*The Annals of Applied Probability*
2017, Vol. 27, No. 4, 2455–2514
DOI: 10.1214/16-AAP1264


# ASYMPTOTIC LOWER BOUNDS FOR OPTIMAL TRACKING: A LINEAR PROGRAMMING APPROACH

BY JIATU CAI*, MATHIEU ROSENBAUM† AND PETER TANKOV*

*Universtité Paris Diderot (Paris 7)* and Ecole Polytechnique†*

We consider the problem of tracking a target whose dynamics is modeled by a continuous Itô semi-martingale. The aim is to minimize both deviation from the target and tracking efforts. We establish the existence of asymptotic lower bounds for this problem, depending on the cost structure. These lower bounds can be related to the time-average control of Brownian motion, which is characterized as a deterministic linear programming problem. A comprehensive list of examples with explicit expressions for the lower bounds is provided.

## CONTENTS



Received December 2015; revised September 2016.
*MSC2010 subject classifications.* 93E20.
*Key words and phrases.* Optimal tracking, asymptotic lower bound, occupation measure, linear programming, singular control, impulse control, regular control.

**1. Introduction.** We consider the problem of tracking a target whose dynamics $(X_t^\circ)$ is modeled by a continuous Itô semi-martingale defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ with values in $\mathbb{R}^d$ such that

$$dX_t^\circ = b_t\, dt + \sqrt{a}_t\, dW_t, \qquad X_0^\circ = 0. \tag{1.1}$$

Here, $(W_t)$ is a $d$-dimensional Brownian motion and $(b_t)$, $(a_t)$ are predictable processes with values in $\mathbb{R}^d$ and $\mathcal{S}_+^d$, the set of $d \times d$ symmetric positive definite matrices, respectively. An agent observes $X_t^\circ$ and adjusts her position in order to follow $X_t^\circ$. However, she has to pay certain intervention costs for position adjustments. The objective is to stay close to the target $X_t^\circ$ while minimizing the tracking efforts. This problem arises naturally in various situations such as discretization of option hedging strategies [16, 17, 52], management of an index fund [33, 48], control of exchange rate [12, 45], portfolio selection under transaction costs [2, 31, 49, 53], trading under market impact [4, 42, 44] or illiquidity costs [46, 51]. The aim of this paper is to describe precisely and in full generality how the problem of optimally tracking the continuous Itô semi-martingale $X^\circ$ can be reduced to the local problem of controlling a Brownian motion in the setting when the cost of position adjustments is small. This simplifies the original problem and in many cases leads to an explicit solution.

More precisely, let $(Y_t^\psi)$ be the position of the agent determined by the control $\psi$, with $Y_t^\psi \in \mathbb{R}^d$ and let $(\overline{X}_t)$ be the deviation of the agent from the target $(X_t^\circ)$, so that

$$\overline{X}_t = -X_t^\circ + Y_t^\psi. \tag{1.2}$$

Let $H_0(\overline{X})$ be the penalty functional for the deviation from the target and $H(\psi)$ the cost incurred by the control $\psi$ up to a finite horizon $T$ (see below for precise definitions). Then the aim of the agent is to minimize the aggregate cost:

$$J(\psi) = H_0(\overline{X}) + H(\psi) \tag{1.3}$$

over a set of admissible strategies to be described in detail below.

The functional $J(\psi)$ is a random variable, and the set of real-valued random variables is only partially ordered, so that the problem of minimizing $J(\psi)$ over the set of admissible strategies does not make sense a priori. Even if one is interested in minimizing the expectation of $J(\psi)$, this problem rarely admits an explicit solution. In this paper, we propose an asymptotic framework where the tracking costs are small and derive a pathwise lower bound for (1.3) under this setting.

More precisely, we introduce a parameter $\varepsilon$ tending to zero and consider a family of cost functionals $H^\varepsilon(\psi)$. For example, we can have $H^\varepsilon(\psi) = \varepsilon^{\beta_\psi} H(\psi)$ for some constant $\beta_\psi$, but different components of the cost functional may also scale with $\varepsilon$ at different rates. For each $\varepsilon > 0$, we define the controlled deviation process:

$$\overline{X}_t^\varepsilon = -X_t^\circ + Y_t^{\psi^\varepsilon}, \tag{1.2-$\varepsilon$}$$

and the aggregate cost functional

$$J^\varepsilon(\psi^\varepsilon) = H_0(\overline{X}^\varepsilon) + H^\varepsilon(\psi^\varepsilon). \tag{1.3-$\varepsilon$}$$

The main result of this paper states that there exists $\beta^* > 0$ explicitly determined by the cost structure $H_0$ and $(H^\varepsilon)_{\varepsilon>0}$ such that for all $\delta > 0$ and any sequence of admissible strategies $\{\psi^\varepsilon, \varepsilon > 0\}$, we have

$$\lim_{\varepsilon\to 0+} \mathbb{P}\left[\frac{1}{\varepsilon^{\beta^*}} J^\varepsilon(\psi^\varepsilon) \geq \int_0^T I_t\, dt - \delta\right] = 1, \tag{1.4}$$

where $I_t$ is the optimal cost of the time-average (ergodic) control problem of Brownian motion with parameters frozen at time $t$ (see the statements of Theorems 2.1 and 3.1–3.4 for a precise formulation for different control types).

The small-cost framework therefore allows to obtain pathwise asymptotic bounds on the functional $J^\varepsilon$ in terms of the time-average control problem of Brownian motion (TACBM) with constant parameters, and thus gives a sense to the problem of pathwise minimization of the tracking costs. In many practically important cases (see Examples 4.3–4.7), we are able to solve explicitly the TACBM problem. Moreover, in a companion paper [13], we show that for the examples where the TACBM problem admits an explicit solution, the lower bound is tight (see Remark 3.1).

Our result enables us to revisit the asymptotic lower bounds for the discretization of hedging strategies in [16, 17]. In these papers, the lower bounds are deduced by using subtle inequalities. Here, we show that these bounds can be interpreted in a simple manner through the time-average control problem of Brownian motion.

The TACBM problem also arises in the study of utility maximization under transaction costs; see [2, 44, 49, 53]. This is not surprising since at first order, these problems and the tracking problem are essentially the same; see Section 5. In the above references, the authors derive the PDE associated to the first-order correction to the maximal expected utility in the frictionless market, which turns out to be the HJB equation associated to the time-average control of Brownian motion. Inspired by [36] and [41], our approach, based on weak convergence of empirical occupation measures, is very different from the PDE-based method and enables us to treat more general situations. Contrary to [36], where the lower bound holds under expectation, we obtain pathwise lower bounds. Compared to [41], we are able to treat impulse control and general dynamics for the target.

The paper is organized as follows. In the rest of the Introduction, we describe precisely the control strategies of the agent and the corresponding cost functionals, and introduce the necessary notation. In Section 2, we introduce our asymptotic framework, state the main result and provide a heuristic derivation of the lower bound in the case of combined regular and impulse control. Various extensions are then discussed in Section 3. In Section 4, we provide an accurate definition for the time-average control of Brownian motion using a relaxed martingale formulation and collect a comprehensive list of explicit solutions in dimension one. The connection with utility maximization with small market frictions is made in Section 5 and the proofs are given in Sections 6, 7 and 8.

*Deviation penalty and cost functionals.* As is usually done in the literature (see, e.g., [32, 43]), we consider a penalty $H_0(\overline{X})$ for the deviation from the target of additive form:

$$H_0(\overline{X}) = \int_0^T r_t D(\overline{X}_t)\,dt,$$

where $(r_t)$ is a random weight process and $D(x)$ a deterministic function. For example, we can take $D(x) = \langle x, \Sigma^D x\rangle$ where $\Sigma^D$ is positive definite and $\langle\cdot,\cdot\rangle$ is the inner product in $\mathbb{R}^d$. On the other hand, depending on the nature of the costs, the agent can either control her speed at all times or jump towards the target instantaneously. The control $\psi$ and the cost functional $H(\psi)$ belong to one of the following classes:

1. *Impulse control*. There is a fixed cost component for each action, so the agent has to intervene in a discrete way. The class $\mathcal{A}^I$ of admissible controls contains all sequences $\{(\tau_j, \xi_j), j \in \mathbb{N}^*\}$ where $\{\tau_j, j \in \mathbb{N}^*\}$ is a strictly increasing sequence of stopping times representing the jump times and satisfying $\lim_{j\to\infty}\tau_j = +\infty$, and for each $j \in \mathbb{N}^*$, $\xi_j \in \mathbb{R}^d$ is a $\mathcal{F}_{\tau_j}$-measurable random vector representing the size of $j$th jump. The position of the agent is given by[1]

$$Y_t = \sum_{0<\tau_j\le t} \xi_j$$

and the cumulated cost is then given by

$$H(\psi) = \sum_{0<\tau_j\le T} k_{\tau_j} F(\xi_j),$$

where $(k_t)$ is a random weight process and $F(\xi) > 0$ is the cost of a jump with size $\xi \neq 0$. If we take $k_t = 1$ and $F(\xi) = \sum_{i=1}^d \mathbb{1}_{\{\xi^i\neq 0\}}$ where $\xi^i$ is the ith component of $\xi$, then $H(\psi)$ represents the total number of actions on each component over the time interval $[0, T]$; see [16, 17]. If $F(\xi) = \sum_{i=1}^d \mathbb{1}_{\{\xi^i\neq 0\}} + \sum_{i=1}^d P_i|\xi^i|$ where $P_i \ge 0$, we say that the cost has a fixed component and a proportional component.

[1] We assume that $Y_0 = 0$ to simplify notation only. Our arguments and results remain unchanged if the agent starts from a "legacy value" $Y_0 \neq 0$.

2. *Singular control*. If the cost is proportional to the size of the jump, then infinitesimal displacement is also allowed and it is natural to model $(Y_t)$ by a process with bounded variation. In this case, the class $\mathcal{A}^S$ of admissible controls contains all couples $(\gamma, \varphi)$ where $\varphi$ is a progressively measurable increasing process with $\varphi_{0-} = 0$, which represents the cumulated amount of intervention and $\gamma$ is a progressively measurable process with $\gamma_t \in \Delta := \{n \in \mathbb{R}^d | \sum_{i=1}^d |n^i| = 1\}$ for all $t \geq 0$, which represents the distribution of the control effort in each direction. In other words, $\varphi_t = \sum_{i=1}^d \|Y^i\|_t$ where $\|\cdot\|$ denotes the absolute variation of a process, and $\gamma_t^i$ is the Radon–Nikodym derivative of $Y_t^i$ with respect to $\varphi_t$. The position of the agent is given by

$$Y_t = \int_0^t \gamma_s \, d\varphi_s,$$

and the corresponding cost is usually given as (see, e.g., [31, 53])

$$H(\psi) = \int_0^T h_t P(\gamma_t) \, d\varphi_t,$$

where $(h_t)$ is a random weight process and we take (for example) $P(\gamma) = \langle P, |\gamma| \rangle$ with $P \in \mathbb{R}_+^d$ and $|\gamma| = (|\gamma^1|, \ldots, |\gamma^d|)^\top$. The vector $P = (P_1, \ldots, P_d)^\top$ represents the coefficients of proportional costs in each direction.

3. *Regular control.* Most often, the process $(Y_t)$ is required to be absolutely continuous with respect to time; see, for example, [44, 51] among many others. In this case, the class $\mathcal{A}^R$ of admissible controls contains all progressively measurable integrable processes $u$ with values in $\mathbb{R}^d$, representing the speed of the agent, the position of the agent is given by

$$Y_t = \int_0^t u_s \, ds,$$

and the cost functional is

$$H(\psi) = \int_0^T l_t Q(u_t) \, dt,$$

where $(l_t)$ is a random weight process and, for example, $Q(u) = \langle u, \Sigma^Q u \rangle$ with $\Sigma^Q$ a positive definite matrix. Comparing to the case of singular control where the control variables are $(\gamma_t)$ and $(\varphi_t)$; here, we optimize over $(u_t)$.

4. *Combined control.* It is possible that several types of control are available to the agent. In that case, $\psi = (\psi_1, \ldots, \psi_n)$ where for each $i$, $\psi_i$ belongs to one of the classes introduced before. For example, in the case of combined regular and impulse control (see [45]), the position of the agent is given by

$$Y_t = \sum_{0<\tau_j \leq t} \xi_j + \int_0^t u_s \, ds,$$

while the cost functional is given by

$$H(\psi) = \sum_{0<\tau_j\le T} k_{\tau_j} F(\xi_j) + \int_0^T l_t Q(u_t)\,dt.$$

Similarly, one can consider other combinations of controls.

1.1. *Notation*. For a complete, separable, metric space $S$, we define $C(S)$ the set of continuous functions on $S$, $C_b(S)$ the set of bounded, continuous functions on $S$, $\mathcal{M}(S)$ the set of finite nonnegative Borel measures on $S$ and $\mathcal{P}(S)$ the set of probability measures on $S$. The sets $\mathcal{M}(S)$ and $\mathcal{P}(S)$ are equipped with the topology of weak convergence. Finally, $C_0^2(\mathbb{R}^d)$ denotes the set of twice continuously differentiable real functions on $\mathbb{R}^d$ with compact support, equipped with the norm:

$$\|f\|_{C_0^2} = \|f\|_\infty + \sum_{i=1}^d \|\partial_i f\|_\infty + \sum_{i,j=1}^d \|\partial_{ij}^2 f\|_\infty.$$

**2. Tracking with combined regular and impulse control.** Instead of giving directly a general result, which would lead to a cumbersome presentation, we focus in this section on the tracking problem with combined regular and impulse control. This allows us to illustrate our key ideas. Other situations, such as singular control, are discussed in Section 3.

2.1. *Assumptions and statement of the main result*. In the case of combined regular and impulse control, a tracking strategy $(u,\tau,\xi)$ is given by a progressively measurable process $u=(u_t)_{t\ge 0}$ with values in $\mathbb{R}^d$ and $(\tau,\xi)=\{(\tau_j,\xi_j), j\in\mathbb{N}^*\}$, with $(\tau_j)$ an increasing sequence of stopping times and $(\xi_j)$ a sequence of random variables with values in $\mathbb{R}^d$ such that $\xi_j$ is $\mathcal{F}_{\tau_j}$-measurable for $j=1,2,\dots$. The process $(u_t)$ represents the speed of the agent. The stopping time $\tau_j$ represents the timing of jth jump and $\xi_j$ the size of the jump. The tracking error obtained by following the strategy $(u,\tau,\xi)$ is given by

$$\overline{X}_t = -X_t^\circ + \int_0^t u_s\,ds + \sum_{0<\tau_j\le t} \xi_j.$$

We recall that $X^\circ$ is a continuous Itô semi-martingale of the form (1.1). From now on, we require that its coefficients satisfy the following assumption.

ASSUMPTION 2.1 (Model). The processes $(a_t)$ and $(b_t)$ are continuous and $(a_t)$ is positive definite on $[0,T]$.

At any time, the agent is paying a cost for maintaining the speed $u_t$ and each jump $\xi_j$ incurs a positive cost. We are interested in the following type of cost functional:

$$J(u,\tau,\xi) = \int_0^T \big(r_t D(\overline{X}_t) + l_t Q(u_t)\big)\,dt + \sum_{j:0<\tau_j\le T} \big(k_{\tau_j} F(\xi_j) + h_{\tau_j} P(\xi_j)\big),$$

where $T \in \mathbb{R}^+$ and $(r_t)$, $(l_t)$, $(k_t)$ and $(h_t)$ are random weight processes satisfying the following assumption.

ASSUMPTION 2.2 (Optimization criterion). The parameters of the cost functional $(r_t)$, $(l_t)$, $(k_t)$ and $(h_t)$ are continuous and positive on $[0,T]$.

REMARK 2.1. The continuity property in Assumptions 2.1 and 2.2 is essential for our method since it allows to reduce our optimization problem to a sequence of local problems with constant coefficients. Note that the random weight processes $(r_t)$, $(l_t)$, $(k_t)$ and $(h_t)$ are not assumed to be adapted.

The cost functions $D$, $Q$, $F$, $P$ are deterministic lower semicontinuous functions which satisfy the following homogeneity property:

$$\begin{aligned} D(\varepsilon x) &= \varepsilon^{\zeta_D} D(x), \qquad & Q(\varepsilon u) &= \varepsilon^{\zeta_Q} Q(u), \\ F(\varepsilon\xi) &= \varepsilon^{\zeta_F} F(\xi), \qquad & P(\varepsilon\xi) &= \varepsilon^{\zeta_P} P(\xi), \end{aligned} \tag{2.1}$$

for any $\varepsilon > 0$ and

$$\zeta_D > 0, \qquad \zeta_Q > 1, \qquad \zeta_F = 0, \qquad \zeta_P = 1.$$

Note that here we slightly extend the setting of the previous section by introducing two functions $F$ and $P$ which typically represent the fixed and the proportional costs (applied to impulse controls), respectively.

In this paper, we essentially have in mind the case where

$$D(x) = \langle x, \Sigma^D x\rangle, \qquad Q(u) = \langle u, \Sigma^Q u\rangle,$$

$$F(\xi) = \sum_{i=1}^d F_i \mathbb{1}_{\{\xi^i \neq 0\}}, \qquad P(\xi) = \sum_{i=1}^d P_i |\xi^i|,$$

with $F_i, P_i \in \mathbb{R}_+$ such that

$$\min_i F_i > 0, \tag{2.2}$$

and $\Sigma^D, \Sigma^Q \in \mathcal{S}_+^d$. Note that in this situation, we have $\zeta_D = \zeta_Q = 2$.

We now explain our asymptotic setting where the costs are small. To this end, we introduce a family of optimization functionals $(J^\varepsilon)_{\varepsilon>0}$. For each $\varepsilon > 0$, the functional $J^\varepsilon$ is defined by

$$\begin{aligned} J^\varepsilon(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon) = {} & \int_0^T \big(r_t D(\overline{X}_t^\varepsilon) + \varepsilon^{\beta_Q} l_t Q(u_t^\varepsilon)\big)\,dt \\ & + \sum_{j:0<\tau_j^\varepsilon\le T} \big(\varepsilon^{\beta_F} k_{\tau_j^\varepsilon} F(\xi_j^\varepsilon) + \varepsilon^{\beta_P} h_{\tau_j^\varepsilon} P(\xi_j^\varepsilon)\big), \end{aligned} \tag{2.3}$$

where $\beta_Q$, $\beta_F$ and $\beta_P$ are positive constants and

$$\overline{X}_t^\varepsilon = -X_t^\circ + \int_0^t u_s^\varepsilon \, ds + \sum_{j:0<\tau_j^\varepsilon \le t} \xi_j^\varepsilon.$$

Our first main result is a precise asymptotic relation between $J^\varepsilon$ and the time-average impulse-regular control problem of Brownian motion with constant parameters. In the case of combined impulse and regular control considered in this section, the dynamics of the controlled Brownian motion is given by

$$X_s = \sqrt{a} W_s + \int_0^s u_\nu \, d\nu + \sum_{0<\tau_j \le s} \xi_j, \tag{2.4}$$

and the time-average control problem can be formulated as

$$\begin{aligned} I^E(a, r, l, k, h) = \inf_{(\tau_j, \xi_j, u)} \varlimsup_{S\to\infty} \frac{1}{S} \mathbb{E}\biggl[&\int_0^S \bigl(rD(X_s) + lQ(u_s)\bigr) \, ds \\ &+ \sum_{0<\tau_j \le S} \bigl(kF(\xi_j) + hP(\xi_j)\bigr)\biggr], \end{aligned} \tag{2.5}$$

where the subscript $E$ stands for expectation and the infimum is computed over $u \in \mathcal{A}^R$ and $(\tau_j, \xi_j) \in \mathcal{A}^I$.

At the level of generality that we are interested in, we need to consider a relaxed formulation of the above control problem, as a linear programming problem on the space of measures. Following [37], we introduce for each control $(u, \xi)$ the occupation measures:

$$\mu_t(H_1) = \frac{1}{t} \mathbb{E} \int_0^t \mathbb{1}_{H_1}(X_s, u_s) \, ds, \qquad H_1 \in \mathcal{B}(\mathbb{R}^d \times \mathbb{R}^d),$$

$$\rho_t(H_2) = \frac{1}{t} \mathbb{E} \sum_{0<\tau_j \le t} \mathbb{1}_{H_2}(X_{s-}, \xi_j), \qquad H_2 \in \mathcal{B}(\mathbb{R}^d \times \mathbb{R}^d).$$

Assuming that these measures do not depend on time, we get that

$$\begin{aligned} &\varlimsup_{S\to\infty} \frac{1}{S} \mathbb{E}\biggl[\int_0^S \bigl(rD(X_s) + lQ(u_s)\bigr) \, ds + \sum_{0<\tau_j \le S} \bigl(kF(\xi_j) + hP(\xi_j)\bigr)\biggr] \\ &\quad = \int_{\mathbb{R}^d \times \mathbb{R}^d} \bigl(rD(x) + lQ(u)\bigr) \mu(dx \times du) \\ &\qquad + \int_{\mathbb{R}^d \times \mathbb{R}^d} \bigl(kF(\xi) + hP(\xi)\bigr) \rho(dx \times d\xi). \end{aligned}$$

On the other hand, by Itô's formula, for any $f \in C_0^2$,

$$\begin{aligned} f(X_t) = f(X_0) &+ \sqrt{a} \int_0^t f'(X_s) \, dW_s + \int_0^t A^a f(X_s, u_s) \, ds \\ &+ \sum_{0<\tau_j \le t} Bf(X_{\tau_j -}, \xi_j), \end{aligned} \tag{2.6}$$

where

$$A^a f(x,u) = \frac{1}{2}\sum_{i,j} a_{ij}\partial_{ij}^2 f(x) + \langle u, \nabla f(x)\rangle,$$
$$Bf(x,\xi) = f(x+\xi) - f(x).$$

Taking expectation in (2.6) and assuming that the law of $X_t$ does not depend on $t$, we see that under adequate integrability conditions the measures $\mu$ and $\rho$ satisfy the constraint:

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} A^a f(x,u)\mu(dx\times du) + \int_{\mathbb{R}^d\times\mathbb{R}^d} Bf(x,\xi)\rho(dx\times d\xi) = 0. \tag{2.7}$$

Therefore, the time-average control problem of Brownian motion (2.4)–(2.5) is closely related to the problem of computing

$$\begin{aligned} I(a,r,l,k,h) = \inf_{\mu,\rho} &\int_{\mathbb{R}^d\times\mathbb{R}^d} \big(rD(x) + lQ(u)\big)\mu(dx\times du) \\ &+ \int_{\mathbb{R}^d\times\mathbb{R}^d} \big(kF(\xi) + hP(\xi)\big)\rho(dx\times d\xi), \end{aligned} \tag{2.8}$$

with $(\mu,\rho) \in \mathcal{P}(\mathbb{R}^d\times\mathbb{R}^d)\times\mathcal{M}(\mathbb{R}^d\times\mathbb{R}^d\setminus\{0\})$ satisfying the constraint (2.7) for any $f\in C_0^2$. We will see in Section 4.2 that this is essentially an equivalent characterization of the time-average control problem (2.4)–(2.5). In Example 4.6, we consider a particular case for which $I$ and the optimal solution $\mu^*$, $\rho^*$ can be explicitly determined. To give a precise formulation of our first main result, we introduce two additional assumptions.

ASSUMPTION 2.3 (Regularity of linear programming). The function $I = I(a,r,l,k,h)$ defined by (2.7)–(2.8) is measurable.

ASSUMPTION 2.4 (Asymptotic framework). The cost functionals are lower semicontinuous functions which satisfy the homogeneity property (2.1) and there exists $\beta > 0$ such that

$$\frac{\beta_F}{\zeta_D + 2 - \zeta_F} = \frac{\beta_P}{\zeta_D + 2 - \zeta_P} = \frac{\beta_Q}{\zeta_D + \zeta_Q} = \beta. \tag{2.9}$$

Moreover, $\inf_{\|x\|=1} D(x) > 0$, $\inf_{\|u\|=1} Q(u) > 0$ and $\inf_{\|\xi\|=1} F(\xi) > 0$.

REMARK 2.2. Let us comment briefly on the above assumptions. Assumption 2.3 is necessary to avoid pathological cases. In most examples, the function $I$ is continuous (see Examples 4.3–4.7). Assumption 2.4 ensures that all components of the cost functional have similar order of magnitude in the limit.

Second, we introduce the following notion.

DEFINITION 2.1. Let $\{Z, (Z^\varepsilon)_\varepsilon, \varepsilon > 0\}$ be random variables on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We say that $Z^\varepsilon$ is asymptotically bounded from below by $Z$ in probability if

$$\forall \delta > 0, \qquad \lim_{\varepsilon\to 0} \mathbb{P}[Z^\varepsilon > Z - \delta] = 1.$$

We write $\liminf_{\varepsilon\to 0} Z^\varepsilon \geq_p Z$.

We now give the version of our main result for the case of combined regular and impulse control.

THEOREM 2.1 (Asymptotic lower bound for combined regular and impulse control). *Under Assumptions* 2.1, 2.2, 2.3 *and* 2.4, *we have*

$$\liminf_{\varepsilon\to 0} \frac{1}{\varepsilon^{\beta\zeta_D}} J^\varepsilon(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon) \geq_p \int_0^T I(a_t, r_t, l_t, k_t, h_t)\, dt, \tag{2.10}$$

*for any sequence of admissible tracking strategies* $\{(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon) \in \mathcal{A}^R \times \mathcal{A}^S, \varepsilon > 0\}$.

By Lemma D.1, the following corollary holds.

COROLLARY 2.1. *We have*

$$\liminf_{\varepsilon\to 0} \frac{1}{\varepsilon^{\zeta_D\beta}} \mathbb{E}[J^\varepsilon(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon)] \geq \mathbb{E}\left[\int_0^T I(a_t, r_t, l_t, k_t, h_t)\, dt\right].$$

2.2. *Heuristic derivation of the main result*. We start with the expression (2.3) for the family of optimization functionals. The key step is to decompose the tracking problem into a sequence of local problems. More precisely, let $\{t_k^\varepsilon = k\delta^\varepsilon, k = 0, 1, \dots, K^\varepsilon\}$ be a partition of the interval $[0, T]$ with $\delta^\varepsilon \to 0$ as $\varepsilon \to 0$. Then we can write

$$\begin{aligned}
J^\varepsilon(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon) &= \sum_{k=0}^{K^\varepsilon-1} \biggl(\int_{t_k^\varepsilon}^{t_k^\varepsilon+\delta^\varepsilon} \bigl(r_t D(\overline{X}_t^\varepsilon) + \varepsilon^{\beta_Q} l_t Q(u_t^\varepsilon)\bigr)\, dt \\
&\qquad + \sum_{j: t_k^\varepsilon < \tau_j^\varepsilon \leq t_k^\varepsilon+\delta^\varepsilon} \bigl(\varepsilon^{\beta_F} k_{\tau_j^\varepsilon} F(\xi_j^\varepsilon) + \varepsilon^{\beta_P} h_{\tau_j^\varepsilon} P(\xi_j^\varepsilon)\bigr)\biggr) \\
&= \sum_{k=0}^{K^\varepsilon-1} \frac{1}{\delta^\varepsilon} \biggl(\int_{t_k^\varepsilon}^{t_k^\varepsilon+\delta^\varepsilon} \bigl(r_t D(\overline{X}_t^\varepsilon) + \varepsilon^{\beta_Q} l_t Q(u_t^\varepsilon)\bigr)\, dt \\
&\qquad + \sum_{j: t_k^\varepsilon < \tau_j^\varepsilon \leq t_k^\varepsilon+\delta^\varepsilon} \bigl(\varepsilon^{\beta_F} k_{\tau_j^\varepsilon} F(\xi_j^\varepsilon) + \varepsilon^{\beta_P} h_{\tau_j^\varepsilon} P(\xi_j^\varepsilon)\bigr)\biggr)(t_{k+1}^\varepsilon - t_k^\varepsilon) \\
&= \sum_{k=0}^{K^\varepsilon-1} j_{t_k^\varepsilon}^\varepsilon (t_{k+1}^\varepsilon - t_k^\varepsilon),
\end{aligned}$$

with

$$j_{t_k^\varepsilon}^\varepsilon = \frac{1}{\delta^\varepsilon}\biggl(\int_{t_k^\varepsilon}^{t_k^\varepsilon+\delta^\varepsilon} \bigl(r_t D(\overline{X}_t^\varepsilon) + \varepsilon^{\beta_Q} l_t Q(u_t^\varepsilon)\bigr)\,dt \\ + \sum_{j:t_k^\varepsilon<\tau_j^\varepsilon\le t_k^\varepsilon+\delta^\varepsilon} \bigl(\varepsilon^{\beta_F} k_{\tau_j^\varepsilon} F(\xi_j^\varepsilon) + \varepsilon^{\beta_P} h_{\tau_j^\varepsilon} P(\xi_j^\varepsilon)\bigr)\biggr).$$

As $\varepsilon$ tends to zero, we expect to have

$$J^\varepsilon(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon) \simeq \int_0^T j_t^\varepsilon\,dt.$$

We are hence led to study $j_t^\varepsilon$ as $\varepsilon \to 0$, which is closely related to the time-average control problem of Brownian motion. To see this, consider the following rescaling of $\overline{X}^\varepsilon$ over the horizon $(t, t+\delta^\varepsilon]$:

$$\widetilde{X}_s^{\varepsilon,t} = \frac{1}{\varepsilon^\beta}\overline{X}_{t+\varepsilon^{2\beta}s}^\varepsilon, \qquad s \in (0, T^\varepsilon],$$

where $T^\varepsilon = \varepsilon^{-2\beta}\delta^\varepsilon$ and $\beta > 0$ is to be determined. We use the superscript $t$ to indicate that the scaled systems correspond to the horizon $(t, t+\delta^\varepsilon]$. Then the dynamics of $\widetilde{X}^{\varepsilon,t}$ is given by (see [50], Proposition V.1.5)

$$\widetilde{X}_s^{\varepsilon,t} = \widetilde{X}_0^{\varepsilon,t} + \int_0^s \widetilde{b}_\nu^{\varepsilon,t}\,d\nu + \int_0^s \sqrt{\widetilde{a}_\nu^{\varepsilon,t}}\,d\widetilde{W}_\nu^{\varepsilon,t} + \int_0^s \widetilde{u}_\nu^{\varepsilon,t}\,d\nu + \sum_{0<\widetilde{\tau}_j^{\varepsilon,t}\le s} \widetilde{\xi}_j^\varepsilon,$$

with

$$\widetilde{b}_s^{\varepsilon,t} = -\varepsilon^\beta b_{t+\varepsilon^{2\beta}s}, \qquad \widetilde{a}_s^{\varepsilon,t} = a_{t+\varepsilon^{2\beta}s}, \qquad \widetilde{W}_s^{\varepsilon,t} = -\frac{1}{\varepsilon^\beta}(W_{t+\varepsilon^{2\beta}s} - W_t),$$

and

$$\widetilde{u}_s^{\varepsilon,t} = \varepsilon^\beta u_{t+\varepsilon^{2\beta}s}^\varepsilon, \qquad \widetilde{\xi}_j^\varepsilon = \frac{1}{\varepsilon^\beta}\xi_j^\varepsilon, \qquad \widetilde{\tau}_j^{\varepsilon,t} = \frac{1}{\varepsilon^{2\beta}}(\tau_j^\varepsilon - t) \vee 0.$$

Note that $(\widetilde{W}_s^{\varepsilon,t})$ is a Brownian motion with respect to $\widetilde{\mathcal{F}}_s^{\varepsilon,t} = \mathcal{F}_{t+\varepsilon^{2\beta}s}$. Abusing notation slightly, we write

$$d\widetilde{X}_s^{\varepsilon,t} = \widetilde{b}_s^{\varepsilon,t}\,ds + \sqrt{\widetilde{a}_s^{\varepsilon,t}}\,d\widetilde{W}_s^{\varepsilon,t} + \widetilde{u}_s^{\varepsilon,t}\,ds \\ + d\biggl(\sum_{0<\widetilde{\tau}_j^{\varepsilon,t}\le s} \widetilde{\xi}_j^\varepsilon\biggr), \qquad s \in (0, T^\varepsilon]. \tag{2.11}$$

Using the homogeneity properties (2.1) of the cost functions, we obtain

$$j_t^\varepsilon = \frac{1}{T^\varepsilon}\biggl(\int_0^{T^\varepsilon} \bigl(\varepsilon^{\beta\zeta_D} r_{t+\varepsilon^{2\beta}s} D(\widetilde{X}_s^{\varepsilon,t}) + \varepsilon^{\beta_Q-\zeta_Q\beta} l_{t+\varepsilon^{2\beta}s} Q(\widetilde{u}_s^{\varepsilon,t})\bigr)\,ds \\ + \sum_{0<\widetilde{\tau}_j^{\varepsilon,t}\le T^\varepsilon} \bigl(\varepsilon^{\beta_F-(2-\zeta_F)\beta} k_{t+\varepsilon^{2\beta}\widetilde{\tau}_j^{\varepsilon,t}} F(\widetilde{\xi}_j^\varepsilon) + \varepsilon^{\beta_P-(2-\zeta_P)\beta} h_{t+\varepsilon^{2\beta}\widetilde{\tau}_j^{\varepsilon,t}} P(\widetilde{\xi}_j^\varepsilon)\bigr)\biggr)$$

$$\simeq \frac{1}{T^\varepsilon}\biggl(\int_0^{T^\varepsilon} \bigl(\varepsilon^{\beta\zeta_D} r_t D\bigl(\widetilde{X}_s^{\varepsilon,t}\bigr) + \varepsilon^{\beta_Q - \zeta_Q\beta} l_t Q\bigl(\widetilde{u}_s^{\varepsilon,t}\bigr)\bigr)\,ds$$
$$+ \sum_{0<\widetilde{\tau}_j^{\varepsilon,t}\le T^\varepsilon} \bigl(\varepsilon^{\beta_F-(2-\zeta_F)\beta} k_t F\bigl(\widetilde{\xi}_j^\varepsilon\bigr) + \varepsilon^{\beta_P-(2-\zeta_P)\beta} h_t P\bigl(\widetilde{\xi}_j^\varepsilon\bigr)\bigr)\biggr).$$

The second approximation can be justified by the continuity of the cost coefficients $r_t$, $l_t$, $k_t$ and $h_t$.

Now, if there exists $\beta > 0$ such that (2.9) is satisfied, then we have

$$j_t^\varepsilon \simeq \varepsilon^{\beta\zeta_D} I_t^\varepsilon,$$

with

$$I_t^\varepsilon = \frac{1}{T^\varepsilon}\biggl(\int_0^{T^\varepsilon} \bigl(r_t D\bigl(\widetilde{X}_s^{\varepsilon,t}\bigr) + l_t Q\bigl(\widetilde{u}_s^{\varepsilon,t}\bigr)\bigr)\,ds + \sum_{0<\widetilde{\tau}_j^{\varepsilon,t}\le T^\varepsilon} \bigl(k_t F\bigl(\widetilde{\xi}_j^\varepsilon\bigr) + h_t P\bigl(\widetilde{\xi}_j^\varepsilon\bigr)\bigr)\biggr). \tag{2.12}$$

By suitably choosing $\delta^\varepsilon$, we have

$$\delta^\varepsilon \to 0, \qquad T^\varepsilon \to \infty.$$

It follows that $\widetilde{b}_s^{\varepsilon,t} \simeq 0$ and $\widetilde{a}_s^{\varepsilon,t} \simeq a_t$ for $s \in (0, T^\varepsilon]$. Therefore, we expect that, as $\varepsilon \to 0$, the process (2.11) behaves approximately as a controlled Brownian motion with diffusion matrix $a_t$, and the functional $I_t^\varepsilon$ satisfies

$$I_t^\varepsilon \gtrsim I^{\mathrm{PW}}(a_t, r_t, l_t, k_t, h_t), \tag{2.13}$$

where $I^{\mathrm{PW}}$ is a deterministic function defined by

$$I^{\mathrm{PW}}(a,r,l,k,h) = \inf_{(u,\tau,\xi)} \limsup_{S\to\infty} \frac{1}{S}\biggl[\int_0^S \bigl(rD(X_s) + lQ(u_s)\bigr)\,ds + \sum_{0\le\tau_j\le S} \bigl(kF(\xi_j) + hP(\xi_j)\bigr)\biggr], \tag{2.14}$$

where the subscript PW stands for pathwise and the infimum is computed over admissible control strategies $u \in \mathcal{A}^R$ and $(\tau, \xi) \in \mathcal{A}^I$, and

$$X_s = \sqrt{a} W_s + \int_0^s u_r\,dr + \sum_{0\le\tau_j\le s} \xi_j. \tag{2.15}$$

We may also expect (see, e.g., [8, 25, 26]) that the optimal value $I^{\mathrm{PW}}$ corresponding to the long-term average pathwise criterion is equal to optimal value of the long-term average expected criterion:

$$I^E(a,r,l,k,h) = \inf_{(u,\tau,\xi)} \limsup_{S\to\infty} \frac{1}{S}\mathbb{E}\biggl[\int_0^S \bigl(rD(X_s) + lQ(u_s)\bigr)\,ds + \sum_{0\le\tau_j\le S}\bigl(kF(\xi_j) + hP(\xi_j)\bigr)\biggr], \tag{2.16}$$

where, once again, $u \in \mathcal{A}^R$ and $(\tau,\xi) \in \mathcal{A}^I$. Therefore, we expect that as $\varepsilon \to 0$:

$$J^\varepsilon \simeq \int_0^T j_t^\varepsilon\,dt \simeq \varepsilon^{\beta\zeta_D}\int_0^T I_t^\varepsilon\,dt \gtrsim \varepsilon^{\beta\zeta_D}\int_0^T I_t^E\,dt,$$

where $I_t^E = I^E(a_t, r_t, l_t, k_t, h_t)$.

REMARK 2.3. The approach of weak convergence is classical for proving inequalities similar to (2.13), in particular in the study of heavy traffic networks (see [40], Section 9, for an overview). The usual weak convergence theorems enable one to show that the perturbed system converges in the Skorokhod topology to the controlled Brownian motion as $\varepsilon$ tends to zero. However, since the time horizon tends to infinity, this does not immediately imply the convergence of time-average cost functionals like $I_t^\varepsilon$.

In [41], the authors consider pathwise average cost problems for controlled queues in the heavy traffic limit, where the control term is absolutely continuous. They use the empirical "functional occupation measure" on the canonical path space and characterize the limit as a controlled Brownian motion. The same method has also been used in [11] in the study of single class queueing networks.

However, this approach cannot be applied directly to singular/impulse controls for which the tightness of the occupation measures is difficult to establish. In fact, the usual Skorokhod topology is not suitable for the impulse control term:

$$\widetilde{Y}_t^\varepsilon := \sum_{0<\widetilde{\tau}_j^\varepsilon\le t} \widetilde{\xi}_j^\varepsilon.$$

Indeed, in the case of singular/impulse control, the component $\{\widetilde{Y}^\varepsilon\}$ is generally not tight under the Skorokhod topology. For example (see [39], page 72), consider the family $(Y_t^\varepsilon)$ where the function $Y_t^\varepsilon$ equals zero for $t < 1$ and jumps upward by an amount $\sqrt{\varepsilon}$ at times $1 + i\varepsilon$, $i = 0, 1, \dots, \varepsilon^{-1/2}$ until it reaches the value unity. The natural limit of $Y^\varepsilon$ is of course $\mathbb{1}_{\{t\ge 1\}}$ but this sequence is not tight in the Skorokhod topology. The nature of this convergence is discussed in [34] and a corresponding topology is provided in [28].

This difficulty could be avoided by introducing a random time change after which the (suitably interpolated) control term becomes uniformly Lipschitz, and hence converges under the Skorokhod topology. This technique is used in [9, 10, 39] to study the convergence of controlled queues with discounted costs.

This seems to be a possible alternative way to extend the approach of [41] to singular/impulse controls. Nevertheless, the analysis would probably be quite involved.

Instead of proving the tightness of control terms $(\widetilde{Y}_t^\varepsilon)$ in weaker topologies, we shall use an alternative characterization of the time-average control problem of Brownian motion. In [36], the authors characterize the time-average control of a Jackson network in the heavy traffic limit as the solution of a linear program. The use of occupation measure on the state space instead of the path space turns out to be sufficient to describe the limiting stochastic control problem. However, the optimization criterion is not pathwise.

In this paper, we use a combination of the techniques in [36] and [41] to obtain pathwise lower bounds.

**3. Extensions of Theorem 2.1 to other types of control.** In this section, we consider the case of combined regular and singular control and those with only one type of control. In particular, we will see that in the presence of singular control, the operator $B$ is different. Formally, we could give similar results for the combination of all three controls or even in the presence of several controls of the same type with different cost functions and scaling properties. To avoid cumbersome notation, we restrict ourselves to the cases meaningful in practice, which are illustrated by explicit examples in Section 4.

3.1. *Combined regular and singular control.* When the fixed cost component is absent, that is $F=0$, impulse control and singular control can be merged. In that case, the natural way to formulate the tracking problem is to consider a strategy $(u^\varepsilon, \gamma^\varepsilon, \varphi^\varepsilon)$ with $u^\varepsilon$ a progressively measurable process as before, $\gamma_t^\varepsilon \in \Delta$ and $(\varphi_t^\varepsilon)$ a possibly discontinuous nondecreasing process such that

$$\overline{X}_t^\varepsilon = -X_t^\circ + \int_0^t u_s^\varepsilon\, ds + \int_0^t \gamma_s^\varepsilon\, d\varphi_s^\varepsilon,$$

and

$$J^\varepsilon(u^\varepsilon, \gamma^\varepsilon, \varphi^\varepsilon) = \int_0^T \big(r_t D(\overline{X}_t^\varepsilon) + \varepsilon^{\beta_Q} l_t Q(u_t^\varepsilon)\big)\, dt + \int_0^T \varepsilon^{\beta_P} h_t P(\gamma_t^\varepsilon)\, d\varphi_t^\varepsilon.$$

Using similar heuristic arguments as those in the previous section, we are led to consider the time-average control of Brownian motion with combined regular and singular control:

$$\begin{aligned} I(a,r,l,h) = \inf_{(u,\gamma,\varphi)} \limsup_{S\to\infty} \frac{1}{S}\mathbb{E}\bigg[\int_0^S \big(rD(X_s) + lQ(u_s)\big)\, ds \\ + \int_0^S hP(\gamma_s)\, d\varphi_s\bigg], \end{aligned} \tag{3.1}$$

where

$$X_s = \sqrt{a}W_s + \int_0^s u_r\, dr + \int_0^s \gamma_r\, d\varphi_r. \tag{3.2}$$

The corresponding linear programming problem is given by

$$I(a,r,l,h) = \inf_{(\mu,\rho)} \int_{\mathbb{R}^d\times\mathbb{R}^d} \big(rD(x)+lQ(u)\big)\mu(dx\times du) + \int_{\mathbb{R}^d\times\Delta\times\mathbb{R}^+_\delta} hP(\gamma)\rho(dx\times d\gamma\times d\delta), \tag{3.3}$$

with $(\mu,\rho)\in\mathcal{P}(\mathbb{R}^d\times\mathbb{R}^d)\times\mathcal{M}(\mathbb{R}^d\times\Delta\times\mathbb{R}^+)$ satisfying the following constraint:

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} A^a f(x,u)\mu(dx\times du) + \int_{\mathbb{R}^d\times\Delta\times\mathbb{R}^+_\delta} Bf(x,\gamma,\delta)\rho(dx\times d\gamma\times d\delta)=0, \qquad \forall f\in C_0^2(\mathbb{R}^d), \tag{3.4}$$

where

$$A^a f(x,u) = \frac{1}{2}\sum_{ij} a_{ij}\partial^2_{ij} f(x) + \langle u,\nabla f(x)\rangle,$$

$$Bf(x,\gamma,\delta) = \begin{cases} \langle\gamma,\nabla f(x)\rangle, & \delta=0,\\ \delta^{-1}\big(f(x+\delta\gamma)-f(x)\big), & \delta>0.\end{cases}$$

We introduce a version of Assumption 2.4 adapted to the present setting.

ASSUMPTION 3.1 (Asymptotic framework). The cost functionals $D$, $Q$ and $P$ are lower semicontinuous functions which satisfy the homogeneity property (2.1) and there exists $\beta>0$ such that

$$\frac{\beta_P}{\zeta_D+2-\zeta_P} = \frac{\beta_Q}{\zeta_D+\zeta_Q} = \beta. \tag{3.5}$$

Moreover, $\inf_{\|x\|=1} D(x)>0$, $\inf_{\|u\|=1} Q(u)>0$ and $\inf_{\gamma\in\Delta} P(\gamma)>0$.

We have the following theorem.

THEOREM 3.1 (Asymptotic lower bound for combined regular and singular control). *Assume that $I(a,r,l,h)$ is measurable, that the parameters $(r_t)$, $(l_t)$ and $(h_t)$ are continuous and positive, that Assumption 2.1 holds true and that Assumption 3.1 is satisfied for some $\beta>0$. Then*

$$\liminf_{\varepsilon\to 0} \frac{1}{\varepsilon^{\beta\zeta_D}} J^\varepsilon\big(u^\varepsilon,\gamma^\varepsilon,\varphi^\varepsilon\big) \ge_P \int_0^T I(a_t,r_t,l_t,h_t)\,dt, \tag{3.6}$$

*for any sequence of admissible tracking strategies $\{(u^\varepsilon,\gamma^\varepsilon,\varphi^\varepsilon)\in\mathcal{A},\varepsilon>0\}$.*

Adapting the proofs of Theorem 2.1 and Theorem 3.1 in an obvious way, we easily obtain the following bounds when only one control is present.

3.2. *Impulse control.* Consider a family of strategies $(\tau^\varepsilon, \xi^\varepsilon)_{\varepsilon>0} \subset \mathcal{A}$, together with the associated family of aggregate cost functionals:

$$J^\varepsilon(\tau^\varepsilon, \xi^\varepsilon) = \int_0^T r_t D(\overline{X}_t^\varepsilon)\,dt + \sum_{0<\tau_j^\varepsilon \le T} \bigl(\varepsilon^{\beta_F} k_{\tau_j^\varepsilon} F(\xi_j^\varepsilon) + \varepsilon^{\beta_P} h_{\tau_j^\varepsilon} P(\xi_j^\varepsilon)\bigr),$$

and deviation processes

$$\overline{X}_t^\varepsilon = -X_t^\circ + \sum_{0<\tau_j^\varepsilon \le t} \xi_j^\varepsilon.$$

Once again, we introduce the version of Assumption 2.4 adapted to the present context.

ASSUMPTION 3.2 (Asymptotic framework). The cost functionals $D$, $F$ and $P$ are lower semicontinuous functions which satisfy the homogeneity property (2.1) and there exists $\beta > 0$ such that

$$\frac{\beta_F}{\zeta_D + 2 - \zeta_F} = \frac{\beta_P}{\zeta_D + 2 - \zeta_P} = \beta. \tag{3.7}$$

Moreover, $\inf_{\|x\|=1} D(x) > 0$, and $\inf_{\|\xi\|=1} F(\xi) > 0$.

The following theorem holds true.

THEOREM 3.2 (Asymptotic lower bound for impulse control). *Let $I = I(a, r, k, h)$ be given by*

$$I(a, r, k, h) = \inf_{(\mu,\rho)} \int_{\mathbb{R}^d} r D(x)\mu(dx) + \int_{\mathbb{R}^d\times\mathbb{R}^d\setminus\{0\}} \bigl(kF(\xi) + hP(\xi)\bigr)\rho(dx \times d\xi),$$

*with $(\mu, \rho) \in \mathcal{P}(\mathbb{R}^d) \times \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d \setminus \{0\})$ satisfying the following constraint:*

$$\int_{\mathbb{R}^d} A^a f(x)\mu(dx) + \int_{\mathbb{R}^d\times\mathbb{R}^d\setminus\{0\}} Bf(x, \xi)\rho(dx \times d\xi) = 0, \qquad \forall f \in C_0^2(\mathbb{R}^d),$$

*where*

$$A^a f(x) = \frac{1}{2}\sum_{i,j} a_{ij}\partial_{ij}^2 f(x), \qquad Bf(x, \xi) = f(x + \xi) - f(x).$$

*Assume that $I(a, r, k, h)$ is measurable, that the parameters $(r_t)$, $(k_t)$ and $(h_t)$ are continuous and positive, that Assumption 2.1 holds true, and that Assumption 2.4 is satisfied for some $\beta > 0$. Then*

$$\liminf_{\varepsilon\to 0} \frac{1}{\varepsilon^{\beta\zeta_D}} J^\varepsilon(\tau^\varepsilon, \xi^\varepsilon) \ge_p \int_0^T I(a_t, r_t, k_t, h_t)\,dt.$$

See Example 4.5 for a closed form solution of $I$.

3.3. *Singular control*. Consider a family of strategies $(\gamma^\varepsilon, \varphi^\varepsilon)_{\varepsilon>0} \subset \mathcal{A}$ together with the associated family of aggregate cost functionals:

$$J^\varepsilon(\gamma^\varepsilon, \varphi^\varepsilon) = \int_0^T r_t D(\overline{X}_t^\varepsilon)\,dt + \int_0^T \varepsilon^{\beta_P} h_t P(\gamma_t^\varepsilon)\,d\varphi_t^\varepsilon,$$

and deviation processes

$$\overline{X}_t^\varepsilon = -X_t^\circ + \int_0^t \gamma_s^\varepsilon\,d\varphi_s^\varepsilon.$$

Assumption 2.4 now takes the following form.

ASSUMPTION 3.3 (Asymptotic framework). The cost functionals $D$ and $P$ are lower semicontinuous functions which satisfy the homogeneity property (2.1) with $\inf_{\|x\|=1} D(x) > 0$ and $\inf_{\gamma\in\Delta} P(\gamma) > 0$. We define $\beta > 0$ by

$$\frac{\beta_P}{\zeta_D + 2 - \zeta_P} = \beta. \tag{3.8}$$

The following theorem holds true.

THEOREM 3.3 (Asymptotic lower bound for singular control). *Let $I = I(a, r, h)$ be given by*

$$I(a,r,h) = \inf_{(\mu,\rho)} \int_{\mathbb{R}^d} rD(x)\mu(dx) + \int_{\mathbb{R}^d\times\Delta\times\mathbb{R}_\delta^+} hP(\gamma)\rho(dx\times d\gamma\times d\delta),$$

*with $(\mu, \rho) \in \mathcal{P}(\mathbb{R}^d) \times \mathcal{M}(\mathbb{R}^d \times \Delta \times \mathbb{R}^+)$ satisfying the following constraint*:

$$\int_{\mathbb{R}^d} A^a f(x)\mu(dx) + \int_{\mathbb{R}^d\times\Delta\times\mathbb{R}_\delta^+} Bf(x,\gamma,\delta)\rho(dx\times d\gamma\times d\delta) = 0, \qquad \forall f \in C_0^2(\mathbb{R}^d),$$

*where*

$$A^a f(x) = \frac{1}{2}\sum_{ij} a_{ij}\partial_{ij}^2 f(x),$$

$$Bf(x,\gamma,\delta) = \begin{cases} \langle \gamma, \nabla f(x)\rangle, & \delta = 0,\\ \delta^{-1}\big(f(x+\delta\gamma) - f(x)\big), & \delta > 0.\end{cases}$$

*Assume that $I(a, r, h)$ is measurable, that the parameters $(r_t)$ and $(h_t)$ are continuous and positive, that Assumption 2.1 holds true and that Assumption 2.4 is satisfied for some $\beta > 0$. Then*

$$\liminf_{\varepsilon\to 0} \frac{1}{\varepsilon^{\beta\zeta_D}} J^\varepsilon(\gamma^\varepsilon, \varphi^\varepsilon) \geq_p \int_0^T I(a_t, r_t, h_t)\,dt.$$

See Example 4.4 for a closed form solution of $I$.

3.4. *Regular control.* Consider a family of strategies $(u^\varepsilon)_{\varepsilon>0} \subset \mathcal{A}$ together with the associated family of aggregate cost functionals

$$J^\varepsilon(u^\varepsilon) = \int_0^T \big(r_t D(\overline{X}_t^\varepsilon) + \varepsilon^{\beta_Q} l_t Q(u_t^\varepsilon)\big)\,dt,$$

and deviation processes

$$\overline{X}_t^\varepsilon = -X_t^\circ + \int_0^t u_s^\varepsilon\,ds.$$

Assumption 2.4 takes the following form.

ASSUMPTION 3.4 (Asymptotic framework). The cost functionals $D$ and $Q$ are lower semicontinuous functions which satisfy the homogeneity property (2.1) with $\inf_{\|x\|=1} D(x) > 0$ and $\inf_{\|u\|=1} Q(u) > 0$. We define $\beta > 0$ by

$$\frac{\beta_Q}{\zeta_D + \zeta_Q} = \beta. \tag{3.9}$$

The following theorem holds true.

THEOREM 3.4 (Asymptotic lower bound for regular control). *Let $I = I(a, r, l)$ be given by*

$$I(a, r, l) = \inf_\mu \int_{\mathbb{R}^d \times \mathbb{R}^d} \big(r D(x) + l Q(u)\big)\mu(dx, du),$$

*with $\mu \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ satisfying the following constraint*:

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} A^a f(x, u)\mu(dx, du) = 0, \qquad \forall f \in C_0^2(\mathbb{R}^d),$$

*where*

$$A^a f(x, u) = \frac{1}{2}\sum_{i,j} a_{ij}\partial_{ij}^2 f(x) + \langle u, \nabla f(x)\rangle.$$

*Assume that $I(a, r, l)$ is measurable, that the parameters $(r_t)$ and $(l_t)$ are continuous and positive on $[0, T]$, that Assumption 2.1 holds true and that Assumption 2.4 is satisfied for some $\beta > 0$. Then*

$$\liminf_{\varepsilon\to 0} \frac{1}{\varepsilon^{\beta\zeta_D}} J^\varepsilon(u^\varepsilon) \ge_p \int_0^T I(a_t, r_t, l_t)\,dt.$$

See Example 4.3 for a closed form solution of $I$.

REMARK 3.1 (Upper bound). It is natural to wonder whether the lower bounds in our theorems are tight and if it is the case, what are the strategies that attain them. In the companion paper [13], we show that for the examples provided in Section 4, there are closed form strategies attaining asymptotically the lower bounds. For instance, in the case of combined regular and impulse control, it means that there exist $(u^{\varepsilon,*}, \tau^{\varepsilon,*}, \xi^{\varepsilon,*}) \in \mathcal{A}$ such that

$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon^{\beta\zeta_D}} J^{\varepsilon}\big(u^{\varepsilon,*}, \tau^{\varepsilon,*}, \xi^{\varepsilon,*}\big) \to_p \int_0^T I(a_t, r_t, l_t, k_t, h_t)\, dt.$$

These optimal strategies are essentially time-varying versions of the optimal strategies for the time-average control of Brownian motion.

**4. Interpretation of lower bounds and examples.** Our goal in this section is to provide a probabilistic interpretation of the lower bounds in Theorems 2.1, 3.1, 3.2, 3.3 and 3.4, which are expressed in terms of linear programming. In particular, we want to connect them with the time-average control problem of Brownian motion. To our knowledge, there is no general result available for the equivalence between time-average control problem and linear programming. Partial results exist in [8, 22, 35, 36, 38] but do not cover all the cases we need. Here, we provide a brief self-contained study enabling us to also treat the cases of singular/impulse controls and their combinations with regular control. We first introduce controlled martingale problems and show that they can be seen as a relaxed version of the controlled Brownian motion (2.4). Then we formulate the time-average control problem in this martingale framework. We finally show that this problem has an equivalent description in terms of infinite dimensional linear program. While essential ingredients and arguments for obtaining these results are borrowed from [35] and [37], we provide sharp conditions which guarantee the equivalence of these two formulations.

4.1. *Martingale problem associated to controlled Brownian motion.* In [2, 44, 49, 53], the authors obtain a HJB equation in the first-order expansion for the value function of the utility maximization problem under transaction costs, which essentially provides a lower bound for their control problems. They mention a connection between the HJB equation and the time-average control problem of Brownian motion; see also [23]. Here, we wish to rigorously establish an equivalence between the linear programs in our lower bounds and the time-average control of Brownian motion. This leads us to introduce a relaxed version for the controlled Brownian motion. We shall see that the optimal costs for all these formulations coincide in the examples provided in the next section.

We place ourselves in the setting of [37], from which we borrow and rephrase several elements, and assume that the state space $E$ and control spaces $U$ and $V$ are complete, separable, metric spaces. For a complete, separable metric space $S$, we define $\mathcal{L}(S)$ the set of nonnegative Borel measures $\Gamma$ on $S \times [0, \infty)$ such that

$\Gamma(S\times[0,t])<\infty$ for all $t\geq 0$. For any such measure and for any $t>0$, $\Gamma_t$ denotes the restriction of $\Gamma$ to $S\times[0,t]$. Consider an operator $A:\mathcal{D}\subset C_b(E)\to C(E\times U)$ and an operator $B:\mathcal{D}\subset C_b(E)\to C(E\times V)$.

DEFINITION 4.1 (Controlled martingale problem). A triplet $(X,\Lambda,\Gamma)$ with $(X,\Lambda)$ an $E\times\mathcal{P}(U)$-valued process and $\Gamma$ an $\mathcal{L}(E\times V)$-valued random variable is a solution of the controlled martingale problem for $(A,B)$ with initial distribution $\nu_0\in\mathcal{P}(E)$ if there exists a filtration $(\mathcal{F}_t)$ such that the process $(X_t,\Lambda_t,\Gamma_t)_{t\geq 0}$ is $(\mathcal{F}_t)$-progressive, $X_0$ has distribution $\nu_0$ and for every $f\in\mathcal{D}$,

$$\begin{aligned}f(X_t)&-\int_0^t\int_U Af(X_s,u)\Lambda_s(du)\,ds\\&-\int_{E\times V\times[0,t]}Bf(x,v)\Gamma(dx\times dv\times ds)\end{aligned}\tag{4.1}$$

is an $\mathcal{F}_t$-martingale.

We now consider two specific cases for the operators $A$ and $B$, which will be relevant in order to express our lower bounds. Furthermore, we explain why these specific choices of $A$ and $B$ are connected to combined regular and singular/impulse control of Brownian motion.

EXAMPLE 4.1 (Combined regular and impulse control of Brownian motion). Let $\mathcal{D}=C_0^2(\mathbb{R}^d)\oplus\mathbb{R}$ and define $A:\mathcal{D}\to C(E\times U)$ and $B:\mathcal{D}\to C(E\times V)$ by

$$Af(x,u)=\frac{1}{2}\sum_{i,j}a_{ij}\partial^2_{ij}f(x)+\langle u,\nabla f(x)\rangle,\tag{4.2}$$

$$Bf(x,\xi)=f(x+\xi)-f(x).\tag{4.3}$$

Here, $E=\mathbb{R}^d$, $U=\mathbb{R}^d$ and $V=\mathbb{R}^d\setminus\{0\}$. We call any solution of this martingale problem the combined regular and impulse control of Brownian motion.

Indeed, consider the following process:

$$X_t=X_0+\sqrt{a}W_t+\int_0^t u_s\,ds+\sum_{0<\tau_j\leq t}\xi_j,$$

where $W$ is a standard Brownian motion on a filtered probability space $(\Omega,\mathcal{F},(\mathcal{F})_t,\mathbb{P})$, $(u_t)$ is a $(\mathcal{F}_t)$-progressively measurable process, $(\tau_j)$ a sequence of $(\mathcal{F}_t)$-stopping times and $(\xi_j)$ a sequence of random variables with values in $\mathbb{R}^d$ such that $\xi_j$ is $\mathcal{F}_{\tau_j}$-measurable for $j=1,2,\ldots$. Define

$$N_t=\sum_j\mathbb{1}_{\{\tau_j\leq t\}},\qquad \xi_t=\xi_j,\qquad t\in(\tau_{j-1},\tau_j].$$

Then for any $f \in \mathcal{D}$, by Itô's formula,

$$f(X_t) - \int_0^t Af(X_s, u_s)\,ds - \int_0^t Bf(X_{s-}, \xi_{s-})\,dN_s = f(X_0) - \int_0^t \nabla f(X_s)^\top \sqrt{a}\,dW_s,$$

which is a martingale. Let

$$\Lambda_t = \delta_{u_t}(du),$$
$$\Gamma\big(H \times [0,t]\big) = \int_0^t \mathbb{1}_H(X_{s-}, \xi_{s-})\,dN_s, \qquad H \in \mathcal{B}\big(\mathbb{R}^d \times \mathbb{R}^d \setminus \{0\}\big).$$

Then $(X, \Lambda, \Gamma)$ solves the martingale problem $(A, B)$ with initial distribution $\mathcal{L}(X_0)$.

EXAMPLE 4.2 (Combined regular and singular control of Brownian motion). Take $\mathcal{D} = C_0^2(\mathbb{R}^d) \oplus \mathbb{R}$ and define

$$Af(x,u) = \frac{1}{2}\sum_{i,j} a_{ij}\partial_{ij}^2 f(x) + \langle u, \nabla f(x)\rangle, \tag{4.4}$$

$$Bf(x,\gamma,\delta) = \begin{cases} \langle \gamma, \nabla f(x)\rangle, & \delta = 0, \\ \delta^{-1}\big(f(x+\delta\gamma) - f(x)\big), & \delta > 0. \end{cases} \tag{4.5}$$

Here, $E = \mathbb{R}^d$, $U = \mathbb{R}^d$ and $V = \Delta \times \mathbb{R}_\delta^+$. Any solution of this martingale problem is called combined regular and singular control of Brownian motion.

Indeed, let $X$ be given by

$$X_t = X_0 + \sqrt{a}W_t + \int_0^t u_s\,ds + \int_0^t \gamma_s\,d\varphi_s,$$

with $u$ a progressively measurable process, $\gamma_s \in \Delta$ and $\varphi_s$ nondecreasing. By Itô's formula, we have

$$f(X_t) - \int_0^t Af(X_s, u_s)\,ds - \int_0^t Bf(X_{s-}, \gamma_s, \delta\varphi_s)\,d\varphi_s = f(X_0) - \int_0^t \nabla f(X_s)^\top \sqrt{a}\,dW_s,$$

which is a martingale for any $f \in \mathcal{D}$. Let

$$\Lambda_t = \delta_{u_t}(du),$$
$$\Gamma\big(H \times [0,t]\big) = \int_0^t \mathbb{1}_H(X_{s-}, \gamma_s, \delta\varphi_s)\,d\varphi_s, \qquad H \in \mathcal{B}\big(\mathbb{R}^d \times \Delta \times \mathbb{R}_\delta^+\big).$$

Then $(X, \Lambda, \Gamma)$ solves the martingale problem $(A, B)$ with initial distribution $\mathcal{L}(X_0)$.

4.2. *Time-average control of Brownian motion*. Now we formulate a relaxed version of the time-average control problem of Brownian motion in terms of a controlled martingale problem. This generalizes [22, 35] to combined regular and singular/impulse control of martingale problems; see also [38]. Recall that $A$ and $B$ are two operators where $A : \mathcal{D} \subset C_b(E) \to C(E \times U)$ and $B : \mathcal{D} \subset C_b(E) \to C(E \times V)$. Consider two Borel measurable functions $C_A : E \times U \to \mathbb{R}_+$ and $C_B : E \times V \to \mathbb{R}_+$.

PROBLEM 4.1 (Martingale formulation of time-average control problem). The time-average control problem under the martingale formulation is given by

$$\begin{aligned} I^M = \inf_{(X,\Lambda,\Gamma)} \limsup_{t\to\infty} \frac{1}{t}\mathbb{E}\biggl[\int_0^t \int_U C_A(X_s,u)\Lambda_s(du)\,ds \\ + \int_{E\times V\times[0,t]} C_B(x,v)\Gamma(dx\times dv\times ds)\biggr], \end{aligned} \tag{4.6}$$

where the inf is taken over all solutions of the martingale problem $(A, B)$ with any initial distribution $\nu_0 \in \mathcal{P}(E)$.

Now, let $(X, \Lambda, \Gamma)$ be any solution of the martingale problem with operators $A$ and $B$. Define $(\mu_t, \rho_t) \in \mathcal{P}(E \times U) \times \mathcal{M}(E \times V)$ as

$$\mu_t(H_1) = \frac{1}{t}\mathbb{E}\biggl[\int_0^t \int_U \mathbb{1}_{H_1}(X_s,u)\Lambda_s(du)\,ds\biggr], \tag{4.7}$$

$$\rho_t(H_2) = \frac{1}{t}\mathbb{E}\bigl[\Gamma\bigl(H_2 \times (0,t]\bigr)\bigr], \tag{4.8}$$

for $H_1 \in \mathcal{B}(E \times U)$ and $H_2 \in \mathcal{B}(E \times V)$. Then the average cost up to time $t$ in (4.6) can be expressed as

$$\int_{E\times U} C_A(x,u)\mu_t(dx\times du) + \int_{E\times V} C_B(x,v)\rho_t(dx\times dv).$$

On the other hand, for $f \in \mathcal{D}$, (4.1) defines a martingale. Taking the expectation, we obtain

$$\begin{aligned} \frac{1}{t}\mathbb{E}\biggl[\int_0^t \int_U Af(X_s,u)\Lambda_s(du)\,ds + \int_{E\times V\times(0,t]} Bf(x,v)\Gamma(dx\times dv\times ds)\biggr] \\ = \frac{1}{t}\mathbb{E}\bigl[f(X_t) - f(X_0)\bigr]. \end{aligned}$$

When $t$ tends to infinity, the right-hand side of the above equality tends to zero, which means that

$$\int_{E\times U} Af(x,u)\mu_t(dx\times du) + \int_{E\times V} Bf(x,v)\rho_t(dx\times dv) \to 0 \qquad \text{as } t \to \infty.$$

This leads us to introduce the following linear programming problem.

PROBLEM 4.2 [Linear programming (LP) formulation of time-average control]. The time-average control problem under the LP formulation is given by

$$I = \inf_{(\mu,\rho)} c(\mu,\rho), \tag{4.9}$$

with

$$\begin{aligned} &c : \mathcal{P}(E \times U) \times \mathcal{M}(E \times V) \to \mathbb{R}_+, \\ &(\mu,\rho) \mapsto \int_{E\times U} C_A(x,u)\mu(dx \times du) + \int_{E\times V} C_B(x,v)\rho(dx \times dv), \end{aligned} \tag{4.10}$$

where the inf is computed over all $\mu \in \mathcal{P}(E \times U)$ and $\rho \in \mathcal{M}(E \times V)$ satisfying the constraint

$$\begin{aligned} &\int_{E\times U} Af(x,u)\mu(dx \times du) \\ &\qquad + \int_{E\times V} Bf(x,v)\rho(dx \times dv) = 0, \qquad \forall f \in \mathcal{D}. \end{aligned} \tag{4.11}$$

We now present the theorem which connects linear programming and time-average control of Brownian motion. To this end, we first recall a condition on the operators $A$ and $B$, used in [37]. We refer the reader to [37] for an explanation of different components of this assumption.

ASSUMPTION 4.1 (Condition 1.2 in [37]). (i) $1 \in \mathcal{D}$ and $A1 = 0$, $B1 = 0$.

(ii) There exist $\psi_A \in C(E \times U)$, $\psi_B \in C(E \times V)$, and constants $a_f, b_f$ depending on $f \in \mathcal{D}$ such that

$$\begin{aligned} &|Af(x,u)| \le a_f \psi_A(x,u), \\ &|Bf(x,v)| \le b_f \psi_B(x,v), \qquad \forall x \in E, u \in U, v \in V. \end{aligned} \tag{4.12}$$

(iii) Defining $(A_0, B_0) = \{(f, \psi_A^{-1}Af, \psi_B^{-1}Bf) : f \in \mathcal{D}\}$, $(A_0, B_0)$ is separable, in the sense that there exists a countable collection $\{g_k\} \subset \mathcal{D}$ such that $(A_0, B_0)$ is contained in the bounded pointwise closure of the linear span of $\{(g_k, A_0 g_k, B_0 g_k) = (g_k, \psi_A^{-1}Ag_k, \psi_B^{-1}Bg_k)\}$.

(iv) For each $u \in U$, the operators $A_u$ and $B_u$ defined by $A_u f(x) = Af(x,u)$ and $B_u f(x) = Bf(x,u)$ are pre-generators (see [37], page 4).

(v) $\mathcal{D}$ is closed under multiplication and separates points.

THEOREM 4.1 (Equivalence between $I^M$ and $I$). *Assume that*:

1. (*Condition on the operators A and B*) *The operators A and B satisfy Assumption* 4.1.

2. (*Condition on the cost function $C_A$*) *The cost function $C_A$ is nonnegative and inf-compact*, *that is*, $\{(x,u) \in E \times U | C_A(x,u) \le c\}$ *is a compact set for each* $c \in \mathbb{R}_+$. *In particular*, $C_A$ *is lower semicontinuous*.

3. (*Condition on cost function* $C_B$) *The cost function* $C_B$ *is nonnegative and lower semi-continuous. Moreover,* $C_B$ *satisfies*

$$\inf_{(x,v)\in E\times V} C_B(x,v) > 0. \tag{4.13}$$

4. (*Relation between operators and cost functions*) *There exist constants* $\theta$ *and* $0<\beta<1$ *such that*

$$\psi_A(x,u)^{1/\beta} \le \theta\big(1+C_A(x,u)\big), \qquad \psi_B(x,v)^{1/\beta} \le \theta C_B(x,v), \tag{4.14}$$

*for* $\psi_A$ *and* $\psi_B$ *given by* (4.12).

*Then the two formulations above of the time-average control problem are equivalent in the sense that*

$$I^M = I.$$

PROOF. The idea of the proof is the same as in [35], with a key ingredient provided by [37], Theorem 1.7.

We first show that $I \le I^M$. Given any solution $(X,\Lambda,\Gamma)$ of the martingale problem, we define the occupation measures as (4.7) and (4.8). Without loss of generality, we can assume that $\limsup_{t\to\infty} c(\mu_t,\rho_t) < \infty$ where $c$ is defined in (4.10) (otherwise we would have $I^M = \infty \ge I$). We will show that $I \le \limsup_{t\to\infty} c(\mu_t,\rho_t)$.

We consider the one-point compactification $\overline{E\times V} = E\times V \cup \{\infty = (x_\infty, v_\infty)\}$ and extend $C_B$ to $\overline{E\times V}$ by

$$C_B(x_\infty, v_\infty) = \liminf_{(x,v)\to(x_\infty,v_\infty)} C_B(x,v) > 0,$$

where the last inequality is guaranteed by (4.13). Since $C_B$ is lower semicontinuous, the level sets $\{(x,v)\in \overline{E\times V} | C_B(x,v)\le c\}$ are compact. By Lemma C.1, we deduce that $c$ is a tightness function on $\mathcal{P}(E\times U)\times\mathcal{M}(\overline{E\times V})$. So the family of occupation measures $(\mu_t,\rho_t)_{t\ge 0}$ is tight if $\rho_t$ is viewed as a measure on $\overline{E\times V}$. It follows that the family of occupation measures indexed by $t$ is relatively compact.

Let $(\mu,\overline{\rho}) \in \mathcal{P}(E\times U)\times\mathcal{M}(\overline{E\times V})$ be any limit point of $(\mu_t,\rho_t)$ with canonical decomposition $\bar\rho = \rho + \theta_{\bar\rho}\delta_\infty$. We claim that $(\mu,\rho)$ satisfies the linear constraint (4.11). Indeed, by (4.12) and (4.14), we have

$$|Af|^{1/\beta} \le \theta_f(1+C_A), \qquad |Bf|^{1/\beta} \le \theta_f C_B,$$

where $\theta_f$ is a nonnegative real number depending on $f$. Then $\sup_t c(\mu_t,\rho_t) < \infty$ implies that $Af$ and $Bf$ are uniformly integrable with respect to $\mu$ and $\bar\rho$. We therefore have

$$\int_{E\times U} Af\, d\mu + \int_{\overline{E\times V}} \mathbb{1}_{E\times V} Bf\, d\bar\rho = \lim_{t\to\infty} \int_{E\times U} Af\, d\mu_t + \int_{\overline{E\times V}} \mathbb{1}_{E\times V} Bf\, d\rho_t.$$

The right-hand side being equal to zero, we obtain

$$\int_{E\times U} Af\,d\mu + \int_{E\times V} Bf\,d\rho = 0.$$

Since $C_A$ and $C_B$ are lower semi-continuous, it follows that (see [15], Theorem A.3.12)

$$I \le c(\mu,\rho) \le c(\mu,\bar{\rho}) \le \liminf_{t\to\infty} c(\mu_t,\rho_t) \le \limsup_{t\to\infty} c(\mu_t,\rho_t).$$

As the choice of the solution of the controlled martingale problem $(X,\Lambda,\Gamma)$ is arbitrary, we conclude that $I \le I^M$.

We now show that $I^M \le I$. Given any $(\mu,\rho)$ satisfying the linear constraint (4.11) such that $c(\mu,\rho)<\infty$, Theorem 1.7 in [37] together with the condition on the operators $A$ and $B$ provides the existence of a stationary solution $(X,\Lambda,\Gamma)$ for the martingale problem $(A,B)$ with marginals $(\mu,\rho)$, hence $I^M \le I$. $\square$

REMARK 4.1. Compared with the results in [8, 35], no near-monotone condition is necessary for the singular component. This is due to the fact that in the linear constraint (4.11), $\rho$ belongs to $\mathcal{M}(E\times V)$ instead of $\mathcal{P}(E\times V)$.

We now give natural examples for which $I = I^M$.

COROLLARY 4.1 (Time-average control of Brownian motion with quadratic costs). *Assume that $C_A$ is given by*

$$C_A(x,u) = r\langle x, \Sigma^D x\rangle + l\langle u, \Sigma^Q u\rangle, \qquad \forall (x,u)\in\mathbb{R}^d\times\mathbb{R}^d, \tag{4.15}$$

*where $\Sigma^D, \Sigma^Q \in \mathcal{S}_d^+$. Then for both following cases, we have $I^M = I$:*

1. *The operators $A$ and $B$ are given by* (4.2)–(4.3) *and the cost function $C_B$ is given by*

$$C_B(x,\xi) = k\sum_{i=1}^d F_i \mathbb{1}_{\{\xi^i\neq 0\}} + h\langle P, |\xi|\rangle, \qquad (x,\xi)\in\mathbb{R}^d\times\mathbb{R}^d\setminus\{0\}, \tag{4.16}$$

*with* $\min_i F_i > 0$.

2. *The operators $A$ and $B$ are given by* (4.4)–(4.5) *and the cost function $C_B$ is given by*

$$C_A(x,\gamma,\delta) = h\langle P, |\gamma|\rangle, \qquad (x,\gamma,\delta)\in\mathbb{R}^d\times\Delta\times\mathbb{R}_\delta^+, \tag{4.17}$$

*with* $\min_i P_i > 0$.

PROOF. We consider the impulse case. First, we show that $A$ and $B$ satisfy [37], Condition 1.2. (i) It is clear that $\mathbb{1}_{\mathbb{R}^d}\in\mathcal{D}$, and $A\mathbb{1}_{\mathbb{R}^d}=0$, $B\mathbb{1}_{\mathbb{R}^d}=0$. (ii) Define

$$\psi_A(x,u) = 1\vee\sum_{i=1}^d |u_i|, \qquad \psi_B(x,\xi) = 1,$$

then (4.12) is satisfied. (iii) Since $C_0^2(\mathbb{R}^d)$ equipped with $\|\cdot\|_{C_0^2}$ is separable, the third condition is satisfied. (iv) $A_u f(x) = Af(x,u)$ and $B_\xi f(x) = Bf(x,\xi)$ satisfy the positive maximum principle, so they are dissipative. It is obvious that they verify [37], (1.10). Hence, they are pre-generators. (v) Obvious. Second, since $C_A$ and $C_B$ are l.s.c. and $\min_i F_i > 0$, the conditions on $C_A$ and $C_B$ are verified. Third, (4.14) holds with $\beta = 1/2$.

The proof for the singular case is similar. Note that $\min_i P_i > 0$ is necessary for (4.13) and the second bound in (4.14) to hold. □

4.3. *Explicit examples in dimension one*. We collect here a comprehensive list of examples in dimension one for which explicit solutions are available. Most of these results exist already in the literature under the classical SDE formulation (see, e.g., [2, 14, 17, 22, 24–26, 44, 53]), but Examples 4.6 and 4.7 are presented here for the first time. The basic idea is to solve explicitly the HJB equation corresponding to the time-average control problem and apply a verification theorem. Similar methods apply under the linear programming framework. However, for completeness we provide in Section 8 detailed proofs tailored to the formulation in terms of linear programming. In fact, we prove only the case of combined regular and impulse control, that is, Example 4.6. The proofs for the other examples are similar, and hence omitted.

EXAMPLE 4.3 (Regular control of Brownian motion). Let $r, l > 0$ and consider the following linear programming problem:

$$I(a,r,l) = \inf_\mu \int_{\mathbb{R}\times\mathbb{R}} (rx^2 + lu^2)\mu(dx,du), \tag{4.18}$$

where $\mu \in \mathcal{P}(\mathbb{R}\times\mathbb{R})$ satisfies

$$\int_{\mathbb{R}\times\mathbb{R}} \left(\frac{1}{2}af''(x) + uf'(x)\right)\mu(dx,du) = 0, \qquad \forall f \in C_0^2(\mathbb{R}). \tag{4.19}$$

By Corollary 4.1, this is equivalent to the time-average control of Brownian motion with quadratic costs in the sense of Problem 4.1.

Let us explain heuristically how to obtain the optimal solution (a rigorous verification argument for the linear programming formulation is provided in Section 8). Roughly speaking, Problem 4.1 describes the following dynamics:

$$dX_t^u = \sqrt{a}\,dW_t + u_t\,dt.$$

The optimization objective is

$$\inf_{(u_t)\in\mathcal{A}} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T \left(r(X_t^u)^2 + lu_t^2\right)dt\right],$$

where the set $\mathcal{A}$ of admissible controls contains all progressively measurable processes $(u_t)$ such that

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\big[(X_T^u)^2\big] < \infty.$$

Consider the associated HJB equation for the couple $(w, I^V)$[2]

$$\inf_{u\in\mathbb{R}} \frac{1}{2}aw''(x) + uw'(x) + lu^2 + rx^2 - I^V = 0.$$

It is easy to to see (see also [44], equation (3.18)) that it is solved explicitly by the couple:

$$w(x) = \sqrt{rl}x^2, \qquad I^V = \sqrt{a^2rl}.$$

Now, let $(u_t)$ be an admissible control and apply Itô's formula to $w(X_T)$:

$$w(X_T) = w(X_0) + \int_0^T \left(\frac{1}{2}aw''(X_t) + u_t w'(X_t)\right) dt + \int_0^T w'(X_t)\, dW_t.$$

It follows that

$$\begin{aligned} \int_0^T (rX_t^2 + lu_t^2)\, dt &\ge \int_0^T \left(I^V - \frac{1}{2}aw''(X_t) - u_t w'(X_t)\right) dt \\ &= TI^V + w(X_0) - w(X_T) + \int_0^T w'(X_t)\, dW_t. \end{aligned} \tag{4.20}$$

Taking expectation, dividing by $T$ on both sides, and using the admissibility conditions, we obtain

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T (rX_t^2 + lu_t^2)\, dt\right] \ge I^V.$$

To show that $I^V$ is indeed the optimal cost, it is enough to show that equality holds in (4.20) for the optimal feedback control given by

$$u^*(x) = -\frac{w'(x)}{2l} = -\theta x, \qquad \theta = \sqrt{\frac{r}{l}}.$$

Therefore, the optimally controlled process is an Ornstein–Uhlenbeck process

$$dX^* = \sqrt{a}\, dW_t - \theta X_t^*\, dt.$$

Naturally, the stationary distribution of $(X^*, u^*(X_t^*))$ is the solution $\mu^*$ of the linear programming problem. We have the following result.

[2]See [7] for an interpretation of this equation and related uniqueness results.

PROPOSITION 4.1. *The solution of* (4.18)–(4.19) *is given by*

$$I(a, r, l) = \sqrt{a^2 r l},$$

*and the optimum is attained by*

$$\mu^*(dx, du) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} dx \otimes \delta_{\{-\theta x\}}(du),$$

*where*

$$\sigma^2 = \frac{a}{2\theta}, \qquad \theta = \sqrt{\frac{r}{l}}.$$

EXAMPLE 4.4 (Singular control of Brownian motion). For any parameters $r, h > 0$, consider the following linear programming problem:

$$I(a, r, h) = \inf_{(\mu,\rho)} \int_{\mathbb{R}} r x^2 \mu(dx) + \int_{\mathbb{R}\times\{\pm 1\}\times\mathbb{R}_\delta^+} h|\gamma|\rho(dx, d\gamma, d\delta), \tag{4.21}$$

where $\mu \in \mathcal{P}(\mathbb{R})$ and $\rho \in \mathcal{M}(\mathbb{R} \times \{\pm 1\} \times \mathbb{R}_\delta^+)$ satisfy

$$\begin{aligned} &\int_{\mathbb{R}} \frac{1}{2} a f''(x)\mu(dx) \\ &\quad + \int_{\mathbb{R}\times\{\pm 1\}\times\mathbb{R}_\delta^+} \gamma f'(x)\rho(dx, d\gamma, d\delta) = 0, \qquad \forall f \in C_0^2(\mathbb{R}). \end{aligned} \tag{4.22}$$

By Corollary 4.1, this is equivalent to the time-average control of Brownian motion with quadratic deviation penalty and proportional costs in the sense of Problem 4.1.

The dynamics of the solution of the controlled martingale problem is heuristically

$$dX_t = \sqrt{a}\, dW_t + \gamma_t\, d\varphi_t,$$

where $\gamma_t \in \{\pm 1\}$ and $\varphi_t$ is a non-decreasing process. The optimization objective is

$$\inf_{(\gamma_t,\varphi_t)\in\mathcal{A}} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T r X_t^2\, dt + h\varphi_T\right].$$

The associated HJB equation is

$$\left(\frac{1}{2} a w''(x) + r x^2 - I^V\right) \wedge \left(\inf_{\gamma\in\{\pm 1\}} \gamma w'(x) + h\right) = 0.$$

An explicit solution for $w$ is provided in [53] (see also [14, 26, 32]):

$$w(x) = \begin{cases} Ax^4 + Bx^2, & -U \le x \le U, \\ w(-U) + h(-U - x), & x < -U, \\ w(U) + h(x - U), & x > U, \end{cases} \tag{4.23}$$

with

$$A = -\frac{1}{6}\frac{r}{a}, \qquad B = \frac{I}{a},$$

and

$$I = \left(\frac{3}{4} a r^{1/2} h\right)^{2/3}, \qquad U = \left(\frac{3}{4} a r^{-1} h\right)^{1/3}.$$

The optimally controlled process is

$$dX_t^* = \sqrt{a}\, dW_t + d\varphi_t^- - d\varphi_t^+,$$

where $\varphi^{\pm}$ are the local times keeping $X_t^* \in [-U, U]$ such that

$$\int_0^t \mathbb{1}_{\{X_s^* \neq -U\}}\, d\varphi_s^- + \int_0^t \mathbb{1}_{\{X_s^* \neq U\}}\, d\varphi_s^+ = 0.$$

In other words, this is a Brownian motion with reflection on the interval $[-U, U]$. The optimal solution $(\mu^*, \rho^*)$ is the stationary distribution of $X_t^*$ and the limit of boundary measures

$$\frac{1}{t}\int_0^t \big(\delta_{(U,-1,0)}\, d\varphi_s^+ + \delta_{(-U,1,0)}\, d\varphi_s^-\big),$$

as $t \to \infty$. We have the following result.

PROPOSITION 4.2. *The solution of* (4.21)–(4.22) *is given by*

$$I(a, r, h) = \left(\frac{3}{4} a r^{1/2} h\right)^{2/3},$$

*and the optimum is attained by*

$$\mu^*(dx) = \frac{1}{2U}\mathbb{1}_{[-U,U]}(x)\, dx, \tag{4.24}$$

$$\rho^*(dx, d\gamma, d\delta) = \frac{a}{2U}\left(\frac{1}{2}\delta_{(-U,1,0)} + \frac{1}{2}\delta_{(U,-1,0)}\right), \tag{4.25}$$

*where*

$$U = \left(\frac{3}{4} a r^{-1} h\right)^{1/3}.$$

EXAMPLE 4.5 (Impulse control of Brownian motion). For any parameters $r, k > 0$ and $h \geq 0$, consider the following linear programming problem:

$$I(a, r, k, h) = \inf_{(\mu,\rho)} \int_{\mathbb{R}} r x^2 \mu(dx) + \int_{\mathbb{R}\times\mathbb{R}\setminus\{0_\xi\}} \big(k + h|\xi|\big)\rho(dx, d\xi), \tag{4.26}$$

where $\mu \in \mathcal{P}(\mathbb{R})$ and $\rho \in \mathcal{M}(\mathbb{R} \times \mathbb{R} \setminus \{0\})$ satisfy

$$\begin{aligned}(4.27)\qquad &\int_{\mathbb{R}} \frac{1}{2} a f''(x)\mu(dx) \\ &\quad + \int_{\mathbb{R}\times\mathbb{R}\setminus\{0\}} \big(f(x+\xi) - f(x)\big)\rho(dx, d\xi) = 0, \qquad \forall f \in C_0^2(\mathbb{R}).\end{aligned}$$

By Corollary 4.1, this is equivalent to the time-average control problem of Brownian motion in the sense of Problem 4.1.

The associated HJB equation is (see also [2, 14, 17, 25])

$$\left(\frac{1}{2} a w''(x) + r x^2 - I^V\right) \wedge \Big(\inf_{\xi \in \mathbb{R}\setminus\{0\}} \big(w(x+\xi) + k + h|\xi|\big) - w(x)\Big) = 0.$$

Let $\theta_1, \theta_2$ and $0 < \widetilde{x}^* < x^*$ be solutions of the following polynomial system:

$$\begin{cases} 6a\theta_1 + r = 0, \\ P(x^*) - P(\widetilde{x}^*) = k + h(x^* - \widetilde{x}^*), \\ P'(x^*) = h, \qquad P'(\widetilde{x}^*) = h, \end{cases}$$

where $P(x) = \theta_1 x^4 + \theta_2 x^2$. Let $U = x^*$. We can show that the solution of the HJB equation is given by

$$w(x) = \begin{cases} P(x), & |x| \le U, \\ w(U) + h(|x| - U), & |x| > U, \end{cases}$$

and

$$I^V = a\theta_2.$$

Let $\xi^* = x^* - \widetilde{x}^*$. The optimally controlled process is a Brownian motion on the interval $[-U, U]$, which jumps to $\pm U \mp \xi^*$ when reaching the boundary point $\pm U$. Such processes have been studied in [5, 18]. We have the following result.

PROPOSITION 4.3. *The solution of* (4.26)–(4.27) *is given by*

$$I(a, r, k, h) = a\theta_2,$$

*and the optimum is attained by*

$$\begin{aligned}(4.28)\qquad \mu^* &= p(x)\,dx, \\ p(x) &= \begin{cases} \dfrac{1}{(x^*)^2 - (\widetilde{x}^*)^2}(x + x^*), & -x^* \le x < -\widetilde{x}^*, \\ \dfrac{1}{x^* + \widetilde{x}^*}, & -\widetilde{x}^* \le x \le \widetilde{x}^*, \\ \dfrac{1}{(x^*)^2 - (\widetilde{x}^*)^2}(x^* - x), & \widetilde{x}^* < x \le x^*, \end{cases}\end{aligned}$$

$$(4.29)\qquad \rho^* = \frac{a}{(x^*)^2 - (\widetilde{x}^*)^2}\left(\frac{1}{2}\delta_{(x^*, \widetilde{x}^* - x^*)} + \frac{1}{2}\delta_{(-x^*, -\widetilde{x}^* + x^*)}\right),$$

*which correspond to the stationary distribution and boundary measure of Brownian motion with jumps from the boundary on the interval* $[-U, U]$.

EXAMPLE 4.6 (Combined regular and impulse control of Brownian motion). For any parameters $r, l, k > 0$ and $h \geq 0$, consider the following linear programming problem:

$$
\begin{aligned}
I(a, r, l, k, h) = \inf_{(\mu,\rho)} \int_{\mathbb{R}\times\mathbb{R}} &(rx^2 + lu^2)\mu(dx, du) \\
&+ \int_{\mathbb{R}\times\mathbb{R}\setminus\{0_\xi\}} (k + h|\xi|)\rho(dx, d\xi),
\end{aligned}
\tag{4.30}
$$

where $\mu \in \mathcal{P}(\mathbb{R} \times \mathbb{R})$ and $\rho \in \mathcal{M}(\mathbb{R} \times \mathbb{R} \setminus \{0_\xi\})$ satisfy

$$
\begin{aligned}
&\int_{\mathbb{R}\times\mathbb{R}} \left(\frac{1}{2}af''(x) + uf'(x)\right)\mu(dx, du) \\
&\quad + \int_{\mathbb{R}\times\mathbb{R}\setminus\{0_\xi\}} \big(f(x + \xi) - f(x)\big)\rho(dx, d\xi) = 0, \qquad \forall f \in C_0^2(\mathbb{R}).
\end{aligned}
\tag{4.31}
$$

By Corollary 4.1, this is equivalent to the time-average control problem of Brownian motion in the sense of Problem 4.1.

The corresponding HJB equation is

$$
\begin{aligned}
&\left(\frac{1}{2}aw''(x) + \inf_u\big(uw'(x) + lu^2\big) + rx^2 - I^V\right) \\
&\quad \wedge \left(\inf_\xi\big(w(x + \xi) + k + h|\xi|\big) - w(x)\right) = 0.
\end{aligned}
$$

In Section 8, we show that this equation admits a classical solution:

$$
w(x) = \begin{cases} (rl)^{1/2}x^2 - 2al \ln {}_1F_1\left(\dfrac{1-\iota}{4}; \dfrac{1}{2}; \left(\dfrac{r}{a^2 l}\right)^{1/2} x^2\right), & |x| \leq U, \\ w(U) + h\big(|x| - U\big), & |x| > U, \end{cases}
\tag{4.32}
$$

where ${}_1F_1$ is the Kummer confluent hypergeometric function (see Appendix A) and $\xi^*$ and $U$ are such that $0 < \xi^* < U$ and

$$
w\big(\pm U \mp \xi^*\big) + k + h\big|\xi^*\big| - w(\pm U) = 0.
$$

Moreover,

$$
I^V = \iota a\sqrt{rl},
$$

for some $\iota \in (0, 1)$.

Let $u^*$ be defined as

$$
u^*(x) = \operatorname*{Argmin}_u Aw(x, u) + C_A(x, u) = -\frac{w'(x)}{2l}.
$$

The optimally controlled process is given by

$$dX_t^* = \sqrt{a}\,dW_t + u^*(X_t^*)\,dt + d\Big(\sum_{\tau_j \le t} \big(\mathbb{1}_{\{X^*_{\tau_j-}=-U\}}\xi^* - \mathbb{1}_{\{X^*_{\tau_j-}=U\}}\xi^*\big)\Big).$$

We have the following result.

PROPOSITION 4.4. *The solution of* (4.30)–(4.31) *is given by*

$$I(a,r,l,k,h) = \iota a\sqrt{rl},$$

*and the optimum is attained by* $(\mu^*, \rho^*)$ *where*

$$\mu^*(dx,du) = p^*(x)\,dx \otimes \delta_{u^*(x)}(du),$$
$$\rho^*(dx,d\xi) = \big(\rho_-^*\delta_{(-U,\xi^*)} + \rho_+^*\delta_{(U,-\xi^*)}\big)(dx,d\xi),$$

*with* $p^*(x) \in C^0([L,U]) \cap C^2([L,U] \setminus \{L+\xi^*(L), U+\xi^*(U)\})$ *solution of*

$$\begin{cases} \frac{1}{2}ap''(x) - \big(u^*(x)p(x)\big)' = 0, \qquad x \in (-U,U) \setminus \{-U+\xi^*, U-\xi^*\}, \\ p(-U) = p(U) = 0, \\ \frac{1}{2}ap'\big((-U)+\big) = p'\big((-U+\xi^*)-\big) - \frac{1}{2}ap'\big((-U+\xi^*)+\big), \\ \frac{1}{2}ap'(U-) = p'\big((U-\xi^*)+\big) - \frac{1}{2}ap'\big((U-\xi^*)-\big), \\ \int_{-U}^{U} p(x) = 1, \end{cases}$$

*and* $\rho_-^*, \rho_+^* \in \mathbb{R}^+$ *given by*

$$\rho_-^* = \frac{1}{2}ap'\big((-U)+\big), \qquad \rho_+^* = -\frac{1}{2}ap'(U-).$$

*Note that* $\mu^*$ *and* $\rho^*$ *are the stationary distribution and boundary measure of the optimally controlled process* $X_t^*$.

EXAMPLE 4.7 (Combined regular and singular control of Brownian motion). For any parameters $r, l, h > 0$, consider the following linear programming problem:

$$I = \inf_{(\mu,\rho)} \int_{\mathbb{R}\times\mathbb{R}} (rx^2 + lu^2)\mu(dx,du) + \int_{\mathbb{R}\times\{\pm 1\}\times\mathbb{R}_\delta^+} h|\gamma|\rho(dx,d\gamma,d\delta), \tag{4.33}$$

where $\mu \in \mathcal{P}(\mathbb{R}\times\mathbb{R})$ and $\rho \in \mathcal{M}(\mathbb{R}\times\{\pm 1\}\times\mathbb{R}_\delta^+)$ satisfy

$$\begin{aligned} &\int_{\mathbb{R}\times\mathbb{R}} \Big(\frac{1}{2}af''(x) + uf'(x)\Big)\mu(dx,du) \\ &\quad + \int_{\mathbb{R}\times\{\pm 1\}\times\mathbb{R}_\delta^+} \gamma f'(x)(dx,d\gamma,d\delta) = 0, \qquad \forall f \in C_0^2(\mathbb{R}). \end{aligned} \tag{4.34}$$

Again by Corollary 4.1, this is equivalent to the time-average control problem of Brownian motion in the sense of Problem 4.1.

The constant $I^V$ and $w : \mathbb{R} \to \mathbb{R}$ exist with the same expressions as in Example 4.6. Similarly, we have the following result.

PROPOSITION 4.5. *The solution of* (4.33)–(4.34) *is given by*

$$I = \iota a\sqrt{rl},$$

*and the optimum is attained by* $(\mu^*, \rho^*)$ *where*

$$\mu^*(dx, du) = p^*(x)\, dx \otimes \delta_{u^*(x)}(du),$$
$$\rho^*(dx, d\gamma, d\delta) = \big(\rho^*_- \delta_{(-U,1,0)} + \rho^*_U \delta_{(U,-1,0)}\big)(dx, d\gamma, d\delta),$$

*with* $p^*(x) \in C^2([-U, U])$ *solution of*

$$\begin{cases} \frac{1}{2} a p''(x) - \big(u^*(x) p(x)\big)' = 0, & x \in (-U, U), \\ \frac{1}{2} a p'(x) + u^*(x) p(x) = 0, & x \in \{-U, U\}, \\ \int_{-U}^{U} p(x) = 1, & \end{cases} \tag{4.35}$$

*and* $\rho^*_-, \rho^*_+ \in \mathbb{R}^+$ *given by*

$$\rho^*_- = \frac{1}{2} a p(-U), \qquad \rho^*_+ = \frac{1}{2} a p(U). \tag{4.36}$$

REMARK 4.2 (Higher dimension examples). To our knowledge, examples with closed-form solutions for time-average control of Brownian motion in higher dimension are not available except [2, 17, 21, 29, 44].

**5. Relation with utility maximization under market frictions.** As we have already mentioned, the lower bound (1.4) appears also in the study of impact of small market frictions in the framework of utility maximization; see [2, 19, 30, 31, 42, 44, 49, 53]. In this section, we explain heuristically how to relate utility maximization under small market frictions to the problem of tracking. It should be pointed out that we are just making connections between these two problems and no equivalence is rigorously established.

We follow the presentation in [31] and consider the classical utility maximization problem:

$$u(t, w_t) = \sup_{\varphi} \mathbb{E}\big[U\big(w_T^{t,w_t}\big)\big],$$

with

$$w_s^{t,w_t} = w_t + \int_t^s \varphi_u\, dS_u,$$

where $\varphi$ is the trading strategy. The market dynamics is an Itô semi-martingale

$$dS_t = b_t^S\,dt + \sqrt{a_t^S}\,dW_t.$$

In the frictionless market, we denote by $\varphi_t^*$ the optimal strategy and by $w_t^*$ the corresponding wealth process. As mentioned in [31], the indirect marginal utility $u'(t, w_t^*)$ evaluated along the optimal wealth process is a martingale density, which we denote by $Z_t$:

$$Z_t = u'(t, w_t^*).$$

Note that $S$ is a martingale under $\mathbb{Q}$ with

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{Z_T}{Z_0}.$$

One also defines the indirect risk tolerance process $R_t$ by

$$R_t = -\frac{u'(t, w_t^*)}{u''(t, w_t^*)}.$$

Consider the exponential utility function as in [30], that is,

$$U(x) = -e^{-px}, \qquad p > 0.$$

Then we have

$$R_t = R = \frac{1}{p}.$$

In a market with proportional transaction costs, the portfolio dynamics is given by

$$w_s^{t,w_t,\varepsilon} = w_t^\varepsilon + \int_t^s \varphi_u^\varepsilon\,dS_u - \int_t^s \varepsilon h_u\,d\|\varphi^\varepsilon\|_u,$$

where $h_t$ is a random weight process and $\varphi_t^\varepsilon$ a process with finite variation. The control problem is then

$$u^\varepsilon(t, w_t) = \sup_{\varphi^\varepsilon} \mathbb{E}\big[U\big(w_T^{t,w_t,\varepsilon}\big)\big].$$

When the cost $\varepsilon$ is small, we can expect that $\varphi_t^\varepsilon$ is close to $\varphi_t^*$ and set

$$\Delta w_T^\varepsilon := w_T^{0,w_0,\varepsilon} - w_T^* = \int_0^T (\varphi_t^\varepsilon - \varphi_t^*)\,dS_t - \varepsilon \int_0^T h_t\,d\|\varphi^\varepsilon\|_t.$$

Then, neglecting terms of higher order, we have heuristically,

$$\begin{aligned}
u^\varepsilon - u &= \mathbb{E}\big[U\big(w_T^* + \Delta w_T^\varepsilon\big)\big] - \mathbb{E}\big[U\big(w_T^*\big)\big] \\
&\simeq \mathbb{E}\Big[U'\big(w_T^*\big)\Delta w_T^\varepsilon + \frac{1}{2}U''\big(w_T^*\big)\big(\Delta w_T^\varepsilon\big)^2\Big]
\end{aligned}$$

$$
\begin{aligned}
&= -u'(w_0)\mathbb{E}^{\mathbb{Q}}\Big[-\Delta w_T^{\varepsilon} + \frac{1}{2R}(\Delta w_T^{\varepsilon})^2\Big] \\
&\simeq -u'(w_0)\mathbb{E}^{\mathbb{Q}}\Big[\varepsilon\int_0^T h_t\, d\|\varphi^{\varepsilon}\|_t + \frac{1}{2R}\Big(\int_0^T (\varphi_t^{\varepsilon} - \varphi_t^*)\, dS_t\Big)^2\Big] \\
&= -u'(w_0)\mathbb{E}^{\mathbb{Q}}\Big[\varepsilon\int_0^T h_t\, d\|\varphi^{\varepsilon}\|_t + \int_0^T \frac{a_t^S}{2R}(\varphi_t^{\varepsilon} - \varphi_t^*)^2\, dt\Big].
\end{aligned}
$$

Similarly, in a market with fixed transaction costs $\varepsilon k_t$ (see [2]), the portfolio dynamics is given by

$$
w_s^{t,w_t,\varepsilon} = w_t^{\varepsilon} + \int_t^s \varphi_u^{\varepsilon}\, dS_u + \Psi_s - \varepsilon \sum_{t<\tau_j^{\varepsilon}\le s} k_{\tau_j^{\varepsilon}} F(\xi_j^{\varepsilon}), \qquad \varphi_t^{\varepsilon} = \sum_{0<\tau_j^{\varepsilon}\le t} \xi_j^{\varepsilon},
$$

and we have

$$
u^{\varepsilon} - u \simeq -u'(w_0)\mathbb{E}^{\mathbb{Q}}\Big[\varepsilon \sum_{0<\tau_j^{\varepsilon}\le T} k_{\tau_j^{\varepsilon}} F(\xi_j^{\varepsilon}) + \int_0^T \frac{a_t^S}{2R}(\varphi_t^{\varepsilon} - \varphi_t^*)^2\, dt\Big].
$$

Finally, in a market with linear impact $\varepsilon l_t$ on price, see [44, 51], the portfolio dynamics is given by

$$
w_s^{t,w_t,\varepsilon} = w_t^{\varepsilon} + \int_t^s \varphi_u^{\varepsilon}\, dS_u - \varepsilon \int_t^s l_u (u_u^{\varepsilon})^2\, du, \qquad \varphi_t^{\varepsilon} = \int_0^t u_t^{\varepsilon}\, dt,
$$

and we have

$$
u^{\varepsilon} - u \simeq -u'(w_0)\mathbb{E}^{\mathbb{Q}}\Big[\varepsilon \int_0^T l_t (u_t^{\varepsilon})^2\, dt + \int_0^T \frac{a_t^S}{2R}(\varphi_t^{\varepsilon} - \varphi_t^*)^2\, dt\Big].
$$

To sum up, utility maximization under small market frictions is heuristically equivalent to the tracking problem if the deviation penalty is set to be

$$
r_t D(x) = \frac{a_t^S}{2R} x^2. \tag{5.1}
$$

Defining the certainty equivalent wealth loss $\Delta^{\varepsilon}$ by

$$
u^{\varepsilon} =: u(w_0 - \Delta^{\varepsilon}),
$$

it follows that

$$
\frac{1}{\varepsilon^{\beta\zeta_D}} \Delta^{\varepsilon} \simeq \mathbb{E}^{\mathbb{Q}}\Big[\int_0^T I_t\, dt\Big], \tag{5.2}
$$

see also [31], equation (3.4), and [44], page 18.

REMARK 5.1 (Higher dimension and general utility function). For the case of higher dimension and general utility function, one should set

$$
r_t D(x) = \frac{1}{2R_t}\langle x, a_t^S x\rangle. \tag{5.3}
$$

In other words, utility maximization under market frictions can be approximated at first order by the problem of tracking with quadratic deviation cost (5.3). Thus, one can establish a connection between the tracking problem and the utility maximization problems in [2, 20, 21, 44].

REMARK 5.2 (General cost structures). When there are multiple market frictions with comparable impacts, the choice of deviation penalty is the same as (5.3) and one only needs to adjust the cost structure. Our results apply directly in these cases; see [20, 42]. For example, in the case of trading with proportional cost and linear market impact (see [42]), the local problem is the time-average control of Brownian motion with cost structure:

$$C_A(x,u) = rx^2 + lu^2 + h|u|.$$

Indeed, equations (4.3)–(4.5) in [42] give rise to a verification theorem for the HJB equation of the time-average control problem of Brownian motion under this cost structure.

REMARK 5.3 (Nonzero interest rate). In the case of nonzero interest rate, the correspondence should be written as

$$\frac{1}{\varepsilon^{\beta\zeta_D}}\Delta^{\varepsilon} \simeq \mathbb{E}^{\mathbb{Q}}\left[\int_0^T e^{-r} I_t \, dt\right], \tag{5.4}$$

where $e^{-rt}S_t$ is a $\mathbb{Q}$-martingale and the tracking problem is defined by (5.3). For example, the right-hand side of (5.4) is the probabilistic representation under Black–Scholes model of equation (3.11) in [54].

REMARK 5.4 (Optimal consumption over infinite horizon). In [49, 53], the authors consider the problem of optimal consumption over infinite horizon under small proportional costs. Their results can be related to the tracking problem in the same way, that is,

$$\frac{1}{\varepsilon^{\beta\zeta_D}}\Delta^{\varepsilon} \simeq \mathbb{E}^{\mathbb{Q}}\left[\int_0^\infty e^{-r} I_t \, dt\right],$$

where $e^{-rt}S_t$ is a $\mathbb{Q}$-martingale and the tracking problem is defined by (5.3).

**6. Proof of Theorem 2.1.** This section is devoted to the proof of Theorem 2.1. In Section 6.1, we first rigorously establish the arguments outlined in Section 2.2, showing that it is enough to consider a small horizon $(t, t+\delta^{\varepsilon}]$. Then we prove Theorem 2.1 in Section 6.2. Our proof is inspired by the approaches in [36] and [41]. An essential ingredient is Lemma 6.3, whose proof is given in Section 6.3.

6.1. *Reduction to local time-average control problem.* We first show that, to obtain (2.10), it is enough to study the rescaled local version of the optimization functional (note that the parameters $r_t, l_t, k_t, h_t$ are frozen at time $t$):

$$
\begin{aligned}
I_t^\varepsilon = \frac{1}{T^\varepsilon}\biggl(\int_0^{T^\varepsilon\wedge(T-t)\varepsilon^{-2\beta}} & \bigl(r_t D(\widetilde{X}_s^{\varepsilon,t}) + l_t Q(\widetilde{u}_s^{\varepsilon,t})\bigr)\,ds \\
&+ \sum_{0<\widetilde{\tau}_j^{\varepsilon,t}\le T^\varepsilon\wedge(T-t)\varepsilon^{-2\beta}} \bigl(k_t F(\widetilde{\xi}_j^\varepsilon) + h_t P(\widetilde{\xi}_j^\varepsilon)\bigr)\biggr),
\end{aligned}
\tag{6.1}
$$

where

$$
d\widetilde{X}_s^{\varepsilon,t} = \widetilde{b}_s^{\varepsilon,t}\,ds + \sqrt{\widetilde{a}_s^{\varepsilon,t}}\,d\widetilde{W}_s^{\varepsilon,t} + \widetilde{u}_s^{\varepsilon,t}\,ds + d\biggl(\sum_{0<\widetilde{\tau}_j^{\varepsilon,t}\le s} \widetilde{\xi}_j^\varepsilon\biggr),
\tag{6.2}
$$

with $T^\varepsilon = \varepsilon^{-2\beta}\delta^\varepsilon$ and $\delta^\varepsilon \in \mathbb{R}_+$ depending on $\varepsilon$ in such a way that

$$
\delta^\varepsilon \to 0, \qquad T^\varepsilon \to \infty,
\tag{6.3}
$$

as $\varepsilon \to 0$. We can simply put $\delta^\varepsilon = \varepsilon^\beta$.

*Localization.* Since we are interested in convergence results in probability, under Assumptions 2.1 and 2.2, it is enough to consider the situation where the following assumption holds.

ASSUMPTION 6.1. There exists a positive constant $M \in \mathbb{R}_+^*$ such that

$$
\sup_{(t,\omega)\in[0,T]\times\Omega} \bigl|a_t(\omega)\bigr| \vee r_t(\omega)^{\pm1} \vee l_t(\omega)^{\pm1} \vee h_t(\omega)^{\pm1} \vee k_t(\omega)^{\pm1} < M < \infty.
$$

Furthermore, $X^\circ$ is a martingale ($b_t \equiv 0$).

Indeed, set $T_m = \inf\{t > 0, \sup_{s\in[0,t]} |b_s| \vee |a_s| \vee r_s^{\pm1} \vee l_s(\omega)^{\pm1} \vee h_s^{\pm1} \vee k_s^{\pm1} \le m\}$. Then we have $\lim_{m\to\infty} \mathbb{P}[T_m = T] = 1$. By standard localization procedure, we can assume that all the parameters are bounded as in Assumption 6.1. Let

$$
\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\biggl\{-\int_0^T a_t^{-1} b_t\,dW_t - \frac{1}{2}\int_0^T b_t^\top a_t^{-2} b_t\,dt\biggr\},
$$

then by Girsanov theorem, $X^\circ$ is a martingale under $\mathbb{Q}$. Since $\mathbb{Q}$ is equivalent to $\mathbb{P}$, we only need to prove (2.10) under $\mathbb{Q}$. Consequently, we can assume that $X^\circ$ is a martingale without loss of generality.

From now on, we will suppose that Assumption 6.1 holds.

*Locally averaged cost.*

LEMMA 6.1. *Under Assumption 6.1, we have almost surely*

$$\liminf_{\varepsilon\to 0}\frac{1}{\varepsilon^{\zeta_D\beta}}J^\varepsilon \ge \liminf_{\varepsilon\to 0}\int_0^T I_t^\varepsilon\,dt.$$

PROOF. We introduce an auxiliary cost functional:

$$\begin{aligned}\bar{J}^\varepsilon = \int_0^T \biggl(\frac{1}{\delta^\varepsilon}\int_t^{(t+\delta^\varepsilon)\wedge T}\bigl(r_s D(\overline{X}_s^\varepsilon)+l_s Q(u_s^\varepsilon)\bigr)\,ds \\ +\frac{1}{\delta^\varepsilon}\sum_{t<\tau_j^\varepsilon\le(t+\delta^\varepsilon)\wedge T}\bigl(\varepsilon^{\beta_F}k_{\tau_j^\varepsilon}F(\xi_j^\varepsilon)+\varepsilon^{\beta_P}h_{\tau_j^\varepsilon}P(\xi_j^\varepsilon)\bigr)\biggr)\,dt.\end{aligned}$$

Note that the parameters $r,l,k,h$ inside the integral are not frozen at $t$. Using Fubini's theorem, we have

$$\begin{aligned}\bar{J}^\varepsilon = \int_0^T \biggl(\frac{s}{\delta^\varepsilon}\mathbb{1}_{\{s<\delta^\varepsilon\}}+\mathbb{1}_{\{s>\delta^\varepsilon\}}\biggr)\bigl(r_s D(\overline{X}_s^\varepsilon)+l_s Q(u_s^\varepsilon)\bigr)\,ds \\ +\sum_{0<\tau_j^\varepsilon\le T}\biggl(\frac{\tau_j^\varepsilon}{\delta^\varepsilon}\mathbb{1}_{\{\tau_j^\varepsilon<\delta^\varepsilon\}}+\mathbb{1}_{\{\tau_j^\varepsilon>\delta^\varepsilon\}}\biggr)\bigl(\varepsilon^{\beta_F}k_{\tau_j^\varepsilon}F(\xi_j^\varepsilon)+\varepsilon^{\beta_P}h_{\tau_j^\varepsilon}P(\xi_j^\varepsilon)\bigr).\end{aligned}$$

Hence,

$$0\le\frac{1}{\varepsilon^{\zeta_D\beta}}(J^\varepsilon-\bar{J}^\varepsilon)\le\delta^\varepsilon I_0^\varepsilon, \tag{6.4}$$

where $I_0^\varepsilon$ is given by (6.1) with $t=0$.

On the other hand, we have

$$\begin{aligned}\widetilde{J}^\varepsilon := \int_0^T I_t^\varepsilon\,dt = \frac{1}{\varepsilon^{\zeta_D\beta}}\int_0^T\biggl(\frac{1}{\delta^\varepsilon}\int_t^{(t+\delta^\varepsilon)\wedge T}\bigl(r_t D(\overline{X}_s^\varepsilon)+l_t Q(u_s^\varepsilon)\bigr)\,ds \\ +\frac{1}{\delta^\varepsilon}\sum_{t<\tau_j^\varepsilon\le(t+\delta^\varepsilon)\wedge T}\bigl(\varepsilon^{\beta_F}k_t F(\xi_j^\varepsilon)+\varepsilon^{\beta_P}h_t P(\xi_j^\varepsilon)\bigr)\biggr)\,dt,\end{aligned}$$

where the parameters are frozen at time $t$. It follows that

$$\biggl|\frac{1}{\varepsilon^{\zeta_D\beta}}\bar{J}^\varepsilon-\widetilde{J}^\varepsilon\biggr|\le M\cdot w(r,l,k,h;\delta^\varepsilon)\cdot\biggl(\frac{1}{\varepsilon^{\zeta_D\beta}}\bar{J}^\varepsilon\wedge\widetilde{J}^\varepsilon\biggr), \tag{6.5}$$

where $w$ denotes the modulus of continuity and $M$ is the constant in Assumption 6.1. Note that we have $w(r,l,k,h;\delta)\to 0+$ as $\delta\to 0+$ by the continuity of $r,l,k,h$. Combining (6.4) and (6.5), the inequality follows. □

*Reduction to local problems*. Using the previous lemma, we can reduce the problem to the study of local problems as stated below.

LEMMA 6.2 (Reduction). *For the proof of Theorem 2.1, it is enough to show that*

$$\liminf_{\varepsilon\to 0}\mathbb{E}\big[I_t^\varepsilon\big]\geq \mathbb{E}[I_t]. \tag{6.6}$$

PROOF. In view of Lemma 6.1, to obtain Theorem 2.1, we need to prove that

$$\liminf_{\varepsilon\to 0}\int_0^T I_t^\varepsilon\,dt \geq_p \int_0^T I_t\,dt.$$

By Assumption 2.3, Lemma D.1 and Fatou's lemma, it is enough to show that

$$\liminf_{\varepsilon\to 0}\mathbb{E}\big[YI_t^\varepsilon\big]\geq \mathbb{E}[YI_t],$$

for any positive random variable $Y$ bounded away from zero and from above. Up to a change of notation $r_t\to Yr_t$, $l_t\to Yl_t$, $h_t\to Yh_t$ and $k_t\to Yk_t$ (note that this is allowed since we do not require $r_t$, $l_t$, $h_t$ and $k_t$ to be adapted), it suffices to show (6.6). $\square$

6.2. *Proof of Theorem* 2.1. After Section 6.1, it suffices to prove (6.6) where $I_t^\varepsilon$ is given by (6.1)–(6.2). In particular, we can assume that

$$\sup_{0<\varepsilon<\varepsilon_0}\mathbb{E}\big[I_t^\varepsilon\big]<\infty, \tag{6.7}$$

for some $\varepsilon_0$ small enough.

Combining ideas from [36, 41], we first consider the empirical occupation measures of $(\widetilde{X}_s^{\varepsilon,t})$. Define the following random occupation measures with natural inclusion:

$$\mu_t^\varepsilon=\frac{1}{T^\varepsilon}\int_0^{T^\varepsilon}\delta_{\{(\widetilde{X}_s^{\varepsilon,t},\widetilde{u}_s^{\varepsilon,t})\}}\,ds\in\mathcal{P}\big(\mathbb{R}^d\times\mathbb{R}^d\big),$$

$$\rho_t^\varepsilon=\frac{1}{T^\varepsilon}\sum_{0<\widetilde{\tau}_j^{\varepsilon,t}\leq T^\varepsilon}\delta_{\{(\widetilde{X}^{\varepsilon,t}_{\widetilde{\tau}_j^{\varepsilon,t}-},\widetilde{\xi}_j^\varepsilon)\}}\in\mathcal{M}\big(\mathbb{R}^d\times\mathbb{R}^d\setminus\{0\}\big)\hookrightarrow\mathcal{M}\big(\overline{\mathbb{R}^d\times\mathbb{R}^d\setminus\{0\}}\big),$$

where for a space $E$, $\overline{E}=E\cup\{\infty\}$ denotes the one-point compactification of $E$. Note that for $\bar{m}\in\mathcal{M}(\overline{E})$ we have the canonical decomposition:

$$\bar{m}(\mathrm{de})=m(\mathrm{de})+\theta\delta_\infty(\mathrm{de}),$$

with $m\in\mathcal{M}(E)$ and $\theta\in\mathbb{R}^+$. Second, we define $c_t:\Omega\times\mathcal{P}(\mathbb{R}^d\times\mathbb{R}^d)\times\mathcal{M}(\overline{\mathbb{R}^d\times\mathbb{R}^d\setminus\{0\}})\to\mathbb{R}$,

$$\begin{aligned}(\omega,\mu,\bar{\rho})\mapsto&\int_{\mathbb{R}^d\times\mathbb{R}^d}\big(r_t(\omega)D(x)+l_t(\omega)Q(u)\big)\mu(dx\times du)\\&+\int_{\overline{\mathbb{R}^d\times\mathbb{R}^d\setminus\{0\}}}\big(k_t(\omega)F(\xi)+h_t(\omega)P(\xi)\big)\bar{\rho}(dx\times d\xi),\end{aligned}$$

where the cost functions $F$ and $P$ are extended to $\overline{\mathbb{R}^d \times \mathbb{R}^d \setminus \{0\}}$ by setting

$$F(\infty_{(x,\xi)}) = \inf_{\xi \in \mathbb{R}^d \setminus \{0\}} F(\xi) > 0, \qquad P(\infty_{(x,\xi)}) = 0, \tag{6.8}$$

with $\infty_{(x,\xi)}$ the point of compactification for $\mathbb{R}^d \times \mathbb{R}^d \setminus \{0\}$. Note that the functions $F$ and $P$ remain l.s.c. on the compactified space, which is an important property we will need in the following. Moreover, for any $\bar{\rho} = \rho + \theta_{\bar{\rho}} \delta_{\infty_{(x,\xi)}}$, we have

$$c_t(\omega, \mu, \bar{\rho}) \geq c_t(\omega, \mu, \rho). \tag{6.9}$$

Now we have

$$I_t^\varepsilon = c_t(\mu_t^\varepsilon, \rho_t^\varepsilon), \tag{6.10}$$

and we can write (6.7) as

$$\sup_{0<\varepsilon<\varepsilon_0} \mathbb{E}[c_t(\mu_t^\varepsilon, \rho_t^\varepsilon)] < \infty \tag{6.11}$$

for some $\varepsilon_0$ small enough.

The following lemma, proved in Section 6.3, is the key of the demonstration for Theorem 2.1.

LEMMA 6.3 (Tightness and limiting behavior of occupation measures). *Assume that* (6.11) *holds, then*:

1. *The sequence* $\{(\mu_t^\varepsilon, \rho_t^\varepsilon)\}$ *is tight as a sequence of random variables with values in* $\mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d) \times \mathcal{M}(\overline{\mathbb{R}^d \times \mathbb{R}^d \setminus \{0\}})$, *equipped with the topology of weak convergence. In particular*, $\{\mathbb{Q}^{(\mu_t^\varepsilon, \rho_t^\varepsilon)}\}$ *is relatively compact in* $\mathcal{P}^{\mathbb{P}}(\Omega \times \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d) \times \mathcal{M}(\overline{\mathbb{R}^d \times \mathbb{R}^d \setminus \{0\}}))$, *equipped with the topology of stable convergence* (*see Appendix* D *for details on stable convergence and for additional notation*).
2. *Let* $\mathbb{Q}_t \in \mathcal{P}^{\mathbb{P}}(\Omega \times \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d) \times \mathcal{M}(\overline{\mathbb{R}^d \times \mathbb{R}^d \setminus \{0\}}))$ *be any stable limit of* $\{\mathbb{Q}^{(\mu_t^\varepsilon, \rho_t^\varepsilon)}\}$ *with disintegration form*:

$$\mathbb{Q}_t(d\omega, d\mu, d\bar{\rho}) = \mathbb{P}(d\omega) \mathbb{Q}_t^\omega(d\mu, d\bar{\rho}), \tag{6.12}$$

*and*

$$\begin{aligned} S(a) = \Big\{ &(\mu, \bar{\rho}) \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d) \times \mathcal{M}(\overline{\mathbb{R}^d \times \mathbb{R}^d \setminus \{0\}}), \bar{\rho} = \rho + \theta_{\bar{\rho}} \delta_\infty \\ &\textit{with } \rho \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d \setminus \{0\}) \textit{ and } \theta_{\bar{\rho}} \in \mathbb{R}^+, \int_{\mathbb{R}^d \times \mathbb{R}^d} A^a f(x,u) \mu(dx,du) \\ &+ \int_{\mathbb{R}^d \times \mathbb{R}^d \setminus \{0\}} Bf(x,\xi) \rho(dx,d\xi) = 0, \forall f \in C_0^2(\mathbb{R}^d) \Big\}, \end{aligned}$$

*where* $A^a$ *and* $B$ *are given by* (4.2) *and* (4.3). *Then we have, for* $\mathbb{P}$-*almost all* $\omega$,

$$\mathbb{Q}_t^\omega[(\mu, \bar{\rho}) \in S(a_t(\omega))] = 1. \tag{6.13}$$

Since the cost functional $c_t$ is lower semi-continuous, we have, using the notation introduced in Lemma 6.3, and recalling in particular that $\mathbb{E}^{\mathbb{Q}_t}$ stands for the expectation with respect to the probability measure $\mathbb{Q}_t$ in (6.12):

$$\begin{aligned}
\liminf_{\varepsilon\to 0}\mathbb{E}[I_t^\varepsilon] &\overset{(6.10)}{=} \liminf_{\varepsilon\to 0}\mathbb{E}\big[c_t\big(\mu_t^\varepsilon,\rho_t^\varepsilon\big)\big] \\
&\overset{(D.1)}{\geq} \mathbb{E}^{\mathbb{Q}_t}\big[c_t(\omega,\mu,\bar{\rho})\big] \\
&\overset{(6.9)}{\geq} \mathbb{E}^{\mathbb{Q}_t}\big[c_t(\omega,\mu,\rho)\big] \\
&\overset{(6.12)}{=} \int_\Omega \mathbb{P}(d\omega)\int_{\mathcal{P}(\mathbb{R}\times\mathbb{R})\times\mathcal{M}(\overline{\mathbb{R}\times\mathbb{R}\setminus\{0\}})} c_t(\omega,\mu,\rho)\mathbb{Q}_t^\omega(d\mu,d\bar{\rho}) \\
&\overset{(6.13)}{=} \int_\Omega \mathbb{P}(d\omega)\int_{S(a_t(\omega))} c_t(\omega,\mu,\rho)\mathbb{Q}_t^\omega(d\mu,d\bar{\rho}) \\
&\geq \int_\Omega \mathbb{P}(d\omega)\inf_{(\mu,\rho)\in S(a_t(\omega))} c_t(\omega,\mu,\rho).
\end{aligned}$$

Finally, by definition of $I$, we have

$$\liminf_{\varepsilon\to 0}\mathbb{E}\big[I_t^\varepsilon\big] \geq \mathbb{E}\big[I(a_t,r_t,l_t,k_t,h_t)\big].$$

6.3. *Proof of Lemma* 6.3. First, we show the tightness of $\{(\mu^\varepsilon,\rho^\varepsilon)\}$ in $\mathcal{P}(\mathbb{R}^d\times\mathbb{R}^d)\times\mathcal{M}(\overline{\mathbb{R}^d\times\mathbb{R}^d\setminus\{0\}})$. A common method is to use tightness functions (see Appendix C).

Recall that the cost functions $F$ and $P$ are extended to $\overline{\mathbb{R}^d\times\mathbb{R}^d\setminus\{0\}}$ by (6.8) such that $c$ is lower semi-continuous. Moreover, $c$ is a tightness function under Assumption 6.1; see Appendix C or [15], page 309.

Consequently, if (6.11) holds, the family of random measures $\{(\mu^\varepsilon,\rho^\varepsilon)\}$ is tight. Furthermore, by Proposition D.1, we have

$$\big(\mu_t^\varepsilon,\rho_t^\varepsilon\big) \to_{\text{stable}} \mathbb{Q}_t \in \mathcal{P}^{\mathbb{P}}\big(\mathcal{P}\big(\mathbb{R}^d\times\mathbb{R}^d\big)\times\mathcal{M}\big(\overline{\mathbb{R}^d\times\mathbb{R}^d\setminus\{0\}}\big)\big),$$

up to a subsequence, with

$$\mathbb{Q}_t(d\omega,d\mu,d\bar{\rho}) = \mathbb{P}(d\omega)\mathbb{Q}_t^\omega(d\mu,d\bar{\rho}).$$

For the rest of the lemma, we use a combination of the arguments in [36] and [41]. Recall that

$$A^a f(x,u) = \frac{1}{2}\sum_{i,j} a_{ij}\partial_{ij}^2 f(x) + u^\top\nabla f(x), \qquad Bf(x,\xi) = f(x+\xi) - f(x).$$

For $f\in C_0^2(\mathbb{R}^d)$, define

$$\begin{aligned}
\Psi_t^f(\omega,\mu,\bar{\rho}) := &\int_{\mathbb{R}^d\times\mathbb{R}^d} A^{a_t(\omega)} f(x,u)\mu(dx\times dx) \\
&+ \int_{\mathbb{R}^d\times\mathbb{R}^d\setminus\{0\}} Bf(x,\xi)\rho(dx,d\xi), \qquad \bar{\rho} = \rho + \theta_{\bar{\rho}}\delta_\infty.
\end{aligned}$$

Note that $\Psi_t^f$ is well defined since $\rho \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^d \setminus \{0\})$. Then we claim that

$$0 \le \mathbb{E}^{\mathbb{Q}_t}\big[\big|\Psi_t^f(\omega, \mu, \bar{\rho})\big|\big] \le \lim_{\varepsilon \to 0} \mathbb{E}\big[\big|\Psi_t^f\big(\omega, \mu_t^\varepsilon(\omega), \rho_t^\varepsilon(\omega)\big)\big|\big] = 0. \tag{6.14}$$

The second inequality follows from point 2 of Proposition D.1. For the third equality, we apply Itô's formula to $f(\widetilde{X}_{T^\varepsilon}^{\varepsilon,t})$ [recall that the dynamics of $\widetilde{X}^{\varepsilon,t}$ is given by (2.11)] and obtain that

$$\begin{aligned} f(\widetilde{X}_{T^\varepsilon}^{\varepsilon,t}) = {} & f(\widetilde{X}_{0+}^{\varepsilon,t}) + \int_0^{T^\varepsilon} f'(\widetilde{X}_s^{\varepsilon,t})\sqrt{\widetilde{a}_s^{\varepsilon,t}}\, d\widetilde{W}_s^{\varepsilon,t} \\ & + \int_0^{T^\varepsilon} \frac{1}{2}\sum_{ij} \widetilde{a}_{ij,s}^{\varepsilon,t}\partial_{ij}^2 f(\widetilde{X}_s^{\varepsilon,t})\, ds + \int_0^{T^\varepsilon} \sum_i \widetilde{u}_{i,s}^{\varepsilon,t}\partial_i f(\widetilde{X}_s^{\varepsilon,t})\, ds \\ & + \sum_{0<\widetilde{\tau}_j^{\varepsilon,t} \le T^\varepsilon} \big(f(\widetilde{X}_{\widetilde{\tau}_j^{\varepsilon,t}-}^{\varepsilon,t} + \widetilde{\xi}_j^\varepsilon) - f(\widetilde{X}_{\widetilde{\tau}_j^{\varepsilon,t}-}^{\varepsilon,t})\big). \end{aligned}$$

Combining the definitions of $\mu^\varepsilon$, $\rho^\varepsilon$ and $\Psi_t^f$, we have

$$\begin{aligned} & \mathbb{E}\big[\big|\Psi_t^f\big(\omega, \mu_t^\varepsilon(\omega), \rho_t^\varepsilon(\omega)\big)\big|\big] \\ & \quad \le \frac{1}{T^\varepsilon}\mathbb{E}\big[\big|f(\widetilde{X}_{T^\varepsilon}^{\varepsilon,t}) - f(\widetilde{X}_{0+}^{\varepsilon,t})\big|\big] + \frac{1}{T^\varepsilon}\mathbb{E}\bigg[\bigg|\int_0^{T^\varepsilon} f'(\widetilde{X}_s^{\varepsilon,t})\sqrt{\widetilde{a}_s^{\varepsilon,t}}\, d\widetilde{W}_s^{\varepsilon,t}\bigg|\bigg] \\ & \qquad + \frac{1}{T^\varepsilon}\mathbb{E}\bigg[\int_0^{T^\varepsilon} \frac{1}{2}\sum_{ij}\big|\widetilde{a}_{ij,s}^{\varepsilon,t} - \widetilde{a}_{ij,0}^{\varepsilon,t}\big|\big|\partial_{ij}^2 f(\widetilde{X}_s^{\varepsilon,t})\big|\, ds\bigg]. \end{aligned}$$

By Assumptions 2.1 and 6.1, the continuity of $(a_s)$ and dominated convergence, the term on the right hand side converges to zero. Therefore, (6.14) holds.

By definition of $\mathbb{Q}_t^\omega$, (6.14) and Fubini's theorem, we have

$$0 = \mathbb{E}^{\mathbb{Q}_t}\big[\big|\Psi_t^f(\omega, \mu, \bar{\rho})\big|\big] = \mathbb{E}^{\mathbb{P}}\big[\mathbb{E}^{\mathbb{Q}_t^\omega}\big[\big|\Psi_t^f(\omega, \mu, \bar{\rho})\big|\big]\big].$$

Hence, we have for $\mathbb{P}$-almost all $\omega$, $\mathbb{E}^{\mathbb{Q}_t^\omega}[|\Psi_t^f(\omega, \mu, \bar{\rho})|] = 0$. Let $D$ be a countable dense subset of $C_0^2$. Since $D$ is countable, we have for $\mathbb{P}$-almost all $\omega$, $\mathbb{E}^{\mathbb{Q}_t^\omega}[|\Psi_t^f(\omega, \mu, \bar{\rho})|] = 0$ for all $f \in D$. Fix $\omega \in \Omega \setminus \mathcal{N}$ for which the property holds. Again by the same argument, we have for $\mathbb{Q}_t^\omega$-almost all $(\mu, \bar{\rho})$, $\Psi_t^f(\omega, \mu, \bar{\rho}) = 0$ for all $f \in D$. Since $D$ is dense in $C_0^2$, $\Psi_t^f(\omega, \mu, \bar{\rho}) = 0$ holds for $f \in C_0^2$.

## 7. Proof of Theorem 3.1.

The rescaled process $(\widetilde{X}_s^{\varepsilon,t})$ is given by

$$d\widetilde{X}_s^{\varepsilon,t} = \widetilde{b}_s^{\varepsilon,t}\, ds + \sqrt{\widetilde{a}_s^{\varepsilon,t}}\, d\widetilde{W}_s^{\varepsilon,t} + \widetilde{u}_s^{\varepsilon,t}\, ds + \widetilde{\gamma}_s^{\varepsilon,t}\, d\widetilde{\varphi}_s^{\varepsilon,t},$$

with

$$\widetilde{\gamma}_s^{\varepsilon,t} = \gamma_{t+\varepsilon^{2\beta}s}^\varepsilon, \qquad \widetilde{\varphi}_s^{\varepsilon,t} = \frac{1}{\varepsilon^\beta}\big(\varphi_{t+\varepsilon^{2\beta}s}^\varepsilon - \varphi_t^\varepsilon\big).$$

The empirical occupation measure of singular control $\nu_t^\varepsilon$ is defined by

$$\nu_t^\varepsilon = \frac{1}{T^\varepsilon}\int_0^{T^\varepsilon} \delta_{\{(\widetilde{X}_s^{\varepsilon,t},\widetilde{\gamma}_s^{\varepsilon,t},\Delta\widetilde{\varphi}_s^{\varepsilon,t})\}}\, d\widetilde{\varphi}_s^{\varepsilon,t},$$

while $\mu_t^\varepsilon$ is defined in the same way as previously.

Define the cost functional $c_t : \Omega \times \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d) \times \mathcal{M}(\overline{\mathbb{R}^d \times \Delta \times \mathbb{R}_\delta^+}) \to \mathbb{R}$:

$$\begin{aligned}(\omega,\mu,\bar{\nu}) \mapsto & \int_{\mathbb{R}^d\times\mathbb{R}^d} \bigl(r_t(\omega)D(x) + l_t(\omega)Q(u)\bigr)\mu(dx\times du) \\ & + \int_{\overline{\mathbb{R}^d\times\Delta\times\mathbb{R}_\delta^+}} h_t(\omega)P(\gamma)\bar{\nu}(dx\times d\gamma\times d\delta),\end{aligned}$$

where $\overline{\mathbb{R}^d \times \Delta \times \mathbb{R}_\delta^+}$ denotes the one-point compactification of $\mathbb{R}^d \times \Delta \times \mathbb{R}_\delta^+$. Then we can show a similar version of Lemma 6.3 and prove Theorem 3.1 with the operator $B$ replaced by (4.5), the key ingredient being that the functional $c$ is a tightness function.

## 8. Proof of Propositions 4.1–4.5.

In this section, we prove Propositions 4.1–4.5. First, we provide a verification argument tailored to the linear programming formulation in $\mathbb{R}^d$. Second, we give full details for the proof of Proposition 4.4. The proofs in the remaining cases are exactly the same hence omitted.

8.1. *Verification theorem in* $\mathbb{R}^d$. Consider $A : \mathcal{D} \to C(\mathbb{R}^d \times \mathbb{R}^d)$ and $B : \mathcal{D} \to C(\mathbb{R}^d \times V)$ with $\mathcal{D} = C_0^2(\mathbb{R}^d)$. The operator $A$ is given by

$$Af(x,u) = \frac{1}{2}\sum_{ij} a_{ij}\partial_{ij}^2 f(x) + \sum_i u_i \partial_i f(x), \qquad f \in C_0^2(\mathbb{R}_2^d).$$

The operator $B$ is given by

$$Bf(x,\xi) = f(x+\xi) - f(x),$$

if $V = \mathbb{R}^d \setminus \{0\}$, and by

$$Bf(x,\gamma,\delta) = \begin{cases}\langle \gamma, \nabla f(x)\rangle, & \delta = 0,\\ \delta^{-1}\bigl(f(x+\delta\gamma) - f(x)\bigr), & \delta > 0,\end{cases}$$

if $V = \Delta \times \mathbb{R}_\delta^+$.

Let $C_A : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^+$ and $C_B : \mathbb{R}^d \times V \to \mathbb{R}^+$ be two cost functions. We consider the following optimization problem:

$$\begin{aligned} I &= \inf_{(\mu,\rho)} c(\mu,\rho) \\ &:= \inf_{(\mu,\rho)} \biggl\{\int_{\mathbb{R}^d\times\mathbb{R}^d} C_A(x,u)\mu(dx,du) + \int_{\mathbb{R}^d\times V} C_B(x,v)\rho(dx,dv)\biggr\},\end{aligned} \tag{8.1}$$

where $(\mu,\rho)\in\mathcal{P}(\mathbb{R}^d\times\mathbb{R}^d)\times\mathcal{M}(\mathbb{R}^d\times V)$ satisfies

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} Af(x,u)\mu(dx,du) + \int_{\mathbb{R}^d\times V} Bf(x,v)\rho(dx,dv)=0, \qquad \forall f\in C_0^2(\mathbb{R}^d). \tag{8.2}$$

LEMMA 8.1 (Verification). *Let $w\in C^1(\mathbb{R}^d)\cap C^2(\mathbb{R}^d\setminus N)$ so that $Aw$ is well defined point-wisely for $x\notin N$ and $Bw$ is well defined for $x\in\mathbb{R}^d$. Assume that*:

1. *For each $(\mu,\rho)\in\mathcal{P}(\mathbb{R}^d\times\mathbb{R}^d)\times\mathcal{M}(\mathbb{R}^d\times V)$ satisfying (8.2) and $c(\mu,\rho)<\infty$, we have $\mu(N\times\mathbb{R}^d)=0$.*
2. *There exists $w_n\in C_0^2(\mathbb{R}^d)$ such that*

$$Aw_n(x,u)\to Aw(x,u), \qquad \forall (x,u)\in\mathbb{R}^d\setminus N\times\mathbb{R}^d,$$
$$Bw_n(x,v)\to Bw(x,v), \qquad \forall (x,v)\in\mathbb{R}^d\times V,$$

*and there exist $\theta\in\mathbb{R}^+$ such that*

$$|Aw_n(x,u)|\le\theta\big(1+C_A(x,u)\big), \qquad \forall (x,u)\in\big(\mathbb{R}^d\setminus N\big)\times\mathbb{R}^d,$$
$$|Bw_n(x,v)|\le\theta C_B(x,v), \qquad \forall (x,v)\in\mathbb{R}^d\times V.$$

3. *There exists a constant $I^V\in\mathbb{R}$ such that*

$$\inf_{u\in\mathbb{R}^d} Aw(x,u)+C_A(x,u)\ge I^V, \qquad x\in\mathbb{R}^d\setminus N, \tag{8.3}$$

$$\inf_{v\in V} Bw(x,v)+C_B(x,v)\ge 0, \qquad x\in\mathbb{R}^d. \tag{8.4}$$

*Then we have $I\ge I^V$.*

*If there exists $(\mu^*,\rho^*)$ satisfying the LP constraint and*

$$Aw(x,u)+C_A(x,u)=I^V, \qquad \mu^*\text{-a.e.}, \tag{8.5}$$

$$Bw(x,v)+C_B(x,v)=0, \qquad \rho^*\text{-a.e.} \tag{8.6}$$

*then we have $I=I^V$. Moreover, the optimum is attained by $(\mu^*,\rho^*)$ and we call $(w,I^V)$ the solution of the HJB equation associated to the linear programming problem.*

PROOF OF LEMMA 8.1. Let $(\mu,\rho)$ be any pair satisfying (8.2) and $c(\mu,\rho)<\infty$. We have

$$\begin{aligned}&\int_{\mathbb{R}^d\times\mathbb{R}^d} Aw(x,u)\mu(dx,du)+\int_{\mathbb{R}^d\times V} Bw(x,v)\rho(dx,dv)\\ &\quad=\int_{\mathbb{R}^d\times\mathbb{R}^d} Aw_n(x,u)\mu(dx,du)+\int_{\mathbb{R}^d\times V} Bw_n(x,v)\rho(dx,dv)\\ &\quad=0.\end{aligned}$$

The first term is well defined since $Aw$ is defined $\mu$-everywhere. The second equality follows from the second condition and the dominated convergence theorem. Hence,

$$c(\mu,\rho)=\int_{\mathbb{R}^d\times\mathbb{R}^d}\big(C_A(x,u)+Aw(x,u)\big)\mu(dx,du)$$
$$+\int_{\mathbb{R}^d\times V}\big(C_B(x,v)+Bw(x,v)\big)\rho(dx,dv)\geq I^V,$$

where the last inequality is due to (8.3)–(8.4), and the equality holds for $(\mu,\rho)=(\mu^*,\rho^*)$ if and only if (8.5)–(8.6) are satisfied. □

Therefore, finding an explicit solution of a linear programming is possible if we can determine a suitable solution pair $(w, I^V)$ from (8.5)–(8.6).

8.2. *Verification of Proposition* 4.4. In this section, we provide an explicit solution of the following linear programming problem:

$$I(a,r,l,k,h)=\inf_{(\mu,\rho)}\int_{\mathbb{R}\times\mathbb{R}}\big(rx^2+lu^2\big)\mu(dx,du)$$
$$+\int_{\mathbb{R}\times\mathbb{R}\setminus\{0\}}\big(k+h|\xi|\big)\rho(dx,d\xi), \tag{8.7}$$

where $\mu\in\mathcal{P}(\mathbb{R}\times\mathbb{R})$ and $\rho\in\mathcal{M}(\mathbb{R}\times\mathbb{R}\setminus\{0\})$ satisfy

$$\int_{\mathbb{R}\times\mathbb{R}}\left(\frac{1}{2}af''(x)+uf'(x)\right)\mu(dx,du)$$
$$+\int_{\mathbb{R}\times\mathbb{R}\setminus\{0\}}\big(f(x+\xi)-f(x)\big)\rho(dx,d\xi)=0, \tag{8.8}$$

for any $f\in C_0^2(\mathbb{R})$. The following lemma, whose proof is given in the Appendix, and Theorem 8.1 establish the existence of the solution pair $(w, I^V)$ of the HJB equation corresponding to (8.7)–(8.8).

LEMMA 8.2 (Solution of the HJB equation for combined regular and impulse control). *There exist $U>\xi^*>0$, $I^V>0$, and $w\in C^1(\mathbb{R})\cap C^2(\mathbb{R}\setminus\{U,U\})$ such that*

$$Aw\big(x,u^*(x)\big)+C_A\big(x,u^*(x)\big)=I^V,\qquad x\in(-U,U), \tag{8.9}$$

$$Bw\big(x,-\operatorname{sgn}(x)\xi^*\big)+C_B\big(x,-\operatorname{sgn}(x)\xi^*\big)=0,\qquad x\in\{-U,U\}, \tag{8.10}$$

*where $u^*(x)$ is defined by*

$$u^*(x):=\operatorname*{Argmin}_{u\in\mathbb{R}}Aw(x,u)+C_A(x,u)=-\frac{w'(x)}{2l}. \tag{8.11}$$

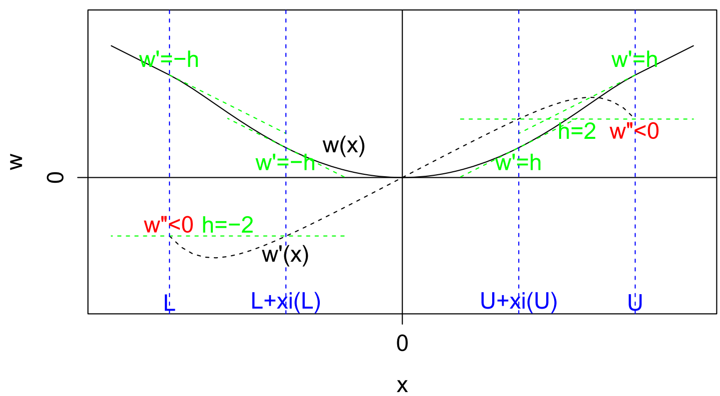


FIG. 1. *Solution of the HJB equation for combined regular and impulse control*

*More precisely, we have* (*cf. Figure* 1)

$$w(x) = \begin{cases} (rl)^{1/2}x^2 - 2al \ln {}_1F_1\left(\frac{1-\iota}{4}; \frac{1}{2}; \left(\frac{r}{a^2 l}\right)^{1/2} x^2\right), & |x| \leq U, \\ w(U) + h(|x| - U), & |x| > U, \end{cases} \tag{8.12}$$

*where* ${}_1F_1$ *is the Kummer confluent hypergeometric function* (*see Appendix* A) *and*

$$I^V = \iota\sqrt{a^2 rl},$$

*for some* $\iota \in (0, 1)$. *Moreover, $w$ satisfies the following conditions*:

$$w'(x) = w'(x - \operatorname{sgn}(x)\xi^*) = \operatorname{sgn}(x)h, \qquad x \in \{-U, U\}, \tag{8.13}$$

$$w''(x) < 0, \qquad x \in \{-U, U\}, \tag{8.14}$$

$$w'(x) \in \begin{cases} (-\infty, -h), & -U < x < -U + \xi^*, \\ (-h, h), & -U + \xi^* < x < U - \xi^*, \\ (h, \infty), & U - \xi^* < x < U, \end{cases} \tag{8.15}$$

*and* $w$, $\xi^*$, $U$ *and* $I^V$ *depend continuously on the parameters* $(a, r, k, h)$.

REMARK 8.1. Equations (8.9) and (8.10) correspond essentially to (8.5) and (8.6). The interval $(-U, U)$ is called continuation region. Equation (8.13) is the so-called "smooth-fit" condition and guarantees that $w$ is a $C^1$ function. Equations (8.14) and (8.15) characterize the growth of the derivatives of $w$ and will be useful in the proof of Theorem 8.1.

Proposition 4.4 is a direct consequence of the following theorem.

THEOREM 8.1 (Combined regular and impulse control). *For any parameters* $a, r, l, k > 0$ *and* $h \geq 0$, *we have*:

1. *The pair* $(w, I^V)$ *in Lemma* 8.2 *is the solution of the HJB equation corresponding to* (8.7)–(8.8) *in the sense of Lemma* 8.1. *In particular, the optimal cost of* (8.7)–(8.8) *is given by* $I = I^V$.

2. *Let* $p^*(x) \in C^0([-U, U]) \cap C^2((-U, U) \setminus \{-U + \xi^*, U - \xi^*\})$ *be a solution of*

$$\begin{cases} \frac{1}{2} a p''(x) - \big(u^*(x) p(x)\big)' = 0, \qquad x \in (-U, U) \setminus \{-U + \xi^*, U - \xi^*\}, \\ p(-U) = p(U) = 0, \\ \frac{1}{2} a p'\big((-U)+\big) = p'\big((-U + \xi^*)-\big) - \frac{1}{2} a p'\big((-U + \xi^*)+\big), \\ \frac{1}{2} a p'(U-) = p'\big((U - \xi^*)+\big) - \frac{1}{2} a p'\big((U - \xi^*)-\big), \\ \int_{-U}^{U} p(x) = 1, \end{cases} \tag{8.16}$$

*write* $\rho_-^*, \rho_+^* \in \mathbb{R}^+$ *for*

$$\rho_-^* = \frac{1}{2} a p'\big((-U)+\big), \qquad \rho_+^* = -\frac{1}{2} a p'(U-), \tag{8.17}$$

*and recall that* $u^*$ *is given by* (8.11). *Then the optimum of* (8.7)–(8.8) *is attained by*

$$\begin{aligned} \mu^*(dx, du) &= p^*(x)\, dx \otimes \delta_{u^*(x)}(du), \\ \rho^*(dx, d\xi) &= \rho_-^* \delta_{(-U, \xi^*)} + \rho_+^* \delta_{(U, -\xi^*)}. \end{aligned} \tag{8.18}$$

PROOF. 1. Consider the function $w$ defined in Lemma 8.2. First, we show that the three conditions in Lemma 8.1 are satisfied by $w$.

(i) Note that $N = \{-U, U\}$. For any $(\mu, \rho)$ satisfying the LP constraint, we show that $\mu(\{x\} \times \mathbb{R}) = 0, \forall x \in \mathbb{R}$. In particular, $\mu(N \times \mathbb{R}) = 0$. Indeed, let $f_n \in C_0^2(\mathbb{R})$ be a sequence of test functions such that $f_n''(z) \to \mathbb{1}_{\{x\}}(z)$ for all $z$, $\|f_n\|_\infty \vee \|f_n'\|_\infty \to 0$ and there exists $\theta \in \mathbb{R}^+$ such that

$$\begin{aligned} \big|A f_n(x, u)\big| &\le \theta\big(1 + C_A(x, u)\big), \\ \big|B f_n(x, \xi)\big| &\le \theta C_B(x, \xi) \qquad \forall x \in \mathbb{R}, u \in \mathbb{R}, \xi \in \mathbb{R} \setminus \{0\}. \end{aligned}$$

For example, let $\varphi \in C_0^2$ with $\varphi''$ being a piece-wise linear function such that $\varphi''(\pm\infty) = \varphi''(-1) = \varphi''(1) = \varphi''(3) = \varphi''(5) = 0$, $\varphi''(0) = \varphi''(4) = 1$ and $\varphi''(2) = -2$ and take $f_n(z) = \frac{1}{n^2} \varphi(n(z - x))$. Since $c(\mu, \rho) < \infty$, we have by dominated convergence theorem

$$\mu\big(\{x\} \times \mathbb{R}\big) = \lim_n \int A f_n(z, u) \mu(dz \times du) + \int B f_n(z, \xi) \rho(dz \times d\xi) = 0.$$

(ii) Let $\varphi_n \in C_0^2$ be a sequence of functions such that $\varphi_n(x) = 1$ for $|x| \le n$ and

$$\sup_n \|\varphi_n\|_{C_0^2} := \|\varphi_n\|_\infty \vee \|\varphi_n'\|_\infty \vee \|\varphi_n''\|_\infty < \infty.$$

Let $w_n = w\varphi_n$. Then $w_n$ is $C^2$ except at $\{-U, U\}$ and is of compact support. For each $n$, $w_n$ satisfies also the LP constraint

$$\int Aw_n\, d\mu + \int Bw_n\, d\rho = 0.$$

Indeed, let $\varphi_\delta$ be any convolution kernel and $w_{n,\delta} := w_n * \varphi_\delta$. So $w_{n,\delta}$ satisfies the LP constraint. Moreover, $Aw_{n,\delta} \to Aw_n$ for $x \notin \{-U, U\}$, $Bw_{n,\delta} \to Bw_n$ for any $(x, u)$ and $\sup_\delta \|w_{n,\delta}\|_{C_0^2} < \|w_n\|_{C_0^2} < \infty$. By dominated convergence, $w_n$ satisfies the LP constraint. Finally, a direct computation shows that for some constants $\theta$ and $\theta'$,

$$|Aw_n| \le \theta' \|\varphi_n\|_{C_0^2}\big(|w| + |w'| + |w''|\big) \le \theta(1 + C_A),$$

$$|Bw_n| \le 2\|\varphi_n\|_\infty |w_n| \le \theta C_B.$$

So the second condition is satisfied.

(iii) By (8.9) and (8.14), we have $Aw + C_A \ge 0$ for $x \notin \{-U, U\}$. By (8.10)–(8.15) and definition of $w$ outside $[U, U]$, we have $Bw + C_B \ge 0$.

By Lemma 8.1, we then conclude that $I = I^V$.

2. We need to show that $\mu^*$ and $\rho^*$ satisfy the LP constraint. Assume that $\mu^*$ and $\rho^*$ are given by (8.18), then by integration by parts, the LP constraint holds if $p^*(x)$ is solution of (8.16). It is easy to see that the latter admits a unique solution. □

## APPENDIX A: KUMMER CONFLUENT HYPERGEOMETRIC FUNCTION ${}_1F_1$

We collect here some properties of the Kummer confluent hypergeometric function ${}_1F_1$ which are useful to establish the existence of solutions of the HJB equations associated to combined control problems in dimension one. Recall that ${}_1F_1$ is defined as

$$ {}_1F_1(a, b, z) = \sum_{k=0}^{\infty} \frac{(a)_k}{(b)_k} \frac{z^k}{k!}, \tag{A.1}$$

with $(a)_k$ the Pochhammer symbol.

LEMMA A.1. *We have the following properties*:

1. *The function* ${}_1F_1$ *admits the following integral representation*:

$${}_1F_1(a, b, z) = \frac{\Gamma(b)}{\Gamma(b-a)\Gamma(a)} \int_0^1 e^{zt} t^{a-1}(1-t)^{b-a-1}\, dt.$$

*It is an entire function of* $a$ *and* $z$ *and a meromorphic function of* $b$.

2. *We have*

$$\frac{\partial}{\partial z}{}_1F_1(a,b,z)=\frac{a}{b}{}_1F_1(a+1,b+1,z),$$

$$\frac{\partial}{\partial a}{}_1F_1(a,b,z)=\sum_{k=0}^{\infty}\frac{(a)_k}{(b)_k}\frac{z^k}{k!}\sum_{p=0}^{k-1}\frac{1}{p+a}.$$

3. *We have*

$$(a+1)z{}_1F_1(a+2,b+2,z)$$
$$+(b+1)(b-z){}_1F_1(a+1,b+1,z)-b(b+1){}_1F_1(a,b,z)=0.$$

4. *We have*

$${}_1F_1(a,b,z)=\frac{\Gamma(a)}{\Gamma(b-a)}e^{i\pi a}z^{-a}\left(1+O\left(\frac{1}{|z|}\right)\right)+\frac{\Gamma(b)}{\Gamma(a)}e^{z}z^{a-b}\left(1+O\left(\frac{1}{|z|}\right)\right),$$

*as* $z\to\infty$.

5. *Consider the Weber differential equation*:

$$w''(x)-\left(\frac{1}{4}x^2+\theta\right)w(x)=0. \tag{A.2}$$

*The even and odd solutions of this equation are given*, *respectively*, *by*

$$\widetilde{w}(x;\theta)=e^{-\frac{1}{4}x^2}{}_1F_1\left(\frac{1}{2}\theta+\frac{1}{4},\frac{1}{2},\frac{1}{2}x^2\right), \tag{A.3}$$

$$\bar{w}(x;\theta)=xe^{-\frac{1}{4}x^2}{}_1F_1\left(\frac{1}{2}\theta+\frac{3}{4},\frac{3}{2},\frac{1}{2}x^2\right). \tag{A.4}$$

PROOF. See [1, 3]. □

## APPENDIX B: PROOF OF LEMMA 8.2

We first look for $w$ in the continuation region $(-U,U)$. Define (the change of variable comes from [47], page 260)

$$w(x):=-2al\ln\widetilde{w}\left(\frac{x}{\alpha};-\frac{\iota}{2}\right),\qquad \alpha^2=\frac{1}{2}a\left(r^{-1}l\right)^{1/2},\qquad \iota=\frac{I^V}{(a^2rl)^{1/2}},$$

where $\widetilde{w}$ is the odd solution (A.3) of the Weber differential equation (A.2). Then $w$ satisfies the following ODE:

$$\frac{1}{2}aw''(x)-\frac{1}{4l}\left(w'(x)\right)^2+rx^2=I^V,$$

which is exactly (8.9). Hence, we conjecture that the solution in the continuation region $(-U, U)$ is given by

$$\begin{aligned} w(x) &= -2al \ln\left(e^{-\frac{1}{4}x^2/\alpha^2} {}_1F_1\left(\frac{1-\iota}{4}, \frac{1}{2}, \frac{1}{2\alpha^2}x^2\right)\right) \\ &= (rl)^{1/2}x^2 - 2al \ln\left({}_1F_1\left(\frac{1-\iota}{4}, \frac{1}{2}, \frac{1}{2\alpha^2}x^2\right)\right). \end{aligned}$$

Now we show that there exist suitable values $U$, $\xi(U)$ and $\iota$ such that $0 \leq U + \xi(U) \leq U$, $\iota \in (0, 1)$ and Conditions (8.10)–(8.15) are satisfied. Let

$$h(x; \iota) := \frac{\partial w}{\partial x} = 2(rl)^{1/2}\left(1 - (1-\iota)g\left(\frac{1}{2\alpha^2}x^2; \iota\right)\right)x,$$

with

$$g(z; \iota) = \frac{{}_1F_1(\frac{1-\iota}{4} + 1, \frac{1}{2} + 1, z)}{{}_1F_1(\frac{1-\iota}{4}, \frac{1}{2}, z)}.$$

We have the following lemma.

LEMMA B.1. *The function $g(z; \iota)$ satisfies*

$$g(z; \iota) \to \begin{cases} 1, & z \to 0^+, \\ \dfrac{2}{1-\iota}, & z \to +\infty, \end{cases}$$

*and*

$$g'(z; \iota) > 0, \qquad \forall z \in [0, +\infty).$$

PROOF. The limits of $g(z; \iota)$ follow from the asymptotic behavior of ${}_1F_1$ (Property 4 in Lemma A.1). To show that $g$ is increasing, we use Properties 2 and 3 in Lemma A.1 and obtain

$$g'(z) = g(z)\left(1 - \frac{1-\iota}{2}g(z)\right) + \frac{1}{2z}(1 - g(z)).$$

Note that $g'(0) > 0$, so $g > 1$ near $x = 0$. Since $g'(z) > 0$ for $g(z) = 1$ and $g'(z) < 0$ for $g(z) = \frac{2}{1-\iota}$, $g(z)$ cannot leave the band $[1, \frac{2}{1-\iota}]$. $\square$

We now state a second lemma (cf. Figure 2).

LEMMA B.2. *The function $h(x; \iota)$ satisfies the following properties*:

1. *For $\iota \in (0, 1)$, we have*

$$\frac{h(x; \iota)}{x} \to \begin{cases} 2(rl)^{1/2}\iota, & x \to 0+, \\ -2(rl)^{1/2}, & x \to +\infty. \end{cases}$$

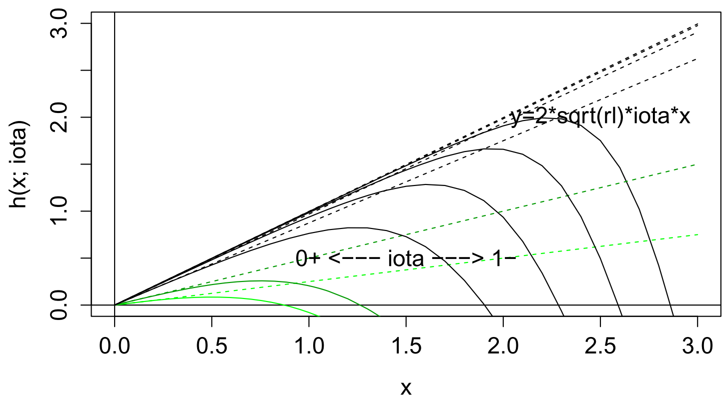


FIG. 2. *Qualitative behavior of* $h(x;\iota)$.

*Let* $0 < \bar{x}_\iota < \infty$ *be the first zero of* $h(x;\iota)$. *We have*

$$h'(x;\iota) = \begin{cases} 2(rl)^{1/2}\iota > 0, & x = 0, \\ -2(rl)^{1/2}2(1-\iota)\bar{x}_\iota g'(\bar{z}_\iota) < 0, & x = \bar{x}_\iota, \end{cases}$$

*where* $\bar{z}_\iota = \frac{1}{2\alpha^2}\bar{x}_\iota^2$ *and*

$$h''(x;\iota) < 0, \qquad x \in [0, \bar{x}_\iota].$$

2. *For* $x \in (0,\infty)$, *we have*

$$\frac{\partial}{\partial \iota} h(x;\iota) > 0$$

*and*

$$h(x;\iota) \to 2(rl)^{1/2}x, \qquad \iota \to 1-.$$

*We have*

$$\bar{x}_\iota = O(\iota^{1/2}), \qquad \iota \to 0+,$$

*and hence*

$$\max_{x\in[0,\bar{x}_\iota]} h(x;\iota) \to 0, \qquad \iota \to 0+.$$

PROOF. *For fixed* $\iota$. The asymptotic behavior of $h$ follows from Lemma B.1. The second property is clear since

$$\begin{aligned} h'(x;\iota) &= 2(rl)^{1/2}\big(-2(1-\iota)g'(z)z + \big(1-(1-\iota)g(z)\big)\big) \\ &= 2(rl)^{1/2}\big(-2(1-\iota)g'(z)z\big) + \frac{h(x)}{x}, \end{aligned}$$

with $z = \frac{1}{2\alpha^2}x^2$. Finally, we get

$$h''(x;\iota) = \frac{d}{dz}h'(z;\iota)z'(x) = -2(rl)^{1/2}(1-\iota)\big(3g'(z) + 2g''(z)z\big)z'(x).$$

Furthermore, we have

$$\begin{aligned}
&3g'(z) + 2g''(z)z \\
&\quad = 3g'(z) + 2z\left(g'(z)\big(1-(1-\iota)g(z)\big) + \frac{1}{2z^2}\big(g(z) - 1 - zg'(z)\big)\right) \\
&\quad = 2zg'(z)\big(1-(1-\iota)g(z)\big) + \frac{1}{z}\big(g(z)-1\big) + 2g'(z)
\end{aligned}$$

which is strictly positive term by term for $x \in [0, \bar{x}_\iota]$.

*For fixed* $x$. The limit of $h$ as $\iota \to 1-$ follows from the fact that ${}_1F_1(a,b,z)$ is entire in $a$ [equation (A.1) and Property 1 in Lemma A.1]. Now we show that $h$ is monotone in $\iota$. Let $G := \partial_\iota {}_1F_1$, we have

$$\begin{aligned}
\partial_\iota h(x;\iota) &= -2al\frac{\partial}{\partial\iota}\frac{\partial}{\partial x}\ln {}_1F_1\left(\frac{1-\iota}{4}; \frac{1}{2}; \frac{1}{2\alpha^2}x^2\right) \\
&= -2al\frac{\partial}{\partial x}\frac{\partial}{\partial\iota}\ln {}_1F_1\left(\frac{1-\iota}{4}; \frac{1}{2}; \frac{1}{2\alpha^2}x^2\right) \\
&= \frac{1}{2}al\frac{\partial}{\partial x}\frac{G}{{}_1F_1}\left(\frac{1-\iota}{4}; \frac{1}{2}; \frac{1}{2\alpha^2}x^2\right) \\
&= \frac{1}{2}alz'(x)\frac{{}_1F_1\frac{\partial}{\partial z}G - G\frac{\partial}{\partial z}{}_1F_1}{{}_1F_1^2}\left(\frac{1-\iota}{4}; \frac{1}{2}; z\right),
\end{aligned}$$

with $z(x) = \frac{1}{2\alpha^2}x^2$. It is enough to show that the last term is positive. Using the series representation of ${}_1F_1$ and $G$ (an Property 2 in Lemma A.1), write

$${}_1F_1 = \sum_{k=0}^{\infty} f_k z^k, \qquad G = \sum_{k=0}^{\infty}\beta_k f_k z^k,$$

with

$$\beta_k = \sum_{p=0}^{k-1}\frac{1}{p+a}, \qquad a = \frac{1-\iota}{4}.$$

We get

$$\begin{aligned}
{}_1F_1\frac{\partial}{\partial x}G - G\frac{\partial}{\partial x}{}_1F_1 &= \left(\sum_{i=0}^{\infty} f_i z^i\right)\left(\sum_{j=0}^{\infty}(j+1)\beta_{j+1}f_{j+1}z^j\right) \\
&\quad - \left(\sum_{i=0}^{\infty}(i+1)f_{i+1}z^i\right)\left(\sum_{j=0}^{\infty}\beta_j f_j z^j\right)
\end{aligned}$$

$$= \sum_{k=0}^{\infty}\Big(\sum_{i+j=k} f_i(j+1)\beta_{j+1}f_{j+1}\Big)z^k - \sum_{k=0}^{\infty}\Big(\sum_{i+j=k}(i+1)f_{i+1}\beta_j f_j\Big)z^k.$$

Then the coefficient of $z^k$ is given by

$$\begin{aligned}
&\sum_{i+j=k} f_i(j+1)\beta_{j+1}f_{j+1} - \sum_{i+j=k}(i+1)f_{i+1}\beta_j f_j\\
&\quad= \sum_{i+j=k}(j+1)(\beta_{j+1}-\beta_i)f_i f_{j+1}\\
&\quad= \sum_{1\le i<j+1\le k}\big((j+1)(\beta_{j+1}-\beta_i)f_i f_{j+1} - i(\beta_i-\beta_{j+1})f_{j+1}f_i\big)\\
&\qquad+\sum_{j+1=k+1}(\cdots)+\sum_{i=j+1}(\cdots)\\
&\quad= \sum_{1\le i<j+1\le k}(j+1-i)(\beta_{j+1}-\beta_i)f_i f_{j+1} + \sum_{j+1=k+1}(\cdots)+\sum_{i=j+1}(\cdots).
\end{aligned}$$

This term is positive since $\beta_k$ is increasing in $k$. Hence, $\partial_\iota h > 0$. Thus, $h(x;\iota)$ is increasing in $\iota$ for fixed $x \in \mathbb{R}^+$.

From the relation between $g$ and $h$, $\bar{z}_\iota$ is the first solution of

$$1-(1-\iota)g(z)=0, \qquad z>0.$$

Moreover, we have

$$g(z) = 1 + \Big(1 - \frac{1-\iota}{2}\Big)z + o(z), \qquad z \to 0+.$$

Then uniformly on $\iota$, $g(z)$ is bounded from below by $1+\frac{1}{3}z$ on $[0, z_0]$, hence

$$\bar{z}_\iota \le \frac{3\iota}{1-\iota} = O(\iota), \qquad \iota \to 0+.$$

Finally, we have

$$\max_{[0,\bar{x}_\iota]} h(x;\iota) \le 2(rl)^{1/2}\bar{x}_\iota \to 0, \qquad \iota \to 0+. \qquad \square$$

PROPOSITION B.1. *For any parameters $r, l, h > 0$ and $k \ge 0$, there exist $\iota \in (0,1)$ and $0 \le U+\xi \le U$ such that*

$$\int_{U+\xi}^{U} h(x;\iota)\,dx = k - h\xi,$$

$$h(U;\iota) = h,$$

$$h(U+\xi;\iota)=h,$$

$$h(x;\iota)\in\begin{cases}(0,h), & 0\le x\le U+\xi,\\ (h,\infty), & U+\xi\le x\le U,\end{cases}$$

$$h''(x;\iota)<0,\qquad 0\le x\le U.$$

*Moreover*, $(\iota,\xi,U)$ *depends continuously on* $(r,l,k,h)$.

PROOF. *Existence*. Let $k>0$. Since $h(x;\iota)$ is monotone in $\iota$ and $h(x;\iota)\to 2(rl)^{1/2}x$ as $\iota\to 1-$, there exists $\underline{\iota}=\underline{\iota}(h)\geq 0$ such that

$$(\underline{\iota},1)=\{\iota\in(0,1),\ \text{there exists exactly two solutions } U_\iota+\xi_\iota \text{ and } U_\iota \text{ on } [0,\bar{x}_\iota]\}.$$

We have

$$h'(U_\iota+\xi_\iota;\iota)>0,\qquad h'(U_\iota)<0,$$

so by the implicit function theorem, $U_\iota$ and $U_\iota+\xi_\iota$ depend continuously on $\iota$. Define

$$K(\iota)=\int_{U_\iota+\xi_\iota}^{U_\iota}h(x;\iota)\,dx.$$

Then $K$ is continuous in $\iota$ and

$$\lim_{\iota\to\underline{\iota}}K(\iota)=0,\qquad \lim_{\iota\to 1-}K(\iota)=\infty.$$

Hence, there exists $\iota(h,k)\in(\underline{\iota}(h),1)$ such that $K(\iota(h,k))=k$. The remaining property of $h$ is easily verified.

If $k=0$, then there exists exactly one $\iota(h)\in(0,1)$ such that the maximum of $h(x;\iota)$ is $h$ and is attained by $U_\iota$ such that

$$h'(U_\iota;\iota)=0.$$

Since $h''(U_\iota;\iota)<0$, $U_\iota$ depends continuously on $\iota$ by the implicit function theorem.

*Continuous dependence.* Since $\xi$ and $U$ depend continuously on $\iota$, it suffices to show that $\iota$ depends continuously on the parameters $a,r,h,k,l$. To see this, note that $\iota=\iota(a,l,r,h,k)$ is determined by

$$K(\iota;a,r,l,k,h)=k.$$

But, we have

$$\frac{\partial}{\partial\iota}K(\iota;a,r,l,k,h)>0.$$

Thus, $\iota$ depends continuously on the parameters by the implicit function theorem. $\square$

PROOF OF LEMMA 8.2. Extend the function $w$ in Proposition B.1 to $\mathbb{R}$ by

$$w(x) = \begin{cases} w(|x|), & |x| \le U, \\ w(U) + h(|x| - U), & |x| > U. \end{cases}$$

Then (8.9)–(8.14) hold. By (8.10) and (8.13), we have $w \in C^1(\mathbb{R}) \cap C^2(\mathbb{R} \setminus \{U, U\})$. $\square$

## APPENDIX C: TIGHTNESS FUNCTION

For more details on the following results, see [15], Appendix A.3 and [6].

DEFINITION C.1. A measurable function $g$ on a metric space taking values on the extended real line $g : (E, d) \to \mathbb{R} \cup \{\infty\}$ is a tightness function if:

1. $\inf_{x\in E} g(x) > -\infty$.
2. $\forall M < \infty$, the level set $\{x \in E | g(x) \le M\}$ is a relatively compact subset of $(E, d)$.

LEMMA C.1. *If $g$ is a tightness function on a Polish space $E$, then*:

1. *The function $G(\mu) = \int_E g(x)\mu(dx)$ is a tightness function on $\mathcal{P}(E)$.*
2. *If in addition $g \ge \delta$ where $\delta$ is a positive constant and $E$ is compact, then $G(\mu) = \int_E g(x)\mu(dx)$ is a tightness function on $\mathcal{M}(E)$.*

PROOF. Note that $\mathcal{M}(E)$ is a metric space, thus sequential compactness is equivalent to relative compactness (see [15], page 303, for the metric). For the first property, see [15], page 309. For the second property, we consider the level set $\{\mu \in \mathcal{M}(E) | G(\mu) \le M\}$ and let $\{\mu_n\}$ be any sequence in the level set. By [6], Theorem 8.6.2, it is enough to show that:

1. The sequence of nonnegative real numbers $\mu_n(E)$ is bounded.
2. The family $\{\mu_n\}$ is tight.

Since $g \ge \delta$, we have $\mu(E) \le G(\mu)/\delta \le M/\delta$. Hence the first condition is true. On the other hand, for any $\varepsilon > 0$, we consider $\mu_n/\mu_n(E) \in \mathcal{P}(E)$. Then $G(\mu_n/\mu_n(E)) \le M/\mu_n(E) \le M/\varepsilon$, if $\mu_n(E) > \varepsilon$. Since $G$ is a tightness function, we deduce that $\{\mu_n | \mu_n(E) > \varepsilon\}$ is tight. Therefore, $\{\mu_n\}$ is tight and the second condition follows. $\square$

## APPENDIX D: CONVERGENCE IN PROBABILITY, STABLE CONVERGENCE

Let $(\Omega, \mathcal{F})$ be a measurable space and $(E, \mathcal{E})$ a Polish space where $\mathcal{E}$ is the Borel algebra of $E$. Define

$$\overline{\Omega} = \Omega \times E, \qquad \overline{\mathcal{F}} = \mathcal{F} \otimes \mathcal{E}.$$

Let $B_{\mathrm{mc}}(\overline{\Omega})$ be the set of bounded measurable functions $g$ such that $z \mapsto g(\omega, z)$ is a continuous mapping for any $\omega \in \Omega$. Let $M_{\mathrm{mc}}(\overline{\Omega})$ be the set of finite positive measures on $(\overline{\Omega}, \overline{\mathcal{F}})$, equipped with the weakest topology such that

$$\mu \mapsto \int_{\overline{\Omega}} g(\omega, z)\mu(d\omega, dz),$$

is continuous for any $g \in B_{\mathrm{mc}}(\overline{\Omega})$.

We fix a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$. Let $\mathcal{P}^{\mathbb{P}}(\Omega \times E, \mathcal{F} \otimes \mathcal{E}) \subset M_{\mathrm{mc}}(\overline{\Omega})$ be the set of probability measures on $\overline{\Omega}$ with marginal $\mathbb{P}$ on $\Omega$, equipped with the induced topology from $M_{\mathrm{mc}}(\overline{\Omega})$. Note that $\mathcal{P}^{\mathbb{P}}(\Omega \times E, \mathcal{F} \otimes \mathcal{E})$ is a closed subset of $M_{\mathrm{mc}}(\overline{\Omega})$. For any random variable $Z$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we define

$$\mathbb{Q}^Z(d\omega, dz) := \mathbb{P}(d\omega) \otimes \delta_{Z(\omega)}(dz) \in \mathcal{P}^{\mathbb{P}}(\Omega \times E, \mathcal{F} \otimes \mathcal{E}).$$

DEFINITION D.1 (Stable convergence). Let $\{Z^\varepsilon, \varepsilon > 0\}$ be random variables defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in the Polish space $(E, \mathcal{E})$. We say that $Z^\varepsilon$ converges stably in law to $\mathbb{Q} \in \mathcal{P}^{\mathbb{P}}(\Omega \times E, \mathcal{F} \otimes \mathcal{E})$, written $Z^\varepsilon \to_{\mathrm{stable}} \mathbb{Q}$, if $\mathbb{Q}^{Z^\varepsilon} \to \mathbb{Q}$ in $M_{\mathrm{mc}}(\overline{\Omega})$.

We use the following properties in our proofs.

PROPOSITION D.1. *Let $\{Z^\varepsilon, \varepsilon > 0\}$ be random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $(E, \mathcal{E})$:*

1. *We have $Z^\varepsilon \to_{stable} \mathbb{Q} \in \mathcal{P}^{\mathbb{P}}(\Omega \times E, \mathcal{F} \otimes \mathcal{E})$ if and only if*

$$\mathbb{E}[Yf(Z^\varepsilon)] \to \mathbb{E}^{\mathbb{Q}}[Yf(z)],$$

*for all bounded random variables $Y$ on $(\Omega, \mathcal{F})$ and all bounded continuous functions $f \in C_b(E, \mathbb{R})$.*

2. *Assume that $Z^\varepsilon \to_{stable} \mathbb{Q} \in \mathcal{P}^{\mathbb{P}}(\Omega \times E, \mathcal{F} \otimes \mathcal{E})$. Then*

$$\liminf_{\varepsilon \to 0} \mathbb{E}[g(\omega, Z^\varepsilon)] \geq \mathbb{E}^{\mathbb{Q}}[g(\omega, z)], \tag{D.1}$$

*for any $g$ bounded from below with lower semi-continuous section $g(\omega, \cdot)$ on $E$.*

3. *Let $Z$ be a random variable defined on $(\Omega, \mathcal{F}, \mathbb{P})$. We have*

$$Z^\varepsilon \to_p Z \quad \Leftrightarrow \quad Z^\varepsilon \to_{stable} \mathbb{Q}^Z.$$

4. *The sequence $\{\mathbb{Q}^{Z^\varepsilon}, \varepsilon > 0\}$ is relatively compact in $\mathcal{P}^{\mathbb{P}}(\Omega \times E, \mathcal{F} \otimes \mathcal{E})$ if and only if $\{Z^\varepsilon, \varepsilon > 0\}$ is relatively compact as subset of $\mathcal{P}(E)$. In particular, if $E$ is compact, then $\mathcal{P}^{\mathbb{P}}(\Omega \times E, \mathcal{F} \otimes \mathcal{E})$ is compact.*

PROOF. 1. This is a direct consequence of [27], Proposition 2.4.

2. This is generalization of the Portmanteau theorem, see [27], Proposition 2.11.

3. The $\Rightarrow$ implication is obvious. Let us prove the other. Consider $F(\omega, z) = |Z(\omega) - z| \wedge 1$. On the one hand, we have $\mathbb{E}[F(\omega, Z^\varepsilon)] \to \mathbb{E}^{\mathbb{Q}^Z}[F(\omega, z)] = 0$ by definition. On the other hand, for any $\delta \in (0, 1)$ we have

$$\mathbb{P}\big[|Z^\varepsilon - Z| > \delta\big] \le \mathbb{E}\big[F\big(\omega, Z^\varepsilon\big) > \delta\big] \le \frac{\mathbb{E}[F(\omega, Z^\varepsilon)]}{\delta},$$

by Markov inequality. We deduce that $Z^\varepsilon \to_p Z$.

4. See [27], Theorem 3.8 and Corollary 3.9. □

LEMMA D.1. *Let $\{Z, Z^\varepsilon, \varepsilon > 0\}$ be* positive *random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $Z$ is integrable. Then the following properties are equivalent*:

(i) *For any positive random variable $Y$ on $(\Omega, \mathcal{F}, \mathbb{P})$, bounded away from zero and from above*,

$$\liminf_{\varepsilon \to 0} \mathbb{E}\big[Y Z^\varepsilon\big] \ge \mathbb{E}[Y Z].$$

(ii)

$$\liminf_{\varepsilon \to 0} Z^\varepsilon \ge_p Z.$$

(iii) *For any sequence $\{\varepsilon_n\}$ we can pick a subsequence $\{\varepsilon_{n_m}\}$ such that*

$$\liminf_{m \to \infty} Z^{\varepsilon_{n_m}} \ge Z \tag{D.2}$$

*almost surely.*

PROOF. (i) $\Rightarrow$ (ii). Let $\delta > 0$ be any real number and, without loss of generality, let $\{Z^\varepsilon\}$ be a minimizing sequence of $\mathbb{P}[Z^\varepsilon > Z - \delta]$ as $\varepsilon \to 0$. Considering the one-point compactification $\mathbb{R}_+ \cup \{\infty\}$, we can assume that $Z^\varepsilon$ converge stably to $\mathbb{Q} \in \mathcal{P}(\Omega \times (\mathbb{R}_+ \cup \{\infty\}))$ with canonical realization $\bar{Z}$. Then we have

$$\mathbb{E}[Y \bar{Z}] \ge \limsup_{\varepsilon \to 0} \mathbb{E}\big[Y Z^\varepsilon\big] \ge \mathbb{E}[Y Z],$$

where the first inequality comes from the fact that $z \mapsto z$ is u.s.c. on $\mathbb{R}_+ \cup \{\infty\}$ and [27], Proposition 2.11. Since $Y$ is arbitrary, we conclude that $\bar{Z} \ge Z$. Then by stable convergence of $Z^\varepsilon$ to $\bar{Z}$, we have $\mathbb{P}[Z^\varepsilon > Z - \delta] \to \mathbb{P}[\bar{Z} > Z - \delta] = 1$.

(ii) $\Rightarrow$ (i). Statement (ii) is equivalent to convergence of $(Z - Z^\varepsilon)^+$ to zero in probability. By the dominated convergence theorem, this implies that for every bounded positive $Y$,

$$\liminf_{\varepsilon\to 0} \mathbb{E}\bigl[Y\bigl(Z - Z^\varepsilon\bigr)^+\bigr] = 0,$$

and so

$$\liminf_{\varepsilon\to 0} \mathbb{E}\bigl[Y\bigl(Z - Z^\varepsilon\bigr)\bigr] \le 0,$$

which proves the implication.

(ii) $\iff$ (iii). Since (D.2) is equivalent to convergence of $(Z - Z^{\varepsilon_{n_m}})^+$ to zero almost surely, the equivalence follows by standard results. $\square$

**Acknowledgments.** The first author would like to thank Professor Thomas Kurtz for helpful advice and for sharing with us the document [38]. We are grateful to the two anonymous referees for their very constructive and insightful remarks.

## REFERENCES

[1] ABRAMOWITZ, M. and STEGUN, I. A., eds. (1972). *Handbook of Mathematical Functions*. Dover, New York.

[2] ALTAROVICI, A., MUHLE-KARBE, J. and SONER, H. M. (2015). Asymptotics for fixed transaction costs. *Finance Stoch*. **19** 363–414. MR3320325

[3] ANCARANI, L. U. and GASANEO, G. (2008). Derivatives of any order of the confluent hypergeometric function ${}_1F_1(a, b, z)$ with respect to the parameter $a$ or $b$. *J. Math. Phys*. **49** 063508, 16. MR2432916

[4] BANK, P., SONER, H. M. and VOSS, M. (2017). Hedging with temporary price impact. *Math. Financ. Econ*. **11** 215–239. MR3604450

[5] BEN-ARI, I. and PINSKY, R. G. (2009). Ergodic behavior of diffusions with random jumps from the boundary. *Stochastic Process. Appl*. **119** 864–881. MR2499861

[6] BOGACHEV, V. I. (2007). *Measure Theory. Vol. II*. Springer, Berlin. MR2267655

[7] BORKAR, V. and BUDHIRAJA, A. (2004/05). Ergodic control for constrained diffusions: Characterization using HJB equations. *SIAM J. Control Optim*. **43** 1467–1492. MR2124281

[8] BORKAR, V. S. and GHOSH, M. K. (1988). Ergodic control of multidimensional diffusions. I. The existence results. *SIAM J. Control Optim*. **26** 112–126. MR0923306

[9] BUDHIRAJA, A. and GHOSH, A. P. (2006). Diffusion approximations for controlled stochastic networks: An asymptotic bound for the value function. *Ann. Appl. Probab*. **16** 1962–2006. MR2288710

[10] BUDHIRAJA, A. and GHOSH, A. P. (2012). Controlled stochastic networks in heavy traffic: Convergence of value functions. *Ann. Appl. Probab*. **22** 734–791. MR2953568

[11] BUDHIRAJA, A., GHOSH, A. P. and LEE, C. (2011). Ergodic rate control problem for single class queueing networks. *SIAM J. Control Optim*. **49** 1570–1606. MR2817491

[12] CADENILLAS, A. and ZAPATERO, F. (2000). Classical and impulse stochastic control of the exchange rate using interest rates and reserves. *Math. Finance* **10** 141–156.

[13] CAI, J., ROSENBAUM, M. and TANKOV, P. (2016). Asymptotic optimal tracking: Feedback strategies. Preprint. Available at arXiv:1603.09472.

[14] DAI, J. G. and YAO, D. (2013). Brownian inventory models with convex holding cost, Part 1: Average-optimal controls. *Stoch. Syst*. **3** 442–499.

[15] DUPUIS, P. and ELLIS, R. S. (2011). *A Weak Convergence Approach to the Theory of Large Deviations* **902**. Wiley, New York.

[16] FUKASAWA, M. (2011). Asymptotically efficient discrete hedging. In *Stochastic Analysis with Financial Applications*. *Progress in Probability* **65** 331–346. Birkhäuser/Springer Basel AG, Basel. MR3050797

[17] GOBET, E. and LANDON, N. (2014). Almost sure optimal hedging strategy. *Ann. Appl. Probab.* **24** 1652–1690.

[18] GRIGORESCU, I. and KANG, M. (2002). Brownian motion on the figure eight. *J. Theoret. Probab.* **15** 817–844. MR1922448

[19] GUASONI, P. and WEBER, M. (2012). Dynamic trading volume. Available at www.ssrn.com.

[20] GUASONI, P. and WEBER, M. (2015). Nonlinear price impact and portfolio choice. Available at www.ssrn.com.

[21] GUASONI, P. and WEBER, M. (2015). Rebalancing multiple assets with mutual price impact. Available at www.ssrn.com.

[22] HELMES, K., STOCKBRIDGE, R. H. and ZHU, C. (2014). Impulse control of standard Brownian motion: long-term average criterion. *System Modeling and Optimization*.

[23] HYND, R. (2012). The eigenvalue problem of singular ergodic control. *Comm. Pure Appl. Math.* **LXV** 649–682.

[24] JACK, A. and ZERVOS, M. (2006). Impulse and absolutely continuous ergodic control of one-dimensional Itô diffusions. In *From Stochastic Calculus to Mathematical Finance*: *The Shiryaev Festschrift* (Y. Kabanov, R. Lipster and J. Stoyanov, eds.) 295–314. Springer, Berlin.

[25] JACK, A. and ZERVOS, M. (2006). Impulse control of one-dimensional Itô diffusions with an expected and a pathwise ergodic criterion. *Appl. Math. Optim.* **54** 71–93. MR2227624

[26] JACK, A. and ZERVOS, M. (2006). A singular control problem with an expected and a pathwise ergodic performance criterion. *J. Appl. Math. Stoch. Anal.* Art. ID 82538, 19. MR2237170

[27] JACOD, J. and MÉMIN, J. (1981). Sur un type de convergence intermédiaire entre la convergence en loi et la convergence en probabilité. In *Séminaire de Probabilités XV* 1979/80 529–546.

[28] JAKUBOWSKI, A. (1997). A non-Skorohod topology on the Skorohod space. *Electron. J. Probab.* **2** 1–21.

[29] JANEČEK, K. and SHREVE, S. E. (2010). Futures trading with transaction costs. *Illinois J. Math.* **54** 1239–1284 (2012). MR2981847

[30] KALLSEN, J. and LI, S. (2013). Portfolio optimization under small transaction costs: A convex duality approach. Preprint. Available at arXiv:1309.3479.

[31] KALLSEN, J. and MUHLE-KARBE, J. (2015). The general structure of optimal investment and consumption with small transaction costs. *Math. Finance* 1–38.

[32] KARATZAS, I. (1983). A class of singular stochastic control problems. *Adv. in Appl. Probab.* **15** 225–254. MR0698818

[33] KORN, R. (1999). Some applications of impulse control in mathematical finance. *Math. Methods Oper. Res.* **50** 493–518. MR1731297

[34] KURTZ, T. G. (1991). Random time changes and convergence in distribution under the Meyer–Zheng conditions. *Ann. Probab.* 1010–1034.

[35] KURTZ, T. G. and STOCKBRIDGE, R. H. (1998). Existence of Markov controls and characterization of optimal Markov controls. *SIAM J. Control Optim.* **36** 609–653.

[36] KURTZ, T. G. and STOCKBRIDGE, R. H. (1999). Martingale problems and linear programs for singular control. In 37*th annual Allerton Conference on Communication Control and Computing*.

[37] KURTZ, T. G. and STOCKBRIDGE, R. H. (2001). Stationary solutions and forward equations for controlled and singular martingale problems. *Electron. J. Probab.* **6** 1–52.

[38] KURTZ, T. G. and STOCKBRIDGE, R. H. (2015). Linear programming formulations of singular stochastic control problems. Personal communication.
[39] KUSHNER, H. J. (2001). *Heavy Traffic Analysis of Controlled Queueing and Communication Networks* **47**. Springer, Berlin.
[40] KUSHNER, H. J. (2014). A partial history of the early development of continuous-time nonlinear stochastic systems theory. *Automatica* **50** 303–334.
[41] KUSHNER, H. J. and MARTINS, L. F. (1993). Limit theorems for pathwise average cost per unit time problems for controlled queues in heavy traffic. *Stoch. Int. J. Probab. Stoch. Process.* **42** 25–51.
[42] LIU, R., MUHLE-KARBE, J. and WEBER, M. (2014). Rebalancing with linear and quadratic costs. Preprint. Available at arXiv:1402.5306.
[43] MENALDI, J. L., ROBIN, M. and TAKSAR, M. I. (1992). Singular ergodic control for multidimensional Gaussian processes. *Math. Control Signals Systems* **5** 93–114.
[44] MOREAU, L., MUHLE-KARBE, J. and SONER, H. M. (2015). Trading with small price impact. *Math. Finance* **27** 350–400.
[45] MUNDACA, G. and OKSENDAL, B. (1997). Optimal stochastic intervention control with application to the exchange rate. *J. Math. Econom.* **29** 225–243.
[46] NAUJOKAT, F. and WESTRAY, N. (2011). Curve following in illiquid markets. *Math. Financ. Econ.* **4** 299–335.
[47] ØKSENDAL, B. and SULEM, A. (2005). *Applied Stochastic Control of Jump Diffusions*. Springer, Berlin. MR2109687
[48] PLISKA, S. R. and SUZUKI, K. (2004). Optimal tracking for asset allocation with fixed and proportional transaction costs. *Quant. Finance* **4** 233–243. MR2055959
[49] POSSAMAÏ, D., SONER, H. M. and TOUZI, N. (2015). Homogenization and asymptotics for small transaction costs: The multidimensional case. *Comm. Partial Differential Equations* 1–42.
[50] REVUZ, D. and YOR, M. (1999). *Continuous Martingales and Brownian Motion*. Springer, Berlin.
[51] ROGERS, L. C. and SINGH, S. (2010). The cost of illiquidity and its effects on hedging. *Math. Finance* **20** 597–615.
[52] ROSENBAUM, M. and TANKOV, P. (2014). Asymptotically optimal discretization of hedging strategies with jumps. *Ann. Appl. Probab.* **24** 1002–1048.
[53] SONER, H. M. and TOUZI, N. (2013). Homogenization and asymptotics for small transaction costs. *SIAM J. Control Optim.* **51** 2893–2921.
[54] WHALLEY, A. and WILMOTT, P. (1997). An asymptotic analysis of an optimal hedging model for option pricing with transaction costs. *Math. Finance* **7** 307–324.

J. CAI
P. TANKOV
LABORATOIRE DE PROBABILITÉS
ET MODÈLES ALÉATOIRES
UNIVERSITÉ PARIS DIDEROT (PARIS 7)
AVENUE DE FRANCE CASE 7012
75205 PARIS CEDEX 13
FRANCE
E-MAIL: cai@math.univ-paris-diderot.fr
tankov@math.univ-paris-diderot.fr

M. ROSENBAUM
CENTRE DE MATHÉMATIQUES APPLIQUÉES
ECOLE POLYTECHNIQUE
91128 PALAISEAU CEDEX
FRANCE
E-MAIL: mathieu.rosenbaum@upmc.fr