of attractive couplings $|U|/t$, is the simplest way to determine the pairing pseudogap temperature $T^*$. For $T < T^*$, singlet pairing correlations lead to a suppression of the $\chi(T)$. The peak in $\chi$ is thus a measure of the pairing scale $T^*$.

Berezinskii-Kosterlitz-Thouless (BKT) Transition: At still lower temperatures there is a transition to a superfluid phase in the attractive Hubbard model for $\rho \neq 1$. To determine the transition temperature $T_c$ we compute the superfluid density $\rho_s(T)$, defined in terms of the transverse current-current correlation function [8]. It is well-known from BKT theory that the superfluid density has a universal jump at $T_c$ with $\rho_s(T_c^-) = 2 T_c / \pi$. Strictly speaking, the $\rho_s$ computed on a finite lattice overestimates the superfluid density in the thermodynamic limit. Thus the $T_c$ estimated from Fig. 3(b) is an upper bound on the true BKT $T_c$.

We emphasize the nonmonotonic dependence of the BKT $T_c$ on $|U|/t$ in the inset to Fig. 2(a) of the main paper. For a filling of $\rho = 0.7$, we find a maximum $T_c \approx 0.17t$ as a function of attraction at $|U|/t \approx 5$. This non-monotonicity, though expected on general grounds (as explained in the paper), was not seen in any previous calculation.

---