nature physics

Article

https://doi.org/10.1038/s41567-022-01871-x

Quantized fractional Thouless pumping of solitons

Received: 14 January 2022

Accepted: 4 November 2022

Published online: 12 January 2023

Check for updates

Marius Jürgensen [1], Sebabrata Mukherjee [1,2], Christina Jörg [1] & Mikael C. Rechtsman [1]

In many contexts, interaction between particles gives rise to emergent phenomena. An example is the fractional quantum Hall effect, where the interaction between electrons leads to fractionally quantized Hall conductance. In photonic systems, the nonlinear response of an ambient medium mediates the interaction between photons, and, in the mean-field limit, these dynamics are described by the nonlinear Schrödinger (also called Gross–Pitaevskii) equation. It was recently shown that at weak nonlinearity, soliton motion in nonlinear Thouless pumps—a dimensionally reduced implementation of a Chern insulator—could be quantized to the Chern number, because solitons track the single-band Wannier function throughout the pumping cycle. Here using arrays of coupled optical waveguides, we show that a sufficiently strong nonlinearity fractionally quantizes the motion of solitons. Specifically, we find that the soliton follows maximally localized multi-band Wannier functions and therefore returns to itself only after multiple cycles of the Thouless pump—but displaced by an integer number of unit cells—leading to a rich fractional plateau structure describing soliton motion. Our results represent an example of emergent behaviour in topologically non-trivial systems in the presence of interactions.

In the integer quantum Hall effect, a gas of non-interacting electrons can be rigorously shown to exhibit integer-quantized Hall conductance[1–3], fixed to a topological invariant—the Chern number. It was therefore a surprise when plateaux of fractional conductance appeared in the experiment of Tsui, Stormer and Gossard[4]. There, the interaction between electrons played the key role, giving rise to the formation of fractionally charged quasiparticles[5–7]. In addition to electronic systems, topological phenomena have also been predicted and observed in bosonic systems, including for microwave photons[8], optical photons[9–11], ultracold atoms[12–14] and mechanical systems[15,16], among others. Compared to their electronic counterparts, photonic systems offer rich design flexibility and new physics[17], such as the inclusion of non-Hermiticity[18], arbitrary driving and nonlinearity[19]. However, photons interact extremely weakly in free space, so photon–photon interactions must be mediated by an ambient medium. The first experiments creating photonic two-particle Laughlin states used repulsive interactions mediated by Rydberg atoms in a twisted cavity[20]. A conceptually different approach has been taken by treating the interactions of many photons in the mean-field limit using nonlinearity. This approach has led to the prediction and observation of various new topological phenomena, such as topological bulk[21,22] and edge solitons[23–25], nonlinearly induced topological insulators[26], and other non-Hermitian and nonlinear phenomena[27,28].

One class of topological models, specifically suitable for studies in photonics due to the design flexibility available in fabricated structures, are Thouless pumps[29,30]. These models are 1 + 1-dimensional reductions of Chern insulators, which use a time-periodic modulation to emulate a wavevector dimension of the related two-dimensional Chern insulator. Given a uniform band occupation and adiabatic driving, the displacement per period in a Thouless pump is dictated by

[1]Department of Physics, The Pennsylvania State University, University Park, PA, USA. [2]Department of Physics, Indian Institute of Science, Bangalore, India. e-mail: marius@psu.edu; mcrworld@psu.edu

https://doi.org/10.1038/s41567-022-01871-x

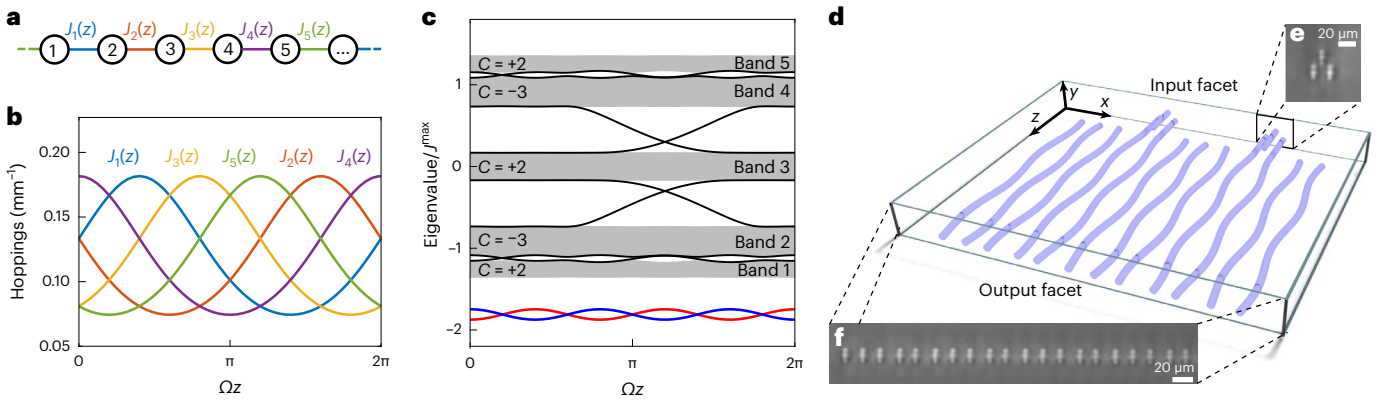


Fig. 1 | Model for quantized fractional Thouless pumping. a, Illustration of the off-diagonal AAH model, with five sites per unit cell and z-dependent hoppings $J_n(z)$ between nearest-neighbour sites. **b**, Modulation of the hopping strength over one period. **c**, Bandstructure (instantaneous energy eigenvalues) of the Hamiltonian, showing five bands (grey) with Chern numbers $C = \{2, -3, 2, -3, 2\}$, end states (black lines) crossing the bandgaps, and nonlinear eigenvalues (red and blue lines) of instantaneous solitons that are pumped by a fraction $f = -1/2$. **d**, Schematic of the implementation of the model in arrays of evanescently coupled waveguides. Only two waveguides per unit cell extend to the input facet, with an additional waveguide fabricated on top to transform a single-site excitation into an effective two-site excitation. **e**,**f**, White-light micrographs showing the input (**e**) and output (**f**) facets.

the Chern number of the occupied band. Thouless pumps have been studied in various systems (for example, refs. 31–38). The inclusion of nonlinearity into Thouless pumps has recently led to the discovery of quantized pumping of solitons: throughout the pumping cycle, the soliton's position follows the instantaneous single-band Wannier function[39,40] of the band, from which it bifurcates at low power. Its displacement is thus given by the relevant band's Chern number, despite lacking the notion of a uniformly filled band[41]. Quantization stems from the fact that the soliton comes back to itself—modulo a translation by an integer number of unit cells due to translation invariance—after each period.

In this Article we theoretically predict and experimentally demonstrate quantized fractional Thouless pumping of solitons, where the strength of the nonlinearity exceeds the relevant bandgap. We experimentally observe fractional soliton pumping by a fraction of $f = -1/2$ in arrays of evanescently coupled waveguides with Kerr nonlinearity. Fractional pumping occurs as the soliton follows the maximally localized multi-band Wannier functions and returns to itself—modulo a translation by an integer number of unit cells—only after multiple periods. Finally, we numerically show how tuning the strength of the nonlinearity leads to multiple plateaux of integer and fractionally quantized displacement within one Thouless pump model.

The nonlinear Aubry–André–Harper model

Experimentally, we realize a nonlinear Thouless pump by focusing high-peak-power laser pulses into arrays of single-mode evanescently coupled waveguides. Due to the Kerr effect, the refractive index becomes intensity-dependent, and the propagation of photons in the system is described by the discrete nonlinear Schrödinger equation[42–48]

$$i\frac{\partial}{\partial z}\psi_m(z) = \sum_n H_{mn}(z)\psi_n(z) - g|\psi_m(z)|^2\psi_m(z) \qquad (1)$$

where m,n denote lattice sites, $\psi_n(z)$ is the amplitude of the wavefunction at propagation distance z for site n, $H_{mn}(z)$ is a z-dependent tight-binding Hamiltonian describing a topological Thouless pump, and $g > 0$ describes the strength of the focusing Kerr nonlinearity. In waveguide systems, z plays the role of a temporal coordinate. Equation (1) is also known as the Gross–Pitaevskii equation, which describes interacting bosons in a Bose–Einstein condensate in the mean-field limit[49,50]. Therefore, our results are not restricted to photonics, but hold for a range of interacting and nonlinear bosonic systems[51–55]. In waveguides, the nonlinearity describes an effective interaction between photons mediated by the ambient material. To treat experiment and theory on the same footing, we define $P = \sum_n|\psi_n(z=0)|^2$, and refer to the strength of nonlinearity as a dimensionless quantity, gP/J^{max}, where J^{max} is the largest hopping value in the Hamiltonian.

We illustrate fractional Thouless pumping in an off-diagonal Aubry–André–Harper (AAH) model[31,56–58] with five sites per unit cell and zero on-site detuning (Fig. 1a). Its nearest-neighbour couplings $J_n(z)$ are periodically modulated in z, and the Hamiltonian is given by $H_{mn}(z) = -J_m(z)\delta_{n,m+1} - J_{m-1}(z)\delta_{n,m-1}$. Figure 1b shows the strength of the hoppings over one period, as used in the experiment. The bandstructure of this model is depicted in Fig. 1c and has five bands with Chern numbers $C = \{2, -3, 2, -3, 2\}$. A schematic illustration of the implementation in an array of evanescently coupled waveguides is shown in Fig. 1d, where the periodic modulation of the distance between neighbouring waveguides changes the evanescent hopping strength. In the experiment the position of waveguide n is $x_n(z) = nd + \delta\cos(\Omega z + \frac{4\pi}{5}n - \frac{6\pi}{20})$, where d defines the average separation between waveguides, δ is the spatial modulation strength, and Ω is the modulation frequency. Throughout this work, we use $d = 17.25$ μm and $\delta = 1$ μm. Figure 1d also schematically shows an input region, where only two waveguides per unit cell are extended all the way to the input facet with an additional waveguide on top (Fig. 1e). We use this 'triple coupler' to transform a single-site excitation of the upper waveguide into an effective two-site excitation of the two lower waveguides. White-light images of the input and output facets are shown in Fig. 1e,f, respectively.

Quantized fractional Thouless pumping

Quantized pumping in linear Thouless pumps ($g = 0$) requires uniform band occupation and adiabatic driving. Under these conditions, the displacement per period is dictated by the Chern number of the occupied band. The scenario is very different in the nonlinear domain, which we explain as follows. Nonlinear systems allow for spatially localized eigenstates, so-called solitons[42,59–62]. In our work we are only concerned with bright discrete spatial solitons, for which the focusing nonlinearity balances the spatial diffraction. For decreasing power, solitons can be traced back to the (linear) band from which they bifurcate (Methods and Extended Data Fig. 1). Recently, it has been shown that low-power solitons in Thouless pumps move (that is, are pumped) according to the Chern number of the band from which they bifurcate, despite non-uniform band occupation[39–41], because they follow the position of the instantaneous single-band Wannier functions.

https://doi.org/10.1038/s41567-022-01871-x

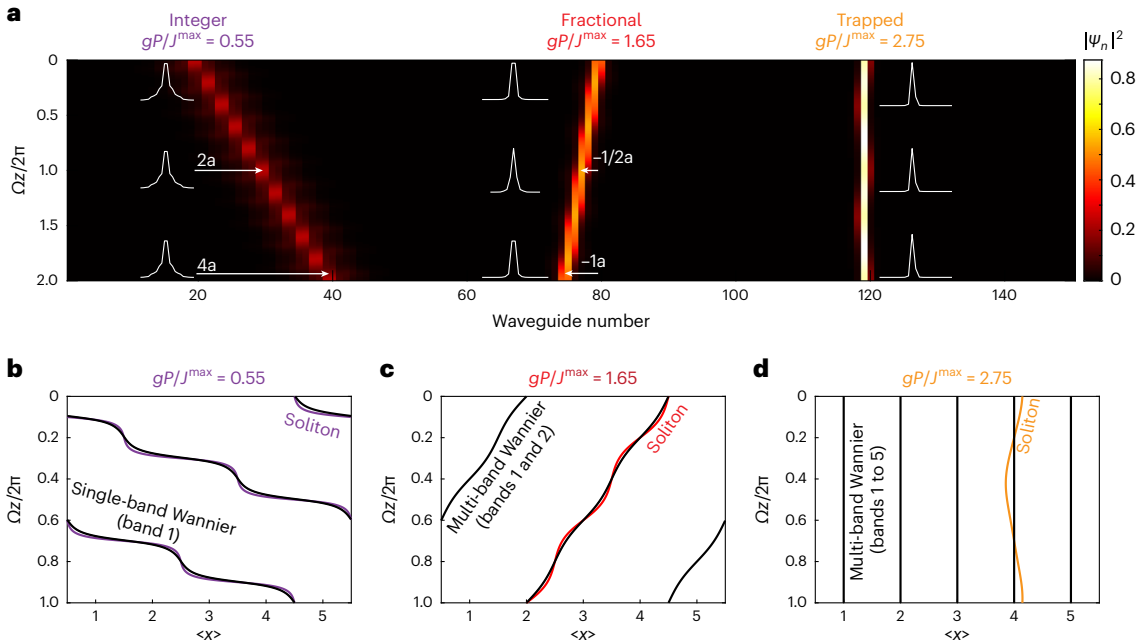


Fig. 2 | Theory of quantized fractional Thouless pumping. a, Individually normalized wavefunctions of instantaneous solitons (gP/J^{max} = 0.55, 1.65 and 2.75, from left to right), calculated for two periods. The insets show the shape of the solitons at z = 0 and after each full period. Arrows denote the displacement of the soliton after one and two periods in units of the lattice vector a. Notice that the shape of the fractionally pumped soliton (in the centre) changes after each period, and its displacement is only 1/2 of a unit cell per period. **b**–**d**, Comparison between the centre-of-mass trajectory of the relevant instantaneous Wannier functions (black) of the linear model and the instantaneous solitons (purple, red) projected into one unit cell (with sites 1 to 5). The low-power soliton (gP/J^{max} = 0.55; shown in **b** in purple) follows the single-band Wannier function of the lowest band. The trajectory of the fractionally pumped soliton (gP/J^{max} = 1.65; shown in **c** in red) follows the maximally localized multi-band Wannier function calculated for the two lowest bands combined. The position of the trapped soliton (gP/J^{max} = 2.75; shown in **d** in orange) follows the multi-band Wannier functions calculated for all five bands combined.

Integer quantization occurs because, after each period, the soliton returns to its initial state, apart from a translation by an integer number of unit cells. Throughout this Article, Wannier functions are defined for the linear (non-interacting) model.

In the adiabatic limit, the propagation of a stable soliton can be examined by calculating the instantaneous soliton for each z-slice. We confirm the existence of an adiabatic limit for soliton pumping numerically (Extended Data Fig. 2), which agrees with the results of refs. 63,64. We focus on the propagation of a soliton that, at z = 0 and $P \to 0$, bifurcates from the lowest band. Figure 2a (left) shows one example of a low-power soliton (gP/J^{max} = 0.55) for two pumping periods. After each period, the soliton's wavefunction returns to its initial wavefunction (as shown by the insets), only translated by two unit cells, as dictated by the Chern number of the band from which the soliton bifurcates (the band has Chern number 2). The propagation of a soliton for higher power (gP/J^{max} = 1.65), which shows fractional pumping, is displayed in the centre of Fig. 2a. In this case, the soliton after one period is clearly different from the soliton at z = 0. The wavefunction is not peaked on two sites, but instead on a single site, and its centre-of-mass displacement is −1/2 unit cells. Only after two periods is the soliton's wavefunction identical to the initial one modulo a translation by one unit cell in the negative direction (leftward). This behaviour goes hand in hand with the presence of two soliton solutions that are degenerate at certain points in the pump cycle (these are the nonlinear eigenvalues plotted in Fig. 1c in red and blue), as we explain in the following. For even higher power, the soliton is trapped (Fig. 2a, right side). Although we show only one example of each regime in Fig. 2, quantization of soliton transport occurs over a wide range of nonlinearity and has a rich plateau structure, as shown below.

To explain this behaviour, and especially the origin of fractional pumping, we plot in Fig. 2b–d the centre-of-mass positions of the instantaneous solitons from Fig. 2a together with the position of the relevant maximally localized instantaneous Wannier functions, projected into a single unit cell. The latter are identical to the gauge-invariant eigenvalues of (potentially multi-band) Wilson loops.

As proved in refs. 39,40, the stable low-power soliton follows (with small deviations) the position of the Wannier functions of the band from which it bifurcates. Hence, the gauge-invariant positions of the single-band Wannier function for each z-slice dictates the path of the low-power (gP/J^{max} = 0.55) soliton as depicted in Fig. 2b. As the two lowest bands of the AAH model are only separated by a small bandgap, the underlying assumption that the soliton's dynamics are determined by a single band is no longer justified for increasing power. Instead, an effective description has to take into account the lowest two bands, for which their multi-band Wannier functions have to be calculated, which means that there are two Wannier functions per unit cell. Importantly, multi-band Wannier functions are typically quite different in shape compared with each of the single-band Wannier functions (see, for example, ref. 65 for details). Although the centre-of-mass positions of multi-band Wannier functions are not gauge-invariant, the positions of the maximally localized Wannier functions are unique, as in the single band case. In Fig. 2c we show that, for gP/J^{max} = 1.65, the soliton follows the position of the instantaneous maximally localized multi-band Wannier functions. As those are displaced by only 1/2 of a unit cell per cycle (and transform into each other), the soliton pumps fractionally. With further increasing power, the soliton becomes trapped and follows the instantaneous maximally localized multi-band Wannier functions of all bands, which are delta functions for each site in a tight-binding model, shown in Fig. 2d. Thus, the trapped regime can be thought of as the trivial limit of multi-band pumping.

Although the winding of the trajectories in Fig. 2b,c looks deceptively similar, there are distinct differences between integer and fractional pumping. First, although at each z-slice there is only one position value in Fig. 2b, there are two in Fig. 2c, as the number of

https://doi.org/10.1038/s41567-022-01871-x

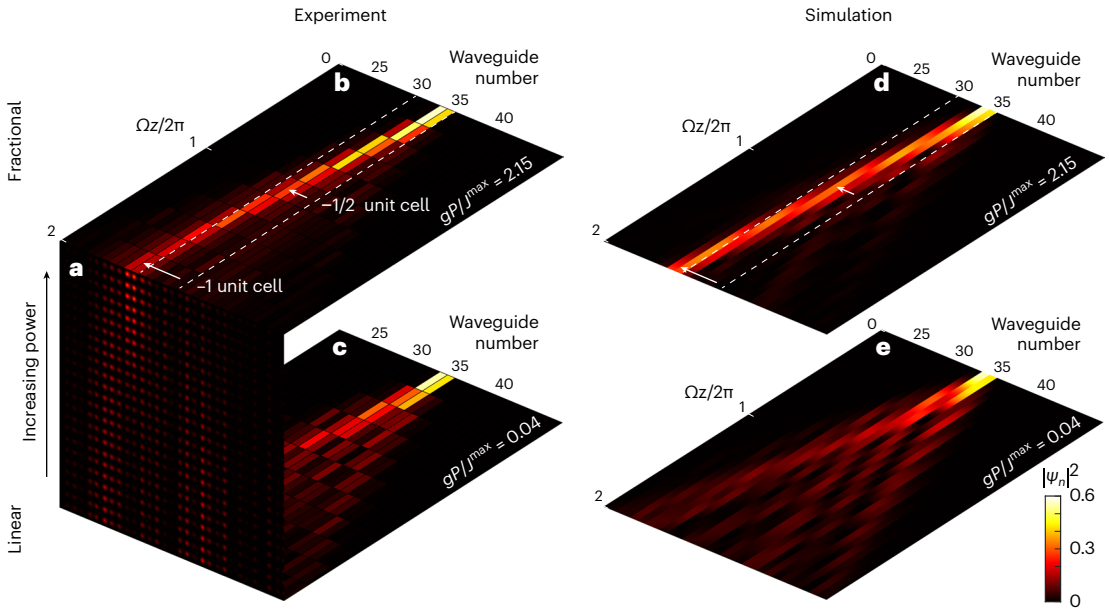


Fig. 3 | Experimental observation of an $f = -1/2$ fractionally pumped soliton. a, Stacked images of the output facet, showing the intensities in the waveguide modes after propagation for two periods. Each row was imaged separately at different input power. The bottom row shows the output facet for linear propagation ($gP/J^{max} = 0.04$). The next row was taken for $gP/J^{max} = 0.09$ and then for each row the power was increased in equal steps of $gP/J^{max} \approx 0.09$ until the top row, with a maximum input power of $gP/J^{max} = 2.15$. **b**,**c**, Normalized integrated intensities in each waveguide for different z-slices for an input power of $gP/J^{max} = 0.04$ (**c**) and $gP/J^{max} = 2.15$ (**b**). The white dashed lines mark one unit cell. After one period, the soliton is peaked on a single site, and its centre of mass has shifted by half a unit cell. After two periods, the soliton is peaked on two sites and displaced by −1 unit cell. **d**,**e**, Tight-binding propagation simulations corresponding to **b** and **a**, including propagation losses and using the measured initial two-site excitation. For direct comparison with the experiment, the plotted intensities are normalized for each z. A comparison of the propagation for all power values is shown in Supplementary Animation 1.

Wannier functions per unit cell equals the number of participating bands. Second, the winding of the Wannier function in Fig. 2b over one period is equal to the Chern number of the band ($C_1 = 2$) and dictates the pumping. In Fig. 2c, the combined winding of both Wannier functions gives the combined Chern number of both bands, $C_{1-2} = +2 - 3 = -1$. In a conventional fermionic Thouless pump with a Fermi level in the bandgap above the second band, both multi-band Wannier states are simultaneously occupied and therefore only integer-quantized pumping occurs. Remarkably, for the soliton it is possible to track a single maximally localized multi-band Wannier function along its trajectory (as shown in Fig. 2c), and the soliton is therefore pumped by a fraction after one period and by an integer after two periods. Fractional pumping is therefore an intrinsically nonlinear effect, because the projection of the multi-band Wannier function onto the linear energy eigenstates (that is, the occupation of the Bloch states) changes with z. In contrast, any linear pumping process with non-degenerate bands has—in the adiabatic limit—a constant occupation by definition.

Observing that the soliton follows the maximally localized multi-band Wannier function, whose combined displacement is given by the sum of the Chern numbers C_i of the respective bands denoted by i, we label the fractionally pumped soliton by the fraction f that describes the average displacement per period:

$$f = \frac{\sum_{i=1}^{N_b} C_i}{N_b}, \qquad (2)$$

where N_b denotes the number of participating bands. In this way, the numerator defines the number of unit cells by which the soliton is pumped in the x-direction before returning to the same wavefunction, and the denominator defines the number of pump cycles (in z) over which this process occurs. For the cases shown in Fig. 2b–d, $f = 2/1 = 2$, $f = (2 - 3)/2 = -1/2$ and $f = (2 - 3 + 2 - 3 + 2)/5 = 0$, respectively. From equation (2) it is clear that bandgap closings (including topological transitions) within the group of participating bands will not change f and therefore do not change the soliton's fractionally quantized displacement (Supplementary Animation 2).

Experimental observation of fractional pumping

We experimentally observed quantized fractional Thouless pumping in evanescently coupled waveguide arrays with Kerr nonlinearity. The waveguides were fabricated by means of femtosecond direct laser writing in borosilicate glass[66,67]. Straight waveguides showed propagation losses of (0.33 ± 0.02) dB cm^{-1} (Extended Data Fig. 3), and we measured no additional nonlinear losses (Methods and Extended Data Fig. 4). To excite the system, we focused high-power laser pulses into the waveguides; these were temporally stretched to 2 ps and down-chirped[22]. This configuration minimizes the generation of new wavelengths via self-phase modulation while reaching the necessary degree of nonlinearity. Maintaining a narrow spectrum is essential, as the hopping constant is a function of wavelength. In our experiment, the spectrum at the position of the pumped soliton broadened to 14 nm (Extended Data Fig. 5) for maximum input power and propagation distance. Because the coupling constants between waveguides vary minimally over this range of wavelengths[22], equation (1) describes our system well, provided we include unavoidable propagation losses.

We measured a fractional pumping of $f = -1/2$, meaning that the soliton is pumped to the left by one unit cell after two periods. For our experiment, it is crucial to efficiently excite the soliton, whose wavefunction is mainly peaked on two sites for $z = 0$, as shown in Fig. 2a. To facilitate this, our sample contained a 5-mm-long input region, in which only two waveguides per unit cell extend all the way to the input facet, together with one additional waveguide on top (also Fig. 1d,e). This 'triple coupler' converts a single-site excitation of the upper waveguide into an effective two-site excitation of the lower waveguides (Methods and Extended Data Fig. 6).

https://doi.org/10.1038/s41567-022-01871-x

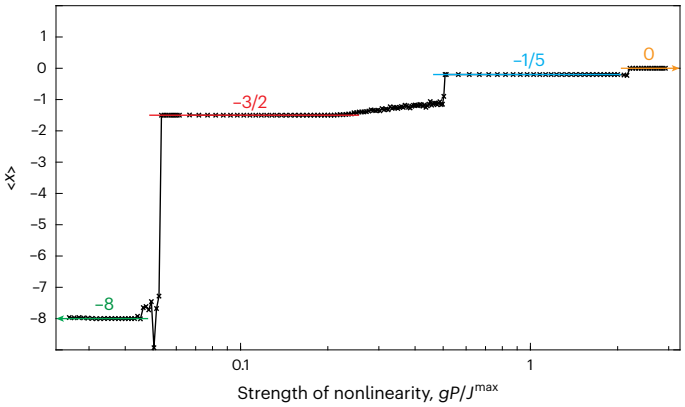


Fig. 4 | Average centre-of-mass displacement per period, $\langle x \rangle$, calculated for an off-diagonal AAH model with 13 bands. For low power ($gP/J^{max} < 0.04$), soliton motion is determined by the integer-quantized Chern number of the band from which the soliton bifurcates ($C = -8$). With increasing nonlinearity, plateaux of fractionally quantized displacement of −3/2 and −1/5 appear. For $gP/J^{max} > 2.2$, the soliton is trapped and the average displacement is zero. Note that data points for $gP/J^{max} < 0.06$ and $gP/J^{max} > 0.06$ are calculated differently to ensure adiabaticity (Supplementary Information).

We detected the fractional pumping behaviour of the soliton at high power by measuring the intensity distribution of the waveguide modes at the output facet as a function of the input power for a lattice with 12 unit cells. Figure 3a shows the normalized mode intensities at the output facet after two periods. Each row of modes corresponds to an individual measurement, and the input power increases from bottom to top. For the lowest input power ($gP/J^{max} = 0.04$; Fig. 3a, bottom row), the system behaves linearly, and the intensity diffracts widely in the waveguide array. For increasing input power, the nonlinearity counteracts the diffraction, and less spreading is visible. At the maximum input power of $gP/J^{max} = 2.15$ (Fig. 3a, top row), the intensity at the output facet is localized mainly to two waveguides, one unit cell away from the excited waveguides. This is the signature of the $f = -1/2$ fractionally pumped soliton after two periods.

To further verify the fractional pumping behaviour, we mapped out the propagation of the soliton by repeatedly cutting the sample (further information is provided in Methods) and imaging the output facet. The normalized integrated intensities per mode (which are equivalent to $|\psi_n|^2$) are shown over two periods in Fig. 3c for linear propagation ($gP/J^{max} = 0.04$) and in Fig. 3b for $gP/J^{max} = 2.15$. Corresponding numerical simulations using the experimentally measured mode intensities of the effective two-site excitation and including realistic losses are shown in Fig. 3e for the linear and Fig. 3d for the fractionally pumped case. In Supplementary Animation 1 we provide a comparison between theory and experiment for additional input power values; this animation clearly shows the transition from linear diffraction to soliton formation. Although here we only present measurements for the fractionally pumped soliton of one unit cell, we found the same behaviour in all other measured unit cells (Extended Data Fig. 7). We expect that the solitons that we observe here would also be formed under spatial modulation instability, starting from a broad input beam. The solitons that form spontaneously would then exhibit quantized motion in the course of the pump.

We point out that our waveguide system does not behave perfectly adiabatically; it is sufficiently adiabatic to observe the fractionally pumped but not the integer-pumped soliton, which has weaker confinement (Methods and Extended Data Fig. 1). Imperfect adiabaticity leads to radiation from the soliton into linearly diffracting modes. In the experiment, the contrast is further lowered by the tails of the laser pulses, which have lower intensity and thus behave more linearly. Nonetheless, our experiment clearly shows a soliton displaced to the left by one unit cell after two periods (the amplitude overlap with an instantaneous soliton is 70%), with the characteristic shape of being localized on two sites after two periods and localized on just one site after one period (see also the instantaneous soliton in the centre of Fig. 2a). We also quantify the soliton's centre-of-mass displacement using a higher order norm: $\langle x_8 \rangle = \sum_n n|\psi_n|^8 / \sum_n |\psi_n|^8$. This quantity reduces the influence of the experimentally unavoidable linearly propagating background in the calculation of the centre of mass. These results are shown in Extended Data Fig. 7: the experimentally observed wavefunction clearly follows the position of the numerically calculated instantaneous soliton with a plateau of fractional displacement per period.

Multiple fractional plateaux

Finally, we demonstrate numerically that multiple plateaux of integer and fractionally quantized pumping can occur within one Thouless pump model for increasing nonlinearity. We numerically solve equation (1) for an off-diagonal AAH model with 13 sites per unit cell (Extended Data Fig. 8 shows the bandstructure, and more details on the model and propagation parameters are provided in Supplementary Information) using periodic boundary conditions. The initial excitation is chosen to be a power-dependent, instantaneous soliton that bifurcates from the lowest band. In Fig. 4 we plot the average centre-of-mass displacement per period $\langle x \rangle$, showing four plateaux of quantized displacement. At low power the centre-of-mass displacement is integer-quantized, as dictated by the Chern number of the band from which the soliton bifurcates: $C = -8$. With increasing nonlinearity, the two lowest bands participate in soliton pumping, resulting in a fractional pumping of $f = (-8 + 5)/2 = -3/2$. At even higher power, another fractionally pumped soliton emerges, which is described by the participation of the five lowest bands: $f = (-8 + 5 + 5 - 8 + 5)/5 = -1/5$. Finally, at very high power, the soliton is trapped as the strong nonlinearly induced on-site detuning effectively detaches the soliton from the lattice. Therefore, the soliton's displacement per period is zero[41]. In our framework this corresponds to the Chern number average of all bands, which is known to be zero in tight-binding models. In Extended Data Fig. 9 we show that—similar to Fig. 2b–d—all pumped solitons follow the centre-of-mass positions of the respective maximally localized multi-band Wannier functions.

Summary and outlook

In summary, we have theoretically predicted fractional pumping of solitons in nonlinear photonic Thouless pumps and experimentally observed a fractional $f = -1/2$ displacement. Furthermore, we have numerically shown the occurrence of a rich structure of multiple quantized plateaux of integer and fractional displacement. This is evocative of features of the fractional quantum Hall effect, which has recently been studied in similar but two-dimensional models with small bandgaps[68,69]. The connection is particularly compelling in that both phenomena seem to require degenerate ground states (the solitons residing below the lowest band are nonlinear ground states and their eigenvalues are degenerate at certain points within the pumping cycle). That said, our results here occur for effectively attractive bosons described by a localized mean-field single-particle wavefunction. This result implies that the fractionalization of transport in interacting topological systems is perhaps more general than was previously understood in the context of the fractional quantum Hall effect.

While finalizing the experiments, the authors became aware of a related work[70].

Online content

Any methods, additional references, Nature Portfolio 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/s41567-022-01871-x.

References

1. Klitzing, K. V., Dorda, G. & Pepper, M. New method for high-accuracy determination of the fine-structure constant based on quantized Hall resistance. *Phys. Rev. Lett.* **45**, 494–497 (1980).
2. Thouless, D. J., Kohmoto, M., Nightingale, M. P. & den Nijs, M. Quantized Hall conductance in a two-dimensional periodic potential. *Phys. Rev. Lett.* **49**, 405–408 (1982).
3. Simon, B. Holonomy, the quantum adiabatic theorem, and Berry's phase. *Phys. Rev. Lett.* **51**, 2167–2170 (1983).
4. Tsui, D. C., Stormer, H. L. & Gossard, A. C. Two-dimensional magnetotransport in the extreme quantum limit. *Phys. Rev. Lett.* **48**, 1559–1562 (1982).
5. Laughlin, R. B. Anomalous quantum Hall effect: an incompressible quantum fluid with fractionally charged excitations. *Phys. Rev. Lett.* **50**, 1395–1398 (1983).
6. Jain, J. K. Composite-fermion approach for the fractional quantum Hall effect. *Phys. Rev. Lett.* **63**, 199–202 (1989).
7. Jain, J. K. *Composite Fermions* (Cambridge Univ. Press, 2007).
8. Wang, Z., Chong, Y., Joannopoulos, J. D. & Soljačić, M. Observation of unidirectional backscattering-immune topological electromagnetic states. *Nature* **461**, 772–775 (2009).
9. Raghu, S. & Haldane, F. D. M. Analogs of quantum-Hall-effect edge states in photonic crystals. *Phys. Rev. A* **78**, 033834 (2008).
10. Rechtsman, M. C. et al. Photonic Floquet topological insulators. *Nature* **496**, 196–200 (2013).
11. Hafezi, M., Mittal, S., Fan, J., Migdall, A. & Taylor, J. M. Imaging topological edge states in silicon photonics. *Nat. Photon.* **7**, 1001–1005 (2013).
12. Atala, M. et al. Direct measurement of the Zak phase in topological Bloch bands. *Nat. Phys.* **9**, 795–800 (2013).
13. Jotzu, G. et al. Experimental realization of the topological Haldane model with ultracold fermions. *Nature* **515**, 237–240 (2014).
14. Aidelsburger, M. et al. Measuring the Chern number of Hofstadter bands with ultracold bosonic atoms. *Nat. Phys.* **11**, 162–166 (2015).
15. Süsstrunk, R. & Huber, S. D. Observation of phononic helical edge states in a mechanical topological insulator. *Science* **349**, 47 (2015).
16. Nash, L. M. et al. Topological mechanics of gyroscopic metamaterials. *Proc. Natl Acad. Sci. USA* **112**, 14495–14500 (2015).
17. Ozawa, T. et al. Topological photonics. *Rev. Mod. Phys.* **91**, 015006 (2019).
18. Zeuner, J. M. et al. Observation of a topological transition in the bulk of a non-Hermitian system. *Phys. Rev. Lett.* **115**, 040402 (2015).
19. Smirnova, D., Leykam, D., Chong, Y. & Kivshar, Y. Nonlinear topological photonics. *Appl. Phys. Rev.* **7**, 021306 (2020).
20. Clark, L. W., Schine, N., Baum, C., Jia, N. & Simon, J. Observation of Laughlin states made of light. *Nature* **582**, 41–45 (2020).
21. Lumer, Y., Plotnik, Y., Rechtsman, M. C. & Segev, M. Self-localized states in photonic topological insulators. *Phys. Rev. Lett.* **111**, 243905 (2013).
22. Mukherjee, S. & Rechtsman, M. C. Observation of Floquet solitons in a topological bandgap. *Science* **368**, 856–859 (2020).
23. Ablowitz, M. J., Curtis, C. W. & Ma, Y.-P. Linear and nonlinear traveling edge waves in optical honeycomb lattices. *Phys. Rev. A* **90**, 023813 (2014).
24. Leykam, D. & Chong, Y. D. Edge solitons in nonlinear-photonic topological insulators. *Phys. Rev. Lett.* **117**, 143901 (2016).
25. Mukherjee, S. & Rechtsman, M. C. Observation of unidirectional solitonlike edge states in nonlinear Floquet topological insulators. *Phys. Rev. X* **11**, 041057 (2021).
26. Maczewsky, L. J. et al. Nonlinearity-induced photonic topological insulator. *Science* **370**, 701–704 (2020).
27. Xia, S. et al. Nontrivial coupling of light into a defect: the interplay of nonlinearity and topology. *Light Sci. Appl.* **9**, 147 (2020).
28. Xia, S. et al. Nonlinear tuning of PT symmetry and non-Hermitian topological states. *Science* **372**, 72–76 (2021).
29. Thouless, D. J. Quantization of particle transport. *Phys. Rev. B* **27**, 6083–6087 (1983).
30. Niu, Q. & Thouless, D. Quantised adiabatic charge transport in the presence of substrate disorder and many-body interaction. *J. Phys. A Math. Gen.* **17**, 2453–2462 (1984).
31. Kraus, Y. E., Lahini, Y., Ringel, Z., Verbin, M. & Zilberberg, O. Topological states and adiabatic pumping in quasicrystals. *Phys. Rev. Lett.* **109**, 106402 (2012).
32. Lohse, M., Schweizer, C., Zilberberg, O., Aidelsburger, M. & Bloch, I. A Thouless quantum pump with ultracold bosonic atoms in an optical superlattice. *Nat. Phys.* **12**, 350–354 (2016).
33. Nakajima, S. et al. Topological Thouless pumping of ultracold fermions. *Nat. Phys.* **12**, 296–300 (2016).
34. Lohse, M., Schweizer, C., Price, H. M., Zilberberg, O. & Bloch, I. Exploring 4D quantum Hall physics with a 2D topological charge pump. *Nature* **553**, 55–58 (2018).
35. Ma, W. et al. Experimental observation of a generalized Thouless pump with a single spin. *Phys. Rev. Lett.* **120**, 120501 (2018).
36. Fedorova, Z., Qiu, H., Linden, S. & Kroha, J. Observation of topological transport quantization by dissipation in fast Thouless pumps. *Nat. Commun.* **11**, 3758 (2020).
37. Cerjan, A., Wang, M., Huang, S., Chen, K. P. & Rechtsman, M. C. Thouless pumping in disordered photonic systems. *Light Sci. Appl.* **9**, 178 (2020).
38. Grinberg, I. H. et al. Robust temporal pumping in a magneto-mechanical topological insulator. *Nat. Commun.* **11**, 974 (2020).
39. Jürgensen, M. & Rechtsman, M. C. Chern number governs soliton motion in nonlinear Thouless pumps. *Phys. Rev. Lett.* **128**, 113901 (2022).
40. Mostaan, N., Grusdt, F. & Goldman, N. Quantized topological pumping of solitons in nonlinear photonics and ultracold atomic mixtures. *Nat. Commun.* **13**, 5997 (2022).
41. Jürgensen, M., Mukherjee, S. & Rechtsman, M. C. Quantized nonlinear Thouless pumping. *Nature* **596**, 63–67 (2021).
42. Christodoulides, D. N. & Joseph, R. I. Discrete self-focusing in nonlinear arrays of coupled waveguides. *Opt. Lett.* **13**, 794–796 (1988).
43. Eisenberg, H., Silberberg, Y., Morandotti, R., Boyd, A. & Aitchison, J. Discrete spatial optical solitons in waveguide arrays. *Phys. Rev. Lett.* **81**, 3383–3386 (1998).
44. Fleischer, J. W., Segev, M., Efremidis, N. K. & Christodoulides, D. N. Observation of two-dimensional discrete solitons in optically induced nonlinear photonic lattices. *Nature* **422**, 147–150 (2003).
45. Christodoulides, D. N., Lederer, F. & Silberberg, Y. Discretizing light behaviour in linear and nonlinear waveguide lattices. *Nature* **424**, 817–823 (2003).
46. Kivshar, Y. S. & Agrawal, G. P. *Optical Solitons: From Fibers to Photonic Crystals* (Academic Press, 2003).
47. Lederer, F. et al. Discrete solitons in optics. *Phys. Rep.* **463**, 1–126 (2008).
48. Kevrekidis, P. G. *The Discrete Nonlinear Schrödinger Equation: Mathematical Analysis, Numerical Computations and Physical Perspectives* Vol. 232 (Springer, 2009).
49. Dalfovo, F., Giorgini, S., Pitaevskii, L. P. & Stringari, S. Theory of Bose-Einstein condensation in trapped gases. *Rev. Mod. Phys.* **71**, 463–512 (1999).
50. Pitaevskii, L. & Stringari, S. *Bose-Einstein Condensation and Superfluidity* Vol. 164 (Oxford Univ. Press, 2016).
51. Donley, E. A. et al. Dynamics of collapsing and exploding Bose-Einstein condensates. *Nature* **412**, 295–299 (2001).
52. Abo-Shaeer, J. R., Raman, C., Vogels, J. M. & Ketterle, W. Observation of vortex lattices in Bose-Einstein condensates. *Science* **292**, 476–479 (2001).

53. Deng, H., Weihs, G., Santori, C., Bloch, J. & Yamamoto, Y. Condensation of semiconductor microcavity exciton polaritons. *Science* **298**, 199–202 (2002).
54. Kasprzak, J. et al. Bose-Einstein condensation of exciton polaritons. *Nature* **443**, 409–414 (2006).
55. Balili, R., Hartwell, V., Snoke, D., Pfeiffer, L. & West, K. Bose-Einstein condensation of microcavity polaritons in a trap. *Science* **316**, 1007–1010 (2007).
56. Harper, P. G. Single band motion of conduction electrons in a uniform magnetic field. *Proc. Phys. Soc. A* **68**, 874 (1955).
57. Aubry, S. & André, G. Analyticity breaking and Anderson localization in incommensurate lattices. *Ann. Isr. Phys. Soc.* **3**, 18 (1980).
58. Ke, Y. et al. Topological phase transitions and thouless pumping of light in photonic waveguide arrays. *Laser Photon. Rev.* **10**, 995–1001 (2016).
59. Askaryan, G. Effect of the gradient of a strong electromagnetic ray on electrons and atoms. *Zhur. Eksptl'. i Teoret. Fiz* **42**, 1567–1570 (1962).
60. Chiao, R. Y., Garmire, E. & Townes, C. H. Self-trapping of optical beams. *Phys. Rev. Lett.* **13**, 479–482 (1964).
61. Ablowitz, M. J. & Segur, H. *Solitons and the Inverse Scattering Transform* (SIAM, 1981).
62. Stegeman, G. I. & Segev, M. Optical spatial solitons and their interactions: universality and diversity. *Science* **286**, 1518–1523 (1999).
63. Liu, J., Wu, B. & Niu, Q. Nonlinear evolution of quantum states in the adiabatic regime. *Phys. Rev. Lett.* **90**, 170404 (2003).
64. Pu, H., Maenner, P., Zhang, W. & Ling, H. Y. Adiabatic condition for nonlinear systems. *Phys. Rev. Lett.* **98**, 050406 (2007).
65. Vanderbilt, D. *Berry Phases in Electronic Structure Theory: Electric Polarization, Orbital Magnetization and Topological Insulators* (Cambridge Univ. Press, 2018).
66. Davis, K. M., Miura, K., Sugimoto, N. & Hirao, K. Writing waveguides in glass with a femtosecond laser. *Opt. Lett.* **21**, 1729–1731 (1996).
67. Szameit, A. & Nolte, S. Discrete optics in femtosecond-laser-written photonic structures. *J. Phys. B At. Mol. Opt. Phys.* **43**, 163001 (2010).
68. Luo, W.-W., He, A.-L., Wang, Y.-F., Zhou, Y. & Gong, C.-D. Bosonic fractional Chern insulating state at integer fillings in a multiband system. *Phys. Rev. B* **104**, 115126 (2021).
69. Schoonderwoerd, L., Pollmann, F. & Möller, G. Interaction-driven plateau transition between integer and fractional Chern insulators. Preprint at https://arxiv.org/abs/1908.00988v2 (2022).
70. Fu, Q., Wang, P., Kartashov, Y. V., Konotop, V. V. & Ye, F. Nonlinear Thouless pumping: solitons and transport breakdown. *Phys. Rev. Lett.* **128**, 154101 (2022).

Methods

Properties of discrete solitons and their bifurcation from bands at low power

Solitons may be localized in space as a result of the balance between nonlinearity and diffraction (spatial solitons) or in time as a result of the balance between nonlinearity and dispersion (temporal solitons). In the main text we are exclusively working with localized discrete spatial solitons, for which the nonlinearity balances the spatial discrete diffraction. Intuitively, spatial solitons can be understood as eigenstates of their self-induced potential. In Extended Data Fig. 1 we show that in the 1D lattices with which we work, at low power, solitons bifurcate from a linear energy band. Therefore, such solitons can be assigned to that band. In particular, we use the five-site AAH model, described in the main text, at $\Omega z/2\pi = 0.3$. The calculated nonlinear eigenvalue and the corresponding soliton wavefunction are shown in Extended Data Fig. 1a,b, respectively. Clearly, for low power, the nonlinear eigenvalue approaches the bottom of a single linear energy band (in this case the lowest). Additionally, the soliton's projection onto the linear energy eigenstates, as displayed in Extended Data Fig. 1c, shows an occupation of only the low energy states. With increasing power the nonlinear eigenvalue departs from the energy band, and nonlinear bifurcations can occur. One such nonlinear bifurcation at non-zero power, marked by dashed lines in Extended Data Fig. 1, shows three distinct features: (1) the slope of the nonlinear eigenvalue changes abruptly as a function of the input power; (2) the spatial shape of the soliton's wavefunction changes notably; (3) the soliton starts to occupy energy eigenstates from the next higher band. This bifurcation leads to the transition from the integer to the fractionally pumped soliton. The same bifurcation is also shown in Supplementary Fig. 1, where it is embedded into the pumping cycle to show how a new trajectory emerges that leads to fractional pumping.

Adiabaticity of integer and fractional soliton propagation

The quantization of transport in Thouless pumps becomes exact in the adiabatic limit ($\Omega \to 0$). In this section we show numerically that (1) we may approach an adiabatic limit for solitons in Thouless pumps but (2) that our experiment is only sufficiently adiabatic to observe the fractionally but not the integer-quantized soliton.

We numerically propagate an instantaneous soliton (solved for at $z = 0$) over one full period for increasing period L. Extended Data Fig. 2a shows the calculated overlap, $|\sum_n \Psi_n^*(L)\phi_n(L)|^2$, between the instantaneous soliton $\phi_n(L)$ and the propagated wavefunction $\Psi_n(L)$, after one period. The deviation from perfect overlap is shown in Extended Data Fig. 2b on a log–log scale, which demonstrates the presence of an adiabatic limit. These results agree with the adiabatic time-evolution studies of nonlinear systems in refs. 63,64. Extended Data Fig. 2 clearly indicates that the lower-power (integer-pumped) soliton needs slower driving frequencies compared to the higher-power (fractionally pumped) soliton to be similarly adiabatic and experimentally observable. Extended Data Fig. 2a confirms that our experiment is sufficiently adiabatic to detect the signature of the fractionally pumped soliton.

Single-band and multi-band Wannier states

As described in the main text, we find that soliton motion in the fractional pumping case tracks maximally localized multi-band Wannier functions, whose motion dictates the fraction by which pumping occurs. We describe the construction of single and multi-band Wannier functions below. Single-band Wannier states are defined as $|w_{R,\alpha}\rangle = \frac{1}{\sqrt{N}}\sum_k e^{-ikR}|\Psi_{k,\alpha}\rangle$, where $|\Psi_{k,\alpha}\rangle$ denotes Bloch states, N is the number of unit cells, R is the lattice vector, k is the wavevector and α is the band index. These Wannier states are orthogonal and form a basis analogously to Bloch states. As single-band Wannier functions are an equal superposition of all Bloch states of a single band, their occupation of the band in question is uniform, and integer-quantized pumping in Thouless pumps is expected for such an excitation. The freedom of choosing independent phase factors for each Bloch state $|\Psi_{k,\alpha}\rangle$ (gauge freedom) allows for the construction of exponentially localized Wannier states. In one dimension, maximally localized Wannier states can be conveniently calculated as the eigenvectors of the projected position operator $\hat{P}\hat{X}\hat{P}$, where $\hat{P} = \sum_k |\Psi_{k,\alpha}\rangle\langle\Psi_{k,\alpha}|$ projects into the relevant band and $\hat{X} = \mathrm{diag}(e^{i2\pi\hat{x}/Na})$ is the Resta position operator with $\hat{x}$ the standard position operator. For a single-band Wannier state the centre-of-mass position is gauge-independent.

Analogously, multi-band Wannier states are defined as $|w_{R,\tilde{\beta}}\rangle = \frac{1}{\sqrt{N}}\sum_k e^{-ikR}|\tilde{\Psi}_{k,\tilde{\beta}}\rangle$, where, importantly, $|\tilde{\Psi}_{k,\tilde{\beta}}\rangle$ are not energy eigenstates, but rather $|\tilde{\Psi}_{k,\tilde{\beta}}\rangle = \sum_\alpha U_{\alpha,\tilde{\beta}}(k)|\Psi_{k,\alpha}\rangle$ where $U_{\alpha,\tilde{\beta}}(k)$ is a unitary matrix that defines the degree of mixing of the Bloch states of different bands within the multi-band Wannier state. For multi-band Wannier states, $\tilde{\beta}$ loses its meaning as a band index, but merely enumerates the multi-band Wannier states ranging from one to the number of bands involved. As multi-band Wannier states cannot be identified with a single band, but only with a group of bands, they show non-uniform (and non-universal) occupation of the bands involved, and generally no quantization is expected in linear Thouless pumping for such an excitation. The possibility of different gauge choices—different choices of $U_{\alpha,\tilde{\beta}}(k)$—leads to different multi-band Wannier states and gauge-dependent centres of mass. Only exponentially localized multi-band Wannier functions have gauge-independent centres of mass (which can be identified as gauge-independent Wilson loop eigenvalues). Numerically, maximally localized multi-band Wannier states in one dimension can be calculated as eigenstates of the projected position operator, when projecting into the subspace of multiple bands. Multi-band Wannier states can be calculated for groups of energetically separated bands, but must be used (instead of single-band Wannier functions) for degenerate bands. The number of Wannier states per unit cell is equal to the number of bands involved, as shown in Fig. 2b–d. Further details on single-band and multi-band Wannier functions are available in ref. 65.

Waveguide fabrication

The waveguides were fabricated using femtosecond laser writing[66,67]. We created a permanent refractive-index change by focusing high-power pulses (~400 nJ pulse energy) into a borosilicate glass sample (Corning Eagle XG). The sample was mounted on a high-precision x–y–z translation stage (Aerotech) and translated once through the focal point with a speed of 8 mm s^{-1}. The 260-fs laser pulses were emitted by an Yb-doped fibre laser system (Menlo BlueCut) using a repetition rate of 500 kHz at a central wavelength of 1,030 nm. The mode field diameters of the laser mode were 4.6 mm and 5.8 mm in the horizontal and vertical directions. To shape the refractive-index profile of the waveguides, we used slit-beam shaping[71] with a slit width of 1.8 mm. The pulses were focused into the sample at a depth of ~100 µm from the surface using a lens with NA = 0.4 (Thorlabs A110TM-B).

Sample cutback

To have access to the wavefunction at different z-positions, we cut the sample. We made ten cuts in total, each time cutting away ~7.1 mm, which is 1/10th of the maximum propagation distance. After each cut we polished the output facet. The uncertainty in the cut position was 0.2 mm. To cut the sample, we defined a break line by transversely moving the sample through the laser focus with identical parameters as for waveguide fabrication, except a higher pulse energy of ~800 nJ. We repeated the transverse movement in 10-µm-height steps. The sample was then mechanically broken along this line and subsequently polished using a commercially available polishing device (Krelltech FLex Waveguide Polisher). During the polishing of the third cut, a crack appeared in the sample close to the output facet. The crack was cut away with the fourth cut. Due to the crack, for measurements taken after the third cut, we observed ~10% lower transmission than expected

for propagation without the crack. As the additional loss occurred only a short distance away from the output facet, we do not expect substantial changes in the output pattern due to the crack. We included the measured data in Extended Data Fig. 3b,c without any further manipulation, but excluded it from the fit of the propagation losses.

Linear waveguide characterization

We characterized the linear properties of the waveguides using directional couplers. The couplers consisted of two straight waveguides separated by a distance s from one another. We varied the length of the second waveguide and measured the transferred intensity at the output facet as a function of the coupling length. We extracted the coupling constant for each separation by fitting the analytic transfer function, including an on-site detuning. The resulting values for the coupling constant and the on-site detuning are shown in Extended Data Fig. 3a,b, respectively. Because of the exponential mode confinement, the coupling function is expected to decay exponentially with the separation. An exponential fit of the coupling function results in $J(s)/\mathrm{mm}^{-1} = 6.672\exp(-234.8s/\mathrm{mm})$. We used this coupling function for the simulations shown in Figs. 1–3, Extended Data Figs. 1 and 2 and for Supplementary Fig. 1.

The on-site refractive-index difference between neighbouring waveguides (thus detuning between their propagation constants) is due to the writing procedure: the fabrication of the first waveguide affects the environment and therefore the fabrication of the second waveguide. As our experiments were performed in the bulk of the system, with each waveguide fabricated next to another, and with separations varying only between 16.25 µm and 18.25 µm, the on-site detuning was small and we neglected it in the simulations. We point out that including this small on-site detuning (fitted to be $V(s)/\mathrm{mm}^{-1} = 0.291 - 0.0124s/\mu\mathrm{m}$) does not close the bandgaps of the off-diagonal AAH model shown in Fig. 1c. Thus, neither does it change the topological properties nor affect our experimental results of fractional pumping.

While cutting the sample to map out the propagation of the soliton in the Thouless pump model, we additionally measured the output power of straight waveguides for an input power of 1 mW. The measured transmission for one waveguide is shown in Extended Data Fig. 3c for propagation lengths longer than 19 mm. No values were taken for a propagation distance of ~40 mm, as the sample did not break uniformly at the position where the straight waveguides were located. To fit the propagation losses we excluded the measured values for a propagation distance of ~55 mm, as this value was taken with a small crack in the sample. The resulting exponential fits for seven straight waveguides resulted in a propagation loss of 0.33 ± 0.02 dB cm^{-1}, where the error is given as one standard deviation of the individual fits.

Nonlinear waveguide characterization and measurements

We used the same set-up for all nonlinear measurements. A schematic (not including neutral density filters, additional mirrors and stages) is shown in Extended Data Fig. 4a. A commercial laser system (Menlo BlueCut) was used to emit 270-fs pulses at a repetition rate of 5 kHz (tunable); these were adjusted in power using a combination of a half-wave plate and a polarization beamsplitter. A pair of gratings (Thorlabs GR25-0610) down-chirped and temporally stretched the horizontally polarized pulses to 2 ps (ref. 22). The pulses were then focused into the waveguides using a lens (Thorlabs AC127-030-B-ML). A second lens (Thorlabs AC254-040-B-ML) imaged the output facet of the glass sample onto a complementary metal–oxide semiconductor camera (Thorlabs CS165MU). We simultaneously focused the output into a fibre (Thorlabs P1-980A-FC-1) using a third lens (Thorlabs AC064-015-B-ML) to measure the spectrum with an optical spectrum analyzer (Anritsu MS9740 A). We used flip-mirrors to measure the time-averaged input (output) power before the first lens (after the second lens) with a photodiode power sensor (Thorlabs S120C).

We verified the absence of nonlinear losses (for example, via multi-photon absorption) by analysing the input–output power dependence. Extended Data Fig. 4b depicts the measurement for three straight waveguides. Our measurements show a linear dependence until reaching an input power of 6.0 mW, at which the waveguides are damaged. Accordingly, we only used input powers of <6.0 mW, where nonlinear losses do not play a role.

We calibrated g (equation (1)) by fitting the output intensities of a directional coupler with known separation $s = 17$ µm and coupling length $l = 6$ mm. We included an on-site detuning of $V = 0.080$ mm^{-1} (Extended Data Fig. 3b) and used a hopping constant of $J = 0.131$ mm^{-1}. Extended Data Fig. 4c shows the experimentally measured intensities, I_1 and I_2, in waveguides 1 and 2, for an excitation of waveguide 1. The corresponding output fit, using least-square fitting and including propagation losses, resulted in $g = 0.068$ mm^{-1} per mW of input power, and is plotted in Extended Data Fig. 4c (black).

To accurately describe our experiment via equation (1), it is important that the generation of new wavelengths via self-phase modulation is minimal, as the evanescent coupling is wavelength-dependent. We expect maximum self-phase modulation for focusing the maximum input power into a single waveguide. Extended Data Fig. 5a shows that the spectrum broadens from 6 nm at low power to ~25 nm at maximum input power, where this width is defined as the width at which 76% of the intensity (equivalent to the full-width at half-maximum of a Gaussian) is found. More relevant for our experiment is the generation of new wavelengths by the fractionally pumped soliton. We measured the spatially resolved spectrum at the excitation position (Extended Data Fig. 5b) and at the soliton position (Extended Data Fig. 5c). Clearly, the self-phase modulation and the spectral broadening is notably reduced to a maximum width of 14 nm, as the fractionally pumped soliton is localized on multiple sites and therefore the power per waveguide is substantially lower. We furthermore point out that the wavelengths with the largest deviation from the central wavelength of 1,030 nm are generated close to the output facet and therefore do not strongly change the output pattern.

Triple coupler

As it is experimentally more convenient to use single-site excitations, we used the first 5 mm of the sample to transform a single-site excitation into an effective two-site excitation to efficiently excite the soliton. We achieved this by writing an auxiliary 5-mm-long waveguide on top of two waveguides that are part of the Thouless pump model (Extended Data Fig. 6a and also Fig. 1d). Extended Data Fig. 6b shows the measured intensities in the three waveguides after 5 mm for increasing input power. Due to the large coupling constant between the waveguides, the effect of a nonlinearly induced on-site index change does not substantially change the output ratios. The slight imbalance between the intensities in the left and right waveguides are due to an on-site detuning stemming from the fabrication procedure. As the modulation in our system is not perfectly adiabatic, it turns out that this imbalance is favourable to excite the fractionally pumped soliton more effectively (with less background radiation) compared to a strict 50/50 excitation.

Further quantification of the pumping behaviour of the soliton

To verify the soliton's pumping behaviour in our experiment we (1) quantified the overlap between the measured and the numerically calculated instantaneous soliton wavefunction; (2) evaluated the position of the soliton during the pumping process and compared it with the position of an instantaneous soliton; (3) verified the existence of a plateau of fractionally quantized pumping.

We calculated the overlap of the measured output intensities ($|\Psi_n^{(\mathrm{Exp})}|^2$) with the theoretically calculated instantaneous soliton wavefunction ($\Psi_n^{(\mathrm{S})}$). As we do not have access to the phase information in the experiment, we calculated the amplitude overlap, $\sum_n |\Psi_n^{(\mathrm{S})}||\Psi_n^{(\mathrm{Exp})}|$.

This is a valid measure as the instantaneous soliton has a uniform phase profile. Extended Data Fig. 7a shows that at linear propagation ($gP/J^{max} = 0.04$; shown in blue) the amplitude overlap decays to 30% after two periods. At high power ($gP/J^{max} = 2.15$; shown in red), the amplitude overlap of the experimentally measured soliton with the instantaneous soliton is substantially larger, at 70% after two periods. Imperfect overlap is caused by radiation into linear (diffracting) modes due to non-adiabaticity, and the experimentally unavoidable linear propagation of the lower power tails of the input pulse.

We further quantified the pumping behaviour of the soliton by calculating its position during the pumping process and compared it with the position of the theoretically calculated instantaneous soliton. Instead of calculating the mean displacement of the intensity (which is strongly affected by the diffracting linear background modes and therefore not meaningful), we calculated a higher order of the centre of mass, $\langle x_8 \rangle$, where $\langle x_8 \rangle = \sum_n n|\Psi_n|^8 / \sum_n |\Psi_n|^8$. This quantity reduces the influence of the linear (low power) background for the calculation of the centre of mass and therefore better describes the position of the soliton. Our results are displayed in Extended Data Fig. 7b and show that the experimentally observed soliton (shown in red) closely follows the predicted trajectory of the numerically calculated instantaneous soliton (shown in grey) with fractional displacement per period.

Finally, we verified the existence of a plateau of fractionally quantized transport in the experiment. Extended Data Fig. 7c shows the higher-order centre-of-mass displacement after two periods for increasing input power. At an input power of $gP/J^{max} \approx 1.4$, the plateau forms and persists until the maximum input power that we can achieve in our experiment ($gP/J^{max} = 2.15$).

Off-diagonal AAH model with 13 sites

The hoppings of the off-diagonal AAH model with 13 sites per unit cell (as used for Fig. 4) are given by $J_n = K + \kappa \cos(\Omega z + \frac{10\pi}{13} n + \frac{2\pi}{13})$ with $K = 1/b$ and $\kappa = 0.95/b$, where b is an arbitrary length. The on-site potential is set to zero. Extended Data Fig. 8 shows the bandstructure of the model, with 13 bands that are clustered into groups of bands lying energetically close together. The Chern numbers of the bands can be evaluated by calculating the winding of the single-band Wannier functions as displayed in Extended Data Fig. 9a, resulting in Chern numbers $C = \{-8, 5, 5, -8, 5, 5, -8, 5, 5, -8, 5, 5, -8\}$, ordered from bottom to top. Similar to Fig. 2, the position of the instantaneous solitons for $gP/J^{max} = 0.04$, 0.10, 0.78 and 3.08 follow closely the position of the (multi-band) Wannier functions (Extended Data Fig. 9b–g).

Data availability

Data that support the findings of this study are available from the corresponding author upon reasonable request. Source data are provided with this paper.

References

71. Ams, M., Marshall, G. D., Spence, D. J. & Withford, M. J. Slit beam shaping method for femtosecond laser direct-write fabrication of symmetric waveguides in bulk glasses. *Opt. Express* **13**, 5676–5681 (2005).

Acknowledgements

We acknowledge fruitful discussions with S. Gopalakrishnan and S. Vaidya. We further acknowledge the support of the ONR YIP programme under award no. N00014-18-1-2595, ONR-MURI programme N00014-20-1-2325, the AFOSR-MURI programme FA9550-22-1-0339, as well as the Packard Foundation (fellowship no. 2017-66821). C.J. gratefully acknowledges funding from the Alexander von Humboldt Foundation within the Feodor-Lynen Fellowship programme. Numerical calculations were performed on the Pennsylvania State University's Institute for Computational and Data Sciences' Roar supercomputer.

Author contributions

M.J. designed, fabricated and cut the devices and carried out all measurements. M.J. characterized the devices with input from C.J. S.M. designed and built the experimental beam shaping and fabrication set-up. The theoretical investigation was carried out by M.J. M.J. and M.C.R. wrote the manuscript, with input from C.J. M.C.R. supervised the project.

Competing interests

The authors declare no competing interests.

Additional information

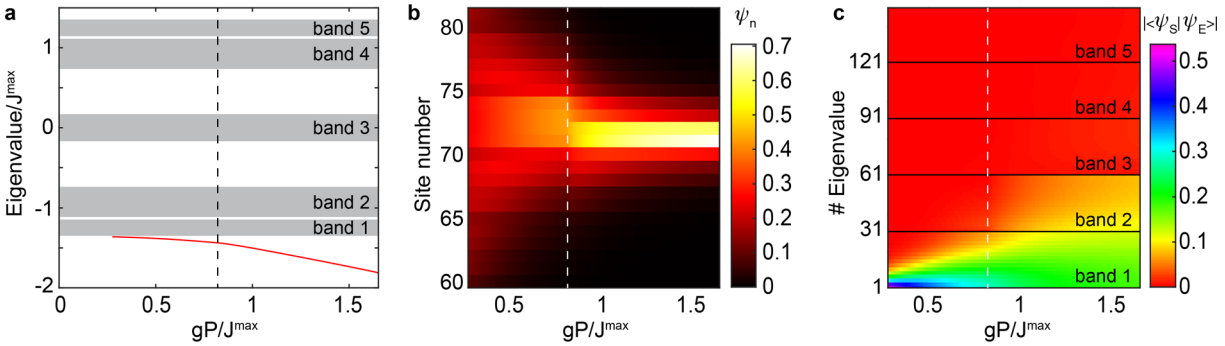
Extended data is available for this paper at https://doi.org/10.1038/s41567-022-01871-x.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41567-022-01871-x.

Correspondence and requests for materials should be addressed to Marius Jürgensen or Mikael C. Rechtsman.

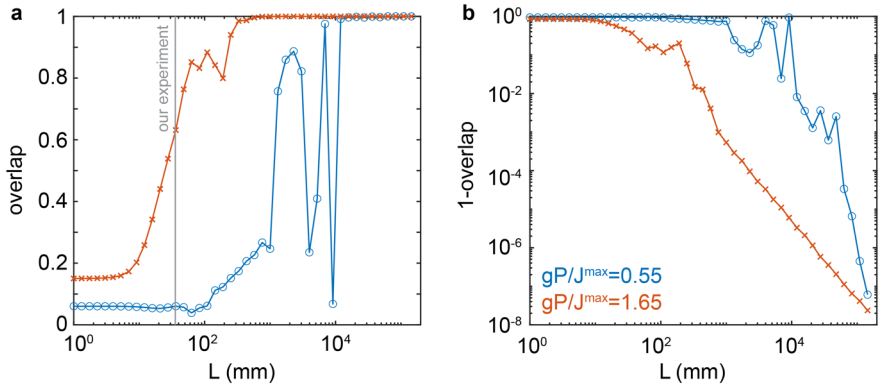
Peer review information *Nature Physics* thanks Yiming Pan and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

https://doi.org/10.1038/s41567-022-01871-x



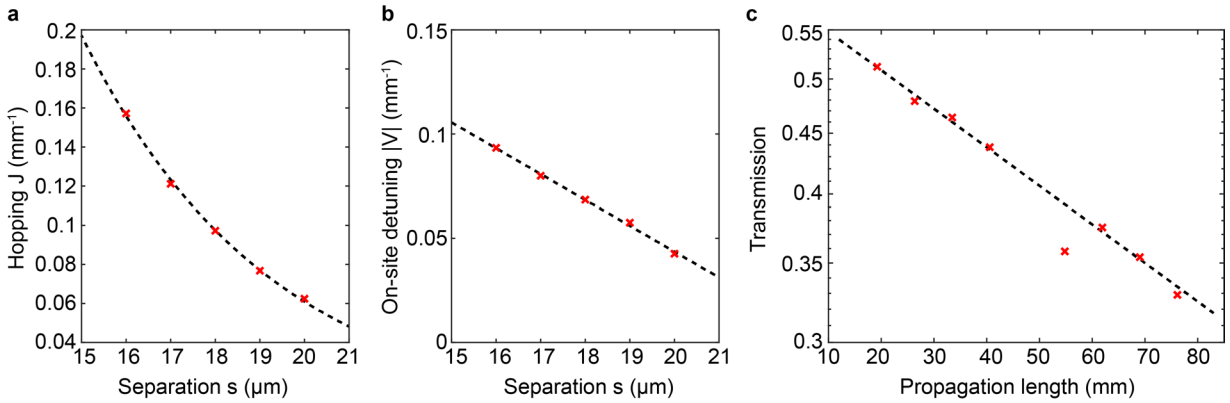
Extended Data Fig. 1 | Properties of solitons. a. Linear energy bands (grey) for the model described in the main text at $\Omega z/2\pi$=0.3 together with the nonlinear eigenvalue of a soliton (red) for increasing nonlinearity. Dashed line marks a nonlinear bifurcation point, with a corresponding slope change. **b.** Wavefunction of the soliton from **a**. **c.** Projection of the soliton wavefunction onto the linear energy eigenstates. The eigenstates are sorted via their eigenvalues from bottom (lowest) to top (highest).

https://doi.org/10.1038/s41567-022-01871-x



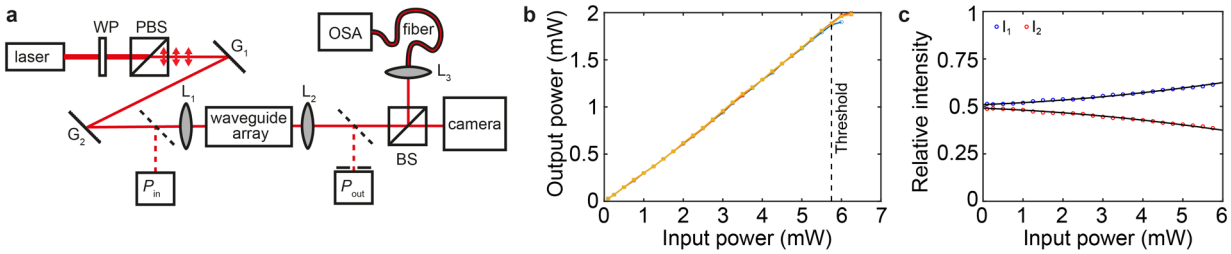
Extended Data Fig. 2 | Adiabaticity of soliton propagation. a. Overlap of a numerically propagated soliton with the instantaneous soliton after one pump cycle with period L. **b**. Same data as **a**, but plotted as deviation from perfect overlap on a log-log scale confirming the existence of an adiabatic limit. The blue (orange) line depicts to gP/J^{max}=0.55 (gP/J^{max}=1.65), which corresponds to the integer (fractionally) pumped soliton in the adiabatic limit.

https://doi.org/10.1038/s41567-022-01871-x



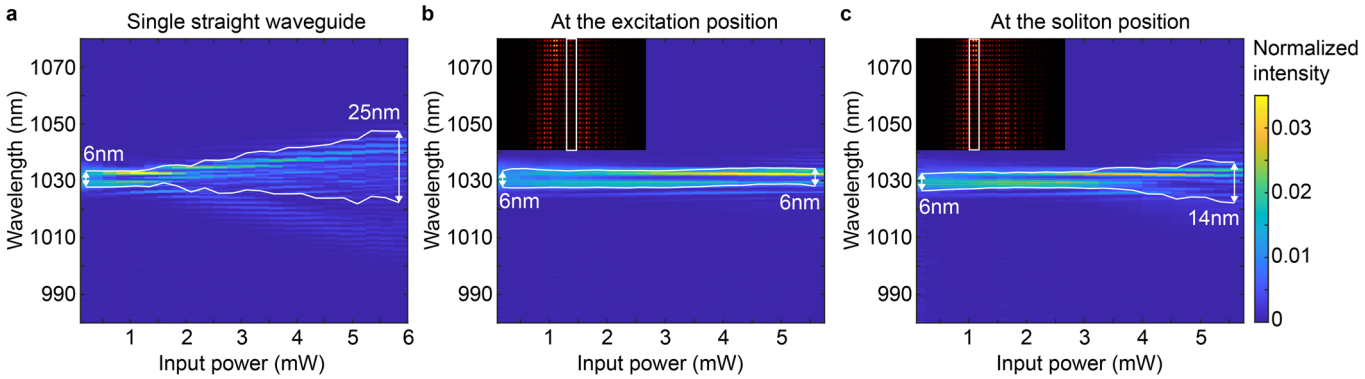
Extended Data Fig. 3 | Linear waveguide characterization. a. Hopping constant *J* as a function of separation *s* between waveguides. Each *J* is calculated via fitting 15 couplers with varying coupling length. Dashed line shows the exponential fit for the coupling function. **b**. On-site detuning in propagation constant for two waveguides written next to each other as a function of separation *s*. **c**. Transmission measurements of a straight waveguide including in- and out-coupling losses. Y-axis is plotted in log scale. The transmission measured for a propagation of ≈ 55 mm is excluded from the exponential fit (dashed line) due to a crack in the sample creating additional losses. The resulting propagation loss is (0.33 ± 0.02)dB/cm.

https://doi.org/10.1038/s41567-022-01871-x



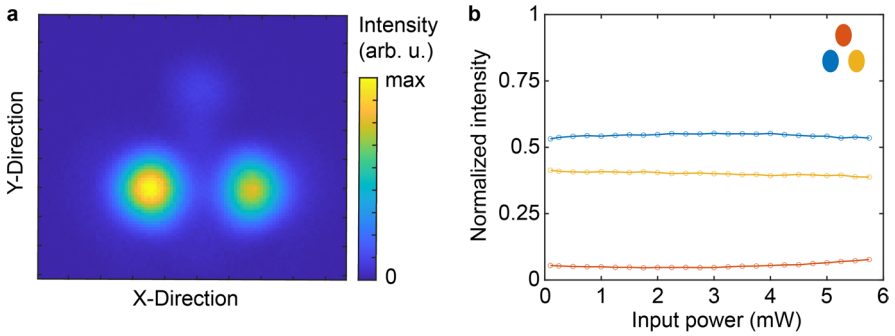
Extended Data Fig. 4 | Nonlinear waveguide characterization. a. Setup for nonlinear characterization of the waveguides including a pair of gratings (G1 and G2) to stretch and down-chirp the excitation pulse. The output facet is imaged onto a camera and simultaneously focused into a fiber-coupled optical spectrum analyzer (OSA). **b**. Measured output power for three straight waveguides as a function of the input power showing a linear dependence and hence no nonlinear absorption for input powers less than 6 mW. The threshold indicates the maximum power used in the experiments to prevent damage to the sample. **c**. Relative intensity in the two waveguides (I_1 and I_2) of a directional coupler as a function of power. Black lines are a least square fit, resulting in g=0.068/mm per mW input power.

https://doi.org/10.1038/s41567-022-01871-x



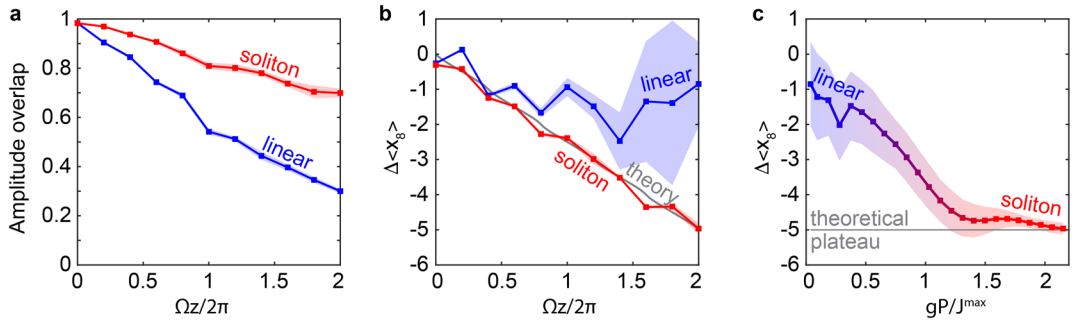
Extended Data Fig. 5 | Spectral analysis. a. Measured normalized spectrum after propagation through a 76.15 mm long, straight waveguide in the uncut sample for increasing input power. The white lines mark the width in which 76% of the intensity is found (corresponding to the full-width-half-maximum of a Gaussian). **b,c**. Similar to **a** but spatially resolved and for an array showing fractional pumping. The insets show the waveguide modes at the output facet for increasing input power (from bottom to top). The white rectangle marks the spatial $1/e^2$ width for which the spectrum is measured.

https://doi.org/10.1038/s41567-022-01871-x



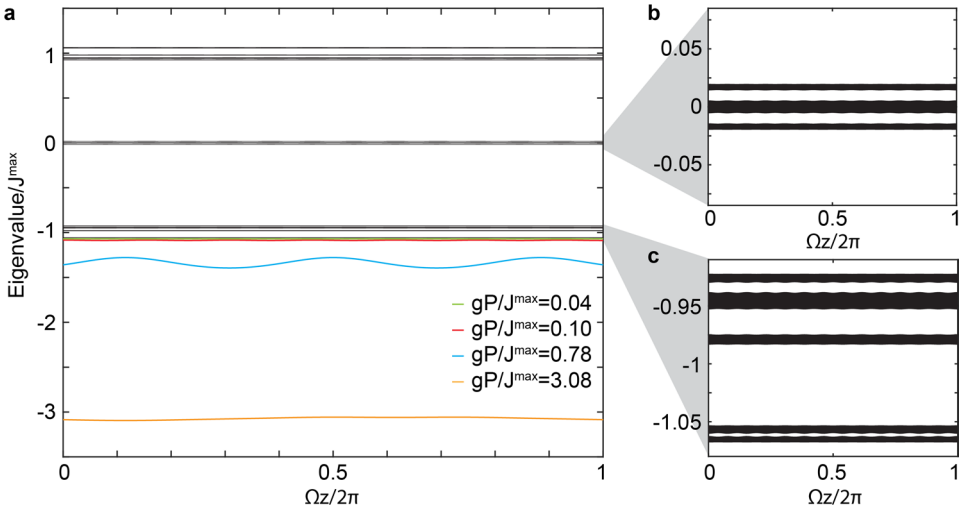
Extended Data Fig. 6 | Triple coupler. a. Intensity distribution at the output facet for a coupler consisting of three waveguides for low input power. The coupler extends over a length of 5 mm and the top waveguide is excited. **b**. Normalized output intensities of a triple coupler as a function of the input power. The inset defines the color-coding. Due to the high coupling constant between the waveguides over a short propagation distance of 5 mm, the effects of nonlinearity are negligible.

https://doi.org/10.1038/s41567-022-01871-x



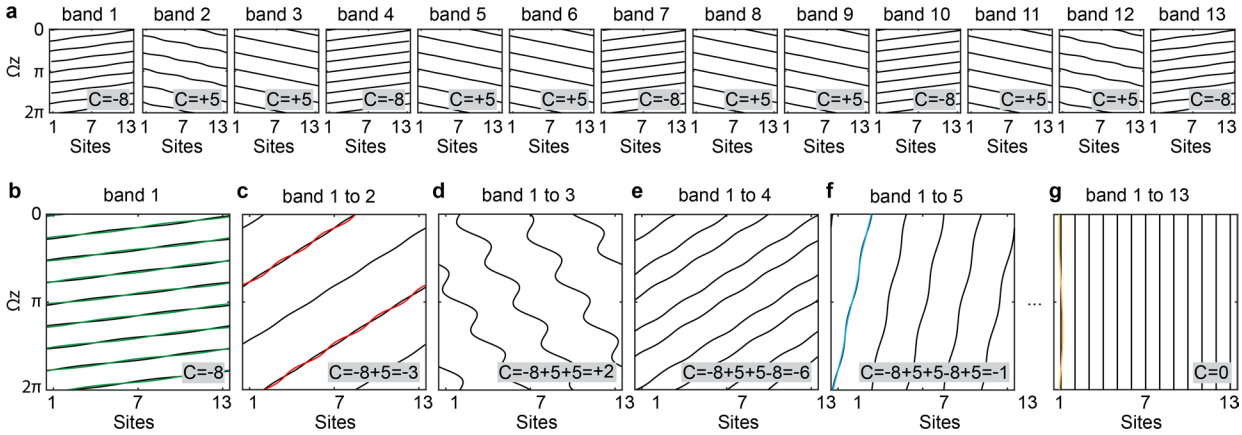
Extended Data Fig. 7 | Verification of soliton's pumping behavior in the experiment. a. Amplitude overlap, $\sum_n |\Psi_n^{(\mathrm{Exp.})}||\Psi_n^{(\mathrm{S})}|$, between the measured output wavefunction, $\Psi_n^{(\mathrm{Exp.})}$, and the numerically-calculated instantaneous soliton wavefunction, $\Psi_n^{(\mathrm{S})}$, as a function of propagation length for linear propagation (gP/J^{max}=0.04; shown in blue) and soliton propagation (gP/J^{max}=2.15; shown in red). **b**. Center of mass displacement of the experimentally observed soliton and linear propagation (calculated using a higher order norm to suppress linear background effects; see text). Gray line indicates the numerically-calculated displacement of an instantaneous fractionally pumped soliton. **c**. Experimentally observed center of mass displacement after two periods for increasing input power showing a plateau at high input power. Gray line indicates the expected theoretical displacement of five sites (one unit cell) after two periods for the fractionally pumped soliton. Solid lines with squares show mean values, and shaded areas show one standard deviation for independently measured soliton propagation in eight different unit cells of the same lattice.

https://doi.org/10.1038/s41567-022-01871-x



Extended Data Fig. 8 | Band structure of a 13 site AAH model. a. Additionally to the 13 linear energy bands (black) the nonlinear eigenvalues of four pumped solitons with gP/J^{max}=0.04, 0.10, 0.78 and 3.08, which are part of the plateaux in Fig. 4 in the main text, are shown in green, red, blue and yellow, respectively. **b** Zoom-in onto the central group of three energy bands. **c** Zoom-in onto the lowest group of five energy bands.

https://doi.org/10.1038/s41567-022-01871-x



Extended Data Fig. 9 | Wannier function positions of a 13 site AAH model. **a**. Center of mass position of single-band Wannier functions calculated for each of the 13 bands individually over one period and projected into a single unit cell. The number of windings is equal to the Chern number C of the corresponding band. **b**-**g** similar to **a** but for multi-band Wannier functions in **c**-**g**. The position of pumped solitons from the plateaus of Fig. 4 in the main text at nonlinearities gP/J^{max}=0.04, 0.10, 0.78 and 3.08, are shown in green, red, blue and yellow, respectively.