



Magnetohydrodynamic Particle-in-cell Simulations of the Cosmic-Ray Streaming Instability: Linear Growth and Quasi-linear Evolution

Xue-Ning Bai^{1,2}, Eve C. Ostriker³, Illya Plotnikov^{3,4}, and James M. Stone³

¹Institute for Advanced Study, Tsinghua University, Beijing 100084, People's Republic of China; xbai@tsinghua.edu.cn

²Center for Astrophysics, Department of Astronomy, Tsinghua University, Beijing 100084, People's Republic of China

³Department of Astrophysical Sciences, Peyton Hall, Princeton University, Princeton, NJ 08544, USA

⁴Université de Toulouse, UPS-OMP, IRAP, 9 av. Colonel Roche, BP 44346, F-31028 Toulouse Cedex 4, France

Received 2019 February 26; revised 2019 April 3; accepted 2019 April 3; published 2019 May 3

Abstract

The gyro-resonant cosmic-ray (CR) streaming instability is believed to play a crucial role in CR transport, leading to the growth of Alfvén waves at small scales that scatter CRs, and impacts the interaction of CRs with the interstellar medium (ISM) on large scales. However, extreme scale separation ($\lambda \ll \text{pc}$), low CR number density ($n_{\text{CR}}/n_{\text{ISM}} \sim 10^{-9}$), and weak CR anisotropy ($\sim v_A/c$) pose strong challenges for proper numerical studies of this instability on a microphysical level. Employing the recently developed magnetohydrodynamic particle-in-cell method, which has unique advantages to alleviate these issues, we conduct 1D simulations that quantitatively demonstrate the growth and saturation of the instability in the parameter regime consistent with realistic CR streaming in the large-scale ISM. Our implementation of the δf method dramatically reduces Poisson noise and enables us to accurately capture wave growth over a broad spectrum equally shared between left- and right-handed Alfvén modes. We are also able to accurately follow the quasi-linear diffusion of CRs subsequent to wave growth, which is achieved by employing phase randomization across periodic boundaries. Full isotropization of the CRs in the wave frame requires the pitch angles of most CRs to efficiently cross 90° and can be captured in simulations with relatively high wave amplitude and/or spatial resolution. We attribute this crossing to nonlinear wave-particle interaction (rather than mirror reflection) by investigating individual CR trajectories. We anticipate that our methodology will open up opportunities for future investigations that incorporate additional physics.

Key words: cosmic rays – instabilities – magnetohydrodynamics (MHD) – methods: numerical – plasmas

1. Introduction

Cosmic rays (CRs) are an essential constituent of the Galaxy. With energy density comparable to or exceeding other components of the interstellar medium (ISM), CRs likely play an important role ISM heating, chemistry, and dynamics over timescales of $\sim 10^3$ – 10^8 yr, as well as galaxy formation and evolution over Gyr timescales (see, e.g., the reviews of Ferrière 2001; Naab & Ostriker 2017; Zweibel 2017). The dominant CR population is protons, with $\sim \text{GeV}$ particles representing the peak of the CR energy distribution; the typical CR energy density of $\sim 1 \text{ eV cm}^{-3}$ corresponds to a number density $n_{\text{CR}} \sim 10^{-9} \text{ cm}^{-3}$ (e.g., Grenier et al. 2015).

It is expected that CRs are primarily produced in collisionless shocks, most likely from young supernova remnants (e.g., Baade & Zwicky 1934), via diffusive shock acceleration (e.g., Krymskii 1977; Bell 1978; Blandford & Ostriker 1978; Drury 1983). They are then scattered across the Galaxy, primarily following magnetic field lines in the turbulent ISM and eventually escaping after a few million yr (Ginzburg & Syrovatskii 1964).

Being a nonthermal high-energy particle population, the CRs only infrequently collide with the ISM particles (this is greatest at low energies, causing ionization and heating), but the CRs and bulk ISM gas indirectly interact as they both couple to the interstellar magnetic field. Through this indirect coupling, CRs can exert pressure forces on the ISM gas that are dynamically important. Moreover, being a (trans-)relativistic plasma component, CRs in the ISM are highly buoyant, with a natural tendency to escape from the Galaxy. As a result, the CR component can possibly contribute to large-scale instabilities

and dynamo activity (Parker 1966, 1992) and to driving galactic winds (Ipavich 1975; Breitschwerdt et al. 1991; Zirakashvili et al. 1996; Everett et al. 2008; see also below).

At a microscopic level, the physics of the interaction between CRs and background thermal ISM gas is extremely rich. Considered as passive test particles, CRs can be scattered diffusively by background ISM turbulence at the gyro-radius scale (e.g., Jokipii 1966; Schlickeiser & Miller 1998; Yan & Lazarian 2002), which is important for understanding CR transport. For typical $\sim \text{GeV}$ CR particles, the gyro-radius for an ISM magnetic field of a few μG is on the order of 10^{12} cm —i.e., a microparsec. Considering the huge dynamic range from the energy-containing scales of ISM turbulence to the relevant wavelength (a factor of 10^8 in length), as well as the anisotropy of the magnetohydrodynamic (MHD) cascade at subparsec scales (Goldreich & Sridhar 1995), it is considered unlikely for background turbulence to be the dominant source of CR scattering for GeV particles (e.g., Zweibel 2013). There are even some suggestions from spectral signatures that only higher-energy ($\gtrsim 200 \text{ GeV}$) particles are affected by background ISM turbulence (Blasi et al. 2012).

Anisotropy in the distribution of CRs can destabilize Alfvén waves. Of particular interest and relevance is the CR streaming instability (CRSI; Kulsrud & Pearce 1969; Wentzel 1974; Skilling 1975a, 1975b, 1975c). Instability occurs when the bulk drift speed of the CRs exceeds the Alfvén speed in the background plasma. The instability drives the growth of Alfvén waves at the cost of free energy from CR anisotropy, and in the absence of wave damping, the resulting scattering of particles off the waves will effectively reduce the CR drift speed to the Alfvén speed. This process limits the bulk flow speed of the

CRs and simultaneously leads to energy and momentum exchange between the CRs and background gas. As a consequence, CRSI plays a fundamental role in the interaction between the CRs and ISM gas and is the key element behind the picture of CR self-confinement, as well as CR-driven winds (see, e.g., Zweibel 2013, 2017; Amato & Blasi 2018 for reviews). On much larger scales in galaxy clusters, the CRSI may also be important in regulating heating (e.g., Guo & Oh 2008; Enßlin et al. 2011; Fujita & Ohira 2011; Wiener et al. 2013, 2018; Jacob & Pfrommer 2017; Wiener & Zweibel 2018). Here, however, we shall focus on the parameter regime appropriate for CR interactions with the ISM.

Studies of CR interactions with the ISM on large scales, whether analytic or numerical, treat the CRs as fluidlike (e.g., Breitschwerdt et al. 1991, 1993; Enßlin et al. 2007; Sharna et al. 2010; Pfrommer et al. 2017; Jiang & Oh 2018; Thomas & Pfrommer 2019), generally adopting the assumption of streaming, diffusion, or some combination. This is necessary given the huge scale separation between the macroscopic scale of astrophysical systems and the tiny gyro-scale of the CRs, but this approach is subject to the prescriptions adopted to treat the microphysical effects of CR diffusion and streaming.

In recent years, there has been a resurgence of interest in CR-driven galactic winds. While most numerical studies to date have adopted diffusion prescriptions (e.g., Jubelgas et al. 2008; Booth et al. 2013; Hanasz et al. 2013; Salem & Bryan 2014; Salem et al. 2014; Girichidis et al. 2016, 2018; Pakmor et al. 2016; Simpson et al. 2016; Jacob et al. 2018), there have also been studies that compare wind outcomes when streaming versus diffusive prescriptions are adopted for the CRs (Ruszkowski et al. 2017; Wiener et al. 2017). The dynamics of winds in which CRs stream at the Alfvén speed have distinctive features, since the effective sound speed of gas increases as its density decreases, enabling steady-state CR-driven acceleration of cool gas to high velocities in typical galactic potentials (Mao & Ostriker 2018). Most recently, several studies of wind driving have considered even more sophisticated treatment of the CR fluid, implementing prescriptions for wave damping and for decoupling of CRs from cold ISM components (Recchia et al. 2016; Farber et al. 2018; Holguin et al. 2018). While all of the above studies highlight the important role played by CRs, it is also evident that different microphysical prescriptions can yield dramatically different outcomes for wind properties. Currently, however, microphysical coefficients that are adopted in simulations are motivated on phenomenological grounds or based on classical analytic theory of the CRSI, diffusion, and wave damping. Development of a modern theoretical foundation for CR fluid treatments, with coefficients calibrated from numerical studies that directly follow the microphysics, will clearly aid progress in this field.

At microscopic level, our current understanding of the CRSI is limited to quasi-linear theory. We note that on the one hand, the instability can become nonlinear when CR streaming is sufficiently strong and eventually transition to the Bell instability (Bell 2004) in more extreme environments, such as super-Alfvénic shocks in supernova remnants. On the other hand, there are additional uncertainties related to mechanisms that lead to wave damping (see discussions in Section 7.2), many of which again require numerical studies.

Numerical studies of the CRSI at a microphysical level per force must involve particle-in-cell (PIC) methods and are

inherently challenging because of the extreme ratios between the density of thermal and nonthermal particles and the scales relevant for the background thermal plasma compared to CR gyro-radii. Another difficulty is that the weak level of CR anisotropy (of order a few times v_A/c) must be accurately represented, demanding a huge number of computational particles (see Section 3). Both of these issues make it very challenging for conventional PIC methods to study the CRSI, especially in the regime relevant for the bulk of the ISM away from CR injection sites (but see Holcomb & Spitkovsky 2018 and further discussion in Section 7.3).

Recently, we developed a new hybrid MHD-PIC method (Bai et al. 2015), implemented within the *Athena* code package (Stone et al. 2008). By treating the background thermal plasma as a fluid described by MHD, we greatly alleviate the issue of scale separation between the thermal gas and CRs. Using this method, we are able to properly capture the basic properties of the CRSI, particularly in astrophysically relevant/realistic parameter regimes. For this purpose, in our first study, we simplify by ignoring wave damping and concentrating on 1D simulations that provide the highest possible resolution. By employing several novel techniques, we demonstrate that our method can accurately reproduce the linear growth of CSRI and subsequently follow the quasi-linear diffusion (QLD) of particles by scattering off of waves generated by the CRSI. We anticipate that this effort will pave the way for future studies that incorporate additional physics under a variety of environments.

This paper is organized as follows. We describe the basic theory of the CRSI in Section 2. In Section 3, we present the techniques and methodology employed in our CRSI simulations using the MHD-PIC approach. Our simulation setup is described in Section 4. In Section 5, we present simulation results from two fiducial runs. This is followed by a parameter study in Section 6. Our results are discussed in a broader context in Section 7, before we conclude in Section 8. Additional numerical aspects of our simulations are described in the Appendices.

2. Basic Theory

In this section, we briefly review the theory of the CRSI, which largely follows the description in Kulsrud (2005). We focus on the growth and saturation of the instability. While wave damping is another important ingredient for the CRSI, it will be addressed in our follow-up work and is not considered here.

2.1. Physical Mechanism

Without rigorous mathematical derivation, the basic physics of the CRSI can be understood schematically as illustrated in Figure 1. Consider a static background thermal plasma with constant density ρ_0 and constant magnetic field $\mathbf{B} = B_0\hat{x}$ with $B_0 > 0$. The Alfvén speed is thus $v_A = B_0/\sqrt{4\pi\rho_0}$. Coexisting with the background plasma is a population of CRs. Let the initial CR velocity distribution be isotropic in a frame that drifts relative to the background gas at velocity v_D , with the drift velocity along the background field \hat{x} . Without loss of generality, let $v_D > 0$ (along the direction of the magnetic field). Let $f_0(\mathbf{v})$ or $f_0(\mathbf{p})$ denote the initial particle velocity or momentum distribution function, monotonically decreasing with increasing p . Note that individual CR particle speeds

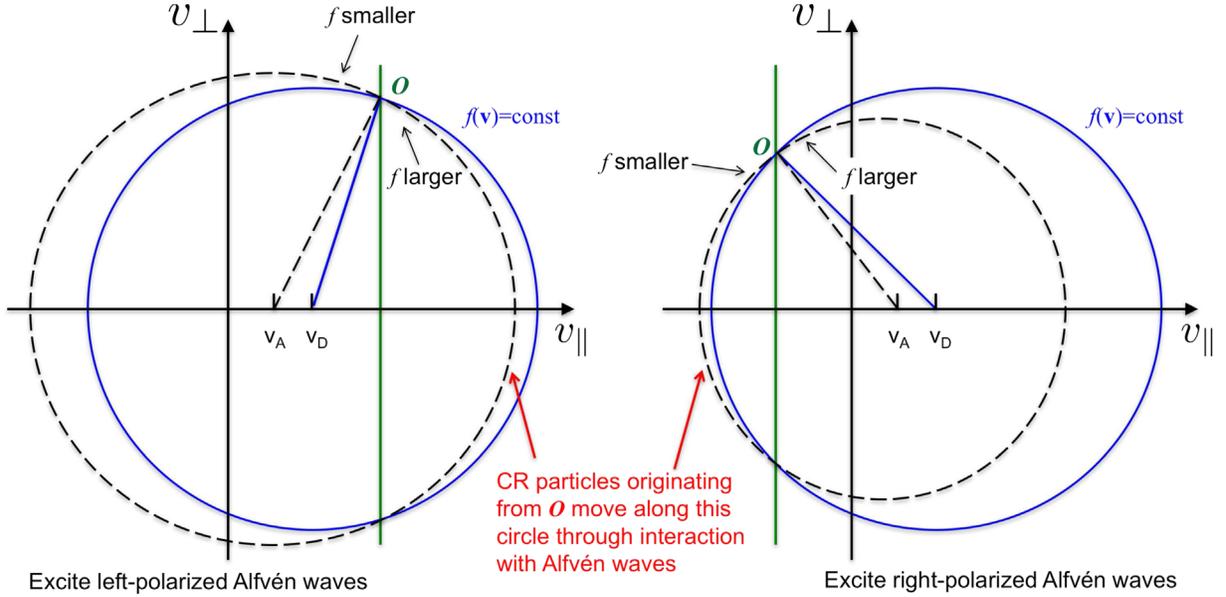


Figure 1. Schematic illustration of the basic physics of the CRSI. When the bulk CR drift velocity v_D exceeds the Alfvén speed v_A , forward-traveling CR particles excite left-polarized, forward-propagating Alfvén waves, whereas backward-traveling CR particles excite right-polarized, forward-propagating Alfvén waves. We are taking particles located near a position O in velocity space as an example. Also note that the sizes of the blue and black circles should be comparable but have to be exaggerated in this schematic figure. See Section 2 for details.

greatly exceed v_D ; hence, approximately half of the particles travel forward/backward with respect to the gas. For particles with mass m , charge q , and Lorentz factor γ , the cyclotron frequency Ω_c and gyrofrequency Ω read

$$\Omega_c = \frac{qB_0}{mc}, \quad \Omega = \Omega_c/\gamma. \quad (1)$$

The CR particles can experience gyro-resonance with circularly polarized Alfvén waves when the pattern of particle gyration matches that of the waves and the frequency of the waves experienced by the particles matches their gyrofrequency. To satisfy the first condition, forward-traveling CRs can resonate with left-handed, forward-propagating Alfvén waves, whereas backward-traveling CRs can resonate with right-handed, forward-propagating Alfvén waves (see Appendix A for our adopted conventions for left-/right-handedness). To satisfy the second condition, the parallel velocity of a particle v_{\parallel} must satisfy

$$\omega - kv_{\parallel} = \pm\Omega \approx -kv_{\parallel}, \quad (2)$$

where ω and k are the frequency and wavevector of the Alfvén wave (we only consider wavevector k along the magnetic field), and the \pm sign corresponds to right-/left-polarized modes, respectively. Note that $\omega = kv_A$, while typically, the CR velocity (and its projection) $v_{\parallel} \gg v_A$. We can thus omit ω in the above equation. In other words, the CR particles travel so rapidly that Alfvén waves simply appear as a static pattern.

Now consider CR particles in velocity space decomposed into components parallel and perpendicular to the background magnetic field (v_{\parallel} , v_{\perp}) and pick a representative particle located at point O shown in Figure 1. To facilitate the illustration, we now assume that particles are nonrelativistic (but it can be easily generalized to relativistic CRs), and hence $f_0(\mathbf{v})$ is constant along a circle centered on $(v_D, 0)$ (blue solid circle in Figure 1). Clearly, f_0 is larger (smaller) inside (outside) the circle. For linear Alfvén waves, their associated electric field

vanishes in the wave frame. Because magnetic fields do no work on particles, each particle trajectory in velocity space must follow a circle centered on $(v_A, 0)$ (black dashed circle in Figure 1). Wave-particle interaction generally leads to a QLD of particle pitch angle toward removing the gradient of $f_0(\mathbf{v})$ across the dashed circle near point O . We see that when $v_D > v_A$, for forward-traveling CR particles near O (left panel), the net outcome should be that more particles with higher v_{\parallel} on the right-hand side of point O diffuse along the dashed circle toward smaller v_{\parallel} . In net, when $v_D > v_A$, the diffusion (via resonant interactions) of particles irreversibly gives momentum to the left-handed Alfvén waves, feeding their growth. Similarly, for backward-traveling CR particles (right panel), more particles on the right side of O diffuse along the dashed circle to the left side, with the change in v_{\parallel} feeding the growth of right-handed Alfvén waves. We can also easily see that if $v_D < v_A$, instead of wave excitation, the Alfvén waves will be damped.

This discussion illustrates how the CRSI is driven by the anisotropy of the CRs. At the same time, it suggests that CRSI growth is sustained by QLD of the CRs, which produces secular and irreversible changes in the background plasma. This means that for a numerical method to capture linear growth of CRSI, it must properly capture QLD as well, which is an important fact to bear in mind when we discuss numerical methods in Section 3.

2.2. Linear Instability of Waves

We now proceed to a more quantitative description of the CRSI. Its derivation follows from standard procedures, as detailed in Kulsrud (2005). Here we only summarize the necessary results needed for setting up and analyzing the simulations.

For convenience, in the discussions below and throughout this paper, we assume that the masses of CR and background ions are identical (being m). As will be seen in Section 3.1, the

value of m has no physical significance and can always be absorbed into a factor for the charge-to-mass ratio. We thus drop it and express the momentum and energy of individual CR particles as

$$\mathbf{p} = \gamma \mathbf{v}, \quad E_p = \gamma c^2, \quad (3)$$

where \mathbf{v} is a particle's velocity and γ is its Lorentz factor. In other words, we effectively take $m = 1$. However, we do retain the symbol m in certain places for better clarity.

As defined earlier, we use $f_0(\mathbf{p}) = f_0(p)$ to denote the initial CR distribution function in the frame moving at v_D , where the distribution is isotropic, with the total CR number density given by $n_{\text{CR}} = \int 4\pi p^2 f_0(p) dp$. For Alfvén waves with perturbations proportional to $\exp[i(kx - \omega t)]$, the CR-modified dispersion relation reads

$$\omega^2 = k^2 v_A^2 \mp \frac{n_{\text{CR}}}{n_i} \Omega_c (\omega - kv_D) [(1 - Q_1) \pm iQ_2], \quad (4)$$

where n_i is the density of background ions. The role of CRs is reflected in the two-dimensionless factors

$$Q_1(k) \equiv \int dp \frac{4\pi p^2 f_0(p)}{n_{\text{CR}}} \left(\frac{p_{\text{res}}}{2p} \log \left| \frac{1 + p/p_{\text{res}}}{1 - p/p_{\text{res}}} \right| \right),$$

$$Q_2(k) \equiv \frac{\pi}{2} \int_{p_{\text{res}}}^{\infty} dp \frac{4\pi p^2 f_0(p)}{n_{\text{CR}}} \left(\frac{p_{\text{res}}}{p} \right), \quad (5)$$

where the dependence of Q_1 and Q_2 on k is reflected in the resonant momentum

$$p_{\text{res}}(k) = \Omega_c / k, \quad (6)$$

which is the minimum CR momentum to resonate with an Alfvén wave with wavenumber k (at zero pitch angle).⁵

Physically, the term associated with 1 in the $(1 - Q_1)$ factor is due to the background CR current along \mathbf{B}_0 . The Q_1 term is from the nonresonant response of CR particles to the waves. It is straightforward to show that Q_1 approaches 1 in the long-wavelength limit and monotonically decreases to zero in the short-wavelength limit. Thus, the real part of the CR contribution is bounded. The imaginary part, namely, the Q_2 term, results from the resonant response of CR particles to the waves. Note that at a given k , only particles with $p > p_{\text{res}}$ contribute to Q_2 , as expected.

In the limit $n_{\text{CR}}/n_i \ll 1$, the real part of the dispersion relation (Equation (4)) is clearly dominated by the $k^2 v_A^2$ term; hence, the term proportional to $(1 - Q_1)$ can be neglected. In this case, the dispersion relation largely gives normal Alfvén waves, which can slowly grow or damp owing to the imaginary part of the dispersion relation (i.e., the Q_2 term from the resonant response). This is the general CRSI originally derived (Kulsrud & Pearce 1969), where the forward-propagating modes, both left- and right-handed, grow at the same rates,

⁵ The dispersion relation shown here is equivalent to Equations (11)–(13) in Amato & Blasi (2009), with slight differences in sign conventions. The terms Q_1 and Q_2 arise from integrals containing df/dp (see also Equation (69) in Chapter 12 of Kulsrud 2005); this is the mathematical embodiment of the situation depicted in Figure 1, in which a gradient in f leads to wave growth. We have performed integration by parts to simplify the results, and df/dp no longer appears.

given by

$$\Gamma(k) = \frac{1}{2} \frac{n_{\text{CR}}}{n_i} \Omega_c \left(\frac{v_D}{v_A} - 1 \right) Q_2(k), \quad (7)$$

and hence the growing modes are essentially linearly polarized. Clearly, instability occurs only when $v_D > v_A$. Moreover, the maximum growth rate is of the order $(n_{\text{CR}}/n_i)\Omega_c$ at wavelengths that maximize Q_2 (to reach order unity), which typically correspond to the resonant scale with the lowest-energy CRs. Backward-propagating modes are damped.

More specifically, the initial distribution function is commonly considered to be a power law with

$$f_0(p) \propto p^{-4+\alpha}, \quad p_0 \leq p \leq p_{\text{max}}, \quad (8)$$

and $f_0(p) = 0$ otherwise. Fermi acceleration gives $\alpha = 0$ (e.g., Bell 1978; Blandford & Ostriker 1978). In typical ISM conditions, the Galactic CR population corresponds to $\alpha \approx 0.7$ for CR particle energy beyond $\sim \text{GeV}$, and $n_{\text{CR}}/n_i \sim 10^{-9}$. The Q_2 factor is given by

$$Q_2(k) = \frac{\pi}{2} \left(\frac{1 + \alpha}{2 + \alpha} \right) p_{\text{res}} \frac{p_{\text{min}}^{-(2+\alpha)} - p_{\text{max}}^{-(2+\alpha)}}{p_0^{-(1+\alpha)} - p_{\text{max}}^{-(1+\alpha)}}, \quad (9)$$

where $p_{\text{min}} = \text{Min}[\text{Max}(p_{\text{res}}, p_0), p_{\text{max}}]$. The maximum growth rate is achieved at $p_{\text{res}} = p_0$, or

$$k = k_0 \equiv \Omega_c / p_0. \quad (10)$$

This corresponds to the resonant wavelength for particles with $p = p_0$ (and zero pitch angle). In the limit $p_{\text{max}} \rightarrow \infty$, we have

$$Q_2(k) = \begin{cases} \frac{\pi}{2} \left(\frac{1 + \alpha}{2 + \alpha} \right) \left(\frac{kp_0}{\Omega_c} \right)^{1+\alpha}, & k < k_0, \\ \frac{\pi}{2} \left(\frac{1 + \alpha}{2 + \alpha} \right) \left(\frac{\Omega_c}{kp_0} \right), & k > k_0. \end{cases} \quad (11)$$

Taking into account the term from $(1 - Q_1)$ breaks the degeneracy between the left- and right-handed modes, especially when n_{CR}/n_i increases. In particular, it makes the right-handed modes grow faster and more dominant toward short wavelengths (see Figure 2), and eventually, it smoothly transitions to the nonresonant Bell mode at $k \gtrsim k_0$ (Bell 2004). This transition occurs when (Amato & Blasi 2009)

$$\frac{n_{\text{CR}}}{n_i} \gtrsim \frac{v_A^2}{p_0 v_D}. \quad (12)$$

For the simulations in this work, we fiducially adopt $v_D/v_A = 2$, $p_0/v_A = 300$, $\alpha = 0.5$, and $n_{\text{CR}}/n_i = 10^{-4}$ to 10^{-3} . The resulting dispersion relation is shown in Figure 2, which displays all features discussed above. Note that for application to realistic large-scale ISM conditions, a very low value of low $n_{\text{CR}}/n_i \sim 10^{-9}$ is appropriate, such that the left-/right-handed modes are degenerate. However, in practice, the values of n_{CR}/n_i we use in the simulations have to be much higher in order to follow wave growth and particle QLD within a reasonable computational time. We see that for our fiducial parameters, the transition to nonresonant modes occurs when $n_{\text{CR}}/n_i \gtrsim 10^{-3}$.

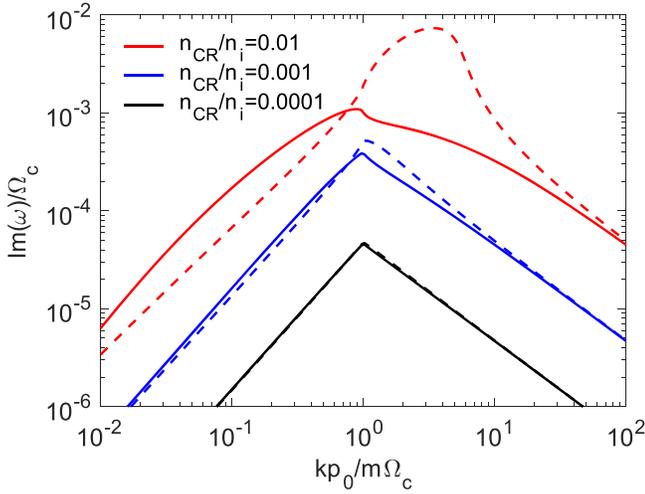


Figure 2. Linear dispersion relation of the CRSI with a truncated power-law CR distribution function (Equation (8)), with fiducial parameters $v_D = 2v_A$, $p_0/m = 300v_A$, and $s = 4.5$. Three values of n_{CR}/n_i are adopted, as shown with red (10^{-2}), blue (10^{-3}), and black (10^{-4}) lines. The growth rates for left- (solid) and right- (dashed) handed modes are shown separately. Note the gradual development and dominance of the nonresonant Bell mode at $n_{CR}/n_i \gtrsim 10^{-3}$. We focus on the regime with smaller n_{CR}/n_i , being more realistic in the bulk of the ISM.

2.3. QLD of Particles

The growth of the CRSI creates a spectrum of Alfvén waves. The intensity of the waves at wavenumber k is described by $I(k)$, so that

$$\int I(k) dk = \left\langle \frac{\delta B^2}{B_0^2} \right\rangle, \quad (13)$$

where angle brackets represent spatial averages. In Appendix A, we describe the procedures to decompose parallel-propagating Alfvén waves into different modes (left-/right-handed, forward/backward propagation), from which we obtain $I(k)$ for each mode.

In a spatially uniform medium (as in our simulations), QLD leads to pitch angle scattering in the frame of the Alfvén waves, with pitch angle θ more commonly replaced by $\mu \equiv \cos \theta$. We use subscript “w” to denote quantities measured in this frame. The evolution of f_w , the CR distribution function in the wave frame in a homogeneous system, is described by (e.g., Jokipii 1966; Kulsrud & Pearce 1969)

$$\frac{\partial f_w}{\partial t} = \frac{\partial}{\partial \mu_w} \left[\frac{1 - \mu_w^2}{2} \nu(\mu_w) \frac{\partial f_w}{\partial \mu_w} \right], \quad (14)$$

where

$$\nu(\mu_w) = \pi \Omega k_{\text{res}} I(k_{\text{res}}), \quad (15)$$

and

$$k_{\text{res}} \equiv \Omega_c / (p_w \mu_w) \quad (16)$$

is the wavenumber of the resonant waves.⁶

⁶ In calculating ν from Equation (15), we use $I(|k_{\text{res}}|)$ and choose the intensity for left-/right-handed waves for positive/negative μ_w . Also note that our normalization of $I(k)$ only accounts for the magnetic energy spectrum (it will be doubled for the total energy spectrum). Because of these normalization choices, our expression for ν does not have a factor of one-fourth.

Our simulations are conducted in the frame where the initial CR distribution is isotropic, whose relative velocity to the wave frame is $\Delta v \equiv v_D - v_A$. Using subscripts “ \parallel ” and “ \perp ” to denote components parallel and perpendicular to background field \mathbf{B}_0 , and to first order of $\Delta v \ll c$, we have for particle momenta

$$p_{w,\perp} = p_{\perp}, \quad p_{w,\parallel} = p_{\parallel} + \gamma \Delta v. \quad (17)$$

With conservation of phase-space volume, it follows that the distribution functions $f = f_w$. The conversion relations between the two frames read

$$p_w = p \left(1 + \mu \frac{\Delta v}{v} \right), \quad \mu_w = \mu + (1 - \mu^2) \frac{\Delta v}{v}, \quad (18)$$

where $v \sim c$ is the velocity corresponding to momentum p . In the quasi-linear regime, the distribution function only deviates slightly from f_0 ; hence, we can write $f(p, \mu) = f_0(p) + \delta f(p, \mu)$. Transforming to the wave frame, and to the leading order of $\delta f/f_0$ and $\Delta v/c$, we obtain

$$f_w(p_w, \mu_w) = f(p, \mu) \approx f_0(p_w) + \delta f(p_w, \mu_w) - \mu_w \frac{\partial f_0}{\partial \ln p} \frac{\Delta v}{v}. \quad (19)$$

In the wave frame, we have

$$\frac{\partial f_w}{\partial \mu_w} \approx \frac{\partial \delta f}{\partial \mu_w} - \frac{\partial f_0}{\partial \ln p} \frac{\Delta v}{v}. \quad (20)$$

In the fully saturated state, we expect $\partial f_w / \partial \mu_w = 0$; hence, the pitch angle distribution should satisfy

$$\frac{\partial \delta f_{\text{sat}}}{\partial \mu_w} \approx \frac{\partial f_0}{\partial \ln p} \frac{\Delta v}{v} \Rightarrow \delta f_{\text{sat}} \approx \mu_w \frac{\partial f_0}{\partial \ln p} \frac{\Delta v}{v}. \quad (21)$$

The amplitude of the Alfvén waves in the saturated state can also be estimated. The net momentum density acquired by the CRs after saturation is reached is

$$\Delta \mathcal{P}_{\text{CR}} = \int \mu p (\delta f_{\text{sat}}) d^3 p = \frac{4\pi}{3} \int \frac{\partial f_0}{\partial \ln p} \frac{\Delta v}{v} p^3 dp. \quad (22)$$

Note that with $\partial f_0 / \partial p < 0$, this implies that if initially, $\Delta v = v_D - v_A > 0$, in saturation, the CR distribution must have acquired a momentum deficit, $\Delta \mathcal{P}_{\text{CR}} < 0$. In the limit that all CRs are relativistic, this yields

$$\Delta \mathcal{P}_{\text{CR}} \approx -\frac{4}{3} \frac{n_{\text{CR}} \Delta v}{c} \langle p \rangle, \quad (23)$$

where $\langle p \rangle \equiv \int f_0 p d^3 p / \int f_0 d^3 p$. For the nonrelativistic CR limit, the momentum reduction is $\Delta \mathcal{P}_{\text{CR}} \approx -m n_{\text{CR}} \Delta v$.

The deficit in CR momentum must have been transferred to forward-propagating Alfvén waves. Note that there is an equipartition in kinetic and magnetic energies in Alfvén waves, with the effective momentum density in waves (the Poynting flux divided by v_A^2 ; see Kulsrud 2005) given by

$$\mathcal{P}_{\text{wave}} = \rho v_A \left\langle \frac{\delta B^2}{B_0^2} \right\rangle. \quad (24)$$

Following the discussion in Kulsrud (see also Wentzel 1974), this momentum density represents the growth of Maxwell and Reynolds stresses in waves needed to transfer the CR

momentum to the bulk gas and is exhibited as a small fractional change in the momentum of the bulk gas.

Equating $|\Delta\mathcal{P}_{\text{CR}}|$ with $\mathcal{P}_{\text{wave}}$, we can obtain the expected saturation level for magnetic fluctuations. For ultrarelativistic and nonrelativistic cases, we simply have

$$\frac{\delta B^2}{B_0^2} \approx \frac{4 \langle p \rangle n_{\text{CR}} \Delta v}{3 mc n_i v_A}, \text{ or } \approx \frac{n_{\text{CR}} \Delta v}{n_i v_A}. \quad (25)$$

The real situation (as we consider) is, of course, somewhere in between nonrelativistic and ultrarelativistic, but the dependence of wave amplitude on the basic problem parameters is clear. These relations show that the saturation level of the instability is limited by the initial free energy in CR streaming. The above upper limit for magnetic field amplitudes would be reduced in the presence of damping.

The timescale for the CRs to relax to a state of isotropy in the wave frame is characterized by the scattering frequency ν , which depends on both the pitch angle and the wave amplitudes. This means that scattering across different pitch angle ranges can take different amounts of time. Still, one may define a characteristic scattering frequency for the bulk CR population as

$$\nu_{\text{QLD}} \sim \nu(p, \mu) \sim \Omega \frac{\delta B^2}{B_0^2}. \quad (26)$$

Note that wave growth is eventually quenched as the gradient in distribution function is smeared out through QLD. This suggests that another way to estimate the wave saturation amplitude is by equating the above ν_{QLD} with Γ from Equation (7), which is indeed similar (to order of magnitude) to the more rigorous estimate discussed above.

2.4. Reflection across 90° Pitch Angle

In Equation (16), when $\mu_m \rightarrow 0$ (i.e., the pitch angle approaches 90°), $k_{\text{res}} \rightarrow \infty$, i.e., the resonant wavelength becomes infinitely small. There is little energy in the waves generated from the CRSI toward the shortest wavelength, leading to a bottleneck for particles to be scattered across this 90° pitch angle. This is a well-known problem, and without crossing the 90° pitch angle, the CRSI would saturate prematurely far from fully utilizing the free energy in the system.

This problem has been discussed extensively in the literature. Within quasi-linear theory, relaxing the magnetostatic approximation (Equation (2)) allows for additional resonant interactions when both forward- and backward-propagating waves are present (Schlickeiser 1989), covering a pitch angle of $\Delta\theta \sim 2v_A/c$ around 90°.

Magnetic mirroring is more commonly invoked in overcoming the 90° problem (e.g., Felice & Kulsrud 2001). With a magnetic moment $M \equiv p_{\perp}^2/2B$ being an adiabatic invariant, a particle will experience a backward mirror force when its guiding center travels along a positive gradient in total field strength. For a particle pitch angle sufficiently close to 90°, this mirror force can lead to reflection, thus directly jumping over the 90° barrier. With total field strength $B = \sqrt{B_0^2 + \delta B^2} \approx B_0(1 + \delta B^2/2B_0^2)$, the critical pitch angle

cosine below which mirror reflection can occur is thus given by

$$\mu_{\text{mir}} \approx \frac{1}{\sqrt{2}} \frac{\delta B}{B_0}. \quad (27)$$

One thus expects mirror reflection to take over when QLD manages to scatter particles to pitch angle $\mu \sim \mu_{\text{mir}}$.

Note that mirror reflection requires slow changes in field strength over many particle gyro-orbits. Given that a broad range of wave spectra is excited by the CRSI, including those at small scales ($\lesssim k_{\text{res}}^{-1}$) for particles of a given momentum, we can also imagine the opposite situation, where changes in field strength occur more abruptly. For instance, an abrupt change of perpendicular field by δB_{dis} would change the direction of particle motion by $\delta B_{\text{dis}}/B_0$, and in the meantime, the change of particle pitch angle is also on the order of $\sim \delta B_{\text{dis}}/B_0$. Depending on the gyro-phase, particles have a good chance to be reflected when $\mu \lesssim \delta B_{\text{dis}}/B_0$.

The (extreme) scenario outlined above falls into the category of nonlinear wave-particle interaction: considering particle orbits in perturbed fields (rather than unperturbed fields, as in quasi-linear theory) gives resonance broadening (e.g., Dupree 1966; Völk 1973; Achterberg 1981) and alleviates the 90° problem. It is more effective toward larger wave amplitudes and eventually enters the regime of strong MHD turbulence (Yan & Lazarian 2008). As we will show, our analysis suggests that this effect is primarily responsible for overcoming the 90° barrier; for more discussion, see Section 5.4.

3. Numerical Method

There are several challenges that confront numerical study of the CRSI. First, there is substantial physical scale separation, which is exhibited as follows.

1. The separation between microscopic scales of the background plasma, including the electron and ion skin depths (c/ω_{pe} or c/ω_{pi} that full- or hybrid-PIC simulations must resolve, where ω_{pe} and ω_{pi} are the electron and ion plasma frequencies) and the resonant scale of the CRSI (\sim gyro-radii of the lowest-energy, transrelativistic CRs). The ratio of the two scales is at least of the order c/v_A , which is prohibitively large for conventional PIC simulations if one were to approach realistic conditions.
2. For QLD of particles, individual CR particles must encounter a sufficient number of independent wave packets, so as to experience random-walk behavior in pitch angle evolution.
3. The issue with particles scattering across the 90° pitch angle. Particles with pitch angle near 90° have resonant wavelengths much shorter than their gyro-radii, implying that the most unstable wavelength of the CRSI must be very finely resolved.

The MHD-PIC approach uniquely alleviates the first scale separation issue by bypassing the microscopic plasma scales (ion skin depth) of the background thermal plasma. However, the remaining two issues remain challenging. Clearly, very high resolution is needed to capture scattering across the 90° pitch angle. In the meantime, when employing periodic boundary conditions, a sufficiently long simulation domain that covers many resonant scales is essential so that particles do not experience the same wave packets upon reentering the

domain. This issue will be discussed further in Section 3.4 and Appendix C.

In addition to the above challenges of physical scale separation, another challenge in modeling the regime of interest for the large-scale ISM is the huge ratio between the number densities of CR particles and background ISM particles, $n_{\text{CR}}/n_i \sim 10^{-9}$. The CR and ISM pressures, however, are generally within an order of magnitude of each other. For modeling the large-scale ISM with its extremely small n_{CR}/n_i , the MHD-PIC method is advantageous in comparison to full- or hybrid-PIC, in that there is no need to represent the background plasma with individual particles.

The third major challenge is that capturing the resonant condition still requires very large number of CR particles. For a given resonant mode, particles of a given momentum can (if capable) contribute only at a specific pitch angle. Correspondingly, only a tiny fraction of particles can potentially be in resonance with a given MHD wave mode. This issue may be seen from another viewpoint. The level of anisotropy in the CR distribution function is of the order $v_D/c \ll 1$. Such a weak level of anisotropy must be accurately represented by the angular distribution of particles, again demanding a huge number of particles in each cell.

Even though the MHD-PIC method is able to concentrate all particles in the CRs, properly capturing CRSI is still challenging when $v_D/c \ll 1$. To overcome this issue, we employ the δf method, which dramatically improves the signal-to-noise ratio and allows us to successfully simulate the CRSI with a substantially reduced number of particles (see Section 3.2 and Appendix B). Below, we describe our methodology in detail.

3.1. Formulation

We solve CR-modified MHD equations in the conservative form⁷

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}_g) = 0, \quad (28)$$

$$\begin{aligned} \frac{\partial(\rho \mathbf{v}_g)}{\partial t} + \nabla \cdot (\rho \mathbf{v}_g^T \mathbf{v}_g - \mathbf{B}^T \mathbf{B} + P^*) \\ = - \left(q_{\text{CR}} n_{\text{CR}} \mathcal{E} + \frac{\mathbf{J}_{\text{CR}}}{c} \times \mathbf{B} \right), \end{aligned} \quad (29)$$

$$\frac{\partial E}{\partial t} + \nabla \cdot [(E + P^*) \mathbf{v}_g - (\mathbf{B} \cdot \mathbf{v}_g) \mathbf{B}] = -\mathbf{J}_{\text{CR}} \cdot \mathcal{E}, \quad (30)$$

where $P^* \equiv P_g + B^2/2$, $\mathcal{E} \equiv -\mathbf{v}_g \times \mathbf{B}/c$ is the electric field, \mathbf{I} is the identity tensor, and the total energy density of the gas is defined as

$$E = \frac{P_g}{\Gamma - 1} + \frac{1}{2} \rho v_g^2 + \frac{1}{2} B^2. \quad (31)$$

In the above, ρ , \mathbf{v}_g , and P_g are gas density, velocity, and pressure, and Γ is the adiabatic index. Note that we have adopted the units where magnetic permeability is unity so that factors of $(4\pi)^{-1/2}$ that would otherwise appear with the magnetic field are eliminated.

⁷ These equations are simplified from those derived in Bai et al. (2015) in that terms related to the CR-modified Hall effect are dropped. As discussed there, these terms are important only when $(n_{\text{CR}} v_D)/(n v_A) \gtrsim 1$, corresponding to very strong CR streaming. In our simulations, this ratio is $\lesssim 10^{-3}$, and in reality, the ratio is several orders of magnitude lower.

The CR number density n_{CR} and current density \mathbf{J}_{CR} are defined as

$$\begin{aligned} n_{\text{CR}}(t, \mathbf{x}) &= \int f(t, \mathbf{x}, \mathbf{p}) d^3 \mathbf{p}, \\ \mathbf{J}_{\text{CR}}(t, \mathbf{x}) &= q_{\text{CR}} \int \mathbf{v} f(t, \mathbf{x}, \mathbf{p}) d^3 \mathbf{p}, \end{aligned} \quad (32)$$

where q_{CR} is the individual CR particle charge, \mathbf{p} and \mathbf{v} are the momentum and velocity of individual CR particles, and $f(t, \mathbf{x}, \mathbf{p})$ is the local CR momentum distribution function.

While the MHD formulation is nonrelativistic, the CR particles can be relativistic. We define an artificial speed of light \mathbb{C} for the CR particles, and the overall formulation is consistent as long as $\mathbb{C} \gg v_A$ (typical MHD velocities). For an individual particle j , we have $\mathbf{p}_j = \gamma_j \mathbf{v}_j$, with the Lorentz factor given by

$$\gamma_j = \frac{\sqrt{\mathbb{C}^2 + p_j^2}}{\mathbb{C}} = \frac{\mathbb{C}}{\sqrt{\mathbb{C}^2 - v_j^2}}. \quad (33)$$

The particle equation of motion reads

$$\frac{d\mathbf{p}_j}{dt} = \left(\frac{q}{mc} \right)_j (c \mathcal{E} + \mathbf{v}_j \times \mathbf{B}). \quad (34)$$

Note that, as mentioned earlier, we have dropped the individual particle mass in the definition of particle momentum. This mass is absorbed to the factor $q/(mc)$, representing particle charge-to-mass ratio. Also note that the physical speed of light c has no significance in any equations where it appears: inspection of Equations (29), (30), (32), and (34) shows that c appears only in the combinations \mathbf{J}_{CR}/c , $c \mathcal{E}$, and q/mc and, for numerical purposes, can be absorbed into the parameter (q/mc) .

In our simulations, ideal MHD equations for the background thermal plasma are solved using the Athena MHD code (Stone et al. 2008), which is a higher-order Godunov code with constrained transport to enforce the divergence-free condition of the magnetic field. The corner transport upwind (Gardiner & Stone 2005, 2008) method is adopted for time integration. We use the Roe Riemann solver (Roe 1981) and third-order reconstruction in characteristic variables. The CR particles are implemented as Lagrangian particles, and the coupling between CRs and the background plasma is handled by adding source terms in the gas momentum and energy updates, as indicated on the right-hand sides of Equations (29) and (30). The CR particle equation of motion (Equation (34)) is solved by the standard Boris integrator (Boris 1970). A standard triangular-shaped cloud (TSC) scheme (Birdsall & Langdon 2005) is used for interpolating grid quantities to particle locations and depositing particle quantities back to the grid (e.g., to evaluate n_{CR} and \mathbf{J}_{CR} in Equation (32)). Details of the CR implementation are described in Bai et al. (2015).

3.2. The δf Method

The original CR implementation in Bai et al. (2015) interprets individual particles as representing the full distribution function $f(\mathbf{x}, \mathbf{p})$ and evaluates physical quantities such as n_{CR} and \mathbf{J}_{CR} directly according to Equation (32). We refer to this approach as the “full- f method,” which can be subject to large Poisson noise.

In the case where the distribution function is close to some equilibrium distribution $f_0(\mathbf{x}, \mathbf{p})$, one can use the fact that f_0 is already known analytically and employ individual particles as Lagrangian markers taken to represent the difference, δf , between f_0 and the full distribution function f . This is known as the δf method (e.g., Dimits & Lee 1993; Parker & Lee 1993; Hu & Krommes 1994; Denton & Kotschenreuther 1995; Kunz et al. 2014). The basis of the δf method is the Liouville theorem, which requires that the full distribution function f be constant along particle trajectories in phase space (i.e., characteristics).

To implement the δf method, we first record the initial value of f at $t = 0$ for all particles (which is essentially f_0). Then, at every time t , we assign a weight w_j to each particle j given by

$$w_j \equiv \frac{\delta f(t, \mathbf{x}_j(t), \mathbf{p}_j(t))}{f(t, \mathbf{x}_j(t), \mathbf{p}_j(t))} = 1 - \frac{f_0(\mathbf{x}_j(t), \mathbf{p}_j(t))}{f(0, \mathbf{x}_j(0), \mathbf{p}_j(0))}. \quad (35)$$

Physical quantities such as the CR number density and current density are obtained by

$$\begin{aligned} n_{\text{CR}}(t, \mathbf{x}) &= n_{\text{CR},0} + \int \delta f(t, \mathbf{x}, \mathbf{p}) d^3\mathbf{p} \\ &\simeq n_{\text{CR},0} + \sum_{j=1}^{N_p} w_j S(\mathbf{x} - \mathbf{x}_j), \end{aligned} \quad (36)$$

$$\begin{aligned} \mathbf{J}_{\text{CR}}(t, \mathbf{x}) &= \mathbf{J}_{\text{CR},0} + q_{\text{CR}} \int \mathbf{v} \delta f(t, \mathbf{x}, \mathbf{p}) d^3\mathbf{p} \\ &\simeq \mathbf{J}_{\text{CR},0} + q_{\text{CR}} \sum_{j=1}^{N_p} w_j \mathbf{v}_j S(\mathbf{x} - \mathbf{x}_j), \end{aligned} \quad (37)$$

where $S(\mathbf{r})$ is the shaping function used in particle interpolation (i.e., the TSC scheme), and $n_{\text{CR},0}$ and $\mathbf{J}_{\text{CR},0}$ are obtained analytically from f_0 . By contrast, the full- f method corresponds to setting $w_j = 1$, $n_{\text{CR},0} = 0$, and $\mathbf{J}_{\text{CR},0} = 0$. In this way, the δf method dramatically reduces the Poisson noise of the background particle distribution, allowing the signal from the δf part to be substantially boosted.

In practice, we have found that the δf method is essential. Without employing it (i.e., using the full- f method), and using our fiducial simulation parameters, we barely observe the development of the CRSI, even using $\sim 10^4$ particles per cell, and the system is almost entirely dominated by Poisson noise (see Appendix D for more information).

We note that with the δf method, exact conservation of the total (gas and CR) momentum and energy, achieved in the full- f method, is lost. This is inevitable, since the formulation of the δf method is intrinsically nonconservative. However, the benefit from low noise is of overwhelming importance for the CRSI problem, and in practice, we find that the error in total momentum and energy is negligible throughout all of our simulations.

3.3. The κ Distribution

While the CRSI is commonly analyzed with f_0 being a simple truncated power law (Equation (8)), this is incompatible with the δf method. The reason is that the δf method requires f_0 to be finite at all p , since it serves as the normalization factor in the weighting function, Equation (35). To avoid divergence, we modify f_0 to be a κ distribution (e.g., see Summers & Thorne 1991 and references

therein for motivation),

$$f_0(p) = \frac{n_{\text{CR}}}{(\pi \kappa p_0^2)^{3/2}} \frac{\Gamma(\kappa + 1)}{\Gamma(\kappa - \frac{1}{2})} \left[1 + \frac{1}{\kappa} \left(\frac{p}{p_0} \right)^2 \right]^{-(\kappa+1)}, \quad (38)$$

which has the property that $f(p)$ is a constant when $p \ll p_0$, and $f(p) \propto p^{-2(\kappa+1)}$ as $p \gg p_0$.

The dispersion relation remains the same, except with the factors of $Q_1(k)$ and $Q_2(k)$ in Equation (5) replaced by

$$\begin{aligned} Q_1 &= \frac{\Gamma(\kappa + 1)}{s_0^3 (\pi \kappa)^{3/2} \Gamma(\kappa - \frac{1}{2})} \int_0^\infty \frac{2\pi s \cdot ds}{[1 + (s/s_0)^2/\kappa]^{\kappa+1}} \\ &\times \log \left| \frac{1+s}{1-s} \right|, \end{aligned} \quad (39)$$

$$Q_2 = \frac{\sqrt{\pi}}{\kappa^{3/2}} \frac{\Gamma(\kappa + 1)}{\Gamma(\kappa - \frac{1}{2})} \frac{1}{s_0 [1 + 1/(\kappa s_0^2)]^\kappa}, \quad (40)$$

where the dependence on k is expressed in $s_0 \equiv p_0/p_{\text{res}} = kp_0/m\Omega_c$. The Q_1 integral can be evaluated numerically, and in the limit $n_{\text{CR}}/n_i \ll 1$, the growth rate can be obtained analytically by substituting Equation (40) into Equation (7).

In Figure 3 (top panel), we show the linear growth rate in the same way as in Figure 2 but for f_0 being a κ distribution. We see that in the long-wavelength limit ($p_{\text{res}} \gg p_0$ or $s_0 \ll 1$), we have $Q_2 \propto s_0^{2\kappa-1}$. In the short-wavelength limit, we have $Q_2 \propto s_0^{-1}$. The results are consistent with a power-law distribution function in the same limits, with $\alpha = 2(\kappa + 1)$. The growth rate varies smoothly in between the two limits, with the peak growth rate only slightly reduced compared to the truncated power-law case. It is this analytical growth rate that we aim to test with further study.

In the bottom panel of Figure 3, we further show the analytical growth rate for different drift velocities at fixed $n_{\text{CR}}/n_i = 10^{-4}$. Increasing the drift velocity has a similar effect as increasing n_{CR}/n_i , which increases the overall growth rate (roughly linearly). It also makes the right-handed mode stand out, which becomes more notable for $v_D \gtrsim 10v_A$. In this paper, we restrict our explorations to $v_D \lesssim 8v_A$.

3.4. Simulation Box Size and Phase Randomization

Even equipped with the δf methods and a κ distribution function, it is difficult to accurately recover the linear growth of the CRSI with a numerical simulation. Over a large suite of tests with a wide range of numerical setups, we have found that sustained growth of the CRSI at the rate predicted by theory is only achieved when the simulation box size is extremely long. Otherwise, the desired growth rates are found only at very early stages, and they become erroneous shortly afterward (we demonstrate this in Appendix C).

Further adjusting the numerical speed of light \mathbb{C} reveals that in order to properly recover growth rates, the required simulation box length L should be such that the time for the CR particles to traverse the simulation domain, $t_{\text{cross}} = L/(\mu\mathbb{C})$, exceeds the growth timescale for CRSI, $1/\Gamma_{\text{max}}$. Given that $\Gamma_{\text{max}} \sim (n_{\text{CR}}/n_i)\Omega_c$, and using Equation (16), this

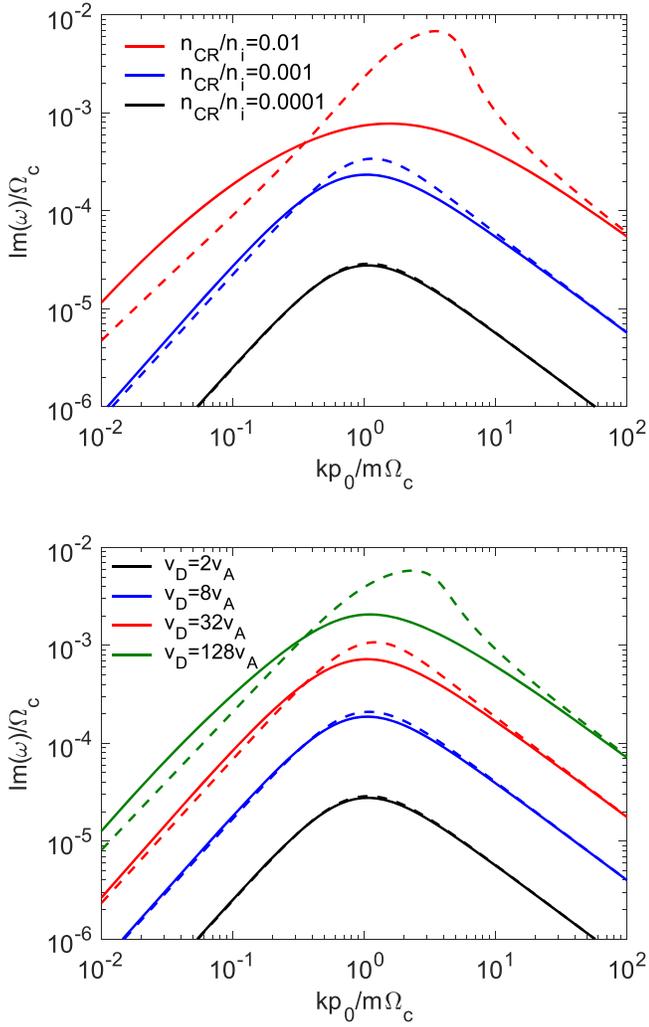


Figure 3. Linear dispersion relation of the CRSI for a κ distribution (Equation (38)) with $\kappa = 1.25$ with fiducial parameters $n_{\text{CR}}/n_i = 10^{-4}$, $v_D = 2v_A$, and $p_0/m = 300v_A$. The top panel varies n_{CR}/n_i , while the bottom panel varies v_D/v_A , as indicated in the legends. The growth rates for the left- (solid) and right- (dashed) handed modes are shown separately.

may be further translated to

$$L > L_{\min} \equiv \frac{\mathbb{C}\mu}{\Gamma_{\max}} \sim \frac{n_i}{n_{\text{CR}}} \frac{k_{\text{res}}^{-1}}{\gamma}. \quad (41)$$

This phenomenon is related to the validity of the random phase approximation for QLD discussed at the beginning of Section 3, which feeds instability growth. For CRSI, CR particles should constantly experience different wave packets throughout the growth process. With wave growth, we may regard the wave packets as being altered over the timescale of $\sim \Gamma_{\max}^{-1}$. With the fastest particles traveling at \mathbb{C} along the background field, a simulation box longer than L_{\min} guarantees that all particles experience “new” wave packets after traversing the entire box and reenter from the other side.

However, for reasonable simulation values of n_i/n_{CR} , the required L_{\min} could easily reach a few hundred or thousand times the resonant/most unstable wavelengths (using $n_i/n_{\text{CR}} \sim 10^9$ as in the bulk ISM, the number could be billions). The condition of $L \gtrsim L_{\min}$ thus would impose a prohibitive requirement for simulating the CRSI, even in 1D.

We resolve this issue by realizing that the effect of the long simulation box can be equivalently achieved by directly randomizing particle phases when a particle crosses the periodic boundary; both approaches allow particles to effectively see different wave packets at all times to conform to the random phase approximation. In our simulations, upon boundary crossing, we fix a particle’s total velocity, as well as its parallel velocity along the background magnetic field, whereas we randomize its gyro-phase. By this simple approach, we completely eliminate the requirement of $L > L_{\min}$. Also, under the δf method, the phase randomization procedure does not introduce more errors to momentum/energy conservation.

4. Simulation Setup and Diagnostics

We set up the simulations in the frame of the “unperturbed” CRs, where their background distribution function is isotropic, $f_0(\mathbf{p}) = f_0(p)$. This is also the f_0 for the δf method, with background CR current $\mathbf{J}_{\text{CR},0} = 0$. We set the background density to $\rho_0 = 1$ and pressure $P_0 = 0.6$, with an ideal gas equation of state with adiabatic index $\Gamma = 5/3$ (so that sound speed $c_s = \sqrt{\Gamma P_0/\rho_0} = 1$). While thermodynamics has no effect on the system, we adopt the adiabatic rather than isothermal equation of state to better characterize energy conservation of the system (see Section 6.1.1). In 1D (along the \hat{x} axis), the background magnetic field is set to $B_0 = 1$ along \hat{x} , so that the Alfvén speed $v_A = B_0/\sqrt{\rho_0} = 1$.

For all simulations, we initialize the background gas to a velocity $\mathbf{v}_0 = -v_D \hat{x}$. Since we are working in the frame where the CRs are initially isotropic, this means that the CR distribution streams to the right (along the direction of the mean magnetic field) with speed v_D relative to the gas. On top of the background state, we initialize the system with a spectrum of Alfvén waves propagating along background \mathbf{B}_0 . For each wavenumber $|k|$, there are four modes corresponding to forward and backward propagation and left and right polarization. The wavenumber coverage ranges from $k = 2\pi/L_x$ to $k = 2\pi/(2\Delta x)$ (except that the longest wave is initiated with zero amplitude), where $L_x = N_x \Delta x$ is the simulation domain size, and N_x and Δx are the number of grid cells and cell size in \hat{x} . Correspondingly, a total of $4(N_x/2 - 1)$ modes are initialized. We choose the wave amplitudes to be such that

$$I(k) = A^2/k, \quad (42)$$

so that wave energy is equally distributed in logarithmic k space, and the total wave energy integrated over 1 dex in k is a fraction $2A^2 \ln 10$ of the background field energy (see Equation (13)); the factor of 2 accounts for velocity perturbations). Each mode is initialized with a random phase.

We choose code units such that individual CR particles have a cyclotron frequency $\Omega_c = 1$ in the background field B_0 . This corresponds to the factor $q/mc = 1$ in Equation (34). Thus, our code unit of length is $d_i = v_A/\Omega_c = 1$. We note that, while irrelevant to MHD, d_i is equivalent to the ion skin depth c/ω_{pi} for the background plasma. Our MHD-PIC framework typically applies to scales greater than d_i .

The CR particles are injected with $f_0(p)$ being a κ distribution. We choose $\kappa = 1.25$ throughout this work, corresponding to $f_0(p) \propto p^{-4.5}$, at large p or $\alpha = 0.5$ in Equation (8). In practice, we divide the momentum space into eight bins ranging from $0.01p_0$ to $100p_0$, with half a dex per bin. We inject an equal number (N_p) of CR particles per cell per

bin so as to guarantee a sufficient number of particles in each bin (otherwise, there would be too few particles toward lower and higher energies). Within each bin in each cell, particles are sampled according to $f_0(p)$ in that momentum range and symmetrically distributed in pitch angle and phase. More specifically, once a set of random momentum p , pitch angle θ , and gyro-phase ϕ is generated, four particles are injected at the cell center whose momenta are $(p \cos \theta, p \sin \theta \cos \phi, p \sin \theta \sin \phi)$, and its permutations with $\theta \rightarrow \pi - \theta$ and $\phi \rightarrow \pi + \phi$.

We note that particles with $p < 0.01p_0$ and $p > 100p_0$ are not included in the simulation, whereas we still use a background distribution function f_0 that covers the entire momentum space. This will introduce some inconsistencies. However, since these momenta are sufficiently far from the peak of f_0 , the waves that they drive are typically beyond what can be accommodated in our simulation box (and, moreover, the simulation time is typically not long enough for these modes to grow).

4.1. Choice of Parameters

The ISM in our Galaxy spans a wide range of conditions in various phases. In an average sense, n_{CR}/n_i is on the order of $\sim 10^{-9}$, $v_A/c \sim 10^{-5}$ – 10^{-4} , and the bulk of the energy is in GeV CR particles that are mildly relativistic. While the parameters that we adopt are by no means fully realistic (this would be unachievable), we aim to achieve as much scale separation as possible to mimic conditions in the large-scale ISM.

We choose the artificial speed of light $\mathbb{C}/v_A = 300$ and a characteristic momentum in the CR distribution of $p_0/v_A = 300 = \mathbb{C}/v_A$. With this choice, Equation (33) yields $\gamma_0 = \sqrt{2}$, so as to mimic the fact that the dominant CR population in the ISM is transrelativistic \sim GeV particles.

The fact that $\mathbb{C} = 300$ is larger than $v_A = 1$ by more than 2 orders of magnitude allows for sufficient separation, making the transformation between different reference frames well within the nonrelativistic limit (e.g., a drift velocity as large as $\sim 10v_A$ only amounts to about 3% of \mathbb{C}). With our choices, the wavelength for the most unstable mode is given by $\lambda_m \approx 2\pi/k_0 = 2\pi p_0/\Omega_c \approx 1885$ in code units.

Based on Equation (7) for the growth rate of CRSI, the two main physical parameters are the CR number density ratio n_{CR}/n_i and the CR drift velocity v_D/v_A . As already seen in Figure 3, given our choices of \mathbb{C} and p_0 , the density ratio n_{CR}/n_i has to be well below 10^{-3} to make the left- and right-handed modes grow at about equal rates (as in standard CRSI). We thus choose $n_{\text{CR}}/n_i = 10^{-4}$ as fiducial and vary it within $[10^{-5}, 10^{-3}]$. Note that further smaller values would make the waves grow too little within a reasonable computational time. We take $v_D = 2v_A$ for our fiducial model and also consider higher drift velocities up to $8v_A$.⁸

⁸ In terms of energy density, we have $E_{\text{CR}}/E_B \sim 45$ under the fiducial parameters, which is much larger than the order-unity value that applies through most of the ISM. We note, however, that while the total energy density of CRs matters for large-scale ISM dynamics, it is not directly relevant for the CRSI because the free energy of the CRSI comes from CR anisotropy (i.e., streaming) rather than total CR kinetic energy. Nonetheless, with $E_{\text{CR}}/E_B \sim (n_{\text{CR}}/n_i)(\mathbb{C}/v_A)^2$, our run with $n_{\text{CR}}/n_i = 10^{-5}$ better approaches the realistic limit. One can also reduce \mathbb{C}/v_A to 100 or smaller to reduce this ratio, although this would limit the range of v_D that can be more reliably covered by simulations.

The remainder are numerical parameters. Fiducially, we choose a simulation box size $L/d_i = 9.6 \times 10^4$, resolved by 9600 cells ($\Delta x/d_i = 10$). The fiducial box is ~ 50 times the most unstable wavelength, which is more than sufficient for particles with $p \sim p_0$, but it better accommodates more energetic particles with $p \gtrsim 10p_0$. Note that without implementing phase randomization, from Equation (41), the box size L_{min} would have to be ~ 300 times longer than our fiducial choice. We also conduct a simulation with a shorter box size for comparison.

The level of noise in the simulation depends on the number of particles. Fiducially, we choose $N_p = 256$ particles per momentum bin per cell (in total, 2048 per cell), but we also consider a case with fewer particles. The initial wave amplitude is fiducially chosen to be $A = 10^{-4}$. Note that A should not be too small, otherwise the initial evolution would be noise-dominated. It should also not be too large, which would trigger artificial QLD of particles from the beginning. A more detailed discussion about noise in our simulations is given in Appendix B.

In Table 1, we list all simulation runs presented in this paper. We will primarily focus on fiducial run Fid and M3 (higher n_{CR}/n_i). For other simulations, we generally vary just one parameter from run Fid at a time and examine the role of both physical and numerical parameters. Note that in the simulations, almost all computational cost is spent on the CR parts, given the large number of particles per cell. For our fiducial run, a typical time step is $6 \times 10^{-2}\Omega_c^{-1}$, and the entire run to $t = 10^6\Omega_c^{-1}$ costs about 4×10^4 CPU hours on Intel Xeon E5-2690 v4 CPUs.

4.2. Diagnostics

We first measure the linear growth rate of the CRSI. Over constant time intervals, we decompose the perpendicular magnetic field and velocity profiles into four Alfvén modes and obtain $I(k, t)$ for each mode (see Appendix A). A linear fit is then performed to $\ln[I(k, t)]$, with the slope being $2\Gamma(k)$. This is done within the early phases of evolution before QLD starts to modify the CR distribution function.

Another diagnostic is to directly measure the CR distribution function. Thanks to the δf method, this again can be done with great precision. In the simulation frame, we simply compute $f(p, \mu) = f_0(p) + \delta f(p, \mu)$, where f_0 is analytic and δf can be obtained in the same way, binning particles onto a $p - \mu$ grid using the weighting scheme in Equation (35).

We can further combine the above information, aiming to test QLD directly. This involves transforming $f(p, \mu)$ into the wave frame using Equation (19) and then solving Equation (14), with $I(k)$ feeding into the calculation. The results can be compared with the measured CR distribution function.

Finally, we have tracked a (small) subset of particles in certain runs over certain time intervals; we combine this information with the field data to investigate how particles overcome the 90° barrier.

5. Simulation Results: Run Fid and M3

In Figure 4, we show the evolutionary history of the total wave energy for the first five runs listed in Table 1. In this section, we first focus on the two main simulation runs, Fid and M3, and will also briefly discuss run M5. With wave amplitude $A = 10^{-4}$, the total wave energy integrated over the wave spectrum in k space and summed over all four modes is about 3.2×10^{-7} (compared to 0.5 for the background magnetic field energy density). The

Table 1
List of Main Simulation Runs

Run	v_D/v_A	n_{CR}/n_i	Domain Size $L_x (d_i)$	Domain Size L_x/λ_m	Resolution (d_i per cell)	N_p per bin (per cell)	Runtime (Ω_c^{-1})
Fid	2.0	1.0×10^{-4}	9.6×10^4	51	10	256	10^6
M3	2.0	1.0×10^{-3}	9.6×10^4	51	10	256	10^6
M5	2.0	1.0×10^{-5}	3.2×10^4	17	10	256	1.5×10^6
vD4	4.0	1.0×10^{-4}	9.6×10^4	51	10	256	10^6
vD8	8.0	1.0×10^{-4}	9.6×10^4	51	10	256	10^6
Fid-Hires	2.0	1.0×10^{-4}	6.4×10^4	34	4	256	10^6
Fid-Short	2.0	1.0×10^{-4}	3.2×10^4	17	10	256	10^6
Fid-Np64	2.0	1.0×10^{-4}	9.6×10^4	51	10	64	10^6
Fid-Np1024	2.0	1.0×10^{-4}	9.6×10^4	51	10	1024	2×10^5

Note. Fixed parameters: $C/v_A = 300$, $p_0/mv_A = 300$, $\kappa = 1.25$, and initial wave amplitude $A = 10^{-4}$. Note that the most unstable wavelength is $\lambda_m = 2\pi p_0/(m\Omega_c) \approx 1885d_i$ for all models.

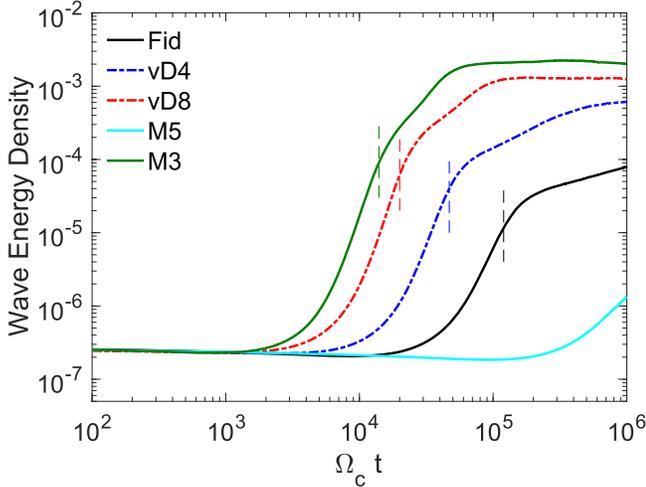


Figure 4. Time evolution of total wave energy in runs Fid (black), M3 (green), and M5 (cyan), as well as runs vD4 (blue dashed–dotted) and vD8 (red dashed–dotted). All are in code units. Vertical dashed lines mark the end of the linear growth phase.

wave energy slowly decreases at first, which is due to numerical damping of high- k modes at grid scale, as well as damping of backward-propagating modes. Shortly afterward, exponential growth of the CRSI becomes clear, leading to a rapid increase of wave energy. This continues to time $t \sim 10^5$ for run Fid and $t \sim 10^4$ for run M3, after which wave growth slows down as QLD starts to substantially modify the particle distribution function. Subsequent growth is slower and characterized by a combination of saturation of the fastest-growing modes excited by low-energy particles and the linear growth of longer-wavelength modes resonant with (less abundant) higher-energy particles. Eventually, for run M3, the wave energy plateaus, marking system saturation. The wave energy density in the saturated state for run M3 is comparable to the estimates in Equation (25). However, for run Fid, in which the saturated energy density should simply be 1 order of magnitude smaller, this is not yet reached.

Clearly, the wave energy evolution is much better discussed on a k -by- k basis and should be accompanied by the evolution of the particle distribution function. These will be addressed in the following subsections for individual evolutionary phases.

5.1. Linear Wave Growth

In Figure 5, we show the Alfvén wave spectrum (in wave intensity $kI(k)$) for the four modes (forward/backward

propagation, left/right polarization) at some early times during the linear phases for the two runs. The initial wave intensity of all modes has $kI(k) = 10^{-8}$ being constant, with $\delta v/v_A \sim \delta B/B_0$ at the level of a few $\times 10^{-4}$. Modes with $k \gtrsim$ a few $\times 10^{-2}$ are numerically damped. With very low numerical noise, linear growth of all other forward-propagating modes is observed from the beginning, whereas backward-propagating modes always damp. The fastest-growing mode, resonant with particles near $p = p_0$, is found to be near $k = k_m \approx 3.3 \times 10^{-3} = \Omega_c/p_0$, as expected.

More quantitatively, we measure the linear growth rate of forward-propagating modes on a k -by- k basis. This is done within $t = 3 \times 10^4 \Omega_c^{-1}$ for run Fid and $t = 10^4 \Omega_c^{-1}$ for run M3. We also consider run M5 and measure its growth rate within $t = 5 \times 10^5 \Omega_c^{-1}$. The results are shown in Figure 6 and compared with analytical growth rates.

Overall, even though the modes are very densely populated, the numerical growth rate matches the analytical results remarkably well. For runs Fid and M5, both the left- and right-handed modes grow at about the same rate, whereas for run M3, the right-handed mode grows slightly faster, all being consistent with theoretical expectations. We stress that phase randomization is essential in achieving these results, as is further discussed in Appendix C.

The growth of the CRSI is better captured toward low k , whereas for high k , instability growth competes with wave damping by numerical dissipation. As a result, the measured growth rates cut off at some $k \sim k_{\text{cut}}$, which decreases from run M3 to run M5 as the instability growth rate decreases by about 2 orders of magnitude. In fact, because of the small CRSI growth rate, k_{cut} corresponds to about 20, 30, and 45 grid cells for runs M3, Fid, and M5. This again indicates that much higher numerical resolution is needed to study the CRSI compared with typical pure MHD instabilities.

Note that for run M3, although the difference in growth rate between the left- and right-handed modes is small, the resulting difference in wave amplitudes can be significant after a few e -folding times. This is already evident in the bottom right panel of Figure 5: within one e -folding time, the intensity of the right-handed mode is already about twice that of the left-handed mode. This will affect subsequent QLD of particles to be discussed later.

As a side note, we have also conducted simulations with drift speed v_D between zero and v_A . We find that all wave modes are damped, and damping is the fastest at $k \sim k_m$. In

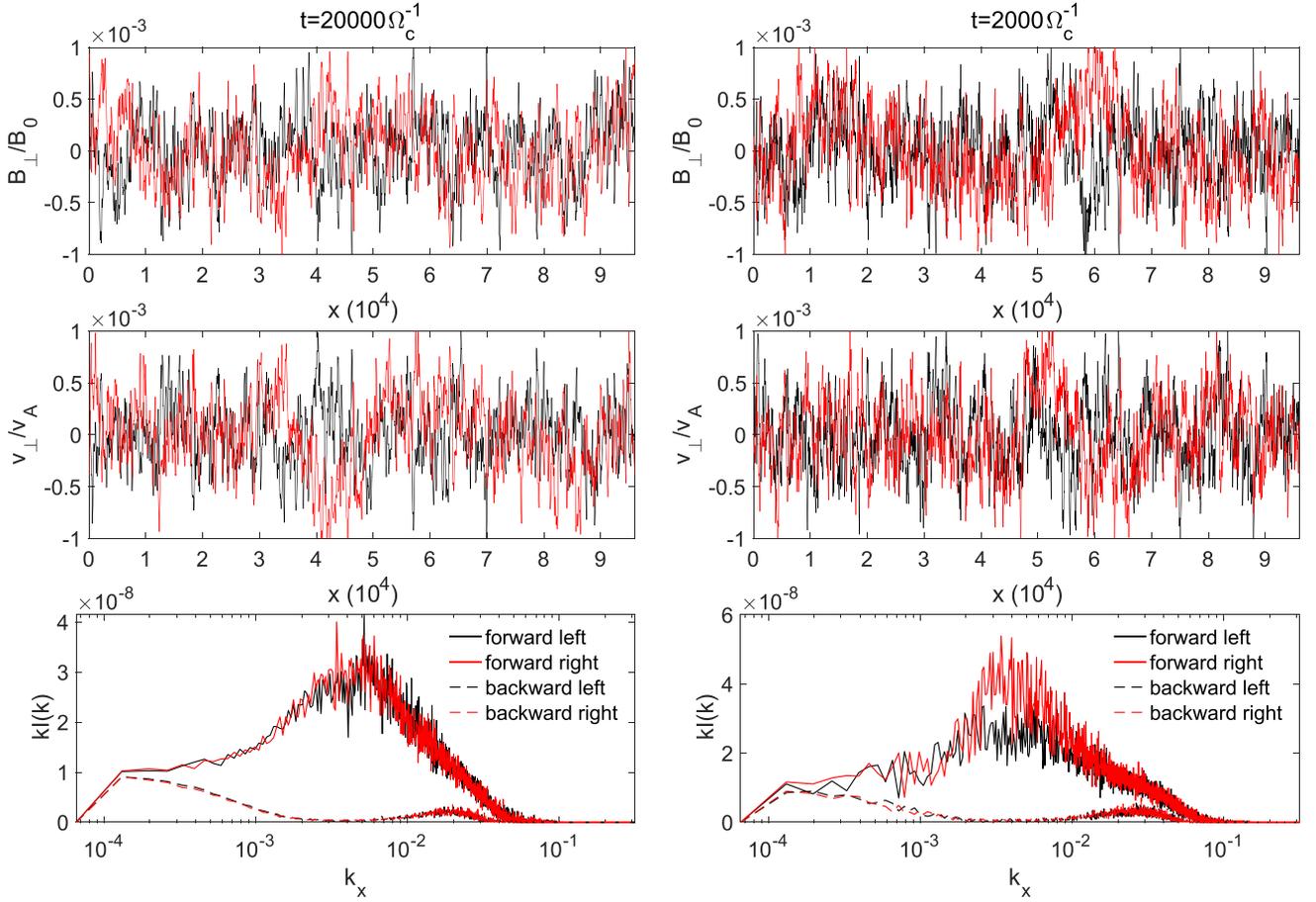


Figure 5. Snapshot of the linear growth phase in runs Fid (left) and M3 (right). The top and middle panels show the profiles of the perpendicular components of the magnetic field and velocity (black and red for the y and z components), normalized to B_0 and v_A . The bottom panel shows the wave spectra (dimensionless, in $kI(k)$) for forward- (solid) and backward- (dashed) propagating modes. Left-/right-handed branches are marked in black/red.

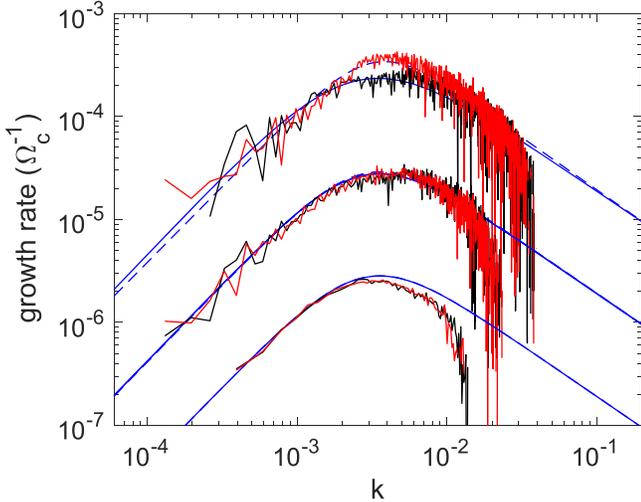


Figure 6. Linear growth rate of the CRSI measured in runs M3 (top), Fid (middle), and M5 (bottom), where black/red curves mark the left-/right-handed polarization. Blue solid/dashed lines mark the analytical growth rate expected from a κ distribution for the left-/right-handed modes.

the case of $v_D = 0$, forward-/backward-propagating modes damp at the same rate, whereas for $0 < v_D < v_A$, damping of forward-propagating modes is slower than that for backward-propagating modes. They all agree well with the dispersion relations.

5.2. Quasi-linear Evolution for Run Fid

Past the linear stage, QLD starts to modify the particle distribution functions and drives the particle distribution toward isotropy in the wave frame. In Figure 7, we show the time evolution of the particle distribution function ($\delta f/f_0$) via four successive snapshots of our fiducial run. For each snapshot, the color scale represents occupation at a given p and μ . We have also marked with dashed lines the loci in phase space resonating with the same wavelength:

$$p\mu = \Omega_c/k = \text{const.} \quad (43)$$

From Figure 7, it is clear that over time, a deficit of particles develops near $\mu \sim 1$, while an excess develops near $\mu \sim -1$. This change in δf reflects the tendency of QLD to scatter CRs toward a new distribution that is isotropic in a frame that is moving to the left with respect to the initial CR distribution. Figure 7 also shows that an excess and deficit of particles develop, respectively, to the right and left of $\mu = 0$; we discuss this further below.

In Figure 8, we show the wave spectrum for the same snapshots as in Figure 7. In these plots, we mark with dotted vertical lines the k values for the corresponding dashed lines in Figure 7. Over time, the spectrum of forward-propagating waves grows at all but the lowest and highest k , with comparable amplitudes for left and right polarizations.

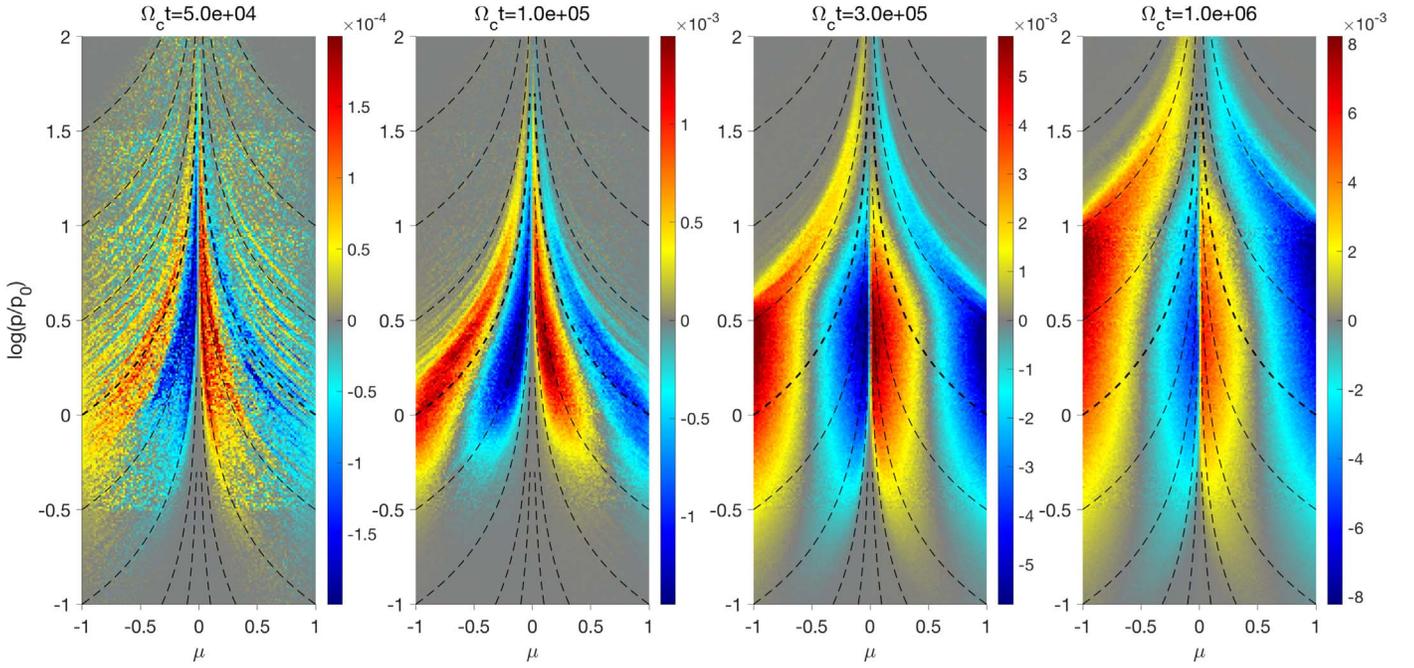


Figure 7. The 2D distribution function $\delta f/f_0$ in the lab (simulation) frame at four snapshots in run Fid. The dashed lines are contours in momentum space that are resonant with the same wave ($p\mu = \text{const}$), with the thick line marking $p\mu = p_0$. Note that the color ranges are different among different panels.

5.2.1. Overall Evolution

At $t \sim 3\text{--}10 \times 10^4 \Omega_c^{-1}$, QLD has started to make appreciable changes to the particle distribution function. The initial QLD is the fastest around particles in resonance with the fastest-growing mode at $k \sim k_m = 3.3 \times 10^{-3}$, as can be clearly identified in the second panel of Figure 7.⁹ As discussed in Section 2.1, the way QLD acts to isotropize the CR distribution is by scattering forward-traveling CRs from higher to lower (positive) parallel velocities and backward-traveling CRs from lower to higher (negative) parallel velocities. This is exactly what we see in Figure 7, with excesses and deficits on the left and right sides of the thick black dashed line.

With some CRs resonant with these waves giving up their free energy through QLD, the growth of the most unstable mode slows down. More slowly growing modes at longer wavelengths have not yet grown to sufficient amplitudes to yield appreciable QLD to particles in resonance with them at this time, and thus still grow at the same rate. As a result, the peak of the wave spectrum shifts toward longer wavelengths, as seen from the time sequence in Figure 8.

Such wave growth then leads to further QLD for more energetic particles, as can be seen in the last two panels of Figure 7. By the end of our simulation at $t = 10^6 \Omega_c^{-1}$, all particles with $p \lesssim 10p_0$ have undergone substantial QLD, whereas for particles with $p > 10p_0$, only the ones with a smaller pitch angle have undergone some QLD. While the trend should continue over time, we are also limited by the simulation box size, which corresponds to 50 times the most unstable wavelength. Modes available for scattering particles with $p > 10p_0$ are fairly limited and discretized, precluding following QLD accurately (the discreteness can already be

seen in the last panel of Figure 7). We thus do not run this simulation for longer.

Across $\mu = 0$, there is a substantial excess (deficit) of particles with $\mu > 0$ ($\mu < 0$) for $p_0 \lesssim p \lesssim 10p_0$, which clearly illustrates the 90° pitch angle problem. Evidently, redistribution of the particle pitch angle occurs only on either side of $\mu = 0$, but particles do not scatter across $\mu = 0$. In part, the failure to scatter across $\mu = 0$ is due to a lack of wave power at large k_x ; particles with small $|\mu|$ would only be able to resonate and scatter off of short-wavelength waves. In our simulation, $k_x > 0.1$ waves have very low amplitude due to damping (see Figure 8; waves at $k_x > \pi N_x/L_x$ are not resolved at all). In addition, for run Fid, the amplitude $\delta B/B_0$ is insufficient to trigger reflection near $m = 0$ for the bulk CR particles. We further discuss the circumstances that lead to “premature” saturation in the next subsection.

Particles with $p \lesssim p_0$ have also undergone QLD, which is caused by shorter-wavelength modes with $k > k_m$. These modes initially grow slower than the fastest-growing mode at $k = k_m$, but they catch up toward later times and eventually saturate at similar amplitudes, as can be seen from Figure 8. Correspondingly, the evolution in the distribution function also propagates toward the lower p side, as seen in Figure 7. However, this is eventually limited by numerical dissipation. Given our resolution, we already see from Figure 6 that modes with $k \gtrsim 6k_m$ grow much slower than theoretical expectations due to numerical dissipation. Considering the scattering across the entire pitch angle range, and given that even particles with $p \sim p_0$ are subject to premature saturation, we do not discuss particles with $p < p_0$ in general.

We have also examined run M5. The overall evolution is largely similar to run Fid: the evolutionary stage at time t in run M5 is similar to that at time $t/10$ in run Fid, which is also visible from Figure 4. This is reasonable, given the exact factor of 10 difference in wave growth rates. The QLD starts to modify the CR distribution function after $t \sim 3 \times 10^5 \Omega_c^{-1}$,

⁹ The initial wave spectrum also leads to some QLD that slightly affects the appearance of the distribution function (owing to contributions from high- k modes) at early times, $t \lesssim 5 \times 10^4 \Omega_c^{-1}$.

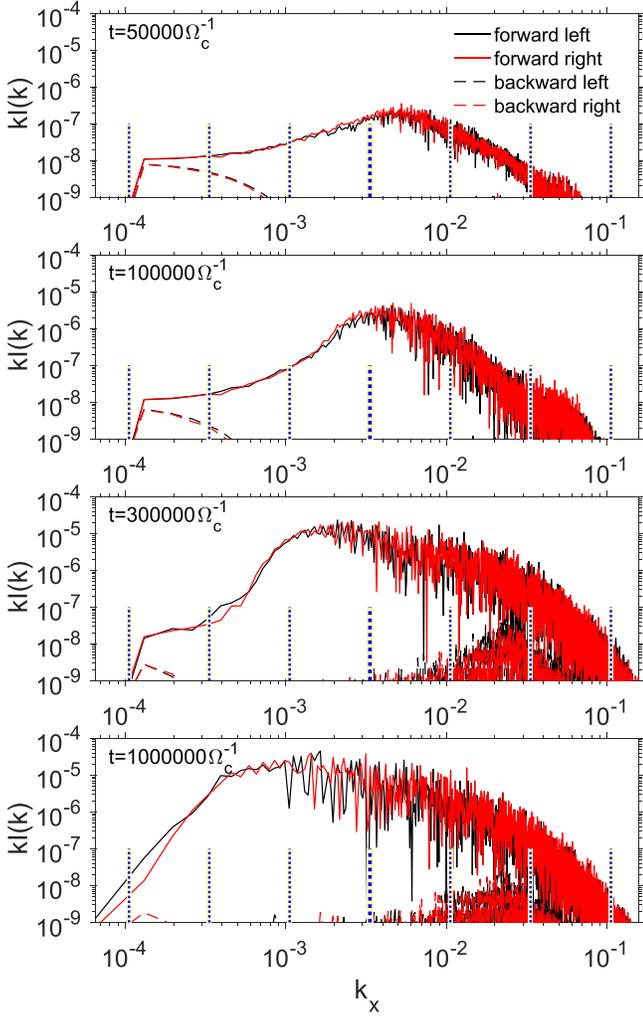


Figure 8. Wave spectrum at different times during QLD in run Fid. The vertical dotted lines mark the wavenumbers that particles located at the black dashed lines in Figure 7 can resonate with.

and by the end of the simulation at $t = 1.5 \times 10^6 \Omega_c^{-1}$, QLD is already close to being completed for particles with $p \sim 1 - 3p_0$, except that they are stuck at $\mu = 0$. Given even smaller wave amplitudes in run M5 than in run Fid, we do not expect that further evolution would help with this 90° problem.

5.2.2. Quantitative Analysis

To see how well our simulation results compare with the quasi-linear theory, we transform the distribution function to the wave frame using Equation (19). For each particle bin (p_w, μ_w) at a given snapshot, we can find its resonant wavenumber k_{res} and corresponding $I(k_{\text{res}})$ and compute the scattering frequency ν from Equation (15). Then, we solve the QLD Equation (14) over time numerically, with $I(k)$ being updated over a time interval of $\Delta t = 500 \Omega_c^{-1}$ (which is the frequency of simulation data output).

In Figure 9, we show four snapshots in this process. Note that for better comparison, we have chosen snapshots 1, 3, and 4 to correspond to snapshots 2, 3, and 4 in Figures 7 and 8. The scattering frequency gradually increases to the level of $\nu_{\text{max}} \sim 10^{-4} \Omega_c$ as the waves grow, and its appearance in $p - \mu$ space again shows the characteristic features corresponding to $p\mu = \text{const}$ from the resonance condition (modulo

that ν decreases toward larger p as particles become relativistic). With short-wavelength modes being damped by numerical dissipation, the scattering frequency for small values of $|p\mu|$ has a cutoff, leaving a gap around $\mu = 0$ that is broader toward smaller p . This causes the 90° barrier for QLD, where no particles can cross $\mu = 0$, as is the case when solving Equation (14).

In the bottom panels, we show the particle distribution function in the wave frame. Initially, $f_w - f_0$ is simply given by Equation (19) taking $\delta f = 0$, which is linear in μ_w . Note that for particles with different momenta, both $\partial \ln f_0 / \partial \ln p$ and v_w can be different, yielding different slopes. With the κ distribution, the slope is smaller for smaller p and approaches a constant $2(\kappa + 1)\Delta v/c$ for relativistic particles (thus, the red and green dotted lines overlap). Subsequent evolution has already been discussed in Section 5.2.1 in the simulation frame, which is straightforward to translate to the wave frame. Several important results and features are worth noticing.

First, at early time ($t \lesssim 2 \times 10^5 \Omega_c^{-1}$), evolution of the numerical distributions follows the QLD prediction remarkably well. This gives us confidence that our simulations properly capture QLD, thanks to the phase randomization technique that we employed.

Second, QLD drives the local slope of $f_w - f_0$ toward zero on either side of $\mu = 0$ (see Equations (20) and (21)); piecewise flattening is achieved at later stages of evolution, reached by lower-energy particles first. This marks the saturation of the CRSI in the local momentum space.

Third, particles undergoing QLD tend to get stuck at $\mu = 0$. Eventually, the CRSI saturates because $f_w = \text{const}$ on both sides of $\mu = 0$ for each p , although there is a discontinuous jump at $\mu = 0$. This is the hallmark of premature saturation of QLD.

Fourth, solving the QLD equation allows us to well reproduce the saturated state for lower-energy particles ($p \lesssim 3p_0$), but this is not true for higher-energy particles ($p \gtrsim 10p_0$). Given that the only missing ingredient in the QLD Equation (14) is the process that governs the crossing of the 90° pitch angle, this indicates that the $\mu = 0$ crossing does happen (slowly) for higher-energy particles in our simulations.

Given the above results, we now check the requirements for QLD to scatter particles to the critical pitch angle, such that the condition for mirror reflection (Equation (27)) can be satisfied. Seen from Figure 4, we have the rms $\delta B/B_0 \sim 9 \times 10^{-3}$, giving $\mu_{\text{mir}} \sim 6 \times 10^{-3}$. Within our simulation time of $10^6 \Omega_c^{-1}$, from Equation (14), we need $kI(k) \gtrsim 10^{-6}$ for effective QLD. Reading from Figure 8, this gives a maximum wavenumber $k_{\text{cut}} \sim 8k_m$. Based on resonance conditions, the minimum momentum for which QLD can scatter the particle to reach $\mu = \mu_{\text{mir}}$ is then

$$p_{\text{min}} \sim \frac{k_m}{\mu_{\text{mir}} k_{\text{cut}}} p_0 \sim 20p_0. \quad (44)$$

In other words, the waves in our simulation only have sufficient power to scatter particles with $p \gtrsim p_{\text{min}}$ to reach pitch angle μ_{mir} , potentially allowing the mirror effect to yield reflection (and deviate from the pure QLD result). However, from the bottom right panel of Figure 9, representing the last snapshot of the simulation, we see that even for p as small as $\sim 3p_0$, deviation from QLD can be identified, though just by a very small amount. Reading off the figure, deviation from QLD

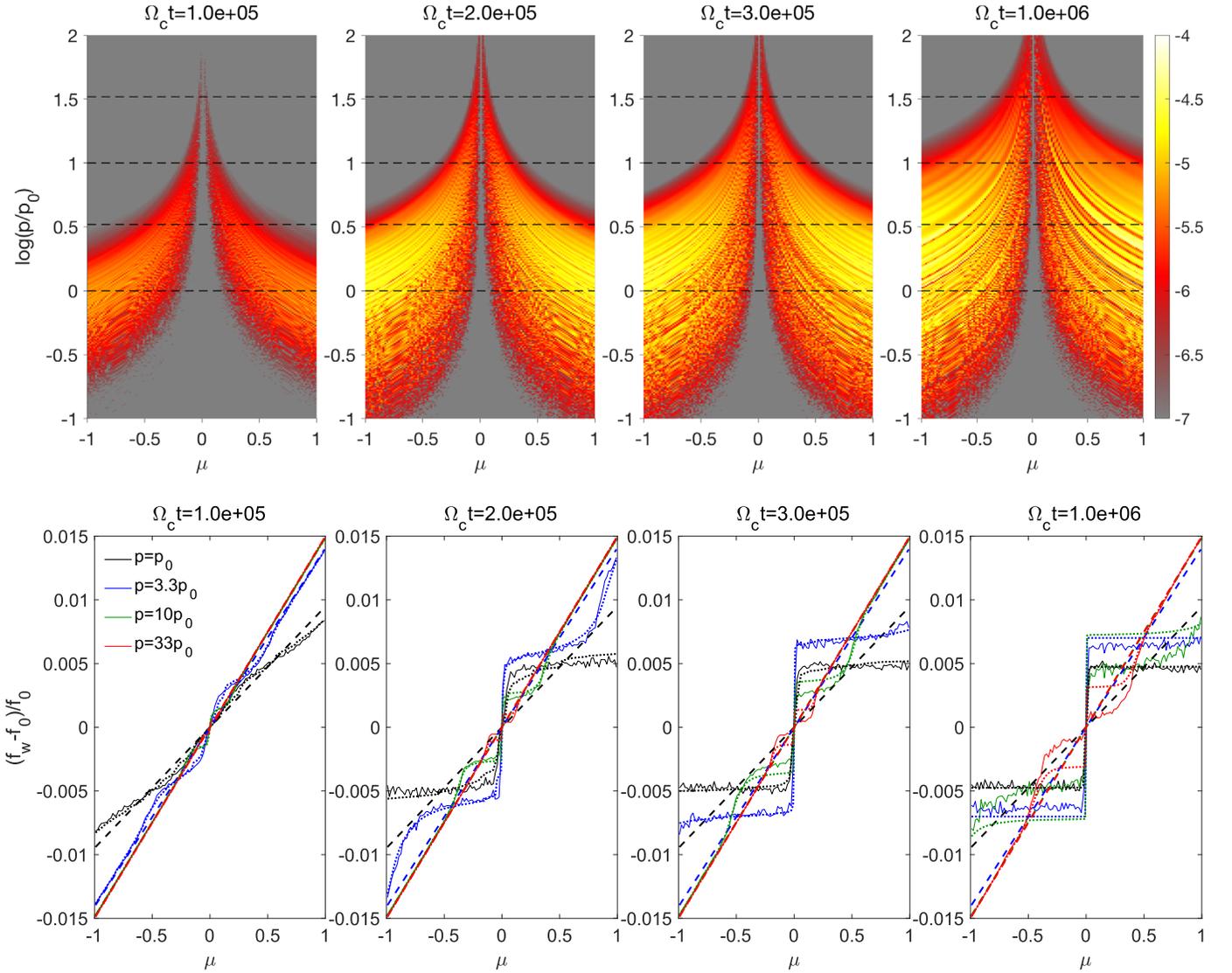


Figure 9. Top: four sequential snapshots of the scattering frequency ν of QLD (see Equation (15)) in momentum space in run Fid. Shown in color is $\log_{10}(\nu/\Omega_c)$. Bottom: particle distribution function in the wave frame f_w at the same four snapshots as in the top panel. Shown are $(f_w - f_0)/f_0$ for four representative momenta as a function of pitch angle cosine μ , indicated by different colors (see legend). These momenta are also marked as black dashed lines in the top panels. Solid lines are measured from the simulations, thick dashed lines mark the initial distribution, and dotted lines are reconstructed distribution functions by solving the QLD Equation (14) in time.

appears to begin for p_{\min} between $3p_0$ and $10p_0$, which is in tension with the expectation from mirror reflection (Equation (44)). Further discussion of this issue is deferred to Section 5.4, where a more detailed study based on run M3 is presented.

5.2.3. Evolution of CR Drift Velocity

While it is generally convenient to use the reduction of CR bulk drift velocity to describe how much CR particles have undergone scattering through the CRSI, this drift velocity is clearly energy-dependent. We can evaluate the energy-dependent drift velocity $v_d(p)$ (to distinguish from the initial CR bulk drift velocity v_D) in the simulation frame as

$$v_d(p) = \frac{\int f(p, \mu) v(p) \mu d\mu}{\int f(p, \mu) d\mu}, \quad (45)$$

where $v(p)$ is the velocity that corresponds to momentum p , and $f = f_0 + \delta f$. Values of $v_d(p)$ can be simply obtained by integrating over a horizontal line in Figure 7. We can further integrate the distribution function over p to obtain the full drift velocity $v_{d,\text{full}}$.

In the left panel of Figure 10, we show the results for four different representative particle momenta from $p = p_0$ to $33p_0$. Based on the previous discussion, the results are straightforward to interpret. The initial drift velocity in the simulation frame is, by definition, zero. Low-energy CRs, which are responsible for driving faster wave growth, undergo more rapid initial QLD and a corresponding reduction their drift speed; higher-energy CRs take much longer to catch up. However, the CRSI and drift reduction of low-energy CRs saturates prematurely, leading to only a modest reduction of v_d , whereas higher-energy CRs ($p \gtrsim 10p_0$) can overcome the $\mu = 0$ barrier and are on the way to achieving complete isotropy. The overall average drift $v_{d,\text{full}}$ decreases over time. Contribution to the

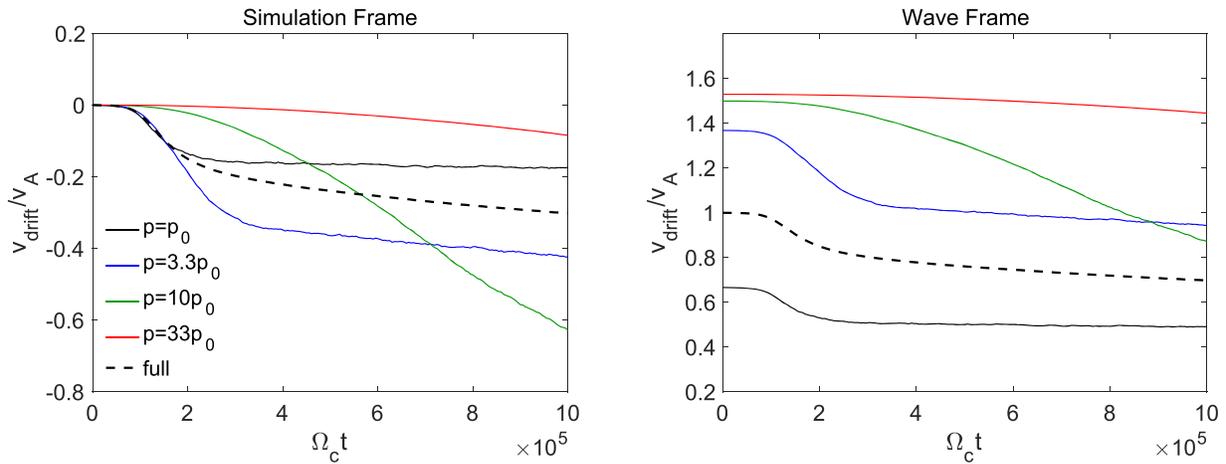


Figure 10. Time evolution of the mean particle drift velocity in different momentum bins, together with the full drift velocity (see legend), for run Fid. We show velocities measured in the simulation frame (left) and wave frame (right). The dashed line represents the mean drift velocity averaged over the entire particle population.

overall reduction is mostly from lower-energy particles with momenta between $\sim p_0$ and $\sim 3p_0$ at early times and higher-energy particles with $p \sim 10p_0$ at late times.

The drift velocity is frame-dependent. Similar to the above procedure, we can measure the drift velocity in the wave frame $v_{d,w}(p)$, and the results are shown in the right panel of Figure 10. While the velocity difference between the two frames is just v_A , initial particle drift speeds within individual momentum bins do not equal v_A . This is because after a frame transformation, particles with a given momentum p come from a range of momenta near p from the original frame, making the results depend on $d \ln f_0 / d \ln p$ (see Equation (19)). Only the full drift speed $v_{d,w,\text{full}}$ precisely equals v_A , as expected. On the other hand, in a fully saturated state, we would expect $v_{d,w}(p) = 0$ for all p . However, we clearly see from the figure that low-energy particles are stuck at nonzero $v_{d,w}(p)$, whereas higher-energy particles require a longer time to achieve saturation.

5.3. Quasi-linear Evolution for Run M3

With n_{CR}/n_i a factor of 10 larger, run M3 evolves much more rapidly than run Fid. We choose four representative snapshots at $\Omega_c t = 10^4, 3 \times 10^4, 10^5$, and 9×10^5 and analyze the results.

5.3.1. Overall Evolution

In Figure 11, we show the particle distribution function at these snapshots, while Figure 12 shows the corresponding wave intensity spectra. As discussed earlier, right-handed modes grow faster near $k \sim k_m$. This continues to $t \gtrsim 10^4 \Omega_c^{-1}$. Since these modes are resonant with backward-traveling CRs, we see that at early times, QLD proceeds much more rapidly for $\mu < 0$ particles resonant with these waves, leading to an initial asymmetry in the particle distribution function.

With asymmetric wave growth and QLD, backward-traveling CRs have used up more free energy, thus slowing down the growth of right-handed waves. The left-handed modes, on the other hand, continue to grow normally. This allows the left-handed modes to eventually catch up with the right-handed modes, and we see that by the time $t = 3 \times 10^4 \Omega_c^{-1}$, the left- and right-handed modes at all wavelengths have similar amplitudes. At the same time, the distribution functions for

forward- and backward-traveling CRs also tend to become more symmetric about $\mu = 0$. However, the $\mu = 0$ barrier remains at this stage for low-energy particles.

Between $\Omega_c t = 3 \times 10^4$ and 10^5 , we see that crossing of the $\mu = 0$ barrier occurs, paving the way for full saturation. The release of more free energy has led to some further wave growth. This process continues until the end of our simulation, with QLD continuing to extend to higher and higher energies.

5.3.2. More Detailed Analysis toward Saturation

As before, for run M3, we compute the CR drift velocity in four different momentum bins for both the simulation frame and the wave frame, and we show the results in Figure 13. We see that in the simulation frame, the particles of all four momentum bins eventually flatten at different levels of $v_d(p)$. In the wave frame, they all reach $v_d(p)$ equal to zero,¹⁰ indicating that they are fully isotropized and have achieved complete saturation.

This saturation is more evident in Figure 14, showing a time sequence of the full distribution function in the wave frame, with the same snapshots as in Figures 11 and 12. We can clearly identify the isotropization of particles in the later two snapshots, extending from the range of $p_0 \lesssim p \lesssim 10p_0$ at $t = 10^5 \Omega_c^{-1}$ to covering most of the particle population at the end of the simulation. Because of this, and with most of the particle population having a momentum close to p_0 , we see that the full drift speed $v_{d,\text{full}}$ averaged over all particle populations shown is reduced almost exactly by v_A (Figure 13).

For run M3, we have also computed the scattering frequency according to Equation (15) and solved the QLD Equation (14) based on output data, and we show the results at four different snapshots in Figure 15. The asymmetry in the scattering of forward-/backward-traveling CRs at early times ($t \sim 10^4 \Omega_c^{-1}$) can be clearly identified in the top left panel. At later times, we can see by comparing with Figure 11 that the extent to which QLD has proceeded in momentum space is strongly related to the region where scattering frequency is high.

¹⁰ The reason that the saturated drift velocity in the wave frame is not exactly zero in Figure 13 is purely numerical, depending on the resolution employed in binning particle momentum during simulation output (due to the sensitive dependence of f_0 on p). Finer binning makes the results closer to zero but with more noise, and vice versa.

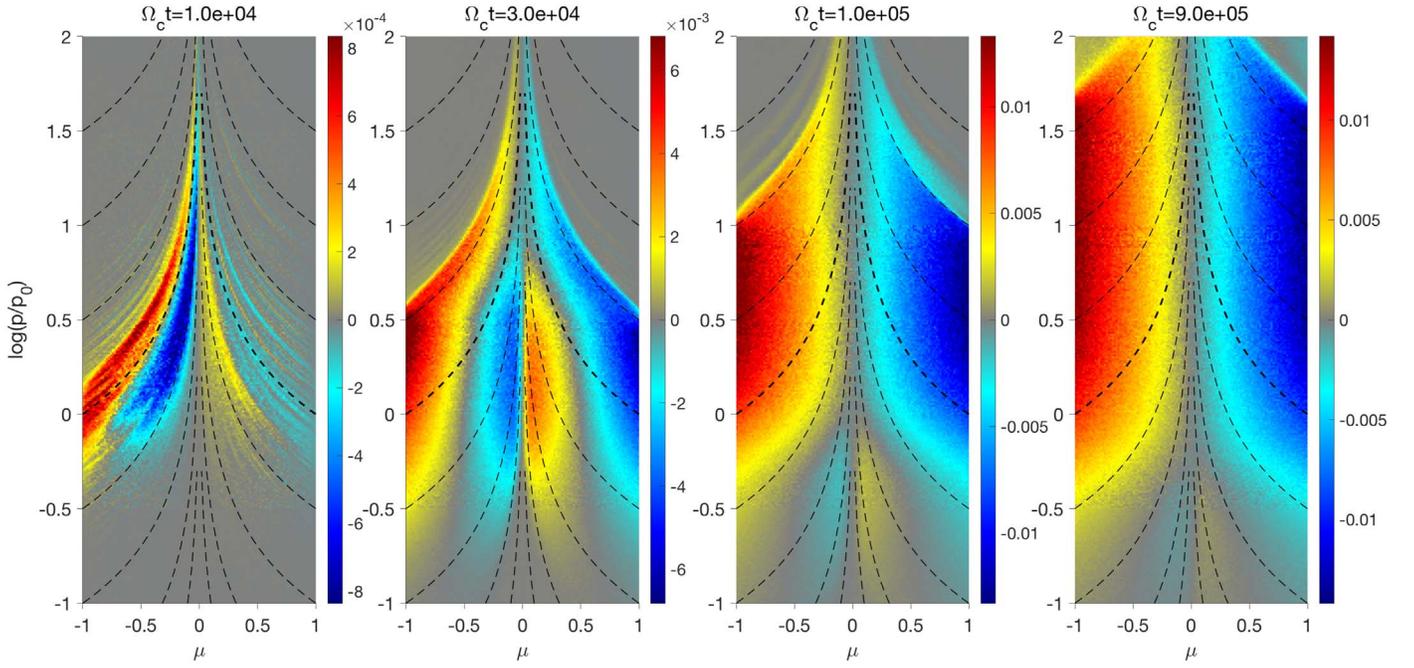


Figure 11. Same as Figure 7 but for run M3, showing the 2D distribution function of $\delta f/f_0$ in the simulation frame at four snapshots.

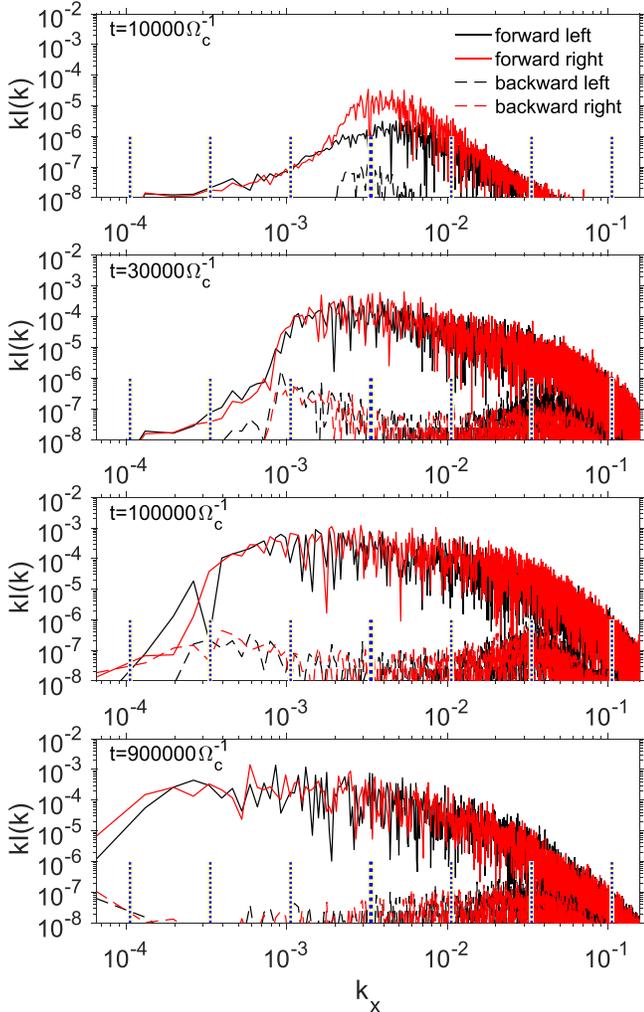


Figure 12. Same as Figure 8 but for run M3, showing the wave spectrum at the same four snapshots as in Figure 11.

The bottom panels of Figure 15 give more quantitative information about how quasi-linear evolution has progressed over time. For lower-energy particles (with $p \sim p_0 - 3p_0$), evolution proceeds rapidly and can initially be well reproduced by solving the QLD equation. During subsequent evolution, particle distribution functions in fixed momentum bins are largely flat (due to rapid QLD) on both sides, with a jump at $\mu = 0$. The jump gradually decreases, as particles get scattered across $\mu = 0$, which is the rate-limiting process for particle isotropization. For higher-energy particles ($p \gtrsim 10p_0$), on the other hand, the crossing of $\mu = 0$ is more rapid (leading to substantial deviation from directly solving the QLD equation). The rate-limiting process for isotropization then becomes QLD itself, especially at large μ .

A key result from our run M3 is efficient particle crossing of the $\mu = 0$ barrier. We again first examine whether this is consistent with mirror reflection. From Figure 4, we see that by the time of $t = 6 \times 10^4 \Omega_c^{-1}$, the magnetic energy density approaches a saturated value ($\delta B^2/B_0^2 \sim 2 \times 10^{-3}$), corresponding to $\delta B/B_0 \sim 0.045$. This gives the mirror reflection threshold of $\mu_{\text{mir}} \sim 0.032$. Within a simulation time of $\sim 10^5 \Omega_c^{-1}$, we need $kI(k) \sim 10^{-5}$ for effective QLD. This is achieved for k up to $k_{\text{cut}} \sim 10k_m$ by looking at Figure 12 at $t = 10^5 \Omega_c^{-1}$. According to the same calculation as in Equation (44), we find that mirror reflection can be achieved for particles with $p \gtrsim p_{\text{min}} = (k_m/k_{\text{cut}}) \mu_{\text{mir}}^{-1} \sim 3p_0$. However, we see that by this time, particles with $p \gtrsim p_0$ are already fully isotropized, and lower-energy particles have also undergone partial isotropization. Again, we conclude that mirror reflection appears insufficient to explain the detailed evolution of the momentum distributions.

5.4. Overcoming the 90° Barrier

To reveal the mechanism behind particle crossing of the 90° pitch angle in run M3, we have randomly selected a subsample of particles (1280 per energy bin) and closely followed their trajectories for a time interval of $200 \Omega_c^{-1}$ starting from $t_0 = 6 \times 10^4 \Omega_c^{-1}$. We count the number of particles that have

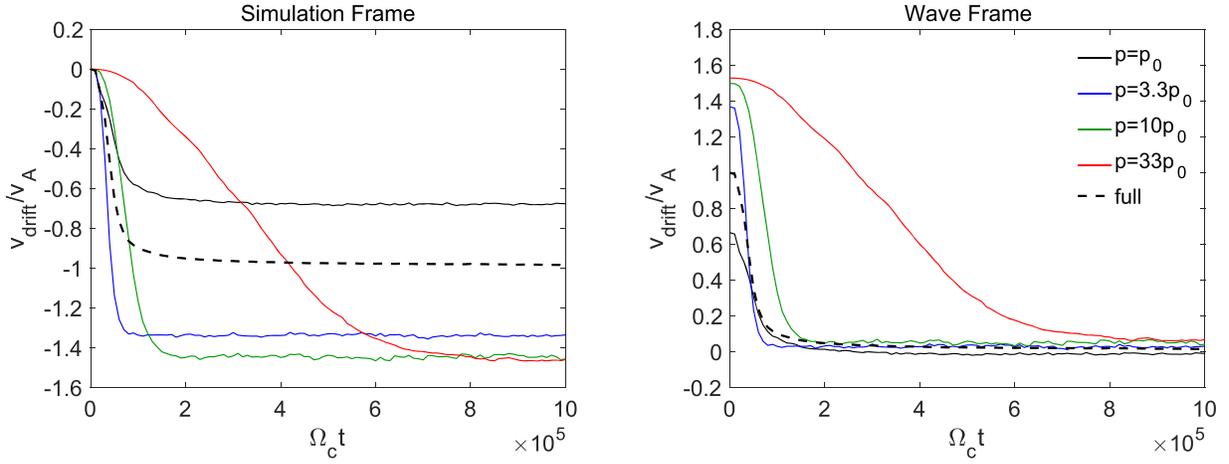


Figure 13. Same as Figure 10 but for run M3, showing the evolution of the mean particle drift velocity in different momentum bins, together with the full drift velocity (see legend).

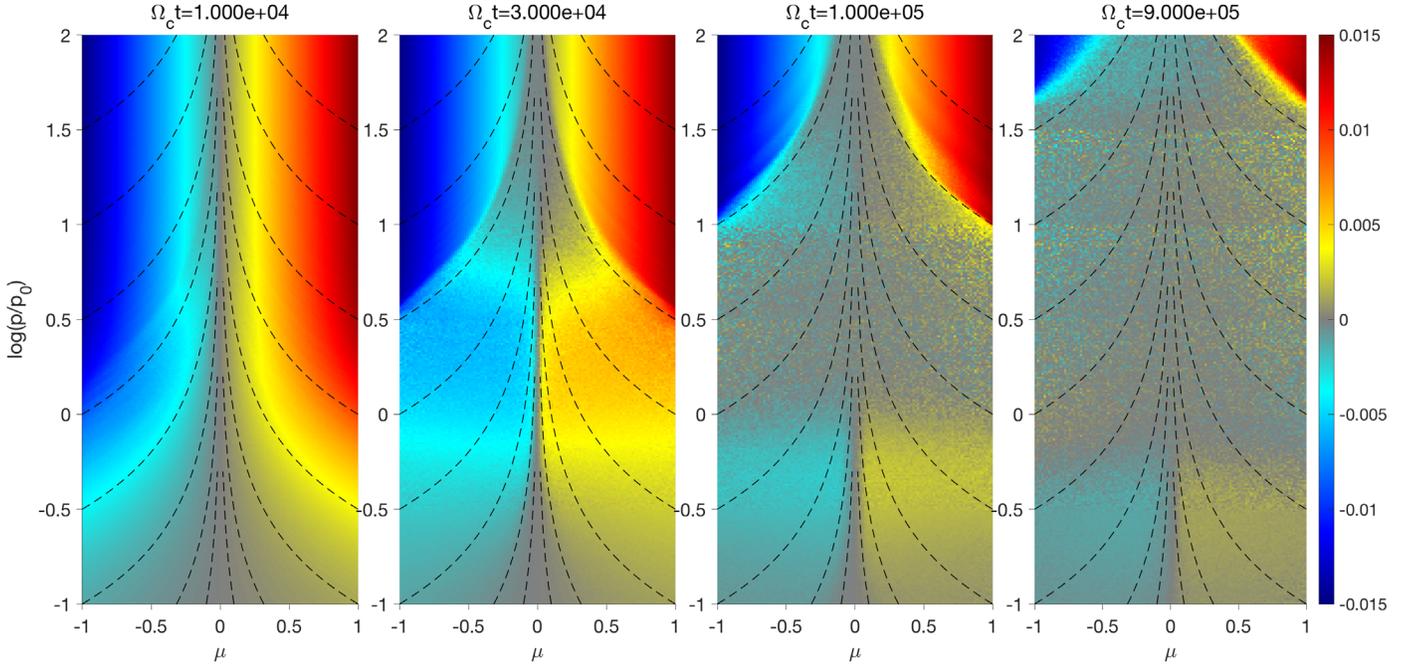


Figure 14. The 2D distribution function $\delta f_w/f_0$ in the wave frame for four snapshots from run M3. The dashed lines are contours in momentum space with the same resonance condition.

undergone reflection across 90° over this interval. Here we consider the particle pitch angle in the wave frame that is relative to the full magnetic field vector (rather than background field), which is marked by μ'_w . To be counted, we require μ'_w before and after this time interval to be beyond $\pm v_A/C \sim 0.0033$ (chosen somewhat arbitrarily, but this does not qualitatively affect the counting result). Among the four central momentum bins from $p = 0.1p_0$ to $10p_0$, we find that 9, 13, 20, and 23 particles are reflected. We looked at the trajectories of all of these particles and choose to show four particles with characteristic behaviors from each momentum bin in Figure 16.

5.4.1. Representative Cases

For the first particle with $p \approx 0.2p_0$ that undergoes a reflection, we see that, surprisingly, the field strength it

experiences is almost constant. Quantitatively, the perpendicular field strength over this time interval is $\delta B \in (0.0398, 0.0417)$. The range it spans translates to $\mu_{\text{mir}} = 1.3 \times 10^{-3}$. While the particle does evolve quasi-adiabatically with many gyrations (which is typical for lower-energy particles given their small gyro-radii and lack of power at small scales), its μ'_w changes by a factor of several larger than the expected μ_{mir} . Checking the magnetic field profile in the bottom panel, we see that the particle is traveling against a shallow gradient in B_z . The range of B_z experienced by the particle $\in (-0.081, -0.041)$, which changes the magnetic field orientation by ~ 0.04 . While this is not proof that the particle must be reflected, it can well accommodate the change in particle μ'_w , which is less than 5×10^{-3} .

For the second particle with $p \approx 0.9p_0$, its higher velocity allows it to traverse larger distances over the course of reflection,

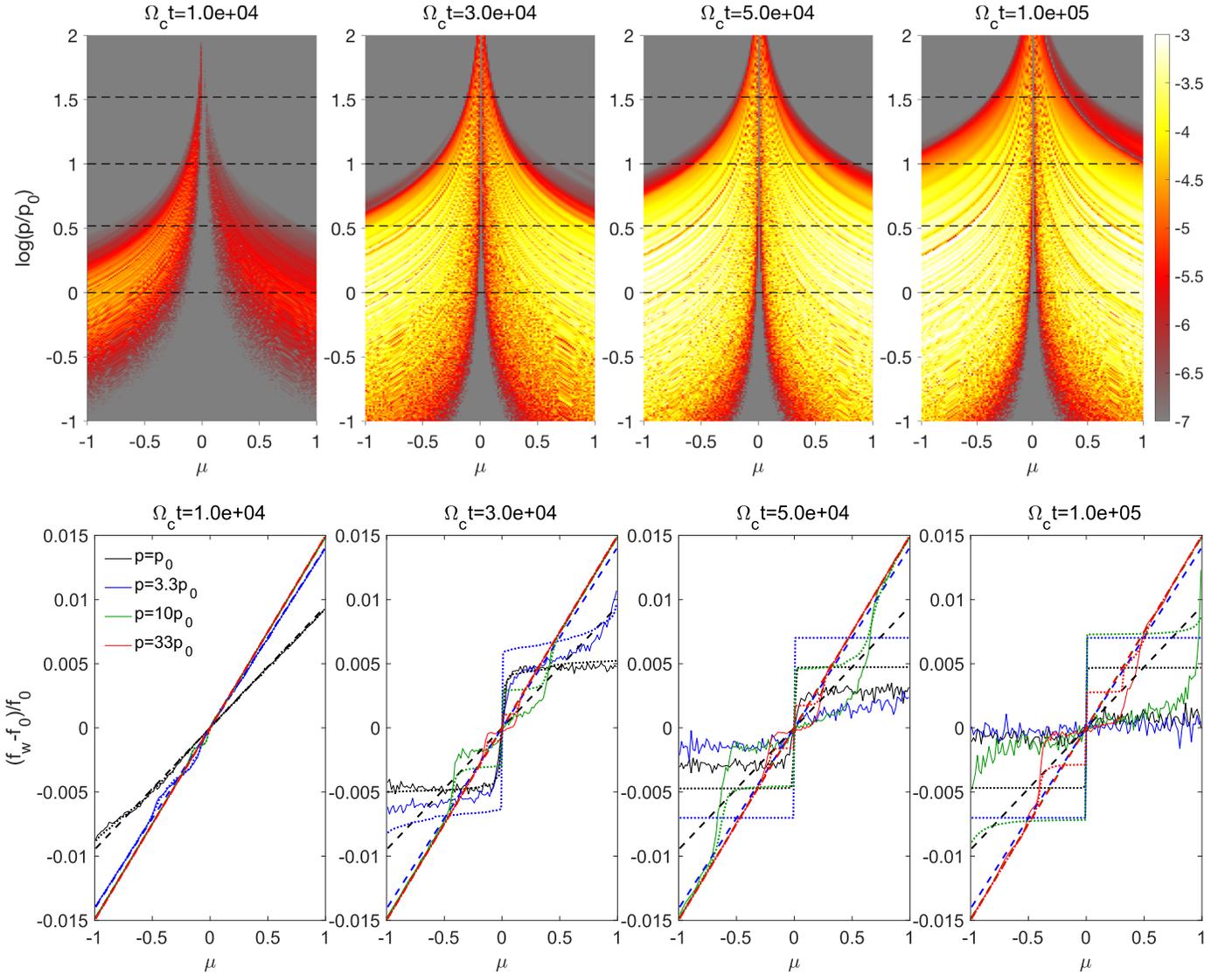


Figure 15. Same as Figure 9 but for run M3 at four different snapshots.

and it experiences more rapid magnetic fluctuations. We have again checked that variations in total field strength are insufficient for mirror reflection. Reflection occurs again when the particle travels against the gradient of individual field components, and in this time, we have B_y increasing rapidly and B_z decreasing more gradually. These variations change the magnetic field orientation by ~ 0.05 within just a few particle gyrations, which is again more than sufficient to account for the change in particle pitch angle by ~ 0.02 .

In the remaining two examples, reflection occurs within two gyro-orbits. These cases are more common among relatively energetic particles: because of their larger gyro-radii, they more often experience abrupt changes in magnetic field. These changes average out to the leading order in quasi-linear theory, except when they lead to reflection. We see that the threshold pitch angle for reflection can be higher for these particles and reach as large as $\mu \sim 0.1$. This is accompanied by abrupt changes in B_y and B_z at comparable levels within the distance that the particle travels in one gyration.

5.4.2. Nonlinear Wave–Particle Interaction

Overall, the four examples roughly cover the range of behaviors of particles during reflection. They can be interpreted under the framework of nonlinear wave–particle interaction (Völk 1973).

Reflection is most effective (almost instant) for the last two cases in Figure 16, for which abrupt changes in fields occur on a scale shorter than the distance that particle travels in one gyro-time. In this regard, we can discuss the critical pitch angle below which nonlinear wave–particle interaction takes over to reflect particles as follows. For a particle with momentum p (and velocity v), let μ_{crit} be this critical pitch angle. It travels by $L_{\text{crit}} \sim 2\pi v \mu_{\text{crit}} / \Omega$ in one gyro-time, corresponding to a wavenumber $k_{\text{crit}} = \Omega / (v \mu_{\text{crit}}) = \Omega_c / (p \mu_{\text{crit}})$. Reflection occurs when an abrupt change in magnetic field orientation is comparable to μ_{crit} . Noting that this depends on the spectrum of waves at scales below some $\zeta(k_{\text{crit}})^{-1}$, we arrive at a relation

$$\mu_{\text{crit}}^2 \sim C \int_{k_{\text{crit}}/\zeta}^{\infty} I(k) dk, \quad (46)$$

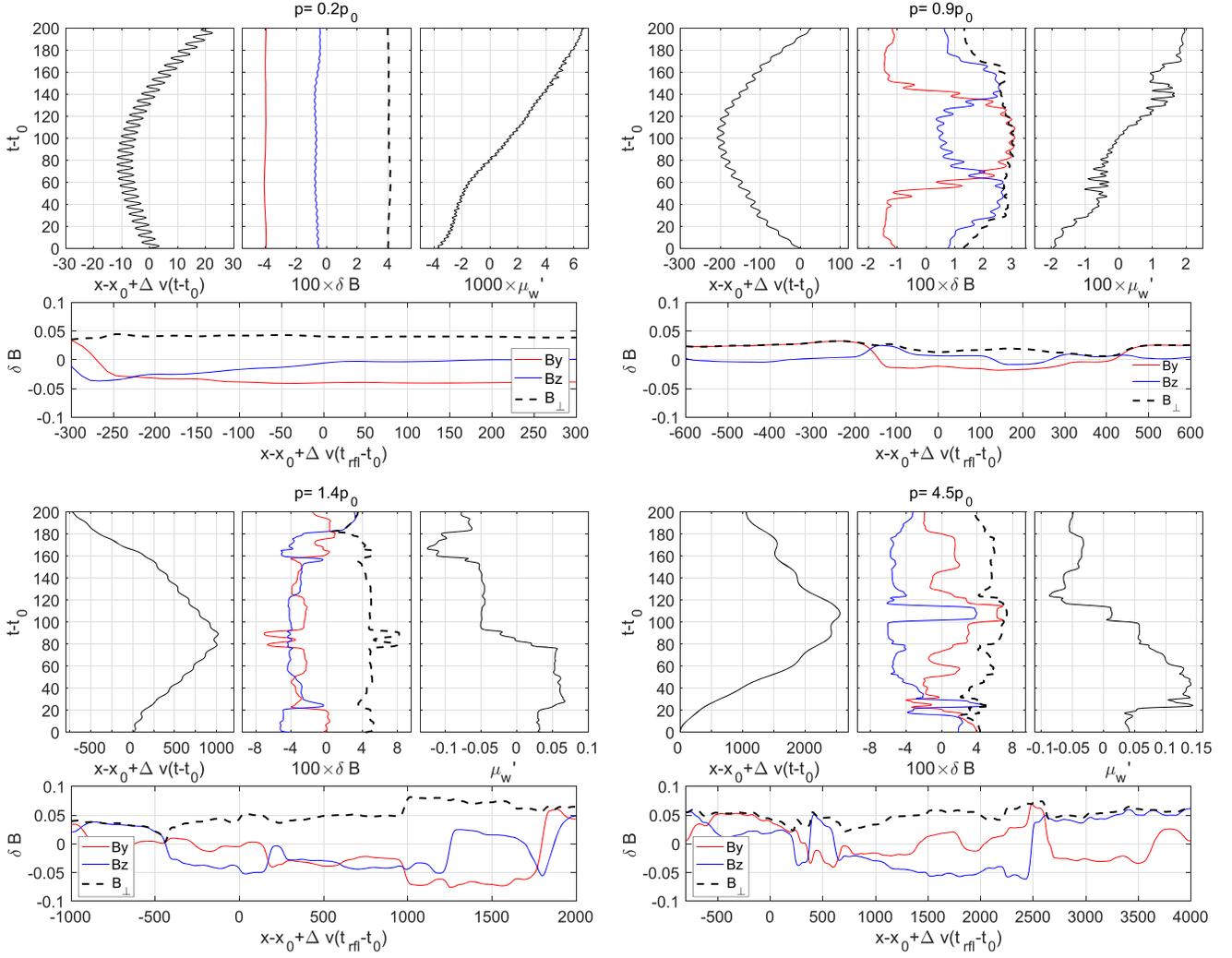


Figure 16. Four typical reflection events for a particle with $p = 0.23p_0$ (top left), $0.92p_0$ (top right), $1.39p_0$ (bottom left), and $4.52p_0$ (bottom right). They are selected right after time $t_0 = 6 \times 10^4 \Omega_c^{-1}$. Among each group of plots, the top panels show the time evolution of the particle position in the wave frame, the perpendicular field strength (individual components in colored lines and total in black dashed lines) experienced by the particle, and its pitch angle cosine in the wave frame relative to the total field. The bottom panel shows the perturbed magnetic field profiles at the time of reflection.

where there is a square on the left-hand side because $\mu_{\text{crit}} \sim \delta B/B_0$, while the right-hand side corresponds to its square. We have also include dimensionless factors ζ and C that encapsulate the roughness of the model.

The solution to this equation gives μ_{crit} , which exhibits the following features.

1. Higher wave intensity leads to smaller μ_{crit} , because more wave power at smaller scales makes reflection easier.
2. Higher particle momentum leads to larger μ_{crit} , because a larger gyro-scale accommodates more wave power to build up abrupt field profiles.
3. If there is a lack of power at small scales, there can be no solution, meaning that direct reflection in a couple gyro-orbits is unlikely.

Clearly, all of these features qualitatively agree with our findings. They are also justified through our discussions in Section 5.4.1.

To be more quantitative, we note that while the integral of $I(k)$ over k gives the rms fluctuations, actual fluctuations can easily

exceed the rms value, and deviations by up to 3σ are not rare. We thus take $C \sim 3^2$ to account for the upper end of the fluctuations. We may further take $\zeta \sim 4$, considering that waves with wavelengths up to $4L_{\text{crit}}$ can contribute significantly to the local slope within this length scale of L_{crit} (i.e., sinusoidal variations have a roughly constant maximum slope in just one-fourth of the periodic domain consecutively). With these numbers, and taking $p = 4.5p_0$ for run M3, we can derive $\mu_{\text{crit}} \sim 0.1$ (which gives $k_{\text{crit}}/\zeta \sim 1.8 \times 10^{-3}$, and the integral gives 0.0011).

Equation (46) closely resembles Equation (54) of Völk (1973), with their $v_{\text{osc}} \sim 0.25 - 0.3(v\mu_{\text{crit}})$, interpreted as the rms velocity oscillation amplitude of a particle in “the fluctuation field that contributes to its diffusion.” The effective diffusion coefficient at $\mu = 0$ is then given by $\sim 3\Omega v_{\text{osc}}^2$.

The discussions above apply to instant reflection upon encountering abrupt field gradients. This corresponds to the $n = 1$ term in Equation (48) of Völk (1973). As one lowers particle momentum, μ_{crit} decreases, and eventually no solution can be found. This situation applies to lower-energy particles in Figure 16. More generally, higher-order effects

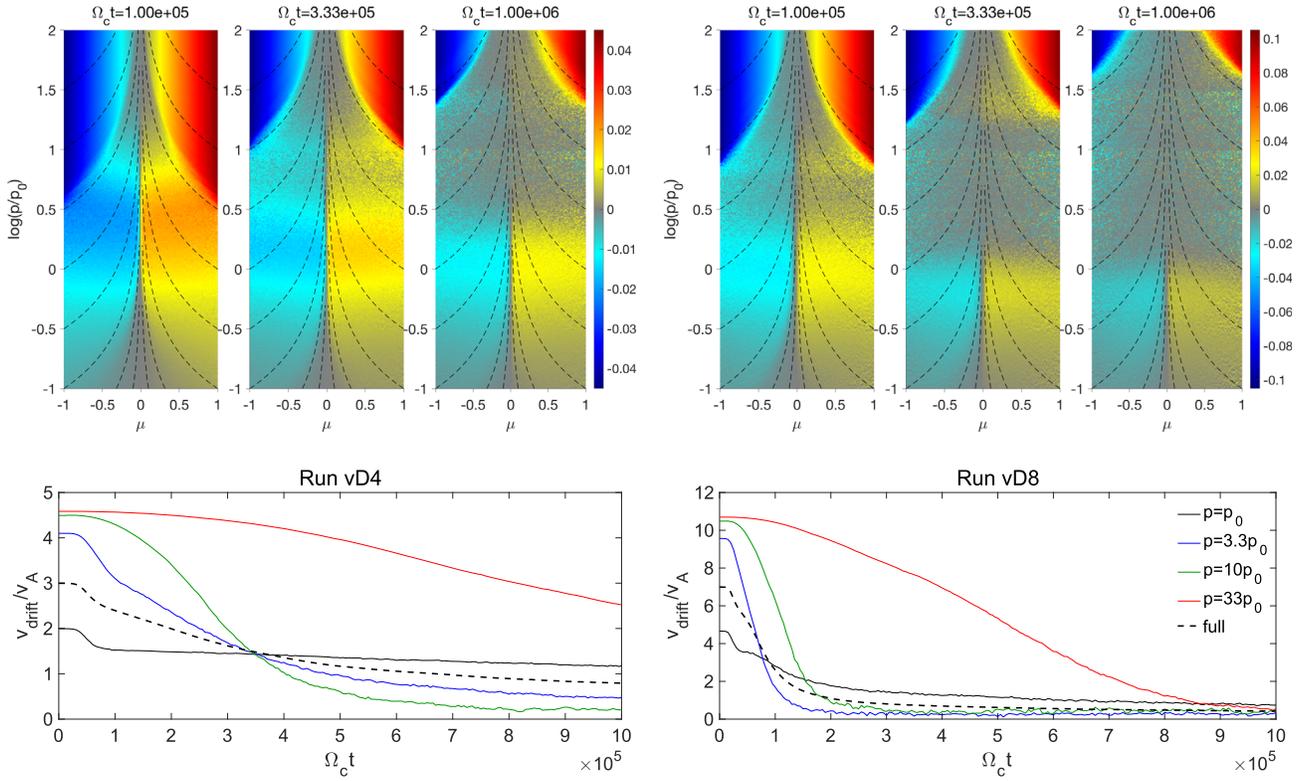


Figure 17. Distribution function and drift velocities for runs vD4 (left) and vD8 (right). Top panels: 2D distribution function $\delta f_w/f_0$ in the wave frame at three snapshots. Dashed lines are contours in momentum space with the same resonance condition. Bottom panels: time evolution of particle drift velocity in the wave frame for different momentum bins, as well as the full drift velocity (see legend).

enter (corresponding to the $n > 1$ terms), allowing particles to be reflected over multiple gyrations in a fluctuating field. Being less effective, it is consistent with fewer crossing particles found in our particle counting in lower-energy bins.

6. Simulation Results: Parameter Study

Our fiducial run selects a parameter range that is physically more realistic but numerically challenging, with small n_{CR}/n_i and low CR drift velocity v_D , leading to a relatively small growth rate and low wave saturation level. In this section, we vary the parameters around our fiducial run and conduct a two-part parameter study. In Section 6.1, we examine the saturation of the CRSI when v_D is increased to higher values. In Section 6.2, we discuss the results of varying solely numerical parameters for our run Fid.

6.1. Dependence on Drift Speeds

With v_D increased to $4v_A$ and $8v_A$, we find that we recover the linear growth rates very well (not shown), with the left- and right-handed modes growing at comparable rates (as expected). Figure 4 includes the time evolution of wave energy for runs vD4 and vD8 (dashed–dotted lines). They initially follow very similar evolutionary paths as our run Fid, except that everything (linear growth and QLD) proceeds faster. More notably, the later evolution of these two runs is more similar to run M3, where the wave energy eventually plateaus, indicating that saturation is close to completion.

In Figure 17, we show results related to the quasi-linear evolution of runs vD4 and vD8. The overall outcomes are in between those of run Fid and run M3. Crossing 90° becomes easier as v_D increases. By the end of the simulations, particles

with $4p_0 \lesssim p \lesssim 20p_0$ are roughly fully isotropized in run vD4. For run vD8, the range enlarges to $p_0 \lesssim p \lesssim 40p_0$.¹¹ Lower-energy particles are in the process of (but have not yet completed) isotropization. Isotropization is limited by the rate of 90° crossing, which is much slower for these lower-energy particles (see discussion in Section 5.4).

6.1.1. The Saturation Level

At this point, we can gather all simulation runs discussed so far to address the saturation level of the CRSI as a function of v_D . The saturation level is measured in terms of wave energy density. With equipartition of magnetic and kinetic energies, it can be simply represented by $\langle \delta B^2/B_0^2 \rangle$ in dimensionless form. Theoretically, the predicted wave energy saturation level is obtained by equating Equations (22) and (24) and is proportional to $(n_{\text{CR}}/n_i)\Delta v$. Integrating the derivative of the distribution function numerically, the result is shown as the dashed line in Figure 18.

The final wave energy density obtained from our runs Fid, vD4, vD8, M3, and Fid-hires (discussed in the next subsection) are also shown in Figure 18. Even though runs M3 and vD8 are close to complete saturation, their final wave energy densities appear to be far from the expected level. This is because numerical dissipation, while very low in the Athena MHD code, is still nonnegligible for the extremely long simulation

¹¹ Two numerical artifacts can be identified in Figure 17. The banded horizontal feature in the top panels is due to the division into eight particle momentum bins. Within each bin, there is more noise toward higher energies because there are fewer particles. One may also find that particle mean drift velocities at various energies do not reach zero, even they are theoretically expected to be fully isotropized (bottom panels). This is related to coarse particle binning during outputs (mentioned earlier in Section 5.3.2).

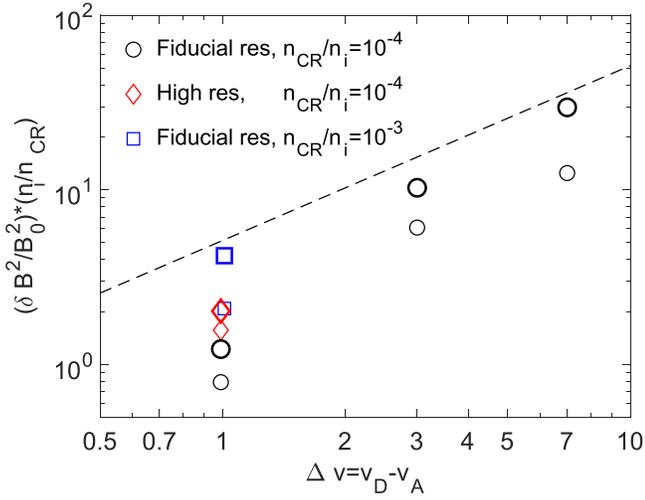


Figure 18. Saturation level measured from our simulations, in terms of wave energy density normalized by CR fraction. Black circles are from runs Fid, vD4, and vD8; red diamonds are from run Fid-hires; and blue squares are from run M3. Small/thin symbols are measured based on the total wave energy density, while large/thick symbols are measured from the same runs but corrected for numerical dissipation. The dashed line indicates the expected saturation level assuming all CRs are fully isotropized, obtained by setting Equations (22) and (24) equal.

time necessary to reach saturation.¹² With an adiabatic equation of state, we can capture the change in gas internal energy over time (thanks to the excellent energy conservation properties in the gas component). We find that it increases steadily as result of numerical dissipation and reaches a level that is comparable to the overall wave energy density by the end of the runs. When we add this energy lost to dissipation to the direct wave energy, we find much more reasonable saturation levels, as shown with large/thick symbols in Figure 18. In particular, run Fid remains at a level below the full saturation prediction because particles become “stuck” at 90° , whereas runs vD4 and vD8 are closer to the expected saturation level as more particles are fully isotropized. Run M3, which, based on its distribution function, is clearly isotropized, is in excellent agreement with theoretical expectations for the wave amplitude.

6.2. Numerical Parameters

Here we examine our runs Fid-hires, Fid-short, Fid-Np64, and Fid-Np1024, in which we keep the physical parameters the same and vary only the numerical parameters (note that for run Fid-hires, we have slightly reduced the box size to save computational cost). In Figure 19, we show the time evolution of the wave energy density of these runs. Overall, these three runs proceed very similarly to run Fid.¹³ Here QLD proceeds slightly slower in run Fid-short, but it catches up to the other runs shortly after. Run Fid-hires eventually grows to reach larger wave amplitudes, which we will show later is due to more efficient crossing of the 90° pitch angles.

In Figure 20, we show the wave spectrum from these runs at time $t = 2 \times 10^5 \Omega_c^{-1}$. The overall shapes of the spectra are comparable to each other. The high-resolution run has less numerical dissipation, allowing the wave spectra to extend to

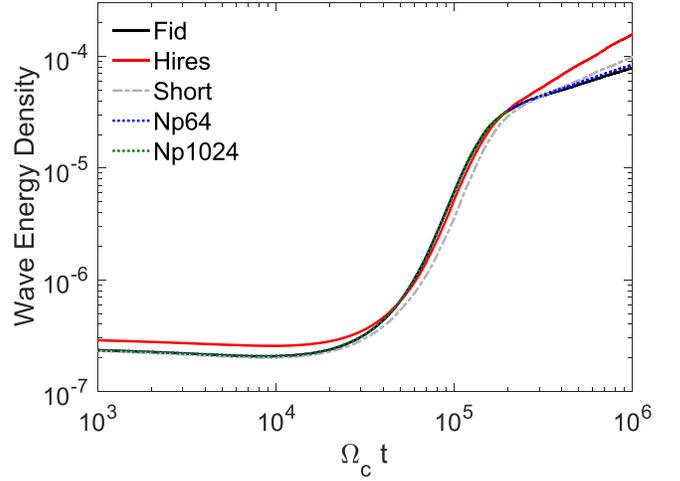


Figure 19. Time evolution of total wave intensity for runs with different numerical parameters.

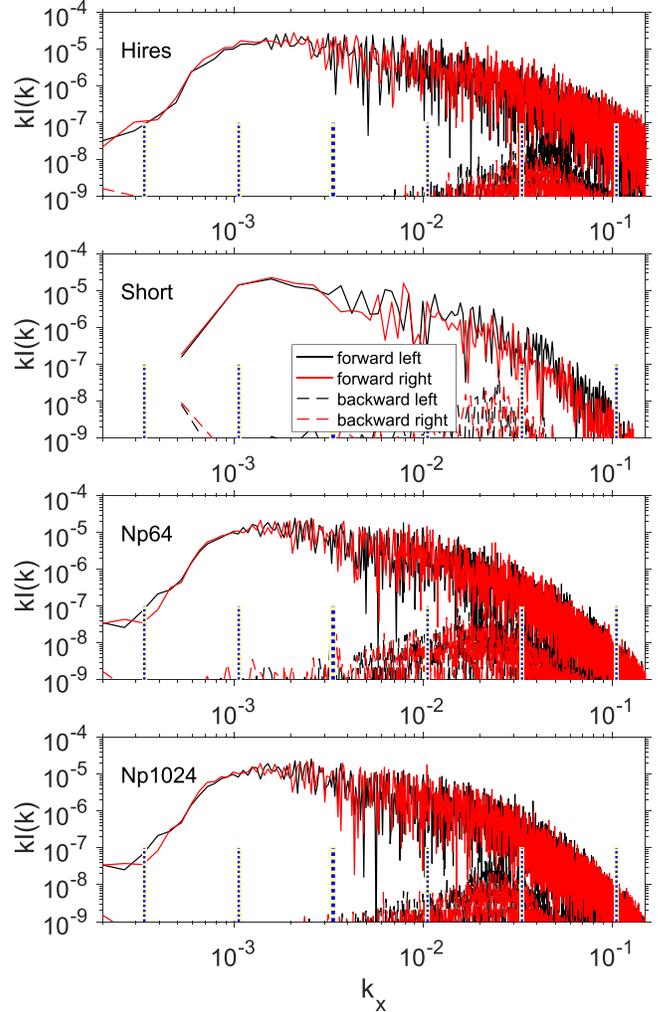


Figure 20. Wave spectrum at time $t = 2 \times 10^5 \Omega_c^{-1}$ from runs Fid-hires, Fid-short, Fid-Np64, and Fid-Np1024. The vertical dotted lines are identical to those in Figure 8 (the location of the thick line indicates the most unstable wavenumber).

larger k . Run Fid-short has a box size that is only one-third of the fiducial box (but still long enough to fit ~ 17 of the most unstable modes); hence, the modes are a lot more sparsely sampled, especially at low- k . This is likely related to its initially

¹² Dissipation of low-amplitude waves imposed in the initial conditions is also evident at early times in Figure 4.

¹³ All of these runs have the same initial wave amplitude $A = 10^{-4}$ (see Equation (42)). Run Fid-hires thus has more initial wave energy because of the larger spectral range it covers.

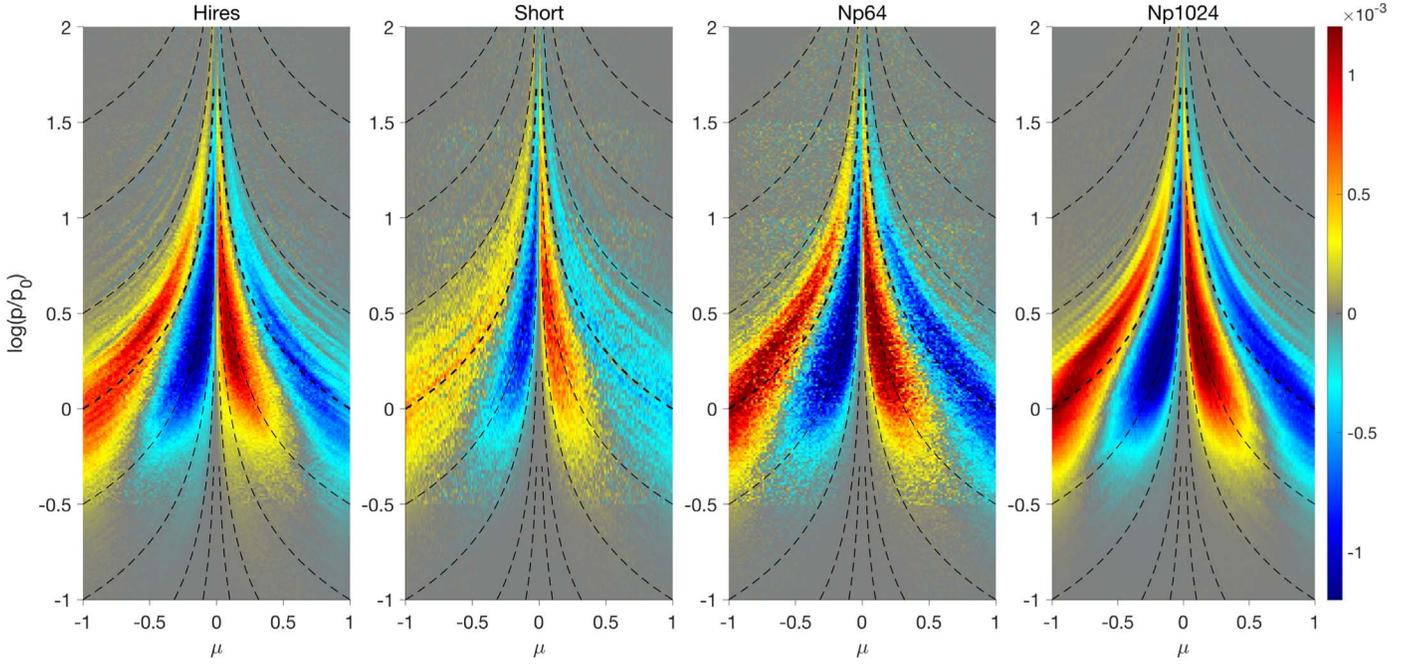


Figure 21. The 2D distribution function $\delta f/f_0$ in the lab (simulation) frame at time $t = 10^5 \Omega_c^{-1}$ for runs Fid-hires, Fid-short, Fid-Np64, and Fid-Np1024. The dashed lines are contours in momentum space that are resonant with the same wave ($p\mu = \text{const}$), with the thick line marking $p\mu = p_0$. The color limits are identical among the four panels.

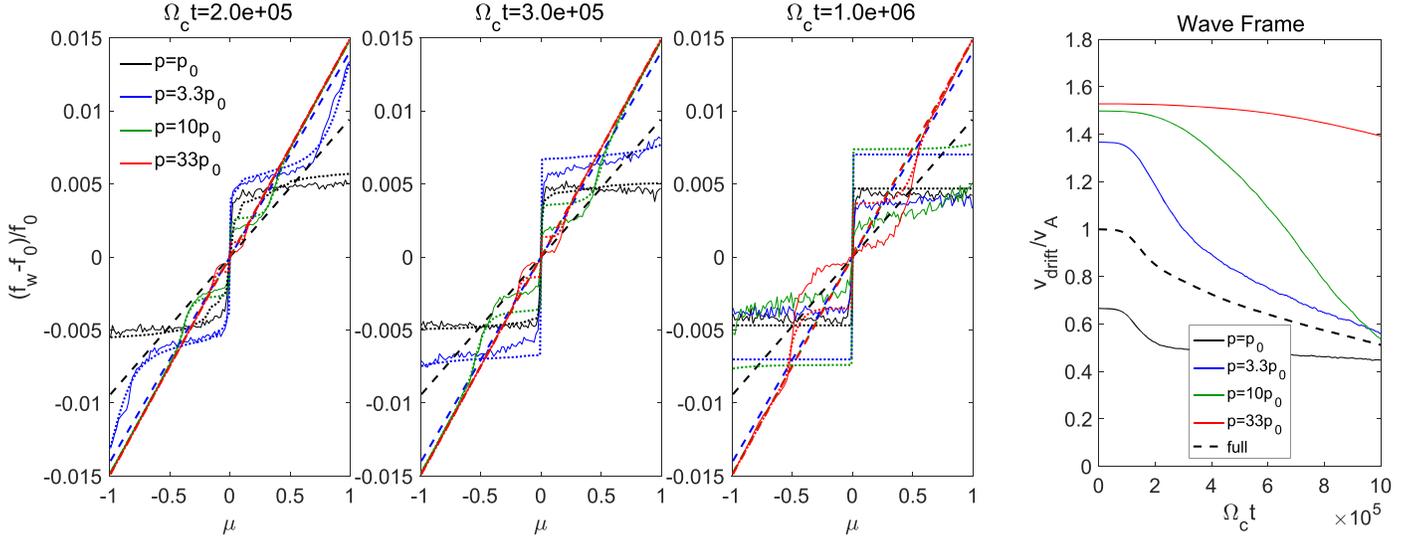


Figure 22. Results from run Fid-hires. Left three panels: snapshots of particle distribution function in the wave frame $(f_w - f_0)/f_0$ for four representative momenta as a function of pitch angle cosine $\mu_w \approx \mu$. Solid lines are measured from the simulations, dashed lines indicate the initial distribution, and dotted lines are reconstructed distribution functions obtained by solving the QLD Equation (14) in time. Right panel: time evolution of particle drift velocity in the same four momentum bins in the wave frame (solid lines). The thick dashed line represents the mean drift velocity averaged over the entire particle population.

slower quasi-linear evolution. We have also tried simulations with one-eighth of the fiducial box size (not shown); while evolution is smooth, the lack of power at larger scales starts to adversely affect the evolution of higher-energy particles.

Figure 20 shows that there are backward-propagating modes excited in all of these runs. The amplitude of these modes reflects the level of noise. Clearly, inserting more particles helps reduce the noise. The bottom line for properly choosing the particle number is that the amplitude of such backward-propagating modes should always be much smaller than the

forward-propagating modes driven by the CRSI. In this regard, our run Fid-Np64 is acceptable, but more caution should be exercised if one were to use even fewer particles.

In Figure 21, we show the particle distribution functions in the wave frame at time $t = 10^5$. These may be compared with Figure 7. Overall, runs Fid-short and Fid-Np64 are more noisy, whereas run Fid-Np1024 appears exceptionally smooth. Our run Fid lies in between. Despite these differences in noise levels, subsequent QLD does not appear to be affected.

Of particular interest is the high-resolution run Fid-hires, where waves grow to substantially higher amplitudes. In Figure 22, we show three snapshots of the particle distribution function in the wave frame for four representative momentum bins. These frames are to be contrasted with Figure 9 for the same snapshots. Clearly, in this run, a substantial fraction of particles with $p \gtrsim 3p_0$ has managed to cross the 90° barrier. This fraction steadily increases with time and, were we to run the simulation longer (e.g., to $t = 2 \times 10^6$), the particles in these momentum bins would essentially be fully isotropized. This trend can also be seen from the last panel, which shows the time evolution of the mean CR drift velocity: v_d for particles with $p \gtrsim 3p_0$ are steadily decreasing toward zero. Clearly, crossing the 90° barrier is easier compared to run Fid. This is natural because higher resolution allows more small-scale modes to be better preserved, thus enhancing the chance for nonlinear wave-particle interaction to reflect the particles (see Section 5.4.2). In Figure 18, we see that the wave energy density in our run Fid-hires is also a factor of several closer to the expected saturation level. We expect that increasing the resolution further (which would be much more computationally costly) would eventually lead to full isotropization of the bulk CR populations.

In sum, our exploration of numerical parameters suggests that full saturation of the CRSI (ideal case, no damping) can be achieved as long as there is sufficient numerical resolution. The required resolution is mainly limited by the demands imposed by crossing the 90° pitch angle, which is sensitive to wave amplitudes at small scales. Similar results to our fiducial model can be achieved when one reduces the number of particles and/or the box size at some modest level (\sim a factor of 3).

7. Discussion

7.1. CR Feedback to Background Gas

Through the CRSI and QLD, CRs transfer momentum to background gas, which serves as an important dynamical feedback mechanism in a broad range of astrophysical environments. In the following, we take run vD8 as an example and discuss CR feedback in terms of momentum and energy.

Momentum feedback is manifested in the top panel of Figure 23, which shows changes in the mean gas and CR momentum density ($\Delta \mathcal{P}_g$ and $\Delta \mathcal{P}_{CR}$). In our simulation frame, the initial CR momentum is zero, and the initial gas momentum is negative. Over time, the CR momentum density decreases, accompanied by an increase in gas momentum density of exactly the same amount. Physically, the back-reaction from the CR isotropization process effectively gives rise to a “parallel force” F_{\parallel} . This parallel force is mediated by forward-propagating Alfvén waves and transfers CR momentum to the gas (e.g., Thomas & Frommer 2019). The amount of momentum transferred to the gas is in fact the same as the “effective” wave momentum defined in Equation (24), which is the momentum needed to excite such forward-propagating waves (Kulsrud 2005). By the end of our run vD8, although the evolution of gas and CR momenta are not yet fully saturated, both are approaching the expected saturation level indicated by the dashed lines, obtained from integrating Equation (22). In the case of run M3 (not shown here), we observe that saturation

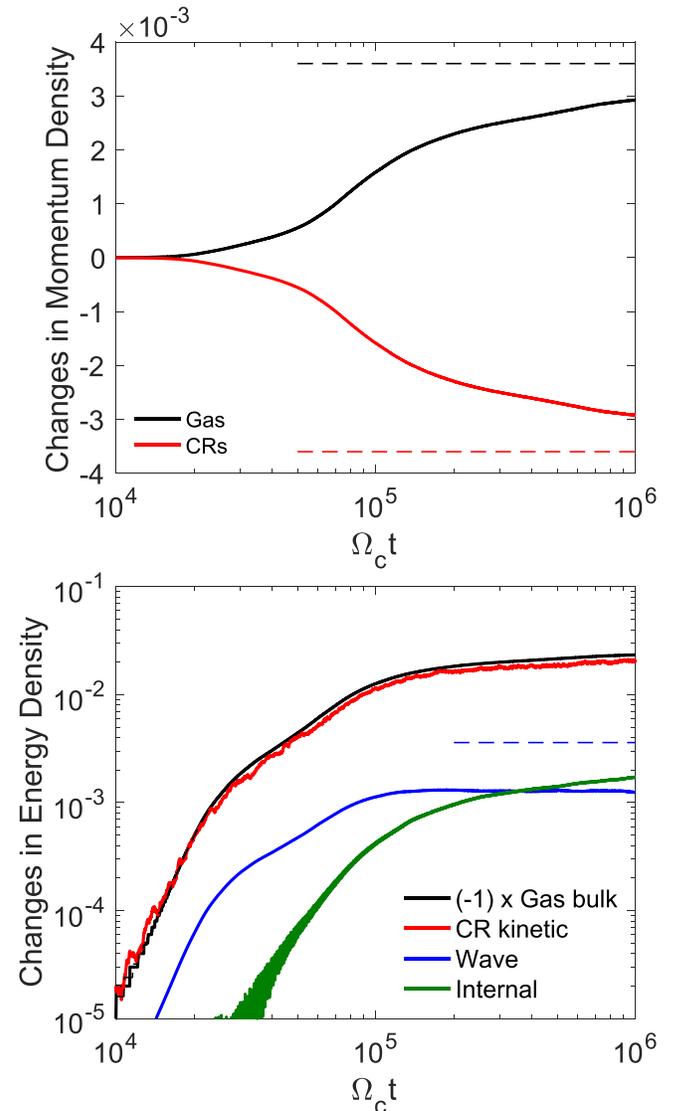


Figure 23. Time evolution of changes in momentum density (top) and energy density (bottom) in various components (labeled in the legends) in run vD8. The horizontal dashed lines mark the expected saturation levels (top: momenta; bottom: wave energy). See Section 7.1 for more details.

of gas and CR momenta is achieved. That is, the mean CR streaming velocity relative to the gas is reduced from $2v_A$ to v_A (Figure 13 shows that the final mean CR velocity is zero in the wave frame), and the gas momentum density is increased to balance the loss from CRs.¹⁴

Energy feedback can be discussed by decomposing the energy budget into various components. These include gas bulk kinetic energy $E_{k,bulk} = \langle \rho v_x^2 / 2 \rangle$, representing gas motion parallel to the background field; wave energy E_{wave} ; CR kinetic energy E_{CR} ; and internal energy of the gas $E_{internal} = P_g / (\Gamma - 1)$. The time evolution of these quantities

¹⁴ As mentioned in Section 3.2, exact momentum conservation is not achieved using the δf method. For run vD8, we have found that the error in total momentum at the end of the run is at the level of 3×10^{-6} , about a fraction 10^{-3} of the momentum change in the CR and gas components. The total error in energy at the end state is $\lesssim 10^{-3}$, which is much less than the changes in bulk gas/CR energies, though it might start to affect the accuracy in estimating wave dissipation. The fractional error in momentum/energy is typically comparable or smaller in other runs.

is shown in the bottom panel of Figure 23, where we only plot their changes relative to the initial condition. The overall energy conservation is expressed as

$$E_{k,\text{bulk}} + E_{\text{CR}} + E_{\text{wave}} + E_{\text{internal}} = \text{const.} \quad (47)$$

Note that $E_{k,\text{bulk}}$ and E_{CR} are frame-dependent. Here it suffices to report the results in our simulation frame.

In the simulation frame, the gas bulk energy density decreases over time (as the bulk gas velocity becomes less negative). This is associated by the work done by the same parallel force F_{\parallel} that changes the gas momentum. In fact, we can exactly reproduce the black solid line in the bottom panel of Figure 23 using F_{\parallel} inferred from acceleration in the gas (from the top panel), with $\Delta E_{k,\text{bulk}} \approx v_g \Delta \mathcal{P}_g$. (Since v_g only changes by a tiny fraction over the process, it is approximately taken to be constant here.) This same force also does opposite work on the CRs, increasing their total kinetic energy. Some of this energy is used to excite and amplify the waves, leading to their continued growth. In the meantime, the waves are slowly dissipated numerically, increasing the internal energy of the gas.¹⁵

We emphasize that the CRSI itself does not lead to heating of background gas. Heating only occurs through internal dissipation in the gas, which is generally accompanied by wave damping. This fact is reflected in the more recent formulation of the CR-(magneto)hydrodynamics of Thomas & Pfrommer (2019), though not necessarily correctly modeled in some earlier studies. If wave growth is balanced by wave damping in a quasi-steady state (not achieved in our present simulations), then the heating rate is simply determined by the rate of wave growth/damping.

7.2. Toward More Realistic Simulations

There are several directions to further improve the realism of our simulations. In particular, there are several wave-damping mechanisms to be incorporated, including ion-neutral damping, nonlinear Landau damping, and turbulent damping of waves.

Ion-neutral damping, or ambipolar damping, is well known (e.g., Kulsrud & Pearce 1969; Soler et al. 2016). It is present when background gas is partially ionized. The ions are directly coupled and respond to magnetic fields, but the neutrals hardly feel the magnetic fields and can possess a different bulk velocity from the ions. Ion-neutral collisions give friction and lead to damping of Alfvén waves in the ion fluid, especially at small scales. Given that crossing the $\mu = 0$ barrier is sensitive to the spectrum of waves at small scales, ambipolar damping may cause severe challenge to the isotropization of the CRs. This effect will be incorporated and studied in greater detail in a follow-up paper.

Nonlinear Landau damping is a result of wave-particle interaction at the Landau resonance between background thermal particles and a beat wave formed by two propagating waves (Lee & Völk 1973) and is more effective in high- β plasmas. Our MHD-PIC formulation does not capture the kinetic physics of the background plasma and hence cannot capture nonlinear Landau damping. However, this is not necessarily a limitation: we would lose the advantage of

MHD-PIC by resolving the kinetic physics, and, in principle, this damping mechanism can be incorporated in the form of subgrid physics.

Turbulent damping is due to the distortion of CR-generated Alfvén waves when they collide with oppositely directed wave packets originating from background turbulence. As a result, the wave energy cascades into smaller scales and is ultimately dissipated (Farmer & Goldreich 2004). This damping rate has only been calculated approximately (e.g., see the more recent work of Lazarian 2016). Incorporating this effect is fully compatible with our MHD-PIC framework and can be studied in future multidimensional simulations.

For this first study, we focused on the classical theory with 1D simulations with a uniform background density and magnetic field, which enable us to confirm resolution and box size requirements. Multidimensional extensions are straightforward and underway, although they are much more computationally expensive. Such extensions will allow modes propagating obliquely with respect to the background field to be investigated and open up more parameter space, as the relative importance of thermal, magnetic, and CR pressures and background turbulence can play certain roles.

Closely related to the CRSI is another type of gyroresonance instability, which is driven by CR pressure anisotropy (e.g., Lazarian & Beresnyak 2006) and can be easily triggered in compressible MHD flows (when CR pressure anisotropy exceeds $\sim v_A/c$). It excites circularly polarized Alfvén waves propagating in opposite directions, which then isotropize the CRs via QLD. Recently, Lebiga et al. (2018) studied this instability numerically based on the MHD-PIC framework in the full- f approach (with relatively strong anisotropy) and found good agreement with quasi-linear theory. In the absence of CR streaming, the CR distribution function is always symmetric about $\mu = 0$; hence, crossing the 90° pitch angle is not required for CR isotropization. In reality, CR streaming and pressure anisotropy likely coexist. This would excite a wide range of waves propagating in both directions, calling for further investigation.

Finally, in any simulation with periodic boundary conditions (even with arbitrarily high resolution such that crossing the 90° pitch angle is not a limitation), there is a finite free energy of the system based on the initial energy and momentum densities of gas and CRs. In the absence of damping (numerical or physical), growth of Alfvén waves would halt when QLD has reduced the mean velocity difference between the gas and the CRs to v_A . In reality, at a given location, the CRSI is typically driven by a large-scale CR pressure gradient. For instance, CRs are continuously generated from the galactic midplane region and diffusively propagate downhill through the gradient to escape from the galaxy. Energy and momentum are continually added to the system. This configuration likely yields continued wave growth balanced by damping to reach a steady state, a situation that has yet to be addressed with more realistic boundary conditions.

7.3. Comparison with Conventional PIC Simulations

Our MHD-PIC approach to studying the CRSI strongly contrasts with conventional full- or hybrid-PIC approaches that are widely used in modeling highly nonlinear plasma problems. The full-PIC approach treats charged particles of all kinds as particles. It resolves all scales down to the electron skin depth (eventually the Debye length); hence, all plasma wave modes

¹⁵ While damping in our simulations is purely numerical, we can see that it is negligible in the early phases of linear and quasi-linear evolution. Only toward later stages, where saturation is largely achieved (and hence there is little wave growth), does numerical dissipation become more noticeable. Thus, our results are largely unaffected by numerical dissipation.

are self-consistently captured along with relevant wave–particle interactions. The hybrid-PIC approach treats electrons as a massless conducting fluid. By compromising the electron-scale physics, it alleviates the stringent resolution requirement of the full-PIC method, while it can still capture the ion-scale physics well. For the CRSI problem, employing full- or hybrid-PIC methods means that background ISM fluid must be represented by particles, hence resolving the corresponding microscopic scales. This causes two major drawbacks: substantially enhanced numerical cost and PIC noise due to the finite number of particles representing the background plasma.

As discussed in Section 3, there is substantial scale separation in the CRSI problem between the CR gyro-radii and the ion skin depth (and below). Conducting a hybrid-PIC simulation using the same parameter settings as our fiducial simulations would require more than an order of magnitude more cells, and another order of magnitude more cells would be needed in the full-PIC case. The numerical cost in the full-PIC case is further exaggerated by the ion-to-electron mass ratio, requiring a factor of m_i/m_e more time steps per CR gyro-period. Moreover, the number of background particles per cell needs to be sufficiently high in order to reduce the noise floor as compared to the signal level of growing waves and avoid artificial QLD of the CRs by such noise.

Despite these difficulties, Holcomb & Spitkovsky (2018) recently presented a 1D study of the CRSI using full-PIC simulations. The typical parameters of their simulations are $m_i/m_e = 100$, $v_A/c = 0.1$ (note in full-PIC, $c = \mathbb{C}$ in our notation), 10/100 cells per electron/ion skin depth, $N_{\text{ppc}} = 50 - 1000$ (all plasma particles equally divided among background and CR species), $n_{\text{CR}}/n_i \in [2 \times 10^{-4}, 2 \times 10^{-2}]$, and v_D/v_A from 1.4 (low anisotropy) to 7.9 (high anisotropy).¹⁶ They considered two different CR distribution functions, ring and power law, with the former representing a monoenergetic CR distribution in the drift frame. They have successfully reproduced the peak growth rate in runs with the ring distribution and, similar to our results, achieved full isotropization in runs with relatively large n_{CR}/n_i and/or CR drift velocity.

We note that the large ratio of v_A/c adopted in Holcomb & Spitkovsky (2018) makes the resonant condition (Equation (2)) asymmetric with respect to forward- and backward-traveling particles, thus modifying the standard CRSI dispersion relation for left-/right-handed modes in different ways. Together with the choices of large n_{CR}/n_i ratio, drift velocity being mildly relativistic, and use of a ring distribution, these parameters represent substantial compromises to alleviate the issue of scale separation and particle noise. As a result, reasonable agreement with linear theory was found mainly in the special case of the ring distribution and power-law distribution with exaggerated high anisotropy and high n_{CR}/n_i values. These more extreme run parameters (approaching the Bell regime), together with the asymmetry in the left-/right-handed modes caused by large v_A/c , make the left-handed modes in these simulations grow at much reduced rates compared with the right-handed modes, leading to asymmetric scattering for forward-/backward-traveling CRs. While this situation might apply near strong shocks where the Bell instability operates, the parameter regime is not representative of typical conditions in the ISM.

¹⁶ Let us recall that in the present study (Fid run), the parameters are $m_i/m_e \rightarrow \infty$, $v_A/\mathbb{C} = 3.3 \times 10^{-3}$, 10 ion skin depths per cell, $N_{\text{ppc}} = 2048$ (CRs only), $n_{\text{CR}}/n_i = 10^{-4}$, and $v_D/v_A = 2$.

Additionally, they have suggested mirror reflection as being responsible for crossing the 90° pitch angle. However, the reflection event shown in their Figure 10 in fact appears more consistent with nonlinear wave–particle interaction, as we have demonstrated in this work.

Perhaps the chief advantage of full- and hybrid-PIC methods over the MHD-PIC approach is the ability to capture nonlinear Landau damping, as the background plasma is treated kinetically. However, the physics of nonlinear Landau damping is highly subtle, and damping relies on second-order longitudinal electric fields resulting from a beat wave formed from two circularly polarized Alfvén waves (Lee & Völk 1973). It is conceivable that capturing such effects must require an extremely low level of background noise, and hence an exceedingly large number of background plasma particles: another major challenge yet to be resolved.

Finally, we comment that several of the difficulties in simulating the CRSI outlined in Section 3 are common for all variations of PIC methods. The δf approach coupled with a κ distribution that we have adopted could also be incorporated in full- and hybrid-PIC simulations, likely leading to substantial improvements in future studies of the CRSI.

8. Summary and Conclusions

In this paper, we have presented the first numerical studies of the CRSI using the MHD-PIC method. We use 1D simulations in a periodic box and an idealized setup (no damping, a reduced speed of light \mathbb{C}) for a range of model and numerical parameters. We have been able to accurately reproduce the theoretically predicted linear growth rate, follow quasi-linear evolution as CR particles interact with CSRI-driven Alfvén waves, and identify the mechanism responsible for CR particles crossing the 90° pitch angle. In the course of this study, we have developed techniques to investigate CRSI and QLD with substantial precision, paving the way for future studies with more realistic setups and incorporating more physics.

Unlike conventional PIC methods, the MHD-PIC method has unique advantages in bypassing the microscopic scales of the background thermal plasma, allowing one to focus on the interaction between the CRs and background gas around the gyro-resonant scale. However, the CRSI still involves substantial scale separation, ranging from scales much smaller than to scales much larger than the CR gyro-resonant scale. Covering this large range of scales is necessary to properly capture the crossing of the 90° pitch angle and QLD. Moreover, the extremely weak level of anisotropy ($\sim v_A/c$) in the CR distribution function over a wide energy range makes it challenging to properly represent with a sufficiently large number of particles.

To address these challenges, we have extended the MHD-PIC formulation to use a δf (instead of full- f) representation of the CR particles on top of a known initial equilibrium distribution function f_0 , which dramatically reduces the Poisson noise. While our fiducial simulations employ as many as 2048 particles per cell to achieve better precision, similar results can be achieved with fewer particles, and we recommend testing the level of Poisson noise in a non-CR-streaming simulation to determine the optimal number of particles. Using the conventional full- f representation, while the signature of the CRSI can be identified, the high noise level precludes any quantitative measurement of its overall properties.

The δf method requires the background distribution function f_0 to be continuous, which is incompatible with f_0 being a truncated power law. We have therefore chosen f_0 to be a κ distribution, and the corresponding dispersion relation of the CRSI is analytically derived. We note that under realistic large-scale ISM conditions (n_{CR}/n_i extremely small and moderate v_D/v_A), the left- and right-handed Alfvén waves grow at equal rates. The Bell-type modes, in which the right-handed modes grow faster, emerge when n_{CR}/n_i and/or v_D/v_A become large. We have carefully chosen the parameters to ensure sufficient scale separation $v_A/c = 1/300$ and that the left- and right-handed modes growth at about the same rates (except for run M3, where there is a small difference), so that the conditions are as close to the ISM as possible. Our particle population also covers a wide range of CR momenta from $0.01p_0$ to $100p_0$, with $p_0 = c$ being the characteristic momentum of the κ distribution.

We have found that when employing a standard periodic boundary condition, properly capturing the linear growth rate requires an excessively long simulation box, corresponding to the box crossing time of a typical particle approaching the wave growth time. We attribute this to the fact that the growth of the CRSI is driven by QLD, which requires particles to encounter a sufficient number of independent wave packets, and is related to the random phase approximation commonly employed in describing QLD. To address this issue, we implemented a phase randomization technique for particles that cross the periodic boundaries, which allows us to achieve even better results using much smaller simulation boxes (box size $\lesssim 50$ instead of $\gtrsim 1000$ of the most unstable wavelength).

Equipped with these techniques, we are able not only to accurately reproduce predicted CRSI linear growth rates over a broad range of wave spectra but also to accurately follow the quasi-linear evolution of the system. We observe that QLD proceeds to isotropize the CR distribution function f_w in the frame of forward-propagating Alfvén waves, so that $\partial f_w / \partial \mu_w$ approaches zero on either side of the pitch angle cosine $\mu_w \approx \mu = 0$. With an increasingly flat distribution of particle pitch angles, wave growth slows down and eventually stops. As longer-wavelength modes that resonate with higher-energy CRs grow more slowly, wave growth proceeds from the most unstable to longer wavelengths (and, in principle, also to shorter wavelengths, though we lack resolution for these). In parallel, quasi-linear evolution proceeds from low- to higher-energy CRs as waves become available.

Whether full isotropization of CRs in the wave frame is achieved depends on whether the particles manage to overcome the $\mu = 0$ barrier (crossing the 90° pitch angle). The mechanisms involved go beyond quasi-linear theory. While mirror reflection is commonly invoked as the main mechanism enabling crossing $\mu = 0$, we have found that reflection generally occurs via nonlinear wave-particle interaction, with particles encountering abrupt changes in field orientation. Reflection is more easily achieved for higher-energy particles and facilitated by the presence of waves at wavelengths much smaller than the gyro-resonance scale (see Section 5.4 for details).

In our study, efficient crossing of the $\mu = 0$ barrier is achieved in simulation runs with strong wave growth (runs M3 and vD8). For these runs, full saturation is reached, with the mean CR drift velocity reduced to the Alfvén speed. The final particle distribution functions, as well as saturation amplitudes

of the waves, are in excellent agreement with theoretical expectations. Our fiducial run, with a relatively low $n_{\text{CR}}/n_i = 10^{-4}$ and small initial drift speed $v_D = 2v_A$, becomes stuck in a state of premature saturation: while the distribution is flat on each side of $\mu = 0$ within most momentum bins, the bulk of the CR population fails to overcome the $\mu = 0$ barrier. The inability of particles to breach the $\mu = 0$ barrier is primarily due to numerical dissipation of low-amplitude waves at small scales, together with the absence of even shorter-wavelength waves due to limited resolution. We have further found that enhancing numerical resolution can substantially alleviate the situation. Thus, the proper choice of numerical resolution for the given physical parameters is also crucial for saturation.

The CRs provide direct momentum feedback to the background gas, which accelerates in accordance with momentum conservation of the gas-CR system. Wave power is directly proportional to the momentum transferred from the CRs to the gas. In terms of energy feedback, the work done by the CRs leads to gas acceleration and wave growth, but no direct heating. Gas heating is only achieved through wave damping, which is purely numerical in our simulations.

As a first study, we have restricted ourselves to 1D simulations in a periodic box in the ideal MHD limit. Our explorations have identified and clarified the numerical methods necessary to properly capture the essential mechanisms of the CRSI and paved the way for future investigations to incorporate more realistic physics. These may include implementing various damping mechanisms, extensions to multiple dimensions, and employing more realistic boundary conditions.

We thank Cole Holcomb, Matt Kunz, Peng Oh, and Anatoly Spitkovsky for useful discussions. X.N.B. acknowledges support by the Youth Thousand Talent program. The work of E.C.O. was supported by grant 510940 from the Simons Foundation. The work of I.P. was supported by NSF grants PHY-1804048 and PHY-1523261 and facilitated by the Max-Planck/Princeton Center for Plasma Physics. Computation of this work has been conducted at computational facilities at the Princeton Institute for Computational Science and Engineering (PICSciE) and from GyroTech Co. Ltd.

Appendix A Decomposition of Alfvén Waves

Given that modifications from the CRs to the Alfvén waves are tiny in this context, it suffices to consider pure linear Alfvén waves. At a fixed wavelength, there are four independent Alfvén modes, depending on left-/right-handed polarization and the direction of propagation relative to the fluid rest frame. Here we take the perspective that the left-/right-handedness is defined from the receiver’s point of view. That is, for a wave propagating toward an observer (relative to the fluid rest frame), the perturbed electric (and magnetic) field vector rotates in a counterclockwise (clockwise) sense for a left- (right-) handed wave.

Let δv and $\delta \mathbf{B}$ be perturbed transverse velocities and magnetic fields and take the perturbations in the form of $\exp[i(kx - \omega t)]$. Note that k can have both signs indicating the direction of propagation, whereas ω always has a positive real

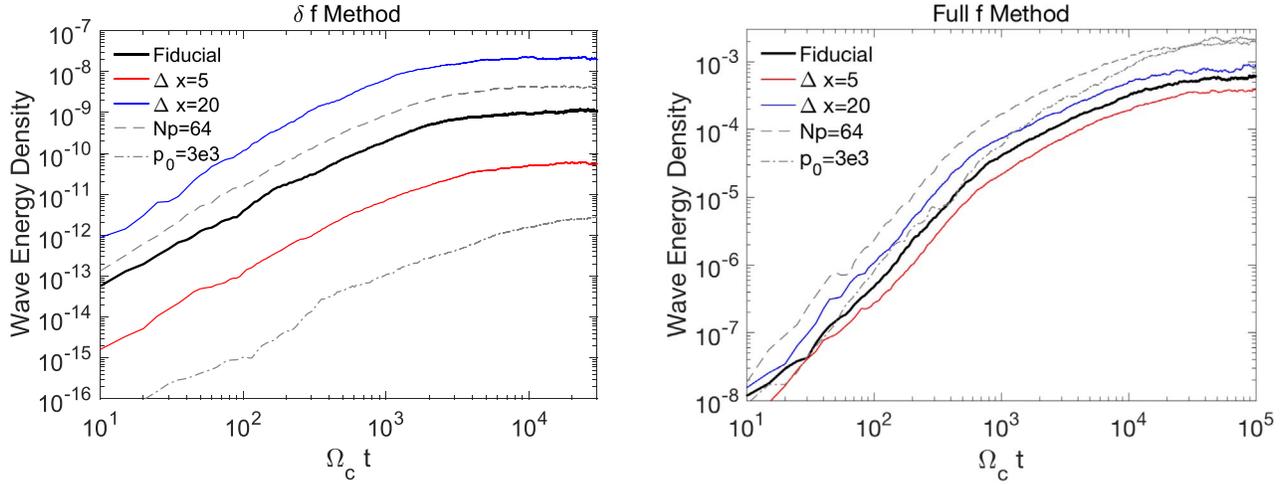


Figure 24. Time evolution of system noise (in terms of total wave energy density) in the zero-drift simulations described in Appendix B. The left and right panels show the results from the δf method and the full- f method, respectively. Thick black lines in both plots correspond to simulations with identical setup to the fiducial simulation, except that $v_D = 0$, and $n_{CR}/n_i = 10^{-3}$ in the δf case while $n_{CR}/n_i = 10^{-4}$ in the full- f case. Other lines marked in the legend indicate runs where one of the parameters is varied while all others are unchanged.

part. The eigenvectors for Alfvén waves satisfy

$$\frac{\delta \mathbf{v}}{v_A} = -\text{sgn}(k) \frac{\delta \mathbf{B}}{B_0}. \quad (48)$$

The sign of B_0 is implicitly contained in the above, whereas v_A is considered to be positive. In the following, we always take $B_0 > 0$ without a loss of generality (for $B_0 < 0$, sign changes must be employed in many of the subsequent descriptions) and consider forward propagation to be along the direction of \mathbf{B}_0 (positive \hat{x}). Under this convention, the \pm sign of k corresponds to forward/backward propagation. The left-/right-handed modes have $\delta B_y = \mp i \delta B_z$ and $\delta v_y = \mp i \delta v_z$ for forward-propagating waves. The opposite sign should be adopted in the above for backward-propagating waves.

Based on the above description, the natural way to decompose circularly polarized Alfvén modes is to consider the following combination:

$$C^\pm(x) = \frac{1}{2} \left[\frac{\delta v_y \pm i \delta v_z}{v_A} - \text{sgn}(k) \frac{\delta B_y \pm i \delta B_z}{B_0} \right]. \quad (49)$$

For instance, for forward-propagating ($k > 0$), right-handed ($\delta B_y = i \delta B_z$) waves, we choose the plus sign. A discrete Fourier transform of $C^\pm(x)$,

$$W^\pm(k_i) = \frac{1}{N} \sum_{n=0}^{N-1} C^\pm(x_n) e^{-ik_i x_n}, \quad \text{where } k_i = \frac{2\pi}{L_x} i, \quad (50)$$

gives the wave amplitude and phase (which is what is being plotted in the figures). The wave intensity is then given by

$$I^\pm(k) = |W^\pm(k)|^2 \frac{L_x}{2\pi}. \quad (51)$$

To initialize a wave spectrum with $I^\pm(k) = A^2/k$, we can simply choose $|W^\pm(k)| = A \sqrt{\Delta k/|k|}$, where $\Delta k \equiv 2\pi/L_x$.

Appendix B Level of Noise in the δf and Full- f Methods

In this appendix, we study the noise properties in our simulations. We do so by conducting simulations with an

almost identical setup to our fiducial runs, except that we change the CR drift velocity to $v_D = 0$ and start from an initial amplitude that is much smaller with $A = 10^{-12}$. We also use a shorter simulation box, $L = 24,000$ (which we verify has no influence on the result). In this case, we only expect wave damping to occur, but Poisson noise from particle sampling will stand out and causes fluctuations that overwhelm the initially low-amplitude waves.

In Figure 24, we show the time evolution of the mean fluctuation energy density in the system, which is given by $(\rho v_\perp^2 + B_\perp^2)/2$. This is done for both the δf method (left) and the full- f method (right). Note that we have chosen $n_{CR}/n_i = 10^{-3}$ in tests using the δf method (a smaller value would take a longer time to reach steady state), whereas we use $n_{CR}/n_i = 10^{-4}$ in the full- f case (a larger value would make the noise somewhat too large). Clearly, we see that even with a higher n_{CR}/n_i , the noise with the δf method is dramatically smaller than that in the full- f method by more than 5 orders of magnitude (this would be more than 6 orders of magnitude if they had the same n_{CR}/n_i).

The energy density of the waves in this test results from the Poisson noise in the particles. We can see that in both cases, reducing the number of particles by a factor of 4 leads to a factor of ~ 4 increase in wave energy density (or a factor $\sqrt{4}$ in wave amplitude). The wave spectrum (not shown) is consistent with white noise (with $I(k)$ being flat) for all modes with a cutoff at high k due to numerical damping. The spectrum at intermediate k grows and saturates faster, whereas modes at the low- k end grow slower, explaining the secular trend at late times.

Interestingly, the noise level also depends on resolution, but the dependence is different for the full- f and δf methods. For the δf method, we find a sensitive dependence, which scales roughly as Δx^4 , and higher spatial resolution can greatly reduce Poisson noise. On the other hand, for the full- f method, the dependence is much weaker and sublinear. In addition, we find that using a larger p_0 while keeping other parameters fixed also gives very different noise levels. Enlarging p_0 by a factor of 10, noise is dramatically reduced by a factor of ~ 300 when using the δf method. Noise is enhanced by a factor of ~ 3 – 4 when

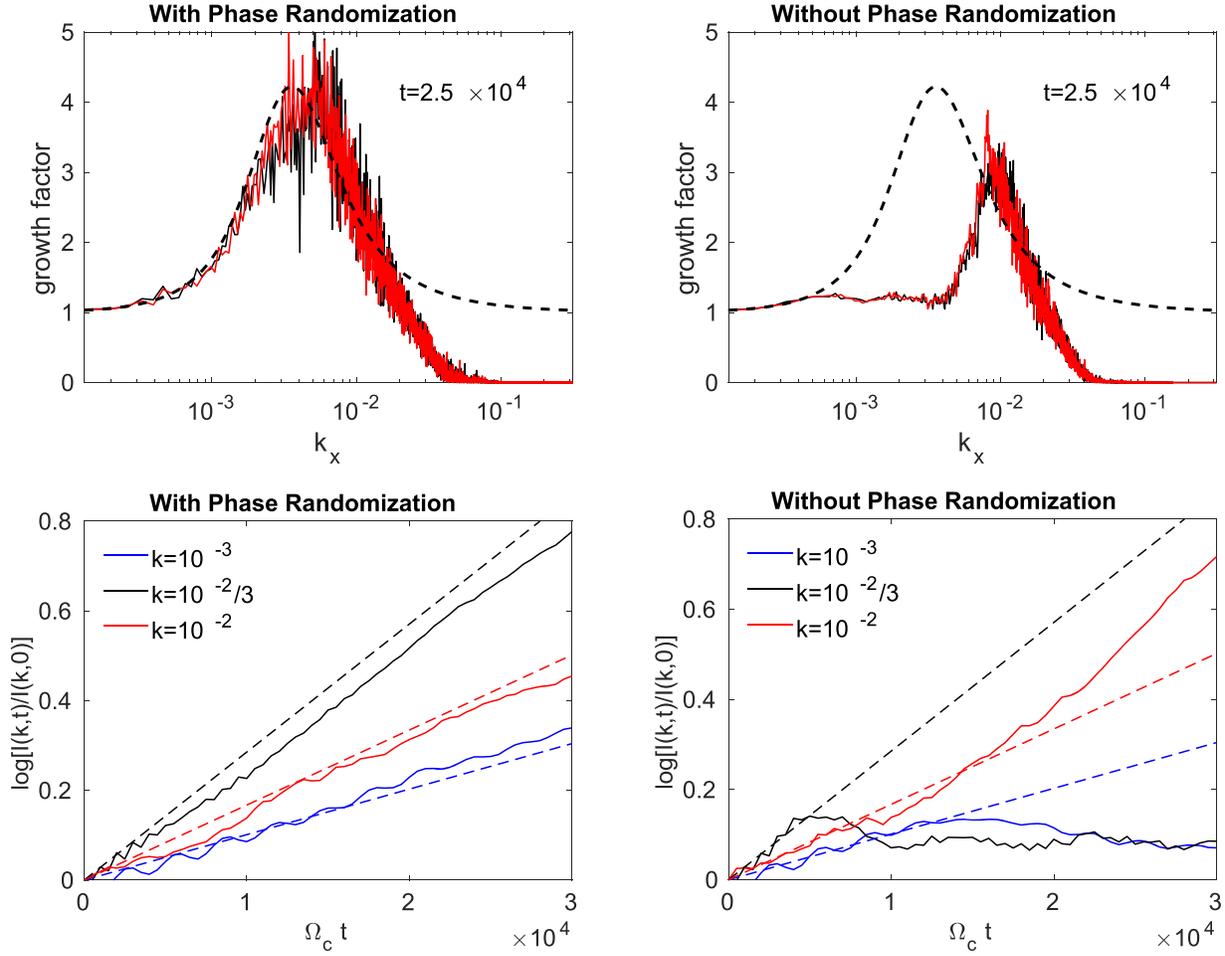


Figure 25. Comparison between fiducial simulations with (left) and without (right) implementing phase randomization across periodic boundaries. The top plots show the growth factor of wave intensity $I(k)$ between $t = 0$ and $2.5 \times 10^4 \Omega_c^{-1}$, with red/black lines marking left-/right-handed modes. The dashed line shows the growth factor expected from linear theory. The bottom plots show the time history of the growth factor for individual modes with different wavenumbers k_x (indicated in the legend). The dashed lines of the same color indicate the expected growth factor evolution.

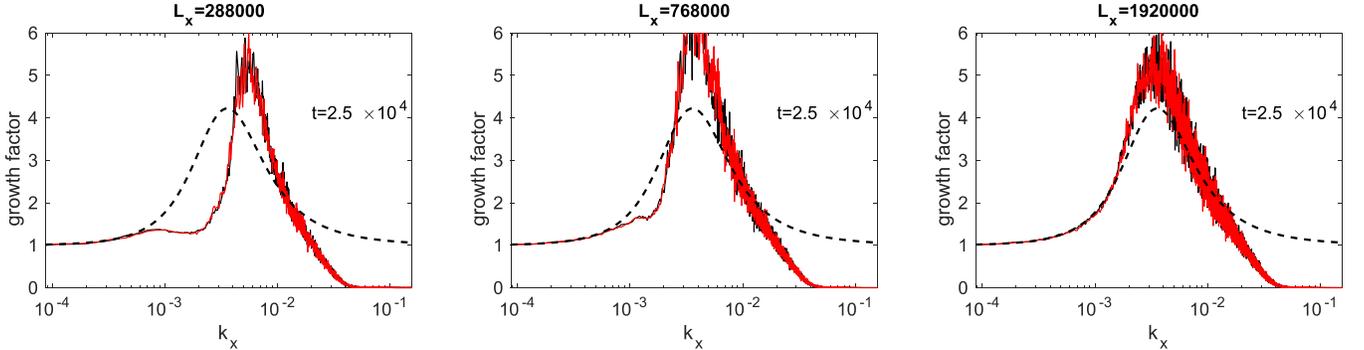


Figure 26. Growth factor of wave intensity $I(k)$ at time $t = 2.5 \times 10^4$ for three runs without phase randomization. Red/black colors mark left-/right-handed modes.

using the full- f method. We do not aim to further study and explain these trends but simply report them here for reference.

Appendix C

Random Phase Approximation and Simulation Box Size

In this appendix, we show that simulations without implementing phase randomization across periodic boundaries fail to properly capture the linear growth rate of the CRSI unless the simulation box size is extremely long. This failure reflects

the fact that the random phase approximation must be satisfied for particles to properly follow QLD, which drives the CRSI.

We first show in Figure 25 a comparison between our fiducial run with and without phase randomization. At some early time (less than one e -folding time), we plot the growth factors across the wave spectrum. Despite some slight deviation, our fiducial run with phase randomization nicely matches the theoretical expectation (except for numerical damping at large k). Without phase randomization, we see that while the fastest growth rate is not different from the theoretical

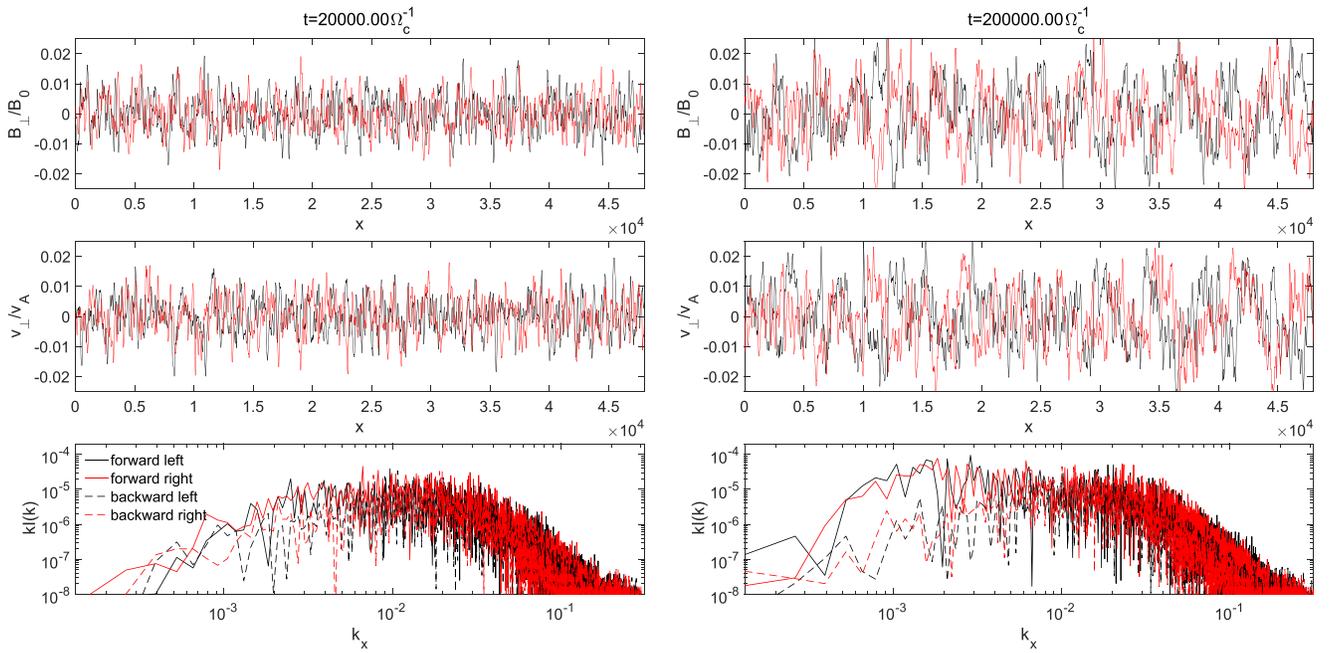


Figure 27. Wave patterns and their Fourier power spectra $[kI(k)]$ at two snapshots of the fiducial simulation using the full- f method: left for $t = 2 \times 10^4 \Omega_c^{-1}$ and right for $t = 2 \times 10^5 \Omega_c^{-1}$.

expectation by much, the wave spectrum has a significant offset, with the most unstable wavelength being about a factor of 3 shorter.

Further looking at the time history of the wave intensity of three representative modes (with the fastest-growing mode $k = (1/3) \times 10^{-2}$ shown in black) in the bottom panels, we see that with phase randomization, wave growth is clearly exponential. Without randomizing the phases, the exponential growth is only maintained for up to $t \sim 5000 \Omega_c^{-1}$, after which the wave amplitudes wildly vary.

We then conducted a series of simulations using fiducial parameters but in different box sizes and without phase randomization. The box sizes are $L_x = 2.88 \times 10^5$, 7.68×10^5 , and 1.92×10^6 . They are 3, 8, and 20 times the length of the fiducial box that we use (0.96×10^5), while they are run for a shorter amount of time (about $(2.5\text{--}4) \times 10^4 \Omega_c^{-1}$) just within the expected linear growth phase. Note that even for the longest box that we have attempted, the box crossing time for a relativistic particle is $1.92 \times 10^6 / 300 = 6.4 \times 10^3 \Omega_c^{-1}$, which is about a factor of $\sim 4\text{--}5$ smaller than the growth time of the CRSI.

In Figure 26, we again show the spectrum of growth factors for these runs. Clearly, the offset in the fastest-growing wavelength from the theoretical expectation gets smaller with increasing box size and eventually vanishes in the run with the longest box size. There is still some deviation in the absolute growth rate, though only by a small factor (the numerical growth rate is accurate within $\sim 10\%$). We thus consider this longest box size run to be marginally appropriate to study the CRSI.

Appendix D Comparison with the Full- f Method

To demonstrate the advantage of the δf method, we show in this appendix results obtained from the full- f approach. The simulation setup is identical to our run Fid, except that the δf weighting scheme is turned off and the box size is halved. Phase

randomization is also included. To further reduce Poisson noise, we use $N_p = 2048$ particles per bin per cell (in total, 16,384 per cell). According to the study in Appendix B, the level of noise is reduced by a factor of 8 in wave intensity, with wave energy density $\sim 10^{-4}$, or $\delta B^2/B_0^2 \lesssim 5 \times 10^{-5}$, in the form of Alfvén waves propagating in both directions. This level of noise would yield a rate of QLD, according to Equation (15), that is comparable to the maximum wave growth rate of $\sim 3 \times 10^{-5} \Omega_c$. The simulation is run to time $t = 2.25 \times 10^5 \Omega_c^{-1}$.

In Figure 27, we show the two snapshots of the wave pattern and spectrum. We see that the noise level indeed quickly reaches the level of $\delta B/B \lesssim 10^{-2}$, causing appreciable QLD of particles. Note that since Alfvén waves in both directions are present due to noise, CR energy would no longer be conserved in the (forward-propagating) wave frame, compromising the physics of the CRSI. As a result, we can only barely identify the signature of the CRSI in the wave spectrum at an early time of $t = 2 \times 10^4$. At a later time, forward-propagating waves are more easily seen to be amplified. However, these waves are at longer wavelengths (with $k \sim 10^{-3}$ instead of $\Omega_c/p_0 = 3.3 \times 10^{-3}$) where intrinsic numerical noise is smaller, allowing the CRSI to better stand out. The growth rates at these longer wavelengths are somewhat too noisy to measure with sufficient precision to compare with theoretical values, and through our analysis, the particles responsible for wave growth must have already undergone some QLD from pure noise.

We have further inspected the evolution of the distribution function, and we show in Figure 28 this distribution function for the last snapshot of the run. Note that because the binning is based on the full- f prescription (i.e., direct particle counting in each bin), it is much noisier than what we had before under the δf prescription. On the other hand, we can find signatures of isotropization. In fact, particles at $p \sim 3p_0$ are already nearly isotropized, and particles with $p \gtrsim 10p_0$ are still in the early stages of isotropization. For particles with $p \lesssim p_0$, accumulation near $\mu \sim 0$ is also clearly visible, as they are stuck and cannot cross the 90° pitch angle. While these features are qualitatively consistent with expectations, QLD proceeds much faster than that obtained

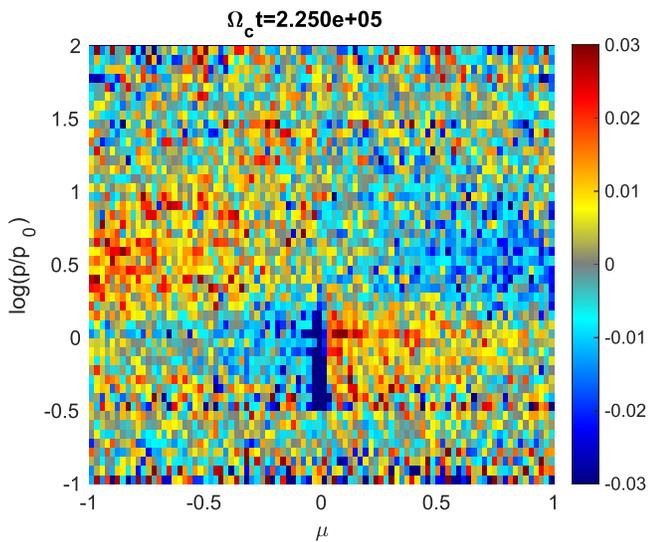


Figure 28. Distribution function $(f - f_0)/f_0$ at the last snapshot $t = 2.25 \times 10^5 \Omega_c^{-1}$ from the fiducial simulation using the full- f method.

using the δf method owing to substantial enhancement of scattering through noise. In particular, recall from discussions in Section 5.2 that isotropization is never achieved when using the δf method.

Overall, we conclude that while the full- f method can capture some basic properties of the CRSI, numerical noise adversely affects its evolution at all stages, leading to severe discrepancies at a quantitative (and sometimes even qualitative) level, which precludes making any accurate/reliable measurement. The situation can only get worse under more realistic conditions (low n_{CR}/n_i and small v_D).

ORCID iDs

Xue-Ning Bai <https://orcid.org/0000-0001-6906-9549>
 Eve C. Ostriker <https://orcid.org/0000-0002-0509-9113>
 Ilya Plotnikov <https://orcid.org/0000-0002-0074-4048>

References

- Achterberg, A. 1981, *A&A*, **98**, 161
 Amato, E., & Blasi, P. 2009, *MNRAS*, **392**, 1591
 Amato, E., & Blasi, P. 2018, *AdSpR*, **62**, 2731
 Baade, W., & Zwicky, F. 1934, *PNAS*, **20**, 259
 Bai, X.-N., Caprioli, D., Sironi, L., & Spitkovsky, A. 2015, *ApJ*, **809**, 55
 Bell, A. R. 1978, *MNRAS*, **182**, 147
 Bell, A. R. 2004, *MNRAS*, **353**, 550
 Birdsall, C. K., & Langdon, A. B. 2005, *Plasma Physics Via Computer Simulation* (London: Taylor and Francis)
 Blandford, R. D., & Ostriker, J. P. 1978, *ApJL*, **221**, L29
 Blasi, P., Amato, E., & Serpico, P. D. 2012, *PhRvL*, **109**, 061101
 Booth, C. M., Agertz, O., Kravtsov, A. V., & Gnedin, N. Y. 2013, *ApJL*, **777**, L16
 Boris, J. P. 1970, in Proc. Fourth Conf. on Numerical Simulation Plasmas, ed. I. Haber et al. (Washington, DC: Naval Research Laboratory), 3
 Breitschwerdt, D., McKenzie, J. F., & Voelk, H. J. 1991, *A&A*, **245**, 79
 Breitschwerdt, D., McKenzie, J. F., & Voelk, H. J. 1993, *A&A*, **269**, 54
 Denton, R. E., & Kotschenreuther, M. 1995, *JCoPh*, **119**, 283
 Dimits, A. M., & Lee, W. W. 1993, *JCoPh*, **107**, 309
 Draine, B. T. (ed.) 2011, *Physics of the Interstellar and Intergalactic Medium* (Princeton, NJ: Princeton Univ. Press)
 Drury, L. O. 1983, *RPPH*, **46**, 973
 Dupree, T. H. 1966, *PhFl*, **9**, 1773
 Enßlin, T., Pfrommer, C., Miniati, F., & Subramanian, K. 2011, *A&A*, **527**, A99
 Enßlin, T. A., Pfrommer, C., Springel, V., & Jubelgas, M. 2007, *A&A*, **473**, 41
 Everett, J. E., Zweibel, E. G., Benjamin, R. A., et al. 2008, *ApJ*, **674**, 258

- Farber, R., Ruszkowski, M., Yang, H.-Y. K., & Zweibel, E. G. 2018, *ApJ*, **856**, 112
 Farmer, A. J., & Goldreich, P. 2004, *ApJ*, **604**, 671
 Felice, G. M., & Kulsrud, R. M. 2001, *ApJ*, **553**, 198
 Ferrière, K. M. 2001, *RvMP*, **73**, 1031
 Fujita, Y., & Ohira, Y. 2011, *ApJ*, **738**, 182
 Gardiner, T. A., & Stone, J. M. 2005, *JCoPh*, **205**, 509
 Gardiner, T. A., & Stone, J. M. 2008, *JCoPh*, **227**, 4123
 Ginzburg, V. L., & Syrovatskii, S. I. 1964, *The Origin of Cosmic Rays* (New York: Macmillan)
 Girichidis, P., Naab, T., Hanasz, M., & Walch, S. 2018, *MNRAS*, **479**, 3042
 Girichidis, P., Naab, T., Walch, S., et al. 2016, *ApJL*, **816**, L19
 Goldreich, P., & Sridhar, S. 1995, *ApJ*, **438**, 763
 Grenier, I. A., Black, J. H., & Strong, A. W. 2015, *ARA&A*, **53**, 199
 Guo, F., & Oh, S. P. 2008, *MNRAS*, **384**, 251
 Hanasz, M., Lesch, H., Naab, T., et al. 2013, *ApJL*, **777**, L38
 Holcomb, C., & Spitkovsky, A. 2018, *ApJ*, submitted (arXiv:1811.01951)
 Holguin, F., Ruszkowski, M., Lazarian, A., Farber, R., & Yang, H.-Y. K. 2018, arXiv:1807.05494
 Hu, G., & Krommes, J. A. 1994, *PhPl*, **1**, 863
 Ipavich, F. M. 1975, *ApJ*, **196**, 107
 Jacob, S., Pakmor, R., Simpson, C. M., Springel, V., & Pfrommer, C. 2018, *MNRAS*, **475**, 570
 Jacob, S., & Pfrommer, C. 2017, *MNRAS*, **467**, 1449
 Jiang, Y.-F., & Oh, S. P. 2018, *ApJ*, **854**, 5
 Jokipii, J. R. 1966, *ApJ*, **146**, 480
 Jubelgas, M., Springel, V., Enßlin, T., & Pfrommer, C. 2008, *A&A*, **481**, 33
 Krymskii, G. F. 1977, *DoSSR*, **234**, 1306
 Kulsrud, R., & Pearce, W. P. 1969, *ApJ*, **156**, 445
 Kulsrud, R. M. 2005, *Plasma Physics for Astrophysics* (Princeton, NJ: Princeton Univ. Press)
 Kunz, M. W., Stone, J. M., & Bai, X.-N. 2014, *JCoPh*, **259**, 154
 Lazarian, A. 2016, *ApJ*, **833**, 131
 Lazarian, A., & Beresnyak, A. 2006, *MNRAS*, **373**, 1195
 Lebiga, O., Santos-Lima, R., & Yan, H. 2018, *MNRAS*, **476**, 2779
 Lee, M. A., & Völk, H. J. 1973, *Ap&SS*, **24**, 31
 Mao, S. A., & Ostriker, E. C. 2018, *ApJ*, **854**, 89
 Naab, T., & Ostriker, J. P. 2017, *ARA&A*, **55**, 59
 Pakmor, R., Pfrommer, C., Simpson, C. M., & Springel, V. 2016, *ApJL*, **824**, L30
 Parker, E. N. 1966, *ApJ*, **145**, 811
 Parker, E. N. 1992, *ApJ*, **401**, 137
 Parker, S. E., & Lee, W. W. 1993, *PhFIB*, **5**, 77
 Pfrommer, C., Pakmor, R., Schaal, K., Simpson, C. M., & Springel, V. 2017, *MNRAS*, **465**, 4500
 Recchia, S., Blasi, P., & Morlino, G. 2016, *MNRAS*, **462**, 4227
 Roe, P. L. 1981, *JCoPh*, **43**, 357
 Ruszkowski, M., Yang, H.-Y. K., & Zweibel, E. 2017, *ApJ*, **834**, 208
 Salem, M., & Bryan, G. L. 2014, *MNRAS*, **437**, 3312
 Salem, M., Bryan, G. L., & Hummel, C. 2014, *ApJL*, **797**, L18
 Schlickeiser, R. 1989, *ApJ*, **336**, 264
 Schlickeiser, R., & Miller, J. A. 1998, *ApJ*, **492**, 352
 Sharma, P., Colella, P., & Martin, D. F. 2010, *SIAM J. Sci. Comp.*, **32**, 3564
 Simpson, C. M., Pakmor, R., Marinacci, F., et al. 2016, *ApJL*, **827**, L29
 Skilling, J. 1975a, *MNRAS*, **172**, 557
 Skilling, J. 1975b, *MNRAS*, **173**, 245
 Skilling, J. 1975c, *MNRAS*, **173**, 255
 Soler, R., Terradas, J., Oliver, R., & Ballester, J. L. 2016, *A&A*, **592**, A28
 Stone, J. M., Gardiner, T. A., Teuben, P., Hawley, J. F., & Simon, J. B. 2008, *ApJS*, **178**, 137
 Summers, D., & Thorne, R. M. 1991, *PhFIB*, **3**, 1835
 Thomas, T., & Pfrommer, C. 2019, *MNRAS*, **485**, 2977
 Uhlig, M., Pfrommer, C., Sharma, M., et al. 2012, *MNRAS*, **423**, 2374
 Völk, H. J. 1973, *Ap&SS*, **25**, 471
 Wentzel, D. G. 1974, *ARA&A*, **12**, 71
 Wiener, J., Oh, S. P., & Guo, F. 2013, *MNRAS*, **434**, 2209
 Wiener, J., Pfrommer, C., & Oh, S. P. 2017, *MNRAS*, **467**, 906
 Wiener, J., & Zweibel, E. G. 2018, arXiv:1812.02179
 Wiener, J., Zweibel, E. G., & Oh, S. P. 2018, *MNRAS*, **473**, 3095
 Yan, H., & Lazarian, A. 2002, *PhRvL*, **89**, 281102
 Yan, H., & Lazarian, A. 2008, *ApJ*, **673**, 942
 Zirakashvili, V. N., Breitschwerdt, D., Ptuskin, V. S., & Voelk, H. J. 1996, *A&A*, **311**, 113
 Zweibel, E. G. 2013, *PhPI*, **20**, 055501
 Zweibel, E. G. 2017, *PhPI*, **24**, 055402