Supporting Information

Hruby et al. 10.1073/pnas.1720415115

Lattice Calibrations

We calibrate the lattice depth along the x direction by amplitude modulation spectroscopy, observing the position of the lowest three Bloch bands (1). The lattice depths along the y and z directions are calibrated via Raman–Nath diffraction (2). The lattice depths are calibrated separately for each experiment, and we obtain the following parameters. In the metastability measurement, $V_x = 1.56(5) \times V_z, V_y = 30.7(1.6)\ E_R^{671}$, and $V_z = 11.1(7)\ E_R^{785}$. In the hysteresis measurement, $V_x = 1.12(3) \times V_z$, $V_y = 26.2(1.1)\ E_R^{671}$ when the y lattice is present and $V_x = 1.02(4) \times V_z$ when the y lattice is switched off, and V_z ranges from $5.7(4)\ E_R^{785}$ to $12.9(2)\ E_R^{785}$. In the phase diagram measurement, $V_x = 1.27(11) \times V_z$, $V_y = 26.2(1.1)\ E_R^{671}$, and V_z ranges from $4.0(5)\ E_R^{785}$ to $12.9(2)\ E_R^{785}$. Errors on the lattice depths in the y and z directions incorporate uncertainties from the calibration and residual offsets on the photodiodes.

Detuning Calibrations

The BEC couples to two linearly polarized TEM_{00} eigenmodes of the cavity, which are tilted by $\alpha = 22^\circ$ with respect to the y and z axes. The resonance frequencies of the eigenmodes are separated due to birefringence by $\delta_B/2\pi = 2.2$ MHz. The detuning Δ_c refers to the lower-lying resonance frequency of the mainly z-polarized mode, and the x lattice is detuned by $2\pi \times 30$ MHz from this mode (Fig. S1). In every experimental repetition, after atomic absorption pictures are taken, we scan the frequency of the x lattice across the cavity resonance and fit the resulting photon signal with a Lorentzian. We deduce a SD of $\Delta_c/2\pi$ of 0.3 MHz.

Magnetic Fields and Gradients

We apply a magnetic gradient field levitating the atomic cloud. In addition, we operate the experiment at a magnetic offset field large enough to achieve a good separation between the atomic hyperfine levels to avoid Raman-assisted spin flips induced by the presence of the lattices and the cavity. We use a magnetic field of $B \approx 130$ G oriented along the z axis and obtain a Zeeman splitting of about $\Delta E/h \approx 90$ MHz, well above the maximum cavity detuning of $\Delta_c^{max}/2\pi = -53$ MHz.

Trapping Frequencies

In our system, the cloud is magnetically levitated and subject to a crossed far off-resonant dipole trap. In the absence of optical lattices, we calculate the trapping frequencies in all three directions and find $(\omega_x, \omega_y, \omega_z) = 2\pi \times (96, 38, 49)$ Hz, respectively. When we include a 671-nm blue-detuned y lattice of depth $V_y = 30\ E_R^{671}$, which is comparable to our experimental parameters, together with an increased dipole trap depth, we calculate trap frequencies of $(\omega_x, \omega_y, \omega_z) = 2\pi \times (116, 38, 67)$ Hz. In the case of lattice depths comparable to the maximum lattice depths used in the experiment ($V_x = V_z = 14\ E_R^{785}$, $V_y = 30\ E_R^{671}$) we calculate trap frequencies of $(\omega_x, \omega_y, \omega_z) = 2\pi \times (219, 221, 193)$ Hz. Deconfinement due to changing zero-point energies is taken into account. We compare our calculations with experimental data and find good agreement. We estimate an error of about 10% resulting primarily from uncertainties in the determination of beam waists at the position of the atoms.

Extraction of the Even–Odd Particle Imbalance Θ and Site Offset δ_{off} from the Measured Photon Flux

We obtain the imbalance Θ from the mean intracavity photon n_{ph} number via

$$\Theta = \left| \sum_{i \in e} \langle \hat{n}_i \rangle - \sum_{i \in o} \langle \hat{n}_i \rangle \right| = \sqrt{n_{ph} \frac{\Delta_c^2}{\eta^2 M_0^2}} \frac{1}{F(\Delta_c)}, \quad [S1]$$

with

$$F(\Delta_c) = \Delta_c \left| \frac{\cos^2\alpha}{\Delta_c' - \delta_B + i\kappa} + \frac{\sin^2\alpha}{\Delta_c' + i\kappa} \right| \overset{|\Delta_c| \gg \kappa, |\delta|, \delta_B}{\approx} 1. \quad [S2]$$

$F(\Delta_c)$ takes into account the two linearly polarized TEM_{00} eigenmodes of the cavity. The effective two-photon Rabi frequency is given by $\eta/2\pi = 2.99\sqrt{V_z/\hbar}\sqrt{\mathrm{Hz}}$, the spatial overlap of the interference lattice provided by the cavity mode and the z lattice with the Wannier-function $W_i(x,z)$ of an atom localized at lattice site i is given by $M_0 = \int\int \mathrm{d}x\,\mathrm{d}z\ W_i^*(x,z)\cos(kx)\cos(kz)\,W_i(x,z)$, the cavity decay rate is $\kappa/2\pi = 1.25$ MHz, and $\Delta_c' = \Delta_c - \delta$ takes into account the dispersively shifted cavity resonance, where δ corresponds to the dispersive shift with a maximum shift per atom of $U_0/2\pi = -56.3$ Hz for each of the two cavity modes. A moving average of window size 4 ms is used on all photon data except for the phase diagram in Fig. 6 where the window size is 10 ms. Note that technical noise on the photon detector is converted into an imbalance Θ. Due to the dependence of Θ on Δ_c and V_z, the background noise causes a noticeable signal far from cavity resonance and contributes to the small but nonzero imbalance visible in Figs. 2*D*, *Left*; 3 *A*, *C*, and *D*; and 4*B*, *Left* and to the imbalance visible in Fig. 6, *Lower Left*. For further details see ref. 3.

The energy offset δ_{off} between even and odd sites is related to the strength of the dynamic checkerboard lattice depth formed by the z lattice and the light scattered into the cavity. It is defined as

$$\delta_{off} = 4\eta M_0 \sqrt{n_{ph}}. \quad [S3]$$

Strength of Effective Atom–Atom Interactions of Global Range

Taking both cavity modes into account, U_l is given by

$$U_l = -K|\eta M_0|^2 \left[\frac{(\Delta_c' - \delta_B)\cos^2\alpha}{(\Delta_c' - \delta_B)^2 + \kappa^2} + \frac{\Delta_c' \sin^2\alpha}{\Delta_c'^2 + \kappa^2} \right] \overset{|\Delta_c| \gg \kappa, |\delta|, \delta_B}{\approx} -K|M_0|^2 \frac{\eta^2}{\Delta_c} \propto \frac{V_z}{\Delta_c} s. \quad [S4]$$

We take the number of lattice sites to be the number of atoms, $K = N$; for details see ref. 3.

Derivation of the Extended Bose–Hubbard Toy Model

Our system is well described by a Bose–Hubbard Hamiltonian with additional global-range interactions of the form (3, 4)

$$\frac{\hat{\mathcal{H}}}{\hbar} = -t \sum_{<i,j>} \left(\hat{b}_i^\dagger \hat{b}_j + h.c. \right) + \frac{U_s}{2} \sum_{i \in e,o} \hat{n}_i (\hat{n}_i - 1) - \sum_{i \in e,o} V_i \hat{n}_i - \frac{U_l}{K} \left(\sum_{i \in e} \hat{n}_i - \sum_{i \in o} \hat{n}_i \right)^2, \quad [S5]$$

where t is the nearest-neighbor tunneling rate, V_i is the site-dependent harmonic trapping potential, $\hat{b}_i$ and $\hat{b}_i^\dagger$ are the bosonic annihilation and creator operators at site i, and $\hat{n}_i = \hat{b}_i^\dagger \hat{b}_i$ is the corresponding number operator. In our toy model, we

assume the limit of zero tunneling and neglect the harmonic trapping potential. We consider a total of N atoms to be distributed among a fixed number of lattice sites $K = N$, independent of the strength of global-range interactions U_l. This is experimentally realistic for deep lattices where an atomic wavepacket cannot spread more than a few lattice sites during the experiment due to very small tunneling t.

In the limit of $t = 0$, the eigenstates of the system are the number states, and we can replace all of the number operators in Eq. **S5** by the corresponding average values, $\langle \hat{n}_i \rangle = n_i$. We introduce the imbalance Θ as

$$\Theta = \langle \hat{\Theta} \rangle = \left\langle \sum_{i \in \mathsf{e}} \hat{n}_i - \sum_{i \in \mathsf{o}} \hat{n}_i \right\rangle = \sum_{i \in \mathsf{e}} n_i - \sum_{i \in \mathsf{o}} n_i. \quad \textbf{[S6]}$$

The atomic configuration of least energy for a given imbalance Θ corresponds to part of the system being in a CDW state, namely the fraction $f_{\mathsf{CDW}} = |\Theta|/K$, while all other atoms are in an MI state. The CDW is characterized by $n_i^{\mathsf{even}} = 2$ and $n_i^{\mathsf{odd}} = 0$ for even and odd sites, respectively, while $n_i = 1$ on all lattice sites in the MI state. The energy of such a state is

$$\langle \hat{\mathcal{H}} \rangle = N\varepsilon = \frac{1}{2} f_{\mathsf{CDW}} U_\mathsf{s} K - \frac{U_\mathsf{l}}{K}\Theta^2 = \frac{1}{2}\frac{|\Theta|}{K} U_\mathsf{s} K - \frac{U_\mathsf{l}}{K}\Theta^2. \quad \textbf{[S7]}$$

With $N = K$, we obtain

$$\varepsilon = \frac{1}{2} U_\mathsf{s} \frac{|\Theta|}{N} - U_\mathsf{l}\left(\frac{|\Theta|}{N}\right)^2. \quad \textbf{[S8]}$$

The system changes its ground state from an MI state with no imbalance to a CDW state with maximum imbalance ($|\Theta|/N = 1$) at $U_\mathsf{l}/U_\mathsf{s} = 1/2$ (Fig. 1*A*). We use this critical point to calculate the energy barrier per particle between the MI and the CDW state which is defined as $E^{\mathsf{barrier}} = E_m - E_g$. Here $E_g = 0$ is the ground-state energy and $E_m = U_\mathsf{s}/8$ is the maximum energy as a function of Θ at $|\Theta|/N = 1/2$. For $(V_\mathsf{x}, V_\mathsf{y}, V_\mathsf{z}) = (13\ E_\mathsf{R}^{785}, 26\ E_\mathsf{R}^{671}, 13\ E_\mathsf{R}^{785})$, we obtain an energy barrier of $E^{\mathsf{barrier}}/h = 260$ Hz which is much larger than the single-particle tunneling rate $t/2\pi = 46$ Hz.

The presence of a trapping potential can lower the energy of a state of intermediate imbalance and possibly reduce the height of the energy barrier between the MI and CDW phases. This reduction in energy can be as large as 600 Hz at the edge of the central 2D layer. Assuming such a situation to be present everywhere in the system, the energy barrier per particle is reduced to about $E^{\mathsf{barrier}}/h = 180$ Hz, which is still significantly larger than the tunneling rate $\sqrt{2}t$ (5).

Calculation of Atomic Density Distributions

Number of 2D Layers. We calculate the number of 2D layers based on the measured atom number and the calculated trap frequencies, following ref. 6. Since the lattice along the y direction is very deep, we assume the atom number in each 2D layer to be fixed.

Maximum Lattice Filling. Following ref. 7, we calculate the atomic density as a function of μ/U_s and t/U_s in the grand canonical ensemble, where μ is the chemical potential. Using the local density approximation and calculated trapping frequencies, we obtain the full density distribution of the atomic cloud which is used to estimate the maximum filling n_i.

Number of Surface Atoms N_{surf}. In the metastability measurement the 2D lattice has different strengths in the x and z directions ($V_\mathsf{x} = 17.3\ E_\mathsf{R}^{785}$, $V_\mathsf{z} = 11.1\ E_\mathsf{R}^{785}$). We estimate N_{surf} from the calculated atomic density distribution in a balanced 2D square lattice around the average lattice depth $\overline{V} = \frac{1}{2}(V_\mathsf{x} + V_\mathsf{z})$. We obtain $N_{\mathrm{surf}} = (4\text{–}8) \times 10^3$ atoms at $\overline{V} = (15\text{–}13)\ E_\mathsf{R}^{785}$, respectively.

Evaluation of the Metastability Measurement

The data are taken in a range of final detunings of $-36 \leq \Delta_\mathsf{c}^\mathsf{f}/2\pi \leq -16$ MHz with an interval of 0.5 MHz, amounting to a total of 41 datasets. For every $\Delta_\mathsf{c}^\mathsf{f}$, the experiment is repeated 13–22 times. In each repetition we start with a detuning ramp in the time interval $0 < T < 20$ ms, followed by a free evolution at $20 \leq T < 70$ ms. The imbalance $\overline{\Theta}$ is obtained as the mean of the imbalance Θ in the time interval $50 < T < 60$ ms. The two distinct imbalance distributions are highlighted by coloring data with $0 < \overline{\Theta} < 7 \times 10^3$ atoms in orange and data with $\overline{\Theta} > 7 \times 10^3$ atoms in green (Fig. 2).

At each final detuning $\Delta_\mathsf{c}^\mathsf{f}$, we take the mean and SD of data in the orange and green region separately, and we obtain Fig. 2*A*. To quantify the gap between the two states, we consider the final detuning region where we find states with both small and large imbalance simultaneously; i.e., -24 MHz $\leq \Delta_\mathsf{c}^\mathsf{f}/2\pi \leq -19.5$ MHz. We consider data above and below $\overline{\Theta} = 7 \times 10^3$ atoms separately and take the mean and SD. The difference defines the gap between the two states, which has a height of $5.2(1.4) \times 10^3$ atoms. In another representation of the same data, we split the imbalance data of each $\Delta_\mathsf{c}^\mathsf{f}$ into 22 bins of bin size 700 atoms and construct a histogram as a function of $\overline{\Theta}$ and $\Delta_\mathsf{c}^\mathsf{f}$ (Fig. 2*B*). To obtain mean counts $\overline{\mathsf{Counts}}$ as shown in Fig. 2*C*, we generate a histogram with respect to $\overline{\Theta}$ of data in the orange region of Fig. 2*B*, where we normalize counts by the respective sample size (469), and we repeat this procedure for data in the green region which has a sample size of 181. This way the obtained histogram becomes independent of the exact sample size in each state, as the sample size is sensitive to the scan region of final detunings.

Hysteresis Measurement: Lattice and Detuning Ramps

The BEC is initially prepared in a crossed far-off resonant dipole trap. Then a strong y lattice is ramped within 100 ms to a final depth of V_y, where the ramp follows an S shape of form $V(T) = V_0\left[3\left(\frac{T}{T_0}\right)^2 - 2\left(\frac{T}{T_0}\right)^3\right]$. Here V_0 is the final lattice depth, T is time, and T_0 is the total duration of the ramp. The y lattice cuts the cloud into weakly coupled 2D layers. The subsequent sequence of amplitude and detuning ramps is shown in Fig. S2. First, the square lattice in the x–z direction is applied using another S-shaped amplitude ramp of 50-ms duration, finishing at depths V_x and V_z. Then the z-lattice detuning $\Delta_\mathsf{c}/2\pi$ is swept from -53 MHz to -13 MHz within a variable time of $\tau = (30\text{–}150)$ ms using a ramp which varies U_l linearly in time. The ramp has the form $\Delta_\mathsf{c}(T) = \left[\left(\frac{1}{\Delta_\mathsf{c}(\tau)} - \frac{1}{\Delta_\mathsf{c}(0)}\right)\frac{T}{\tau} + \frac{1}{\Delta_\mathsf{c}(0)}\right]^{-1}$, where $\Delta_\mathsf{c}(0)$ and $\Delta_\mathsf{c}(\tau)$ represent the initial and final detuning, respectively. Subsequently, the detuning $\Delta_\mathsf{c}/2\pi$ is swept back to -53 MHz, using an inverted ramp of the same duration. Finally, the square lattice is ramped down within 50 ms, using another S-shaped ramp.

Hysteresis Loops: Data Evaluation and Comparison of Different Ramp Times

Extraction of Thresholds. In the hysteresis measurement, we extract the threshold for the onset of an imbalance Θ during ramp I and the threshold where Θ vanishes again during ramp II. We define both of these thresholds as the point where the intracavity photon signal is 20 times higher than the mean background level. The background level is obtained by averaging the photon signal over 50 ms while all lattices are switched off. As a result of this method, the imbalances Θ at the threshold positions are of different magnitude (orange and green diamonds in Fig. 3*A*). The experiment is repeated at least three times for every lattice depth V_z, and the corresponding averaged thresholds for

the z-lattice depth V_z and detuning Δ_c are shown by orange and green diamonds in Fig. 6.

Hysteresis Loop Definition. We show closed hysteresis loops of the imbalance Θ as a function of global-range interaction strength U_l in Fig. 3*A*. The loop naturally closes at high U_l (right side of Fig. 3*A*) where the detuning ramp is inverted. At low U_l (left side of Fig. 3*A*) we plot data down to the point where the two curves cross. We consider only crossing points which happen below the thresholds of the creation and disappearance of an imbalance. This additional condition is needed to exclude crossings happening in the middle of the hysteresis loop due to, e.g., heating; such a case is visible in Fig. 3*D*. To reduce noise, we average U_l using an averaging window of $2\pi \times 20$ Hz to find this crossing point.

Hysteresis Area. The hysteresis area A is obtained by integrating the imbalance Θ as a function of U_l during ramp II and subtracting this signal from the integrated curve during ramp I. We define a normalized hysteresis area as the ratio of hysteresis area A and a factor A_{max}. Here A_{max} is a fixed constant which defines the maximum possible hysteresis area, i.e., the product of the total number of atoms and the maximum strength of U_l. The hysteresis area shown in Figs. 3*B* and Fig. S3 is the average of at least three repetitions for every lattice depth V_z. To ensure comparability of the data, we use hysteresis area data only for those lattice depths for which the averaged maximum imbalance Θ_{max} satisfies the constraint that $\Theta_{max} \geq \Theta_{z,max} - \Delta\Theta_{z,max}$, where $\Theta_{z,max}$ is the average of the maximum imbalance obtained for the case of deepest lattices and $\Delta\Theta_{z,max}$ is the corresponding SD.

At small U_l, changes in the interaction strength stem from ramping the lattice depth V_z which changes both U_l and U_s. However, a large fraction of the hysteresis loop is occurring during the frequency ramps where U_l is varying while U_s stays constant. Taking for example the case of a frequency ramp of duration $\tau = 80$ ms as shown in Fig. 3, U_s is reduced by less than 9% at the point where the hysteresis loop closes for small U_l. At ramp times of $\tau = (30-50)$ ms, this reduction in U_s increases to 23%.

We note that we do not use U_l/U_s as an x axis for the extraction of a hysteresis area as it does not allow a direct comparison between the case of strong short-range interactions U_s when all 3D lattices are present and the case where the y lattice is switched off to reduce U_s.

Hysteresis Area as a Function of Ramp Time τ. A study of the hysteresis area is shown in Fig. S3. In all cases, we observe a qualitatively comparable behavior as in Fig. 3*B* where increasing interactions increase the observed hysteresis area. Heating from the presence of optical lattices reduces the overall signal with increasing ramp time, leading to a negative hysteresis area clearly visible in Fig. S3 *C* and *D*. The difference in the hysteresis area between the case of strong and weak U_s (with and without the y lattice, respectively) is nearly the same for different ramp times.

Imbalance Dynamics: Data Evaluation

To quantify the position, duration, and height of the fast jump as shown in Fig. 4, we use the following definition of an effective derivative,

$$\left.\frac{d\Theta}{dt}(T)\right|_{\xi} = \frac{1}{\xi}\left[\max\left[\Theta\left(T-\frac{\xi}{2}:T+\frac{\xi}{2}\right)\right] - \min\left[\Theta\left(T-\frac{\xi}{2}:T+\frac{\xi}{2}\right)\right]\right], \quad \textbf{[S9]}$$

where max and min yield the maximum and minimum value of Θ within a time interval of $\pm\xi/2$ around the time T, and we use $\xi = 4$ ms. This effective derivative helps to improve the signal-to-noise ratio. The fast jump is then associated with a maximum in the amplitude of the effective derivative. We fit the signal from the effective derivative with a Gaussian in a time window of ± 10 ms around the fast jump. The central position of the Gaussian fit, t_0, is used to extract the position of the fast jump in terms of the z-lattice depth and the detuning Δ_c (Fig. 6). The FWHM of the Gaussian represents the duration of the jump ξ_{jump} and is used to extract the jump height

$$h_{jump} = \Theta(T_0 + \xi_{jump}/2) - \Theta(T_0 - \xi_{jump}/2).$$

To extract T_0, ξ_{jump}, and h_{jump} we consider only those experimental realizations where the fast jump occurs when all of the external parameters are kept constant after the quench, and we obtain 54 such realizations. Formally, this constraint is defined as $T_0 - \xi_{jump}/2 > T_{const}$, where T_{const} is the time from which all external parameters are kept constant. The height $\Delta\Theta$ and duration ΔT of the fast jump stated in the main text are obtained by averaging all individual data of h_{jump} and ξ_{jump}.

Repeating the extraction procedure described above with reduced time interval ξ or reduced moving average window size, we observe shorter durations of the step at the cost of a reduced signal-to-noise ratio. The value provided in the main text is thus an upper bound on the actual step duration.

We obtain the tunneling time in a double well in the following way. We consider the two states $|1,1\rangle$ and $|2,0\rangle$ resonantly coupled by the tunneling $\sqrt{2}t$, where $\sqrt{2}$ accounts for bosonic enhancement. Starting in the state $|1,1\rangle$, the system reaches the state $|2,0\rangle$ within the tunneling time.

Phase of the Light Field. Using our heterodyne detection we also extract the time phase of the light field scattered into the cavity with respect to the lattice in the z direction (8). Because of residual phase drifts of the heterodyne setup, we cannot relate the phase signals between consecutive experimental runs. To improve clarity of the phase signal shown in Fig. 4*C* a mean offset phase is subtracted in each realization to remove these shot-to-shot phase drifts. The mean offset phase is obtained by time averaging of the phase signal from 20 ms $\leq T \leq$ 65 ms in the metastability measurement and from 40 ms $\leq T \leq$ 80 ms in the hysteresis measurement. Here $T = 0$ ms corresponds to the initial time (0 ms) in Fig. 4.

Extraction of the Change in Excitation Energy ΔE from the Measured Photon Flux

We extract the change in excitation energy ΔE during the imbalance jump (*ii*), where τ_s counts the time since the beginning of the jump. At the beginning of the jump, superfluid surface atoms account for an initial imbalance of $\Theta(\tau_s = 0\text{ ms})$. We assume that the imbalance stemming from these surface atoms stays approximately constant during the jump. As the imbalance stemming from reordering bulk atoms increases with time τ_s the site offset δ_{off} also increases, reducing the excitation energy of all previously imbalanced atoms. From the measured imbalance $\Theta(\tau_s)$ and site offset $\delta_{off}(\tau_s)$ we obtain

$$\frac{\Delta E(\tau_s)}{h} = \underbrace{\int_{\tau_s'=0}^{\tau_s}\left[\delta_{off}(\tau_s) - \delta_{off}(\tau_s')\right]\frac{d\Theta}{d\tau_s'}d\tau_s'}_{\text{bulk}} + \underbrace{\left[\delta_{off}(\tau_s) - \delta_{off}(0)\right]\Theta(0)}_{\text{surface}}. \quad \textbf{[S10]}$$

The result is shown in Fig. 5*C*. We observe a released energy of $2\pi \times 7.7(2.1)$ MHz during the imbalance jump (*ii*), where surface atoms account for $2\pi \times 5.7(1.2)$ MHz and bulk atoms account for $2\pi \times 2.0(1.0)$ MHz.

Energy Stored in the Superfluid Surface and the Insulating Bulk During the Imbalance Jump

In the superfluid surface, energy can be stored in the form of phononic excitations. As a first estimate, we use Bogoliubov theory in a 2D lattice to calculate the bandwidth for the phonons in the static lattice, which then reads (9)

$$\varepsilon^{\mathrm{kin}}_{\mathrm{surf}}/\hbar = \sqrt{4(t_{\mathrm{x}}+t_{\mathrm{z}})\Big(4(t_{\mathrm{x}}+t_{\mathrm{z}})+2nU_{\mathsf{s}}\Big)}. \qquad \textbf{[S11]}$$

Here we take repulsive interactions in the form of $U_{\mathsf{s}}/2\pi = 2.2(1)$ kHz into account, resulting from the overall 3D lattice potential. Tunneling in the (x, z) direction is $(t_{\mathrm{x}}, t_{\mathrm{z}})/2\pi = (15, 57)$ Hz, and we assume a lattice filling of $n = 0.5$. Using the calculated number of surface atoms as $N_{\mathrm{surf}} = 6(2) \times 10^3$ atoms, we obtain the kinetic energy which can be stored in the surface as $E^{\mathrm{kin}}_{\mathrm{surf}}/h = \Big(N_{\mathrm{surf}}\,\varepsilon^{\mathrm{kin}}_{\mathrm{surf}}\Big)/h = 5.0(1.7)$ MHz. We note that Bogoliubov theory is strictly valid only in the limit of weak interactions $2nU_{\mathsf{s}} << 4(t_{\mathrm{x}} + t_{\mathrm{z}})$. In the regime of strong interactions, the atoms can sustain excitations which are particle- and hole-like and whose energy is proportional to U_{s}. Hence Bogoliubov theory provides a lower bound on the excitation energy ($\propto \sqrt{U_{\mathsf{s}}}$) that can be stored in the surface atoms.

In the insulating bulk, kinetic energy can be stored in the band of particle-hole excitations which exist on top of the $|2,0\rangle$ state. These particle-hole excitations are formed by all atoms which are not in the superfluid shells and not in the $|2,0\rangle$ state. We find the number of excitations as $N_{\mathrm{exc}} = N - N_{\mathrm{surf}} - N_{|2,0\rangle} = N - \Theta^{\mathrm{max}}_{\mathrm{jump}} = 16(3) \times 10^3$ atoms, where the imbalance at the end of the imbalance jump (*ii*) is given by $\Theta^{\mathrm{max}}_{\mathrm{jump}} = 9(1) \times 10^3$ atoms and the total number of atoms is $N = 25(2) \times 10^3$ atoms. We estimate the amount of kinetic energy which each excitation would have to carry from the reduction of excitation energy ΔE during the jump (*ii*). We obtain an energy per excited particle of about $130(70)$ Hz, which is on the order of the single-particle tunneling rates.

Number of Photons Scattered During the Imbalance Jump

We estimate the total number of photons incoherently scattered from the z lattice—off the atoms—into the single-cavity mode. In the bad-cavity limit we can consider a quasi-stationary intracavity light field. In this limit, scattering of photons into the cavity mode balances photon loss through the cavity mirrors. Photons leave the cavity at a rate given by the inverse cavity lifetime of $2\kappa = 2 \times 2\pi \times 1.25$ MHz. Here we neglect the low rate of scattering of incoherent cavity photons back into the z lattice as they will not exhibit bosonic enhancement. We observe an average mean intracavity photon number of $n_{\mathsf{ph}} = 0.18(2)$ during the time of the imbalance jump (*ii*) of $\Delta T = 4.3(6)$ ms. The scattering rate into the cavity then becomes $2.8(3) \times 10^6$ photons per second and the number of scattered photons during the jump is about $12(3) \times 10^3$ photons.

Phase Diagram Measurement: Data Evaluation

Phase Diagram Measurement. To construct the phase diagram of the system we follow ref. 3 with the difference that we now prepare a BEC of $16(1) \times 10^3$ atoms instead of $42(4) \times 10^3$ atoms. Due to the lower atom number, states with nonzero imbalance Θ are created closer to resonance with respect to ref. 3. The wavelength of the square lattice is now 784.7 nm instead of 785.3 nm previously while all other parameters are comparable.

Contrary to the metastability measurement and the hysteresis measurement, the detuning Δ_{c} is kept constant throughout each experimental sequence. We start with a BEC and slowly ramp up the lattice depth in all three directions. Then, all trapping potentials are abruptly switched off and absorption pictures of the atomic cloud are taken after 7 ms of ballistic expansion. We obtain the BEC fraction from a bimodal fit of the atomic density distribution and the maximum imbalance from the maximum photon flux leaking out of the cavity. To construct the phase diagram, this experiment is repeated at different detunings Δ_{c} and final lattice depths V_{z}.

Data Evaluation. Each data point of the phase diagram is taken on average four times. We obtain the phase boundary between states with and without spatial coherence for each detuning Δ_{c} from the position of a kink in the BEC fraction as a function of the lattice depth V_{z} (Fig. S4), which we associate with the loss of superfluidity and the formation of an insulating phase (10). We use a multiple-line fit to find the kink position. For each detuning Δ_{c} we estimate the SD of the kink position using a bootstrapping algorithm which resamples the data 10^4 times. The samples are constructed by taking out of the four experimental iterations one data point in the BEC fraction at random for each lattice depth V_{z}. The samples are then fitted individually and a histogram of the resulting kink positions is constructed (Fig. S4). We obtain the position of the kink and the 1σ SD shown in Fig. 6 from a Gaussian fit to the histogram.

We obtain information on the creation of an imbalance Θ by detecting photons leaking from the cavity with a heterodyne setup (11). In each experimental repetition we take a single data point of Θ after all lattices are ramped up and just before taking atomic absorption pictures. The photon data are resampled together with the measured z-lattice depth V_{z} to reduce noise. An averaging window of 10 ms is used. We deduce the phase diagram from the imbalance Θ and the transition between states with and without spatial coherence, where we use the criteria established in ref. 3: The superfluid region (SF) shows spatial coherence but no imbalance, the lattice supersolid region (SS) shows spatial coherence and a nonzero imbalance, the MI region shows no spatial coherence and no imbalance, and the CDW region shows no spatial coherence but a nonzero imbalance.

1. Stöferle T, Moritz H, Schori C, Köhl M, Esslinger T (2004) Transition from a strongly interacting 1D superfluid to a Mott insulator. *Phys Rev Lett* 92:130403.
2. Morsch O, Oberthaler M (2006) Dynamics of Bose-Einstein condensates in optical lattices. *Rev Mod Phys* 78:179–215.
3. Landig R, et al. (2016) Quantum phases from competing short- and long-range interactions in an optical lattice. *Nature* 532:476–479.
4. Dogra N, Brennecke F, Huber SD, Donner T (2016) Phase transitions in a Bose-Hubbard model with cavity-mediated global-range interactions. *Phys Rev A* 94:023632.
5. Panas J, Kauch A, Byczuk K (2017) Spectral properties and phase diagram of correlated lattice bosons in an optical cavity within bosonic dynamical mean-field theory. *Phys Rev B* 95:115105.
6. Pedri P, et al. (2001) Expansion of a coherent array of Bose-Einstein condensates. *Phys Rev Lett* 87:220401.
7. Dhar A, Singh M, Pai RV, Das BP (2011) Mean-field analysis of quantum phase transitions in a periodic optical superlattice. *Phys Rev A* 84:033631.
8. Baumann K, Mottl R, Brennecke F, Esslinger T (2011) Exploring symmetry breaking at the Dicke quantum phase transition. *Phys Rev Lett* 107:140402.
9. van Oosten D, van der Straten P, Stoof HTC (2001) Quantum phases in an optical lattice. *Phys Rev A* 63:053601.
10. Jiménez-García K et al. (2010) Phases of a two-dimensional Bose gas in an optical lattice. *Phys Rev Lett* 105:110401.
11. Landig R, Brennecke F, Mottl R, Donner T, Esslinger T (2015) Measuring the dynamic structure factor of a quantum gas undergoing a structural phase transition. *Nat Commun* 6:7046.

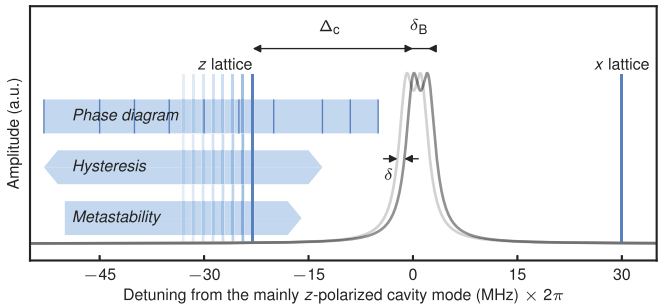

Fig. S1. Various detunings used in the experiment. The resonances of the two linearly polarized TEM_{00} modes of the empty cavity are shown by the dark gray line, and their resonance frequencies are separated by $\delta_B/2\pi = 2.2$ MHz due to birefringence. The full width at half maximum (FWHM) of each resonance is $2\kappa/2\pi = 2.5$ MHz. Coupling of atoms to the cavity shifts the cavity resonance by the dispersive shift δ down in frequency (light gray line). The *z* lattice is detuned by a variable amount of Δ_c from the lower-lying resonance frequency of the mainly *z*-polarized mode of the empty cavity, shown by the vertical blue line on the left half of the figure, where the neighboring light blue lines illustrate the scan direction. The *x* lattice is detuned by $2\pi \times 30$ MHz from the same mode, shown by the vertical blue line on the right. Horizontal arrows depict the scan directions and ranges of the different experiments, and the short vertical blue lines indicate where the phase diagram data are taken.

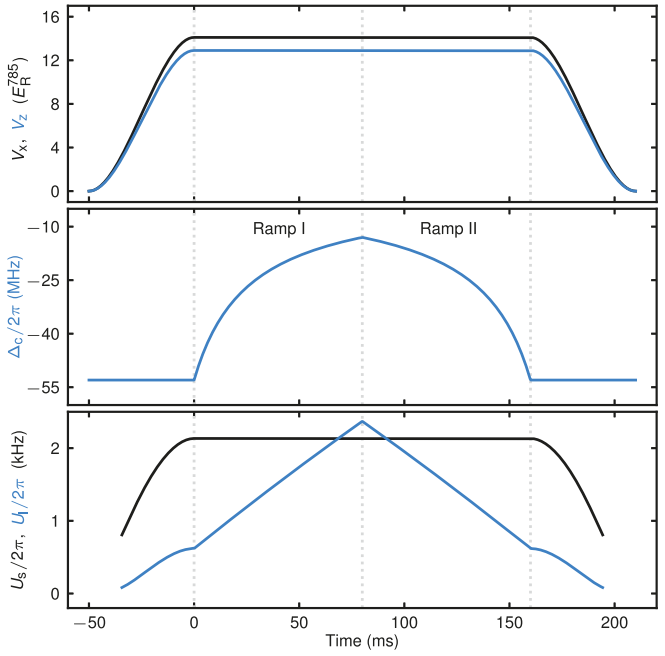

Fig. S2. Temporal sequence of lattice amplitude and detuning ramps in the hysteresis measurement. Time $T = [-50, 0]$ ms: The square lattice in the $x-z$ direction is ramped to a depth of (V_x, V_z) (*Top*) at a constant detuning Δ_c (*Middle*), predominantly increasing U_s but also U_l (*Bottom*). $T = [0, 80]$ ms: The detunig Δ_c is ramped toward resonance (ramp I) to vary U_l linearly in time while U_s is kept constant. [80, 160] ms: The ramp in the detuning is inverted and Δ_c is brought back to the starting point (ramp II). [160, 210] ms: The square lattice in the $x-z$ direction is ramped down again.

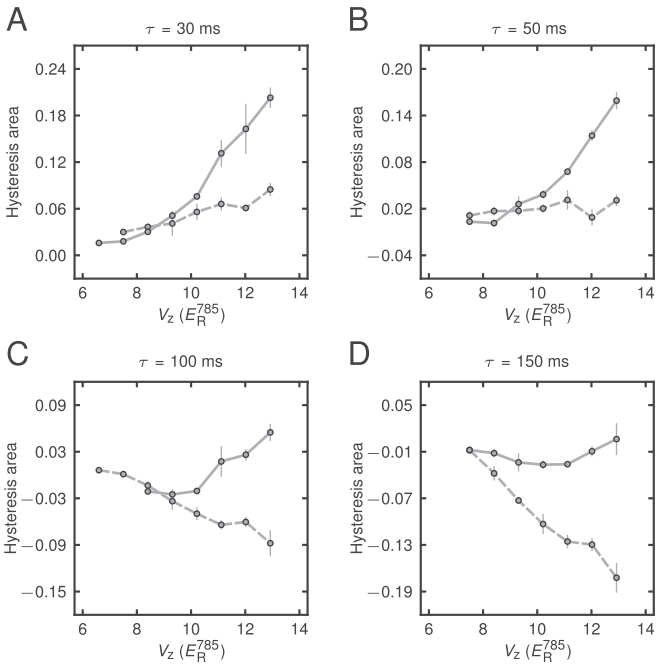

Fig. S3. Hysteresis area as a function of the ramp time τ. We obtain the hysteresis area for ramp times (*A*) $\tau = 30$ ms, (*B*) $\tau = 50$ ms, (*C*) $\tau = 100$ ms, and (*D*) $\tau = 150$ ms in the same way as in Fig. 3*B*. Solid lines represent the case where lattices in all three directions are applied, while dashed lines represent the case where the *y* lattice is switched off to reduce U_s. Data points represent statistical means and errors are SD.

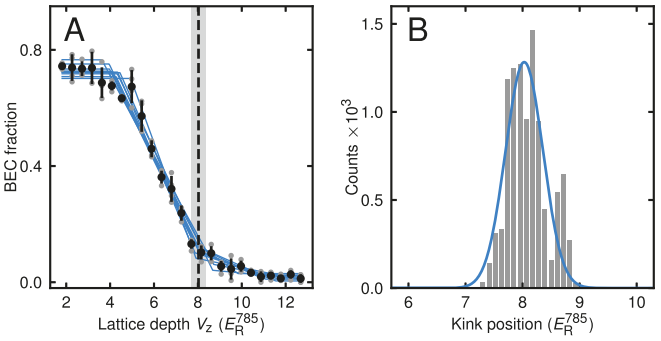

Fig. S4. Extraction of the phase boundary between states with and without spatial coherence. (*A*) The BEC fraction as a function of the *z*-lattice depth V_z is shown for a detuning of $\Delta_c/2\pi = -13$ MHz. We observe a kink in the BEC fraction and we use a multiline fit to extract its position. The data are resampled 10^4 times using a bootstrapping method and fitted separately to estimate the 1σ SD (gray area) around the kink position (dashed line). The blue lines represent the fit results of 10 random samples. (*B*) Histogram of the kink position resulting from resampling the data. We fit a Gaussian to the histogram and extract the kink position and the 1σ SD from this fit.