On a kinetic Elo rating model for players with dynamical strength

Bertram Düring

Department of Mathematics

Joint work with
M. Torregrossa and M.-T. Wolfram
Interacting many-agent systems in socio-economics

Examples: wealth distribution in an economy, opinion formation, crowd dynamics, ...
Examples: wealth distribution in an economy, opinion formation, crowd dynamics, ...

Features:

- large number of interacting agents
- model of full system not tractable
- quantities of interest are aggregates
- dynamics!
- emergent behaviour, self-organisation
Examples: wealth distribution in an economy, opinion formation, crowd dynamics, ...

Features:

- large number of interacting agents
- model of full system not tractable
- quantities of interest are aggregates
- dynamics!
- emergent behaviour, self-organisation

⇒ mathematical tools from kinetic theory
Kinetic models for socio-economic systems

Conceptual approach (e.g. [Pareschi&Toscani, 2015], ...):

▶ describe dynamics of system by microscopic interactions among agents
▶ perform many interactions (analytically or numerically)
▶ observe emergent behaviour, patterns in macroscopic distribution of agents
▶ derive partial differential equations (Boltzmann, Fokker-Planck-type) which (approximatively) govern the time-evolution of the density
Kinetic models for socio-economic systems

Conceptual approach (e.g. [Pareschi&Toscani, 2015], ...):

- describe dynamics of system by microscopic interactions among agents
- perform many interactions (analytically or numerically)
- observe emergent behaviour, patterns in macroscopic distribution of agents
- derive partial differential equations (Boltzmann, Fokker-Planck-type) which (approximatively) govern the time-evolution of the density

Benefits:

- more (analytically and numerically) tractable model
- understanding role of parameters in the microscopic interactions for emergent behaviour
- PDE: nonlinear, anisotropic, nonlocal, degenerate
Elo rating for zero-sum games

- Rating system developed by physicist Arpad Elo to determine relative skill levels of players in zero-sum games
- Originally used for chess
- Also for online gaming, table tennis, ...
- Multiplayer games: football, basketball, ..
- 2018: FIFA world ranking to use Elo system
Elo rating for zero-sum games

- rating system developed by physicist Arpad Elo to determine relative skill levels of players in zero-sum games
- originally used for chess
- also for online gaming, table tennis, ...
- multiplayer games: football, basketball, ...
- 2018: FIFA world ranking to use Elo system
- each player assigned rating number which may change as games played
- difference in rating between two players should predict outcome of a game
- players with same rating who play each other should have same probability of winning/loosing
- difference between ratings determines number of points gained or lost after a game
Elo rating: Junca & Jabin (2015)

**Continuous variables:**
- Strength $\rho$ (fixed, unobservable)
- Rating $R$ (variable, observable)

Following each game, ratings are adjusted

\[
R^*_{i} = R_{i} + \gamma (S_{ij} - b(R_{i} - R_{j})) \\
R^*_{j} = R_{j} + \gamma (-S_{ij} - b(R_{j} - R_{i}))
\]

- Random variable $S_{ij} \in \{-1, 1\}$: score result of the game
- Function $b$ moderates extreme differences, e.g. $b(z) = \tanh(c z)$ with some $c > 0$
- Assume mean score $\langle S_{ij} \rangle = b(\rho_{i} - \rho_{j})$
- Speed of adjustment $\gamma > 0$

Effect:
- Player with high rating wins against player with a low rating $\Rightarrow$ ratings change little
- Player with low rating wins against highly rated player $\Rightarrow$ ratings are strongly adjusted
Elo rating: Junca & Jabin (2015)

continuous variables: strength $\rho$ (fixed, unobservable)
rating $R$ (variable, observable)

Following each game, ratings are adjusted

$$R_i^* = R_i + \gamma(S_{ij} - b(R_i - R_j))$$
$$R_j^* = R_j + \gamma(-S_{ij} - b(R_j - R_i))$$
Elo rating: Junca & Jabin (2015)

**continuous variables:** strength \( \rho \) (fixed, unobservable) 
rating \( R \) (variable, observable)

Following each game, ratings are adjusted

\[
R_i^* = R_i + \gamma(S_{ij} - b(R_i - R_j)) \\
R_j^* = R_j + \gamma(-S_{ij} - b(R_j - R_i))
\]

- random variable \( S_{ij} \in \{-1, 1\} \): score result of the game
- function \( b \) moderates extreme differences, e.g.
  \( b(z) = \tanh(cz) \) with some \( c > 0 \)
- assume mean score \( \langle S_{ij} \rangle = b(\rho_i - \rho_j) \)
- speed of adjustment \( \gamma > 0 \)
Elo rating: Junca & Jabin (2015)

**continuous variables**: strength $\rho$ (fixed, unobservable)

**rating** $R$ (variable, observable)

Following each game, ratings are adjusted

\[
R_i^* = R_i + \gamma(S_{ij} - b(R_i - R_j))
\]

\[
R_j^* = R_j + \gamma(-S_{ij} - b(R_j - R_i))
\]

- random variable $S_{ij} \in \{-1, 1\}$: score result of the game
- function $b$ moderates extreme differences, e.g.
  \[
b(z) = \tanh(cz) \text{ with some } c > 0
  \]
- assume mean score $\langle S_{ij} \rangle = b(\rho_i - \rho_j)$
- speed of adjustment $\gamma > 0$

**Effect:**

- player with high rating wins against player with a low rating $\implies$ ratings change little
- player with low rating wins against highly rated player $\implies$ ratings are strongly adjusted
Elo rating: Junca & Jabin (2015)

Question: Is the rating fair, i.e. do ‘\( R_i \rightarrow \rho_i \)’ over time?
Elo rating: Junca & Jabin (2015)

**Question:** Is the rating fair, i.e. do ‘$R_i \rightarrow \rho_i$’ over time? Kinetic equation: distribution of players $f = f(r, t)$ with respect to ratings satisfies

$$\partial_t f(r, t) + \partial_r (a(f) f) = 0$$

with $a(f) = \int_{\mathbb{R}^2} w(r - r')(b(\rho - \rho') - b(r - r')) f(t, r', \rho') d\rho' dr'$

and given interaction rate function $w(r - r')$
Elo rating: Junca & Jabin (2015)

**Question:** Is the rating fair, i.e. do ‘$R_i \rightarrow \rho_i$’ over time?

Kinetic equation: distribution of players $f = f(r, t)$ with respect to ratings satisfies

$$\partial_t f(r, t) + \partial_r (a(f) f) = 0$$

with $a(f) = \int_{\mathbb{R}^2} w(r - r')(b(\rho - \rho') - b(r - r')) f(t, r', \rho') d\rho' dr'$

and given interaction rate function $w(r - r')$

**Long time behaviour:**

- $w = 1$ (‘all-play-all’ tournament): ratings converge exponentially fast to intrinsic strengths
- $w$ with local interactions: ratings may not converge to intrinsic strengths, rating fails to give a fair representation of the player’s strength distribution
Elo rating: learning effects

**continuous variables:**
- strength $\rho$ (variable, unobservable)
- rating $R$ (variable, observable)

Following each game, ratings are adjusted:

$$
R^*_i = R_i + \gamma (S_{ij} - b(R_i - R_j))$$
$$
R^*_j = R_j + \gamma (-S_{ij} - b(R_j - R_i))$$

and players learn:

$$
\rho^*_i = \rho_i + \gamma h (\rho_j - \rho_i) + \eta$$
$$
\rho^*_j = \rho_j + \gamma h (\rho_i - \rho_j) + \tilde{\eta}
$$

where $\eta$, $\tilde{\eta}$ are random variables with mean zero.
Elo rating: learning effects

**continuous variables**: strength $\rho$ *(variable, unobservable)*

**rating $R$** *(variable, observable)*

Following each game, ratings are adjusted

$$R_i^* = R_i + \gamma(S_{ij} - b(R_i - R_j))$$

$$R_j^* = R_j + \gamma(-S_{ij} - b(R_j - R_i))$$

and players learn

$$\rho_i^* = \rho_i + \gamma h(\rho_j - \rho_i) + \eta$$

$$\rho_j^* = \rho_j + \gamma h(\rho_i - \rho_j) + \tilde{\eta}$$

where $\eta, \tilde{\eta}$ are random variables with mean zero.
Elo rating: learning effects

**continuous variables**: strength $\rho$ *(variable, unobservable)*  
**rating** $R$ *(variable, observable)*

Following each game, ratings are adjusted

\[ R_i^* = R_i + \gamma(S_{ij} - b(R_i - R_j)) \]
\[ R_j^* = R_j + \gamma(-S_{ij} - b(R_j - R_i)) \]

and players learn

\[ \rho_i^* = \rho_i + \gamma h(\rho_j - \rho_i) + \eta \]
\[ \rho_j^* = \rho_j + \gamma h(\rho_i - \rho_j) + \tilde{\eta} \]

where $\eta, \tilde{\eta}$ are random variables with mean zero.
We consider two main effects:

- **learning by interaction**: we assume each player learns in a game, however players with lower strength benefit more. Possible choice \( h_1(\rho_j - \rho_i) = 1 + b(\rho_j - \rho_i) \)

- **gain/loss of self-confidence**: assume gain/loss of stronger player is the same as that of the weaker one, e.g. \( h_2(\rho_j - \rho_i) = S_{ij}[1 - \tanh^2(\rho_j - \rho_i)] \)
Choice of learning mechanism

We consider two main effects:

- **learning by interaction**: we assume each player learns in a game, however players with lower strength benefit more. Possible choice $h_1(\rho_j - \rho_i) = 1 + b(\rho_j - \rho_i)$

- **gain/loss of self-confidence**: assume gain/loss of stronger player is the same as that of the weaker one, e.g. $h_2(\rho_j - \rho_i) = S_{ij}[1 - \tanh^2(\rho_j - \rho_i)]$

With parameters $\alpha, \beta$ we have in summary

$$h(\rho_j - \rho_i) = \alpha h_1(\rho_j - \rho_i) + \beta h_2(\rho_j - \rho_i)$$
Some properties of the interaction

Preservation of total value of the rating pointwise and in mean,

\[ \langle R_i^* + R_j^* \rangle = R_i + R_j. \]
Some properties of the interaction

Preservation of total value of the rating pointwise and in mean,

$$\langle R_i^* + R_j^* \rangle = R_i + R_j.$$ 

Evolution of total strength depends is not affected by the function $h_2$, since

$$\langle \rho_j^* + \rho_j^* \rangle - (\rho_j + \rho_j) = 2\gamma\alpha.$$
Some properties of the interaction

Preservation of total value of the rating pointwise and in mean,

\[ \langle R_i^* + R_j^* \rangle = R_i + R_j. \]

Evolution of total strength depends is not affected by the function \( h_2 \), since

\[ \langle \rho_j^* + \rho_j^* \rangle - (\rho_j + \rho_j) = 2\gamma \alpha. \]

\( \therefore \) constant increase of strength of population
Boltzmann type equation

Distribution function $f_\gamma = f_\gamma(\rho, R, t)$ satisfies

$$
\frac{d}{dt} \int_\Omega \phi(\rho_i, R_j) f_\gamma(\rho_i, R_i, t) \, d\rho_i dR_i
$$

$$
= \frac{1}{2} \left\langle \int_\Omega \int_\Omega \left( \phi(\rho^*_i, R^*_j) + \phi(\rho^*_j, R^*_j) - \phi(\rho_i, R_i) - \phi(\rho_j, R_j) \right)
\times w(R_i - R_j) f_\gamma(\rho_i, R_i, t) f_\gamma(\rho_j, R_j, t) \, d\rho_j dR_j d\rho_i dR_i \right\rangle
$$

where $\phi(\cdot)$ is a (smooth) test function
Fokker-Planck limit

Rescaling $t' = \gamma t$, in the quasi-invariant limit

$\gamma \to 0$, $\sigma_\eta \to 0$ such that $\frac{\sigma_\eta^2}{\gamma} =: \sigma^2$ is fixed
Fokker-Planck limit

Rescaling $t' = \gamma t$, in the quasi-invariant limit

$\gamma \to 0, \sigma_\eta \to 0$ such that $\frac{\sigma_\eta^2}{\gamma} =: \sigma^2$ is fixed

we obtain the Fokker-Planck equation

$$
\frac{\partial f(\rho, R, t)}{\partial t} + \frac{\partial}{\partial R} \left( a[f] f(\rho, R, t) \right) + \frac{\partial}{\partial \rho} \left( c[f] f(\rho, R, t) \right) - \frac{\sigma^2}{2} d[f] \frac{\partial^2}{\partial \rho^2} f(\rho, R, t) = 0
$$

where

$$
a[f] = \int_{\mathbb{R}^2} w(R - R_j) \left( b(\rho - \rho_j) - b(R - R_j) \right) f(\rho_j, R_j, t) \, d\rho_j dR_j
$$

$$
c[f] = \int_{\mathbb{R}^2} w(R - R_j) \left( \alpha h_1(\rho_j - \rho) + \beta \langle h_2(\rho_j - \rho) \rangle \right) f(\rho_j, R_j, t) \, d\rho_j dR_j
$$

$$
d[f] = \int_{\mathbb{R}^2} w(R - R_j) f(\rho_j, R_j, t) \, d\rho_j dR_j
$$
Shifted Fokker-Planck equation

We want to study steady states of the distribution
\[ g(\rho, R, t) = f(\rho + H(\rho, R, t), R, t) \]
where \( H \) is given by
\[ \frac{\partial H(\rho, R, t)}{\partial t} = \int R^2 \alpha w(R - R_j) f(\rho_j, R_j, t) d\rho_j dR_j. \]

\[ \Rightarrow \] ensures mean value is preserved in time.

The evolution equation for \( g(\rho, R, t) \) is
\[ \frac{\partial g}{\partial t} + \frac{\partial}{\partial R} (a[g]g) + \frac{\partial}{\partial \rho} (\tilde{c}[g]g) - \sigma^2 \frac{d[g]}{2 \partial^2 \rho^2} g = 0, \]
where
\[ \tilde{c}[g] = \int R^2 \alpha b(\rho_j - \rho) + \beta \langle h^2 (\rho_j - \rho) \rangle w(R - R_j) g(\rho_j, R_j, t) d\rho_j dR_j. \]
We want to study steady states of the distribution

\[ g(\rho, R, t) = f(\rho + H(\rho, R, t), R, t) \]

where \( H \) is given by

\[ \frac{\partial H(\rho, R, t)}{\partial t} = \int_{\mathbb{R}^2} \alpha w(R - R_j) f(\rho_j, R_j, t) \, d\rho_j dR_j. \]

\( \Rightarrow \) ensures mean value is preserved in time.
Shifted Fokker-Planck equation

We want to study steady states of the distribution

\( g(\rho, R, t) = f(\rho + H(\rho, R, t), R, t) \) where \( H \) is given by

\[
\frac{\partial H(\rho, R, t)}{\partial t} = \int_{\mathbb{R}^2} \alpha w(R - R_j) f(\rho_j, R_j, t) \, d\rho_j dR_j.
\]

\( \tilde{c}[g] \) ensures mean value is preserved in time.

The evolution equation for \( g(\rho, R, t) \) is

\[
\frac{\partial g}{\partial t} + \frac{\partial}{\partial R}(a[g]g) + \frac{\partial}{\partial \rho}(\tilde{c}[g]g) - \frac{\sigma^2}{2}d[g] \frac{\partial^2}{\partial \rho^2}g = 0,
\]

where

\[
\tilde{c}[g] = \int_{\mathbb{R}^2} (\alpha b(\rho_j - \rho) + \beta \langle h_2(\rho_j - \rho) \rangle) w(R - R_j) g(\rho_j, R_j, t) \, d\rho_j dR_j.
\]
We consider the following problem on a bounded domain \( \Omega \subset \mathbb{R}^2 \), with no-flux boundary condition

\[
\frac{\partial g}{\partial t} + \frac{\partial}{\partial R}(a[g]g) + \frac{\partial}{\partial \rho}(\tilde{c}[g]g) - \frac{\sigma^2}{2}d[g] \frac{\partial^2}{\partial \rho^2}g = 0, \quad \text{in } \Omega \times (0, T),
\]

\[
\frac{\partial}{\partial \nu}g = 0 \quad \text{on } \partial \Omega,
\]

\[
g(\rho, R, 0) = g_0(\rho, R) \quad \text{in } \Omega.
\]
Let \( \Omega \subset \mathbb{R}^2 \) bounded Lipschitz domain.

**Theorem**

Let \( g_0 \in H^1(\Omega) \) and \( 0 \leq g_0 \leq M_0 \) for some \( M_0 > 0 \) and assume \( h_1, \langle h_2 \rangle, b \in L^\infty(\Omega) \cap C^2(\Omega) \). Then there exists a weak solution \( g \in L^2(0, T; H^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega)) \), satisfying \( 0 \leq g \leq M_0 e^{\lambda t} \) for all \( (\rho, R) \in \Omega, t > 0 \), with a constant \( \lambda > 0 \) depending on the functions \( h_1, \langle h_2 \rangle, b \) and \( w \).
Sketch of the proof

The proof involves several steps:

- **Step 0**: regularised, truncated problem, adding $\mu \Delta g(\rho, R, t), \mu > 0$
- **Step 1**: solution of linearised, regularised problem; definition of fixed point operator
- **Step 2**: uniform $L^\infty$ bounds and existence of fixed point (Leray-Schauder)
- **Step 3**: uniform $H^1$ bound (independent of $\mu$)
- **Step 4**: limit $\mu \to 0$ (Aubin-Lions lemma)
Define the energy $E_2(t) = \int_{\mathbb{R}^2} (\rho - R)^2 g(\rho, R, t) \, d\rho dR$. 

$\Rightarrow$ indicates concentration in neighbourhood of diagonal.
Long-time behaviour of solutions

Define the energy $E_2(t) = \int_{\mathbb{R}^2} (\rho - R)^2 g(\rho, R, t) \, d\rho dR$.
At least for $w = 1$ we can compute

$$\frac{d}{dt} E_2(t)$$

$$= - \int_{\mathbb{R}^4} (R - R_j) b(R - R_j) g(\rho, R, t) g(\rho_j, R_j, t) \, d\rho_j dR_j d\rho dR$$

$$- \int_{\mathbb{R}^4} (\rho - \rho_j) b(\rho - \rho_j) g(\rho, R, t) g(\rho_j, R_j, t) \, d\rho_j dR_j d\rho dR$$

$$- \alpha \int_{\mathbb{R}^4} (\rho - \rho_j) b(\rho - \rho_j) g(\rho, R, t) g(\rho_j, R_j, t) \, d\rho_j dR_j d\rho dR$$

$$- 2\beta \int_{\mathbb{R}^4} (\rho - \rho_j) \langle h_2(\rho - \rho_j) \rangle g(\rho, R, t) g(\rho_j, R_j, t) \, d\rho_j dR_j d\rho dR$$

$$+ \sigma^2$$

$\Rightarrow$ indicates concentration in neighbourhood of diagonal
Long-time behaviour of solutions

Define the energy $E_2(t) = \int_{\mathbb{R}^2} (\rho - R)^2 g(\rho, R, t) \, d\rho dR$.

At least for $w = 1$ we can compute

$$\frac{d}{dt} E_2(t)$$

$$= - \int_{\mathbb{R}^4} (R - R_j) b(R - R_j) g(\rho, R, t) g(\rho_j, R_j, t) \, d\rho_j dR_j d\rho dR$$

$$- \int_{\mathbb{R}^4} (\rho - \rho_j) b(\rho - \rho_j) g(\rho, R, t) g(\rho_j, R_j, t) \, d\rho_j dR_j d\rho dR$$

$$- \alpha \int_{\mathbb{R}^4} (\rho - \rho_j) b(\rho - \rho_j) g(\rho, R, t) g(\rho_j, R_j, t) \, d\rho_j dR_j d\rho dR$$

$$- 2\beta \int_{\mathbb{R}^4} (\rho - \rho_j) \langle h_2(\rho - \rho_j) \rangle g(\rho, R, t) g(\rho_j, R_j, t) \, d\rho_j dR_j d\rho dR$$

$$+ \sigma^2$$

$\leadsto$ indicates concentration in neighbourhood of diagonal
Numerical results: all-play-all

Direct Monte Carlo simulation method: $N = 5000$ players

Steady state (top view) – no diffusion

Steady state (top view) – diffusion $\nu = 0.025$
Numerical steady states for Fokker-Planck equation

no diffusion

with diffusion
Numerical results: all-play-all

Energy decay for Fokker-Planck equation

\[ E_2(t) = \int_{\mathbb{R}^2} (\rho - R)^2 g(\rho, R, t) \, d\rho dR \]
Numerical results: competition with similar rating

Consider two groups of players:

- first group is underrated, all players have rating $R = 0.2$, but $\rho \in \mathcal{N}(0.75, 0.1)$
- second group is overrated, with rating $R = 0.9$ and uniform distribution in $\rho$

Choose $\alpha = 0.1$ and $\beta = 0$. 
Numerical results: competition with similar rating

Consider two groups of players:

- first group is underrated, all players have rating $R = 0.2$, but $\rho \in \mathcal{N}(0.75, 0.1)$
- second group is overrated, with rating $R = 0.9$ and uniform distribution in $\rho$

Choose $\alpha = 0.1$ and $\beta = 0$. 
Consider two groups of players:

- first group is underrated, all players have rating $R = 0.2$, but $\rho \in \mathcal{N}(0.75, 0.1)$
- second group is overrated, with rating $R = 0.9$ and uniform distribution in $\rho$

Choose $\alpha = 0.1$ and $\beta = 0$.

.assignment initial ratings is a delicate issue
Numerical results: competition with similar rating

Consider two groups of players:

- first group is underrated, all players have rating $R = 0.2$, but $\rho \in \mathcal{N}(0.75, 0.1)$
- second group is overrated, with rating $R = 0.9$ and uniform distribution in $\rho$

Choose $\alpha = 0.1$ and $\beta = 0.05$. 

Numerical results: competition with similar rating

Consider two groups of players:
- first group is underrated, all players have rating \( R = 0.2 \), but \( \rho \in \mathcal{N}(0.75, 0.1) \)
- second group is overrated, with rating \( R = 0.9 \) and uniform distribution in \( \rho \)

Choose \( \alpha = 0.1 \) and \( \beta = 0.05 \).
Numerical results: competition with similar rating

Consider two groups of players:

- first group is underrated, all players have rating \( R = 0.2 \), but \( \rho \in \mathcal{N}(0.75, 0.1) \)
- second group is overrated, with rating \( R = 0.9 \) and uniform distribution in \( \rho \)

Choose \( \alpha = 0.1 \) and \( \beta = 0.05 \).

\[ \rightarrow \text{adapted learning mechanism leads to convergence of the ratings} \]
Assume one player is playing unfair, e.g. through cheating, doping or bribing of referees.
Numerical results: foul play

Assume one player is playing unfair, e.g. through cheating, doping or bribing of referees.

outcome of every microscopic game which involves this player is biased in their favour
Numerical results: foul play

Assume one player is playing unfair, e.g. through cheating, doping or bribing of referees.

Outcome of every microscopic game which involves this player is biased in their favour.

Ratings and strength of all players except the first one converge around diagonal. The cheating player (indicated by a star) ends up with a higher rating.
Summary

- Elo rating system for games
  - Boltzmann-type, Fokker-Planck-type limit equations
- well-posedness
- long-time behaviour: convergence to players’ strength
- assigning initial ratings is delicate


THANK YOU!
Summary

- Elo rating system for games
  - Boltzmann-type, Fokker-Planck-type limit equations
- well-posedness
- long-time behaviour: convergence to players’ strength
- assigning initial ratings is delicate

Summary

- Elo rating system for games
  - Boltzmann-type, Fokker-Planck-type limit equations
- well-posedness
- long-time behaviour: convergence to players’ strength
- assigning initial ratings is delicate


THANK YOU!