On the Limitations of the Elo, Real-World Games are Transitive, not Additive

Quentin Bertrand
Mila and Université de Montréal

Wojciech Marian Czarnecki
VoyLab

Gauthier Gidel
Mila and Université de Montréal
Canada CIFAR AI Chair

Abstract

The Elo score has been extensively used to rank players by their skill or strength in competitive games such as chess, go, or StarCraft II. The Elo score implicitly assumes games have a strong additive—hence transitive—component. In this paper, we investigate the challenge of identifying transitive components in games. As a starting point, we show that the Elo score provably fails to extract the transitive component of some elementary transitive games. Based on this observation, we propose an alternative ranking system that properly extracts the transitive components in these games. Finally, we conduct an in-depth empirical validation on real-world game payoff matrices: it shows significant prediction performance improvements compared to the Elo score.

1 INTRODUCTION

Accurately evaluating and ranking players in games has been a concern for decades. For instance, in the context of chess, Elo (1961) stated that "The rating system shall provide: as close an estimate as possible of the current playing strength of an individual as computed from his performance in competition with other players". The goal of the Elo score was to rank the players by their "skill" or "strength". In this model, the probability of a confrontation outcome between two players is an increasing function of the difference in their Elo scores.

More recently, a surge of interest in evaluation and ranking for games has been triggered by the successes of multi-agent reinforcement learning agents in increasingly complex environments (Vinyals et al., 2019; Berner et al., 2019; Liu et al., 2021). Thus, having accurate and scalable ways to evaluate agents becomes crucial to monitoring agents' learning, selecting the best agents, and performing evolutionary algorithms (Jaderberg et al., 2019). This interest yields alternatives and extensions of the Elo score (Herbrich et al., 2006; Minka et al., 2018) which take into account the sequential aspect of the games (Cardoso et al., 2019), to quantify potential uncertainty on the Elo score (Glickman, 2001; Herbrich et al., 2006; Sismanis, 2010). In addition, several regularizations can be added to the Elo score to improve its generalization (Sismanis, 2010) and sample complexity (Yan et al., 2022).

Self-play learning algorithms can now easily beat humans at games such as chess, shogi, or go (Silver et al., 2018; Bansal et al., 2018; Jaderberg et al., 2019), which are considered mostly transitive (if a beats b, and b beats c, then a beats c). However, designing learning algorithms for highly non-transitive games such as StarCraft II[1] is much more challenging (Lanctot et al., 2017; Balduzzi et al., 2018, 2019). For instance, Vinyals et al. (2019) explains that self-play "is more forgetful" when trained to play StarCraft II. To "avoid cycles", Vinyals et al. (2019) considered a form of play against a population of agents "by playing against all previous players in the league".

Even though cyclic behaviors are observed in practice: in the games Age of Empires II (Horrigan, 2022), StarCraft II, (Vinyals et al., 2017) or Dota 2 (Berner et al., 2019). It seems that for many players, the transitive property is still valid (see for instance Figure 7a in Section 4). That is why the Elo score has been used to rank participants for multiple games and sports, from chess to basketball (NBA) or football (FIFA).

These notions of cyclicity and transitivity remain relatively qualitative and usually require expert domain knowledge to be assessed. To understand why some games are more cyclic than others, we propose quantifying the amount of cyclicity and transitivity in real-world games. More precisely, given n players and their empirical payoff matrix $(\mathrm{P}_{ij})_{1\leq i,j\leq n}$,[2] we propose new scores to answer the question:

[1] StarCraft II is considered to have cyclic components: a beats b, who beats c, who beats a.

[2] We would like to emphasize that our goal differs from standard empirical game theory (Walsh et al., 2003; Wellman, 2006) which aims at approximating the Nash of the game.

Proceedings of the 26th International Conference on Artificial Intelligence and Statistics (AISTATS) 2023, Valencia, Spain. PMLR: Volume 206.

can one quantify the amount of transitivity of a zero-sum two-player game from empirical data?

The contributions of this work are the following:

- We propose a disc decomposition (Theorem 2) which yields a quantitative and tractable definition of the amount of transitivity of a real-world game.
- We compute this quantity for several real-world games, including chess and StarCraft II, which yields better predictions for new matchup outcomes.

The paper is organized as follows. Section 2 provides some background on zero-sum games and the Elo score. The proposed disc decomposition, based on the normal decomposition of skew-symmetric matrices, is developed in Section 3. Experiments on payoff matrices coming from real-world games played by bots and humans are provided in Section 4. All proofs can be found in the appendix.

Notation Capital bold letters denote matrices, and lowercase bold letters denote vectors. The set of integers from 1 to n is denoted $[n]$. The sigmoid function is written as σ, and its inverse is written as logit. The binary cross entropy $(x, \hat{x}) \mapsto -x \log(\hat{x}) - (1-x)\log(1-\hat{x})$ is denoted bce.

Related Work We consider Balduzzi et al. (2018, 2019) to be our closest related works. Relying on combinatorial Hodge theory (Jiang et al., 2011), Balduzzi et al. (2018) proposed an extension of the Elo score, called m-Elo, which can express potential cyclic components. Given a payoff matrix $\mathbf{P}$ (Definition 2.1) they impose a transitive component $\mathbf{u}^{\mathrm{Trans}} \triangleq (\sum_j \mathrm{logit}(\mathbf{P})_{ij})_{i\in[n]}$, then they compute the normal decomposition (Theorem 1) of the matrix $\mathrm{logit}(\mathbf{P}) - (\mathbf{u}^{\mathrm{Trans}}\mathbf{1}^\top - \mathbf{1}\mathbf{u}^{\mathrm{Trans}\top})$. Unlike the m-Elo, we do not impose the transitive component to correspond to an Elo score. In this work, we instead propose to directly compute the normal decomposition of $\mathrm{logit}(\mathbf{P})$ and provide a principled result (Theorem 2) to interpret its main component in terms of transitivity of the induced disc game. Moreover, we show how to handle missing and infinite (when the probability of winning is 0 or 1 the logit is infinite) entries (Section 3.3). Balduzzi et al. (2019) proposed to compute the normal decomposition of $2\mathbf{P} - 1$ and visualize it as 2-dim embeddings but did not provide theorems or insights to interpret it. For ourselves we leverage the symmetric zero-sum game structure, on the other hand, previous works from the pairwise comparison community (Shah and Wainwright, 2017) relied on threshold singular value decomposition (Chatterjee, 2015) to provide statistical guarantees on the payoff matrix estimation.

While we focus on player strength evaluations for better matchup predictions, a related line of work consists of "only" ranking players without matchup predictions. Czarnecki et al. (2020) proposed to compute the Nash equilibria of empirical games to cluster players into level sets of "strength". However, when the game has more than two players or is not zero-sum, the computation of the Nash equilibrium is in a class of problem complexity called PPAD-complete (Chen et al., 2009; Daskalakis et al., 2009) which is considered intractable. Motivated by this intractability result Omidshafiei et al. (2019); Rashid et al. (2021) proposed a tractable ranking technique (α-rank) theoretically grounded in the dynamical system theory.

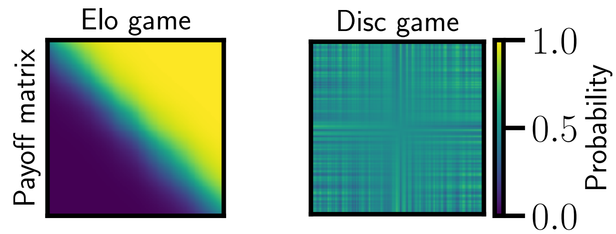


Figure 1: **Elo and Cyclic Disc Games.** Payoff matrices for an Elo (left) and a disc (right) game. For the Elo game (left) one can see that player 1 beats everyone: $\mathrm{P}_{1j} > 1/2$ for all $j \in [n]\backslash\{1\}$. Player 2 beats everyone except player 1, etc. Conversely, there is no total order for the disc game (right).

2 BACKGROUND

2.1 Symmetric Zero-Sum Games

In this section, we first define symmetric zero-sum games (Definition 2.1), fully transitive and cyclic games (Definition 2.2). Then we provide two examples of games, the Elo *game* (Example 1) and the disc game (Example 2).

Definition 2.1. *We define a symmetric zero-sum game through a probability matrix* $\mathbf{P} \in \mathbb{R}^{n\times n}$, *such that for all* $i, j \in [n]$, $\mathrm{P}_{ij} \in [0,1]$ *and* $\mathrm{P}_{ij} = 1 - \mathrm{P}_{ji}$. *We call a player an index* i *of this matrix.*

Using Definition 2.1 (i) $\mathrm{P}_{ij} > 0.5$ can be seen as i beats j (ii) $\mathrm{P}_{ij} < 0.5$ corresponds to i is beaten by j (iii) $\mathrm{P}_{ij} = 0.5$ is a tie. For many real-world games, one can have access to large databases of human game outcomes: the symmetric zero-sum game payoff matrix can be estimated empirically. Along with this paper, we will use empirical payoff and symmetric zero-sum games interchangeably. We argue that real-world zero-sum games lie on a spectrum between being fully transitive (without any cycle) and being fully cyclic (all the players belong to the same cycle).

Definition 2.2. (Fully Cyclic and Transitive). *The game* $\mathbf{P}$ *is said to be* fully transitive *if for all players* i, j, k, *if* $\mathrm{P}_{ij} > 0.5$ *and* $\mathrm{P}_{jk} > 0.5$ *then* $\mathrm{P}_{ik} > 0.5$. *The game* $\mathbf{P}$ *is said to be* fully cyclic *if there exists an ordering* γ *such that* $\mathrm{P}_{\gamma(1)\gamma(2)} > 0.5, \ldots, \mathrm{P}_{\gamma(n-1)\gamma(n)} > 0.5$ *and* $\mathrm{P}_{\gamma(n)\gamma(1)} > 0.5$.

The prototypical examples of fully transitive games are Elo games (Example 1, Figure 1 left): each player is assigned one score (the Elo), and the probability of the outcome between two players is an increasing function (a sigmoid) of the difference in the Elo scores.

Example 1. (Elo Game Balduzzi et al. 2019). *Let* $\mathbf{u} \in \mathbb{R}^n$, $\mathrm{P}_{ij} = \sigma(\mathrm{u}_i - \mathrm{u}_j)$, *for all* $i, j \in [n]$.

One can easily show that Elo games (Example 1) are fully transitive. However, the "reciprocate" is false in general: in Section 2.2 we provide an example (Example 3) of a fully transitive game for which the Elo score fails to correctly rank players. On the other side of the spectrum, the typical example of a fully cyclic game is the cyclic disc game (Example 2, Figure 1 right), where each player i is assigned two scores $(\mathrm{u}_i, \mathrm{v}_i)$.

Example 2. (Disc Game, Balduzzi et al. 2019). *Let* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, $\mathrm{P}_{ij} = \sigma(\mathrm{u}_i \mathrm{v}_j - \mathrm{v}_i \mathrm{u}_j)$, *for all* $i, j \in [n]$.

For example, if for all $i \in [n]$, $\mathrm{u}_i = 1$, the disc game is an Elo game (Example 1) and is transitive. One the other hand, if $\mathbf{u} = (\cos \frac{2\pi i}{n})_{i \in [n]}, \mathbf{v} = (\sin \frac{2\pi i}{n})_{i \in [n]}$, then the disc game is fully cyclic (Figure 1, right). The main contribution of Section 3.1 is to show that a disc game can be nothing else but fully transitive or fully cyclic (Proposition 1).

2.2 The Elo Score and its Limitations

In this section, we recall the definition of the stationary Elo score (Equation (1)). Then we recall one usual issue with the Elo score. Finally, we recall that the Elo score can fail on some transitive games (Example 3 and Figure 2).

Recalls on the Elo score For zero-sum symmetric empirical games, Elo (1978) proposed a rating system able to predict the probability of the outcome of a game between two agents. Given two agents i and j, with a respective Elo score of u_i and u_j, the probability of i beating j under the Elo model is $\mathbb{P}(i \text{ beats } j) = \sigma(\alpha(\mathrm{u}_i - \mathrm{u}_j))$, where $\alpha > 0$ is a scaling factor which brings the values of $\mathbf{u}$ in a range which is easier to grasp for humans (for simplicity α is set to 1). In a stationary regime, for an empirical payoff matrix $\mathbf{P}$, the[3] Elo score $\mathrm{u}_i^{\mathrm{Elo}}$ of each player i is defined as following (Balduzzi et al., 2018, Prop. 1), with bce being the binary cross entropy,

$$\mathbf{u}^{\mathrm{Elo}} = \underset{\mathbf{u}}{\mathrm{argmin}} \sum_{i,j} \mathrm{bce}(\mathrm{P}_{ij}, \sigma(\mathrm{u}_i - \mathrm{u}_j)) \ . \tag{1}$$

Issues with the Elo score A first issue with the Elo score is that it assumes that the modeled game is additive (in the logit space): $\mathrm{logit}(\mathrm{P}_{ij}) + \mathrm{logit}(\mathrm{P}_{jk}) = \mathrm{logit}(\mathrm{P}_{ik})$. Hence the modeled game should be transitive, which is not always the case for real-world games (such as rock-paper-scissor or StarCraft II, for instance). Another issue with the Elo score is the following: even if a game is transitive, the Elo score

[3]The Elo score is not unique and is defined up to a constant.

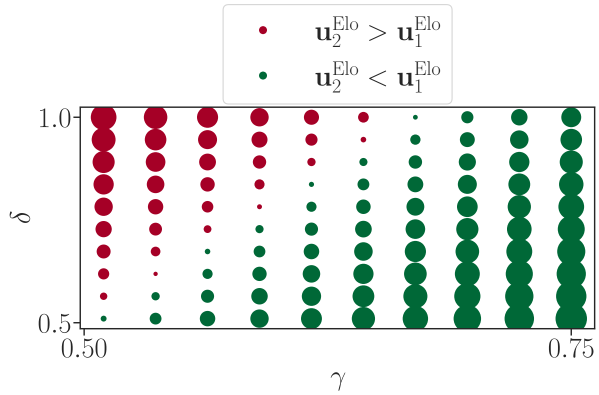


Figure 2: **Elo Score Fails to Rank Players for Some Transitive Games.** This figure displays the difference $\mathrm{u}_1^{\mathrm{Elo}} - \mathrm{u}_2^{\mathrm{Elo}}$ between the Elo scores of $\mathrm{u}_1^{\mathrm{Elo}}$ of player 1, and $\mathrm{u}_2^{\mathrm{Elo}}$ of player 2 (computed using the stationary Elo score in Equation (1)). The difference of the Elo scores is displayed for multiple probability matrices $\mathbf{P}^{(\gamma,\delta)}$ of the transitive game (Example 3). Red dots indicate that $\mathrm{u}_2^{\mathrm{Elo}} > \mathrm{u}_1^{\mathrm{Elo}}$ and green dots indicate $\mathrm{u}_1^{\mathrm{Elo}} > \mathrm{u}_2^{\mathrm{Elo}}$, the size of the dots is proportional to $|\mathrm{u}_1^{\mathrm{Elo}} - \mathrm{u}_2^{\mathrm{Elo}}|$.

can "fail" at ranking the players correctly. Indeed, there are some situations where the Elo score, a single scalar variable, is not expressive enough to predict the outcome of future confrontations.

Example 3. *Here we define a family of three-player transitive games for all* $\gamma, \delta \in (0.5, 1]$. *Contrary to Elo games (Example 1), outcome probabilities might be non-additive:*

$$\mathbf{P}^{(\gamma,\delta)} = \begin{pmatrix} 0.5 & \gamma & \gamma \\ 1-\gamma & 0.5 & \delta \\ 1-\gamma & 1-\delta & 0.5 \end{pmatrix} \ .$$

Example 3 describes the payoff matrix of a game that is transitive for $\gamma, \delta \in (0.5, 1]$, however when γ is close to 0.5 and δ is close to 1—i.e., the second player slightly loses against the first one and significantly wins against the third one—Elo score fails to assign scores which yield correct matchup predictions between players. Figure 2 displays the set of values γ, δ for which the Elo score fails (in red) and succeeds (in green) to correctly estimate the probability of winning between the first and the second players of the game $\mathbf{P}^{(\gamma,\delta)}$ (Example 3). Despite the game being transitive (player 1 beats player 2 and 3 and player 2 beats player 3), there exists a significant range of values γ, δ, for which the Elo score assigns a larger score to player 2 than player 1, and thus wrongly predicts the outcome of the confrontation.

Connection with Stochastically Transitive Models Bradley-Terry-Luce and Thurstone models (Tutz, 1986; Atkinson et al., 1998; Negahban et al., 2012; Shah and Wainwright, 2017) are a generalization of the Elo model. There are used to stochastically approximate transitive models of a specific type of what we called symmetric zero-sum games with a strong transitive component. Similar limitations with the Elo score that the ones we present in Example 3 were previously mentioned by Shah and Wainwright (2017, Figure 1a). Moreover, Shah and Wainwright (2017) showed that there exist some transitive matrices (called SST matrices) which are poorly approximated by one-dimensional parametric models. This result motivates our main contribution, which corresponds to a multi-dimensional parametric model approximating a matrix that corresponds to a symmetric zero-sum game.

Failures of the Elo score to correctly 'rank' players for a transitive game (Example 3 and Figure 2) call for models which can handle a larger class of real-world games, going beyond Elo games. In Section 3 we propose a disc decomposition, which can correctly rank players on Example 3 (see Figure 3).

3 PROPOSED APPROACH

We now present our main contribution, which is threefold:

- First in Section 3.1 we thoroughly study the disc game, $\mathrm{P}_{ij} = \sigma(\mathrm{u}_i \mathrm{v}_j - \mathrm{v}_i \mathrm{u}_j)$, $i, j \in [n]$ (Example 2). We show that depending on the values of $\mathbf{u}$ and $\mathbf{v}$, the disc game is either fully transitive or fully cyclic (Proposition 1).
- Then, in Section 3.2 we show that any empirical matrix from real-world zeros-sum games can be decomposed as a sum of disc games, with *at most one transitive disc game* (Theorem 2). This motivates our disc rating system which corresponds to the extraction of the empirical matrix's main component (disc game). If this main component corresponds to a transitive (resp. cyclic) disc game, then the original game is mainly transitive (resp. cyclic).
- Finally in Section 3.3 we provide the optimization details needed to compute the proposed scores (Algorithm 2).

Given a payoff matrix $\mathbf{P}$ (Definition 2.1), the following diagram summarizes the proposed approach

$$\mathbf{P} \xrightarrow{\text{logit}} \underbrace{\mathbf{A}}_{\text{skew-symmetric}} \xrightarrow[\text{decomposition (Thm. 1)}]{\text{Truncated normal}} \hat{\mathbf{A}} \xrightarrow{\text{sigmoid}} \hat{\mathbf{P}}.$$

3.1 Detailed Study of the Disc Game

Section 2.1 provides examples of $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ values for which the disc game (Example 2) is *fully transitive* or *fully cyclic*. In this section, we show that there are no other options: a disc game is either be *fully cyclic* or *fully transitive*, depending on the of the values of $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.

Proposition 1. (A Disc Game is Fully Cyclic or Fully Transitive). *Let* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ *and* $\mathrm{P}_{ij} = \sigma(\mathrm{u}_i \mathrm{v}_j - \mathrm{v}_i \mathrm{u}_j)$, $i, j \in [n]$ *be a disc game. Let* $U := \mathrm{hull}\{(\mathrm{u}_i, \mathrm{v}_i)\}$ *be the convex hull of the players. If* $(0,0)$ *is* not *in the border of* U,[4] *then the disc game is either fully transitive or fully cyclic. Precisely, the disc game is*

1. *fully cyclic if and only if* $(0,0) \in \mathrm{Int}(U)$,
2. *fully transitive if and only if* $(0,0) \notin \mathrm{Int}(U)$.

The main takeaway from this result is that the convex hull of the players U is the key object to determine if the disc game is transitive or cyclic. The dichotomy happens around the origin. The next proposition states that if a disc game is transitive, then $\mathbf{u}$ and $\mathbf{v}$ can be reparametrized.

Proposition 2. (Reparametrization). *Let* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ *and* $\mathrm{P}_{ij} := \sigma(\mathrm{u}_i \mathrm{v}_j - \mathrm{v}_i \mathrm{u}_j)$, $i, j \in [n]$ *be a* ***transitive*** *disc game. Then there exists a reparametrization of the disc game* $(\tilde{\mathrm{u}}_i, \tilde{\mathrm{v}}_i)$ *such that, for all* $i \in [n]$, $\tilde{\mathrm{v}}_i > 0$, *and*

$$\sigma(\mathrm{u}_i \mathrm{v}_j - \mathrm{v}_i \mathrm{u}_j) = \mathrm{P}_{ij} = \sigma(\tilde{\mathrm{u}}_i \tilde{\mathrm{v}}_j - \tilde{\mathrm{v}}_i \tilde{\mathrm{u}}_j) \ . \tag{2}$$

The proof of this result relies on Proposition 1 and the hyperplane separation theorem (Boyd and Vandenberghe, 2004, Example 2.20). Because the reparametrization $(\tilde{\mathbf{u}}_i, \tilde{\mathrm{v}}_i)$ yields the same payoffs, it corresponds to the same disc game. Proposition 2 means that, without any loss of generality, one can consider that $\mathrm{v}_i > 0$, $i \in [n]$ for any transitive disc games.

In the next section, we will see that any empirical payoff $\mathbf{P}$ (Definition 2.1) can be transformed into a skew-symmetric matrix $\mathbf{A}$ and decomposed as a sum of disc games. We will then use Proposition 1 on the main disc game to assess the transitivity/cyclicity of the original empirical game $\mathbf{A}$.

3.2 Disc Decomposition

First, we recall a standard result on the decomposition of skew-symmetric matrices (Theorem 1): any real skew-symmetric matrix can be decomposed as a sum of matrices of the form $\mathbf{u}\mathbf{v}^\top - \mathbf{v}\mathbf{u}^\top$, $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. Combining this result with Proposition 1, we finally show that zero-sum game payoff matrices have at most a single transitive disc game component (Theorem 2).

Theorem 1. (Normal Decomposition, Greub 1975). *Suppose* $\mathbf{A} \in \mathbb{R}^{n \times n}$ *is such that* $\mathrm{A} = -\mathrm{A}^\top$. *Then, with* $k = \lfloor n/2 \rfloor$, *there exists* $\lambda_1 \geq \ldots \geq \lambda_k$ *and* $(\mathbf{u}^{(l)}, \mathbf{v}^{(l)}) \in \mathbb{R}^n \times \mathbb{R}^n$, $1 \leq l \leq k$, *such that* $(\mathbf{u}^{(l)}, \mathbf{v}^{(l)})$ *is an orthogonal family and* $\mathbf{A} = \sum_{l=1}^{k} (\mathbf{u}^{(l)} \mathbf{v}^{(l)\top} - \mathbf{v}^{(l)} \mathbf{u}^{(l)\top})$.

A proof of Theorem 1 can be found in Francinou et al. (2008, Sec. 2.5). Theorem 1 is sometimes referred to as the Schur decomposition of skew-symmetric matrices (Balduzzi et al., 2018, Prop. 2).

In the context of symmetric zero-sum games, the first takeaway of Theorem 1 is that such a game is a sum of disc

[4] if it is the case we can, for instance, apply an infinitesimal perturbation on the points $(\mathrm{u}_i, \mathrm{v}_i)$.

Algorithm 1 Alternate Minimization

input : $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{us}, \mathbf{vs} \in \mathbb{R}^{n \times k}$, $l \in \mathbb{N}$
init : $\mathbf{u}, \mathbf{v} \neq 0_n$
for $k = 1, 2 \ldots,$ **do**
$\mathbf{u} \leftarrow \text{argmin}_{\mathbf{u}} \mathcal{L}(\mathbf{P}, \mathbf{A} + \mathbf{u}\mathbf{v}^\top - \mathbf{v}\mathbf{u}^\top) + \sum_{m=1}^{l-1} \frac{\langle \mathbf{u}, \mathbf{vs}_{:m} \rangle^2}{\|\mathbf{vs}_{:m}\|^2} + \sum_{m=1}^{l-1} \frac{\langle \mathbf{u}, \mathbf{us}_{:m} \rangle^2}{\|\mathbf{us}_{:m}\|^2}$
$\mathbf{v} \leftarrow \text{argmin}_{\mathbf{v}} \mathcal{L}(\mathbf{P}, \mathbf{A} + \mathbf{u}\mathbf{v}^\top - \mathbf{v}\mathbf{u}^\top) + \frac{\langle \mathbf{u}, \mathbf{v} \rangle^2}{\|\mathbf{u}\|^2} + \sum_{m=1}^{l-1} \frac{\langle \mathbf{v}, \mathbf{vs}_{:m} \rangle^2}{\|\mathbf{vs}_{:m}\|^2} + \sum_{m=1}^{l-1} \frac{\langle \mathbf{v}, \mathbf{us}_{:m} \rangle^2}{\|\mathbf{us}_{:m}\|^2}$
return $\mathbf{u}, \mathbf{v}$

games. The second consequence (Theorem 2) is less obvious but maybe even more important: among all these disc games *at most one* is transitive.

Theorem 2. (An Empirical Game has at most a Single Transitive Disc Game). *Let* $\mathbf{P}$ *be the payoff matrix of a symmetric zero-sum game, and let* $(\mathbf{u}^{(l)}, \mathbf{v}^{(l)}) \in \mathbb{R}^n \times \mathbb{R}^n$ *be the normal decomposition of the skew-symmetric matrix* $\text{logit}(\mathbf{P})$ *(Theorem 1), then there exists* at most *one pair* $(\mathbf{u}^{(l)}, \mathbf{v}^{(l)}) \in \mathbb{R}^n \times \mathbb{R}^n$ *such that the disc game defined by* $\mathrm{P}_{ij} = \sigma(\mathrm{u}_i^{(l)} \mathrm{v}_j^{(l)} - \mathrm{v}_i^{(l)} \mathrm{u}_j^{(l)})$*,* $i, j \in [n]$*, is transitive.*

Theorem 2 provides us with multiple insights. If one has access to an empirical game $\mathbf{P}$, and one can compute the normal decomposition (Theorem 2) of the skew-symmetric matrix $\text{logit}(\mathbf{P})$, then the largest value λ_1 and its associated vectors $(\mathbf{u}^{(1)}, \mathbf{v}^{(1)})$ encapsulate information on the largest component. If the disc game associated with the largest component is transitive, we will say that the considered empirical game based on $\mathbf{P}$ is transitive. Note that Theorem 2 can be generalized to any invertible functions f which transforms zero-sum game payoff matrices into skew-symmetric matrices, such as for instance $f : \mathbf{P} \mapsto 2\mathbf{P} - 1$.

Section 3.3 details main component computation when dealing with real data (*e.g.*, missing and inexact entries).

3.3 Computational Details

Now we provide the details to compute the proposed score. Consider an empirical game $(\mathrm{P}_{ij})_{\leq i,j \leq n}$. One can compute the decomposition of $\text{logit}(\mathbf{P})$ from Theorem 1 by solving the following optimization problem, with $\mathcal{L} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $\mathcal{L} : x, \hat{x} \mapsto \|\text{logit}(x) - \hat{x}\|^2$

$$\min_{\mathbf{u}^{(k)}, \mathbf{v}^{(l)} \in \mathbb{R}^n} f(\mathbf{u}, \mathbf{v}) \triangleq \mathcal{L}(\mathbf{P}, \textstyle\sum_{l=1}^{k} \mathbf{u}^{(l)} {\mathbf{v}^{(l)}}^\top - \mathbf{v}^{(l)} {\mathbf{u}^{(l)}}^\top)$$

$$\triangleq \sum_{i,j,l} \mathcal{L}(\mathrm{P}_{ij}, \mathrm{u}_i^{(l)} \mathrm{v}_j^{(l)} - \mathrm{v}_i^{(l)} \mathrm{u}_j^{(l)}) \ , \qquad (3)$$

$$\text{s. t.} \quad \langle \mathbf{u}^{(l)}, \mathbf{u}^{(m)} \rangle = \langle \mathbf{v}^{(l)}, \mathbf{v}^{(m)} \rangle = 0 \ , \qquad (4)$$

$$\langle \mathbf{u}^{(l)}, \mathbf{v}^{(l)} \rangle = 0 \, , \qquad 1 \leq l < m \leq k \ . \qquad (5)$$

The main challenge of this optimization problem is to maintain the orthogonality constraint between the components. We propose Algorithm 2 which sequentially finds the l^{th} main components $(\mathbf{u}^{(l)}, \mathbf{v}^{(l)})$ for $l \in [k]$. The l^{th} pair of components is found with Algorithm 1 which maintains the orthogonality constraints by using a penalty term (inspired by the algorithm of Gemp et al. (2021) which computes PCA for large-scale problems). Note that *in practice, empirical probability matrices can have* 0 *or* 1 *entries (for instance if a player* i *always loses again player* j*), for which* logit *is not defined.* That's why in the experiments (Section 4) we use the following loss function $\mathcal{L} : x, \hat{x} \mapsto \text{bce}(x, \sigma(\hat{x}))$ (where bce is the binary cross entropy).

Missing Entries Payoff matrices coming from real-world games played by humans usually contain missing entries: one does not have access to the matchups between all the players. For instance, on the `Lichess` website, there usually are few confrontations between low-ranked and high-ranked players. In other words, one only partially has access to P_{ij}, for (i, j) in a given set of pairs of players $\mathcal{D}^{\text{obs}} \subset [n] \times [n]$. Note that the optimization problem formulations based on Equation (3) can handle missing entries (by summing only on the available entries) as in Candès and Recht (2008); Candès and Plan (2009). With $k = 1$ pair of components, instead of the problem defined in Equation (18), the proposed disc decomposition with a partial set of observation $\mathcal{D}^{\text{obs}}$ reads:

$$\underset{\mathbf{u}, \mathbf{v}, \text{ s.t. } \mathbf{u}^\top \mathbf{v} = 0}{\text{argmin}} \sum_{(i,j) \in \mathcal{D}^{\text{obs}}} \mathcal{L}(\mathrm{P}_{ij}, \mathrm{u}_i \mathrm{v}_j - \mathrm{v}_i \mathrm{u}_j) \ .$$

Interpretation of the Disc Decomposition in the Case $k = 1$ Let $(\mathbf{u}^{\text{Disc}}, \mathbf{v}^{\text{Disc}})$ be the first pair ou component of the disc decomposition (Equation (3)). The probability of player i beating player j is given by

$$\mathbb{P}(i \text{ beats } j) = \sigma(\mathrm{u}_i^{\text{Disc}} \mathrm{v}_j^{\text{Disc}} - \mathrm{v}_i^{\text{Disc}} \mathrm{u}_j^{\text{Disc}}) \ . \qquad (6)$$

If it occurs that for all $j \in [n]$, $\mathrm{v}_j^{\text{Disc}} > 0$, which corresponds to a transitive game (see Theorem 2 and Proposition 2), then the outcome probability can be written as $\sigma(\mathrm{u}_i^{\text{Disc}} \mathrm{v}_j^{\text{Disc}} - \mathrm{v}_i^{\text{Disc}} \mathrm{u}_j^{\text{Disc}}) = \sigma(\mathrm{v}_i^{\text{Disc}} \mathrm{v}_j^{\text{Disc}} (\mathrm{u}_i^{\text{Disc}} / \mathrm{v}_i^{\text{Disc}} - \mathrm{u}_j^{\text{Disc}} / \mathrm{v}_j^{\text{Disc}}))$. For a player i, the ratio $\tilde{\mathrm{u}}_i^{\text{Disc}} \triangleq \mathrm{u}_i^{\text{Disc}} / \mathrm{v}_i^{\text{Disc}}$ can be interpreted as its *strength*, and v_i as its *consistency* to beat lower-rated players (and be beaten by higher-rated players):

$$\mathbb{P}(i \text{ beats } j) = \sigma(\mathrm{v}_i^{\text{Disc}} \mathrm{v}_j^{\text{Disc}} (\tilde{\mathrm{u}}_i^{\text{Disc}} - \tilde{\mathrm{u}}_j^{\text{Disc}})) \ . \qquad (7)$$

If $\tilde{\mathrm{u}}_i^{\text{Disc}} > \tilde{\mathrm{u}}_j^{\text{Disc}}$ then i beats j. The larger $\mathrm{v}_i^{\text{Disc}}$, the larger the probability to win against a lower-rated player ($\tilde{\mathrm{u}}_i^{\text{Disc}} > \tilde{\mathrm{u}}_j^{\text{Disc}}$), but the larger the probability to lose against a higher rated player ($\tilde{\mathrm{u}}_i^{\text{Disc}} < \tilde{\mathrm{u}}_j^{\text{Disc}}$). This consistency score can be seen as a way to correct the Elo score which implicitly assumes that the strength of each player is linearly comparable in the logit space. For instance, one implicit bias of the Elo score is that if i beats j with probability P_{ij} and j beats k with probability P_{jk} then i beats k with

Algorithm 2 Compute Disc Decomposition

input : $k \in \mathbb{N}$ (# of pairs of components)
init : $\mathbf{A} = 0_{n\times n}$,
$\mathbf{us} = 0_{n\times k}$, $\mathbf{vs} = 0_{n\times k}$
for $l = 1, 2, \ldots, k$ **do**
  $\mathbf{u}^{(l)}, \mathbf{v}^{(l)} =$ Algorithm 1$(\mathbf{A}, \mathbf{us}, \mathbf{vs}, l)$
  $\mathbf{A} \leftarrow \mathbf{A} + \mathbf{u}^{(l)}\mathbf{v}^{(l)\top} - \mathbf{v}^{(l)}\mathbf{u}^{(l)\top}$
  $\mathbf{us}_{:l}, \mathbf{vs}_{:l} = \mathbf{u}^{(l)}, \mathbf{v}^{(l)}$
return $(\mathbf{u}^{(l)}, \mathbf{v}^{(l)})_{1\leq l\leq k}$

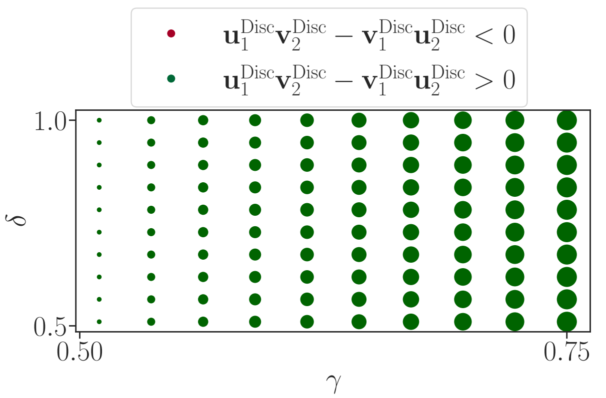


Figure 3: **Disc Decomposition Manages to Rank Players for some Transitive Games.** Let $\mathbf{u}^{\text{Disc}}, \mathbf{v}^{\text{Disc}}$ be the first pair of components computed using the stationary disc decomposition ($k = 1$ in Equation (3)). The logit of the probability of winning, $\mathrm{u}_1^{\text{Disc}}\mathrm{v}_2^{\text{Disc}} - \mathrm{v}_1^{\text{Disc}}\mathrm{u}_2^{\text{Disc}}$, is displayed for multiples probability matrices of the transitive game (Example 3). Red dots indicate that $\mathrm{u}_1^{\text{Disc}}\mathrm{v}_2^{\text{Disc}} - \mathrm{v}_1^{\text{Disc}}\mathrm{u}_2^{\text{Disc}} < 0$, and green dots $\mathrm{u}_1^{\text{Disc}}\mathrm{v}_2^{\text{Disc}} - \mathrm{v}_1^{\text{Disc}}\mathrm{u}_2^{\text{Disc}} > 0$. The size of the dots is proportional to $|\mathrm{u}_1^{\text{Disc}}\mathrm{v}_2^{\text{Disc}} - \mathrm{v}_1^{\text{Disc}}\mathrm{u}_2^{\text{Disc}}|$. Contrary to the Elo score (Figure 2), the proposed disc decomposition can correctly rank players 1 and 2.

probability $\operatorname{logit}(\mathrm{P}_{ik}) = \operatorname{logit}(\mathrm{P}_{ij}) + \operatorname{logit}(\mathrm{P}_{jk})$. As illustrated in Example 3 if such a property of additivity in the logit space is not occurring in the data, the Elo score may have trouble even predicting the right ranking for the players (Figure 2). Conversely, when including a notion of *consistency* in our score we manage to correctly predict the ranking between the players in Example 3 (see Figure 3.)

Online Update The main focus of this work is to study the intrinsic characteristics of the game at a given time t. This is opposed to the line of work that considers the games' sequential aspect, quantifying the Elo score uncertainty with a Bayesian model (TrueSkill, Herbrich et al. 2006), or relying on the more recent games (Sismanis, 2010). However, an update rule can still be derived and interpreted for the proposed disc decomposition. The stationary version of the proposed disc decomposition is written in Equation (3), the online version (Jabin and Junca, 2015) of the disc decomposition writes, with a step size $\eta > 0$, $\mathrm{S}^t_{i,j}$ the result of the confrontation between players i and j, and $\hat{\mathrm{P}}^t_{ij} = \sigma(\mathrm{v}^t_i \mathrm{v}^t_j (\tilde{\mathrm{u}}^t_i - \tilde{\mathrm{u}}^t_j))$:

$$\tilde{\mathrm{u}}_i^{t+1} = \tilde{\mathrm{u}}_i^t + \eta(\mathrm{S}^t_{i,j} - \hat{\mathrm{P}}^t_{ij})\mathrm{v}^t_i \mathrm{v}^t_j \ , \tag{8}$$
$$\mathrm{v}_i^{t+1} = \mathrm{v}_i^t + \eta(\mathrm{S}^t_{i,j} - \hat{\mathrm{P}}^t_{ij})\mathrm{v}^t_j (\tilde{\mathrm{u}}^t_i - \tilde{\mathrm{u}}^t_j) \ . \tag{9}$$

As for the usual Elo score, the increase in the "strength" $\tilde{\mathrm{u}}_i$ after a confrontation is proportional to $\mathrm{S}^t_{i,j} - \hat{\mathrm{P}}^t_{ij}$. The "consistency" v_i increases when beating lower-rated players and decreases when losing against lower-rated players. Informally, this quantity encompasses how much one should trust the current Elo score and scales up or down the "strength" update.

4 EXPERIMENTS

In this section, we perform experiments [5] on real payoff matrices from Czarnecki et al. (2020) and real chess and StarCraft II data. For each type of data, we propose two kinds of experiments. First, we qualitatively compare and interpret the first pair of components the proposed disc decomposition ($k = 1$ in Equation (3)), directly in the probability space P_{ij} ($\mathcal{L}(x, \hat{x}) \mapsto \|x - \hat{x} - \mathbf{1}/2\|^2$ in Equation (3)) or in the logit space $\operatorname{logit}\mathrm{P}_{ij}$ ($\mathcal{L} : x, \hat{x} \mapsto \operatorname{bce}(x, \sigma(\hat{x}))$ in Equation (3)). When possible, we reparametrized the disc decomposition $\mathbf{u}$ and $\mathbf{v}$ as in Proposition 2. Then some entries of the payoff matrix are hidden (as well as the symmetric entries), and each model is trained and evaluated to predict the missing entries (20% of the dataset). For this second experiment, we compare the following models:

- The usual Elo score (Elo, 1978), which relies on the assumption that the game is purely transitive.
- A variant of the Elo score with a quadratic loss (as in Elo ++, Sismanis 2010).
- Balduzzi et al. (2018) which has $2k + 1$ parameters ($k \in \{1, 2, 3\}$), where the payoff matrix is decomposed as $\operatorname{grad}(\mathbf{A}) + \operatorname{rot}(\mathbf{A})$, and then $\operatorname{rot}(\mathbf{A})$ is approximated by the Schur decomposition.
- Balduzzi et al. (2019), where the payoff matrix is directly approximated by the Schur decomposition in the probability space (*i.e.*, $2\mathbf{P} - 1$), which has $2k$ parameters.
- The proposed disc decomposition (Algorithm 2) which has $2k$ parameters ($k \in \{1, 2, 3\}$ pairs of components).

One can find the mathematical details of each model in Appendix A.2. Note that among the compared methods for matchups probability estimation, only Elo, Elo++, and the proposed disc decomposition (in the transitive case) are ranking methods.

4.1 Data from Czarnecki et al. (2020)

Comments on Figure 4 First, we investigate the visualization and the performance of some payoff matrices

[5] Code can be found at https://github.com/QB3/discrating.

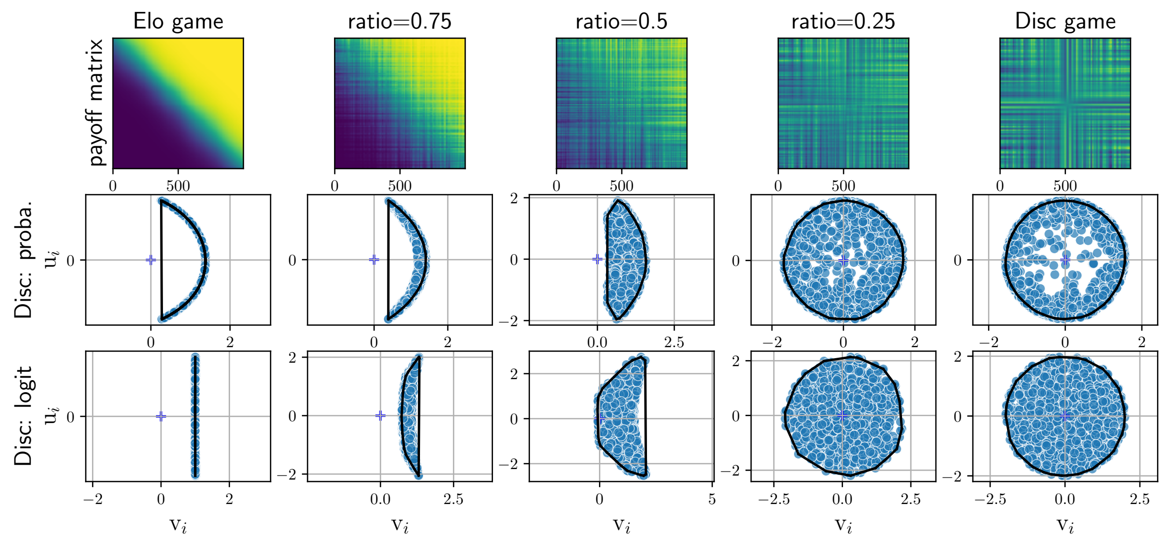


Figure 4: **From Elo to Disc Game.** Payoff matrices (top) and visualization using Algorithm 1 in the probability space (middle) and logit space (bottom), for multiple games with payoff $\mathrm{logit}(\mathbf{P}) = \mathrm{ratio} \cdot \mathrm{logit}(\mathbf{P}^{\mathrm{Elo}}) + (1 - \mathrm{ratio}) \cdot \mathrm{logit}(\mathbf{P}^{\mathrm{Disc}})$.

Table 1: Prediction performance (MSE) on unseen data interpolating from a pure Elo game to a disc game: $\mathrm{logit}(\mathbf{P}) = \mathrm{ratio} \cdot \mathrm{logit}(\mathbf{P}^{\mathrm{Elo}}) + (1 - \mathrm{ratio}) \cdot \mathrm{logit}(\mathbf{P}^{\mathrm{Disc}})$.

		Elo game		ratio = 0.75		ratio = 0.5		ratio = 0.25		Disc game	
Model	# Param.	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
Elo	n	1.6e−10	1.6e−10	1.0e−2	1.0e−2	3.6e−2	3.6e−2	6.3e−2	6.3e−2	7.9e−2	7.9e−2
Elo ++	n	2.1e−11	**2.1e−11**	1.0e−2	1.0e−2	3.6e−2	3.6e−2	6.3e−2	6.3e−2	7.9e−2	7.9e−2
Disc decomposition (ours, k=1)	$2n$	2.1e−10	2.2e−10	9.7e−3	9.8e−3	3.4e−2	3.4e−2	1.2e−2	1.2e−2	2.5e−6	2.6e−6
Balduzzi et al. (2019)	$2n$	2.7e−3	2.7e−3	1.0e−2	1.0e−2	3.4e−2	3.5e−2	1.4e−2	1.4e−2	4.5e−3	4.6e−3
m-Elo (Balduzzi et al. 2018, k=1)	$3n$	2.7e−3	2.7e−3	1.0e−2	1.0e−2	3.4e−2	3.5e−2	1.4e−2	1.4e−2	4.5e−3	4.6e−3
Balduzzi et al. (2019)	$4n$	1.3e−4	1.3e−4	3.9e−3	4.0e−3	5.8e−3	6.0e−3	4.5e−3	4.6e−3	8.5e−4	8.8e−4
Disc decomposition (ours, k=2)	$4n$	1.6e−10	1.7e−10	6.5e−4	**6.8e−4**	2.4e−3	**2.5e−3**	5.5e−4	**5.6e−4**	4.5e−7	**4.6e−7**

from Czarnecki et al. (2020). Figure 4 (and Figure 9 in Appendix B) display the payoff matrices (top row), the representation of Balduzzi et al. (2019) (middle row), and the proposed disc decomposition (bottom row) for games taken from Czarnecki et al. (2020). In these representations, one point represents one player: each player i is summarized by two scores $(\mathrm{v}_i, \mathrm{u}_i)$, which correspond to the coordinate of each point. Figure 4 displays the representations for multiple payoff matrices: from a pure Elo game (Example 1), to a pure cyclic disc game Example 2, through an average of the payoff matrices in the log space: $\mathrm{logit}(\mathbf{P}) = \mathrm{ratio} \cdot \mathrm{logit}(\mathbf{P}^{\mathrm{Elo}}) + (1 - \mathrm{ratio}) \cdot \mathrm{logit}(\mathbf{P}^{\mathrm{Disc}})$. For a pure Elo game (top left), one can see that the proposed model recovers perfectly a transitive game (bottom left): the model recovers $\mathrm{v}_i = 1$ for all the players i. On the other side of the spectrum, for a disc game (top right), the proposed method can find a disc game representation.

Comments on Table 1 It shows the prediction performances (MSE on unseen data) for each of these payoff matrices. Whereas the Elo score can predict almost perfectly the Elo game, the Elo score fails as soon as the Elo game assumption is violated. Conversely, the proposed disc decomposition can correctly predict the Elo game and the disc game. The proposed model can better predict future outcomes, from the Elo game to the disc game. For the Elo game, Elo, Elo ++ and our method yield similar performances ($\sim 10^{-10}$). For such small orders of magnitude, we believe that the difference between the methods is mostly due to numerical errors.

Figure 9 (in Appendix B) shows the representations for multiple real games played by computers (machine learning algorithms) from Czarnecki et al. (2020), including games considered transitive (Go and AlphaStar) and games considered cyclic (Blotto and Kuhn-Poker). One can see that the convex hull of the proposed representation does not contain 0 for games considered as mostly transitive: this validates Proposition 1.

Comments on Figures 5 and 6 Figure 5 shows the mean squared error (MSE) between the estimated probabilities and the empirical ones. More precisely, for each compared method, it shows the MSE of the Elo score divided by the MSE of the method, hence the larger the better. The predictions from Figure 5 show that the proposed representation

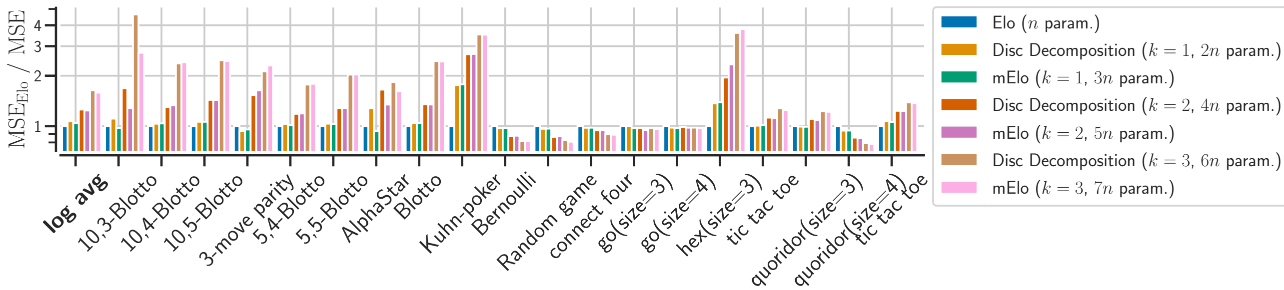


Figure 5: **Prediction Performances (the Higher, the Better) on Unseen Data.** The prediction performances on unseen data of the Elo (which has one parameter), the mElo (Balduzzi et al. 2018, which has $(2k+1)n$ parameters), and the proposed disc decomposition (Algorithm 2, which has $2kn$ parameters) are compared on a wide range of payoff matrices from Czarnecki et al. (2020). With fewer parameters, the proposed disc decomposition yields better mean squared error than Balduzzi et al. (2018).

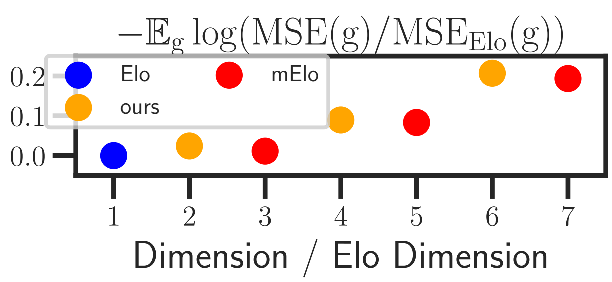


Figure 6: **Normalized (Log)-Average Prediction Performances on all the Datasets from Czarnecki et al. (2020) (the Higher, the Better).** Using fewer parameters, the proposed disc decomposition yields better prediction performances for unseen matchups.

yields better predictions for the outcome probabilities. The improvements of the proposed disc decomposition are more significant for games considered cyclic, such as Blottos or Kuhn-Poker. 'Bernoulli' and 'Random Game' are generated at random. Since there is only noise in these games, the larger the number of parameters, the larger the overfitting (as observed in Figure 5). We currently do not have explanations for why a larger number of parameters does not yield better predictions on 'Connect Four' and 'quoridor(size= 4)' games. In Figure 6, the averaged MSE across all the games from Czarnecki et al. (2020) shows that, with fewer parameters, the proposed disc decomposition achieves better performances than (Balduzzi et al., 2018).

4.2 Human Data: Lichess and StarCraft (Figure 7)

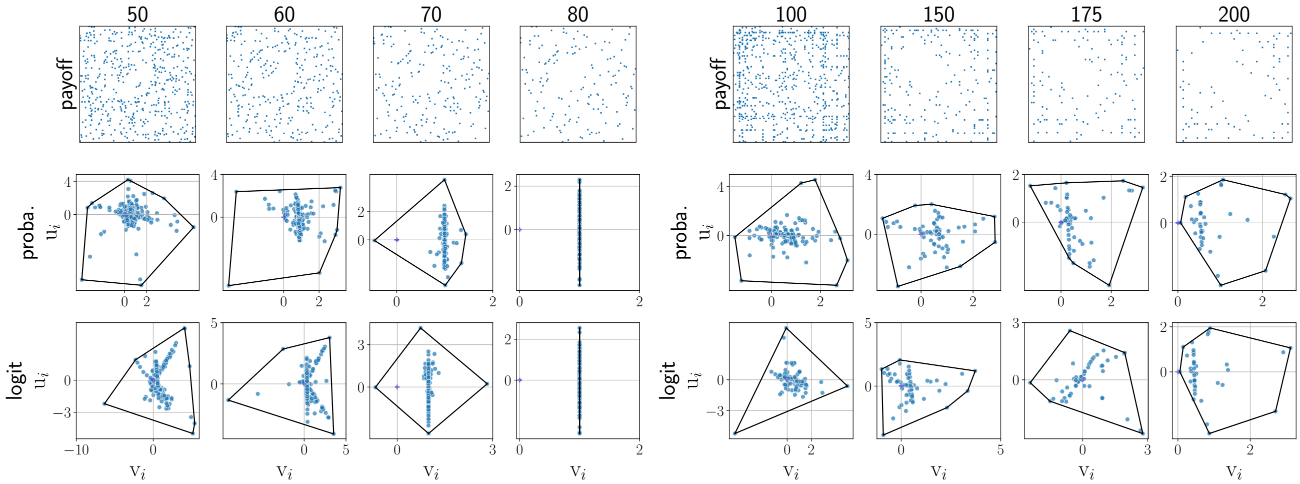
For the game of chess, we used the Lichess elite database[6]. Lichess is an open-source platform allowing one to play chess online against other players, and the Lichess elite database consists of "all (standard) games from Lichess filtered to only keep games by players rated 2400+ against players rated 2200+". For our experiments, we used the data from August 2019 to May 2020, which contains more than 4.7 million games between more than 40 000 players. For StarCraft, we used the aligulac data [7] from tournament with more than 1.7 millions of games, between 20 000 players. In StarCraft, each confrontation consists of multiple "games". We aggregated the games independently of the "race" used by the players. In both cases, the data we are using is public, anonymized, and only concerns non-sensitive information (matchup results). Elite players usually play online against others more often than amateur ones. Hence the number of matchups is larger for elite pairs of players, which yields a better estimation of the payoff matrix. In other words, it makes sense to restrict ourselves to "elite" players.

Influence of the Number of Matchups Real-world data has a lot of missing entries; hence the matchups payoff matrices (see Figure 7a) are incomplete. In Figure 7 a blue square indicates a confrontation between the players of the corresponding row and column. To decrease the noise in the estimation of the outcome probability, we only kept the probabilities coming from a large number of matchups between pairs of players. For chess (Figure 7a) and StarCraft (Figure 7b), we plotted the obtained representations as a function of this number of matchups.

Comments on Figure 7 It displays the proposed disc decomposition computed with Algorithm 1 in the probability space (middle) and logit space (bottom), for Lichess (Figure 7a) and StarCraft (Figure 7b). One can see that once the probability matrix is accurate enough (Figure 7a, right, more than 80 confrontations) the proposed method recovers a fully transitive game ($v_i > 0$). The proposed representation also yields a strong transitive component for 60 and 70 confrontations. For StarCraft, even by taking only the entries on the probability matrix with more than

[6] https://lichess.org/team/lichess-elite-database

[7] http://aligulac.com/



(a) **Real Lichess Data.** (b) **Real StarCraft II Data.**

Figure 7: Proposed disc decomposition as a function of the number of matchups for chess and StarCraft II.

200 confrontations, one does not recover a representation as transitive as for chess. One interesting conclusion from this experiment is that our score naturally learns from human data that chess and StarCraft are transitive games. To conclude, as expected from expert knowledge, chess seems to be more transitive than StarCraft (on the studied datasets).

5 CONCLUSION

In this work, we first showed some limitations of the Elo score for transitive and cyclic games. Then we studied in detail the disc game: we showed that depending on $(\mathbf{u}, \mathbf{v})$, the game is either cyclic or transitive. Based on this example, we proposed disc decomposition, which can recover the usual Elo score if the game is an Elo game and extend the Elo score to transitive non-Elo games and cyclic games. The theoretical results were extensively validated on real data.

Limitation The main weakness of this work is that it does not take into account a potential structure in the game. It is a strength and a weakness. On the one hand, it is a very general method that only requires matchup results between players. On the other hand, a lot of information could be leveraged from the specificity of the game player, e.g., its (a priori) transitive aspect (as in Balduzzi et al. 2019) or its sequential aspect (as in Herbrich et al. 2006 or Sismanis 2010). A second issue is that the proposed approach is less straightforward to interpret than the Elo. While the Elo assigns only one scalar per player, which can easily yield a ranking interpretation, the proposed scores can be harder to interpret if the game is not transitive.

Societal Impact Our work is primarily methodological: we do not see potential negative societal impacts.

Acknoledgments QB would like to thank Samsung Electronics Co., Ldt. for funding this research. GG is supported by an IVADO grant.

References

Donald R Atkinson, Bruce E Wampold, Susana M Lowe, Linda Matthews, and Hyun-Nie Ahn. Asian american preferences for counselor characteristics: Application of the Bradley-Terry-Luce model to paired comparison data. *The Counseling Psychologist*, 26(1):101–123, 1998. (Cited on page 3)

David Balduzzi, Karl Tuyls, Julien Perolat, and Thore Graepel. Re-evaluating evaluation. *Advances in Neural Information Processing Systems*, 31, 2018. (Cited on pages 1, 2, 3, 4, 6, 7, 8, and 13)

David Balduzzi, Marta Garnelo, Yoram Bachrach, Wojciech Czarnecki, Julien Perolat, Max Jaderberg, and Thore Graepel. Open-ended learning in symmetric zero-sum games. In *International Conference on Machine Learning*, pages 434–443. PMLR, 2019. (Cited on pages 1, 2, 3, 6, 7, 9, and 13)

Trapit Bansal, Jakub Pachocki, Szymon Sidor, Ilya Sutskever, and Igor Mordatch. Emergent complexity via multi-agent competition. *ICLR*, 2018. (Cited on page 1)

Christopher Berner, Greg Brockman, Brooke Chan, Vicki Cheung, Przemysław Dębiak, Christy Dennison, David Farhi, Quirin Fischer, Shariq Hashme, Chris Hesse, et al. Dota 2 with large scale deep reinforcement learning. *arXiv preprint arXiv:1912.06680*, 2019. (Cited on page 1)

Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004. (Cited on pages 16 and 4)

Emmanuel J. Candès and Yaniv Plan. Matrix completion with noise. 2009. (Cited on page 5)

Emmanuel J. Candès and Benjamin Recht. Exact matrix completion via convex optimization. 2008. (Cited on page 5)

Adrian Rivera Cardoso, Jacob Abernethy, He Wang, and Huan Xu. Competing against equilibria in zero-sum games with evolving payoffs. *ICML*, 2019. (Cited on page 1)

Sourav Chatterjee. Matrix estimation by universal singular value thresholding. 2015. (Cited on page 2)

Xi Chen, Xiaotie Deng, and Shang-Hua Teng. Settling the complexity of computing two-player nash equilibria. *Journal of the ACM (JACM)*, 2009. (Cited on page 2)

Wojciech M. Czarnecki, Gauthier Gidel, Brendan Tracey, Karl Tuyls, Shayegan Omidshafiei, David Balduzzi, and Max Jaderberg. Real world games look like spinning tops. *Advances in Neural Information Processing Systems*, 2020. (Cited on pages 2, 6, 7, 8, 12, 13, 14, and 15)

Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. The complexity of computing a nash equilibrium. *SIAM Journal on Computing*, 2009. (Cited on page 2)

Arpad Elo. *The USCF rating system*. 1961. (Cited on page 1)

Arpad Elo. The rating of Chess players, past and present (Arco, New York). 1978. (Cited on pages 3, 6, and 12)

FIFA. FIFA elo rating. https://web.archive.org/web/20180612141237/https://resources.fifa.com/image/upload/revision-of-the-fifa-coca-cola-world-ranking.pdf?cloudid=jgxjkdrj1jfwyunjbkha. (Cited on page 1)

Serge Francinou, Hervé Gianella, and Serge Nicolas. *Exercices de mathématiques oraux x-ens algèbre*, volume 3. CASSINI, 2008. (Cited on pages 16 and 4)

Ian Gemp, Brian McWilliams, Claire Vernade, and Thore Graepel. Eigengame: PCA as a nash equilibrium. *ICLR*, 19, 2021. (Cited on page 5)

Mark E Glickman. Dynamic paired comparison models with stochastic variances. *Journal of Applied Statistics*, 28 (6):673–689, 2001. (Cited on page 1)

Werner H. Greub. *Linear algebra*, volume 23. Springer Science & Business Media, 4th edition, 1975. (Cited on pages 16 and 4)

Ralf Herbrich, Tom Minka, and Thore Graepel. Trueskill™: a bayesian skill rating system. *Advances in neural information processing systems*, 19, 2006. (Cited on pages 1, 6, and 9)

Matthew Horrigan. Playing for the legend in the age of empires II online community. *Press Start*, 8(1):20–40, 2022. (Cited on page 1)

Pierre-Emmanuel Jabin and Stéphane Junca. A continuous model for ratings. *SIAM Journal on Applied Mathematics*, 75(2):420–442, 2015. (Cited on page 6)

Max Jaderberg, Wojciech M. Czarnecki, Iain Dunning, Luke Marris, Guy Lever, Antonio Garcia Castaneda, Charles Beattie, Neil C. Rabinowitz, Ari S Morcos, Avraham Ruderman, et al. Human-level performance in 3d multiplayer games with population-based reinforcement learning. *Science*, 364(6443):859–865, 2019. (Cited on page 1)

Xiaoye Jiang, Lek-Heng Lim, Yuan Yao, and Yinyu Ye. Statistical ranking and combinatorial hodge theory. *Mathematical Programming*, 127(1):203–244, 2011. (Cited on page 2)

Marc Lanctot, Vinicius Zambaldi, Audrunas Gruslys, Angeliki Lazaridou, Karl Tuyls, Julien Pérolat, David Silver, and Thore Graepel. A unified game-theoretic approach to multiagent reinforcement learning. *Advances in neural information processing systems*, 30, 2017. (Cited on page 1)

Dong C. Liu and Jorge Nocedal. On the limited memory bfgs method for large scale optimization. *Mathematical programming*, 45(1):503–528, 1989. (Cited on pages 12 and 13)

Siqi Liu, Guy Lever, Zhe Wang, Josh Merel, SM Eslami, Daniel Hennes, Wojciech M. Czarnecki, Yuval Tassa, Shayegan Omidshafiei, Abbas Abdolmaleki, et al. From motor control to team play in simulated humanoid football. *arXiv preprint arXiv:2105.12196*, 2021. (Cited on page 1)

Tom Minka, Ryan Cleven, and Yordan Zaykov. Trueskill 2: An improved bayesian skill rating system. *Technical Report*, 2018. (Cited on page 1)

NBA. NBA Elo rating. https://fivethirtyeight.com/features/how-we-calculate-nba-elo-ratings/. (Cited on page 1)

Sahand Negahban, Sewoong Oh, and Devavrat Shah. Iterative ranking from pair-wise comparisons. *Advances in Neural Information Processing Systems*, 25, 2012. (Cited on page 3)

Shayegan Omidshafiei, Christos Papadimitriou, Georgios Piliouras, Karl Tuyls, Mark Rowland, Jean-Baptiste Lespiau, Wojciech M. Czarnecki, Marc Lanctot, Julien Perolat, and Remi Munos. α-rank: Multi-agent evaluation by evolution. *Scientific reports*, 9(1):1–29, 2019. (Cited on page 2)

Tabish Rashid, Cheng Zhang, and Kamil Ciosek. Estimating α-rank by maximizing information gain. *arXiv preprint arXiv:2101.09178*, 2021. (Cited on page 2)

Nihar B Shah and Martin J Wainwright. Simple, robust and optimal ranking from pairwise comparisons. *The Journal of Machine Learning Research*, 18(1):7246–7283, 2017. (Cited on pages 2, 3, and 4)

David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, et al. A general reinforcement learning algorithm that masters Chess, Shogi, and Go through self-play. *Science*, 362 (6419):1140–1144, 2018. (Cited on page 1)

Yannis Sismanis. How I won the" chess ratings-elo vs the rest of the world" competition. *arXiv preprint arXiv:1012.4571*, 2010. (Cited on pages 1, 6, 9, 12, and 13)

Gerhard Tutz. Bradley-Terry-Luce models with an ordered response. *Journal of mathematical psychology*, 30(3):306–316, 1986. (Cited on page 3)

Oriol Vinyals, Timo Ewalds, Sergey Bartunov, Petko Georgiev, Alexander Sasha Vezhnevets, Michelle Yeo, Alireza Makhzani, Heinrich Küttler, John Agapiou, Julian Schrittwieser, et al. Starcraft II: A new challenge for reinforcement learning. *arXiv preprint arXiv:1708.04782*, 2017. (Cited on page 1)

Oriol Vinyals, Igor Babuschkin, Wojciech M. Czarnecki, Michaël Mathieu, Andrew Dudzik, Junyoung Chung, David H. Choi, Richard Powell, Timo Ewalds, Petko Georgiev, et al. Grandmaster level in starcraft ii using multi-agent reinforcement learning. *Nature*, 2019. (Cited on page 1)

Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, et al. Scipy 1.0: fundamental algorithms for scientific computing in python. *Nature methods*, 17(3):261–272, 2020. (Cited on pages 12 and 13)

William E Walsh, David C Parkes, and Rajarshi Das. Choosing samples to compute heuristic-strategy nash equilibrium. In *International Workshop on Agent-Mediated Electronic Commerce*, pages 109–123. Springer, 2003. (Cited on page 1)

Michael P Wellman. Methods for empirical game-theoretic analysis. In *AAAI*, pages 1552–1556, 2006. (Cited on page 1)

Xue Yan, Yali Du, Binxin Ru, Jun Wang, Haifeng Zhang, and Xu Chen. Learning to identify top Elo ratings: A dueling bandits approach. *arXiv preprint arXiv:2201.04480*, 2022. (Cited on page 1)

The appendix is organized as follows: Appendix A provides computational details on Algorithms 1 and 2. Appendix B provides additional experiments on data from Czarnecki et al. (2020) and StarCraft data. Finally, Appendix C contains the proofs of Propositions 1 and 2, Theorems 1 and 2.

A ADDITIONAL COMPUTATIONAL DETAILS

A.1 Computation Details

Gradient Computation In order to compute the proposed representation one needs to solve the following optimization problem ($k = 1$ in Equation (3))

$$\underset{\mathbf{u},\mathbf{v},\text{ s.t. }\mathbf{u}^\top\mathbf{v}=0}{\operatorname{argmin}} \; f(\mathbf{u},\mathbf{v}) \triangleq \sum_{i=1}^{n}\sum_{j=1}^{n} \mathcal{L}(\mathrm{P}_{ij}, \mathrm{u}_i \mathrm{v}_j - \mathrm{v}_i \mathrm{u}_j) \ . \tag{10}$$

Hence one needs to compute the gradient of f with respect to $\mathbf{u}$ and $\mathbf{v}$. Bellow is provided the formula for these gradients

$$\nabla_{\mathbf{u}_k} f(\mathbf{u},\mathbf{v}) = \sum_{i,j} \mathrm{v}_j \mathcal{L}'(\mathrm{P}_{ij}, \mathrm{u}_i\mathrm{v}_j - \mathrm{v}_i\mathrm{u}_j)\mathbb{1}_{i=k} - \mathrm{v}_i \mathcal{L}'(\mathrm{P}_{ij}, \mathrm{u}_i\mathrm{v}_j - \mathrm{v}_i\mathrm{u}_j)\mathbb{1}_{j=k} \ , \tag{11}$$

$$= \sum_{j} \mathrm{v}_j \mathcal{L}'(\mathbf{P}_{k,j}, \mathbf{u}_k\mathrm{v}_j - \mathbf{v}_k\mathrm{u}_j) - \sum_{i} \mathrm{v}_i \mathcal{L}'(\mathbf{P}_{i,k}, \mathrm{u}_i\mathbf{v}_k - \mathrm{v}_i\mathbf{u}_k) \ , \tag{12}$$

$$\nabla_{\mathbf{u}} f(\mathbf{u},\mathbf{v}) = \mathcal{L}'(\mathbf{P}, \mathbf{u}\mathbf{v}^\top - \mathbf{v}\mathbf{u}^\top)\mathbf{v} - \mathcal{L}'(\mathbf{P}, \mathbf{u}\mathbf{v}^\top - \mathbf{v}\mathbf{u}^\top)^\top\mathbf{v} \ . \tag{13}$$

With similar derivations, one can obtain

$$\nabla_{\mathbf{v}} f(\mathbf{u},\mathbf{v}) = \mathcal{L}'(\mathbf{P}, \mathbf{u}\mathbf{v}^\top - \mathbf{v}\mathbf{u}^\top)^\top\mathbf{u} - \mathcal{L}'(\mathbf{P}, \mathbf{u}\mathbf{v}^\top - \mathbf{v}\mathbf{u}^\top)\mathbf{u} \ . \tag{14}$$

For instance, if $\mathcal{L}(\cdot, x) = \|\cdot - \sigma(x)\|^2$, then $\mathcal{L}'(\cdot, x) = (1-\sigma(x))\sigma(x)(\sigma(x) - \cdot)$. Once the gradients are computed, one can solve the problem in 10 using alternate minimization in $\mathbf{u}$ and $\mathbf{v}$: each optimization subproblem can be solved using the `scipy` implementation (Virtanen et al., 2020) of the l-BFGS algorithm (Liu and Nocedal, 1989).

Missing Entries Payoff matrices coming from real-world games played by humans usually contain missing entries: one does not have access to the matchups between all the players. For instance, on the `Lichess` website, there usually are few confrontations between low-ranked and high-ranked players. In other words, one only partially has access to P_{ij}, for (i,j) in a given set of pairs of players $\mathcal{D}^{\text{obs}} \subset [n] \times [n]$. Instead of the problem defined in Equation (10), the proposed disc decomposition with a partial set of observation $\mathcal{D}^{\text{obs}}$ reads:

$$\underset{\mathbf{u},\mathbf{v},\text{ s.t. }\mathbf{u}^\top\mathbf{v}=0}{\operatorname{argmin}} \; f(\mathbf{u},\mathbf{v}) \triangleq \sum_{(i,j)\in\mathcal{D}^{\text{obs}}} \mathcal{L}(\mathrm{P}_{ij}, \mathrm{u}_i \mathrm{v}_j - \mathrm{v}_i \mathrm{u}_j) \ . \tag{15}$$

A.2 Detailed on the Compared Methods

Table 1 compares the performance on unseen data of the following models for payoff matrices from Czarnecki et al. (2020)

- The usual Elo score (Elo, 1978)

$$\mathbf{u}^{\text{Elo}} \in \underset{\mathbf{u}}{\operatorname{argmin}} \sum_{i,j} \mathcal{L}(\mathrm{P}_{ij}, \sigma(\mathrm{u}_i - \mathrm{u}_j)), \ \text{ with } \mathcal{L} : (x, \hat{x}) \mapsto \operatorname{bce}(x, \hat{x}) \ . \tag{16}$$

- Elo ++ (Sismanis, 2010), without the sequential aspect

$$\mathbf{u}^{\text{Elo++}} \in \underset{\mathbf{u}}{\operatorname{argmin}} \sum_{i,j} \mathcal{L}(\mathrm{P}_{ij}, \sigma(\mathrm{u}_i - \mathrm{u}_j)), \ \text{ with } \mathcal{L} : (x, \hat{x}) \mapsto \frac{1}{2}\|x - \hat{x}\|^2 \ . \tag{17}$$

- This work (logit space)

$$(\mathbf{u}^{\text{Disc}}, \mathbf{v}^{\text{Disc}}) \in \underset{\mathbf{u},\mathbf{v},\text{ s.t. }\mathbf{u}^\top\mathbf{v}=0}{\operatorname{argmin}} \sum_{i,j} \mathcal{L}(\mathrm{P}_{ij}, \mathrm{u}_i\mathrm{v}_j - \mathrm{v}_i\mathrm{u}_j), \ \text{ with } \mathcal{L} : (x, \hat{x}) \mapsto \operatorname{bce}(x, \sigma(\hat{x})) \ . \tag{18}$$

- This work (probability space), which can be seen as a representation proposed in Balduzzi et al. (2019)

$$(\mathbf{u}^{\mathrm{Disc}}, \mathbf{v}^{\mathrm{Disc}}) \in \underset{\mathbf{u}, \mathbf{v}, \text{ s.t. } \mathbf{u}^\top \mathbf{v}=0}{\operatorname{argmin}} \sum_{i,j} \mathcal{L}(\mathrm{P}_{ij}, \mathrm{u}_i \mathrm{v}_j - \mathrm{v}_i \mathrm{u}_j), \text{ with } \mathcal{L} : (x, \hat{x}) \mapsto \frac{1}{2}\|x - \frac{1}{2} - \hat{x}\|^2 \ . \tag{19}$$

- m-Elo (Balduzzi et al., 2018), with $\bar{\mathbf{P}} \triangleq (\frac{1}{n}\sum_j \mathrm{P}_{ij})_{1 \leq i \leq n}$

$$(\mathbf{u}^{\mathrm{m-Elo}}, \mathbf{v}^{\mathrm{m-Elo}}) \in \underset{\mathbf{u}, \mathbf{v}, \text{ s.t. } \mathbf{u}^\top \mathbf{v}=0}{\operatorname{argmin}} \sum_{i,j} \mathcal{L}(\mathrm{P}_{ij} - (\bar{\mathbf{P}}_i - \bar{\mathbf{P}}_j), \mathrm{u}_i \mathrm{v}_j - \mathrm{v}_i \mathrm{u}_j) \ , \tag{20}$$

  with $\mathcal{L} : (x, \hat{x}) \mapsto \frac{1}{2}\|x - \frac{1}{2} - \hat{x}\|^2$.

This work and the m-Elo are computed using Algorithm 2. Each optimization problem in the alternate minimization (Algorithm 1) is solved using the `scipy` (Virtanen et al., 2020) implementation of l-BFGS (Liu and Nocedal, 1989).

B ADDITIONAL EXPERIMENTS

B.1 Data from Czarnecki et al. (2020)

Figures 8 and 9 are similar to Figure 4, but the Nash clustering visualization from Czarnecki et al. (2020) has been added. For each payoff matrix, Czarnecki et al. (2020) successively compute Nash equilibria of empirical games to cluster players into level sets of "strength". For each dot i, v_i corresponds to the clustered index and u_i to the fraction of the population beaten by the cluster (this is different from our representation where each dot corresponds to a player). The number of clusters can be interpreted as a measure of transitivity: the larger the number of clusters, the more transitive the game.

B.2 Prediction Performance for StarCraft data

Figure 10 shows the prediction performance on unseen entries for StarCraft II confrontations. As observed in Sismanis (2010), adding an extra parameter makes the models significantly overfit. That is why a regularization parameter was added to prevent overfitting:

$$(\mathbf{u}^{\mathrm{Disc}}, \mathbf{v}^{\mathrm{Disc}}) \in \underset{\mathbf{u}, \mathbf{v}, \text{ s.t. } \mathbf{u}^\top \mathbf{v}=0}{\operatorname{argmin}} \sum_{i,j} \mathcal{L}(\mathrm{P}_{ij}, \mathrm{u}_i \mathrm{v}_j - \mathrm{v}_i \mathrm{u}_j) + \frac{\lambda}{2} \sum_i (\mathrm{v}_i - 1)^2 \ , \tag{21}$$

where the regularization parameter λ is chosen using cross-validation.

One can see on Figure 10 that the proposed regularized disc rating can yield better predictions than the usual Elo score. But improvements are not as impressive as for data from Czarnecki et al. 2020 (Figure 5). This might be due to a more difficult estimation of the payoff matrix: one only has access to an incomplete and potentially noisy estimation of the payoff matrix.

C PROOFS OF THEOREMS AND PROPOSITIONS

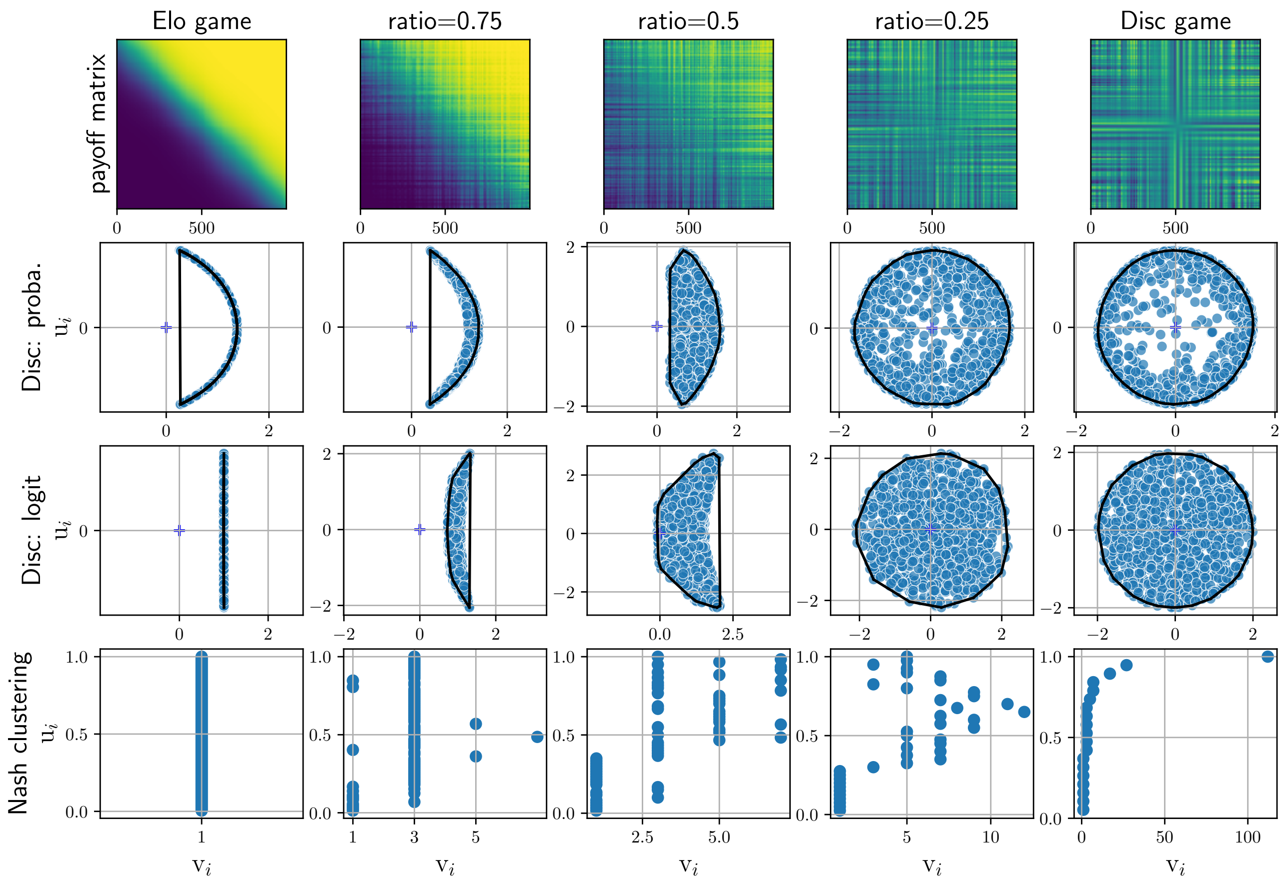


Figure 8: **From Elo to Disc Game.** Payoff matrices (top) and visualization using Algorithm 1 in the probability space (second row), logit space (third row), or Czarnecki et al. (2020) (bottom), for multiple games with payoff $\mathrm{logit}(\mathbf{P}) = \mathrm{ratio} \cdot \mathrm{logit}(\mathbf{P}^{\mathrm{Elo}}) + (1 - \mathrm{ratio}) \cdot \mathrm{logit}(\mathbf{P}^{\mathrm{Disc}})$.

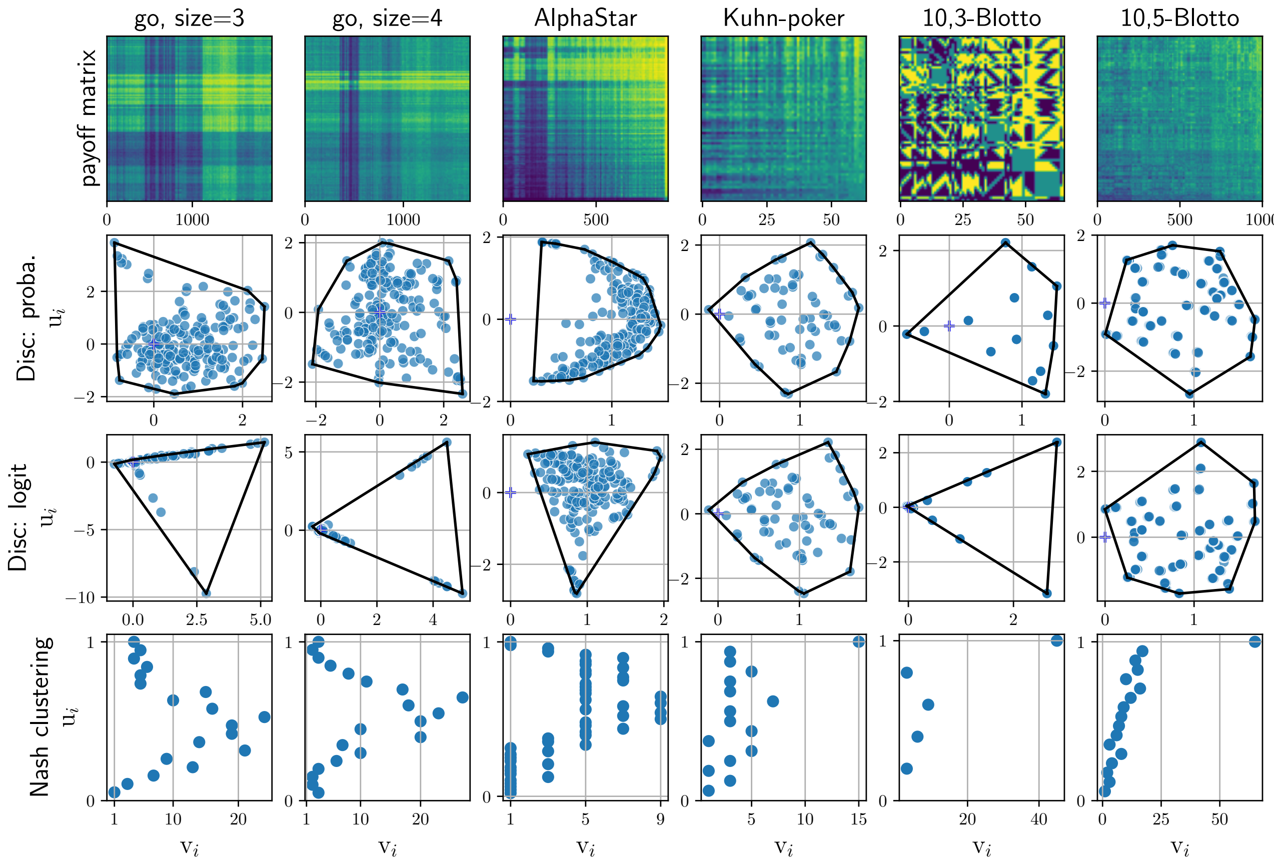


Figure 9: **Transitive and Cyclic Games.** Payoff matrices (top) and visualization using Algorithm 1 in the probability space (second row), logit space (third row), or Czarnecki et al. (2020) (bottom), for multiple games with payoff from Czarnecki et al. (2020).

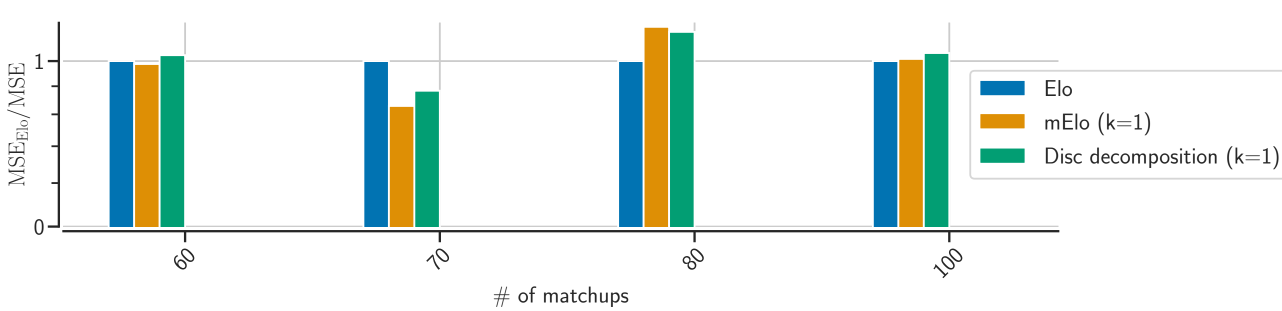


Figure 10: Prediction performances as a function of the number of matchups for the StarCraft data.

Proposition 1. (A Disc Game is Fully Cyclic or Fully Transitive). *Let* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ *and* $\mathrm{P}_{ij} = \sigma(\mathrm{u}_i \mathrm{v}_j - \mathrm{v}_i \mathrm{u}_j)\,,\ i,j \in [n]$ *be a disc game. Let* $U := \mathrm{hull}\{(\mathrm{u}_i, \mathrm{v}_i)\}$ *be the convex hull of the players. If* $(0,0)$ *is* not *in the border of* U,[8] *then the disc game is either fully transitive or fully cyclic. Precisely, the disc game is*

1. *fully cyclic if and only if* $(0,0) \in \mathrm{Int}(U)$,
2. *fully transitive if and only if* $(0,0) \notin \mathrm{Int}(U)$.

Proof. Let us start this proof by showing that there always exists a cycle a size 3 in a larger cycle.

Lemma 1 A cycle of size n implies a cycle of size 3. *If there exists* $\mathrm{u}_1, \dots, \mathrm{u}_n \in U$ *such that* $\mathrm{u}_1 \xrightarrow{\text{beats}} \dots \xrightarrow{\text{beats}} \mathrm{u}_n \xrightarrow{\text{beats}} \mathrm{u}_1$, *then there exists* $u, v, w \in U$ *such that* $u \xrightarrow{\text{beats}} v \xrightarrow{\text{beats}} w \xrightarrow{\text{beats}} u$.

Proof. We will use a recursive argument. It is true for $n = 3$. Let us assume it is true for a given $n \in \mathbb{N}$, then given $\mathrm{u}_1 \xrightarrow{\text{beats}} \dots \xrightarrow{\text{beats}} \mathrm{u}_{n+1} \xrightarrow{\text{beats}} u_1$ we have two possible cases

1. $\mathrm{u}_1 \xrightarrow{\text{beats}} \mathrm{u}_n \xrightarrow{\text{beats}} \mathrm{u}_{n+1} \xrightarrow{\text{beats}} \mathrm{u}_1$: In that case, we have proved the result.
2. $\mathrm{u}_1 \xleftarrow{\text{beats}} \mathrm{u}_n \xrightarrow{\text{beats}} \mathrm{u}_{n+1} \xrightarrow{\text{beats}} \mathrm{u}_1$. In that case, it means that $\mathrm{u}_1 \xrightarrow{\text{beats}} \dots \xrightarrow{\text{beats}} \mathbf{u}_n \xrightarrow{\text{beats}} \mathrm{u}_1$ and thus we can use the recurrence hypothesis that there exists a cycle of size 3 inside a cycle of size n.

□

We can now show Proposition 1.

The idea of the proof is to use the polar coordinates for all the points $(\mathrm{u}_i, \mathrm{v}_i) = (r_i, \theta_i) \in \mathbb{R}^2$. We can notice that $\mathrm{P}_{ij} = \sigma(\mathrm{u}_i \mathrm{v}_j - \mathrm{v}_i \mathrm{u}_j) = \sigma(r_i r_j \sin(\theta_i - \theta_j))$. Thus $\mathrm{P}_{ij} > 1/2$ if and only if $\theta_j + \pi > \theta_i > \theta_j$. If the game is cyclic it means that there exists a cycle of size 3. Without any loss of generality let us assume $\theta_1 = 0$. Then we have that $\pi > \theta_2 > 0$ and $-\pi < \theta_3 < 0$. Finally the fact that $\theta_3 < \theta_2 + \pi$ implies that either $\theta_2 > \pi/2$ or $\theta_3 < \pi/2$. Thus this forms a triangle that contains $(0,0)$. Thus we can show that $(0,0)$ is in the interior of the convex set defined by three points of the cycle. If $(0,0)$ is in the interior of U then we can find a cycle. □

Proposition 2. (Reparametrization). *Let* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ *and* $\mathrm{P}_{ij} := \sigma(\mathrm{u}_i \mathrm{v}_j - \mathrm{v}_i \mathrm{u}_j)\,,\ i,j \in [n]$ *be a* ***transitive*** *disc game. Then there exists a reparametrization of the disc game* $(\tilde{\mathrm{u}}_i, \tilde{\mathrm{v}}_i)$ *such that, for all* $i \in [n]$, $\tilde{\mathrm{v}}_i > 0$, *and*

$$\sigma(\mathrm{u}_i \mathrm{v}_j - \mathrm{v}_i \mathrm{u}_j) = \mathrm{P}_{ij} = \sigma(\tilde{\mathrm{u}}_i \tilde{\mathrm{v}}_j - \tilde{\mathrm{v}}_i \tilde{\mathrm{u}}_j)\ . \tag{22}$$

Proof. To prove this result we will use Proposition 1. Since the disc game is considered transitive and that $(0,0)$ is not in the border of U, we know that $(0,0) \notin U = \mathrm{hull}\{(\mathrm{u}_i, \mathrm{v}_i)\}$. Thus, by the hyperplane separation theorem (Boyd and Vandenberghe, 2004, Example 2.20) there exists a direction $a \in \mathbb{R}^2\,,\ \|a\|_2 = 1$ such that, $\langle a, [\mathrm{u}_i, \mathrm{v}_i] \rangle > 0\,,\ \forall i \in [n]$. Setting $b := [a_2, -a_1]$, we get that (a, b) is an orthonormal basis of $\mathbb{R}^2$. Let $(\tilde{\mathrm{u}}_i, \mathbf{v}_i) := (\langle a, [\mathrm{u}_i, \mathrm{v}_i] \rangle, \langle b, [\mathrm{u}_i, \mathrm{v}_i] \rangle)$ be the coordinate of $[\mathrm{u}_i, \mathrm{v}_i]$ in this new basis. Finally, we just need to remark that $\mathrm{u}_i \mathrm{v}_j - \mathrm{v}_i \mathrm{u}_j$ does not depend on the choice of basis. It is because, $\mathrm{u}_i \mathrm{v}_j - \mathrm{v}_i \mathrm{u}_j = \sigma(\|OM_i\| \|OM_j\| \sin \widehat{M_i O M_j})$ where $M_i \in \mathbb{R}^2$ corresponds to the point with coordinate $[(\mathrm{u}_i, \mathrm{v}_i)]$. □

Theorem 1. (Normal Decomposition, Greub 1975). *Suppose* $\mathbf{A} \in \mathbb{R}^{n \times n}$ *is such that* $\mathrm{A} = -\mathrm{A}^\top$. *Then, with* $k = \lfloor n/2 \rfloor$, *there exists* $\lambda_1 \geq \ldots \geq \lambda_k$ *and* $(\mathbf{u}^{(l)}, \mathbf{v}^{(l)}) \in \mathbb{R}^n \times \mathbb{R}^n$, $1 \leq l \leq k$, *such that* $(\mathbf{u}^{(l)}, \mathbf{v}^{(l)})$ *is an orthogonal family and* $\mathbf{A} = \sum_{l=1}^k (\mathbf{u}^{(l)} \mathbf{v}^{(l)\top} - \mathbf{v}^{(l)} \mathbf{u}^{(l)\top})$.

Proof. Proof of Theorem 1 can be found in Francinou et al. (2008, Sec. 2.5) or Greub (1975, §8.16). □

Theorem 2. (An Empirical Game has at most a Single Transitive Disc Game). *Let* $\mathbf{P}$ *be the payoff matrix of a symmetric zero-sum game, and let* $(\mathbf{u}^{(l)}, \mathbf{v}^{(l)}) \in \mathbb{R}^n \times \mathbb{R}^n$ *be the normal decomposition of the skew-symmetric matrix* $\mathrm{logit}(\mathbf{P})$ *(Theorem 1), then there exists* at most *one pair* $(\mathbf{u}^{(l)}, \mathbf{v}^{(l)}) \in \mathbb{R}^n \times \mathbb{R}^n$ *such that the disc game defined by* $\mathrm{P}_{ij} = \sigma(\mathrm{u}_i^{(l)} \mathrm{v}_j^{(l)} - \mathrm{v}_i^{(l)} \mathrm{u}_j^{(l)})$, $i, j \in [n]$, *is transitive.*

[8] if it is the case we can, for instance, apply an infinitesimal perturbation on the points $(\mathrm{u}_i, \mathrm{v}_i)$.

Proof. The proof of Theorem 2 relies on

- The fact that $(\mathbf{u}^{(l)}, \mathbf{v}^{(l)})$ are orthogonal.
- The reparametrization property (Proposition 2).

The normal decomposition $(\mathbf{u}^{(l)}, \mathbf{v}^{(l)})$ is composed of orthogonal vectors ($\mathbf{v}^{(l)}$ are orthogonal to each others and to $\mathbf{u}^{(l)}$).

Suppose that there exist two transitive pairs of components, $(\mathbf{u}^{(1)}, \mathbf{v}^{(1)})$ and $(\mathbf{u}^{(2)}, \mathbf{v}^{(2)})$. Using Proposition 2 we consider the reparametrization such that $\mathrm{v}_i^{(1)} > 0$, for all $i \in [n]$, and $\mathrm{v}_i^{(2)} > 0$, for all $i \in [n]$, hence one has $\mathbf{v}^{(1)\top}\mathbf{v}^{(2)} > 0$. In addition, since $\mathbf{v}^{(1)}$ and $\mathbf{v}^{(2)}$ are orthogonal we have that $\mathrm{v}_i^{(1)\top}\mathrm{v}_i^{(2)} = 0$, which yields a contradiction. Hence there exists at most one transitive pair of components. □

References for the Appendix

Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004. (Cited on pages 16 and 4)

Serge Francinou, Hervé Gianella, and Serge Nicolas. *Exercices de mathématiques oraux x-ens algèbre*, volume 3. CASSINI, 2008. (Cited on pages 16 and 4)

Werner H. Greub. *Linear algebra*, volume 23. Springer Science & Business Media, 4th edition, 1975. (Cited on pages 16 and 4)