



HAL
open science

Large ranking games with non-observable diffusion control

Stefan Ankirchner, Nabil Kazi-Tani, Julian Wendt, Chao Zhou

► **To cite this version:**

Stefan Ankirchner, Nabil Kazi-Tani, Julian Wendt, Chao Zhou. Large ranking games with non-observable diffusion control. 2021. hal-03138716

HAL Id: hal-03138716

<https://hal.science/hal-03138716>

Preprint submitted on 11 Feb 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.

Large ranking games with non-observable diffusion control

S. Ankirchner ^{*} N. Kazi-Tani [†] J. Wendt [‡] C. Zhou [§]

February 11, 2021

We consider a symmetric game with the following features: each player can control the fluctuation intensity of an individual dynamic state process up to some exponentially distributed random time; no player can observe the other players' actions; the players whose final states are among the best $\alpha \in (0, 1)$ of all states receive a fixed prize. Within the mean-field limit version of the game we compute an explicit equilibrium, a threshold strategy that consists in choosing the maximal fluctuation intensity when the state is below a given threshold, and the minimal intensity else. We show that for large n the n -tupel of the threshold strategy provides an approximate Nash-equilibrium of the n -player game. We also derive the rate at which the approximate equilibrium reward and the best response reward converge to each other, as the number of players n tends to infinity.

2020 MSC : 91A16, 91A07, 93E20.

Keywords : Diffusion control; game with non-observable actions; rank-based reward; mean-field limit; oscillating Brownian motion.

1. Introduction

We start by describing a game that models, in a stylized form, the competition among many agents who can each control the fluctuation intensity

^{*}Stefan Ankirchner, Institute for Mathematics, University of Jena, Ernst-Abbe-Platz 2, 07743 Jena, Germany. *Email*: s.ankirchner@uni-jena.de.

[†]Nabil Kazi-Tani, Univ Lyon, Université de Lyon 1, Laboratoire de Sciences Actuarielle et Financière, ISFA, 50 Avenue Tony Garnier, F-69007 Lyon, France. *Email*: nabil.kazi-tani@univ-lyon1.fr.

[‡]Julian Wendt, Institute for Mathematics, University of Jena, Ernst-Abbe-Platz 2, 07743 Jena, Germany. *Email*: julian.wendt@uni-jena.de.

[§]Chao Zhou, Department of Mathematics, City University of Hong Kong, Kowloon Tong, Hong Kong. On leave from Department of Mathematics, National University of Singapore. *Email*: chaozhou@cityu.edu.hk.

of a state until an individual random time horizon and who receive some benefit if their final state is among the highest $\alpha \in (0, 1)$. Our game model is generic and allows for various interpretations: it can for instance describe competition among research and development teams [9], among animals for mating partners [25], or among cryptocurrencies miners [29].

Throughout let n be a large natural number representing the number of players. Let $W = (W^1, \dots, W^n)$ be an n -dimensional Brownian motion on a complete probability space (Ω, \mathcal{F}, P) . In addition, let τ_1, \dots, τ_n be an independent family of random variables that are exponentially distributed with common parameter $\lambda \in (0, \infty)$. We assume that each τ_i , $i \in \{1, \dots, n\}$, is independent of the Brownian motion W . For every $i \in \{1, \dots, n\}$ let $(\mathcal{F}_t^i)_{t \geq 0}$ be the filtration generated by the Brownian motion W^i and augmented by the P -null sets in \mathcal{F} .

Let $0 < \sigma_1 \leq \sigma_2$. The set of strategies of player i , denoted by \mathcal{A}^i , consists of all (\mathcal{F}_t^i) -progressively measurable processes $H : \Omega \times [0, \infty) \rightarrow [\sigma_1, \sigma_2]$. Given that player i chooses the strategy H^i , her final state satisfies

$$X_{\tau_i}^{i, H^i} = \int_0^{\tau_i} H_s^i dW_s^i.$$

To abbreviate notation we set $\bar{\mathcal{A}}_n = \mathcal{A}^1 \times \dots \times \mathcal{A}^n$.

We suppose that each player aims at maximizing the probability of her own final state to be greater than or equal to the empirical $(1 - \alpha)$ -quantile of all players' final states. More precisely, for $H = (H^1, \dots, H^n) \in \bar{\mathcal{A}}_n$ let

$$\mu^{n, H} = \frac{1}{n} \sum_{i=1}^n \delta_{X_{\tau_i}^{i, H^i}} \quad (1)$$

be the empirical distribution of the players' final states. We define $q(\mu^{n, H}, 1 - \alpha) = \inf\{r \in \mathbb{R} : \mu^{n, H}((-\infty, r]) \geq 1 - \alpha\}$. Each player wants to maximize the probability $P(X_{\tau_i}^{i, H^i} \geq q(\mu^{n, H}, 1 - \alpha))$. Note that $X_{\tau_i}^{i, H^i} \geq q(\mu^{n, H}, 1 - \alpha)$ if and only if not more than α of all final states are greater than the final state of player i .

We stress that the assumptions imply that no player can observe the other players' strategy processes while controlling her own state process X^i . Hence, the other players' states are not observable, which mathematically means that for any $t \geq 0$ and i, j with $i \neq j$ the sigma-field $\sigma(X_s^j : s \leq t)$ is not contained in \mathcal{F}_t^i . The non-observability of actions means that one can interpret the game as a static game with a single period only: at time zero each player i chooses a strategy H^i and thus determines the distribution of her final state; at time one each player draws independently from the chosen distribution and the best α are rewarded with the prize.

It is standard to predict or explain the players' behaviour in terms of a Nash equilibrium, which here is defined as follows.

Definition 1.1. A tuple $H = (H^1, \dots, H^n) \in \bar{\mathcal{A}}_n$ is called a Nash equilibrium if for all $i \in \{1, \dots, n\}$ we have

$$P(X_{\tau_i}^{i, H^i} \geq q(\mu^{n, H}, 1 - \alpha)) \geq \sup_{G \in \mathcal{A}^i} P(X_{\tau_i}^{i, G} \geq q(\mu^{n, (H_{-i}, G)}, 1 - \alpha)),$$

where $(H_{-i}, G) = (H^1, \dots, H^{i-1}, G, H^{i+1}, \dots, H^n)$.

Notice that the expectation of any player's outcome is always zero, whatever control is chosen. However, other distributional properties such as the variance, the skewness and the position of quantiles do depend on the strategy. For example, a larger H^i leads to a larger variance of the final outcome $X_{\tau_i}^i$. Moreover, if player i chooses large values for H^i when her state is small and small values when the state is large, then the distribution of $X_{\tau_i}^i$ has a negative skew. In particular, a state-dependent choice of the fluctuation intensity allows to influence the position of quantiles, and hence the probability of receiving the reward. So the game helps in understanding how a rank-based reward determines the distributional properties of the players' states and by which dynamic strategies they can be implemented.

The discontinuous payoff and the general information structure make it difficult to compute an exact Nash equilibrium for the game, or even to show existence of an equilibrium by using abstract arguments. Our idea is to compute an approximate Nash equilibrium by considering the mean-field limit of the game. The mean-field game approach, introduced by [19], [20], [21] and [15], [14], is based on the observation that in a symmetric game with many similar players, the empirical distribution of the realized players' states can be approximated with the state distribution of a *single* player. Thus, the Nash equilibrium with many players can be approximated by solving a single agent control problem where the reward depends on the distribution of the state.

It turns out that the mean-field version of the game described above possesses an explicit equilibrium strategy (see Section 2). This allows further to show that the symmetric n -tuple is an approximate Nash equilibrium of the n -player game (see Section 3). More precisely, the difference of expected payoff under the approximate equilibrium strategies and the best response payoff converges to zero as the number of players n tends to infinity.

Results and comparison to the literature

We obtain that, within the mean-field limit version of the game, there exists a symmetric equilibrium that consists in choosing the maximal fluctuation intensity when the state is below a given threshold, and the minimal intensity else. We refer to such a strategy as a threshold strategy. The threshold of the mean-field equilibrium strategy is chosen such that it coincides with the $(1 - \alpha)$ -quantile of the resulting final state distribution. The threshold can be

identified as a fixed point of a mapping from the set of threshold strategies to itself. Fixed point arguments are a common tool for determining equilibria in mean-field games (see, e.g., [6] and [7]).

Controlling the state with a threshold strategy entails that the state dynamics are a so-called oscillating Brownian motion (OBM), a process introduced in [17] and studied in [24] from a control theory perspective without interaction. The distribution of OBM is explicitly known at deterministic times. This allows us to compute the threshold of the mean-field equilibrium in closed form.

In the mean-field equilibrium the distribution of the final state is negatively skewed and the $(1 - \alpha)$ -quantile is farer to the right than when using a strategy with constant fluctuations. Choosing the maximal fluctuation intensity below the threshold increases risk, measured, e.g., in terms of the variance of the final state. Thus, the results show that the rank-based reward incentivizes players to introduce a negative skew, and a high variance.

We do not compute an exact Nash equilibrium for the n player game. We believe that it is far more difficult to determine an exact equilibrium than an approximate one based on a mean-field limit. Indeed, one can show that already the 2 player version of the game does not, in general, possess a symmetric Nash equilibrium in threshold strategies; a counter-example will be presented in the dissertation of the third author.

In general, games with rank-based rewards are hard to analyze because of discontinuities in the reward function. A workaround is to consider rewards depending continuously on the average state of all players. Espinosa and Touzi [11] consider a game where each player aims at maximizing the expected utility of her income depending on the deviation the player's state from the average state. They characterize a Nash equilibrium in terms of a system of coupled backward stochastic differential equations (BSDEs). [10] establishes a Nash equilibrium in a market with N agents with the performance criteria of relative wealth level when the mean return rate is unobservable. Each investor has a heterogeneous prior belief on the return rate of the risky asset. By a separation result and a martingale argument, they show that the optimal investment strategy under this setting can be characterized by a fully-coupled linear FBSDE system. Two sets of deep neural networks are used for first computing each investor's estimate of the mean return rate and then numerically solving the FBSDEs. [13] uses a probabilistic approach to study an N -player exponential utility game and a mean field exponential utility game. Each player manages two stocks; one is driven by an individual shock and the other is driven by a common shock. Moreover, each player is concerned not only with her own terminal wealth but also with the relative performance of her competitors. A multi-dimensional FBSDE with quadratic growth (a mean-field FBSDEs, respectively) is derived to characterize an equilibrium of the N -player game (of the mean field game, respectively).

The idea to use mean-field limits for analyzing differential games with rank-based rewards has been employed already in the articles [3],[4], [27] and [5]. All these articles consider versions of a rank-based mean-field game, where agents influence the drift rate of the state, but not the diffusion rate.

The articles [3], [4], [27] provide conditions guaranteeing that the mean-field games approximate the corresponding n -player game for large n ; and they also determine the rate at which the expected reward of the mean-field approximation and the best response converge to each other, as $n \rightarrow \infty$. Motivated by these articles, in Section 3 we also compute a convergence rate for the game at hand. In the recent paper [2], the authors analyze a mean field game where the diffusion coefficient can be controlled and show that the solution of the problem is characterized by a McKean-Vlasov second order backward SDE. Note that it is not possible to directly apply the results of [2] to our setting, since the particular form of our criteria given as a probability in Definition 1.1 would make the terminal condition of the related second order BSDE discontinuous, which is not allowed in [2].

2. The mean-field approximation

In this section we first describe the mean-field version of the game introduced in Section 1. We then prove that there exists an equilibrium in closed form.

Let B be a Brownian motion on a complete probability space (Ω, \mathcal{F}, P) and τ be an exponentially distributed random variable with parameter $\lambda > 0$ independent from B . Furthermore, let $(\mathcal{F}_t)_{t \geq 0}$ be the augmented Brownian filtration, generated by B and augmented by the P -null sets in \mathcal{F} .

We denote by \mathcal{M} the set of all progressively measurable control processes $(\beta_t)_{t \geq 0}$ taking values in $[\sigma_1, \sigma_2]$. Given that an agent chooses the control function $\beta \in \mathcal{M}$, her state process is given by $X_t^\beta := \int_0^t \beta_s dB_s$.

Remark 2.1. All feedback controls with a feedback function $m : \mathbb{R} \rightarrow [\sigma_1, \sigma_2]$ of bounded variation are contained in \mathcal{M} . Indeed, since m is uniformly bounded away from zero there exists a weak solution to the SDE

$$dX_t = m(X_t)dB_t, \quad X_0 = 0, \quad (2)$$

according to Theorem 2.6.1 in [18], and pathwise uniqueness applies according to results in [26]. Hence, there exists a unique strong solution X^m to (2) (cf. section 5.3 in [16]), and $(m(X_t^m))_{t \geq 0}$ belongs to \mathcal{M} .

Let $\alpha \in (0, 1)$ and denote by $q(X_\tau^\beta, 1 - \alpha)$ the $(1 - \alpha)$ -quantile of the random variable X_τ^β , i.e. $q(X_\tau^\beta, 1 - \alpha) = \inf\{r \in \mathbb{R} : P(X_\tau^\beta \leq r) \geq 1 - \alpha\} = \inf\{r \in \mathbb{R} : P(X_\tau^\beta \geq r) \leq \alpha\}$.

For all $\beta \in \mathcal{M}$ the random variable X_τ^β is absolutely continuous w.r.t. the Lebesgue measure on \mathbb{R} , see Proposition A.4 below. Hence,

$$P(X_\tau^\beta \geq q(X_\tau^\beta, 1 - \alpha)) = \alpha.$$

Definition 2.2. We call $\beta^* \in \mathcal{M}$ an equilibrium strategy of the mean-field game if

$$P(X_\tau^{\beta^*} \geq q(X_\tau^{\beta^*}, 1 - \alpha)) = \max_{\beta \in \mathcal{M}} P(X_\tau^\beta \geq q(X_\tau^\beta, 1 - \alpha)) \quad (3)$$

Before proving that there exists an equilibrium strategy, we first show that it is enough to consider control processes of feedback type, that are constant equal to the minimal volatility σ_1 if the state is greater than or equal to a given threshold b , and that are constant equal to the maximal volatility σ_2 else. We refer to such controls as threshold controls. More precisely, the threshold control with barrier $b \in \mathbb{R}$ is defined by

$$m_b(x) = \begin{cases} \sigma_1, & \text{if } x \geq b, \\ \sigma_2, & \text{else.} \end{cases}$$

In the following we simply write X^b for the state process X^{m_b} . Note that $(m_b(X_t^b))_{t \geq 0} \in \mathcal{M}$ since m_b is of bounded variation.

We have the following result:

Lemma 2.3. *Let $b \in \mathbb{R}$. Then*

$$P(X_\tau^b \geq b) = \max_{\beta \in \mathcal{M}} P(X_\tau^\beta \geq b). \quad (4)$$

Proof. This follows from [24], Remark 8. Alternatively, see Section 6 in [1]. \square

Lemma 2.3 shows that the strategy maximizing the probability of being above $q(X_\tau^\beta, 1 - \alpha)$ at time τ is the threshold control with threshold $q(X_\tau^\beta, 1 - \alpha)$. The following theorem provides an explicit equilibrium strategy.

Theorem 2.4. *The threshold strategy m_{b^*} with threshold*

$$b^* = \begin{cases} -\frac{\sigma_2}{\sqrt{2\lambda}} \log\left(\frac{\sigma_1 + \sigma_2}{\sigma_2} \alpha\right), & \text{if } \alpha < \frac{\sigma_2}{\sigma_1 + \sigma_2}, \\ \frac{\sigma_2}{\sqrt{2\lambda}} \log\left(\frac{\sigma_1 + \sigma_2}{\sigma_1} (1 - \alpha)\right), & \text{if } \alpha \geq \frac{\sigma_2}{\sigma_1 + \sigma_2}, \end{cases} \quad (5)$$

is an equilibrium strategy for the mean-field game. It holds $q(X_\tau^{b^}, 1 - \alpha) = b^*$.*

Notice that in order to prove Theorem 2.4 it is enough to find a fixed point of the function

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad b \mapsto q(X_\tau^b, 1 - \alpha).$$

Indeed, if $f(b) = b$, then $b = q(X_\tau^b, 1 - \alpha)$. Lemma 2.3 further implies that $P(X_\tau^b \geq q(X_\tau^b, 1 - \alpha)) = \max_{\beta \in \mathcal{M}} P(X_\tau^\beta \geq q(X_\tau^b, 1 - \alpha))$, and hence that m_b is an equilibrium strategy.

For threshold controls, the controlled state process is a so-called oscillating Brownian motion (OBM). Note that the transition densities of an OBM are explicitly known. For the reader's convenience we recall the following result.

Lemma 2.5. *For all $x \in \mathbb{R}$ let Y^x be the solution of the SDE $dY_t = m_0(Y_t)dB_t$ with initial condition $Y_0 = x$. Then, for all $t > 0$, the random variable Y_t^x has a density $p(t, x, y)$ with respect to the Lebesgue measure, given by*

$$p(t, x, y) = \begin{cases} \frac{2\sigma_1}{\sigma_2(\sigma_1 + \sigma_2)} \frac{1}{\sqrt{2\pi t}} e^{-\left(\frac{x}{\sigma_1} - \frac{y}{\sigma_2}\right)^2 \frac{1}{2t}}, & \text{for } x \geq 0, y < 0, \\ \frac{2\sigma_2}{\sigma_1(\sigma_1 + \sigma_2)} \frac{1}{\sqrt{2\pi t}} e^{-\left(\frac{y}{\sigma_1} - \frac{x}{\sigma_2}\right)^2 \frac{1}{2t}}, & \text{for } x < 0, y \geq 0, \\ \frac{1}{\sigma_1 \sqrt{2\pi t}} \left(e^{-\frac{(y-x)^2}{2\sigma_1^2 t}} + \frac{\sigma_2 - \sigma_1}{\sigma_1 + \sigma_2} e^{-\frac{(y+x)^2}{2\sigma_1^2 t}} \right), & \text{for } x \geq 0, y \geq 0, \\ \frac{1}{\sigma_2 \sqrt{2\pi t}} \left(e^{-\frac{(y-x)^2}{2\sigma_2^2 t}} + \frac{\sigma_1 - \sigma_2}{\sigma_1 + \sigma_2} e^{-\frac{(y+x)^2}{2\sigma_2^2 t}} \right), & \text{for } x < 0, y < 0. \end{cases}$$

Proof. This follows from Theorem 1 in Keilson, Wellner [17]; see also Theorem 4 in [24] and Section 2 in Lejay, Pigato [22] (beware of the typo in the last case). \square

As mentioned above, for any $b \in \mathbb{R}$ the state process X^b controlled by the threshold control m_b is just an OBM. Using Lemma 2.5 we observe that X_t^b possesses the probability density function $p(t, -b, x - b)$.

We now turn to the proof of Theorem 2.4.

Proof of Theorem 2.4. Suppose first that $\alpha \in (0, \frac{\sigma_2}{\sigma_1 + \sigma_2})$. Using Lemma 2.5 we have for all $b > 0$

$$\begin{aligned} P(X_\tau^b \geq b) &= P(Y_\tau^{-b} \geq 0) = \int_0^\infty \int_0^\infty \lambda e^{-\lambda t} p(t, -b, y) dy dt \\ &= \frac{2\sigma_2 \lambda}{\sigma_1(\sigma_1 + \sigma_2)} \int_0^\infty e^{-\lambda t} \int_0^\infty \frac{1}{\sqrt{2\pi t}} e^{-\left(\frac{y}{\sigma_1} + \frac{b}{\sigma_2}\right)^2 \frac{1}{2t}} dy dt \\ &= \frac{2\sigma_2 \lambda}{\sigma_1 + \sigma_2} \int_0^\infty e^{-\lambda t} \Phi\left(-\frac{b}{\sigma_2 \sqrt{t}}\right) dt \\ &= \frac{2\sigma_2 \lambda}{\sigma_1 + \sigma_2} \mathcal{L}\left[\Phi\left(-\frac{b}{\sigma_2 \sqrt{\cdot}}\right)\right](\lambda), \end{aligned}$$

where \mathcal{L} denotes the Laplace transformation. Notice that $\Phi\left(-\frac{b}{\sigma_2 \sqrt{t}}\right) = \frac{1}{2} \operatorname{erfc}\left(\frac{b}{\sigma_2 \sqrt{2t}}\right)$ and $\mathcal{L}\left[\operatorname{erfc}\left(\frac{b}{\sigma_2 \sqrt{2\cdot}}\right)\right](\lambda) = \frac{1}{\lambda} e^{-\frac{b}{\sigma_2} \sqrt{2\lambda}}$ (see e.g. chapter 8 in [12]). Hence,

$$P(X_\tau^b \geq b) = \frac{\sigma_2}{\sigma_1 + \sigma_2} e^{-\frac{b}{\sigma_2} \sqrt{2\lambda}}.$$

The right-hand side of the previous equation is smaller than or equal to $\frac{\sigma_2}{\sigma_1 + \sigma_2}$. Moreover, $P(X_\tau^b \geq b) = \alpha$ if and only if

$$b = -\frac{\sigma_2}{\sqrt{2\lambda}} \log \left(\frac{\sigma_1 + \sigma_2}{\sigma_2} \alpha \right).$$

In this case, $q(X_\tau^b, 1 - \alpha) = b$; in other words b is a fixed point of the mapping f .

Now suppose that $\alpha \in [\frac{\sigma_2}{\sigma_1 + \sigma_2}, 1)$. Again with Lemma 2.5 we can conclude that for every $b \leq 0$ we have as above

$$\begin{aligned} P(X_\tau^b \geq b) &= P(Y_\tau^{-b} \geq 0) = \int_0^\infty \int_0^\infty \lambda e^{-\lambda t} p(t, -b, y) dy dt \\ &= \lambda \int_0^\infty e^{-\lambda t} \frac{1}{\sqrt{2\pi\sigma_1^2 t}} \left(\int_0^\infty \left(e^{-\frac{(y+b)^2}{2\sigma_1^2 t}} + \frac{\sigma_2 - \sigma_1}{\sigma_1 + \sigma_2} e^{-\frac{(y-b)^2}{2\sigma_1^2 t}} \right) dy \right) dt \\ &= \lambda \int_0^\infty e^{-\lambda t} \left(\Phi \left(-\frac{b}{\sigma_2 \sqrt{t}} \right) + \frac{\sigma_2 - \sigma_1}{\sigma_1 + \sigma_2} \Phi \left(\frac{b}{\sigma_2 \sqrt{t}} \right) \right) dt \\ &= \lambda \int_0^\infty e^{-\lambda t} \left(1 - \frac{2\sigma_1}{\sigma_1 + \sigma_2} \Phi \left(\frac{b}{\sigma_2 \sqrt{t}} \right) \right) dt \\ &= 1 - \frac{2\sigma_1 \lambda}{\sigma_1 + \sigma_2} \mathcal{L} \left[\Phi \left(\frac{b}{\sigma_2 \sqrt{\cdot}} \right) \right] (\lambda) \\ &= 1 - \frac{\sigma_1}{\sigma_1 + \sigma_2} e^{\frac{b}{\sigma_2} \sqrt{2\lambda}}. \end{aligned}$$

Since $b \leq 0$, the right-hand side of the previous inequality is greater than or equal to $\frac{\sigma_2}{\sigma_1 + \sigma_2}$. Moreover, $P(X_\tau^b \geq b) = \alpha$ if and only if

$$b = \frac{\sigma_2}{\sqrt{2\lambda}} \log \left(\frac{\sigma_1 + \sigma_2}{\sigma_1} (1 - \alpha) \right).$$

□

Remark 2.6. We briefly analyze how the threshold (5) of the equilibrium strategy depends on the model parameters.

a) First notice that b^* increases as α decreases, and

$$\lim_{\alpha \downarrow 0} b^* = \infty, \quad \lim_{\alpha \uparrow 1} b^* = -\infty.$$

Thus, reducing the proportion α of best players who receive a reward implies that the players choose a higher threshold up to which they control their process with the maximal diffusion rate. In other words, intensifying competition leads to riskier strategies.

This observation is in line with results in [30] and [31], obtained within models, where agents can control the time horizon of comparison. It is

shown that as the number of agents increases and hence the proportion of agents receiving a reward decreases, the agents choose in expectation larger stopping rules and hence increase the standard deviation of the states at the comparison time.

b) Note that b^* is proportional to $\sqrt{E[\tau]} = 1/\sqrt{\lambda}$, and hence to the standard deviation of the final state $X_\tau^{b^*}$. If $\alpha < \frac{\sigma_2}{\sigma_1 + \sigma_2}$, then increasing the run time of the game encourages players to choose higher thresholds.

3. Approximate Nash equilibrium of the n player game

We now come back to the n -player game version of Section 1. We show that the n -tuple of the mean-field equilibrium threshold strategy is an ε -approximate equilibrium in the n -player game, with ε converging to zero as n tends to infinity.

We first rigorously define what we mean by an ε -Nash equilibrium for the game at hand (cf., e.g., the similar definition in [4]).

Definition 3.1. Let $\varepsilon > 0$. A tuple $H = (H^1, \dots, H^n) \in \bar{\mathcal{A}}_n$ is called ε -Nash equilibrium of the n -player game if for all $i \in \{1, \dots, n\}$ we have

$$P(X_{\tau_i}^{i,H^i} \geq q(\mu^{n,H}, 1 - \alpha)) + \varepsilon \geq \sup_{G \in \mathcal{A}^i} P(X_{\tau_i}^{i,G} \geq q(\mu^{n,(H_{-i},G)}, 1 - \alpha)),$$

where (H_{-i}, G) , $\mu^{n,H}$ and $\mu^{n,(H_{-i},G)}$ are defined as in Section 1.

In the following we denote by a_i the threshold strategy of player i with threshold b^* defined as in (5). More precisely, $a_i = (m_{b^*}(X_t^{i,*}))_{t \geq 0}$, where $(X_t^{i,*})$ is the strong solution of $dX_t = m_{b^*}(X_t)dW_t^i$, $X_0 = 0$.

Moreover, let μ be the distribution of the equilibrium state $X_\tau^{b^*}$ of the mean-field game and denote by F_μ its distribution function. Note that for each player i the final state $X_{\tau_i}^{i,*}$ has also distribution μ .

Finally, let \mathcal{W} denote the Lambert W function, i.e. \mathcal{W} is defined as the unique inverse function of the mapping $[0, \infty) \rightarrow \mathbb{R}, x \mapsto xe^x$.

Our main result in this section is as follows:

Theorem 3.2. *There exists a sequence $\varepsilon_n \geq 0$ with $\lim_{n \rightarrow \infty} \varepsilon_n = 0$ such that $a = (a_1, \dots, a_n)$ is an ε_n -Nash equilibrium of the n -player game. We can choose $\varepsilon_n \in \mathcal{O}(\sqrt{\mathcal{W}(n)}n^{-\frac{1}{2}})$.*

We remark that Theorem 3.2 is formulated in a way similar to Theorem 4.2 of [8]. In order to prove Theorem 3.2 we recall the following result for the approximation of sample quantiles (see e.g. Lemma 4.3 in [28] or [23, p. 468]). This result is also known as *Smirnov's lemma*. We state it in a version adapted to our setting.

Lemma 3.3. *Let $a = (a_1, \dots, a_n)$ and $(\alpha_n)_{n \in \mathbb{N}}$ be a sequence with $\alpha_n \in (0, 1)$ and $\lim_{n \rightarrow \infty} \sqrt{n}(\alpha_n - \alpha) = 0$. Then there exists a constant $C > 0$ such that for all $n \geq 1$*

$$\sup_{x \in \mathbb{R}} \left| P(q(\mu^{n,a}, 1 - \alpha_n) \leq x) - \Phi \left(\frac{\sqrt{n} F_\mu(x) - (1 - \alpha)}{\sqrt{\alpha(1 - \alpha)}} \right) \right| \leq \frac{C}{\sqrt{n}}.$$

Proof. The random variables $X_{\tau_1}^{1,*}, \dots, X_{\tau_n}^{n,*}$ are independent and identically distributed with distribution function F_μ , because each strategy a_i is based on the feedback function m_{b^*} applied to the state i and the times τ_1, \dots, τ_n are independent and identically distributed. Hence, we can apply Lemma 4.3 in [28] and obtain the result. \square

As a consequence of Lemma 3.3 we obtain the following result:

Lemma 3.4. *Let $a = (a_1, \dots, a_n)$ and $(\alpha_n)_{n \in \mathbb{N}}$ be a sequence with $\alpha_n \in (0, 1)$ and $\lim_{n \rightarrow \infty} \sqrt{n}(\alpha_n - \alpha) = 0$. Then there exists a $C > 0$ such that*

$$P \left(|q(\mu^{n,a}, 1 - \alpha_n) - q(\mu, 1 - \alpha)| > C \sqrt{\frac{\mathcal{W}(n)}{n}} \right) \in \mathcal{O} \left(\frac{1}{\sqrt{n}} \right).$$

We postpone the proof of this lemma to the Appendix and now prove Theorem 3.2.

Proof of Theorem 3.2. We compare $a = (a_1, \dots, a_n)$ with the tuple where player i deviates from a_i by choosing a strategy $G \in \mathcal{A}^i$. To this end, we write $\tilde{a} = (a_{-i}, G) = (a_1, \dots, a_{i-1}, G, a_{i+1}, \dots, a_n)$.

Then we have

$$P(X_{\tau_i}^{i,G} \geq q(\mu^{n,\tilde{a}}, 1 - \alpha)) - P(X_{\tau_i}^{i,*} \geq q(\mu^{n,a}, 1 - \alpha)) = A_n + B_n + C_n, \quad (6)$$

where

$$\begin{aligned} A_n &:= P(X_{\tau_i}^{i,G} \geq q(\mu^{n,\tilde{a}}, 1 - \alpha)) - P(X_{\tau_i}^{i,G} \geq q(\mu, 1 - \alpha)), \\ B_n &:= P(X_{\tau_i}^{i,G} \geq q(\mu, 1 - \alpha)) - P(X_{\tau_i}^{i,*} \geq q(\mu, 1 - \alpha)), \\ C_n &:= P(X_{\tau_i}^{i,*} \geq q(\mu, 1 - \alpha)) - P(X_{\tau_i}^{i,*} \geq q(\mu^{n,a}, 1 - \alpha)). \end{aligned}$$

Note that $q(\mu, 1 - \alpha) = b^*$ and that a_i is the control in \mathcal{A}^i maximizing the probability to be greater than or equal to b^* at time τ_i . Therefore, $P(X_{\tau_i}^{i,G} \geq q(\mu, 1 - \alpha)) \leq P(X_{\tau_i}^{i,*} \geq q(\mu, 1 - \alpha))$ (see Lemma 2.3), and hence

$$B_n \leq 0. \quad (7)$$

We next show that C_n converges to zero. To this end, we use that the sequence $X_{\tau_1}^{1,*}, \dots, X_{\tau_n}^{n,*}$ is i.i.d., and hence that $\mu^{n,a}$ converges weakly to μ . Moreover, due to Lemma 3.4 there exists a constant $K_1 > 0$ such that

$$P \left(|q(\mu^{n,a}, 1 - \alpha) - q(\mu, 1 - \alpha)| > K_1 \sqrt{\frac{\mathcal{W}(n)}{n}} \right) \in \mathcal{O} \left(\frac{1}{\sqrt{n}} \right). \quad (8)$$

Therefore, we can choose two sequences $u_n := q(\mu, 1 - \alpha) + K_1 \sqrt{\frac{\mathcal{W}(n)}{n}}$ and $l_n := q(\mu, 1 - \alpha) - K_1 \sqrt{\frac{\mathcal{W}(n)}{n}}$ controlling the speed of convergence. Note that we can write

$$|C_n| \leq P(q(\mu^{n,a}, 1 - \alpha) \notin [l_n, u_n]) + P(X_{\tau_i}^{i,*} \in [l_n, u_n]). \quad (9)$$

Lemma A.3 below implies that we can estimate the speed of convergence of $P(X_{\tau_i}^{i,*} \in [l_n, u_n])$. In particular, there exists a constant $K_2 > 0$ such that for n sufficiently large

$$P(X_{\tau_i}^{i,*} \in [l_n, u_n]) \leq K_2(u_n - l_n) = 2K_1K_2\sqrt{\frac{\mathcal{W}(n)}{n}}. \quad (10)$$

Thus, (9) together with (8) and (10) implies

$$|C_n| \in \mathcal{O}\left(\sqrt{\frac{\mathcal{W}(n)}{n}}\right). \quad (11)$$

The convergence of A_n to zero as $n \rightarrow \infty$ follows from similar arguments. Notice that the state $X_{\tau_i}^{i,G}$ of player i is not necessarily independent of the other players' states $X_{\tau_j}^{j,*}, j \neq i$, and can also have a different distribution. We can, however, replace the quantile $q(\mu^{n,\bar{a}}, 1 - \alpha)$ by a quantile that does not depend on G . To this end, let $\mu^{n-1,a} := \frac{1}{n-1} \sum_{j \neq i} \delta_{X_{\tau_j}^{j,*}}$ be the empirical measure without player i , and observe that

$$q(\mu^{n,\bar{a}}, 1 - \alpha) \geq q\left(\mu^{n-1,a}, 1 - \frac{n}{n-1}\alpha\right), \quad n \geq 2, \quad (12)$$

where we set $q(\mu^{n-1,a}, 1 - \frac{n}{n-1}\alpha) := -\infty$ if $\frac{n}{n-1}\alpha \geq 1$. (Proof of (12): Suppose that $r \geq q(\mu^{n,\bar{a}}, 1 - \alpha)$. Then at most αn players are better than r , or equivalently, at least $(1 - \alpha)n$ players are worse than r . In the $(n - 1)$ -player game without player i then also at most αn players are better than r , or equivalently, at least $(1 - \alpha)n - 1$ players are worse than r . This means $\mu^{n-1,a}((-\infty, r]) \geq 1 - \frac{n}{n-1}\alpha$ and $r \geq q\left(\mu^{n-1,a}, 1 - \frac{n}{n-1}\alpha\right)$. Choosing $r = q(\mu^{n,\bar{a}}, 1 - \alpha)$ yields the assertion.)

Since the sequence $(X_{\tau_j}^{j,*})_{j \neq i}$ is i.i.d., the empirical measure $\mu^{n-1,a}$ converges weakly to μ . Then Lemma 3.4, with $\alpha_n := \frac{1}{n} + \mathbb{1}_{(0,1)}\left(\frac{n}{n-1}\alpha\right)\left(\frac{n}{n-1}\alpha - \frac{1}{n}\right)$ for $n \geq 2$, implies that there is a K_3 such that

$$P\left(\left|q\left(\mu^{n-1,a}, 1 - \frac{n}{n-1}\alpha\right) - q(\mu, 1 - \alpha)\right| > K_3\sqrt{\frac{\mathcal{W}(n-1)}{n-1}}\right) \in \mathcal{O}\left(\sqrt{\frac{\mathcal{W}(n)}{n}}\right), \quad (13)$$

as $n \rightarrow \infty$. Now, consider again the two sequences $(u_n)_{n \in \mathbb{N}}$ and $(l_n)_{n \in \mathbb{N}}$ defined as above by replacing the constant K_1 with K_3 . We obtain with (12) that, for $n \geq 2$,

$$\begin{aligned} A_n &\leq P\left(X_{\tau_i}^{i,G} \geq q\left(\mu^{n-1,a}, 1 - \frac{n}{n-1}\alpha\right)\right) - P\left(X_{\tau_i}^{i,G} \geq q(\mu, 1 - \alpha)\right) \\ &\leq P\left(q\left(\mu^{n-1,a}, 1 - \frac{n}{n-1}\alpha\right) \notin [l_{n-1}, u_{n-1}]\right) + P\left(X_{\tau_i}^{i,G} \in [l_{n-1}, u_{n-1}]\right). \end{aligned} \quad (14)$$

To determine the convergence rate of the second term on the right-hand side of (14), note that we have to find a bound uniform in G . By Lemma A.3 below we have

$$P\left(X_{\tau_i}^{i,G} \in [l_{n-1}, u_{n-1}]\right) \leq K_4(u_{n-1} - l_{n-1}) = K_4 \sqrt{\frac{\mathcal{W}(n-1)}{n-1}}$$

The bound on the right-hand side is independent of the control G and converges to zero as $n \rightarrow \infty$. Therefore, we obtain with (13) and (14) that

$$\sup_{G \in \mathcal{A}^i} A_n \in \mathcal{O}\left(\sqrt{\frac{\mathcal{W}(n)}{n}}\right). \quad (15)$$

Finally, by combining (6) with (7), (9) and (15) we arrive at

$$\sup_{G \in \mathcal{A}^i} \left(P\left(X_{\tau_i}^{i,G} \geq q\left(\mu^{n,\tilde{a}}, 1 - \alpha\right)\right)\right) \leq P\left(X_{\tau_i}^{i,*} \geq q\left(\mu^{n,a}, 1 - \alpha\right)\right) + K_5 \sqrt{\frac{\mathcal{W}(n)}{n}},$$

for some constant $K_5 > 0$, and hence the proof is complete. \square

Remark 3.5. The choice of the convergence rate $\sqrt{\frac{\mathcal{W}(n)}{n}}$ in the proof of Theorem 3.2 is indeed optimal. Firstly, in Lemma 3.4 we cannot expect a better convergence rate than the Berry-Esseen rate $\mathcal{O}(n^{-\frac{1}{2}})$. Secondly, any rate faster than $\sqrt{\mathcal{W}(n)}$ in the definition of $(u_n)_{n \in \mathbb{N}}$ and $(l_n)_{n \in \mathbb{N}}$, slows the convergence of those sequences down. Lastly, any rate slower than $\sqrt{\mathcal{W}(n)}$ results in a faster convergence rate for $(u_n)_{n \in \mathbb{N}}$ and $(l_n)_{n \in \mathbb{N}}$, but slows the convergence of the first term on the right-hand side of (9) down (see proof of Lemma 3.4).

A. Appendix

A.1. Proof of Lemma 3.4

Proof of Lemma 3.4. Let F_μ denote the cumulative distribution function of the equilibrium measure $\mu = P_{X_t^{b^*}}$. Note that $X_t^{b^*}$ has the probability density

$f^*(t, x) := p(t, -b^*, x - b^*)$, with p denoting the probability density function of the OBM, given in Lemma 2.5. More explicitly, $f^*(t, \cdot)$ is given by

$$f^*(t, x) = \begin{cases} \frac{2\sigma_1}{\sigma_2(\sigma_1 + \sigma_2)} \frac{1}{\sqrt{2\pi t}} e^{-\left(\frac{x}{\sigma_2} + b^* \left(\frac{1}{\sigma_1} - \frac{1}{\sigma_2}\right)\right)^2 \frac{1}{2t}}, & \text{for } b^* \leq 0, x < b^*, \\ \frac{1}{\sigma_1 \sqrt{2\pi t}} \left(e^{-\frac{x^2}{2\sigma_1^2 t}} + \frac{\sigma_2 - \sigma_1}{\sigma_1 + \sigma_2} e^{-\frac{(x - 2b^*)^2}{2\sigma_1^2 t}} \right), & \text{for } b^* \leq 0, x \geq b^*, \\ \frac{1}{\sigma_2 \sqrt{2\pi t}} \left(e^{-\frac{x^2}{2\sigma_2^2 t}} + \frac{\sigma_1 - \sigma_2}{\sigma_1 + \sigma_2} e^{-\frac{(x - 2b^*)^2}{2\sigma_2^2 t}} \right), & \text{for } b^* > 0, x < b^*, \\ \frac{2\sigma_2}{\sigma_1(\sigma_1 + \sigma_2)} \frac{1}{\sqrt{2\pi t}} e^{-\left(\frac{x}{\sigma_1} + b^* \left(\frac{1}{\sigma_2} - \frac{1}{\sigma_1}\right)\right)^2 \frac{1}{2t}}, & \text{for } b^* > 0, x \geq b^*, \end{cases} \quad t > 0.$$

At first, we show that f^* is locally bounded from below at the point $x = b^* = q(\mu, 1 - \alpha)$, and that one can bound the difference $F_\mu(y) - F_\mu(x)$: Fix $t > 0$. We consider the two cases $b^* \leq 0$ and $b^* > 0$.

1. case $b^* \leq 0$: Let $\varepsilon > 0$. Then we have that for $b^* - \varepsilon \leq x < b^*$

$$f^*(t, x) \geq \frac{2\sigma_1}{\sigma_2(\sigma_1 + \sigma_2)} \frac{1}{\sqrt{2\pi t}} e^{-\left(\frac{b^* - \varepsilon}{\sigma_1} - \frac{\varepsilon}{\sigma_2}\right)^2 \frac{1}{2t}} \geq \frac{2\sigma_1}{\sigma_2(\sigma_1 + \sigma_2)} \frac{1}{\sqrt{2\pi t}} e^{-\frac{(b^* - \varepsilon)^2}{2\sigma_1^2 t}},$$

and for $b^* \leq x \leq b^* + \varepsilon$

$$f^*(t, x) \geq \frac{2\sigma_2}{\sigma_1 + \sigma_2} \frac{1}{\sigma_1 \sqrt{2\pi t}} e^{-\frac{(-b^* + \varepsilon)^2}{2\sigma_1^2 t}} \geq \frac{2\sigma_1}{\sigma_2(\sigma_1 + \sigma_2)} \frac{1}{\sqrt{2\pi t}} e^{-\frac{(b^* - \varepsilon)^2}{2\sigma_1^2 t}}.$$

2. case $b^* > 0$: Let $\varepsilon > 0$. Then we have for $b^* - \varepsilon \leq x < b^*$

$$f^*(t, x) \geq \frac{1}{\sigma_2 \sqrt{2\pi t}} \left(e^{-\frac{(b^*)^2}{2\sigma_2^2 t}} - \frac{\sigma_2 - \sigma_1}{\sigma_1 + \sigma_2} e^{-\frac{(b^*)^2}{2\sigma_2^2 t}} \right) \geq \frac{2\sigma_1}{\sigma_2(\sigma_1 + \sigma_2)} \frac{1}{\sqrt{2\pi t}} e^{-\frac{(b^* + \varepsilon)^2}{2\sigma_1^2 t}},$$

and for $b^* \leq x \leq b^* + \varepsilon$

$$f^*(t, x) \geq \frac{2\sigma_2}{\sigma_1(\sigma_1 + \sigma_2)} \frac{1}{\sqrt{2\pi t}} e^{-\left(\frac{b^* + \varepsilon}{\sigma_2} + \frac{\varepsilon}{\sigma_1}\right)^2 \frac{1}{2t}} \geq \frac{2\sigma_1}{\sigma_2(\sigma_1 + \sigma_2)} \frac{1}{\sqrt{2\pi t}} e^{-\frac{(b^* + \varepsilon)^2}{2\sigma_1^2 t}}.$$

Summing up, we see that if $b^* \leq 0$ we have for any $\varepsilon > 0$ and $x, y \in [b^* - \varepsilon, b^* + \varepsilon], x \leq y$,

$$\begin{aligned} F_\mu(y) - F_\mu(x) &= \int_0^\infty \lambda e^{-\lambda t} (P(X_t^{b^*} \leq y) - P(X_t^{b^*} \leq x)) dt \\ &= \int_0^\infty \lambda e^{-\lambda t} \int_x^y f^*(t, z) dz dt \\ &\geq \frac{2\sigma_1 \lambda}{\sigma_2(\sigma_1 + \sigma_2)} \frac{y - x}{\sqrt{2\pi}} \int_0^\infty e^{-\lambda t} \frac{1}{\sqrt{t}} e^{-\frac{(b^* - \varepsilon)^2}{2\sigma_1^2 t}} dt \\ &= \frac{\sigma_1 \sqrt{2\lambda}}{\sigma_2(\sigma_1 + \sigma_2)} e^{-\frac{b^* + \varepsilon}{\sigma_1} \sqrt{2\lambda}} (y - x). \end{aligned} \quad (16)$$

In the last step we have used that the Laplace transform of the function $\frac{1}{\sqrt{t}}e^{-\frac{a^2}{4t}}$ is given by $\sqrt{\frac{\pi}{\lambda}}e^{-a\sqrt{\lambda}}$ for $a > 0$ (see e.g. chapter 8 in [12]).

If $b^* > 0$ we have for any $\varepsilon > 0$ and $x, y \in [b^* - \varepsilon, b^* + \varepsilon], x \leq y$, that

$$F_\mu(y) - F_\mu(x) \geq \frac{\sigma_1\sqrt{2\lambda}}{\sigma_2(\sigma_1 + \sigma_2)}e^{-\frac{b^* + \varepsilon}{\sigma_1}\sqrt{2\lambda}}(y - x). \quad (17)$$

Now, let $u_n := q(\mu, 1 - \alpha) + \varepsilon_n, l_n := q(\mu, 1 - \alpha) - \varepsilon_n$ with $\varepsilon_n := \frac{1}{C_1}\sqrt{\frac{\mathcal{W}(n)}{n}}$ and $C_1 := \frac{\sigma_1}{\sigma_2(\sigma_1 + \sigma_2)}\sqrt{\frac{\lambda}{2}}e^{-\frac{|b^*|}{\sigma_1}\sqrt{2\lambda}}$. (16) and (17) imply that for $x, y \in [l_n, u_n], x \leq y$, and for n sufficiently large, we have that

$$F_\mu(y) - F_\mu(x) \geq 2C_1e^{-\frac{\sqrt{2\lambda}}{\sigma_1 C_1}}\sqrt{\frac{\mathcal{W}(n)}{n}}(y - x) \geq C_1(y - x), \quad (18)$$

since $\frac{\mathcal{W}(n)}{n} \xrightarrow{n \rightarrow \infty} 0$. Moreover, Lemma 3.3 implies that

$$\sup_{x \in \mathbb{R}} \left| P(q(\mu^{n,a}, 1 - \alpha_n) \leq x) - \Phi\left(\frac{\sqrt{n}F_\mu(x) - (1 - \alpha)}{\sqrt{\alpha(1 - \alpha)}}\right) \right| \leq \frac{C_2}{\sqrt{n}}, \quad (19)$$

for $n \geq 1$ and some $C_2 > 0$. Combining (18) and (19) we observe for n sufficiently large that

$$\begin{aligned} & P(q(\mu^{n,a}, \alpha_n) \notin [l_n, u_n]) \\ &= 1 - (P(q(\mu^{n,a}, 1 - \alpha_n) < u_n) - P(q(\mu^{n,a}, 1 - \alpha_n) \leq l_n)) \\ &\leq 1 - \left(\Phi\left(\frac{\sqrt{n}F_\mu(u_n) - (1 - \alpha)}{\sqrt{\alpha(1 - \alpha)}}\right) - \Phi\left(\frac{\sqrt{n}F_\mu(l_n) - (1 - \alpha)}{\sqrt{\alpha(1 - \alpha)}}\right) \right) + \frac{2C_2}{\sqrt{n}} \\ &= \Phi\left(-\frac{\sqrt{n}F_\mu(u_n) - (1 - \alpha)}{\sqrt{\alpha(1 - \alpha)}}\right) + \Phi\left(-\frac{\sqrt{n}(1 - \alpha) - F_\mu(l_n)}{\sqrt{\alpha(1 - \alpha)}}\right) + \frac{2C_2}{\sqrt{n}} \\ &\leq 2\Phi\left(-\frac{\sqrt{n}\frac{C_1\varepsilon_n}{\sqrt{\alpha(1 - \alpha)}}}{\sqrt{\alpha(1 - \alpha)}}\right) + \frac{2C_2}{\sqrt{n}} \\ &\leq 2\Phi\left(-\frac{\sqrt{\mathcal{W}(n)}}{\sqrt{\alpha(1 - \alpha)}}\right) + \frac{2C_2}{\sqrt{n}} \\ &\leq 2\Phi\left(-\sqrt{\mathcal{W}(n)}\right) + \frac{2C_2}{\sqrt{n}}, \end{aligned}$$

since $\frac{1}{\sqrt{\alpha(1 - \alpha)}} \geq 1$. One can show that $\Phi(-x) \in \mathcal{O}\left(\frac{1}{x}e^{-\frac{x^2}{2}}\right)$ as $x \rightarrow \infty$. Therefore, there is a $C_3 > 0$ such that for n large enough

$$P(q(\mu^{n,a}, \alpha_n) \notin [l_n, u_n]) \leq C_3\frac{e^{-\frac{\mathcal{W}(n)}{2}}}{\sqrt{\mathcal{W}(n)}} + \frac{2C_2}{\sqrt{n}} = (C_3 + 2C_2)\frac{1}{\sqrt{n}}.$$

□

A.2. Auxiliary results

In this section we aim at finding an upper bound for the probability $P(X_\tau^{x,\beta} \in [l, u])$ independent of the control β chosen. This task is motivated by the proof of Theorem 3.2, where this estimate is needed. In particular, we need an explicit formula for this bound, given in (21) below, to determine the speed of convergence. The results stated here can be found in [1] and are proven there in full detail.

We consider the following set-up: Let B be a Brownian motion on a complete probability space (Ω, \mathcal{F}, P) and τ be an exponentially distributed random variable with parameter $\lambda > 0$ independent of B . Furthermore, let $(\mathcal{F}_t)_{t \geq 0}$ be the augmented Brownian filtration, generated by B and augmented by the P -null sets in \mathcal{F} .

Let $0 < \sigma_1 \leq \sigma_2$ and $l, u \in \mathbb{R}$, $l < u$. We denote by $\widetilde{\mathcal{M}}$ the set of all $(\mathcal{F}_t)_{t \geq 0}$ -progressively measurable control processes $(\beta_t)_{t \geq 0}$ taking values in the interval $[\sigma_1, \sigma_2]$. For $\beta \in \widetilde{\mathcal{M}}$ and any initial value $x \in \mathbb{R}$ we consider state processes $X^{x,\beta}$, given by

$$X_t^{x,\beta} = x + \int_0^t \beta_s dB_s, \quad t \geq 0.$$

Note that similar to Remark 2.1, all feedback controls $(m(X_t^{x,m}))_{t \geq 0}$ with $m : \mathbb{R} \rightarrow [\sigma_1, \sigma_2]$ being of bounded variation are contained in $\widetilde{\mathcal{M}}$. Here $X^{x,m}$ denotes the solution to the SDE

$$dX_t = m(X_t)dB_t, \quad X_0 = x. \quad (20)$$

Let $\hat{m} : \mathbb{R} \rightarrow [\sigma_1, \sigma_2]$ be a feedback function, given by

$$\hat{m}(x) = \begin{cases} \sigma_1, & x \in [l, u], \\ \sigma_2, & \text{else.} \end{cases}$$

We set

$$F(l, u, x) = \begin{cases} \frac{e^{-\frac{\sqrt{2\lambda}}{\sigma_2}(l-x)}}{\frac{\sigma_1}{\sigma_2} \coth\left(\frac{u-l}{2\sigma_1}\sqrt{2\lambda}\right) + 1}, & x < l, \\ 1 - \frac{\cosh\left(\frac{2x-(u+l)}{2\sigma_1}\sqrt{2\lambda}\right)}{\cosh\left(\frac{u-l}{2\sigma_1}\sqrt{2\lambda}\right) + \frac{\sigma_2}{\sigma_1} \sinh\left(\frac{u-l}{2\sigma_1}\sqrt{2\lambda}\right)}, & x \in [l, u], \\ \frac{e^{-\frac{\sqrt{2\lambda}}{\sigma_2}(x-u)}}{\frac{\sigma_1}{\sigma_2} \coth\left(\frac{u-l}{2\sigma_1}\sqrt{2\lambda}\right) + 1}, & x > u. \end{cases} \quad (21)$$

We obtain the following uniform estimate:

Proposition A.1. For any $\beta \in \widetilde{\mathcal{M}}$ we have

$$P(X_\tau^{x,\beta} \in [l, u]) \leq F(l, u, x).$$

Moreover, $F(l, u, x) = P(X_\tau^{x,\hat{m}} \in [l, u])$.

One can prove this proposition by solving the control problem of maximizing the probability $P(X_\tau^{x,\beta} \in [l, u])$ over all controls $\beta \in \widetilde{\mathcal{M}}$. To this end, one can perform a classical verification with the function F to show that it is equal to the value function v . It is easy to verify that $F(l, u, \cdot)$ is concave on $[l, u]$ and convex otherwise. We refer to Proposition 2.3 in [1] for a complete proof. In particular, one can show the following:

Proposition A.2. We have

$$v(x) := \sup_{\beta \in \widetilde{\mathcal{M}}} P(X_\tau^{x,\beta} \in [l, u]) = F(l, u, x), \quad x \in \mathbb{R},$$

and $(\hat{m}(X_t^{x,\hat{m}}))_{t \geq 0}$ is an optimal control.

To determine the speed of convergence of $P(X_\tau^{x,\beta} \in [l, u])$ when $u - l \rightarrow 0$ we need the following estimate. It also entails the absolute continuity of the random variable $X_\tau^{x,\beta}$.

Lemma A.3. We have, for any $\beta \in \widetilde{\mathcal{M}}$,

$$P(X_\tau^{x,\beta} \in [l, u]) \leq \frac{\sigma_2}{\sigma_1^2} \sqrt{\frac{\lambda}{2}} e^{\frac{\sqrt{2\lambda}}{\sigma_1}(u-l)} (u - l). \quad (22)$$

In particular, this estimate holds true for $F(l, u, x)$.

Proposition A.4. Let $\beta \in \widetilde{\mathcal{M}}$. Then $X_\tau^{x,\beta}$ is absolutely continuous w.r.t. the Lebesgue measure on \mathbb{R} .

Proof of Lemma A.3. If $x \notin [l, u]$ we have

$$\begin{aligned} P(X_\tau^{x,\beta} \in [l, u]) &\leq F(l, u, x) \\ &\leq \frac{1}{\frac{\sigma_1}{\sigma_2} \coth\left(\frac{u-l}{2\sigma_1}\sqrt{2\lambda}\right) + 1} \leq \frac{\sigma_2}{\sigma_1} \frac{1}{\coth\left(\frac{u-l}{2\sigma_1}\sqrt{2\lambda}\right) + 1} \\ &= \frac{\sigma_2}{\sigma_1} \frac{1}{2 + \frac{2}{\exp\left(\frac{u-l}{\sigma_1}\sqrt{2\lambda}\right) - 1}} = \frac{\sigma_2}{2\sigma_1} \frac{\exp\left(\frac{u-l}{\sigma_1}\sqrt{2\lambda}\right) - 1}{\exp\left(\frac{u-l}{\sigma_1}\sqrt{2\lambda}\right)} \\ &\leq \frac{\sigma_2}{2\sigma_1} \left(e^{\frac{\sqrt{2\lambda}}{\sigma_1}(u-l)} - 1 \right) \leq \frac{\sigma_2}{\sigma_1^2} \sqrt{\frac{\lambda}{2}} e^{\frac{\sqrt{2\lambda}}{\sigma_1}(u-l)} (u - l). \end{aligned}$$

If $x \in [l, u]$ we have

$$\begin{aligned}
P(X_\tau^{x,\beta} \in [l, u]) &\leq F(l, u, x) \leq 1 - \frac{1}{\cosh\left(\frac{u-l}{2\sigma_1}\sqrt{2\lambda}\right) + \frac{\sigma_2}{\sigma_1} \sinh\left(\frac{u-l}{2\sigma_1}\sqrt{2\lambda}\right)} \\
&= \frac{\cosh\left(\frac{u-l}{2\sigma_1}\sqrt{2\lambda}\right) + \frac{\sigma_2}{\sigma_1} \sinh\left(\frac{u-l}{2\sigma_1}\sqrt{2\lambda}\right) - 1}{\cosh\left(\frac{u-l}{2\sigma_1}\sqrt{2\lambda}\right) + \frac{\sigma_2}{\sigma_1} \sinh\left(\frac{u-l}{2\sigma_1}\sqrt{2\lambda}\right)} \\
&\leq \cosh\left(\frac{u-l}{2\sigma_1}\sqrt{2\lambda}\right) + \frac{\sigma_2}{\sigma_1} \sinh\left(\frac{u-l}{2\sigma_1}\sqrt{2\lambda}\right) - 1 \\
&\leq \frac{\sigma_2}{\sigma_1} \left(\cosh\left(\frac{u-l}{2\sigma_1}\sqrt{2\lambda}\right) + \sinh\left(\frac{u-l}{2\sigma_1}\sqrt{2\lambda}\right) - 1 \right) \\
&= \frac{\sigma_2}{\sigma_1} \left(e^{\frac{\sqrt{2\lambda}}{2\sigma_1}(u-l)} - 1 \right) \leq \frac{\sigma_2}{\sigma_1^2} \sqrt{\frac{\lambda}{2}} e^{\frac{\sqrt{\lambda}}{\sqrt{2}\sigma_1}(u-l)} (u-l) \\
&\leq \frac{\sigma_2}{\sigma_1^2} \sqrt{\frac{\lambda}{2}} e^{\frac{\sqrt{2\lambda}}{\sigma_1}(u-l)} (u-l).
\end{aligned}$$

□

Proof of Proposition A.4. Let $N \subseteq \mathbb{R}$ be a Lebesgue null set and $\varepsilon > 0$. Then N can be covered by a countable union of intervals $[l_n, u_n]$, $n \in \mathbb{N}$, such that $\sum_{n=1}^{\infty} |u_n - l_n| < \varepsilon$. We have, using Lemma A.3,

$$\begin{aligned}
P(X_\tau^{x,\beta} \in N) &\leq \sum_{n=1}^{\infty} P(X_\tau^{x,\beta} \in [l_n, u_n]) \leq \sum_{n=1}^{\infty} G(l_n, u_n, x) \\
&\leq \sum_{n=1}^{\infty} \frac{\sigma_2}{\sigma_1^2} \sqrt{\frac{\lambda}{2}} e^{\frac{\sqrt{2\lambda}}{\sigma_1}(u_n - l_n)} (u_n - l_n) \\
&\leq \frac{\sigma_2}{\sigma_1^2} \sqrt{\frac{\lambda}{2}} e^{\frac{\sqrt{2\lambda}}{\sigma_1}\varepsilon} \sum_{n=1}^{\infty} (u_n - l_n) < \frac{\sigma_2}{\sigma_1^2} \sqrt{\frac{\lambda}{2}} e^{\frac{\sqrt{2\lambda}}{\sigma_1}\varepsilon} \varepsilon.
\end{aligned}$$

Since the choice of $\varepsilon > 0$ was arbitrary, we obtain that $P(X_\tau^{x,\beta} \in N) = 0$. □

Acknowledgement.

Support from the German Research Foundation through the project AN 1024/5-1 is gratefully acknowledged. Chao Zhou's research is supported by Singapore MOE (Ministry of Educations) AcRF Grants R-146- 000-271-112 and R-146-000-284-114 as well as NSFC Grant No. 11871364.

References

- [1] S. Ankirchner and J. Wendt. A sharp upper bound for the expected interval occupation time of Brownian martingales. *HAL preprint hal-03138433*.
- [2] A. Barrasso and N. Touzi. Controlled diffusion Mean Field Games with common noise, and McKean-Vlasov second order backward SDEs. *arXiv preprint arXiv:2005.07542*, 2020.
- [3] E. Bayraktar, J. Cvitanić, and Y. Zhang. Large tournament games. *Ann. Appl. Probab.*, 29(6):3695–3744, 2019.
- [4] E. Bayraktar and Y. Zhang. A rank-based mean field game in the strong formulation. *Electron. Commun. Probab.*, 21:Paper No. 72, 12, 2016.
- [5] E. Bayraktar and Y. Zhang. Terminal ranking games. *Forthcoming, Mathematics of Operations Research*, 2019.
- [6] R. Carmona and F. Delarue. *Probabilistic theory of mean field games with applications. I*, volume 83 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2018. Mean field FBSDEs, control, and games.
- [7] R. Carmona and F. Delarue. *Probabilistic theory of mean field games with applications. II*, volume 84 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2018. Mean field games with common noise and master equations.
- [8] R. Carmona and D. Lacker. A probabilistic weak formulation of mean field games and applications. *Ann. Appl. Probab.*, 25(3):1189–1231, 2015.
- [9] J. P. Choi. Dynamic R&D competition under "hazard rate" uncertainty. *The RAND Journal of Economics*, pages 596–610, 1991.
- [10] C. Deng, X. Su, and C. Zhou. Relative wealth concerns with partial information and heterogeneous priors. *arXiv preprint arXiv:2007.11781*, 2020.
- [11] G.-E. Espinosa and N. Touzi. Optimal investment under relative performance concerns. *Math. Finance*, 25(2):221–257, 2015.
- [12] G. B. Folland. *Fourier analysis and its applications*. Pacific Grove, CA: Wadsworth & Brooks/Cole Advanced Books & Software, 1992.
- [13] G. Fu, X. Su, and C. Zhou. Mean field exponential utility game: A probabilistic approach. *arXiv preprint arXiv:2006.07684*, 2020.
- [14] M. Huang, P. E. Caines, and R. P. Malhamé. Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized ϵ -Nash equilibria. *IEEE Trans. Automat. Control*, 52(9):1560–1571, 2007.

- [15] M. Huang, R. P. Malhamé, and P. E. Caines. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.*, 6(3):221–251, 2006.
- [16] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, New York, NY, 1998.
- [17] J. Keilson and J. A. Wellner. Oscillating Brownian motion. *J. Appl. Probability*, 15(2):300–310, 1978.
- [18] N. V. Krylov. *Controlled Diffusion Processes*, volume 14. Springer Science & Business Media, 1980.
- [19] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. I. Le cas stationnaire. *C. R. Math. Acad. Sci. Paris*, 343(9):619–625, 2006.
- [20] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. II. Horizon fini et contrôle optimal. *C. R. Math. Acad. Sci. Paris*, 343(10):679–684, 2006.
- [21] J.-M. Lasry and P.-L. Lions. Mean field games. *Jpn. J. Math.*, 2(1):229–260, 2007.
- [22] A. Lejay and P. Pigato. Statistical estimation of the oscillating Brownian motion. *Bernoulli*, 24(4B):3568–3602, 2018.
- [23] C. Ma and J. Robinson. 17 approximations to distributions of sample quantiles. In *Order Statistics: Theory & Methods*, volume 16 of *Handbook of Statistics*, pages 463 – 484. Elsevier, 1998.
- [24] J. M. McNamara. Optimal control of the diffusion coefficient of a simple diffusion process. *Math. Oper. Res.*, 8(3):373–380, 1983.
- [25] J. M. McNamara and O. Leimar. *Game theory in biology: concepts and frontiers*. Oxford University Press, USA, 2020.
- [26] S. Nakao. On the pathwise uniqueness of solutions of one-dimensional stochastic differential equations. *Osaka Journal of Mathematics*, 9(3):513–518, 1972.
- [27] M. Nutz and Y. Zhang. A mean field competition. *Math. Oper. Res.*, 44(4):1245–1263, 2019.
- [28] M. L. Puri and S. S. Ralescu. Limit theorems for random central order statistics. In *Adaptive statistical procedures and related topics (Upton, N. Y., 1985)*, volume 8 of *IMS Lecture Notes Monogr. Ser.*, pages 447–475. Inst. Math. Statist., Hayward, CA, 1986.
- [29] A. Sapirshtein, Y. Sompolinsky, and A. Zohar. Optimal selfish mining strategies in bitcoin. In J. Grossklags and B. Preneel, editors, *Financial Cryptography and Data Security*, pages 515–532, Berlin, Heidelberg, 2017. Springer Berlin Heidelberg.
- [30] C. Seel and P. Strack. Gambling in contests. *J. Econom. Theory*, 148(5):2033–2048, 2013.

- [31] C. Seel and P. Strack. Continuous time contests with private information. *Math. Oper. Res.*, 41(3):1093–1107, 2016.