



HAL
open science

Pareto Optima for a Class of Singular Control Games

Rama Cont, Xin Guo, Renyuan Xu

► **To cite this version:**

Rama Cont, Xin Guo, Renyuan Xu. Pareto Optima for a Class of Singular Control Games. 2020.
hal-03049246

HAL Id: hal-03049246

<https://hal.science/hal-03049246v1>

Preprint submitted on 9 Dec 2020

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.

Pareto Optima for a Class of Singular Control Games

In memory of Mark H Davis, mentor and friend

Rama Cont

Xin Guo

Renyuan Xu

2020

Abstract

We study a class of N -player stochastic differential games of singular control, motivated by the study of a dynamic model of interbank lending with benchmark rates. We describe Pareto optima for this game and show how they may be achieved through the intervention of a regulator, whose policy is a solution to a singular stochastic control problem. Pareto optima are characterized in terms of the solution to a new class of Skorokhod problems with piecewise-continuous free boundary.

Pareto optimal policies are shown to correspond to the enforcement of endogenous bounds on interbank lending rates. Analytical comparison between Pareto optima and Nash equilibria for the case of two players allows to quantify the impact of regulatory intervention on the stability of the interbank rate.

Keywords: LIBOR rate, interbank markets, stochastic differential game, singular stochastic control, Pareto optimum, Nash equilibrium, Skorokhod problem.

Contents

1	Introduction	2
1.1	A model of interbank lending with benchmark rates	2
1.2	A class of stochastic differential games of singular control	3
2	Mathematical formulation of the game	5
3	Regulator's problem	6
4	Pareto-optimal policies	14
4.1	Optimal policy for the regulator	14
4.2	Pareto-optimal policy	21
4.3	Pareto-optimal policies for interbank lending	22
5	Explicit solution for $N = 2$ players	23
5.1	Pareto-optimal solution for $N = 2$	23
5.2	Pareto optimum vs Nash equilibrium	26

1 Introduction

The market for interbank lending offers an interesting example of strategic interaction among financial institutions in which players react to an *average* of the action of other players. One of the widely commented features of the interbank market is the fixing mechanism for interbank benchmark interest rates, the most well-known example of which is the London Interbank Offer Rate (LIBOR) which plays a central role in financial markets. Historically these benchmarks have not been negotiated rates but a ‘trimmed’ average of quotes collected daily from major banks. At every date t , participating banks contribute a quote X_t^i representing their offered rate; a calculation agent then ‘trims’ the tails of the distribution by removing the highest and lowest quotes and computes the value of the benchmark rate \bar{X}_t as a weighted average of the remaining non-discarded quotes (Avellaneda & Cont, 2010). The resulting benchmark rate \bar{X}_t –the LIBOR rate– then serves as a reference for the valuation of interbank loans and debt contracts, as well as many other financial contracts indexed on the benchmark rate. A deviation (spread) of a bank’s rate from the benchmark may lead to a perception of credit risk and loss of market share -if the spread is positive- or an opportunity cost if the spread is negative, thus incentivizing banks to align their offered rates with the benchmark.

This mechanism leads to strategic interactions among market participants in a dynamic setting, where interactions are mediated through an average action, or more generally through the distribution of actions of other participants and has been criticized for its vulnerability to manipulations (Avellaneda & Cont, 2010), which have been extensively documented (H.M. Treasury, 2012; Duffie & Stein, 2015). One of the lessons from the manipulation of LIBOR and other benchmarks is that insufficient attention had been paid to incentives, strategic interactions, mechanism design and the role of the regulator in such markets.

1.1 A model of interbank lending with benchmark rates

We shall now describe a stylized model of interbank rates which represents interactions among banks in terms of a stochastic dynamic game. Consider first an exogenous process r_t representing a rate set by the central bank, with respect to which banks will position their lending rates. r_t is typically modeled as a mean-reverting diffusion process driven by a Brownian motion B representing random macroeconomic shocks. Each bank i quotes a rate r_t^i at a ‘spread’ X_t^i with respect to the reference rate r_t : $r_t^i = r_t + X_t^i$. Bank i is also affected by the macroeconomic shocks represented by B but may control its rate r_t^i through positive or negative adjustments to its spread X_t^i , which we may represent by a pair $(\xi^{i,+}, \xi^{i,-})$ of non-decreasing processes.

$$dX_t^i = \sigma^i dB_t + d\xi_t^{i,+} - d\xi_t^{i,-}. \quad (1.1)$$

The benchmark (‘LIBOR’) rate L_t is then defined as a weighted average of these offered rates:

$$L_t = r_t + \bar{X}_t, \quad \bar{X}_t = \sum_{i=1}^N a_i X_t^i \quad \text{and} \quad a_i \geq 0, \quad \sum_{i=1}^N a_i = 1.$$

Note that the ‘drift’ term in the dynamics (1.1) originates from the control. One may also consider an additional drift term $\mu^i dt$ in the uncontrolled dynamics, a positive drift corresponding to a bank whose creditworthiness is gradually deteriorating, leading to a steady increase of its spread.

We now turn to the incentives and costs faced by banks. Each bank i receives interest income from its lending activity, at rate r_t^i . The interest income of the bank over a short period $[t, t + dt]$ is $r_t^i Q_t^i dt$ where $Q_t^i > 0$ is the volume of lending activity (loan volume). Given that the bank can borrow at the interbank $L_t = r_t + \bar{X}_t$, this represents an *opportunity cost* of $(\bar{X}_t - X_t^i) Q_t^i dt$. In a competitive lending market, the market share Q_t^i of bank i will be a decreasing function $q_i(\cdot)$ of its spread $r_t^i - L_t = X_t^i - \bar{X}_t$

relative to the benchmark rate: $Q_t^i = q_i(X_t^i - \bar{X}_t)$. Assuming an inter-temporal discount rate of $\rho > 0$, this leads to a running cost term

$$\int_0^\infty e^{-\rho t} (\bar{X}_t - X_t^i) q_i(X_t^i - \bar{X}_t) dt.$$

For example, a linear dependence $q_i(x) = -\kappa_i x$, where $\kappa_i > 0$ represents the sensitivity of loan volume to the interest rate, leads to a quadratic running cost $\int_0^\infty \kappa_i \exp^{-\rho t} (\bar{X}_t - X_t^i)^2 dt$.

These considerations only pertain to the relative costs of bank simultaneously engaging in borrowing and lending. Other constraints prevent the banks from deviating from the reference rate beyond a certain level; these are often ‘soft’, rather than hard (i.e., inequality), constraints and may be modeled by a penalty on $|r_t^i|$, or equivalently a running cost $f_i(X_t^i)$ where f_i is centered at some reference value and increases fast enough (e.g., quadratically) at infinity. As an example we shall use $f_i(x) = \nu_i(x - s_0)^2$ with $\nu_i > 0$.

The benchmark fixing mechanism described above may be incorporated in the model through a cost term associated with the control $(\xi^{i,+}, \xi^{i,-})$. Recall that the LIBOR is computed as a *trimmed* average of quotes, discarding the highest and lowest ‘outliers’. This means an offered rate X^i will not be taken into account if it lies too far from the mean. In absence of collusion between banks, this mechanism discourages them from making large daily adjustments to their offered rates, as a large upward or downward adjustment may result in their quote being disregarded in the benchmark calculation. This may be modeled through a cost term which penalizes the size of the adjustment e.g., $K_i^+ d\xi_t^{i,+} + K_i^- d\xi_t^{i,-}$, with $K_i^+, K_i^- > 0$, where $1/K_i^+$ (resp. $1/K_i^-$) represents a typical distance $(X^i - \bar{X})_+$ (resp. $(\bar{X} - X^i)_+$) beyond which quotes are discarded. For instance one can take $K_i^+ = K_i^- = 1/\gamma$ where γ represents a measure of dispersion (interquartile range or multiple of standard deviation) of the quote distribution. The case of an asymmetric penalty $K_i^+ > K_i^-$ (resp. $K_i^+ < K_i^-$) is useful to model the case of a bank i systematically quoting above (or below) the benchmark. This leads to an objective function

$$J^i(\mathbf{x}; \boldsymbol{\xi}) = \mathbb{E} \left[\int_0^\infty e^{-\rho t} ((\bar{X}_t - X_t^i) q_i(X_t^i - \bar{X}_t) dt + \nu_i |X_t^i - s_0|^2 dt + K_i^+ d\xi_t^{i,+} + K_i^- d\xi_t^{i,-}) \mid \mathbf{X}_{0-} = \mathbf{x} \right] \quad (1.2)$$

for bank i , where the control variable is a pair of non-decreasing processes $\xi_t^{i,+}, \xi_t^{i,-}$ representing the rate adjustments of bank i and the expectation is taken with respect to the law of the controlled process (1.1). The controls $\xi^{i,+}, \xi^{i,-}$ are in general allowed to be right-continuous with left limits (càdlàg) which allows for jumps, representing *impulses*, as well as continuous adjustments to the rates. Such controls are called singular controls (Karatzas, 1982) and have been used for analyzing optimal investment policy and option pricing and hedging problems with transaction costs (Davis & Norman, 1990; Davis et al., 1993; Kallsen & Muhle-Karbe, 2015, 2017; Zariphopoulou, 1992).

In the case where $a_i = \frac{1}{N}$, $q_i = q_j$, $\nu_i = \nu_j$ and $K_i^\pm = K_j^\pm$ for $i \neq j$, the payoff structure is symmetric under permutation of indices and this can be formulated as mean field game, which was studied in (Lasry & Lions, 2007; Guo & Xu, 2019). However we shall not need this assumption and will treat below the case of a more general, not necessarily symmetric, cost function $h^i(\mathbf{X}_t)$.

1.2 A class of stochastic differential games of singular control

Motivated by the example above, we study a class of N -player stochastic differential games, where each player $i = 1, \dots, N$ controls a diffusive process X_t^i through (positive or negative) $\boldsymbol{\xi}^i := (\xi^{i,+}, \xi^{i,-})$ additive control terms

$$dX_t^i = \mu^i dt + \boldsymbol{\sigma}^i \cdot d\mathbf{B}_t + d\xi_t^{i,+} - d\xi_t^{i,-}, \quad X_{0-}^i = x^i. \quad (1.3)$$

and a seeks to minimize the sum of a discounted running cost and a proportional cost of intervention

$$J^i(\mathbf{x}; \boldsymbol{\xi}) = \mathbb{E} \left[\int_0^\infty e^{-\rho t} (h^i(\mathbf{X}_t) dt + K_i^+ d\xi_t^{i,+} + K_i^- d\xi_t^{i,-}) \mid \mathbf{X}_{0-} = \mathbf{x} \right].$$

The first two terms in (1.3) correspond to the 'baseline' (uncontrolled) diffusion dynamics, and the last two term correspond to the control $\boldsymbol{\xi}^i = (\xi^{i,+}, \xi^{i,-})$, modeled as a pair of non-decreasing càdlàg processes, leading to a *singular control* problem (Karatzas, 1982) for each player. Nash equilibria for such singular control games have been studied in (Guo & Xu, 2019). In the present work, we focus on Pareto-optimal outcomes.

Contribution. The present work is a study of Pareto-optimal policies for the class of stochastic singular control games considered above, motivated by the interbank lending problem. We relate the Pareto optima of this game to the solution of a 'regulator's problem', characterized as a singular stochastic control problem which we study in detail. The regularity analysis of the value function, following the approach of Soner & Shreve (1989), for the regulator's problem enables us to characterize the optimal controls for this problem and subsequently the Pareto-optimal policies for the N -player game.

We obtain a description of Pareto-optimal policies in terms of a multidimensional Skorokhod problem for a 'regulated diffusion' in a bounded region whose boundary is piece-wise smooth with possible corners. The state process follows a diffusion process in the interior, and the control intervenes only at the boundary to reflect it back into the interior.

Finally, we derive explicit descriptions of Pareto-optimal policies when $N = 2$. This complements the existing literature on Nash equilibrium for stochastic two player games (De Angelis & Ferrari, 2018; Dianetti & Ferrari, 2020; Hernandez-Hernandez et al., 2015; Kwon & Zhang, 2015). Analytical comparison between the Pareto-optimal and the Nash equilibrium solutions demonstrates the role of regulator in the interbank lending game.

Our analysis provides insights for regulatory intervention on the interbank market. In particular, it allows us to quantify the impact of a regulator on the stability of the benchmark rate.

Relation with previous literature. Stylized mean-field models of interbank borrowing and lending have been considered by Carmona et al. (2013) and Sun (2018), who focus on Nash equilibria and general stochastic controls.

A related strand of literature consists of studies on central bank interventions on interest rates and exchange rates using an *impulse control* approach (Bensoussan et al., 2012; Cadenillas & Zapatero, 2000; Jeanblanc-Picqué, 1993). In these approaches, interventions are associated with a fixed cost and thus finite in number. The singular control framework adopted here seems more natural for modeling situations such as interbank markets where participants intervene continuously and where the cost of intervention is proportional to the action rather than fixed. Singular controls allow for discontinuities, so include impulse control as a special case. However as we will observe in Section 4, Pareto-optimal solutions do not include impulse and instead involve only continuous controls.

Nash equilibria for stochastic games of singular control have been studied by Chiarolla et al. (2013); De Angelis & Ferrari (2018); Dianetti & Ferrari (2020); Hernandez-Hernandez et al. (2015); on the other hand, there are few studies of Pareto-optimal strategies for such games. Aïd et al. (2017) considered a two-player game in an impulse control framework between a representative energy consumer and a representative electricity producer, and derived an asymptotic Pareto-optimal policy. Fischer & Livieri (2016) solved explicitly a mean-variance portfolio optimization problem with N stocks. Ferrari et al. (2017) and Wang & Ewald (2010) considered the problem of public good contribution and analyzed the Pareto-optimal policy for the N -player stochastic game under the framework of regular control and singular control, respectively.

The analysis of Pareto optima in stochastic games is often done by studying an auxiliary N -dimensional stochastic control problem. This approach can be traced back to the economic literature on mechanism design and social welfare optimization in Bator (1957) and Coleman (1979). The mathematical challenge lies in analyzing the associated high-dimensional Hamilton–Jacobi–Bellman (HJB) equations and characterizing the optimal control policy.

Outline. The remainder of the paper is organized as follows. Section 2 presents the mathematical formulation of the N -player stochastic differential game, and describes its relation with the auxiliary control problem. Section 3 provides detailed analysis of the auxiliary control problem and the construction of the optimal strategies. Section 4 characterizes the Pareto optima in terms of a sequence of Skorokhod problems. Implications of our analysis for the interbank lending problem are discussed in Section 4.3. Section 5 provides explicit solutions in the case $N = 2$, and compares it with the Nash equilibrium.

2 Mathematical formulation of the game

In this section, we describe the mathematical framework of the N -player game.

Controlled dynamics. Let $(X_t^i)_{t \geq 0} \in \mathbb{R}$ denote the state of player i at time t , $1 \leq i \leq N$. In the absence of controls, $\mathbf{X}_t := (X_t^1, \dots, X_t^N) \in \mathbb{R}^N$ follows

$$\mathbf{X}_t = \mathbf{X}_0 + \boldsymbol{\mu}t + \boldsymbol{\sigma}\mathbf{B}_t, \quad \mathbf{X}_0 = (x^1, \dots, x^N), \quad (2.1)$$

where $\mathbf{B} := (B^1, \dots, B^D) \in \mathbb{R}^D$ is a D -dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$, and $\boldsymbol{\mu} := (\mu_1, \dots, \mu_N) \in \mathbb{R}^N$ and $\boldsymbol{\sigma} := (\sigma_{ij})_{1 \leq i \leq N, 1 \leq j \leq D} \in \mathbb{R}^{N \times D}$ are constants with $\boldsymbol{\sigma}\boldsymbol{\sigma}^T \succeq \lambda \mathbf{I}$ for some $\lambda > 0$.

When player i chooses a control $\boldsymbol{\xi}^i := (\xi^{i,+}, \xi^{i,-})$ from an admissible control set \mathcal{U}_N^i , then X_t^i evolves as

$$dX_t^i = \mu^i dt + \boldsymbol{\sigma}^i \cdot d\mathbf{B}_t + d\xi_t^{i,+} - d\xi_t^{i,-}, \quad X_{0-}^i = x^i. \quad (2.2)$$

Here $\boldsymbol{\xi}^i = (\xi^{i,+}, \xi^{i,-})$ is a pair of non-decreasing càdlàg processes and $\boldsymbol{\sigma}^i$ is the i^{th} row of the volatility matrix $\boldsymbol{\sigma}$. We will denote by $\mathbb{P}^{\mathbf{x}}$ the law of the process (2.2) and $\mathbb{E}_{\mathbf{x}}$ the expectation with respect to this law.

Admissible controls \mathcal{U}_N^i . The admissible control set for player i is defined as

$$\mathcal{U}_N^i = \left\{ (\xi_t^{i,+}, \xi_t^{i,-})_{t \geq 0} \mid \xi_t^{i,+} \text{ and } \xi_t^{i,-} \text{ are } \mathcal{F}_t\text{-progressively measurable, càdlàg non-decreasing,} \right. \\ \left. \text{with } \mathbb{E} \left[\int_0^\infty e^{-\rho t} d\xi_t^{i,+} \right] < \infty, \mathbb{E} \left[\int_0^\infty e^{-\rho t} d\xi_t^{i,-} \right] < \infty, \xi_{0-}^{i,+} = 0, \xi_{0-}^{i,-} = 0 \right\}. \quad (2.3)$$

Objective functions. Each player i chooses a control $(\xi^{i,+}, \xi^{i,-})$ in \mathcal{U}_N^i to minimize

$$J^i(\mathbf{x}; \boldsymbol{\xi}) = \mathbb{E}_{\mathbf{x}} \int_0^\infty e^{-\rho t} [h^i(\mathbf{X}_t) dt + K_i^+ d\xi_t^{i,+} + K_i^- d\xi_t^{i,-}]. \quad (\mathbf{N}\text{-player})$$

Here $\rho > 0$ is a constant discount factor, $K_i^+, K_i^- > 0$ are the cost of controls, and $h^i(\mathbf{x}) : \mathbb{R}^N \rightarrow \mathbb{R}_+$ is the running cost function.

In this paper, we focus on characterizing Pareto optima of the game (**N-player**) subject to the dynamics (2.2).

Definition 1 (Pareto optimality). $\boldsymbol{\xi}^* \in \mathcal{U}_N := (\mathcal{U}_N^1, \dots, \mathcal{U}_N^N)$ is a Pareto-optimal policy for the game (**N-player**) if and only if there does not exist $\boldsymbol{\xi} \in \mathcal{U}_N$ such that

$$\forall i \in \{1, \dots, N\}, \quad J^i(\mathbf{x}; \boldsymbol{\xi}) \leq J^i(\mathbf{x}; \boldsymbol{\xi}^*); \quad \text{and} \quad \exists j \in \{1, \dots, N\}, \quad J^j(\mathbf{x}; \boldsymbol{\xi}) < J^j(\mathbf{x}; \boldsymbol{\xi}^*).$$

Pareto optima correspond to *efficient* outcomes of a game, which may or may not result from decentralized optimization by N players. The intervention of a regulator may be necessary to enforce a Pareto-optimal policy.

3 Regulator's problem

To study Pareto optima for game (**N-player**), we consider a 'welfare function' defined as an aggregated cost:

$$\begin{aligned} J(\mathbf{x}; \boldsymbol{\xi}) &= \sum_{i=1}^N L_i J^i(\mathbf{x}, \boldsymbol{\xi}) \\ &= \mathbb{E}_{\mathbf{x}} \int_0^{\infty} e^{-\rho t} \left[H(\mathbf{X}_t) dt + \sum_{i=1}^N L_i K_i^+ d\xi_t^{i,+} + \sum_{i=1}^N L_i K_i^- d\xi_t^{i,-} \right], \end{aligned} \quad (3.1)$$

where the dynamics of \mathbf{X}_t is given by (2.2), and

$$H(\mathbf{x}) := \sum_{i=1}^N L_i h^i(\mathbf{x}), \quad \text{with } L_i > 0 \text{ and } \sum_{i=1}^N L_i = 1. \quad (3.2)$$

We will show that Pareto optima of (**N-player**) correspond to solutions of the following auxiliary stochastic control problem

$$v(\mathbf{x}) = \min_{\boldsymbol{\xi} \in \mathcal{U}_N} J_L(\mathbf{x}; \boldsymbol{\xi}), \quad (\mathbf{Regulator})$$

which may be interpreted as the problem facing a market regulator seeking to optimize the aggregate cost (3.1).

To ensure that the game is well defined, the following assumptions will be made throughout the paper, unless otherwise specified:

Assumptions. There exist $C > c > 0$ such that

A1. $\forall \mathbf{x} \in \mathbb{R}^N, 0 \leq H(\mathbf{x}) \leq C(1 + \|\mathbf{x}\|^2).$

A2. $\forall \mathbf{x}, \mathbf{x}' \in \mathbb{R}^N, |H(\mathbf{x}) - H(\mathbf{x}')| \leq C(1 + \|\mathbf{x}\| + \|\mathbf{x}'\|)\|\mathbf{x} - \mathbf{x}'\|.$

A3. $H(\mathbf{x}) \in \mathcal{C}^2(\mathbb{R}^N)$, H is convex, with $0 < c \leq \partial_{\mathbf{z}}^2 H(\mathbf{x}) \leq C$ for all unit direction $\mathbf{z} \in \mathbb{R}^N$.

For example, for the payoff described in the interbank lending problem in Section 1.1,

$$H(\mathbf{x}) = \sum_{i=1}^N L_i \left[\kappa_i \left(x^i - \sum_{j \neq i} a_j x^j \right)^2 + \nu_i (x^i)^2 \right] \quad \text{with } \kappa_i, \nu_i > 0. \quad (3.3)$$

Then H satisfies **A1-A3** for any choice of weight $L_i > 0$.

We shall first analyze the regularity property of the value function v , which is necessary for establishing subsequently the existence and uniqueness of the optimal control. We then see that the optimal control for (**Regulator**) yields a Pareto-optimal policy for game (**N-player**).

The regularity analysis of the value function involves several steps. The first step is to show that the value function for **(Regulator)** is a viscosity solution to the following HJB equation

$$\max\{\rho u - \mathcal{L}u - H(\mathbf{x}), \beta(\nabla u) - 1\} = 0, \quad (3.4)$$

with the operator $\mathcal{L} = \frac{1}{2} \sum_{i,j=1}^N \sigma^i \cdot \sigma^j \partial_{x^i x^j}^2 + \sum_{i=1}^N \mu^i \partial_{x^i}$, and

$$\beta(\mathbf{q}) = \max_{1 \leq i \leq N} \left[\left(\frac{q^i}{L_i K_i^-} \right)^+ \vee \left(\frac{q^i}{L_i K_i^+} \right)^- \right], \quad (3.5)$$

where $\mathbf{q} := (q^1, \dots, q^N)$, $(a)^+ = \max\{0, a\}$ and $(a)^- = \max\{0, -a\}$ for any $a \in \mathbb{R}$.

The second step is to show that the value function for **(Regulator)** is $\mathcal{W}_{loc}^{2,\infty}$. The third step is to show that the HJB equation (3.4) has a unique $\mathcal{W}_{loc}^{2,\infty}$ solution.

Let us start with the following property of the value function v for the control problem **(Regulator)**.

Proposition 2. *Under Assumptions A1-A2, there exists $K > 0$ such that,*

$$(i) \quad 0 \leq v(\mathbf{x}) \leq K(1 + \|\mathbf{x}\|^2), \quad \forall \mathbf{x} \in \mathbb{R}^N;$$

$$(ii) \quad |v(\mathbf{x}) - v(\mathbf{x}')| \leq K(1 + \|\mathbf{x}\| + \|\mathbf{x}'\|)\|\mathbf{x} - \mathbf{x}'\|, \quad \forall \mathbf{x}, \mathbf{x}' \in \mathbb{R}^N.$$

Proof. First, $v(\mathbf{x}) \geq 0$ is clear by the non-negativity of $H(\mathbf{x})$. Moreover, by the property that $\sigma\sigma^T \succeq \lambda \mathbf{I}$ with $\lambda > 0$, it follows from a known estimate and the martingale argument (Menaldi & Robin, 1983, (2.15)) that the solution $\{\tilde{\mathbf{X}}_t\}_{t \geq 0} := \{\mathbf{x} + \boldsymbol{\mu}t + \boldsymbol{\sigma}\mathbf{B}_t\}_{t \geq 0}$ with $\boldsymbol{\xi} = \mathbf{0}$ satisfies

$$\mathbb{E}_{\mathbf{x}} \int_0^\infty e^{-\rho t} \|\tilde{\mathbf{X}}_t\|^2 dt \leq K(1 + \|\mathbf{x}\|^2), \quad \forall \mathbf{x} \in \mathbb{R}^N,$$

for some constant $K > 0$. By Assumption A1, there exists a constant $K > 0$ such that

$$v(\mathbf{x}) \leq J(\mathbf{x}, \mathbf{0}) \leq K(1 + \|\mathbf{x}\|^2), \quad \forall \mathbf{x} \in \mathbb{R}^N.$$

Thus (i) of Proposition 2 is established.

For each fixed $\mathbf{x} \in \mathbb{R}^N$, let

$$\mathcal{U}_{\mathbf{x}} = \{\boldsymbol{\xi} \in \mathcal{U} : J(\mathbf{x}, \boldsymbol{\xi}) \leq J(\mathbf{x}, \mathbf{0})\}. \quad (3.6)$$

By Assumption A1,

$$\mathbb{E}_{\mathbf{x}} \int_0^\infty e^{-\rho t} \|\mathbf{X}_t\|^2 dt \leq K(1 + \|\mathbf{x}\|^2), \quad \forall \mathbf{x} \in \mathbb{R}^N, \boldsymbol{\xi} \in \mathcal{U}_{\mathbf{x}}. \quad (3.7)$$

For $\boldsymbol{\xi} \in \mathcal{U}_{\mathbf{x}}$, it is easy to verify that

$$\mathbb{E}_{\mathbf{x}} \int_0^\infty e^{-\rho t} \|\boldsymbol{\xi}_t\|^2 dt \leq K(1 + \|\mathbf{x}\|^2), \quad (3.8)$$

and

$$|v(\mathbf{x}) - v(\mathbf{x}')| \leq \sup \{|J(\mathbf{x}; \boldsymbol{\xi}) - J(\mathbf{x}'; \boldsymbol{\xi})| : \boldsymbol{\xi} \in \mathcal{U}_{\mathbf{x}} \cup \mathcal{U}_{\mathbf{x}'}\}, \quad \forall \mathbf{x}, \mathbf{x}' \in \mathbb{R}^N.$$

Meanwhile,

$$|J(\mathbf{x}; \boldsymbol{\xi}) - J(\mathbf{x}'; \boldsymbol{\xi})| \leq \mathbb{E} \int_0^\infty e^{-\rho t} |H(\mathbf{X}_t^{\mathbf{x}}) - H(\mathbf{X}_t^{\mathbf{x}'})| dt.$$

Statement (ii) for v follows by Assumption **A2**, along with the facts that $\mathbf{X}_t^{\mathbf{x}} - \mathbf{X}_t^{\mathbf{x}'} = \mathbf{x} - \mathbf{x}'$ and that for any $\boldsymbol{\xi} \in \mathcal{U}_{\mathbf{x}} \cup \mathcal{U}_{\mathbf{x}'}$,

$$\begin{aligned}\mathbb{E}_{\mathbf{x}} \int_0^\infty e^{-\rho t} \|\mathbf{X}_t^{\mathbf{x}}\| dt &\leq K(1 + \|\mathbf{x}\| + \|\mathbf{x}'\|), \\ \mathbb{E}_{\mathbf{x}'} \int_0^\infty e^{-\rho t} \|\mathbf{X}_t^{\mathbf{x}'}\| dt &\leq K(1 + \|\mathbf{x}\| + \|\mathbf{x}'\|).\end{aligned}\tag{3.9}$$

In fact, if $\boldsymbol{\xi} \in \mathcal{U}_{\mathbf{x}}$, (3.9) follows immediately from (3.8) by the Hölder inequality. Meanwhile, if $\boldsymbol{\xi} \in \mathcal{U}_{\mathbf{x}'}$, (3.9) holds because

$$\|\mathbf{X}_t^{\mathbf{x}}\| \leq \|\mathbf{X}_t^{\mathbf{x}'}\| + \|\mathbf{x} - \mathbf{x}'\| \leq \|\mathbf{X}_t^{\mathbf{x}'}\| + \|\mathbf{x}\| + \|\mathbf{x}'\|.$$

□

Next, we establish the viscosity property of the value function in the following sense.

Definition 3 (Continuous viscosity solution). *The value function v for problem (**Regulator**) is a continuous viscosity solution to (3.4) on \mathbb{R}^N if*

- $\forall \mathbf{x}_0 \in \mathbb{R}^N, \forall \phi \in \mathcal{C}^2(\mathbb{R}^N)$ such that \mathbf{x}_0 is a local minimum of $(v - \phi)(\mathbf{x})$ with $v(\mathbf{x}_0) = \phi(\mathbf{x}_0)$,

$$\max\{\rho\phi - \mathcal{L}\phi - H(\mathbf{x}), \beta(\nabla\phi) - 1\} \geq 0.$$
- $\forall \mathbf{x}_0 \in \mathbb{R}^N, \forall \phi \in \mathcal{C}^2(\mathbb{R}^N)$ such that \mathbf{x}_0 is a local maximum of $(v - \phi)(\mathbf{x})$ with $v(\mathbf{x}_0) = \phi(\mathbf{x}_0)$,

$$\max\{\rho\phi - \mathcal{L}\phi - H(\mathbf{x}), \beta(\nabla\phi) - 1\} \leq 0.$$

Theorem 4 (Viscosity solution). *Under Assumptions **A1** and **A3**, the value function v to the control problem (**Regulator**) is a continuous viscosity solution of the HJB equation (3.4).*

Proof. The convexity of v follows from the joint convexity of $J(\mathbf{x}; \boldsymbol{\xi})$ in the following sense:

$$J(\theta\mathbf{x} + (1 - \theta)\mathbf{x}'; \theta\boldsymbol{\xi} + (1 - \theta)\boldsymbol{\xi}') \leq \theta J(\mathbf{x}; \boldsymbol{\xi}) + (1 - \theta)J(\mathbf{x}'; \boldsymbol{\xi}'),\tag{3.10}$$

for any $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^N$ and any $\boldsymbol{\xi}, \boldsymbol{\xi}' \in \mathcal{U}_N$. The convexity of J in $(\mathbf{x}; \boldsymbol{\xi})$ is then obvious since $\mathbf{X}_t^{\mathbf{x}}$ depends linearly on $(\mathbf{x}, \boldsymbol{\xi})$ and the set \mathcal{U}_N and the function H are both convex.

We now show that v is both a viscosity super-solution and sub-solution to the HJB equation (3.4).

Sub-solution. Consider the following controls: $\xi_t^{i,-} = 0$ and

$$\xi_t^{i,+} = \begin{cases} 0, & t = 0, \\ \eta^{i,+}, & t \geq 0, \end{cases}$$

where $0 \leq \eta^{i,+} \leq \epsilon$. Define the exit time

$$\tau_\epsilon := \inf\{t \geq 0, \mathbf{X}_t \notin \bar{B}_\epsilon(\mathbf{x}_0)\}.$$

Note that \mathbf{X} has at most one jump at $t = 0$ and is continuous on $[0, \tau_\epsilon)$. By the dynamic programming principle,

$$\phi(\mathbf{x}_0) = v(\mathbf{x}_0) \leq \mathbb{E}_{\mathbf{x}_0} \int_0^{\tau_\epsilon \wedge h} e^{-\rho t} \left[H(\mathbf{X}_t) dt + \sum_{i=1}^N L_i K_i^+ d\xi_t^{i,+} \right] + \mathbb{E}_{\mathbf{x}_0} [e^{-\rho(\tau_\epsilon \wedge h)} \phi(\mathbf{X}_{\tau_\epsilon \wedge h})].\tag{3.11}$$

Applying Itô's formula to the process $e^{-\rho t}\phi(\mathbf{X}_t)$ between 0 and $\tau_\epsilon \wedge h$, and taking expectation, we obtain

$$\begin{aligned} \mathbb{E}_{\mathbf{x}_0} [e^{-\rho(\tau_\epsilon \wedge h)}\phi(\mathbf{X}_{\tau_\epsilon \wedge h})] &= \phi(\mathbf{x}_0) + \mathbb{E}_{\mathbf{x}_0} \left[\int_0^{\tau_\epsilon \wedge h} e^{-\rho t} (-\rho\phi + \mathcal{L}\phi)(\mathbf{X}_t) dt \right] \\ &+ \mathbb{E}_{\mathbf{x}_0} \left[\sum_{0 \leq t \leq \tau_\epsilon \wedge h} [\phi(\mathbf{X}_t) - \phi(\mathbf{X}_{t-})] \right]. \end{aligned} \quad (3.12)$$

Combining (3.11) and (3.12), we have

$$\begin{aligned} \mathbb{E}_{\mathbf{x}_0} \left[\int_0^{\tau_\epsilon \wedge h} e^{-\rho t} (\rho\phi - \mathcal{L}\phi - H)(\mathbf{X}_t) dt \right] &- \mathbb{E}_{\mathbf{x}_0} \left[\int_0^{\tau_\epsilon \wedge h} e^{-\rho t} \left(\sum_{i=1}^N L_i K_i^+ d\xi_t^{i,+} \right) \right] \\ &- \mathbb{E}_{\mathbf{x}_0} \left[\sum_{0 \leq t \leq \tau_\epsilon \wedge h} \phi(\mathbf{X}_t) - \phi(\mathbf{X}_{t-}) \right] \leq 0. \end{aligned} \quad (3.13)$$

- Taking first $\eta^{i,+} = 0$ for all $i = 1, 2, \dots, N$, i.e., $\xi^{i,+} = \xi^{i,-} = 0$, we see that \mathbf{X} is continuous and that only the first term in the LHS of (3.13) is nonzero. Dividing the above inequality (3.13) by h and letting $h \rightarrow 0$, then by the dominated convergence theorem,

$$\rho\phi(\mathbf{x}_0) - \mathcal{L}\phi(\mathbf{x}_0) - H(\mathbf{x}_0) \leq 0.$$

- Now, by taking $\eta^{i,+} > 0$ and $\eta^{j,+} = 0$ for $j \neq i$ in (3.13), and noting that $\xi^{i,+}$ and \mathbf{X} jump only at $t = 0$ with size $\eta^{i,+}$, we get

$$\left[\int_0^{\tau_\epsilon \wedge h} e^{-\rho t} (\rho\phi - \mathcal{L}\phi - H)(\mathbf{X}_t) dt \right] - L_i K_i^+ \eta^{i,+} - \phi(\mathbf{x}_0 + \eta^{i,+} \mathbf{e}_i) + \phi(\mathbf{x}_0) \leq 0.$$

Taking $h \rightarrow 0$, then dividing by $\eta^{i,+}$ and letting $\eta \rightarrow 0$, we have

$$-L_i K_i^+ \leq \partial_{x^i} \phi(\mathbf{x}).$$

- Meanwhile, taking an admissible control such that $\xi^{i,+} = 0$ and

$$\xi_t^{i,-} = \begin{cases} 0, & t = 0, \\ \eta^{i,-}, & t \geq 0, \end{cases}$$

where $0 \leq \eta^{i,-} \leq \epsilon$. By a similar argument, we have

$$\forall i = 1, 2, \dots, N, \quad \partial_{x^i} \phi(\mathbf{x}) \leq L_i K_i^-.$$

This proves the sub-solution viscosity property

$$\max\{\rho\phi - \mathcal{L}\phi - H(x), \beta(\nabla\phi) - 1\} \leq 0.$$

Super-solution. This part is proved by contradiction. Suppose the claim is not true. Then there exist $\mathbf{x}_0 \in \mathbb{R}^N$, $\epsilon > 0$, $\phi(\mathbf{x}) \in \mathcal{C}^2(\mathbb{R}^N)$ with $\phi(\mathbf{x}_0) = v(\mathbf{x}_0)$, $v \geq \phi$ in $\bar{B}_\epsilon(\mathbf{x}_0)$ and $\nu > 0$ such that for all $\mathbf{x} \in \bar{B}_\epsilon(\mathbf{x}_0)$,

$$\rho\phi(\mathbf{x}_0) - \mathcal{L}\phi(\mathbf{x}_0) - H(\mathbf{x}_0) \leq -\nu, \quad (3.14)$$

and

$$-L_i K_i^+ + \nu \leq \partial_{x^i} \phi \leq L_i K_i^- - \nu. \quad (3.15)$$

Given any admissible control ξ , consider the exit time $\tau_\epsilon = \inf\{t \geq 0, \mathbf{X}_t \notin \bar{B}_\epsilon(\mathbf{x}_0)\}$. Applying Itô's formula (Meyer, 1976, Theorem 21) to $e^{-\rho t} \phi(\mathbf{x})$ and any semi-martingale $\{\mathbf{X}_t\}_{t \geq 0}$ under admissible control $(\xi^{i,+}, \xi^{i,-})_{i=1}^N$ leads to

$$\begin{aligned} \mathbb{E}_{\mathbf{x}_0} [e^{-\rho \tau_\epsilon} \phi(\mathbf{X}_{\tau_\epsilon-})] &= \phi(\mathbf{x}_0) + \mathbb{E}_{\mathbf{x}_0} \left[\int_0^{\tau_\epsilon} e^{-\rho t} (-\rho \phi + \mathcal{L} \phi)(\mathbf{X}_t) dt \right] \\ &+ \mathbb{E}_{\mathbf{x}_0} \left[\int_0^{\tau_\epsilon} e^{-\rho t} \sum_{i=1}^N \partial_{x^i} \phi(\mathbf{X}_t) [(d\xi_t^{i,+})^c - (d\xi_t^{i,-})^c] \right] \\ &+ \mathbb{E}_{\mathbf{x}_0} \left[\sum_{0 \leq t < \tau_\epsilon} e^{-\rho t} [\phi(\mathbf{X}_t) - \phi(\mathbf{X}_{t-})] \right]. \end{aligned}$$

Note that for all $0 \leq t < \tau_\epsilon$, $\mathbf{X}_t \in \bar{B}_\epsilon(\mathbf{x}_0)$. Then, by (3.14), and noting that $\Delta X_t^i = \Delta \xi_t^{i,+} - \Delta \xi_t^{i,-}$, we have for all $0 \leq t < \tau_\epsilon$,

$$\phi(\mathbf{X}_t) - \phi(\mathbf{X}_{t-}) = \sum_{i=1}^N \Delta X_t^i \int_0^1 \partial_{x^i} \phi(\mathbf{X}_t + z \Delta \mathbf{X}_t) dz \leq \sum_{i=1}^N [(L_i K_i^- - \nu) \Delta \xi_t^{i,+} + (L_i K_i^+ - \nu) \Delta \xi_t^{i,-}].$$

Similarly,

$$\phi(\mathbf{X}_t) - \phi(\mathbf{X}_{t-}) \geq \sum_{i=1}^N [-(L_i K_i^- - \nu) \Delta \xi_t^{i,-} - (L_i K_i^+ - \nu) \Delta \xi_t^{i,+}]. \quad (3.16)$$

In light of relations (3.14)-(3.16),

$$\begin{aligned} \mathbb{E}_{\mathbf{x}_0} [e^{-\rho \tau_\epsilon} \phi(\mathbf{X}_{\tau_\epsilon-})] &\geq \phi(\mathbf{x}_0) + \mathbb{E}_{\mathbf{x}_0} \left[\int_0^{\tau_\epsilon} e^{-\rho t} (-H + \nu)(\mathbf{X}_t) dt \right] \\ &+ \mathbb{E}_{\mathbf{x}_0} \left[\int_0^{\tau_\epsilon-} e^{-\rho t} \sum_{i=1}^N -(L_i K_i^+ - \nu) d\xi_t^{i,+} - (L_i K_i^- - \nu) d\xi_t^{i,-} \right] \\ &= \phi(\mathbf{x}_0) - \mathbb{E}_{\mathbf{x}_0} \int_0^{\tau_\epsilon} e^{-\rho t} \left[H(\mathbf{X}_t) dt + \sum_{i=1}^N L_i K_i^+ d\xi_t^{i,+} + \sum_{i=1}^N L_i K_i^- d\xi_t^{i,-} \right] \\ &+ \sum_{i=1}^N (\mathbb{E}_{\mathbf{x}_0} [e^{-\rho \tau_\epsilon} L_i K_i^+ \Delta \xi_{\tau_\epsilon}^{i,+}] + \mathbb{E}_{\mathbf{x}_0} [e^{-\rho \tau_\epsilon} L_i K_i^- \Delta \xi_{\tau_\epsilon}^{i,-}]) \\ &+ \nu \left\{ \mathbb{E}_{\mathbf{x}_0} \left[\int_0^{\tau_\epsilon} e^{-\rho t} dt \right] + \mathbb{E}_{\mathbf{x}_0} \left[\int_0^{\tau_\epsilon-} e^{-\rho t} (d\xi_t^{i,+} + d\xi_t^{i,-}) \right] \right\}. \quad (3.17) \end{aligned}$$

Note that $\mathbf{X}_{\tau_\epsilon-} \in \bar{B}_\epsilon(\mathbf{x}_0)$, $\mathbf{X}_{\tau_\epsilon}$ is either on the boundary $\partial B_\epsilon(\mathbf{x}_0)$ or out of $\bar{B}_\epsilon(\mathbf{x}_0)$. However, there is some random variable δ valued in $[0, 1]$ such that

$$\mathbf{x}_\delta = \mathbf{X}_{\tau_\epsilon-} + \delta \Delta \mathbf{X}_{\tau_\epsilon} = \mathbf{X}_{\tau_\epsilon-} + \delta (\Delta \xi_{\tau_\epsilon}^+ - \Delta \xi_{\tau_\epsilon}^-) \in \partial \bar{B}_\epsilon(\mathbf{x}_0).$$

Then similar to (3.16), we have

$$\phi(\mathbf{x}_\delta) - \phi(\mathbf{X}_{\tau_\epsilon-}) \geq \delta \sum_{i=1}^N [-(L_i K_i^- - \nu) \Delta \xi_{\tau_\epsilon}^{i,-} - (L_i K_i^+ - \nu) \Delta \xi_{\tau_\epsilon}^{i,+}]. \quad (3.18)$$

Note that $\mathbf{X}_{\tau_\epsilon} = \mathbf{x}_\delta + (1 - \delta)(\Delta\xi_t^+ - \Delta\xi_t^-)$, thus

$$v(\mathbf{x}_\delta) \leq (1 - \delta) \sum_{i=1}^N (L_i K_i^+ \Delta\xi_{\tau_\epsilon}^{i,+} + L_i K_i^- \Delta\xi_{\tau_\epsilon}^{i,-}) + v(\mathbf{X}_{\tau_\epsilon}). \quad (3.19)$$

Recalling that $v(\mathbf{x}_\delta) \geq \phi(\mathbf{x}_\delta)$, inequalities (3.18)-(3.19) imply

$$(1 - \delta) \sum_{i=1}^N (L_i K_i^+ \Delta\xi_{\tau_\epsilon}^{i,+} + L_i K_i^- \Delta\xi_{\tau_\epsilon}^{i,-}) + v(\mathbf{X}_{\tau_\epsilon}) \geq \phi(\mathbf{X}_{\tau_\epsilon-}) + \delta \sum_{i=1}^N [-(L_i K_i^- - \nu) \Delta\xi_{\tau_\epsilon}^{i,-} - (L_i K_i^+ - \nu) \Delta\xi_{\tau_\epsilon}^{i,+}].$$

Therefore,

$$\sum_{i=1}^N ((L_i K_i^+ - \delta\nu) \Delta\xi_{\tau_\epsilon}^{i,+} + (L_i K_i^- - \delta\nu) \Delta\xi_{\tau_\epsilon}^{i,-}) + v(\mathbf{X}_{\tau_\epsilon}) \geq \phi(\mathbf{X}_{\tau_\epsilon-}).$$

Plugging the last inequality into (3.17), along with $\phi(\mathbf{x}_0) = v(\mathbf{x}_0)$, yields

$$\begin{aligned} & \mathbb{E}_{\mathbf{x}_0} e^{-\tau_\epsilon} \left[\sum_{i=1}^N ((L_i K_i^+ - \delta\nu) \Delta\xi_{\tau_\epsilon}^{i,+} + (L_i K_i^- - \delta\nu) \Delta\xi_{\tau_\epsilon}^{i,-}) + v(\mathbf{X}_{\tau_\epsilon}) \right] \\ & \geq v(\mathbf{x}_0) - \mathbb{E}_{\mathbf{x}_0} \int_0^{\tau_\epsilon} e^{-\rho t} \left[H(\mathbf{X}_t) dt + \sum_{i=1}^N L_i K_i^+ d\xi_t^{i,+} + \sum_{i=1}^N L_i K_i^- d\xi_t^{i,-} \right] \\ & + \sum_{i=1}^N (\mathbb{E}_{\mathbf{x}_0} [e^{-\rho\tau_\epsilon} L_i K_i^+ \Delta\xi_{\tau_\epsilon}^{i,+}] + \mathbb{E}_{\mathbf{x}_0} [e^{-\rho\tau_\epsilon} L_i K_i^- \Delta\xi_{\tau_\epsilon}^{i,-}]) \\ & + \nu \left\{ \mathbb{E}_{\mathbf{x}_0} \left[\int_0^{\tau_\epsilon} e^{-\rho t} dt \right] + \mathbb{E}_{\mathbf{x}_0} \left[\int_0^{\tau_\epsilon-} e^{-\rho t} (d\xi_t^{i,+} + d\xi_t^{i,-}) \right] \right\}. \end{aligned}$$

Hence

$$\begin{aligned} & \mathbb{E}_{\mathbf{x}_0} e^{-\tau_\epsilon} v(\mathbf{X}_{\tau_\epsilon}) + \mathbb{E}_{\mathbf{x}_0} \int_0^{\tau_\epsilon} e^{-\rho t} \left[H(\mathbf{X}_t) dt + \sum_{i=1}^N L_i K_i^+ d\xi_t^{i,+} + \sum_{i=1}^N L_i K_i^- d\xi_t^{i,-} \right] \\ & \geq v(\mathbf{x}_0) + \nu \left\{ \mathbb{E}_{\mathbf{x}_0} \left[\int_0^{\tau_\epsilon} e^{-\rho t} dt \right] + \mathbb{E}_{\mathbf{x}_0} \left[\int_0^{\tau_\epsilon-} e^{-\rho t} (d\xi_t^{i,+} + d\xi_t^{i,-}) \right] + \delta \mathbb{E}_{\mathbf{x}_0} [e^{-\tau_\epsilon} \Delta\xi_{\tau_\epsilon}^{i,+} + e^{-\tau_\epsilon} \Delta\xi_{\tau_\epsilon}^{i,-}] \right\}. \end{aligned}$$

We now claim that there exists a constant $g_0 > 0$ such that for all admissible control ξ ,

$$\mathbb{E}_{\mathbf{x}_0} \left[\int_0^{\tau_\epsilon} e^{-\rho t} dt \right] + \mathbb{E}_{\mathbf{x}_0} \left[\int_0^{\tau_\epsilon-} e^{-\rho t} (d\xi_t^{i,+} + d\xi_t^{i,-}) \right] + \delta \mathbb{E}_{\mathbf{x}_0} [e^{-\tau_\epsilon} \Delta\xi_{\tau_\epsilon}^{i,+} + e^{-\tau_\epsilon} \Delta\xi_{\tau_\epsilon}^{i,-}] \geq g_0. \quad (3.20)$$

Indeed, one can always find some constant G_0 such that the \mathcal{C}^2 function

$$\psi(\mathbf{x}) = G_0((\mathbf{x} - \mathbf{x}_0)^2 - \epsilon^2)$$

satisfies

$$\begin{cases} \min_i \{\rho\psi - \mathcal{L}\psi + 1, 1 - |\partial_{x_i}\psi|\} \geq 0, & \text{on } \overline{B}_\epsilon(\mathbf{x}_0), \\ \psi = 0 & \text{on } \partial\overline{B}_\epsilon(\mathbf{x}_0). \end{cases}$$

Applying Meyer's version of Itô's formula (Meyer, 1976, Theorem 21) to $e^{-\rho t}\psi(\mathbf{x})$ and any semimartingale $\{\mathbf{X}_t\}_{t \geq 0}$ under admissible control $(\xi^{i,+}, \xi^{i,-})_{i=1}^N$ leads to

$$\mathbb{E}_{\mathbf{x}_0} [e^{-\rho\tau_\epsilon} \psi(\mathbf{X}_{\tau_\epsilon-})] \leq \psi(\mathbf{x}_0) + \mathbb{E}_{\mathbf{x}_0} \left[\int_0^{\tau_\epsilon} e^{-\rho t} dt \right] + \sum_{i=1}^N \mathbb{E}_{\mathbf{x}_0} \left[\int_0^{\tau_\epsilon-} e^{-\rho t} (d\xi_t^{i,+} + d\xi_t^{i,-}) \right]. \quad (3.21)$$

Since $\psi_{x^i}(\mathbf{x}_0) \geq -1$ for all $i = 1, 2, \dots, N$,

$$\psi(\mathbf{X}_{\tau_\epsilon^-}) - \psi(\mathbf{x}_\delta) \geq -\nabla\psi(\mathbf{X}_{\tau_\epsilon^-} - \mathbf{x}_\delta) \geq -\delta \sum_{i=1}^N \Delta \xi_{\tau_\epsilon}^{i,-},$$

which, combined with (3.21), yields

$$\begin{aligned} & \mathbb{E}_{\mathbf{x}_0} \left[\int_0^{\tau_\epsilon} e^{-\rho t} dt \right] + \sum_{i=1}^N \mathbb{E}_{\mathbf{x}_0} \left[\int_0^{\tau_\epsilon^-} (d\xi_t^{i,+} + d\xi_t^{i,-}) \right] + \mathbb{E}_{\mathbf{x}_0} \left[e^{-\rho\tau_\epsilon} \delta \sum_{i=1}^N \Delta \xi_{\tau_\epsilon}^{i,-} \right] \\ & \geq \mathbb{E}_{\mathbf{x}_0} [e^{-\rho\tau_\epsilon} \psi(\mathbf{x}_\delta)] - \psi(\mathbf{x}_0) = G_0 \epsilon^2. \end{aligned}$$

Hence (3.20) holds with $g_0 = G_0 \epsilon^2$. □

We can further show that the value function is the unique and $\mathcal{W}_{loc}^{2,\infty}(\mathbb{R}^N)$ solution to the HJB equation (3.4).

Theorem 5 (Regularity and uniqueness.). *Under Assumptions **A3**, the value function v to the control problem (**Regulator**) is the unique $\mathcal{W}_{loc}^{2,\infty}(\mathbb{R}^N)$ solution to the HJB equation (3.4). In addition, there exists $K > 0$ such that*

$$0 \leq \partial_{\mathbf{z}}^2 v(\mathbf{x}) \leq K, \text{ a.e. for } \mathbf{x} \in \mathbb{R}^N, \quad (3.22)$$

for any second order directional derivative $\partial_{\mathbf{z}}^2$. Finally, v is $\mathcal{C}^{4,\alpha}$ in the continuation region

$$\mathcal{C}_N := \{\mathbf{x} \mid \beta(\nabla v(\mathbf{x})) < 1\}, \quad (3.23)$$

and \mathcal{C}_N is bounded.

Proof. To prove $v \in \mathcal{W}_{loc}^{2,\infty}$, let B be any open ball and let $\psi \in C_0^\infty(\mathbb{R}^N)$ be any test function with a support contained in B . Since $(\Delta x^i)^{-2} \delta_i^2 v(\mathbf{x})$ is bounded on B for $|\Delta x^i| \leq 1$, there is a sequence $\eta_k \rightarrow 0+$ as $k \rightarrow \infty$ such that, denoting by g_k the result of replacing Δx^i by η_k in $(\Delta x^i)^{-2} \delta_i^2 v(\mathbf{x})$, we have $g_k \rightarrow Q$ weakly in $L^p(B)$ for some p with $1 < p < \infty$. It is then easy to see that

$$\int_{\mathbb{R}^N} \psi(\mathbf{x}) Q(\mathbf{x}) d\mathbf{x} = \int_{\mathbb{R}^N} \partial_{x^i}^2 \psi v(\mathbf{x}) d\mathbf{x}, \quad \forall \psi \in C_0^\infty(B), \quad (3.24)$$

where $Q = \partial_{x^i}^2 v$. The existence and local boundedness of second order derivatives is now immediate: for $k = 1, 2, \dots, N$, let \mathbf{e}_k denote the unit vector in the direction of the positive x_k axis; for any fixed $i \neq j$ with $1 \leq i, j \leq N$, let \mathbf{y} be a new coordinate whose axis points to the $\frac{\mathbf{e}_i + \mathbf{e}_j}{\sqrt{2}}$ direction, then $\partial_{x_i x_j}^2 v = \partial_{\mathbf{y}}^2 v - \frac{1}{2}(\partial_{x^i}^2 v + \partial_{x^j}^2 v)$.

To show that v is the unique solution to HJB, we proceed by a contradiction argument. Suppose v_1 and v_2 are two non-negative solutions. Let \mathbf{y}_0 be the point where v_2 attains its minimum value. Given $\delta > 0$, define

$$\phi_\delta(\mathbf{x}) := v_1(\mathbf{x}) - v_2(\mathbf{x}) - \delta \|\mathbf{x} - \mathbf{y}_0\|^2, \quad \forall \mathbf{x} \in \mathbb{R}^N.$$

The function ϕ_δ attains its maximum at some $x_\delta \in \mathbb{R}^N$ and

$$0 = \nabla \phi_\delta(\mathbf{x}_\delta) = \nabla v_1(\mathbf{x}) - \nabla v_2(\mathbf{x}) - 2\delta(\mathbf{x}_\delta - \mathbf{y}_0). \quad (3.25)$$

This leads to

$$\nabla v_1(\mathbf{x}_\delta) = \nabla v_2(\mathbf{x}_\delta) + 2\delta(\mathbf{x}_\delta - \mathbf{y}_0).$$

Consequently,

$$1 \geq \beta(\nabla v_1(\mathbf{x}_\delta)) = \beta(\nabla v_2(\mathbf{x}_\delta) + 2\delta(\mathbf{x}_\delta - \mathbf{y}_0)).$$

Since y_0 is the minimal point of v_2 , we have

$$\nabla v_2(\mathbf{x}_\delta) \cdot (\mathbf{x}_\delta - \mathbf{y}_0) \geq 0.$$

This means that either $\beta(\Delta v_2(\mathbf{x}_\delta)) < 1$, or for any $i \in \arg \max \beta(\nabla v_2(\mathbf{x}_\delta))$, $(\mathbf{x}_\delta - \mathbf{y}_0)^i = 0$. Suppose the latter, then by (3.25),

$$0 = \partial_{x^i} v_1(\mathbf{x}_\delta) - \partial_{x^i} v_2(\mathbf{x}_\delta) - 2\delta(\mathbf{x}_\delta - \mathbf{y}_0)^i.$$

Hence

$$\partial_{x^i} v_1(\mathbf{x}_\delta) = \partial_{x^i} v_2(\mathbf{x}_\delta) = 0,$$

for $i \in \arg \max \beta(\nabla v_2(\mathbf{x}_\delta))$. This implies $\beta(\nabla v_2(\mathbf{x}_\delta)) < 1$. Meanwhile from (3.4), we know

$$\Delta v_2(\mathbf{x}_\delta) = v_2(\mathbf{x}_\delta) - H(\mathbf{x}_\delta).$$

By Bony's maximum principle (Lions, 1983),

$$\begin{aligned} 0 &\geq \liminf \operatorname{ess}_{\mathbf{x} \rightarrow \mathbf{x}_\delta} \Delta \phi_\delta(\mathbf{x}) \\ &= \liminf \operatorname{ess}_{\mathbf{x} \rightarrow \mathbf{x}_\delta} \Delta v_1(\mathbf{x}) - \Delta v_2(\mathbf{x}) - 4\delta \\ &\geq v_1(\mathbf{x}_\delta) - v_2(\mathbf{x}_\delta) - 4\delta. \end{aligned}$$

It follows that for any $\mathbf{x} \in \mathbb{R}^N$,

$$v_1(\mathbf{x}) - v_2(\mathbf{x}) = \phi_\delta(\mathbf{x}) + \delta \|\mathbf{x} - \mathbf{y}_0\|^2 \leq \phi_\delta(\mathbf{x}_\delta) + \delta \|\mathbf{x} - \mathbf{y}_0\|^2 \leq \delta(4 + \|\mathbf{x} - \mathbf{y}_0\|^2).$$

Letting $\delta \rightarrow 0$, we have $v_1(\mathbf{x}) \leq v_2(\mathbf{x})$. Similarly, we have $v_2(\mathbf{x}) \leq v_1(\mathbf{x})$.

Since $|\partial_{x^i} v(\mathbf{x})| \leq L_i \max\{K_i^+, K_i^-\}$ ($i = 1, 2, \dots, N$) on \mathbb{R}^N but H grows at least quadratically by Assumption **A3**, \mathcal{C}_N must be bounded.

Finally, let B be any open ball such that $\bar{B} \in \mathcal{C}_N$. By Theorem 6.13 in (Gilbarg & Trudinger, 2015), the Dirichlet problem in B ,

$$\begin{cases} \rho \tilde{v} - \mathcal{L} \tilde{v} = H(\mathbf{x}), & \forall \mathbf{x} \in B, \\ \tilde{v} = v, & \forall \mathbf{x} \in \partial B, \end{cases} \quad (3.26)$$

has a solution $\tilde{v} \in \mathcal{C}^0(\bar{B}) \cap \mathcal{C}^{2,\alpha}(B)$. In particular, $\tilde{v} - v \in \mathcal{W}^{2,\infty}(B)$, therefore by (3.26), $\tilde{v} - v \in \mathcal{W}_0^{1,2}(B)$. By Theorem 8.9 of (Gilbarg & Trudinger, 2015), $v = \tilde{v}$ in B , thus $v \in \mathcal{C}^{2,\alpha}(B)$. By Theorem 6.17 of (Gilbarg & Trudinger, 2015), $v \in \mathcal{C}^{4,\alpha}(B)$ thus $v \in \mathcal{C}^{4,\alpha}(\mathcal{C}_N)$ for all $\alpha \in (0, 1)$.

To prove (3.22), let $\Delta_i \mathbf{x} := (0, \dots, 0, \Delta x^i, 0, \dots, 0)$ be the N -dimensional row vector with the i -th entry being Δx^i for $i = 1, 2, \dots, N$. For any function $F : \mathbb{R}^N \rightarrow \mathbb{R}$, define the second difference of F in the x^i direction by

$$\delta_i^2 F(\mathbf{x}) = F(\mathbf{x} + \Delta_i \mathbf{x}) + F(\mathbf{x} - \Delta_i \mathbf{x}) - 2F(\mathbf{x}). \quad (3.27)$$

It is easy to check that

$$\delta_i^2 v(\mathbf{x}) \leq \sup\{\delta_i^2 J(\mathbf{x}; \boldsymbol{\xi}) : \boldsymbol{\xi} \in \mathcal{U}_\mathbf{x}\}. \quad (3.28)$$

Since $H \in \mathcal{C}^2(\mathbb{R}^N)$, for $\mathbf{x} \in \mathbb{R}^N$,

$$\delta_i^2 H(\mathbf{x}) = (\Delta x^i)^2 \int_0^1 \int_{-\lambda}^{\lambda} \partial_{x^i}^2 H(x^1, \dots, x^i + \mu \Delta x^i, \dots, x^N) d\mu d\lambda. \quad (3.29)$$

By Assumption **A3**,

$$\delta_i^2 H(\mathbf{x}) \leq K(\Delta x^i)^2 \int_0^1 \int_{-\lambda}^{\lambda} d\mu d\lambda = (\Delta x^i)^2 K. \quad (3.30)$$

Hence

$$0 \leq \delta_i^2 v(\mathbf{x}) \leq K(\Delta x^i)^2, \quad \mathbf{x} \in \mathbb{R}^N, |\Delta x^i| \leq 1. \quad (3.31)$$

The lower bound of (3.31) follows from the convexity of v by Theorem 4. □

Remark 6. The proof of Theorem 5 is inspired by the approach in (Soner & Shreve, 1989, Theorem 4.5) and (Williams et al., 1994, Theorem 3.1). In (Soner & Shreve, 1989), the following HJB equation (3.32) (See Eqn. 3.1 in (Soner & Shreve, 1989)) was studied for an N -dimensional control problem

$$\max \left\{ \rho u - \mathcal{L}u - H(\mathbf{x}), \sqrt{\sum_{i=1}^N (\partial_{x^i} u)^2} - 1 \right\} = 0. \quad (3.32)$$

Comparing the gradient constraints in (3.32) with (3.4), it is clear that the operator β in (3.4) is less regular than $\|\nabla u\|_2$ in (3.32) as $\|\nabla u(\cdot)\|_2$ has smoother and gradual changes in the state space \mathbb{R}^N . In contrast, β in (3.4) involves a maximum operator as a result of game interactions.

The very same HJB equation (3.4) has appeared in Menaldi & Taksar (1989) for analyzing the convergence of finite variation controls from controls of bounded velocity. However, no characterization of the optimal control nor regularity analysis has been studied.

4 Pareto-optimal policies

The regularity analysis of the value function for problem (**Regulator**) enables us to establish the existence and the uniqueness of its optimal control, for any given weight (L_1, \dots, L_N) such that $L_i > 0$ and $\sum_{i=1}^N L_i = 1$ (Section 4.1). The optimal control in (**Regulator**) is then shown to lead to a Pareto-optimal policy for game (**N-player**) (Theorem 12) for each choice of weights (L_1, \dots, L_N) .

4.1 Optimal policy for the regulator

To ensure the uniqueness of the Pareto-optimal policy, we impose the following assumption on the value function v .

A4. The diagonal dominates the row/column in the Hessian matrix $\nabla^2 v$. That is,

$$\partial_{x^i}^2 v(\mathbf{x}) > \sum_{j \neq i} |\partial_{x^i x^j}^2 v(\mathbf{x})|, \quad \forall i, = 1, 2, \dots, N \quad \text{and} \quad \mathbf{x} \in \bar{\mathcal{C}}_N. \quad (4.1)$$

Note that a similar assumption was used in (Gomes et al., 2010, Assumption 3) to analyze Nash equilibrium strategies. This assumption guarantees that the reflection direction of the Skorokhod problem is not parallel to the boundary, and that the controlled dynamics are continuous when $\mathbf{x} \in \mathcal{C}_N$. Assumption **A4** may be relaxed using techniques of Kruk (2000) to deal with possible jumps at the reflection boundary.

Given this additional assumption and the regularity of the value function, we are now ready to characterize the Pareto-optimal policy to game (**N-player**).

We shall show that when $\mathbf{x} \in \bar{\mathcal{C}}_N$, the optimal policy can be constructed by formulating and solving a sequence of Skorokhod problems with piecewise \mathcal{C}^1 boundaries, and then by passing to the limit of this sequence of ϵ -optimal policies. We shall also show that the reflection field of the Skorokhod problem can be extended to the entire state space under appropriate conditions, completing the construction of the Pareto-optimal policy when \mathbf{x} is outside $\bar{\mathcal{C}}_N$.

Optimal policy for $\mathbf{x} \in \bar{\mathcal{C}}_N$. First, recall the definition of the Skorokhod problem Ramanan (2006).

Definition 7 (Skorokhod problem). *Let G be an open domain in \mathbb{R}^N with $S = \partial G$. Let $\Gamma(a, b) = \{x \in \mathbb{R}^N : |x - a| = b\}$. To each point $\mathbf{x} \in S$, we will associate a set $\mathbf{r}(\mathbf{x}) \subset \Gamma(0, 1)$ called the directions of reflection. We say that a continuous process*

$$\boldsymbol{\xi}_t = \int_0^t \mathbf{N}_s d\eta_s, \quad (4.2)$$

with $\eta_t = \bigvee_{[0,t]} \boldsymbol{\xi}$, is a solution to a Skorokhod problem with data $(\mathbf{x} + \boldsymbol{\mu}t + \boldsymbol{\sigma}\mathbf{B}_t, G, \mathbf{r}, \mathbf{x})$ if

- (a) $|\mathbf{N}_t| = 1$, η_t is continuous and nondecreasing;
- (b) the process $\mathbf{X}_t = \mathbf{x} + \boldsymbol{\mu}t + \boldsymbol{\sigma}\mathbf{B}_t + \int_0^t \mathbf{N}_s d\eta_s$ satisfies $\mathbf{X}_t \in \bar{G}$, $0 \leq t < \infty$, a.s.;
- (c) for every $0 \leq t < \infty$,

$$\eta_t = \int_0^t \mathbf{1}_{(\mathbf{X}_s \in \partial G, \mathbf{N}_s \in \mathbf{r}(\mathbf{X}_s))} d\eta_s.$$

Now let us introduce some notations for the Skorokhod problem associated with the continuation region \mathcal{C}_N defined in (3.23). By definition,

$$\mathcal{C}_N = \{\mathbf{x} \mid \beta(\nabla v(\mathbf{x})) < 1\} = \bigcap_{j=1}^{2N} G_j, \quad (4.3)$$

where for $i = 1, 2, \dots, N$,

$$G_i = \{\mathbf{x} \mid \partial_{x^i} v(\mathbf{x}) < L_i K_i^-\}, \quad G_{i+N} = \{\mathbf{x} \mid \partial_{x^i} v(\mathbf{x}) > -L_i K_i^+\}. \quad (4.4)$$

Denote $\mathcal{S} = \partial\mathcal{C}_N$ as the boundary of \mathcal{C}_N , denote $I(\mathbf{x}) = \{j \mid \mathbf{x} \notin G_j, j = 1, 2, \dots, 2N\}$ as the boundary that \mathbf{x} lines on, and define the vector field γ_j on each face G_j as

$$\gamma_i = -\mathbf{e}_i, \quad \gamma_{i+N} = \mathbf{e}_i, \quad (4.5)$$

where $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)$ with the i^{th} component being 1. Then the directions of the reflection is defined as

$$\mathbf{r}(\mathbf{x}) = \left\{ \sum_{j \in I(\mathbf{x})} c_j \gamma_j(\mathbf{x}) : c_i \geq 0 \text{ and } \left| \sum_{j \in I(\mathbf{x})} c_j \gamma_j(\mathbf{x}) \right| = 1 \right\}. \quad (4.6)$$

Theorem 8 (ϵ -policy). *Assume Assumptions **A1-A4** holds and $\mathbf{x} \in \mathcal{C}_N$. For any $\epsilon > 0$, there exists a unique solution to the Skorokhod problem with data $(\mathbf{x} + \boldsymbol{\mu}t + \boldsymbol{\sigma}\mathbf{B}_t, \mathcal{C}_\epsilon, \mathbf{r}_\epsilon, \mathbf{x})$ is an ϵ -optimal (admissible) policy of the control problem (**Regulator**) with*

$$\boldsymbol{\xi}_t^\epsilon = \int_0^t \mathbf{N}_s^\epsilon \cdot d\eta_s^\epsilon, \quad (4.7)$$

and $\mathbf{N}_s^\epsilon = \mathbf{r}_\epsilon(\mathbf{X}_s^\epsilon)$ on \mathcal{S}_ϵ , where $\mathbf{X}_t^\epsilon = \mathbf{x} + \boldsymbol{\mu}t + \boldsymbol{\sigma}\mathbf{B}_t + \boldsymbol{\xi}_t^\epsilon$. That is,

$$J(\mathbf{x}, \boldsymbol{\xi}^\epsilon) < v(\mathbf{x}) - C_0\epsilon,$$

for some constant C_0 that is independent of ϵ . Here $\mathcal{C}_\epsilon \subseteq \mathcal{C}$ has piecewise smooth boundaries.

Proof. The proof consists two steps. We first construct an approximation \mathcal{C}_ϵ of \mathcal{C}_N with piecewise \mathcal{C}^1 boundaries. Clearly, if $\partial\mathcal{C}_N$ itself is piecewise \mathcal{C}^1 , the $\mathcal{C}_\epsilon = \mathcal{C}_N$. We then show that the solution to the Skorokhod problem with piecewise smooth boundary provides an ϵ -policy to the regulator's control problem.

Step 1: Skorokhod problem with piecewise smooth boundary. Let $\phi^\delta(x) \in C^\infty(\mathbb{R}^N, \mathbb{R}_+)$ be such that $\phi^\delta(\mathbf{x}) = 0$ for $|\mathbf{x}| \geq \delta$ and

$$\int_{\mathbb{R}^N} \phi^\delta(\mathbf{x}) d\mathbf{x} = 1. \quad (4.8)$$

Since $v \in \mathcal{W}_{loc}^{2,\infty}(\mathbb{R}^N)$, consider a regularization of $v(\mathbf{x})$ via ϕ^ϵ , such that

$$v^\delta(\mathbf{x}) = \phi^\delta * v(\mathbf{x}). \quad (4.9)$$

The boundedness of v , ∇v , D^2v on $B_R(0)$, with $\bar{\mathcal{C}}_N \subset B_{R-1}(0)$, implies that H^δ , v^δ are bounded uniformly on $\bar{\mathcal{C}}_N$ for $\delta < 1$, and

$$v^\delta \rightarrow v, \quad \nabla v^\delta \rightarrow \nabla v, \quad H^\delta \rightarrow H \quad \text{uniformly in } \bar{\mathcal{C}}_N.$$

Denote $K_{\max} = \max_{i=1,2,\dots,N} \{L_i K_i^+, L_i K_i^-\}$, $K_{\min} = \min_{i=1,2,\dots,N} \{L_i K_i^+, L_i K_i^-\}$ and recall K in (3.22) such that $0 \leq \partial_{\mathbf{z}}^2 v(\mathbf{x}) \leq K$ for any second order directional derivative $\partial_{\mathbf{z}}^2$. Then, for any $\epsilon_k \in (0, \frac{1}{4})$, there exists $\delta_k := \delta_k(\epsilon_k) \in (0, \frac{\epsilon_k K_{\min}}{K})$ such that for all $\delta \in [0, \delta_k]$, $\|\nabla v^\delta - \nabla v\|_1 < K_{\min} \epsilon_k$. Take a non-negative and non-increasing sequence $\{\epsilon_k\}_k$ such that $\lim_{k \rightarrow \infty} \epsilon_k = 0$. Denote $w^{\delta_k}(\mathbf{x}) = \beta(\nabla v^{\delta_k}(\mathbf{x}))$ and $\mathcal{C}_{\epsilon_k} := \{\mathbf{x} \mid w^{\delta_k}(\mathbf{x}) < 1 - 2\epsilon_k\} = \cap_{j=1}^{2N} G_j^{\epsilon_k}$, where $i = 1, 2, \dots, N$,

$$\begin{aligned} G_i^{\epsilon_k} &= \{\mathbf{x} \mid \partial_{x^i} v^{\delta_k}(\mathbf{x}) < (1 - 2\epsilon_k) L_i K_i^-\}, \\ G_{i+N}^{\epsilon_k} &= \{\mathbf{x} \mid \partial_{x^i} v^{\delta_k}(\mathbf{x}) > (-1 + 2\epsilon_k) L_i K_i^+\}. \end{aligned} \quad (4.10)$$

Since $\|\nabla v^{\delta_k} - \nabla v\|_1 < K_{\min} \epsilon_k$ in \mathcal{C}_N and by the definition in (4.10), we have $\mathcal{C}_{\epsilon_k} \subset \mathcal{C}_N$. Also notice that $\partial G_j^{\epsilon_k} \cap \bar{\mathcal{C}}_{\epsilon_k} \in \mathcal{C}^2$ because v^{δ_k} is smooth. Now, take any ϵ from the sequence $\{\epsilon_k\}_k$, and denote $\mathcal{S}_\epsilon = \partial\mathcal{C}_\epsilon$ as the boundary of \mathcal{C}_ϵ , and $I_\epsilon(\mathbf{x}) = \{j \mid \mathbf{x} \notin G_j^\epsilon, j = 1, 2, \dots, 2N\}$. Define the vector field γ_j on each face G_j^ϵ as (4.5) and the directions of reflection by

$$\mathbf{r}_\epsilon(\mathbf{x}) = \left\{ \sum_{j \in I_\epsilon(\mathbf{x})} c_j \gamma_j(\mathbf{x}) : c_i \geq 0 \text{ and } \left| \sum_{j \in I_\epsilon(\mathbf{x})} c_j \gamma_j(\mathbf{x}) \right| = 1 \right\}. \quad (4.11)$$

When $\epsilon = 0$, denote $I(\mathbf{x}) := I_0(\mathbf{x})$ and $\mathbf{r}(\mathbf{x}) := \mathbf{r}_0(\mathbf{x})$ for the index set and reflection cone of region \mathcal{C}_N , respectively. Then define the normal direction on face G_j^ϵ as n_j^ϵ ($j = 1, 2, \dots, 2N$) with

$$\begin{aligned} n_i^\epsilon &= -\frac{\nabla(\partial_{x^i} v^\delta)}{\|\nabla(\partial_{x^i} v^\delta)\|_2}, \\ n_{i+N}^\epsilon &= \frac{\nabla(\partial_{x^i} v^\delta)}{\|\nabla(\partial_{x^i} v^\delta)\|_2}, \quad i = 1, 2, \dots, N. \end{aligned}$$

Note that the normal direction n_j^ϵ ($j = 1, 2, \dots, 2N$) is well-defined by the construction of (4.10).

Next we want to show that $n_i^\epsilon \cdot \gamma_i = \frac{\partial_{x^i}^2 v^\delta}{\|\nabla(\partial_{x^i} v^\delta)\|_2} > 0$ and $n_{i+N}^\epsilon \cdot \gamma_{i+N} = \frac{\partial_{x^i}^2 v^\delta}{\|\nabla(\partial_{x^i} v^\delta)\|_2} > 0$ for $i = 1, 2, \dots, N$.

To do so, we want to show that $B_\delta(\mathbf{x}) \in \mathcal{C}_N$ for $\mathbf{x} \in \mathcal{S}_\epsilon$. Note that $(-1 + 2\epsilon)L_i K_i^+ \leq \partial_{x^i} v(\mathbf{x}) \leq (1 - 2\epsilon)L_i K_i^-$ for $\mathbf{x} \in \bar{\mathcal{C}}_\epsilon$. For any $\mathbf{y} \in B_\delta(\mathbf{x})$, $|\partial_{x^i} v(\mathbf{x}) - \partial_{x^i} v(\mathbf{y})| \leq K\|\mathbf{x} - \mathbf{y}\| \leq K\delta \leq \epsilon K_{\min}$. Therefore, $(-1 + \epsilon)L_i K_i^+ \leq (-1 + 2\epsilon)L_i K_i^+ - \epsilon K_{\min} \leq \partial_{x^i} v(\mathbf{y}) \leq (1 - 2\epsilon)L_i K_i^- + \epsilon K_{\min} \leq (1 - \epsilon)L_i K_i^-$. Thus, $\mathbf{y} \in \mathcal{C}_N$ for all $\mathbf{y} \in B_\delta(\mathbf{x})$ and $\mathbf{x} \in \mathcal{S}_\epsilon$. Moreover, under Assumption **A4**, $\partial_{x^i}^2 v^\delta(\mathbf{x}) = \int_{\mathbf{y} \in B_\delta(\mathbf{x})} \partial_{x^i}^2 v(\mathbf{x}) \phi^\delta(\mathbf{x} - \mathbf{y}) d\mathbf{y} > 0$ for all $\mathbf{x} \in \mathcal{S}_\epsilon$.

Furthermore, at each point $\mathbf{x} \in \mathcal{S}_\epsilon$, there exists $\gamma \in \mathbf{r}_\epsilon(\mathbf{x})$ pointing into \mathcal{C}_ϵ . This is because there is no $\mathbf{x} \in \partial\mathcal{C}_\epsilon$ such that $i, i + N \in I_\epsilon(\mathbf{x})$ for all $i = 1, 2, \dots, N$. This implies $|I_\epsilon(\mathbf{x})| \leq N$ for all $\mathbf{x} \in \partial\mathcal{C}_\epsilon$. Now Assumption **A4** implies the following condition (3.8) in Dupuis & Ishii (1993): the existence of scalars $b_j \geq 0$ $j \in I_\epsilon(\mathbf{x})$, such that

$$b_j \langle \gamma_j(\mathbf{x}), n_j(\mathbf{x}) \rangle > \sum_{k \in I_\epsilon(\mathbf{x}) \setminus \{i\}} b_k |\langle \gamma_k(\mathbf{x}), n_k(\mathbf{x}) \rangle|.$$

Here we can simply take $b_j = 1$ for all $j \in I_\epsilon(\mathbf{x})$. Therefore, by (Dupuis & Ishii, 1993, Theorem 4.8 and Corollary 5.2), there exists a unique strong solution to the Skorokhod problem with data $(\{\mathbf{x} + \boldsymbol{\mu}t + \boldsymbol{\sigma}B_t\}_{t \geq 0}, \mathcal{C}_\epsilon, \mathbf{r}_\epsilon, \mathbf{x})$.

Step 2. ϵ -optimal policy. Now we shall show that the solution to the Skorokhod problem with data $(\{\mathbf{x} + \boldsymbol{\mu}t + \boldsymbol{\sigma}B_t, \mathcal{C}_\epsilon, \mathbf{r}_\epsilon, \mathbf{x})$ is an ϵ -optimal policy of the control problem (**Regulator**) with

$$\boldsymbol{\xi}_t^\epsilon = \int_0^t \mathbf{N}_s^\epsilon \cdot d\eta_s^\epsilon, \quad (4.12)$$

and $\mathbf{N}_s^\epsilon \in \mathbf{r}_\epsilon(\mathbf{X}_s^\epsilon)$ on \mathcal{S}_ϵ , with $\mathbf{X}_t^\epsilon = \mathbf{x} + \boldsymbol{\mu}t + \boldsymbol{\sigma}B_t + \boldsymbol{\xi}_t^\epsilon$. By (Dupuis & Ishii, 1993, Theorem 4.8), \mathbf{X}^ϵ is a continuous process. Since $v \in C^{4,\alpha}(\mathcal{C}_N)$, applying the Itô formula to the continuous semi-martingale yields

$$\begin{aligned} v(\mathbf{x}) &= \mathbb{E}_{\mathbf{x}} \int_0^\infty e^{-\rho t} [H(\mathbf{X}_t^\epsilon) dt + \nabla v(\mathbf{X}_t^\epsilon) \cdot \mathbf{N}_t^\epsilon d\eta_t^\epsilon] \\ &\geq \mathbb{E}_{\mathbf{x}} \int_0^\infty e^{-\rho t} [H(\mathbf{X}_t^\epsilon) dt + (1 - 3\epsilon) [(N_t^\epsilon)^+ \cdot \mathbf{K}_L^+ + (N_t^\epsilon)^- \cdot \mathbf{K}_L^+] d\eta_t^\epsilon] \\ &= \mathbb{E}_{\mathbf{x}} \int_0^\infty e^{-\rho t} [H(\mathbf{X}_t^\epsilon) dt + [(N_t^\epsilon)^+ \cdot \mathbf{K}_L^+ + (N_t^\epsilon)^- \cdot \mathbf{K}_L^+] d\eta_t^\epsilon] \\ &\quad - 3\epsilon \mathbb{E}_{\mathbf{x}} \int_0^\infty e^{-\rho t} [(N_t^\epsilon)^+ \cdot \mathbf{K}_L^+ + (N_t^\epsilon)^- \cdot \mathbf{K}_L^+] d\eta_t^\epsilon \\ &\geq \mathbb{E}_{\mathbf{x}} \int_0^\infty e^{-\rho t} [H(\mathbf{X}_t^\epsilon) dt + [(N_t^\epsilon)^+ \cdot \mathbf{K}_L^+ + (N_t^\epsilon)^- \cdot \mathbf{K}_L^+] d\eta_t^\epsilon] \\ &\quad - 3\epsilon K_{\max} \mathbb{E}_{\mathbf{x}} \int_0^\infty e^{-\rho t} d\eta_t^\epsilon, \end{aligned} \quad (4.13)$$

where $\mathbf{N}^\epsilon(\mathbf{x}) \in \mathbf{r}_\epsilon(\mathbf{x})$ on \mathcal{S}_ϵ ,

$$\mathbf{K}_L^+ := (L_1 K_1^+, \dots, L_N K_N^+), \quad \mathbf{K}_L^- := (L_1 K_1^-, \dots, L_N K_N^-), \quad \text{and} \quad K_{\max} = \max_{1 \leq i \leq N} \{L_i K_i^+, L_i K_i^-\}. \quad (4.14)$$

The first inequality of (4.13) holds since $\|\nabla v^\delta - \nabla v\|_{L_1} < K_{\min} \epsilon$ for $\delta \in [0, \delta(\epsilon)]$ and (4.10). Moreover, there exists constant $C > 0$ such that $\mathbb{E}_{\mathbf{x}} \left[\int_0^\infty e^{-\rho t} d\eta_t^\epsilon \right] \leq C$ for all $\epsilon < \frac{1}{2}$. Hence

$$v(\mathbf{x}) \geq J(\mathbf{x}; \boldsymbol{\xi}_i^\epsilon) - 3\epsilon C K_{\max}.$$

□

Proposition 9. *Assume Assumptions A1-A4 hold, when $\mathbf{x} \in \mathcal{C}_N$. The optimal policy acts only on $\partial \mathcal{C}_N$, and its reflection direction is in $\mathbf{r}(\mathbf{x})$.*

Proof. Take the smooth function ϕ^ϵ in (4.8) and the smooth version of value function v^ϵ in (4.9). Let $H^\epsilon(\mathbf{x}) = \phi^\epsilon * H(\mathbf{x})$. From the HJB Equation (3.4),

$$\rho v - \mathcal{L}v \leq H, \quad \beta(\nabla v) \leq 1 \text{ in } \mathbb{R}^N,$$

and

$$\rho v^\epsilon - \mathcal{L}v^\epsilon \leq H^\epsilon, \quad \beta(\nabla v^\epsilon) \leq 1 \text{ in } \mathbb{R}^N. \quad (4.15)$$

Letting $T > 0$ and applying Meyer's version of Itô's formula (Meyer, 1976, Theorem 21) to $e^{-\rho t} v^\epsilon(\mathbf{x})$ and any semi-martingale $\{\mathbf{X}_t\}_{t \geq 0}$ under admissible control $(\xi^{i,+}, \xi^{i,-})_{i=1}^N$ yield

$$\begin{aligned} \mathbb{E}_{\mathbf{x}} [e^{-\rho T} v^\epsilon(\mathbf{X}_T)] &= v^\epsilon(\mathbf{x}) + \mathbb{E}_{\mathbf{x}} \int_0^T e^{-\rho t} (\mathcal{L}v^\epsilon - \rho v^\epsilon)(\mathbf{X}_t) dt \\ &+ \mathbb{E}_{\mathbf{x}} \int_0^T e^{-\rho t} \nabla v^\epsilon(\mathbf{X}_t) \cdot d\xi_t \\ &+ \mathbb{E}_{\mathbf{x}} \int_0^T \sum_{0 \leq t < T} e^{-\rho t} (v^\epsilon(\mathbf{X}_t) - v^\epsilon(\mathbf{X}_{t-}) - \nabla v^\epsilon \cdot (\mathbf{X}_t)(\boldsymbol{\xi}_t - \boldsymbol{\xi}_{t-})), \end{aligned}$$

with the last term coming from the jumps of \mathbf{X}_t . By (4.15),

$$\begin{aligned} &\mathbb{E}_{\mathbf{x}} [e^{-\rho T} v^\epsilon(\mathbf{X}_T)] + \mathbb{E}_{\mathbf{x}} \int_0^T e^{-\rho t} H^\epsilon(\mathbf{X}_t) dt - \mathbb{E}_{\mathbf{x}} \int_0^T e^{-\rho t} \nabla v^\epsilon(\mathbf{X}_t) \cdot d\xi_t \\ &+ \mathbb{E}_{\mathbf{x}} \int_0^T \sum_{0 \leq t < T} e^{-\rho t} (-v^\epsilon(\mathbf{X}_t) + v^\epsilon(\mathbf{X}_{t-}) + \nabla v^\epsilon(\mathbf{X}_t) \cdot (\boldsymbol{\xi}_t - \boldsymbol{\xi}_{t-})) \geq v^\epsilon(\mathbf{x}). \end{aligned} \quad (4.16)$$

Moreover, H^ϵ , v^ϵ are bounded uniformly on $\bar{\mathcal{C}}_N$ for $\epsilon < 1$ because v , ∇v , $D^2 v$ are bounded on $B_R(0)$, with $\bar{\mathcal{C}}_N \subset B_{R-1}(0)$, thus

$$v^\epsilon \rightarrow v, \quad \nabla v^\epsilon \rightarrow \nabla v, \quad H^\epsilon \rightarrow H \quad \text{uniformly in } \bar{\mathcal{C}}_N.$$

Meanwhile, for $\forall \mathbf{x} \in \mathcal{C}_N$,

$$v(\mathbf{x}) = \mathbb{E}_{\mathbf{x}} \int_0^\infty e^{-\rho t} [H(\mathbf{X}_t^*) dt + [(\mathbf{N}_t^*)^+ \cdot \mathbf{K}_L^+ + (\mathbf{N}_t^*)^- \cdot \mathbf{K}_L^-] d\eta_t^*], \quad (4.17)$$

where $\mathbf{X}_t^* = \mathbf{x} + \boldsymbol{\mu}t + \boldsymbol{\sigma}\mathbf{B}_t + \boldsymbol{\xi}_t^*$ with $\boldsymbol{\xi}_t^* := \int_0^t \mathbf{N}_s^* d\eta_s^*$ the optimal control, and \mathbf{K}_L^+ and \mathbf{K}_L^- are defined in (4.14). In particular,

$$\mathbb{E}_{\mathbf{x}} \int_0^\infty e^{-\rho t} d\eta_t^* < \infty, \quad (4.18)$$

which leads to

$$\mathbb{E}_{\mathbf{x}} \int_0^T e^{-\rho t} [(\mathbf{N}_t^*)^+ \cdot \mathbf{K}_L^+ + (\mathbf{N}_t^*)^- \cdot \mathbf{K}_L^-] d\eta_t^* < \infty.$$

By the bounded convergence theorem and (4.16),

$$\begin{aligned} & \mathbb{E}_{\mathbf{x}} [e^{-\rho T} v(\mathbf{X}_T^*)] + \mathbb{E}_{\mathbf{x}} \int_0^T e^{-\rho t} H(\mathbf{X}_t^*) dt - \mathbb{E}_{\mathbf{x}} \int_0^T e^{-\rho t} \nabla v(\mathbf{X}_t^*) \cdot \mathbf{N}_t^* d\eta_t \\ & + \mathbb{E}_{\mathbf{x}} \int_0^T \sum_{0 \leq t < T} e^{-\rho t} (-v(\mathbf{X}_t^*) + v^\epsilon(\mathbf{X}_{t-}^*) + \nabla v(\mathbf{X}_t^*) \cdot \mathbf{N}_t^*(\eta_t^* - \eta_{t-}^*)) \geq v(\mathbf{x}). \end{aligned} \quad (4.19)$$

The last term on the left-hand side is nonpositive because of convexity of v , hence

$$\mathbb{E}_{\mathbf{x}} [e^{-\rho T} v(\mathbf{X}_T^*)] + \mathbb{E}_{\mathbf{x}} \int_0^T e^{-\rho t} H(\mathbf{X}_t^*) dt - \mathbb{E}_{\mathbf{x}} \int_0^T e^{-\rho t} \nabla v(\mathbf{X}_t^*) \cdot \mathbf{N}_t^* d\eta_t^* \geq v(\mathbf{x}).$$

Letting $T \rightarrow \infty$, by the boundedness of \mathbf{X}_t^* , $\beta(\nabla v) \leq 1$, $|\mathbf{N}_t^*| = 1$, and (4.18),

$$\mathbb{E}_{\mathbf{x}} \int_0^T e^{-\rho t} H(\mathbf{X}_t^*) dt - \mathbb{E}_{\mathbf{x}} \int_0^T e^{-\rho t} \nabla v(\mathbf{X}_t^*) \cdot \mathbf{N}_t^* d\eta_t^* \geq v(\mathbf{x}).$$

Along with (4.17), we have

$$0 \geq \mathbb{E}_{\mathbf{x}} \int_0^\infty e^{-\rho t} ([\nabla v(\mathbf{X}_t^*) + \mathbf{K}_L^+] \cdot (\mathbf{N}_t^*)^+ d\eta_t^* + [-\nabla v(\mathbf{X}_t^*) + \mathbf{K}_L^-] \cdot (\mathbf{N}_t^*)^- d\eta_t^*).$$

Given $\beta(\nabla v) \leq 1$, we have $-K_i^+ \leq v_{x^i}(\mathbf{x}) \leq K_i^-$, $\forall \mathbf{x} \in \mathbb{R}^N$ and $i = 1, 2, \dots, N$. Hence

$$0 \geq \mathbb{E}_{\mathbf{x}} \int_0^\infty e^{-\rho t} ([\nabla v(\mathbf{X}_t^*) + \mathbf{K}_L^+] \cdot (\mathbf{N}_t^*)^+ d\eta_t^* + [-\nabla v(\mathbf{X}_t^*) + \mathbf{K}_L^-] \cdot (\mathbf{N}_t^*)^- d\eta_t^*) \geq 0.$$

This implies $d\eta_t^* = 0$ when $\beta(\nabla v(\mathbf{X}_t^*)) < 1$ a.e. in t . Also, when $d\eta_t^* \neq 0$, $\mathbf{N}_t^*(\mathbf{x}) \in \mathbf{r}(\mathbf{x})$ for $\mathbf{x} \in \mathcal{S}$ a.e. for $t \in [0, \infty)$, where the reflection cone $\mathbf{r}(\mathbf{x})$ is defined in (4.6). □

Now we are ready to establish the main theorem when $\mathbf{x} \in \bar{\mathcal{C}}_N$.

Theorem 10 (Existence and uniqueness of optimal controls). *Take $\mathbf{x} \in \bar{\mathcal{C}}_N$ and assume **A1- A4**. Then there exists a unique optimal control $\boldsymbol{\xi}^*$ to problem (**Regulator**), which is a solution to the Skorokhod problem (7) with data $(\mathbf{x} + \boldsymbol{\mu}t + \boldsymbol{\sigma}\mathbf{B}_t, \mathcal{C}_N, \mathbf{r}, \mathbf{x})$ such that $\mathbf{X}_t^* \in \bar{\mathcal{C}}_N$.*

Proof. The existence of the optimal control to the control problem (**Regulator**) follows from an appropriate modification of Theorem 4.5 and Corollary 4.11 in (Menaldi & Taksar, 1989), as below.

First, $J(\mathbf{x}; \boldsymbol{\xi}^\epsilon) \rightarrow v(\mathbf{x})$ as $\epsilon \rightarrow 0$, then $\boldsymbol{\xi}_t^\epsilon(\omega)$ converges under measure m_T on $([0, T] \times \Omega, \mathcal{B}[0, T] \times \mathcal{F})$ for any $T > 0$, which equals to the product of Lebesgue measure and \mathbb{P} . Furthermore, there exists a unique optimal policy $\boldsymbol{\xi}^*$ which is the limit of a subsequence of $\{\boldsymbol{\xi}^\epsilon\}_\epsilon$.

If $(\mathbf{N}^{\epsilon_k}, \eta^{\epsilon_k})$ is a sequence of ϵ_k -optimal policies for \mathbf{x} and $\lim_{k \rightarrow \infty} \epsilon_k \rightarrow 0$, then one can extract a subsequence $\epsilon_{k'}$ such that

$$\boldsymbol{\xi}_t^{\epsilon_{k'}} = \int_0^t \mathbf{N}_s^{\epsilon_{k'}} d\eta_s^{\epsilon_{k'}} \rightarrow \boldsymbol{\xi}_t^*, \quad (4.20)$$

under $\text{Leb} \times \mathbb{P}$ for almost all (t, ω) , where Leb is the Lebesgue measure on $[0, \infty)$.

By the analysis in Theorem 8, there exists a sequence of ϵ_k -optimal policy and $\epsilon_k \rightarrow 0$ when $k \rightarrow \infty$. Therefore, the optimal control exists. Let

$$A = \{\omega \mid \mathbf{X}_t^{\epsilon_{k'}}(\omega) \in \bar{\mathcal{C}}_{\epsilon_{k'}} \text{ for all } 0 \leq t < \infty \text{ and all } k' \geq 0\},$$

then by definition (4.12), $P(A) = 1$. Also define

$$B = \{\omega \mid \mathbf{X}_t^{\epsilon_{k'}} \rightarrow \mathbf{X}_t \text{ a.e. Leb on } [0, \infty)\},$$

then by (4.20), $P(B) = 1$. For all $\omega \in A \cap B$, since $\bar{\mathcal{C}}_N$ is closed,

$$\mathbf{X}_t(\omega) \in \bar{\mathcal{C}}_N \text{ Leb a.e. on } [0, \infty).$$

Properties **(b)** and **(c)** of the Skorokhod problem (Definition 7) follow from Theorem 8 and Proposition 9, respectively. By Assumption **A4**, for any $\mathbf{x} \in \partial\mathcal{C}_N$ and $\boldsymbol{\gamma}(\mathbf{x}) \in \mathbf{r}(\mathbf{x})$, $\boldsymbol{\gamma}(\mathbf{x})$ is not parallel to $\partial\mathcal{C}_N$ at \mathbf{x} . Hence, property **(a)** holds, i.e., the optimal control is continuous.

It remains to show the uniqueness of the optimal control. This is done by a contradiction argument. Suppose that there are two optimal controls $\{\boldsymbol{\xi}^*\}_{t \geq 0}$ and $\{\boldsymbol{\xi}^{**}\}_{t \geq 0}$ such that $\boldsymbol{\xi}^* \neq \boldsymbol{\xi}^{**}$ almost surely. Let $\{\mathbf{X}_t^*\}_{t \geq 0}$ and $\{\mathbf{X}_t^{**}\}_{t \geq 0}$ be the corresponding trajectories. Let $\boldsymbol{\xi}_t = \frac{\boldsymbol{\xi}_t^* + \boldsymbol{\xi}_t^{**}}{2}$ and $\mathbf{X}_t = \frac{\mathbf{X}_t^* + \mathbf{X}_t^{**}}{2}$. Then by Assumption **A3**,

$$\begin{aligned} v(\mathbf{x}) - J(\mathbf{x}; \boldsymbol{\xi}_t) &= \frac{(J(\mathbf{x}; \boldsymbol{\xi}^*) + J(\mathbf{x}; \boldsymbol{\xi}^{**}))}{2} - J(\mathbf{x}; \boldsymbol{\xi}) \\ &\geq \mathbb{E}_{\mathbf{x}} \int_0^\infty e^{-\rho t} \left[\frac{H(\mathbf{X}_t^*) + H(\mathbf{X}_t^{**})}{2} - H\left(\frac{\mathbf{X}_t^* + \mathbf{X}_t^{**}}{2}\right) \right] dt > 0. \end{aligned}$$

Therefore $v(\mathbf{x}) > J(\mathbf{x}; \boldsymbol{\xi})$, which contradicts the optimality of $\{\boldsymbol{\xi}_t^*\}_{t \geq 0}$ and $\{\boldsymbol{\xi}_t^{**}\}_{t \geq 0}$. Hence the uniqueness of the optimal control. \square

Optimal policy for $\mathbf{x} \notin \bar{\mathcal{C}}_N$. When $\mathbf{x} \notin \bar{\mathcal{C}}_N$, the optimal policy is to jump immediately to some point $\hat{\mathbf{x}} \in \bar{\mathcal{C}}_N$ and then follows the optimal policy in $\bar{\mathcal{C}}_N$. We will need the following assumption so that the reflection field of the Skorokhod problem is extendable to the \mathbb{R}^N plane (Dupuis & Ishii, 1991). Note that when $N = 2$, **A5** follows directly from Assumptions **A1-A3**.

A5. There is a map $\pi : \mathbb{R}^N \rightarrow \mathcal{C}_N$ satisfying $\pi(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \mathcal{C}_N$ and $\pi(\mathbf{x}) - \mathbf{x} \in \mathbf{r}(\pi(\mathbf{x}))$.

This assumption was also adopted in (Dupuis & Ishii, 1991, Assumption 3.1).

Theorem 11. *Given **A1-A3**, and **A5**. For any $\mathbf{x} \notin \bar{\mathcal{C}}_N$, there exists an optimal policy π such that $\pi(\mathbf{x}) \in \partial\mathcal{C}_N$ at time 0 and*

$$v(\mathbf{x}) = v(\pi(\mathbf{x})) + l(\mathbf{x} - \pi(\mathbf{x})),$$

with $l(\mathbf{y}) = \sum_i l_i(y_i)$, where

$$l_i(y_i) = \begin{cases} L_i K_i^- y_i, & \text{if } y_i \geq 0, \\ -L_i K_i^+ y_i, & \text{if } y_i < 0. \end{cases} \quad (4.21)$$

Proof. Notice that $l(\mathbf{y})$ is convex and

$$l_i(y_i) = \max_{-L_i K_i^+ \leq k \leq L_i K_i^-} \{ky_i\} = \max\{-L_i K_i^+ y_i, L_i K_i^- y_i\} \text{ for } y_i \in \mathbb{R}.$$

Here we define two linear approximations which correspond to the lower and the upper bounds of the value function $v(\mathbf{x})$, respectively.

For $\mathbf{x} \notin \bar{\mathcal{C}}_N$, define

$$\begin{aligned} u_1(\mathbf{x}) &= v(\pi(\mathbf{x})) + \nabla v(\pi(\mathbf{x})) \cdot (\mathbf{x} - \pi(\mathbf{x})), \\ u_2(\mathbf{x}) &= v(\pi(\mathbf{x})) + l(\mathbf{x} - \pi(\mathbf{x})). \end{aligned} \tag{4.22}$$

Then $u_2(\mathbf{x}) \geq v(\mathbf{x})$ by the sub-optimality of the policy, and $u_1(\mathbf{x}) \leq v(\mathbf{x})$ by convexity. Thus,

$$u_1(\mathbf{x}) \leq v(\mathbf{x}) \leq u_2(\mathbf{x}). \tag{4.23}$$

We now show $u_1(\mathbf{x}) = u_2(\mathbf{x})$. By Assumption **A5**, u_1 and u_2 in (4.22) can be rewritten as

$$\begin{aligned} u_1(\mathbf{x}) &= v(\pi(\mathbf{x})) + \nabla v(\pi(\mathbf{x})) \cdot d(\pi(\mathbf{x})) \|\mathbf{x} - \pi(\mathbf{x})\|, \\ u_2(\mathbf{x}) &= v(\pi(\mathbf{x})) + \mathbf{P}(\pi(\mathbf{x})) \cdot d(\pi(\mathbf{x})) \|\mathbf{x} - \pi(\mathbf{x})\|, \end{aligned}$$

where $d(\pi(\mathbf{x})) \in \mathbf{r}(\pi(\mathbf{x}))$ and $\mathbf{P}(\mathbf{x}) = (P_1, \dots, P_N)(\mathbf{x})$, with

$$P_i(\mathbf{x}) = L_i K_i^+ \mathbf{1}(\partial_{x^i} v(\mathbf{x}) < 0) + L_i K_i^- \mathbf{1}(\partial_{x^i} v(\mathbf{x}) > 0).$$

Therefore $u_1(\mathbf{x}) = u_2(\mathbf{x})$. □

4.2 Pareto-optimal policy

Pareto-optimal policy for (**N-player**) can now be constructed from the optimal control for problem (**Regulator**) according to the following.

Theorem 12. *The optimal control of problem (**Regulator**) yields a Pareto-optimal policy for the game (**N-player**).*

Proof. The proof is straightforward. To see this, take the payoff function J^i in (**N-player**), $v(\mathbf{x})$ the value function in (**Regulator**), and the optimal control $\boldsymbol{\xi}^* := (\boldsymbol{\xi}^{1*}, \dots, \boldsymbol{\xi}^{N*})$, if exists, to problem (**Regulator**), then for any $\boldsymbol{\xi} := (\boldsymbol{\xi}^1, \dots, \boldsymbol{\xi}^N) \in \mathcal{U}_N$ and L_i , with $L_i > 0$, $\sum_{i=1}^N L_i = 1$,

$$\sum_{i=1}^N L_i J^i(\mathbf{x}; \boldsymbol{\xi}) \geq v(\mathbf{x}), \tag{4.24}$$

where value $v(\mathbf{x})$ is reached when player i takes the control $\boldsymbol{\xi}_t^{i*}$ ($i = 1, 2, \dots, N$).

If there is another $\boldsymbol{\xi}' := (\boldsymbol{\xi}^{1'}, \dots, \boldsymbol{\xi}^{N'}) \in \mathcal{U}_N$ and $k \in \{1, \dots, N\}$ such that

$$J^k(\mathbf{x}; \boldsymbol{\xi}^{1'}, \dots, \boldsymbol{\xi}^{N'}) < J^k(\mathbf{x}; \boldsymbol{\xi}^{1*}, \dots, \boldsymbol{\xi}^{N*}),$$

then given $L_i > 0$ for all i , there must exists $j \in \{1, \dots, N\}$ such that

$$J^j(\mathbf{x}; \boldsymbol{\xi}^{1'}, \dots, \boldsymbol{\xi}^{N'}) > J^j(\mathbf{x}; \boldsymbol{\xi}^{1*}, \dots, \boldsymbol{\xi}^{N*}).$$

Hence the control $\boldsymbol{\xi}^*$ is a Pareto-optimal policy by definition. □

Combining Theorems 10, 11 and 12 yields the following result which summarizes the structure of the set of Pareto optima:

Theorem 13 (Pareto-optimal policies). *Under Assumptions **A1-A5**, for any set of weights $\mathbf{L} = (L_1, \dots, L_N)$ with $L_i > 0$ and $\sum_{i=1}^N L_i = 1$, the unique solution $\xi_{\mathbf{L}} \in \mathcal{U}_N$ to the regulator's problem (**Regulator**) yields a Pareto-optimal policy for the game (**N-player**).*

The analytical structure of the continuation region (4.3) and the Pareto-optimal policy suggest the following description: \mathbf{X}_t evolves according to the uncontrolled diffusion process inside the interior of \mathcal{C}_N and when it hits boundary at a point belonging to ∂G_i or ∂G_{i+N} , then bank i will adjust its rate to push it back continuously inside \mathcal{C}_N . In particular the optimal policies lead to continuous controls ξ^i i.e. no impulses are used.

4.3 Pareto-optimal policies for interbank lending

Let us now translate these results in the setting of the interbank lending model described in Section 1.1.

Theorem 13 implies that Pareto optima for the interbank lending market may be described in terms of policies of a regulator facing the optimization problem (**Regulator**) with an aggregate payoff function (3.2) representing a weighted average of payoffs of individual banks. Under a Pareto-optimal policy, the interbank rates may be described as a ‘regulated diffusion’ in a bounded region \mathcal{C}_N defined by (4.3). The boundedness of \mathcal{C}_N implies that from the payoff structure (1.2) emerges an *endogenous* bound on the interbank rates: the Pareto-optimal policy guides \mathbf{X}_t to remain confined in the bounded region \mathcal{C}_N , which implies in particular that both the benchmark and the spreads X^i remain bounded. In the context of the LIBOR mechanism, this can be seen as the impact of ‘trimmed’ averaging, which is the origin of the terms K_i^+ , K_i^- , as explained in Section 1.1: as banks internalize the risk of being ‘outliers’ in the benchmark fixing, they confine their rates to a bounded region.

The process \mathbf{X}_t diffuses in the interior of \mathcal{C}_N , following the random shocks banks are subjected to, and is pushed into the interior when it reaches the boundary. More precisely, the boundary $\partial\mathcal{C}_N$ is composed of $2N$ ‘faces’ corresponding to the saturation of the constraints in (4.4). Edges correspond to intersections of two or more faces. When \mathbf{X}_t reaches a point $\mathbf{x} \in \partial\mathcal{C}_N$, action is taken by all banks i such that $\mathbf{x} \notin G_i \cup G_{i+N}$: if $\mathbf{x} \notin G_i$ then X^i is reduced i.e. $d\xi^{i,-} > 0$ and if $\mathbf{x} \notin G_{i+N}$ then X^i is increased i.e. $d\xi^{i,+} > 0$. When \mathbf{X}_t reaches the interior of such a face, only bank i adjusts its rate in order to continuously push back \mathbf{X}_t to the interior. Similarly, if \mathbf{X}_t reaches an edge, two or more banks need to simultaneously adjust their rates. The rate at which such simultaneous adjustments occur is given by the *intersection local time* (Rosen, 1987) of (X^1, \dots, X^N) on the boundary. Therefore Pareto-optimal policy rarely leads to more than one bank’s rate to be adjusted; a simultaneous rate adjustment by several banks is most likely *not* associated with a Pareto-optimal policy and is thus the signature of a non-optimal behavior by banks.

We also note that, although the admissible controls allow for ‘impulses’ i.e. discontinuous adjustments of rates, such impulses are not optimal and Pareto-optimal policies correspond to continuously pushing the process to the interior. The only possible discontinuity in Pareto-optimal policies (discussed in Theorem 11) is the case of an initial impulse at $t = 0$ to bring the initial condition into \mathcal{C}_N , which we may interpret as the entry of a new bank into the interbank market.

The set of all such Pareto optima is parameterized by the set of allocations $L = (L_1, \dots, L_N)$ with $L_i > 0$ and $\sum_{i=1}^N L_i = 1$. These allocations lead to different outcomes across banks. A natural choice is to take L_i proportional to the market share (or loan volume) of bank i ; (3.2) then represents an aggregate wealth maximization problem and this policy leads to the same pro-rata cost across banks. As is clear from (3.5), choosing a higher weight L_i leads to a tighter control on the rates of bank i .

5 Explicit solution for $N = 2$ players

We now study in more detail the structure of the optimal strategies for the case of $N = 2$. Our analytical results illustrate the difference between Nash equilibria and Pareto optima and demonstrate the impact of regulatory intervention in this game.

5.1 Pareto-optimal solution for $N = 2$

For the special case of $N = 2$, we can derive explicitly its Pareto-optimal solution. For ease of exposition, we shall assume the following conditions in the case of $N = 2$.

B1. $a_1 = a_2$ and $L_1 = L_2$. In other words, the regulator allocates equal weights to the banks.

B2. $h^1(x^1, x^2) = h^2 = h^1(x^1, x^2) = h(x^1 - x^2)$, $h \in C^3(\mathbb{R})$ is symmetric, and there exist $0 < c < C$ such that $c < h'' < C$, and h'' is non-decreasing and bounded away from 0.

B3. $\mu^1 = \mu^2 = 0$, $K_1^+ = K_1^- =: K_1 > 0$ and $K_2^+ = K_2^- =: K_2 > 0$.

Note that Assumption **B2** is more general than Assumptions **A1-A3**. As a result, we will see in Proposition 16 that the non-action region may not necessarily be bounded and the Pareto-optimal policy for the game may not be unique with fixed weights $L_1 = L_2$.

Under Assumption **B3**, the rates X_t^1 and X_t^2 are assumed to be

$$X_t^i = \sigma^i \cdot d\mathbf{B}_t + d\xi_t^{i,+} - d\xi_t^{i,-}, \text{ with } x_{0-}^i = x^i, \quad i = 1, 2. \quad (5.1)$$

The regulator's value function $v(x^1, x^2)$ is

$$\begin{aligned} v(x^1, x^2) &= \inf_{(\xi^1, \xi^2) \in \mathcal{U}_2} J(x^1, x^2, \xi^1, \xi^2) = \inf_{(\xi^1, \xi^2) \in \mathcal{U}_2} \frac{1}{2} [J^1(x^1, x^2, \xi^1, \xi^2) + J^2(x^1, x^2, \xi^1, \xi^2)] \\ &= \inf_{(\xi^1, \xi^2) \in \mathcal{U}_2} \mathbb{E}_{(x^1, x^2)} \left[\int_0^\infty e^{-\rho t} \left(h(X_t^1 - X_t^2) dt + \frac{K_1}{2} d\xi_t^{1,+} + \frac{K_1}{2} d\xi_t^{1,-} + \frac{K_2}{2} d\xi_t^{2,+} + \frac{K_2}{2} d\xi_t^{2,-} \right) \right], \end{aligned} \quad (5.2)$$

subject to (5.1).

Lemma 14. *Assume $K_2 < K_1$ and **B1-B3**. Then for any $(\xi^{1*}, \xi^{2*}) \in \arg \inf_{(\xi^1, \xi^2) \in \mathcal{U}_2} J(x^1, x^2, \xi^1, \xi^2)$,*

$$(\xi_t^{1,+*}, \xi_t^{1,-*}) = (0, 0) \text{ for any } t \geq 0 \text{ a.s..}$$

Proof. The statement is proved by contradiction. Assume there exists an optimal policy $(\xi^{1*}, \xi^{2*}) \in \arg \inf_{(\xi^1, \xi^2) \in \mathcal{U}_2} J(x^1, x^2, (\xi^1, \xi^2))$ and $t_0 \geq 0$ such that

$$\xi_{t_0}^{1*} > 0.$$

Since $\xi^{1,+*}$ is a non-decreasing process, we have $\xi_t^{1,+*} > 0$ for all $t \geq t_0$. Now construct the following admissible policy $(\bar{\xi}^1, \bar{\xi}^2)$ such that, $\forall t \geq 0$,

$$\begin{cases} \bar{\xi}_t^{2,-} = \xi_t^{1*} + \xi_t^{2*}, \\ \bar{\xi}_t^{1,+} = 0, \\ \bar{\xi}_t^{1,-} = \xi_t^{1*}, \quad \bar{\xi}_t^{2,+} = \xi_t^{2*}. \end{cases} \quad (5.3)$$

Then

$$J(x^1, x^2, \xi^{1*}, \xi^{2*}) - J(x^1, x^2, \bar{\xi}^1, \bar{\xi}^2) = \mathbb{E}_{(x^1, x^2)} \left[\int_0^\infty e^{-\rho t} \frac{K_1 - K_2}{2} d\xi_t^{1*} \right] > 0,$$

which contradicts the optimality of the control process (ξ^{1*}, ξ^{2*}) . \square

We now show that solving the control problem (5.1)-(5.2) is equivalent to the following control problem (5.4)-(5.5) when $K_1 > K_2$,

$$\begin{aligned} u(y) &= \inf_{\boldsymbol{\eta} \in \mathcal{U}_1} \widehat{J}(y, \boldsymbol{\eta}) \\ &= \inf_{\boldsymbol{\eta} \in \mathcal{U}_1} \mathbb{E}_y \left[\int_0^\infty e^{-\rho t} \left(h(Y_t) dt + \frac{K_2}{2} d\eta_t^+ + \frac{K_2}{2} d\eta_t^- \right) \right], \end{aligned} \quad (5.4)$$

subject to

$$dY_t = (\boldsymbol{\sigma}^1 - \boldsymbol{\sigma}^2) \cdot d\mathbf{B}_t - d\eta_t^+ + d\eta_t^-, \quad \text{with } Y_{0-} = y. \quad (5.5)$$

Lemma 15 (Equivalence). *Assume **B1-B3** and $K_1 > K_2$, then*

(i) $v(x^1, x^2) = u(x^1 - x^2)$;

(ii) *If $(\boldsymbol{\xi}^{1*}, \boldsymbol{\xi}^{2*}) \in \arg \inf_{(\boldsymbol{\xi}^1, \boldsymbol{\xi}^2) \in \mathcal{U}_2} J(x^1, x^2, (\boldsymbol{\xi}^1, \boldsymbol{\xi}^2))$, then $(\xi_t^{1*,+}, \xi_t^{1*,-}) = (0, 0) \forall t$ a.s., and $\boldsymbol{\xi}^{2*} \in \arg \inf_{\boldsymbol{\eta} \in \mathcal{U}_1} \widehat{J}(x^1 - x^2, \boldsymbol{\eta})$;*

(iii) *If $\boldsymbol{\eta}^* \in \arg \inf_{\boldsymbol{\eta} \in \mathcal{U}_1} \widehat{J}(x^1 - x^2, \boldsymbol{\eta})$, then $((\mathbf{0}, \mathbf{0}), \boldsymbol{\eta}) \in \arg \inf_{(\boldsymbol{\xi}^1, \boldsymbol{\xi}^2) \in \mathcal{U}_2} J(x^1, x^2, (\boldsymbol{\xi}^1, \boldsymbol{\xi}^2))$.*

Proof. By Lemma 14, $(\xi_t^{1*,+}, \xi_t^{1*,-}) = (0, 0)$ for any $t \geq 0$ a.s.. Therefore, we can consider a smaller class of admissible control set where $(\xi_t^{1*,+}, \xi_t^{1*,-}) = (0, 0) \forall t \geq 0$ and $\boldsymbol{\xi}^2 \in \mathcal{U}_1$. Note that with $(\xi_t^{1*,+}, \xi_t^{1*,-}) = (0, 0)$, we have

$$X_t^1 - X_t^2 = (\boldsymbol{\sigma}^1 - \boldsymbol{\sigma}^2) \cdot \mathbf{B}_t - \xi_t^{2,+} + \xi_t^{2,-} + (x^1 - x^2), \quad (5.6)$$

and

$$J(x^1, x^2, \boldsymbol{\xi}^1, \boldsymbol{\xi}^2) = \mathbb{E}_{(x^1, x^2)} \left[\int_0^\infty e^{-\rho t} \left(h(X_t^1 - X_t^2) dt + \frac{K_2}{2} d\xi_t^{2,+} + \frac{K_2}{2} d\xi_t^{2,-} \right) \right]. \quad (5.7)$$

Clearly problem (5.6)-(5.7) is equivalent to the one-dimensional control problem (5.4)-(5.5) with $y = x^1 - x^2$. Hence the claim. \square

Proposition 16 (Pareto-optimal solution when $N = 2$). *Assume **B1-B3**, then*

(i) *If $K_1 = K_2 = K$, then the following control yields one Pareto-optimal policy to game (5.1)-(5.2):*

$$\begin{aligned} \boldsymbol{\xi}_t^{1*} &= (\xi_t^{1*,+}, \xi_t^{1*,-}) = \left(0, \max \left\{ 0, \max_{0 \leq u \leq t} \left\{ (x^1 - x^2) + (\boldsymbol{\sigma}^1 - \boldsymbol{\sigma}^2) \cdot \mathbf{B}_u + \xi_u^{2*,-} - c_1 \right\} \right\} \right), \\ \boldsymbol{\xi}_t^{2*} &= (\xi_t^{2*,+}, \xi_t^{2*,-}) = \left(0, \max \left\{ 0, \max_{0 \leq u \leq t} \left\{ -(x^1 - x^2) + (\boldsymbol{\sigma}^2 - \boldsymbol{\sigma}^1) \cdot \mathbf{B}_u + \xi_u^{1*,-} - c_1 \right\} \right\} \right), \end{aligned} \quad (5.8)$$

where c_1 is the unique solution to

$$\frac{\tilde{\sigma}}{\sqrt{2\rho}} \tanh \left(\frac{\sqrt{2\rho}}{\tilde{\sigma}} x \right) = \frac{p_1'(x) - \frac{K}{2}}{p_1''(x)}, \quad (5.9)$$

and

$$p_1(x) = \mathbb{E} \left[\int_0^\infty e^{-\rho t} h(x + \tilde{\sigma} B_t) dt \right], \quad (5.10)$$

with $\tilde{\sigma} = \sqrt{\sum_{i=1}^2 \sum_{j=1}^D \sigma_{ij}^2}$. The associated Pareto-optimal value is

$$v(x^1, x^2) = \begin{cases} -\frac{\tilde{\sigma}^2 p_1''(c_1) \cosh\left(\frac{(x^1 - x^2) \sqrt{2\rho}}{\tilde{\sigma}}\right)}{2\rho \cosh\left(\frac{c_1 \sqrt{2\rho}}{\tilde{\sigma}}\right)} + p_1(x^1 - x^2), & 0 \leq x^1 - x^2 \leq c_1, \\ v(x^2 + c_1, x^2) + \frac{K_2}{2}(x^1 - x^2 - c_1), & x^1 - x^2 \geq c_1, \\ v(-x^1, -x^2), & x^1 - x^2 < 0. \end{cases} \quad (5.11)$$

(ii) If $K_1 > K_2$ then the following control yields a Pareto-optimal policy to game (5.1)-(5.2),

$$\xi_t^{1*} = (0, 0), \quad \text{and} \quad \xi_t^{2*} = (\xi_t^{2*,+}, \xi_t^{2*,-}) \quad \text{with} \quad (5.12)$$

$$\xi_t^{2*,-} = \max \left\{ 0, \max_{0 \leq u \leq t} \left\{ -(x^1 - x^2) + (\sigma^2 - \sigma^1) \cdot \mathbf{B}_u + \xi_u^{2*,+} - \tilde{c}_1 \right\} \right\}, \quad (5.13)$$

$$\xi_t^{2*,+} = \max \left\{ 0, \max_{0 \leq u \leq t} \left\{ (x^1 - x^2) + (\sigma^1 - \sigma^2) \cdot \mathbf{B}_u + \xi_u^{2*,-} - \tilde{c}_1 \right\} \right\}, \quad (5.14)$$

where \tilde{c}_1 is the unique solution to

$$\frac{\tilde{\sigma}}{\sqrt{2\rho}} \tanh\left(\frac{\sqrt{2\rho}}{\tilde{\sigma}} x\right) = \frac{p_1'(x) - \frac{K_2}{2}}{p_1''(x)}, \quad (5.15)$$

and the associated Pareto-optimal value is

$$v(x^1, x^2) = \begin{cases} -\frac{\tilde{\sigma}^2 p_1''(\tilde{c}_1) \cosh\left(\frac{(x^1 - x^2) \sqrt{2\rho}}{\tilde{\sigma}}\right)}{2\rho \cosh\left(\frac{\tilde{c}_1 \sqrt{2\rho}}{\tilde{\sigma}}\right)} + p_1(x^1 - x^2), & 0 \leq x^1 - x^2 \leq \tilde{c}_1, \\ v(x^2 + \tilde{c}_1, x^2) + \frac{K_2}{2}((x^1 - x^2) - \tilde{c}_1), & x^1 - x^2 \geq \tilde{c}_1, \\ v(-x^1, -x^2), & x^1 - x^2 < 0. \end{cases} \quad (5.16)$$

Remark 17. Note that under **B1-B3**, the Pareto-optimal policy is no longer unique with fixed $L_1 = L_2 = \frac{1}{2}$. For instance, when $K_1 = K_2 = K$, the following control yields another Pareto-optimal policy with the same value function defined in (5.11):

$$\begin{aligned} \xi_t^{1*} &= (\xi_t^{1*,+}, \xi_t^{1*,-}) = (0, 0) \quad \text{and} \quad \xi_t^{2*} = (\xi_t^{2*,+}, \xi_t^{2*,-}), \quad \text{with} \\ \xi_t^{2*,-} &= \max \left\{ 0, \max_{0 \leq u \leq t} \left\{ -(x^1 - x^2) + (\sigma^2 - \sigma^1) \cdot \mathbf{B}_u + \xi_u^{2*,+} - c_1 \right\} \right\}, \\ \xi_t^{2*,+} &= \max \left\{ 0, \max_{0 \leq u \leq t} \left\{ (x^1 - x^2) + (\sigma^1 - \sigma^2) \cdot \mathbf{B}_u + \xi_u^{2*,-} - c_1 \right\} \right\}. \end{aligned} \quad (5.17)$$

Remark 18. According to the Pareto-optimal policy, the optimally controlled dynamics X_t^{1*} and X_t^{2*} are such that $\mathbb{P}(\|X_t^{1*} - X_t^{2*}\| \leq c_1, \forall t \geq 0) = 1$. This suggests that there should be mechanism, such as ‘trimming’, to maintain the dispersion of rates within a certain range. In addition, this solution form indicates that it is socially optimal for the more efficient bank (i.e., the one with the lower cost of control) to take the lead in lending rate adjustment. Consequently, the other bank is a free rider in the game.

Proof. First let us prove the case when $K_1 > K_2$. By Lemma 15, it is sufficient to focus on the single-agent problem (5.4)-(5.5) with $y = x^1 - x^2$. Following the standard analysis (Beneš et al., 1980; Karatzas, 1982), the HJB equation for the one-dimensional control problem follows (5.4)-(5.5) is

$$\max \left\{ \rho u(x) - h(x) - \frac{\tilde{\sigma}^2}{2} u''(x), u'(x) - \frac{K_2}{2}, -u'(x) - \frac{K_2}{2} \right\} = 0. \quad (5.18)$$

There is a \mathcal{C}^2 solution (Beneš et al., 1980; Karatzas, 1982) given by

$$u(x) = \begin{cases} -\frac{\tilde{\sigma}^2 p_1''(\tilde{c}_1) \cosh\left(\frac{x\sqrt{2\rho}}{\tilde{\sigma}}\right)}{2\rho \cosh\left(\tilde{c}_1 \frac{\sqrt{2\rho}}{\tilde{\sigma}}\right)} + p_1(x), & 0 \leq x \leq \tilde{c}_1, \\ u(\tilde{c}_1) + \frac{K_2}{2}(x - \tilde{c}_1), & x \geq \tilde{c}_1, \\ u(-x), & x < 0, \end{cases} \quad (5.19)$$

where \tilde{c}_1 is the unique positive solution to (5.15) and $p_1(x)$ is defined as in (5.10). The corresponding control of the regulator is a bang-bang type such that (5.13)-(5.14) hold. Furthermore, it is easy to see that $v(x^1, x^2) := u(x^1 - x^2)$, with $u(x)$ defined in (5.19), is indeed the value function of problem (5.2).

Next when $K_1 = K_2$, $\xi^{1,+}$ and $\xi^{2,-}$ controls Y_t in the same direction with the same cost. The same holds for $\xi^{2,+}$ or $\xi^{1,-}$, hence the Pareto-optimal policy (5.8) and (5.17). \square

5.2 Pareto optimum vs Nash equilibrium

We now use the above analytical results to compare the Pareto-optimal strategies with the Nash equilibrium strategies, whose definition we recall:

Definition 19 (Nash equilibrium). $\boldsymbol{\eta} = (\eta^1, \dots, \eta^N) \in \mathcal{U}_N$ is a Nash equilibrium strategy of the stochastic game (**N-Player**), if for any $i = 1, \dots, N$, $\mathbf{X}_{0-} = \mathbf{x}$, and any $(\boldsymbol{\eta}^{-i}, \xi^i) \in \mathcal{U}_N$, the following inequality holds,

$$J^i(\mathbf{x}; \boldsymbol{\eta}) \leq J^i(\mathbf{x}; (\boldsymbol{\eta}^{-i}, \xi^i)).$$

$v^i(\mathbf{x}) := J^i(\mathbf{x}; \boldsymbol{\eta})$ is called the Nash equilibrium value for player i associated with $\boldsymbol{\eta}$.

Proposition 20 (Pareto-optimal vs Nash equilibrium solutions when $N = 2$). *Assume **B1-B3** and $K_1 = K_2 = K$, then*

(i) *The following controls give a Nash equilibrium policy to game (5.1)-(5.2):*

$$\begin{aligned} (\eta_t^{1,+}, \eta_t^{1,-}) &= \left(0, \max \left\{ 0, \max_{0 \leq u \leq t} \left\{ (x^1 - x^2) + (\boldsymbol{\sigma}^1 - \boldsymbol{\sigma}^2) \cdot \mathbf{B}_u + \eta_u^{2,-} - c_2 \right\} \right\} \right), \\ (\eta_t^{2,+}, \eta_t^{2,-}) &= \left(0, \max \left\{ 0, \max_{0 \leq u \leq t} \left\{ -(x^1 - x^2) + (\boldsymbol{\sigma}^2 - \boldsymbol{\sigma}^1) \cdot \mathbf{B}_u + \eta_u^{1,-} - c_2 \right\} \right\} \right), \end{aligned} \quad (5.20)$$

where $c_2 > 0$ is the unique positive solution to

$$\frac{\tilde{\sigma}}{\sqrt{2\rho}} \tanh\left(\frac{\sqrt{2\rho}}{\tilde{\sigma}} x\right) = \frac{p_1'(x) - K}{p_1''(x)}, \quad (5.21)$$

with p_1 defined in (5.10). The value functions v^1 and v^2 corresponding to the Nash equilibrium $(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2)$ defined in (5.20) are

$$v^1(x^1, x^2) = \begin{cases} v^1(x^2 - c_2, x^2), & x^1 \leq x^2 - c_2, \\ -\frac{\tilde{\sigma}^2 p_1''(c_2) \cosh\left(\frac{\sqrt{2\rho}}{\tilde{\sigma}}(x^1 - x^2)\right)}{2\rho \cosh\left(c_2 \frac{\sqrt{2\rho}}{\tilde{\sigma}}\right)} + p_1(x^1 - x^2), & x^2 - c_2 \leq x^1 \leq x^2 + c_2, \\ K(x^1 - x^2 - c_2) + v^1(x^2 + c_2, x^2), & x^1 \geq x^2 + c_2, \end{cases} \quad (5.22)$$

and

$$v^2(x^1, x^2) = \begin{cases} v^2(x^1, x^1 - c_2), & x^2 \leq x^1 - c_2, \\ -\frac{\tilde{\sigma}^2 p_1''(c_2) \cosh\left(\frac{\sqrt{2\rho}}{\tilde{\sigma}}(x^2 - x^1)\right)}{2\rho \cosh\left(c_2 \frac{\sqrt{2\rho}}{\tilde{\sigma}}\right)} + p_1(x^2 - x^1), & x^1 - c_2 \leq x^2 \leq x^1 + c_2, \\ K(x^2 - x^1 - c_2) + v^2(x^1, x^1 + c_2), & x^2 \geq x^1 + c_2; \end{cases} \quad (5.23)$$

- (ii) $c_2 > c_1$, where c_1 is the unique positive solution to (5.9) and c_2 is the unique positive solution to (5.21).

That is, Pareto-optimal policy yields tighter threshold for spread, hence game is more stable than that under the Nash equilibrium policy. (See Figure 1.)

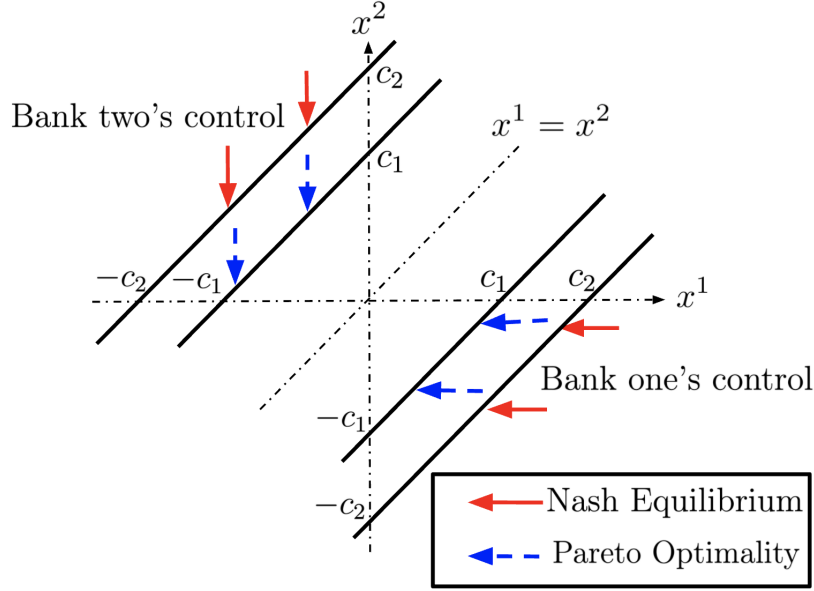


Figure 1: Comparison: Nash and Pareto ($K_1 = K_2$).

Proof. Similar to the derivation in (Guo & Xu, 2019), we have the following Quasi-variational inequalities for the Nash equilibrium of game (5.1) with J^1 and J^2 and $K_1 = K_2 = K$,

$$\left\{ \begin{array}{l} \max \left\{ \rho v^i(x^1, x^2) - h(x^1 - x^2) - \frac{\tilde{\sigma}^2}{2} \left(\partial_{x^1}^2 v^i(x^1, x^2) + \partial_{x^2}^2 v^i(x^1, x^2) \right), \right. \\ \quad \left. \partial_{x^i} v^i(x^1, x^2) - K, -\partial_{x^i} v^i(x^1, x^2) - K \right\} = 0, \\ \quad \quad \quad \text{on } \{(x^1, x^2) : -K < \partial_{x^j} v^j(x^1, x^2) < K\}, \\ \partial_{x^i} v^i(x^1, x^2) = 0, \quad \quad \quad \text{on } \{(x^1, x^2) : \partial_{x^j} v^j(x^1, x^2) = K \text{ or } \partial_{x^j} v^j(x^1, x^2) = -K\}, \end{array} \right.$$

for $i \neq j$ and $1 \leq i, j \leq 2$. Moreover, one can show that (5.22)-(5.23) are the solution to (5.24). By verification theorem (Guo & Xu, 2019, Theorem 3), some detailed calculations can verify that (5.22)-(5.23) are the game values associated with the Nash equilibrium policy (5.20).

Now we provide the proof for Claim (ii). Define $g(x) = \frac{\tilde{\sigma}}{\sqrt{2\rho}} \tanh\left(\frac{\sqrt{2\rho}}{\tilde{\sigma}}x\right)$, $g_1(x) = \frac{p_1'(x) - \frac{K}{2}}{p_1''(x)}$ and $g_2(x) = \frac{p_1'(x) - K}{p_1''(x)}$, where p_1 is defined in (5.10). Then $g(0) = 0$, $g'(x) > 0$ for any $x \in \mathbb{R}^+$, and $\lim_{x \rightarrow \infty} g(x) = \frac{\tilde{\sigma}}{\sqrt{2\rho}}$.

Thanks to Assumption **(B2)**,

$$0 < \frac{c}{\rho} \leq p_1''(x) = \mathbb{E} \int_0^\infty e^{-\rho t} h''(x + \tilde{\sigma} B_t) dt \leq \frac{C}{\rho}.$$

The function $p_1'(x)$ is negative at $x = 0$ and increases monotonically to ∞ on \mathbb{R}^+ . Hence there exists a unique positive zero c_0 . Moreover, for any $x > c_0$, $g_1'(x) = 1 - \frac{p_1''(x)}{p_1''(x)} g_1(x) \geq 1$. This is because $p_1'''(x) \leq 0$ for $x \geq 0$. We conclude that there exists a unique point $c_0 < c_1 < \infty$ such that $g(c_1) = g_1(c_1)$.

Now apply similar analysis to c_2 , which is the unique solution to $g(x) = g_2(x)$ such that $0 < c_2 < \infty$. Notice that, $g_1(x) - g_2(x) = \frac{K}{2p_1''(x)} > 0$ because $p_2''(x) > 0$. Hence $c_2 > c_1$. \square

References

- Aïd, R., Basei, M., & Pham, H. (2017). The coordination of centralised and distributed generation. *arXiv preprint arXiv:1705.01302*.
- Avellaneda, M., & Cont, R. (2010). Transparency in over-the-counter interest rate derivatives markets. *Finance Concepts*.
- Bator, F. M. (1957). The simple analytics of welfare maximization. *The American Economic Review*, 47(1), 22–59.
- Beneš, V. E., Shepp, L. A., & Witsenhausen, H. S. (1980). Some solvable stochastic control problems. *Stochastics: An International Journal of Probability and Stochastic Processes*, 4(1), 39–83.
- Bensoussan, A., Long, H., Perera, S., & Sethi, S. (2012). Impulse control with random reaction periods: a central bank intervention problem. *Operations Research Letters*, 40(6), 425–430.
- Cadenillas, A., & Zapatero, F. (2000). Classical and impulse stochastic control of the exchange rate using interest rates and reserves. *Mathematical Finance*, 10(2), 141–156.
- Carmona, R., Fouque, J.-P., & Sun, L.-H. (2013). *Mean field games and systemic risk* (Working Paper).
- Chiarolla, M. B., Ferrari, G., & Riedel, F. (2013). Generalized Kuhn–Tucker conditions for N-firm stochastic irreversible investment under limited resources. *SIAM Journal on Control and Optimization*, 51(5), 3863–3885.
- Coleman, J. L. (1979). Efficiency, utility, and wealth maximization. *Hofstra L. Rev.*, 8, 509.
- Davis, M. H., & Norman, A. R. (1990). Portfolio selection with transaction costs. *Mathematics of Operations Research*, 15(4), 676–713.
- Davis, M. H., Panas, V. G., & Zariphopoulou, T. (1993). European option pricing with transaction costs. *SIAM Journal on Control and Optimization*, 31(2), 470–493.
- De Angelis, T., & Ferrari, G. (2018). Stochastic nonzero-sum games: a new connection between singular control and optimal stopping. *Advances in Applied Probability*, 50(2), 347–372.
- Dianetti, J., & Ferrari, G. (2020). Nonzero-sum submodular monotone-follower games: existence and approximation of Nash equilibria. *SIAM Journal on Control and Optimization*, 58(3), 1257–1288.
- Duffie, D., & Stein, J. C. (2015). Reforming LIBOR and other financial market benchmarks. *Journal of Economic Perspectives*, 29(2), 191–212.
- Dupuis, P., & Ishii, H. (1991). On Lipschitz continuity of the solution mapping to the Skorokhod problem, with applications. *Stochastics: An International Journal of Probability and Stochastic Processes*, 35(1), 31–62.
- Dupuis, P., & Ishii, H. (1993). SDEs with oblique reflection on nonsmooth domains. *The annals of Probability*, 554–580.
- Ferrari, G., Riedel, F., & Steg, J.-H. (2017). Continuous-time public good contribution under uncertainty: a stochastic control approach. *Applied Mathematics & Optimization*, 75(3), 429–470.
- Fischer, M., & Livieri, G. (2016). Continuous time mean-variance portfolio optimization through the mean field approach. *ESAIM: Probability and Statistics*, 20, 30–44.

- Gilbarg, D., & Trudinger, N. S. (2015). *Elliptic Partial Differential Equations of Second Order*. Springer.
- Gomes, D. A., Mohr, J., & Souza, R. R. (2010). Discrete time, finite state space mean field games. *Journal de Mathématiques Pures et Appliquées*, *93*(3), 308–328.
- Guo, X., & Xu, R. (2019). Stochastic games for fuel follower problem: N versus MFG. *SIAM Journal on Control and Optimization*, *57*(1), 659–692.
- Hernandez-Hernandez, D., Simon, R. S., Zervos, M., et al. (2015). A zero-sum game between a singular stochastic controller and a discretionary stopper. *Annals of Applied Probability*, *25*(1), 46–80.
- H.M. Treasury. (2012). *The Wheatley Review of LIBOR: Final Report* (Tech. Rep.). Author. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/191762/wheatley_review_libor_finalreport_280912.pdf
- Jeanblanc-Picqué, M. (1993). Impulse control method and exchange rate. *Mathematical Finance*, *3*(2), 161–177.
- Kallsen, J., & Muhle-Karbe, J. (2015). Option pricing and hedging with small transaction costs. *Mathematical Finance*, *25*(4), 702–723.
- Kallsen, J., & Muhle-Karbe, J. (2017). The general structure of optimal investment and consumption with small transaction costs. *Mathematical Finance*, *27*(3), 659–703.
- Karatzas, I. (1982). A class of singular stochastic control problems. In *Stochastic differential systems* (pp. 312–319). Springer.
- Kruk, L. (2000). Optimal policies for N-dimensional singular stochastic control problems part I: the Skorokhod problem. *SIAM Journal on Control and Optimization*, *38*(5), 1603–1622.
- Kwon, H., & Zhang, H. (2015). Game of singular stochastic control and strategic exit. *Mathematics of Operations Research*, *40*(4), 869–887.
- Lasry, J.-M., & Lions, P.-L. (2007). Mean field games. *Japanese Journal of Mathematics*, *2*(1), 229–260.
- Lions, P.-L. (1983). A remark on Bony maximum principle. *Proceedings of the American Mathematical Society*, *88*(3), 503–508.
- Menaldi, J.-L., & Robin, M. (1983). On some cheap control problems for diffusion processes. *Transactions of the American Mathematical Society*, *278*(2), 771–802.
- Menaldi, J.-L., & Taksar, M. I. (1989). Optimal correction problem of a multidimensional stochastic system. *Automatica*, *25*(2), 223–232.
- Meyer, P. A. (1976). Martingales locales changement de variables, formules exponentielles. In *Séminaire de Probabilités X Université de Strasbourg* (pp. 291–331). Springer.
- Ramanan, K. (2006). Reflected diffusions defined via the extended Skorokhod map. *Electronic journal of probability*, *11*, 934–992.
- Rosen, J. (1987). Joint continuity of the intersection local times of Markov processes. *Annals of Probability*, *15*, 659–675.
- Soner, H. M., & Shreve, S. E. (1989). Regularity of the value function for a two-dimensional singular stochastic control problem. *SIAM Journal on Control and Optimization*, *27*(4), 876–907.

- Sun, L.-H. (2018). Systemic risk and interbank lending. *Journal of Optimization Theory and Applications*, 179(2), 400–424.
- Wang, W.-K., & Ewald, C.-O. (2010). Dynamic voluntary provision of public goods with uncertainty: a stochastic differential game model. *Decisions in Economics and Finance*, 33(2), 97–116.
- Williams, S., Chow, P., & Menaldi, J. (1994). Regularity of the free boundary in singular stochastic control. *Journal of Differential Equations*, 111(1), 175 - 201. doi: <https://doi.org/10.1006/jdeq.1994.1080>
- Zariphopoulou, T. (1992). Investment-consumption models with transaction fees and Markov-chain parameters. *SIAM Journal on Control and Optimization*, 30(3), 613–636.