

# CONTRACTIVITY OF WASSERSTEIN DISTANCE AND OTHER FUNCTIONALS FOR THE LANDAU EQUATION

F.-U. CAJA-LOPEZ, M. G. DELGADINO, M.-P. GUALDANI, M. TASKOVIC

**ABSTRACT.** Following the breakthrough work of Guillen and Silvestre [1], that shows that the Fisher information is monotonically decreasing for solutions to the homogeneous Landau equation, we study, for the same equation, the monotonicity properties of other physically relevant functionals. In the case of Maxwellian molecules, we show that the relative  $L^2$  norm with respect to the equilibrium decays exponentially fast in time and is monotonically decreasing after some time. Moreover, still for the Maxwellian case, we provide a novel and short quantitative proof of time monotonicity of the entropic Wasserstein metric.

For soft potentials, we show that the Wasserstein metric is contractive, conditional to  $L^1(0, T; L^p(\mathbb{R}^3))$  bound for the solution. This result provides an alternative proof of the Fournier and Fournier-Guerin uniqueness theorem in [2] [3].

## 1. INTRODUCTION

In this paper, we consider the space-homogeneous Landau equation

$$(1.1) \quad \partial_t f = \operatorname{div}_v \int_{\mathbb{R}^d} \alpha(|v-w|) |v-w|^2 K(v-w) (f(w) \nabla_v f(v) - f(v) \nabla_w f(w)) dw,$$

with  $K(z) := \operatorname{Id} - \frac{z \otimes z}{|z|^2}$  being the orthogonal projection to  $\langle z \rangle^\perp$  and  $\alpha : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ . The most relevant physical case is  $d = 3$  and  $\alpha(r) = r^{-3}$ , which corresponds to Coulomb interactions of electrons, and arises in mathematical modeling of electrons in plasma. The Landau equation is one of the most used mathematical models in plasma physics. From a mathematical point of view, the presence of nonlinear and non-local terms makes the analysis quite challenging. After decades of being an open problem, well posedness of (1.1) with smooth data has been recently settled in the breakthrough work of Guillen-Silvestre [1]. The proof relies on a novel a-priori estimate which shows that the Fisher information is non-increasing in time. Subsequently, global well-posedness for less regular initial data was shown in [4, 5, 6, 7]. The surprising monotonicity of the Fisher information hints at the possibility of additional functionals that either decay monotonically or stay bounded.

In a preliminary attempt to answer this question, we focus here on three specific functionals. The first one is the relative  $L^2$  distance between any solution to (1.1) and the Maxwellian equilibrium. The other two are the classical 2-Wasserstein distance and the entropic Wasserstein distance between two solutions to (1.1). We summarize our main results in three theorems.

The first theorem shows that in the case of Maxwell molecules ( $\alpha \equiv 1$ ), the relative  $L^2$  distance decays exponentially fast in time. Moreover, this distance is monotonically decreasing after

---

*Key words and phrases.* Landau equation, Maxwellian molecules, entropic Wasserstein distance, relative  $L^2$ -norm, exponential decay, contractivity.

a short initial time. This transient period *measures* how far the initial data is from a radially symmetric configuration. If the initial data is *almost* radially symmetric, the relative  $L^2$  norm is monotone decreasing for all times.

**Theorem 1.** *Let  $f$  be a solution to (1.1) with  $\alpha \equiv 1$  and initial data  $f_{in}$  such that  $\int f_{in}(1, v, |v|^2) dv = (1, 0, d)$ . Let  $m$  be the corresponding unique equilibrium solution,  $m(v) := (2\pi)^{-\frac{d}{2}} e^{-\frac{1}{2}|v|^2}$ . Then the relative  $L^2$ -norm*

$$E(t) := \int_{\mathbb{R}^d} \left( \frac{f}{m} - 1 \right)^2 m dv$$

*decays exponentially fast in time,*

$$E(t) \leq c_0 e^{-2(d-1)t} E(0),$$

*where  $c_0$  depends only on the dimension  $d$ . Moreover,  $E(t)$  is monotonically decreasing for any  $t \geq t_0$  with  $t_0 = \max \left\{ \frac{1}{4d} \log \left( \frac{2 \sup_i (\lambda_i^{in} - 1)}{d-1} \right), 0 \right\}$ , where  $\lambda_i^{in}$  are the eigenvalues of the matrix  $\mathbb{P}_{ij}^{in} = \int v_i v_j f_{in} dv$ .*

**Remark.** If  $f_{in}$  satisfies  $\lambda_i^{in} \leq \frac{d+1}{2}$  for all  $i = 1, \dots, d$ , then  $t_0 = 0$ , namely the relative  $L^2$  norm decreases from the initial time. In particular this applies to radially symmetric initial data, which have  $\lambda_i^{in} = 1$ .

Our second result shows that, still for the Maxwell molecules case, the Landau equation is contractive with respect to the entropic 2-Wasserstein distance. Before we define the distance, let's recall the definition of Boltzmann entropy  $\mathcal{H}(f)$  and relative entropy  $\mathcal{H}(f|g)$  between two probability measures  $f$  and  $g$  as

$$\mathcal{H}(f) := \int f \log f dx, \quad \mathcal{H}(f|g) := \int \left( \frac{f}{g} \log \frac{f}{g} - \frac{f}{g} + 1 \right) g dx.$$

The entropic 2-Wasserstein distance is defined as

$$(1.2) \quad \mathcal{T}_\varepsilon(\mu, \nu) = \inf_{\Pi \in \Gamma(\mu, \nu)} \left\{ \iint |v - w|^2 d\Pi(v, w) + \varepsilon \mathcal{H}(\Pi|\mu \otimes \nu) \right\},$$

where  $\mu, \nu$  are two probability measures over  $\mathbb{R}^d$ , and  $\Gamma(\mu, \nu)$  stands for the set of transport plans, which are probability measures over  $\mathbb{R}^d \times \mathbb{R}^d$  with marginals  $\mu$  and  $\nu$ , see [8] for further details. Before we state our second main theorem, we recall a useful expression introduced in [1] for the matrix  $|v - w|^2 K(v - w)$ :

$$|v - w|^2 K(v - w) = \sum_{k=1}^3 b_k(v - w) \otimes b_k(v - w),$$

with

$$b_1(v - w) := \begin{pmatrix} 0 \\ w_3 - v_3 \\ v_2 - w_2 \end{pmatrix}, \quad b_2(v - w) := \begin{pmatrix} v_3 - w_3 \\ 0 \\ w_1 - v_1 \end{pmatrix}, \quad b_3(v - w) := \begin{pmatrix} w_2 - v_2 \\ v_1 - w_1 \\ 0 \end{pmatrix}.$$

In the following of the manuscript we denote by  $\tilde{b}_k$  the vector

$$\tilde{b}_k := \begin{pmatrix} b_k \\ -b_k \end{pmatrix}.$$

Our second main result is summarized in the following theorem:

**Theorem 2.** *Let  $f, g$  be solutions to (1.1) for  $\alpha \equiv 1$  in  $\mathbb{R}^3$ . Then  $\mathcal{T}_\varepsilon(f, g)$  is non-increasing for any  $\varepsilon \geq 0$ . Moreover, the dissipation can be expressed as*

$$(1.3) \quad \frac{d}{dt} \mathcal{T}_\varepsilon(f, g) = \varepsilon \underbrace{\frac{d}{dt} \mathcal{H}(f)}_{\leq 0} - \frac{1}{2} \sum_k \int_0^1 \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \left| \nabla(\tilde{b}_k \cdot \nabla U) \right|^2 P dv dw ds \leq 0,$$

where  $P : [0, 1] \rightarrow \mathcal{P}(\mathbb{R}^3 \times \mathbb{R}^3)$  is the Wasserstein geodesic between  $F(v, w) = f(v)f(w)$  and  $G(v, w) = g(v)g(w)$  with vector field given by  $\nabla U : [0, 1] \rightarrow L^2(\mathbb{R}^3 \times \mathbb{R}^3, dP)^{\otimes 6}$ .

The Geodesic  $P$  solves a continuity equation with  $\nabla U$  as the driving vector field, and  $U$  solves a Hamilton-Jacobi-Bellman equation. See Section 5 for the precise statements and explicit expression of  $U$ .

We provide two different proofs of the monotonicity of  $\mathcal{T}_\varepsilon$ . The first one closely follows the symmetrization argument introduced in [1] (see Section 4). The second proof relies on the duality between entropic optimal transport and Hamilton-Jacobi-Bellman equations; thanks to the duality argument, we are able to find the explicit expression of the dissipation (1.3). This approach is developed in Section 5.

We remark that the contractivity of the 2-Wasserstein distance  $\mathcal{T}$  for Maxwellian molecules was already known [9, 10]. Even more classically, we mention the work of Tanaka [11], which first showed the monotonicity of the 2-Wasserstein metric for the Boltzmann equation. We also mention [12] which produces a tailored 2-Wasserstein like metric on which the Landau equation arises as the gradient flow of the entropy; understanding if the Landau equation is contractive in this natural metric is an interesting open problem. Here we provide a shorter and more intuitive proof of the monotonicity of  $\mathcal{T}$ . The expression of the dissipation on the right hand side of (1.3) is at the authors' best knowledge new.

The last main result concerns soft potentials and the contractivity of the 2-Wasserstein  $\mathcal{T}(f_t, g_t)$  distance. Recall that  $\mathcal{T}(f_t, g_t)$  is defined in (1.2) with  $\varepsilon = 0$ .

**Theorem 3.** *Let  $f_t, g_t$  be two smooth solutions to (1.1) in  $\mathbb{R}^3$  with  $\alpha(r) = r^\gamma$ .*

(i) *Let  $-3 < \gamma < 0$ . Assume that  $f_t, g_t \in L_t^1 L_v^p$  for  $p > \frac{3}{3+\gamma}$ .*

*There exist constants  $C_{\gamma,p} > 0$  and  $0 < \beta_{\gamma,p} < 1$  such that*

$$(1.4) \quad \mathcal{T}(f_t, g_t) \leq \mathcal{T}(f_{in}, g_{in}) \exp \left( C_{\gamma,p} \int_0^t \left( \|f_s\|_{L^p(\mathbb{R}^3)}^{\beta_{\gamma,p}} + \|g_s\|_{L^p(\mathbb{R}^3)}^{\beta_{\gamma,p}} \right) ds \right).$$

(i) *Let  $\gamma = -3$ . Assume  $f_t, g_t \in L_t^1 L_v^\infty$ , and initial data such that  $\mathcal{T}(f_{in}, g_{in}) < e^{-4}$ . Then there exist constants  $C > 0$  and  $\tau_0 > 0$  such that for  $t < \tau_0$  :*

$$(1.5) \quad \mathcal{T}(f_t, g_t) \leq \mathcal{T}(f_{in}, g_{in})^{e^{-\beta(t)}}, \quad \beta(t) := C \int_0^t \max \left( \|f_s\|_{L^\infty(\mathbb{R}^3)} + \|g_s\|_{L^\infty(\mathbb{R}^3)}, 1 \right) ds.$$

The contractive estimates for soft potentials of Theorem 3 match the estimates of Fournier and Guerin [2] and Fournier [3]. Indeed, our result shows that uniqueness of solution for (1.1) for Coulomb potential holds provided the solution belongs to  $L^1(0, \tau, L^\infty(\mathbb{R}^3))$  for small

$\tau > 0$ . Our proof does not rely on stochastic methods, but solely uses the symmetrization properties of the distance  $\mathcal{T}$  and of the Landau operator.

It turns out that Theorem 3 holds for a wider range of potentials. More precisely, it holds for any general potential  $\alpha(r)$  that satisfies

$$\alpha'(r)^2 \leq Kr^{\gamma-2}\alpha(r),$$

for some positive constant  $K$  and some  $\gamma \in [-3, 0]$ .

**Organization of the paper.** The rest of the paper is organized as follows. Section 2 collects some important properties of the Landau operator with Maxwellian molecules. Section 3 contains the proof of Theorem 1. Theorem 2 is proven using convexity properties in Section 4 and via duality arguments in Section 5. Finally, Section 6 contains the proof of Theorem 3.

## 2. THE COLLISION OPERATOR FOR MAXWELLIAN MOLECULES

The proof of Theorem 1 ( $\alpha \equiv 1$ ) starts with the Landau equation (1.1) rewritten in the following form:

$$(2.1) \quad \partial_t f = \operatorname{div} \left( \left[ |v|^2 K(v) + d \operatorname{Id} - \mathbb{P}(t) \right] \nabla f + (d-1)vf \right),$$

where  $\mathbb{P}(t)$  is the diagonal matrix formed by

$$(2.2) \quad \lambda_i(t) = \int v_i^2 f(v) dv, \quad \lambda_i(t) = 1 + e^{-4dt}(\lambda_i(0) - 1).$$

We briefly show how to derive the expression (2.1) assuming

$$\int f_{\text{in}}(1, v, |v|^2) dv = \int f(1, v, |v|^2) dv = (1, 0, d).$$

This has already been done in [14], but we present a slightly modified proof for completeness, as these computations will be used in the proof of Theorem 1.

**Proposition 1.** *Let  $f \geq 0$  such that  $\int f(1, v, |v|^2) dv = (1, 0, d)$ . The matrix*

$$A[f](v, t) := \int_{\mathbb{R}^d} K(v-w)|v-w|^2 f(w, t) dw$$

*in the Maxwell molecules case  $\alpha \equiv 1$  can be rewritten as*

$$A[f](v, t) = |v|^2 K(v) + d \operatorname{Id} - \mathbb{P}(t)$$

*where  $\mathbb{P}(t)$  is the matrix with entries  $P_{ij}(t) := \int w_i w_j f(w, t) dw$ .*

*Moreover, the Landau equation (1.1) reduces to*

$$(2.3) \quad \partial_t f = \operatorname{div} \left( \left[ |v|^2 K(v) + d \operatorname{Id} - \mathbb{P}(t) \right] \nabla f + (d-1)vf \right).$$

*Proof.* Direct computations show that

$$A[f] = \operatorname{Id} \int |v-w|^2 f(w) dw - \int (v-w) \otimes (v-w) f(w) dw.$$

For the first term we have

$$\int |v - w|^2 f(w) dw = \left[ |v|^2 \int f(w) dw - 2v \cdot \int wf(w) dw + \int |w|^2 f(w) dw \right] = (|v|^2 + d).$$

For the second:

$$\begin{aligned} \int (v - w) \otimes (v - w) f(w) dw &= v \otimes v \int f(w) dw - 2v \otimes \int wf(w) dw + \int (w \otimes w) f(w) dw \\ &= (v \otimes v) + \int (w \otimes w) f(w) dw = (v \otimes v) + \mathbb{P}(t). \end{aligned}$$

Hence we have

$$A[f] = |v|^2 \left( \text{Id} - \frac{v \otimes v}{|v|^2} \right) + d \text{Id} - \mathbb{P}(t) = |v|^2 K(v) + d \text{Id} - \mathbb{P}(t).$$

This also provides a very simple expression for

$$\text{div}(A[f]) = \text{div}(|v|^2 \text{Id} - v \otimes v) = [2v - (d+1)v] = -(d-1)v.$$

Since (1.1) can be expressed in terms of  $A[f]$  as

$$\partial_t f = \text{div}(A[f] \nabla f - f \text{div} A[f]),$$

substituting  $A[f]$  and  $\text{div}(A[f])$  above we obtain (2.3).  $\square$

**Remark.** If the solution is rotationally symmetric  $f(t, v) = f(t, |v|^2)$ , equation (2.3) reduces to the classical Fokker Planck equation. Let  $F$  be an antiderivative of  $f$ , for each  $t$  and  $i \neq j$  we have

$$P_{ij}(t) = \int w_i w_j f(|w|^2) dw = \int w_i \partial_{w_j} \left( \frac{F(|w|^2)}{2} \right) dw = 0.$$

For the case  $i = j$ , due to radial symmetry, the area under  $w_i^2 f(|w|^2)$  is the same for every  $i$ , so

$$P_{ii}(t) = \int w_i^2 f(|w|^2) dw = \frac{1}{d} \int |w|^2 f(|w|^2) dw = 1.$$

This shows that  $\mathbb{P}(t) = \text{Id}$ . Moreover, we have

$$A[f] \nabla f(v) = (|v|^2 K(v) + (d-1)\text{Id}) 2v f'(|v|^2) = (d-1) 2v f'(|v|^2) = (d-1) \nabla f(v)$$

and the Landau equation for Maxwell molecules reduces to

$$\partial_t f = (d-1) \text{div}(\nabla f(v) + v f(v)).$$

Next, we show how  $A[f]$  evolves in time:

**Proposition 2.** *Let  $f$  be a solution to (2.3) with  $\int f(1, v, |v|^2) dv = (1, 0, d)$ . The matrix  $\mathbb{P}$  with entries  $P_{ij}(t) = \int w_i w_j f(w) dw$  satisfies the ode*

$$(2.4) \quad \mathbb{P}'(t) = -4d(\mathbb{P}(t) - \text{Id}),$$

with explicit solution

$$\mathbb{P}(t) = \text{Id} + e^{-4dt} (\mathbb{P}_{in} - \text{Id}).$$

*Proof.* Using the equation for  $f$ , and  $\operatorname{div}(A[f]) = -(d-1)v$  we deduce

$$\begin{aligned} \frac{d}{dt} \int w_i w_j f(w) dw &= \int w_i w_j \operatorname{div}_v \left( A[f] \nabla f dw - \operatorname{div}_v(A[f]) f \right) dw \\ &= - \int \langle \nabla(w_i w_j), A[f] \nabla f \rangle dw + \int \langle \nabla(w_i w_j), \operatorname{div}(A[f]) \rangle f dw \\ &= \int \operatorname{div}(A[f] \nabla(w_i w_j)) f dw - (d-1) \int \langle \nabla(w_i w_j), w \rangle f dw. \end{aligned}$$

The first term simplifies as

$$\begin{aligned} \operatorname{div}(A[f] \nabla(w_i w_j)) &= \operatorname{div}(A[f]) \cdot \nabla(w_i w_j) + \operatorname{tr}(A[f] D^2(w_i w_j)) \\ &= -(d-1)w \cdot \nabla(w_i w_j) + \operatorname{tr}(A[f] D^2(w_i w_j)). \end{aligned}$$

Note that  $D^2(w_i w_j) = \Delta_{ij} + \Delta_{ji}$ , where  $\Delta_{ij}$  is the matrix with all zeroes except a one at position  $(i, j)$ . In general, if  $A$  is a matrix then  $A\Delta_{ij}$  has all of its columns equal to 0 except the  $j$ -th, where you find the  $i$ -th column of  $A$ . Therefore, if the matrix  $A$  is symmetric, we have  $\operatorname{tr}(AD^2(w_i w_j)) = 2A_{ij}$ , which shows:

$$\operatorname{div}(A[f] \nabla(w_i w_j)) = -(d-1)w \cdot (\vec{e}_i w_j + \vec{e}_j w_i) + 2A_{ij}[f] = -2(d-1)w_i w_j + 2A_{ij}[f].$$

Similarly, we have

$$\int \langle \nabla(w_i w_j), w \rangle f dw = 2 \int w_i w_j f(w) dw,$$

which shows

$$(2.5) \quad \frac{d}{dt} \int w_i w_j f(w) dw = 2 \int A_{ij}[f] f(w) dw - 4(d-1) \int w_i w_j f(w) dw.$$

Now we use the following expression for  $A[f]$

$$\begin{aligned} A[f] &= |v|^2 K(v) + d \operatorname{Id} - \mathbb{P}(t) \\ &= |v|^2 K(v) + \int |w|^2 \left( \operatorname{Id} - \frac{w \otimes w}{|w|^2} \right) f(w) dw, \end{aligned}$$

which we plug into (2.5), yielding

$$\begin{aligned} 2 \int A_{ij}[f](w) f(w) dw &= 2 \int \left( |w|^2 K(w) + \int |\xi|^2 K(\xi) f(\xi) d\xi \right) f(w) dw \\ &= 4d \operatorname{Id} - 4\mathbb{P}(t), \end{aligned}$$

which finishes the proof.  $\square$

**Remark.** The previous formula allows one to simplify the expression of  $\mathbb{P}(t)$ . Let  $f_{in}$  be a general initial data with second moment matrix  $\mathbb{P}_{in}$ . Let  $R$  be the rotation matrix such that  $R \mathbb{P}_{in} R^t = \operatorname{diag}(\lambda_1^{in}, \dots, \lambda_d^{in})$ . Then, the rotated initial data  $\tilde{f}_{in}(v) := f_{in}(R^{-1}v)$  has diagonal second moment matrix:

$$\mathbb{P}_{in} = \int (v \otimes v) f_{in}(R^{-1}v) dv = \int (Rv) \otimes (Rv) f_{in}(v) dv = R \mathbb{P}(0) R^t = \operatorname{diag}(\lambda_1^{in}, \dots, \lambda_d^{in}).$$

Without loss of generality, in Section 3 we will consider initial data such that  $\int w_i w_j f_{in} dw = 0$  if  $i \neq j$ . Then  $\mathbb{P}_{in}$  is a diagonal matrix and, in view of (2.4),  $\mathbb{P}(t)$  is also diagonal with elements

$\lambda_i(t) = 1 + e^{-4dt} (\lambda_i^{\text{in}} - 1)$ . Because of conservation of the second moment we have that for all times  $t \geq 0$ :

$$\text{tr}(\mathbb{P}(t)) = \sum_{i=1}^d \lambda_i(t) = d.$$

### 3. EXPONENTIAL DECAY OF RELATIVE $L^2$ NORM

We start the proof Theorem 1 with the following computation:

**Lemma 1.** *Let  $f$  be a solution of (2.3) with initial data such that  $\mathbb{P}_{\text{in}}$  is diagonal. We have*

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \int \left( \frac{f}{m} - 1 \right)^2 m dv &\leq - (d-1) \int \left| \nabla \left( \frac{f}{m} \right) \right|^2 m dv \\ &\quad + \sum_i (\lambda_i(t) - 1) \int \left( \left[ \partial_{v_i} \left( \frac{f}{m} \right) \right]^2 - \frac{1}{2} \frac{f^2}{m^2} v_i^2 \right) m dv, \end{aligned}$$

where  $\lambda_i(t) = \int w_i^2 f(w) dw = 1 + e^{-4dt} (\lambda_i^{\text{in}} - 1)$ .

*Proof.* We rewrite (2.3) as

$$\partial_t f = \text{div}(|v|^2 K(v) \nabla f) + d \text{div} \left( m \nabla \left( \frac{f}{m} \right) \right) - \text{div}(fv + \mathbb{P} \nabla f).$$

Because of conservation of mass,  $\int f^2/m dv$  and  $\int (f/m - 1)^2 m dv$  only differ by a constant, so

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \int \left( \frac{f}{m} - 1 \right)^2 m dv &= \int \frac{f}{m} \partial_t f dv \\ &= - \int \left\langle K(v) |v|^2 \nabla f(v), \nabla(f/m) \right\rangle dv \\ &\quad - d \int |\nabla(f/m)|^2 m dv + \int \langle fv + \mathbb{P} \nabla f, \nabla(f/m) \rangle dv. \end{aligned}$$

Since  $K(v) \nabla m = 0$ , the first term in the last expression is negative:

$$\left\langle K(v) |v|^2 \nabla f(v), \nabla(f/m) \right\rangle = \left\langle |v|^2 K(v) \nabla(f/m), \nabla(f/m) \right\rangle m \geq 0.$$

Therefore, we have

$$\frac{1}{2} \frac{d}{dt} \int \left( \frac{f}{m} - 1 \right)^2 m dv \leq -d \int |\nabla(f/m)|^2 m dv + I_1 + I_2,$$

where  $I_1 := \int f v \cdot \nabla(f/m) dv$  and  $I_2 := \int \langle \mathbb{P} \nabla f, \nabla(f/m) \rangle dv$ . Expanding  $\nabla(f/m)$  in  $I_1$  yields

$$\begin{aligned} I_1 &= \int f v \cdot \frac{\nabla f + fv}{m} dv = \int \frac{|v|^2}{m} f^2 dv + \int \frac{v}{m} \cdot \widehat{f \nabla f} dv \\ &= \int \frac{|v|^2}{m} f^2 dv - \frac{1}{2} \int f^2 \left( \frac{d + |v|^2}{m} \right) dv \\ &= \frac{1}{2} \int \frac{f^2}{m} (|v|^2 - d) dv. \end{aligned}$$

The identity  $\nabla f = m \nabla(f/m) - fv$  yields

$$\begin{aligned} I_2 &= \int \langle \mathbb{P}(m \nabla(f/m) - fv), \nabla(f/m) \rangle dv \\ &= \int \langle \mathbb{P} \nabla(f/m), \nabla(f/m) \rangle m dv - \int \langle \mathbb{P} v, f \nabla(f/m) \rangle dv. \end{aligned}$$

We rewrite the last term as

$$\begin{aligned} - \int \langle \mathbb{P} v, f \nabla(f/m) \rangle dv &= -\frac{1}{2} \int \mathbb{P} v \cdot \nabla(f^2/m^2) m dv \\ &= \frac{1}{2} \int \frac{f^2}{m^2} (m \operatorname{div}(\mathbb{P} v) - \mathbb{P} v \cdot v m) dv \\ &= \frac{d}{2} \int \frac{f^2}{m} dv - \frac{1}{2} \int \frac{f^2}{m} \langle \mathbb{P} v, v \rangle dv, \end{aligned}$$

since  $\operatorname{tr}(\mathbb{P}(t)) = d$ . This yields

$$\begin{aligned} I_1 + I_2 &= \frac{1}{2} \int \frac{f^2}{m} (|v|^2 - d) dv + \int \langle \mathbb{P} \nabla(f/m), \nabla(f/m) \rangle m dv \\ &\quad + \frac{d}{2} \int \frac{f^2}{m} dv - \frac{1}{2} \int \frac{f^2}{m} \langle \mathbb{P} v, v \rangle dv \\ &= \frac{1}{2} \int \frac{f^2}{m} |v|^2 dv - \frac{1}{2} \sum_i \lambda_i \int \frac{f^2}{m} v_i^2 dv + \sum_i \lambda_i \int [\partial_{v_i}(f/m)]^2 m dv \\ &= -\frac{1}{2} \sum_i (\lambda_i - 1) \int \frac{f^2}{m} v_i^2 dv + \sum_i (\lambda_i - 1) \int [\partial_{v_i}(f/m)]^2 m dv + \int |\nabla(f/m)|^2 m dv, \end{aligned}$$

which proves the lemma.  $\square$

In the next result we show a gaussian weighted Poincare inequality, which is a (slightly) sharper version of the one found in [15].

**Lemma 2.** *Let  $m(x) = (2\pi)^{-\frac{d}{2}} e^{-\frac{|x|^2}{2}}$  be the Gaussian weight on  $\mathbb{R}^d$ . Then for any smooth function  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$  we have*

$$\int_{\mathbb{R}^d} |x_i|^2 \varphi^2(x) m(x) dx \leq 2 \int_{\mathbb{R}^d} |\partial_{x_i}(x) \varphi|^2 m(x) dx + 2 \int_{\mathbb{R}^d} \varphi^2(x) m(x) dx,$$

Moreover, if  $\varphi$  has zero average  $\int_{\mathbb{R}^d} \varphi(x)m(x) dx = 0$  then we have

$$\int_{\mathbb{R}^d} \varphi^2(x)m(x) dx \leq \int |\nabla \varphi(x)|^2 m(x) dx.$$

*Proof.* We consider the basis of Hermite polynomials  $\{H^\alpha\}_{\alpha \in \mathbb{N}^d}$  defined by

$$H_\alpha(x) = \frac{(-1)^{|\alpha|}}{\sqrt{\alpha!}} e^{\frac{|x|^2}{2}} \partial_\alpha e^{-\frac{|x|^2}{2}}(x) = \text{He}_{\alpha_1}(x_1) \cdots \text{He}_{\alpha_d}(x_d),$$

where  $\alpha = (\alpha_1, \dots, \alpha_d)$  is a multi-index and  $\alpha! = \prod_{i=1}^d \alpha_i!$ . The Hermite polynomials form a complete orthonormal basis with respect to the Gaussian measure  $m(x)$ . Any smooth function  $\varphi$  can be expressed as

$$\varphi(x) = \sum_{\alpha} a_{\alpha} H_{\alpha}(x),$$

where  $a_{\alpha} = \int_{\mathbb{R}^d} \varphi(x) H_{\alpha}(x) m(x) dx$  are the coefficients of the expansion. We start with the first inequality. We have

$$\begin{aligned} \int_{\mathbb{R}^d} |x_1|^2 \varphi^2(x) m(x) dx &= \int_{\mathbb{R}^d} \left| \sum_{\alpha} a_{\alpha} x_1 H_{\alpha} \right|^2 m(x) dx \\ &= \int_{\mathbb{R}^d} \left| \sum_{\alpha} a_{\alpha} x_1 \text{He}_{\alpha_1}(x_1) \cdots \text{He}_{\alpha_d}(x_d) \right|^2 m(x) dx \\ &= \int_{\mathbb{R}^d} \left| \sum_{\alpha} a_{\alpha} (\text{He}_{\alpha_1+1}(x_1) + \text{He}'_{\alpha_1}(x_1)) \cdots \text{He}_{\alpha_d}(x_d) \right|^2 m(x) dx, \end{aligned}$$

where we have used the identity  $x \text{He}_n = \text{He}_{n+1} + \text{He}'_n$ . We thus have

$$\begin{aligned} \int_{\mathbb{R}^d} |x_1|^2 \varphi^2(x) m(x) dx &\leq 2 \int_{\mathbb{R}^d} \left| \sum_{\alpha} a_{\alpha} \text{He}_{\alpha_1+1}(x_1) \cdots \text{He}_{\alpha_d}(x_d) \right|^2 m(x) dx \\ &\quad + 2 \int_{\mathbb{R}^d} \left| \sum_{\alpha} a_{\alpha} \text{He}'_{\alpha_1}(x_1) \cdots \text{He}_{\alpha_d}(x_d) \right|^2 m(x) dx \\ &= 2 \int_{\mathbb{R}^d} |\varphi(x)|^2 m(x) dx + 2 \int_{\mathbb{R}^d} |\partial_{x_1} \varphi(x)|^2 m(x) dx. \end{aligned}$$

The Poincare inequality follows in a similar manner, using the identity  $\text{He}'_n = n \text{He}_{n-1}$ .  $\square$

We are now ready to prove Theorem 1:

*Proof of Theorem 1.* We start with Proposition 1, using that  $\lambda_i - 1 = e^{-4dt}(\lambda_i^{\text{in}} - 1)$ :

$$\begin{aligned} \frac{1}{2}E'(t) &\leq -(d-1) \int |\nabla(f/m)|^2 m dv + \sum_i (\lambda_i - 1) \int \left( [\partial_{v_i}(f/m)]^2 - \frac{1}{2} \frac{f^2}{m^2} v_i^2 \right) m dv \\ &= -(d-1) \int |\nabla(f/m)|^2 m dv + e^{-4dt} \sum_i (\lambda_i^{\text{in}} - 1) \int \left( [\partial_{v_i}(f/m)]^2 - \frac{1}{2} \frac{f^2}{m^2} v_i^2 \right) m dv \\ &= \sum_{i=1}^d \left[ -(d-1 - e^{-4dt}(\lambda_i^{\text{in}} - 1)) \int [\partial_{v_i}(f/m)]^2 m dv - \frac{1}{2} e^{-4dt}(\lambda_i^{\text{in}} - 1) \int \frac{f^2}{m} v_i^2 dv \right]. \end{aligned}$$

Because the  $\lambda_i^{\text{in}}$  are nonnegative, and their sum equals to  $d$ , at most  $d-1$  of them are greater than one and at most  $d-1$  of them are less than one. Also, the maximum value of  $(\lambda_i^{\text{in}} - 1)$  is  $d-1$  and the minimum is  $-1$ . Now, we note that

$$\begin{aligned} -\frac{1}{2} \sum_i (\lambda_i - 1) \int \left( \frac{f}{m} - 1 \right)^2 v_i^2 m dv &= -\frac{1}{2} \sum_i (\lambda_i - 1) \int \left( \frac{f^2}{m} v_i^2 - 2f v_i^2 + m v_i^2 \right) dv \\ &= -\frac{1}{2} \sum_i (\lambda_i - 1) \int \frac{f^2}{m} v_i^2 dv + \sum_i (\lambda_i - 1) \lambda_i, \end{aligned}$$

since  $\sum_i (\lambda_i - 1) = 0$ . Moreover  $\sum_i (\lambda_i - 1) \lambda_i = \sum_i (\lambda_i - 1)^2 \geq 0$ , which implies

$$-\frac{1}{2} \sum_i (\lambda_i - 1) \int \frac{f^2}{m} v_i^2 dv \leq -\frac{1}{2} \sum_i (\lambda_i - 1) \int \left( \frac{f}{m} - 1 \right)^2 v_i^2 m dv.$$

Going back to the original estimate, we have that

$$\begin{aligned} \frac{1}{2}E'(t) &\leq \sum_{i=1}^d \left[ -(d-1 - e^{-4dt}(\lambda_i^{\text{in}} - 1)) \int [\partial_{v_i}(f/m)]^2 m dv \right. \\ (3.1) \quad &\quad \left. - \frac{1}{2} e^{-4dt}(\lambda_i^{\text{in}} - 1) \int \left( \frac{f}{m} - 1 \right)^2 v_i^2 m dv \right]. \end{aligned}$$

For the  $\lambda_i^{\text{in}}$  greater or equal 1, the quantity in the square bracket is nonpositive, and we just trivially bound

$$\begin{aligned} &-(d-1 - e^{-4dt}(\lambda_i^{\text{in}} - 1)) \int [\partial_{v_i}(f/m)]^2 m dv - \frac{1}{2} e^{-4dt}(\lambda_i^{\text{in}} - 1) \int \left( \frac{f}{m} - 1 \right)^2 m v_i^2 dv \\ &\leq -(d-1 - e^{-4dt}(\lambda_i^{\text{in}} - 1)) \int [\partial_{v_i}(f/m)]^2 m dv \leq 0. \end{aligned}$$

For the terms with  $\lambda_i^{\text{in}} < 1$  in (3.1) we use lemma 2 (for  $\varphi = f/m - 1$ ) to obtain

$$\begin{aligned} &-(d-1 - e^{-4dt}(\lambda_i^{\text{in}} - 1)) \int [\partial_{v_i}(f/m)]^2 m dv + \frac{1}{2} e^{-4dt}(1 - \lambda_i^{\text{in}}) \int \left( \frac{f}{m} - 1 \right)^2 m v_i^2 dv \\ &\leq -(d-1) \int [\partial_{v_i}(f/m)]^2 m dv + e^{-4dt}(1 - \lambda_i^{\text{in}}) \int \left( \frac{f}{m} - 1 \right)^2 m dv. \end{aligned}$$

Assume now the worst case scenario in which only one  $\lambda_i^{\text{in}}$  is greater or equal than 1 and the rest are all less or equal than 1. Without loss of generality let  $\lambda_1^{\text{in}} \geq 1$  and  $\lambda_j^{\text{in}} \leq 1$  for  $j \geq 2$ .

We have that

$$\begin{aligned}
& \frac{1}{2} \frac{d}{dt} \int \left( \frac{f}{m} - 1 \right)^2 m dv \leq -(d-1 - e^{-4dt}(\lambda_1^{in} - 1)) \int [\partial_{v_1}(f/m)]^2 m dv \\
& - \sum_{i=2}^d (d-1) \int [\partial_{v_i}(f/m)]^2 m dv + \sum_{i=2}^d e^{-4dt}(1 - \lambda_i^{in}) \int \left( \frac{f}{m} - 1 \right)^2 m dv \\
& \leq -(d-1 - e^{-4dt}(\lambda_1^{in} - 1)) \int |\nabla(f/m)|^2 m dv + e^{-4dt}(\lambda_1^{in} - 1) \int \left( \frac{f}{m} - 1 \right)^2 m dv \\
& \leq -(d-1 - 2e^{-4dt}(\lambda_1^{in} - 1)) \int \left( \frac{f}{m} - 1 \right)^2 m dv,
\end{aligned}$$

where we have used the fact that  $\sum_{i=2}^d (1 - \lambda_i^{in}) = (\lambda_1^{in} - 1)$  and the Gaussian Poincare inequality. The ODE

$$y' \leq -2(d-1 - 2e^{-4dt}(\lambda_1^{in} - 1))y$$

has solution

$$y(t) \leq y_0 e^{-2(d-1)t + \frac{(\lambda_1^{in}-1)}{d}(1-e^{-4dt})}$$

and this finishes the proof.  $\square$

One can also repeat the computations of Lemma 1 with the test function  $\varphi = (f/m)^{p-1}$  and obtain

$$\begin{aligned}
& \frac{1}{p} \frac{d}{dt} \int (f/m)^p m dv \\
& \leq -\frac{4(p-1)}{p^2} \sum_i \left[ (d - \lambda_i) \int \left( \partial_{v_i}(f/m)^{\frac{p}{2}} \right)^2 m dv + \frac{p}{4}(\lambda_i - 1) \int (f/m)^p v_i^2 m dv \right].
\end{aligned}$$

As before, the quantity inside the bracket is nonpositive when  $\lambda_i \geq 1$ . For the case  $\lambda_i < 1$  we can use Lemma 2 with  $\varphi = (f/m)^{\frac{p}{2}}$ , which shows

$$\int (f/m)^p v_i^2 m dv \leq 2 \int \left( \partial_{v_i}(f/m)^{\frac{p}{2}} \right)^2 m dv + 2 \int (f/m)^p m dv,$$

yielding

$$\begin{aligned}
& \frac{1}{p} \frac{d}{dt} \int (f/m)^p m dv \leq -\frac{4(p-1)}{p^2} \sum_{\lambda_i < 1} (d - \lambda_i) \int \left( \partial_{v_i}(f/m)^{\frac{p}{2}} \right)^2 m dv \\
& + \frac{2(p-1)}{p} \sum_{\lambda_i < 1} (1 - \lambda_i) \int \left( \partial_{v_i}(f/m)^{\frac{p}{2}} \right)^2 m dv \\
& + \frac{2(p-1)}{p} \sum_{\lambda_i < 1} (1 - \lambda_i) \int (f/m)^p m dv \\
& \leq - \sum_{\lambda_i < 1} \left( d - \lambda_i - \frac{p}{2}(1 - \lambda_i) \right) \int \left( \partial_{v_i}(f/m)^{\frac{p}{2}} \right)^2 m dv \\
& + \frac{2(p-1)(d-1)}{p} \sup_i (1 - \lambda_i) \int (f/m)^p m dv.
\end{aligned}$$

Remembering that  $\lambda_i(t) = 1 + e^{-4dt} (\lambda_i^{\text{in}} - 1)$ , we have an inequality of the form

$$E'_p(t) \leq c_0 e^{-4dt} E_p(t), \quad c_0 = 2(p-1)(d-1) \sup_i (1 - \lambda_i^{\text{in}})$$

for  $t \geq t_0 := \frac{1}{4d} \log \left( \frac{p+2}{2(d-1)} \sup_i (1 - \lambda_i) \right)$ , which shows that

$$E_p(t) \leq \tilde{c}_0 E_p(t_0), \quad \tilde{c}_0 = \exp \left( \frac{c_0}{4d} e^{-4dt_0} \right).$$

In summary, we have proven the following:

**Proposition 3.** *Let  $f$  be solution to (2.3) with  $\int f_{\text{in}}(1, v, |v|^2) dv = (1, 0, d)$ . Let  $m$  be the corresponding unique equilibrium solution,  $m(v) := (2\pi)^{-\frac{d}{2}} e^{-\frac{1}{2}|v|^2}$  and consider the relative  $L^p$  norm  $E_p(t) := \int (f/m)^p m dv$ . Then there exists a constant  $C$  depending on  $d$ ,  $p$  and  $\lambda_i^{\text{in}}$  such that*

$$E_p(t) \leq C E_p(t_0), \quad t \geq t_0 := \frac{1}{4d} \log \left( \frac{p+2}{2(d-1)} \sup_i (1 - \lambda_i) \right).$$

#### 4. CONTRACTIVITY OF ENTROPIC 2-WASSERSTEIN DISTANCE

In this section we show that  $\mathcal{T}_\varepsilon(f, g)$  is decreasing as time grows for  $f$  and  $g$  any solutions of the Landau equation for Maxwell molecules. To prove monotonicity of  $\mathcal{T}_\varepsilon(f, g)$  we closely follow the strategy used in [1] for the Fisher information. Let  $q(f)$  be the Landau operator ( $\alpha \equiv 1$ )

$$(4.1) \quad q(f) := \operatorname{div}_v \int_{\mathbb{R}^3} |v-w|^2 K(v-w) (\nabla_v - \nabla_w) [f(v)f(w)] dw,$$

and define

$$(4.2) \quad Q(F) := (\operatorname{div}_v - \operatorname{div}_w) \left[ |v-w|^2 K(v-w) (\nabla_v - \nabla_w) F \right].$$

When  $F = f(v)f(w)$ , the operator  $Q(F)$  is connected with  $q(f)$  through an integration:

$$\int Q(f(v)f(w)) dw = q(f).$$

In the following we will denote with  $(f \otimes f)(v, w)$  the product

$$(f \otimes f)(v, w) := f(v)f(w).$$

The next lemma states that the first variation of the Wasserstein distance between  $f$  and  $g$  along the Landau flow can be symmetrized, and the resulting expression is the first variation of the Wasserstein distance in double variables along a linear problem.

**Proposition 4.** *Consider the operators  $q$ ,  $Q$  given in (4.1), (4.2). Let  $F = f \otimes f$ ,  $G = g \otimes g$ . We have*

$$(4.3) \quad \langle \mathcal{T}'_\varepsilon(f, g), (q(f), q(g)) \rangle = \frac{1}{2} \langle \mathcal{T}'_\varepsilon(F, G), (Q(F), Q(G)) \rangle.$$

*Proof.* To prove (4.3) we use the following dual formulation (see for instance [8]):

$$\mathcal{T}_\varepsilon(f, g) = \varepsilon \mathcal{H}(f) + \mathcal{J}(f, g), \quad \mathcal{J}(f, g) := \sup \left\{ \int u_1 g - \underbrace{\int u_0 f : \partial_s u + \frac{1}{2} |\nabla u|^2 + \frac{\varepsilon}{2} \Delta u = 0}_{\text{bHJB equation}} \right\},$$

where  $\mathcal{H}(f)$  is the entropy of  $f$  and  $u$  verifies the bHJB equation. As a first step, we show that the entropy functional verifies the desired equality. On the one hand, we have

$$\begin{aligned} \langle \mathcal{H}'(f), q(f) \rangle &= \int (\log f(v) + 1) q(f) dv = \int q(f) \log f(v) dv \\ &= - \iint \left\langle |v-w|^2 K(v-w) \frac{\nabla f(v)}{f(v)}, (\nabla_v - \nabla_w)(f(v)f(w)) \right\rangle dv dw \\ &= -\frac{1}{2} \iint \left\langle |v-w|^2 K(v-w) \left( \frac{\nabla f(v)}{f(v)} - \frac{\nabla f(w)}{f(w)} \right), (\nabla_v - \nabla_w)(f(v)f(w)) \right\rangle dv dw \\ &= -\frac{1}{2} \iint |v-w|^2 \left| K(v-w) \left( \frac{\nabla f(v)}{f(v)} - \frac{\nabla f(w)}{f(w)} \right) \right|^2 f(v)f(w) dv dw. \end{aligned}$$

On the other hand, we have

$$\begin{aligned} \langle \mathcal{H}'(F), Q(F) \rangle &= \iint (\log F(v,w) + 1) Q(F)(v,w) dv dw \\ &= \iint (\log f(v) + \log f(w)) Q(F)(v,w) dv dw \\ &= - \iint \left\langle \frac{\nabla f(v)}{f(v)} - \frac{\nabla f(w)}{f(w)}, |v-w|^2 K(v-w)(\nabla_v - \nabla_w)(f(v)f(w)) \right\rangle dv dw \\ &= 2 \langle \mathcal{H}'(f), q(f) \rangle. \end{aligned}$$

Now we show that  $\mathcal{J}(f, g)$  lifts in the same way as the entropy by relating the optimal  $u$  to the optimal plan. It is known that the optimal plan between  $f$  and  $g$  is of the form

$$\pi^*(v, w) = \exp \left\{ \varepsilon^{-1} \tilde{\phi}_\varepsilon(v) + \varepsilon^{-1} \tilde{\psi}_\varepsilon(w) \right\} r_{\frac{\varepsilon}{2}}(v, w),$$

where  $r_t(v, w)$  denotes the heat kernel at the time  $t$  and  $\tilde{\phi}_\varepsilon, \tilde{\psi}_\varepsilon$  are the Schrodinger potentials, which converge to the Kantorovich potentials  $\tilde{\phi}, \tilde{\psi}$  for the cost  $|v-w|^2$  as  $\varepsilon \rightarrow 0$ . Moreover,  $\pi^*$  is the only plan  $\pi \in \Pi(f, g)$  such that  $\frac{d\pi}{dr_{\frac{\varepsilon}{2}}}$  is tensorized and the Kantorovich potentials provide the optimal  $u$  for the dynamic formulation  $\mathcal{J}(f, g)$ , given by

$$u = \varepsilon \log \left[ h_{\frac{(1-s)\varepsilon}{2}} e^{\varepsilon^{-1} \tilde{\psi}_\varepsilon} \right],$$

where  $h_t$  denotes the heat flow so that  $\partial_t(h_t f) = \Delta(h_t f)$  and  $h_0 = \text{Id}$ . If we double the variables and consider  $F = f \otimes f$ ,  $G = g \otimes g$  then the optimal plan would be  $\pi^* \otimes \pi^*$  and the

optimal solution for  $\mathcal{J}(F, G)$  would be  $u(v) + u(w)$  giving:

$$\begin{aligned}
\langle \mathcal{J}'(F, G), (Q(F), Q(G)) \rangle &= \iint \left[ (u_1(v) + u_1(w)) Q(G) - (u_0(v) + u_0(w)) Q(F) \right] dv dw \\
&= 2 \iint u_1(v) Q(G) dv dw - 2 \iint u_0(v) Q(F) dv dw \\
&= 2 \int u_1(v) \operatorname{div}_v \int |v - w|^2 K(v - w) (\nabla_v - \nabla_w) G dw dv \\
&\quad - 2 \int u_0(v) \operatorname{div}_v \int |v - w|^2 K(v - w) (\nabla_v - \nabla_w) F dw dv \\
&= 2 \int u_1(v) q(g) dv - 2 \int u_0(v) q(f) dv \\
&= 2 \langle \mathcal{J}'(f, g), (q(f), q(g)) \rangle,
\end{aligned}$$

which finishes the proof.  $\square$

What the above proposition shows is that computing  $\frac{d}{dt} \mathcal{T}_\varepsilon(f, g)$  is equivalent to computing  $\frac{d}{dt} \mathcal{T}_\varepsilon(F, G)|_{t=0}$  where  $F$  and  $G$  solve the linear problem

$$\partial_t U = Q(U)$$

with initial conditions respectively  $f(v)f(w)$  and  $g(v)g(w)$ . Therefore in what follows we show that

$$\frac{d}{dt} \mathcal{T}_\varepsilon(F, G)|_{t=0} \leq 0.$$

As a first step, we rewrite the operator  $Q(\cdot)$  defined in (4.2) as composition of two linear operators. We recall the vector fields  $b_k$

$$(4.4) \quad b_1(v - w) = \begin{pmatrix} 0 \\ w_3 - v_3 \\ v_2 - w_2 \end{pmatrix}, \quad b_2(v - w) = \begin{pmatrix} v_3 - w_3 \\ 0 \\ w_1 - v_1 \end{pmatrix}, \quad b_3(v - w) = \begin{pmatrix} w_2 - v_2 \\ v_1 - w_1 \\ 0 \end{pmatrix},$$

and  $\tilde{b}_k(v - w) = \begin{pmatrix} b_k \\ -b_k \end{pmatrix}$ . The vectors  $b_k$  span the subspace  $\langle v - w \rangle^\perp$  and verify

$$b_1 \otimes b_1 + b_2 \otimes b_2 + b_3 \otimes b_3 = |v - w|^2 K(v - w).$$

Using that  $\tilde{b}_k(v - w)$  is divergence free and perpendicular to  $\begin{pmatrix} v - w \\ w - v \end{pmatrix}$  it is easy to obtain the expression

$$(4.5) \quad Q(F) = \sum_{k=1}^3 \tilde{b}_k \cdot \nabla \left( \tilde{b}_k \cdot \nabla F \right) = \sum_{k=1}^3 \left( L_{\tilde{b}_k} \circ L_{\tilde{b}_k} \right) F,$$

where

$$L_b(F) := b \cdot \nabla F.$$

Here the key observation is that the flow of the vector field  $\tilde{b}_k(v - w)$  induces an isometry in  $\mathbb{R}^6$ . That is,  $\begin{pmatrix} v(t) \\ w(t) \end{pmatrix}$ ,  $\begin{pmatrix} \bar{v}(t) \\ \bar{w}(t) \end{pmatrix}$  are solutions of the ODE  $\begin{pmatrix} \dot{v} \\ \dot{w} \end{pmatrix} = \tilde{b}_k(v - w)$  then:

$$\frac{1}{2} \frac{d}{dt} \left| \begin{pmatrix} v \\ w \end{pmatrix} - \begin{pmatrix} \bar{v} \\ \bar{w} \end{pmatrix} \right|^2 = 0.$$

Consequently, the entropic 2-Wasserstein distance will remain unchanged if we transport two densities along  $\tilde{b}_k$ , as stated in the following lemma:

**Lemma 3.** *Consider two functions  $F_k(v, w, t)$ ,  $G_k(v, w, t)$  satisfying the equations*

$$(4.6) \quad \begin{cases} \partial_t F_k = L_{\tilde{b}_k} F_k = \tilde{b}_k(v - w) \cdot \nabla F_k \\ F_k(v, w, 0) = F(v, w) \end{cases}, \quad \begin{cases} \partial_t G_k = L_{\tilde{b}_k} G_k = \tilde{b}_k(v - w) \cdot \nabla G_k \\ G_k(v, w, 0) = G(v, w) \end{cases}.$$

*Then  $t \mapsto \mathcal{T}_\varepsilon(F_k, G_k)$  is constant, which implies  $\langle \mathcal{T}'_\varepsilon(F_k, G_k), (L_{\tilde{b}_k} F_k, L_{\tilde{b}_k} G_k) \rangle = 0$ .*

*Proof.* Firstly note that we can write (4.6) as continuity equations using that  $\tilde{b}_k$  are divergence free:

$$\partial_t F_k + \operatorname{div}(-\tilde{b}_k F_k) = 0, \quad \partial_t G_k + \operatorname{div}(-\tilde{b}_k G_k) = 0.$$

Then it is known (see [17, Chap. 4]) that  $F_k(\cdot, t) = (X_t)_\# F$ ,  $G_k(\cdot, t) = (X_t)_\# G$  where  $X_t(v, w)$  represents the flow of the ODE  $\begin{pmatrix} \dot{v} \\ \dot{w} \end{pmatrix} = -\tilde{b}_k(v - w)$ . By the previous Remark, we see that  $X_t$  is an isometry in  $\mathbb{R}^6$ , which shows

$$\mathcal{T}_\varepsilon(F_k(\cdot, t), G_k(\cdot, t)) = \mathcal{T}_\varepsilon((X_t)_\# F, (X_t)_\# G) = \mathcal{T}_\varepsilon(F, G). \quad \square$$

Now we are ready to prove the monotonicity of  $\mathcal{T}_\varepsilon(f, g)$  where  $f, g$  are solutions of the Landau equation for Maxwell molecules.

*Proof of Theorem 2.* Due to Proposition 4, it is enough to show that  $\frac{d}{dt}|_{t=0} \mathcal{T}_\varepsilon(F, G) \leq 0$  where  $F, G$  are solutions of the lifted equations  $\partial_t F = Q(F)$ ,  $\partial_t G = Q(G)$ . We rewrite this as

$$\langle \mathcal{T}'_\varepsilon(F, G), (Q(F), Q(G)) \rangle \leq 0.$$

Consider for  $k = 1, 2, 3$  solutions of

$$\begin{cases} \partial_t F_k(v, w, t) = L_{\tilde{b}_k}(F_k) \\ F_k(v, w, 0) = F(v, w) \end{cases}, \quad \begin{cases} \partial_t G_k(v, w, t) = L_{\tilde{b}_k}(G_k) \\ G_k(v, w, 0) = G(v, w) \end{cases}.$$

Since  $\mathcal{T}_\varepsilon(F_k, G_k)$  is constant, we deduce that

$$\begin{aligned} 0 &= \frac{d^2}{dt^2} \Big|_{t=0} \mathcal{T}_\varepsilon(F_k, G_k) \\ &= \langle \mathcal{T}'_\varepsilon(F, G), (\partial_{tt} F_k(0), \partial_{tt} G_k(0)) \rangle + \langle \mathcal{T}''_\varepsilon(F, G)(\partial_t F_k(0), \partial_t G_k(0)), (\partial_t F_k(0), \partial_t G_k(0)) \rangle \\ &= \langle \mathcal{T}'_\varepsilon(F, G), ((L_{\tilde{b}_k} \circ L_{\tilde{b}_k})F, (L_{\tilde{b}_k} \circ L_{\tilde{b}_k})G) \rangle + \langle \mathcal{T}''_\varepsilon(F, G)(L_{\tilde{b}_k} F, L_{\tilde{b}_k} G), (L_{\tilde{b}_k} F, L_{\tilde{b}_k} G) \rangle. \end{aligned}$$

Then, the expression (4.5) for the operator  $Q$  shows

$$\begin{aligned} \langle \mathcal{T}'_\varepsilon(F, G), (Q(F), Q(G)) \rangle &= \sum_{k=1}^3 \left\langle \mathcal{T}'_\varepsilon(F, G), ((L_{\tilde{b}_k} \circ L_{\tilde{b}_k})F, (L_{\tilde{b}_k} \circ L_{\tilde{b}_k})G) \right\rangle \\ (4.7) \quad &= - \sum_{k=1}^3 \left\langle \mathcal{T}''_\varepsilon(F, G)(L_{\tilde{b}_k}F, L_{\tilde{b}_k}G), (L_{\tilde{b}_k}F, L_{\tilde{b}_k}G) \right\rangle \leq 0, \end{aligned}$$

where we have used that  $\mathcal{T}_\varepsilon$  is linearly convex, as evidenced by the dual formulation (4.3).  $\square$

**Remark.** Inequality (4.7) would provide an explicit expression of the Wasserstein dissipation if one has an explicit formula for  $\mathcal{T}''_\varepsilon(F, G)$ . For  $\varepsilon > 0$  finding this explicit expression is still an open problem. For  $\varepsilon = 0$  the first and second author have formally proven in [18] that

$$\begin{aligned} - \sum_{k=1}^3 \left\langle \mathcal{T}''(F, G)(L_{\tilde{b}_k}(F), L_{\tilde{b}_k}(G)), (L_{\tilde{b}_k}(F), L_{\tilde{b}_k}(G)) \right\rangle \\ (4.8) \quad = -2 \sum_{k=1}^3 \int \left| D^2\Phi(v, w)^{-\frac{1}{2}} \nabla \xi_k^f(v, w) - D^2\Psi(\bar{v}, \bar{w})^{-\frac{1}{2}} \nabla \xi_k^g(\bar{v}, \bar{w}) \right|^2 \Pi_* \left[ \begin{pmatrix} v \\ w \end{pmatrix}, \begin{pmatrix} \bar{v} \\ \bar{w} \end{pmatrix} \right], \end{aligned}$$

where  $\xi_k^f, \xi_k^g$  solve

$$\begin{aligned} (4.9) \quad -\operatorname{div}\left(f(v)f(w)D^2\Phi(v, w)^{-1}\nabla\xi_k^f(v, w)\right) &= \tilde{b}_k(v-w) \cdot \nabla(f(v)f(w)), \\ -\operatorname{div}\left(g(v)g(w)D^2\Psi(v, w)^{-1}\nabla\xi_k^g(v, w)\right) &= \tilde{b}_k(v-w) \cdot \nabla(g(v)g(w)), \end{aligned}$$

and  $\Phi, \Psi$  are the convex functions mapping  $(\nabla\Phi)_\#(f \otimes f) = g \otimes g$ ,  $(\nabla\Psi)_\#(g \otimes g) = f \otimes f$ . Formula (4.8) shows that the distance to the Maxwellian  $\mathcal{T}(f, m)$  is strictly decreasing in time unless  $f = m$ . Indeed, assume  $\frac{d}{dt}\mathcal{T}(f, m) = 0$ , then we have that

$$0 = \frac{d}{dt}\mathcal{T}(f, m) = -\frac{1}{2} \sum_{k=1}^3 \int \left\langle D^2\Phi(v, w) \nabla \xi_k^f, \nabla \xi_k^f \right\rangle f(v)f(w) dv dw,$$

where we are using that  $\nabla \xi_k^m \equiv 0$ . This shows that  $\nabla \xi_k^f \equiv 0$  as well. Then, remembering that  $b_k(v-w)$ ,  $k = 1, 2, 3$  spans  $(v-w)^\perp$ , (4.9) implies that

$$K(v-w)(f(w)\nabla f(v) - f(v)\nabla f(w)) = 0.$$

The above identity is satisfied if and only if  $f = m$  (among the functions with mass 1, second moment 3 and first momentum equal to 0).

**Remark.** The method proposed in this section is not suitable when  $\alpha \not\equiv 1$ . As noted in [1], the flow of the ODE

$$\begin{pmatrix} \dot{v} \\ \dot{w} \end{pmatrix} = \sqrt{\alpha(|v-w|)} \tilde{b}_k(v-w),$$

no longer generates an isometry and Lemma 3 does not hold. The next section provides an alternative method that adapts to any  $\alpha$ .

## 5. PROOF OF THEOREM 2: SECOND METHOD

In this section we prove Theorem 2 using a different method. During the proof we will introduce tools that will be also useful for the proof of Theorem 3; therefore we will keep the term  $\alpha$  in the computations.

*Intuition using the heat equation.* Before starting our computation on the Landau equation, we illustrate our intuition by presenting our method applied to the heat equation. The argument is based on the dual formulation of the metric (see [8]):

$$(5.1) \quad \mathcal{T}(f, g) = \sup \left\{ \int u_1 g \, dv - \underbrace{\int u_0 f \, dv : \partial_s u + \frac{1}{2} |\nabla u|^2 = 0}_{\text{HJ equation}} \right\}.$$

We have the following proposition:

**Proposition.** *Consider two nonnegative solutions in  $\mathbb{R}^d$  of the heat equation  $\partial_t f = \Delta f$ ,  $\partial_t g = \Delta g$ , both with mass 1. Then the Wasserstein distance  $\mathcal{T}(f, g)$  is decreasing with dissipation*

$$\frac{d}{dt} \mathcal{T}(f, g) = - \int_0^1 \int_{\mathbb{R}^d} |D^2 u|^2 \rho(x) \, dx \, ds,$$

where  $\rho(x, s)$ ,  $0 \leq s \leq 1$  is an interpolation between  $f$  and  $g$  given by the equations

$$(5.2) \quad \partial_s \rho + \operatorname{div}(\rho \nabla u) = 0, \quad \rho_0 = f, \quad \rho_1 = g, \quad \partial_s u + \frac{1}{2} |\nabla u|^2 = 0.$$

**Remark 1.** *The previous result is analogous to the estimate of contractivity for finite dimensional gradient flows. Given a smooth function  $F : \mathbb{R}^d \rightarrow \mathbb{R}$ , consider two solutions of the gradient flow  $\dot{x} = -\nabla F(x)$ ,  $\dot{y} = -\nabla F(y)$ . Then the squared Euclidean distance  $|x - y|^2$  is decreasing with dissipation*

$$\frac{1}{2} \frac{d}{dt} |x - y|^2 = -\langle \nabla F(p_0), \dot{p}_0 \rangle + \langle \nabla F(p_1), \dot{p}_1 \rangle = - \int_0^1 D^2 F(p_s)[\dot{p}_s, \dot{p}_s] \, ds,$$

where  $p_s = sx + (1-s)y$ ,  $0 \leq s \leq 1$  is the constant speed geodesic between  $x$  and  $y$ .

*Proof.* We use the dual formulation of the Wasserstein distance (5.1) It is known that the optimal  $u$  provides an interpolation between  $f$  and  $g$  given by the equation (5.2). Therefore, we have

$$\begin{aligned} \frac{d}{dt} \mathcal{T}(f, g) &= \int_{\mathbb{R}^d} u_1 \Delta g \, dx - \int_{\mathbb{R}^d} u_0 \Delta f \, dx \\ &= \int_0^1 \frac{d}{ds} \int_{\mathbb{R}^d} \Delta u \rho \, dx \, ds \\ &= - \int_0^1 \int_{\mathbb{R}^d} \left( \frac{1}{2} \Delta |\nabla u|^2 \rho + \Delta u \operatorname{div}(\rho \nabla u) \right) \, dx \, ds \\ &= - \int_0^1 \int_{\mathbb{R}^d} \left( \frac{1}{2} \Delta |\nabla u|^2 - \nabla \Delta u \cdot \nabla u \right) \rho \, dx \, ds \\ &= -\frac{1}{2} \int_0^1 \int_{\mathbb{R}^d} |D^2 u|^2 \rho \, dx \, ds, \end{aligned}$$

where we have used Bochner's formula  $\frac{1}{2}\Delta|\nabla u|^2 = \nabla\Delta u \cdot \nabla u + |D^2u|^2$ .  $\square$

The 2-Wasserstein contractivity of the heat equation is a classical result. In fact, Jordan, Kinderlehrer and Otto showed in [13] how the heat equation could be understood as a gradient flow for the entropy functional in the 2-Wasserstein space. Contractivity without a rate formally follows by the displacement convexity of the entropy functional.

Let's now use the same method for the Landau equation. We start by computing  $\frac{d}{dt}\mathcal{T}(f, g)$  ( $\varepsilon = 0$ ) where  $f, g$  are solutions of the Landau equation for a general  $\alpha$ . We recall that  $\mathcal{T}(f, g)$  is defined in (5.1). Moreover, the supremum is attained at a function  $u$  which provides the constant speed geodesic  $\rho_s(v) = \rho(v, s)$ ,  $0 \leq s \leq 1$  between  $f$  and  $g$  through the continuity equation

$$(5.3) \quad \partial_s\rho + \operatorname{div}(\rho\nabla u) = 0, \quad \rho_0 = f, \quad \rho_1 = g.$$

**Lemma 4.** *Let  $f, g$  be two smooth solutions to (1.1), we have*

$$\begin{aligned} \frac{d}{dt}\mathcal{T}(f, g) &= -\frac{1}{2}\sum_{k=1}^3 \int_0^1 \iint \left| \nabla(\tilde{b}_k \cdot \nabla u) \right|^2 \alpha\rho dv dw ds \\ &\quad + \frac{1}{2}\sum_{k=1}^3 \int_0^1 \iint \frac{\alpha'}{r} (\tilde{b}_k \cdot \nabla(\tilde{b}_k \cdot \nabla u)) (\hat{n} \cdot \nabla u) \rho dv dw ds. \end{aligned}$$

*Proof.* Using Proposition 4 and the dual formulation (5.1) we see that

$$\frac{d}{dt}\mathcal{T}(f, g) = \frac{1}{2} \iint u_1 Q(G) dv dw - \frac{1}{2} \iint u_0 Q(F) dv dw,$$

where  $F(v, w) = f(v)f(w)$ ,  $G(v, w) = g(v)g(w)$  and  $u(v, w, s)$  is a solution in  $\mathbb{R}^6$  of the Hamilton-Jacobi equation from (5.1). We note that the previous expression can be rewritten as

$$\frac{d}{dt}\mathcal{T}(f, g) = \frac{1}{2} \int_0^1 \frac{d}{ds} \iint u Q(\rho) dv dw ds,$$

where  $\rho$  is the Wasserstein geodesic interpolating  $F$  and  $G$ . To lighten notation, we drop the integration with respect to  $s$  and focus on rewriting  $\frac{d}{ds} \iint u Q(\rho) dv dw ds$ . For this purpose, we use the decomposition

$$Q(F) = \alpha \sum_{k=1}^3 \tilde{b}_k \cdot \nabla(\tilde{b}_k \cdot \nabla F),$$

where from now on  $\nabla$  and  $\operatorname{div}$  denote gradient and divergence with respect to  $v, w$ . We will repeatedly use that  $\alpha\tilde{b}_k$  is divergence free, which shows that  $Q$  is self-adjoint:

$$\begin{aligned} \frac{d}{ds} \iint u Q(\rho) dv dw &= \frac{d}{ds} \iint Q(u)\rho dv dw \\ &= - \sum_k \iint \left[ \alpha\tilde{b}_k \cdot \nabla(\tilde{b}_k \cdot \nabla u) \operatorname{div}(\rho\nabla u) + \frac{1}{2}\alpha\tilde{b}_k \cdot \nabla(\tilde{b}_k \cdot \nabla|\nabla u|^2)\rho \right] dv dw. \end{aligned}$$

Introducing  $r = |v - w|$  and  $\hat{n} = \begin{pmatrix} v - w \\ w - v \end{pmatrix}$ , we see that  $\nabla\alpha(r) = \frac{\alpha'(r)}{r}\hat{n}$  so that

$$(5.4) \quad \begin{aligned} \frac{d}{dt}\mathcal{T}(f, g) &= \frac{1}{2} \sum_k \iint \left[ \nabla(\tilde{b}_k \cdot \nabla(\tilde{b}_k \cdot \nabla u)) \cdot \nabla u - \frac{1}{2} \tilde{b}_k \cdot \nabla(\tilde{b}_k \cdot \nabla |\nabla u|^2) \right] \alpha \rho \\ &\quad + \frac{1}{2} \sum_k \iint \frac{\alpha'}{r} (\tilde{b}_k \cdot \nabla(\tilde{b}_k \cdot \nabla u)) (\hat{n} \cdot \nabla u) \rho. \end{aligned}$$

Our next goal is to rewrite (5.4) as

$$(5.5) \quad \begin{aligned} \frac{d}{dt}\mathcal{T}(f, g) &= -\frac{1}{2} \sum_{k=1}^3 \int_0^1 \iint \left| \nabla(\tilde{b}_k \cdot \nabla u) \right|^2 \alpha \rho \, dv \, dw \, ds \\ &\quad + \frac{1}{2} \sum_{k=1}^3 \int_0^1 \iint \frac{\alpha'}{r} (\tilde{b}_k \cdot \nabla(\tilde{b}_k \cdot \nabla u)) (\hat{n} \cdot \nabla u) \rho \, dv \, dw \, ds. \end{aligned}$$

The last term in (5.4) already matches (5.5), so we focus on the first term of the right hand side of (5.4). We decompose it using the basis vectors

$$\tilde{e}_i = \begin{pmatrix} e_i \\ e_i \end{pmatrix}, \quad \hat{e}_i = \begin{pmatrix} e_i \\ -e_i \end{pmatrix}, \quad 1 \leq i \leq 3,$$

where  $e_i$  form the canonical basis of  $\mathbb{R}^3$ . More precisely, we use the decompositions

$$\begin{aligned} \nabla(\tilde{b}_k \cdot \nabla(\tilde{b}_k \cdot \nabla u)) \cdot \nabla u &= \frac{1}{2} \sum_{i=1}^3 \left( \tilde{e}_i \cdot \nabla(\tilde{b}_k \cdot \nabla(\tilde{b}_k \cdot \nabla u)) \right) (\tilde{e}_i \cdot \nabla u) \\ &\quad + \frac{1}{2} \sum_{i=1}^3 \left( \hat{e}_i \cdot \nabla(\tilde{b}_k \cdot \nabla(\tilde{b}_k \cdot \nabla u)) \right) (\hat{e}_i \cdot \nabla u), \end{aligned}$$

and

$$\tilde{b}_k \cdot \nabla(\tilde{b}_k \cdot \nabla|\nabla u|^2) = \frac{1}{2} \sum_{i=1}^3 \tilde{b}_k \cdot \nabla(\tilde{b}_k \cdot \nabla(\hat{e}_i \cdot \nabla u)^2) + \frac{1}{2} \sum_{i=1}^3 \tilde{b}_k \cdot \nabla(\tilde{b}_k \cdot \nabla(\hat{e}_i \cdot \nabla u)^2).$$

This allows us to rewrite (5.4) as

$$(5.6) \quad \begin{aligned} \frac{d}{dt}\mathcal{T}(f, g) &= \frac{1}{4} \sum_{i,k} \int_0^1 [I_1(\tilde{e}_i, k) + I_2(\tilde{e}_i, k)] \, ds + \frac{1}{4} \sum_{i,k} \int_0^1 [I_1(\hat{e}_i, k) + I_2(\hat{e}_i, k)] \, ds \\ &\quad + \frac{1}{2} \sum_k \int_0^1 \iint \frac{\alpha'}{r} (\tilde{b}_k \cdot \nabla(\tilde{b}_k \cdot \nabla u)) (\hat{n} \cdot \nabla u) \rho \, ds, \end{aligned}$$

where we set for a fixed vector  $e \in \mathbb{R}^6$  the quantities

$$\begin{aligned} I_1(e, k) &:= \iint e \cdot \nabla(\tilde{b}_k \cdot \nabla(\tilde{b}_k \cdot \nabla u)) (e \cdot \nabla u) \alpha \rho, \\ I_2(e, k) &:= -\frac{1}{2} \iint \tilde{b}_k \cdot \nabla(\tilde{b}_k \cdot \nabla[(e \cdot \nabla u)^2]) \alpha \rho. \end{aligned}$$

The advantage of using  $\tilde{e}_i$ ,  $\hat{e}_i$  over the canonical basis is the simplicity of the Lie brackets  $[\tilde{e}_i, \tilde{b}_k]$ ,  $[\hat{e}_i, \tilde{b}_k]$ , which we will call commutators from now on. We remind the reader that if

$a = a(v)$ ,  $b = b(v)$  are vector fields then Lie bracket  $[a, b]$  is another vector field that verifies the identity

$$a \cdot \nabla(b \cdot \nabla u) - b \cdot \nabla(a \cdot \nabla u) = [a, b] \cdot \nabla u$$

for any smooth function  $u$ . The components of  $[a, b]$  are defined as

$$[a, b]_i = \sum_j (a_j \partial_{v_j} b_i - b_j \partial_{v_j} a_i) = a \cdot \nabla b_i - b \cdot \nabla a_i.$$

We calculate, all the possible commutators. We have that  $[\tilde{e}_i, \tilde{b}_k] = 0$  for every  $i$  and  $k$ , and that  $[\hat{e}_i, \tilde{b}_k]$  satisfies:

$[\hat{e}_i, \tilde{b}_k]$	$\tilde{b}_1$	$\tilde{b}_2$	$\tilde{b}_3$
$\hat{e}_1$	0	$-2\hat{e}_3$	$2\hat{e}_2$
$\hat{e}_2$	$2\hat{e}_3$	0	$-2\hat{e}_1$
$\hat{e}_3$	$-2\hat{e}_2$	$2\hat{e}_1$	0

Now we rewrite  $I_1(e, k)$  by commuting  $e$ ,  $\tilde{b}_k$  and integrating by parts keeping in mind that  $\operatorname{div}(\alpha \tilde{b}_k) = 0$ :

$$\begin{aligned} I_1(e, k) &= - \iint (e \cdot \nabla (\tilde{b}_k \cdot \nabla u))^2 \alpha \rho - \iint (e \cdot \nabla (\tilde{b}_k \cdot \nabla u)) (\tilde{b}_k \cdot \nabla \rho) (e \cdot \nabla u) \alpha \\ &\quad + \iint ([e, \tilde{b}_k] \cdot \nabla (\tilde{b}_k \cdot \nabla u)) (e \cdot \nabla u) \alpha \rho + \iint (e \cdot \nabla (\tilde{b}_k \cdot \nabla u)) ([e, \tilde{b}_k] \cdot \nabla u) \alpha \rho. \end{aligned}$$

Similarly,  $I_2(e, k)$  integrating by parts and commuting  $\tilde{b}_k$ ,  $e$ :

$$I_2(e, k) = \iint (e \cdot \nabla (\tilde{b}_k \cdot \nabla u)) (e \cdot \nabla u) (\tilde{b}_k \cdot \nabla \rho) \alpha - \iint ([e, \tilde{b}_k] \cdot \nabla u) (e \cdot \nabla u) (\tilde{b}_k \cdot \nabla \rho) \alpha.$$

Note that the second to last term cancels with one of  $I_1(e, k)$ , yielding

$$\begin{aligned} I_1(e, k) + I_2(e, k) &= - \iint (e \cdot \nabla (\tilde{b}_k \cdot \nabla u))^2 \alpha \rho - \iint ([e, \tilde{b}_k] \cdot \nabla u) (e \cdot \nabla u) (\tilde{b}_k \cdot \nabla \rho) \alpha \\ &\quad + \iint ([e, \tilde{b}_k] \cdot \nabla (\tilde{b}_k \cdot \nabla u)) (e \cdot \nabla u) \alpha \rho + \iint (e \cdot \nabla (\tilde{b}_k \cdot \nabla u)) ([e, \tilde{b}_k] \cdot \nabla u) \alpha \rho. \end{aligned}$$

Now, to compute (5.6) we must set  $e = e_i$  and sum in  $i$ . Using that  $[\tilde{e}_i, \tilde{b}_k] = 0$  for all  $i, k$ , we have

$$\frac{1}{4} \sum_i [I_1(\tilde{e}_i, k) + I_2(\tilde{e}_i, k) + I_1(\hat{e}_i, k) + I_2(\hat{e}_i, k)] = -\frac{1}{2} \iint |\nabla (\tilde{b}_k \cdot \nabla u)|^2 \alpha \rho$$

$$\begin{aligned} (5.7) \quad &\quad - \frac{1}{4} \sum_{i=1}^3 \iint ([\hat{e}_i, \tilde{b}_k] \cdot \nabla u) (\hat{e}_i \cdot \nabla u) (\tilde{b}_k \cdot \nabla \rho) \alpha \\ (5.8) \quad &\quad + \frac{1}{4} \sum_{i=1}^3 \iint ([\hat{e}_i, \tilde{b}_k] \cdot \nabla (\tilde{b}_k \cdot \nabla u)) (\hat{e}_i \cdot \nabla u) \alpha \rho \end{aligned}$$

$$(5.9) \quad + \frac{1}{4} \sum_{i=1}^3 \iint (\hat{e}_i \cdot \nabla (\tilde{b}_k \cdot \nabla u)) ([\hat{e}_i, \tilde{b}_k] \cdot \nabla u) \alpha \rho.$$

Now observe that (5.7) is symmetric in  $\hat{e}_i$ ,  $[\hat{e}_i, \tilde{b}_k]$ . Since the table of commutators is anti-symmetric, we deduce that actually (5.7) vanishes. A similar reasoning shows that (5.8) and (5.9) vanish. Plugging all this information into (5.6) shows that

$$(5.10) \quad \begin{aligned} \frac{d}{dt} \mathcal{T}(f, g) = & -\frac{1}{2} \sum_{k=1}^3 \int_0^1 \iint \left| \nabla (\tilde{b}_k \cdot \nabla u) \right|^2 \alpha \rho \, dv \, dw \, ds \\ & + \frac{1}{2} \sum_{k=1}^3 \int_0^1 \iint \frac{\alpha'}{r} (\tilde{b}_k \cdot \nabla (\tilde{b}_k \cdot \nabla u)) (\hat{n} \cdot \nabla u) \rho \, dv \, dw \, ds. \end{aligned}$$

This concludes the proof.  $\square$

With these preliminary computations for  $\varepsilon = 0$ , we are ready to prove Theorem 2:

*Proof. Step 1.* We show that the operator

$$Q(u) := (\operatorname{div}_v - \operatorname{div}_w) \left[ |v - w|^2 K(v - w) (\nabla_v u - \nabla_w u) \right]$$

commutes with the Laplacian  $\Delta_{v,w}$ , namely

$$Q\Delta_{v,w}u = \Delta_{v,w}Qu.$$

We use the vectors  $\tilde{e}_i = \begin{pmatrix} e_i \\ e_i \end{pmatrix}$ ,  $\hat{e}_i = \begin{pmatrix} e_i \\ -e_i \end{pmatrix}$  and  $\tilde{b}_k = \begin{pmatrix} b_k \\ -b_k \end{pmatrix}$ , noting that

$$\Delta u = \frac{1}{2} \sum_i \tilde{e}_i \cdot \nabla (\tilde{e}_i \cdot \nabla u) + \frac{1}{2} \sum_i \hat{e}_i \cdot \nabla (\hat{e}_i \cdot \nabla u), \quad Qu = \sum_k \tilde{b}_k \cdot \nabla (\tilde{b}_k \cdot \nabla u).$$

We have

$$(5.11) \quad \begin{aligned} \Delta Qu &= \frac{1}{2} \sum_{i,k} \tilde{e}_i \cdot \nabla \left( \tilde{e}_i \cdot \nabla \left( \tilde{b}_k \cdot \nabla (\tilde{b}_k \cdot \nabla u) \right) \right) + \frac{1}{2} \sum_{i,k} \hat{e}_i \cdot \nabla \left( \hat{e}_i \cdot \nabla \left( \tilde{b}_k \cdot \nabla (\tilde{b}_k \cdot \nabla u) \right) \right) \\ (5.12) \quad Q\Delta u &= \frac{1}{2} \sum_{i,k} \tilde{b}_k \cdot \nabla \left( \tilde{b}_k \cdot \nabla (\tilde{e}_i \cdot \nabla (\tilde{e}_i \cdot \nabla u)) \right) + \frac{1}{2} \sum_{i,k} \tilde{b}_k \cdot \nabla \left( \tilde{b}_k \cdot \nabla (\hat{e}_i \cdot \nabla (\hat{e}_i \cdot \nabla u)) \right). \end{aligned}$$

Therefore, we may begin by finding an expression for the commutator

$$\begin{aligned} [L_e, L_b] u &:= L_e L_b u - L_b L_e u \\ &= e \cdot \nabla (e \cdot \nabla (b \cdot \nabla (b \cdot \nabla u))) - b \cdot \nabla (b \cdot \nabla (e \cdot \nabla (e \cdot \nabla u))), \end{aligned}$$

where  $e$ ,  $b$  are vectors in  $\mathbb{R}^3$  which may depend on  $v$  and  $w$  and  $L_e u := e \cdot \nabla (e \cdot \nabla u)$ ,  $L_b u := b \cdot \nabla (b \cdot \nabla u)$ . To do this we commute the order of differentiation to make the sequence  $[e, b, b, e]$ .

$$L_e L_b u = e \cdot \nabla (b \cdot \nabla (b \cdot \nabla (e \cdot \nabla u))) + e \cdot \nabla ([e, b] \cdot \nabla (b \cdot \nabla u)) + e \cdot \nabla (b \cdot \nabla ([e, b] \cdot \nabla u)).$$

On the other hand, we have

$$L_b L_e u = e \cdot \nabla (b \cdot \nabla (b \cdot \nabla (e \cdot \nabla u))) - b \cdot \nabla ([e, b] \cdot \nabla (e \cdot \nabla u)) - [e, b] \cdot \nabla (b \cdot \nabla (e \cdot \nabla u)).$$

From here we deduce the expression

$$\begin{aligned} [L_e, L_b] u &= e \cdot \nabla (b \cdot \nabla ([e, b] \cdot \nabla u)) + [e, b] \cdot \nabla (b \cdot \nabla (e \cdot \nabla u)) \\ &\quad + b \cdot \nabla ([e, b] \cdot \nabla (e \cdot \nabla u)) + e \cdot \nabla ([e, b] \cdot \nabla (b \cdot \nabla u)). \end{aligned}$$

To notice cancellations we rewrite the expressions to be symmetric in  $e$  and  $[e, b]$ . Commuting:

$$\begin{aligned} [L_e, L_b] u &= e \cdot \nabla (b \cdot \nabla ([e, b] \cdot \nabla u)) + [e, b] \cdot \nabla (b \cdot \nabla (e \cdot \nabla u)) \\ &\quad + b \cdot \nabla ([e, b] \cdot \nabla (e \cdot \nabla u)) + b \cdot \nabla (e \cdot \nabla ([e, b] \cdot \nabla u)) \\ &\quad + e \cdot \nabla ([e, b], b] \cdot \nabla u) + [e, b] \cdot \nabla ([e, b] \cdot \nabla u). \end{aligned}$$

Plugging this into (5.11, 5.12) and using that  $[\tilde{e}_i, \tilde{b}_k] = 0$  yields

$$\begin{aligned} \Delta Qu - Q\Delta u &= \frac{1}{2} \sum_{i,k} \left[ L_{\hat{e}_i}, L_{\tilde{b}_k} \right] u \\ (5.13) \quad &= \frac{1}{2} \sum_{i,k} \left( \hat{e}_i \cdot \nabla \left( \tilde{b}_k \cdot \nabla \left( [\hat{e}_i, \tilde{b}_k] \cdot \nabla u \right) \right) + [\hat{e}_i, \tilde{b}_k] \cdot \nabla \left( \tilde{b}_k \cdot \nabla (\hat{e}_i \cdot \nabla u) \right) \right) \\ (5.14) \quad &+ \frac{1}{2} \sum_{i,k} \left( \tilde{b}_k \cdot \nabla \left( [\hat{e}_i, \tilde{b}_k] \cdot \nabla (\hat{e}_i \cdot \nabla u) \right) + \tilde{b}_k \cdot \nabla \left( \hat{e}_i \cdot \nabla \left( [\hat{e}_i, \tilde{b}_k] \cdot \nabla u \right) \right) \right) \\ (5.15) \quad &+ \frac{1}{2} \sum_{i,k} \left( \hat{e}_i \cdot \nabla \left( [[\hat{e}_i, \tilde{b}_k], \tilde{b}_k] \cdot \nabla u \right) + [\hat{e}_i, \tilde{b}_k] \cdot \nabla \left( [\hat{e}_i, \tilde{b}_k] \cdot \nabla u \right) \right). \end{aligned}$$

Expressions (5.13) and (5.14) vanish after summing in  $i$ . The third expression (5.15) vanishes in a similar manner. For example, observe that for  $k = 1$  we have

$$[[\hat{e}_2, \tilde{b}_1], \tilde{b}_1] = 2[\hat{e}_2, \tilde{b}_1] = -4\hat{e}_2, \quad [[\hat{e}_3, \tilde{b}_1], \tilde{b}_1] = -2[\hat{e}_3, \tilde{b}_1] = -4\hat{e}_3,$$

which shows that with  $k = 1$  fixed

$$(5.15) = -2\hat{e}_2 \cdot \nabla (\hat{e}_2 \cdot \nabla u) + 2\hat{e}_2 \cdot (\hat{e}_2 \cdot \nabla u) - 2\hat{e}_3 \cdot \nabla (\hat{e}_3 \cdot \nabla u) + 2\hat{e}_3 \cdot (\hat{e}_3 \cdot \nabla u) = 0.$$

A similar cancellation happens for other values of  $k$ , finishing the proof of *Step 1*.

*Step 2.* Now we proceed with the proof of expression (1.3) for the dissipation of entropic optimal transport. We use the following dual formulation

$$\mathcal{T}_\varepsilon(f, g) = \varepsilon \mathcal{H}(f) + \sup \left\{ \int u_1 g - \int u_0 f : \partial_s u + \frac{1}{2} |\nabla u|^2 + \frac{\varepsilon}{2} \Delta u = 0 \right\},$$

where  $\mathcal{H}$  denotes the usual entropy. Moreover, the optimal  $u$  provides an interpolation between  $f$  and  $g$  through the equation

$$(5.16) \quad \partial \rho_s + \operatorname{div}(\rho \nabla u) = \frac{\varepsilon}{2} \Delta \rho, \quad \rho_0 = f, \quad \rho_1 = g.$$

The interpolant  $\rho$  can be written explicitly in terms of the Schrödinger potentials and the heat kernel, although we make no use of this relation here. See [8] for more details. Now, if  $\partial_t f = q(f)$ ,  $\partial_t g = q(g)$  are solutions of the Landau equation for  $\alpha \equiv 1$  then, thanks to

Proposition 4, we have

$$\begin{aligned} 2 \frac{d}{dt} [\mathcal{T}_\varepsilon(f, g)] &= \langle \mathcal{T}'_\varepsilon(F, G), (Q(F), Q(G)) \rangle \\ &= \varepsilon \frac{d}{dt} \mathcal{H}(F) + \iint u_1 Q(G) - \iint u_0 Q(F) \\ &= \varepsilon \frac{d}{dt} \mathcal{H}(F) + \iint Q(u_1) G - \iint Q(u_0) F. \end{aligned}$$

Consider the interpolation between  $F$  and  $G$  of the form (5.16) and denote it  $\rho(v, w, s)$ ,  $0 \leq s \leq 1$ . Then we have

$$\begin{aligned} 2 \frac{d}{dt} \mathcal{T}_\varepsilon(f, g) &= -\varepsilon D(F) + \iint Q(u_1) G - \iint Q(u_0) F \\ &= \varepsilon \frac{d}{dt} \mathcal{H}(F) + \int_0^1 \iint Q(\partial_s u) \rho + Q(u) \partial_s \rho \, ds \\ &= \varepsilon \frac{d}{dt} \mathcal{H}(F) + \int_0^1 \iint \left[ \nabla Q(u) \cdot \nabla u - \frac{1}{2} Q(|\nabla u|^2) \right] \rho \, ds + \frac{\varepsilon}{2} \int_0^1 \iint [\Delta Q u - Q \Delta u] \rho \, ds \\ &= \varepsilon \frac{d}{dt} \mathcal{H}(F) - \sum_k \int_0^1 \iint \left| \nabla (\tilde{b}_k \cdot \nabla u) \right|^2 \rho \, ds, \end{aligned}$$

yielding the expression for the dissipation (1.3) and finishing the proof of Theorem 2.  $\square$

## 6. PROOF OF THEOREM 3

We now prove the contractive estimates for soft potentials using the previously derived expression of  $\frac{d}{dt} \mathcal{T}(f, g)$  in Lemma 4.

*Proof of Theorem 3.* The starting point is the expression for the dissipation of  $\mathcal{T}$  found in Lemma 4. We first consider the second term and apply Young's inequality. We get

$$\begin{aligned} \frac{1}{2} \iint \frac{\alpha'}{r} \left( \tilde{b}_k \cdot \nabla (\tilde{b}_k \cdot \nabla u) \right) (\hat{n} \cdot \nabla u) \rho \, dv \, dw &\leq \frac{1}{4} \iint \left( \tilde{b}_k \cdot \nabla (\tilde{b}_k \cdot \nabla u) \right)^2 \frac{\alpha}{r^2} \rho \, dv \, dw \\ &\quad + \frac{1}{4} \iint \frac{(\alpha')^2}{\alpha} |\hat{n} \cdot \nabla u|^2 \rho \, dv \, dw. \end{aligned}$$

To bound the first term, we make use of the following identity (see [1]) for functions  $G : \mathbb{R}^6 \rightarrow \mathbb{R}$ :

$$|\nabla G|^2 = \frac{1}{2} \sum_i (\tilde{e}_i \cdot \nabla G)^2 + \frac{(\hat{n} \cdot \nabla G)^2}{2r^2} + \sum_k \frac{(\tilde{b}_k \cdot \nabla G)^2}{2r^2}.$$

In particular, this shows that

$$\frac{1}{4} \int_0^1 \iint \left( \tilde{b}_k \cdot \nabla (\tilde{b}_k \cdot \nabla u) \right)^2 \frac{\alpha}{r^2} \rho \, dv \, dw \, ds \leq \frac{1}{2} \int_0^1 \iint \left| \nabla (\tilde{b}_k \cdot \nabla u) \right|^2 \alpha \rho \, dv \, dw \, ds,$$

Plugging both estimates in (5.10) and denoting by

$$K := \sup_{r>0} \frac{\alpha'(r)^2}{\alpha(r)r^{\gamma-2}}$$

we get

$$\frac{d}{dt}\mathcal{T}(f,g) \leq \frac{3K}{4} \int_0^1 \iint |v-w|^{\gamma-2} |\hat{n} \cdot \nabla u|^2 \rho \, dv \, dw \, ds.$$

It is now necessary to introduce  $\rho(v,w) \equiv \rho(v)\rho(w)$ ,  $u(v,w) \equiv u(v) + u(w)$ , which yields the expression

$$\begin{aligned} \frac{d}{dt}\mathcal{T}(f,g) &\leq \frac{3K}{4} \int_0^1 \iint |v-w|^{\gamma-2} |(v-w) \cdot (\nabla u(v) - \nabla u(w))|^2 \rho(v)\rho(w) \, dv \, dw \, ds \\ &\leq \frac{3K}{4} \int_0^1 \iint |v-w|^\gamma |\nabla u(v) - \nabla u(w)|^2 \rho(v)\rho(w) \, dv \, dw \, ds. \end{aligned}$$

We now focus on the case  $-3 < \gamma < 0$ , and we simply bound

$$|\nabla u(v) - \nabla u(w)|^2 \leq 2(|\nabla u(v)|^2 + |\nabla u(w)|^2)$$

and reverse the role of  $v, w$  in one of the terms to obtain

$$\begin{aligned} \frac{d}{dt}\mathcal{T}(f,g) &\leq 3K \int_0^1 \iint |v-w|^\gamma |\nabla u(v)|^2 \rho(v)\rho(w) \, dv \, dw \, ds \\ &= 3K \int_0^1 \int |\nabla u(v)|^2 \rho(v) \underbrace{\left( \int \rho(w) |v-w|^\gamma \, dw \right)}_{:=(I_\gamma\rho)(v)} \, dv \\ &\leq 3K \|I_\gamma\rho\|_{L^\infty(\mathbb{R}^3 \times [0,1])} \mathcal{T}(f,g). \end{aligned}$$

Next, we bound the singular integral  $I_\gamma\rho$  by splitting the integral in a ball of radius  $R$  and its complement:

$$\begin{aligned} (I_\gamma\rho)(v) &= \int_{B_R(v)} |v-w|^\gamma \rho(w) \, dw + \int_{\mathbb{R}^d \setminus B_R(v)} |v-w|^\gamma \rho(w) \, dw \\ &\leq \left( \int_{B_R} |w|^{\gamma p'} \rho(w) \, dw \right)^{\frac{1}{p'}} \|\rho\|_{L^p(\mathbb{R}^3)} + R^\gamma \\ &= \left( \frac{4\pi(p-1)}{p(3+\gamma)-3} \right)^{1-\frac{1}{p}} R^{\gamma+\frac{3}{p'}} \|\rho\|_{L^p(\mathbb{R}^3)} + R^\gamma, \end{aligned}$$

where  $p > \frac{3}{3+\gamma}$ . By choosing  $R = \|\rho\|_{L^p}^{-\frac{p'}{3}}$  we obtain the following bound:

$$\|I_\gamma\rho\|_{L^\infty(\mathbb{R}^3 \times [0,1])} \leq C_{\gamma,p} \|\rho\|_{L^p(\mathbb{R}^3)}^{\beta_{p,\gamma}}, \quad C_{\gamma,p} = \left( \frac{4\pi(p-1)}{p(3+\gamma)-3} \right)^{1-\frac{1}{p}} + 1, \quad \beta_{p,\gamma} = -\frac{\gamma p}{3(p-1)}.$$

Recall that the functional  $\rho \mapsto \int \rho^p dv$  is convex with respect to the 2-Wasserstein distance, see [19]. Interpolation yields:

$$\begin{aligned} \|\rho_s\|_{L^p(\mathbb{R}^3)} &= \left( \int \rho_s(v)^p dv \right)^{\frac{1}{p}} \leq \left( (1-s) \int f(v)^p dv + s \int g(v)^p dv \right)^{\frac{1}{p}} \\ &\leq (1-s)^{\frac{1}{p}} \|f\|_{L^p(\mathbb{R}^3)} + s^{\frac{1}{p}} \|g\|_{L^p(\mathbb{R}^3)} \leq \|f\|_{L^p(\mathbb{R}^3)} + \|g\|_{L^p(\mathbb{R}^3)}. \end{aligned}$$

In summary, we have shown that for  $\alpha(r) = r^\gamma$ ,  $-3 < \gamma < 0$ , and  $p > \frac{3}{3+\gamma}$ , there exists  $C_{\gamma,p} > 0$  and  $0 < \alpha_{\gamma,p} < 1$  such that

$$\frac{d}{dt} \mathcal{T}(f_t, g_t) \leq C_{\gamma,p} K \left( \|f_t\|_{L^p(\mathbb{R}^3)}^{\beta_{\gamma,p}} + \|g_t\|_{L^p(\mathbb{R}^3)}^{\beta_{\gamma,p}} \right) \mathcal{T}(f_t, g_t).$$

This proves Theorem 3 for  $-3 < \gamma < 0$ .

Next, we study the Coulomb case  $\gamma = -3$ , which requires a special adaptation of the previous computation since  $|v-w|^{-3}$  is not locally integrable. We need to use the regularity of  $u$ , which can be expressed through the Hopf-Lax formula as

$$(6.1) \quad u_s(v) = -\frac{1}{s} \phi_s^*(v) + \frac{|v|^2}{2s}, \quad \phi_s(v) = s\phi(v) + (1-s) \frac{|v|^2}{2}.$$

Since  $\phi^*$  is convex, this shows  $D^2 u_s \leq \frac{1}{s} \text{Id}$ . On the other hand, it is easy to see that  $\tilde{u}(s, v) = -u(1-s, v)$  also satisfies a similar inequality. In fact, we have the bounds

$$-\frac{1}{1-s} \text{Id} \leq D^2 u_s \leq \frac{1}{s} \text{Id},$$

which implies that  $|\nabla u(v) - \nabla u(w)| \leq \frac{1}{\min(s, 1-s)} |v-w|$ . We consider a parameter  $0 < \varepsilon < \frac{1}{2}$  to be optimized and note that  $\frac{d}{dt} \mathcal{T}(f_t, g_t)$  is bounded by

$$\frac{3K}{4} \int_0^1 [\min(s, 1-s)]^{-2\varepsilon} \iint |v-w|^{-3+2\varepsilon} |\nabla u(v) - \nabla u(w)|^{2(1-\varepsilon)} \rho(v) \rho(w) dv dw ds.$$

Using the inequality  $(|\nabla u(v)| + |\nabla u(w)|)^{2(1-\varepsilon)} \leq 2^{1-2\varepsilon} (|\nabla u(v)|^{2(1-\varepsilon)} + |\nabla u(w)|^{2(1-\varepsilon)})$  and swapping the roles of  $v, w$  in the second resulting term, we see that

$$\begin{aligned} \frac{d}{dt} \mathcal{T}(f_t, g_t) &\leq 3K \int_0^1 [\min(s, 1-s)]^{-2\varepsilon} \iint |v-w|^{-3+2\varepsilon} |\nabla u(v)|^{2(1-\varepsilon)} \rho(v) \rho(w) dv dw ds \\ &= 3K \int_0^1 [\min(s, 1-s)]^{-2\varepsilon} \int |\nabla u(v)|^{2(1-\varepsilon)} \rho(v) (I_{-3+2\varepsilon} \rho)(v) dv ds, \end{aligned}$$

where  $(I_\gamma \rho)(v) = \int \rho(w) |v-w|^\gamma dw$  as before. This yields

$$\begin{aligned} \frac{d}{dt} \mathcal{T}(f_t, g_t) &\leq 3K \|I_{-3+2\varepsilon} \rho\|_{L^\infty(\mathbb{R}^3 \times [0,1])} \int_0^1 [\min(s, 1-s)]^{-2\varepsilon} \int |\nabla u(v)|^{2(1-\varepsilon)} \rho(v) dv ds \\ &\leq 3K \|I_{-3+2\varepsilon} \rho\|_{L^\infty(\mathbb{R}^3 \times [0,1])} \int_0^1 [\min(s, 1-s)]^{-2\varepsilon} \left( \int |\nabla u(v)|^2 \rho(v) dv \right)^{1-\varepsilon} ds. \end{aligned}$$

Next, we use that  $(\rho, u)$  are the constant-speed geodesic between  $f$  and  $g$  in the 2-Wasserstein distance. In particular, this means that  $\int |\nabla u_s(v)|^2 \rho_s(v) dv$  is independent of  $s$  and equal to

$\mathcal{T}(f, g)$ . Then, using that  $\int_0^1 [\min(s, 1-s)]^{-2\varepsilon} ds = \frac{4^{1-\varepsilon}}{1-2\varepsilon}$ , we are left with

$$(6.2) \quad \frac{d}{dt} \mathcal{T}(f, g) \leq \frac{12K}{1-2\varepsilon} \|I_{-3+2\varepsilon}\rho\|_{L^\infty} \mathcal{T}(f, g)^{1-\varepsilon}.$$

Finally, we need to optimize in  $\varepsilon$ . As a first step, we carefully bound the singular integral with respect to  $\varepsilon$ . Given  $R > 0$  we have

$$\begin{aligned} (I_{-3+2\varepsilon}\rho)(v) &= \int_{B_R(v)} \frac{\rho(w)}{|v-w|^{3-2\varepsilon}} dw + \int_{\mathbb{R}^3 \setminus B_R(v)} \frac{\rho(w)}{|v-w|^{3-2\varepsilon}} dw \\ &\leq \|\rho\|_{L^\infty(\mathbb{R}^3)} \int_{B_R} \frac{dw}{|w|^{3-2\varepsilon}} + R^{-3+2\varepsilon} = \frac{2\pi}{\varepsilon} \|\rho\|_{L^\infty(\mathbb{R}^3)} R^{2\varepsilon} + R^{-3+2\varepsilon}. \end{aligned}$$

Setting  $R = \|\rho\|_{L^\infty(\mathbb{R}^3)}^{-\frac{1}{3}} \varepsilon^{\frac{1}{3}}$  yields that

$$(6.3) \quad \|I_{-3+2\varepsilon}\rho\|_{L^\infty(\mathbb{R}^d)} \leq (2\pi+1) \|\rho\|_{L^\infty(\mathbb{R}^3)}^{1-\frac{2\varepsilon}{3}} \varepsilon^{-1+\frac{2\varepsilon}{3}} \leq (2\pi+1) \max(\|\rho\|_{L^\infty(\mathbb{R}^3)}, 1) \varepsilon^{-1}.$$

If we plug (6.3) in (6.2) and assume that  $0 < \varepsilon < \frac{1}{4}$  we obtain

$$\frac{d}{dt} \mathcal{T}(f, g) \leq 24(2\pi+1)K \max(\|f\|_{L^\infty(\mathbb{R}^3)} + \|g\|_{L^\infty(\mathbb{R}^3)}, 1) \varepsilon^{-1} [\mathcal{T}(f, g)]^{1-\varepsilon}.$$

We choose  $\varepsilon$  that minimizes  $\varepsilon^{-1} \mathcal{T}(f, g)^{1-\varepsilon}$ :

$$\varepsilon = \begin{cases} \frac{-1}{\log \mathcal{T}(f, g)}, & \text{if } \mathcal{T}(f, g) \leq e^{-4} \\ \frac{1}{4}, & \text{if } \mathcal{T}(f, g) \geq e^{-4} \end{cases}$$

In the case  $\mathcal{T}(f, g) \leq e^{-4}$ , the bound is given by

$$\frac{d}{dt} \mathcal{T}(f, g) \leq -24(2\pi+1)K \max(\|f\|_{L^\infty(\mathbb{R}^3)} + \|g\|_{L^\infty(\mathbb{R}^3)}, 1) \mathcal{T}(f, g) \log \mathcal{T}(f, g),$$

and the proof is finished by solving the ode. Moreover  $\mathcal{T}(f, g) = 0$  if  $f_{in} = g_{in}$  from Osgood uniqueness criteria.  $\square$

#### ACKNOWLEDGMENTS

MGD is partially supported by NSF Grant DMS-2205937. MPG is partially supported by NSF Grant DMS-2206677 and DMS-2511625. MT is supported by NSF grant DMS-2206187. The authors would like to thank Nestor Guillen, Jun Kitagawa, and Luis Silvestre for the enlightening discussions.

#### REFERENCES

- [1] Nestor Guillen and Luis Silvestre. The Landau equation does not blow up. *Acta Mathematica* Vol. 234, nr. 2, 315–375, 2025.
- [2] Nicolas Fournier and Hélène Guérin. Well-posedness of the spatially homogeneous Landau equation for soft potentials. *Journal of Functional Analysis*, 256(8):2542–2560, 2009.
- [3] Nicolas Fournier. Uniqueness of bounded solutions for the homogeneous Landau equation with a Coulomb potential. *Communications in Mathematical Physics*, 299(e):765–782, 2010.
- [4] William Golding, Maria Gualdani, and Amelie Loher. Global smooth solutions to the Landau-Coulomb equation in  $L^{3/2}$ . *Arch. Ration. Mech. Anal.* 249 (2025), no. 3, Paper No. 34.
- [5] Laurent Desvillettes, William Golding, , Maria Gualdani, and Amelie Loher. Production of the Fisher information for the Landau-Coulomb equation with  $L^1$  initial data. <https://arxiv.org/pdf/2410.10765.pdf>.

- [6] Sehyun Ji. Dissipation estimates of the Fisher information for the Landau equation. <https://arxiv.org/abs/2410.09035>, 2024.
- [7] Ling-Bing He, Jie Ji, and Yue Luo. Existence, uniqueness and smoothing estimates for spatially homogeneous Landau-Coulomb equation in  $H^{-\frac{1}{2}}$  space with polynomial tail. <https://arxiv.org/abs/2412.07287>, 2025.
- [8] Nicola Gigli and Luca Tamanini. Benamou–Brenier and duality formulas for the entropic cost on RCD(K, N) spaces. *Probability Theory and Related Fields*, 2020.
- [9] Nicolas Fournier and Arnaud Guillin. From a Kac-like particle system to the Landau equation for hard potentials and Maxwell molecules. *Ann. Sci. Éc. Norm. Supér. (4)* 50 (2017), no. 1, 157–199.
- [10] Kleber Carrapatoso. Propagation of chaos for the spatially homogeneous Landau equation for Maxwellian molecules. *Kinet. Relat. Models* 9 (2016), no. 1, 1–49.
- [11] Tanaka Hiroshi. Probabilistic treatment of the Boltzmann equation of Maxwellian molecules. *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* 46 (1978/79), no. 1, 67–105.
- [12] José A. Carrillo, Matias G. Delgadino, Laurent Desvillettes, and Jeremy S.-H. Wu. The Landau equation as a Gradient Flow. *Analysis & PDE*, 17(4):1331–1375, 2024.
- [13] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker–Planck equation. *SIAM journal on mathematical analysis*, 29(1):1–17, 1998.
- [14] Laurent Desvillettes and Cédric Villani. On the spatially homogeneous Landau equation for hard potentials part I: existence, uniqueness and smoothness. *Communications in Partial Differential Equations*, 2007.
- [15] Jean Dolbeault and Bruno Volzone. Improved Poincaré inequalities. *Nonlinear Analysis: Theory, Methods and Applications*, 75(16):5985–6001, 2012.
- [16] Cédric Villani. *Topics in optimal transportation*, volume 58. American Mathematical Society, 2003.
- [17] Filippo Santambrogio. *Optimal transport for applied mathematicians*, volume 55. Springer, 2015.
- [18] Francisco U. Caja-Lopez and Matias G. Delgadino. The second variation of the Wasserstein distance. *In preparation*, 2025.
- [19] Robert J McCann. A convexity principle for interacting gases. *Advances in mathematics*, 128(1):153–179, 1997.

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF TEXAS AT AUSTIN (USA)

*Email address:* funai.caja@utexas.edu

*Email address:* matias.delgadino@utexas.edu

*Email address:* gualdani@math.utexas.edu

DEPARTMENT OF MATHEMATICS, EMORY UNIVERSITY (USA)

*Email address:* maja.taskovic@emory.edu