


# Gravitational Turbulence: the Small-Scale Limit of the Cold Dark Matter Power-Spectrum

Yonadav Barry Ginat <sup>1,2</sup> Michael L. Nastac,<sup>1,3</sup> Robert J. Ewart,<sup>1,4</sup> Sara Konrad,<sup>5</sup> Matthias Bartelmann,<sup>5</sup> and Alexander A. Schekochihin<sup>1,6</sup>

<sup>1</sup>*Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Parks Road, Oxford, OX1 3PU, United Kingdom*

<sup>2</sup>*New College, Holywell Street, Oxford, OX1 3BN, United Kingdom\**

<sup>3</sup>*St. John's College, St. Giles, Oxford, OX1 3JP, United Kingdom*

<sup>4</sup>*Balliol College, Broad Street, Oxford, OX1 3BJ, United Kingdom*

<sup>5</sup>*Institute for Theoretical Physics, Heidelberg University, Germany*

<sup>6</sup>*Merton College, Merton Street, Oxford, OX1 4JD, United Kingdom*

The matter power-spectrum,  $P(k)$ , is one of the fundamental quantities in the study of large-scale structure in cosmology. Here, we study its small-scale asymptotic limit, and show that for cold dark matter,  $P(k)$  has a universal  $k^{-d}$  asymptotic scaling with the wave-number  $k$ , for  $k \gg k_{\text{nl}}$ , in  $d$  spatial dimensions. We offer a theoretical explanation for this scaling, based on a non-perturbative analysis of the system's phase-space structure. Gravitational collapse is shown to drive a turbulent phase-space cascade of the quadratic Casimir invariant, where linear and non-linear time scales are balanced. A parallel is drawn to the phenomenon of Batchelor turbulence in hydrodynamics, where large scales mix smaller ones via tides. We also derive the  $k^{-d}$  scaling by expressing  $P(k)$  as a phase-space integral in the framework of kinetic field theory, which is analysed by the saddle-point method; the dominant critical points of this integral are precisely those where the time scales are balanced. The coldness of the dark-matter distribution function—its non-vanishing only on a 3D sub-manifold of phase-space, underpins both approaches. The theory is accompanied by 1D Vlasov-Poisson simulations, which confirm it.

## I. INTRODUCTION

One of the basic observable quantities in large-scale structure is the two-point correlation function of the overdensity field, whose Fourier transform is the power spectrum,  $P(k, t)$  [e.g., 1]. The two-point correlation function is of fundamental importance, for it allows us to probe theories of the early Universe, dark matter, inflation, and to study gravity [1–4]. In this paper, we will explore the small-scale asymptotic behaviour of  $P(k)$  (we omit the  $t$  argument when it is not confusing to do so), in the limit  $k \gg k_{\text{nl}}$ , where the (inverse) non-linear scale  $k_{\text{nl}}$  is defined by  $\int_{k_{\text{nl}}}^{\infty} P_{\text{lin}}(k) k^2 dk = 2\pi^2 \delta_c^2$  [5], with  $\delta_c$  denoting the spherical-collapse threshold. The small-scale limit of  $P(k)$  is theoretically important for the understanding of the gravitational  $N$ -body problem in the large- $N$  limit [6], and the formation of large-scale structure, non-linear clustering and self-similarity [1, 7], but also for the understanding of the nature of dark matter and gravitational back-reaction of small scales on large ones—both relativistic [8] and in the context of the effective field theory of large-scale structure [9] (for reviews and references see, e.g., [5, 10, 11]) and the general bias expansion [2]. Even before modifying gravity, it is important to know what non-linear phenomena occur in the standard theory. Data from cosmological dark-matter-only simulations are consistent with  $P(k, t)$  developing a  $k^{-d}$  tail at small scales [e.g., 12, fig. 6] in 3D, and in 1D, as shown

in refs. [13, figure 6] or [14, figure 1] (cf. [15]). The emergence of a power-law tail—and the simplicity of its exponent—hints that a fundamental physical reason for it must exist, ultimately stemming from the nature of the gravitational interaction of cold dark matter. Here, we will describe the mechanism that produces this asymptotic scaling, by studying the mass distribution in the velocity space as well as in position space.

We restrict ourselves to the strictly collisionless case where the particle mass  $m \rightarrow 0$ , while the total particle number  $N \rightarrow \infty$ , so that  $M \equiv Nm$  remains fixed (and so does the volume). In this limit the phase-space distribution of particles is well-described by the Vlasov equation [1, 13, 16–19]:<sup>1</sup>

$$\frac{\partial f}{\partial \eta} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} + \mathbf{g} \cdot \frac{\partial f}{\partial \mathbf{v}} = 0, \quad (1)$$

where  $f$  is the distribution function (the one-point probability density in phase-space),  $\eta$  is the conformal time, defined by  $dt = a d\eta$ , where  $a$  is the scale-factor of the background (which is taken to be a Friedmann-Lemaître-Robertson-Walker space-time),  $\mathbf{x}$  and  $\mathbf{v}$  are the co-moving position and velocity, and the gravitational field  $\mathbf{g}$  is self-consistently derived from Poisson's equation

<sup>1</sup> The collisionless Vlasov equation, of course, ignores dissipation via collisions (or equivalently, finite- $N$  effects). However, we will find that due to turbulence, such dissipation will inevitably be accessed; this point will be discussed further in §A 2.

\* yb.ginat@physics.ox.ac.uk

tion

$$\nabla \cdot \mathbf{g} \equiv -\nabla^2 \Phi = -\frac{4\pi G}{a^{d-2}} \left[ \int f(\mathbf{x}, \mathbf{v}, \eta) d^d v - a^d \bar{\rho}_m \right], \quad (2)$$

where  $\bar{\rho}_m$  is the background matter density and  $\Phi$  is the gravitational potential. This system of equations applies well on scales much smaller than the horizon, for particles much slower than  $c$ . Henceforth, we will work in general spatial dimensions  $d \in \{1, 2, 3\}$ . Dark matter is assumed to be cold initially, with a thermal velocity  $v_{\text{th}} \rightarrow 0$ , such that the initial gravitational potential energy is much larger than the initial thermal energy (this assumption is excellent for our Universe—see, e.g., [3, 20–25]). Clearly, despite being cold, dark matter is inherently kinetic, and cannot be described by merely fluid equations adequately on non-linear scales, because these equations cease to be valid after streams cross [e.g. 19, 26, 27].

Below we will derive the  $k \gg k_{\text{nl}}$  limit of  $P(k)$  in two complementary ways. First, we will study the problem in phase-space, and show that the small-scale asymptotics arise when the time scales involved in the Vlasov equation balance with each other in a particular way, to be explained below. Two ingredients will comprise this argument: the balance of time scales, and the conservation of the second Casimir invariant [28–32]

$$C_2 \equiv \frac{1}{V} \iint f^2 d^d x d^d v, \quad (3)$$

where  $V$  is the spatial volume. This invariant is sometimes referred to as ‘enstrophy’, ‘phasesstrophy’, or ‘ $f$ -strophy’. The Vlasov equation conserves an infinite number of phase-space invariants—not only  $C_2$ —but we will use  $C_2$  here because it is directly related to the power-spectrum (*vide infra*). We will show that the spectrum can be predicted if small-scale structure is understood as resulting from a turbulent cascade of  $C_2$  from large scales to small ones.

For the second approach, we use non-perturbative kinetic field theory (KFT) [for a review, see 33] for cosmic structure formation in dark matter. Here,  $P(k)$  is expressed as an integral over the initial particle positions and velocities (weighted by the initial-condition probability distribution) of the characteristic function of the displacement field; this integral will have an explicit  $k$ -dependence, and we will utilise this to perform an asymptotic saddle-point analysis.

The two approaches complement each other, both relying on the same assumptions, but highlighting their rôles in different ways. We remark that using phase-space expressions ensures that the validity of our results extends beyond the bounds of configuration-space-based approaches, such as Lagrangian [15, 34, 35] or Eulerian techniques [26]; in particular, it is regular at stream-crossing, and accounts for free streaming automatically. Indeed, as we already mentioned above, the inherently kinetic phenomenon of multi-scale structure of the distribution of dark matter, developing by virtue of a strongly

non-linear interaction, suggests that a type of turbulence in phase-space is involved. In turbulent phase-space dynamics, there is a flux of  $C_2$  from large scales to small ones [31], and we will show below that  $C_2$  cascades to smaller scales by gravitational collapse here, too. As early as [35] it was realised that the phenomena of gravitation and turbulence might be linked—here we make the analogy precise and characterise this gravitational turbulence.

The rest of this paper is organised as follows: in §II we formulate the first approach, based on the Vlasov-Poisson system, and show how the concept of a ‘phase-space cascade’ can be used to derive the small-scale asymptotics of the phase-space power-spectrum, from which  $P(k)$  may be computed. In §III we derive the same asymptotic scaling of  $P(k)$  again, but via a saddle-point analysis of an integral expression for  $P(k)$ . We test our theory by comparing it with numerical Vlasov-Poisson simulations throughout the paper. The main conclusions are discussed in §IV and summarised in §V.

## II. PHASE-SPACE TURBULENCE

We start by describing the initial condition—a single, cold stream—and show what its evolution looks like, and then derive the equation that governs the phase-space Fourier transform of  $f$  in §II B. We will then analyse the time scales involved in the Vlasov-Poisson system in §II C, and show in §II D that a turbulent flux of  $C_2$  to smaller scales characterises the dynamics. To do that, we will derive a transport equation for the integrand of  $C_2$  with a source term; in §II E the latter will be found to receive contributions from all larger scales, because of the Jeans instability [36]. This in turn will allow us to find the asymptotic scaling of  $P(k)$  in §II F. We will present simulation results throughout this section, to test the theory.

### A. Cold streams

We assume that the system has an initial condition consisting of a super-position of streams, each one of the form:

$$f(\eta = 0, \mathbf{x}, \mathbf{v}) = f_{\text{in}}(\mathbf{x}, \mathbf{v}) \equiv \frac{\rho_{\text{in}}(\mathbf{x})}{(2\pi v_{\text{th}}^2)^{d/2}} e^{-\frac{[\mathbf{v} - \mathbf{u}_{\text{in}}(\mathbf{x})]^2}{2v_{\text{th}}^2}}, \quad (4)$$

where  $v_{\text{th}}^2 \ll \min \{u_{\text{in}}^2, \int |\Phi_{\text{in}}| f_{\text{in}} d^d x d^d v\}$ , where  $\Phi_{\text{in}}$  is derived from  $\rho_{\text{in}}$ ; that is, the initial condition is very cold—there is little thermal energy and it is negligible in comparison with the gravitational potential energy of the system or the kinetic energy of mean flows.  $\rho_{\text{in}}$  and  $\mathbf{u}_{\text{in}}$  are typically Gaussian random fields [1, 2]. We also choose the co-ordinates so that  $\int \rho(\mathbf{x}) \mathbf{u}(\mathbf{x}) d^d x = 0$ , where  $\rho(\mathbf{x})$  is the density and  $\rho \mathbf{u} \equiv \int \mathbf{v} f d^d v$ . The initial distribution  $f_{\text{in}}$  is essentially a single stream, and,

in fact, its Maxwellian shape may be approximated by a Dirac delta-function. The thermal speed  $v_{\text{th}}$  is chosen to be the smallest scale in the problem, so that the entire analysis of this paper focuses on the limit where  $sv_{\text{th}} \ll 1$ , where  $s$  is the Fourier conjugate to velocity (see §II B). The distribution function remains a collection of streams as it evolves in time, by Liouville’s theorem: locally almost everywhere in phase-space (in a sufficiently small phase-space neighbourhood of almost any point  $(\mathbf{x}_0, \mathbf{v}_0)$  where  $f \neq 0$ ), it can be written as  $\rho(\eta, \mathbf{x})\delta^{\text{D}}(\mathbf{v} - \mathbf{u}(\eta, \mathbf{x})) + O(v_{\text{th}})$ , where  $\delta^{\text{D}}$  is the Dirac delta function; it remains so as long as collisions or finite- $N$  effects may be ignored.

To study the evolution of the initial condition (4), we have conducted a suite of Vlasov-Poisson simulations in 1D on a Minkowski background ( $a(\eta) = 1$ ), using the Gkeyll code [37], which is originally a Vlasov solver for kinetic plasmas; by setting the vacuum permittivity to  $\varepsilon_0 < 0$  we can use the code to study gravity as opposed to electrostatics. Details on the method and simulation setup may be found in appendix A. The simulation’s units are chosen such that  $\tau_0^{-2} \equiv 4\pi GM/L = 1$  (unit of time),<sup>2</sup>  $k_0 \equiv 2\pi/L = 1$  (unit of length) and  $v_0 \equiv (k_0\tau_0)^{-1} = 1$  (unit of speed), where  $L$  is the length of the simulation ‘box’ (with periodic boundary conditions) and  $M$  is the total mass in the box.

The time-evolution of a cold system of three streams—three copies of equation (4)—is displayed in figure 1. This figure shows that each stream collapses quickly, by rotating and twisting in phase-space. Evidently, this motion generates small-scale structure, which we consider to be a type of turbulence in phase-space. As usual, in order to characterise the turbulence and its spectrum of fluctuations, one must identify which invariant quantity cascades to small scales, what the cascade’s time scale is, and also the flux of that invariant quantity [38]. We will do so in §§II D, II C, and II E respectively.

That the turbulence is in phase-space and not merely in position space is clear: already at  $t = 4\tau_0$ , the system can no longer be described as a single stream almost everywhere, so standard cosmological perturbation theory would already be inadequate, and the second velocity cumulant  $v_{\text{rms}} \equiv \left[ \int dv (v - u(x))^2 f(x, v)/\rho(x) \right]^{1/2}$  is much larger than  $v_{\text{th}}$ , and is of order  $v_0$ . However, in phase-space, we see that the topology of a three single lines is preserved—as one expects from Liouville’s theorem [e.g., 39]. The large- $s$  limit consists of  $1 \ll sv_{\text{rms}}$ .

The fact that the system is a collection of streams implies that *locally in phase-space*, one can describe each stream with fluid equations: inserting  $\rho(\eta, \mathbf{x})\delta^{\text{D}}(\mathbf{v} - \mathbf{u}(\eta, \mathbf{x}))$  into the Vlasov equation gives continuity and Euler equations, and the divergence of the latter yields the Raychaudhuri equation [40] which describes gravitational collapse under the Jeans instability [19, 34, 36, 41–47]. Written in the frame of reference

that moves with a given stream, it reads

$$\frac{d\theta}{d\eta} + \mathcal{H}\theta + \frac{\theta^2}{3} + \sigma^{ij}\sigma_{ij} = \nabla \cdot \mathbf{g} + \frac{\Sigma}{\rho} + 2\omega_i\omega^i, \quad (5)$$

where  $\mathcal{H}$  is the conformal Hubble constant,  $\theta \equiv \partial_i u^i$  is the stream’s divergence,  $\boldsymbol{\omega} \equiv \nabla \times \mathbf{u}/2$  is its vorticity,  $\sigma_{ij} \equiv (\partial_i u_j + \partial_j u_i - 2\theta\delta_{ij}/3)/2$  is its shear,  $d/d\eta$  is the Lagrangian derivative along the stream,  $\Sigma/\rho$  is a pressure-related term that is  $\propto v_{\text{th}}^2$ , and  $\mathbf{g}$  is the total gravitational field felt by the stream. We will use equation (5) below to gauge the time scale of gravitational collapse,  $\tau_c$ , which is the time scale for the decrease of  $\theta$ .

## B. Batchelor approximation

We need to characterise a turbulence, which is inherently a multi-scale process, so it is more convenient to study equations (1)-(2) in Fourier space.

### 1. Fourier transform

Let us define the Fourier transform by (marked by a circumflex)

$$\hat{f}(\mathbf{k}, \mathbf{s}) \equiv \iint f(\mathbf{x}, \mathbf{v}) e^{i\mathbf{k}\cdot\mathbf{x} - i\mathbf{s}\cdot\mathbf{v}} d^d x d^d v, \quad (6)$$

and similarly for other functions of  $(\mathbf{x}, \mathbf{v})$ . For  $d > 1$ , we denote  $k \equiv |\mathbf{k}|$ ,  $s \equiv |\mathbf{s}|$ , etc. Under this Fourier transform the Vlasov-Poisson system becomes

$$\frac{\partial \hat{f}}{\partial \eta} + \mathbf{k} \cdot \frac{\partial \hat{f}}{\partial \mathbf{s}} + i\mathbf{s} \cdot \int \frac{d^d k'}{(2\pi)^d} \hat{\mathbf{g}}(\mathbf{k}') \hat{f}(\mathbf{k} - \mathbf{k}', \mathbf{s}) = 0, \quad (7)$$

$$ak^2 \hat{\Phi} = -4\pi G \hat{\rho}, \quad (8)$$

where  $\hat{\rho}$  is the Fourier-transformed density. The second Casimir invariant, defined by equation (3), is given in Fourier space by Parseval’s theorem:

$$C_2 = \frac{1}{(2\pi)^{2d} V} \iint |\hat{f}|^2 d^d k d^d s. \quad (9)$$

The integrand  $|\hat{f}|^2$  is directly related to the power-spectrum: let the *phase-space power spectrum* be

$$\hat{F}(\mathbf{k}, \mathbf{s}) \equiv \langle |\hat{f}|^2 \rangle, \quad (10)$$

where  $\langle \cdot \rangle$  is an ensemble average over many random realisations of the initial conditions. The density power spectrum is then  $P(k) = \hat{F}(\mathbf{k}, 0)$ .

### 2. Flux of $C_2$

Let us now see how the integrand of the second Casimir invariant in equation (9) evolves: multiplying equation (7) by  $\hat{f}^*$  and taking the real part, we find

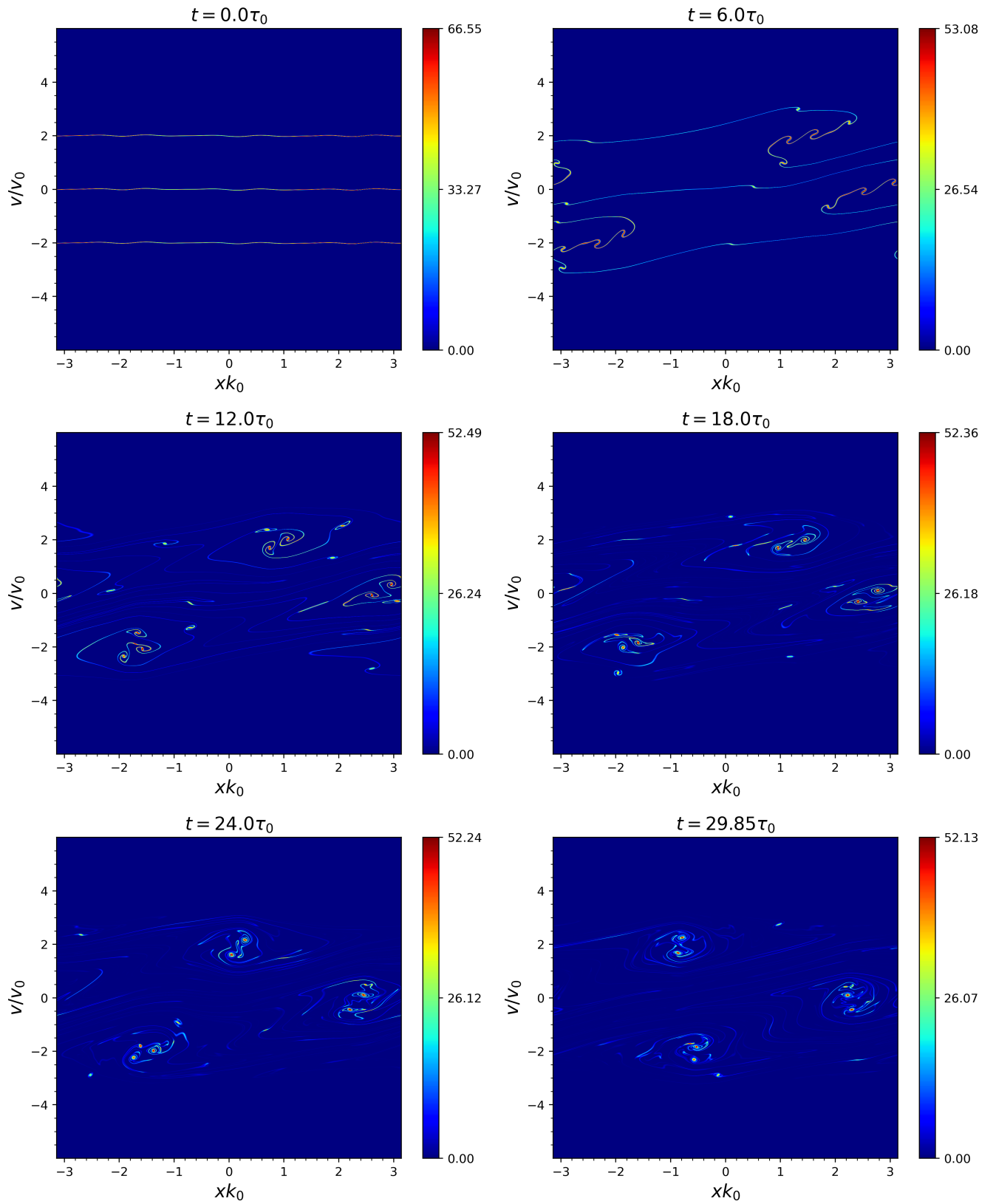


FIG. 1. Colour plots of the distribution function showing the time evolution of three cold streams. See text and appendix A for details.

$$\frac{\partial|\hat{f}|^2}{\partial\eta} + \mathbf{k} \cdot \frac{\partial|\hat{f}|^2}{\partial\mathbf{s}} + \text{is} \cdot \int \frac{d^d k'}{(2\pi)^d} \left[ \hat{\mathbf{g}}(\mathbf{k}') \hat{f}^*(\mathbf{k}, \mathbf{s}) \hat{f}(\mathbf{k} - \mathbf{k}', \mathbf{s}) - \hat{\mathbf{g}}^*(\mathbf{k}') \hat{f}(\mathbf{k}, \mathbf{s}) \hat{f}^*(\mathbf{k} - \mathbf{k}', \mathbf{s}) \right] = 0. \quad (11)$$

To proceed, we need to assume that  $\mathbf{g}$  is a smooth field: let  $\delta g_r$  be the amplitude of a typical fluctuation in  $\mathbf{g}$  on scales  $r = 1/k$  and smaller, *viz.*

$$\delta g_r^2 \equiv \int_{k'r > 1} \frac{d^d k'}{(2\pi)^d} |\hat{\mathbf{g}}(\mathbf{k}')|^2. \quad (12)$$

Integrating equation (11) over all  $\mathbf{k}$  and  $\mathbf{s}$  yields the conservation of  $C_2$ , as it should. We take  $\mathbf{g}$  to be a sufficiently continuous field, so that for a sufficiently large  $k$  (small  $r$ ),  $\delta g_r \sim \kappa r^\lambda + \text{h.o.t.}$ , with Hölder exponent  $\lambda \leq 1$ , and  $\kappa$  a coefficient. A smooth gravitational field has  $\lambda = 1$ , because one may Taylor-expand  $\mathbf{g}$ , meaning that the fluctuations are dominated by tidal forces (and  $\kappa$  has dimensions of  $[\text{time}]^{-2}$ ).<sup>3</sup> We show in appendix B that if one were to assume  $\lambda < 1$  in  $d \leq 3$  spatial dimensions and repeat the analysis of §II, one would find a power spectrum inconsistent with equation (2) (cf. [32]), whence  $\lambda = 1$ .

Let us now take the  $k \rightarrow \infty$  limit of equation (11). Then, for a smooth gravitational field ( $\lambda = 1$ ), the last term on the left-hand side is dominated by two contributions: (i)  $k' \ll k$  and (ii)  $|\mathbf{k} - \mathbf{k}'| \ll k$ . These are sometimes referred to as ‘squeezed’ triangles, while case (i) is also known as the Batchelor limit [31, 48, 49]. Taylor-expanding  $\hat{f}(\mathbf{k} - \mathbf{k}', \mathbf{s})$  in the Batchelor limit turns the non-linear term in (11) into

$$\begin{aligned} & \text{is} \cdot \int \frac{d^d k'}{(2\pi)^d} \left[ \hat{\mathbf{g}}(\mathbf{k}') \hat{f}^*(\mathbf{k}, \mathbf{s}) \left( \hat{f}(\mathbf{k}, \mathbf{s}) - \frac{\partial \hat{f}}{\partial k^i} q^i \right) - \hat{\mathbf{g}}^*(\mathbf{k}') \hat{f}(\mathbf{k}, \mathbf{s}) \left( \hat{f}^*(\mathbf{k}, \mathbf{s}) - \frac{\partial \hat{f}^*}{\partial k^i} k'^i \right) \right] + \text{h.o.t.} \\ & \simeq -\text{i} \int \frac{d^d k'}{(2\pi)^d} s^j \left[ \hat{g}_j(\mathbf{k}') k'^i \hat{f}^*(\mathbf{k}, \mathbf{s}) \frac{\partial \hat{f}}{\partial k^i} - \hat{g}_j^*(\mathbf{k}') k'^i \hat{f}(\mathbf{k}, \mathbf{s}) \frac{\partial \hat{f}^*}{\partial k^i} \right] \\ & = -\text{i} \int \frac{d^d k'}{(2\pi)^d} s^j \hat{g}_j(\mathbf{k}') k'^i \left[ \hat{f}^*(\mathbf{k}, \mathbf{s}) \frac{\partial \hat{f}}{\partial k^i} + \hat{f}(\mathbf{k}, \mathbf{s}) \frac{\partial \hat{f}^*}{\partial k^i} \right] \equiv -\text{i} s^j \Phi_j^i \frac{\partial |\hat{f}|^2}{\partial k^i}. \end{aligned} \quad (13)$$

The transition from the first line to the second in (13) is correct because the leading-order terms are proportional to  $\int d^d k' \hat{\mathbf{g}}(\mathbf{k}') = 0$ , in the centre-of-mass frame.

Thus, the non-linear term reduces to a tidal interaction, and it encapsulates the idea that on a given small scale, corresponding to wave-number  $k$ , the dominant effect of the gravitational field is nothing but the tidal forces: the distribution function  $f$  at small scales (large  $k$ ) is distorted by the gravitational (tidal) field at the same (large) energy-containing scale—the matrix  $\Phi_j^i = -\text{i} \delta^{in} \partial_n \partial_j \Phi$  is the Hessian matrix of the gravitational potential, i.e. the tidal matrix.

### 3. Ensemble average

We now have

$$\frac{\partial|\hat{f}|^2}{\partial\eta} + \mathbf{k} \cdot \frac{\partial|\hat{f}|^2}{\partial\mathbf{s}} - \text{i} s^j \Phi_j^i \frac{\partial|\hat{f}|^2}{\partial k^i} = -(\text{ii}), \quad (14)$$

where (ii) represents the last term of equation (11) in the limit (ii), i.e.  $|\mathbf{k} - \mathbf{k}'| \ll k$ . This equation is a transport equation in Fourier space, with a source  $-(\text{ii})$ . Taking the average of equation (14) yields an evolution equation for the power spectrum  $\hat{F}$ :

$$\frac{\partial \hat{F}}{\partial \eta} + \mathbf{k} \cdot \frac{\partial \hat{F}}{\partial \mathbf{s}} - \text{i} s^j \frac{\partial}{\partial k^i} \langle \Phi_j^i |\hat{f}|^2 \rangle = \hat{S}, \quad (15)$$

where

$$\begin{aligned} \hat{S} &= \iint d^d x_1 d^d v_1 d^d x_2 d^d v_2 e^{\text{i}\mathbf{k} \cdot \mathbf{x} - \text{i}\mathbf{s} \cdot \mathbf{v}} \\ &\times \left\langle f(\mathbf{x}_2, \mathbf{v}_2) [\mathbf{g}(\mathbf{x}_1) - \mathbf{g}(\mathbf{x}_2)] \cdot \frac{\partial \bar{f}(\mathbf{v}_1)}{\partial \mathbf{v}} \right\rangle, \end{aligned} \quad (16)$$

where we have shortened  $\bar{f} \equiv \int f d^d x / V$ , and  $\mathbf{x} = \mathbf{x}_1 - \mathbf{x}_2$ ,  $\mathbf{v} = \mathbf{v}_1 - \mathbf{v}_2$ . We will estimate  $\hat{S}$  in §II E below.

The three-point correlation function  $\langle \Phi_j^i |\hat{f}|^2 \rangle$  is composed of  $\Phi_j^i$ , which is by construction a large-scale quantity, multiplied by  $|\hat{f}|^2$ , which depends on  $\mathbf{k}$ . We contend that, as  $\Phi_j^i$  varies only on large scales, it will vary much less from one ensemble realisation to another than  $|\hat{f}|^2$ . Furthermore, the non-linear term in the Vlasov equation affects the evolution more in over-dense regions, for

<sup>2</sup> Assuming the Poisson equation (2) implies that  $G$  has different dimensions in 1D from 3D.

<sup>3</sup> The case  $\lambda > 1$  is also smooth, but highly atypical, where the tidal forces vanish; while this could happen in isolated points with exactly zero over-density, we ignore it here.

it remains small in under-dense ones; the realisations where  $\mathbf{x} = 0$  is a Jeans-unstable over-dense region are therefore those that dominate this ensemble average, so  $\text{itr} \Phi_j^i > 0$ , and generically (in  $d > 1$ )  $i\Phi_j^i$  has both positive and negative eigenvalues.<sup>4</sup>

The time scales appearing on the left-hand side of equation (15) are the advection (linear) time scale associated with  $\mathbf{k} \cdot \frac{\partial}{\partial \mathbf{s}}$ , and the non-linear (gravitational) time scale associated with  $s^j \Phi_j^i \frac{\partial}{\partial k^i} \sim s \delta g_r$ . We will discuss these time scales presently in §II C. Then, §II D describes the left-hand side of (15), while §II E describes the right-hand side.

### C. Critical balance

There are two (conformal) time scales in the Vlasov equation (1): the linear time scale,  $\tau_p \sim r/\delta v$ , and the non-linear time scale  $\tau_g \sim \delta v/\delta g_r$ , where  $\delta g_r$  is defined in equation (12) and  $\delta v$  is a velocity-difference scale. In the Fourier-transformed Vlasov equation (7) the linear (or phase-mixing) time scale is

$$\tau_p \equiv \frac{s}{k} \quad (17)$$

whereas the non-linear (gravitational) time scale is

$$\tau_g \equiv \frac{1}{s \delta g_r}, \quad (18)$$

with  $r = 1/k$ . In the Batchelor approximation (14), the non-linear time scale is  $\tau_g^{-1} = s \|\Phi_j^i\|/k$ , which is the same as (18) (up to an order-unity constant).

In the highly non-linear régime, these two time scales must balance each other: if  $\tau_p$  were much shorter than  $\tau_g$ , so that only the first two terms of equation (7) dominated, phase mixing would drive velocity gradients up until  $\tau_g$  shrank to the same order of magnitude. The converse is also true: if  $\tau_g$  were much smaller, then gravitational collapse would drive spatial gradients of  $f$  up, until  $\tau_p$  &  $\tau_g$  matched. This so-called *critical balance* may be thought of as a type of dominant-balance asymptotic argument for the Vlasov equation, where one allows the system enough time to establish this balance (see for instance [31, 32, 50–52] for examples of critical balance in various areas of physics).

Critical balance (setting  $\tau_p \sim \tau_g$  with  $\delta g_r \sim \kappa r$  as expected for a smooth field) implies that for a given length-scale  $r = 1/k$ , there exists a corresponding (inverse) velocity-scale

$$s_c(k) \equiv \frac{k}{\sqrt{\kappa}}. \quad (19)$$

Likewise, for a given  $s$  we define

$$k_c(s) \equiv \sqrt{\kappa} s. \quad (20)$$

Additionally, if the two time scales are equal, their shared value defines a so-called *critical-balance time*.

The argument that the two time scales must balance applies only to modes well inside the horizon, such that  $k \gg \mathcal{H}/c$  (where  $\mathcal{H}$  is the conformal Hubble constant): the conformal time in any asymptotically de-Sitter cosmology (i.e. one with a positive cosmological constant) is bounded from above by some value  $\eta_{\max}$  [53]. Therefore, for modes with  $k$  too small, linear phase mixing can only generate velocity gradients up to  $s_{\max} \simeq k\eta_{\max}$ . If the amplitude of  $\hat{f}(k, s_{\max})$  is not large enough for the non-linear term to become important, then the evolution of such a mode will always be primarily linear. Here we are interested in the  $k \rightarrow \infty$  limit, so it is safe to ignore this nuance. Additionally, due to the finite age of the Universe and hierarchical structure formation, the decrease of  $\tau_g$  until it matches  $\tau_p$  might not have happened yet for all values of  $k$ , as structures on the largest scales have yet to collapse. Again, this does not affect the  $k \rightarrow \infty$  limit, and we may simply take  $k \gg \max\{k_{\text{nl}}, \mathcal{H}/c\}$ . In a sense, from the point of view of the turbulence discussed below,  $\max\{k_{\text{nl}}, \mathcal{H}/c\}$  serves as the ‘outer scale’ of the system.

### D. Phase-space cascade

Let us now see what kind of flow in  $(\mathbf{k}, \mathbf{s})$  space is engendered by equation (15), ignoring  $\hat{S}$  until §II E: the positive eigenvalues of  $i\Phi_j^i$  will drive a rotation in  $(\mathbf{k}, \mathbf{s})$ , where small-scale velocity structure interchanges with small-scale spatial structure, while the positive eigenvalues will drive a flow of both to ever smaller scales. In fact, one can analyse equation (14) directly—before ensemble-averaging—while still ignoring the right-hand-side. As  $\Phi_j^i$  generically depends on time, this analysis is true locally in time (and space, on scales  $\lesssim k^{-1}$ ), but by critical balance, the long mode  $\Phi_j^i$  cannot vary on a time-scale shorter than the critical-balance time. We therefore approximate it as constant (in which case there are analytical solutions), but the qualitative features described here—namely, a rotation in  $(\mathbf{k}, \mathbf{s})$  and a flow to larger values—remain true for a time-dependent  $\Phi_j^i$ .

Consider a positive eigenvalue of  $\Phi_j^i$ . If  $\mathbf{s}_+$  is its corresponding eigenvector, then for  $\mathbf{s} \parallel \mathbf{s}_+$ , equation (14) simply reduces to a transport equation under the action of a harmonic oscillator potential, i.e. a rotation in the  $(\mathbf{k}, \mathbf{s})$  space. This ensures that the large- $s$  structure in the initial condition is transported to large  $k$ , and *vice versa*. The negative eigenvalues ensure that there is a flow to ever smaller scales, because then the solution is a linear combination of hyperbolic functions (cf. [32]).

In 1D, it would appear naïvely that there is only phase-space rotation when  $\mathbf{x} = 0$  is over-dense, because there

<sup>4</sup> The existence of a positive eigenvalue is guaranteed by Poisson’s equation.

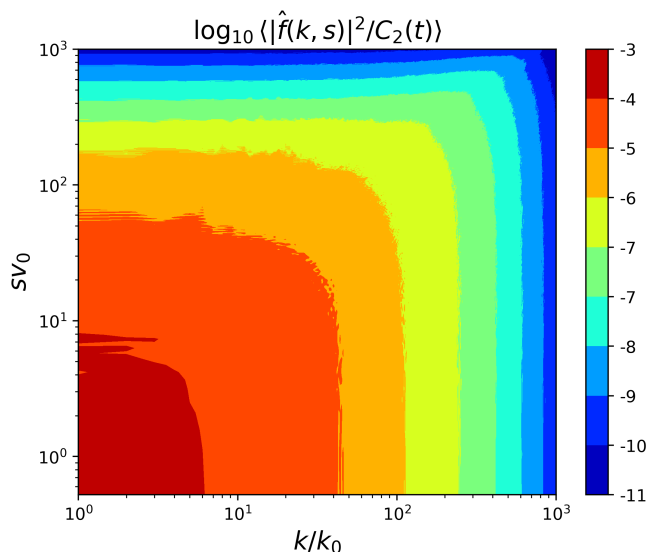


FIG. 2. A contour plot of the time-averaged power-spectrum  $\langle |\hat{f}|^2(t, k, s)/C_2(t) \rangle$ , for the same simulation as in figure 1 (note that  $C_2$  decays because of collisions and finite-grid effects, the power spectrum has to be normalised to it at every time).

is only one, necessarily positive, eigenvalue. This, however, is a singular case. As the system evolves matter moves around, and  $\mathbf{x} = 0$  generically changes from being over-dense to under-dense (alternatively, it does so for different realisations of the initial conditions), and therefore the sign of  $\Phi_j^i$  also changes. Thus, there is a temporal sequence of phase-space rotations and stretching—essentially, differential phase-space rotation, leading to the generation of small-scale structure (cf. appendix C 2).

Consequently, phase-space rotation transports small-scale structure from  $s$  to  $k$  and back, and that is supplemented by an order-unity smaller (since the positive eigenvalue is necessarily the largest) cascade of power to increasingly larger  $k$  and  $s$  (similar to the plasma echo joint with a cascade in [31]).<sup>5</sup> Hence, equation (14) describes a phase-space turbulence, where  $C_2$  cascades to smaller scales by larger-scale tidal fields, along the critical-balance line.

To test this conclusion with the simulations, we calculated time-averaged  $\hat{F}$ , as a function of  $|k|$ ,  $|s|$  (with the negative values folded on the positive ones), for the simulation shown in figure 1. This is plotted in figure 2, which shows a contours time-average (as a proxy for an ensemble-average)  $\langle |\hat{f}|^2/C_2(t) \rangle_{\text{Time}} \sim \hat{F}$ . They are arranged in a rectangular shape: the critical-balance line

(§II C) is nothing but the diagonal of this rectangle, and there is a slight excess along the line, due to the flow along it. This rectangle is brought about by the aforementioned rotation in the  $(k, s)$  space, conjoined with the effect of the positive eigenvalues of  $\Phi_j^i$ , which stretch structures along the critical-balance line.

### E. Source-term behaviour for a cold system

Having described the homogeneous version of equation (14), let us now include the source term.

#### 1. Phase-space flux

Equation (15) is a conservation equation with a source, of the form

$$\frac{\partial \hat{F}}{\partial \eta} + \nabla_{\mathbf{k}, \mathbf{s}} \cdot \mathbf{\Gamma} = \hat{S}. \quad (21)$$

Here  $\nabla_{\mathbf{k}, \mathbf{s}}$  is a 6-dimensional phase-space gradient, and  $\mathbf{\Gamma}$  is a 6-dimensional flux, whose components are

$$\Gamma_i^{\mathbf{s}} = k_i \hat{F} \quad (22)$$

$$\Gamma_i^{\mathbf{k}} = -is^j \langle \Phi_{ij} |\hat{f}|^2 \rangle. \quad (23)$$

As described in §II D, there is a flow of  $C_2$  to larger values of  $k$  and  $s$ . Let us calculate the phase-space flux,  $\mathcal{F}^{\mathbf{s}}$ , flowing through a sphere in  $(\mathbf{k}, \mathbf{s})$ -space of radius  $s$  (or  $k_c(s)$ , where  $s_c(k_{\text{nl}}) \ll s \ll v_{\text{th}}^{-1}$ ). In steady state, integrating equation (15) over  $\mathbf{k}$  yields

$$\frac{\partial}{\partial \mathbf{s}} \cdot \int \mathbf{\Gamma}^{\mathbf{s}} d^d \mathbf{k} = \int \hat{S} d^d \mathbf{k}. \quad (24)$$

Integrating both sides over the ball in  $\mathbf{s}$ -space with radius  $s$  and using Gauss' theorem yields

$$\mathcal{F}^{\mathbf{s}} \equiv \iint \hat{s}^i \Gamma_i^{\mathbf{s}} d^d \mathbf{k} d^{d-1} \hat{s} = \iint_{|\mathbf{k}| < k_c(s), |\mathbf{s}'| < s} \hat{S} d^d \mathbf{k} d^d \mathbf{s}', \quad (25)$$

The volume of the ball is  $\sim s^d$ , so if  $\int d^d \mathbf{k} \hat{S} \sim \text{const}$ , then the integral would scale like  $s^d$ , if the sign of the constant is independent of  $s$ ; this is indeed the case, as we will show now, by virtue of the gravitational-collapse instability. If we instead wish to estimate the flux in  $\mathbf{k}$ ,  $\mathcal{F}^{\mathbf{k}}$ , the same happens, except that now  $\mathcal{F}^{\mathbf{k}}$  will scale like  $s_c(k)^d \propto k^d$ , by critical balance.

#### 2. Jeans instability of a stream

We are now finally in a position to estimate the  $(\mathbf{k}, \mathbf{s})$  dependence of  $\hat{S}$ . We will do that for the cold case described in §II A above; the coldness of the distribution—its being a collection of streams—implies that each

<sup>5</sup> This behaviour is generic: chaos—the exponential separation of nearby trajectories in phase-space—combined with Liouville's theorem, necessitates the formation of structure on increasingly smaller scales.

stream, in its own reference frame, is Jeans unstable. This in turn implies that in the integral on the right-hand side of equation (25), all scales add up coherently, as we will show below.

As remarked in §II A above, our analysis focuses on the limit where  $sv_{\text{th}} \ll 1 \ll sv_{\text{rms}}$ , or, equivalently,  $k_{\text{nl}} \ll k \ll k_c(v_{\text{th}}^{-1})$ , where  $k_c(s)$  is defined in equation (20). On these scales,  $f$  is just a collection of streams, so when zooming in on one of them, and applying the Raychaudhuri equation (5) in its frame of reference, one finds that the stream will continue to collapse on all scales until  $k$  is large enough so that the collapse time  $\tau_c$  matches  $(kv_{\text{th}})^{-1}$ . This is because all the terms in equation (5) are of dimension  $[\text{time}]^{-2}$ , with  $\nabla \cdot \mathbf{g} \sim \tau_g^{-2}$ , and both last terms on the right-hand-side are  $O(k^2 v_{\text{th}}^2)$ ,<sup>6</sup> while the  $\mathcal{H}\theta$  term is completely negligible, for  $k \gg \mathcal{H}/c$ . Equation (5) then reduces to

$$\frac{d\theta}{dt} = -\frac{\theta^2}{3} + \nabla \cdot \mathbf{g} + O(k^2 v_{\text{th}}^2), \quad (26)$$

which means that for  $s_c(k)v_{\text{th}} \ll 1$ , the gravity term  $\nabla \cdot \mathbf{g}$  dominates and there is nothing to stop a collapse, on a time scale  $\tau_g$ , so  $\tau_c = \tau_g$ . The Jeans instability occurs on every stream individually, in its own reference frame; this is not dissimilar to the instabilities described in, e.g., [41, 44]. Indeed, a linear stability analysis of a spatially-homogeneous distribution gives a Jeans-instability growth rate [36]

$$\tau_c^{-2} = \tau_g^{-2} - k^2 v_{\text{th}}^2, \quad (27)$$

which is essentially the same as  $\tau_c^{-2}$  inferred from the Raychaudhuri equation (26). Furthermore, in the limit  $\tau_g kv_{\text{th}} \ll 1$ ,  $\tau_c$  is scale-independent, because  $\tau_g$  is—by critical balance and equation (18). For more on the collapse time, see also appendix D.

This means that the system experiences collapse on all scales smaller than the outer scale and larger than  $1/k_c(v_{\text{th}}^{-1})$ , so the source  $\hat{S}$  at a scale  $k \equiv 1/r \ll k_c(v_{\text{th}}^{-1})$  receives contributions from all  $k' < 1/r$ . The contribution to the source from scale  $\mathbf{k}$  is  $\int \hat{S} d^d s$ : by the above discussion, every scale collapses at the same rate, because nothing in equation (26) depends on  $\mathbf{k}$ , *provided that the system is cold on these scales*, i.e.,  $s_c(k)v_{\text{th}} \ll 1$ ;<sup>7</sup> besides, the collapse process, by its very nature, brings large-scale structure to smaller scales. We thus posit that  $\int \hat{S} d^d s$  is a constant, so that

$$\int_0^{1/r} d^d k \int d^d s \hat{S} \sim \varepsilon [s_c(1/r)]^d = \varepsilon r^{-d} \kappa^{-d/2}, \quad (28)$$

where  $\varepsilon$  is  $(k, s)$ -independent. We give a dimensional argument for equation (28) in appendix C, whose outcome

<sup>6</sup> Indeed, if initially the system had  $\omega^i = 0$ , then, in the limit  $v_{\text{th}} = 0$ , this will stay this way locally in phase-space [40].

<sup>7</sup> This holds for the  $\mathbf{s}$ -integrated source because (26) is a fluid equation.

is consistent with the physical reasoning laid out above. This is only valid, of course, when  $(kv_{\text{th}})^{-1}$  is much longer than the collapse time and  $sv_{\text{th}} \ll 1$ .

## F. Spectra from $C_2$ cascade

We have argued that two time scales  $\tau_p$  and  $\tau_g$  must balance each other, so the rate of change of  $\hat{F}$  in equation (15) must also be the same as  $\tau_p$  or  $\tau_g$  (for the other two terms in that equation are balanced critically). Now let us use that to find the small-scale asymptotics of the power spectrum  $\hat{F}$ . To leading order in the large- $(k, s)$  limit, one can parameterise it by

$$\hat{F}(k, s) \sim \begin{cases} F_1 k^\gamma s^\xi, & \text{if } s \ll s_c(k) \ll v_{\text{th}}^{-1}, k_{\text{nl}} \ll k \\ F_2 s^\delta k^\sigma, & \text{if } v_{\text{rms}}^{-1} \ll s \ll v_{\text{th}}^{-1}, k \ll k_c(s), \end{cases} \quad (29)$$

for some  $\gamma, \delta \leq 0$ ,  $\xi, \sigma \in \mathbb{R}$ . For  $P(k)$  to be defined,  $\hat{F}$  must be independent of  $s$  in the limit  $s \ll s_c$ ,  $k_c(s) \ll k \ll k_c(v_{\text{th}}^{-1})$ , whence  $\xi = 0$ ;  $\sigma = 0$  for the same reason ( $\bar{f}$  must have a defined velocity variance). Let us find  $\gamma$  and  $\delta$ .

To find  $\gamma$ , consider the amount of  $C_2$  on a scale  $k^{-1}$  (or smaller), which is just the variance of  $f$  over all scales up to  $k = 1/r$ , *viz.*

$$\delta f_r^2 = \frac{1}{(2\pi)^{2d}} \int_{1/r}^\infty d^d k \int d^d s \hat{F}. \quad (30)$$

The flux of  $C_2$  (§II E 1) is given by equation (28)—it is  $\varepsilon s_c^d$ , so the rate of change of  $\delta f_r^2$  is

$$\frac{\delta f_r^2}{\tau_c} \sim \varepsilon s_c(k)^d, \quad (31)$$

by equation (15), where the time-derivative is  $\partial/\partial\eta \sim 1/\tau_c$ . Therefore,

$$\delta f_r^2 \sim \varepsilon \tau_c s_c^d = \frac{\varepsilon s_c^{d+1}}{k} \sim \varepsilon k^d \kappa^{-(d+1)/2}. \quad (32)$$

The  $s$  integral in (30) is dominated by  $s \leq s_c$  if  $\hat{F}$  declines sufficiently rapidly with  $s$  (i.e. if  $\delta \leq -d$ , which will be verified momentarily). Inserting  $\hat{F} \sim F_1 k^\gamma$  into equation (30) and then equating with (32) yields

$$\varepsilon k^d \kappa^{-(d+1)/2} \sim F_1 s_c^d(k) k^{\gamma+d}. \quad (33)$$

For  $s_c(k) \propto k$ , this is solved by  $\gamma = -d$ , which implies that  $P(k) \sim \varepsilon k^{-d}/\sqrt{\kappa}$ . By critical balance,  $\tau_c \sim 1/\sqrt{\kappa}$ .

One can also find  $\hat{F}$  in the large  $s$  limit with  $k$  fixed, i.e., the exponent  $\delta$ , in a similar manner. This time, the analogue of equation (31) for the flux of  $C_2$  to higher velocity scales is

$$\frac{\delta f_v^2}{\tau_c} \sim \varepsilon s^d, \quad (34)$$



where  $\delta f_v^2 \equiv (2\pi)^{-2d} \int_{1/\delta v}^{\infty} d^d s \int d^d k \hat{F}$  is the variance of  $f$  over all velocity scales up to  $s \equiv 1/\delta v$ . Therefore, by setting  $\tau_c = \tau_p$ ,

$$\delta f_v^2 \sim \varepsilon \frac{s}{k_c(s) s^{-d}} = \varepsilon \frac{s^d}{\sqrt{\kappa}}. \quad (35)$$

If  $\hat{F} \sim F_2 s^\delta$  in this limit, then a similar calculation to the one above yields

$$F_2 k_c^d(s) s^{\delta+d} \sim \varepsilon \frac{s^d}{\sqrt{\kappa}}, \quad (36)$$

whence  $\delta = -d$ .

We have thus obtained that for cold dark matter,

$$\hat{F}(k, s) \sim \varepsilon \tau_c \times \begin{cases} k^{-d}, & \text{if } s \ll s_c(k) \ll v_{\text{th}}^{-1}, k_{\text{nl}} \ll k \\ \tau_c^d s^{-d}, & \text{if } v_{\text{rms}}^{-1} \ll s \ll v_{\text{th}}^{-1}, k \ll k_c(s). \end{cases} \quad (37)$$

These are the leading-order asymptotics: they are true for large  $k$  at  $s \rightarrow 0$ , and large  $s$  at  $k \rightarrow 0$ ,<sup>8</sup> and this is the main result of this paper. In general,  $\hat{F}$  can also depend on the angle between  $\mathbf{s}$  and  $\mathbf{k}$ , *viz.*, on  $\mathbf{k} \cdot \mathbf{s}$ . This angular dependence arises only at the next order, because it must not exist at  $s = 0$ , or at  $k = 0$ .<sup>9</sup>

In appendix D, we show that exactly the same scaling is obtained for an Einstein-de-Sitter background (where the scale-factor is  $a(\eta) \propto \eta^2$ ), which has an explicit similarity symmetry, and hence an explicit way of defining  $\tau_c$ .

Let us see whether the theoretical asymptotic scalings (37) are reproduced in our 1D Vlasov-Poisson simulation. The power spectra of the system illustrated in figure 1 are shown in figure 3: a  $k^{-1}$  power-law establishes itself quickly, persisting up to  $k \sim 100k_0$ , which is of the order of  $k_c(v_{\text{th}}^{-1})$ , as expected. These are just  $|\hat{f}|^2$ , averaged over a short time-interval around the time-step when they were computed, not ensemble-averaged, but the time average is a proxy for the latter. We use a short time window for averaging because the system is not stationary, and we re-scale  $|\hat{f}|^2$  by  $C_2(t)$  at every time-step (which decays due to collisions and finite-grid effects). However, in figure 4, which displays cleaner power-laws in the appropriate range of  $k$  and  $s$ , we do average over the entire simulation. While the system is not stationary,

this only affects the amplitude—not the overall scaling—and figure 3 shows that the re-scaling by  $C_2$  corrects for that; hence, we use figure 4 as an approximation for  $\hat{F}$ .

As the system is sourced by the gravitational-collapse instability, which was shown in §II E to continue until  $(k v_{\text{th}})^{-1}$  matches the collapse time, the scaling of the power spectrum must be truncated at  $v_{\text{th}}$ . To test the theory further, it is necessary to see whether it is indeed the case that the asymptotics (37) persist until  $v_{\text{th}}$  is reached. We ran identical simulations, differing only by the value of  $v_{\text{th}}$ , to verify this. The result is presented in figure 5: the left column has  $v_{\text{th}} = 0.005v_0$  while the right column has  $v_{\text{th}} = 0.025v_0$ . The  $s$ -spectra in the bottom row show that indeed, the  $s^{-1}$  power-law is truncated at a lower value of  $s$ , by a factor that matches the ratio of  $v_{\text{th}}$  of both runs. The  $s^{-2}$  (or  $k^{-2}$ ) scaling of  $\hat{F}$  at  $s v_{\text{th}} \gg 1$  (or  $k \gg k_c(v_{\text{th}}^{-1})$ ) is discussed in appendix A.

The  $k^{-d}$  asymptotic of the density power-spectrum may be derived by an altogether different, yet systematic approach—by performing an asymptotic analysis of the an integral expression for  $P(k)$  and examining its critical points. We do so in §III below, which is self-contained.

### III. SADDLE-POINT APPROACH

In this section, we focus on 3 spatial dimensions. For pure gravitational Newtonian evolution of  $N$  identical particles, the power-spectrum is exactly given by [e.g., 33]

$$P(k, t) \propto \prod_{n=1}^N \int d^3 q_n d^3 p_n \mathcal{P}(\{q\}, \{p\}) e^{i\mathbf{k} \cdot [\mathbf{x}_1(t) - \mathbf{x}_2(t)]}, \quad (38)$$

where  $(\mathbf{p}_n, \mathbf{q}_n)$  is the initial phase-space position of particle  $n$  and  $(\mathbf{x}_n(t), \mathbf{v}_n(t))$  is the phase-space position of particle  $n$  at time  $t$ , and  $\mathcal{P}(\{q\}, \{p\})$  is the joint probability distribution of the initial phase-space positions  $(\{q\}, \{p\}) \equiv \{(\mathbf{p}_n, \mathbf{q}_n)\}_{n=1}^N$  of all particles. This equation is permutation-invariant, and, therefore, the choice of two particles is arbitrary.

For *cold* dark matter with Gaussian initial conditions, the initial distribution  $\mathcal{P}$  is

$$\mathcal{P}(\{q\}, \{p\}) = \frac{V^{-N} \mathcal{C}(\{q\}, \{p\})}{\sqrt{(2\pi)^{3N} \det C_{pp}^N}} e^{-\frac{\{p\}^T (C_{pp}^N)^{-1} \{p\}}{2}}, \quad (39)$$

where  $C_{pp}^N = C_{pp}^N(\{q\})$  is the  $3N \times 3N$  covariance matrix of  $\{p\}$ , and  $\mathcal{C}$  encapsulates initial density-density and density-momentum correlations [54], whose functional dependence on particle positions depends on the cosmology (we take a  $\Lambda$ CDM background). Below we will require the following properties of  $C_{pp}^N$ , derived in appendix E, based on the assumption  $v_{\text{th}} \rightarrow 0$  (so the conclusion again will be valid at  $k \ll k_c(v_{\text{th}}^{-1})$  as in §II): (i)  $\Sigma \simeq Aq^2$  at small  $q$  for some order-unity matrix  $A$ , (ii)  $|\mathbf{a}|^2 \Sigma \sim O(1)$  in this limit, and (iii) at  $q \rightarrow \infty$ , the

<sup>8</sup> This scaling is marginal, in that the region up to  $s = s_c$  in the integral in equation (30) turns out to be as large at the region  $s > s_c$ ; this, however, does not invalidate the conclusion, because if both are dominant, then they both contribute  $\sim s_c^d$ , so the rest of the argument still goes through.

<sup>9</sup> A repetition of the calculation presented in this sub-section with  $\lambda < 1$  (defined below equation (12)) yields that there does not exist a solution to the analogues of (33) and (34), when conjoined with Poisson's equation,  $k \delta g_r \sim \delta \rho_r$ , whence  $\lambda < 1$  is not consistent with a self-gravitating system, whose gravitational field is not dominated by an external potential. See appendix B for details.

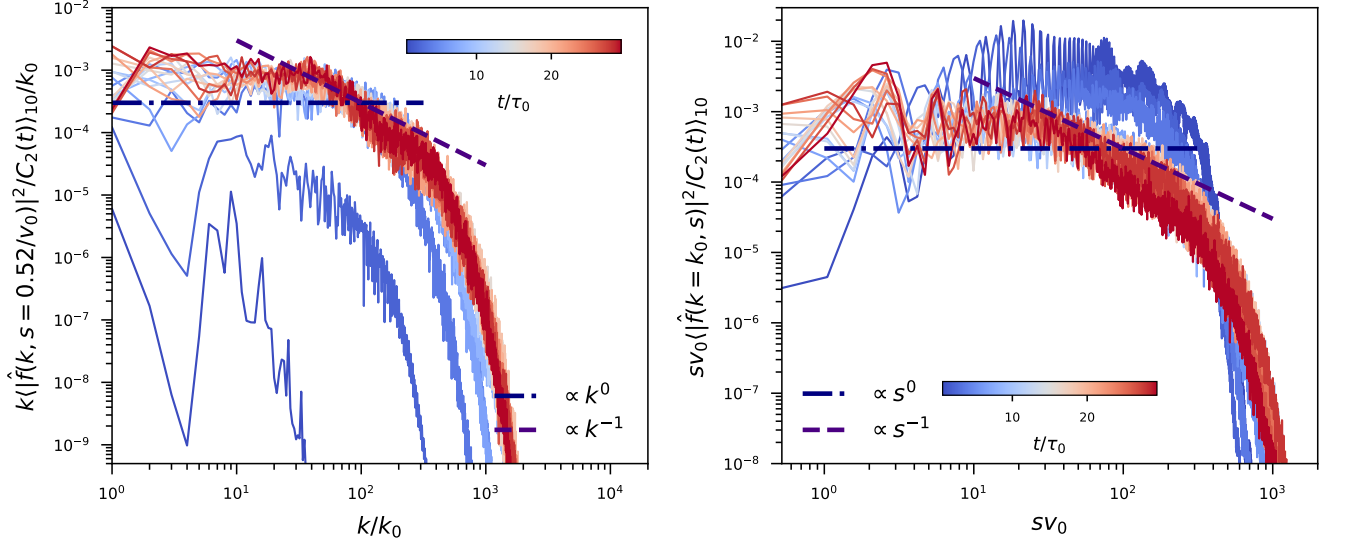


FIG. 3. The time evolution of the power spectra  $\langle |\hat{f}|^2 \rangle_{10}$  for the simulation illustrated in figure 1. These were produced by time-averaging  $|\hat{f}|^2$  over 10 simulation outputs (corresponding to time windows of  $\pm 0.68\tau_0$ ). *Left panel*: the  $k$ -spectrum (at the fundamental  $s$  mode) at different times, multiplied by  $k$ . *Right panel*: the  $s$ -spectrum at the fundamental  $k$  mode. Both spectra are compensated by the expected asymptotics (multiplied by  $k$  and  $s$  respectively). A  $k^{-1}$  power-law establishes itself quickly, persisting until  $k \sim 200k_0$ , which is of the order of  $k_c(v_{\text{th}}^{-1})$ , as expected; similarly, the  $s^{-1}$  power-law persists until  $sv_0 \sim 200$ . See appendix A for plots of  $\langle |\hat{f}|^2 \rangle_{10}$  at some individual times.

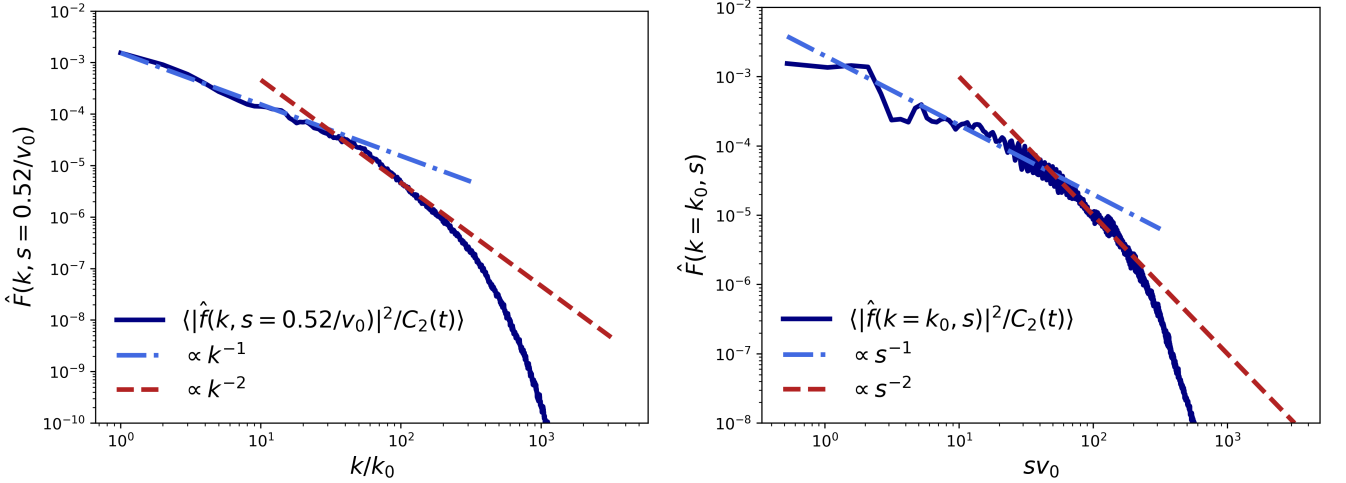


FIG. 4. The time-averaged spectra for the same set-up as in figure 1, as proxies for the ensemble-averages (see text). Unlike figure 3, these power spectra are uncompensated.

correlation matrix  $C_{pp}^N$  tends to a constant, i.e.  $\Sigma$  is order unity in the limit  $q \rightarrow \infty$ .

#### A. Strategy to obtain the small-scale asymptotics

The usual procedure to obtain the power spectrum  $P(k)$  from eq. (38) would involve integrating out particles  $3, \dots, N$ , leaving only a 12-dimensional integral, over the phase-sub-space of particles 1 and 2. We will, how-

ever, go the other way round, and integrate *first* over the relative position and momentum of this pair, and only then over all other particles—we will see that this integration order is well-suited to deriving the asymptotics of  $P(k, t)$  as  $k \rightarrow \infty$ . To do so, we change variables from  $\mathbf{q}_1$  and  $\mathbf{q}_2$  to  $\mathbf{q} \equiv \mathbf{q}_1 - \mathbf{q}_2$  and  $\mathbf{Q} \equiv (\mathbf{q}_1 + \mathbf{q}_2)/2$ , and to their conjugate momenta,  $\mathbf{p}$  and  $\mathbf{P}$ , respectively. As  $\{p\}^T (C_{pp}^N)^{-1} \{p\}$  is quadratic in  $\{p\}$ , we can make the

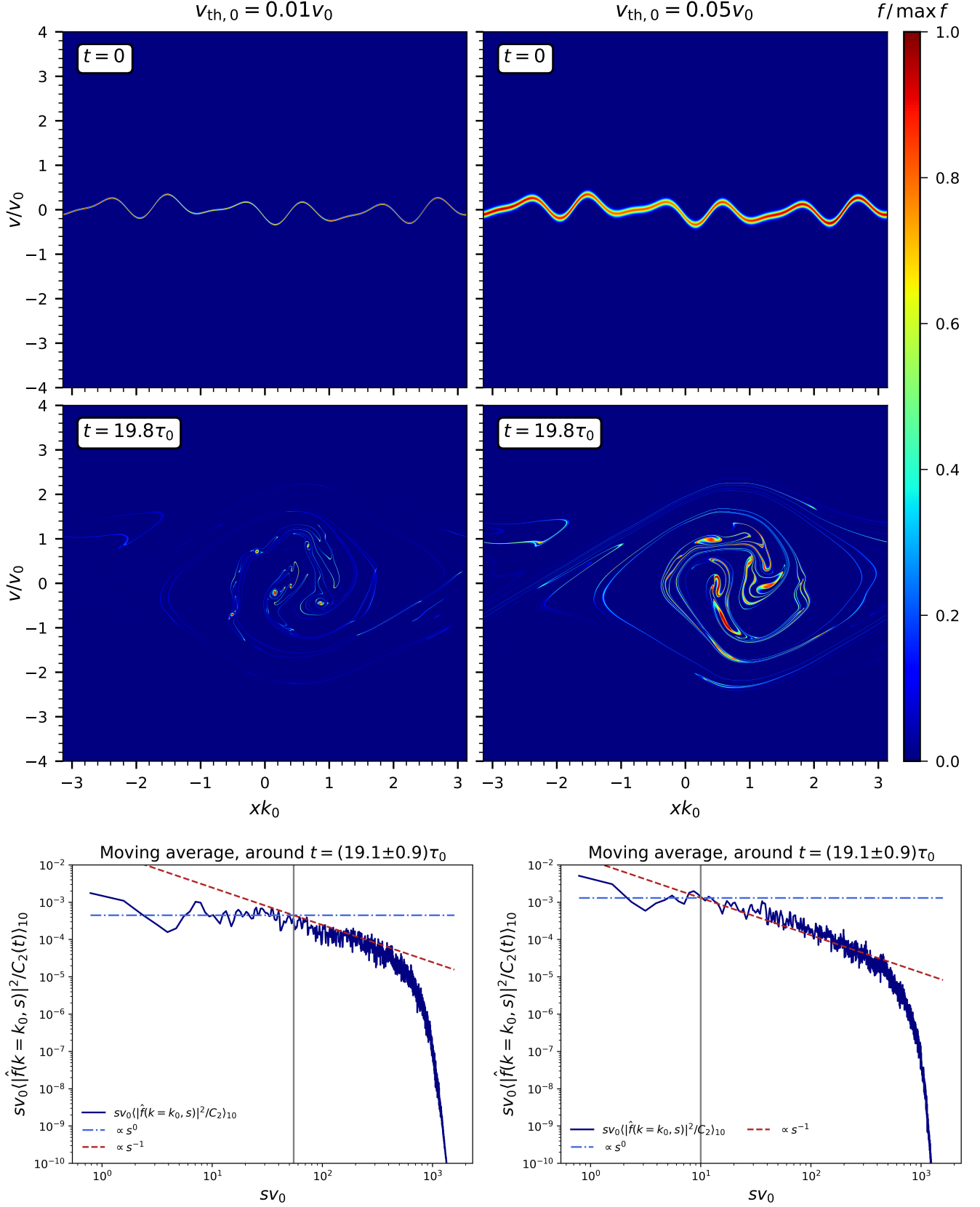


FIG. 5. A comparison between the evolution and the spectra of two systems, with identical initial conditions, save for the value of  $v_{th}$ . *Left panels*: a cold case,  $v_{th} = 0.01v_0$ ; *right panels*: a warmer case with  $v_{th} = 0.05v_0$ . The plots in the bottom row show the  $s$  spectra at the end of the simulations, compensated by  $s$ . In the warmer case, the spectrum ceases to scale as  $s^{-1}$  at around  $sv_0 \sim 10$ , while in the colder one it does so at  $sv_0 \sim 50$  (marked by grey lines). See text and appendix A for details.

$\mathbf{p}$  dependence explicit, *viz.*,

$$\begin{aligned} \{p\}^T (C_{pp}^N)^{-1} \{p\} &\equiv \mathbf{p}^T \Sigma^{-1}(\mathbf{q}) \mathbf{p} \\ -2\mathbf{a}(\mathbf{q}, \mathbf{Q}, \mathbf{P}, \{(q, p)\}_3^N) \cdot \mathbf{p} &- 2B(\mathbf{q}, \mathbf{Q}, \mathbf{P}, \{(q, p)\}_3^N), \end{aligned} \quad (40)$$

where  $\mathbf{a}$  is a linear function of momenta and  $B$  a quadratic one.

Using Duhamel's principle for Hamilton's equations of motion, one has [e.g., 33]

$$\mathbf{i}\mathbf{k} \cdot [\mathbf{x}_1(t) - \mathbf{x}_2(t)] = \mathbf{i}\mathbf{q} \cdot \mathbf{k} + \mathbf{i}g_{qp}(t, t_{\text{initial}}) \mathbf{p} \cdot \mathbf{k} + \mathbf{i}k\psi_I, \quad (41)$$

where  $g_{qp}$  is a 'propagator' that encapsulates linear evolution (defined below), and the 'interaction term' is

$$\begin{aligned} k\psi_I(\mathbf{P}, \mathbf{Q}, \{(q, p)\}_3^N) \\ \equiv \mathbf{k} \cdot \int_0^t dt' g_{qp}(t', t_{\text{initial}}) (\mathbf{F}_1(t') - \mathbf{F}_2(t')), \end{aligned} \quad (42)$$

$\mathbf{F}_n$  being the additional acceleration of particle  $n$ , relative to its motion in the part already included in  $g_{qp}$ , caused by all other particles. For example, if  $g_{qp} = t - t_{\text{initial}}$ , one re-obtains regular Newtonian dynamics. Using input from first-order cosmological perturbation theory (or, alternatively, re-summed kinetic field theory [55]), we take  $g_{qp}(t) \equiv [D_+(t) - D_+(t_{\text{initial}})]/\dot{D}_+$ , where  $D_+$  is the  $\Lambda$ CDM growth factor, and then ignoring the  $\psi_I$  term in (41) just yields the Zel'dovich approximation, whose asymptotics were studied by [56].

## B. Possible saddle points

Using the above changed variables, we define the exponent

$$\varphi \equiv -\frac{1}{2} \mathbf{p}^T \Sigma^{-1} \mathbf{p} + \mathbf{a} \cdot \mathbf{p} + \mathbf{i}\mathbf{q} \cdot \mathbf{k} + \mathbf{i}g_{qp} \mathbf{p} \cdot \mathbf{k} + \mathbf{i}k\psi_I + B, \quad (43)$$

which, using equations (40) and (41), turns the integral for  $P(k, t)$  into

$$P(k, t) \propto \prod_{n=3}^N \int \frac{d^3 q_n d^3 p_n d^3 Q d^3 P d^3 q d^3 p}{V^N \sqrt{(2\pi)^{3N}} \det C_{pp}^N} \mathcal{C}(\{q\}, \{p\}) e^\varphi. \quad (44)$$

We start by integrating over  $\mathbf{p}, \mathbf{q}$  in equation (38), and apply the saddle-point approximation. This is a movable-saddle problem, so it requires care in handling [57].<sup>10</sup> The function  $\psi_I$  is a smooth function of  $\mathbf{q}$  and  $\mathbf{p}$ , because it arises from the Hamiltonian flow in phase-space, generated by a smooth gravitational potential (recall that we

neglect collisions and that time is bounded). This assertion is just the statement that  $\lambda = 1$  as before. Letting  $\psi_q \equiv \partial\psi_I/\partial\mathbf{q}$  and  $\psi_p \equiv \partial\psi_I/\partial\mathbf{p}$ , the exponent (43) is stationary when

$$\begin{aligned} -\Sigma^{-1} \mathbf{p} + \mathbf{a} + \mathbf{i}g_{qp} \mathbf{k} + \mathbf{i}k\psi_p &= 0 \\ -\frac{1}{2} p^i \frac{\partial(\Sigma^{-1})_{ij}}{\partial\mathbf{q}} p^j + p_i \frac{\partial a^i}{\partial\mathbf{q}} + \frac{\partial B}{\partial\mathbf{q}} &+ \mathbf{i}\mathbf{k} + \mathbf{i}k\psi_q = 0. \end{aligned} \quad (45)$$

Let us parameterise the solution as

$$\mathbf{p} = k^\alpha \mathbf{c}, \quad \mathbf{q} = k^\beta \mathbf{d}, \quad (46)$$

where  $\mathbf{c}, \mathbf{d} \in \mathbb{C}^3$  are order unity as  $k \rightarrow \infty$ . Together with  $\alpha, \beta$ , they are to be determined by equations (45), by seeking a dominant balance, i.e., such a balance that the exponent (43) has the weakest  $k$ -dependence around the stationary point that solves equations (45).

*A priori*, equations (45) could permit a balance that is independent of the initial condition distribution, i.e., that same balance would exist for uniform initial conditions. But in that case, Liouville's theorem—changing variables from the initial phase-space positions  $\{(\mathbf{q}, \mathbf{p})\}^N$  to the current ones  $\{(\mathbf{x}, \mathbf{v})\}^N$ —ensures that this yields a contribution to  $P(k)$  proportional to  $\delta^D(\mathbf{k})$ , whence this balance is not dominant. Conversely, the dominant balance must involve the first terms on the left-hand sides of equations (45). If  $\beta > 0$ , then, as  $\Sigma \xrightarrow{q \rightarrow \infty} \text{const}$ , the first term in the first of equations (45) is proportional to  $p$ . As this must play a part in a dominant balance, this implies that  $\alpha = 1$ . Substituting such a saddle point into equation (43) gives an exponentially suppressed contribution,  $\sim O[\exp(\alpha - k^2)]$  to  $P(k)$  at most.

Thus stationary points with  $\beta < 0$  are dominant, as they could contribute a power-law tail to  $P(k)$ . For  $\beta < 0$ , equations (45) imply  $1 = \alpha - 2\beta$  and  $1 = 2\alpha - 3\beta$ , whence  $\alpha = \beta = -1$ . This is consistent, since generically

$$\lim_{\substack{(q,p) \sim k^{-1} \\ k \rightarrow \infty}} (|\psi_q|, |\psi_p|) = O(1), \quad (47)$$

and so the interaction term is potentially as important as the linear term; that both are equally important is essentially a statement of critical balance (cf. §II C).

## C. Evaluation of the asymptotics

Having proven that the asymptotic expansion of the power-spectrum is given, up to exponentially small (in  $k$ ) contributions, by its dominant saddle point at  $q, p \sim k^{-1}$ , one is allowed to replace  $\varphi$  with its expansion at small  $q, p$ . This yields

$$\begin{aligned} \varphi \simeq -\frac{1}{2} \mathbf{p}^T \Sigma^{-1} \mathbf{p} + \mathbf{a} \cdot \mathbf{p} + \mathbf{i}\mathbf{q} \cdot \mathbf{k} + \mathbf{i}g_{qp} \mathbf{p} \cdot \mathbf{k} \\ + \mathbf{i}k [\mathbf{p} \cdot \psi_p(0) + \mathbf{q} \cdot \psi_q(0)] + B(q=0) + o(1). \end{aligned} \quad (48)$$

<sup>10</sup> Also bear in mind that the saddle point may be complex, but that is innocuous, since the exponent can be continued analytically to the complex plane.

With this asymptotic approximation for  $\varphi$ , we need to integrate equation (44) over  $\mathbf{p}$  and  $\mathbf{q}$  to obtain its asymptotic scaling with  $k$ . Observe, that  $\det C_{pp}^N$  factorises into a  $\det \Sigma$  multiplied by normalisation factors for the other

momenta variables. The only other pieces that still depend on  $\mathbf{Q}$ ,  $\mathbf{P}$ , and the initial positions and momenta of the other particles are  $\mathcal{C}$ ,  $\mathbf{a}$ ,  $B$ , and  $\psi_{q,p}$ . Thus,

$$P(k, t) \simeq \left\langle \int \frac{d^3 q d^3 p \mathcal{C}}{[(2\pi)^3 \det \Sigma]^{1/2}} \exp \left[ -\frac{1}{2} \mathbf{p}^T \Sigma^{-1} \mathbf{p} + \mathbf{a} \cdot \mathbf{p} + i\mathbf{q} \cdot \mathbf{k} + ig_{qp} \mathbf{p} \cdot \mathbf{k} \right] e^{ik(\mathbf{q} \cdot \psi_q(0) + \mathbf{p} \cdot \psi_p(0))} \right\rangle, \quad (49)$$

where average is over the position and momentum of the centre of mass of particles 1 and 2, as well as the positions and momenta of all the other particles ( $B$  is absorbed into the averaging). Integrating over  $\mathbf{p}$ , we get

$$P(k, t) \simeq \left\langle \int d^3 q \mathcal{C} \exp \left\{ -\frac{k^2}{2} \left[ g_{qp} \hat{\mathbf{k}} + \psi_p(0) - \frac{i\mathbf{a}}{k} \right]^T \Sigma(\mathbf{q}) \left[ g_{qp} \hat{\mathbf{k}} + \psi_p(0) - \frac{i\mathbf{a}}{k} \right] + i\mathbf{q} \cdot \mathbf{k} + ik\mathbf{q} \cdot \psi_q(0) \right\} \right\rangle. \quad (50)$$

This integration does not introduce any power of  $k$  because both  $d^3 p$  and  $\det \Sigma^{1/2}$  are proportional to  $k^{-3}$  at the vicinity of the stationary point, so the two cancel. Recalling that the saddle point is at  $\mathbf{q} = \mathbf{c}/k$ , where  $\mathbf{c}$  is an order-unity constant, we change variables to  $\mathbf{y} = k\mathbf{q}$ , whence

$$P(k, t) \simeq \frac{1}{k^3} \left\langle \int d^3 y \mathcal{C} e^{-\frac{k^2}{2} [g_{qp} \hat{\mathbf{k}} + \psi_p(0) - \frac{i\mathbf{a}}{k} \mathbf{a}(\mathbf{y}/k)]^T \Sigma(\mathbf{y}/k) [g_{qp} \hat{\mathbf{k}} + \psi_p(0) - \frac{i\mathbf{a}}{k} \mathbf{a}(\mathbf{y}/k)] + i\mathbf{y} \cdot [\hat{\mathbf{k}} + \psi_q(0)]} \right\rangle. \quad (51)$$

Here, the integral  $d^3 q$ , as opposed to the momentum integral above, is not compensated by any function that scales like  $k^{-3}$ , so it is this integration that yields the  $k^{-3}$  scaling, which emerges from the  $q \sim k^{-1}$  at the stationary point.

As  $\Sigma(q) \simeq Aq^2 + \text{h.o.t.}$  at small  $q$ , we find that the coefficient multiplying  $k^{-3}$  is of order unity ( $\mathcal{C}$  is evaluated at  $\mathbf{p} = \mathbf{d}/k$ ,  $\mathbf{q} = \mathbf{c}/k$ ). Writing  $\Sigma^{nm} \simeq S_{ij}^{nm} q^i q^j$  and  $\lambda^i = g_{qp} \hat{k}^i + \psi_p^i(0)$ , we have that the exponent in (51) is

$$\begin{aligned} & \left[ g_{qp} \hat{\mathbf{k}} + \psi_p(0) - \frac{i\mathbf{a}}{k} \mathbf{a}(\mathbf{y}/k) \right]^T \Sigma(\mathbf{y}/k) \left[ g_{qp} \hat{\mathbf{k}} + \psi_p(0) - \frac{i\mathbf{a}}{k} \mathbf{a}(\mathbf{y}/k) \right] + i\mathbf{y} \cdot [\hat{\mathbf{k}} + \psi_q(0)] \\ & = S_{ij}^{nm} \lambda_n \lambda_m y^i y^j - 2i \left[ g_{qp} \hat{\mathbf{k}} + \psi_p(0) \right]^T S_a \mathbf{y} - |\tilde{\mathbf{a}}|^2 + i\mathbf{y} \cdot [\hat{\mathbf{k}} + \psi_q(0)], \end{aligned} \quad (52)$$

where  $\tilde{\mathbf{a}} \equiv \lim_{q \rightarrow 0} \sqrt{\Sigma} \mathbf{a}$  and  $S_a \mathbf{q} \equiv \lim_{q \rightarrow 0} \Sigma \mathbf{a}$ , whence the leading-order term becomes

$$P(k) \simeq \frac{(2\pi)^{3/2}}{k^3} \left\langle \mathcal{C} \sqrt{\det S} e^{|\tilde{\mathbf{a}}|^2} e^{-[\hat{\mathbf{k}} + \psi_q(0) + S_a^T (g_{qp} \hat{\mathbf{k}} + \psi_p(0))]^T S^{-1} [\hat{\mathbf{k}} + \psi_q(0) + S_a^T (g_{qp} \hat{\mathbf{k}} + \psi_p(0))]/2} \right\rangle. \quad (53)$$

with  $S_{ij} \equiv S_{ij}^{nm} \lambda_n \lambda_m$ . The inner product  $S_{ij} y^i y^j$  is nothing but the matrix  $\Sigma$ , contracted with  $\lambda^i - i\mathbf{a}^i/k$  and its conjugate.

We have therefore established that for collisionless, cold, dark matter in the non-relativistic limit, the asymptotic behaviour of the power spectrum as  $k \rightarrow \infty$  is dominated by the contribution of the saddle point at  $q, p \sim k^{-1}$ . This yields  $P(k) \sim k^{-3}$ , with an order-unity coefficient; that this coefficient is non-zero follows from §II.

## IV. DISCUSSION

### A. The importance of being cold

The  $k^{-d}$  power spectrum arose from a turbulent cascade in phase-space, where the gravitational field was treated as a smooth field with the Batchelor approxima-

tion; this is a cold, gravitational, phase-space analogue of Batchelor turbulence in hydrodynamics [48], where a tracer is advected by a large-scale, smooth velocity field. A crucial step in the derivation of the source term in §II E was that the source was a constant multiplied by  $\hat{F}$ , at  $s_c(k)$ ,  $s \ll v_{\text{th}}^{-1}$ , i.e., that it accumulated when integrating to give  $\varepsilon s^d$ . This was also reflected in §III, where the initial conditions were special for cold dark matter, with the particular functional form of the  $C_{pp}^N$  matrix,  $\Sigma \sim Aq^2$ , reflecting that if two particles start out at exactly the same spatial position, they will remain together for ever.

When the system is not cold, these assumptions fail. Indeed, the source term would not accumulate when integrated over  $k$  and  $s$  up to small, non-linear scales, because

at  $sv_{\text{th}}^{-1} \gg 1$ ,  $s_c(k)v_{\text{th}} \gg 1$ , there is no instability to drive the collapse, so one would expect  $\iint \hat{S}dkds = \text{const}$ .

The coldness has two implications: first, it means that locally in phase-space, the distribution function is always a stream, and therefore, locally, the Raychaudhuri equation (5) applies. This, in turn, means that the stream is unstable to gravitational collapse, with a collapse time-scale, set by  $\tau_g \sim 1/\sqrt{G\delta\rho}$ , i.e., precisely the time-scale to give equation (31). Secondly,  $\tau_c$  is scale-independent, so the contributions to the integrated source (25) add up coherently. Together, these two features imply that the source has the shape in equation (28); this then enables the estimate of  $\hat{F}$  in §II F.

While in a collisionless setting  $v_{\text{th}}$  would not change with time,<sup>11</sup> in reality finite- $N$  effects do increase  $v_{\text{th}}$ . Additionally, baryons would influence the matter power spectrum significantly on small scales, probably leading to a deviation from the power-law scalings found here. The approach used in this paper does not apply once interactions with baryons are included, and this is its main limitation. Another limitation is that it only applies in-so-far as dark matter behaves as a *cold*, classical *phase-space* fluid—other dark-matter candidates might follow different scalings, and we defer this to future work.

Were it not for baryons, a full theory of the constituent particle of dark matter should predict both the initial value of  $v_{\text{th}}$  and the functional form of the collision operator; this should allow one to find how  $C_2$  changes with time, and thence how  $v_{\text{th}}$  evolves. Measuring the break in the dark-matter power-spectrum—the transition from  $k^{-3}$  to a different power-law (in  $d = 1$  to a  $k^{-2}$  scaling as in figure 4)—would hypothetically allow one to find the present-day value of  $v_{\text{th}}$ , which is directly connected to the nature of dark matter (if indeed it held that  $u_\nu \ll v_{\text{th}}$ , where the collisional velocity-scale  $u_\nu$  is defined in appendix A). For example, for WIMPs, which decouple from the photons when non-relativistic, one would have

$$v_{\text{th}}(z) = c \frac{T(z)}{T_{\text{kd}}} \sqrt{\frac{2k_{\text{B}}T_{\text{kd}}}{mc^2}} \quad (54)$$

$$\approx 3.3 \times 10^{-12} c(1+z) \left[ \frac{10 \text{ MeV}}{T_{\text{kd}}} \right]^{1/2} \left[ \frac{\text{GeV}}{m} \right]^{1/2},$$

where  $T(z)$  is the photon temperature at redshift  $z$  and  $T_{\text{kd}}$  is the kinetic-decoupling temperature [20, 21, 58]. This is by far too small to be observed practically, and

<sup>11</sup> Because  $C_2 \sim \|f\|_\infty^2 v_{\text{th}}^d$ , and both  $C_2$  and  $\|f\|_\infty$  are conserved (in fact, one can define  $v_{\text{th}} \equiv [C_2 \|f\|_\infty^{-2}]^{1/d} / \sqrt{\pi}$  as a measure of the thinness of the distribution). There is of course a change in  $\|f\|_\infty = \max f$  because of the Universe's expansion, but this is very slow for the large- $k$  limit, where  $k \gg \mathcal{H}/c$ .

moreover, it corresponds to (cf. appendix D)

$$k_c(v_{\text{th}}^{-1}) \sim \delta_{\text{typ}}^{1/2} \frac{H^{-1}}{v_{\text{th}}} \quad (55)$$

$$\approx \frac{10^8 h(z)}{1+z} \delta_{\text{typ}}^{1/2} \text{Mpc}^{-1} \left[ \frac{m}{\text{GeV}} \right]^{1/2} \left[ \frac{T_{\text{kd}}}{10 \text{ MeV}} \right]^{1/2},$$

where the Hubble constant is  $H(z) = 100h(z) \text{ km s}^{-1} \text{ Mpc}^{-1}$  (this value is of the same order as the free-streaming scale [20] for a typical over-density  $\delta_{\text{typ}}$  of order unity).

## B. Some like it hot: similarities with plasma physics

Phase-space Batchelor cascade was recently proposed to be the universal régime of (plasma) Vlasov-Poisson turbulence at Debye and sub-Debye scales [31, 32], and has been numerically verified in 1D simulations of turbulence driven by external forcing [32] and the two-stream instability [59]. Like in the cold-dark-matter turbulence presented here, the (electric) field fluctuations are spatially smooth, so the phase-space mixing of the distribution function is dominated by the outer scale fields. However, unlike cold-dark-matter turbulence, which is sourced at every Jeans-unstable scale, the cascade in the plasma case is one of constant  $C_2$  flux, because  $C_2$  is only sourced at the outer scale; this changes the scalings of the power spectrum and of the field spectrum. This situation is analogous to the gravitational phase-space turbulence at  $s \gg v_{\text{th}}^{-1}$  in the simulations presented in figure 5. We will investigate this régime further in future work.

## C. Implications for dark-matter haloes

The  $P(k) \sim k^{-3}$  scaling (in 3D) derived here for the non-linear power-spectrum sheds some light on universal properties of dark matter haloes; within the hierarchical clustering paradigm [60], the small-scale limit of the power-spectrum is dominated by the 1-halo term,

$$P_{1h}(k) = \int dM \frac{dn}{dM} [R^3 \bar{\delta} \tilde{u}(kR)]^2, \quad (56)$$

where  $dn/dM$  is the halo mass function,  $R(M) \equiv R_{200}(M)/c(M)$ ,  $\bar{\delta}$  is the amplitude of the halo density profile,  $c(M)$  is its concentration,  $R_{200}$  is the radius where the density is 200 times the critical density of the Universe, and  $\tilde{u}$  is the Fourier transform of the normalised density profile  $u(x/R_s)$  [61]. If we define  $\nabla_c \equiv -d \ln c / d \ln M$  and  $\nabla_n \equiv M \frac{d}{dM} \left[ \ln \frac{dn}{dM} \right]$ , both at  $M \rightarrow 0$ , then, as  $k \rightarrow \infty$

$$P_{1h}(k) \sim k^\gamma, \quad \gamma = \frac{-3(3 + \nabla_n)}{(1 + 3\nabla_c)}, \quad (57)$$

provided the halo density-profile  $\tilde{u}(w)$  decreases sufficiently fast with  $w$ . Setting  $\gamma = -3$  yields a relation between the universal concentration-to-mass relation and the halo mass function, *viz.*

$$\nabla_n = 3\nabla_c - 2. \quad (58)$$

## V. SUMMARY

In this paper we have described a physical mechanism that produces the  $k^{-d}$  asymptotic scaling of the dark-matter power spectrum naturally. This was done in two ways: by expressing  $P(k)$  as a phase-space integral and analysing it with a stationary-phase method (§III), and via a phenomenological study of a critically-balanced phase-space cascade, akin to Batchelor turbulence, for a cold, collisionless, self-gravitating system (§II). Both methods are phase-space based, and so remain valid even after streams cross. The fact that the phase-space distribution function is cold (i.e. only non-zero in a  $d$ -dimensional sub-manifold of phase-space) was crucial to both approaches. Gravitational collapse sources a cascade of the quadratic Casimir invariant. The cascade is sustained by the joint action of phase mixing and tidal forces (which are a smooth Batchelor-like field), that transfer phase-space structure into ever smaller scales. The balance between linear free streaming and tidal forces also appears in the saddle-point argument of §III. Usually in turbulent systems there is an inertial range [62], where there is a constant-flux cascade of an invari-

ant. Gravitational turbulence is unique in that the flux is not constant over the range of scales of interest, and yet there is a universal scaling régime of the phase-space power spectrum.

We have also conducted 1D Vlasov-Poisson simulations, which confirmed the theory.

The derivation of the small-scale asymptotics of the dark-matter power-spectrum may allow for a non-trivial test of effective field theories of the large scales, both by imposing these asymptotics on them, or by using  $\hat{F}$  found here as a closure for these theories. We intend to investigate this in future work.

## ACKNOWLEDGMENTS

We wish to thank Michael Barnes, William Clarke, and Anatoly Spitkovsky for helpful discussions, and are grateful to James Juno for help and advice with the Gkeyll code. Y.B.G. was supported in part by a Leverhulme Trust International Professorship Grant to S. Sondhi (No. LIP-2020-014). Y.B.G. and A.A.S. were supported in part by the Simons Foundation via a Simons Investigator Award to A.A.S.. M.L.N. was supported by a Clarendon Scholarship at Oxford. R.J.E. was supported by a UK EPSRC studentship. S.K. and M.B. were supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC 2181/1-390900948 (the Heidelberg STRUCTURES Excellence Cluster). This research was supported in part by NSF grant PHY-2309135 to the Kavli Institute for Theoretical Physics (KITP).

- 
- [1] P. J. E. Peebles, *The Large-Scale Structure of the Universe* (Princeton University Press, Princeton, N.J., 1980).
- [2] V. Desjacques, D. Jeong, and F. Schmidt, Large-scale galaxy bias, *Phys. Rep.* **733**, 1 (2018), arXiv:1611.09787 [astro-ph.CO].
- [3] N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, *et al.* (Planck Collaboration), Planck 2018 results. VI. Cosmological parameters, *A&A* **641**, A6 (2020), arXiv:1807.06209 [astro-ph.CO].
- [4] M. Tegmark and M. Zaldarriaga, Separating the early universe from the late universe: Cosmological parameter estimation beyond the black box, *Phys. Rev. D* **66**, 103508 (2002), arXiv:astro-ph/0207047 [astro-ph].
- [5] G. Cabass, M. M. Ivanov, M. Lewandowski, M. Mirbabayi, and M. Simonović, Snowmass white paper: Effective field theories in cosmology, *Physics of the Dark Universe* **40**, 101193 (2023), arXiv:2203.08232 [astro-ph.CO].
- [6] D. Heggie and P. Hut, *The Gravitational Million-Body Problem: A Multidisciplinary Approach to Star Cluster Dynamics* (Cambridge University Press, Cambridge, 2003).
- [7] A. J. S. Hamilton, P. Kumar, E. Lu, and A. Matthews, Reconstructing the primordial spectrum of fluctuations of the Universe from the observed nonlinear clustering of galaxies, *ApJ* **374**, L1 (1991).
- [8] Y. B. Ginat, Multiple-scales approach to the averaging problem in cosmology, *J. Cosmology Astropart. Phys.* **2021**, 049 (2021), arXiv:2005.03026 [gr-qc].
- [9] D. Baumann, A. Nicolis, L. Senatore, and M. Zaldarriaga, Cosmological non-linearities as an effective fluid, *J. Cosmology Astropart. Phys.* **2012**, 051 (2012), arXiv:1004.2488 [astro-ph.CO].
- [10] T. Baldauf, Effective field theory of large-scale structure, in *Effective Field Theory in Particle Physics and Cosmology: Lecture Notes of the Les Houches Summer School: Volume 108, July 2017* (Oxford University Press, 2020).
- [11] M. M. Ivanov, Effective field theory for large-scale structure, in *Handbook of Quantum Gravity*, edited by C. Bambi, L. Modesto, and I. Shapiro (Springer Nature Singapore, Singapore, 2023).
- [12] V. Springel, R. Pakmor, A. Pillepich, R. Weinberger, D. Nelson, L. Hernquist, M. Vogelsberger, S. Genel, P. Torrey, F. Marinacci, and J. Naiman, First results from the IllustrisTNG simulations: matter and galaxy clustering, *MNRAS* **475**, 676 (2018), arXiv:1707.03397

- [astro-ph.GA].
- [13] A. Taruya and S. Colombi, Post-collapse perturbation theory in 1D cosmology—beyond shell-crossing, *MNRAS* **470**, 4858 (2017), arXiv:1701.09088 [astro-ph.CO].
- [14] A. Halle, S. Colombi, and S. Peirani, Phase-space structure analysis of self-gravitating collisionless spherical systems, *A&A* **621**, A8 (2019), arXiv:1701.01384 [astro-ph.GA].
- [15] S.-F. Chen and M. Pietroni, Asymptotic expansions for Large Scale Structure, *J. Cosmology Astropart. Phys.* **2020**, 033 (2020), arXiv:2002.11357 [astro-ph.CO].
- [16] I. H. Gilbert, Collisional relaxation in stellar systems, *ApJ* **152**, 1043 (1968).
- [17] P.-H. Chavanis, Kinetic theory of long-range interacting systems with angle-action variables and collective effects, *Physica A* **391**, 3680 (2012), arXiv:1107.1475 [cond-mat.stat-mech].
- [18] D. Lazarovici and P. Pickl, A mean field limit for the Vlasov-Poisson system, *Arch. Ration. Mech. Anal.* **225**, 1201 (2017).
- [19] C. Rampf, Cosmological Vlasov-Poisson equations for dark matter, *Rev. Mod. Plasma Phys.* **5**, 10 (2021).
- [20] A. M. Green, S. Hofmann, and D. J. Schwarz, The power spectrum of SUSY-CDM on subgalactic scales, *MNRAS* **353**, L23 (2004), arXiv:astro-ph/0309621 [astro-ph].
- [21] A. Loeb and M. Zaldarriaga, Small-scale power spectrum of cold dark matter, *Phys. Rev. D* **71**, 103520 (2005), arXiv:astro-ph/0504112 [astro-ph].
- [22] M. Kunz, S. Nesseris, and I. Sawicki, Constraints on dark-matter properties from large-scale structure, *Phys. Rev. D* **94**, 023510 (2016), arXiv:1604.05701 [astro-ph.CO].
- [23] D. B. Thomas, M. Kopp, and C. Skordis, Constraining the properties of dark matter with observations of the Cosmic Microwave Background, *ApJ* **830**, 155 (2016), arXiv:1601.05097 [astro-ph.CO].
- [24] D. Gilman, S. Birrer, A. Nierenberg, T. Treu, X. Du, and A. Benson, Warm dark matter chills out: constraints on the halo mass function and the free-streaming length of dark matter with eight quadruple-image strong gravitational lenses, *MNRAS* **491**, 6077 (2020), arXiv:1908.06983 [astro-ph.CO].
- [25] S. Ilić, M. Kopp, C. Skordis, and D. B. Thomas, Dark matter properties through cosmic history, *Phys. Rev. D* **104**, 043520 (2021), arXiv:2004.09572 [astro-ph.CO].
- [26] N. Menci, An Eulerian perturbation approach to large-scale structures: extending the adhesion approximation, *MNRAS* **330**, 907 (2002), arXiv:astro-ph/0111228 [astro-ph].
- [27] E. Kozlikin, R. Lilow, F. Fabis, and M. Bartelmann, A first comparison of Kinetic Field Theory with Eulerian Standard Perturbation Theory, *J. Cosmology Astropart. Phys.* **2021**, 035 (2021), arXiv:2012.05812 [astro-ph.CO].
- [28] G. Knorr, Time asymptotic statistics of the Vlasov equation, *J. Plasma Phys.* **17**, 553 (1977).
- [29] P. H. Diamond, S.-I. Itoh, and K. Itoh, *Modern Plasma Physics* (Cambridge University Press, Cambridge, 2010).
- [30] M. Lesur and P. H. Diamond, Nonlinear instabilities driven by coherent phase-space structures, *Phys. Rev. E* **87**, 031101 (2013).
- [31] M. L. Nastac, R. J. Ewart, W. Sengupta, A. A. Schekochihin, M. Barnes, and W. D. Dorland, Phase-space entropy cascade and irreversibility of stochastic heating in nearly collisionless plasma turbulence, *Phys. Rev. E* **109**, 065210 (2024), arXiv:2310.18211 [physics.plasm-ph].
- [32] M. L. Nastac, R. J. Ewart, J. Juno, M. Barnes, and A. A. Schekochihin, Universal fluctuation spectrum in Vlasov-Poisson turbulence, forthcoming (2024).
- [33] S. Konrad and M. Bartelmann, Kinetic field theory for cosmic structure formation, *Rivista Nuovo Cimento* **45**, 737 (2022), arXiv:2202.11077 [astro-ph.CO].
- [34] Y. B. Zel'dovich, Gravitational instability: an approximate theory for large density perturbations., *A&A* **500**, 13 (1970).
- [35] S. F. Shandarin and Y. B. Zel'dovich, The large-scale structure of the universe: Turbulence, intermittency, structures in a self-gravitating medium, *Rev. Mod. Phys.* **61**, 185 (1989).
- [36] J. H. Jeans, The stability of a spherical nebula, *Phil. Trans. Roy. Soc. London A* **199**, 1 (1902).
- [37] J. Juno, A. Hakim, J. TenBarge, E. Shi, and W. Dorland, Discontinuous galerkin algorithms for fully kinetic plasmas, *J. Comput. Phys.* **353**, 110 (2018).
- [38] U. Frisch, *Turbulence: The Legacy of A. N. Kolmogorov* (Cambridge University Press, Cambridge, 1995).
- [39] V. I. Arnold, V. V. Kozlov, and A. I. Neishtadt, *Mathematical Aspects of Classical and Celestial Mechanics*, 3rd ed., Encyclopaedia of Mathematical Sciences, Vol. 3 (Springer-Verlag, Berlin, 2006) [Dynamical systems. III], Translated from the Russian original by E. Khukhro.
- [40] E. Bertschinger and B. Jain, Gravitational instability of cold matter, *ApJ* **431**, 486 (1994), arXiv:astro-ph/9307033 [astro-ph].
- [41] M. Hénon, Numerical experiments on the stability of spherical stellar systems, *A&A* **24**, 229 (1973).
- [42] J. A. Fillmore and P. Goldreich, Self-similar gravitational collapse in an expanding universe, *ApJ* **281**, 1 (1984).
- [43] E. Bertschinger, Self-similar secondary infall and accretion in an Einstein-de Sitter universe, *ApJS* **58**, 39 (1985).
- [44] J. Barnes, J. Goodman, and P. Hut, Dynamical instabilities in spherical stellar systems, *ApJ* **300**, 112 (1986).
- [45] J. Binney, Discreteness effects in cosmological N-body simulations, *MNRAS* **350**, 939 (2004), arXiv:astro-ph/0311155 [astro-ph].
- [46] S. Colombi and J. Touma, Vlasov-Poisson in 1D: waterbags, *MNRAS* **441**, 2414 (2014).
- [47] M. White, The Zel'dovich approximation, *MNRAS* **439**, 3630 (2014), arXiv:1401.5466 [astro-ph.CO].
- [48] G. K. Batchelor, Small-scale variation of convected quantities like temperature in turbulent fluid. Part 1. General discussion and the case of small conductivity, *J. Fluid Mech.* **5**, 113 (1959).
- [49] K. R. Sreenivasan, Turbulent mixing: A perspective, *PNAS* **116**, 18175 (2019).
- [50] P. Goldreich and S. Sridhar, Toward a theory of interstellar turbulence. II. Strong Alfvénic turbulence, *ApJ* **438**, 763 (1995).
- [51] S. V. Nazarenko and A. A. Schekochihin, Critical balance in magnetohydrodynamic, rotating and stratified turbulence: towards a universal scaling conjecture, *J. Fluid Mech.* **677**, 134 (2011), arXiv:0904.3488 [physics.flu-dyn].
- [52] A. A. Schekochihin, MHD turbulence: a biased review, *J. Plasma Phys.* **88**, 155880501 (2022), arXiv:2010.00699 [physics.plasm-ph].
- [53] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time*, Cambridge Monographs on Math-



- ematical Physics (Cambridge University Press, Cambridge, 1973).
- [54] F. Fabis, E. Kozlikin, R. Lilow, and M. Bartelmann, Kinetic field theory: exact free evolution of Gaussian phase-space correlations, *J. Stat. Mech.* **4**, 043214 (2018), arXiv:1710.01611 [cond-mat.stat-mech].
- [55] R. Lilow, F. Fabis, E. Kozlikin, C. Viermann, and M. Bartelmann, Resummed Kinetic Field Theory: general formalism and linear structure growth from Newtonian particle dynamics, *J. Cosmology Astropart. Phys.* **2019**, 001 (2019), arXiv:1809.06942 [astro-ph.CO].
- [56] S. Konrad, Y. B. Ginat, and M. Bartelmann, On the asymptotic behaviour of cosmic density-fluctuation power spectra of cold dark matter, arXiv e-prints, arXiv:2202.08059 (2022), arXiv:2202.08059 [astro-ph.CO].
- [57] N. Bleistein and R. A. Handelsman, *Asymptotic Expansions of Integrals*, 2nd ed. (Dover Publications, Inc., New York, 1986).
- [58] E. W. Kolb and M. S. Turner, *The Early Universe* (CRC Press, Taylor & Francis, London, 1990).
- [59] R. J. Ewart, M. L. Nastac, P. J. Bilbao, T. Silva, L. O. Silva, and A. A. Schekochihin, Relaxation to universal non-Maxwellian equilibria in a collisionless plasma, arXiv e-prints, arXiv:2409.01742 (2024), arXiv:2409.01742 [physics.plasm-ph].
- [60] W. H. Press and P. Schechter, Formation of galaxies and clusters of galaxies by self-similar gravitational condensation, *ApJ* **187**, 425 (1974).
- [61] C.-P. Ma and J. N. Fry, Deriving the nonlinear cosmological power spectrum and bispectrum from analytic dark matter halo profiles and mass functions, *ApJ* **543**, 503 (2000), arXiv:astro-ph/0003343 [astro-ph].
- [62] A. Kolmogorov, The local structure of turbulence in incompressible viscous fluid for very large Reynolds' numbers, *Akad. Nauk SSSR Doklady* **30**, 301 (1941).
- [63] A. Hakim, M. Francisquez, J. Juno, and G. W. Hammett, Conservative discontinuous Galerkin schemes for nonlinear Dougherty-Fokker-Planck collision operators, *J. Plasma Phys.* **86**, 905860403 (2020), arXiv:1903.08062 [physics.comp-ph].
- [64] S. Konrad and M. Bartelmann, On the asymptotic behaviour of cosmic density-fluctuation power spectra, *MNRAS* **515**, 2578 (2022), arXiv:2110.07427 [astro-ph.CO].
- [65] S. D. M. White, Prompt cusp formation from the gravitational collapse of peaks in the initial cosmological density field, *MNRAS* **517**, L46 (2022), arXiv:2207.13565 [astro-ph.CO].
- [66] M. Davis and P. J. E. Peebles, On the integration of the BBGKY equations for the development of strongly nonlinear clustering in an expanding universe, *ApJS* **34**, 425 (1977).

## Appendix A: Simulation methods

Here we describe the details on the Vlasov-Poisson simulations performed in this paper, whose results were shown in §II and figures 1-5. The simulations were conducted using the Gkeyll code<sup>12</sup>, which is an Eulerian solver in phase-space, originally designed for the Vlasov-Maxwell system. Gkeyll uses a discontinuous Galerkin algorithm to estimate distribution functions, which conserves energy exactly [37]. We set the vacuum permittivity  $\varepsilon_0 = -1$ , which turns electrostatic interactions into gravitational ones—and, for unit particle charges it is equivalent to choosing units such that  $G = 1/(4\pi)$ . The simulations we performed were 1D in position space (so the phase-space is 2D), with periodic spatial boundary conditions; the units of length were such that the box-size was  $L = 2\pi$ . We always assume that the particle mass is unity, which implies that the system's total mass is normalised to  $M = 2\pi$ . These three choices specify the units for the simulations, and are equivalent to choosing  $\tau_0 = k_0 = v_0 = 1$  (these are defined in §II A). Thus, the outer scale  $k_{\text{nl}} \sim k_0$  is of order unity.

### 1. Initial conditions

We simulated a single-species distribution function, with an initial condition (4), where we chose  $v_{\text{th}} = 0.005v_0$  or  $0.01v_0$ , and

$$\rho_{\text{in}}(x) \frac{L}{M} = 1 - \sum_{n=1}^5 \frac{a_n}{k_n} \sin(k_n x + \phi_n) \quad (\text{A1})$$

$$u_{\text{in}}(x) = v_0 \sum_{n=1}^5 b_n \cos(k_n x + \phi_n), \quad (\text{A2})$$

where  $\phi_n \in [0, 2\pi]$  are random phases,  $\frac{a_n}{k_0} \sim U[0, 0.2]$  and  $k_n = k_0 U_n$ , with  $U_n \in \{1, 2, \dots, 10\}$ —a uniformly-distributed random integer. For runs with one initial stream (figures 5 and 7),  $b_n = a_n/k_0$ , while for runs with three initial streams (figure 1), we set  $b_n = 0.05a_n/k_0$  and tripled the above initial conditions by shifting  $u_{\text{in}}(x)$  by  $\pm 2v_0$

<sup>12</sup> <https://gkeyll.readthedocs.io/>

and duplicating for two additional streams. We ensured that the resolution was sufficiently fine that a Maxwellian of width  $v_{\text{th}}$  would still be resolved: we used an  $N_x \times N_v$  phase-space grid with  $N_x = N_v = 4032$  for the single-stream initial condition, and  $N_x = N_v = 7680$  for the multi-stream one.

## 2. Collisions

One would like to run collisionless simulations, but the finite resolution of the code induces an unavoidable effective collisionality due to the grid. Real dark matter also has a finite number of particle—and finite- $N$  effects also induce a effective collisionality. We have also added a weak collision operator  $\left(\frac{\partial f}{\partial t}\right)_c \propto \nu$ , where  $\nu$  is the collision frequency to improve convergence. The collision operator used was the Dougherty (LBO) collision operator—a type of Fokker-Planck operator in velocity-space, *viz.*

$$\left(\frac{\partial f}{\partial t}\right)_c = \nu \frac{\partial}{\partial v} \left[ (v - u) f + v_i^2 \frac{\partial f}{\partial v} \right], \quad (\text{A3})$$

where  $u(x)$  is first velocity moment of  $f$ . Details (and the definition of  $v_i$ ) can be found in [63] which describes its implementation in Gkeyll. The inevitability of some collisions—whether due to the grid, finite- $N$  effects, or a collision operator—implies the existence of another scale in the problem. We used a Dorland number of  $\text{Do}^{-1} \equiv \nu\tau_0 = 10^{-6}$  for all runs, save for figure 7, where  $\text{Do} = 10^5$ . The collision time scale is [31, 63]

$$\tau_\nu \equiv \frac{1}{\nu s^2 v_{\text{rms}}^2}, \quad (\text{A4})$$

where  $v_{\text{rms}}^2$  is the second velocity cumulant. A relevant velocity scale is therefore

$$u_\nu = \sqrt{\nu\tau_c} v_{\text{rms}}, \quad (\text{A5})$$

where  $\tau_c$  is described in §II, which by critical balance (see §II C) also gives a collision length

$$l_\nu = \sqrt{\nu\tau_c} L \sim \sqrt{\nu\tau_c} k_{\text{nl}}^{-1}. \quad (\text{A6})$$

For a sufficiently small  $\nu$ , one can have<sup>13</sup>

$$k_{\text{nl}} \ll k \ll k_c(v_{\text{th}}^{-1}) \ll l_\nu^{-1} \quad (\text{A7})$$

$$v_{\text{rms}}^{-1} \ll s \ll v_{\text{th}}^{-1} \ll u_\nu^{-1}, \quad (\text{A8})$$

where  $k_c(s)$  is defined by equation (20). While this hierarchy justifies neglecting collisions in this paper, as phase-space structure cascades to ever smaller scales by (15), the collisional scales, where structure dissipates, are eventually reached [31, 32]. Likewise, in reality a shot-noise floor  $\hat{F} \approx V/N$  is eventually reached, too.

## 3. Results

In the set-up described above, our expectation for the phase-space power-spectrum  $\hat{F}(k, s)$  in equation (37) becomes

$$\hat{F}(k, s) \sim \begin{cases} F_1 k^{-1}, & \text{if } \max\{k_{\text{nl}}, k_c(s)\} \ll k \ll \min\{k_c(v_{\text{th}}^{-1}), l_\nu^{-1}\} \\ F_2 s^{-1}, & \text{if } \max\{v_{\text{rms}}^{-1}, s_c(k)\} \ll s \ll \min\{v_{\text{th}}^{-1}, u_\nu^{-1}\}. \end{cases} \quad (\text{A9})$$

We are not concerned in the paper with asymptotics of the power spectrum  $\hat{F}$  (defined in (10)) at values of  $s$  and  $k$  above  $v_{\text{th}}^{-1}$  (respectively  $k_c(v_{\text{th}}^{-1})$ )—but still less than the collision scale. But if one extrapolates the findings of [32]—who explored turbulence in the régime where  $v_{\text{th}}$  is of order unity in electrostatic plasmas—from electrostatics

<sup>13</sup> In reality, it might be that  $u_\nu \gtrsim v_{\text{th}}$ , for both parameters depend sensitively on the nature of the dark-matter constituent parti-

cle(s). But as  $1/s$  is larger than both, none of the conclusions of the this paper are invalidated by this.

to gravity (ignoring any possible differences), while noting that above  $v_{\text{th}}^{-1}$  gravitational collapse halts by equation (26)—then one should have, for  $v_{\text{th}} \gg u_\nu$ ,

$$\hat{F}(k, s) \sim \begin{cases} F_1 k^{-1}, & \text{if } \max\{k_{\text{nl}}, k_c(s)\} \ll k \ll k_c(v_{\text{th}}^{-1}) \\ F_2 s^{-1}, & \text{if } \max\{v_{\text{rms}}^{-1}, s_c(k)\} \ll s \ll v_{\text{th}}^{-1} \\ F_3 k^{-2}, & \text{if } \max\{k_{\text{nl}}, k_c(s), k_c(v_{\text{th}}^{-1})\} \ll k \ll l_\nu^{-1} \\ F_4 s^{-2}, & \text{if } \max\{v_{\text{rms}}^{-1}, s_c(k), v_{\text{th}}^{-1}\} \ll s \ll u_\nu^{-1}; \end{cases} \quad (\text{A10})$$

the spectrum is truncated exponentially at  $k > l_\nu^{-1}$  or  $s > u_\nu^{-1}$ .

Figure 1 shows an example time-evolution of the system, with  $v_{\text{th}} = 0.005v_0$ , and  $\text{Do} = 10^6$ . This corresponds to  $u_\nu \sim 10^{-3}v_0$ , whence the hierarchy of scales is satisfied. The grid  $s$  is  $N_x = N_v = 7680$ , and the velocity box-size is  $12v_0$ , so the Nyquist velocity-scale is  $s_{\text{Ny}}v_0 \approx 2011$ . This figure shows that the stream collapses quickly, by rotating and twisting in phase-space. In figure 3 we show the evolution of the power-spectra, and some snapshots (beginning, middle and end) are given in 6. Figure 7 is similar to figure 1 but shows the time-evolution of one initial stream, with  $\text{Do} = 10^5$ . It has the same qualitative features.

Because of numerical noise,  $f$  can become slightly negative in very confined areas. This is because the discontinuous Galerkin algorithm that Gkeyll uses is not positivity-preserving, and errors arise from over-shooting due to large derivatives. That is why we have added a small collision operator, which smooths large velocity derivatives, and alleviates this effect. In any case, these regions do not cause the simulation to become unstable, because they are isolated and  $|f|$  is still very small there; the total mass occupied by negative  $f$  is  $\iint f\Theta(-f)dxdv < 10^{-5}M$  for the simulation in figure 1 and  $< 0.02M$  for figure 7. Therefore, this does not invalidate any of our conclusions in this paper, and we have set  $f \mapsto f\Theta(f)$  for the purpose of plotting figures 1,5 and 7.

## Appendix B: Non-smooth gravitational-field fluctuations

Let us prove that a Hölder exponent  $\lambda < 1$  is inconsistent with critical balance in conjunction with Poisson's equation. By that equation and the Paley-Wiener theorem (which works for  $\lambda < 1$ ),  $P(k) \sim k^{2-d}\delta g_r^2$ .  $\delta g_r \sim k^{-\lambda}$ , so  $P(k) \sim k^{2-d-2\lambda}$ . If  $P(k) \sim F_1 k^\gamma$ , then  $\gamma = 2 - 2\lambda - d$ . By critical balance, this means that  $s_c(k) \propto k^{(1+\lambda)/2}$ , so the collapse time is  $\tau_c \sim k^{(\lambda-1)/2}$ . As the (integrated) source is  $\varepsilon s_c^d$ , the  $C_2$  budget equation (31) then implies that  $\gamma = -d$ , whence  $\lambda = 1$ , which is a contradiction. A proof that the Vlasov-Poisson system and critical balance are incompatible with  $\lambda < 1$  for a constant source is given in §III.F of [32].

## Appendix C: Estimates for the source

In this appendix we will discuss  $\iint \hat{S}$  in more detail, and give an argument that it scales like  $\varepsilon s_c(k)^d$ , to complement §II E.

### 1. Integrated source

Consider the integral (25). As remarked in §II E 2, by dimensional analysis, the magnitude of the contribution of each is proportional to the collapse time, which is nothing but the critical-balance time (i.e. independent of  $k$  and  $s$ ). The sign of the proportionality coefficient can, *prima facie*, depend on scale, but in §C 2 below we argue that it doesn't. Let us present a dimensional argument that each scale gives a scale-independent contribution: by analogy with equation (25), the integrated source is

$$\mathcal{F}^{\mathbf{k}} = (2\pi)^{-2d} \int^{1/r} d^d k \int d^d s \hat{S} \sim \int_r^\infty d^d x \int d^d v \left\langle [\mathbf{g}(1) - \mathbf{g}(2)] \cdot \frac{\partial}{\partial \mathbf{v}} (f(2)\bar{f}(1)) \right\rangle + (1 \leftrightarrow 2). \quad (\text{C1})$$

Crucially, the product  $[\mathbf{g}(1) - \mathbf{g}(2)] \cdot \partial/\partial \mathbf{v}$  is the (inverse) collapse time, proportional to  $1/\tau_c$ . Furthermore, its sign must also be independent of scale: as long as  $\max\{s, s_c(k)\}v_{\text{th}} \ll 1$ , nothing in equation (26) depends on scale (see §C 2 below for further justification that its sign is independent of scale) so the proportionality constant is truly constant.

To estimate the source, we use the following estimates:  $|\mathbf{g}(1) - \mathbf{g}(2)|^2 \sim G^2 \hat{F}(k, 0)k^{d-2}s_c^d(k)$ ,  $f(2)^2 \sim \hat{F}(\mathbf{k}, \mathbf{s})k^d s^d$ ,  $\bar{f}(1)^2 \sim \hat{F}(0, s)s^d k_c(s)^d$ , where  $k \sim 1/x$ ,  $s \sim 1/v$ . Hence, for  $s_c(r^{-1}) \ll 1/r$ , by Poisson's equation and equation (16)

(cf. equation (30) below too)

$$\left\langle \tilde{\mathbf{g}} \cdot \frac{\partial}{\partial \mathbf{v}} (f(2)\bar{f}(1)) \right\rangle \sim G s^{3d/2+1} \hat{F}(k, 0) k^{3d/2-1} \sqrt{\hat{F}(0, 0)}. \quad (\text{C2})$$

Integrating over  $v$  up to the critical-balance line gives a factor of  $s^{-d}$ , leaving<sup>14</sup>

$$G s_c^{d/2+1} \hat{F}(k, 0) k^{3d/2-1} \sqrt{\hat{F}(0, 0)}; \quad (\text{C3})$$

after integration over  $x$  on scales larger than  $r$ , if one parameterises  $\hat{F}(k, 0) \sim F_1 k^\gamma$ , one obtains

$$\mathcal{F}^{\mathbf{k}} \sim F_1 \tau_c^{d+1} G \sqrt{\hat{F}(0, 0)} r^{-\gamma-2d}. \quad (\text{C4})$$

If  $\gamma = -d$  then this is  $\propto r^{-d}$ . A similar calculation yields the analogous result  $\mathcal{F}^{\mathbf{s}} \sim \varepsilon s^d$ .

Fortunately, we know that the density power spectrum of cold dark matter assumes its  $k^{-d}$  shape very quickly initially [15, 56, 64], when matter is still in a régime where the Zel'dovich approximation is accurate, whence one can safely take initially  $\gamma = -d$ . We will show in §II F below that if indeed  $\iint \hat{S} \sim \varepsilon s^d$ , then this gives a  $k^{-d}$  (or  $s^{-d}$ ) power spectrum. This power spectrum in turn keeps producing a source with just the right scaling, to ensure that  $\hat{F}$  retains the same  $k^{-d}$  power-law. This is an *a posteriori* justification for  $\iint \hat{S} \sim \varepsilon s^d$ .

## 2. Sign

To show that the sign is indeed independent of scale, it will be easier to work in real space. So far, the entire analysis was done in  $d$  spatial dimensions for generality, but it is easiest to estimate the source term in  $d = 1$ ; this generalises to  $d$  dimensions.

First, we will show that locally, gravitational collapse occurs about every local maximum of  $f$  that is also in an over-dense phase-space region. Let  $w_0 = (x_0, v_0)$  be a local maximum of  $f$ . In the limit  $v_{\text{th}} \rightarrow 0$ , there exists a phase-space neighbourhood of  $w_0$ , where  $f = \rho(x)h(v - u(x)) + O(v_{\text{th}})$ , where  $\rho(x) = \rho_0 - \rho_2(x - x_0)^2$ ,  $u(x) = v_0 + v_1 x$ , and  $h$  is a sharply-peaked function (around 0), as is visible from figure 1 or 7, and was explained in §II A. Besides, the gravitational field can be decomposed into a sum of terms: one,  $g_{\text{in}}$ , arising from the stream that passes through  $w_0$ , and the external field,  $g_{\text{ext}}$ :  $g \equiv g_{\text{in}} + g_{\text{ext}}$ . Without loss of generality we can assume that  $x_0 = 0$ ,  $v_0 = 0$ ,  $g(x_0) = 0$ ,  $v_1 > 0$ ,  $\rho_2 > 0$ , whence by Poisson's equation

$$g_{\text{in}}(x) = -4\pi G x (\rho_0 - \rho_2 x^2/3). \quad (\text{C5})$$

As we are considering a small neighbourhood of  $w_0$ , the external field is just given by the tidal approximation around  $w_0$ , i.e.  $g_{\text{ext}} = -g_1 x$ , where  $g_1 > 0$  because we have assumed that  $x_0$  is an over-dense region.

The time-evolution of this system is just that of an anharmonic oscillator (cf. [65]), i.e. it is a rotation in phase-space whose frequency is  $\omega^2 = 4\pi G \rho_0 + g_1$ , plus a correction that depends on energy due to the anharmonicity,  $\propto G \rho_2 (\omega^2 x^2 + v^2)$ , which is small near  $w_0 = (0, 0)$ .<sup>15</sup> This is by definition the process of gravitational collapse, which is what we had to show.<sup>16</sup>  $\rho_0$  and  $\rho_2$  in general change with time, but that does not modify the conclusions qualitatively. Additionally,  $f$  increases along the stream towards  $w_0$ .

Now let us proceed to describe the source term, i.e. the integrand of equation (25), whose sign we wish to estimate. By equation (16) it is

$$\left\langle (g(1) - g(2)) \frac{\partial (\bar{f}(1)\delta f(2))}{\partial [v(1) - v(2)]} \right\rangle + (1 \leftrightarrow 2), \quad (\text{C6})$$

where the shorthand notation is  $(n) = (x_n, v_n)$ . We are interested in a particular scale, where the distance between  $x_1$  and  $x_2$  is  $x \sim k^{-1}$  and the distance between  $v_1$  and  $v_2$  is  $v \sim s^{-1}$ . Thus, we consider two points on a collapsing eddy—a

<sup>14</sup> Crucially, the  $sv_{\text{th}} \ll 1$  limit is needed here, otherwise the integral would not have accumulated contributions up to  $s$ , but rather only up to  $v_{\text{th}}^{-1}$ . At larger  $s$  the behaviour changes entirely [cf. 32].

<sup>15</sup> The differential phase-space rotation generates small-scale structure and a cascade to smaller scales; cf. §III D.

<sup>16</sup> In generalising this to  $d > 1$ , an over-dense region does not guarantee that the tidal forces will always be directed towards  $w_0$ . However, there will be at least one direction where they will be. Thus, collapse will occur at least along one direction, whereupon eventually  $\rho_0$  will become sufficiently large.

single stream which is collapsing around some local maximum of  $f$ —whose phase-space  $s$  is  $xv \sim (ks)^{-1}$ . See for example the  $t = 12\tau_0$  panel of figure 1 or the  $t = 8\tau_0$  panel of figure 7, which display larger-scale collapsing structures which in turn contain similar smaller ones. By ergodicity, the ensemble average includes shifting the centre-of-mass of the eddy and averaging over its position. Therefore, without loss of generality, we assume that the eddy is centred at the origin of phase-space. As mentioned above, on small scales the eddy comprises a collapsing stream, whence  $g(x)$  points inwards, i.e.  $\text{sgn}(g(x)) = -\text{sgn}(x)$ . Additionally, the derivative  $\partial(\bar{f}(1)f(2))/\partial v$  has two contributions: one,  $\bar{f}(1)\frac{\partial f(2)}{\partial v}$  is correlated with the direction of  $g$ : we choose local phase-space co-ordinates  $w_{\parallel}$  and  $w_{\perp}$ , where  $w_{\parallel}$  is parallel to the stream, and  $w_{\perp}$  is perpendicular to it. The phase-space gradient of  $f$  is mostly along  $w_{\perp}$ , of course, but from symmetry reasons, its velocity component changes sign for in-coming and out-flowing particles, and thus changes sign with respect to  $g$ , so its contribution to the any ensemble-average vanishes.

The other contribution, from  $w_{\parallel}$ , does not: imagine that point (1) is at  $x_1 > 0$ ,  $v_1 > 0$ , while point (2) has  $x_2 < 0$ ,  $v_2 < 0$ . Then  $g(1) - g(2) < 0$ , and the  $w_{\parallel}$  component of  $\frac{\partial f(2)}{\partial v_2}$  is negative, so  $\frac{\partial f(2)}{\partial v} > 0$ . This point, therefore, gives a negative contribution to (C6). Now imagine that  $x_1 > 0$ ,  $v_1 < 0$ , and  $x_2 < 0$ , but  $v_2 < 0$ . Then  $g(1) - g(2) < 0$  still. The  $w_{\parallel}$  component of  $\frac{\partial f(2)}{\partial v_2}$  is again negative, because the stream gradient increases in a clockwise fashion, whence  $\frac{\partial f(2)}{\partial v} > 0$ . Thus, all options for the positions of points (1) and (2) give the same sign, if they are in the same eddy. By the symmetry between the two particles, all other configurations have the same sign.

On the other hand, the second term,  $f(2)\frac{\partial \bar{f}(1)}{\partial v}$  changes sign rapidly, because  $f$  describes a stream, so  $\bar{f}$  is a collection of sharp peaks in velocity-space, with a rapidly-oscillating derivative, which comes from the phase-space gradient of  $f$  along  $w_{\perp}$ . There is also a contribution from the direction along  $w_{\parallel}$ , which leads to an overall increase in  $\bar{f}$  towards the eddy's centre-of-mass velocity. Now, this  $w_{\parallel}$ -component also fluctuates with respect to the sign of  $g$ , for it is positive for in-falling particles, while it is negative for particles which are faster than the eddy's centre-of-mass.

Hence, the only non-fluctuating contribution to the source term, from an eddy of  $s (ks)^{-1}$ , has a fixed sign. The sign, moreover, is independent of the scale, and the fluctuating parts will drop out when averaging and integrating over the various scales to get equation (25). Consequently, the source term does have a definite sign, and the integrand in equation (25) adds up coherently.

#### Appendix D: Einstein-de-Sitter universe

Although equation (37) is valid for  $\Lambda$ CDM, and in general any cosmology where there exists a separation of scales between  $\mathcal{H}^{-1}$  and the non-linear scale  $k_{\text{nl}}$ , in the special case of an Einstein-de-Sitter (EdS) cosmology, there exists a scale-invariance [e.g., 66] which provides additional insight. We focus in this appendix on  $d = 3$  for simplicity. This scale-invariance involves a (conformal) time scale,

$$\tau_{\text{sc}} \equiv \frac{2}{H_0} \left( \frac{\pi}{2\Omega_m^{1/2} H_0 t_0} \right)^{1/3} \left( \frac{3}{5\delta_{\text{rms}}(k)} \right)^{1/2} \quad (\text{D1})$$

(which is just the collapse conformal time in the standard spherical-collapse model), with  $H_0$  and  $t_0$  referring to the present-day Hubble constant and cosmic time, and  $\delta_{\text{rms}}(k)$  is the root-mean-square amplitude of the initial dark-matter density fluctuation at scale  $k$ , normalised by the linear growth rate at the initial time. If the initial over-density power-spectrum is  $P_{\text{in}}(k) = P_0(k/k_p)^n$  for some pivot scale  $k_p$ , then

$$\tau_{\text{sc}} = \frac{2\sqrt{3}}{\sqrt{5}H_0} \left( \frac{\pi}{4\Omega_m^{1/2} H_0 t_0} \right)^{1/3} \left( \frac{k_p^n}{P_0 k^{3+n}} \right)^{1/4}. \quad (\text{D2})$$

Requiring a dominant balance between the collapse time-scale  $\tau_{\text{sc}}$  and the linear  $\tau_p$  implies

$$s_c(k) \propto k^{(1-n)/4}. \quad (\text{D3})$$

Consequently, the analogue of equation (33) here would imply

$$k^{4+\gamma} \sim s_c(k) \quad (\text{D4})$$

or  $\gamma = -(15+n)/4$ . In a steady state, the power spectrum must be invariant under the collapse process, and so in a self-sustaining scenario,  $\gamma = n$ , which is solved by  $n = -3$ .<sup>17</sup>

Note, that the collapse time in equation (D1) is precisely the collapse-time  $\tau_c$  in §IIE for  $n = -3$ .

<sup>17</sup> One may wish to relax the steady-state assumption, and set  $n =$

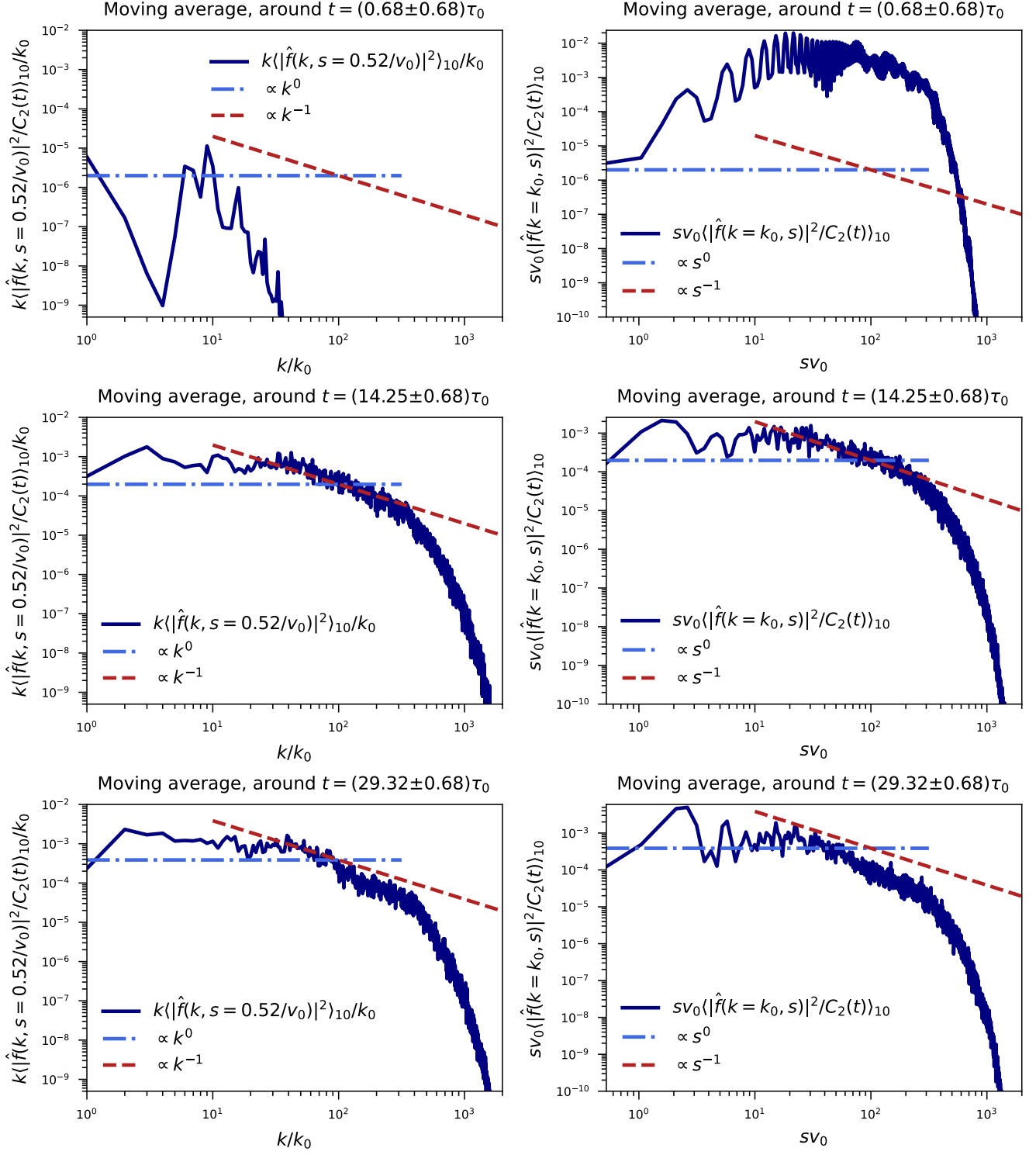


FIG. 6. Same as figure 3, but highlighting three times for clarity. See text for details.

$n_s - 4$ ; this yields  $\gamma = -(11 + n_s)/4$ , almost indistinguishable from  $-3$ . However, as the calculation in this paper is in the fully non-linear régime, it is appropriate to set  $n = -3$  here, for it is

already established that this asymptotic scaling develops already in the Zel'dovich approximation [56, 64].

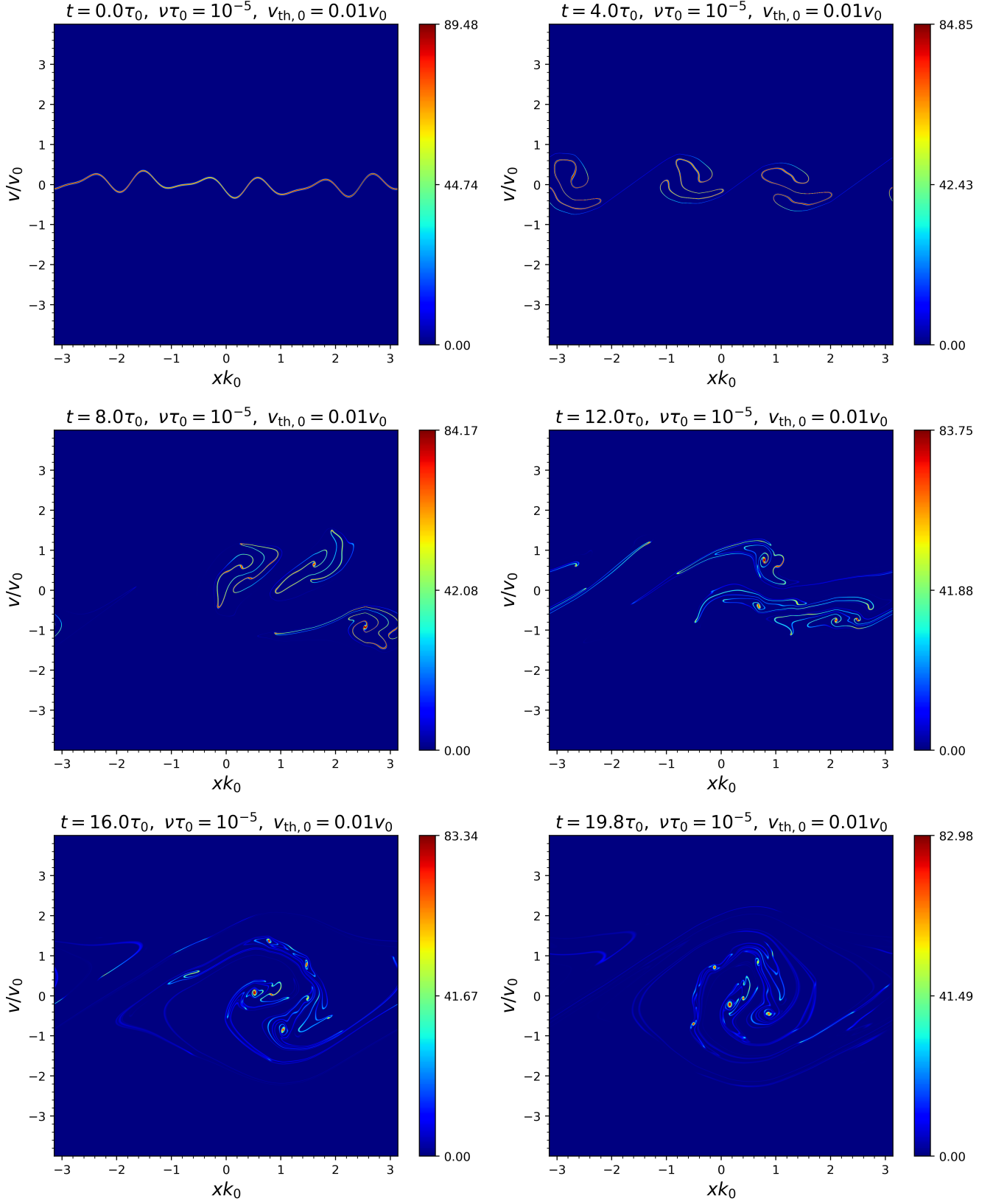


FIG. 7. Colour plots of the distribution function showing the time evolution of a cold stream with  $v_{th} = 0.01v_0$ , and  $Do = 10^5$ . See text for details.

### Appendix E: Initial momentum-correlations

The purpose of this appendix is to prove that  $\Sigma \sim Aq^2$  as  $q \rightarrow 0$ , and that  $\Sigma \mathbf{a} \sim O(q)$ , for §III. What follows holds for  $v_{\text{th}} = 0$ . Let us write

$$\{p\}^T (C_{pp}^N)^{-1} \{p\} \equiv \mathbf{p}^T \Sigma^{-1}(\mathbf{q}) \mathbf{p} - 2\mathbf{a} \left( \mathbf{q}, \mathbf{Q}, \mathbf{P}, \{q\}_3^N, \{p\}_3^N \right) \cdot \mathbf{p} - 2B \left( \mathbf{q}, \mathbf{Q}, \mathbf{P}, \{q\}_3^N, \{p\}_3^N \right), \quad (\text{E1})$$

as in equation (40). Since  $2\mathbf{p}_1 = \mathbf{p} + 2\mathbf{P}$  and  $2\mathbf{p}_2 = 2\mathbf{P} - \mathbf{p}$ , the part involving  $\mathbf{a}$  comes solely from the cross-correlations between  $\mathbf{p}_{1,2}$  and  $\mathbf{p}_3, \dots, \mathbf{p}_N$ , *viz.*

$$\begin{aligned} \mathbf{p}_1^T (C_{pp}^N)^{-1}_{1n} \mathbf{p}_n + \mathbf{p}_2^T (C_{pp}^N)^{-1}_{2n} \mathbf{p}_n &= \mathbf{P} [(C_{pp}^N)^{-1}_{1n} + (C_{pp}^N)^{-1}_{2n}] \mathbf{p}_n + \frac{1}{2} \mathbf{p} [(C_{pp}^N)^{-1}_{1n} - (C_{pp}^N)^{-1}_{2n}] \mathbf{p}_n \\ &+ \mathbf{P} [(C_{pp}^N)^{-1}_{11} - (C_{pp}^N)^{-1}_{22}] \mathbf{p} + 2\mathbf{P} [(C_{pp}^N)^{-1}_{11} + (C_{pp}^N)^{-1}_{21} + (C_{pp}^N)^{-1}_{12} + (C_{pp}^N)^{-1}_{22}] \mathbf{P}, \end{aligned} \quad (\text{E2})$$

where  $(C_{pp}^N)^{-1}_{mn}$  specifies the part of  $(C_{pp}^N)^{-1}$  that pertains to particles  $m$  and  $n$ , and an Einstein summation convention is implied; this is because  $(C_{pp}^N)^{-1}_{21} = (C_{pp}^N)^{-1}_{12}$ .

Now,  $(C_{pp}^N)^{-1}_{mn}$  is a function of the positions of the particles only. When  $q$  is small,  $\mathbf{q}_1$  and  $\mathbf{q}_2$  both tend to  $\mathbf{Q}$ , and by the indistinguishability of the constituent particles,

$$(C_{pp}^N)^{-1}_{1n}(\mathbf{q}_1 = \mathbf{Q}, \mathbf{q}_2 = \mathbf{Q}, \mathbf{q}_3, \dots, \mathbf{q}_N) = (C_{pp}^N)^{-1}_{2n}(\mathbf{q}_1 = \mathbf{Q}, \mathbf{q}_2 = \mathbf{Q}, \mathbf{q}_3, \dots, \mathbf{q}_N) \quad (\text{E3})$$

$$(C_{pp}^N)^{-1}_{11}(\mathbf{q}_1 = \mathbf{Q}, \mathbf{q}_2 = \mathbf{Q}, \mathbf{q}_3, \dots, \mathbf{q}_N) = (C_{pp}^N)^{-1}_{22}(\mathbf{q}_1 = \mathbf{Q}, \mathbf{q}_2 = \mathbf{Q}, \mathbf{q}_3, \dots, \mathbf{q}_N) \quad (\text{E4})$$

Consequently, for small  $q$  and finite  $\mathbf{Q}, \dots, \mathbf{q}_N, \mathbf{P}, \dots, \mathbf{p}_N$ , we have

$$[(C_{pp}^N)^{-1}_{1n} - (C_{pp}^N)^{-1}_{2n}] \sim O(q) \quad (\text{E5})$$

$$[(C_{pp}^N)^{-1}_{11} - (C_{pp}^N)^{-1}_{22}] \sim O(q). \quad (\text{E6})$$

In the limit  $q \rightarrow 0$ , therefore, the matrix  $(C_{pp}^N)^{-1}$ , *qua* a linear operator in momentum-space, splits into a block matrix in the basis  $(\mathbf{p}, \mathbf{P}, \mathbf{p}_3, \dots, \mathbf{p}_N)$ , with one block for  $\mathbf{p}$  and another for  $\mathbf{P}, \mathbf{p}_3, \dots, \mathbf{p}_N$ . Hence its inverse  $C_{pp}^N$  is also a block matrix in this basis. Concretely,

$$C_{pp}^N = \left( \begin{array}{c|cc} (C_{pp}^N)_{11} + (C_{pp}^N)_{22} - (C_{pp}^N)_{12} - (C_{pp}^N)_{21} & 0 & (C_{pp}^N)_{1n} - (C_{pp}^N)_{2n} \\ \hline 0 & (C_{pp}^N)_{11} + (C_{pp}^N)_{22} + (C_{pp}^N)_{12} + (C_{pp}^N)_{21} & (C_{pp}^N)_{1n} + (C_{pp}^N)_{2n} \\ (C_{pp}^N)_{1n} - (C_{pp}^N)_{2n} & (C_{pp}^N)_{1n} + (C_{pp}^N)_{2n} & (C_{pp}^N)_{nm} \end{array} \right). \quad (\text{E7})$$

The matrix  $\Sigma^{-1}$  is the top-left block of  $(C_{pp}^N)^{-1}$ , *i.e.*

$$\Sigma = [(C_{pp}^N)_{11} + (C_{pp}^N)_{22} - (C_{pp}^N)_{12} - (C_{pp}^N)_{21}] - [(C_{pp}^N)_{1n} - (C_{pp}^N)_{2n}] (C_{pp}^N)^{-1}_{nm} [(C_{pp}^N)_{1m} - (C_{pp}^N)_{2m}] \quad (\text{E8})$$

This, in conjunction with results on the initial correlation sub-matrix of  $C_{pp}^N$  pertaining to particles 1 and 2 [33, 56, 64] to the effect that when  $v_{\text{th}} = 0$ ,  $(C_{pp}^N)_{11} + (C_{pp}^N)_{22} - (C_{pp}^N)_{12} - (C_{pp}^N)_{21} \sim O(q)^2$  (with the coefficient determined by the cosmology, which we take to be standard  $\Lambda$ CDM), implies that indeed  $\Sigma \sim Aq^2$  at small  $q$ . Besides

$$a_k \propto -(\Sigma^{-1})_{kn} [(C_{pp}^N)_{1n} - (C_{pp}^N)_{2n}] (C_{pp}^N)^{-1}_{nm} \mathbf{p}_m. \quad (\text{E9})$$

Additionally, the results of [33, §3.1.2], in conjunction with equation (E7), imply that the entire  $C_{pp}^N$  matrix tends to a constant as  $q \rightarrow \infty$ , which means that  $\Sigma$ ,  $\mathbf{a}$  and  $B$  become  $O(1)$  in that limit, too.