



**HAL**  
open science

## Contractions to update Elo ratings for round-robin tournaments

Stéphane Junca

► **To cite this version:**

Stéphane Junca. Contractions to update Elo ratings for round-robin tournaments. 2021. hal-03286591

**HAL Id: hal-03286591**

**<https://hal.science/hal-03286591>**

Preprint submitted on 14 Jul 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Contractions to update Elo ratings for round-robin tournaments

Stéphane JUNCA

## Abstract

A popular rating system to measure player strength by pairwise comparison is the Elo's system. This system is widely used to rank players in chess, go and online games. Assuming the independence of the outcomes, it is a Markov chain. The convergence of Elo ratings is studied when the update after each round-robin tournament is a contraction. An important case is the linear updating. The contractive assumption is discussed and compared to the non contractive usual Elo.

**Keywords:** Pairwise comparisons, Elo, Markov chain, iterated random functions, contractivity.

*Mathematics Subject Classification.* 60J20, 91D30.

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Convergence</b>	<b>5</b>
<b>3</b>	<b>Linear recurrence</b>	<b>6</b>
3.1	Vectorial linear Markov chain . . . . .	6
3.2	The linear Elo . . . . .	8
3.3	Why is the linear Elo not used? . . . . .	14
<b>4</b>	<b>Nonlinear recurrence</b>	<b>17</b>
4.1	Contractor in $L^\infty$ . . . . .	17
4.2	The nonlinear contractive Elo . . . . .	19
4.3	On the contractivity . . . . .	22

## 1 Introduction

The Elo rating system is widely used for games and mainly today for online games. It can be used to rank any objects when a pairwise comparison is available [5, 16]. Its simplicity and efficiency is a success of this ranking. In this paper, the Elo's system is mathematically studied under a very constraining assumption on the contractivity of the update. This contractivity is false for the original Elo's system but it is 'almost' true in practice. This point is discussed below in sections 3.3, 4.3. Moreover, the contractivity is an important tool to give some mathematical results, in particular for the fundamental linear model, in general to obtain the necessary

stability conditions for the Elo's system. It is worth saying that Arpad Elo first tried his rating system with the linear model [9, 16].

**The Elo model** The Elo rating system is first presented for 2 players. Let two players  $i$  and  $j$  with rating  $R_i^n$  and  $R_j^n$  at time  $n$ ,  $\mathbb{S}_{ij}^n \in [0, 1]$  the result of the match, Arpad Elo stated the following computations after each match to update the ratings,

$$R_i^{n+1} = R_i^n + K(\mathbb{S}_{ij}^{n+1} - \mathbf{b}(R_i^n - R_j^n)). \quad (1.1)$$

Arpad Elo considered the case where the result of the match at time  $n + 1$  is  $\mathbb{S}_{ij}^{n+1} = 0$  or  $1$  and  $\mathbf{b}(R_i^n - R_j^n)$  is the *expected score* for two players as a function of the difference of the ratings at the previous time. The function  $\mathbf{b}(\cdot)$  has usually to take values in  $[0, 1]$ . If the difference of rating is important, the logical result, the win of the stronger player, has to affect slightly the ratings. Conversely, if the weaker player wins, then, the ratings has to be more modified.

There is another interpretation of the function  $\mathbf{b}(\cdot)$ . If  $\mathbf{b}(\cdot) \equiv 0$  in (1.1) then the rating diverges towards  $+\infty$ . This is not interesting because we want to measure the strength of the player by a finite number. In some sense, the ratings are expected to converge if the strength of the players are constant. Thus, the function  $\mathbf{b}(\cdot)$  can be seen as a damp to stabilize the basic random walk model  $\mathbf{b}(\cdot) \equiv 0$ . As a stabilisator,  $\mathbf{b}(\cdot)$  can take values on  $\mathbb{R}$  as it is done below. The function can also be seen as a bonus or malus function of the result of the outcome. But  $\mathbf{b}(\cdot)$  has some minimal constraint to satisfies. For instance,  $\mathbf{b}(\cdot)$  is increasing, at least non decreasing,

$$\frac{d\mathbf{b}}{dx} \geq 0. \quad (1.2)$$

Arpad Elo was a physicist and stated that the natural following conservation law is fulfilled,

$$R_i^{n+1} + R_j^{n+1} = R_i^n + R_j^n. \quad (1.3)$$

Since, for all  $i, j$  and  $n$ , the scores satisfy the symmetry,

$$\mathbb{S}_{ij}^{n+1} + \mathbb{S}_{ji}^{n+1} = 1, \quad (1.4)$$

The function  $\mathbf{b}(\cdot)$  has to satisfies the same symmetry for all  $x$ ,

$$\mathbf{b}(x) + \mathbf{b}(-x) = 1, \quad (1.5)$$

which means that  $\mathbf{b}(0) = 1/2$  and  $x \mapsto \mathbf{b}(x) - 1/2$  is an odd function.

For the function  $\mathbf{b}(\cdot)$  Arpad Elo chose a centered distribution function of a normal law [9]. Now the logistic Elo is mainly used in practice.

$$\mathbf{b}(x) = \frac{1}{1 + 10^{-x/400}}. \quad (1.6)$$

It is motivated for 3 or more players to keep a transitivity relation beetwen the rating of players. If the player 1 is stronger than the player 2 and the player 2 is stronger than the palyer 3 then the player 1 is stronger than the palyer 3. A quantitative and more precise notion of transitivity leads to a logistic law [1, 4, 6, 19]. This important point is not more detailed here because other functions  $\mathbf{b}(\cdot)$  than the logistic function (1.6) are considered.

Other bonus-malus function  $\mathbf{b}$  can be used. For instance a linear function and more generally a non decreasing function fullfilling the symmetry relation (1.5).

The  $K$ -factor has to be small in some sense. Not too big to not drastically change the rating and not too small to expect that if the rating of a player is not the good one then the process converges fastly enough to the new rating representing the real strength of the player. Jeff Sonas studied numerically the important question of the size of  $K$  [19]. The size of  $K$  will appear below to be crucial to obtain convergence theorems.

Usually the rating is bigger than 1000 thus  $K \ll 1000$ . In this way, if the rating are normalized to be of order  $O(1)$  replacing  $R_i^n$  by  $R_i^n/1000$  in (1.1), the new factor is  $K/1000$  becomes really small, and accordingly modifying the  $\mathbf{b}(\cdot)$  function.

Arpad Elo expected that the rating of each player represents well the strength of players after many games. Mathematically that means, in some sense, that the rating's converges to a law centered at the true value representing the skill of the player. In other words,  $R_k^n$  is expected to converge in law towards  $R_k^\infty$  for  $k = 1, 2$ ,  $\mathbb{E}(R_k^\infty)$  corresponds to the real strength of player  $k$ , and,  $\mathbf{b}(\mathbb{E}(R_i^\infty) - \mathbb{E}(R_j^\infty))$  is the probability that  $i$  beats  $j$ .

Notice that the model (1.1) is invariant under translation. If  $R_i^n$  satisfies (1.1) for all  $n$ ,  $R_j^n$  the similar relation and 'r' be a constant then,  $R_i^n + r$ ,  $R_j^n + r$  also satisfy the same relation (1.1). Thus, the pertinent information is the relative strength  $R_i^n - R_j^n$ .

In all this paper only bounded independent (w.r.t.  $n$ ) score  $\mathbb{S}_{ij}^n$  are considered with the same law as  $\mathbb{S}_{ij}$  with  $\mathbb{S}_{ij} = 0$  or  $1$  or with continuous values when the score can be more quantified than a win or a loss,  $\mathbb{S}_{ij} \in [0, 1]$  or  $-M \leq \mathbb{S}_{ij} \leq M$ ,

$$\sup_{i,j} |\mathbb{S}_{ij}| \leq M < +\infty. \quad (1.7)$$

Keeping the same law means that the strength of the players is supposed to be constant. The independance of the score with respect to the time means that players are not affected by the previous score.

**Elo assumption on the expected score:** Assume that the strength of the players is constant, no learning effect as in [8]. The score expectation of an encounter only depends on the difference of the theoritical ratings  $\rho_i - \rho_j$  [14],

$$\mathbb{E}(\mathbb{S}_{ij}) = \mathbf{b}(\rho_i - \rho_j), \quad \forall n \in \mathbb{N} \quad (\text{Elo Assumption}) \quad (1.8)$$

Indeed, the notion of theoritical rating or true strength is natural but generally unknow.

Furthermore, how one number can be describe the various skills of a chess player? However, to rank players only one number is needed. In the Elo's system it is implicitly expected that  $R_i^n$  is a random variable fluctuating around the 'real strength' for large  $n$ , so,

$$\rho_i := \lim_{n \rightarrow +\infty} \mathbb{E}(R_i^n). \quad (1.9)$$

The existence of the limit is already a mathematical problem and needs some stability of the Elo's system.

In this paper, the more general case where the score expectation only depends on the player without any notion of true rating is considered:

$$\mathbb{E}(\mathbb{S}_{ij}) = s_{ij} \quad (1.10)$$

Indeed, for a Bernoulli model,  $\mathbb{S}_{ij} = 0$  or  $1$ ,  $p_{ij} = P(\mathbb{S}_{ij} = 1)$ , the matrix of scores reads,

$$s_{ij} = p_{ij}. \tag{1.11}$$

It is not assumed that the mean scores  $s_{ij}$  satisfy the Elo assumption (1.8). This is more realistic. It is well known that "circular" situations can occur, for instance:

- Player number 1 is stronger than player 2, i.e.  $p_{12} > 1/2$ .
- Player number 2 is stronger than player 3, i.e.  $p_{23} > 1/2$ .
- Player number 3 is stronger than player 1, i.e.  $p_{31} > 1/2$ .

More complex "cycles" can be studied with more players [15]. Nevertheless, it is expected that the Elo rating system ranks the players after enough time. An important question is to try to identify the 'real strength'  $\rho_i$  and to verify that the 'real strength' ranks well all the players.

**Round-robin tournaments** The fairest tournament to rank  $N$  players is the round-robin tournament when all meet all. The following evaluation of the rating after each tournament is studied.

$$R_i^{n+1} - R_i^n = K \sum_j (\mathbb{S}_{ij}^{n+1} - b(R_i^n - R_j^n)). \tag{1.12}$$

The integer  $n$  means after  $n$  tournaments. On internet, the rating is usually updated after each match. A numerical study [?] seems to show that the result is similar for a large number of matches.

The important contractivity assumption is discussed in the section 4.3. It is not true (but 'almost' true) for the usual Elo model. It is mandatory for the linearized model.

The results obtained in the contractive setting are,

- condition on the size of  $K$ ,
- convergence in law with exponential rate.
- Elo ratings are bounded.
- uniqueness of the invariant measure.
- explicit study of the linear case:
  - The Elo's system ranks well the players.
  - The Elo assumption that the expected score depends on the difference of the rating is not true.

The paper is organized as follows. The convergence in law is expounded in Section 2. Section 3 deals with the linear case. The limit and the exponential rate of convergence is explicit. Some assumptions on the Elo's system are tested. Section 4 deals with the nonlinear case. The sections 3.3 and 4.3 present the main drawback of the contractivity assumption and discuss some advantages.

## 2 Convergence

The Elo vectorial Markov chain (1.12) for  $N$  players reads,

$$R^{n+1} = R^n + K(S^{n+1} - B(R^n)) = G(R^n) + K S^{n+1}, \quad (2.1)$$

with vectorial notations,

$$\begin{aligned} R^n &= (R_1^n, \dots, R_N^n)^T, \\ \mathbf{1} &= (1, \dots, 1)^T \in \mathbb{R}^N, \end{aligned} \quad (2.2)$$

$$\begin{aligned} r &= (r_1, \dots, r_N)^T, \\ \mathbb{B}_{ij}(r) &= \mathbf{b}(r_i - r_j), \end{aligned} \quad (2.3)$$

$$B(r) = \mathbb{B}(r) \mathbf{1}, \quad (2.4)$$

$$G(r) = r - K B(r), \quad (2.5)$$

$$S^{n+1} = \mathbb{S}^{n+1} \mathbf{1}, \quad (2.6)$$

and the convention for all  $i$

$$\mathbb{S}_{ii}^{n+1} = \frac{1}{2}, \quad (2.7)$$

which corresponds to  $\mathbf{b}(0) = 1/2$  and allow us to extend the sum (1.12) also for  $j = i$  without changing the total sum.

Under contrative assumptions (2.8) on the bonus-malus function the following convergence holds.

### Theorem 2.1 (Bounds and exponential convergence in law)

If  $R^0$  is the initial random vector which is bounded in  $L^\infty(\Omega)$  and independent of the scores,  $(\mathbb{S}^n)$  is a sequence of random independent matrixes with the same bounded law as  $\mathbb{S}$  (1.7), and if the following contractance assumptions are fulfilled:

$$K(N-1) \sup \mathbf{b}' < 1, \quad 0 < \inf \mathbf{b}', \quad (2.8)$$

then, with

$$\gamma = \max(|1 - K N \sup b'|, |1 - K N \inf b'|) < 1, \quad (2.9)$$

the Markov chain  $(R^n)_n$  converges in law, with the exponential rate  $\gamma$ , to the law of the random variable

$$R^\infty + \frac{1}{N}(R^0 \cdot \mathbf{1}) \mathbf{1}$$

where  $R^\infty \perp \mathbf{1}$ ,  $R^\infty$  is independent of  $R^0$  and its law  $\mu^\infty$  is the vector valued measure supported on the hyperplane  $\mathbf{1}^\perp$  **uniquely** determined by:

$$\mu^\infty = \left( (Id - K B) \# \mu^\infty \right) \star \lambda_S, \quad (2.10)$$

where  $\#$  is the push-forward of a measure by an application,  $\star$  denotes the convolution of measures, and  $\lambda_S$  is the  $S$ ' law. The support of  $\mu^\infty$  is bounded. Let the vector  $r^\infty$  uniquely defined by  $r^\infty \perp \mathbf{1}$  and  $r^\infty = G(r^\infty) + \mathbb{E}(\mathbb{S} \mathbf{1})$  then

$$\|R^\infty - r^\infty\|_\infty \leq \frac{K}{1-\gamma} \|(\mathbb{S} - \mathbb{E}(\mathbb{S})) \mathbf{1}\|_\infty. \quad (2.11)$$

This theorem can be seen in the framework of ‘contractive iterated random functions’ [7]. The point here is that the map  $Id - KB$  is not contractive on  $\mathbb{R}^N$  but on the hyperplane  $\mathbb{I}^\perp$ .

Another point is the compact support of the limit law which is again a consequence of the contractivity (2.8) and it is not true for the used Elo model (1.6) where  $\inf b' = 0$  [15].

“The convergence of the sequence  $(R^n)$  towards a random variable  $R^\infty$  is not expected [7]. Else, by (2.1),  $R^\infty$  has to satisfy:  $B(R^\infty) = S$ . This is false (except the non interesting case when  $S$  is a constant).

Notice that, a priori, the limit depends on the factor  $K$ .

The linear case is first presented with more precise results. The limit in law is ‘explicit’, the optimal contractive assumption is provided, and the Elo assumptions (1.8) are studied. After, the convergence in law for the nonlinear case is given.

### 3 Linear recurrence

The linear case only means that the bonus-malus function is linear,

$$\mathbf{b}(r) = \frac{1}{2} + Lr. \quad (3.1)$$

In practice,  $L$  is quite small,  $L = b'(0) = \ln(10)/1600 \sim 10^{-3}$ . For small  $r$  it corresponds to the expectation to win if the ratings difference is  $r$ . It is an approximation for  $|r| < 200$ . Conversely, for  $r$  too big,  $|r| > 1/L \sim 1000$ ,  $|\mathbf{b}(r)| > 1$  so it does not correspond to a probability. Indeed, it is a malus (resp. a bonus) to have a too big (resp. a too small) Elo. Thus, after the encounter, independently on the outcome, there is a loss (resp. win) of Elo points. It is a reason why the linear function  $\mathbf{b}$  is not used. This discussion is continued in Section 3.3. Nevertheless, usually players have a rating difference  $\leq 400 \ll 1000$ , so this linear model is interesting to try to understand such typical situations. Moreover, Arpad Elo himself had started to study the linear Elo [9, 16]. Notice that the Elo ranks well the players in figures 1, 2, 3, 4.

The linear case is also an interesting simplest case to test or unvalid some usual assumptions on the Elo’s system presented in the introduction.

This section is organized as follows. First, a general result on linear Markov chains is given [7, 17]. Second, the special case of the linear Elo rating is studied. Third, some assumed behaviors of the Elo rating system are checked on the linear case.

#### 3.1 Vectorial linear Markov chain

Consider a general vectorial linear iterative Markov chain [7],

$$R^{n+1} = AR^n + S^{n+1}, \quad (3.2)$$

where  $R^0$  is given,  $R^n \in \mathbb{R}^N$ ,  $A \in M_{n,n}(\mathbb{R})$ ,  $(S^n)$  a random independent vectorial sequence in  $\mathbb{R}^N$  with the same law as  $S$ , and  $R^0$  is independent of this sequence. This is a well known subject [7]. The convergence result is recalled with a simple proof to be self-contained. A more general result can be found and proved in [7].

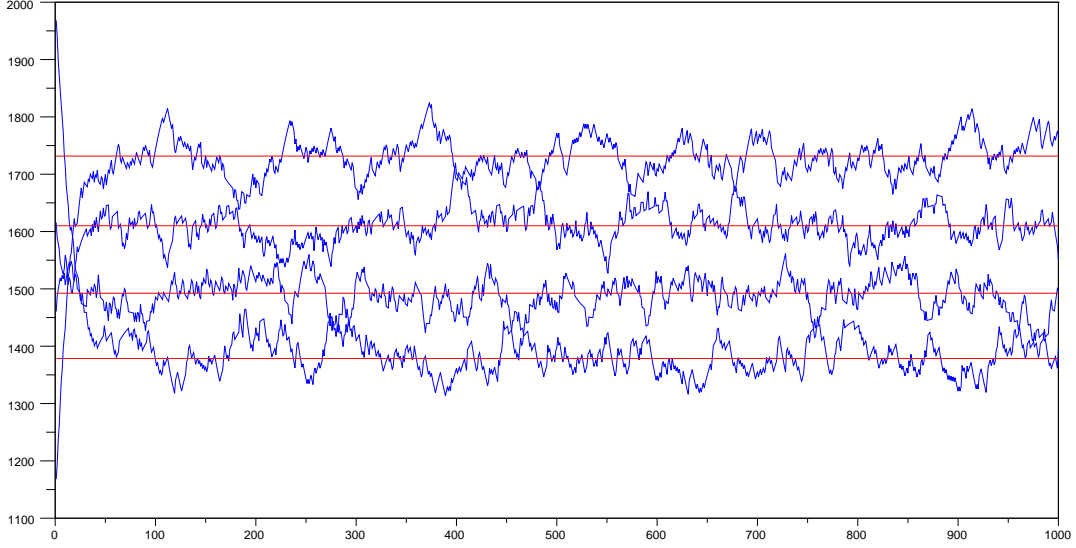


Figure 1: 4 players, number of tournaments on the horizontal axis, ratings on the vertical axis, true strength in red.

**Theorem 3.2 ( Limit bounded law, expectation, covariance)**

If the spectral radius of  $A$  satisfies  $\rho(A) < 1$ ,  $R^0$  and  $S$  is a bounded random vector,  $R^0, (S^n)$  are independent, then  $(R^n)$  converges in law towards  $\tilde{R}^\infty$  with a geometric rate,

$$R^n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \tilde{R}^\infty = \sum_{n=0}^{+\infty} A^n S^n, \quad (3.3)$$

$$\|\tilde{R}^\infty\| \leq \frac{\|S\|}{1 - \|A\|} \quad (3.4)$$

$$\sup_n \|R^n\| \leq \|R^0\| + \frac{\|S\|}{1 - \|A\|}, \quad (3.5)$$

$$\lim_{n \rightarrow +\infty} \mathbb{E}(R^n) = \mathbb{E}(\tilde{R}^\infty) = (Id - A)^{-1} \mathbb{E}(S), \quad (3.6)$$

$$\|Cov(\tilde{R}^\infty)\| \leq \frac{1}{1 - \|A\|^2} \|Cov(S)\|, \quad (3.7)$$

for a vector norm  $\|\cdot\|$  on  $\mathbb{R}^N$  such that the induced matrix norm satisfies

$$\|A\| := \sup_{R \neq 0} \frac{\|AR\|}{\|R\|} < 1.$$

The matrix  $Cov(X)$  is the covariance matrix of the random vector  $X$ ,

$$Cov(X) = \mathbb{E}([X - \mathbb{E}(X)][X - \mathbb{E}(X)]^T).$$

We cannot expect a convergence almost everywhere of the Markov chain. Indeed, if  $R^n$  converges a.e. then  $S^n$  also converges a.e. by (3.2). This convergence



contradicts the independance of the sequence  $S^n$  (except in the trivial case when  $S$  is almost surely constant).

Notice that the law of  $R^n$  is uniformly compactly supported.

**Proof :** Equation (3.2) yields:

$$R^{n+1} = A^{n+1}R^0 + \sum_{j=0}^n A^j S^{n+1-j}$$

Since  $\rho(A) < 1$  and  $R^0$  is bounded,  $A^{n+1}R^0$  converges with a geometric rate towards the vector 0. By the independance of the sequence  $(S^n)$  and  $R^0$ ,  $R^{n+1}$  has the same law as  $\tilde{R}^{n+1}$ :

$$R^{n+1} \stackrel{\mathcal{L}}{=} \tilde{R}^{n+1} = A^{n+1}R^0 + \sum_{j=0}^n A^j S^j.$$

$\tilde{R}^n$  converges uniformly towards  $\tilde{R}^\infty$  with a geometric rate.  $R^{n+1}$  and  $\tilde{R}^{n+1}$  have the same law, thus  $\|R^{n+1}\| = \|\tilde{R}^{n+1}\| \leq \|R^0\| + \sum_{j=0}^n \|A\|^j \|S\|$ , and the uniform bounds (3.4), (3.5) follow.

The convergence in law and the previous uniform bound (3.5) imply the convergence of the esperance and the variance, so we can use (3.2) or the summation of the geometric series,

$$\begin{aligned} \mathbb{E}(\tilde{R}^\infty) &= A \mathbb{E}(\tilde{R}^\infty) + \mathbb{E}(S), \\ \text{Cov}(\tilde{R}^\infty) &= A \text{Cov}(\tilde{R}^\infty) A^T + \text{Cov}(S), \\ \text{Cov}(\tilde{R}^\infty) &= \sum_{n=0}^{+\infty} A^n \text{Cov}(S) (A^T)^n, \end{aligned}$$

which concludes the proof. □

The limit is generally not a discrete law. Consider, this simple scalar example where  $S$  follows a symmetric Bernoulli law  $\mathcal{B}(1/2)$  and:

$$R^{n+1} = \frac{1}{2}(R^n + S^n), \quad R^0 \equiv 0.$$

The limit law is simply the uniform law on  $[0, 1]$ ,  $\mathcal{U}([0, 1])$ . An interesting discussion on the properties of the limit law in dimension 1 can be found in the review [7].

## 3.2 The linear Elo

Now the linear Elo model (3.1) is studied. This function satisfies (1.5), it is an increasing function and it fits well the usual bonus function (1.6) with  $L \approx 0.001$ , but the function  $\mathbf{b}(\cdot)$  is not bounded.

A key role is played by the constant vector  $\mathbf{1}$ . It is related to the conservation of the total sum of Elo ratings. It is invariant by the matrix  $A$  involved in this linear Markov chain. Thus, the matrix  $A$  is not contractive on  $\mathbb{R} \mathbf{1}$ . Hopefully it can be contractive on the hyperplane  $\mathbf{1}^\perp$ .

The important Elo assumptions (1.8) are discussed on this simplest model by computing the ‘real strengths’  $\rho_i$ ,  $i = 1, \dots, N$ , (1.9).

Let us start by some invariance properties also valid for the nonlinear case (1.12).

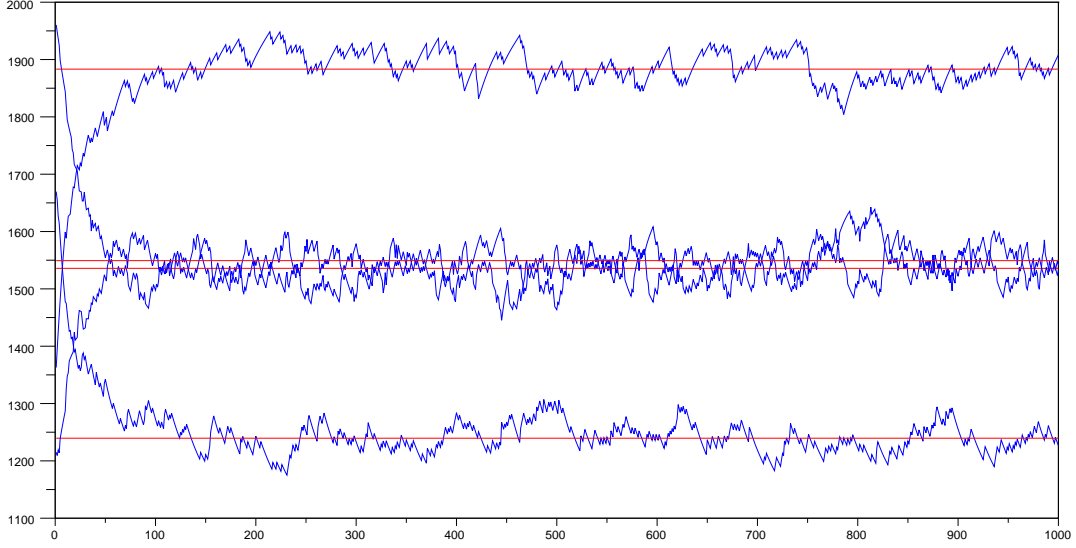


Figure 2: 4 players including two players with close strengths.

**Proposition 3.1 (Invariances and the constant vector  $\mathbf{1}$ )**

- *Invariance under translation: for any  $r \in \mathbb{R}$ ,*

$$R^n + r \mathbf{1}$$

*satisfies recurrent system (1.12). That is to say we can translate all the ratings by the same constant  $r$ :  $\underline{R}_i^n = R_i^n + r$ .*

- *Conservation of the total rating: for all  $n \in \mathbb{N}$ ,*

$$R^n \cdot \mathbf{1} = R^0 \cdot \mathbf{1},$$

*that is to say* 
$$\sum_{i=1}^N R_i^n = \sum_{i=1}^N R_i^0.$$

**Proof :** The following proof does not use the linearity of  $\mathbf{b}(\cdot)$  and is still valid for a nonlinear  $\mathbf{b}(\cdot)$ .

The Elo update (1.1) in vectorial form (2.1) can be rewritten with the difference,

$$R^{n+1} - R^n = K(S^{n+1} - \mathbb{B}) \mathbf{1}. \tag{3.8}$$

The matrix  $\mathbb{B}$  only involves the differences of ratings, so  $\underline{R}^n$  satisfies the same relation and the invariance by translation follows.

The matrix  $\mathbb{S}$  and  $\mathbb{B}$  have the same symmetry from (1.4), (1.5).

$$\mathbb{S} + \mathbb{S}^\top = \mathbb{U} = \mathbb{B} + \mathbb{B}^\top, \tag{3.9}$$

with  $\mathbb{U}_{ij} = 1$  for all  $i$  and  $j$ . That means that  $\mathbb{S} - (1/2)\mathbb{U}$  and  $\mathbb{B} - (1/2)\mathbb{U}$  are skew-symmetric.

Moreover  $\mathbb{U} = \mathbb{I} \mathbb{I}^\top$ , so the useful identities follow,

$$\mathbb{I}^\top (\mathbb{S} + \mathbb{S}^\top) \mathbb{I} = \mathbb{I}^\top \mathbb{U} \mathbb{I} = (\mathbb{I}^\top \mathbb{I})^2 = N^2 = 2 \mathbb{I}^\top \mathbb{S} \mathbb{I}, \quad (3.10)$$

$$\mathbb{I}^\top (\mathbb{B} + \mathbb{B}^\top) \mathbb{I} = N^2 = 2 \mathbb{I}^\top \mathbb{B} \mathbb{I}, \quad (3.11)$$

since  $\mathbb{I}^\top \mathbb{S} \mathbb{I} = \mathbb{I}^\top \mathbb{S}^\top \mathbb{I}$ ,  $\mathbb{I}^\top \mathbb{B} \mathbb{I} = \mathbb{I}^\top \mathbb{B}^\top \mathbb{I}$  are scalar quantities.

Thus, multiplying (3.8) on the left by  $\mathbb{I}^\top$  yields,

$$\mathbb{I}^\top (R^{n+1} - R^n) = K \mathbb{I}^\top (\mathbb{S}^{n+1} - \mathbb{B}) \mathbb{I} = K(N^2 - N^2)/2 = 0,$$

and the conservation of  $\mathbb{I}^\top R^n$  follows.

Notice that only the symmetry of  $\mathbb{S}$  and  $\mathbb{B}$  are used, so the proof is valid for the nonlinear case.  $\square$

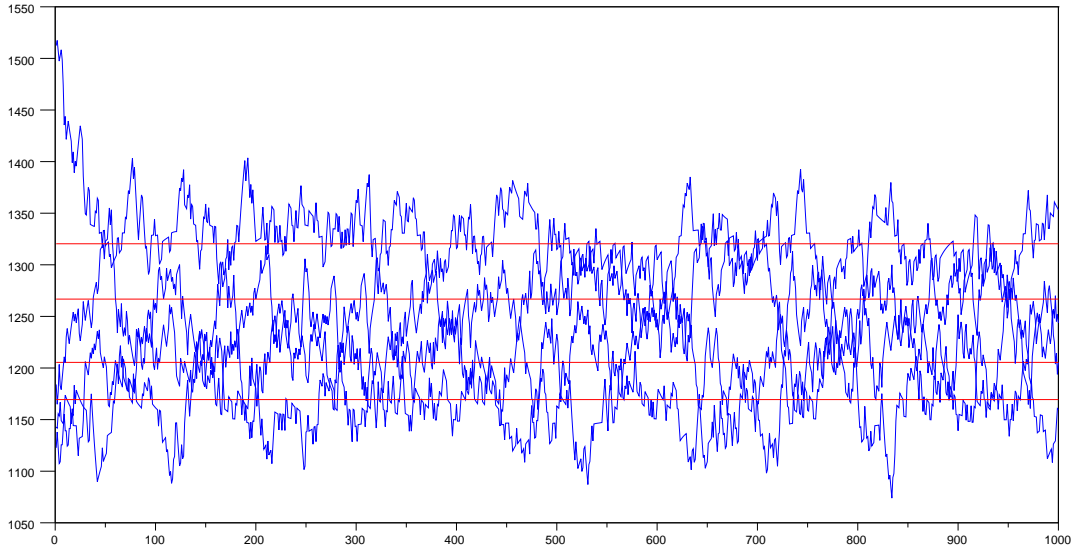


Figure 3: 4 players with close strengths. Notice that the strongest player loses points at the start because their initial ranking is too high.

Now, the vectorial function  $B(\cdot)$  is affine since  $\mathbf{b}(\cdot)$  is linear and be rewritten,

$$B(r) = ((1/2)\mathbb{U} + \hat{B}r) \mathbb{I},$$

with the the constant matrix  $\hat{B}$  defined by,

**Lemma 3.1**  $\hat{B} = L(N Id - \mathbb{U})$  where where  $\mathbb{U}$  is the the matrix defined by  $\mathbb{U}_{ij} \equiv 1$ .

**Proof :** For all  $i$ , the following computations yield the announced result,

$$\begin{aligned} \mathbb{B}_{ij}(r) &= 1/2 + L(r_i - r_j) \\ (B(r))_i &= (\mathbb{B}(r) \mathbb{I})_i = N/2 + \sum_{j=1}^N L(r_i - r_j) = N/2 + L \left( N r_i - \sum_{j=1}^N r_j \right) \\ &= \left( \left( (1/2)\mathbb{U} + L(N Id - \mathbb{U})r \right) \mathbb{I} \right)_i. \end{aligned}$$

$\square$

**Lemma 3.2** On  $\mathbb{I}^\perp$ , the matrix  $\hat{B}$  is the homothety  $LN Id$ .

$(LN)^{-1}\hat{B}$  is the matrix of the orthogonal projection on  $\mathbb{I}^\perp$  and  $\text{Ker}\hat{B} = \mathbb{R}\mathbf{1}$ .

**Proof :** The spectrum of  $\mathbb{U}$  is  $\{N, 0\}$  and the multiplicity of 0 is  $N - 1$ , since  $\text{rank}\mathbb{U} = 1$  and  $\mathbb{U}\mathbf{1} = (\mathbf{1}\mathbf{1}^\top)\mathbf{1} = N\mathbf{1}$ . In other words,  $N^{-1}\mathbb{U}$  is simply the orthogonal projector on  $\mathbb{R}\mathbf{1}$ . Thus  $\mathbb{U}$  is diagonalizable with spectrum  $\{0, N\}$  and 0 has the multiplicity  $N - 1$ . Then,  $\hat{B}$  is also diagonalizable. The spectrum of  $\hat{B}$  is  $\{0, LN\}$  and the multiplicity of  $LN$  is  $N - 1$ . Thus  $\hat{B}$  is the matrix of the linear map on  $\mathbb{R}^N$  such that the restriction on  $\mathbb{I}^\perp$  is simply  $LN Id$ .  $\square$

Now, the asymptotic behavior of  $R^n$  is stated in the linear case.

**Theorem 3.3 (Asymptotic law, expectation and covariance)** Assume  $R^0$  is a random bounded initial vector independent to the scores and

$$KLN < 2, \quad (3.12)$$

then,  $R^n$  converges in law with the geometric rate  $\gamma$ ,  $|\gamma| < 1$ ,

$$\gamma = 1 - KLN, \quad (3.13)$$

$$R^n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \frac{1}{N}(R^0 \cdot \mathbb{1})\mathbf{1} + R^\infty, \quad (3.14)$$

$$\mathbb{I} \perp R^\infty := K \sum_{n=0}^{\infty} \gamma^n S^n, \quad (3.15)$$

$$\mathbb{E}(R_i^\infty) = \frac{1}{LN} \sum_{j \neq i} s_{ij}, \quad (3.16)$$

$$\text{Cov}(R^\infty) = \frac{K^2}{1 - \gamma^2} \text{Cov}(S). \quad (3.17)$$

Note that the asymptotic expectation does not depend on  $K$ . On the contrary the standard deviation depends on  $K$ .

A consequence of the contractive assumption (3.12) is a bound of the number of players for a round-robin tournament, for  $K = 20$  used for amator chess players by the FIDE (international chess federation)  $N < 2/(KL) = 100$ , with  $K = 20$  and  $L = 1/1000$ , and for professional players (chess masters)  $K = 10$  so  $N < 200$ . In practice, the number of players for round robin-tournament is around 10 so the contractance of the update is assured.

If  $KLN = 1$ , then  $\gamma = 0$  and the convergence is reached at  $N = 1$ ,  $R^\infty = KS$ . This is a discrete random vector if  $S$  is discrete. In general a not discrete variable is expected for  $\gamma \neq 0$  [7].

**Proof :** The framework of linear Markov chain (3.2) is used with  $A = Id - KB$  and  $S^{n+1}$  is replaced by  $KS^{n+1}$ . Unfortunately,  $A\mathbf{1} = \mathbf{1}$  so, the spectral radius  $\rho(A) = 1$  and Theorem 3.2 cannot be used directly. But, the restriction of  $A$  on  $\mathbb{I}^\perp$  is simply  $(1 - KLN)Id_{\mathbb{I}^\perp} = \gamma Id_{\mathbb{I}^\perp}$ , and,  $Id_{\mathbb{I}^\perp} - A_{\mathbb{I}^\perp} = (KLN)Id_{\mathbb{I}^\perp}$ , Lemma 3.2. Thanks to the invariance by translation, Proposition 3.1,  $R^n$  can be replaced by  $R^n - \frac{1}{N}(R^0 \cdot \mathbf{1})\mathbf{1}$ . Now  $R^n \perp \mathbf{1}$  for all  $n$  and the contractivity of  $A$  on  $\mathbb{I}^\perp$  can be used. On  $\mathbb{I}^\perp$ , the limit of the rating is performing thanks to Theorem 3.2. The induced euclidean norm can be used on  $\mathbb{I}^\perp$  since  $\|A_{\mathbb{I}^\perp}\|_2 = |\gamma| < 1$ .  $\square$

**Individual Variance** The asymptotic variance of each rating depends on  $K$  and the number of players. It has the order  $K^2N$  due to size  $N \times N$  of the score matrix  $\mathbb{S}$  and the bound of the independent coefficients (1.7),

$$\begin{aligned} \text{Var}(S_i) &= \sum_{j \neq i} \text{Var}(S_{ij}) \leq (N-1)M^2 \\ \text{Var}(R_i^\infty) &= \frac{K^2}{1-\gamma^2} \text{Var}(S_i) \leq K^2N \frac{M^2}{1-\gamma^2}. \end{aligned} \quad (3.18)$$

Moreover, the contractivity assumption  $KL N < 2$  also yields to,

$$\text{Var}(R_i^\infty) \leq K \frac{2M^2}{L(1-\gamma^2)}.$$

Another way to bound the individual variance from Inequality (3.18) and finally eliminating  $K$  is,

$$\begin{aligned} \text{Var}(R_i^\infty) &\leq K^2N \frac{M^2}{1-\gamma^2} = \frac{K^2N^2L^2}{NL^2} \frac{M^2}{(1-\gamma^2)} = \frac{1}{N} \frac{(1-\gamma)^2}{(1-\gamma^2)} \frac{2M^2}{L^2}, \\ &= \frac{2}{N} \frac{1-\gamma}{1+\gamma} \left(\frac{M}{L}\right)^2 \leq \frac{2}{N} \left(\frac{M}{L}\right)^2. \end{aligned}$$

That means that bigger is  $N$ , smaller is  $K$  and smaller is the variance.

From Inequality (3.18), the variance can be seen of order  $K$  by eliminating  $N$  using the contractivity assumption,  $N < 2/(KL)$ ,

$$\text{Var}(R_i^\infty) \leq K \frac{2M^2}{L(1-\gamma^2)}.$$

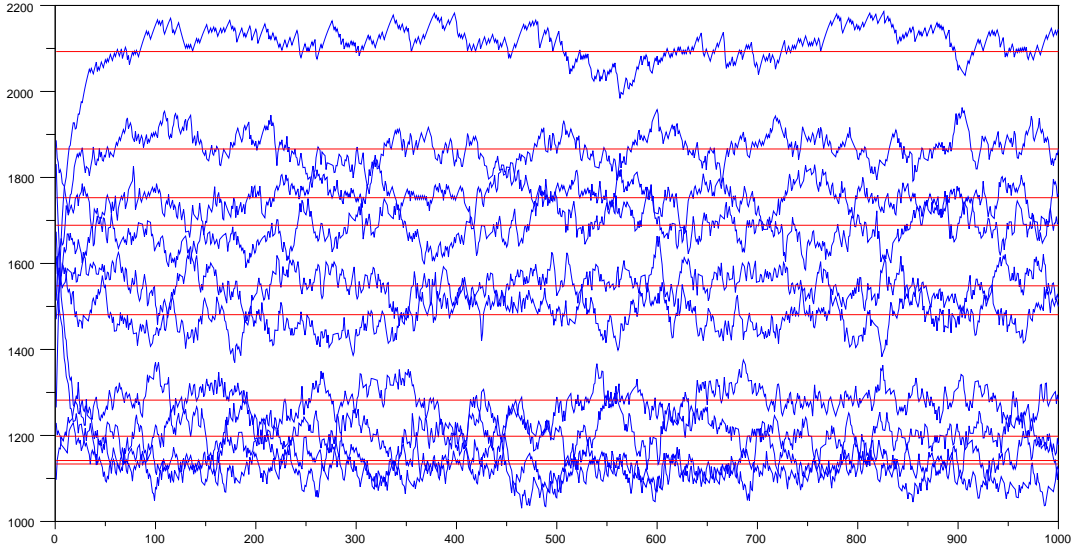


Figure 4: 10 players.

**Convergence towards the real strengths** The Elo assumption (1.8) is first supposed. That reads for  $\mathbb{E}(S_{ij})$  (1.10),

$$s_{ij} = \mathbf{b}(\rho_i - \rho_j) = L(\rho_i - \rho_j) \quad (3.19)$$

Notice that this assumption is very constraining. They are  $(N-1)N/2$  unknowns for the quantities  $(s_{ij})$  since they satisfy the symmetry (1.4). There are only  $N-1$  independent parameters  $\rho_i$  since  $\rho \perp \mathbf{1}$  or the model is invariant by translation. Therefore the nonlinear system (3.19),  $1 \leq i < j \leq N$  to define  $\rho_i$  (up to a translation) is quite overdetermined, except for the case  $N=2$ . Nevertheless, in this case (3.19) and for the linear model the ratings converge towards the  $\rho_i$  (up to a translation).

**Proposition 3.2 (About Elo assumption)** *Assume  $\rho \perp \mathbf{1}$ . If the matrix  $(s_{ij})_{ij}$  lives in the  $N-1$  vectorial subspace of matrixes defined by (3.19), that is to say the Elo assumption (1.8) is true, then, the expectation of the limit rating law is the theoretical Elo rating,  $\mathbb{E}(R_i^\infty) = \rho_i$  for all  $i$ .*

**Proof :** The limit law is  $R^\infty$  given by Theorem 3.3. It suffices to apply the equality (3.16),

$$\begin{aligned} \mathbb{E}(R_i^\infty) &= \frac{1}{LN} \sum_{j \neq i} s_{ij} = \frac{1}{LN} \sum_{j \neq i} L(\rho_i - \rho_j) \\ &= \frac{1}{N} \sum_{j=1}^N (\rho_i - \rho_j) = \rho_i - \frac{1}{N} (\rho \cdot \mathbf{1}) = \rho_i. \end{aligned}$$

□

Without the Elo assumption (1.8), the true skill is naturally given by the expected total score for a round-robin tournament.

$$s_{i\cdot} := \frac{1}{N-1} \sum_{j \neq i} s_{ij}. \quad (3.20)$$

These quantities rank the players. Indeed, under the contractive assumption (3.12) the limit rating is a biased estimation of  $s_{i\cdot}$ . Therefore the mean ratings rank well the players.

**Proposition 3.3 (Convergence towards the true skill)**

*Assume  $KL N < 2$ , then, the following convergence of the empirical rating holds*

$$\rho_i := \lim_{n \rightarrow +\infty} \mathbb{E}(R_i^n) = \frac{1}{L} \left(1 - \frac{1}{N}\right) s_{i\cdot} + \frac{1}{N} (\mathbb{E}(R^0) \cdot \mathbf{1}). \quad (3.21)$$

**Proof :** This is again a corollary of Theorem 3.3. From (3.15), the following equality holds

$$\rho_i = \mathbb{E}(R_i^\infty) + (\mathbb{E}(R^0) \cdot \mathbf{1})/N.$$

Now, as in the proof of Proposition (3.2) with equality (3.16), we have

$$L \mathbb{E}(R_i^\infty) = \frac{1}{N} \sum_{j \neq i} s_{ij} = \frac{N-1}{N} s_{i\cdot}.$$

which concludes the proof. □

### 3.3 Why is the linear Elo not used?

In this section, the score belongs to  $[0, 1]$  ( $M = 1$ ) to discuss the probabilistic interpretation of the function  $\mathbf{b}$ . Arpad Elo first assumed that the score is 0 or 1, [9]. In this case the score is a Bernoulli law and  $\mathbb{E}(S_{ij}^n) = p_{ij}$  is the constant probability such that  $i$  beats  $j$ .

For the linear Elo, the function  $\mathbf{b}$  is not confined in  $[0, 1]$ . Thus the interpretation as a probability to win is lost when  $\mathbf{b}(\Delta r) > 1$  where  $\Delta r$  is the difference of the rating between the strongest and the weakest player. The critical value is approximately the prime number 347,

$$\Delta r = \frac{1}{2L} \approx 347.44, \quad (3.22)$$

$\mathbf{b}(347) < 1 < \mathbf{b}(348)$  for the linearization of the logistic (1.6) at the origin. By symmetry,  $\mathbf{b}$  becomes negative for too negative  $\Delta r$ ,  $\mathbf{b}(-348) < 0 < \mathbf{b}(-347)$ .

Nevertheless, in this section, It is shown that the linear Elo is quite good for difference of ratings less than 347 Linear arguments are implicitly used for the initialisation of the rating, and the Elo performance. A main other drawback is the loss of Elo points for a too strong winner. However, further reflections shows that such a loss occurs only marginally for two players. In fact, it is shown below that the interval

$$I_A = \left( -\frac{1}{2L}, +\frac{1}{2L} \right) \approx ] - 347, 347[, \quad (3.23)$$

is an attractive set for the difference of the ratings and almost a stable set. Thus such dramatic events as  $\mathbf{b}(R_i^n - R_j^n)$  is not a probability or the winner losses Elo points are not so important.

**Linear approximation** The error between the linear and the logistic function  $\mathbf{b}$  is computed for different values of  $\Delta r$ ,  
error =  $(1/2 + L\Delta r) - (1/(1 + 10^{-\Delta r/400}))$ .

$\Delta r$	100	200	300	347	400	500	1000
error $\approx$	0.004	0.03	0.08	0.12	0.17	0.27	0.94

The logistic function (1.6) is quite flat and the linear approximation is good enough for  $|r| < 300$  with  $L = \mathbf{b}'(0) = \ln(10)/1600 \approx 0.0014$ .

**Elo initialisation** Due to the slow convergence of the Elo it is important to have a good initial Elo for a new chess player. Before having an Elo, the new player plays  $\nu$  games against rated players. Let  $R_1, \dots, R_\nu$  the ratings of the rated player encountered,  $S_i$  the score in  $\{0, 1/2, 1\}$ ,  $\bar{R} = (R_1 + \dots + R_\nu)/\nu$  the mean rating and  $\bar{S} = (S_1 + \dots + S_\nu)/\nu$  the mean score, the initial Elo  $R^0 = \bar{R} + b^{(-1)}(\bar{S})$ . In other words, the initialisation of the Elo  $R^0$  is uniquely defined by the relation,

$$\bar{S} = \mathbf{b}(R^0 - \bar{R}). \quad (3.24)$$

Implicitely, using means score and means rating, it is a linear way to compute the initial Elo after  $\nu$  games. It coincides with the limit rating in the linear case if the new player always has the same result in an infinite serie of round-robin tournaments with the same players (3.21).

**Elo performance** The Elo performance for a tournament is exactly computed as the Elo initialisation (3.24). Thus, again, it is a linear argument assuming implicitly that the bonus-malus function is linear.

**Loss of points for the winner** This becomes possible and it is a real drawback for the linear Elo. In the following computations the loss of Elo points is computed with  $K = 20$  for the highest rating which is the strongest player and wins the match but loses Elo points.

$\Delta r =$	347	400	500	600	700	800	900	1000
loss $\approx$	0	- 1.5	- 4	- 7	- 10	-13	-16	-19

For professional chess players (International Master)  $K = 10$  and the loss is the half. In the next paragraph, it is explained a scenario where a such big loss cannot occur.

**Bounded stable set** Consider first the case of only two players  $i$  and  $j$ . Let

$$D^n = R_i^n - R_j^n$$

the difference of rating. A stable interval is, with  $K = 20$ ,

$$I_S = \left( -\frac{1}{2L} - 2K, +\frac{1}{2L} + 2K \right) \approx ] - 387, 387[, \quad (3.25)$$

**Proposition 3.4 (2 players)** Assume the stability condition  $2KL < 1$ . If  $D^0 \in I_A$  then  $D^n$  belongs to  $I_S$  for all  $n \in \mathbb{N}$ .

Notice that the stability condition  $2KL < 1$  for  $I_S$  is stronger than the contractivity condition (3.12) which reads for two players ( $N = 2$ ),  $KL < 1$ .

The Elo is expressed with  $D^n$  thanks to the mass conservation (1.3),

$$R_i^n = m_0 + D^n/2,$$

where the total mass is  $2m_0 = R_i^0 + R_j^0$ . Thus  $R_i^n$  belongs to translated interval  $m_0 + (1/2)I_S$  for all  $n$ .

Moreover, when  $D^n$  belongs to  $I_S$  but outside  $I_A$ , the dynamics is attractive to come back in  $I_A$ . There is a particular case when  $D^n$  is exactly at the boundary of  $I_A$ ,  $D^{n_0} = 1/(2L)$  and for  $n > n_0$  the player  $i$  always wins then  $D^n$  stays at the boundary. In fact, the dramatic loss of points for the winner only occurs when  $D^n$  belongs to the set  $I_S - I_A$ .

**Proof :** By symmetries ( $D^n$ ) satisfies the recurrence relation [15],

$$D^{n+1} - D^n = 2K \left( S_{ij}^{n+1} - \mathbf{b}(D^n) \right)$$

Assume that  $D^n$  belongs to  $I_A$  and  $D^n$  is positive. The negative case can be handle in a similar way. There are two cases.

If  $D^n$  belongs to  $I_A$  then  $0 < b(D^n) < 1$  and  $D^{n+1} \leq D^n + 2K \leq 1/(2L) + 2K$ , i.e.  $D^{n+1} \in I_S$ .

If  $1/(2L) \leq D^n < 1/(2L) + 2K$ , that is  $D^n$  belongs to  $I_S$  but outside  $I_A$ , then  $1 \leq b(D^n) < 1 + 2KL \leq 2$  and  $-4K = 2K(0 - 2) < D^{n+1} - D^n$ , then



$-1/(2L) - 2K \leq 1/(2L) - 4K \leq D^n - 4K < D^{n+1} \leq D^n$ . Thus  $D^{n+1}$  cannot go outside  $I_S$ .

Finally, by induction,  $D^n$  stays in  $I_S$  for all  $n$ .  $\square$

The extension for  $N$  players is now exposed.

**Proposition 3.5 ( $N$  players)** *Let  $Nm_0 = R_1^0 + \dots + R_N^0$  be the total mass,*

$$J_A = m_0 + \left(-\frac{1}{2L}, +\frac{1}{2L}\right), \quad J_S = m_0 + \left(-\frac{1}{2L} - NK, +\frac{1}{2L} + NK\right).$$

*The set  $J_S$  is stable under condition  $KLN < \phi = (1 + \sqrt{5})/2 \approx 1.6$ .*

*Moreover, under the strong contractive assumption  $KLN < 1$ , there exists  $n_0 \geq 0$  such that, for all  $n \geq n_0$  and all  $i \in \{1, \dots, N\}$ ,  $R_i^n \in J_S$ .*

*In fact, all  $R_i^n$  approach  $J_A$  when  $n \rightarrow +\infty$ ,*

$$\limsup_{n \rightarrow +\infty} \max_{i=1}^N |R_i^n - m_0| \leq \frac{1}{2L}. \quad (3.26)$$

The mean mass  $m_0$  has a central role due to the mass conservation.

The contractive assumption used here is stronger than condition (3.12).

This proposition suggests that it is more stable to update the Elo after each match instead of each round-robin tournament since  $I_S$  is smaller than  $J_S$ . See also [15].

**Proof :** The recursion formula (1.12) is rewritten using the linearity of the function **b**. Incidentally, it gives another proof of the convergence in law for the linear case. For this purpose, introduce the notations, the total score

$$\bar{S}_i^{n+1} = \frac{S_{i1}^{n+1} + \dots + S_{iN}^{n+1}}{N}.$$

Notice that the sum is extended also for  $j = i$ . In this way, the mean score appears in the recursion formul. The mean score is quite fundamental since it is the constant  $m_0$  due to the mass conservation. Furthermore, the term for  $j = i$  does not add something since  $S_{ii}^{n+1} - b(0) = 1/2 - 1/2 = 0$ .

$$R_i^{n+1} - R_i^n = K \sum_{j=1}^N (S_{ij}^{n+1} - \mathbf{b}(R_i^n - R_j^n)).$$

Now, the mean quantities are used,

$$R_i^{n+1} - R_i^n = NK \left( \bar{S}_i^{n+1} - \frac{1}{2} - L(R_i^n - m_0) \right). \quad (3.27)$$

This formula can again be written in two ways. The first way is

$$R_i^{n+1} - R_i^n = NK \left( \bar{S}_i^{n+1} - \mathbf{b}(R_i^n - m_0) \right). \quad (3.28)$$

There is a simple interpretation of (3.28). It is like the player  $i$  plays against only one player having the mean rating, which is constant due to the mass conservation, and has the mean score  $\bar{S}_i^{n+1} \in [0, 1]$ . The constant  $K$  is replaced by  $NK$  reflecting that  $(N - 1)$  matches are played (the match  $j = i$  does not exist).

The second way is

$$R_i^{n+1} - m_0 = \gamma(R_i^n - m_0) + NK \left( \bar{S}_i^{n+1} - 1/2 \right), \quad \gamma = 1 - KLN. \quad (3.29)$$

The constant  $\gamma \in ]-1, 1[$  under the contractive assumption  $KLN < 2$ . The formula (3.29) is powerful enough to prove again Theorem (3.3).

The stability of  $J_S$  is first proven with formula (3.27) and assumption  $KLN < \phi$ . So, assume that  $R_i^n \in J_S$ . Two cases are considered,  $R_i^n \in J_A$  or  $R_i^n \notin J_A$

When all the ratings  $R_i^n$  belong to  $J_A$  then  $L|R_i^n - m_0| < 1/2$ , and  $|R_i^{n+1} - R_i^n| \leq NK$ , i.e.  $R_i^{n+1} \in J_S$ .

Now, consider that  $R_i^n$  does not belong to  $J_A$ . For instance  $R_i^n > m_0 + 1/(2L)$ , then  $L(R_i^n - m_0) > 1/2$  and  $R_i^{n+1} < R_i^n$ . Then  $R_i^{n+1}$  decreases. We have to check that it does not decrease too much and leaves  $J_S$ . Indeed, it stays in  $J_S$  thanks to the following computations,

$$\begin{aligned} R_i^{n+1} &\geq R_i^n - NK|1/2 + L(R_i^n - m_0)| \geq R_i^n - NK|1/2 + 1/2 + KLN| \\ &> (m_0 + 1/(2L) + NK) - KN(1 + \phi) \\ &= m_0 + 1/(2L) - \phi NK > m_0 - 1/(2L) - NK, \end{aligned}$$

since the last inequality is equivalent to  $(\phi - 1)KN \leq 1/L$ , but  $KLN < \phi$  and  $(\phi - 1)\phi = 1$ , ( $\phi^2 = \phi + 1$ ). .

Now, the attractivity of  $J_A$  (3.26) is proven under the stronger condition  $KLN \leq 1$ , so  $0 \leq \gamma < 1$ . Notice that  $|\bar{S}_i^{n+1} - 1/2| \leq 1/2$  since  $0 \leq \bar{S}_i^{n+1} \leq 1$ . Thus, Formula (3.29) yields,

$$|R_i^{n+1} - m_0| \leq \gamma|R_i^n - m_0| + NK/2.$$

Such inequality allows to pass to the limit sup and yields,

$$\limsup |R_i^n - m_0| \leq \frac{1}{1 - \gamma} NK/2 = \frac{NK}{2KLN} = \frac{1}{2L}.$$

In this last line the fact that there is  $\gamma$  instead of  $|\gamma|$  is crucial to the last inequality, else a bigger interval appears. Thus, (3.26) is proved with  $0 \leq \gamma < 1$ . As a consequence, there exists  $n_0$  such that all the rating  $R_i^{n_0}$  belong to  $J_S$ . Due to the stability of  $J_S$  they stay in  $J_S$  and continue to approach  $J_A$ .  $\square$

## 4 Nonlinear recurrence

For the nonlinear case, an elementary proof of the convergence and the uniqueness of the equilibria is given. This is a case well known in probability [7]. The interesting case is the study of the Elo where the map is not contractive on the whole space due to the invariance by translation of the Elo. Nevertheless, it is shown that the Elo mapping is a contractor on the hyperplane  $\mathbb{1}^\perp$  when the bonus-malus function  $\mathbf{b}$  is a contractor and  $K$  is not too big. This is enough to apply a classic result on iterated contractive random mappings.

### 4.1 Contractor in $L^\infty$

A more general result can be found in [7] using other norms and less restrictive assumption on the contractive mapping. For our purpose, the following result is now presented, self-contained and used for the contractive Elo right after.

**Theorem 4.4 [Contraction mapping]** Let  $(S^n)$  be a sequence of independent bounded random vector in  $\mathbb{R}^d$  with the same law  $S$ ,  $G(\cdot)$  a contractive map on  $\mathbb{R}^d$  with  $0 \leq \gamma < 1$  be its Lipschitz constant,  $R^0$  be a random bounded variable independent to  $(S^n)$  and the Markov chain  $R^n$  defined as follow,

$$R^{n+1} = G(R^n) + S^{n+1}. \quad (4.1)$$

Then the sequence  $(R^n)$  converges in law towards a bounded random vector  $R^\infty$  such that its law  $\mu^\infty$  is the vector valued measure uniquely determined by,

$$\mu^\infty = (G\#\mu^\infty) \star \lambda_S,$$

where  $\lambda_S$  is the measure of the  $S$  law.

Furthermore, let  $r^\infty$  be the constant vector uniquely defined by  $r^\infty = G(r^\infty) + \mathbb{E}(S)$  then,

$$\|R^\infty(\omega) - r^\infty\|_{L^\infty(\Omega, \mathbb{R}^d)} \leq \frac{1}{1-\gamma} \|S - \mathbb{E}(S)\|_\infty. \quad (4.2)$$

Notice that, we cannot expect the convergence of the sequence  $R^n$ , else from the recurrence relation that means that  $S^n$  converges. Since  $(S^n)$  is i.i.d, that means that  $S^n$  is constant.

**Remark 4.1** Although, the limit expectation is unknown, (4.2) shows that the limit law is close to a fixed point, closer if all the  $d$  standard deviation  $(S_i)_{1 \leq i \leq d}$  are small.

**Proof :** The idea is to prove the convergence of another sequence  $(W_n)$  such that  $W_n$  has the same law as  $R^n$  by reversing the time [7] as for the linear case.

$$V_{k+1}^{n+1} = G(V_k^n) + S^{n+1-k}, \quad k \leq n \quad \text{and} \quad V_0^n = R^0.$$

Since  $(S^{n+1-k})_{n \leq N}$  and  $(S^{n+1})_{n \leq N}$  have the same law,  $W_{n+1} = V_{n+1}^{n+1}$  has the same law as  $R^{n+1}$ . It remains to prove the convergence of the sequence  $(W_n)$ . For  $k = n, n-1, \dots, 1$ , since the term  $S^{n+1-k}$  is time reversed, the simplifications of the source term follows,

$$\begin{aligned} V_{k+1}^{n+1} - V_k^n &= G(V_k^n) - G(V_{k-1}^n), \\ |V_{k+1}^{n+1} - V_k^n| &\leq \gamma |V_k^n - V_{k-1}^n|, \\ |W_{n+1} - W_n| &= |V_{n+1}^{n+1} - V_n^n| \leq \gamma^n |V_1^{n+1} - V_0^n| \\ &\leq \gamma^n |G(R^0) + S^n - R^0| \leq \gamma^n C, \end{aligned}$$

where the constant  $C \leq \|R^0\|_\infty + \|G(R^0)\|_\infty + \|S\|_\infty$ . Thus  $(W_n)$  converges uniformly and  $(R^n)$  converges in law. Passing to the limit in the relation (4.1) gives the equation for the limit law (4.2).

The uniqueness is the consequence of the previous convergence and the contractance. The proof of the uniqueness is simpler than the proof of convergence. Assume there is two equilibria. Takes  $R^0$  and  $\tilde{R}^0$  two random variables associated to these two equilibria. Let  $R^n$  and  $\tilde{R}^n$  be the associated sequence at time  $n$  with the same sequence  $(S)$  in (4.1) but with different initial random variables  $R^0$  and  $\tilde{R}^0$ . Thus  $R^n$  and  $\tilde{R}^n$  have the same law has their initial data. Moreover,  $R^n - \tilde{R}^n$  satisfies an homogenous recurrence relation,  $R^{n+1} - \tilde{R}^{n+1} = G(R^n) - G(\tilde{R}^n)$  so

$|R^{n+1} - \tilde{R}^{n+1}| \leq \gamma |R^n - \tilde{R}^n|$  and  $|R^n - \tilde{R}^n| \leq \gamma^n |R^0 - \tilde{R}^0|$ . By contractance and the uniform bound,  $R^n - \tilde{R}^n \rightarrow 0$  uniformly. Thus the limit laws are equal.

It turns to get the  $L^\infty$  estimate (4.2). Summing up the inequality  $|W_n - W_{n-1}| \leq \gamma^{n-1} |G(R^0) + S^n - R^0|$  yields the upper bound

$$|W_n - W_0| \leq |W_n - W_{n-1}| + \dots + |W_1 - W_0| \leq |G(R^0) + S^n - R^0| / (1 - \gamma).$$

The limit is independent of the initial data  $R^0$  due to the previous uniqueness result. With  $R^0 = r^\infty$ , since  $W^0 = R^0$ , and taking advantage that  $r^\infty$  is a fixed point, the inequality becomes,

$$\begin{aligned} |W_n - r^\infty| &\leq |G(r^\infty) + \mathbb{E}(S) - r^\infty + (S^n - \mathbb{E}(S))| / (1 - \gamma) \\ &= |S^n - \mathbb{E}(S)| / (1 - \gamma) \leq \|S - \mathbb{E}(S)\|_\infty / (1 - \gamma). \end{aligned}$$

Passing to the limit the same bound is valid for  $R^\infty$ ,

$|R^\infty - r^\infty| \leq \|S - \mathbb{E}(S)\|_\infty / (1 - \gamma)$  which concludes the proof.  $\square$

## 4.2 The nonlinear contractive Elo

To use the general theorem 4.4 on contractive map, we recall that the Markov chain associated to the Elo (1.12) can be rewritten in the vectorial form (2.1) as follows:

$$\begin{aligned} R^{n+1} &= G(R^n) + K S^n, \\ G(r) &= r - K B(r). \end{aligned}$$

The contractive property of  $G(\cdot)$  depends on the properties on  $B(\cdot)$ . So, the map  $B(\cdot)$  is first studied. Notice that  $B(\cdot)$  is not injective on  $\mathbb{R}^N$  since

$$\mathbb{B}(\mathbb{R} \mathbf{1}) \equiv \frac{1}{2} \mathbf{U}, \quad B(\mathbb{R} \mathbf{1}) \equiv \frac{N}{2} \mathbf{1}, \quad (4.3)$$

from (3.9). However  $B(\cdot)$  is a bijection on  $\mathbf{1}^\perp$  under the assumption,

$$\inf_{x \in \mathbb{R}} \mathbf{b}'(x) > 0. \quad (4.4)$$

### Proposition 4.6 (Monotonicity of the vectorial map $B(\cdot)$ on $\mathbf{1}^\perp$ )

If (4.4) is satisfied then, the map  $B(\cdot)$  is a diffeomorphism from  $\mathbf{1}^\perp$  to  $\mathbf{1}^\perp$ ,  $B(\mathbf{1}^\perp) = \mathbf{1}^\perp$ . and the map  $B(\cdot)$  is monotonic and coercive on  $\mathbf{1}^\perp$ ,

$$(B(r) - B(r')) \cdot (r - r') \geq \frac{(\inf \mathbf{b}')}{2} (\|r - r'\|_{\mathbf{1}^\perp})^2, \quad \forall r, r' \in \mathbb{R}^N, \quad (4.5)$$

where  $(\|r\|_{\mathbf{1}^\perp})^2 = \sum_{i=1}^N \sum_{j=1}^N (r_i - r_j)^2$  is an euclidean norm on  $\mathbf{1}^\perp$ .

The map  $B(\cdot)$  is a gradient,

$$B(r) = \nabla \psi(r_1, \dots, r_N), \quad (4.6)$$

$$\psi(r) = \frac{1}{2} \left( \sum_i \sum_j \beta(r_i - r_j) - N \sum_{i=1}^N r_i \right), \quad \beta(x) = \int_0^x \mathbf{b}(y) dy.$$

**Proof :** First,  $B(\cdot)$  is constant on  $\mathbb{R} \mathbf{1}$  (4.3) since  $\mathbf{b}(0) = 1/2$ . It can be seen as the general symmetry relation (3.9) which yields  $B(t \mathbf{1}) = (N/2) \mathbf{1}$  for all  $t \in \mathbb{R}$  as in (3.11) for the linear case.

Second, the coercive estimate (4.5) is performed on the scalar product  $s$  defined by,

$$s := (B(r) - B(r')) \cdot (r - r') = \sum_{i=1}^N \sum_{j=1}^N (\mathbf{b}(r_i - r_j) - \mathbf{b}(r'_i - r'_j))(r_i - r'_i).$$

Exchanging  $i$  and  $j$  and using the symmetry relation (1.5), another formula is obtained for  $s$ , as in [14],

$$\begin{aligned} s &= \sum_{i=1}^N \sum_{j=1}^N (\mathbf{b}(r_j - r_i) - \mathbf{b}(r'_j - r'_i))(r_j - r'_j) \\ &= \sum_{i=1}^N \sum_{j=1}^N ((1 - \mathbf{b}(r_i - r_j)) - (1 - \mathbf{b}(r'_i - r'_j)))(r_j - r'_j) \\ &= - \sum_{i=1}^N \sum_{j=1}^N (\mathbf{b}(r_i - r_j) - \mathbf{b}(r'_i - r'_j))(r_j - r'_j) \end{aligned}$$

Adding the two expressions of  $s$  yields,

$$2s = \sum_{i=1}^N \sum_{j=1}^N (\mathbf{b}(r_i - r_j) - \mathbf{b}(r'_i - r'_j))((r_i - r'_i) - (r_j - r'_j))$$

The monotonicity on  $\mathbf{b}$  with  $\delta = \inf \mathbf{b}' > 0$  yields to a lower bound for  $2s$ ,

$$\begin{aligned} (\mathbf{b}(x) - \mathbf{b}(y))(x - y) &\geq \delta(x - y)^2 \\ 2s &= \sum_{i=1}^N \sum_{j=1}^N (\mathbf{b}(r_i - r_j) - \mathbf{b}(r'_i - r'_j))((r_i - r_j) - (r'_i - r'_j)) \\ &\geq \sum_{i=1}^N \sum_{j=1}^N \delta(r_i - r'_i - (r_j - r'_j))^2 = \delta \|r - r'\|_{\mathbf{1}^\perp}^2, \end{aligned}$$

which gives the coercive estimate (4.5)

Third,  $\|r\|_{\mathbf{1}^\perp}$  is an euclidean norm on  $\mathbf{1}^\perp$  associated with the scalar product,

$$(r, r')_{\mathbf{1}^\perp} := \sum_{i=1}^N \sum_{j=1}^N (r_i - r_j)(r'_i - r'_j).$$

The bilinear form is non negative and vanishes on  $\mathbb{R} \mathbf{1}$ . On  $\mathbf{1}^\perp$ , for  $r \neq 0_{\mathbf{1}^\perp}$ ,  $\|r\|_{\mathbf{1}^\perp} > 0$  since the condition  $0 = \mathbf{1}^\top r = \sum r_i$  and  $r \neq 0$  insures that there exist  $i$  and  $j$  such that  $r_i \neq r_j$ .

Fourth,  $B(\cdot)$  is a gradient,

$$\begin{aligned} 2 \frac{\partial \psi}{\partial r_k} &= \left( \sum_{j=1}^N (\beta'(r_k - r_j) - \beta'(r_j - r_k)) \right) - N \\ &= \left( \sum_{j=1}^N (\mathbf{b}(r_k - r_j) + (\mathbf{b}(r_k - r_j) - 1)) \right) - N = 2B_k(r). \end{aligned}$$

Finally, the diffeomorphism property is a consequence of the monotonicity.  $\square$

Now, the contractivity is obtained under the assumption  $\inf \mathbf{b}' > 0$  (4.4).

**Proposition 4.7 (Contractive Elo on  $\mathbb{I}^\perp$ )**

If assumption (4.4) is fulfilled and

$$K(N-1) \sup \mathbf{b}' < 1, \quad (4.7)$$

then,  $G(\cdot)$  is a contractive map on  $\mathbb{I}^\perp$ .

Notice that  $G(\cdot)$  is not contractant on  $\mathbb{R}^d$  since  $G(t \mathbf{1}) = (t - KN/2) \mathbf{1}$  for all  $t \in \mathbb{R}$ , (4.3). Indeed, the optimal Lipschitz constant of  $G(\cdot)$  on  $\mathbb{R}^N$  is 1.

The contractivity condition (4.7) is similar as the linear case (3.12) (and exactly the same for  $N = 2$ ).

**Proof :** The differential of  $G(\cdot)$ ,  $DG$  is a symmetric matrix from (4.6),

$$DG = Id - K \nabla^2 \psi. \quad (4.8)$$

The symmetry of  $DB$  can be checked directly and the formula will be used later,

$$\begin{aligned} \frac{\partial B_i}{\partial r_i} &= \sum_j \mathbf{b}'(r_i - r_j), & \forall i \\ \frac{\partial B_i}{\partial r_j} &= -\mathbf{b}'(r_i - r_j), & \forall j \neq i. \end{aligned}$$

The function  $\mathbf{b}'$  is even since  $x \mapsto \mathbf{b}(x) - 1/2$  is an odd function.

Notice that  $DB \mathbf{1} = 0$  from the previous formula or from the invariance of the hyperplane  $\mathbb{I}^\perp$  under  $B(\cdot)$  and also  $G(\cdot)$ . On  $\mathbb{I}^\perp$ ,  $DB$  correspond to a positive quadratic form (4.5). Now, the eigenvalues of  $DB$  and then  $DG$  has to be estimated. Let  $\lambda$  be an eigenvalue of  $DB$  with the associated eigenvector  $v \in \mathbb{I}^\perp$ ,  $v \neq 0$ . Let us assume  $v_i = \max_j v_j > 0$ , else change  $v$  by  $-v$ . Notice that there exists  $k \neq i$  such that  $v_i > v_k$  since  $v \notin \mathbb{R} \mathbf{1}$  (not all components can be equal). Using  $(v_i - v_k) > 0$  and  $\inf \mathbf{b}' > 0$  yield,

$$\lambda v_i = \sum_{j \neq i} b'(r_i - r_j)(v_i - v_j) \geq b'(r_i - r_j)(v_i - v_k) > 0,$$

thus  $\lambda > 0$ , which is also a consequence of (4.5). Now, more precise estimates are obtained on  $\lambda$  with  $L = \sup \mathbf{b}'$  (4.10),

$$\begin{aligned} DB_{ii} = (e_i, DB e_i) &\leq \lambda \leq \max_i \sum_j |\partial_j DB_{ij}|, \\ \sum_{j \neq i} b'(r_i - r_j) &\leq \lambda \leq 2 \max_i \sum_{j \neq i} b'(r_i - r_j) \\ (N-1) \inf \mathbf{b}' &\leq \lambda \leq 2L(N-1). \end{aligned} \quad (4.9)$$

The associated eigenvalue of  $DG$  is  $1 - K\lambda \in ]1 - 2KL(N-1), 1[$ . Therefore  $G(\cdot)$  is contractant on  $\mathbb{I}^\perp$  if  $1 - 2KL(N-1) > -1$  which concludes the proof.  $\square$

We can now prove the convergence in law of the ratings  $(R_i^n)$  when  $n$  goes towards  $+\infty$  under the contractive assumption (2.8), that is the Theorem 2.1.

**Proof :** To use the contractivity of  $G(\cdot)$  on the hyperplane  $\mathbb{I}$ , thanks the invariance by translation, Proposition 3.1 valid in the nonlinear case,  $R_i^n$  is replaced by

$$\mathring{R}_i^n = R_i^n - (1/N) \sum_j R_j^0$$

without affecting the stochastic process. Now  $\mathring{R}^0 \perp \mathbb{I}$ . Moreover,  $\mathring{R}^n \perp \mathbb{I}$  for all  $n$  due to the conservation of the total sum of ratings, Proposition 3.1. The conditions (2.8) allow to use Propositions 4.7 and Theorem 4.4 and to conclude the proof.  $\square$

### 4.3 On the contractivity

The contractivity is an important tool to study of the convergence of the linearized model or general nonlinear models. But this assumption is not fulfilled for the Elo's system used in practice. First, it is explained why the usual Elo rating system is not contractive. Second some consequences are discussed. In particular many comments valid for the linear case in Section (3.3) can be extended in the nonlinear case.

It first turns to present the non contractivity of the usual Elo model. The Elo model with the logistic function  $\mathbf{b}(\cdot)$  (1.6) is not contractive. Let us explain this point for two players  $i$  and  $j$  as in the introduction. The two equations (1.1) -one for  $i$  and one for  $j$ - to update the Elo can be rewritten as a single equation for the difference of the rating, using the symmetries of the scores and the bonus function  $\mathbf{b}(\cdot)$ , [15],

$$D_{ij}^{n+1} = g(D_{ij}^n) + 2K \mathbb{S}_{ij}^{n+1}, \quad D_{ij}^n = R_i^n - R_j^n, \quad g(x) = x - 2K \mathbf{b}(x).$$

The question is the *contractance* or not of the function  $g(\cdot)$ .

The function  $\mathbf{b}(\cdot)$  is in practice the increasing function (1.6) satisfying,

$$L := \sup \mathbf{b}' = \mathbf{b}'(0) = \ln 10 / (4 \times 400) = 0,001439116... \simeq 10^{-3} \quad (4.10)$$

$$0 = \inf \mathbf{b}'. \quad (4.11)$$

Therefore, the Lipschitz constant of  $g$  is  $\max(|1 - 2KL|, 1) \geq 1$ . The function  $g$  is never contractive. At least, the condition  $KL \leq 1$  is natural to avoid instability. However, on any compact set  $[-A, A]$  and assuming  $KL < 1$ , the Lipschitz constant of  $g$  is  $0 < \max(|1 - 2KL|, 1 - K\mathbf{b}'(A)) < 1$ . So, if the ratings belongs in the same bounded set then  $g$  is contractive. The chess players genrally believe that the Elo is bounded. In practice, it is. But theoretically, it is not [15]. It can growth like  $\ln n$  but with very small propabilities [15]. Nevertheless, with a contractivity in average [7], it is possible to prove the convergence in law for two players [2, 3]. So understanding the contractivity conditions can be useful to deal with the non contractive case.

There is a worse consequence on the condition (4.4), the function is unbounded,  $\lim_{\pm\infty} \mathbf{b}(x) = \pm\infty$ . It is clear at  $+\infty$  and the other limit comes from the skew-symmetry of  $\mathbf{b}$ , (1.5). Thus, the interpretation of  $\mathbf{b}(\rho_i - \rho_j)$  as the expected score is loss when  $\mathbf{b}(\rho_i - \rho_j) \notin [0, 1]$ .

The main draw backs of this unboundeness is the possible loss of points for the winner. It is possible that a very strong player will win against a very weak one and lose Elo points. This point is discussed in section 3.3 for the linear Elo. It

is explained that this case occurs when the difference of ratings are too large and that this case is marginal for two players due to confinement of the ratings near an attractive zone.

Indeed this drawback is already a drawback in the chess world. Players know that playing against lower rated player is dangerous, because they have to win. The loss is very costly. And professional chess players have asked for a smaller  $K$ -factor ( $K = 10$ ) to risk less precious Elo points in such an encounter. This leads to another interesting problem, how take into account different species of players with a different  $K$ -factor. In particular the conservation of total sum of the Elo ratings is lost. This is already a problem for the Glicko an extension of the Elo by Glickman [10, 11, 12, 13]. Problems like this show that there are a lot of unanswered questions about Elo's system. On the other hand, the linear case and the nonlinear contractive one is a step towards better understanding the Elo's system.

## References

- [1] Aldous, David: Elo ratings and the sports model: a neglected topic in applied probability? *Statist. Sci.* 32 (2017), no. 4, 616–629.
- [2] Avdeev, Vadim A.: Stationary distribution of the player rating in the Elo model with one adversary. (Russian) *Diskret. Mat.* 26 (2014), no. 4, 3–14; translation in *Discrete Math. Appl.* 25 (2015), no. 3, 121–130.
- [3] Avdeev, Vadim A.: Local contractibility of a process of player rating variation in the Elo model with one adversary. (Russian) *Diskret. Mat.* 27 (2015), no. 1, 3–21; translation in *Discrete Math. Appl.* 25 (2015), no. 5, 261–276.
- [4] Bradley, Ralph Allan; Terry, Milton E.: Rank analysis of incomplete block designs. I. The method of paired comparisons. *Biometrika* 39 (1952), 324–345.
- [5] Cattelan, Manuela: Models for paired comparison data: a review with emphasis on dependent data. *Statist. Sci.* 27 (2012), no. 3, 412–433.
- [6] Chetrite, Raphael; Diel, Roland; Lerasle, Matthieu: The number of potential winners in Bradley-Terry model in random environment. *Ann. Appl. Probab.* 27 (2017), no. 3, 1372–1394.
- [7] Diaconis, Persi; Freedman, David: Iterated random functions. *SIAM Rev.* 41, No. 1, 45–76 (1999).
- [8] Düring, Bertram; Torregrossa, Marco; Wolfram, Marie-Therese: Boltzmann and Fokker-Planck equations modelling the Elo rating system with learning effects. *J. Nonlinear Sci.* 29 (2019), no. 3, 1095–1128.
- [9] Elo, Arpad E.: *The Rating of Chessplayers, Past and Present*. Arco Pub., Inc. New York, (1978), 206 p.
- [10] Glickman, Mark Evan: *Paired comparison models with time-varying parameters*, Ph.D. dissertation, Department of Statistics, Harvard University, (1993).
- [11] Glickman, Mark Evan: *Rating Competitors in Online Games*. The ISBA Newsletter, 6 (1999), 8–9.
- [12] Glickman, M. E.; Jones, A. C.: *Rating the chess rating system*. *Chance*, 12 (1999), 2, 21–28.
- [13] Glickman, Mark Evan: *Dynamic paired comparison models with stochastic variances*. *J. Appl. Stat.* 28 (2001), no. 6, 673–689.



- [14] Jabin Pierre-Emmanuel; Junca, Stéphane: *A continuous model for ratings*. SIAM J. Appl. Math. 75 (2015), no. 2, 420-442.
- [15] Krifa, Adrien; Spinelli, Florian; Junca, Stéphane: On the convergence of the Elo rating system for a Bernoulli model and round-robin tournaments. Master project, Université Côte d'Azur, 2021, hal-03286065.
- [16] Langville, Amy N.; Meyer, Carl D.: *Who's # 1? The science of rating and ranking*. Princeton University Press, Princeton, NJ, 2012. xvi+247 pp.
- [17] Meyn, Sean P.; Tweedie, Richard L.: *Markov Chains and Stochastic Stability*, Communications and Control Engineering Series. Springer, 1993. xvi+ 548 pp.
- [18] Sonas, Jeff: Chessmetrics. <http://www.chessmetrics.com>.
- [19] Wikipedia: Elo rating system [Online; accessed 19-February-2013]. Jeff Sonas [Online; accessed 13-May-2021].

LJAD, Inria & CNRS, Université Côte d'Azur, Nice, France

<https://cv.archives-ouvertes.fr/stephane-junca>

[stephane.junca@univ-cotedazur.fr](mailto:stephane.junca@univ-cotedazur.fr)

July 14, 2021