High-frequency heating of the solar wind triggered by low-frequency turbulence

Jonathan Squire, Romain Meyrand, Matthew W. Kunz, Lev Arzamasskiy, Alexander A. Schekochihin and Eliot Quataert

The fast solar wind’s high speeds and non-thermal features require that considerable heating occurs well above the Sun’s surface. Two leading theories seem incompatible: low-frequency ‘Alfvénic’ turbulence, which transports energy outwards and is observed ubiquitously by spacecraft but seems insufficient to explain the observed dominance of ion over electron heating; and high-frequency ion-cyclotron waves, which explain the non-thermal heating of ions but lack an obvious source. Here we argue that the recently proposed ‘helicity barrier’ effect, which limits electron heating by inhibiting the turbulent cascade of energy to the smallest scales, can unify these two paradigms. Our six-dimensional simulations show how the helicity barrier causes the large-scale energy to grow through time, generating small parallel scales and high-frequency ion-cyclotron-wave heating from low-frequency turbulence, while simultaneously explaining various other long-standing observational puzzles. The predicted causal link between plasma expansion and the ion-to-electron heating ratio suggests that the helicity barrier could contribute to key observed differences between fast and slow wind streams.

The basic mechanisms that heat the solar corona and accelerate the solar wind remain mysterious despite intensive study over many decades. A successful theory must explain how energy contained in photospheric motions and magnetic fields can be liberated to cause extreme and sudden heating of the coronal plasma, along with its acceleration to velocities well in excess of the escape velocity of the Sun. Adding to the complexity, the coronal plasma is collisionless—the mean-free path of protons can be large compared with the largest observed structures—meaning that it can be far out of local thermal equilibrium. This freedom opens up a wide array of channels for plasma heating: ions might be heated more than electrons (or vice versa), or particles might gain energy preferentially in a particular direction with respect to the local magnetic field. Such differences can have pronounced macroscopic consequences.

The dominant heating mechanism(s) must be consistent with an extensive array of measurements taken both remotely, from the low corona itself, and in situ, from spacecraft spread throughout the solar wind. In fast wind streams, these data indicate that the heating must be spatially extended out to several solar radii to drive observed wind speeds. The mechanism must preferentially heat protons over electrons, while heating heavier ions (for example, α-particles) even more effectively; it must heat protons in the direction perpendicular to the local magnetic field more than in the parallel direction to explain temperature anisotropies; and its features and/or after-effects should be observable in the measured field fluctuations and particle distribution, particularly at the low altitudes now being explored by the Parker Solar Probe (PSP).

One paradigm that can—at least in principle—satisfy the above requirements is heating through Alfvénic turbulence. Low-frequency Alfvénic motions in the low corona are observed to contain sufficient energy to power the wind, and there are well-developed theories for how such motions become turbulent following reflection from large-scale density gradients. This turbulence transfers energy into successively smaller-scale motions perpendicular to the magnetic field (larger $k_\perp$, where $k_\perp$ is the inverse perpendicular scale), ultimately dissipating to heat the plasma. The difficulty is that most theories predict that in the strongly magnetized limit relevant to the solar corona (the low-$\beta$ limit, where $\beta$ is the ratio of thermal to magnetic pressure), such low-frequency, high-$k_\perp$ structures dissipate to heat predominantly electrons. Other low-frequency plasma motions, such as compressive waves, generally cause parallel heating of ions. Both possibilities are inconsistent with observations. More promisingly, for turbulence of sufficient amplitudes, ‘stochastic heating’ can heat ions through a random walk on ion-gyroscale electric field fluctuations. Although it could plausibly explain key observations, questions remain, such as its possible quenching due to flattening of the distribution function and the influence of cross helicity. Another possibility, that ions are heated by kinetic-Alfvén-wave (KAW) turbulence at sub-gyroradius scales, remains less well understood and may be inefficient at low $\beta$.

In the opposite limit of short field-parallel wavelengths (large $k_\parallel$), high-frequency ion-cyclotron waves (ICWs) provide a simpler mechanism for strong perpendicular ion heating. At wavenumbers approaching $k_\parallel \approx d_i^{-1}$, where $d_i$ is the inertial length, their frequency approaches the ion gyrofrequency, where the cyclotron resonance causes highly efficient energy transfer from electromagnetic fields to ion velocities. ICWs are observed ubiquitously in situ and can suprathermally heat minor ions in a way that is observationally compelling. However, a sufficiently energetic direct solar source of ICWs is highly unlikely, and the Alfvénic cascade does not efficiently transfer energy to small parallel scales, seemingly ruling out their turbulent origin. Although their occurrence in data can be explained by kinetic instabilities, in most theories this implies that they would cool, rather than heat, the plasma; except perhaps in the presence of strong non-thermal particle beams.
Fig. 1 | The time evolution confirms the formation of the helicity barrier. 

**a.** Outward-/inward-propagating fluctuation amplitudes (left y axis) and imbalance (right y axis). 

**b.** Energy budget and heating, illustrated through the ion heating rate (Q) and the small-scale resistive dissipation (εη). The balanced part of the injected energy (−2εz) saturates early by t ≈ 3τs (∼ 3Lc/vs, blue shaded region); this is demonstrated by the saturation of εz (blue line in a) and the resistive dissipation (purple line in b), which absorbs only a small fraction of the input energy εz ≳ 2ε − εη. Eventually, by t ≈ 14τs (orange shaded region), ion heating absorbs the remainder of the injected energy so that Q ≈ εz − εη (brown line in b), halting the growth of εz. A numerical cooling effect has been removed to compute the energy budget (Methods and Extended Data Fig. 1).

If combined, these two heating paradigms—via Alfvénic turbulence or ICWs—can conceivably satisfy the fast-wind heating requirements described above, maintaining an abundant source of perpendicular ion heating well above the solar surface. Here we assess whether a newly discovered effect, termed the helicity barrier14, can fulfill this role by obstructing the dissipation of collisionless Alfvénic turbulence into electron heat. Using six-dimensional, high-resolution, hybrid-kinetic simulations, we explore the effect of the helicity barrier on collisionless turbulent heating, choosing parameters to match the conditions observed in fast wind streams as closely as possible. We assess the relevance of our results to the solar wind by comparing detailed features of the turbulent spectra and ion distribution function to observations from PSP and other spacecraft.

The helicity barrier

Solar-wind turbulence is imbalanced (possessing cross helicity), meaning that it is energetically dominated by Alfvénic structures that propagate outwards from the Sun (designated z; the inward-propagating component is designated ẑ). The theory of highly perpendicular (kx ≫ ky) perturbations in a collisionless plasma predicts that, at scales larger than the ion gyroradius ρi, imbalanced turbulent Alfvénic energy in εz can cascade towards smaller scales (larger ky), as required to heat the plasma.10 In contrast, at sub-gyroradius scales kxρi ≳ 1, magnetic helicity conservation implies the opposite: imbalanced energy can cascade inversely, towards larger scales.13,22,23,24 The effect is directly analogous to two-dimensional hydrodynamic turbulence, where enstrophy conservation causes energy to cascade inversely, creating large-scale vortex structures. However, unlike in hydrodynamics, in low-β plasmas, cross helicity at kxρi ≲ 1 transforms conservatively into magnetic helicity at kxρi ≳ 1 (the system conserves a generalized helicity24). This implies that an imbalanced energy flux arriving at kxρi ≲ 1 from large scales cannot cascade to arbitrarily small scales.

Mathematically, it is helpful to separate the turbulent energy flux ε into components associated with the large-scale outward (∥ε∥) and inward (∥√ε∥) propagating fluctuations: ε = ∥ε∥ + ∥√ε∥. Generalized helicity conservation prohibits the conversion of ∥ε∥ into ∥√ε∥ at any scale (note, however, that at kxρi ≳ 1, εy is associated with a mixture of outward- and inward-propagating KAWs). Because a forward cascade must be balanced (∥ε∥ ≃ εy) at kxρi ≳ 1, only a small portion ∼2ε of the energy flux can cascade to the small scales at which it will heat electrons.13 The rest of the flux, ∼(εy − εy), is stuck—it hits the helicity barrier and thus remains at scales kxρi ≲ 1. If the system is forced continuously, this large-scale energy grows in time with a decreasing parallel correlation length2, as expected from critical balance25. We show that this growth eventually funnels the turbulent energy into a spectrum of ICW fluctuations, heating the ions, which absorb the majority of the energy flux.

Numerical method

Our simulation used the PEGASUS+ code26, which solves the hybrid-kinetic equations with isothermal electrons using the particle-in-cell method. The system was strongly magnetized with mean magnetic field B0 = −0.2Bz. Alfvén speed vA = B0/√4πnmI, and initial ion βi, βi ≡ 8πnkBTi/B0 = 0.3 (where mi and kB Ti = miυ0/2 are the ion mass, number density and temperature, respectively, with υ0 the ion thermal velocity and kB the Boltzmann constant). Perpendicular (x- and y-directed) ion-velocity fluctuations u⊥ and magnetic fluctuations B⊥ were driven at large scales and correlated to create imbalance, with x ≡ u⊥ + B⊥/√4πnmI ≁ x ≡ u⊥ − B⊥/√4πnmI (zerturbations propagate in the +z direction). The energy-injection rate ε and cross-helicity-injection rate εz ≃ εy − εy ≃ 0.9ε were constant in time. Plasma heating was strongly influenced by the amplitude δB⊥/B0 and spectral anisotropy kρi/k0 of fluctuations with kxρi ≲ 1. To reach realistic values without simulating the enormous scale separation of the real solar wind, we used a highly elongated box with dimensions Lx ≈ Lz = 6L⊥; given the box size Lxρi ≡ 2πρi/Lx ≈ 0.05, this gave conditions near kxρi ≲ 1 that were comparable to those observed.19 The elongated domain also implies that the timescales we probed were rapid compared with the solar wind’s outer scales and its expansion rate, justifying the external forcing to represent a turbulent flux from larger scales and our neglect of expansion effects. The simulation’s resolution was Nρi × Nτs = 392 × 2,352 cells, so that the smallest resolved scales had kxρi ≃ πNτs/L⊥t = 10. Most other simulation parameters, including a (hyper-)resistivity ηi that dissipated small-scale magnetic energy, were chosen to match a previous balanced turbulence simulation with βi = 0.3 (ref. 14). This allowed direct comparison with their spectra and heating. Further details are provided in Methods.

Results

The simulation’s evolution through time, shown in Fig. 1, exhibits several features that are expected from the helicity barrier12 but not from other theories of imbalanced turbulence12,17. Figure 1a shows the growth of the root-mean-square amplitudes zrms/vA ≡ (ẑZ)1/2/vA and imbalance (normalized cross helicity) σi, = 2(4πnmIu⊥B⊥)/4πnmIu⊥ + B⊥2 (where (...) denotes a box average). Because ẑZ/vA ≃ 2u⊥ρivA ≃ 2B⊥/B0, the final δB⊥/B0 ≈ 0.26 was nearly twice that of the balanced simulation (δB⊥/B0 ≈ 0.14). We see that z− saturates quickly, by time t ≈ 3τs, whereas saturation of z occurs only after t ≈ 14τs (τx ≡ Lx/vA is the Alfvén time). This is expected: energy in z− can cascade to kxρi ≳ 1 through standard KAW turbulence, whereas most of the energy in z− cannot cascade past kxρi ≲ 1 due to the helicity barrier. The imbalance saturates at σi ≃ 0.98, which, although large, is regularly observed by PSP.14 Figure 1b shows the ion heating rate
Unlike the imbalanced case, does not exhibit a steep transition range in either $k_i$ or $k_p$, which is probably also the case in the solar wind\cite{44}. Note that spectra are adversely affected by particle noise for $k_p \rho \gtrsim 3$ (Methods and Extended Data Figs. 1 and 2). Evidence for the presence of ICWs in the saturated state is shown in Fig. 3. The projected magnetic field lines and electric field (Fig. 3a) reveal the coexistence of $k_d = 1$ parallel structure with the sub-$\rho_i$ striations of KAW turbulence in the perpendicular plane. More quantitatively, Fig. 3b shows the two-dimensional $(k_x, k_y)$ spectrum of $\mathbb{B}_i$ (Methods and Extended Data Fig. 3). Several features are manifest: first, unlike in the analogous balanced-turbulence simulation\cite{30}, the outer-scale parallel correlation length is markedly smaller than $L_p$ because it decreases with increasing amplitude to maintain critical balance, $k_d \rho_i \approx k_z \rho_i$ (the region of maximal spectral power moves upwards in time as $z^+ \gtrsim 1$ grows); second, the cone of maximal spectral power seems to steepen as the turbulence moves to smaller scales, creating small parallel scales faster than the canonical result $k_i \sim k_{\perp}^{-2}$ (or $k_i \sim k_{\perp}^{1/2}$ for aligned turbulence); third, there is a clear spectral bump at $k_d \rho_i \approx 0.8$ and $k_i \rho_i \approx 0.8$, the signature of ICWs. By integrating the energy spectrum over modes with $k_i \rho_i \leq k_p$, we estimated that these ICW modes contain $\approx 1\%$ of the total energy; this fraction grows by a factor of $\approx 50$ from $t \approx 5\tau_i$ to saturation and exceeds that of saturated balanced turbulence by a factor of $\approx 30$. We confirm that these modes are in fact ICWs in Fig. 3c, which shows the normalized magnetic-helicity spectrum $k \mathcal{H}_i / \mathcal{E}_B$, where $\mathcal{H}_k = \langle B^*_i \mathbb{B}_i^* - B^*_i \mathbb{B}_i^* \rangle / k$, binned in $(k_i, k_z)$. As predicted\cite{42} and observed\cite{43}, ICWs (with $k_i \rho_i \lesssim k_p$) are characterized by $k \mathcal{H}_i / \mathcal{E}_B < 0$, whereas $k \mathcal{H}_i / \mathcal{E}_B > 0$ for $k_i \rho_i \gtrsim k_p$. This results from the intrinsic polarization of ICWs and their propagation direction; $k \mathcal{H}_i / \mathcal{E}_B \approx -1$ indicates that ICWs propagate almost exclusively in the $+z$ direction, like the large-scale $z^+$.

Evidence that plasma heating occurs through ICWs is provided in Fig. 4. Quasilinear cyclotron-heating theory\cite{30} is based on the idea that ions and ICWs interact strongly if the wave frequency $\omega_i$ is resonant with the Doppler-shifted ion gyromotion. When ICWs exist across a range of $k_i$, the process flattens the ion distribution function $f(w_i, w_z)$ along specific ‘scattering contours’, which can be computed\cite{45} from $\omega_i$ for waves of a particular $k_i$ ($w_i$ and $w_z$ are the field-perpendicular and field-parallel velocities of ions in the frame moving with the plasma). Theory suggests that, because the scattering contours steepen with increasing $k_i / k_{\perp}$, heating by oblique ICWs generates an $f$ that increases along the scattering contours of parallel ICWs\cite{47}. The consequences are twofold: first, oblique-ICW heating generates parallel ICWs, explaining the dominance of $k_i \ll k_p$ modes in Fig. 3b; second, in a quasi-steady state, $f$ is nearly flat along the parallel ICW scattering contours. We plot these along with $f(w_i, w_z)$, in Fig. 4a. There is exceptionally good agreement at saturation for $w_i \lesssim -w_{\perp i} d_i$ particles, which are those that can resonate with $k_i \rho_i \lesssim 1$ ICWs propagating in the $+z$ direction (Methods). The time evolution is also telling: the quasilinear flattening starts at large $|w_i|$ and moves upwards as time advances, which is expected because there is more power in low-$k_i$ modes that resonate with high-$|w_i|$ ions. Figure 4b shows the perpendicular energy diffusion coefficient $D_{\perp i}^\perp$, computed for $w_i < -w_{\perp i} d_i$, using the time evolution of $f$, (ref. \textsuperscript{46}) and validated by computing $(\mathcal{E}_B, w_i)$ directly from particle trajectories\cite{47}. Quasilinear ICW theory predicts\cite{47} that $D_{\perp i}^\perp \propto w_{\perp i}^{-2}$, as seen in Fig. 4, whereas a stochastically heated plasma has approximately constant $D_{\perp i}^\perp$ for $w_i \sim v_{\mathbb{B}}$. This quantitatively confirms the dominance of quasilinear ICW heating and we find no other evidence for stochastic heating in this simulation, although
it is possible that it could govern saturation under different conditions (for example, lower $\beta$; Methods and Extended Data Fig. 4). Finally, we see a clear flattening of $f(w, w_\perp)$ at $w_\perp \approx v_A$ and small $w_\perp$, which forms a modestly super-Alfvénic beam feature in the direction of dominant wave propagation, similar to those observed in the fast solar wind\cite{1990ApJ...354..947L,1993ApJ...400..295S,1996ApJ...458..778G}. By comparing $\partial f/\partial t$ with $\langle E, w_\perp \rangle$ (Extended Data Fig. 5), we confirm that this arises through Landau damping of Alfvén waves as their phase velocity increases near $k_\rho \rho \approx 1$\cite{2018ApJ...866...76Q}.

**Discussion**

Together, Figs. 3 and 4 provide strong evidence for heating through ICWs generated by imbalanced Alfvénic turbulence regulated by the helicity barrier. Given the excellent match with in situ observations of spectra, helicity and distribution functions, we suggest that the same is true in the fast solar wind, reconciling the paradigms of low-frequency Alfvénic turbulence and ICW heating. If, as suggested above, ICWs are produced by the decrease in outer parallel scale with increasing amplitude, then the turbulence will saturate once it has grown sufficiently for $k_\rho \approx 1$ scales to be reached (where Alfvén waves become ICWs) before the $k_\rho \approx 1$ scales (where Alfvén waves become KAWs). If large-scale fluctuations are critically balanced with spectrum $E_B \propto k_\rho^{-3/2}$, this suggests a critical outer-scale ($-L_\perp$) saturation amplitude for appreciable ICW heating given by $\delta z^+ \approx A \beta^{1/2} (\rho / L_\perp)^{-1/3}$, with our simulation giving a proportionality factor $A \approx 0.2$ (an observationally testable prediction). A corollary is that the energy-injection rate sets only the timescale to reach saturation and not the amplitude, unlike standard viscous damping. Correspondingly—and in contrast to other possible ion-heating mechanisms—if the amplitude of the fluctuations is too small at some time (or in some region of a radially stratified wind), then energy is not deposited into electrons. Instead, the helicity barrier halts the cascade, storing the energy in fluctuations that keep growing, to eventually heat ions through ICWs. This mechanism is expected to become more robust with decreasing $\beta$, at least for $\beta \gg m/e m_e$, electron heating effects are important at $k_\rho \approx 1$. Our finding that the helicity barrier occurs even at modest $\beta (\beta \approx 0.3)$ indicates that it should apply nearly everywhere in the corona. Thus, the ratio of electron to ion heating—an important input to larger-scale models of the solar wind and other astrophysical objects—is simply $Q/Q_e \approx e_i / e$ for saturated imbalanced Alfvénic turbulence, nearly independent of $\beta$ for $m/e m_e \ll \beta \ll 1$.

A heating ratio that is controlled by imbalance is appealing to explain the generation of fast and slow wind streams by reflection-driven Alfvénic turbulence. Fast wind—which emerges from coronal holes with lower expansion factors\cite{2013ApJ...764..177A} and thus less wave reflection and larger $e_i / e$—is observed to have hotter protons, strong minor-ion heating, cooler electrons, larger imbalance and steep transition-range spectra. Slow wind—which arises from less...
ordered fields involving closed loops and/or larger expansion factors, and thus has smaller $\epsilon_{\text{diss}}/\epsilon$, or even $\epsilon_{\text{diss}}/\epsilon \ll 1$ in a closed-field region—observed to have cooler protons, little suprathermal minor-ion heating, hotter electrons and less imbalance, and does not usually exhibit a steep kinetic transition range. Although reflection-driven turbulence models can already reproduce the observed correlation of expansion factor with wind speed based on the radial location of energy deposition\cite{34,35}, the additional physics afforded by the helicity barrier yields interesting implications. First, it would naturally explain the other aforementioned correlations of electron and ion thermodynamics with wind speed. Second, because of the high thermal speeds of electrons, a given quantity of energy deposited into ion heat generally drives a higher asymptotic wind velocity than if electrons are heated\cite{31}. We propose that the helicity barrier could act as a ‘switch’ effect, supplementing other turbulent heating physics: low-expansion regions with a robust barrier would heat predominantly ions at large radii, ideal conditions for generating fast wind speeds; high-expansion regions would break the barrier, depositing energy into electron heat and exacerbating the inefficient acceleration that results from heating at lower radii. Thus, by linking plasma thermodynamics to magnetic-field expansion, it is plausible that the helicity barrier plays an important role in generating the bimodal speed distribution of the solar wind.

Fig. 4 | The evolution of the ion distribution function $f(w_i, w_j)$ shows how ICWs heat ions. a. Structure of $f$ at the times indicated, shown with a logarithmically spaced colour bar and contours. At saturation (right), $f$ aligns nearly perfectly with the scattering contours of parallel ICWs (dashed lines) for $w_i \lesssim -w_j$, which is the region of velocity space where resonant ICWs exist. b. Perpendicular energy diffusion coefficient. Solid lines include only $w_i < -w_j$, the dashed line includes all $w_i$ with the lower values of $D_{th}^\perp$ indicating less heating for $w_i > 0$.

Methods

Hybrid-kinetic simulation method. The equations of the hybrid-kinetic model solved by PEGASUS++ are\cite{34}:

$$\frac{\partial f}{\partial t} + v \cdot \nabla f = \frac{c}{m} \left( E + \frac{v}{c} \times B \right) + \frac{F^\parallel + F^\perp}{m} - \frac{\partial f}{\partial v_i} = 0, \tag{1}$$

$$\frac{\partial B}{\partial t} = -c \nabla \times (E + F^\parallel) + \eta_i \nabla^2 B, \tag{2}$$

$$E = \frac{u_i \times B}{c} + \frac{(\nabla \times B) \times B}{4\pi c n_i} - \frac{\tau_i n_i}{\epsilon n_i}, \tag{3}$$

The kinetic equation (1) is solved in a six-dimensional $(x,v)$ space using a particle-in-cell method, while Faraday’s law equation (2) and the kinetic Ohm’s law equation (3) are solved on a three-dimensional grid. A single ion species of charge $e$ and mass $m_i$ is assumed, $c$ is the speed of light, and $F^\parallel$ and $F^\perp$ are the forcing terms, described below. In equation (3), $n_i = \int f \, dv_i$ and $u_i = \int f \, dv_i \, v_i$ are computed from $f$. Equations (1)–(3) are derived from the two-species Vlasov equation by expanding the electron equation in $m_i/m_e \ll 1$, assuming quasi-neutrality $n_i = n_e = n$ and isothermal electrons (temperature $T_e$). PEGASUS++ uses a second-order accurate predictor–corrector scheme to enforce the kinetic Ohm’s law\cite{31} and has been highly optimized for efficient operation on large supercomputing systems. The hyper-resistivity in equation (3) is not intended to represent a true physical resistivity (it does not contribute to the electric field in the particle push), but is included to dissipate energy in the magnetic field at the smallest grid scales. This is a proxy for electron heating in the model.

Simulation parameters. As discussed in the main text, the simulation domain is elongated by a factor of $6$ ($L_x = L_y = 67.5 d_B$ and $L_z = 404.7 d_B$, but with cube grid cells $N_x = N_y = 392, N_z = 2352$) to realize realistic solar-wind conditions near $k_{\rho/B} = 1$. These conditions can be estimated roughly by taking the outer-scale fluctuations to be in approximate critical balance $\delta B / B_0 \approx L_z / L_x$, and assuming a magnetic spectrum $\delta B (k_\perp) \propto k_\perp^{-3}$ and $k_{\rho/B} \approx k_\perp^{\approx 5}$ in the magnetohydrodynamic (MHD) inertial range\cite{36}. Given the outer perpendicular scale $k_{\rho/B} \approx 2\pi/\delta B_0 L_x \approx 0.05$, this suggests that at $k_{\rho/B} \approx 1$ the spectral anisotropy is $k_{\rho/B}^\perp \approx \tan 86.5^\circ$ with fluctuation amplitude $\delta B / B_0 \approx 0.06$. This is comparable to observed values (in fact, it is somewhat more anisotropic than many measurements; see fig. 1 in ref. 31), justifying the appropriateness of our study to solar wind heating. We also note that the elongated simulation domain implies that the inferred outer scale (the scale at which $\delta B \approx B_0$) has an extremely long turnover time $\tau_{\text{exp}} \approx (L_z / L_x)^2 \tau_A$ (where $\tau_A = L_x / c$ is the volume). Because these outer-scale timescales exceed the duration of the simulation ($\approx 18 t_A$), it is apt to consider the outer-scale forcing in our simulation as representing a turbulent flux of energy arriving from larger scales, even if the outer scales are decaying\cite{37} (as relevant to the solar wind). Similarly, we note that the direct effect of solar-wind expansion is negligible at these scales. A simple estimate can be obtained by matching the simulation ion-inertial scale to that in the solar wind, defining the expansion time as $\tau_{\text{exp}} \equiv R/U$, where $R$ is the heliocentric radius and $U$ is the bulk solar-wind velocity. Using parameters similar to PS9’s first perihelion\cite{38}, which also had $B_0 \approx 0.3$, yielded $\tau_{\text{exp}} \approx 8 \times 10^{-4}$ (1/350 km s$^{-1}$) (R/350 km s$^{-1}$) $^{-1}$ (U/80 km s$^{-1}$) $^{-1}$, illustrating that the expansion of the box is negligible ($\approx 1\%$) over the simulation’s duration.

Other parameters of the simulated domain were chosen to match the balanced simulation of ref. 11, which had a resolution of $N_x \times N_y \times N_z = 200^3 \times 1,200$, a smaller box with $k_{\rho/B} \equiv 2\pi/\delta B_0 L_x \approx 0.1$ and saturated amplitude $\delta B / B_0 \approx 0.14 \approx L_z / L_x = 1/6$. The energy-injection rate was computed from $\epsilon / V = C_\text{inj} \delta B_0 / \tau_A u_i$, with $u_i / c = \tau_A / L_x$ and a Kolmogorov constant $C_\text{inj} = 0.29$ that matches to that measured from the balanced simulation. This implies lower energy injection per unit volume in the unbalanced case, compensating for the slower outer-scale motions in its larger box; stated differently, with this $C_\text{inj}$, turbulence in the larger $k_{\rho/B} = 0.05$ box would saturate with $\delta B / B_0 \approx 0.14$ if the forcing were balanced, which also implies that it would have a smaller amplitude (smaller $k_{\rho/B}$) at $k_{\rho/B} = 1$. The simulation was initialized by randomly drawing particle velocities from a stationary Maxwellian distribution with temperature $T_	ext{inj}$. We used $N_x \times N_y \times N_z = 216$ particles per cell and the full-f method\cite{34}, with particles initially evenly distributed in the cell. In computing the gridded moments of $f$, we used two filter passes to reduce the impact of the particle noise\cite{34}.

The value of the hyper-resistivity, $\eta_i \approx 2.4 \times 10^{-3} d_B^2$, was also chosen to match that of the balanced simulation (here $d_B$ is the ion gyrofrequency). Its value is not intended to represent reality, but just to absorb magnetic energy that cascades to the grid scales of the simulation. We have tested the impact of this
choice by restarting the simulation in the saturated regime (at $t=16.3\tau_b$) with $q_1 \approx 1.6 \times 10^{-3} \rho_i k_{\perp}^2$. This modification extends $E_k(k_\perp)$ to smaller scales where it flattens further, as expected, while changing $c_f$ only slightly and making no noticeable difference to the diagnostics presented in Figs. 3 and 4. We are thus confident that the chosen $q_1$ is appropriate and that $c_f$ is not adversely affected by grid-scale effects.

**Forcing.** The plasma is driven at the largest scales in the box with the forcing terms $F_{\parallel}^0$ and $F_{\perp}^0$ in equations (1) and (2), respectively. Given that our elongated box is designed to represent a small patch of a much larger system, these terms are supposed to crudely mimic the effect of the larger-scale turbulence on our box’s outer scale. As ref. 35, we opted to design the forcing to inject energy and cross helicity at a constant rate through time, which necessarily requires $F_{\parallel}^0$ and $F_{\perp}^0$ to be adjusted to respond to the state of the plasma. We thus defined $F_{\parallel}^0 = f_U T$ and $F_{\perp}^0 = f_B T$, where the forcing function $f_U$ was divergence-free and purely perpendicular to $B$ (no $z$ component), which implies that the forcing is nearly purely Alfvénic (its compressive part is small). $F_{\parallel}^0$ was evolved through an Ornstein–Uhlenbeck process for each mode with $2\pi L_z \leq k \leq 4\pi L_z$, where $L_z$ represents each direction $x$, $y$ and $z$ (the correlation time is $\tau_{\text{corr}} = r_z/2$). At each time step, we computed $f_B$, $F_{\perp}^0$, and $B_{\perp} = (\nabla \times F_{\perp}^0)$ and adjusted the values of $f_U$ and $f_B$ to make the injected energy and cross helicity equal to their desired values ($c_f$ and $c_{\perp}$, respectively). This process involved inverting the curl to compute $F_{\perp}^0$, which was carried out using a Fourier transform; by adding the magnetic force in this way and evolving $B$ using the standard constrained-transport algorithm of PEGASUS++, we ensured that $\nabla \cdot B = 0$ to machine precision.

This forcing method is an extension (to allow for imbalance) of the default routines implemented in the ATHENA code, and has thus been used in a number of previous works (for example, ref. 36). We tested it by separately measuring the energy and cross-helicity injection, which agree almost perfectly with the input values, and by testing the full energy budget of the simulation (see below). A possible concern with this method is that the Alfvén waves can change more rapidly than the spatial form of the force $F$. In the simulation, we saw occasional sudden changes in $f$ and $f_b$ that seem to be caused by the plasma flow and magnetic-field perturbations (which dominantly propagate in the $z$ direction) moving out of phase with the large-scale spatial structure of the forcing (determined by the time-averaged $F$). To address this, we restarted the simulation at $t = 10\tau_b$, with a modified version of the forcing that limited the change in $f$ and $f_b$ across a timestep $\delta t \approx (\pm 1\tau_b/10)$, where the $\pm$ was decided by whether $f$ and $f_b$ were above or below the optimal value that gave the input energy and cross-helicity injection. This reduced the sudden changes (high-frequency power) in $f$ and $f_b$ and at the cost of causing $\varepsilon_E$ and $\varepsilon_B$ to vary considerably (by around $\pm 40\%$) in time. However, this change made no noticeable difference to the heating, distribution function and spectra, lending confidence in the robustness of our results. Finally, we have confirmed that the helicity barrier also forms robustly in the reduced model of ref. 37 when the outer scales are forced with white-in-time noise (as opposed to with constant energy and cross-helicity injection). Thus, we do not expect our results to be particularly sensitive to the design of the forcing.

**Electric-field noise and numerical cooling.** A persistent problem for the particle-in-cell method is the influence of electric field noise, which arises due to random density noise from the finite number of numerical particles. In the full-$f$ hybrid kinetic method with the PEGASUS++ algorithm, a key impact of this noise is a numerical cooling, which slowly drains thermal energy from the system. This was removed from Fig. 1 for clarity, but can be accurately assessed by computing the thermal energy budget, which is shown in Extended Data Fig. 1. The sum of the various contributions to the total rate of change of energy, shown by the thick black line, would equal zero if energy were conserved, but is instead negative, indicating numerical cooling. This feature—which clearly contemplates an urgent need to develop better turbulence models at the low--turbulent timescales still holding for the transition range, which makes the sub-$\rho_i$ fluctuations very small in magnitude. Again, it can only be ameliorated by increasing the number of particles per cell, which yields small gains for large computational expense (the noise scales $\propto N_{\text{tot}}^{-1/2}$).

If one subtracts the noise spectra from the saturated spectra (which is probably a reasonable procedure because the noise power is dominated by modes of high $k$, which, whereas the majority of high-$k$ power at saturation resides at lower $k$, our modifications to $c_f$ and $c_{\perp}$ re-flattening at $k_{\perp} > 2$ and an $k^{-3/2}$ KAW range in $E_k$ (not shown).

Finally, we note that a different preliminary simulation with a value of $\epsilon$ three times larger, four filter passes and $N_{\text{tot}} = 128$ reproduced the key early time features discussed above, including a double-kinked electric field spectrum (this simulation was then used until $t = 5\tau_b$, which gives the largest magnetic-field perturbations and a double kink). In addition to causing the large-scale energy to grow faster in time, the larger $\epsilon$ in this case caused the noise in spectra and the numerical cooling to be proportionally much smaller. We are thus confident that the physical features reported in the main text are robust.

**Measurement of the parallel spectrum.** In Figs. 2 and 3, we measured parallel spectra using a field-line-following method, which we describe here. While structure–function methods are more commonly used to study anisotropic MHD turbulence, the extremely steep parallel spectra (up to $k^{-1}$) caused by the helicity barrier are not well captured by structure functions. Our field-line-following method computes a $(k_\parallel,k_\perp)$ spectrum by first constructing $N_{\text{tot}}$ magnetic field lines by solving $dr/ds = B(r)/|B|$ from line start $s = 0$ to $s = L_z$, where the periodicity of the system allows $L_z = 2L_z$ to be used. The field of interest $X(s)$ (for example, $X = B_\parallel$ or $X = E_\parallel$) is then Fourier-filtered to a given bin in $k_\parallel = \sqrt{k_\parallel^2 + k_\perp^2}$, giving $X_k(s)$, which is then interpolated onto the coordinates $s(r)$ of the previously computed field lines. The spectrum $E_k(r) = \int ds' |X_k'(s')|^2$ is then computed in the $s$ direction, which becomes a single $k_\parallel$ slice of $E_k(k_\parallel,k_\perp)$. Repeating this process across a grid of $k_\parallel$ yields the full 3D spectrum. Note that if this process is applied to compute a $(k_\parallel,k_\perp)$ spectrum (that is, $r(s)$ in the $z$ direction), it yields the same result as computing the spectrum in the standard way with a 3D Fourier transform. Through experimentation, we found that $r_{\text{corr}} = 10L_z$ and $N_{\text{tot}} = N_{\text{tot}}/L_z$, gave high-quality results (although results were almost independent of $L_z$ for $L_z > 2L_z$). In addition, a Hamming filter was used to compute the $k_\parallel$ spectra because the $s$ direction was non-periodic. Because the field-perpendicular plane is assumed to be the $x,y$ plane, the method is valid in the reduced-MHD limit of $L_z/L_z \ll \Delta R$, and seems to relate to the ion-Larmor-scale $k_{\parallel,\perp} \approx 1$ (ref. 38). Our method also recovered the expected extremely steep scaling in $k_\parallel$ ($E_k \propto k_\parallel^{-2}$) predicted by theory at low $k_\parallel$ (see appendix B of ref. 38; not shown). A more quantitative comparison revealed that the most notable difference compared with the structure–function method is that the 2D spectrum is shifted upwards (to larger $k_\parallel$) than the $k_{\parallel,\perp}$ spectrum, which is very similar to the $(k_\parallel,k_\perp)$ spectrum, particularly in the $k_{\parallel,\perp} > L_z$ region of ICWs, so Fig. 3b can be used to compare to a qualitative level.

**Heating through quasilinear diffusion on ICWs.** Computation of the resonance contours for parallel ICWs. Here we describe in more detail the computation of the quantitative scattering contours. Typical Doppler scattering contours (in $k_\parallel$) are shown in Fig. 4. Quasilinear theory is based on the idea that particles interact strongly with a spectrum of ICWs if they satisfy the cyclotron-resonance condition

$$\omega_0 - k_\parallel w_{\|} = \pm 2\Omega_c, \quad (4)$$

Here $\omega_0$ is the wave's frequency and $w_{\parallel}$ is the ion's parallel velocity (both are measured in the fluid frame in which the plasma is stationary); equation (4) expresses the condition for the Doppler-shifted wave frequency to be resonant with the ion's gyromotion, which causes the wave and the particle to interact strongly. In the frame of the background wave frequency $\omega_q$, where $\omega_q = \omega_0$, the magnetic field perturbation is constant in time. This means that the electric field must be purely potential in this frame ($E = -\nabla \Phi$), and a particle should approximately conserve its kinetic energy as it scatters from the wave. Thus, if equation (4) has only one solution for each $w_{\parallel}$—in other words, when there exist only waves of a single $k_\parallel$ (for example, the parallel wave spectrum in this regime)—particles diffuse only along specific scattering contours in the $w_{\parallel}$ plane. We found that solutions (if any, that is, $\omega_0 \neq \omega_0$) are these are semicircles of constant particle energy $E_k$ in the wave's frame $E_{k,\text{wave}} = (w_{\|} - \omega_0 k_\parallel)^2 + w_\perp^2$. More generally, the scattering
contours are defined by the null solutions $\eta(w_\perp, w_i) = \text{const.}$ of the quasilinear dispersion relation:

$$\left[ 1 - \frac{w_i}{w_{\perp}} \frac{\partial}{\partial w_{\perp}} \right] \eta + \frac{w_\perp}{w_{\perp}} \frac{\partial}{\partial w_i} \eta = 0,$$

(5)

where $v_{\phi}(w_i) = v_\phi/k_i$ for the $w_i$ that satisfies the resonant condition (4). If $f_i$ is a decreasing function along the scattering contours, the quasilinear diffusion heats the plasma in the range of $w_\perp$ where there is appreciable power in waves of the corresponding resonant $k_i$; if $f_i$ is an increasing function along the scattering contours, the distribution function will become unstable, growing the wave power in the resonant range of $k_i$.

To compute the scattering contours, we assumed the cold-plasma ICW dispersion relation for parallel propagating waves ($k_i = 0$), which is

$$w_\perp = k_i v_\perp \sqrt{1 - \omega_A^2/\omega_i^2} \Rightarrow v_{\phi}(k_i) = \frac{\omega_A}{2k_i} \left( k_i^2 v_\perp^2 + 4Q^2/ \omega_i - k_i v_\perp \right).$$

(6)

A parametric solution for $\eta$ can be found by solving equations (5) and (6) for $v_{\phi}(w_i)$, and then using this in the solution of equation (5),

$$w_\perp^2 + w_i^2 - 2 i \omega v_{\phi}(w_i) = \text{const.}$$

This yields the contours:

$$w_\perp^2 = \frac{1}{2} \left( 1 + \eta^2 \right) \left( 1 - \frac{\omega_i^2}{\omega_A^2} \right) v_{\phi}^2,$$

(7)

where $\eta = k_i v_\perp/Q_i$ relates to $w_i$ implicitly through equation (4) (the full explicit solution is uninformatically complex).

The quasilinear diffusion process relies on maintaining wave power in the relevant range of $k_i$, but the simulation exhibits a steep drop in the spectrum for $k_i \geq 1$ (Fig. 3). This is probably because linear ICWs become strongly damped at $k_i \approx 1$ at $\beta \equiv 0.3$ (as can be shown by solving the hot-plasma dispersion relation). This implies a minimum $w_i = w_{\perp,\text{lim}}$ above which $f_i$ should tend to flatten along the scattering contours and below which it should not. This is computed from equations (6) and (4) by solving for $w_i$ at $k_i = d_i^2$, giving

$$w_{\perp,\text{lim}} \approx -\frac{1}{2} \left( 3 - \sqrt{3} \right) v_{\phi}.$$  

(8)

The contours (7) and the cutoff (8), which are plotted in Fig. 4a, provide an exceptionally good match to the shape of $f_i$.

Oblique ICWs. The above calculation assumes that only parallel waves exist in the plasma. In the presence of oblique waves, the dependence of $v_{\phi}(k_i)$ on $k_i$ means that a range of scattering contours exist for a given $w_i$. References 2 and 21 provide a compelling argument as to why it is the parallel ICW scattering contours that should determine the form of $f_i$ even if oblique modes provide the primary heating power (see also ref. 21). They note that the dependence of $v_{\phi}(k_i)$ on $k_i$ is such that higher-$k_i$ modes produce scattering contours that are steeper (that is, $\omega_\perp/\omega_i = \text{sh}^2\omega_i/\omega_i$) larger). A quasilinear diffusion process along these contours will thus produce an $f_i$ that is an increasing function along the resonance contours of parallel modes (equation (5)), which will be unstable and generate parallel ICWs. A spectrum of driven oblique modes will therefore generate parallel ICWs in the process of heating—effectively a kinetic mechanism for spectral transfer from oblique to parallel modes—creating an $f_i$ that is almost flat along the parallel ICW resonance contours.

This phenomenology provides a reasonable explanation for the behaviour that we observed in our simulation: in the saturated stage, a clear population of parallel ($k_i = 0$) ICWs appeared (Fig. 3b), even though the turbulent cascade of energy to such modes is probably slow compared with the power input into oblique modes (indeed, the power in parallel modes is quite small earlier in the simulation). However, we caution that there are also other explanations for the different contours, such as differences between the $\beta \approx 0.3$ and cold-plasma ICW dispersion relations or higher-order resonances.) Such a mechanism also implies that the ICW heating process can continue even after $f_i$ becomes perfectly flat along the parallel ICW resonance contours. This is an important feature of this heating process for application to the solar wind, and seems to be what we observe in the saturated state: after $t \approx 1.4\tau$, heating continues but $f_i$ expands outwards across the scattering contours. Potential complications with this scenario arise from higher-order resonances (which are possible with oblique but not with parallel ICWs) and electron damping by ICWs at $k_i < 1$ (ref. 21), but these are generally expected to be unimportant to the overall physics.

Finally, we emphasize that by the end of the simulation, the average thermal anisotropy of the full distribution was only $T_{\perp}/T_\parallel \approx 1.03$. This value was small for two main reasons: first, because $t_{\text{run}} \approx 54 \tau_\text{res}$ was longer than the simulation duration and second, because of the development of the beam, which contributed to the parallel temperature. Given that this $T_{\perp}/T_\parallel$ was far below the usual bi-Maxwellian ICW instability threshold of $T_{\perp}/T_\parallel \approx 1.7$ at $\beta \approx 0.3$ that is often used in observational studies, it is clear that the detailed shape of the distribution function must be considered to understand its stability to parallel ICWs and other wave modes.

**Computation of $D_i^{\perp \perp}$**. The perpendicular energy diffusion coefficient provides a useful quantitative diagnostic of the plasma heating mechanism. It is defined by assuming that

$$\frac{\partial \Phi_i}{\partial t} = \frac{\partial}{\partial w_i} \left( D_i^{\perp \perp} \frac{\partial \Phi_i}{\partial w_i} \right),$$

(9)

with the full expression integrated over a range of $w_i$ (this is method II of ref. 20). We also computed this result, and compared the direct measurement of the perpendicular heating from $D_i^{\perp \perp}$ with that from the perpendicular component of the parallel heating.

**Stochastic heating**. Here we address whether, in addition to quasilinear ICW heating, stochastic ion heating might play an important role in turbulence with a helicity barrier. The mechanism is of particular interest, given its prominence in previous theoretical and observational studies21,33,54. Possible scenarios could involve multiple heating mechanisms operating at a particular time, or a transition from one heating mechanism to another as $f_i$ changes shape through time. We found no evidence for such behaviour in this simulation: $D_i^{\perp \perp}$ seemed to maintain its quasilinear scaling $D_i^{\perp \perp} \propto w_i^2$ throughout, and $f_i$ did not deviate from the scattering contours as it evolved in the saturated state.

We speculate that in this simulation, the lack of stochastic heating is simply a consequence of its small electric potential $\Phi_i$ fluctuations around $k_F \rho_i \approx 1$, which were required to make ion gyro-orbits sufficiently random to cause heating21. In Extended Data Fig. 4, we compare the spectrum of $\Phi_i$ ($E_\phi$) in the imbalanced simulation and the balanced simulation19. $E_\phi(k_i, \eta_i = 1)$ grew modestly during the imbalanced simulation, despite the growth in $E_\phi$ at larger scales, because of the steep drop at $k_F \rho_i < 1$ due to the helicity barrier. Coupled with our larger box size, we see that although the imbalanced simulation saturated with larger amplitude turbulence, its ion-gyroscale $\Phi_i$ fluctuations were smaller than in the balanced run. Given that stochastic heating plays only a modest role in this balanced run (it became subdominant after several turnover times due to flattening of the core of $f_i$ ref. 24), this difference in $\Phi_i$ may be sufficient to render stochastic heating unimportant in imbalanced turbulence at these parameters.

It is unclear whether stochastic heating would play a role in other regimes or over longer timescales. Provided gyroscale fluctuations have sufficient amplitudes, stochastic heating is expected to be more robust at lower $\beta$ because more heating occurs before it is quenched by the flattening of the core of $f_i$ (ref. 24). This conclusion is supported by the $\beta = 1/9$ hybrid simulation of ref. 24. On the other hand, test-particle simulations show a strong reduction in the efficiency of heating of $w \equiv v \phi$, $v \phi (\beta \approx 1)$ particles in imbalanced, compared with balanced, turbulence.6 If a similar reduction occurs also in the $w \equiv v \phi$, low-$\beta$ regime, the effectiveness of stochastic heating in turbulence with a helicity barrier may also be limited. Further work is needed. However, it is worth noting that even if stochastic heating, rather than ICW heating, eventually absorbs the turbulent energy flux, the helicity barrier could remain a key ingredient in solar-wind heating. Just as for ICW heating, the barrier would allow turbulent fluctuations to grow in amplitude sufficiently to enable ion heating, rather than dissipating their energy into electron heating if their amplitude is initially too small.

**Landau damping and the ion beam**. An interesting feature of the ion distribution function shown in Fig. 4 is that the plateau $w_i \approx 0$, $w_\perp \approx \omega_i$. This forms a modestly super-Alfvénic beam with similar properties55 and directionality6 to those observed in the solar wind. Here we present evidence that this feature is a result of
Landau damping of perpendicular Alfvén waves as they become dispersive (speed up) near $k_i\rho_i \sim 1$. Test-particle calculations have shown this process to be highly effective\(^1\). Extended Data Fig. 5 compares the measured parallel heating of the distribution function

$$\left< \frac{\partial f_i}{\partial t} \right> \left( w_j \right) = \int \frac{\partial f_i}{\partial \vec{k} \cdot \vec{E}_i} \frac{1}{2} w_i^2 \frac{\partial f_i}{\partial \vec{k} \cdot \vec{E}_i}$$

(11)

to the parallel heating inferred from the work done by the parallel electric field $\vec{E}_i(w_i\vec{E}_i)$, which is computed during the simulation from particle trajectories. We see a clear peak in both quantities at $w_i \approx 1$. As discussed above (equation (10)), both methods measure plasma heating but they can differ by a total derivative. Their similarity in Extended Data Fig. 5—particularly the similarity in their magnitudes even at different times during the simulation when $(\partial f_i/\partial t)$ differs—suggests that parallel electric field work (that is, Landau damping) is responsible for the formation of the beam.

**Data availability**
The 6D simulations presented in this article generated approximately 30 TB of data. Interested parties are invited to contact the corresponding author to make arrangements for the transfer of those data.

**Code availability**
All analysis scripts presented in this work are available on request from the corresponding author. The PEGASUS++ code will be made publicly available in the near future in conjunction with a detailed publication about its numerical methods. Readers can contact the corresponding author to get updates.

Received: 9 September 2021; Accepted: 1 February 2022; Published online: 24 March 2022

**References**

1. Cranmer, S. R. & Winebarger, A. R. The properties of the solar corona and its connection to the solar wind. *Ann. Rev. Astron. Astrophys.* **57**, 157–187 (2019).
2. Marsch, E. Kinetic physics of the solar corona and solar wind. *Living Rev. Solar Phys.* **3**, 1 (2006).
3. Parker, E. N. Dynamical theory of the solar wind. *Space Sci. Rev.* **4**, 666–708 (1965).
4. Hansteen, V. H. & Leer, E. Coronal heating, densities, and temperatures and solar wind acceleration. *J. Geophys. Res.* **100**, 21577–21594 (1995).
5. Kohl, J. L. et al. First results from the SOHO Ultraviolet Coronagraph Spectrometer. *Sol. Phys.* **175**, 613–644 (1997).
6. Cranmer, S. R., Field, G. B. & Kohl, J. L. Spectroscopic constraints on models of ion cyclotron resonance heating in the polar solar corona and high-speed solar wind. *Astrophys. J.* **518**, 937–947 (1999).
7. Bale, S. D. et al. High speed coronal plasma in an equatorial coronal hole. *Nature* **576**, 237–242 (2019).
8. de Pontieu, B. et al. Chromospheric Alfvénic waves strong enough to power the solar wind. *Science* **318**, 1574–1577 (2007).
9. Tomszyk, S. et al. Alfvén waves in the solar corona. *Science* **317**, 1192–1196 (2007).
10. Velli, M., Grappin, R. & Mangeney, A. Turbulent cascade of incompressible unidirectional Alfvén waves in the interplanetary medium. *Phys. Rev. Lett.* **63**, 1807–1810 (1989).
11. van Ballegooijen, A. A., Asgari-Targhi, M., Cranmer, S. R. & DeLuca, E. E. Heating of the solar chromosphere and corona by Alfvén wave turbulence. *Astrophys. J.* **736**, 3 (2011).
12. Shoda, M., Suzuki, T. K., Asgari-Targhi, M. & Yokoyama, T. Three-dimensional simulation of the fast solar wind driven by compressible magnetohydrodynamic turbulence. *Astrophys. J. Lett.* **880**, L2 (2019).
13. Quataert, E. & Gruzinov, A. Turbulence and particle heating in advection-dominated accretion flows. *Astrophys. J.* **520**, 248–255 (1999).
14. Schekochihin, A. A., Kawazura, Y. & Barnes, M. A. Constraints on ion versus electron heating by plasma turbulence at low beta. *J. Plasma Phys.* **85**, 905850303 (2019).
15. Schekochihin, A. A. et al. Astrophysical gyrokinetic: kinetic and fluid turbulent cascades in magnetized weakly collisional plasmas. *Astrophys. J. Suppl. Ser.* **182**, 310 (2009).
16. Chandran, B. D. G. et al. Perpendicular ion heating by low-frequency Alfvén-wave turbulence in the solar wind. *Astrophys. J.* **720**, 503–510 (2010).
17. Chandran, B. D. G., Dennis, T. J., Quataert, E. & Bale, S. D. Incorporating kinetic physics into a two-fluid solar-wind model with temperature anisotropy and low-frequency Alfvén-wave turbulence. *Astrophys. J.* **743**, 197 (2011).
18. Vech, D., Klein, K. G. & Kasper, J. C. Nature of stochastic ion heating in the solar wind: testing the dependence on plasma beta and turbulence amplitude. *Astrophys. J. Lett.* **850**, L11 (2017).
Articles

49. Marsch, E. et al. Solar wind protons: three-dimensional velocity distributions and derived plasma parameters measured between 0.3 and 1 AU. J. Geophys. Res. 87, 52–72 (1982).

50. Verniero, J. L. et al. Parker Solar Probe observations of proton beams simultaneous with ion-scale waves. Astrophys. J. Suppl. Ser. 248, 5 (2020).

51. Li, X. et al. A kinetic Alfvén wave and the proton distribution function in the fast solar wind. Astrophys. J. Lett. 719, L190–L193 (2010).

52. Wang, Y. M. & Sheeley, N. R. Solar wind speed and coronal flux-tube expansion. Astrophys. J. 355, 726 (1990).

53. Cramer, S. R. in Solar Wind 11/SOHO 16: Connecting Sun and Heliosphere Special Publication 592 (eds Fleck, B. et al.) 159 (ESA, 2005).

54. Chandran, B. D. G. An approximate analytic solution to the coupled problems of coronal heating and solar-wind acceleration. J. Plasma Phys. 87, 905870304 (2021).

55. Byers, J. A., Cohen, B. L., Condit, W. C. & Hanson, J. D. Hybrid simulations of quasineutral phenomena in magnetized plasma. J. Comp. Phys. 27, 363–396 (1978).

56. Goldreich, P. & Sridhar, S. Toward a theory of interstellar turbulence. Strong Alfvénic turbulence. Astrophys. J. 438, 763–775 (1995).

57. Kasper, J. C. et al. Alfvénic turbulence. J. Geophys. Res. Space Phys. 121, 8705870304 (2016).

58. Wang, Y. M. & Sheeley, N. R. Solar wind speed and coronal flux-tube expansion. Astrophys. J. 355, 726 (1990).

59. Lynn, J. W., Parrish, I. J., Quataert, E. & Chandran, B. D. G. Resonant interactions between protons and oblique Alfvén/ion-cyclotron waves. In Twelfth International Solar Wind Conference AIP Conference Series Vol. 1216 (eds Maksimovic, M. et al.) 72–75 (AIP, 2010).

60. Gary, S. P. & Borovsky, J. E. Alfvén–cyclotron fluctuations: linear Vlasov theory. J. Geophys. Res. Space Phys. 109, A06105 (2004).

61. Hellinger, P., Trávníček, P., Kasper, J. C. & Lazarus, A. J. Solar wind proton temperature anisotropy: linear theory and WIND/SWE observations. Geophys. Res. Lett. 33, L09101 (2006).

62. Klein, K. G. & Chandran, B. D. G. Evolution of the proton velocity distribution due to stochastic heating in the near-Sun solar wind. Astrophys. J. 820, 47 (2016).

63. Isenberg, P. A. & Lee, M. A. A dispersive analysis of bispherical pickup ion broadening and heating of charged particles in magnetohydrodynamic turbulence. Astrophys. J. 758, 78 (2012).

64. Cho, J. & Lazarian, A. Simulations of electron magnetohydrodynamic turbulence. Astrophys. J. 701, 236–252 (2009).

65. Meyrand, R., Kanekar, A., Dorland, W. & Schekochihin, A. A. Fluidization of collisionless plasma turbulence. Proc. Natl. Acad. Sci. USA 116, 1185–1194 (2019).

66. Matthaeus, W. H. & Goldstein, M. L. Measurement of the rugged invariants of magnetohydrodynamic turbulence in the solar wind. J. Geophys. Res. 87, 6011–6028 (1982).

67. Isenberg, P. A. & Lee, M. A. A dispersive analysis of bispherical pickup ion distributions. J. Geophys. Res. 101, 11055–11066 (1996).

68. Martinović, M. M. et al. The enhancement of proton stochastic heating in the near-Sun solar wind. Astrophys. J. Suppl. Ser. 246, 30 (2020).

69. He, J. et al. Sunward propagating Alfvén waves in association with sunward drifting proton beams in the solar wind. Astrophys. J. 805, 176 (2015).

Acknowledgements

We thank B. Dorland, B. Chandran and A. Mallet for illuminating discussions. J.S. and R.M acknowledge support from the Royal Society Te Apārangi, New Zealand, through Marsden Fund grant number UOO1727 and Rutherford Discovery Fellowship RDF-U001804. M.W.K. and E.Q. were supported by the Department of Energy through the NSF/DOE Partnership in Basic Plasma Science and Engineering, award numbers DE-SC0019046 and DE-SC0019047, with additional support for E.Q. from a Simons Investigator Award from the Simons Foundation. L.A. acknowledges the support of the Institute for Advanced Study, and the work of A.A.S. was supported in part by UK EPSRC grant number EP/R034737/1. This research was part of the Frontera computing project at the Texas Advanced Computing Center, which is made possible by National Science Foundation award number OAC-1818253. Further computational support was provided by the New Zealand eScience Infrastructure (NeSI) high-performance computing facilities, funded jointly by NeSI’s collaborator institutions and the NZ MBIE, and through the PICSciE-OFF TIGRESS High Performance Computing Center and Visualization Laboratory at Princeton University. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

J.S. and R.M conceived the study. L.A., M.W.K. and J.S. developed the numerical methods and model, with J.S. and L.A. performing the simulations. Data analysis and visualization was carried out by J.S. and L.A., with all authors contributing to general understanding and interpretation of the results. The manuscript was written primarily by J.S. with M.W.K., A.A.S. and E.Q. leading revisions and editing.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41550-022-01624-z.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41550-022-01624-z.

Correspondence and requests for materials should be addressed to Jonathan Squire.

Peer review information Nature Astronomy thanks Munehito Shoda, Christopher Chen and Philip Isenberg for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2022
Extended Data Fig. 1 | Electric field noise and numerical cooling. Contributions to the energy budget per unit volume of the imbalanced simulation from the energy injection (ε/V; dashed red line), increase in thermal energy $Q_i = \frac{\partial}{\partial t} \langle \frac{m v^2_{th}}{2} \rangle$ (blue line), growth rate of mechanical energy $\frac{\partial}{\partial t} \left( \frac{nm u^2_i}{2} + \frac{B^2}{8\pi} \right)$ (green line), and resistive dissipation $\frac{\varepsilon \eta}{V}$ ($V$ is the volume and $\langle \cdots \rangle$ denotes a box average). The black line shows the total energy budget $\text{Total} = \varepsilon/V - \varepsilon \eta / V - \frac{\partial}{\partial t} \langle \frac{m v^2_{th}}{2} \rangle - \frac{\partial}{\partial t} \left( \frac{nm u^2_i}{2} + \frac{B^2}{8\pi} \right)$, which is constant and negative, indicating numerical cooling that is effectively independent of the turbulence or the heating of ions.
Extended Data Fig. 2 | The effect of particle noise on turbulence spectra. Perpendicular ($k_\perp$) spectra of the magnetic field ($E_B$), electric field $E_E$, and KAW-normalized density $E_{n_{KAW}} = \beta_i (1 + 2\beta_i) E_n$ in the saturated state (solid lines) and at very early times (averaged over $t \leq 0.2\tau_A$). The latter is from before the turbulence has developed and is thus a proxy for the noise floor in a given quantity. At the smallest scales, $k_\perp \rho_i \gtrsim 3$, spectra are only modestly above the noise floor and therefore uncertain.
Extended Data Fig. 3 | Measurement of the parallel spectrum. Two-dimensional perpendicular magnetic-field spectrum $E_\parallel (k_{\perp}, k_{\parallel}) = E_{Bz}(k_{\perp}, k_{\parallel}) + E_{By}(k_{\perp}, k_{\parallel})$ from the balanced simulation. The method recovers the large- and small-scale scalings of $k_{\parallel}$ with $k_{\perp}$ measured using structure functions (see fig. 2 of ref. 19), as well as the predicted 2D spectrum in the $k_{\perp}\rho_i < 1$ range (see appendix B of ref. 19).
Extended Data Fig. 4 | Assessment of the influence of stochastic heating. We show perpendicular spectra of the electric potential $\Phi$, computed from the curl free part of $E$. Colored lines show various times from the imbalanced simulation. The black line shows the equivalent balanced simulation, which is averaged over the early period of the simulation (between $t = 3.5\tau_A$ and $t = 4.5\tau_A$) when stochastic-ion heating absorbs the majority of the turbulent energy flux. Despite the larger turbulence amplitude in the imbalanced simulation, the electric-potential fluctuations around $k_\perp \rho_i \sim 1$ - those important for stochastic heating - are smaller.
Extended Data Fig. 5 | Development of the ion beam. We compare the rate of change parallel thermal energy (solid lines; see text) with the work done on particles by the parallel electric field $e\langle w_t E_f \rangle$ (dotted lines). The thick dark-blue lines show the saturated state and the orange-pink lines show $t = 7τ_A$. The similarity of the magnitude and general shape of the two measures of heating suggests that Landau damping is responsible for the formation of the ion beam.