A continuous model for ratings

Pierre-Emmanuel JABIN [*], Stéphane JUNCA [†‡]

December 5, 2014

Abstract

We study rating systems, such as the famous ELO system, applied to a large number of players. We assume that each player is characterized by an intrinsic inner strength and follow the evolution of their rating evaluations by deriving a new continuous model, a kinetic-like equation. We then investigate the validity of the rating systems by looking at their large time behavior as one would ideally expect the rating of each player to converge to their actual strength. The simplistic case when all players interact indeed yields an exponential convergence of the ratings. However, the behavior in the more realistic cases with only local interactions is more complex with several possible equilibria depending on the exact initial distribution of initial ratings and possibly very slow convergence.

Keywords: Kinetic, mean field, ELO, entropies, asymptotic behavior.
Mathematics Subject Classification: 35Q91 (35B40-91D30-82C70)

Contents

[*]CSCAMM and Dept. of Mathematics, University of Maryland, College Park, USA
[†]Université de Nice Sophia-Antipolis, J.A. Dieudonné laboratory, UMR CNRS 6621,Parc Valrose, 06108 Nice, France, junca@unice.fr
[‡]Team Coffee, INRIA Sophia Antipolis Mediterranée, 2004 route des Lucioles - BP 93, 06902 Sophia Antipolis Cedex, France

1 Introduction

1.1 General discussion

We first present the main ideas, goals and conclusions of our investigations in non mathematical language. Precise assumptions, models and results are given in later subsections.

The ELO rating system was first introduced by Arpad Elo, an American physicist and a chess player, for chess competitions [8]. It was adopted by the American chess federation and then the international chess federation FIDE. This system and several variants are now used to rank individual players or teams in many sports, or online games. The Elo system bases its ratings on encounters between players or teams, resulting on a win for one of the player or a tie. An encounter is typically the outcome of a game but nothing prevents the system from being applied to even more general settings; the result of an encounter could be decided by voting for instance.

After either each encounter, or a series of them, the system updates the ratings of each player depending on their wins or losses and on the ratings of the players against whom they competed (the mathematical formulas are given below).

A natural question is the validity of the Elo system: Is the ranking after many interactions between players fair and in which sense? In particular the delicate question of the initial ratings was noticed early on. When a new player enters the system, a rating has to be attributed and cannot be very accurate. As the new player competes, one would hope that its rating would become less dependent on the initial, somewhat arbitrary choice.

A. Elo tried to validate its rating system by numerical studies [8]. Most rigorous mathematical study of this widely used rating system are statistical works by M. Glickman [9, 10, 11, 12] who also tried to improve the Elo system.

Our goal is to investigate the validity of the system in the specific case with a large number of players, which is typical for online games for example. The starting point in our modeling is to assume that each player is characterized by a constant number, her or his intrinsic strength. The outcome of an encounter between two players is a random event where the probability of each winning is entirely determined by the difference of the intrinsic strengths of the players. This is actually equivalent to assuming that there exists an intrinsic transitive ranking of the players: If player A is stronger than (wins more often against) player B and player B stronger than player C then A is stronger than C. The question now becomes whether the rating system can recover and identify this intrinsic ranking as it of course does not have access to the information about the player's intrinsic strengths; it can only observe its consequences on the frequency of wins and losses.

This assumption, though simplifying and not necessarily very realistic, has the advantage of making transparent the study of a rating system: The system will perform well if the rating it calculates for each player is close to the intrinsic strength.

To complete our modeling, we need some rules to determine how players interact. The players do not have the information either about their intrinsic strengths. We hence assume that they base their decision to interact only on the public informations available to everyone, that is the ratings provided by the system. This is likely reasonable for anonymous players online for example. More precisely, we assume that encounters are random and based only on the relative ratings of the two players.

We study two very different situations:

- The "all meet all". This means that two players always have a minimal probability of encounter no matter how far apart their ratings are.
- The local interaction case where two players only have a chance to meet if their ratings are close enough.

Let us now briefly summarize the conclusions of our study which differ depending on which of the above case is assumed. We are able to show that in both cases, the ratings converge, always creating the impression that the rating system works.

In the "all meet all" case, the Elo system performs well: The final ratings indeed coincide the intrinsic strengths of the players and the convergence is exponential producing a good estimate of the players strengths in a minimal time.

However, in the local interactions case, the Elo system may fail and one can obtain ratings that are very different from the players' intrinsic strengths. No rates of convergence are available and in fact meta-stable equilibria can be observed so it is even delicate to determine if the ratings are even yet close to their limits. In addition the limiting ratings intrinsically depend on the initial distribution of the ratings, which is a notorious problem as mentioned before.

Fortunately, our studies also show that it is straightforward to recognize that the system is not performing well. This can only happen if a group of players becomes isolated. In other words, the system can only fail if a gap in the distribution of ratings is observed, which is easy to detect.

In the rest of this introduction, we first briefly introduce the historical Elo model for two players with precise mathematical formula. We then present our new continuous kinetic-like model, valid for a large number of players and interactions. The main results on that model are given in Section 2.

1.2 The Elo model

We explain here the basic idea behind the Elo system with just two players. The new rating of each player depends on the previous rating and the result of the match, increasing for a win and decreasing for a loss.

For two players labelled i and j, Arpad Elo ([8]) proposed formula (1.1) to evaluate their rating at time $n+1$ knowing their rating at time n and the result of the encounter:

$$R_i^{n+1} = R_i^n + K(S_{ij}^n - b(R_i^n - R_j^n)), \tag{1.1}$$

- n is the time,
- R_i^n is the rating of player i at time n,
- S_{ij}^n is the score or result of the encounter at time n between players i and j. The most standard convention is $S_{ij}^n = 0$ for a loss for player i or $S_{ij}^n = 1$ for a win for player i. However, to better emphasize the symmetry of the model, without affecting its mathematical properties, we prefer to consider the *centered score* $S_{ij}^n = \pm 1$ so that $S_{ji}^n = -S_{ij}^n$. In this article, since it leads to the same models, we typically consider the general case $S_{ij}^n \in [0, 1]$ or $S_{ij}^n \in [-1, 1]$ in the centered formulation.
- $b(R_i^n - R_j^n)$ is the predicted mean score, based on the ratings. It is the key ingredient in the system and it further explained below.
- The factor K is a positive constant weighting the change of the rating after an encounter. K is usually relatively small as it is rather natural not to upset too much the ratings based only on one encounter.

To understand better the key ingredient of model (1.1), let us take the expected value of (1.1). Assuming that the ratings at time n have already been determined (so it is in fact the corresponding conditional expectation)

$$\mathbb{E}_n(R_i^{n+1}) = R_i^n + K(\mathbb{E}_n(S_{ij}^n) - b(R_i^n - R_j^n)),$$

where $\mathbb{E}_n$ denotes the expectation knowing the ratings at time n.

In this very simple example, one sees that if the ratings are stationary, more precisely if $\mathbb{E}_n(R_i^{n+1}) = R_i^n$, then

$$\mathbb{E}_n(S_{ij}^n) = b(R_i^n - R_j^n).$$

In that case the ratings provide an exact evaluation of the respective strengths of the players, measured by probability that one will win over the other. The classical model uses an increasing function $b(.)$ which values in $]-1, 1[$. Furthermore, to keep a conservative model

$$R_i^{n+1} + R_j^{n+1} = R_i^n + R_j^n,$$

the function $b(.)$ is usually chosen odd.

1.3 A new continuous rating model

We now assume that each player has an intrinsic strength or real rating ϱ_i and a rating r_i given by the Elo system. For the score S_{ij}^n we assume that the random variables are independent which values in $[-1, 1]$. The independence is a fundamental property leading to a discrete Markov chain model before the derivation of the continuous model.

Furthermore we assume that the mean of the score S_{ij}^n depends only on the difference of the intrinsic strengths

$$\mathbb{E}(S_{ij}^n) = b(\varrho_i - \varrho_j), \tag{1.2}$$

which agrees with the argument in the previous subsection. In fact if in the simple case where S_{ij} only takes the values $\{+1, -1\}$, this means that the probability of player i winning is

$$\mathbb{P}(S_{ij}^n) = \frac{1}{2} + \frac{1}{2} b(\varrho_i - \varrho_j).$$

In practice, a player cannot meet all the other players. A common choice on the Internet is to meet players with close ratings. Consequently we introduce the weight function $w(r-r') \geq 0$ related to the probability that two players with respective rating r and r' will have an interaction.

When the number of players is large $N >> 1$, K is small and when the number of interactions (comparisons) is also large we derive a deterministic kinetic model, see Section 3 for the details. Denoting by $f(t, r, \varrho)$ the density of players with rating r and intrinsic strength ϱ at time t, we find the equation

$$\frac{\partial}{\partial t} f + \frac{\partial}{\partial r} \left(a[f]\ f\right) = 0, \tag{1.3}$$

$$f(t=0, r, \varrho) = f^0(r, \varrho), \tag{1.4}$$

the scalar velocity field $a[f]$ is given by a convolution:

$$a[f] = a[f](t, r, \varrho) = \int_{\mathbb{R}^2} w(r-r') \big(b(\varrho - \varrho') - b(r-r')\big)\, f(t, r', \varrho')\, d\varrho'\, dr'. \tag{1.5}$$

Note that

- r is the variable for the actual, evaluated rating of a player.
- ϱ represents the natural strength of the player. Ideally one would hope that the evaluated rating is close to the strength.
- $b(.)$ is a "bonus-malus" function indicating how the rating of a player changes after an encounter with another player. It is the mean centered score predicted by the Elo model. $b(.)$ is an **odd** increasing function.
- $w(.)$ is an **even** nonnegative function related to the probability that two players with ratings r and r' can interact.

Why Eq. (1.3) is a good model for a rating system and how it can be formally derived is worthy of a discussion in itself, which is done in Section 3. We just mention that the derivation is of mean field type and we summarize here the main assumptions under which Eq. (1.3) is derived

- There is a large number of players having each a large number of encounters.
- The effect of each encounter on the rating is small.
- The probability of winning an encounter depends only on the difference between the intrinsic strengths of the players.
- The effect of each encounter is only to modify the ratings of the players (their intrinsic strength is unaffected).
- The probability of an encounter between two players depends only on their relative ratings.

There are many mathematical studies of physical and social sciences models with mean field interaction and kinetic model, of which we only mention a sample. The general mean field approach to dynamics of interacting particles can be found in [17] (see also [1, 14] for the cases with singular forces).

Even though it is obtained from a very different setting, there are some similarities between our model and kinetic equations for granular media, see for instance

[2, 4, 5]. However, the derivation and structure of the model makes it closer to so-called flocking models, see among many [13]. In particular the study of the asymptotic behaviors of flocking models often faces very similar issues in the cases where the interaction has a finite range. There are some results of convergence in those cases, see for instance [16] with an assumption that all individuals remain connected; and [15] for a more general convergence result but without rates. However, those results for flocking only deal with discrete settings (finite number of individuals). In our case we are able to obtain the convergence for the continuous setting (infinite number of players) though is also considerably more delicate. Notice that the specific structure of the model (in particular the constant intrinsic strengths) is actually different enough from flocking models.

Eq. (1.3) is a kinetic model, with a degenerate dissipation rate when w is compactly supported. We mention [6, 7] and references therein for other examples of study of long time behavior of "degenerate" kinetic equations. Those studies are notoriously difficult, with various approaches used. In [6] for instance, a hypo-coercivity assumption is used. Unfortunately it involves the uniqueness of the equilibrium which does not hold in our case. Our approaches still heavily uses the entropies of the equation, but all of them are needed together with a precise and careful study of the dynamic.

In section 2 we state the main results on the continuous model (1.3). Section 3 provides a formal derivation of our model. We obtain the first a priori estimates for in section 4. Then section 5 deals with the simple but not realistic case when all meet all. The discrete case with local interactions is studied in section 6. Finally, the more general and difficult case is treated in section 7.

2 Main results

Our main results concern model (1.3). It is straightforward to obtain well posedness for Eq. (1.3) under reasonable assumptions, for instance

Theorem 2.1 *Assume that $w,\ b\in C(\mathbb{R})\cap L^\infty(\mathbb{R})$, and $f^0\geq 0$ belongs to $M^1(\mathbb{R}^2)$ then there exists a unique $f\geq 0$ in $L^\infty(\mathbb{R}_+,\ M^1(\mathbb{R}^2))$, solution to (1.3)-(1.4) in the sense of distributions. The corresponding solution satisfies in addition*

$$\int f(t,r,\rho)\,dr\,d\rho=\int f^0(r,\rho)\,dr\,d\rho=m_0,$$
$$\int (r-\rho)\,f(t,r,\rho)\,dr\,d\rho=\int (r-\rho)\,f^0(r,\rho)\,dr\,d\rho=m_1,$$
$$f(t,-,\rho)=\int f(t,r,\rho)\,dr=f^0(-,\rho)=\int f^0(r,\rho)\,dr,$$

and if $f^0\in L^p(\mathbb{R}^2)$ then $f\in L^\infty([0,\ T],\ L^p(\mathbb{R}^2))$ for any $T>0$.

Note that if f is a measure, we make a slight abuse of notation, writing $f\,dr\,d\rho$ instead of the proper $f(t,dr,d\rho)$. We mostly leave the proof of this result to the reader and only briefly indicate in Section 4 how to obtain the additional conservations. We now focus on the behavior for large times of the system. For this one has to be more precise on the structure of the equation.

We assume that the "bonus" function is Lipschitz continuous and strictly increasing on any compact set

$$|b(x) - b(y)| \leq B|x-y|, \qquad \forall \text{ compact } K,\ \exists \beta_K > 0 \text{ s.t. } \inf_K \frac{db}{dr} \geq \beta_K. \tag{2.1}$$

Note that the bonus function introduced by Elo is, after centering

$$b(\Delta r) = 2\left(\frac{1}{1+10^{-\Delta r/400}}\right) - 1,$$

which satisfies (2.1). Of course in practice, many actual systems truncate b.

Ideally, one would like the rating of each individual to converge toward its intrinsic strength. However, there is an obvious invariance by translation in the system: We still obtain a solution if all ratings r are shifted by the same amount. Taking this into account, a natural candidate for the limit of f as $t \longrightarrow \infty$ is

$$f^0(-,\rho)\,\delta(r-\rho-m_1/m_0) = \delta(r-\rho-m_1/m_0)\int f^0(s,\rho)\,ds,$$

where we here denote

$$f^0(-,\rho) = \int f^0(s,\rho)\,ds, \quad m_0 = \int_{\mathbb{R}^2} f^0(r,\rho)\,dr\,d\rho, \quad m_1 = \int_{\mathbb{R}^2} (r-\rho)\,f^0(r,\rho)\,dr\,d\rho.$$

In the case when the probability of encounters never vanish, it is simple to prove this and we even have a rate

Theorem 2.2 *Assume that b satisfies* (2.1)*, that $f^0 \geq 0$ in $M^1(\mathbb{R}^2)$ with compact support and that $w \in C(\mathbb{R}) \cap L^\infty(\mathbb{R})$ with $w > 0$. Then*

$$f(t,.,.) \longrightarrow f^0(-,\rho)\,\delta(r-\rho-m_1/m_0) \text{ in weak}-*\ M^1(\mathbb{R}^2), \quad \text{as } t \to \infty.$$

Moreover, there exists $\lambda > 0$ and $C > 0$ s.t.

$$\left\| f(t,.,.) - f^0(-,\rho)\,\delta(r-\rho-m_1/m_0)\right\|_{W^{-1,1}(\mathbb{R}^2)} \leq C\,e^{-\lambda t}.$$

In that case, the proof is straightforward and only uses one entropy of the system. Unfortunately in practice, it is not realistic to assume that $w > 0$. Indeed two players with very different rankings will not agree to an encounter.

Therefore, it is more realistic to assume that

$$w > 0 \text{ on }]-1,1[, \qquad w(x) = 0 \quad \forall |x| \geq 1. \tag{2.2}$$

Of course the interval $(-1,1)$ was only chosen here for convenience and the results would remain the same for any other open interval.

We start by the case where f^0 is discrete, *i.e.* a sum of Dirac masses

Theorem 2.3 *Assume that b satisfies* (2.1)*, that $w \in C(\mathbb{R})$ satisfies* (2.2) *and that*

$$f^0 = \sum_{i=1}^N \alpha_i\,\delta(r-r_i)\,\delta(\rho-\rho_i), \quad \alpha_i > 0 \quad \forall i.$$

Then there exists $c_1, \ldots, c_N$ s.t.

$$f(t,.,.) \longrightarrow \sum_{i=1}^N \alpha_i\delta(r-\rho_i-c_i)\,\delta(\rho-\rho_i) \text{ in weak}-*\ M^1(\mathbb{R}^2), \quad \text{as } t \to \infty,$$

where moreover either $c_i = c_j$ or $|\rho_i + c_i - \rho_j - c_j| \geq 1$.

In this case, it is already not possible in general to simply identify the limit in terms of m_0, m_1 or f^0. Indeed consider the situation of several groups of players which do not interact as their initial ratings are too far. The ratings in each group would converge to the intrinsic strengths plus a shift. But the shifts for each group could be very different. This is the situation described by the Theorem: If $|\rho_i+c_i-\rho_j-c_j|\geq 1$ then the two players belong to two distinct groups. If $|\rho_i+c_i-\rho_j-c_j|<1$ then they belong to the same group and necessarily $c_i=c_j$.

The proof of Th. 2.3 is still straightforward but requires the use of all entropies (contrary to Th. 2.2). We now turn to the general case which is the main contribution of this paper

Theorem 2.4 *Assume that b satisfies (2.1), that $w\in C^2$ satisfies (2.2) and that $f^0\geq 0$ is compactly supported then there exists distinct constants $c_1,\ldots,c_n$ and $M^1(\mathbb{R})$ measures $h_1,\ldots,h_n$ s.t.*

$$f(t,.,.)\longrightarrow\sum_{i=1}^N h_i(\rho)\delta(r-\rho-c_i)\ \ \textit{in weak}-*\ M^1(\mathbb{R}^2),\quad \textit{as } t\to\infty.$$

Moreover for $i\neq j$, if $|\rho+c_i-\rho'-c_j|<1$ then either $h_i(\rho)=0$ or $h_j(\rho')=0$.

In the last line, when h_i and h_j are measures, the statement should be interpreted to mean that their supports are disjoint.

Again here the final equilibrium cannot be simply identified and we can at most say that if two players interact then they should have the same "shift" c_i. The proof of Th. 2.4 is now much more delicate; in particular the use of the entropies is not enough (because of a strong degeneracy in the dissipation) and it is necessary to use in a precise manner the structure of the equation.

Notations. As is usual, we denote $f'=f(t,r',\rho')$. C denotes a constant which depends only on f^0 and which value may change from line to line.

3 Derivation of the model

In this section we explain how the continuous model was obtained from some simple considerations. The arguments are written in a formal manner but could easily be made rigorous provided the functions b and w are smooth enough.

We start by a discrete model based on simple and direct interactions between a finite (but large) number N of players.

The basic dynamic is described by a Markov chain on the ratings R_i^n of each player i at step n. That chain is parameterized by the theoretical ratings or intrinsic strengths ρ_i of the players.

Between step n and $n+1$, players i and j engage in a match with probability $W_{ij}^n=W_N(R_i^n-R_j^n)$ where W_N can scale with the number of players.

For simplicity and only in this section, the result of the match is a random variable $S_{ij}^n\in\{-1,\ 1\}$, 1 meaning a victory for player i and -1 for player j (hence $S_{ji}^n=-S_{ij}^n$). This case contains all the important ideas and the derivation could easily be extended from it to a more general situation where S_{ij}^n take more values, a draw $S_{ij}^n=0$ for instance or even a continuum $S_{ij}^n\in[-1,\ 1]$.

We assume that the S_{ij}^n are independent random variables: The result of one match does not depend on the result of any previous matches. Moreover we assume that the law of S_{ij}^n depends only on the difference between the theoretical ratings (or intrinsic strengths) of the players ρ_i, ρ_j so that

$$\mathbb{E}(S_{ij}^n) \quad = \quad b(\varrho_i - \varrho_j). \tag{3.1}$$

Denote that M_i^n the random set of j s.t. i and j engage in a match at step n. The ratings are updated at step $n+1$ according to the results of the encounters,

$$R_i^{n+1} \quad = \quad R_i^n + K \sum_{j \in M_i^n} (S_{ij}^n - b(R_i^n - R_j^n)), \tag{3.2}$$

as per the formula initially developed by Arpad Elo. The parameter K is a fixed constant which determines how much one encounter impacts the ratings and is hence usually small.

Note that, of course, (3.2) is only one possible model, corresponding to a rating system where ratings are updated at some fixed intervals (after tournaments, every month, every day...)

There are other systems where ratings are updated after each encounter for example. The corresponding model would be a jump process R_i^t, over a continuous variable t. However, given the scalings chosen here, this would lead to the same limit equation.

The invariance by translation is a key property of the Elo model (3.2). That is to say, for any constant c, we can replace for all i and all n, R_i^n by $R_i^n - c$. The important quantity to determine the issue of an encounter is not the rating itself but the relative rating: $R_i^n - R_j^n$.

To obtain a deterministic version of this model, an estimate of the variance is needed. More precisely, the variance has to stay small. It turns out that model (3.2) admits a simple and crude L^∞ estimate which leads to control the first and the second moment *i.e.* the mean: $r_i^n = \mathbb{E}(R_i^n)$ and the variance:

$$|r_i^n| = |\mathbb{E}(R_i^n)| \leq \sup |R_i^n|, \qquad \mathbb{E}\left((R_i^n - r_i^n)^2\right) \leq (2 \sup |R_i^n|)^2.$$

Since $|S_{ij}^n| \leq 1$ and $|b| \leq 1$ we simply have $|R_i^{n+1}| \leq |R_i^n| + 2KN$ and then

$$|R_i^n| \quad \leq \quad 2nKN. \tag{3.3}$$

Note that since the R_i^n are a Markov chain by taking conditional expectations

$$r_i^{n+1} = r_i^n + \sum_j \mathbb{E}\left[K\, W_N(R_i^n - R_j^n)(b(\rho_i - \rho_j) - b(R_i^n - R_j^n))\right]. \tag{3.4}$$

Assume that $W_N(r) = \bar{W}_N\, w(r)$. Denote $v_i^n = \mathbb{E}\left(|V_i^n|\right)$ where $V_i^n = R_i^n - r_i^n$ and, $v^n = \max_{j \leq N, m \leq n} v_j^m$.

Replacing the random variable R_i^n by its expectation r_i^n we have:

$$\begin{aligned} r_i^{n+1} \quad &= \quad r_i^n + K\bar{W}_N \sum_j \left[w(r_i^n - r_j^n)(b(\rho_i - \rho_j) - b(r_i^n - r_j^n))\right] + \mathcal{O}(z^n), \quad (3.5) \\ z^n \quad &= \quad C\, N\, K\, \bar{W}_N\, v^n. \end{aligned}$$

where the constant denoted by C only depends on the Lipschitz norms of b and w. Throughout the derivation, the constant C may change but C is always independent of n, N, i, j.

Let us rewrite (3.2) with independent Bernoulli variables $\mathbb{I}_{ij}^n$. $\mathbb{I}_{ij}^n = 1$ means that the player number i meets the player number j at step n. We assume that the following equality on the conditional expectation holds

$$\mathbb{E}(\mathbb{I}_{ij}^n \,|\, R_i^n - R_j^n) = W_N(R_i^n - R_j^n)$$

as explained before to get (3.4). (3.2) and a modification of (3.5) up to the order z^n become

$$\begin{aligned} R_i^{n+1} &= R_i^n + K \sum_j \mathbb{I}_{ij}^n (S_{ij}^n - b(R_i^n - R_j^n)), \\ r_i^{n+1} &= r_i^n + K \sum_j \mathbb{E}\left[\mathbb{I}_{ij}^n (b(\rho_i - \rho_j) - b(r_i^n - r_j^n))\right] + \mathcal{O}(z^n) \end{aligned}$$

Therefore comparing R_i^n with r_i^n

$$\begin{aligned} V_i^{n+1} &= V_i^n + K \sum_j \left[\mathbb{I}_{ij}^n \, S_{ij}^n - \mathbb{E}(\mathbb{I}_{ij}^n \, b(\rho_i - \rho_j))\right] \\ &\quad -K \sum_j \left[\mathbb{I}_{ij}^n \, b(R_i^n - R_j^n) - \mathbb{E}(\mathbb{I}_{ij}^n \, b(r_i^n - r_j^n))\right] + \mathcal{O}(z^n) \\ \mathbb{E}V_i^{n+1} &= \mathbb{E}V_i^n + K \sum_j \mathbb{E}\left[\mathbb{I}_{ij}^n (S_{ij}^n - b(\rho_i - \rho_j))\right] \\ &\quad -K \sum_j \mathbb{E}\left[\mathbb{I}_{ij}^n (b(R_i^n - R_j^n) - b(r_i^n - r_j^n))\right] + \mathcal{O}(z^n) \\ &= \mathbb{E}V_i^n + K \sum_j \mathbb{E}\left[\mathbb{I}_{ij}^n (S_{ij}^n - b(\rho_i - \rho_j))\right] + \mathcal{O}(z^n) \\ v_i^n &\leq v_i^0 + \mathbb{E}\left|\sum_{m=0}^{n} K \sum_j \mathbb{I}_{ij}^m \, \left[S_{ij}^m - b(\rho_i - \rho_j)\right]\right| + \mathcal{O}(n\, z^n). \end{aligned}$$

Note that the random variables S_{ij}^m are independent from the $\mathbb{I}_{ij}^m$. Therefore once a realization has been determined for $\mathbb{I}_{ij}^m$, the sum

$$\sum_{m=0}^{n} \sum_j \mathbb{I}_{ij}^m \, \left[S_{ij}^m - b(\rho_i - \rho_j)\right]$$

is a sum of independent random variables with 0 mean and variance of order 1. By the law of large numbers its conditional expectation given the $\mathbb{I}^n_{ij}$ is

$$\begin{aligned}
&\mathbb{E}\left(\left|\sum_{m=0}^n \sum_j \mathbb{I}^m_{ij}\,\left[S^m_{ij}-b(\rho_i-\rho_j)\right]\right|\;\Big|\;(\mathbb{I}^m_{ij})_{j,m}\right)\\
\le\;&\mathbb{E}\left(\left|\sum_{m=0}^n \sum_j \mathbb{I}^m_{ij}\,\left[S^m_{ij}-b(\rho_i-\rho_j)\right]\right|^2\;\Big|\;(\mathbb{I}^m_{ij})_{j,m}\right)^{1/2}\\
=\;&\mathbb{E}\left(\sum_{m=0}^n\sum_{l=0}^n\sum_k\sum_j \mathbb{I}^m_{ij}\,\mathbb{I}^l_{ik}\,\left[S^n_{ij}-b(\rho_i-\rho_j)\right]\,\left[S^m_{ik}-b(\rho_i-\rho_k)\right]\;\Big|\;(\mathbb{I}^m_{ij})_{j,m}\right)^{1/2}\\
\le\;& C\,\left(\sum_{m=0}^n\sum_j \mathbb{I}^m_{ij}\right)^{1/2}.
\end{aligned}\tag{3.6}$$

Consequently

$$\mathbb{E}\left|\sum_{m=0}^n K\sum_j \mathbb{I}^m_{ij}\,\left[S^m_{ij}-b(\rho_i-\rho_j)\right]\right|\le C\,K\,\sqrt{N\,n\,\bar W_N}.$$

Finally we deduce that

$$v^n_i\le v^0_i+C\,K\,\sqrt{n\,N\,\bar W_N}+C\,n\,N\,K\,\bar W_N\,v^n,$$

and then

$$v^n\le v^0+C\,K\,\sqrt{n\,N\,\bar W_N}+C\,n\,N\,K\,\bar W_N\,v^n. \tag{3.7}$$

We are now ready to precise our scalings. We assume that the ratings have been updated often enough on the time interval considered, *i.e.* we look at some time scale where many updates have been performed, for a large number of players and for small changes K in ratings. More precisely

$$n=\frac{t}{dt},\quad N\,K\,\bar W_N=dt,\quad t\sim 1,\quad dt<<1,\quad K<<1,\quad N>>1. \tag{3.8}$$

The estimate (3.7) shows that

$$v^n\le v^0+C\,\sqrt{K}+C\,t\,v^n.$$

since $nK^2N\bar W_N=nK\,dt\sim K$. Given that the initial ratings are deterministic $v^0_i=0$ and choosing $t\le 1/2C$, we get that with a constant C twice the previous constant C

$$v^n\le C\,\sqrt{K},$$

for any $n\le 1/(2C\,dt)$. By reusing the argument starting now from $n=1/(2C\,dt)$,

$$v^n\le C\sqrt{K}\,\exp(C\,n\,dt).$$

Therefore for any time t of order 1, the variance is small. We can consider the mean r_i^n instead the random variable R_i^n. Let us replace this estimate in the equation (3.5) for r_i^n

$$r_i^{n+1} = r_i^n + K\bar{W}_N \sum_j \left[w(r_i^n - r_j^n)(b(\rho_i - \rho_j) - b(r_i^n - r_j^n))\right] + \mathcal{O}\left(dt\,\sqrt{K}\,e^{C\,n\,dt}\right).$$

Summing up from $n = 0$, this implies

$$\begin{aligned} r_i^n =& r_i^0 + dt \sum_{m=0}^{n-1} \frac{N\,K\bar{W}_N}{dt} \frac{1}{N} \sum_j \left[w(r_i^m - r_j^m)(b(\rho_i - \rho_j) - b(r_i^m - r_j^m))\right] \\ &+ \mathcal{O}\left(\sqrt{K}\,e^{C\,n\,dt}\right). \end{aligned}$$

We may define $r_i(t)$ at any time t by taking them piecewise constant and equal to r_i^n at $t = ndt$. Introduce the solution to the ODE in integral form

$$\bar{r}_i(t) = r_i^0 + \int_0^t \frac{1}{N} \sum_j \left[w(\bar{r}_i - \bar{r}_j)(b(\rho_i - \rho_j) - b(\bar{r}_i - \bar{r}_j))\right].$$

Since $N\,K\,\bar{W}_N = dt$, Gronwall's lemma directly implies that

$$\sup_i |r_i(t) - \bar{r}_i(t)| \le C\,(dt + \sqrt{K})\,e^{C\,t}.$$

Now define the empirical measures

$$\begin{aligned} \mu_N(t, r, \rho) &= \frac{1}{N} \sum_i \delta(r - r_i(t))\,\delta(\rho - \rho_i), \\ \bar{\mu}_N(t, r, \rho) &= \frac{1}{N} \sum_i \delta(r - \bar{r}_i(t))\,\delta(\rho - \rho_i). \end{aligned}$$

The distance between the two in weak norm ($W^{-1,1}$ for instance) is $O(dt+\sqrt{K})\,e^{Ct}$. On the other hand, by the definition of the $\bar{r}_i$, it is straightforward to check that $\bar{\mu}_N$ is a measure-valued solution to (1.3) where $a[\mu]$ is defined by (1.5), which we repeat here for convenience

$$a[\mu] = a[\mu](t, r, \varrho) = \int_{\mathbb{R}^2} w(r - r')\left(b(\varrho - \varrho') - b(r - r')\right)\mu(t, dr', d\varrho').$$

Using well known stability results for measure-valued solution when the interaction coefficients are smooth (see [17]), $\bar{\mu}_N$ and hence the empirical measure μ_N converge for large N and small K to a measure μ which solves (1.3).

4 A priori estimates for the continuous model

We detail in this section all the a priori estimates that are used later. The calculations are formal but could easily be made rigorous in the context of Th. 2.1 for instance.

4.1 Invariants

We start by listing the natural invariants of the system

- *Conservation of mass.* Simply integrate Eq. (1.3) to find

$$\frac{d}{dt}\int_{\mathbb{R}^2} f(t,r,\rho)\,dr\,d\rho = 0.$$

Hence

$$m_0 = \int_{\mathbb{R}^2} f(t,r,\rho)\,dr\,d\rho = \int_{\mathbb{R}^2} f^0(r,\rho)\,dr\,d\rho. \tag{4.1}$$

- *Conservation of relative momentum.* This corresponds to the invariance by translation: If $g^0(r,\rho) = f^0(r+r_0,\rho+\rho_0)$, the solution g to (1.3) is $g(t,r,\rho) = f(t,r+r_0,\rho+\rho_0)$. Now multiply Eq. (1.3) by $\lambda\, r + \mu\,\rho$ and integrate

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{R}^2}(\lambda\, r+\mu\,\rho) f\,dr\,d\rho &= \int_{\mathbb{R}^2}\lambda\, a[f]\, f\,dr\,d\rho\\ &= \lambda\int_{\mathbb{R}^4} w(r-r')\,(b(\rho-\rho')-b(r-r'))\,f\,f'\,dr\,d\rho\,dr'\,d\rho' = 0\end{aligned}$$

with $f' = f(r',\rho')$ and by the skew-symmetry in r, ρ and r', ρ' of $b(\rho-\rho') - b(r-r')$. Therefore

$$\int_{\mathbb{R}^2} r\,f\,dr\,d\rho = \int_{\mathbb{R}^2} r\,f^0\,dr\,d\rho, \quad \int_{\mathbb{R}^2}\rho\,f\,dr\,d\rho = \int_{\mathbb{R}^2}\rho\,f^0\,dr\,d\rho. \tag{4.2}$$

- *Conservation of intrinsic strength.* Eq. (1.3) only includes a derivative in r. Denote the marginal distribution with respect to ϱ by

$$f(t,-,\varrho) := \int f(t,r,\varrho)dr,$$

and integrate (1.3) with respect to r to find $\partial_t f(t,-,\rho) = 0$ and hence

$$f(t,-,\rho) = f^0(-,\rho), \quad \forall\rho. \tag{4.3}$$

In the rest of the paper, we simplify the proofs by normalizing those invariants and taking

$$\begin{aligned} m_0 = \int f(t,r,\rho)\,dr\,d\rho = 1,\\ \int r\,f(t,r,\rho)\,dr\,d\rho = \int \rho\,f(t,r,\rho)\,dr\,d\rho = 0.\end{aligned} \tag{4.4}$$

4.2 Energy and entropies

We turn to the dissipated quantities. Through this subsection, we assume that b is non decreasing.

We have a family of entropies: For any φ define

$$E_\varphi = \int \varphi(r-\varrho) f(t,r,\varrho)drd\varrho, \tag{4.5}$$

then E_φ is dissipated for any convex φ

Proposition 4.1 (Entropy decay) *Let ϕ be any $C^1(\mathbb{R})$ convex function then*

$$\frac{d\,E_\varphi}{dt} = -D_\varphi(f(t)) = \int \psi(r,r',\rho,\rho')\,f'\,f\,dr'\,d\varrho'\,dr\,d\varrho \le 0, \tag{4.6}$$

with

$$\psi = \frac{1}{2}(\varphi'(r-\rho)-\varphi'(r'-\rho'))w(r-r')(b(\varrho-\varrho')-b(r-r')) \le 0. \tag{4.7}$$

Proof : The proof is a straightforward calculation. By simply multiplying (1.3) by $\varphi(r-\rho)$ and integrating

$$\frac{d\,E_\varphi}{dt} = \int \varphi'(r-\rho)\,w(r-r')(b(\varrho-\varrho')-b(r-r'))\,f'\,f\,dr'\,d\varrho'\,dr\,d\varrho.$$

Now the skew-symmetry of $b(\varrho-\varrho')-b(r-r')$ gives the formula. To conclude on the sign, observe that since φ' is non decreasing, $\varphi'(r-\rho)-\varphi'(r'-\rho') \ge 0$ iff $r-\rho \ge r'-\rho'$ *i.e.* iff $r-r' \ge \rho-\rho'$ that is if $b(\varrho-\varrho')-b(r-r') \le 0$ as b is non decreasing. □

An important example of entropy if the relative energy

$$E(t) \;=\; \int (r-\varrho)^2 f(t,r,\varrho)drd\varrho. \tag{4.8}$$

As a particular case of Prop. 4.1, one obtains

Proposition 4.2 (Energy decay) *One has that*

$$\frac{d\,E}{dt} = -\int (r-r'-\rho+\rho')\,w(r-r')(b(\varrho-\varrho')-b(r-r'))\,f'\,f\,dr'\,d\varrho'\,dr\,d\varrho \le 0. \tag{4.9}$$

The dissipation of energy is the key to obtain exponential decay of the energy and then exponential convergence to equilibrium.

As a first consequence, we note that Prop. 4.1 implies an uniform control in time of the compact support of f.

Corollary 4.1 (Compact support) *If $f^0(r,\varrho)$ is compactly supported in $S_r \times S_\varrho$ then for all t, $(r,\varrho) \to f(t,r,\rho)$ is compactly supported in $\mathcal{R} \times S_\varrho$ where $\mathcal{R} = \{r'+\varrho'-\varrho'',\, r' \in S_r,\, \varrho',\varrho'' \in S_\varrho\}$.*

For instance if $S_r \subset [-a,a]$ and $S_\varrho \subset [-\alpha,\alpha]$ then $\mathcal{R} \subset [-a-2\alpha, a+2\alpha]$.
Proof : First note that the support of f with respect to ϱ is independent of time from (4.3). If $(r,\varrho) \in S_r \times S_\varrho$ then $r-\varrho$ belongs in $\mathcal{A} = S_r - S_\varrho = \{r-\varrho,\, r \in S_r,\, \varrho \in S_\varrho\}$. Let φ be a convex non negative function which only vanishes on $\mathcal{A}$. Then $E_\varphi(0)=0$ and for all time E_φ vanishes. For each fixed (t,ϱ), $r \to f(t,r,\varrho)$ is compactly supported in $r-\rho \in \mathcal{A}$, i.e. $r \in S_\varrho + \mathcal{A} = \mathcal{R}$. □

5 All meet all: Proof of Th. 2.2

We use here the assumptions of Th. 2.2: b satisfies (2.1) and $w \in C(\mathbb{R}) \cap L^(\infty)$ with $w > 0$ everywhere. By the normalizing (4.4), the expected equilibrium is now simply $f^0(-,\rho)\,\delta(r-\rho)$.

The proofs here use only one entropy inequality: the energy decay.

We start by just proving the convergence of the energy with an exponential rate but without assuming that w is uniformly bounded from below over $\mathbb{R}$.

Proposition 5.3 *If f^0 is compactly supported and w is positive everywhere then the energy associated to the solution of* (1.3) *decays exponentially*

$$E(t) \leq E(0)\exp(-2\underline{w}\,\beta\,t),$$

where β depends only on K the initial support of f^0 and b through assumption (2.1)*, and $\underline{w}$ depends only on w and K.*

Proof : By (4.9)

$$\int_0^\infty D(f)(t)\,dt < \infty,$$

with

$$D(f) = \int (r - r' - \rho + \rho')\,w(r - r')\,(b(\rho - \rho') - b(r - r'))\,f'\,f\,dr'\,dr\,d\rho'\,d\rho.$$

By Corollary 4.1, $K^1 \supset K$, the support of f, is uniformly bounded for all time. By positivity and continuity of w, $w(r - r') \geq \underline{w} > 0$ for all (r, ρ), $(r', \rho') \in K^1$. By assumption (2.1), on the uniform support of f

$$(r - r' - \rho + \rho')\,w(r - r')\,(b(\rho - \rho') - b(r - r')) \geq \beta\,\underline{w}\,|r - r' - \rho + \rho'|^2.$$

Hence

$$D(f) \geq \beta\,\underline{w} \int |r - r' - \rho + \rho'|^2\,f'\,f\,dr'\,dr\,d\rho'\,d\rho.$$

Now note that by (4.4)

$$\int (r - \rho)\,(r' - \rho')\,f'\,f\,dr'\,dr\,d\rho'\,d\rho = 0.$$

Therefore

$$D(f) \geq 2\,\beta\,\underline{w} \int |r - \rho|^2\,f(t, r, \rho)\,dr\,d\rho, \tag{5.1}$$

and

$$\frac{d}{dt}E(t) \leq -2\,\underline{w}\,\beta \int |r - \rho|^2\,f = -2\,\omega\,\beta\,E(t),$$

concluding by Gronwall's lemma. □

With this proposition, we can identify the limit and prove the convergence.

The solution $f(t, ., .)$ is uniformly bounded in M^1. Take any extracted subsequence $f(t_n, ., .)$ converging weak-* to $\widetilde{f}$. By Corollary 4.1, the support is uniformly bounded and one may pass to the limit to obtain

$$\int |r - \rho|^2\,\widetilde{f}(r, \rho)\,dr\,d\rho = 0.$$

Just note that $\int \widetilde{f}\, dr\, d\rho = 1$. This implies that $\widetilde{f}$ is supported on the diagonal $r = \rho$ and that there exists $h(\rho)$ s.t.

$$\widetilde{f} = h(\rho)\, \delta(r - \rho).$$

By (4.3), $\widetilde{f}(-, \rho) = f^0(-, \rho)$ so $h = f^0(-, \rho)$. Therefore $\widetilde{f}$ is completely determined and unique. The whole sequence $f(t, ., .)$ then necessarily converges to this limit, proving the first part of Th. 2.2.

Indeed, the result is stronger. We can prove directly that the convergence has an exponential rate. Now we recall that

$$\|g\|_{W^{-1,1}} = \sup_{\|\phi\|_{W^{1,\infty}} \leq 1} \int \phi\, g.$$

Take any ϕ s.t. $\|\phi\|_{W^{1,\infty}} \leq 1$ and compute using (4.3)

$$\begin{aligned} \int \phi\, (f - f^0(-, \rho)\, \delta(r - \rho))\, dr\, d\rho &= \int (\phi(r, \rho) - \phi(\rho, \rho))\, f(t, r, \rho)\, dr\, d\rho \\ &\leq \int |r - \rho|\, f(t, r, \rho)\, dr\, d\rho. \end{aligned}$$

Consequently by Cauchy-Schwartz

$$\|f - f^0(-, \rho)\, \delta(r - \rho)\|_{W^{-1,1}} \leq \sqrt{E(t)},$$

which concludes the proof of Th. 2.2.

6 Discrete case: Proof of Th. 2.3

In this section we study the asymptotic behavior of ratings when the number of players is finite.

We now assume that b satisfies (2.1), and that $w \in C(\mathbb{R})$ satisfies (2.2). The solution f to (1.3) with a sum of Dirac masses as initial data is of the form

$$f = \sum_{i=1}^{N} \alpha_i\, \delta(t - R_i(t))\, \delta(\rho - \rho_i),$$

where the R_i solve

$$\dot{R}_i = \sum_j \alpha_j\, w(R_i - R_j)\, (b(\rho_i - \rho_j) - b(R_i - R_j)), \tag{6.1}$$

with $R_i(0) = r_i$. Note that the weights α_i and the ρ_i are constant in time. Let us denote $e_i(t) = R_i(t) - \varrho_i$, $i = 1, \cdots, N$.

We start by using the entropy dissipation for the whole family to obtain

Proposition 6.4 *For all $\phi \in C^0(\mathbb{R}, \mathbb{R})$, the function $V[\phi](t) = \sum_{i=1}^{N} \alpha_i \phi(e_i(t))$ has a limit as $t \to \infty$.*

Note that we do not require ϕ to be convex.
Proof : Let first $\phi \in C^1$ be a convex function and note that of course

$$V[\phi] = \int \phi(r-\rho)\, f(t, dr, d\rho).$$

Hence from Prop. 4.1 $V[\phi]$ is non increasing, bounded from below by Corollary 4.1 (uniform support of f): $V[\phi]$ has a limit.

Now if ϕ is concave then $V[\phi] = -V[-\phi]$ also has a limit as $-\phi$ is convex. Any C^2 function is the sum of a convex and a concave function.

Finally by density $V[\phi]$ converges for any continuous function ϕ. More precisely the ratings are uniformly bounded, see Corollary 4.1, so all $(e_i(t))$ stays in a compact set $\mathcal{K}$ for all time. On $\mathcal{K}$, any continuous function can be approximate by a sequence (ϕ_n) of C^2 functions. In particular

$$|V[\phi] - V[\phi_n]| \leq \|\phi - \phi_n\|_{L^\infty(\mathcal{K})},$$

and this bound is hence uniform in time.

As each $V[\phi_n]$ converges, one may define $l_n = \lim_{t\to+\infty} V\,[\phi_n](t)$. By the previous bound, $(l_n)_n$ is a Cauchy sequence and hence converges towards some limit l. We can conclude thanks to the uniform convergence of the sequence (ϕ_n) that $l = \lim_{t\to+\infty} V[\phi](t)$. □

From this it is straightforward to deduce the convergence of the $e_i(t)$. For instance define the empirical measure $g_N = \sum_i \alpha_i \delta(e - e_i(t))$ and note that for any ϕ

$$V[\phi] = \int \phi(e)\, g_N(t, de)$$

converges. This implies the weak-* convergence of g_N and then of each $e_i(t)$ toward a constant denoted c_i.

By the definition of the e_i, we obtain $R_i \to \rho_i + c_i$ and

$$f \longrightarrow \sum_i \alpha_i \delta(r - \rho_i - c_i)\, \delta(\rho - \rho_i).$$

Finally we use once more the energy and the fact that

$$\int_0^\infty D(f)\, dt < \infty.$$

By the previous convergence of f, this shows that

$$\begin{aligned} D\left(\sum_i \alpha_i \delta(r - \rho_i - c_i)\, \delta(\rho - \rho_i)\right) &= 0 \\ = \sum_{i,j} \alpha_i\, \alpha_j w(\rho_i - \rho_j + c_i - c_j)\, (c_i - c_j)\, &(b(\rho_i - \rho_j) - b(\rho_i - \rho_j + c_i - c_j)). \end{aligned}$$

As b is strictly increasing and $w > 0$ on $(-1,\ 1)$, either $c_i = c_j$ or $|\rho_i - \rho_j + c_i - c_j| \geq 1$.

7 Proof of Th. 2.4: General case with local interactions

We finally turn to the most complex setting. Through this section, the assumptions in Th. 2.4 are always supposed to be satisfied.

7.1 The structure of the possible limits

As f is uniformly bounded in $M^1(\mathbb{R}^2)$ and with uniform compact support, it is always possible to extract a weak-* converging subsequence. We denote $\Omega \subset \mathcal{M}^1_+(\mathbb{R}_r \times \mathbb{R}_\varrho)$ be the set of all possible limits of such extracted subsequence: all the weak-* adherence values of f.

In this subsection, we collect all information on Ω provided by the invariants and the entropies. We summarize the results with

Proposition 7.5 *There exist n, $c_1, \cdots, c_n$, $m_1, \cdots, m_n$ depending only on f^0 s.t. for any $\widetilde{f}(r,\varrho) \in \Omega$*

$$\widetilde{f}(r,\varrho) = \sum_{i=1}^{n} \delta(r - (\varrho + c_i))\, h_i(\varrho),$$

for some non negative measures h_i which may depend on $\widetilde{f}$ but satisfy

- *The mass of each nonnegative measure h_i is $\int h_i(\varrho)d\varrho = m_i$,*
- *For any r, there is at most one i such that $\widetilde{f}(r,\varrho) = \delta(r - (\varrho + c_i))\, h_i(\varrho)$ on the vertical strip $S(r) =]r-1, r+1[\times \mathbb{R}_\varrho$. Equivalently for $i \neq j$ then $h_i(\rho)\, h_j(\rho') = 0$ if $|\rho + c_i - \rho' - c_j| < 1$.*
- $\displaystyle\sum_{i=1}^{n} h_i(\varrho) = f^0(-,\varrho)$. *And if $f^0(-,\varrho)$ belongs to L^p, the h_i are also in L^p.*

As one can see, we cannot at this point completely determine every $\widetilde{f} \in \Omega$ and show that it is reduced to one element (there are too many possibilities for the h_i). This will require a more detailed analysis of Eq. (1.3) performed in the next subsection.

We start the proof of Prop. 7.5 by the equivalent of Prop. 6.4 in this extended case

Corollary 7.2 *Define*

$$g(t,r) = \int_{\mathbb{R}} f(t, r+\varrho, \varrho) d\varrho. \tag{7.1}$$

There exists a measure $g_\infty(r) \geq 0$ in $M^1(\mathbb{R}_r)$, depending only on f^0 such that

$$g(t,.) \rightharpoonup g_\infty(r) \quad \text{in } M^1 \text{ weak} * \text{ as } t \to +\infty. \tag{7.2}$$

Proof : As in the proof Prop. 6.4, Prop. 4.1 implies that

$$\int_{\mathbb{R}^2} \phi(r-\rho)\, f(t,r,\rho)\, dr\, d\rho$$

has a limit for any convex ϕ. By linearity this is also true for any concave ϕ and hence any continuous ϕ. By a change of variable this is exactly the convergence of g. □

The next step is to use the entropy dissipation and pass to the limit into it.

Lemma 7.1 (Zero entropy dissipation) *For any* $\widetilde{f} \in \Omega$

$$\int_{\mathbb{R}} \widetilde{f}(r, r+\varrho) d\varrho = g_\infty(r), \qquad \forall \varphi \in C^2 \text{ convex}, \quad D_\varphi(\widetilde{f}) = 0.$$

Proof : We recall that $D_\varphi(f(t)) = -\frac{dE_\varphi}{dt}$. Hence if ϕ is convex, $D_\varphi \geq 0$ and

$$\int_0^\infty D_\phi(f(t))\, dt < \infty.$$

We differentiate $D_\varphi(f(t))$ in (4.6)

$$\frac{d\, D_\varphi(f(t))}{dt} = I(t) = \int \psi(r, \varrho, r', \varrho')\, \partial_t(f' f) dr' d\varrho' dr d\varrho,$$

for some $\psi \in C^1$. Now $\partial_t(f'f) = f'\partial_t(f) + f\partial_t(f') = -f'\partial_r(a[f]f) - f\partial_{r'}(a[f']f')$ thanks to (1.3). Integrating by part and using the uniform control on the support of f provided by Corollary 4.1, we deduce that D_ϕ is Lipschitz in time on $\mathbb{R}_+$. This implies that $D_\phi(f(t)) \to 0$.

Now, if $\widetilde{f} \in \Omega$ there exists a sequence $t_n \to +\infty$ such that $f(t_n) \rightharpoonup \widetilde{f}$. Using again the uniform control on the support, we can pass to the limit in $D_\phi(f(t_n))$ to get $D_\varphi(\widetilde{f}) = 0$.

Simply by passing to the limit in $g(t, r)$ we similarly obtain

$$\int_{\mathbb{R}} \widetilde{f}(r, r+\varrho) d\varrho = g_\infty(r).$$

$\square$

Next we show that $D_\phi(\widetilde{f}) = 0$ implies a very precise structure on $\widetilde{f}$

Lemma 7.2 (Stationary density with zero dissipation) *Let* $\widetilde{f} = \widetilde{f}(r, \varrho)$ *be any nonnegative measure with compact support. If the entropy dissipation* $D_\varphi(\widetilde{f}) = 0$ *vanishes for a strictly convex test function* φ *then there exist a finite number* $c_1, \cdots, c_n \in \mathbb{R}$, $h_i \in \mathcal{M}^1_+(\mathbb{R})$ *with compact support s.t.*

$$\widetilde{f}(r, \varrho) = \sum_{i=1}^n \delta(r - (\varrho + c_i))\, h_i(\varrho). \tag{7.3}$$

Moreover, on any vertical strip $S(r) = (r-1, r+1) \times \mathbb{R}_\varrho$*, there is at most one non vanishing term in the sum* (7.3), i.e.

either $\quad S(r) \cap \mathrm{Supp}(\widetilde{f}) = \emptyset$

or $\quad S(r) \cap \mathrm{Supp}(\widetilde{f}) \subset \{(r', r' - c_i),\, |r - r'| < 1\}$ *for only one index* i.

Proof : Let φ be a convex function. From definitions (4.6) and (4.7), on $|r - r'| < 1$, $\psi > 0$ if $r - r' \neq \varrho - \varrho'$.

Cover the support of $\widetilde{f}$ by a finite number of stripes S_{r_i}. If $r,\ r' \in S_{r_i}$ then $w(r - r') > 0$. This proves that on $S^2_{r_i}$ the product $\widetilde{f}\,\widetilde{f}'$ is supported on $r - \rho = r' - \rho'$. Therefore there exists c_i and h_i s.t. on S_{r_i}, $\widetilde{f} = \delta(r - \rho - c_i)\, h_i(\rho)$. $\square$

To finish the proof of Prop. 7.5, it remains to explain why the c_i and the mass of each h_i is uniquely determined by f^0. For that observe that if $\widetilde{f}$ satisfies (7.3) then

$$\int \widetilde{f}(r + \rho, \rho)\, d\rho = \sum_i \delta(r - c_i) \int h_i(\rho)\, d\rho = g_\infty(r).$$

7.2 Concentration of mass along the diagonals

From the previous subsection, to prove the convergence of f, it suffices to show that the h_i are uniquely determined. The key in doing so is to explain that once some mass is concentrated near one of the diagonals $\{r = \rho + c_i\}$ then nothing moves.

This is done in two steps. First we show that the mass is concentrated along some tubular neighborhood of all the diagonals. In itself this is not enough as there could still be some exchange of mass; from one $\{r = \rho + c_i\}$ to another. In a second time, we rule this out by estimating the sign of the transport velocity $a[f]$ on the boundary of the tubular neighborhood.

Lemma 7.3 (Stability of the expected support)
For all $\varepsilon > 0$, there exists $T > 0$ such that for all $t > T$,

$$\int_{\mathcal{T}_\varepsilon^c} f(t,r,\varrho)drd\varrho < \varepsilon \quad \textit{where} \quad \mathcal{T}_\varepsilon = \bigcup_{i=1}^{n} \{(r,\varrho),\ |r - (\varrho + c_i)| \le \varepsilon\},$$

$\mathcal{T}_\varepsilon^c = \mathbb{R}^2 - \mathcal{T}_\varepsilon$, and $(c_1, \cdots, c_n)$ are given by the support of g_∞.

Proof : Let

$$d(f(t,.,.);\Omega) = \inf_{\widetilde{f}\in\Omega} \sup_{\{|\varphi|\le 1, |\nabla\varphi|\le 1\}} \left| \int (f(t,r,\varrho) - \widetilde{f}(r,\varrho))\varphi(r,\varrho)drd\varrho \right|$$

be the distance from $f(t,.,.)$ to Ω in $W^{-1,1}$. We assert that $\lim_{t\to+\infty} d(f(t,.,.);\Omega) = 0$.

This is a classical compactness argument. By contradiction if this assertion is false then there exists $\varepsilon > 0$ and an increasing sequence (t_n) converging towards $+\infty$ such that $d(f(t_n),\Omega) > \varepsilon$ for all n. Since $f(t,.,.)$ is uniformly bounded and the unit ball in $\mathcal{M}^1$ is compact for the topology induced by the $W^{-1,1}$ norm, up to a subsequence, $f(t_n) \to \widetilde{f}$. Now by taking this $\widetilde{f}$ in the definition of the distance, one obtains a contradiction.

Now we choose a test function to exploit the fact that $d(f(t),\Omega) \to 0$. Let $\varphi(r,\varrho)$ be a nonnegative piecewise affine function with $|\nabla\varphi| = 1$ near $\mathcal{T}_0 = \mathcal{T}_{\varepsilon=0}$, and $\varphi = 0$ on $\mathcal{T}_0$. For such a φ there exists $C > 0$ such that, $\varphi \ge C\varepsilon$ on$\mathcal{T}_\varepsilon^c$ for ε small enough. Now this Lipschitz function φ is fixed.

Note that for any $\widetilde{f} \in \Omega$, $\widetilde{f}$ also vanishes on $\mathcal{T}_0^c$ and hence

$$\begin{aligned}
\int_{\mathcal{T}_\varepsilon^c} f(t,r,\varrho)drd\varrho &\le (C\varepsilon)^{-1} \int_{\mathcal{T}_\varepsilon^c} f(t,r,\varrho)\varphi(r,\varrho)drd\varrho \\
&\le (C\varepsilon)^{-1} \int_{\mathbb{R}^2} f(t,r,\varrho)\varphi(r,\varrho)drd\varrho \\
&= (C\varepsilon)^{-1} \int_{\mathbb{R}^2} (f(t,r,\varrho) - \widetilde{f}(r,\varrho))\varphi(r,\varrho)drd\varrho \\
&\le (C\varepsilon)^{-1} d(f(t,.,.);\Omega).
\end{aligned}$$

We conclude the proof by taking $T > 0$ such that $d(f(t,.,.);\Omega) \le C\varepsilon^2$ for all $t > T$.
□

We now use the specific dynamic along the characteristics: $a[f] = (a_r, a_\varrho) = (a_r, 0)$, i.e. the two components velocity is only horizontal in the plane (r,ϱ).

Proposition 7.6 (Growth of the mass on the tubular sets)
Let $0 < \delta < 1$. For any $\varepsilon > 0$, any $\widetilde{f} \in \Omega$, there exists t_0 s.t. for any r_0, ρ_0 in the support of $\widetilde{f}$

$$\int_{-\sqrt{\varepsilon}}^{\sqrt{\varepsilon}} \int_{-\delta/2}^{+\delta/2} f(t, r_0 + s + y, \rho_0 + s)\, ds\, dy$$

is non decreasing for any $t \geq t_0$.

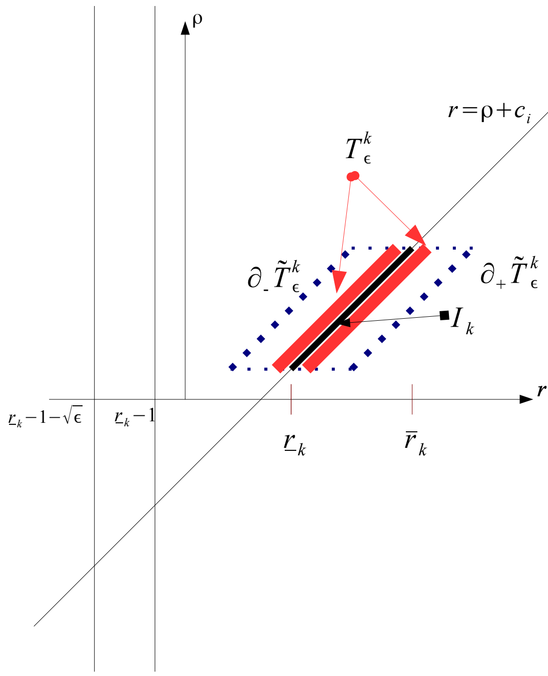

Figure 1: *line* $r = \varrho + c_i$, $\mathcal{T}_\varepsilon^k$, $\widetilde{\mathcal{T}}_\varepsilon^k$, a_+, a_-

Proof : Take the corresponding $\widetilde{f} \in \Omega$. The support of $\widetilde{f}$ is included in $\mathcal{T}_0$. Therefore we introduce a finite number of segment with diameter $\delta < 1$

$$I_k \quad = \quad \left\{ (r_k + s, \varrho_k + s),\ |s| < \frac{\delta}{2} \right\},$$

such that the support of $\widetilde{f}$ belongs to $\displaystyle\bigcup_{k=1}^{K} I_k$. Note that there exists an unique index i such that $r_k = \varrho_k + c_i$ thanks to Proposition 7.5. Therefore in a neighborhood of any I_k, $\widetilde{f}$ is equal to $h_i(\rho)\,\delta(r - (\rho + c_i))$.

Denote $\underline{r}_k = \min\{r,\ (r, \varrho) \in I_k\}$ and $\overline{r}_k = \max\{r,\ (r, \varrho) \in I_k\}$. We may assume that for any $\eta > 0$

$$\int_0^\eta h_i(\underline{r}_k + s)\, ds > 0 \text{ and } \int_0^\eta h_i(\overline{r}_k - s)\, ds > 0. \tag{7.4}$$

Otherwise h_i vanishes in a neighborhood of $\underline{r}_k$ or $\overline{r}_k$ and we can take a smaller interval.

Take $0 < \varepsilon$ small enough. From Lemma 7.3, there exist $t_0 > 0$ such that for all $t \geq t_0$,

$$\int_{\mathcal{T}_\varepsilon^c} f(t, r, \varrho) dr d\varrho < \varepsilon \quad \text{and} \quad d(f(t_0, ., .); \widetilde{f}) \leq \varepsilon^2,$$

with in addition

$$\min_k \int_{I_k} \widetilde{f}(r, \varrho) dr d\varrho > 4\, \varepsilon^{1/4}.$$

We build two "horizontal" neighborhoods of each I_k, Figure 1,

$$\mathcal{T}_\varepsilon^k = \{(r+s, \varrho),\ (r, \varrho) \in I_k,\ |s| < \varepsilon\} \quad \subset \quad \widetilde{\mathcal{T}}_\varepsilon^k = \mathcal{T}_{\sqrt{\varepsilon}}^k.$$

Let us point out here that $\bigcup_k \mathcal{T}_\varepsilon^k \subset \mathcal{T}_\varepsilon$ but that in general one cannot expect equality to hold: there might be several cylinders in $\mathcal{T}_\varepsilon$ that do not correspond to any I_k as they do not intersect the support of this particular $\widetilde{f} \in \Omega$.

Let us first remark that at $t = t_0$

$$\int_{\mathbb{R}^2 \setminus \bigcup_k \widetilde{\mathcal{T}}_\varepsilon^k} f(t, r, \varrho)\, dr\, d\rho < \varepsilon. \tag{7.5}$$

Indeed build again a Lipschitz test function ϕ which vanishes on $\bigcup_k I_k$, satisfies $|\nabla \phi| \leq 1$ and s.t. $\phi \geq \sqrt{\varepsilon}$ out of $\widetilde{\mathcal{T}}_\varepsilon^k$. This is easy since there are only a finite number of I_k and hence it is possible to choose ε small enough s.t. $\widetilde{\mathcal{T}}_\varepsilon^k \cap \widetilde{\mathcal{T}}_\varepsilon^l = \emptyset$ for $k \neq l$. Now

$$\varepsilon^{1/2} \int_{\mathbb{R}^2 \setminus \bigcup_k \widetilde{\mathcal{T}}_\varepsilon^k} f(t, r, \varrho)\, dr\, d\rho \leq \int \phi\, f\, dr\, d\rho \leq d(f, \widetilde{f}) \leq \varepsilon^2.$$

Recall that

$$\min_k \int_{I_k} \widetilde{f}(r, \varrho) dr d\varrho > 4\, \varepsilon^{1/4}.$$

$f(t, ., .)$ has only small masses outside the $\widetilde{f}$ support from (7.5), so, losing at most one $\varepsilon^{1/4}$ for positive ε small enough we also have at time $t = t_0$

$$\begin{aligned} &\min_k \int_{\widetilde{\mathcal{T}}_\varepsilon^k} f(t, r, \varrho) dr d\varrho > 3\, \varepsilon^{1/4}, \\ &\min_k \int_{\mathcal{T}_\varepsilon^k} f(t, r, \varrho) dr d\varrho > 2\, \varepsilon^{1/4}. \end{aligned} \tag{7.6}$$

Indeed the total mass of f outside $\mathcal{T}_\varepsilon$ is less than ε (Lemma 7.3); $\widetilde{f} = 0$ outside $\bigcup_k \mathcal{T}_\varepsilon^k$ and the total mass of $\widetilde{f}$ in $\mathcal{T}_\varepsilon^k$ is larger than $4\, \varepsilon^{1/4}$; and finally $d(f(t_0, ., .);\ \widetilde{f}) \leq \varepsilon^2$ which concludes by taking a test function with support in $\mathcal{T}_\varepsilon^k$ and ε small enough.

Note that at this point in (7.6), we could use $3\, \varepsilon^{1/4}$ for both estimates. However, this will not be possible for $t > t_0$ and we write the inequalities such that they will be valid for all $t > t_0$.

The strategy is to show that as long as (7.5) and (7.6) hold then every $\widetilde{\mathcal{T}}_\varepsilon^k$ is a trapping region and hence (7.5)-(7.6) keep on being satisfied. The reason why we need the two tubular neighborhoods $\mathcal{T}_\varepsilon^k$ and $\widetilde{\mathcal{T}}_\varepsilon^k$ is that we cannot control the sign of $a[f]$ on the boundary of $\mathcal{T}_k$ directly.

As the mass of f in any fixed domain changes continuously in time, there exists a maximal time t_1 s.t. on the time interval $[t_0, t_1)$, (7.5) holds and

$$\min_k \int_{\widetilde{\mathcal{T}}_\varepsilon^k} f(t,r,\varrho)\,dr\,d\varrho > 3\,\varepsilon^{1/4}, \quad \forall t \in [t_0,\ t_1).$$

By Lemma 7.3, this implies again that

$$\min_k \int_{\mathcal{T}_\varepsilon^k} f(t,r,\varrho)\,dr\,d\varrho > 2\,\varepsilon^{1/4}, \quad \forall t \in [t_0,\ t_1),$$

and (7.6) is hence automatically satisfied on $[t_0,\ t_1)$.

Define the left and right boundary of $\widetilde{\mathcal{T}}_\varepsilon^k$:

$$\partial_\pm \widetilde{\mathcal{T}}_\varepsilon^k \quad = \quad \{(r \pm \sqrt{\varepsilon}, \varrho),\ (r,\varrho) \in I_k\}.$$

We prove that on $[t_0, t_1)$ and for every k

$$\pm\, a[f(t)] < 0 \quad \text{on} \quad \partial_\pm \widetilde{\mathcal{T}}_\varepsilon^k. \tag{7.7}$$

As f follows the characteristics given by $a[f]$ this implies that on $[t_0,\ t_1)$, for any $r,\ \rho \in I_k$

$$\int_{-\sqrt{\varepsilon}}^{\sqrt{\varepsilon}} f(t, r+s, \rho)\,ds$$

is non decreasing in time for any $(r,\ \rho) \in I_k$. Hence if $t_1 < \infty$ for every k

$$\int_{\widetilde{\mathcal{T}}_\varepsilon^k} f(t_1,r,\rho)\,dr\,d\rho \geq \int_{\widetilde{\mathcal{T}}_\varepsilon^k} f(t_0,r,\rho)\,dr\,d\rho > 3\,\varepsilon^{1/4}.$$

Moreover, this would imply that

$$\begin{aligned}\int_{\mathbb{R}^2\setminus\bigcup_k \widetilde{\mathcal{T}}_\varepsilon^k} f(t_1,r,\rho)\,dr\,d\rho &= \int f\,dr\,d\rho - \sum_k \int_{\widetilde{\mathcal{T}}_\varepsilon^k} f(t_1,r,\rho)\,dr\,d\rho \\ &\leq \int_{\mathbb{R}^2\setminus\bigcup_k \widetilde{\mathcal{T}}_\varepsilon^k} f(t_0,r,\rho)\,dr\,d\rho < \varepsilon,\end{aligned}$$

and t_1 could not be the maximal time. This would show that $t_1 = \infty$ and conclude the proof of the proposition.

Therefore it only remains to prove (7.7). Choose any k; we perform the computations on the left boundary $\partial_- \widetilde{\mathcal{T}}_\varepsilon^k$ as the estimate is similar on the other, right boundary $\partial_+ \widetilde{\mathcal{T}}_\varepsilon^k$.

Let (r,ϱ) belong to $\partial_- \widetilde{\mathcal{T}}_\varepsilon^k$. Decompose $a[f](t,r,\varrho) = A_+ + A_-$ with

$$\begin{aligned} A_+ = &\int_{|r-r'|<1,\ r'-\varrho' \geq c_i - \sqrt{\varepsilon}} w(r-r')(b(\varrho-\varrho') - b(r-r'))\ f(t,r',\varrho')dr'd\varrho', \\ A_- = &\int_{|r-r'|<1,\ r'-\varrho' < c_i - \sqrt{\varepsilon}} w(r-r')(b(\varrho-\varrho') - b(r-r'))\ f(t,r',\varrho')dr'd\varrho'. \end{aligned}$$

The term A_+ is positive as $b(\varrho-\varrho')-b(r-r')>0$ iff $r'-\varrho'>r-\varrho$ and here $r-\rho=c_i-\sqrt{\varepsilon}$. Moreover, the domain of integration in A_+ contains $\mathcal{T}_\varepsilon^k$ and thus

$$A_+ \geq a_+ = \int_{\mathcal{T}_\varepsilon^k} w(r-r')(b(\varrho-\varrho')-b(r-r'))\, f(t,r',\varrho')dr'd\varrho'.$$

The left boundary $\partial_-\widetilde{\mathcal{T}}_\varepsilon^k$ is at distance at least $\sqrt{\varepsilon}/2$ from $\mathcal{T}_\varepsilon^k$. The function $(b(\varrho-\varrho')-b(r-r'))$ vanishes exactly for $(r',\varrho')\in\partial_-\widetilde{\mathcal{T}}_\varepsilon^k$. As b' is bounded from below, there exists a positive constant c such that $b(\varrho-\varrho')-b(r-r')>c\sqrt{\varepsilon}$ for all $(r',\varrho')\in\mathcal{T}_\varepsilon^k$.

Note that $w(r-r')$ is bounded from below $(r',\varrho')\in\mathcal{T}_\varepsilon^k$ as the diameter of I_k is $\delta<1$. Hence with possibly a different constant c

$$w(r-r')(b(\varrho-\varrho')-b(r-r'))>c\sqrt{\varepsilon} \quad \text{on} \quad \mathcal{T}_\varepsilon^k.$$

Since $\int_{\mathcal{T}_\varepsilon^k} f(t,r,\varrho)drd\varrho > 2\,\varepsilon^{1/4}$, $\forall t\in[t_0,t_1]$, we have

$$a_+ \quad > \quad c\,\varepsilon^{3/4}.$$

Now, we turn to A_-. Since $A_-\leq 0$, we show that $|A_-|$ is smaller than a_+. Decompose $A_-=A_-^1+A_-^2$ with

$$A_-^1 = \int_{(r',\rho')\notin\bigcup_l\tilde{\mathcal{T}}_\varepsilon^l,\; r'-\varrho'<c_i-\sqrt{\varepsilon}} w(r-r')(b(\varrho-\varrho')-b(r-r'))\, f(t,r',\varrho')dr'd\varrho'$$
$$A_-^2 = \int_{(r',\rho')\in\bigcup_l\tilde{\mathcal{T}}_\varepsilon^l,\; r'-\varrho'<c_i-\sqrt{\varepsilon}} w(r-r')(b(\varrho-\varrho')-b(r-r'))\, f(t,r',\varrho')dr'd\varrho'.$$

Now for A_-^1 we simply use (7.5) which shows that

$$|A_-^1|\leq C\int_{(r',\rho')\notin\bigcup_l\tilde{\mathcal{T}}_\varepsilon^l} f(t,r',\varrho')\,dr'\,d\varrho' < C\,\varepsilon.$$

Now, we turn to A_-^2. Take any $r',\ \rho'\in\bigcup_l\widetilde{\mathcal{T}}_\varepsilon^l$ s.t. $w(r-r')>0$ and $r'-\rho'<c_i-\sqrt{\varepsilon}$. Necessarily there exists l s.t. $r',\ \rho'\in\widetilde{\mathcal{T}}_\varepsilon^l$ and hence some j s.t. $|r'-\rho'-c_j|\leq\sqrt{\varepsilon}$. Therefore $i\neq j$. We recall that Proposition 7.5 means that the support of $\widetilde{f}$ is located on line parallel to the lines $r=\rho$. Moreover if $\widetilde{f}$ is never zero on two segments J and J^* on different lines then the projection on the r-axis I and I^* are at distance at least one. That is to say that players with rating in I do not play against players with rating in I^*. Thus by Estimate (7.4) and Prop. 7.5, there cannot be any interaction between I_l and I_k which means that

$$d(\underline{r}_k,\ I_l)\geq 1, \quad d(\overline{r}_k,\ I_l)\geq 1.$$

That is to say that the distance between I_k and I_l is at least one. On the other hand since $w(r-r')>0$ then $|r-r'|<1$. Now $r\in[\underline{r}_k-\sqrt{\varepsilon},\ \overline{r}_k+\sqrt{\varepsilon}]$ and r' is similarly within $\sqrt{\varepsilon}$ of $\{r'',\ \exists\rho'' s.t.(r'',\ \rho'')\in I_l\}$. The only possibility is drawn on the left part on Figure 1:

$$r'\in[\underline{r}_k-1-\sqrt{\varepsilon},\ \underline{r}_k-1+\sqrt{\varepsilon}]\cup[\overline{r}_k+1-\sqrt{\varepsilon},\ \overline{r}_k+1+\sqrt{\varepsilon}]$$

Thus $|A_-^2|$ is weaker than

$$a_- = \int_{r' \in [\underline{r}_k - 1 - \sqrt{\varepsilon}, \underline{r}_k - 1 + \sqrt{\varepsilon}]} w(r - r') \left| b(\varrho - \varrho') - b(r - r') \right| f(t, r', \varrho')\, dr'\, d\varrho',$$

plus the corresponding term with $\overline{r}_k$.

Now w belongs to C^2 with compact support in $[-1,\ 1]$ so $w(-1+s) = O(s^2)$ and hence $w(r-r') = O(\sqrt{\varepsilon})^2 = O(\varepsilon)$ in the previous integral. The expression involving b is bounded in L^∞ on the compact support of f and f is bounded in $\mathcal{M}^1$ so finally

$$a_- \leq C\,\varepsilon, \qquad |A_-^2| \leq C\,\varepsilon.$$

We can conclude and estimate the sign of $a[f]$ on $\partial_- \widetilde{\mathcal{T}}_\varepsilon^k$:

$$\begin{aligned} a[f](t, r, \varrho) &= A_+ + A_- = A_+ - |A_-| \geq a_+ - C\,\varepsilon \\ &\geq c\,\varepsilon^{3/4} - C\,\varepsilon > 0, \end{aligned}$$

for ε small enough. □

7.3 Conclusion of the proof of Theorem 2.4

This is essentially a straightforward consequence of Prop. 7.6. Take $\widetilde{f} \in \Omega$ and any $\varepsilon > 0$. Take t_0 according to Prop. 7.6 and s.t. $\|f(t_0) - \widetilde{f}\|_{W^{-1,1}} \leq \varepsilon^2$. Define $\tilde{T}_\varepsilon$ as the tubular neighborhood of the support of $\widetilde{f}$ of size $\sqrt{\varepsilon}$. Then

$$\int_{\mathbb{R}^2 \setminus \tilde{T}_\varepsilon} f(t_0, r, \rho)\, dr\, d\rho \leq \varepsilon.$$

Therefore by Prop. 7.6 for any $t \geq t_0$

$$\int_{\mathbb{R}^2 \setminus \tilde{T}_\varepsilon} f(t, r, \rho)\, dr\, d\rho \leq \varepsilon.$$

Therefore and still by Prop. 7.6,

$$\|f(t) - \widetilde{f}\|_{W^{-1,1}} \leq \sqrt{\varepsilon}$$

for all $t \geq t_0$.

This proves that Ω is reduced to one unique element and that f converges.

Acknowledgments: We thank Clément Sire for fruitful discussions. P.E. Jabin was partially supported by NSF grant number DMS 1312142 and NSF grant RNMS 1107444.

References

[1] J. Barré, M. Hauray, P.E. Jabin, Stability of trajectories for N-particles dynamics with singular potential. J. Stat. Mech., (2010), doi:10.1088/1742-5468/2010/07/P07005.

[2] D. Benedetto, E. Caglioti, J. A. Carrillo, and M. Pulvirenti. *A non Maxwellian steady distribution for one-dimensional granular media.* J. Stat. Physics 91 (1998), no. 5/6, 979-990.

[3] F. Bouchut, F. Golse, M. Pulvirenti. *Kinetic equations and asymptotic theory.* Series in Applied Mathematics (Paris), 4. Gauthier-Villars, Elsevier, Paris, 2000. x+162 pp.

[4] J. A. Carrillo, R. J. McCann, and C. Villani. *Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates.* Rev. Mat. Iberoamericana 19 (2003), no. 3, 971-1018.

[5] J. A. Carrillo, R. J. McCann, and C. Villani. *Contractions in the 2-Wasserstein length space and thermalization of granular media.* Arch. Rat. Mech. Anal. 179 (2006), no. 2, 217-263.

[6] L. Desvillettes. *Hypocoercivity: the example of linear transport.* Contemp. Math. 409, (2006), 33-53.

[7] L. Desvillettes, F. Salvarani. *Asymptotic behavior of degenerate linear transport equations.* Bull. Sci. Math. 133 (2009), no. 8, 848-858.

[8] A. Elo. *The Rating of Chessplayers, Past and Present.* Arco Publishing, Inc. New York, (1978).

[9] M. Glickman. *Paired comparison models with time-varying parameters*, Ph.D. dissertation, Department of Statistics, Harvard University, (1993).

[10] M. Glickman. *Dynamic paired comparison models with stochastic variances.* J. Appl. Stat. 28 (2001), no. 6, 673-689.

[11] M. Glickman, A. Jones. *Rating the chess rating system.* Chance, 12 (1999), 2, 21–28.

[12] M. Glickman. *Rating Competitors in Online Games.* The ISBA Newsletter, 6 (1999), 8–9.

[13] S.-Y. Ha, E. Tadmor. *From particle to kinetic and hydrodynamic descriptions of flocking.* Kinet. Relat. Models 1 (2008), no. 3, 415-435.

[14] M. Hauray, P.-E. Jabin. *N-particles approximation of the Vlasov equations with singular potential.* Arch. Ration. Mech. Anal. 183 (2007), no. 3, 489-524.

[15] P.-E. Jabin, S. Motsch. *Clustering and asymptotic behavior in opinion formation.* J. Differential Equations 257 (2014), no. 11, 4165-4187.

[16] S. Motsch, E. Tadmor. *A new model for self-organized dynamics and its flocking behavior.* J. Stat. Phys. 144 (2011), no. 5, 923-947.

[17] H. Spohn, Large scale dynamics of interacting particles, *Springer-Verlag Berlin*, 1991.

[18] C. Villani. Topics in optimal transportation. Graduate Studies in Mathematics, 58. American Mathematical Society, Providence, RI, 2003. xvi+370 pp.