

Supplemental Material for *Quantifying Energy Conversion in Higher Order Phase Space Density Moments in Plasmas*

Paul A. Cassak* and M. Hasan Barbhuiya
*Department of Physics and Astronomy and the Center for KINETIC Plasma Physics
 West Virginia University, Morgantown, WV 26506, USA*

Haoming Liang
*Center for Space Plasma and Aeronomic Research
 University of Alabama in Huntsville, Huntsville, AL 35899, USA*

Matthew R. Argall
*Space Science Center, Institute for the Study of Earth, Oceans, and Space
 University of New Hampshire, Durham, NH 03824, USA*
 (Dated: January 20, 2023)

A. Derivation of Kinetic Entropy Evolution Equation

The evolution equation for the kinetic entropy density s_σ defined in Eq. (2) is obtained by taking its partial time derivative and eliminating $\partial f_\sigma / \partial t$ using the Boltzmann equation [1],

$$\frac{\partial f_\sigma}{\partial t} + \mathbf{v} \cdot \nabla f_\sigma + \frac{\mathbf{F}_\sigma}{m_\sigma} \cdot \nabla_v f_\sigma = C[f], \quad (\text{S.1})$$

where \mathbf{F}_σ is the sum of any body forces, ∇_v is the velocity space gradient operator, and $C[f]$ is the inter- and intra-species collision operator, yielding [2-4]

$$\frac{\partial s_\sigma}{\partial t} + \nabla \cdot \mathcal{J}_\sigma = \dot{s}_{\sigma, \text{coll}}, \quad (\text{S.2})$$

where \mathcal{J}_σ is the kinetic entropy density flux

$$\mathcal{J}_\sigma = -k_B \int \mathbf{v} f_\sigma \ln \left(\frac{f_\sigma \Delta^3 r_\sigma \Delta^3 v_\sigma}{N_\sigma} \right) d^3 v \quad (\text{S.3})$$

and $\dot{s}_{\sigma, \text{coll}}$ is the local time rate of change of kinetic entropy density through collisions,

$$\dot{s}_{\sigma, \text{coll}} = -k_B \int C[f] \ln \left(\frac{f_\sigma \Delta^3 r_\sigma \Delta^3 v_\sigma}{N_\sigma} \right) d^3 v. \quad (\text{S.4})$$

Note, there is no term containing body forces such as the electric and magnetic forces in Eq. (S.2) because the force term in Eq. (S.1) identically vanishes in deriving Eq. (S.2). An equivalent form of Eq. (S.2) comes from writing $\mathbf{v} = \mathbf{u}_\sigma + \mathbf{v}'_\sigma$ in Eq. (S.3), which implies $\mathcal{J}_\sigma = s_\sigma \mathbf{u}_\sigma + \mathcal{J}_{\sigma, \text{th}}$, where the thermal kinetic entropy density flux $\mathcal{J}_{\sigma, \text{th}}$ is defined as

$$\mathcal{J}_{\sigma, \text{th}} = -k_B \int \mathbf{v}'_\sigma f_\sigma \ln \left(\frac{f_\sigma \Delta^3 r_\sigma \Delta^3 v_\sigma}{N_\sigma} \right) d^3 v. \quad (\text{S.5})$$

Then, Eq. (S.2) becomes [3]

$$\frac{\partial s_\sigma}{\partial t} + \nabla \cdot (s_\sigma \mathbf{u}_\sigma + \mathcal{J}_{\sigma, \text{th}}) = \dot{s}_{\sigma, \text{coll}}. \quad (\text{S.6})$$

This equation is in the stationary (Eulerian) reference frame.

Here, we manipulate Eq. (S.6) to derive an evolution equation for kinetic entropy per particle s_σ / n_σ in a co-moving (Lagrangian) frame [Eq. (3)]. Using the convective derivative $d/dt = \partial/\partial t + \mathbf{u}_\sigma \cdot \nabla$ and dividing Eq. (S.6) by the density n_σ gives

$$\frac{1}{n_\sigma} \frac{ds_\sigma}{dt} + \frac{s_\sigma}{n_\sigma} (\nabla \cdot \mathbf{u}_\sigma) + \frac{\nabla \cdot \mathcal{J}_{\sigma, \text{th}}}{n_\sigma} = \frac{\dot{s}_{\sigma, \text{coll}}}{n_\sigma}. \quad (\text{S.7})$$

Using the continuity equation $dn_\sigma/dt = -n_\sigma \nabla \cdot \mathbf{u}_\sigma$ (since N_σ is assumed constant), we get

$$\frac{1}{n_\sigma} \frac{ds_\sigma}{dt} - \frac{s_\sigma}{n_\sigma^2} \frac{dn_\sigma}{dt} + \frac{\nabla \cdot \mathcal{J}_{\sigma, \text{th}}}{n_\sigma} = \frac{\dot{s}_{\sigma, \text{coll}}}{n_\sigma}. \quad (\text{S.8})$$

Finally, the two terms on the left are equal to $d(s_\sigma/n_\sigma)/dt$, which completes the derivation of Eq. (3).

We conclude this section with two important notes. First, Eq. (2) is the ‘‘Boltzmann’’ form of kinetic entropy density s_σ . For collisionless systems, any function of f_σ is conserved, so other entropies could be defined [5]. We choose the Boltzmann entropy because it reduces to the ideal fluid entropy density for a system in LTE and, for collisional systems, the total Boltzmann entropy $S_\sigma = \int s_\sigma d^3 r$ obeys an H-theorem (S_σ is non-decreasing in time) for a reasonably defined collision operator [6]. Neither need be the case for other entropies. The present analysis may be redone for other entropies for future work.

Second, the approach we use remains valid even if there is an entropy source in the Boltzmann equation beyond collisions, such as due to boundaries of a finite domain. Such sources can lead to non-conservation of total kinetic entropy $S_\sigma = \int s_\sigma d^3 r$ even in collisionless systems [7], but s_σ is local in space and time and therefore remains well-defined.

* Paul.Cassak@mail.wvu.edu

B. Derivation of Generalized Work Term

Here, we derive Eq. (5a). Dividing both sides of Eq. (4a) by n_σ and taking its total time derivative gives

$$\frac{d}{dt} \left(\frac{s_{\sigma p}}{n_\sigma} \right) = -\frac{k_B}{n_\sigma} \frac{dn_\sigma}{dt}. \quad (\text{S.9})$$

A brief derivation reveals this is equivalent to

$$\frac{d}{dt} \left(\frac{s_{\sigma p}}{n_\sigma} \right) = k_B n_\sigma \frac{d(1/n_\sigma)}{dt}. \quad (\text{S.10})$$

Defining $V_\sigma = 1/n_\sigma$ as the volume per particle, and using $\mathcal{P}_\sigma = n_\sigma k_B \mathcal{T}_\sigma$, the previous equation is equivalent to

$$\frac{d}{dt} \left(\frac{s_{\sigma p}}{n_\sigma} \right) = \frac{1}{\mathcal{T}_\sigma} \frac{d\mathcal{W}_\sigma}{dt}, \quad (\text{S.11})$$

where $d\mathcal{W}_\sigma = \mathcal{P}_\sigma dV_\sigma = \mathcal{P}_\sigma d(1/n_\sigma)$ is the non-LTE generalization of the work per particle done by the system.

To physically interpret this, note $s_{\sigma p}$ is associated with the number of permutations of particles in position space that produce the same macrostate without concern for their velocity [8]. The argument of the natural logarithm in $s_{\sigma p}/n_\sigma = -k_B \ln(n_\sigma \Delta^3 r_\sigma / N_\sigma)$ is always between 0 and 1, so $s_{\sigma p}/n_\sigma$ is non-negative and is a strictly decreasing function of n_σ . Thus, local compression ($d\mathcal{W}_\sigma = \mathcal{P}_\sigma dV_\sigma < 0$) increases n_σ and decreases $s_{\sigma p}/n_\sigma$ [*i.e.*, $d(s_{\sigma p}/n_\sigma)/dt < 0$], while local expansion ($d\mathcal{W}_\sigma = \mathcal{P}_\sigma dV_\sigma > 0$) decreases n_σ and increases $s_{\sigma p}/n_\sigma$ [*i.e.*, $d(s_{\sigma p}/n_\sigma)/dt > 0$].

C. Derivation of Generalized Energy Term

We next derive Eq. (5b). We decompose the velocity space kinetic entropy density $s_{\sigma v}$ in Eq. (4b) as $s_{\sigma v} = s_{\sigma v, \mathcal{E}} + s_{\sigma v, \text{rel}}$, where

$$s_{\sigma v, \mathcal{E}} = -k_B \int f_\sigma \ln \left(\frac{f_{\sigma M} \Delta^3 v_\sigma}{n_\sigma} \right) d^3 v, \quad (\text{S.12})$$

$$s_{\sigma v, \text{rel}} = -k_B \int f_\sigma \ln \left(\frac{f_\sigma}{f_{\sigma M}} \right) d^3 v, \quad (\text{S.13})$$

where $f_{\sigma M}$ is the Maxwellianized distribution of f_σ defined in Eq. (8). The relative entropy is related to the Kullback-Leibler divergence [9] from information theory which is a measure of the statistical difference between a two probability distributions, and has been extensively used in a variety of fields, such as statistical mechanics, applied mathematics, chemistry, biology, quantum information theory, and economics [2, 10–19]. Substituting Eq. (8) into Eq. (S.12) and carrying out straight-forward manipulations gives

$$\frac{s_{\sigma v, \mathcal{E}}}{n_\sigma} = \frac{3}{2} k_B \left[1 + \ln \left(\frac{2\pi k_B \mathcal{T}_\sigma}{m_\sigma (\Delta^3 v)^{2/3}} \right) \right]. \quad (\text{S.14})$$

Its Lagrangian time derivative immediately gives

$$\frac{d}{dt} \left(\frac{s_{\sigma v, \mathcal{E}}}{n_\sigma} \right) = \frac{1}{\mathcal{T}_\sigma} \frac{d\mathcal{E}_{\sigma, \text{int}}}{dt}, \quad (\text{S.15})$$

where $d\mathcal{E}_{\sigma, \text{int}} = (3/2)k_B d\mathcal{T}_\sigma$ is the increment in internal energy per particle. This reproduces Eq. (5b). Thus, $d(s_{\sigma v, \mathcal{E}}/n_\sigma)/dt > 0$ implies the effective temperature increases, while $d(s_{\sigma v, \mathcal{E}}/n_\sigma)/dt < 0$ implies the effective temperature decreases. Physically, $s_{\sigma v}$ is associated with the number of permutations of particles of different velocities in a given position in phase space that produces the same macrostate [8].

D. Derivation of Generalized Heat Term

We now derive Eq. (5c). We find it is advantageous to decompose $\nabla \cdot \mathcal{J}_{\sigma, \text{th}}$ using Eq. (S.5) as

$$\nabla \cdot \mathcal{J}_{\sigma, \text{th}} = (\nabla \cdot \mathcal{J}_{\sigma, \text{th}})_{\mathcal{W}} + (\nabla \cdot \mathcal{J}_{\sigma, \text{th}})_{\mathcal{E}} + (\nabla \cdot \mathcal{J}_{\sigma, \text{th}})_{\text{rel}}, \quad (\text{S.16})$$

where

$$(\nabla \cdot \mathcal{J}_{\sigma, \text{th}})_{\mathcal{W}} = -k_B \int (f_\sigma \mathbf{v}'_\sigma) \cdot \nabla \left[\ln \left(\frac{f_\sigma \Delta^3 r_\sigma \Delta^3 v_\sigma}{N_\sigma} \right) \right] d^3 v, \quad (\text{S.17a})$$

$$(\nabla \cdot \mathcal{J}_{\sigma, \text{th}})_{\mathcal{E}} = -k_B \int \left[\nabla \cdot (\mathbf{v}'_\sigma f_\sigma) \ln \left(\frac{f_{\sigma M} \Delta^3 r_\sigma \Delta^3 v_\sigma}{N_\sigma} \right) \right] d^3 v, \quad (\text{S.17b})$$

and $(\nabla \cdot \mathcal{J}_{\sigma, \text{th}})_{\text{rel}}$ is defined in Eq. (7). The latter has equivalent forms of

$$\begin{aligned} (\nabla \cdot \mathcal{J}_{\sigma, \text{th}})_{\text{rel}} &= -\nabla \cdot (\mathbf{u}_\sigma s_{\sigma v, \text{rel}}) \\ &= -k_B \int \nabla \cdot (\mathbf{v} f_\sigma) \ln \left(\frac{f_\sigma}{f_{\sigma M}} \right) d^3 v, \\ &= -s_{\sigma v, \text{rel}} (\nabla \cdot \mathbf{u}_\sigma) \\ &= -k_B \int (\mathbf{v}'_\sigma \cdot \nabla f_\sigma) \ln \left(\frac{f_\sigma}{f_{\sigma M}} \right) d^3 v, \end{aligned}$$

which may be useful in applications depending on which quantities are easiest to measure in a given system.

We first treat $(\nabla \cdot \mathcal{J}_{\sigma, \text{th}})_{\mathcal{W}}/n_\sigma$. The gradient of the term in brackets in Eq. (S.17a) is $(1/f_\sigma) \nabla f_\sigma$. Using $\mathbf{v}'_\sigma = \mathbf{v} - \mathbf{u}_\sigma$, straight-forward manipulations give

$$\frac{(\nabla \cdot \mathcal{J}_{\sigma, \text{th}})_{\mathcal{W}}}{n_\sigma} = -k_B \nabla \cdot \mathbf{u}_\sigma = -k_B n_\sigma \frac{d(1/n_\sigma)}{dt}, \quad (\text{S.18})$$

where we use the continuity equation to eliminate $\nabla \cdot \mathbf{u}_\sigma$. Therefore, this term describes the non-LTE generalization of heating associated with compression or expansion, with the same form as $d(s_{\sigma p}/n_\sigma)/dt$ in Eq. (S.10) but with the opposite sign. This motivates our use of the \mathcal{W} subscript.

Turning to $(\nabla \cdot \mathcal{J}_{\sigma, \text{th}})_{\mathcal{E}}/n_{\sigma}$, we use Eq. (8) to write Eq. (S.17b) as

$$\frac{(\nabla \cdot \mathcal{J}_{\sigma, \text{th}})_{\mathcal{E}}}{n_{\sigma}} = -\frac{k_B}{n_{\sigma}} \int \nabla \cdot (\mathbf{v}'_{\sigma} f_{\sigma}) \left\{ \ln \left[\left(\frac{n_{\sigma} \Delta^3 r_{\sigma} \Delta^3 v_{\sigma}}{N_{\sigma}} \right) \left(\frac{m_{\sigma}}{2\pi k_B \mathcal{T}_{\sigma}} \right)^{3/2} \right] - \frac{m_{\sigma} v_{\sigma}'^2}{2k_B \mathcal{T}_{\sigma}} \right\} d^3v.$$

The term in square brackets is independent of \mathbf{v} and hence comes out of the integral, and the remaining part of that integral is $\int \nabla \cdot (\mathbf{v}'_{\sigma} f_{\sigma}) d^3v = 0$. Manipulations of the remaining term after integration by parts gives

$$\frac{(\nabla \cdot \mathcal{J}_{\sigma, \text{th}})_{\mathcal{E}}}{n_{\sigma}} = \frac{\nabla \cdot (\mathbf{q}_{\sigma}/\mathcal{T}_{\sigma})}{n_{\sigma}} - \frac{m_{\sigma}}{2n_{\sigma}} \int f_{\sigma} \mathbf{v}'_{\sigma} \cdot \left[\frac{\nabla v_{\sigma}'^2}{\mathcal{T}_{\sigma}} - \frac{v_{\sigma}'^2 \nabla \mathcal{T}_{\sigma}}{\mathcal{T}_{\sigma}^2} \right] d^3v, \quad (\text{S.19})$$

where \mathbf{q}_{σ} is the vector heat flux density defined after Eq. (1). Using index notation and the Einstein summation convention,

$$\mathbf{v}'_{\sigma} \cdot \nabla v_{\sigma}'^2 = v_{\sigma}'_j \frac{\partial (v_{\sigma}'_k v_{\sigma}'_k)}{\partial r_j} = 2v_{\sigma}'_j v_{\sigma}'_k \frac{\partial v_{\sigma}'_k}{\partial r_j} = -2v_{\sigma}'_j v_{\sigma}'_k \frac{\partial u_{\sigma k}}{\partial r_j},$$

where we use $\partial v_{\sigma}'_k / \partial r_j = \partial (v_k - u_{\sigma k}) / \partial r_j = -\partial u_{\sigma k} / \partial r_j$ in the last equality. Carrying out the remaining integrals and simplifying gives

$$\frac{(\nabla \cdot \mathcal{J}_{\sigma, \text{th}})_{\mathcal{E}}}{n_{\sigma}} = \frac{\nabla \cdot \mathbf{q}_{\sigma}}{n_{\sigma} \mathcal{T}_{\sigma}} + \frac{(\mathbf{P}_{\sigma} \cdot \nabla) \cdot \mathbf{u}_{\sigma}}{n_{\sigma} \mathcal{T}_{\sigma}}. \quad (\text{S.20})$$

Comparing the right hand side of Eq. (S.20) with Eq. (1), we see both terms appear directly in the internal energy equation with the opposite sign, motivating the choice of the subscript \mathcal{E} . This term describes heat per particle that changes only the effective temperature. A negative value of $(\nabla \cdot \mathcal{J}_{\sigma, \text{th}})_{\mathcal{E}}/n_{\sigma}$ drives increasing \mathcal{T}_{σ} , and a positive value drives decreasing \mathcal{T}_{σ} .

A consequence of Eq. (S.18) is that $(\nabla \cdot \mathcal{J}_{\sigma, \text{th}})_{\mathcal{W}}/n_{\sigma} = -(\mathcal{P}_{\sigma}/n_{\sigma} \mathcal{T}_{\sigma}) \nabla \cdot \mathbf{u}_{\sigma}$, so

$$\frac{(\nabla \cdot \mathcal{J}_{\sigma, \text{th}})_{\mathcal{W}}}{n_{\sigma}} + \frac{(\nabla \cdot \mathcal{J}_{\sigma, \text{th}})_{\mathcal{E}}}{n_{\sigma}} = \frac{\nabla \cdot \mathbf{q}_{\sigma}}{n_{\sigma} \mathcal{T}_{\sigma}} + \frac{\Pi_{\sigma, jk} \mathcal{D}_{\sigma, jk}}{n_{\sigma} \mathcal{T}_{\sigma}} = -\frac{1}{\mathcal{T}_{\sigma}} \frac{d\mathcal{Q}_{\sigma}}{dt}, \quad (\text{S.21})$$

where we use the known decomposition $(\mathbf{P}_{\sigma} \cdot \nabla) \cdot \mathbf{u}_{\sigma} = \mathcal{P}_{\sigma} (\nabla \cdot \mathbf{u}_{\sigma}) + \Pi_{\sigma, jk} \mathcal{D}_{\sigma, jk}$, with $\Pi_{\sigma, jk} = P_{\sigma, jk} - \mathcal{P}_{\sigma} \delta_{jk}$ being elements of the deviatoric pressure tensor $\mathbf{\Pi}$, $\mathcal{D}_{\sigma, jk} = (1/2)(\partial u_{\sigma j} / \partial r_k + \partial u_{\sigma k} / \partial r_j) - (1/3)\delta_{jk} (\nabla \cdot \mathbf{u}_{\sigma})$ being elements of the traceless symmetric strain rate tensor \mathcal{D} , and δ_{jk} being the Kroenecker delta [3, 20]. This derivation provides the expression given in Eq. (5c).

E. Derivation of Relative Energy Per Particle Rate for a Bi-Maxwellian Plasma

Here we derive the rate of relative energy per particle change $d\mathcal{E}_{\sigma, \text{rel}}/dt$ for a bi-Maxwellian initial phase space

density. Consider a purely collisionless system in which the initial f_{σ} is bi-Maxwellian with a converging bulk flow $\mathbf{u}(\mathbf{r}, t)$. We define z as the parallel direction \parallel , x and y as perpendicular \perp directions, and T_{\perp} and T_{\parallel} as the uniform temperatures in the \perp and \parallel directions. The initial phase space density is

$$f_{biM} = n \left(\frac{m}{2\pi k_B T_{\perp}} \right) \left(\frac{m}{2\pi k_B T_{\parallel}} \right)^{1/2} e^{-m[(v_x - u_x)^2 + (v_y - u_y)^2]/2k_B T_{\perp}} e^{-m(v_z - u_z)^2/2k_B T_{\parallel}} \quad (\text{S.22})$$

where n is the initial number density and the constituent mass is m . The Maxwellianized distribution for this system has the form of Eq. (8) with effective temperature $\mathcal{T} = (2T_{\perp} + T_{\parallel})/3$. Then,

$$\ln \left(\frac{f_{biM}}{f_M} \right) = \ln \left(\frac{\mathcal{T}^{3/2}}{T_{\perp} T_{\parallel}^{1/2}} \right) - \frac{m(v_x'^2 + v_y'^2)}{2k_B} \left[\frac{T_{\parallel} - T_{\perp}}{T_{\perp}(2T_{\perp} + T_{\parallel})} \right] - \frac{m v_z'^2}{2k_B} \left[\frac{2(T_{\perp} - T_{\parallel})}{T_{\parallel}(2T_{\perp} + T_{\parallel})} \right],$$

and a straight-forward derivation using Eq. (6) yields

$$\frac{s_{v, \text{rel}}}{n} = -\frac{3}{2} k_B \ln \left[\frac{2}{3} \left(\frac{T_{\perp}}{T_{\parallel}} \right)^{1/3} + \frac{1}{3} \left(\frac{T_{\parallel}}{T_{\perp}} \right)^{2/3} \right] \quad (\text{S.23})$$

The Lagrangian time derivative of this equation, after some algebra, gives

$$\frac{d}{dt} \left(\frac{s_{v, \text{rel}}}{n} \right) = k_B \left(\frac{T_{\parallel} - T_{\perp}}{2T_{\perp} + T_{\parallel}} \right) \left(\frac{1}{T_{\perp}} \frac{dT_{\perp}}{dt} - \frac{1}{T_{\parallel}} \frac{dT_{\parallel}}{dt} \right). \quad (\text{S.24})$$

The thermal evolution equations in a collisionless gyrotropic system, which follow directly from the second parallel and perpendicular moments of the collisionless Boltzmann equation, are written in terms of parallel and perpendicular pressures as [21, 22]

$$\frac{dP_{\parallel}}{dt} + P_{\parallel} \nabla \cdot \mathbf{u} + 2P_{\parallel} [\hat{\mathbf{z}} \cdot (\hat{\mathbf{z}} \cdot \nabla)] \cdot \mathbf{u} = 0, \quad (\text{S.25a})$$

$$\frac{dP_{\perp}}{dt} + 2P_{\perp} \nabla \cdot \mathbf{u} - P_{\perp} [\hat{\mathbf{z}} \cdot (\hat{\mathbf{z}} \cdot \nabla)] \cdot \mathbf{u} = 0, \quad (\text{S.25b})$$

where $P_{\perp} = nk_B T_{\perp}$ and $P_{\parallel} = nk_B T_{\parallel}$. Substituting these into Eq. (S.24) gives

$$\frac{d}{dt} \left(\frac{s_{v, \text{rel}}}{n} \right) = k_B \left(\frac{T_{\parallel} - T_{\perp}}{2T_{\perp} + T_{\parallel}} \right) \left(-\nabla_{\perp} \cdot \mathbf{u}_{\perp} + 2 \frac{\partial u_z}{\partial z} \right), \quad (\text{S.26})$$

where $\mathbf{u}_{\perp} = \mathbf{u} - \hat{\mathbf{z}} u_z$. Finally, using Eq. (9a) to relate this to $d\mathcal{E}_{\sigma, \text{rel}}/dt$ gives

$$\frac{d\mathcal{E}_{\sigma, \text{rel}}}{dt} = \frac{1}{3} k_B (T_{\parallel} - T_{\perp}) \left(-\nabla_{\perp} \cdot \mathbf{u}_{\perp} + 2 \frac{\partial u_z}{\partial z} \right). \quad (\text{S.27})$$

To interpret this result physically, suppose $T_{\parallel} > T_{\perp}$. First consider a bulk flow profile $\mathbf{u} = \mathbf{u}_{\perp}$ that is isotropically converging in the xy plane. Compression leads

to heating, but only in the perpendicular direction [23]. Thus, f_σ becomes more Maxwellian. From Eq. (S.27), both $T_{\parallel} - T_{\perp}$ and the bulk velocity term are positive, so $d\mathcal{E}_{\sigma,\text{rel}}/dt > 0$, consistent with energy going into higher order moments making f_σ more Maxwellian. Now consider converging bulk flow in the z direction. The compression heats the distribution in the parallel direction, so f_σ becomes more elongated in the parallel direction, *i.e.*, less Maxwellian. From Eq. (S.27), $T_{\parallel} - T_{\perp}$ is positive but the bulk velocity term is negative, so this evolution away from Maxwellianity is consistent with $d\mathcal{E}_{\sigma,\text{rel}}/dt$ being negative. This example illustrates general features: $d\mathcal{E}_{\sigma,\text{rel}}/dt > 0$ is associated with energy conversion making higher moments of f_σ evolve to become more Maxwellian, and vice-versa.

F. Numerical Simulation Methodology

Details of the simulation in addition to what follows are available in Ref. [24]. We use the massively parallel particle-in-cell code p3d [25] to perform simulations that are 3D in velocity-space and 2.5D in position-space. Periodic boundary conditions are used in both spatial directions. The code uses the relativistic Boris particle stepper [26] for stepping particles forward in time, while the trapezoidal leapfrog method [27] is utilized for stepping electromagnetic fields forward in time. The fields have a time-step half of that of the particles. The multi-grid method [28] is used to clean the electric field \mathbf{E} to enforce Poisson's equation every 10 particle time-steps.

Simulation results are presented in normalized units. The reference magnetic field B_0 is the initial asymptotic magnetic field strength. The reference number density n_0 is the peak current sheet number density minus the asymptotic background number density. Length scales are normalized to the ion inertial scale $d_{i0} = c/\omega_{pi0}$, where $\omega_{pi0} = (4\pi n_0 q_i^2/m_i)^{1/2}$ is the ion plasma frequency (in cgs units), q_i is the ion charge, m_i is the ion mass, and c is the speed of light. Velocities are normalized to the Alfvén speed $c_{A0} = B_0/(4\pi m_i n_0)^{1/2}$. Times are normalized to the inverse ion cyclotron frequency $\Omega_{ci0}^{-1} = (q_i B_0/m_i c)^{-1}$. Temperatures are normalized to $m_i c_{A0}^2/k_B$. Current densities are normalized to $cB_0/4\pi d_{i0}$. Reduced phase space densities, with one velocity dimension integrated out, are normalized to n_0/c_{A0}^2 . Energy per particle conversion rates are given in units of $\Omega_{ci0} B_0^2/4\pi n_0$.

The initial condition has two oppositely directed current sheets with drifting Maxwellian initial distributions. The magnetic field profile is a double tanh with no initial out-of-plane (guide) magnetic field. The current sheet thickness is $w_0 = 0.5$, the background density is $n_{up} = 0.2$, and the electron and ion temperatures are $1/12$ and $5/12$, respectively. A magnetic perturbation of amplitude 0.05 seeds an X-line/O-line pair in each of the two current sheets. The simulation system size is $L_x \times L_y = 12.8 \times 6.4$, where x and y are the outflow

and inflow directions, respectively. The speed of light c is 15 and the electron to ion mass ratio is $m_e/m_i = 0.04$. There are 1024×512 grid cells initialized with 25,600 weighted particles per grid (PPG), which is chosen to be very large to decrease particle noise. The grid-length Δ in both directions is 0.0125, which is smaller than the smallest length scale which is the electron Debye length of 0.0176. The time-step Δt is 0.001, which is smaller than the smallest time scale which is the inverse of electron plasma frequency of 0.012. Our choice of these numerical parameters results in a total energy increase by only 0.022% by $t = 14$.

All plots display data from only the lower current sheet at time $t = 13$, when the rate of reconnection is increasing most rapidly in time. To reduce PIC noise for all quantities plotted other than phase space densities, the raw quantities are recursively smoothed four times over a width of four cells, then any temporal or spatial derivatives are carried out, and then the results are again smoothed recursively four times over four cells. For temporal derivatives, the presented data is calculated from a finite difference between $t = 12$ and 14 (on ion cyclotron time scales). The results are compared to those obtained from a finite difference between $t = 12.96$ and 13.04 (electron time scales), and the results are found to differ by less than 5%; this change is deemed inconsequential for the present purposes.

Kinetic entropy is calculated in the simulations employing the implementation from Ref. [29]. Optimization of the velocity-space grid [30] is done by checking the agreement between the kinetic entropy density for electrons s_e calculated by the simulation for various Δv_e and the theoretical value at $t = 0$. We find an optimal Δv_e of 1.33 which leads to an initial agreement to within $\pm 1\%$ in the upstream region and $\pm 3\%$ at the center of the current sheet. For plots of reduced electron phase space densities, we use a domain of size 0.0625×0.0625 centered at the location of interest. Particles are binned with a velocity space bin of size 0.1 in all velocity directions.

G. Additional Comments on the Relation to Previous Work

Here, we put our result in context of previous work on related topics.

- **Energy Conversion in δf_σ Kinetic Theory and Gyrokinetics:** We first compare the present work with previous work on energy conversion in linearized kinetic theory and gyrokinetics [31–34]. Consider a linear expansion of the phase space density about its Maxwellianized distribution, so that $f_\sigma = f_{\sigma M} + \delta f_\sigma$, and $\delta f_\sigma \ll f_{\sigma M}$. A straightforward calculation using Eq. (S.13) reveals that the linearized relative entropy $\delta s_{\sigma v,\text{rel}}$ is

$$\delta s_{\sigma v,\text{rel}} \simeq -k_B \int \frac{(\delta f_\sigma)^2}{2f_{\sigma M}} d^3v. \quad (\text{S.28})$$

In linear theory, the density and temperature in $f_{\sigma M}$ are their equilibrium values, which we call $n_{\sigma 0}$ and $T_{\sigma 0}$, respectively. Then, the linearized equation describing the relative energy increment using Eq. (9a) is

$$\frac{d\mathcal{E}_{\sigma,\text{rel}}}{dt} \simeq T_{\sigma 0} \frac{d(s_{\sigma v,\text{rel}}/n_{\sigma 0})}{dt}. \quad (\text{S.29})$$

Since the equilibrium temperature does not change to low order in linear theory, $n_{\sigma 0}$ and $T_{\sigma 0}$ are constant in time, so integrating Eq. (S.29) in time gives

$$\begin{aligned} \delta\mathcal{E}_{\sigma,\text{rel}} &\simeq \frac{T_{\sigma 0}}{n_{\sigma 0}} \delta s_{\sigma v,\text{rel}} \\ &\simeq -\frac{k_B T_{\sigma 0}}{n_{\sigma 0}} \int \frac{(\delta f_{\sigma})^2}{2f_{\sigma M}} d^3v, \end{aligned} \quad (\text{S.30})$$

where we use Eq. (S.28).

In comparison, the free energy in a δf_{σ} linearized thermodynamic approach [31] [κ_{int} in their Eq. (7)] was derived to be

$$\kappa_{\text{int}} = k_B T_0 \int \frac{(\delta f)^2}{2f_{\sigma M}} d^3v \quad (\text{S.31})$$

and in a gyrokinetic analysis of energy conversion [32, 33], the comparable term from the free energy [the first term of W in Eq. (74) from Ref. [33]] is

$$W = \int d^3r \sum_{\sigma} \int k_B T_{0\sigma} \frac{(\delta f)^2}{2f_{\sigma M}} d^3v. \quad (\text{S.32})$$

Clearly, the linearized relative energy per particle $\delta\mathcal{E}_{\sigma,\text{rel}}$ in Eq. (S.30) is related to the free energy in the δf_{σ} thermodynamic and the gyrokinetic approaches. In particular, $\kappa_{\text{int}} = -n_{\sigma 0} \delta\mathcal{E}_{\sigma,\text{rel}}$ and $W = -\int d^3r \sum_{\sigma} n_{\sigma 0} \delta\mathcal{E}_{\sigma,\text{rel}}$. [Note, the relative entropy term differs from the nonlinear term used in Ref. [35] that reproduces Eq. (S.31) when linearized; theirs is related to \bar{M}_{KP} rather than the relative energy term.] The sign difference is a result of $\delta\mathcal{E}_{\sigma,\text{rel}}$ measuring the energy going into the random energy of the particles, while κ_{int} and W describe energy going into the bulk flow energy and magnetic fields from the particles. Thus, the present work is consistent with previous work, and generalizes these linear approaches for phase space densities arbitrarily far from LTE.

- **Previous Schematics of Energy Conversion:**

We now put the sketch of energy conversion in Fig. 3 in the context of previous sketches about energy conversion in plasmas. It is similar to Fig. 1 in Ref. [20], except theirs is averaged over a closed or periodic domain so the heat flux does not contribute, theirs includes conversion into bulk kinetic energy and electromagnetic energy which are omitted from the present treatment for simplicity,

and ours includes collisions. The key difference is the additional energy conversion channel associated with relative energy and heat that arise from our analysis as another possible energy conversion channel.

Another related sketch is Fig. 4 in Ref. [36], which describes energy conversion in weakly collisional turbulent plasmas. There, electromagnetic fields play a key role in converting energy to non-thermal (non-LTE) energy in the plasma, which ultimately produce irreversible dissipation through the collisions. The present work treats only internal moments of the phase space density, which formally has only indirect input from body forces [which, for example, do not appear in Eq. (1)]. Thus, our result is in many ways complementary to the research done on the field-particle correlation [37]. It would be interesting and important to unite the two approaches in future work.

- **The Velocity Space Cascade and Hermite Expansions of f_{σ} :**

An important approach that has previously been used to study non-LTE energy conversion is to take a local phase space density and expand the velocity space part in Hermite polynomials [38–40]. The coefficients in the expansion provide information about how non-Maxwellian the system is at that location in space and time. In a weakly collisional or collisionless system, many phase space densities develop sharp structures in velocity space, which shows up as a cascade of power into the higher order coefficients in the expansion.

It would be tempting to associate the power in non-LTE modes, called the enstrophy in Ref. [38], with the relative energy per particle in the present analysis, but this association is not possible. The reason is that the enstrophy is a local quantity that can be calculated for any phase space density, but the relative energy per particle is history dependent, so only changes to it can be uniquely determined from the local phase space density at a particular time. A phase space density becoming more non-Maxwellian has an increase in enstrophy, while it corresponds to a decrease in the relative energy per particle because the Maxwellian is the maximum entropy state. While associating the two approaches in this manner is therefore not possible, we do believe there are links between the two approaches which will be pursued in future studies.

- **Energy Conversion Using Other Entropies:**

Recent work quantified non-LTE effects using non-Boltzmann entropies for collisionless plasmas [5, 41]. In Ref. [5], energy conversion was parametrized by moments of integer powers of f_{σ} , which are invariants in collisionless systems. In Ref. [41], it was shown that power law entropies are well-suited for

describing power law tails during non-thermal particle acceleration. As pointed out there, these terms provide information about the shape of the phase space density, so there are some similarities about the aims of the two studies despite their different approaches.

The formulation here using the Boltzmann entropy is related to these invariants, as an expansion of the natural logarithm in powers of f_σ inside the kinetic entropy density $s_\sigma = -k_B \int f_\sigma \ln(f_\sigma \Delta^3 r_\sigma \Delta^3 v_\sigma / N_\sigma) d^3 v$ yields integrals over all integer powers of f_σ , as done in Ref. [5]. Consequently, the form derived here based on Boltzmann entropy without expanding the natural logarithm automatically contains the information about all of the power law invariants for collisionless systems. Ref. [5] is important for identifying how the energy is contained in different individual invariants, which is not possible in the present formulation. However, our results can readily be used for collisional systems even though powers of f_σ are no longer invariants.

- **Extended Irreversible Thermodynamics (EIT):** EIT begins with the kinetic entropy evolution equation [Eq. (S.2)] and employs a perturbative expansion of f_σ , and the terms of higher order represent corrections to the first law of thermodynamics. This is very important because the corrections are in terms of fluid moments of f_σ , so a direct measurement of f_σ is not necessary. The advantage of the present analysis is that all internal moments are retained, so there is no need to be near LTE.

We also point out that the phase space density f_σ inside the natural logarithm in the general expression for $\mathcal{J}_{\sigma,\text{th}}$ [Eq. (S.5)] is expanded about the Maxwellianized distribution $f_{\sigma M}$ in EIT. The lowest order term in this expansion is [3]

$$\mathcal{J}_{\sigma,q} = -k_B \int \mathbf{v}'_\sigma f_\sigma \ln \left(\frac{f_{\sigma M} \Delta^3 r_\sigma \Delta^3 v_\sigma}{N_\sigma} \right) d^3 v. \quad (\text{S.33})$$

A brief derivation using Eq. (8) reveals that $\mathcal{J}_{\sigma,q} = \mathbf{q}_\sigma / \mathcal{T}_\sigma$. In the present study, instead of decomposing f_σ inside $\mathcal{J}_{\sigma,\text{th}}$, we decompose f_σ inside $\nabla \cdot \mathcal{J}_{\sigma,\text{th}}$ as Eq. (S.16). The difference here is that

$\mathcal{J}_{\sigma,q} = \mathbf{q}_\sigma / \mathcal{T}_\sigma$ from Eq. (S.33), so $\nabla \cdot \mathcal{J}_{\sigma,q}$ contains both a $(\nabla \cdot \mathbf{q}_\sigma) / \mathcal{T}_\sigma$ term and a $-(\mathbf{q}_\sigma \cdot \nabla \mathcal{T}_\sigma) / \mathcal{T}_\sigma^2$ term. The latter term is included as an entropy source term in the fluid form of EIT [3]. Eq. (S.20) reveals that $-(\mathbf{q}_\sigma \cdot \nabla \mathcal{T}_\sigma) / \mathcal{T}_\sigma^2$ vanishes exactly when all orders of non-LTE terms are retained so that it should not be retained.

- **Quantum Statistical Mechanics:** There are similarities and differences of our results with a recent independent analysis showing that the quantum first law of thermodynamics can be obtained from the quantum relative entropy [42]. In the classical limit, the density matrix ρ is analogous to the distribution function f_σ / n_σ [43]. The maximally mixed state σ_m , which has the highest entropy, is analogous to the Maxwellianized distribution function $f_{\sigma M} / n_\sigma$. The von Neumann entropy $S(\rho) = -\text{tr}[\rho \ln \rho]$ [44] is decomposed as $S(\rho) = S_{\text{cross}}(\rho) - S_{\text{rel}}(\rho)$, where $S_{\text{cross}}(\rho) = -\text{tr}[\rho \ln \sigma_m]$ is the cross-entropy and $S_{\text{rel}} = \text{tr}[\rho \ln \rho - \rho \ln \sigma_m] = -S(\rho) + S_{\text{cross}}(\rho)$ is the relative entropy [42]. This is similar to the decomposition done here for the velocity space kinetic entropy per particle, so $S_{\text{cross}}(\rho)$ is analogous to $s_{\sigma v,\mathcal{E}} / n_\sigma$ [Eq. (S.12)] and $S_{\text{rel}}(\rho)$ is analogous to $-s_{\sigma v,\text{rel}} / n_\sigma$ [Eq. (6)]. In Ref. [42], the volume of the system was kept fixed for simplicity, so there was no term analogous to the position space entropy term in our analysis. Including this term, which gives rise to work in the classical case, is very straight-forward; indeed, it appears automatically when the phase space density f_σ is employed instead of the distribution function f_σ / n_σ . Undoubtedly the quantum statistical mechanical approach can be generalized to include work done on the system using open quantum mechanics [45].

For the classical case presented here, the physical interpretation of the terms are able to be clearly ascertained. This allows us to help elucidate the physical interpretation of the terms in the quantum statistical mechanics treatment [42]. The time rate of change of the relative quantum entropy is a measure of whether a system is evolving towards or away from the maximally mixed state and the rate at which it does so. Scaling it by the temperature of the state described by σ_m gives the time rate of change in the energy associated with non-equilibrium terms.

[1] L. Boltzmann, Weitere studien über das wärmeleichgewicht unter gasmolekülen, Wiener Berichte **66**, 275 (1872), in (Boltzmann 1909) Vol. I, paper 23.
 [2] B. C. Eu, Boltzmann entropy, relative entropy, and related quantities in thermodynamic space, The Journal of

chemical physics **102**, 7169 (1995).
 [3] D. Jou, G. Lebon, and J. Casas-Vázquez, *Extended Irreversible Thermodynamics*, 4th ed. (Springer, Dordrecht, 2010).
 [4] G. L. Eyink, Cascades and dissipative anomalies in nearly collisionless plasma turbulence, Phys. Rev. X **8**, 041020

- (2018).
- [5] V. Zhdankin, Generalized entropy production in collisionless plasma flows and turbulence, *Phys. Rev. X* **12**, 031011 (2022).
- [6] L. Boltzmann, Über die beziehung dem zweiten haubtsatze der mechanischen wärmetheorie und der wahrrscheinlichkeitsrechnung resp. dem sätzen über das wäärmegleichgewicht, *Wiener Berichte* **76**, 373 (1877), in (Boltzmann 1909) Vol. II, paper 42.
- [7] W. Grandy, Time Evolution in Macroscopic Systems. II. The Entropy., *Foundations of Physics* **34**, 21 (2004).
- [8] H. Liang, P. A. Cassak, S. Servidio, M. A. Shay, J. F. Drake, M. Swisdak, M. R. Argall, J. C. Dorelli, E. E. Scime, W. H. Matthaeus, V. Roytershteyn, and G. L. Delzanno, Decomposition of plasma kinetic entropy into position and velocity space and the use of kinetic entropy in particle-in-cell simulations, *Phys. Plasmas* **26**, 082903 (2019).
- [9] S. Kullback and R. A. Leibler, On information and sufficiency, *The annals of mathematical statistics* **22**, 79 (1951).
- [10] E. T. Jaynes, Information theory and statistical mechanics, in: K. ford, ed., *statistical physics* (Benjamin, New York, 1963) p. 181.
- [11] H. Grad, On boltzmann’s h-theorem, *Journal of the Society for Industrial and Applied Mathematics* **13**, 259 (1965).
- [12] R. J. Diperna, Uniqueness of solutions to hyperbolic conservation laws, *Indiana University Mathematics Journal* **28**, 137 (1979).
- [13] C. M. Dafermos, The second law of thermodynamics and stability, *Archive for Rational Mechanics and Analysis* **70**, 167 (1979).
- [14] V. Vedral, The role of relative entropy in quantum information theory, *Reviews of Modern Physics* **74**, 197 (2002).
- [15] J. C. Robertson, E. W. Tallman, and C. H. Whiteman, Forecasting using relative entropy, *Journal of Money, Credit, and Banking* **37**, 383 (2005).
- [16] A. E. Tzavaras, Relative entropy in hyperbolic relaxation, *Communications in Mathematical Sciences* **3**, 119 (2005).
- [17] M. S. Shell, The relative entropy is fundamental to multi-scale and inverse thermodynamic problems, *The Journal of chemical physics* **129**, 144108 (2008).
- [18] F. Berthelin, A. E. Tzavaras, and A. Vasseur, From discrete velocity boltzmann equations to gas dynamics before shocks, *Journal of Statistical Physics* **135**, 153 (2009).
- [19] J. C. Baez and B. S. Pollard, Relative entropy in biological systems, *Entropy* **18**, 10.3390/e18020046 (2016).
- [20] Y. Yang, W. H. Matthaeus, T. N. Parashar, C. C. Haggerty, V. Roytershteyn, W. Daughton, M. Wan, Y. Shi, and S. Chen, Energy transfer, pressure tensor, and heating of kinetic plasma, *Physics of Plasmas* **24**, 072306 (2017), <https://doi.org/10.1063/1.4990421>.
- [21] G. F. Chew, M. L. Goldberger, and F. E. Low, The boltzmann equation and the one-fluid hydromagnetic equations in the absence of particle collisions, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **236**, 112 (1956).
- [22] M. Hesse and J. Birn, Mhd modeling of magnetotail instability for anisotropic pressure, *Journal of Geophysical Research: Space Physics* **97**, 10643 (1992).
- [23] P. A. Cassak and M. H. Barbhuiya, Pressure–strain interaction. I. On compression, deformation, and implications for Pi-D, *Physics of Plasmas* **29**, 122306 (2022).
- [24] M. H. Barbhuiya and P. A. Cassak, Pressure–strain interaction. III. Particle-in-cell simulations of magnetic reconnection, *Physics of Plasmas* **29**, 122308 (2022).
- [25] A. Zeiler, D. Biskamp, J. F. Drake, B. N. Rogers, M. A. Shay, and M. Scholer, Three-dimensional particle simulations of collisionless magnetic reconnection, *J. Geophys. Res.* **107**, 1230 (2002).
- [26] C. K. Birdsall and A. B. Langdon, *Plasma physics via computer simulation* (Institute of Physics Publishing, Philadelphia, 1991) Chap. 15.
- [27] P. N. Guzdar, J. F. Drake, D. McCarthy, A. B. Hassam, and C. S. Liu, Three-dimensional fluid simulations of the nonlinear drift-resistive ballooning modes in tokamak edge plasmas, *Phys. Fluids B* **5**, 3712 (1993).
- [28] U. Trottenberg, C. W. Oosterlee, and A. Schuller, *Multi-grid* (Academic Press, San Diego, 2000).
- [29] M. R. Argall, M. H. Barbhuiya, P. A. Cassak, S. Wang, J. Shuster, H. Liang, D. J. Gershman, R. B. Torbert, and J. L. Burch, Theory, observations, and simulations of kinetic entropy in a magnetotail electron diffusion region, *Physics of Plasmas* **29**, 022902 (2022).
- [30] H. Liang, P. A. Cassak, M. Swisdak, and S. Servidio, Estimating effective collision frequency and kinetic entropy uncertainty in particle-in-cell simulations, *Journal of Physics: Conference Series* **1620**, 012009 (2020).
- [31] K. Hallatschek, Thermodynamic potential in local turbulence simulations, *Physical review letters* **93**, 125001 (2004).
- [32] G. G. Howes, S. C. Cowley, W. Dorland, G. W. Hammett, E. Quataert, and A. A. Schekochihin, Astrophysical gyrokinetics: basic equations and linear theory, *The Astrophysical Journal* **651**, 590 (2006).
- [33] A. Schekochihin, S. Cowley, W. Dorland, G. Hammett, G. G. Howes, E. Quataert, and T. Tatsuno, Astrophysical gyrokinetics: kinetic and fluid turbulent cascades in magnetized weakly collisional plasmas, *The Astrophysical Journal Supplement Series* **182**, 310 (2009).
- [34] T. Tatsuno, W. Dorland, A. Schekochihin, G. Plunk, M. Barnes, S. Cowley, and G. Howes, Nonlinear phase mixing and phase-space cascade of entropy in gyrokinetic plasma turbulence, *Physical review letters* **103**, 015003 (2009).
- [35] S. Cerri, M. W. Kunz, and F. Califano, Dual phase-space cascades in 3d hybrid-vlasov–maxwell turbulence, *The Astrophysical Journal Letters* **856**, L13 (2018).
- [36] G. G. Howes, A. J. McCubbin, and K. G. Klein, Spatially localized particle energization by landau damping in current sheets produced by strong alfvén wave collisions, *Journal of Plasma Physics* **84** (2018).
- [37] K. G. Klein and G. G. Howes, Measuring collisionless damping in heliospheric plasmas using field–particle correlations, *The Astrophysical Journal Letters* **826**, L30 (2016).
- [38] S. Servidio, A. Chasapis, W. Matthaeus, D. Perrone, F. Valentini, T. Parashar, P. Veltri, D. Gershman, C. Russell, B. Giles, and et al., Magnetospheric multi-scale observation of plasma velocity-space cascade: Hermite representation and theory, *Physical review letters* **119**, 205101 (2017).
- [39] O. Pezzi, S. Servidio, D. Perrone, F. Valentini, L. Sorriso-

- Valvo, A. Greco, W. Matthaeus, and P. Veltri, Velocity-space cascade in magnetized plasmas: Numerical simulations, *Physics of Plasmas* **25**, 060704 (2018).
- [40] O. Pezzi, Y. Yang, F. Valentini, S. Servidio, A. Chasapis, W. H. Matthaeus, and P. Veltri, Energy conversion in turbulent weakly collisional plasmas: Eulerian hybrid vlasov-maxwell simulations, *Physics of Plasmas* **26**, 072301 (2019), <https://doi.org/10.1063/1.5100125>.
- [41] V. Zhdankin, Nonthermal particle acceleration from maximum entropy in collisionless plasmas, arXiv preprint arXiv:2203.13054 (2022).
- [42] S. Floerchinger and T. Haas, Thermodynamics from relative entropy, *Phys. Rev. E* **102**, 052117 (2020).
- [43] J. J. Sakurai, *Modern quantum mechanics; rev. ed.* (Addison-Wesley, Reading, MA, 1994).
- [44] J. Von Neumann, Thermodynamik quantenmechanischer gesamtheiten, *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* **1927**, 273 (1927).
- [45] D. A. Lidar, Lecture notes on the theory of open quantum systems, arXiv preprint arXiv:1902.00967v2 (2020).