Quantized nonlinear Thouless pumping

We illustrate our findings of quantized nonlinear Thouless pumping in an array of coupled waveguides. In the scalar-paraxial regime, the propagation of monochromatic light in the array is governed by the discrete nonlinear Schrödinger equation:

$$i \frac{\partial}{\partial z} \phi_n = \sum_{m} H_{n,m}(z) \phi_m - g|\phi_n|^2 \phi_n$$  \hspace{1cm} (1)

Here, $\phi_n(z)$ is the wavefunction (the envelope of the electric field) in waveguide $n$; $m$ and $n$ run over all waveguides, $H$ is the linear $z$-dependent tight-binding Hamiltonian (for example, describing a linear Thouless pump) and $z$ is the propagation distance, which for Thouless pumps plays the part of a synthetic wavevector dimension. The parameter $g$ describes the strength of the nonlinearity and is positive (negative) for a focusing (defocusing) Kerr nonlinearity. In the case of sufficiently low intensities, the equation reduces to the linear Schrödinger equation.

From an experimental point of view, $g$ is dependent on the nonlinear refractive index coefficient of the underlying material, on the effective area of the waveguide modes, and on the wavelength. This nonlinear Schrödinger equation with $g > 0$ is equivalent to an attractive Gross–Pitaevskii equation describing bosonic interactions in a Bose–Einstein condensate in the mean-field limit. Indeed, the results we obtain below are generically applicable to a range of bosonic wave systems.

Aubry–André–Harper model

We use an off-diagonal implementation of the Aubry–André–Harper (AAH) model with three sites per unit cell labelled A, B and C (Fig. 1a).

The off-diagonal AAH model is described by a tight-binding Hamiltonian with equal on-site potential (which can be set to zero at all lattice sites) and real off-diagonal nearest-neighbour couplings $J_n$ that are periodic functions with modulation frequency $\Omega$. The modulation of the couplings over one period is displayed in Fig. 1b, where the choice of colours corresponds to Fig. 1a. Replacing the general linear Hamiltonian in equation (1) with the AAH model results in:

$$i \frac{\partial}{\partial z} \phi_n = -J_n(z) \phi_{n+1} - J_{n-1}(z) \phi_{n-1} - g|\phi_n|^2 \phi_n$$  \hspace{1cm} (2)

This equation conserves the norm of the solution: $P = \sum_n |\phi_n|^2$.

To emphasize the relationship between a 1 + 1-dimensional pumping model (1 spatial and 1 propagation/temporal dimension) to a two-dimensional Chern insulator, we plot the band structure in Fig. 1c. At each point $z$ within a period, the energy eigenvalues of the instantaneous Hamiltonian (calculated for an array of 30 waveguides with open boundary conditions) are calculated and plotted. The band structure shows three bands with Chern numbers $C = [-1, 2, -1]$ connected via topological end states (red). A schematic illustration of a realization in a one-dimensional array of evanescently coupled waveguides is shown in Fig. 1d. The modulation of the coupling is achieved by periodically modulating the waveguide positions and therefore changing the spatial overlap of neighbouring waveguide modes. The position of waveguide $n$ in the transverse direction is given by $x(z) = n + d \cos(2\pi n/3 + \Omega z + \alpha_0)$ with $d$ being the average separation between two waveguides, $\delta$ the spatial modulation strength and $\alpha_0$ an initial phase. The white-light micrograph in Fig. 1e shows the output facet of a waveguide array with six out of ten unit cells.
Linear and nonlinear evolution

Using this Thouless pump, we demonstrate the differences between linear and nonlinear quantized pumping in three distinct regimes with different powers: (1) a low-power, linear regime in which the wavefunction evolves according to the linear Schrödinger equation; (2) an intermediate-power regime in which we observe the formation of a soliton that is pumped by a fixed number of unit cells during a pumping period; and (3) a high-power regime in which we observe a trapped soliton. We refer to the three regimes as the linear, pumped and trapped regimes, respectively. Numerical propagation simulations of equation (2) with a 4th-order Runge–Kutta scheme are shown in Fig. 2 for three periods. We use periodic boundary conditions; the number of waveguides exceeds the number of waveguides shown.

In the linear regime (Fig. 2a) the excitation is chosen as a Wannier state of the lowest band with Chern number \( C = -1 \) (see Fig. 1c) that uniformly populates all Bloch states within that band (see also Extended Data Fig. 2). This occupation is analogous to a low-temperature Fermionic system with the Fermi level in a bandgap, as is necessary for quantized pumping. Owing to diffraction, the wavefunction spreads during evolution, showing two dominant outer lobes similar to diffraction in a trivial array of straight waveguides. The bulk topological properties of the model are manifested in the transverse displacement of the centre of mass by \( C \) times the lattice constant (see also Fig. 2d).

Characteristic nonlinear behaviour (at a power at which we may clearly see the formation of discrete solitons) is displayed in Fig. 2b and c. In both cases, the excitation is a nonlinear eigenstate (that is, a soliton) of the instantaneous nonlinear Hamiltonian, which bifurcates from the lowest band and is found using Newton’s method. We point out that bifurcation from a band does not imply a uniform band occupation, but instead the projection of the soliton wavefunction onto the linear Bloch states is a strong function of power (see Methods and Extended Data Fig. 1). The difference between the pumped regime (Fig. 2b) and the trapped regime (Fig. 2c) is dictated by the amount of power injected into the system, \( gP/J_{\text{max}} = 1.9 \) and \( gP/J_{\text{max}} = 2.1 \), respectively. In both regimes, a suppression of spatial diffraction due to the focusing Kerr nonlinearity is observed. Although the shape of the wavefunction changes during one period, it remains strongly peaked and reproduces itself after each period, so that its shape resembles the shape of the input state. In the pumped regime (Fig. 2b), the wavefunction travels across the lattice with a displacement identical to the Chern number of the band from which it bifurcates (\( C = -1 \)). After each period, the soliton is displaced by three lattice sites (one unit cell), which can be clearly seen in Fig. 2d. At high power (Fig. 2c), the nonlinearly induced potential effectively decouples the waveguide and the wavefunction is trapped within the single site into which it was injected: the soliton’s centre of mass oscillates around one lattice site (Fig. 2d), but ends up in the starting position after each cycle. The solutions shown in Fig. 2b and c are the pumped and trapped solitons, respectively.

The displacement of the wavefunction’s centre of mass in the linear case is well understood and arises from a uniform occupation of a topologically nontrivial band. The same principle is responsible for quantized charge transport in Fermionic pumps even when disorder and interparticle interactions are present\(^{38}\). In stark contrast, here, the projection of the pumped soliton onto the complete set of linear Bloch states does not show a uniform occupation of any band; indeed, the occupation is a strong function of power. This means that the band occupation (to the extent that it is meaningful in this nonlinear system) is non-uniform and non-universal; thus quantized soliton motion is dictated by the band from which it bifurcates, rather than by the population within each band (see Methods, Extended Data Fig. 2 and Supplementary Video 1). Not surprisingly, propagation with the same input wavefunction but in the linear regime does not show quantized pumping. In the nonlinear regime we find numerically that the input soliton ‘tracks’ the instantaneous localized stable soliton solutions at every value of \( z \) during propagation. In other words, it behaves much as an eigenstate would in a linear time-varying, but adiabatic, system, despite being a fundamentally nonlinear entity. Strictly speaking, these solitons radiate because of the \( z \)-dependence of the Hamiltonian, but we find that the radiated intensity is proportional to \( \Omega^2 \) (see Methods, Extended Data Fig. 3) and therefore becomes negligible for \( \Omega \rightarrow 0 \). This agrees with the results of ref. \(^{39}\), which proposes this as a nonlinear version of the adiabatic theorem. Thus, we can analyse the pumped and trapped solitons in terms of instantaneous nonlinear eigenstates.

Although we focus here on the soliton that bifurcates from the lowest energy band, we show in Extended Data Fig. 2 that there also exist pumped solitons bifurcating from the higher bands that are displaced by the Chern number of the band from which they bifurcate. In a range of models tested thus far, for which a contiguous path exists, the number of unit cells pumped at a suitable degree of nonlinearity is the Chern number of the band from which the solitons bifurcate.

Mechanism of quantized nonlinear pumping

To explain the mechanism of pumping, we examine the position of the centres of mass of instantaneous solitons found at each time slice, \( z \), as shown in Fig. 3a−d. We see that in Fig. 3a, in the pumped regime \( (gP/J_{\text{max}} = 1.5) \), the soliton solutions follow a contiguous path through the lattice. For this power, we find a soliton solution (per unit cell), which is displaced along its path by one unit cell. For higher power \( (gP/J_{\text{max}} = 1.9) \), Fig. 3b), new nonlinear eigenstates emerge via
saddle-node nonlinear bifurcations. (For a classification of bifurcations see, for example, chapter 21 in ref. 40.) However, at this power, these bifurcations do not divert the original path of the soliton. In contrast, for still higher power \((gP/J)^{\text{max}} = 2.1, 2.5\), \(\text{Fig. 3c and d respectively}\), a pitchfork bifurcation of nonlinear eigenstates, associated with a spontaneous symmetry breaking, gives rise to the splitting of the path of the soliton’s centre of mass, causing it to return to the site from which it started at the beginning of the cycle. A depiction of the centres of mass of the bifurcating nonlinear solutions as a function of power are shown in \(\text{Fig. 3e, f, and a clear animation of this process can be seen in Supplementary Video 2}\). The detailed methods for obtaining \(\text{Fig. 3}\) are presented in the Methods.

As we have described, owing to the periodic Hamiltonian, the nonlinear eigenstates at the beginning and end of each pump cycle are identical, and, owing to translation invariance, exist for each unit cell. During adiabatic evolution, the soliton then tracks these eigenstates and when coming back to the beginning of a pump cycle, the soliton is forced to occupy the initial state either in the same unit cell or displaced by an integer number of unit cells. In our model the Hamiltonian takes on the same form after \(1/3\) \((2/3)\) of one full period, including a translation by one \((two)\) sites. This allows us to observe quantized pumping of single sites as well, in addition to integer unit cells.

Although no topological invariants are known for nonlinear systems, we can a posteriori define a topological invariant for nonlinear pumping. To that end, we define an extended unit cell such that the exponential tails of the localized soliton become negligible within it. We then solve for periodic evolution of the soliton therein and include the potential induced by the nonlinearity as a linear potential in the Hamiltonian. With this, we are able to calculate the Chern number (for details see Methods and Extended Data Fig. 4) for the band describing the soliton evolution, which is \(C = -1\) for the cases shown in \(\text{Fig. 3a, b and C = 0 for those in Fig. 3c, d}\). In Extended Data Fig. 5 we show such calculated Chern numbers for the soliton when tuning the system through a nonlinearly induced topological phase transition (by increasing

---

**Fig. 2** Linear and nonlinear propagation in topological Thouless pumps. a. Normalized amplitude of the discrete wavefunction, \(|\psi\rangle\), for a linear evolution over three periods for an input state with uniform excitation of the lowest band (chosen as a maximally localized Wannier state). It develops a discrete diffraction pattern, while its centre of mass is being pumped to the left by three unit cells after three periods. b. Nonlinear evolution for a pumped soliton with a degree of nonlinearity \((gP/J)^{\text{max}} = 1.9\). The excitation is an instantaneous nonlinear eigenstate (that is, a soliton) of the system. c. Same as b but with \((gP/J)^{\text{max}} = 2.1\) and showing a trapped soliton. d. Displacement of the centre of mass for the cases shown in a-c. The parameters for all figures are \(d = 22\ \mu\text{m}, \delta = 2\ \mu\text{m}, \alpha = -2\pi/12\) and \(\Omega = 2\pi/\text{L}\) with \(\text{L} = 8,000\ \text{mm}\).

---

**Fig. 3** Mechanism of nonlinear pumping. a–d. Centres of mass of available soliton solutions at each value of \(z\) in the pump cycle, showing contiguous paths. Black solid (dashed) lines indicate the position of the centre of mass for stable (unstable) nonlinear eigenstates of the instantaneous Hamiltonian.

---

**e, f.** Bifurcation diagrams for the nonlinear eigenstates at \(\Omega = \pi\) (e) and \(\Omega = 4\pi/3\) (f), as a function of power. Blue and red symbols label specific soliton positions of different branches, as shown in a–d. Each lattice consists of 30 sites and the parameters are the same as in Fig. 2.
In separate experiments, we observe nonlinear soliton pumping by one, two and three sites. Figure 4a–c shows the observed waveguide occupancies at the output facet. We use (time-averaged) optical input powers of $(P_I) = \{0.1, 2.0, 3.5, 5.0, 6.0\}$ mW, which we convert into $gP_J / g_{P_J}^{\max}$ for each individual experiment with $g = (0.07 \pm 0.01)$ mm$^{-1}$ per mW of (time-averaged) input power (see Methods and Extended Data Fig. 6). For low power ($gP_J / g_{P_J}^{\max} = 0.1$ and 0.2), linear diffraction is observed and the intensity spreads over several sites. The displacement of the centre of mass is not quantized, because the single-site excitation does not uniformly populate a band. For increasing input power, we observe strong localization of the wavepacket caused by soliton formation (blue arrow): this is the pumped regime. A further increase in the input power causes light to strongly localize in the waveguide into which it was injected (green arrow): this is the trapped regime. This transition from linear diffraction to a pumped soliton and finally to a trapped soliton is the experimental signature of quantized nonlinear pumping.

For the corresponding simulations in Fig. 4d–f, we scaled the coupling function using the linear propagation in the individual array and included realistic optical losses (measured to be approximately 0.7 dB cm$^{-1}$). Owing to the losses, the power thresholds to observe the pumped and trapped soliton are higher than those in the idealized lossless case (Figs. 2 and 3). In Fig. 4, simulations that take into account such loss agree well with the experimental results. The lower contrast in the experiment arises from the fact that the tails of the pulse (in time) propagate linearly and diffract, while only the region of high power in the temporal centre of the pulse is affected by strong nonlinearity$^{44}$. Nonetheless, the pumped soliton and a transition to a trapped soliton is clearly observed in simulation and experiment. Finally, in Extended Data Fig. 8 we also experimentally confirm our theoretical prediction that the trapping of the soliton is due to a pitchfork bifurcation (see also Methods).

**Online content**

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-03688-9.
Methods

Soliton bifurcations

Solitons are localized solutions to the time-independent nonlinear Schrödinger equation, and thus do not change during evolution. They have been observed in a wide variety of platforms, such as arrays of evanescently coupled waveguides. The solitons presented in the main text bifurcate from the lowest-energy band. Indeed, we also find solitons bifurcating from higher-energy bands. Extended Data Fig. 1a shows characteristic bifurcation behaviour of solitons for the AAH model at $z = 0$ (parameters used are identical to Figs. 2 and 3). For low power, the nonlinear eigenvalues are close to the lower band edge from which the solitons bifurcate, and they move deeper into the bandgap with increasing power. With increasing separation from the band edges, the soliton localization increases (Extended Data Fig. 1b). We also examine the projection of the soliton wavefunction onto each linear Bloch state $|\phi_p\rangle$ of the system (and hence the population of each band).

Importantly, the degree of occupation is a strong function of power. Extended Data Fig. 1c depicts the typical behaviour for the soliton that bifurcates from the lowest band. For low power ($gP/J^{\text{max}} \ll 1$), the soliton occupies mostly the energetically lowest-lying state in the band from which it bifurcates. In this regime, the occupation is very clearly non-uniform and therefore different from the occupation required for quantized pumping in a linear Thouless pump. The occupation of eigenstates spreads out more with increasing power, but is never perfectly uniform. This can be clearly seen in the two upper panels in Extended Data Fig. 1c, which depict the occupation of the linear states for the pumped solitons shown in Extended Data Fig. 2g and Fig. 2b (with powers of $gP/J^{\text{max}} = 0.2$ and 1.9, respectively).

Quantized pumping of solitons bifurcating from all bands

In the main text, we show quantized nonlinear pumping for a soliton bifurcating from the lowest band with $gP/J^{\text{max}} = 1.9$. Here, we show that (1) quantized nonlinear Thouless pumping is not restricted to systems with nonlinear effects of the order of the hopping strength, but can also be observed at lower powers; (2) there exist solitons bifurcating from higher bands that are pumped by the respective Chern number of that band; and (3) the occupation of the linear eigenstates is strongly non-uniform over the pumping cycle, in stark contrast to linear Thouless pumping.

All three points are summarized in Extended Data Fig. 2 where we additionally contrast quantized linear and nonlinear Thouless pumping. We use the same model as in the main text: an off-diagonal AAH model with a three-site unit cell and three bands with Chern numbers $C = (-1, 2, -1)$ (see Extended Data Fig. 2d). In the linear regime (Extended Data Fig. 2a–c) the excitation is chosen to be a maximally localized Wannier function for a single band. This ensures uniform band occupation. We then numerically evolve the initial Wannier function over three periods and choose the modulation frequency such that the soliton comes back to its original position (Extended Data Fig. 4c). The $z$ axis then creates the second periodic dimension needed to define a conventional Brillouin zone. Once the wavefunction $\phi_p(z)$ of the pumped soliton is known for all $z$ within one unit cell it can be treated as a linear (sliding) potential $V(z) = gP\phi_p(z)^2$ for the supercell. Owing to the spatial periodicity of the supercell, we are able to define the wavevector $k_s$ (see upper panel in Extended Data Fig. 4d). The system is now somewhat akin to the original model of pumping by Thouless whereby the energy eigenvalues of one band are identical to the nonlinear eigenvalue of the soliton. This allows for the identification of the band describing the soliton motion and the eigenstates of this band (for example, the lowest one for the soliton described in the main text), are described by the soliton wavefunction itself. The corresponding band structure has $3m$ bands and is shown in Extended Data Fig. 4d. Using ref. 47, we numerically calculate the Chern number of that band which describes the motion of the pumped soliton.

Linear and nonlinearly induced topological phase transition

The nonlinear off-diagonal AAH model allows for the examination of soliton propagation while tuning through two different types of topological phase transitions, as follows: (1) a nonlinearly induced topological phase transition in which an increase in the soliton power (increasing nonlinearity) leads to a transition from a pumped to a trapped soliton. In this case the underlying linear model is not changed and therefore also the Chern numbers of the linear bands (from which the soliton bifurcates) are unchanged; and (2) a linearly induced topological phase transition in which the power of the soliton remains unchanged but the pump model is altered (that is, by changing the hoppings) such that it undergoes a linear topological phase transition that changes the Chern numbers of the linear bands.

Confirming adiabatic behaviour in the nonlinear system

To show that nonlinear Thouless pumps, similarly to linear pumps, have an adiabatic regime and therefore show perfectly quantized pumping in the adiabatic limit, we numerically evaluate the intensity radiated by the soliton during propagation. We use a system of 180 sites and numerically calculate the remaining intensity with absorbing boundary conditions on 40 waveguides at each end. Extended Data Fig. 3 illustrates that the absorbed intensity is approximately proportional to the square of the driving period, $P^2$, and therefore that an adiabatic regime exists.

Chern number calculation of the nonlinearly induced potential

As mentioned in the main text, the topological invariant describing a traditional (linear) Thouless pump is the Chern number, which is calculated using Bloch states. For the nonlinear case, no topological invariants are known. Owing to the broken translation invariance, no natural notion of a band structure exists. Hence, a novel type of description is needed. Here, we calculate the Chern number associated with the motion of the pumped soliton a posteriori, after the propagation is known. It therefore has no predictive power but allows us to characterize the motion of the soliton with respect to the underlying linear topology of the pump model. As such, it provides a consistent way of defining a Chern number for nonlinear propagation.

The sequence of steps to calculate the Chern number for nonlinear soliton propagation is shown in Extended Data Fig. 4a. Owing to the exponential confinement of the soliton, we define a spatial supercell (see Extended Data Fig. 4b) containing $m$ original unit cells in which the soliton is nearly entirely localized. The only requirement for $m$ is that it must be sufficiently large for the exponential tails of the soliton to be negligible. We then solve for the time evolution of the pumped soliton within the spatial supercell using periodic boundary conditions such that the soliton comes back to its original position (Extended Data Fig. 4c). The $z$ axis then creates the second periodic dimension needed to define a conventional Brillouin zone. Once the wavefunction $\phi_p(z)$ of the pumped soliton is known for all $z$ within one unit cell it can be treated as a linear (sliding) potential $V(z) = gP\phi_p(z)^2$ for the supercell. Owing to the spatial periodicity of the supercell, we are able to define the wavevector $k_s$ (see upper panel in Extended Data Fig. 4d). The system is now somewhat akin to the original model of pumping by Thouless whereby the energy eigenvalues of one band are identical to the nonlinear eigenvalue of the soliton. This allows for the identification of the band describing the soliton motion and the eigenstates of this band. (For example, the lowest one for the soliton described in the main text), are described by the soliton wavefunction itself. The corresponding band structure has $3m$ bands and is shown in Extended Data Fig. 4d. Using ref. 47, we numerically calculate the Chern number of that band which describes the motion of the pumped soliton.
The first case of a nonlinearly induced topological phase transition is illustrated in Figs. 2 and 3. There, we show how increasing the power \((gP/J)^{\text{max}}\) leads to a transition from a pumped soliton to a trapped soliton via spontaneous symmetry-breaking bifurcations. Owing to its power dependence, this transition is unique to nonlinear systems. Extended Data Fig. 5a shows the Chern number \(C\) (for details of the calculation, see Methods and Extended Data Fig. 4) associated with the soliton for increasing power. In the pumped regime, the Chern number of the soliton is \(C=−1\) and in the trapped regime, it is \(C=0\). Between the pumped and the trapped regime there exists a small topological phase transition region in which no Chern number can be calculated because the propagation of the soliton is not periodic. These results show that the transition from the pumped to the trapped soliton can be associated with a topological phase transition induced by nonlinearity.

For the sake of comparison, we also examine the second case of solitons in a lattice that is undergoing a linearly induced topological phase transition. Such a topological phase transition is accompanied by an energy gap closing and opening in which the Chern numbers of the bands change. To tune the off-diagonal AAH model through a (standard linear) topological phase transition, we parametrize the \((\text{dimensionless})\) nearest-neighbour couplings in the following way:

\[
J_n(z) = J + gP J K \cdot \cos \left( \frac{2n}{3} \zeta + 2\pi/6 \right).
\]

We choose \(K=1\) and for \(J>0.25\) the system is topologically equivalent to the model described in the main text, having Chern numbers of \(C=[1,2,−1]\). For \(J=0.25\) a gap closing occurs, changing the Chern numbers to \(C=[2,−4,2]\) for \(J=0.25\) (ref. 49). We calculate the Chern number associated with the soliton propagation for a soliton (with \(gP/K=0.7\)), which bifurcates from the lowest linear energy band, while sweeping downward from \(J=0.7\) to \(J=0\) and therefore through a linear topological phase transition. Extended Data Fig. 5b shows that the soliton is pumped by the Chern number of the band from which it bifurcates and changes its behaviour at the topological phase transition point, when the Chern number of the lowest band itself changes from \(-1\) to \(+2\). Additionally, in Supplementary Video 3 we visually show the instantaneous solutions of the soliton that bifurcated from the lowest band (with \(gP/K=0.7\)), its centre of mass and the band structure over one cycle, while sweeping from \(J=0.7\) to \(J=0\).

### Nonlinear waveguide characterization

The simplified schematic of the experimental setup used to carry out the nonlinear experiments is depicted in Extended Data Fig. 6a. A commercial Yb-doped fibre laser (Menlo BlueCut) emits pulse trains at a repetition rate of 5 kHz (tunable). The pulses have a temporal width of approximately 260 fs and are spectrally centred at 1,030 nm. We adjust the power using a half-wave plate together with a polarizing beam splitter. The parallel gratings \(G_1\) and \(G_2\) (Thorlabs GR25/0610i) down-chirp and temporally stretch the pulses to 2 ps; see also ref. 49. The 2-ps pulses are focused into the single waveguides using lens \(L_1\) (Thorlabs C280 TMD-B). Lenses \(L_1\) (Thorlabs AC064-015-B-ML) and \(L_2\) (Thorlabs LB1811-B-ML) are used to image the output facet onto a complementary metal–oxide–semiconductor (CMOS) camera (Thorlabs DCC1545M), while the light is simultaneously injected into a fibre-coupled optical spectrum analyser (Anritsu MS9740 A). The simplified schematic neglects additional mirrors and neutral density filters in the setup. We use a photodiode power sensor (Thorlabs S120C) to measure the time-averaged input and output power before \(L_1\) and after \(L_2\), respectively. Extended Data Fig. 6b shows the observed dependence between input and output power for two single, separated waveguides. The linear relationship, visible via the linear fit, suggests that no nonlinear losses, for example, via multi-photon absorption, occur, and therefore such losses can be neglected in our experiments.

The pulses are temporally stretched in order to avoid substantial generation of new wavelengths via self-phase-modulation in the sample. This is essential; if the spectrum is too broad, then the coupling constants between waveguides are not well defined. Extended Data Fig. 6c shows the spectral density of the output intensity after propagation through a 76-mm-long waveguide array (straight waveguides with equal separation of 24 μm). Up to 6 mW input power, which is the largest power used in the experiment, 76% of the intensity (which is equivalent to the intensity found within the FWHM of a Gaussian) is found within a spectral range of 20 nm. The coupling variation due to self-phase-modulation is therefore of the order of the intrinsic coupling uncertainty (see Extended Data Fig. 7) and we can neglect this effect and use equation (1) to theoretically describe our experiments.

To relate the (experimental) time-averaged input power to the nonlinearity used in simulations, \(gP\) (see equation (2)), we launch light into single waveguides in an array of straight waveguides with 24 μm separation. The low hopping constant \((J=0.01\ \text{mm}^{-1})\) at this separation allows us to observe a change of the spatial output intensity pattern for input powers \((P_{\text{in}})\) below 1 mW, for which changes in the spectral shape are minimal (see also Extended Data Fig. 6c). At higher input powers above 1 mW, fitting results are distorted by the (temporal) tails of the pulse, which are governed by linear diffraction. For each input power, we find the corresponding \(gP\) via fitting (least-squares) the numerically obtained output intensity pattern using equation (2) and including propagation losses. The dependence of \(gP(\zeta=0)\) on the input power is shown in Extended Data Fig. 6d for five datasets, resulting in a value of \(g=0.07\pm0.01\ \text{mm}^{-1}\) per mW of (time-averaged) input power.

### Linear waveguide characterization

We measure the coupling strength \(J(d, \lambda)\) as a function of separation, \(d\), between waveguides and wavelength, \(\lambda\). The waveguide arrays are characterized using a commercial supercontinuum source (NKT SuperK COMPACT) with a subsequent filter (NKT SuperK SELECT) to select the desired excitation wavelength. The light is focused into single waveguides in three separate arrays of straight waveguides \((d=20\ \mu m, 22\ \mu m\text{ and }24\ \mu m)\) with 30 waveguides each. To avoid edge effects, we focus only into the central 22 waveguides for \(d=22\ \mu m\) and \(d=24\ \mu m\) and into the central 12 waveguides for \(d=20\ \mu m\). The experimentally obtained output intensity pattern for each excitation is fitted with the theoretical output pattern in waveguide \(n\), given by \(\phi_n(\lambda) = (i)^n\phi_0(\zeta=0)B_n(2\lambda)\) via least-squares fitting. Here, \(L\) is the sample length of 76 mm, \(B_n\) is the Bessel function of first kind of order \(n\), \(\phi_0(\zeta=0)\) is the amplitude of the excited waveguide \((n=0)\) at \(\zeta=0\). Extended Data Fig. 7 shows the mean coupling constants with one standard deviation. With increasing wavelength the optical waveguide mode width increases, leading to a larger overlap between neighbouring modes and therefore higher coupling constants. The linear fits show that within a wavelength range of ±10 nm around a central wavelength of \(\lambda_c=1,030\ \text{nm}\) the relative change \(\Delta J/\Delta J(\lambda_c)\) in the coupling constant is \(\Delta J/\Delta J(\lambda_c) = 5\%, 6\%\) and \(7\%\) for \(d=20\ \mu m, 22\ \mu m\) and \(24\ \mu m\), respectively. This change is comparable to the relative coupling uncertainty of about 7%, 4% and 10% at the central wavelength of \(\lambda_c=1,030\ \text{nm}\). At \(1,030\ \text{nm}\), fitting an exponential function, we obtain \(J(\lambda_c) = 27.21\ \text{exp}(−330\ \text{mm}^{-1} \cdot \lambda)\), with \(d\) in millimetres. In the simulations in Fig. 4d–f, we account for the remaining uncertainty in the coupling by using an additional multiplicative fitting factor for the coupling function, which is evaluated by fitting the linear output intensity pattern. The factors are 1.06, 1.28 and 1.12 for Fig. 4d, e and f, respectively.

### Obtaining solitons and linear stability analysis

The instantaneous nonlinear eigenstates (solitons) in Fig. 3 are obtained using a Newton iteration scheme. For that purpose, we use the FindRoot function in Mathematica® (with a precision of more than 15 digits). This method depends critically on the initial ansatzes. At high power \((gP/J)^{\text{max}}=5\), we use six different ansatzes to evaluate the
relevant soliton eigenstates, belonging to six different branches. Three of the initial ansatzes are localized on a single site, while the other three ansatzes are localized between two sites with equal intensity in two neighbouring sites. We then iteratively use the solitons at higher power as ansatzes for lower power. Additionally, we confirm the convergence by calculating the difference between the soliton obtained via the Newton method, $\Phi_n^{\text{soliton}}$, and the eigenvector $\Phi_n^{ev}$ with the highest overlap, obtained from the nonlinear Hamiltonian $H_{\text{non-lin}} = H_{\text{in}} - g|\Phi_n^{\text{soliton}}|^2$ in the following sense: $\sum_{m\neq n} |\langle \Phi_m^{\text{soliton}} | \Phi_n^{\text{soliton}} \rangle|^2 < 10^{-15}$.

Assuming a static (for example, $z$-independent) Hamiltonian, a system prepared in an unstable nonlinear eigenstate does not guarantee a static evolution of the system, because small fluctuations (even those caused by the limits of numerical precision) can amplify, and thus the soliton has a finite lifetime. We test the stability of solitons using linear stability analysis, which indicates when solitons are linearly unstable. We follow ref. $50$ and start with a nonlinear eigenstate $\Phi^{(0)}(z) = e^{-i\kappa z} \Phi^{(0)}$ for the nonlinear Schrödinger equation:

$$i \frac{d}{dz} \Phi = H \Phi - g|\Phi|^2 \Phi,$$

(4)

where $H$ is the linear tight-binding Hamiltonian of the system, the parameter $g$ describes the strength of the nonlinearity, $\Phi$ is the (discrete) wavefunction and $\Lambda$ describes the eigenvalue of the eigenstate $\Phi^{(0)}$. Compared to equation (1), we have dropped the subscripts, but it is clear that this equation should be understood as a tight-binding matrix equation.

To test for stability, we take the following ansatz for a small perturbation around the solution: $\Phi = e^{-i\varepsilon \kappa z} (\Phi^{(0)} + \varepsilon (\nu(z) + i\omega z))$. Plugging this ansatz into equation (4) and using the fact that $\Phi^{(0)}(z)$ solves the equation, we arrive (after some algebra) to first order in $\varepsilon$ at the following two equations for the real and the imaginary part of the perturbation:

$$\frac{d}{dz} \nu = (-\Lambda + H - 2g|\Phi^{(0)}|^2 + g(\Phi^{(0)})^2) w \equiv L_{w},$$

$$\frac{d}{dz} \omega = (-\Lambda + H - 2g|\Phi^{(0)}|^2 - g(\Phi^{(0)})^2) w \equiv -L_{w}.$$

(5)

These coupled equations are solved by separating the $z$-dependence via $\nu = e^{i\kappa z} \nu \nu$ and $w = e^{i\kappa z} w$, which leads to $\Lambda w = -L w$ and $\Lambda w = L w$. The values of $\kappa^2$ can now be calculated as eigenvalues of the matrix $-L L$. If $\kappa^2$ is positive, then $\kappa$ is real and the perturbations can build up exponentially. In contrast, if $\kappa$ is negative, then $\kappa$ is imaginary and the perturbations are oscillating waves that do not grow exponentially. In the latter case the soliton is linearly stable. In the main text, we identify solitons as stable if $\kappa^2 < 10^{-15}$.

**Measured power dependence of the pitchfork bifurcation**

We experimentally analyse our theoretical prediction that the trapping of the soliton is due to a pitchfork bifurcation (as shown in Fig. 3e). We measure the output power as a function of input power. Once the pumped soliton appears, a further increase of the input power leads to an increase of the peak power of the pumped soliton (when the wavefunction is normalized to the total power) owing to the soliton’s stronger confinement. For a sufficiently high input power, the pitchfork bifurcation is triggered. At this point, the power in the waveguide where the pumped soliton is localized decreases relative to that in the injected waveguide. This point marks the onset of trapping, and thus the appearance of the pitchfork bifurcation. We can test the dependence of this threshold power on the spatial modulation strength $\delta$. We measure this power for different arrays with varying $\delta$, as well as in different unit cells within the same waveguide array (with the same $\delta$). In Extended Data Fig. 8, we plot the experimentally obtained output power required to observe the onset of trapping, as a function of the degree of waveguide modulation, $\delta$. We compare this to the numerically obtained power for which the pitchfork bifurcation occurs. To directly compare the two, we normalize the power to unity at $\delta = 1.5 \mu m$; on this basis, clear agreement between theory and simulation is observed.

**Data availability**

The data that support the findings of this study are available from the corresponding author on reasonable request.

43. Eisenberg, H., Silberberg, Y., Morandotti, R., Boyd, A. & Aitchison, J. Discrete spatial optical solitons in waveguide arrays. *Phys. Rev. Lett.* **81**, 3383 (1998).

44. Christodoulides, D. N. & Joseph, R. I. Discrete self-focusing in nonlinear arrays of coupled waveguides. *Opt. Lett.* **13**, 794–796 (1988).

45. Fleischer, J. W., Segev, M., Eberly, J. & Christodoulides, D. N. Observation of two-dimensional discrete solitons in optically induced nonlinear photonic lattices. *Nature* **422**, 147–150 (2003).

46. Lederer, F., Stegeman, G. I., Christodoulides, D. N., Assanto, G., Segev, M. & Silberberg, Y. Discrete solitons in optics. *Phys. Rep.* **463**, 1–126 (2008).

47. Fukui, T., Hatsugai, Y. & Suzuki, H. Chern numbers in discretized Brillouin zone: efficient method of computing (spin) Hall conductances. *J. Phys. Soc. Jpn* **74**, 1674–1677 (2005).

48. Ke, Y. et al. Topological phase transitions and Thouless pumping of light in photonic waveguide arrays. *Laser Photon. Rev.* **10**, 985–1001 (2016).

49. Wolfram Research, Inc. *Mathematica*. v.12.0 (2019).

50. Kevrekidis, P. G. The Discrete Nonlinear Schrödinger Equation: Mathematical Analysis, Numerical Computations and Physical Perspectives Vol. 232 (Springer Science & Business Media, 2009).

Acknowledgements We acknowledge discussions with A. Cerjan, S. Gopalakrishnan, D. Leekam, O. Zilberberg and P. Kevrekidis. We acknowledge support from the ONR Young Investigator programme under award number N00014-18-1-2595, the ONR-MURI programme N00014-20-1-2325 and the Packard Foundation fellowship, under number 2017-66821. M.J. acknowledges support of the Verner M. Hillman Distinguished Graduate Fellowship at the Pennsylvania State University. Some numerical calculations were performed on the Pennsylvania State University’s Institute for Computational and Data Sciences’ Roar supercomputer.

**Author contributions** M.C.R. initiated and supervised the project; M.J. fabricated and built the experimental beam shaping and fabrication setup; theoretical investigation was carried out by M.J.; M.J. and M.C.R. wrote the manuscript with input from S.M.

**Competing interests** The authors declare no competing interests.

**Additional information**

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41586-021-03688-9.

**Correspondence and requests for materials** should be addressed to M.J. or M.C.R.

**Peer review information** Nature thanks the anonymous reviewers for their contribution to the peer review of this work.

**Reprints and permissions information** is available at http://www.nature.com/reprints.
Extended Data Fig. 1 | Analysis of soliton bifurcation. a, Energy eigenvalues for the linear eigenstates at $z = 0$ (black) showing three bands. Nonlinear eigenvalues for solitons bifurcating from the upper (middle, lower) band are shown in green (blue, red) and are decreasing with increasing power. b, Degree of soliton localization as measured by the Inverse Participation Ratio $\text{IPR} = \sum_i |\psi_i|^4 / \left( \sum_i |\psi_i|^2 \right)^2$ for the solitons shown in a. c, Spectral overlap coefficients between the soliton bifurcating from the lowest band and the linear energy eigenstates are a strong function of power. The two upper panels show the occupation at $gP_J / J_{\text{max}} = 0.2$ and $1.9$ (corresponding to Extended Data Fig. 2g and Fig. 2b, respectively). Note the non-uniform occupation of the lowest band (from which the soliton bifurcates), which is particularly pronounced at low power.
Extended Data Fig. 2 | Higher Chern number pumping and band occupation. a−c, Numerically calculated \( \rho_z \)-evolution for an initial excitation of a maximally localized Wannier state of the upper (middle, lower) band (shown in a (b, c), respectively). The displacement of the centre of mass is dictated by the Chern number of the occupied band. The insets show the projection \( \langle \phi(z) | \phi_{\text{lin}}(z) \rangle \) of the propagating wavefunction \( \phi(z) \) onto the instantaneous linear Bloch states \( \phi_{\text{lin}}(z) \) of the lower (middle, upper) band, ordered from left to right by increasing energy eigenvalue. d, Band structure of the off-diagonal AAH model, showing three bands with Chern numbers \( C = \{-1, 2, -1\} \). End states are shown in red. e−g, Similar to a−c, but with an initial excitation of a nonlinear eigenstate bifurcated from the upper (middle, lower) band (shown in e (f, g), respectively) with \( gP/\gamma_{\text{max}} = 0.2 \). The displacement of the centre of mass is identical to the Chern number of the band from which the soliton bifurcates. The insets show a strongly non-uniform occupation of the linear Bloch states. Parameters for the system are chosen to be identical to those in Fig. 2, except \( L = 4 \times 10^3 \text{ mm} \) with a size of 900 sites for a−c and \( L = 8 \times 10^3 \text{ mm} \) for e−g.
Extended Data Fig. 3 | Adiabatic behaviour in the nonlinear system.
Absorbed intensity (relative to the total intensity) during one driving period in relation to the driving frequency. Blue circles are numerical values and the red line has a slope of $-2$. The parameters for the simulation are 180 sites with absorbing boundary conditions using 40 sites at each end, $d = 24 \mu m$, $d_0 = 2 \mu m$, $\alpha_0 = -2\pi/12$ and $gP J_{max}^3 = 1.9$. 
Extended Data Fig. 4 | Calculation of the Chern number for the pumped soliton. a. Sequence of steps to calculate the Chern number for the nonlinear soliton propagation. b. Spatial supercell for a pumped soliton composed of \(m\) unit cells with periodic boundary conditions (PBC). c. \(z\)-evolution of the discrete soliton wavefunction \(\phi_n(z)\) in a supercell with \(m = 4\) unit cells and \(m\) periods forming a two-dimensional periodic supercell. d. Band structure (lower panel) with 3\(m\) bands calculated using the supercell (upper panel). The eigenstates of one band (here, the lowest) describe the motion of the soliton and its Chern number \(C\) can be calculated in the conventional way (here, \(C = -1\)).
Extended Data Fig. 5 | Topological phase transitions. **a**, Nonlinearly induced topological phase transition. The Chern number associated with the soliton is calculated for increasing power $gP J / \text{max}$. In the grey area, no contiguous path for an adiabatic soliton evolution is found. The red line indicates the Chern number of the lowest band in the linear model from which the soliton bifurcates. **b**, Linearly induced topological phase transition. The Chern number associated with the soliton is calculated as a function of decreasing hopping strength $J$. The red line indicates the linear Chern number of the band from which the soliton bifurcates. The topological phase transition occurs at $J = 0.25$. 
Extended Data Fig. 6 | Nonlinear waveguide characterization. a, Simplified, schematic illustration of the experimental setup, including a half-waveplate (WP) together with a polarizing beam splitter (PBS) to adjust the power of the emitted laser pulses. Two gratings (G1, G2) temporally stretch the pulse to 2 ps. Lens L1 focuses the pulses into single waveguides within the waveguide array. Lenses L2 and L3 image the output facet onto a CMOS camera. Simultaneously, using a further beamsplitter (BS), the light is additionally coupled into a fibre and its spectrum is analysed with an optical spectrum analyser (OSA).
b, Measured input power to output power dependence for two datasets, showing no nonlinear losses due to multi-photon absorption. The black line indicates a linear fit.
c, Spectral distribution of the pulse after propagation of 76 mm in a lattice of straight waveguides with equal separation of 24 μm. The white lines denote the spectral range in which 76% of the intensity (equivalent to FWHM of a Gaussian) are found.
d, Theoretical nonlinear parameter $g_P(z=0)$ versus the experimental time averaged input power. Depicted are values from five different waveguides in a lattice of straight waveguides with a separation of 24 μm. The black line indicates the value of $g = (0.07 \pm 0.01)$ mm$^{-1}$ per mW (time-averaged) input power as the mean value with one standard deviation.
Extended Data Fig. 7 | Characterization of the coupling strength. Coupling constants $J(d, \lambda)$ for wavelength $\lambda$, extracted from waveguide lattices with straight waveguides and equal separation, $d$, between the waveguides. The errorbars show one standard deviation.
Extended Data Fig. 8 | Parameter dependence of pitchfork bifurcation.

Input power (relative to the mean input power at modulation $\delta = 1.5 \mu m$) required for maximum relative intensity in the pumped soliton as a function of spatial modulation strength $\delta$ of the waveguides. Dots in colour are measurements, black lines are the respective mean values with one standard deviation. The black dotted line shows the numerically obtained threshold power for the pitchfork bifurcation point.