Rev. Mat. Iberoam. 41 (2025), no. 2, 745–806
DOI 10.4171/RMI/1511

Controlling unknown linear dynamics with almost optimal regret

Jacob Carruth, Maximilian F. Eggl, Charles Fefferman and Clarence W. Rowley

Abstract. Here and in a companion paper, we consider a simple control problem in which the underlying dynamics depend on a parameter a that is unknown and must be learned. In this paper, we assume that a can be any real number and we do not assume that we have a prior belief about a. We seek a control strategy that minimizes a quantity called the regret. Given any $\varepsilon > 0$, we produce a strategy that minimizes the regret to within a multiplicative factor of $(1 + \varepsilon)$.

1. Introduction

Continuing from [7, 8, 17], we explore a new flavor of adaptive control theory, which we call "agnostic control".

Many works in adaptive control theory attempt to control a system whose underlying dynamics are initially unknown and must be learned from observation. The goal is then to bound REGRET, a quantity defined by comparing our expected cost with that incurred by an opponent who knows the underlying dynamics. Typically, one tries to achieve a regret whose order of magnitude is as small as possible after a long time. Adaptive control theory has extensive practical applications; see, e.g., [4, 5, 9, 22, 30].

In some applications, we do not have the luxury of waiting for a long time. This is the case, e.g., for a pilot attempting to land an airplane following the sudden loss of a wing, as in [6]. Our goal, here and in the companion paper [7], is to achieve the absolute minimum possible regret over a fixed, finite time horizon. This poses formidable mathematical challenges, even for simple model systems.

We will study a one-dimensional, linear model system whose dynamics depend on a single unknown parameter a. When a is large positive, the system is highly unstable. (There is no "stabilizing gain" for all a.) We suppose that the unknown a may be any real number and we do not assume that we are given a Bayesian prior probability distribution for it.

Mathematics Subject Classification 2020: 49N30.
Keywords: optimal control, adaptive control, agnostic control.

Modulo an arbitrarily small increase in regret, we reduce the problem to a Bayesian variant in which the unknown a is confined to a finite set and governed by a prior probability distribution.

For the Bayesian problem, our task is to find a strategy that minimizes the expected cost. This leads naturally to a PDE (a Bellman equation).

In the companion paper [7], we prove that the optimal strategy for Bayesian control is indeed given in terms of the solution of the Bellman equation, and that any strategy significantly different from that optimum incurs a significantly higher cost. We proceed modulo assumptions about existence and regularity of the relevant PDE solutions, for which we lack rigorous proofs. (However, we have obtained numerical solutions, which seem to behave as expected.)

Let us now explain the above in more detail.

1.1. The model system

Our system consists of a particle moving in one dimension, influenced by our control and buffeted by noise. The position of our particle at time t is denoted by $q(t) \in \mathbb{R}$. At each time t, we may specify a "control" $u(t) \in \mathbb{R}$, determined by history up to time t, i.e., by $(q(s))_{s\in[0,t]}$. A "strategy" (aka "policy") is a rule for specifying $u(t)$ in terms of $(q(s))_{s\in[0,t]}$ for each t. We write $\sigma, \sigma', \sigma^*$, etc. to denote strategies. The noise is provided by a standard Brownian motion $(W(t))_{t\geq 0}$.

The particle moves according to the stochastic ODE

$$dq(t) = (aq(t) + u(t))\,dt + dW(t), \quad q(0) = q_0, \tag{1.1}$$

where a and q_0 are real parameters. Due to the noise in (1.1), $q(t)$ and $u(t)$ are random variables.

Over a time horizon $T > 0$, we incur a COST, given[1] by

$$\textsc{Cost} = \int_0^T \{(q(t))^2 + (u(t))^2\}\,dt. \tag{1.2}$$

This quantity is a random variable determined by a, q_0, T and our strategy σ. Here, the starting position q_0 and time horizon T are fixed and known.

We would like to keep our cost as low as possible. We examine several variants of the above control problem, making successively weaker assumptions regarding our knowledge of the parameter a. Those variants are as follows.

Variant I: Classical control

We suppose first that the parameter a is known. We write $\mathcal{J}(\sigma, a; T, q_0)$ to denote the expected COST incurred by executing a given strategy σ. Our task is to pick σ to minimize $\mathcal{J}(\sigma, a; T, q_0)$. As shown in textbooks (e.g., [3]), the optimal strategy σ is given by

$$u(t) = -\kappa(T - t, a)q(t)$$

[1] By rescaling, we can consider seemingly different cost functions of the form $\int_0^T (q^2 + \lambda u^2)$ for $\lambda > 0$.

for a known elementary function κ; see also Section 2 below. We denote this strategy by $\sigma_{\text{opt}}(a)$. It will be important later to note that $\sigma_{\text{opt}}(a)$ satisfies the inequality

$$|u(t)| \le C \max\{a, 1\} \cdot |q(t)| \quad \text{for an absolute constant } C. \tag{1.3}$$

Variant II: Bayesian control

Next, suppose that the parameter a is unknown, but is subject to a given prior probability distribution $d\text{Prob}(a)$ supported in an interval $[-A, A]$. Our goal is now to pick a strategy σ to minimize our expected cost, given by

$$\int_{-A}^{A} \mathcal{J}(\sigma, a; T, q_0)\, d\text{Prob}(a). \tag{1.4}$$

To solve this problem, we first note a major simplification: in principle, a strategy σ is a one-parameter family of functions on an infinite-dimensional space, because it specifies $u(t)$ in terms of the path $(q(s))_{s\in[0,t]}$ for each t. However, one computes (for details, see [7]) that the posterior probability distribution for the unknown a, given past history $(q(s))_{s\in[0,t]}$, is determined by the prior $d\text{Prob}(a)$, together with the two observable quantities

$$\zeta_1(t) = \int_0^t q(s)\, dq(s) - \int_0^t q(s)u(s)\, ds \quad \text{and} \quad \zeta_2(t) = \int_0^t (q(s))^2\, ds. \tag{1.5}$$

Therefore, it is natural to suppose that our optimal strategy $\sigma_{\text{Bayes}}(d\text{Prob})$ takes the form

$$u(t) = \tilde{u}(q(t), \zeta_1(t), \zeta_2(t), t) \tag{1.6}$$

for a function $\tilde{u}$ on $\mathbb{R}^4$.

So, instead of looking for a one-parameter family of functions on an infinite-dimensional space, we merely have to specify a function $\tilde{u}$ of four variables. It is not hard to derive a PDE (the Bellman equation) for the function $\tilde{u}(q, \zeta_1, \zeta_2, t)$ that, according to heuristic reasoning, minimizes our expected cost. We have produced approximate solutions $\tilde{u}$ to the Bellman equation in numerical simulations[2], but we do not have rigorous proofs of existence or regularity. We proceed by imposing the

PDE ASSUMPTION: The Bellman equation has a smooth solution, and the resulting control strategy satisfies the estimate

$$|u(t)| \le C_0 A^{m_0}\,[|q(t)| + 1], \tag{1.7}$$

for constants C_0 and m_0 independent of A.

Since our prior distribution $d\text{Prob}(a)$ is supported in $[-A, A]$, a glance at (1.3) suggests that the optimal Bayesian strategy should satisfy

$$|u(t)| \le C \max\{A, 1\} \cdot |q(t)|.$$

Our numerical simulations appear to confirm this belief. Accordingly, (1.7) seems to be a very safe assumption.

[2]For details on all of the numerical simulations referenced in this paper, we refer the reader to the supplementary material available on our website: https://github.com/meggl23/NumericalAgnosticControl (visited on November 20, 2024).

Under the above PDE Assumption, we prove in [7] that the natural result that the strategy $\sigma_{\mathrm{Bayes}}(d\mathrm{Prob})$ arising from the Bellman equation indeed minimizes the expected cost given by (1.4). Moreover, any strategy that differs significantly from $\sigma_{\mathrm{Bayes}}(d\mathrm{Prob})$ incurs a significantly higher expected cost.

Our results on Bayesian control pave the way for our analysis of agnostic control.

Variant III: Agnostic control for bounded a

We suppose that our parameter a is confined to a bounded interval $[-A, A]$, but is otherwise unknown. In particular, we do not assume that we are given a Bayesian prior probability distribution $d\mathrm{Prob}(a)$. Consequently, we cannot define a notion of expected cost by formula (1.4).

Instead, our goal will be to minimize *worst-case regret*, defined by comparing the performance of our strategy with that of the optimal known-a strategy $\sigma_{\mathrm{opt}}(a)$. We will introduce several variants of the notion of regret.

Let us fix a starting position q_0, a time horizon T, and an interval $[-A, A]$ guaranteed to contain the unknown a. To a given strategy σ, we associate the following functions on $[-A, A]$:

- *Additive Regret*, defined as

$$\mathrm{AR}(\sigma, a) = \mathcal{J}(\sigma, a; T, q_0) - \mathcal{J}(\sigma_{\mathrm{opt}}(a), a; T, q_0) \geq 0.$$

- *Multiplicative Regret* (aka "competitive ratio"), defined as

$$\mathrm{MR}(\sigma, a) = \frac{\mathcal{J}(\sigma, a; T, q_0)}{\mathcal{J}(\sigma_{\mathrm{opt}}(a), a; T, q_0)} \geq 1.$$

- *Hybrid Regret*, defined in terms of a parameter $\gamma > 0$ by setting

$$\mathrm{HR}_\gamma(\sigma, a) = \frac{\mathcal{J}(\sigma, a; T, q_0)}{\mathcal{J}(\sigma_{\mathrm{opt}}(a), a; T, q_0) + \gamma}.$$

Writing $\textsc{Regret}(\sigma, a)$ to denote any one of the above three functions on $[-A, A]$, we define the *worst-case regret*

$$\textsc{Regret}^*(\sigma) = \sup\{\textsc{Regret}(\sigma, a) : a \in [-A, A]\}. \tag{1.8}$$

We seek a strategy σ that minimizes worst-case regret.

The above notions are useful in different regimes. If we expect to pay a large cost, then we care more about multiplicative regret than about additive regret. (If we have to pay 10^9 dollars, we are unimpressed by a saving of 10^5 dollars.) Similarly, if our expected cost is small, then we care more about additive regret than about multiplicative regret. (If we pay only 10^{-5} dollars, we do not care that we might instead pay 10^{-9} dollars.) If we fix γ to be a cost we are willing to neglect, then hybrid regret $\mathrm{HR}_\gamma(\sigma, a)$ provides meaningful information regardless of the order of magnitude of our expected cost.

So far, we have defined three flavors of worst-case regret, and posed the problem of minimizing that regret. The solution to our agnostic control problem is given by the following result, proved in [7].

Theorem 1.1. *Fix* $[-A, A]$, q_0 *and* T (*and* γ, *if we use hybrid regret*). *Assume the* PDE Assumption. *Then*

(I) *There exists a probability measure* $d\mathrm{Prob}_*$ *on* $[-A, A]$ *for which the optimal Bayesian strategy* $\sigma_{\mathrm{Bayes}}(d\mathrm{Prob}_*)$ *minimizes the worst-case regret among all possible strategies.*

Moreover,

(II) *The measure* $d\mathrm{Prob}_*$ *is supported on a finite set* $E \subset [-A, A]$, *where*

(III) E *is precisely the set of points* $a \in [-A, A]$ *at which the function* $[-A, A] \ni a \mapsto \mathrm{REGRET}(\sigma_{\mathrm{Bayes}}(d\mathrm{Prob}_*), a)$ *achieves its maximum.*

So, for optimal agnostic control, we should pretend to believe that the unknown a is confined to a finite set E and governed by the probability distribution $d\mathrm{Prob}_*$, even though in fact we know nothing about a except that it lies in $[-A, A]$.

It is easy to see that a probability distribution $d\mathrm{Prob}_*$ satisfying (II) and (III) gives rise to a Bayesian strategy satisfying (I). The hard part of Theorem 1.1 is the assertion that such a probability measure exists.

Theorem 1.1 lets us search for optimal agnostic strategies: we first guess a finite set E and a probability measure $d\mathrm{Prob}$ concentrated on E. By solving the Bellman equation, we produce the strategy $\sigma = \sigma_{\mathrm{Bayes}}(d\mathrm{Prob})$, which allows us to compute the function $[-A, A] \ni a \mapsto \mathrm{REGRET}(\sigma, a)$. If the maxima of that function occur precisely at the points of E, then σ is the desired optimal agnostic strategy. If the maxima of that function do not occur at the points of E, then we update our guess for E and try again. We have carried this out numerically for several $[-A, A]$, $q_0$and T.

This concludes our discussion of agnostic control for bounded a. Finally, we pass to the most general case.

Variant IV: Fully agnostic control

We make no assumption whatever regarding the unknown a; our a may be any real number, and we are not given a Bayesian prior distribution for it.

Our goal is again to find a strategy that minimizes the *worst-case regret*, defined as in (1.8), except that the sup is now taken over all $a \in \mathbb{R}$.

The main result of this paper is that, with negligible increase in regret, we can reduce matters to agnostic control for bounded a. More precisely, we will prove the following result.

Theorem 1.2. *Fix a time horizon* T *and a nonzero starting position* q_0, *as well as constants* C_0 *and* m_0 (*for which, see* (1.7)). *Then, given* $\varepsilon > 0$, *there exists* $A > 0$ *for which the following holds.*

Let σ *be a strategy for the starting position* q_0 *and time horizon* $T + \varepsilon$. *Suppose* σ *satisfies estimate* (1.7) *for the given* C_0, m_0 *and* A.

Then there exists a strategy σ_* *for the starting position* q_0 *and the time horizon* T, *satisfying the following estimates.*

(A) *For* $a \in [-A, A]$, *we have*

$$\mathcal{J}(\sigma_*, a; T, q_0) \leq \varepsilon + (1+\varepsilon)\ \sup\{\mathcal{J}(\sigma, a'; T+\varepsilon, q_0) : |a' - a| \leq \varepsilon|a|\}.$$

(B) *For $a \notin [-A, A]$, we have*

$$\mathcal{J}(\sigma_*, a; T, q_0) \le \varepsilon + (1 + \varepsilon)\, \mathcal{J}(\sigma_{\mathrm{opt}}(a), a; T, q_0).$$

So, if $a \in [-A, A]$, then σ_* performs almost as well as σ; and if $a \notin [-A, A]$, then σ_* performs almost as well as the optimal known-a strategy $\sigma_{\mathrm{opt}}(a)$.

To apply Theorem 1.2, we take σ to be an optimal agnostic strategy for the starting position q_0 and the time horizon $T + \varepsilon$, assuming a to be confined to the interval $[-(1+\varepsilon)A, +(1+\varepsilon)A]$.

If our *PDE Assumption* holds, then σ satisfies the hypothesis of Theorem 1.2. It's easy to deduce from Theorem 1.2 that the worst-case hybrid regret of the strategy σ_* (for fully agnostic control) is at most $O(\varepsilon)$ percent worse than that of σ (for agnostic control with a confined to $[-(1 + \varepsilon)A, (1 + \varepsilon)A]$).

For any $\gamma > 0$ and any strategy σ for time horizon T and starting position q_0, we let $\mathrm{HR}^*_\gamma(\sigma; T)$ denote the worst-case hybrid regret (over all $a \in \mathbb{R}$) of σ. The above discussion then implies the following.

Corollary 1.3. *Fix constants $T > 0$, $\eta > 0$ and $q_0 \neq 0$. Assume the* PDE Assumption. *Then for any $\varepsilon > 0$, we can construct a strategy σ_{Ag} for time horizon T and starting position q_0 satisfying*

$$\mathrm{HR}^*_\gamma(\sigma_{\mathrm{Ag}}; T) \le (1 + C\varepsilon) \cdot \mathrm{HR}^*_\gamma(\sigma; T + \varepsilon)$$

for any strategy σ for time horizon $T + \varepsilon$ and starting position q_0.

There are variants of Theorem 1.2 and Corollary 1.3 for the case of starting position $q_0 = 0$; see Corollary 2.2.

Recap

Let us summarize what has been achieved.

Our goal is to minimize worst-case regret in the setting where a may be any real number. Modulo an arbitrarily small increase in regret, we may reduce matters to the case in which a is confined to a bounded interval $[-A, A]$. We then look for a probability measure $d\mathrm{Prob}_*$ living on a finite subset $E \subset [-A, A]$ such that the regret of the optimal Bayesian strategy for $d\mathrm{Prob}_*$ is maximized precisely on E. We can calculate the optimal Bayesian strategy for a given prior probability measure by solving a Bellman equation. However, our results are conditional; we have to make an assumption on the existence, size, and smoothness of solutions to the Bellman equation. In numerical simulations, we have produced evidence for our PDE Assumption, and we have produced optimal agnostic strategies for cases in which the unknown a is confined to a bounded interval.

1.2. Ideas from the proof of Theorem 1.2

Recall that, in Theorem 1.2, we are given a strategy σ, a small positive ε, and a large positive A depending on ε. The strategy σ applies to a starting position $q_0 \neq 0$, and to a time horizon $T + \varepsilon$. Our task is to find a strategy σ_* satisfying conditions (A) and (B).

In this introduction, we allow ourselves some inaccuracy in the interest of simplicity. See Sections 5–7 for correct details.

A recurring theme in the proof of Theorem 1.2 is that because the unknown a may be arbitrarily large and positive, the system may be arbitrarily unstable. Consequently, disasters of exponentially small probability may lead to exponentially large expected cost. To prepare the way for the proof of Theorem 1.2, we first examine two regimes in which we can do almost as well as if we knew the value of a.

The large q regime

Suppose q_0 is very large. A glance at (1.1) suggests that the noise $dW(t)$ has only a small effect compared to that of the term $(aq + u)\,dt$. Therefore, after initially setting $u \equiv 0$ and observing $q(t)$ for small t, we quickly arrive at a guess $\bar{a}$ for the unknown a. That guess is probably highly accurate. Moreover, the larger a is, the sooner we can arrive at the guess $\bar{a}$.

Once we have found $\bar{a}$, we can simply play the known-a strategy $\sigma_{\text{opt}}(\bar{a})$ until the end of the game at time T. (If $\bar{a}$ is large positive, then in place of $\sigma_{\text{opt}}(\bar{a})$, we use the strategy in which $u(t) = -2\bar{a}q(t)$, which is equivalent to $\sigma_{\text{opt}}(\bar{a})$ asymptotically for $\bar{a} \gg 1$.)

This "Large q strategy" incurs an expected cost almost as small as that of an opponent who knows a.

A crucial point is that for large positive a, the probability of a significantly inaccurate guess $\bar{a}$ is $O(\exp(-cq_0^2 a))$, while the expected cost if such an error occurs is $O(q_0^2 \exp(CTa))$. Hence, for large q_0, the exponentially tiny probability of disaster overwhelms the exponentially large resulting cost.

The large a regime

As in our discussion of the Large q strategy, a glance at (1.1) suggests that the nosie $dW(t)$ will have negligibly small effect compared to that of the term $(aq + u)\,dt$, provided a is large positive, say, $a \geq A$. This leads to a "naïve large a strategy" in which we initially set $u \equiv 0$, observe $q(t)$ for small t, arrive at a guess $\bar{a}$ for the unknown a, and then play the known-a strategy $\sigma_{\text{opt}}(\bar{a})$ until the game ends at time T.

This time, however, the exponentially small probability of an error of the form $\bar{a} \ll a$ is dominated by the exponentially large cost of the ensuing disaster. Consequently, the naïve large a strategy fails. The cure is to pick some q_0^*, big but not too big, and execute the naïve large a strategy only until the first moment we encounter $|q(t)| = q_0^*$. At that moment, we switch over to the Large q strategy. If we never encounter $|q(t)| = q_0^*$, then we continue with the naïve large a strategy until the end of the game.

The above modification limits the damage arising from the event $\bar{a} \ll a$. We have thus produced a "Large a strategy" whose expected cost is close to that of the optimal known-a strategy $\sigma_{\text{opt}}(a)$ whenever $a \geq A$. When $a < A$, the expected cost of our Large a strategy is $O(A^2)$. That is bigger than we would like, but it is not exponentially large.

Strengthening the given strategy

Next, we consider the strategy σ given in the statement of Theorem 1.2. Recall that σ is assumed to satisfy the estimate (1.7). From that estimate, we see at once that σ disastrously undercontrols in case $a \gg A^{m_0}$, leading to exponentially large expected cost. We can remedy this defect by modifying σ the same way we modified the naïve large a strategy. We pick a q_0^*, large but not too large, and switch over from σ to the Large q strategy as soon as we encounter $|q(t)| = q_0^*$. Thus, we obtain a strategy $\tilde{\sigma}$ that performs almost as well as σ for $a \in [-A, A]$, and avoids exponentially large expected cost if $a > A$.

Like σ, the strategy $\tilde{\sigma}$ starts at position q_0. By rescaling $\tilde{\sigma}$ slightly, we arrive at a strategy $\bar{\sigma}$ with starting position $(1+\varepsilon)q_0$. Like $\tilde{\sigma}$, the strategy $\bar{\sigma}$ performs almost as well as σ when $a \in [-A, A]$, and avoids exponentially large expected cost when $a > A$.

Armed with the Large q and Large a strategies, and the modified strategy $\bar{\sigma}$, we can now describe the strategy σ_* whose existence is asserted by Theorem 1.2. Without loss of generality, we suppose that our nonzero starting position q_0 is positive.

In the strategy σ_*, there are two epochs, a Prologue and a Main Act. During the Prologue, we set $u \equiv 0$. The Prologue ends as soon as we encounter one of the three events

(a) $q(t) = (1+\varepsilon)q_0$,

(b) $q(t) = -q_{\mathrm{rare}}$ for a suitable $q_{\mathrm{rare}} > 0$ (big but not too big),

(c) the end of the game at time T (in which case there is no Main Act).

Event (b) occurs with small probability, regardless of the unknown a. If it does occur, then during the Main Act we play a slight variant of the Large q strategy to bound our losses.

If instead we enter the Main Act via case (a), then by observing how long it took to pass from the initial position q_0 to the position $(1+\varepsilon)q_0$, we obtain a guess $\bar{a}$ for the unknown a. If a is large positive, then as in our discussion of the Large a strategy, our guess $\bar{a}$ is probably highly accurate. Otherwise, $\bar{a}$ is likely not so close to a, but at least $\bar{a}$ probably will not be large positive. Therefore, for large A, our guess $\bar{a}$ will at least tell us whether $a \gtrsim A$ or $a \ll A$. Accordingly, in case (a) we proceed as follows during the Main Act.

- If $\bar{a} > A$, then during the Main Act, we execute the Large a strategy.
- If $\bar{a} \leq A$, then during the Main Act, we execute $\bar{\sigma}$, the improved version of the given strategy σ.

The Main Act lasts until the end of the game at time T.

This completes our description of the strategy σ_*. We hope the reader finds it plausible that our σ_* satisfies conditions (A) and (B) in Theorem 1.2.

We again warn the reader that our discussion in this introduction is somewhat oversimplified. For instance, our basic stochastic ODE (1.1) is not obviously well defined if our strategy allows $u(t)$ to be a discontinuous function of t. We have not even given a rigorous definition of a strategy. Our rigorous discussion starts from scratch in Section 2.

Survey of prior literature

Literature that considers adaptive control of a simple linear system similar to the one considered in this paper commonly consists of one or more of the following features: (i) unknown governing dynamics, (ii) unknown cost function, and (iii) adversarial noise. Examples of such work include [12,15,19,25–27,34] as well as our own prior work [8,17].

Initial work in obtaining regret bounds in the infinite time horizon for the related LQR (linear-quadratic regulator) problem was undertaken in [1], which proved that under certain assumptions, the expected additive regret of the adaptive controller is bounded by $\tilde{O}(\sqrt{T})$. Further progress was made on this problem in [10]. Assuming controllability of the system, the authors gave the first efficient algorithm capable of attaining sublinear additive regret in a single trajectory in the setting of online nonstochastic control. See also the related [29], which obtained sublinear adaptive regret bounds, a stronger metric than standard regret and more suitable for time-varying systems. Additional adaptive control approaches include [13, 14] using the system level synthesis. This expands on ideas in [32], which showed that the ordinary least-squares estimator learns a linear system nearly optimally in one shot. Other work uses Thompson sampling [2, 24] or deep learning [11]. Perhaps most related to the work performed in this study is [23], which designed an online learning algorithm with sublinear expected regret that moves away from episodic estimates of the state dynamics (meaning that no boundedness or initially stabilizing control needed to be assumed).

In [17], the third and fourth authors of the present paper, along with B. Guillén Pegueroles and M. Weber, found regret minimizing strategies for a problem with simple unknown dynamics (a particle moving in one-dimension at a constant, unknown velocity subject to Brownian motion). In [21], along with D. Goswami and D. Gurevich, they generalized these results to an analogous, higher-dimensional system with the addition of sensor noise. In [17], they also posed the problem of finding regret minimizing strategies for the more complicated dynamics (1.1). In [8], the authors of the present paper, along with M. Weber, took the first steps toward resolving this problem. Specifically, we exhibited a strategy for the dynamics (1.1) with bounded multiplicative regret.

Historically, significant work has been undertaken in the closely related “multi-armed bandit” problem; see, for instance, the classic papers [31, 33]. Recent work considering this paradigm includes [35], which used reinforcement learning to obtain dynamic regret whose order of magnitude is optimal, and [16], which studied the more general generalized linear bandits (GLBs) and obtains similar regret bounds.

We finally want to point out the parallel field of adversarial control, where the noise profile is chosen by an adversary instead of randomly. This includes [28], which attained minimum dynamic regret and guaranteed compliance with hard safety constraints in the face of uncertain disturbance realizations using the system level synthesis framework, and [20], which studied the problem of competitive control.

As this list of references is by no means exhaustive and does not do justice to the wealth of studies in the literature, we point the reader to the book [22] and the references therein for a more thorough overview of online control.

We emphasize that our approach in [17], [8], and the present paper differs from the other work cited above in that

- we seek strategies that minimize the worst-case regret for a fixed time horizon T, whereas the literature is mainly concerned with $T \to \infty$.
- Typically, in the literature one assumes either that the dynamics are bounded or that one is given a stabilizing control. We make no such assumptions, and so we must control a system that is arbitrarily unstable.
- However, we achieve the above ambitious goals only for a simple model system.

2. Setup

Except for where we explicitly state otherwise, we adopt the following conventions regarding constants throughout this paper. We write $C, C', C'', \ldots$ to denote positive absolute constants. When we wish to specify that a positive absolute constant is smaller than 1, we write $c, c', c'', \ldots$. Finally, if a constant depends on some quantity X, then we write $C_X, C'_X, c_X, c'_X, \ldots$. The values of these constants may change from line to line.

Whenever we say that a parameter/constant depends on other parameters/constants, we assume that the dependence is continuous unless we explicitly state otherwise.

We let $W(t)$ denote Brownian motion starting at $W(0) = 0$ and normalized so that $\mathrm{E}[(W(t))^2] = t$. We write $(\Omega, \mathcal{F}, \mathrm{Prob})$ for the corresponding probability space, and $\mathrm{E}[X]$ for the expected value of a random variable X. We write ω to denote an arbitrary element of Ω. For $t \in [0, T]$, we write $\mathcal{F}_t$ to denote the sigma algebra determined by the history of the Brownian motion from time 0 until time t.

We introduce a *time horizon* $T > 0$ and a *starting position* $q_0 \neq 0$. We define a *strategy* (for time horizon T and starting position q_0) to be a collection of random variables $q(t, a)$, $u(t, a)$ defined for all $t \in [0, T]$ and $a \in \mathbb{R}$ and satisfying the following:

(S.1) For every $a \in \mathbb{R}$, $q(t, a)$ is a continuous function of t with probability 1 and $u(t, a)$ is an L^2 function of t with probability 1.

(S.2) For every $a \in \mathbb{R}$ and $t \in [0, T]$, the maps $(s, \omega) \mapsto q(s, a, \omega)$ and $(s, \omega) \mapsto u(s, a, \omega)$, defined on $[0, t] \times \Omega$, are measurable as functions on $[0, t] \times (\Omega, \mathcal{F}_t, \mathrm{Prob})$. Intuitively, this means that q and u are determined by the past.

(S.3) For every $a \in \mathbb{R}$,

$$\mathrm{E}\Big[\int_0^T (q(t, a))^2 + (u(t, a))^2 \, dt \Big] < \infty.$$

(S.4) For almost all $\omega \in \Omega$, we have that for all $a, b \in \mathbb{R}$, and for all $t \in [0, T]$, if $q(s, a, \omega) = q(s, b, \omega)$ for all $s \in [0, t]$, then $u(s, a, \omega) = u(s, b, \omega)$ for almost all $s \in [0, t]$. This tells us that u does not depend on the unknown a.

(S.5) For every $a \in \mathbb{R}$ and $t \in [0, T]$, we have

$$q(t, a) = q_0 + W(t) + \int_0^t [a q(\tau, a) + u(\tau, a)] \, d\tau$$

with probability 1.

For a given $a \in \mathbb{R}$, we refer to $q(t, a)$ and $u(t, a)$, respectively, as the *particle trajectory* and the *control variable* for a at time t.

We remark that any strategy for time horizon T and starting position q_0 gives rise to a strategy for time horizon T' and starting position q_0 for any $T' \in (0, T)$ (simply by restricting the time domain).

We will use σ to denote an arbitrary strategy. We then write q^σ and u^σ to denote the families of particle trajectories and control variables associated with σ.

For any strategy σ, we define a random variable

$$\mathrm{Cost}(\sigma, a) = \int_0^T \big((q^\sigma(t, a))^2 + (u^\sigma(t, a))^2 \big) \, dt.$$

This random variable is well-defined because, with probability 1, $q(t,a)$ is a continuous function of t and $u(t,a)$ is an L^2 function of t. We then define the *expected cost* of σ by

$$\mathcal{J}(\sigma, a; T, q_0) = \mathrm{E}[\mathrm{Cost}(\sigma, a)].$$

For any smooth function $v\colon [0,T] \to \mathbb{R}$, we define a strategy σ_v by setting

$$q^{\sigma_v}(t,a) = q_0 + W(t) + \int_0^t (a - v(\tau)) q^{\sigma_v}(\tau, a)\, d\tau, \tag{2.1}$$

$$u^{\sigma_v}(t,a) = -v(t) q^{\sigma_v}(t,a). \tag{2.2}$$

We refer to σ_v as a *simple feedback strategy* with *gain function* v.

For $\alpha \in \mathbb{R}$ and $s \geq 0$, we define

$$\kappa(s,\alpha) = \frac{\tanh(s\sqrt{\alpha^2+1})}{\sqrt{\alpha^2+1} - \alpha \tanh(s\sqrt{\alpha^2+1})}.$$

We let $\sigma_{\mathrm{opt}}(\alpha)$ denote the simple feedback strategy with gain function $t \mapsto \kappa(T-t,\alpha)$. In Section 3, we will show that for any $a \in \mathbb{R}$, the strategy $\sigma_{\mathrm{opt}}(a)$ minimizes the quantity $\mathcal{J}(\sigma, a; T, q_0)$ over all strategies σ. We therefore refer to the family of strategies $(\sigma_{\mathrm{opt}}(\alpha))_{\alpha\in\mathbb{R}}$ as *optimal known-a* (or just known-a) strategies. For ease of notation, we define

$$\mathcal{J}_0(a; T, q_0) = \mathcal{J}(\sigma_{\mathrm{opt}}(a), a; T, q_0);$$

we refer to $\mathcal{J}_0(a; T, q_0)$ as the *optimal expected cost for known a* (for time horizon T and starting position q_0).

Fix a real number C_0 and an integer $m_0 \geq 1$. We say that a strategy σ (for time horizon $\hat{T}$) is *A-bounded* for some $A > 0$ if

$$|u^\sigma(t,a)| \leq C_0 A^{m_0} [|q^\sigma(t,a)| + 1] \quad \text{for all } a \in \mathbb{R}, t \in [0,\hat{T}].$$

Recall that we assume that our starting position q_0 is nonzero.

Theorem 2.1. *Let $\varepsilon > 0$. Then for $A > 0$ sufficiently large depending on ε, T, q_0, C_0 and m_0, the following holds.*

Let σ be an A-bounded strategy for time horizon $T+\varepsilon$ and starting position q_0. Then the strategy σ_* for time horizon T and starting position q_0 specified in Section 5.2 satisfies the following.

(1) *If $a \in [-A, A]$, then*

$$\mathcal{J}(\sigma_*, a; T, q_0) < \varepsilon + (1+\varepsilon) \cdot \sup\{\mathcal{J}(\sigma, b; T+\varepsilon, q_0) : |a-b| < \varepsilon|a|\}.$$

(2) *If $|a| > A$, then*

$$\mathcal{J}(\sigma_*, a; T, q_0) < \varepsilon + (1+\varepsilon) \cdot \mathcal{J}_0(a; T, q_0).$$

We now state a corollary of Theorem 2.1; this is the variant of Theorem 1.2 for starting position $q_0 = 0$ mentioned in the introduction. The proof of this corollary is given in Section 5.4.

Corollary 2.2. *Let $\varepsilon > 0$. Then for $A > 0$ sufficiently large depending on ε, T, C_0 and m_0, the following holds.*

Let σ be an A-bounded strategy for time horizon $T + \varepsilon$ and starting position ε. Then the strategy $\hat{\sigma}_*$ for time horizon T and starting position 0 specified in Section 5.4 satisfies the following.

(1) *If $a \in [-A, A]$, then*

$$\mathcal{J}(\hat{\sigma}_*, a; T, 0) < \varepsilon + (1 + \varepsilon) \cdot \sup\{\mathcal{J}(\sigma, b; T + \varepsilon, \varepsilon) : |a - b| < \varepsilon|a|\}.$$

(2) *If $|a| > A$, then*

$$\mathcal{J}(\hat{\sigma}_*, a; T, 0) < \varepsilon + (1 + \varepsilon) \cdot \mathcal{J}_0(a; T, 0).$$

We remark that Corollary 2.2 implies a variant of Corollary 1.3 for $q_0 = 0$, but we do not state it here.

Our proof of Theorem 2.1 makes use of two additional strategies. These strategies, defined in Sections 6 and 7, respectively, are almost optimal when $|q_0|$ and a are large. Specifically, we prove the following theorems.

Theorem 2.3. *Let $\varepsilon > 0$. Then there exists $q_{\mathrm{big}} \geq 1$, depending (continuously) on ε and T, such that the following is true. For any $q_0 \geq q_{\mathrm{big}}$, let $\mathrm{L}q\mathrm{S}$ be the strategy for time horizon T and starting position q_0 defined in Section 6 depending on ε, T and q_0. Then*

$$\mathcal{J}(\mathrm{L}q\mathrm{S}, a; T, q_0) \leq (1 + \varepsilon) \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \in \mathbb{R}.$$

In our previous paper [8], we exhibited a strategy BR_0 satisfying

$$\mathcal{J}(\mathrm{BR}_0, a; T, 0) < C_T \cdot \mathcal{J}_0(a; T, 0) \quad \text{for any } a \in \mathbb{R}.$$

By making a simple modification to the strategy $\mathrm{L}q\mathrm{S}$ in Theorem 2.3, we can produce, for any $q_0 \in \mathbb{R}$, a strategy BR satisfying

$$\mathcal{J}(\mathrm{BR}; a, T, q_0) < C_{T,q_0} \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \in \mathbb{R};$$

we refer to this as a bounded regret strategy for starting position q_0.

Theorem 2.4. *Let $\varepsilon > 0$ and let $\mathrm{L}a\mathrm{S}$ be the strategy for time horizon T and starting position q_0 defined in Section 7 depending on ε, T and q_0. Then for any $A \geq A_{\min}(\varepsilon, T, q_0)$, we have*

$$\mathcal{J}(\mathrm{L}a\mathrm{S}, a; T, q_0) \leq (1 + \varepsilon) \cdot \mathcal{J}_0(a; T, q_0) \quad \textit{for any } a \geq A$$

and

$$\mathcal{J}(\mathrm{L}a\mathrm{S}, a; T, q_0) \leq C_{T,q_0} \cdot A^2 \quad \textit{for any } a \leq A.$$

3. Preliminary results on strategies

We begin this section by stating some properties of strategies.

Remark 3.1. By symmetry, any strategy σ for time horizon T and starting position q_0 gives rise to a strategy σ_- for time horizon T and starting position $-q_0$ defined by

$$q^{\sigma_-}(t,a) = -q^{\sigma}(t,a) \quad \text{and} \quad u^{\sigma_-}(t,a) = -u^{\sigma}(t,a),$$

and satisfying

$$\mathcal{J}(\sigma, a; T, q_0) = \mathcal{J}(\sigma_-, a; T, -q_0).$$

For the remainder of this paper we assume, without loss of generality, that $q_0 > 0$; this is justified by Remark 3.1.

Remark 3.2 (Rescaling property). Let σ be a strategy for time horizon T and starting position q_0 with particle trajectories $q^{\sigma}(t,a)$ and control variables $u^{\sigma}(t,a)$. Let $\lambda > 0$, and define the rescaled random variables

$$\tilde{q}(t,a) = \lambda \cdot q^{\sigma}\Big(\frac{t}{\lambda^2}, \lambda^2 a\Big), \quad \tilde{u}(t,a) = \frac{1}{\lambda} \cdot u^{\sigma}\Big(\frac{t}{\lambda^2}, \lambda^2 a\Big), \quad \tilde{W}(t) = \lambda \cdot W\Big(\frac{t}{\lambda^2}\Big).$$

Note that $\tilde{W}(t)$ is a standard Brownian motion and that for any $a \in \mathbb{R}$, $t \in [0, \lambda^2 T]$, we have

$$\tilde{q}(t,a) = \lambda q_0 + \tilde{W}(t) + \int_0^t [a\,\tilde{q}(\tau,a) + \tilde{u}(\tau,a)]\,d\tau$$

with probability 1. It follows that $\tilde{q}(t,a)$ and $\tilde{u}(t,a)$ determine a strategy for time horizon $\lambda^2 T$ and starting position λq_0. Denoting this strategy by σ_λ, we have

$$\begin{aligned}
\mathcal{J}(\sigma_\lambda, a; \lambda^2 T, \lambda q_0) &= \mathrm{E}\Big[\int_0^{\lambda^2 T} \big(\tilde{q}^2(t,a) + \tilde{u}^2(t,a)\big)\,dt\Big] \\
&= \mathrm{E}\Big[\int_0^{T} \big(\lambda^4 \cdot (q^{\sigma}(t,\lambda^2 a))^2 + (u^{\sigma}(t,\lambda^2 a))^2\big)\,dt\Big] \le \max\{\lambda^4, 1\} \cdot \mathcal{J}(\sigma, \lambda^2 a; T, q_0).
\end{aligned}$$

We say that σ_λ is a *rescaling* of the strategy σ. Note that the rescaled strategy σ_1 is equal to the original strategy σ.

3.1. Branching strategies

We will often decide to switch from one strategy to another.

To explain how we do that, recall that we denote strategies by σ, σ', $\hat{\sigma}$, etc. We introduce the notion of a *parametrized strategy*, denoted $\sigma(\cdot)$. For each parameter value $\alpha \in \mathbb{R}^d$ (some $d \ge 0$), $\sigma(\alpha)$ is a strategy. (When $d = 0$, a parametrized strategy is just a strategy.) For instance, the strategy $\sigma_{\mathrm{opt}}(\beta)$ for time horizon T and starting position q_0 is a parametrized strategy; β, T and q_0 are the parameters. We remark that the parameter α will often include our guess for the unknown a.

Now suppose we are given parametrized strategies $\sigma_0(\cdot), \sigma_1(\cdot), \ldots, \sigma_N(\cdot)$. We will combine the σ's into a new parametrized strategy $\sigma^{\#}$ (a *branching strategy*). Suppose we are given a parameter value α and that $\sigma_0(\alpha)$ is a strategy for some time horizon T and starting position q_0. The strategy $\sigma^\#(\alpha)$ is then also a strategy for time horizon T and starting position q_0.

For each $a \in \mathbb{R}$, we pick a stopping time $\tau(a,\alpha) \le T$ depending on α and satisfying the following: for any $a, b \in \mathbb{R}$, $t \in [0,T]$, and $\omega \in \Omega$, if $q^{\sigma_0(\alpha)}(s,a,\omega) = q^{\sigma_0(\alpha)}(s,b,\omega)$ for all $s \in [0,t]$, then $\tau(a,\alpha,\omega) > t$ if and only if $\tau(b,\alpha,\omega) > t$. Note that this condition ensures that the stopping times are defined only in terms of q and cannot use knowledge of the unknown a. The strategy $\sigma^\#(\alpha)$ proceeds as follows.

Until time $\tau(a,\alpha)$, we execute the strategy $\sigma_0(\alpha)$. If $\tau(a,\alpha) = T$, then we are done.

If $\tau(a,\alpha) < T$, then we will pick a new parametrized strategy $\hat{\sigma}(\cdot)$ and a new parameter $\hat{\alpha}$. The $\hat{\sigma}(\cdot)$ will be one of our given parametrized strategies $\sigma_1(\cdot), \dots, \sigma_N(\cdot)$. Our choice of $\hat{\sigma}(\cdot)$ and $\hat{\alpha}$ is determined by α, the history up to time $\tau(a,\alpha)$, and the requirement that $\hat{\sigma}(\hat{\alpha})$ is a strategy for starting position $\hat{q}_0 = q(\tau(a,\alpha))$ and some time horizon $\hat{T} \ge (T - \tau(a,\alpha))$.

Once we have picked $\hat{\sigma}(\cdot)$ and $\hat{\alpha}$, we forget the past, regard $t = \tau(a,\alpha)$ as if we were at $t = 0$, and execute the strategy $\hat{\sigma}(\hat{\alpha})$, starting at position $\hat{q}_0$. We stop playing at time T.

Thus, we have combined our parametrized strategies $\sigma_0(\cdot), \sigma_1(\cdot), \dots, \sigma_N(\cdot)$ into a branching (parametrized) strategy $\sigma^\#(\cdot)$.

We may then combine our $\sigma^\#(\cdot)$ with additional parametrized strategies to form further branching strategies.

We note that the strategies σ_* and $\tilde{\sigma}$ constructed in Section 5, the strategy LqS constructed in Section 6, and the strategy LaS constructed in Section 7 are all examples of branching strategies.

3.2. Optimal known-*a* strategies

Let $a \in \mathbb{R}$. In Section 2, we defined the optimal expected cost for known a by

$$\mathcal{J}_0(a; T, q_0) = \mathcal{J}(\sigma_{\mathrm{opt}}(a), a; T, q_0),$$

where $\sigma_{\mathrm{opt}}(a)$ is the simple feedback strategy with gain function

$$\kappa(T-t,\alpha) = \frac{\tanh((T-t)\sqrt{\alpha^2+1})}{\sqrt{\alpha^2+1} - \alpha\tanh((T-t)\sqrt{\alpha^2+1})}; \tag{3.1}$$

recall that we refer to $\sigma_{\mathrm{opt}}(a)$ as an optimal known-a strategy. Observe that, for fixed α, the function κ solves a Riccati equation:

$$\frac{d\kappa}{dt}(t,\alpha) = [1 + 2a\kappa(t,\alpha) - \kappa^2(t,\alpha)].$$

For $t \in [0,T]$, we define

$$K(t,\alpha) = \int_0^t \kappa(s,\alpha)\,ds. \tag{3.2}$$

Clearly, we have (again for fixed α)

$$\frac{dK}{dt}(t,\alpha) = \kappa(t,\alpha).$$

It is well known (see, for example, [3]) that

$$\mathcal{J}_0(a;T,q_0) = \kappa(T,a)\cdot q_0^2 + K(T,a). \tag{3.3}$$

This implies the following remark.

Remark 3.3. Let $0 < q_0' < q_0''$. Then

$$\mathcal{J}_0(a;T,q_0') < \mathcal{J}_0(a;T,q_0'') < \left(\frac{q_0''}{q_0'}\right)^2 \mathcal{J}_0(a;T,q_0') \quad \text{for any } a \in \mathbb{R}.$$

Combining (3.1) and (3.2) and evaluating the resulting integral, we get

$$\begin{aligned} K(T,a) &= \big(a + \sqrt{a^2+1}\big)T + \log\left(\frac{\sqrt{a^2+1}-a}{2\sqrt{a^2+1}}\right) \\ &\quad + \log\left(1 + e^{-2T\sqrt{a^2+1}}\left(\frac{\sqrt{a^2+1}+a}{\sqrt{a^2+1}-a}\right)\right). \end{aligned}$$

Note that for any $\varepsilon > 0$, there exists $A > 0$, depending on ε and the time horizon T, such that:

- For $a > A$, we have

$$|\kappa(T,a) - 2a| < \varepsilon a, \tag{3.4}$$
$$|K(T,a) - 2aT| < \varepsilon aT. \tag{3.5}$$

- For $a < -A$, we have

$$\left|\kappa(T,a) - \frac{1}{2|a|}\right| \le \frac{\varepsilon}{|a|}, \tag{3.6}$$
$$\left|K(T,a) - \frac{T}{2|a|}\right| \le \frac{\varepsilon T}{|a|}. \tag{3.7}$$

From (3.3)–(3.7), we deduce that for any $\varepsilon > 0$, there exists $A > 0$, depending on ε and T, such that

$$|\mathcal{J}_0(a;T,q_0) - 2a(q_0^2+T)| < \varepsilon a(q_0^2+T) \quad \text{when } a > A, \text{ and} \tag{3.8}$$
$$\left|\mathcal{J}_0(a;T,q_0) - \frac{1}{2|a|}(q_0^2+T)\right| < \frac{\varepsilon}{|a|}(q_0^2+T) \quad \text{when } a < -A. \tag{3.9}$$

Now let $\varepsilon > 0$ be arbitrary and introduce $\delta > 0$ sufficiently small depending on ε, T, q_0. Suppose that $T' > 0$ and $q_0' \in \mathbb{R}$ satisfy $|T - T'|, |q_0 - q_0'| < \delta$. We claim that

$$|\mathcal{J}_0(a';T',q_0') - \mathcal{J}_0(a;T,q_0)| < \varepsilon \cdot \mathcal{J}_0(a;T,q_0) \quad \text{for any } a \in \mathbb{R}, |a-a'| < |a|\delta. \tag{3.10}$$

By (3.8) and (3.9) above, there exists $A > 0$, depending on ε and T, such that:

(1) For any $a, a' > A$, we have

$$|\mathcal{J}_0(a';T',q_0') - 2a'((q_0')^2 + T')| < \varepsilon a'((q_0')^2 + T'), \tag{3.11}$$
$$|\mathcal{J}_0(a;T,q_0) - 2a(q_0^2+T)| < \varepsilon a(q_0^2+T). \tag{3.12}$$

(2) For any $a, a' < -A$, we have

$$\left|\mathcal{J}_0(a'; T', q_0') - \frac{1}{2|a'|}((q_0')^2 + T')\right| < \frac{\varepsilon}{|a'|}((q_0')^2 + T'), \tag{3.13}$$

$$\left|\mathcal{J}_0(a; T, q_0) - \frac{1}{2|a|}(q_0^2 + T)\right| < \frac{\varepsilon}{|a|}(q_0^2 + T). \tag{3.14}$$

Combining (3.11) and (3.12), and using the assumptions $|T - T'| < \delta$ and $|q_0 - q_0'| < \delta$, we get

$$|\mathcal{J}_0(a'; T', q_0') - \mathcal{J}_0(a; T, q_0)| < C_{T,q_0} \cdot a \cdot (\varepsilon + \delta) \quad \text{for any } a \geq 2A,\ |a - a'| < \delta a.$$

Taking δ sufficiently small depending on ε, we use (3.12) to deduce that (for A sufficiently large depending on ε and T) we have

$$|\mathcal{J}_0(a'; T', q_0') - \mathcal{J}_0(a; T, q_0)| < C'_{T,q_0} \cdot \varepsilon \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \geq 2A, |a - a'| < \delta a. \tag{3.15}$$

Similarly, we use (3.13) and (3.14) to deduce that

$$|\mathcal{J}_0(a'; T', q_0') - \mathcal{J}_0(a; T, q_0)| < C'_{T,q_0} \cdot \varepsilon \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \leq -2A, |a - a'| < \delta |a|. \tag{3.16}$$

Note that (3.15) and (3.16) imply that for a sufficiently large number $\tilde{A} > A$ depending on ε, T and q_0, we have

$$|\mathcal{J}_0(a'; T', q_0') - \mathcal{J}_0(a; T, q_0)| < \varepsilon \cdot \mathcal{J}(a; T, q_0) \quad \text{for any } |a| \geq 2\tilde{A}, |a - a'| < \delta |a|. \tag{3.17}$$

Next, we note that (3.3) implies that $\mathcal{J}_0(a; T, q_0)$ is of the form

$$\mathcal{J}_0(a; T, q_0) = f_1(a, T)\, q_0^2 + f_2(a, T),$$

for smooth functions $f_1, f_2 \colon \mathbb{R} \times (0, \infty) \to (0, \infty)$ independent of q_0. Therefore (since $\tilde{A}$ is determined by ε, T and q_0), we have

$$|\mathcal{J}_0(a'; T', q_0') - \mathcal{J}_0(a; T, q_0)| < C_{\varepsilon,T,q_0} \cdot \delta \quad \text{for any } |a|, |a'| \leq 3\tilde{A} \text{ and } |a - a'| < \delta |a|$$

and

$$\mathcal{J}_0(a; T, q_0) > c_{\varepsilon,T,q_0} \quad \text{for any } |a| \leq 3\tilde{A}.$$

Combining the last two inequalities gives

$$|\mathcal{J}_0(a'; T', q_0') - \mathcal{J}_0(a; T, q_0)| < C_{\varepsilon,T,q_0} \cdot \delta \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } |a| \leq 2\tilde{A}, |a - a'| < \delta |a|. \tag{3.18}$$

Combining (3.15), (3.16), (3.18), and taking δ sufficiently small depending on ε, T and q_0 proves (3.10).

We summarize the above discussion (specifically (3.3), (3.8), (3.9), and (3.10)) as a lemma.

Lemma 3.4. *The optimal expected cost for known a, $\mathcal{J}_0$, has the following properties*:

(1) *We have*

$$\mathcal{J}_0(a;T,q_0) = \kappa(T,a)\cdot q_0^2 + K(T,a).$$

(2) *For any $\varepsilon > 0$, there exists $A > 0$, depending on ε and T, such that*

$$|\mathcal{J}_0(a;T,q_0) - 2a(q_0^2+T)| < \varepsilon a(q_0^2+T) \quad \textit{when } a > A, \textit{ and} \tag{3.19}$$

$$\left|\mathcal{J}_0(a;T,q_0) - \frac{1}{2|a|}(q_0^2+T)\right| < \frac{\varepsilon}{|a|}(q_0^2+T) \quad \textit{when } a < -A. \tag{3.20}$$

(3) *Let $\varepsilon > 0$. Let $\delta > 0$ be sufficiently small depending on ε, T and q_0, and suppose that $T' > 0$ and $q_0' \in \mathbb{R}$ satisfy $|T-T'| < \delta$ and $|q_0 - q_0'| < \delta$. Then, for any $a \in \mathbb{R}$,*

$$|\mathcal{J}_0(a';T',q_0') - \mathcal{J}_0(a;T,q_0)| < \varepsilon\cdot\mathcal{J}_0(a;T,q_0) \quad \textit{for any } |a-a'| < |a|\delta.$$

The next lemma says that the strategy $\sigma_{\text{opt}}(a)$ is indeed optimal for known a, i.e., for fixed $a \in \mathbb{R}$, the strategy $\sigma_{\text{opt}}(a)$ minimizes the quantity $\mathcal{J}(\sigma,a;T,q_0)$.

Lemma 3.5. *Let σ be an arbitrary strategy for time horizon T and starting position q_0. Then*

$$\mathcal{J}(\sigma,a;T,q_0) \geq \mathcal{J}_0(a;T,q_0) \quad \textit{for any } a \in \mathbb{R}.$$

Proof. Let $N \gg 1$ be a sufficiently large integer depending on a and T; we shall write $c, c', C, C', \ldots$ to denote constants determined by a and T. The symbols $c, c', C, C', \ldots$ may denote different constants in different occurrences.

We set $h = T/N$, and for $\nu = 0,\ldots,N$, we set $t_\nu = \nu h$.

By downward induction on ν, we will show that

$$\begin{aligned}\mathrm{E}\Big[\int_{t_\nu}^{T}((q(s))^2+(u(s))^2)\,ds\,\Big|\mathcal{F}_{t_\nu}\Big]&\cdot(1+h^{1/100})\\ &\geq \kappa(T-t_\nu,a)\,(q(t_\nu))^2 + K(T-t_\nu,a) - h^{1/100}(T-t_\nu).\end{aligned}\tag{3.21}$$

Once we prove (3.21), we take $\nu = 0$ and let $N\to\infty$ to derive the conclusion of the lemma. So our task is to prove (3.21). We begin our induction on ν.

In the base case $\nu = N$, (3.21) holds since $\kappa(0,a) = K(0,a) = 0$.

For the induction step, we fix ν ($0 \leq \nu < N$), and assume the inductive hypothesis

$$\begin{aligned}\mathrm{E}\Big[\int_{t_{\nu+1}}^{T}((q(s))^2+(u(s))^2)\,ds\Big|\mathcal{F}_{t_{\nu+1}}\Big]&\cdot(1+h^{1/100})\\ \geq \kappa(T-t_{\nu+1},a)\cdot(q(t_{\nu+1}))^2 &+ K(T-t_{\nu+1},a) - h^{1/100}(T-t_{\nu+1}).\end{aligned}\tag{3.22}$$

Our goal is then to prove (3.21) assuming (3.22).

Recall that our Brownian motion is denoted by $(W(t))_{t\geq 0}$. We set

$$\Delta W_\nu = W(t_{\nu+1}) - W(t_\nu), \tag{3.23}$$

$$\omega(\nu) = \sup\{|W(t)-W(t_\nu)| : t \in [t_\nu, t_{\nu+1}]\}, \tag{3.24}$$

$$\|u\|_\nu = \Big(\int_{t_\nu}^{t_{\nu+1}} u^2\,ds\Big)^{1/2}, \quad \bar{u}_\nu = \frac{1}{h}\int_{t_\nu}^{t_{\nu+1}} u\,ds. \tag{3.25}$$

For the rest of the proof, we condition on $\mathcal{F}_{t_\nu}$, and we write $\mathrm{E}[\cdots]$ to denote the expectation conditioned on $\mathcal{F}_{t_\nu}$.

We have

$$\mathrm{E}[\Delta W_\nu] = 0, \quad \mathrm{E}[(\Delta W_\nu)^2] = h, \quad \mathrm{E}[(\omega(\nu))^2] \le Ch, \tag{3.26}$$

and

$$h|\bar{u}_\nu|^2 \le \|u\|_\nu^2; \tag{3.27}$$

more generally,

$$\left| \int_{t_\nu}^{t} u \, ds \right| \le h^{1/2} \|u\|_\nu \quad \text{for } t \in [t_\nu, t_{\nu+1}]. \tag{3.28}$$

By the definition of a strategy, we have

$$\begin{aligned} q(t) - q(t_\nu) = [W(t) - W(t_\nu)] + aq(t_\nu)(t - t_\nu) + \int_{t_\nu}^{t} a[q(s) - q(t_\nu)] \, ds \\ + \int_{t_\nu}^{t} u \, ds, \quad \text{for } t \in [t_\nu, t_{\nu+1}]. \end{aligned} \tag{3.29}$$

Setting

$$\operatorname{osc}(\nu) = \sup\{|q(t) - q(t_\nu)| : t \in [t_\nu, t_{\nu+1}]\}, \tag{3.30}$$

we deduce that

$$\operatorname{osc}(\nu) \le \omega(\nu) + |aq(t_\nu)|h + |a| \operatorname{osc}(\nu) \cdot h + h^{1/2} \|u\|_\nu.$$

Since h is less than a small enough constant determined by a, we may absorb the term $|a|\operatorname{osc}(\nu)h$ into the left-hand side above, to conclude that

$$\operatorname{osc}(\nu) \le C\omega(\nu) + Ch \cdot |aq(t_\nu)| + Ch^{1/2} \|u\|_\nu. \tag{3.31}$$

From (3.30) and (3.31), we see that

$$|q(t) - q(t_\nu)| \le C\omega(\nu) + Ch|q(t_\nu)| + Ch^{1/2} \|u\|_\nu \quad \text{for } t \in [t_\nu, t_{\nu+1}],$$

hence

$$\int_{t_\nu}^{t_{\nu+1}} q^2 \, ds \ge (1 - Ch^{1/10})(q(t_\nu))^2 h - Ch^{9/10}(\omega(\nu))^2 - Ch^{19/10} \|u\|_\nu^2. \tag{3.32}$$

Also, (3.29), (3.30) and (3.31) yield

$$q(t_{\nu+1}) = (1 + ah)q(t_\nu) + \Delta W_\nu + \bar{u}_\nu h + \mathrm{ERR}(\nu), \quad \text{with} \tag{3.33}$$

$$|\mathrm{ERR}(\nu)| \le Ch\omega(\nu) + Ch^2|q(t_\nu)| + Ch^{3/2} \|u\|_\nu. \tag{3.34}$$

Using (3.33) and (3.34), we estimate

$$\begin{aligned} (q(t_{\nu+1}))^2 = \big[(1 + ah)q(t_\nu) + \Delta W_\nu + \bar{u}_\nu h]^2 + (\mathrm{ERR}(\nu))^2 \\ + 2[(1 + ah)q(t_\nu) + \Delta W_\nu + \bar{u}_\nu h\big] \cdot (\mathrm{ERR}(\nu)). \end{aligned} \tag{3.35}$$

We have

$$
\begin{aligned}
&|[(1+ah)q(t_\nu)+\Delta W_\nu+\bar{u}_\nu h]\cdot(\mathrm{ERR}(\nu))| \\
&\qquad \le C\big[|q(t_\nu)|+\omega(\nu)+h^{1/2}\|u\|_\nu\big]\cdot\big[h\omega(\nu)+h^2|q(t_\nu)|+Ch^{3/2}\|u\|_\nu\big] \\
&\qquad \le Ch^2|q(t_\nu)|^2+Ch(\omega(\nu))^2+Ch^2\|u\|_\nu^2+Ch^{3/2}|q(t_\nu)|\cdot\|u\|_\nu \\
&\qquad\quad +Ch|q(t_\nu)|\omega(\nu)+Ch^{3/2}\|u\|_\nu\omega(\nu).
\end{aligned}
\tag{3.36}
$$

Into (3.36) we substitute the estimates

$$
\begin{aligned}
h^{3/2}|q(t_\nu)|\cdot\|u\|_\nu &\le Ch^{3/2}|q(t_\nu)|^2+Ch^{3/2}\|u\|_\nu^2,\\
h|q(t_\nu)|\omega(\nu) &\le Ch^{5/4}|q(t_\nu)|^2+Ch^{3/4}(\omega(\nu))^2, \quad \text{and}\\
h^{3/2}\|u\|_\nu\omega(\nu) &\le Ch^{3/2}\|u\|_\nu^2+Ch^{3/2}(\omega(\nu))^2.
\end{aligned}
$$

We find that

$$
\big|[(1+ah)q(t_\nu)+\Delta W_\nu+\bar{u}_\nu h]\cdot\mathrm{ERR}(\nu)\big|\le Ch^{5/4}|q(t_\nu)|^2+Ch^{3/2}\|u\|_\nu^2+Ch^{3/4}(\omega(\nu))^2.
$$

Consequently, (3.35) implies the estimate

$$
\begin{aligned}
(q(t_{\nu+1}))^2 &\ge [(1+ah)q(t_\nu)+\Delta W_\nu+\bar{u}_\nu h]^2-Ch^{5/4}|q(t_\nu)|^2\\
&\quad -Ch^{3/2}\|u\|_\nu^2-Ch^{3/4}(\omega(\nu))^2\\
&=\{(1+ah)^2-Ch^{5/4}\}(q(t_\nu))^2+(\Delta W_\nu)^2+\bar{u}_\nu^2h^2\\
&\quad +2(1+ah)q(t_\nu)\Delta W_\nu+2(1+ah)q(t_\nu)\bar{u}_\nu h+2\Delta W_\nu\bar{u}_\nu h\\
&\quad -Ch^{3/2}\|u\|_\nu^2-Ch^{3/4}(\omega(\nu))^2.
\end{aligned}
\tag{3.37}
$$

Since

$$
\begin{aligned}
|\Delta W_\nu\bar{u}_\nu h| &\le |\Delta W_\nu|\cdot h^{1/2}\|u\|_\nu \le Ch^{1/2}|\Delta W_\nu|^2+Ch^{1/2}\|u\|_\nu^2\\
&\le Ch^{1/2}(\omega(\nu))^2+Ch^{1/2}\|u\|_\nu^2,
\end{aligned}
$$

estimate (3.37) implies the following:

$$
\begin{aligned}
(q(t_{\nu+1}))^2 &\ge \{(1+ah)^2-Ch^{5/4}\}(q(t_\nu))^2+(\Delta W_\nu)^2\\
&\quad +2(1+ah)q(t_\nu)\Delta W_\nu+2(1+ah)q(t_\nu)\bar{u}_\nu h-Ch^{1/2}\|u\|_\nu^2\\
&\quad -Ch^{1/2}(\omega(\nu))^2.
\end{aligned}
\tag{3.38}
$$

Recall that we are conditioning on $\mathcal{F}_{t_\nu}$, so that $q(t_\nu)$ is deterministic. From (3.26) and (3.38), we therefore learn that

$$
\begin{aligned}
\mathrm{E}[(q(t_{\nu+1}))^2] &\ge (1+2ah-Ch^{5/4})(q(t_\nu))^2+h+2h(1+ah)q(t_\nu)\mathrm{E}[\bar{u}_\nu]\\
&\quad -Ch^{1/2}\mathrm{E}[\|u\|_\nu^2]-Ch^{3/2},
\end{aligned}
\tag{3.39}
$$

while (3.26) and (3.32) yield

$$
\mathrm{E}\Big[\int_{t_\nu}^{t_{\nu+1}}q^2\,ds\Big]\ge h(1-Ch^{1/10})(q(t_\nu))^2-Ch^{19/10}\mathrm{E}[\|u\|_\nu^2]-Ch^{19/10}.
\tag{3.40}
$$

We bring the inductive assumption (3.22) into play. Thanks to (3.22), (3.39) and (3.40), we have

$$
\begin{aligned}
&\mathrm{E}\Big[\int_{t_\nu}^{T}(q^2+u^2)\,ds\Big]\cdot(1+h^{1/100})\\
&\quad= \mathrm{E}\Big[\int_{t_\nu}^{t_{\nu+1}} q^2\,ds\Big]\cdot(1+h^{1/100}) + \mathrm{E}\Big[\int_{t_\nu}^{t_{\nu+1}} u^2\,ds\Big]\cdot(1+h^{1/100})\\
&\qquad + \mathrm{E}\Big[\mathrm{E}\Big[\int_{t_{\nu+1}}^{T}(q^2+u^2)\,ds\Big|\mathcal{F}_{t_{\nu+1}}\Big]\cdot(1+h^{1/100})\Big]\\
&\quad\ge \{h(1-Ch^{1/10})(1+h^{1/100})(q(t_\nu))^2 - Ch^{19/10}\,\mathrm{E}[\|u\|_\nu^2] - Ch^{19/10}\}\\
&\qquad + \{\mathrm{E}[\|u\|_\nu^2](1+h^{1/100})\}\\
&\qquad + \{\kappa(T-t_{\nu+1},a)\,\mathrm{E}[(q(t_{\nu+1}))^2] + K(T-t_{\nu+1},a) - h^{1/100}(T-t_{\nu+1})\}\\
&\quad\ge h\Big(1+\frac{1}{2}h^{1/100}\Big)(q(t_\nu))^2 + \mathrm{E}\Big[\|u\|_\nu^2\Big(1+\frac{1}{2}h^{1/100}\Big)\Big] - Ch^{19/10}\\
&\qquad + \kappa(T-t_{\nu+1},a)\Big\{(1+2ah-Ch^{5/4})(q(t_\nu))^2 + h + 2h(1+ah)q(t_\nu)\,\mathrm{E}[\bar{u}_\nu]\\
&\qquad - Ch^{1/2}\,\mathrm{E}[\|u\|_\nu^2] - Ch^{3/2}\Big\} + K(T-t_{\nu+1},a) - h^{1/100}(T-t_{\nu+1}),
\end{aligned}
$$

and consequently,

$$
\begin{aligned}
&\mathrm{E}\Big[\int_{t_\nu}^{T}(q^2+u^2)\,ds\Big](1+h^{1/100})\\
&\qquad\ge \Big\{\kappa(T-t_{\nu+1},a) + h[2a\kappa(T-t_{\nu+1},a)+1] + \frac{1}{4}h^{101/100}\Big\}(q(t_\nu))^2\\
&\qquad\quad + \mathrm{E}[\|u\|_\nu^2 + 2h(1+ah)\kappa(T-t_{\nu+1},a)q(t_\nu)\bar{u}_\nu]\\
&\qquad\quad + \{K(T-t_{\nu+1},a) - h^{1/100}(T-t_{\nu+1}) + \kappa(T-t_{\nu+1},a)h - Ch^{3/2}\}
\end{aligned}
\tag{3.41}
$$

Now, recalling (3.27), we see that

$$
\begin{aligned}
&\|u\|_\nu^2 + 2h(1+ah)\kappa(T-t_{\nu+1},a)q(t_\nu)\bar{u}_\nu\\
&\quad\ge \bar{u}_\nu^2 h + 2h(1+ah)\kappa(T-t_{\nu+1},a)q(t_\nu)\bar{u}_\nu\\
&\quad= h(\bar{u}_\nu + (1+ah)\kappa(T-t_{\nu+1},a)q(t_\nu))^2 - h(1+ah)^2\kappa^2(T-t_{\nu+1},a)(q(t_\nu))^2\\
&\quad\ge -h(1+ah)^2\kappa^2(T-t_{\nu+1},a)(q(t_\nu))^2.
\end{aligned}
$$

Therefore, from (3.41), we have

$$
\begin{aligned}
&\mathrm{E}\Big[\int_{t_\nu}^{T}(q^2+u^2)\,ds\Big]\cdot(1+h^{1/100})\\
&\qquad\ge \Big\{\kappa(T-t_{\nu+1},a) + h[1+2a\kappa(T-t_{\nu+1},a)\\
&\qquad\quad - (1+ah)^2\kappa^2(T-t_{\nu+1},a)] + \frac{1}{4}h^{101/100}\Big\}(q(t_\nu))^2\\
&\qquad\quad + \{K(T-t_{\nu+1},a) - h^{1/100}(T-t_{\nu+1}) + \kappa(T-t_{\nu+1},a)h - Ch^{3/2}\}.
\end{aligned}
\tag{3.42}
$$

Recall that

$$\frac{d}{dt}[\kappa(T-t,a)] = -[1 + 2a\kappa(T-t,a) - \kappa^2(T-t,a)]$$

and that

$$\frac{d}{dt}[K(T-t,a)] = -\kappa(T-t,a).$$

Therefore, the expressions in curly brackets on the right in (3.42) are bounded below, respectively, by $\kappa(T-t_\nu,a)$ and by

$$K(T-t_\nu) - h^{1/100}(T-t_{\nu+1}) - Ch^{3/2}.$$

Consequently, (3.42) implies the estimate

$$\begin{aligned}
&\mathrm{E}\Big[\int_{t_\nu}^{T}(q^2+u^2)\,ds\Big]\cdot(1+h^{1/100}) \\
&\qquad \geq \kappa(T-t_\nu,a)\cdot(q(t_\nu))^2 + K(T-t_\nu,a) - h^{1/100}(T-t_{\nu+1}) - Ch^{3/2} \\
&\qquad \geq \kappa(T-t_\nu,a)\cdot(q(t_\nu))^2 + K(T-t_\nu,a) - h^{1/100}(T-t_\nu).
\end{aligned} \tag{3.43}$$

Recalling that $\mathrm{E}[\cdots]$ here denotes expectation conditioned on $\mathcal{F}_{t_\nu}$, we see that (3.43) is precisely our desired inequality (3.21).

This completes our downward induction on ν, proving the lemma. ∎

3.3. The expected cost of simple feedback strategies

Let $v\colon[0,T]\to\mathbb{R}$ be a smooth function, and let σ_v denote the simple feedback strategy for time horizon T and starting position q_0 with gain function $v\colon[0,T]\to\mathbb{R}$ (see Section 2 for the definition of a simple feedback strategy). We fix some $a\in\mathbb{R}$ and let q denote the particle trajectory $q^{\sigma_v}(\cdot,a)$. Note that q solves the stochastic ODE

$$dq = (a-v)q\,dt + dW, \quad q(0)=q_0, \tag{3.44}$$

and that

$$\mathcal{J}(\sigma_v,a;T,q_0) = \mathrm{E}\Big[\int_0^T q^2(\tau)(1+v^2(\tau))\,d\tau\Big]. \tag{3.45}$$

Define a (smooth) function $\phi\colon[0,T]\to\mathbb{R}$ by

$$\phi(t) = \int_t^T (1+v^2(\tau))\exp\Big(2\int_t^\tau\big(a-v(s)\big)\,ds\Big)\,d\tau;$$

the function ϕ solves the ODE

$$-\phi'(t) = 2\phi(t)(a-v(t)) + 1 + v^2(t), \quad \phi(T)=0. \tag{3.46}$$

By Itô's lemma, we have

$$d(q^2\phi) = [\phi' q^2 + \phi]\,dt + (2q\phi)\,dq.$$

Combining this with (3.44) and (3.46) gives

$$d(q^2\phi) = [-(1+v^2)q^2 + \phi]\,dt + (2\phi q)\,dW. \tag{3.47}$$

With probability 1, we have

$$\int_0^T d(q^2\phi) = q^2(T)\,\phi(T) - q^2(0)\,\phi(0) = -q_0^2\,\phi(0);$$

combining this with (3.47) gives

$$\int_0^T q^2(\tau)(1+v^2(\tau))\,d\tau = \phi(0)\,q_0^2 + \int_0^T \phi(\tau)\,d\tau + \int_0^T 2\phi(\tau)q(\tau)\,dW_\tau. \tag{3.48}$$

Taking expectation and using (3.45), we get

$$\mathcal{J}(\sigma_v, a; T, q_0) = \phi(0)\,q_0^2 + \int_0^T \phi(\tau)\,d\tau.$$

We summarize the above result as a lemma.

Lemma 3.6. *Let $v\colon [0,T] \to \mathbb{R}$ be a smooth function, and let σ_v be the simple feedback strategy with gain function v. Then*

$$\mathcal{J}(\sigma_v, a; T, q_0) = \phi(0)\cdot q_0^2 + \int_0^T \phi(s)\,ds,$$

where

$$\phi(t) = \int_t^T (1+v^2(s))\exp\Big(2\int_t^s (a - v(r))\,dr\Big)\,ds.$$

Let $\alpha \geq 0$ be a real number. We define the *constant gain strategy* $\mathrm{CG}(\alpha)$ to be the simple feedback strategy with the constant gain function $v(t) \equiv \alpha$. We remark that $\mathrm{CG}(\alpha)$ is independent of the time horizon T and the starting position q_0.

Lemma 3.6 then implies that for any $a \in \mathbb{R}$, we have

$$\begin{aligned}\mathcal{J}(\mathrm{CG}(\alpha), a; T, q_0) = (1+\alpha^2)\Big[\frac{1}{2(a-\alpha)}\Big[\Big(\frac{e^{2T(a-\alpha)}-1}{2(a-\alpha)}\Big) - T\Big]\\ + \Big(\frac{e^{2T(a-\alpha)}-1}{2(a-\alpha)}\Big)q_0^2\Big] \quad \text{for any } \alpha \geq 0, \alpha \neq a,\end{aligned} \tag{3.49}$$

and for any $a \geq 0$, we have

$$\mathcal{J}(\mathrm{CG}(a), a; T, q_0) = (1+a^2)\,T\Big(q_0^2 + \frac{1}{2}T\Big). \tag{3.50}$$

Using (3.49) and (3.50), we deduce the following corollary.

Corollary 3.7. *For $\alpha \geq 0$ and $a \in \mathbb{R}$ we have*

$$\mathcal{J}(\mathrm{CG}(\alpha), a; T, q_0) \leq \begin{cases} C_T(1+\alpha^2)(1+q_0^2)\,e^{2T(a-\alpha)} & \text{when } (a-\alpha) > 1/T, \\ C_T(1+\alpha^2)(1+q_0^2) & \text{when } |a-\alpha| \leq 1/T, \\ \frac{(1+\alpha^2)}{2|a-\alpha|}\,(q_0^2 + T) & \text{when } (a-\alpha) < -1/T. \end{cases}$$

For any $\alpha \in \mathbb{R}$, the strategy $\sigma_{\mathrm{opt}}(\alpha)$ is a simple feedback strategy with gain function $\kappa(T-t, \alpha)$. Note that this strategy depends on T, but is independent of q_0. By Lemma 3.4, and because $\kappa(T-t, a) \leq C\max\{1, a\}$ for all $a \in \mathbb{R}$ and $t \in [0,T]$, we have $\mathcal{J}_0(a; T, q_0) \leq C\max\{1, a\}(q_0^2 + T)$. We combine this with Lemma 3.6 to deduce the following corollary.

Corollary 3.8. *For any $a \in \mathbb{R}$, the following holds*:

$$\mathcal{J}(\sigma_{\mathrm{opt}}(\alpha), a; T, q_0) \leq \begin{cases} C \max\{1, a\}(q_0^2 + T) & \textit{if } \alpha = a, \\ C_T(1+\alpha^2)\, e^{2|a|T}(1+q_0^2) & \textit{for any } \alpha \in \mathbb{R}. \end{cases}$$

We remark that the first upper bound in Corollary 3.8 (for $\alpha = a$) is sharp when a is large. The second upper bound is far from sharp unless $a \gg \alpha \gg 1$.

4. The uncontrolled system

Let $T > 0$, $q_0 > 0$, and let $a \in \mathbb{R}$ be arbitrary. In this section, we consider the dynamics

$$dq = (aq)\, dt + dW_t, \quad q(0) = q_0. \tag{4.1}$$

We define a random process

$$X_t = e^{-at} q(t) - q_0 = \int_0^t e^{-as}\, dW_s. \tag{4.2}$$

If $a = 0$, then X_t is standard Brownian motion. If $a \neq 0$, then X_t is a normal random variable with

$$\mathrm{E}[X_t] = 0 \quad \text{and} \quad \mathrm{Var}[X_t] = \frac{1 - e^{-2at}}{2a}\cdot \tag{4.3}$$

For the remainder of this paper, we adopt the convention that $(1 - e^{-2at})/(2a)$ is equal to t when $a = 0$. With this convention in place, (4.3) holds for all $a \in \mathbb{R}$. As a consequence of the above, for any $\delta > 0$ we have

$$\mathrm{Prob}\,(|X_t| > \delta) \leq \begin{cases} C \exp(-c\delta^2 a) & \text{for any } a \geq 0, t \geq 0, \\ C \exp(-c\delta^2/t) & \text{for any } |a|t \leq 1/10. \end{cases} \tag{4.4}$$

Note that these two cases are not mutually exclusive.

We also note that X_t satisfies the reflection principle, i.e., for any $M > 0$ we have

$$\mathrm{Prob}\Big(\sup_{0 \leq s \leq t} X_s \geq M\Big) = 2 \cdot \mathrm{Prob}\,(X_t \geq M). \tag{4.5}$$

This is true because

(1) X_t has almost surely continuous paths,
(2) X_t satisfies the strong Markov property, and
(3) the random variable $(X_{t+t'} - X_t)$ is equal in distribution to $e^{-at} X_{t'}$, which is symmetric.

Examining the proof of the reflection principle for Brownian motion in [18], one sees that properties (1)–(3) are sufficient to prove (4.5).

4.1. Stopping times

We let $\varepsilon > 0$ be a given parameter. We define two stopping times.

- τ_+ is equal to the first time $t \in (0, T)$ for which $q(t) = (1+\varepsilon)q_0$ if such a time exists, and equal to T if no such time exists.
- τ_- is equal to the first time $t \in (0, T)$ for which $q(t) = (1-\varepsilon)q_0$ if such a time exists, and equal to T if no such time exists.

We will make use of the following claim throughout this section.

Claim 4.1. *Let $a \neq 0$ and let $\eta \geq \eta_0$, where $\eta_0 > 1$ is a sufficiently large absolute constant. Define*

$$t_0 = \frac{\eta\varepsilon}{|a|}.$$

Suppose that $t_0 < T$ and that $t_0|a| < 1/10$. Then, provided ε is sufficiently small depending on T, the following hold.

(i) *If $a > 0$, then*

$$\mathrm{Prob}(\tau_+ > t_0) \leq C\exp(-c\varepsilon\eta q_0^2 a).$$

(ii) *If $a < 0$, then*

$$\mathrm{Prob}(\tau_- > t_0) \leq C\exp(-c\varepsilon\eta q_0^2|a|).$$

Proof. We will only prove (i); the proof of (ii) is nearly identical. Assume that $a > 0$. By our assumption that $t_0 < T$, we have

$$\begin{aligned}\mathrm{Prob}(\tau_+ > t_0) &= \mathrm{Prob}(q(t) < (1+\varepsilon)q_0 \text{ for all } t \in [0, t_0]) \\ &\leq \mathrm{Prob}(q(t_0) < (1+\varepsilon)q_0).\end{aligned} \tag{4.6}$$

Let X_t be as in (4.2). If $q(t_0) < (1+\varepsilon)q_0$, then (since we assume $at_0 = \eta\varepsilon < 1/10$) we have

$$X_{t_0} < e^{-at_0}(1+\varepsilon)q_0 - q_0 \leq (1 - c\eta\varepsilon)(1+\varepsilon)q_0 - q_0 \leq -c'\eta\varepsilon q_0$$

provided η is larger than an absolute constant and ε is smaller than an absolute constant. This shows that

$$\mathrm{Prob}(\tau_+ > t_0) \leq \mathrm{Prob}(X_{t_0} \leq -c'\eta\varepsilon q_0);$$

applying (4.4) (and again using the assumption $t_0|a| < 1/10$) proves (i). ∎

Lemma 4.2. *Let $c_0 > 0$ be a sufficiently small absolute constant. Define*

$$t_{\max} = c_0\varepsilon^{1/2}, \quad a_{\text{tiny}} = \varepsilon^{1/2} \quad \text{and} \quad a_{\text{small}} = \varepsilon^{1/4}.$$

Then, provided ε is sufficiently small depending on T, the following hold.

(i) *For any $a \leq a_{\text{tiny}}$,*

$$\mathrm{Prob}((\tau_+ < t_{\max}) \text{ AND } (\tau_+ < \tau_-)) \leq C\exp(-c_\varepsilon q_0^2(|a|+1)).$$

(ii) *For any $a \geq -a_{\text{tiny}}$,*

$$\mathrm{Prob}((\tau_- < t_{\max}) \text{ AND } (\tau_- < \tau_+)) \leq C\exp(-c_\varepsilon q_0^2(|a|+1)).$$

(iii) *For any* $|a| \geq a_{\text{small}}$,

$$\text{Prob}((\tau_+ \geq t_{\max}) \text{ AND } (\tau_- \geq t_{\max})) \leq C \exp(-c_\varepsilon q_0^2 |a|).$$

(iv) *Define a stopping time* $\tilde{\tau} = \min\{\tau_+, \tau_-, t_{\max}\}$. *Then, provided* q_0 *is sufficiently large depending on* ε *and* T, *we have*

$$\mathrm{E}\Big[\int_0^{\tilde{\tau}} q^2(t)\, dt\Big] < C\varepsilon^{1/2} \mathcal{J}_0(a; T, q_0) \quad \textit{for any } a \in \mathbb{R}.$$

Proof. We note that by taking ε sufficiently small depending on T, we can assume that $t_{\max} < T$.

We first claim that

$$\text{Prob}(\tau_+ < t_{\max}) \leq C \exp(-c\,\varepsilon^{3/2} q_0^2) \quad \text{for any } 0 \leq a \leq a_{\text{tiny}}, \tag{4.7}$$

$$\text{Prob}(\tau_- < t_{\max}) \leq C \exp(-c\,\varepsilon^{3/2} q_0^2) \quad \text{for any } -a_{\text{tiny}} \leq a \leq 0. \tag{4.8}$$

Assume that $0 \leq a \leq a_{\text{tiny}}$. Since $t_{\max} < T$, we have

$$\text{Prob}(\tau_+ < t_{\max}) = \text{Prob}(\exists t \in (0, t_{\max}) : q(t) \geq (1+\varepsilon) q_0).$$

Note that we have $0 \leq at \leq c_0\varepsilon$, and thus $e^{-at} \geq (1 - c_0\varepsilon)$ for any $t \in (0, t_{\max})$. Therefore, if $q(t) \geq (1+\varepsilon) q_0$ for any $t \in (0, t_{\max})$, then

$$X_t = e^{-at} q(t) - q_0 \geq (1 - c_0\varepsilon)(1+\varepsilon) q_0 - q_0 > \varepsilon q_0/2$$

provided c_0 and ε are smaller than certain absolute constants. We have shown that

$$\text{Prob}(\tau_+ < t_{\max}) \leq \text{Prob}(\exists t \in (0, t_{\max}) : X_t > \varepsilon q_0/2).$$

Applying (4.5) and then (4.4) (we use that $t_{\max} a \leq c_0\varepsilon < 1/10$) proves (4.7). We omit the proof of (4.8), as it can be easily inferred from the proof of (4.7).

We now claim that

$$\text{Prob}(\tau_- < t_{\max}) \leq C \exp(-c\,\varepsilon^{3/2} q_0^2) \quad \text{for any } 0 \leq a \leq a_{\text{small}}, \tag{4.9}$$

$$\text{Prob}(\tau_+ < t_{\max}) \leq C \exp(-c\,\varepsilon^{3/2} q_0^2) \quad \text{for any } -a_{\text{small}} \leq a \leq 0. \tag{4.10}$$

Assume that $0 \leq a \leq a_{\text{small}}$. If $q(t) \leq (1-\varepsilon) q_0$ for some $t \in (0, t_{\max})$, then

$$X_t \leq e^{-at}(1-\varepsilon) q_0 - q_0 \leq -\varepsilon q_0.$$

This implies that

$$\text{Prob}(\tau_- < t_{\max}) \leq \text{Prob}(\exists t \in (0, t_{\max}) : X_t \leq -\varepsilon q_0).$$

Applying (4.5) and then (4.4) (note that $t_{\max} a < c_0\varepsilon^{3/4} < 1/10$ for any $0 \leq a \leq a_{\text{small}}$) proves (4.9). The proof of (4.10) is very similar; we omit it.

Now assume that $a \geq a_{\text{small}}$. We claim that the following estimates hold:

$$\text{Prob}(\tau_+ \geq t_{\max}) < C \exp(-c\varepsilon q_0^2 a), \tag{4.11}$$

$$\text{Prob}((\tau_- < t_{\max}) \text{ AND } (\tau_- < \tau_+)) < C \exp(-c\varepsilon^2 q_0^2 a). \tag{4.12}$$

Let $\eta_0 > 0$ be the absolute constant in Claim 4.1, and define

$$\tilde{t}_0 = \frac{\eta_0\,\varepsilon}{a}.$$

By our assumption that $a \geq a_{\text{small}}$, we have

$$\tilde{t}_0 \leq \eta_0\,\varepsilon^{3/4} < c_0\,\varepsilon^{1/2} = t_{\max} < T \tag{4.13}$$

provided ε is smaller than an absolute constant. Therefore

$$\operatorname{Prob}(\tau_+ \geq t_{\max}) \leq \operatorname{Prob}(\tau_+ \geq \tilde{t}_0). \tag{4.14}$$

Provided ε is smaller than an absolute constant we have $\tilde{t}_0\,a = \eta_0\,\varepsilon < 1/10$. Together with Claim 4.1, (4.14) implies (4.11).

Define

$$t_0^* = \min\left\{t_{\max}, \frac{1}{10a}\right\}.$$

Observe that $t_0^* < T$, and so we have

$$\operatorname{Prob}(\tau_- < t_0^*) = \operatorname{Prob}(\exists t \in (0, t_0^*) : q(t) < (1-\varepsilon)\,q_0).$$

If $q(t) < (1-\varepsilon)\,q_0$ for some $t \in (0, t_0^*)$, then (since $a > 0$)

$$X_t < e^{-at}(1-\varepsilon)\,q_0 - q_0 < -\varepsilon\,q_0.$$

Therefore

$$\operatorname{Prob}(\tau_- < t_0^*) \leq \operatorname{Prob}(\exists t \in (0, t_0^*) : X_t < -\varepsilon\,q_0).$$

Combining this with (4.5) gives

$$\operatorname{Prob}(\tau_- < t_0^*) \leq 2 \cdot \operatorname{Prob}(X_{t_0^*} < -\varepsilon\,q_0)$$

Using (4.4) and the fact that $a t_0^* < 1/10$ gives

$$\operatorname{Prob}(\tau_- < t_0^*) \leq C\,\exp(-c\,\varepsilon^2 q_0^2/t_0^*). \tag{4.15}$$

We claim that (4.15) implies (4.12) under the assumption that

$$a_{\text{small}} \leq a \leq \frac{1}{10 c_0\,\varepsilon^{1/2}}. \tag{4.16}$$

To see this, note that (4.16) implies that $t_0^* = t_{\max}$, and therefore (4.15) gives

$$\operatorname{Prob}(\tau_- < t_{\max}) \leq C\,\exp(-c\,\varepsilon^{3/2} q_0^2). \tag{4.17}$$

We then use (4.16) to deduce that

$$\operatorname{Prob}(\tau_- < t_{\max}) \leq C\,\exp(-c\,\varepsilon^2 q_0^2\,a).$$

This proves (4.12) assuming (4.16).

We now prove (4.12) assuming that

$$a > \frac{1}{10 c_0\, \varepsilon^{1/2}}. \tag{4.18}$$

Note that this assumption implies that $t_0^* < t_{\max}$. We have

$$\operatorname{Prob}\big((\tau_- < t_{\max}) \text{ AND } (\tau_- < \tau_+)\big) \le \operatorname{Prob}(\tau_- < t_0^*) + \operatorname{Prob}(\tau_+ > t_0^*). \tag{4.19}$$

By (4.15) (and using that assumption (4.18) implies that $t_0^* = (10a)^{-1}$), we have

$$\operatorname{Prob}(\tau_- < t_0^*) \le C \exp(-c\varepsilon^2 q_0^2\, a). \tag{4.20}$$

Observe that $t_0^* < T$ and that $t_0^* a = 1/10$. We can therefore apply Claim 4.1 with $\eta = 1/(10\varepsilon)$ to get

$$\operatorname{Prob}(\tau_+ > t_0^*) \le C \exp(-c\, q_0^2\, a). \tag{4.21}$$

Combining (4.19)–(4.21) proves (4.12) under assumption (4.18). This completes the proof of (4.12).

By mirroring the proofs of (4.11) and (4.12), we can show that for $a \le -a_{\text{small}}$ we have

$$\operatorname{Prob}(\tau_- \ge t_{\max}) < C \exp(-c\varepsilon q_0^2\, |a|), \tag{4.22}$$

$$\operatorname{Prob}\big((\tau_+ < t_{\max}) \text{ AND } (\tau_+ < \tau_-)\big) < C \exp(-c\varepsilon^2 q_0^2\, |a|). \tag{4.23}$$

Combining (4.7), (4.10), (4.23) proves (i). Similarly, combining (4.8), (4.9), (4.12) proves (ii). Equations (4.11) and (4.22) prove (iii).

We now prove (iv). Note that with probability 1 we have that $\tilde\tau \le t_{\max} = c_0\, \varepsilon^{1/2}$ and that $|q(t)| \le (1+\varepsilon) q_0$ for all $t \in (0, \tilde\tau)$. Therefore

$$\mathrm{E}\Big[\int_0^{\tilde\tau} q^2(t)\, dt\Big] < C q_0^2\, \varepsilon^{1/2}. \tag{4.24}$$

By Lemma 3.4, there exists $\tilde A > 0$ depending only on T such that

$$\mathcal{J}_0(a; T, q_0) > \frac{q_0^2}{100\,|a|} \quad \text{for any } a < -\tilde A, \tag{4.25}$$

$$\mathcal{J}_0(a; T, q_0) > c_T\, q_0^2 \quad \text{for any } a > -\tilde A. \tag{4.26}$$

Combining (4.24) and (4.26) gives

$$\mathrm{E}\Big[\int_0^{\tilde\tau} q^2(t)\, dt\Big] < C_T\, \varepsilon^{1/2}\, \mathcal{J}_0(a; T, q_0) \quad \text{for any } a > -\tilde A. \tag{4.27}$$

This proves (iv) when $a > -\tilde A$.

Now assume that $a < -\tilde A$. Let η_0 be as in Claim 4.1 and define

$$\bar t_0 = \frac{\eta_0\, \varepsilon}{|a|}.$$

By taking ε smaller than an absolute constant, we can ensure that $\bar{t}_0 < t_{\max}$ and that $\bar{t}_0\,|a| < 1/10$. Therefore we can apply Claim 4.1 to get

$$\operatorname{Prob}(\tilde{\tau} > \bar{t}_0) \le \operatorname{Prob}(\tau_- > \bar{t}_0) \le C\,\exp(-c\varepsilon q_0^2\,|a|).$$

Recall that with probability 1 we have $|q(t)| \le (1+\varepsilon)q_0$ for all $t \in [0,\tilde{\tau}]$. Therefore

$$\begin{aligned}\mathrm{E}\Big[\int_0^{\tilde{\tau}} q^2(t)\,dt\Big] &< C q_0^2\,\bar{t}_0 + C q_0^2\,t_{\max}\operatorname{Prob}(\tilde{\tau} > \bar{t}_0)\\ &< \frac{C q_0^2\,\varepsilon}{|a|} + C q_0^2\,t_{\max}\exp(-c\varepsilon q_0^2\,|a|).\end{aligned}$$

Taking q_0 sufficiently large depending on ε and using (4.25) gives

$$\mathrm{E}\Big[\int_0^{\tilde{\tau}} q^2(t)\,dt\Big] < C\varepsilon\mathcal{J}_0(a,T,q_0) \quad \text{for any } a < -\tilde{A}. \tag{4.28}$$

Combining (4.27) and (4.28) proves (iv). ∎

We now let $q_0^* > q_0$ be a real number, and we define two additional stopping times:

- τ_+^* is equal to the first time $t \in (0,T)$ for which $q(t) = q_0^*$ if such a time exists, and equal to T if no such time exists.
- τ_-^* is equal to the first time $t \in (0,T)$ for which $q(t) = -q_0^*$ if such a time exists, and equal to T if no such time exists.

Lemma 4.3. *Provided ε is sufficiently small depending on T, the following hold.*

(i) *For any $a \ge -\varepsilon^{-1/2}$,*

$$\operatorname{Prob}(\tau_-^* < \tau_+) \le C_{q_0}\,\varepsilon^{1/4}.$$

(ii) *Define a stopping time*

$$\bar{\tau} = \min\{\tau_+, \tau_-^*\}.$$

Then

$$\mathrm{E}\Big[\int_0^{\bar{\tau}} q^2(t)\,dt\Big] < C_{T,q_0,q_0^*}\,\varepsilon^{1/4} \quad \textit{for any } a \in \mathbb{R}.$$

Proof. We first prove a claim. Define

$$\tilde{t}_1 = \varepsilon\min\{1, 1/|a|\}$$

(when $a = 0$, we set $\tilde{t}_1 = \varepsilon$). We claim that

$$\operatorname{Prob}(\tau_+ > \tilde{t}_1) \le C q_0\,\varepsilon^{1/4} \quad \text{for any } |a| \le \varepsilon^{-1/2}. \tag{4.29}$$

Assume that $|a| \le \varepsilon^{-1/2}$. By taking ε sufficiently small depending on T, we can assume that $\tilde{t}_1 < T$. We have

$$\operatorname{Prob}(\tau_+ > \tilde{t}_1) = \operatorname{Prob}(q(t) < (1+\varepsilon)q_0 \quad \text{for all } t \in [0,\tilde{t}_1]).$$

Note that $|a|\tilde{t}_1 < \varepsilon$, so if $q(t) < (1+\varepsilon)q_0$ for any $t \in [0,\tilde{t}_1]$, then

$$X_t < e^{-at}(1+\varepsilon)q_0 - q_0 < C\varepsilon q_0. \tag{4.30}$$

Therefore

$$\begin{aligned}\operatorname{Prob}(\tau_+ > \tilde{t}_1) &\le \operatorname{Prob}(X_t < C\varepsilon q_0 \text{ for all } t \in [0, \tilde{t}_1]) \\ &= 1 - \operatorname{Prob}(\exists t \in [0, \tilde{t}_1] : X_t > C\varepsilon q_0) = 1 - 2 \cdot \operatorname{Prob}(X_{\tilde{t}_1} > C\varepsilon q_0),\end{aligned} \tag{4.31}$$

where the last equality follows from (4.5). Recall (see (4.3)) that $X_{\tilde{t}_1}$ is a normal random variable with mean 0 and variance $(1 - e^{-2a\tilde{t}_1})/2a$. Also recall that $|a|\tilde{t}_1 < \varepsilon$, so provided ε is smaller than some absolute constant, the variance of $X_{\tilde{t}_1}$ is bounded above and below by constant multiples of $\tilde{t}_1$. Therefore

$$\operatorname{Prob}(X_{\tilde{t}_1} > C\varepsilon q_0) \ge \frac{1}{2} - C\int_0^{C\varepsilon q_0/\tilde{t}_1^{1/2}} e^{-t^2/2}\, dt \ge \frac{1}{2} - \frac{C\varepsilon q_0}{\tilde{t}_1^{1/2}}, \tag{4.32}$$

so

$$\operatorname{Prob}(\tau_+ > \tilde{t}_1) \le \frac{C\varepsilon q_0}{\tilde{t}_1^{1/2}}.$$

Using the assumption $|a| \le \varepsilon^{-1/2}$ and the definition of $\tilde{t}_1$ implies (4.29).

We now prove (i). Note that for any $t^* \in (0, T)$, we have

$$\operatorname{Prob}(\tau_-^* < \tau_+) \le \operatorname{Prob}(\tau_-^* < t^*) + \operatorname{Prob}(\tau_+ > t^*). \tag{4.33}$$

In the event that $q(t) = -q_0^*$ for some $t \in (0, T)$, then

$$X_t = -e^{-at}\, q_0^* - q_0 < -q_0.$$

Therefore,

$$\operatorname{Prob}(\tau_-^* < t^*) = \operatorname{Prob}(\exists t \in (0, t^*) : q(t) = -q_0^*) \le \operatorname{Prob}(\exists t \in (0, t^*) : X_t < -q_0).$$

Equation (4.5) implies

$$\operatorname{Prob}(\tau_-^* < t^*) \le 2 \cdot \operatorname{Prob}(X_{t^*} < -q_0). \tag{4.34}$$

We let $\tilde{t}_1$ be as above; note that by definition $\tilde{t}_1|a| < \varepsilon$ and $1/\tilde{t}_1 > 1/\varepsilon$. Combining (4.34) and (4.4) gives

$$\operatorname{Prob}(\tau_-^* < \tilde{t}_1) \le C\exp(-cq_0^2/\varepsilon) \text{ for } |a| \le \varepsilon^{-1/2}.$$

Combining this with (4.33) and (4.29) proves (i) for $|a| \le \varepsilon^{-1/2}$.

Assume that $a \ge \varepsilon^{-1/2}$. Define

$$\bar{t}_1 = \frac{1}{10a}.$$

Let η_0 be as in Claim 4.1. Taking ε sufficiently small depending on T ensures that $\bar{t}_1 < T$ and that $(10\varepsilon)^{-1} > \eta_0$. We can therefore apply Claim 4.1 with $\eta = (10\varepsilon)^{-1}$ to get

$$\operatorname{Prob}(\tau_+ > \bar{t}_1) \le C\exp(-cq_0^2 a). \tag{4.35}$$

Taking ε smaller than an absolute constant gives $1/\bar{t}_1 > 1$. Combined with (4.34) and (4.4), this gives

$$\operatorname{Prob}(\tau_-^* < \bar{t}_1) \le C\exp(-cq_0^2 a).$$

By (4.33), we have therefore shown that

$$\mathrm{Prob}\,(\tau_-^* < \tau_+) \le C\exp(-c q_0^2 a) \quad \text{for any } a > \varepsilon^{-1/2}.$$

This implies (i) for $a > \varepsilon^{-1/2}$ (here we use the assumption that q_0 is nonzero).

We now prove (ii). We split the analysis into cases.

Case I (Large negative a).

Assume that $a < -\varepsilon^{-1/2}$. Note that the system (4.1) satisfies

$$q(t) = q_0 + W(t) + \int_0^t a\, q(\tau)\, d\tau,$$

with probability 1. Therefore the expected value of the integral of q^2 from time 0 to time T is equal to the expected cost incurred from time 0 to time T by the constant gain strategy CG(0), i.e., the strategy that sets $u = 0$ (see Section 3). Applying Corollary 3.7 with $\alpha = 0$, taking ε sufficiently small depending on T, and using that $\bar{\tau} \le T$ with probability 1, we have

$$\mathrm{E}\Big[\int_0^{\bar{\tau}} q^2(t)\, dt\Big] \le \frac{q_0^2 + T}{2|a|} \quad \text{for any } a < -\varepsilon^{-1/2},$$

and therefore,

$$\mathrm{E}\Big[\int_0^{\bar{\tau}} q^2(t)\, dt\Big] < C_{T,q_0}\, \varepsilon^{1/2} \quad \text{for any } a < -\varepsilon^{-1/2}. \tag{4.36}$$

Case II (Large positive a).

Assume that $a > \varepsilon^{-1/2}$ and let $\bar{t}_1$ be as above. Since $\bar{\tau} < T$ and since $|q(t)| \le \min\{(1+\varepsilon)q_0, q_0^*\}$ for all $t \in (0, \bar{\tau})$ with probability 1, we have

$$\mathrm{E}\Big[\int_0^{\bar{\tau}} q^2(t)\, dt\Big] \le C_{q_0,q_0^*}\,(\bar{t}_1 + T\cdot \mathrm{Prob}\,(\bar{\tau} > \bar{t}_1)). \tag{4.37}$$

Note that our assumption that $a > \varepsilon^{-1/2}$ implies that $\bar{t}_1 < c\varepsilon^{1/2}$. The event $\bar{\tau} > \bar{t}_1$ implies the event $\tau_+ > \bar{t}_1$, and therefore we combine (4.37) with (4.35) to get

$$\mathrm{E}\Big[\int_0^{\bar{\tau}} q^2(t)\, dt\Big] < C_{T,q_0,q_0^*}\, \varepsilon^{1/2} \quad \text{for any } a > \varepsilon^{-1/2} \tag{4.38}$$

(note that we also use the assumption that q_0 is nonzero).

Case III (Bounded a).

Let $\tilde{t}_1$ be as above. Then

$$\mathrm{E}\Big[\int_0^{\bar{\tau}} q^2(t)\, dt\Big] < C_{q_0,q_0^*}\,(\tilde{t}_1 + T\cdot \mathrm{Prob}\,(\bar{\tau} > \tilde{t}_1)) \quad \text{for any } |a| \le \varepsilon^{-1/2}.$$

The event $\bar{\tau} > \tilde{t}_1$ implies the event $\tau_+ > \tilde{t}_1$. Therefore, by (4.29), we have

$$\mathrm{E}\Big[\int_0^{\bar{\tau}} q^2(t)\, dt\Big] < C_{T,q_0,q_0^*}\, \varepsilon^{1/4} \quad \text{for any } |a| \le \varepsilon^{-1/2}. \tag{4.39}$$

Combining (4.36), (4.38), and (4.39) proves (ii). ∎

Lemma 4.4. *Assume $a < 0$ and define $\rho = q_0^*/q_0$ (recall that $q_0^* > q_0$, and so $\rho > 1$). Then*

$$\mathrm{Prob}((\tau_+^* < T) \text{ OR } (\tau_-^* < T)) \leq C_{T,\rho}(1 + q_0^{-2}) \exp(-c_\rho q_0^2 |a|).$$

Proof. Let N be the smallest positive integer such that

$$10|a|T \leq N.$$

Note that

$$N < (1 + 10|a|T). \tag{4.40}$$

Define $\Delta t = T/N$ (observe that our choice of N ensures that $\Delta t|a| \leq 1/10$) and $I_j = [j\Delta t, (j+1)\Delta t]$ for $j = 0, \dots, N-1$. Note that

$$\mathrm{Prob}((\tau_+^* < T) \text{ OR } (\tau_-^* < T)) \leq \sum_{j=0}^{N-1} \mathrm{Prob}(\exists t \in I_j : |q(t)| \geq q_0^*). \tag{4.41}$$

We claim that

$$\mathrm{Prob}(\exists t \in I_j : |q(t)| \geq q_0^*) \leq C \exp(-c_\rho q_0^2 |a|) \quad \text{for any } j = 0, \dots, N-1. \tag{4.42}$$

Combining (4.40)–(4.42) proves the lemma. Therefore, it just remains to establish (4.42).

Since $a < 0$, we have that $e^{-at} \geq e^{-aj\Delta t}$ for all $t \in I_j$. Therefore, in the event that $q(t) \geq q_0^*$,

$$X_t = q(t)\, e^{-at} - q_0 \geq q_0^*\, e^{-at} \left(1 - \frac{1}{\rho} e^{at}\right) \geq c_\rho\, q_0^*\, e^{-aj\Delta t} \tag{4.43}$$

(we have used that $\rho > 1$ and $e^{at} \leq 1$). In the event that $q(t) \leq -q_0^*$,

$$X_t = q(t)\, e^{-at} - q_0 \leq -q_0^*\, e^{-aj\Delta t}. \tag{4.44}$$

Combining (4.43) and (4.44) implies that

$$\mathrm{Prob}(\exists t \in I_j : |q(t)| \geq q_0^*) \leq \mathrm{Prob}(\exists t \in I_j : |X_t| \geq c_\rho\, q_0^*\, e^{-aj\Delta t}). \tag{4.45}$$

Since $I_j \subset [0, (j+1)\Delta t]$, we have

$$\begin{aligned} &\mathrm{Prob}(\exists t \in I_j : |X_t| \geq c_\rho\, q_0^*\, e^{-aj\Delta t}) \\ &\qquad \leq \mathrm{Prob}(\exists t \in [0, (j+1)\Delta t] : |X_t| \geq c_\rho\, q_0^*\, e^{-aj\Delta t}). \end{aligned} \tag{4.46}$$

Equation (4.5) implies that

$$\begin{aligned} &\mathrm{Prob}(\exists t \in [0, (j+1)\Delta t] : |X_t| \geq c_\rho\, q_0^*\, e^{-aj\Delta t}) \\ &\qquad = 2 \cdot \mathrm{Prob}\left(|X_{(j+1)\Delta t}| \geq c_\rho\, q_0^*\, e^{-aj\Delta t}\right). \end{aligned} \tag{4.47}$$

Recall that $X_{(j+1)\Delta t}$ is a normal random variable with mean 0 and standard deviation

$$\left(\frac{e^{2|a|(j+1)\Delta t} - 1}{2|a|}\right)^{1/2}.$$

Recall also that our choice of N ensures that $\Delta t|a| < 1/10$. Therefore $c < e^{-|a|\Delta t} < 1$, and we have

$$\frac{c_\rho\, q_0^*\, e^{|a|j\Delta t}\,|a|^{1/2}}{(e^{2|a|(j+1)\Delta t}-1)^{1/2}} = \frac{c_\rho\, q_0^*\, e^{-|a|\Delta t}\,|a|^{1/2}}{(1-e^{-2|a|(j+1)\Delta t})^{1/2}} \geq c'_\rho\, q_0\, |a|^{1/2}.$$

Therefore, when $|X_{(j+1)\Delta t}| \geq c_\rho q_0^* e^{-aj\Delta t}$, the normal random variable $X_{(j+1)\Delta t}$ is at least $c'_\rho q_0 |a|^{1/2}$ standard deviations from its mean of 0. Thus

$$\text{Prob}\left(|X_{(j+1)\Delta t}| \geq c_\rho\, q_0^*\, e^{-aj\Delta t}\right) \leq C \exp(-c'_\rho\, q_0^2\, |a|). \tag{4.48}$$

Combining (4.45)–(4.48) proves (4.42), finishing the proof of the lemma. ■

4.2. Estimating the parameter *a*

We remain in the setting of the previous section.

In the event that $\tau_+ < T$, we define a random variable

$$\bar{a}_+ = \frac{\log(1+\varepsilon)}{\tau_+}.$$

Note that $\bar{a}_+$ is an estimate for the parameter a in (4.1). Similarly, in the event that $\tau_- < T$, we define a random variable

$$\bar{a}_- = \frac{\log(1-\varepsilon)}{\tau_-}.$$

Let $X \subset (0,\infty)$ be a subset. We will be interested in estimating the probability of events of the form "$\tau_+ < T$ and $\bar{a}_+ \in X$". We will abbreviate such events to "$\bar{a}_+ \in X$"; note that when $\tau_+ = T$, the random variable $\bar{a}_+$ is undefined, and so the event $\bar{a}_+ \in X$ is meaningless. Similarly, for $Y \subset (-\infty, 0)$, we let "$\bar{a}_- \in Y$" denote the event "$\tau_- < T$ and $\bar{a}_- \in Y$".

Lemma 4.5. *Let $A^\# > 1$. Then provided ε is sufficiently small depending on T, we have*

$$\text{Prob}(\bar{a}_+ > A^\#) \leq C \exp(-c_\varepsilon\, q_0^2\, A^\#)$$

for any $|a| \leq A^\#/10$.

Proof. Define

$$t_2 = \frac{\log(1+\varepsilon)}{A^\#}. \tag{4.49}$$

By taking ε sufficiently small depending on T, we ensure that $t_2 < T$. Therefore,

$$\text{Prob}\,(\bar{a}_+ > A^\#) = \text{Prob}\,(\tau_+ < t_2). \tag{4.50}$$

Suppose that $0 \leq a \leq A^\#/10$. Note that for any $t \in [0, t_2)$, we have

$$ta < \frac{a}{A^\#}\log(1+\varepsilon) < \frac{\varepsilon}{10},$$

and thus

$$e^{-at} \geq (1-\varepsilon/10).$$

Consequently, if $q(t) > (1+\varepsilon)q_0$ for any $t \in [0, t_2)$, then

$$X_t = e^{-at}q(t) - q_0 > e^{-at}(1+\varepsilon)q_0 - q_0 \geq c\varepsilon q_0.$$

Now suppose that $-A^{\#}/10 \le a < 0$. If $q(t) > (1+\varepsilon)q_0$ for some $t \in [0, t_2)$, then

$$X_t = e^{|a|t} q(t) - q_0 > \varepsilon q_0.$$

We have shown that

$$\operatorname{Prob}(\tau_+ < t_2) \le \operatorname{Prob}(\exists t \in [0, t_2) : X_t \ge c\varepsilon q_0) \quad \text{for any } |a| \le \frac{A^{\#}}{10}.$$

Combining this with (4.50) and (4.5), we deduce that

$$\operatorname{Prob}(\bar{a}_+ \ge A^{\#}) \le \operatorname{Prob}(\exists t \in [0, t_2) : X_t \ge c\varepsilon q_0) = 2 \cdot \operatorname{Prob}(X_{t_2} \ge c\varepsilon q_0). \tag{4.51}$$

Since $t_2 |a| < C\varepsilon$, we apply (4.4) to get

$$\operatorname{Prob}(X_{t_2} \ge c\varepsilon q_0) \le C \exp(-c\varepsilon^2 q_0^2 / t_2) \le \exp(-c_\varepsilon q_0^2 A^{\#}).$$

Combining this with (4.51) proves the lemma. ∎

Lemma 4.6. *Let $\delta > 0$. Then provided ε is sufficiently small depending on T, the following hold.*

(i) $\operatorname{Prob}(|a - \bar{a}_+| > \delta a) \le C \exp(-c_{\varepsilon,\delta}\, q_0^2 a)$ *for any $a \ge \varepsilon^{1/2}$.*

(ii) $\operatorname{Prob}(|a - \bar{a}_-| > \delta |a|) \le C \exp(-c_{\varepsilon,\delta}\, q_0^2 |a|)$ *for any $a \le -\varepsilon^{1/2}$.*

Proof. We will just prove (i), as essentially the same argument can be used to prove (ii).

Without loss of generality‘, we assume that $\delta < 1/10$. Note that

$$\operatorname{Prob}(|a - \bar{a}_+| > \delta a) = \operatorname{Prob}(\bar{a}_+ > (1+\delta)a) + \operatorname{Prob}(0 < \bar{a}_+ < (1-\delta)a). \tag{4.52}$$

Now define

$$t_3 = \frac{\log(1+\varepsilon)}{(1+\delta)a}.$$

By taking ε sufficiently small depending on T, we ensure that $t_3 < T$, and therefore

$$\operatorname{Prob}(\bar{a}_+ > (1+\delta)a) = \operatorname{Prob}(\tau_+ < t_3).$$

For any $t \in [0, t_3)$, we have

$$at \le at_3 \le \frac{\log(1+\varepsilon)}{1+\delta}.$$

Therefore, in the event that $q(t) > (1+\varepsilon)q_0$ for some $t \in [0, t_3)$, we have

$$X_t = e^{-at} q(t) - q_0 \ge q_0 \left((1+\varepsilon)^{1-1/(1+\delta)} - 1\right) = c_{\varepsilon,\delta}\, q_0.$$

This implies that

$$\operatorname{Prob}(\bar{a}_+ > (1+\delta)a) \le \operatorname{Prob}(\exists t \in (0, t_3) : X_t > c_{\varepsilon,\delta}\, q_0).$$

We use (4.5) to deduce that

$$\operatorname{Prob}(\bar{a}_+ > (1+\delta)a) \le 2 \cdot \operatorname{Prob}(X_{t_3} > c_{\varepsilon,\delta}\, q_0);$$

we then note that $t_3 a < C\varepsilon^{1/2} < 1/10$ and apply (4.4) to get

$$\mathrm{Prob}\,(\bar{a}_+ > (1+\delta)a) \leq C\exp(-c_{\varepsilon,\delta}\, q_0^2\, a). \tag{4.53}$$

Now define

$$t_4 = \frac{\log(1+\varepsilon)}{(1-\delta)a}.$$

By taking ε sufficiently small depending on T, we ensure that $t_4 < T$, and therefore

$$\mathrm{Prob}\,(0 < \bar{a}_+ < (1-\delta)a) \leq \mathrm{Prob}\,(\tau_+ > t_4).$$

If $q(t_4) < (1+\varepsilon)q_0$, then

$$X_{t_4} < e^{-at_4}(1+\varepsilon)q_0 - q_0 = q_0((1+\varepsilon)^{-\delta/(1-\delta)} - 1) < -c_{\varepsilon,\delta}\, q_0.$$

Therefore,

$$\mathrm{Prob}\,(0 < \bar{a}_+ < (1-\delta)a) \leq \mathrm{Prob}\,(X_{t_4} < -c_{\varepsilon,\delta}\, q_0);$$

applying (4.4) gives

$$\mathrm{Prob}\,(0 < \bar{a}_+ < (1-\delta)a) \leq C\exp(-c_{\varepsilon,\delta}\, q_0^2\, a). \tag{4.54}$$

Combining (4.52), (4.53) and (4.54) proves (i). ∎

5. The almost optimal strategy

Throughout this section, we fix a time horizon $T > 0$ and a starting position $q_0 > 0$.

We fix constants C_0 and m_0 as in the definition of an A-bounded strategy in Section 2. Recall that a strategy σ (for time horizon T) is A-bounded for some $A > 0$ if

$$|u^\sigma(t,a)| \leq C_0 A^{m_0}[|q^\sigma(t,a)| + 1] \quad \text{for all } a \in \mathbb{R}, t \in [0,T].$$

Throughout this section, we allow all constants and parameters to depend on C_0 and m_0.

In this section, we prove the following theorem.

Theorem 5.1. *Let $\varepsilon > 0$. Then for $\varepsilon_0 > 0$ sufficiently small depending on ε, and for $A > 0$ sufficiently large depending on ε and ε_0, the following holds.*

Let σ be an A-bounded strategy for time horizon $T + \varepsilon_0$ and starting position q_0. Then the strategy σ_* for time horizon T and starting position q_0 specified in Section 5.2 satisfies the following.

(1) *If $a \in [-A, A]$, then*

$$\mathcal{J}(\sigma_*, a; T, q_0) < \varepsilon + (1+\varepsilon)\cdot \sup\{\mathcal{J}(\sigma, b; T+\varepsilon_0, q_0) : |a-b| < \varepsilon|a|\}.$$

(2) *If $|a| > A$, then*

$$\mathcal{J}(\sigma_*, a; T, q_0) < \varepsilon + (1+\varepsilon)\cdot \mathcal{J}_0(a; T, q_0).$$

We now show that Theorem 5.1 implies Theorem 2.1. Let $\varepsilon > 0$. Let ε_0 be sufficiently small (depending on ε) and let A be sufficiently large (depending on ε,ε_0) so that the conclusion of Theorem 5.1 holds. Note that we can assume that $\varepsilon_0 < \varepsilon$.

Let σ be an A-bounded strategy for time horizon $T + \varepsilon$ and starting position q_0. Since $\varepsilon_0 < \varepsilon$, σ is also an A-bounded strategy for time horizon $T + \varepsilon_0$ and starting position q_0. Moreover,

$$\mathcal{J}(\sigma, a; T + \varepsilon_0, q_0) \leq \mathcal{J}(\sigma, a; T + \varepsilon, q_0) \quad \text{for any } a \in \mathbb{R}. \tag{5.1}$$

Let σ_* denote the strategy defined in Section 5.2. Observe that part (2) of Theorem 2.1 follows immediately from part (2) of Theorem 5.1. To prove part (1) of Theorem 2.1, we simply combine equation (5.1) with Part (2) of Theorem 5.1. This proves Theorem 2.1.

The remainder of Section 5 is devoted to proving Theorem 5.1. For the remainder of Section 5, we assume that we are given $\varepsilon > 0$ smaller than some constant. We let $\varepsilon_0 > 0$ be sufficiently small depending on ε, and $A > 0$ be sufficiently large depending on ε and ε_0.

In Section 5.2, we define the strategy σ_*. First, in Section 5.1, we establish some preliminary results. In Section 5.3, we prove Theorem 5.1.

5.1. Preliminaries

5.1.1. Averting disaster when *a* is large. Let σ be an A-bounded strategy for time horizon $T + \varepsilon_0$ and starting position q_0. Because σ is A-bounded, if $a \gg A$, then with high probability the control u^σ will fail to control the particle, and so the expected cost of σ will be much larger than $\mathcal{J}_0(a; T, q_0)$ (see [7]). We now show how to use Theorem 2.3 to obtain a strategy $\tilde{\sigma}$ that performs almost as well as σ when $|a| \leq A$, and much better than σ when $a \gg A$.

By Theorem 2.3, there exists a number $q_0^* > \max\{1, 2|q_0|\}$, depending on ε, ε_0, T and q_0, but independent of A, such that for any $T' \in [\varepsilon_0, 2T]$, there exists a strategy $\mathrm{L}q\mathrm{S}(T')$ for time horizon T' and starting position q_0^* satisfying

$$\mathcal{J}(\mathrm{L}q\mathrm{S}(T'), a; T', q_0^*) < (1 + \varepsilon) \cdot \mathcal{J}_0(a; T', q_0^*) \quad \text{for any } a \in \mathbb{R}. \tag{5.2}$$

We now define the strategy $\tilde{\sigma}$ for time horizon T and starting position q_0.

We execute the strategy σ from time 0 until time t^*, where t^* is equal to the first time $t \in (0, T)$ for which $|q^\sigma(t)| = q_0^*$, if such a time exists, and t^* is equal to T if no such time exists. Note that t^* is a stopping time.

If $t^* = T$, then $\tilde{\sigma}$ exercises the same control variable as σ at all times $t \in [0, T]$.

If $t^* < T$, then we execute the strategy $\mathrm{L}q\mathrm{S}(T + \varepsilon_0 - t^*)$ (from (5.2)) starting from time t^* and position $|q^{\tilde{\sigma}}(t^*)| = q_0^*$. For any $s \in (0, T)$, we have $(T + \varepsilon_0 - s) > \varepsilon_0$. By taking ε_0 sufficiently small depending on T we ensure that $(T + \varepsilon_0) < 2T$ and therefore that the strategy $\mathrm{L}q\mathrm{S}(T + \varepsilon_0 - s)$ is well-defined for any $s \in (0, T)$.

This concludes the definition of the strategy $\tilde{\sigma}$. We now estimate its expected cost.

If we never encounter $|q^{\tilde{\sigma}}| = q_0^*$, i.e., if $t^* = T$, then the strategies σ and $\tilde{\sigma}$ exercise the same control from time 0 to time T, hence they incur the same cost.

Suppose instead that we first encounter $|q^{\tilde{\sigma}}| = q_0^*$ at some time $t^* < T$. Until time t^*, the strategies σ and $\tilde{\sigma}$ exercise the same control and incur the same cost. Let S denote the expected cost of σ starting from position q_0^* and time t^* and continuing until time $T + \varepsilon_0$, and let $\tilde{S}$ denote the expected cost of $\tilde{\sigma}$ starting from position q_0^* and time t^* and continuing until time T. Note that

$$\tilde{S} = \mathcal{J}(\mathrm{L}q\mathrm{S}(T + \varepsilon_0 - t^*), a; T - t^*, q_0) < \mathcal{J}(\mathrm{L}q\mathrm{S}(T + \varepsilon_0 - t^*), a; T + \varepsilon_0 - t^*, q_0).$$

Therefore, by (5.2),

$$\tilde{S} < (1 + \varepsilon) \cdot \mathcal{J}_0(a; T + \varepsilon_0 - t^*, q_0).$$

By Lemma 3.5,

$$\mathcal{J}_0(a; T + \varepsilon_0 - t^*, q_0) \leq S,$$

and therefore $\tilde{S} \leq (1 + \varepsilon) \cdot S$.

As a consequence of the above discussion, we have

$$\mathcal{J}(\tilde{\sigma}, a; T, q_0) < (1 + \varepsilon) \cdot \mathcal{J}(\sigma, a; T + \varepsilon_0, q_0) \quad \text{for any } a \in \mathbb{R}. \tag{5.3}$$

We now derive another estimate on the expected cost of $\tilde{\sigma}$ – this one is much sharper than (5.3) when $|a| \gg A$.

From time 0 until time t^*, we execute the strategy σ and, by construction, have $|q^{\tilde{\sigma}}(t)| \leq q_0^*$ for all $t \in [0, t^*]$. Since σ is A-bounded, we therefore have

$$|u^{\tilde{\sigma}}(t)| < C_0 A^{m_0} (|q(t)| + 1) \quad \text{for any } t \in [0, t^*).$$

Therefore, from time 0 until time t^*, we incur a cost of at most $C_T A^{2m_0} (q_0^*)^2$ with probability 1.

In the event that $t^* < T$, then from time t^* until time T, we execute the strategy $\mathrm{L}q\mathrm{S}(T + \varepsilon_0 - t^*)$ (starting from position $\pm q_0^*$). Therefore, from time t^* until time T, we incur an expected cost that is at most $(1 + \varepsilon)$ times the expected cost of the optimal known-a strategy $\sigma_{\mathrm{opt}}(a)$ for time horizon $(T + \varepsilon_0 - t^*)$ and starting position q_0^*. By Corollary 3.8, this is at most $C_T (q_0^*)^2 \max\{a, 1\}$. We therefore have

$$\mathcal{J}(\tilde{\sigma}, a; T, q_0) < C_T (q_0^*)^2 \max\{A, a\}^{2m_0} \quad \text{for any } a \in \mathbb{R}.$$

Since q_0^* is determined by ε, ε_0, T and q_0, we have

$$\mathcal{J}(\tilde{\sigma}, a; T, q_0) < C_{T, q_0, \varepsilon, \varepsilon_0} \max\{A, a\}^{2m_0} \quad \text{for any } a \in \mathbb{R}. \tag{5.4}$$

5.1.2. Rescaling strategies. Remark 3.2 implies the following: given $\lambda > 1$ and a strategy σ for time horizon T and starting position q_0, we can define a strategy σ_λ for time horizon $\lambda^2 T$ and starting position λq_0 that satisfies

$$\mathcal{J}(\sigma_\lambda, a; \lambda^2 T, \lambda q_0) \leq \lambda^4 \cdot \mathcal{J}(\sigma, \lambda^2 a; T, q_0) \quad \text{for any } a \in \mathbb{R}. \tag{5.5}$$

Moreover, since $T < \lambda^2 T$, σ_λ is also a strategy for time horizon T and starting position λq_0, and we have

$$\mathcal{J}(\sigma_\lambda, a; T, \lambda q_0) \leq \mathcal{J}(\sigma_\lambda, a; \lambda^2 T, \lambda q_0) \quad \text{for any } a \in \mathbb{R}. \tag{5.6}$$

Combining (5.5) and (5.6), we deduce that

$$\mathcal{J}(\sigma_\lambda, a; T, \lambda q_0) \leq \lambda^4 \cdot \mathcal{J}(\sigma, \lambda^2 a; T, q_0) \quad \text{for any } a \in \mathbb{R}. \tag{5.7}$$

Let σ be an A-bounded strategy for time horizon $T + \varepsilon_0$ and starting position q_0.

By the results of the previous section (specifically, see (5.3) and (5.4)), there exists a strategy $\tilde{\sigma}$ for time horizon T and starting position q_0 satisfying

$$\mathcal{J}(\tilde{\sigma}, a; T, q_0) < (1+\varepsilon) \cdot \mathcal{J}(\sigma, a; T + \varepsilon_0, q_0) \quad \text{for any } a \in \mathbb{R} \tag{5.8}$$

and

$$\mathcal{J}(\tilde{\sigma}, a; T, q_0) < C_{T, q_0, \varepsilon, \varepsilon_0} \cdot \max\{A, a\}^{2m_0} \quad \text{for any } a \in \mathbb{R}. \tag{5.9}$$

By the discussion above, there then exists a strategy $\tilde{\sigma}_{1+\varepsilon_0}$ for time horizon T and starting position $(1+\varepsilon_0)q_0$ satisfying

$$\mathcal{J}(\tilde{\sigma}_{1+\varepsilon_0}, a; T, (1+\varepsilon_0)q_0) < (1+C\varepsilon)\, \mathcal{J}(\sigma, (1+\varepsilon_0)^2 a; T+\varepsilon_0, q_0) \quad \text{for any } a \in \mathbb{R} \tag{5.10}$$

(recall that ε_0 is sufficiently small depending on ε; in particular, to deduce (5.10) we assume that $\varepsilon_0 < \varepsilon$) and

$$\mathcal{J}(\tilde{\sigma}_{1+\varepsilon_0}, a; T, (1+\varepsilon_0)q_0) < C_{T, q_0, \varepsilon, \varepsilon_0} \cdot \max\{A, a\}^{2m_0} \quad \text{for any } a \in \mathbb{R}. \tag{5.11}$$

From equation (5.10), we deduce that

$$\begin{aligned} &\mathcal{J}(\tilde{\sigma}_{1+\varepsilon_0}, a; T, (1+\varepsilon_0)q_0) \\ &\qquad < (1+C\varepsilon) \cdot \sup\{\mathcal{J}(\sigma, b; T+\varepsilon_0, q_0) : |a-b| < \varepsilon|a|\} \quad \text{for any } a \in \mathbb{R} \end{aligned} \tag{5.12}$$

(again taking ε_0 sufficiently small depending on ε).

5.2. Definition of the almost optimal strategy

Let σ be an A-bounded strategy for time horizon $T + \varepsilon_0$ and starting position q_0.

To prove Theorem 5.1, we must exhibit a strategy σ_* for time horizon T and starting position q_0 satisfying

$$\begin{aligned} \mathcal{J}(\sigma_*, a; T, q_0) < \varepsilon + (1+C\varepsilon) \cdot \sup\{\mathcal{J}(\sigma, b; T+\varepsilon_0, q_0) : |b-a| < \varepsilon|a|\} \\ \text{for any } a \in [-A, A] \end{aligned} \tag{5.13}$$

and

$$\mathcal{J}(\sigma_*, a; T, q_0) < \varepsilon + (1+C\varepsilon) \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } |a| > A. \tag{5.14}$$

The definition of σ_* requires a bit of setup.

By Theorem 2.4, provided A is sufficiently large depending on ε, T, q_0, there exists a strategy LaS for time horizon T and starting position $(1+\varepsilon_0)q_0$ such that

$$\mathcal{J}(\text{L}a\text{S}, a; T, (1+\varepsilon_0)q_0) < (1+\varepsilon) \cdot \mathcal{J}_0(a; T, (1+\varepsilon_0)q_0) \quad \text{for any } a > \frac{1}{100}A, \tag{5.15}$$

and

$$\mathcal{J}(\mathrm{L}a\mathrm{S}, a; T, (1+\varepsilon_0)q_0) < C_{T,q_0}\, A^2 \quad \text{for any } a < \frac{1}{100}A. \tag{5.16}$$

By Lemma 3.4, provided ε_0 is sufficiently small depending on ε, T and q_0, we have

$$|\mathcal{J}_0(a; T, (1+\varepsilon_0)q_0) - \mathcal{J}_0(a; T, q_0)| < \varepsilon \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \in \mathbb{R}. \tag{5.17}$$

Combining (5.15) and (5.17) gives

$$\mathcal{J}(\mathrm{L}a\mathrm{S}, a; T, (1+\varepsilon_0)q_0) < (1 + C\varepsilon) \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a > \frac{1}{100}A. \tag{5.18}$$

By Section 5.1.2 (see equations (5.11) and (5.12)), provided ε_0 is sufficiently small depending on ε, T and q_0, there exists a strategy $\bar{\sigma}$ for time horizon T and starting position $(1+\varepsilon_0)q_0$ such that

$$\begin{aligned} &\mathcal{J}(\bar{\sigma}, a; T, (1+\varepsilon_0)q_0) \\ &\qquad < (1 + C\varepsilon) \cdot \sup\{\mathcal{J}(\sigma, b; T + \varepsilon_0, q_0) : |a - b| < \varepsilon|a|\} \quad \text{for any } a \in \mathbb{R} \end{aligned} \tag{5.19}$$

and

$$\mathcal{J}(\bar{\sigma}, a; T, (1+\varepsilon_0)q_0) < C_{T,q_0,\varepsilon,\varepsilon_0} \max\{A, a\}^{2m_0} \quad \text{for any } a \in \mathbb{R}. \tag{5.20}$$

Last, we introduce a constant $q_{\mathrm{rare}} > q_0$, depending on T and q_0, and a strategy BR ("BR" for bounded regret) for time horizon T and starting position $-q_{\mathrm{rare}}$ satisfying

$$\mathcal{J}(\mathrm{BR}, a; T, -q_{\mathrm{rare}}) < C \cdot \mathcal{J}_0(a; T, -q_{\mathrm{rare}}) \quad \text{for any } a \in \mathbb{R}.$$

This is possible due to Theorem 2.3; see, in particular, the discussion of bounded regret strategies following the statement of Theorem 2.3. We then observe, by Remark 3.3, that

$$\mathcal{J}_0(a; T, -q_{\mathrm{rare}}) < C_{T,q_0} \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \in \mathbb{R}.$$

This implies that the strategy BR satisfies

$$\mathcal{J}(\mathrm{BR}, a; T, -q_{\mathrm{rare}}) < C_{T,q_0} \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \in \mathbb{R}. \tag{5.21}$$

We fix the strategies LaS, $\bar{\sigma}$ and BR for the remainder of Section 5. We are now ready to define the strategy σ_*.

The strategy σ_* consists of two epochs: the Prologue and the Main Act. Only the Prologue is guaranteed to occur. In fact, for large negative a, we do not expect the Main Act to occur.

For the remainder of Section 5 we write q and u to denote, respectively, the particle trajectories q^{σ_*} and the control variables u^{σ_*}.

PROLOGUE.

During the Prologue, we set $u = 0$. We define a stopping time τ_1 by setting τ_1 equal to the first time $t \in (0, T)$ for which either $q(t) = (1+\varepsilon_0)q_0$ or $q(t) = -q_{\mathrm{rare}}$ if such a time exists, and setting τ_1 to be equal to T if no such time exists. The Prologue lasts from time 0 until time τ_1. If $\tau_1 = T$, then at time τ_1, the game ends along with the Prologue. If $\tau_1 < T$, then at time τ_1, we enter the Main Act.

MAIN ACT.

Suppose that we enter the Main Act at some time $\tau_1 \in (0, T)$. We either have $q(\tau_1) = (1+\varepsilon_0)q_0$ or $q(\tau_1) = -q_{\mathrm{rare}}$; we proceed differently in each case. If $q(\tau_1) = (1+\varepsilon_0)q_0$, then we define

$$\bar{a} = \frac{\log(1+\varepsilon_0)}{\tau_1}.$$

The value of $\bar{a}$ then determines our strategy during the Main Act. Note that in the event that $\bar{a}$ is defined (i.e., $\tau_1 < T$ and $q(\tau_1) = (1+\varepsilon_0)q_0$), we have that $\bar{a} > 0$ with probability 1. If $\bar{a} \geq A/10$, then we play strategy LaS beginning at time τ_1 and position $q(\tau_1) = (1+\varepsilon_0)q_0$ until the game ends at time T. If $\bar{a} < A/10$, then we play strategy $\bar{\sigma}$ beginning at time τ_1 and position $q(\tau_1) = (1+\varepsilon_0)q_0$ until the game ends at time T. This completes the description of our strategy during the Main Act in the event that $q(\tau_1) = (1+\varepsilon_0)q_0$. If $q(\tau_1) = -q_{\mathrm{rare}}$, then we play strategy BR beginning at time τ_1 and position $q(\tau_1) = -q_{\mathrm{rare}}$ until the end of the game at time T.

This concludes the definition of the strategy σ_*.

5.3. Proof of Theorem 5.1

Recall that the Prologue ends at time τ_1, and that during the Prologue, we set $u = 0$. We define a random variable COST_P to be the cost incurred during the Prologue for a given realization of the noise, i.e.,

$$\mathrm{COST}_P = \int_0^{\tau_1} q^2(t)\,dt.$$

Note that COST_P depends on a.

Taking ε_0 sufficiently small depending on T, we see from Lemma 4.3 that

$$\mathrm{E}[\mathrm{COST}_P] < C_{T,q_0}\,\varepsilon_0^{1/4} \quad \text{for any } a \in \mathbb{R} \tag{5.22}$$

(recall that q_{rare} is determined by T, q_0) and that

$$\mathrm{Prob}\,(q(\tau_1) = -q_{\mathrm{rare}}) \leq C_{q_0}\,\varepsilon_0^{1/4} \quad \text{for any } a \geq -\varepsilon_0^{-1/2}. \tag{5.23}$$

Recall that, if it occurs, the Main Act begins at time $\tau_1 < T$, in which case we have either $q(\tau_1) = (1+\varepsilon_0)q_0$ or $q(\tau_1) = -q_{\mathrm{rare}}$. In the event that the Main Act occurs and $q(\tau_1) = (1+\varepsilon_0)q_0$, we define the random variable

$$\bar{a} = \frac{\log(1+\varepsilon_0)}{\tau_1};$$

the value of $\bar{a}$ then determines our strategy during the Main Act. We will consider events of the form "$\bar{a} \in X$" for $X \subset \mathbb{R}$. This is shorthand for the event "$\tau_1 < T, q(\tau_1) = (1+\varepsilon_0)q_0$, and $\bar{a} \in X$" (recall that $\bar{a}$ is not defined unless $\tau_1 < T$ and $q(\tau_1) = (1+\varepsilon_0)q_0$).

Let $\alpha > 1$. We claim that, provided ε_0 is sufficiently small depending on T, we have the following estimates:

$$\text{If } a > 10\alpha, \text{ then } \mathrm{Prob}\,(\bar{a} < \alpha) < C\exp(-c_{\varepsilon_0}\,q_0^2\,a). \tag{5.24}$$

$$\text{If } |a| < \alpha/10, \text{ then } \mathrm{Prob}\,(\bar{a} > \alpha) < C\exp(-c_{\varepsilon_0}\,q_0^2\,\alpha). \tag{5.25}$$

$$\text{If } a < 0, \text{ then } \mathrm{Prob}\,(\bar{a} > 0) < C_{T,q_0,\varepsilon_0}\exp(-c_{\varepsilon_0}\,q_0^2\,|a|). \tag{5.26}$$

These estimates follow, respectively, from Lemmas 4.6, 4.5, and 4.4 in Section 4.

We define a random variable $\textsc{Cost}_M$ to be the cost incurred during the Main Act for a given realization of the noise, i.e.,

$$\textsc{Cost}_M = \int_{\tau_1}^{T} ((q(t))^2 + (u(t))^2)\, dt.$$

We use the remainder of this section to prove the estimates

$$\mathrm{E}[\textsc{Cost}_M] < C\varepsilon + (1 + C\varepsilon) \cdot \sup\{\mathcal{J}(\sigma, b; T + \varepsilon_0, q_0) : |b - a| < \varepsilon|a|\} \quad \text{for any } |a| \le A \tag{5.27}$$

and

$$\mathrm{E}[\textsc{Cost}_M] < C\varepsilon + (1 + C\varepsilon) \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } |a| > A. \tag{5.28}$$

Combining (5.27), (5.28) with the fact that

$$\mathcal{J}(\sigma_*, a; T, q_0) = \mathrm{E}[\textsc{Cost}_P] + \mathrm{E}[\textsc{Cost}_M]$$

and the estimate (5.22) above, and taking ε_0 sufficiently small depending on ε, T and q_0, we deduce (5.13) and (5.14), which in turn imply the conclusion of Theorem 5.1.

Note that

$$\mathrm{E}[\textsc{Cost}_M] = \mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{\bar{a}>0}] + \mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{q(\tau_1)=-q_{\mathrm{rare}}}]. \tag{5.29}$$

In the event that we enter the Main Act at some time $\tau_1 \in (0, T)$ for which $q(\tau_1) = -q_{\mathrm{rare}}$, we execute the strategy BR starting from time τ_1 and position $-q_{\mathrm{rare}}$. Combining (5.21) and (5.23) gives

$$\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{q(\tau_1)=-q_{\mathrm{rare}}}] < C_{T,q_0} \cdot \varepsilon_0^{1/4} \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \ge -\varepsilon_0^{-1/2}. \tag{5.30}$$

Provided ε_0 is sufficiently small (depending on T), Lemma 3.4 implies that

$$\mathcal{J}_0(a; T, q_0) < C_{T,q_0}\, \varepsilon_0^{1/2} \quad \text{for any } a < -\varepsilon_0^{-1/2};$$

combining this with (5.21) gives

$$\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{q(\tau_1)=-q_{\mathrm{rare}}}] < C_{T,q_0}\, \varepsilon_0^{1/2} \quad \text{for any } a < -\varepsilon_0^{-1/2}. \tag{5.31}$$

We combine (5.30) and (5.31), and take ε_0 sufficiently small depending on ε, T and q_0, to get

$$\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{q(\tau_1)=-q_{\mathrm{rare}}}] < \varepsilon + \varepsilon \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \in \mathbb{R}. \tag{5.32}$$

Throughout the remainder of this section, we will make use of the fact that

$$\mathcal{J}_0(a; T, q_0) < \mathcal{J}(\sigma, a; T + \varepsilon_0, q_0) \quad \text{for any } a \in \mathbb{R}; \tag{5.33}$$

this is a consequence of Lemma 3.5. We note that (5.32) and (5.33) imply that

$$\begin{aligned} &\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{q(\tau_1)=-q_{\mathrm{rare}}}] \\ &\qquad < \varepsilon + \varepsilon \cdot \sup\{\mathcal{J}(\sigma, b; T + \varepsilon_0, q_0) : |b - a| < \varepsilon|a|\} \quad \text{for any } a \in \mathbb{R}. \end{aligned} \tag{5.34}$$

It remains to control the expected value of COST_M when $\bar{a} > 0$. We will consider four cases.

Case I: *a is large, positive.*

Suppose that $a > A$. In the event that $\bar{a} \geq A/10$, we execute the strategy LaS during the Main Act. Therefore, by (5.18), we have

$$\mathrm{E}[\mathrm{COST}_M \cdot \mathbb{1}_{\bar{a} \geq A/10}] < \mathcal{J}(\mathrm{L}a\mathrm{S}, a; T, (1+\varepsilon_0)q_0) < (1 + C\varepsilon) \cdot \mathcal{J}_0(a; T, q_0). \tag{5.35}$$

In the event that $0 < \bar{a} < A/10$, we execute the strategy $\bar{\sigma}$ during the Main Act. Therefore, by (5.20) and (5.24), we have

$$\begin{aligned} \mathrm{E}[\mathrm{COST}_M \cdot \mathbb{1}_{0<\bar{a}<A/10}] &< \mathcal{J}(\bar{\sigma}, a; T, (1+\varepsilon_0)q_0) \cdot \mathrm{Prob}\,(\bar{a} < A/10) \\ &< C_{T,q_0,\varepsilon,\varepsilon_0}\, a^{2m_0}\, \exp(-c_{\varepsilon_0}\, q_0^2\, a). \end{aligned} \tag{5.36}$$

Combining (5.35) and (5.36) and taking A sufficiently large, depending on T, q_0, ε and ε_0, gives

$$\mathrm{E}[\mathrm{COST}_M \cdot \mathbb{1}_{\bar{a}>0}] < (1 + C\varepsilon) \cdot \mathcal{J}_0(a; T, q_0) + \varepsilon \quad \text{for any } a > A. \tag{5.37}$$

Case II: *a is positive, medium-sized.*

Suppose that $\frac{1}{100}A < a < A$.

If the event $\bar{a} \geq A/10$ occurs, then during the Main Act we execute the strategy LaS. By (5.18), we then have

$$\begin{aligned} \mathrm{E}[\mathrm{COST}_M \cdot \mathbb{1}_{\bar{a}>A/10}] &< \mathcal{J}(\mathrm{L}a\mathrm{S}, a; T, (1+\varepsilon_0)q_0) \cdot \mathrm{Prob}\,(\bar{a} \geq A/10) \\ &< (1 + C\varepsilon) \cdot \mathcal{J}_0(a; T, q_0) \cdot \mathrm{Prob}\,(\bar{a} \geq A/10). \end{aligned} \tag{5.38}$$

In the event that $0 < \bar{a} < A/10$, we execute the strategy $\bar{\sigma}$ during the Main Act. By (5.19), we then have

$$\begin{aligned} \mathrm{E}[\mathrm{COST}_M \cdot \mathbb{1}_{0<\bar{a}<A/10}] &< \mathcal{J}(\bar{\sigma}, a; T, (1+\varepsilon_0)q_0) \cdot \mathrm{Prob}\,(\bar{a} < A/10) \\ &< (1 + C\varepsilon) \cdot \sup\{\mathcal{J}(\sigma, b; T + \varepsilon_0, q_0) : |a - b| < \varepsilon|a|\} \\ &\quad \cdot \mathrm{Prob}\,(\bar{a} < A/10). \end{aligned} \tag{5.39}$$

Combining (5.38), (5.39) and (5.33), we get

$$\mathrm{E}[\mathrm{COST}_M \cdot \mathbb{1}_{\bar{a}>0}] < (1 + C\varepsilon) \cdot \sup\{\mathcal{J}(\sigma, b; T + \varepsilon_0, q_0) : |a - b| < \varepsilon|a|\} \tag{5.40}$$

for any $\frac{1}{100}A < a < A$.

Case III: *a is bounded.*

Suppose that $|a| < \frac{1}{100}A$.

In the event that $0 < \bar{a} < A/10$, we execute the strategy $\bar{\sigma}$ during the Main Act. By (5.19), we have

$$\begin{aligned} \mathrm{E}[\mathrm{COST}_M \cdot \mathbb{1}_{0<\bar{a}<A/10}] &< \mathcal{J}(\bar{\sigma}, a; T, (1+\varepsilon_0)q_0) \\ &< (1 + C\varepsilon) \cdot \sup\{\mathcal{J}(\sigma, b; T + \varepsilon_0, q_0) : |a - b| < \varepsilon|a|\}. \end{aligned} \tag{5.41}$$

In the event that $\bar{a} \geq A/10$, we execute the strategy LaS during the Main Act. By (5.16) and (5.25), we therefore have

$$\mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{\bar{a} > A/10}] < C_{T,q_0} A^2 \exp(-c_{\varepsilon_0} q_0^2 A). \tag{5.42}$$

Combining (5.41) and (5.42), and taking A sufficiently large, depending on T, q_0, ε and ε_0, gives

$$\mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{\bar{a} > 0}] < \varepsilon + (1 + C\varepsilon) \cdot \sup\{\mathcal{J}(\sigma, b; T + \varepsilon_0, q_0) : |a - b| < \varepsilon|a|\} \tag{5.43}$$

for any $|a| < A/100$.

Case IV: *a is large, negative.*

Suppose that $a < -\frac{1}{100}A$.

In the event that $\bar{a} > 0$, we execute either the strategy LaS or the strategy $\bar{\sigma}$ during the Main Act. Combining (5.16), (5.20) and (5.26) gives

$$\mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{\bar{a} > 0}] < C_{T,q_0,\varepsilon,\varepsilon_0} |a|^{2m_0} \exp(-c_{\varepsilon_0} q_0^2 |a|).$$

Taking A sufficiently large depending on T, q_0, ε and ε_0 then gives

$$\mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{\bar{a} > 0}] < \varepsilon \quad \text{for any } a < -\frac{A}{100}. \tag{5.44}$$

Combining (5.37), (5.40), (5.43) and (5.44) with (5.29), (5.32) and (5.34) proves (5.27) and (5.28).

5.4. Proof of Corollary 2.2

Let $\varepsilon > 0$ be given, and let A be sufficiently large so that the conclusion of Theorem 1.2 holds. Let σ be an A-bounded strategy for time horizon $T + \varepsilon$ and starting position ε.

Then the strategy σ_* (as in Theorem 1.2) for time horizon T and starting position ε satisfies

$$\mathcal{J}(\sigma_*, a; T, \varepsilon) < \varepsilon + (1 + \varepsilon) \cdot \sup\{\mathcal{J}(\sigma, b; T + \varepsilon, \varepsilon) : |a - b| < \varepsilon|a|\} \quad \text{for any } a \in [-A, A] \tag{5.45}$$

and

$$\mathcal{J}(\sigma_*, a; T, \varepsilon) < \varepsilon + (1 + \varepsilon) \cdot \mathcal{J}_0(a; T, \varepsilon) \quad \text{for any } |a| > A. \tag{5.46}$$

Combining (5.46) with Lemma 3.4 (specifically, parts (1) and (2) of Lemma 3.4), we deduce that, provided A is sufficiently large depending on ε and T, we have

$$\mathcal{J}(\sigma_*, a; T, \varepsilon) < \varepsilon + (1 + C\varepsilon) \cdot \mathcal{J}_0(a; T, 0) \quad \text{for any } |a| > A. \tag{5.47}$$

By Remark 3.1, σ_* gives rise to a strategy for time horizon T and starting position $-\varepsilon$; we also denote this strategy by σ_*, and recall that

$$\mathcal{J}(\sigma_*, a; T, \varepsilon) = \mathcal{J}(\sigma_*, a; T, -\varepsilon).$$

We now construct the strategy $\hat{\sigma}_*$ for time horizon T and starting position 0. We set $u^{\hat{\sigma}_*}(t) = 0$ from time 0 until time $\hat{t}$, where $\hat{t}$ is equal to the first time $t \in (0, T)$ for which $|q^{\hat{\sigma}_*}(t)| = \varepsilon$ if such a time exists, and $\hat{t}$ is equal to T if no such time exists.

If $\hat{t} = T$, then we set $u^{\hat{\sigma}_*} \equiv 0$ until the game ends at time T. In this case, we incur a cost of at most $T\varepsilon$ with probability 1.

If $\hat{t} < T$, then we execute the strategy σ_* starting from time $\hat{t}$ and position $|q(\hat{t})| = \varepsilon$ until the game ends at time T. In this case, our expected cost from time 0 until time T is at most $T\varepsilon + \mathcal{J}(\sigma_*, a; T, \varepsilon)$.

We have defined the strategy $\hat{\sigma}^*$ and shown that

$$\mathcal{J}(\hat{\sigma}^*, a; T, 0) < T\varepsilon + \mathcal{J}(\sigma_*, a; T, \varepsilon).$$

Combining this with (5.45) and (5.47) implies the conclusion of the corollary.

6. The large q strategy

In this section, we prove Theorem 2.3.

Fix a time horizon $T > 0$. Let $\varepsilon > 0$ be given. Note that it suffices to prove Theorem 2.3 under the assumption that ε is sufficiently small depending on T.

We let $t_{\max} \in (0, T)$ denote a real number that we will choose in Section 6.1 depending on ε.

Let $A_1 \geq 1$ be a sufficiently large real number depending on ε and T.

We let $q_{\text{big}} \geq 1$ be a sufficiently large real number depending on ε, A_1 and T. Note that since A_1 is determined by ε and T, q_{big} depends only on ε and T. We fix a starting position $q_0 \geq q_{\text{big}}$.

For $\alpha \in \mathbb{R}$, recall that $\sigma_{\text{opt}}(\alpha)$ denotes the known-a strategy for time horizon T (recall that this strategy is independent of the starting position; see Section 2). The corresponding control variable is

$$u^{\sigma_{\text{opt}}(\alpha)}(t, a) = -\kappa(T - t, \alpha)\, q^{\sigma_{\text{opt}}(\alpha)}(t, a)$$

for any $a \in \mathbb{R}$, $t \in [0, T]$, where

$$\kappa(T - t, \alpha) = \frac{\tanh((T - t)\sqrt{\alpha^2 + 1})}{\sqrt{\alpha^2 + 1} - \alpha \tanh((T - t)\sqrt{\alpha^2 + 1})}.$$

We will occasionally refer to the known-a strategy for an arbitrary time horizon $T' \in (0, T]$; we denote this strategy by $\sigma_{\text{opt}}(\alpha; T')$ for $\alpha \in \mathbb{R}$. We note that $\kappa(T' - t, \alpha)$ is the gain function corresponding to $\sigma_{\text{opt}}(\alpha; T')$, and we remark that

$$0 \leq \kappa(T' - t, \alpha) \leq C \cdot \max\{\alpha, 1\} \quad \text{for all } t \in [0, T'], \alpha \in \mathbb{R}, \tag{6.1}$$

with the constant C independent of T'.

We are now ready to define the strategy LqS. For the remainder of Section 6, we write q and u to denote, respectively, the particle trajectories $q^{\mathrm{L}q\mathrm{S}}$ and the control variables $u^{\mathrm{L}q\mathrm{S}}$.

We note that the strategy LqS consists of two epochs: the Prologue and the Main Act. Both epochs are guaranteed to occur.

PROLOGUE.

We define a stopping time τ_M by setting τ_M equal to the first time $t \in (0, t_{\max})$ for which $q(t) \notin ((1-\varepsilon)q_0, (1+\varepsilon)q_0)$ if such a time exists, and $\tau_M = t_{\max}$ if no such time exists.

The Prologue lasts from time 0 to time τ_M. During the Prologue, we exercise no control, i.e., we set $u \equiv 0$. At time τ_M, we enter the Main Act.

MAIN ACT.

The Main Act lasts from time τ_M until the end of the game at time T. Our strategy during the Main Act depends on what happens during the Prologue. We define some events that capture the possible outcomes of the Prologue:

$$\begin{aligned}
\mathrm{PRO}_+ &= \{q(\tau_M) = (1+\varepsilon)q_0\},\\
\mathrm{PRO}_- &= \{q(\tau_M) = (1-\varepsilon)q_0\},\\
\mathrm{PRO}_{\max} &= \{\tau_M = t_{\max}\}.
\end{aligned}$$

The events PRO_+, PRO_- and $\mathrm{PRO}_{\max}$ partition our probability space.

If the event PRO_+ occurs with $\tau_M = s$ for some $s \in (0, t_{\max})$, then we define

$$\bar{a} = \frac{\log(1+\varepsilon)}{s}.$$

If the event PRO_- occurs with $\tau_M = s$ for some $s \in (0, t_{\max})$, then we define

$$\bar{a} = \frac{\log(1-\varepsilon)}{s}.$$

If PRO_+ (respectively, PRO_-) occurs, then we have $\bar{a} > 0$ (respectively, $\bar{a} < 0$) with probability 1.

If the event $\mathrm{PRO}_{\max}$ occurs, then we leave $\bar{a}$ undefined.

For $X \subset (0, \infty)$, we will sometimes write "$\bar{a} \in X$" as shorthand for the event "$\bar{a} \in X$ and PRO_+". Similarly, for $Y \subset (-\infty, 0)$ we write "$\bar{a} \in Y$" for the event "$\bar{a} \in Y$ and PRO_-".

We now specify our strategy during the Main Act.

Case I (We believe a is small).

If the event $\mathrm{PRO}_{\max}$ occurs, then we set

$$u(t) = -\kappa(T-t, 0) \cdot q(t)$$

during the Main Act. In this case, the expected cost incurred during the Main Act is equal to the expected cost of the known-a strategy $\sigma_{\mathrm{opt}}(0; T - t_{\max})$ for time horizon $(T - t_{\max})$ and starting position $q(t_{\max})$.

Case II (We believe a is large, positive).

If the event PRO_+ occurs and $\bar{a} \geq A_1$, then we set

$$u(t) = -2\bar{a}\, q(t)$$

during the Main Act. Note that this is the control variable corresponding to the constant gain strategy $\mathrm{CG}(2\bar{a})$ (see Section 3) with starting position $(1+\varepsilon)q_0$.

Case III (We believe that a is bounded, positive).

If the event PRO_+ occurs and $0 < \bar{a} < A_1$ and $\tau_M = s$ for some $s \in (0, t_{\max})$, then we set

$$u(t) = -\kappa(T - t, \bar{a}) \cdot q(t)$$

during the Main Act. In this case, the expected cost incurred during the Main Act is equal to the expected cost of the optimal known-a strategy $\sigma_{\text{opt}}(\bar{a}; T - s)$ for time horizon $T - s$ and starting position $(1 + \varepsilon) q_0$.

Case IV (We believe that a is large, negative).

If the event PRO_- occurs and $\bar{a} \le -A_1$, then we set

$$u(t) = 0$$

during the Main Act. Note that this is the control variable corresponding to the constant gain strategy CG(0) with starting position $(1 - \varepsilon) q_0$.

Case V (We believe that a is bounded, negative).

If the event PRO_- occurs and $-A_1 < \bar{a} < 0$ and $\tau_M = s$ for some $s \in (0, t_{\max})$, then we set

$$u(t) = -\kappa(T - s, \bar{a}) \cdot q(t)$$

during the Main Act. In this case, the expected cost incurred during the Main Act is equal to the expected cost of the optimal known-a strategy $\sigma_{\text{opt}}(\bar{a}; T - s)$ for time horizon $T - s$ and starting position $(1 - \varepsilon) q_0$.

This concludes the definition of the strategy LqS.

6.1. The parameter $t_{\max}$

We will now choose the parameter $t_{\max}$. We set

$$t_{\max} = c_0 \varepsilon^{1/2},$$

where c_0 is a sufficiently small absolute constant. We define

$$a_{\text{tiny}} = \varepsilon^{1/2} \quad \text{and} \quad a_{\text{small}} = \varepsilon^{1/4}.$$

Provided ε is sufficiently small depending on T, we apply Lemma 4.2 to deduce the following:

$$\text{Prob}(\text{PRO}_+) \le C \exp(-c_\varepsilon q_0^2 (|a| + 1)) \quad \text{for any } a \le a_{\text{tiny}}, \tag{6.2}$$

$$\text{Prob}(\text{PRO}_-) \le C \exp(-c_\varepsilon q_0^2 (|a| + 1)) \quad \text{for any } a \ge -a_{\text{tiny}}, \tag{6.3}$$

$$\text{Prob}(\text{PRO}_{\max}) \le C \exp(-c_\varepsilon q_0^2 |a|) \quad \text{for any } |a| \ge a_{\text{small}}. \tag{6.4}$$

6.2. Bounding the expected cost

Define random variables

$$\text{COST}_P = \int_0^{\tau_M} q^2(t)\, dt \quad \text{and} \quad \text{COST}_M = \int_{\tau_M}^{T} ((q(t))^2 + (u(t))^2))\, dt,$$

that are, respectively, the costs incurred during the Prologue and Main Act for a given realization of the noise. Note that both random variables depend on a.

Provided q_{big} is sufficiently large depending on ε and T, Lemma 4.2 implies that

$$\mathrm{E}[\mathrm{COST}_P] < C\varepsilon^{1/2}\mathcal{J}_0(a;T,q_0) \quad \text{for any } a \in \mathbb{R}. \tag{6.5}$$

Our goal in the remainder of this section is to control the expected value of COST_M. Recall that we play a different strategy during the Main Act depending on which one of five events occur during the Prologue. We analyze each of these cases separately in Sections 6.2.1–6.2.5. In Section 6.2.6, we prove Theorem 2.3.

6.2.1. COST_M when we believe that a is small. Recall that if the event $\mathrm{PRO}_{\max}$ occurs, then the Main Act begins at time $t_{\max}$, and we set

$$u(t) = -\kappa(T-t,0)\cdot q(t)$$

during the Main Act. In this case, our expected cost during the Main Act is equal to the expected cost of the optimal known-a strategy $\sigma_{\text{opt}}(0; T-t_{\max})$ for time horizon $T-t_{\max}$ starting from position $q(t_{\max}) = \tilde{q}_0$ for some $\tilde{q}_0 \in ((1-\varepsilon)q_0, (1+\varepsilon)q_0)$. By Corollary 3.8, we thus have

$$\mathrm{E}[\mathrm{COST}_M \,|\, \mathrm{PRO}_{\max}] \le C_T\, q_0^2\, e^{2|a|T} \quad \text{for any } a \in \mathbb{R}. \tag{6.6}$$

Combining (6.4) and (6.6) gives

$$\mathrm{E}[\mathrm{COST}_M \cdot \mathbb{1}_{\mathrm{PRO}_{\max}}] < C_T\, q_0^2 \cdot \exp(2|a|T - c_\varepsilon q_0^2 |a|) \quad \text{for any } |a| \ge a_{\text{small}}.$$

Taking q_{big} sufficiently large depending on ε and T, we get

$$\mathrm{E}[\mathrm{COST}_M \cdot \mathbb{1}_{\mathrm{PRO}_{\max}}] < \varepsilon \cdot \min\{1, 1/|a|\} \quad \text{for any } |a| \ge a_{\text{small}}. \tag{6.7}$$

By Lemma 3.4, there exists a constant $\tilde{A} > 1$, depending only on T, such that

$$\mathcal{J}_0(a;T,q_0) > \frac{c_T}{|a|} \quad \text{for any } a < -\tilde{A}, \tag{6.8}$$

$$\mathcal{J}_0(a;T,q_0) > c_T \quad \text{for any } a > \tilde{A}. \tag{6.9}$$

Again by Lemma 3.4,

$$\mathcal{J}_0(a;T,q_0) \ge \int_0^T \kappa(t,a)\,dt > c_T \quad \text{for any } |a| \le \tilde{A}. \tag{6.10}$$

Combining (6.7)–(6.10) gives

$$\mathrm{E}[\mathrm{COST}_M \cdot \mathbb{1}_{\mathrm{PRO}_{\max}}] \le \varepsilon C_T\, \mathcal{J}_0(a;T,q_0) \quad \text{for any } |a| \ge a_{\text{small}}. \tag{6.11}$$

By Lemma 3.6, we have

$$\mathcal{J}(\sigma_{\text{opt}}(\alpha), a; T, q_0) = h(\alpha,a)\, q_0^2 + j(\alpha,a) \quad \text{for any } \alpha \in \mathbb{R}, a \in \mathbb{R},$$

where $h, j : \mathbb{R}^2 \to (0,\infty)$ are smooth functions depending on T but not on q_0. This implies that

$$\mathrm{E}[\mathrm{COST}_M \,|\, \mathrm{PRO}_{\max}] \le h(0,a)\,(1+\varepsilon)^2 q_0^2 + j(0,a)$$

and

$$\mathcal{J}_0(a;T,q_0) = h(a,a)\,q_0^2 + j(a,a). \tag{6.12}$$

Therefore,

$$\begin{aligned}\mathrm{E}[\mathrm{Cost}_M\,|\,\mathrm{Pro}_{\max}] \le \mathcal{J}_0(a;T,q_0) + |j(a,a) - j(0,a)| + h(a,a)\,C\,\varepsilon q_0^2 \\ + C q_0^2\,|h(a,a) - h(0,a)|.\end{aligned}$$

Since h and j are smooth functions (depending on T), we have

$$\begin{aligned}&\mathrm{E}[\mathrm{Cost}_M\,|\,\mathrm{Pro}_{\max}] \\ &\quad \le \mathcal{J}_0(a;T,q_0) + C_T\,q_0^2|a| + h(a,a)\,C\,\varepsilon q_0^2 \quad \text{for any } |a| \le a_{\text{small}}.\end{aligned} \tag{6.13}$$

We combine this with (6.12) to get

$$\mathrm{E}[\mathrm{Cost}_M\,|\,\mathrm{Pro}_{\max}] < (1 + C\varepsilon)\cdot \mathcal{J}_0(a;T,q_0) + C_T\,q_0^2\,\varepsilon^{1/4} \quad \text{for any } |a| \le a_{\text{small}}.$$

Equation (6.12) also implies that

$$\mathcal{J}_0(a;T,q_0) \ge c_T\,q_0^2 \quad \text{for any } |a| \le a_{\text{small}}.$$

We have therefore shown that

$$\begin{aligned}&\mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{\mathrm{Pro}_{\max}}] \\ &\quad \le \mathcal{J}_0(a;T,q_0)\cdot(1 + C_T\,\varepsilon^{1/4})\cdot \mathrm{Prob}\,(\mathrm{Pro}_{\max}) \quad \text{for any } |a| \le a_{\text{small}}.\end{aligned} \tag{6.14}$$

6.2.2. Cost_M when we believe that a is large, positive. Let $\alpha \ge 0$ be a positive real number. Recall that $\mathrm{CG}(\alpha)$ denotes the simple feedback strategy with constant gain function α (see Section 3).

Suppose that the event Pro_+ occurs and that we have $\bar{a} = \tilde{a}$ for some $\tilde{a} \ge A_1$. In this case, we set $u = -2\tilde{a}q$ during the Main Act. We therefore have

$$\mathrm{E}[\mathrm{Cost}_M|\bar{a} = \tilde{a}] \le \mathcal{J}(\mathrm{CG}(2\tilde{a}), a; T, (1+\varepsilon)q_0) \quad \text{for any } \tilde{a} \ge A_1, a \in \mathbb{R}. \tag{6.15}$$

Define

$$A^* = \max\{|a|, A_1\}.$$

Combining (6.15) with Corollary 3.7 implies the following estimates:

$$\mathrm{E}[\mathrm{Cost}_M\,|\,10^n A^* < \bar{a} < 10^{n+1}A^*] < C_T 10^{2n}(A^*)^2 q_0^2 \quad \text{for any } a \in \mathbb{R}, n \ge 0, \tag{6.16}$$

and

$$\mathrm{E}[\mathrm{Cost}_M|A_1 \le \bar{a} < 10|a|] < C_T|a|^2 q_0^2 \quad \text{for any } a < -A_1/10. \tag{6.17}$$

Now let $0 < \delta \ll 1$ be a sufficiently small parameter depending on ε. Using Corollary 3.7 again, we get

$$\begin{aligned}&\mathrm{E}[\mathrm{Cost}_M|(A_1 \le \bar{a} < 10a) \text{ AND } (|a - \bar{a}| > \delta a)] \\ &\qquad\qquad < C_T\,a^2\,e^{2aT}\,q_0^2 \quad \text{for any } a \ge A_1/10,\end{aligned} \tag{6.18}$$

and

$$\text{(6.19)} \quad \mathrm{E}[\mathrm{Cost}_M \,|\, (\bar{a} \geq A_1) \text{ AND } (|\bar{a} - a| < \delta a)] < (1 + 4a^2(1+\delta)^2)\,\frac{q_0^2 + T}{2a(1 - C\delta)} \quad \text{for any } a \geq A_1/10.$$

We assume for now that $a > A_1/10$. Taking A_1 sufficiently large and δ sufficiently small (both depending on ε) in (6.19) gives

$$\mathrm{E}[\mathrm{Cost}_M \,|\, (\bar{a} \geq A_1) \text{ AND } (|\bar{a} - a| < \delta a)] < 2a\,(q_0^2 + T)\,(1 + \varepsilon).$$

Combining this with Lemma 3.4 gives

$$\mathrm{E}[\mathrm{Cost}_M \,|\, (\bar{a} \geq A_1) \text{ AND } (|\bar{a} - a| < \delta a)] < (1 + C\varepsilon)\, \mathcal{J}_0(a, T, q_0)$$

provided A_1 is sufficiently large depending on ε and T. This implies

$$\text{(6.20)} \quad \mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{\bar{a} \geq A_1} \cdot \mathbb{1}_{|a - \bar{a}| < \delta a}] < (1 + C\varepsilon) \cdot \mathcal{J}_0(a; T, q_0) \cdot \mathrm{Prob}\,(\bar{a} \geq A_1) \quad \text{for any } a > A_1/10.$$

We continue to assume that $a > A_1/10$. Taking ε sufficiently small depending on T, we apply Lemma 4.6 to get

$$\text{(6.21)} \qquad \mathrm{Prob}\,(|a - \bar{a}| > \delta a) \leq C \exp(-c_{\varepsilon,\delta}\, q_0^2\, a).$$

Combining (6.21) and (6.18) gives

$$\mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{A_1 \leq \bar{a} < 10a} \cdot \mathbb{1}_{|\bar{a} - a| > \delta a}] \leq C_T\, a^2\, q_0^2 \exp(2a(T - c_{\varepsilon,\delta}\, q_0^2)).$$

Taking q_{big} sufficiently large depending on ε and T (recall that δ is determined by ε), we get

$$\text{(6.22)} \qquad \mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{A_1 \leq \bar{a} < 10a} \cdot \mathbb{1}_{|\bar{a} - a| > \delta a}] < \varepsilon \quad \text{for any } a \geq A_1/10.$$

We continue to assume that $a > A_1/10$. Provided ε is sufficiently small depending on T, Lemma 4.5 gives

$$\mathrm{Prob}\,(\bar{a} > 10^n a) \leq C \exp(-c_\varepsilon\, q_0^2\, 10^n\, a) \quad \text{for any } n \geq 1.$$

Combining this with (6.16) gives

$$\mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{10^{n+1} a > \bar{a} > 10^n a} \cdot \mathbb{1}_{\bar{a} \geq A_1}] \leq C_T\, 10^{2n}\, a^2\, q_0^2 \exp(-c_\varepsilon\, q_0^2\, 10^n\, a) \quad \text{for any } n \geq 1.$$

Taking q_{big} to be sufficiently large depending on ε and T and summing over n gives

$$\text{(6.23)} \qquad \mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{\bar{a} > 10a} \cdot \mathbb{1}_{\bar{a} \geq A_1}] < \varepsilon \quad \text{for any } a > A_1/10.$$

Combining (6.20), (6.22), (6.23), and (6.9), we get

$$\text{(6.24)} \quad \mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{\bar{a} \geq A_1}] < \mathcal{J}_0(a; T, q_0) \cdot (\mathrm{Prob}\,(\bar{a} > A_1) + C_T\, \varepsilon) \quad \text{for any } a > A_1/10.$$

We now assume that $|a| \leq A_1/10$. Lemma 4.5 implies that

$$\text{Prob}\,(\bar{a} \geq 10^n A_1) \leq C \exp(-c_\varepsilon q_0^2 10^n A_1) \quad \text{for any } n \geq 0 \tag{6.25}$$

provided ε is sufficiently small depending on T. We combine (6.25) and (6.16) to get

$$\begin{aligned} \mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{\bar{a} \geq A_1}] &= \sum_{n\geq 0} \mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{10^{n+1}A_1 > \bar{a} \geq 10^n A_1}] \\ &\leq C_T \sum_{n \geq 0} 10^{2n} A_1^2 q_0^2 \exp(-c_\varepsilon q_0^2 10^n A_1). \end{aligned} \tag{6.26}$$

Taking q_{big} sufficiently large depending on ε and T gives

$$\mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{\bar{a} \geq A_1}] \leq \frac{\varepsilon}{A_1} \quad \text{for any } |a| \leq A_1/10. \tag{6.27}$$

Lemma 3.4 implies that there exists $\tilde{A}$, depending only on T, such that

$$\mathcal{J}_0(a; T, q_0) > c_T\,|a|^{-1} \quad \text{for any } a < -\tilde{A}, \tag{6.28}$$

$$\mathcal{J}_0(a; T, q_0) > c_T \quad \text{for any } a > -\tilde{A}. \tag{6.29}$$

Combining (6.27)–(6.29) gives

$$\mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{\bar{a} \geq A_1}] \leq C_T\,\varepsilon\,\mathcal{J}_0(a; T, q_0) \quad \text{for any } |a| \leq A_1/10. \tag{6.30}$$

We now assume that $a < -A_1/10$. Combining (6.16) with Lemma 4.5, we have

$$\begin{aligned} &\mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{10^{n+1}|a| \geq \bar{a} > 10^n|a|}] \\ &\qquad \leq C_T\, 10^{2n}\, |a|^2 q_0^2 \exp(-c_\varepsilon q_0^2 10^n |a|) \quad \text{for any } n \geq 1. \end{aligned}$$

Taking q_{big} sufficiently large depending on ε and T and summing over $n \geq 1$ gives

$$\mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{\bar{a} > 10|a|}] \leq \frac{\varepsilon}{|a|}. \tag{6.31}$$

Combining (6.17) with Lemma 4.2 gives

$$\mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{A_1 \leq \bar{a} < 10|a|}] < C_{T,\varepsilon}\, |a|^2 q_0^2 \exp(-c_\varepsilon q_0^2 |a|).$$

Taking q_{big} sufficiently large depending on ε and T gives

$$\mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{A_1 \leq \bar{a} < 10|a|}] < \frac{\varepsilon}{|a|}. \tag{6.32}$$

Combining (6.28), (6.31) and (6.32) gives

$$\mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{\bar{a} \geq A_1}] < C_T\,\varepsilon\,\mathcal{J}_0(a; T, q_0) \quad \text{for any } a < -A_1/10 \tag{6.33}$$

provided A_1 is sufficiently large depending on T.

Combining (6.30) and (6.33) gives

$$\mathrm{E}[\mathrm{Cost}_M \cdot \mathbb{1}_{\bar{a} \geq A_1}] < C_T\,\varepsilon\,\mathcal{J}_0(a; T, q_0) \quad \text{for any } a \leq A_1/10. \tag{6.34}$$

6.2.3. COST_M when we believe that a is bounded, positive. Suppose that the event PRO_+ occurs and that $\bar{a} = \tilde{a}$ for some $0 < \tilde{a} < A_1$. This determines the time at which the Main Act begins; denote this time by s, and note that $q(s) = (1+\varepsilon)q_0$. In this case, we set as our control variable

$$u(t) = -\kappa(T-t,\tilde{a}) \cdot q(t)$$

during the Main Act (which lasts from time s until time T), and thus

$$\mathrm{E}[\mathrm{COST}_M \,|\, \bar{a} = \tilde{a}] \le \mathcal{J}(\sigma_{\mathrm{opt}}(\tilde{a}), a; T, (1+\varepsilon)q_0) \quad \text{for any } a \in \mathbb{R}, 0 < \tilde{a} < A_1. \tag{6.35}$$

By Lemma 3.6, we have

$$\mathcal{J}(\sigma_{\mathrm{opt}}(\tilde{a}), a; T, (1+\varepsilon)q_0) = h(\tilde{a}, a)(1+\varepsilon)^2 q_0^2 + j(\tilde{a}, a) \quad \text{for any } a, \tilde{a} \in \mathbb{R}$$

and

$$\mathcal{J}_0(a; T, q_0) = h(a,a)\, q_0^2 + j(a,a) \quad \text{for any } a \in \mathbb{R}, \tag{6.36}$$

where h and j are smooth, positive functions (depending on T). We therefore have

$$\begin{aligned}\mathcal{J}(\sigma_{\mathrm{opt}}(\tilde{a}), a; T, (1+\varepsilon)q_0) \le \mathcal{J}_0(a; T, q_0) + |j(\tilde{a},a) - j(a,a)| + C h(a,a)\varepsilon q_0^2 \\ + C q_0^2 |h(\tilde{a},a) - h(a,a)|.\end{aligned}$$

Since h and j are smooth functions, we have

$$\begin{aligned}\mathcal{J}(\sigma_{\mathrm{opt}}(\tilde{a}), a; T, (1+\varepsilon)q_0) \le \mathcal{J}_0(a; T, q_0) + C_{A_1,T}\, q_0^2 |a - \tilde{a}| \\ + C\varepsilon q_0^2 h(a,a) \quad \text{for any } a, \tilde{a} \in [0, 10A_1]\end{aligned}$$

and

$$h(a,a) > c_{A_1,T} \quad \text{for any } a \in [0, 10A_1].$$

Taking $0 < \delta \ll 1$ to be sufficiently small depending on ε, A_1 and T gives

$$\begin{aligned}&\mathcal{J}(\sigma_{\mathrm{opt}}(\tilde{a}), a; T, (1+\varepsilon)q_0) \\ &\quad < \mathcal{J}_0(a; T, q_0) + C\varepsilon q_0^2 h(a,a) \quad \text{for any } |a - \tilde{a}| < \delta a,\ a \in [a_{\mathrm{tiny}}, 10A_1].\end{aligned}$$

Combining this with (6.35), (6.36) gives

$$\begin{aligned}&\mathrm{E}[\mathrm{COST}_M \,|\, 0 < \bar{a} < A_1 \text{ AND } |a - \bar{a}| < \delta a] \\ &\qquad \le (1 + C\varepsilon) \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \in [a_{\mathrm{tiny}}, 10A_1].\end{aligned} \tag{6.37}$$

We have therefore shown

$$\begin{aligned}&\mathrm{E}[\mathrm{COST}_M \cdot \mathbb{1}_{0<\bar{a}<A_1} \cdot \mathbb{1}_{|a-\bar{a}|<\delta a}] \\ &\qquad < (1 + C\varepsilon)\, \mathcal{J}_0(a; T, q_0)\, \mathrm{Prob}\,(0 < \bar{a} \le A_1) \quad \text{for any } a \in [a_{\mathrm{tiny}}, 10A_1].\end{aligned} \tag{6.38}$$

We combine (6.35) and Corollary 3.8 to get

$$\mathrm{E}[\mathrm{COST}_M \,|\, \bar{a} = \tilde{a}] \le C_T\, A_1^2 e^{2|a|T} q_0^2 \quad \text{for any } a \in \mathbb{R}, 0 < \tilde{a} < A_1. \tag{6.39}$$

By Lemma 4.6, we have

$$\mathrm{Prob}\,(|a - \bar{a}| > \delta a) \le C \exp(-c_{\varepsilon,\delta}\, q_0^2 a) \quad \text{for any } a > a_{\mathrm{tiny}}, \tag{6.40}$$

$$\mathrm{Prob}\,(0 < \bar{a} < A_1) \le C \exp(-c_\varepsilon\, q_0^2 a) \quad \text{for any } a \ge 10A_1. \tag{6.41}$$

By (6.2), we have

$$\begin{aligned}\mathrm{Prob}\,(0 < \bar{a} \le A_1) &\le \mathrm{Prob}\,(\textsc{Pro}_+)\\ &\le C\,\exp(-c_\varepsilon\, q_0^2\,(|a|+1)) \quad \text{for any } a \le a_{\text{tiny}}.\end{aligned} \tag{6.42}$$

Combining (6.39), (6.41), and (6.42) gives

$$\begin{aligned}&\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{0<\bar{a}<A_1}]\\ &\qquad < C\,A_1^2 q_0^2 \exp(2|a|T - c_\varepsilon q_0^2\,(|a|+1)) \quad \text{for any } a \notin [a_{\text{tiny}}, 10A_1].\end{aligned}$$

Combining (6.39) and (6.40) gives

$$\begin{aligned}&\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{0<\bar{a}<A_1} \cdot \mathbb{1}_{|a-\bar{a}|>\delta a}]\\ &\qquad < CA_1^2 q_0^2 \exp(2aT - c_{\varepsilon,\delta}\, q_0^2\, a) \quad \text{for any } a \in [a_{\text{tiny}}, 10A_1].\end{aligned}$$

Taking q_{big} to be sufficiently large depending on ε and T (recall that A_1 is determined by ε and T, and δ is determined by ε, A_1 and T), we get

$$\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{0<\bar{a}<A_1}] < \varepsilon \cdot \min\{1, |a|^{-1}\} \quad \text{for any } a \notin [a_{\text{tiny}}, 10A_1], \tag{6.43}$$

$$\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{0<\bar{a}<A_1} \cdot \mathbb{1}_{|a-\bar{a}|>\delta a}] < \varepsilon \quad \text{for any } a \in [a_{\text{tiny}}, 10A_1]. \tag{6.44}$$

As a consequence of Lemma 3.4, we have

$$\min\{|a|^{-1}, 1\} < C_T \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \in \mathbb{R}, \tag{6.45}$$

$$1 < C_T \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \ge 0. \tag{6.46}$$

Therefore (6.43) implies

$$\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{0<\bar{a}<A_1}] < C_T\, \varepsilon\, \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \notin [a_{\text{tiny}}, 10A_1], \tag{6.47}$$

and (6.38) and (6.44) imply

$$\begin{aligned}&\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{0<\bar{a}<A_1}]\\ &\qquad < \mathcal{J}_0(a; T, q_0)\,(\mathrm{Prob}\,(0 < \bar{a} < A_1) + C_T\,\varepsilon) \quad \text{for any } a \in [a_{\text{tiny}}, 10A_1].\end{aligned} \tag{6.48}$$

6.2.4. $\textsc{Cost}_M$ when we believe that a is large, negative. Recall that CG(0) denotes the simple feedback strategy with constant gain function 0. This is simply the strategy in which we set the control variable equal to zero for the entire game.

In the event that $\bar{a} = \tilde{a}$ for some $\tilde{a} \le -A_1$, we play the strategy CG(0) during the Main Act. We therefore have

$$\mathrm{E}[\textsc{Cost}_M \,|\, \bar{a} \le -A_1] \le \mathcal{J}(\mathrm{CG}(0), a; T, q_0) \quad \text{for any } a \in \mathbb{R}. \tag{6.49}$$

Taking A_1 to be sufficiently large depending on T, we apply Corollary 3.7 to get

$$\mathrm{E}[\textsc{Cost}_M \,|\, \bar{a} \le -A_1] \le \frac{1}{2\,|a|}\,(q_0^2 + T) \quad \text{for any } a \le -A_1/10. \tag{6.50}$$

By Lemma 3.5,

$$\frac{1}{2|a|}(q_0^2+T) \le (1+\varepsilon)\,\mathcal{J}_0(a;T,q_0) \quad \text{for any } a \le -A_1/10 \tag{6.51}$$

provided A_1 is sufficiently large depending on ε and T.

Combining (6.50) and (6.51) gives

$$\begin{aligned}\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{\bar a \le -A_1}] \\ \le (1+\varepsilon)\cdot \mathcal{J}_0(a;T,q_0)\cdot \mathrm{Prob}\,(\bar a \le -A_1) \quad \text{for any } a \le -A_1/10.\end{aligned}\tag{6.52}$$

Combining (6.49) and Corollary 3.7 gives

$$\mathrm{E}[\textsc{Cost}_M \,|\, \bar a \le -A_1] \le C_T\, e^{2|a|T}\, q_0^2 \quad \text{for any } a > -A_1/10. \tag{6.53}$$

Combining Lemma 4.6 and (6.3) gives

$$\mathrm{Prob}\,(\bar a \le -A_1) \le C\,\exp(-c_\varepsilon\, q_0^2\,(|a|+1)) \quad \text{for any } a > -A_1/10. \tag{6.54}$$

Inequalities (6.53) and (6.54) imply that

$$\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{\bar a \le -A_1}] \le C_T\, q_0^2\, \exp(2|a|T - c_\varepsilon\, q_0^2\,(|a|+1)) \quad \text{for any } a > -A_1/10.$$

Taking q_{big} sufficiently large depending on ε and T, we get

$$\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{\bar a \le -A_1}] \le \varepsilon \min\{1, |a|^{-1}\} \quad \text{for any } a > -A_1/10.$$

Combining this with (6.45) gives

$$\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{\bar a \le -A_1}] \le C_T\,\varepsilon\,\mathcal{J}_0(a;T,q_0) \quad \text{for any } a > -A_1/10. \tag{6.55}$$

6.2.5. $\textsc{Cost}_M$ when we believe that a is bounded, negative. Proceeding as in the proofs of equations (6.47) and (6.48) in Section 6.2.3, we can show that

$$\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{-A_1 < \bar a < 0}] < C_T\,\varepsilon\,\mathcal{J}_0(a;T,q_0) \quad \text{for any } a \notin [-10A_1, -a_{\mathrm{tiny}}] \tag{6.56}$$

and

$$\begin{aligned}\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{-A_1 < \bar a < 0}] \\ \le \mathcal{J}_0(a;T,q_0)\,(\mathrm{Prob}\,(-A_1 \le \bar a < 0) + C_T\varepsilon) \quad \text{for any } a \in [-10A_1, -a_{\mathrm{tiny}}].\end{aligned}\tag{6.57}$$

6.2.6. Proof of Theorem 2.3. We collect the estimates (6.14), (6.11), (6.24), (6.34), (6.48), (6.47), (6.52), (6.55), (6.57), and (6.56) from Section 6.2. We assume that q_{big} and A_1 are large enough depending on ε and T for all of these estimates to hold.

$$\begin{aligned}
&\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{\textsc{Pro}_{\max}}] \le \mathcal{J}_0(a;T,q_0)\,(\mathrm{Prob}\,(\textsc{Pro}_{\max}) + C_T\,\varepsilon^{1/4}) \quad \text{for any } |a| \le a_{\mathrm{small}},\\
&\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{\textsc{Pro}_{\max}}] \le \varepsilon C_T\,\mathcal{J}_0(a;T,q_0) \quad \text{for any } |a| \ge a_{\mathrm{small}},\\
&\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{\bar a \ge A_1}] \le \mathcal{J}_0(a;T,q_0)\,(\mathrm{Prob}\,(\bar a > A_1) + C_T\,\varepsilon) \quad \text{for any } a > A_1/10,\\
&\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{\bar a \ge A_1}] \le C_T\,\varepsilon\,\mathcal{J}_0(a;T,q_0) \quad \text{for any } a \le A_1/10,\\
&\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{0 < \bar a < A_1}]\\
&\qquad \le \mathcal{J}_0(a;T,q_0)\,(\mathrm{Prob}\,(0 < \bar a \le A_1) + C_T\,\varepsilon) \quad \text{for any } a \in [a_{\mathrm{tiny}}, 10A_1],
\end{aligned}$$

$$
\begin{aligned}
&\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{0<\bar{a}<A_1}] \le C_T \varepsilon \mathcal{J}_0(a;T,q_0) \quad \text{for any } a \notin [a_{\text{tiny}}, 10A_1],\\
&\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{\bar{a}\le -A_1}] \le \mathcal{J}_0(a;T,q_0)(\mathrm{Prob}\,(\bar{a} < -A_1) + \varepsilon) \quad \text{for any } a \le -A_1/10,\\
&\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{\bar{a}\le -A_1}] \le C_T \varepsilon \mathcal{J}_0(a;T,q_0) \quad \text{for any } a > -A_1/10,\\
&\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{-A_1<\bar{a}<0}]\\
&\qquad \le \mathcal{J}_0(a;T,q_0)(\mathrm{Prob}\,(-A_1 \le \bar{a} < 0) + C_T\varepsilon) \quad \text{for any } a \in [-10A_1, -a_{\text{tiny}}],\\
&\mathrm{E}[\textsc{Cost}_M \cdot \mathbb{1}_{-A_1<\bar{a}<0}] < C_T \varepsilon \mathcal{J}_0(a;T,q_0) \quad \text{for any } a \notin [-10A_1, -a_{\text{tiny}}].
\end{aligned}
$$

Combining these estimates proves

$$
\mathrm{E}[\textsc{Cost}_M] \le \mathcal{J}_0(a;T,q_0)\,(1 + C_T \varepsilon^{1/4}) \quad \text{for any } a \in \mathbb{R}. \tag{6.58}
$$

Combining (6.58) with (6.5) and taking q_{big} to be large enough so that (6.5) holds proves that

$$
\mathcal{J}(\mathrm{L}q\mathrm{S}, a; T, q_0) = \mathrm{E}[\textsc{Cost}_P] + \mathrm{E}[\textsc{Cost}_M] \le \mathcal{J}_0(a;T,q_0)(1 + C_T \varepsilon^{1/4}) \quad \text{for any } a \in \mathbb{R}.
$$

This proves Theorem 2.3.

7. The large *a* strategy

In this section we prove Theorem 2.4.

We fix a time horizon $T > 0$ and a starting position $q_0 > 0$.

Let $\varepsilon > 0$ be given and let $A \ge 1$ be a sufficiently large number depending on ε, T, q_0. Without loss of generality we assume that ε is sufficiently small depending on T, q_0.

By Theorem 2.3, there exists a number $q_0^* \ge 2q_0$ depending on q_0 and T and a strategy BR (here “BR” stands for Bounded Regret) such that

$$
\mathcal{J}(\mathrm{BR}, a; T, q_0^*) \le 2 \cdot \mathcal{J}_0(a;T,q_0^*) \quad \text{for any } a \in \mathbb{R}.
$$

By Remark 3.3, we have

$$
\mathcal{J}(\mathrm{BR}, a; T, q_0^*) \le C_{T,q_0} \cdot \mathcal{J}_0(a;T,q_0) \quad \text{for any } a \in \mathbb{R}. \tag{7.1}
$$

The strategy BR for time horizon T and starting position q_0^* gives rise to a strategy BR_- for time horizon T and starting position $-q_0^*$ satisfying

$$
\mathcal{J}(\mathrm{BR}_-, a; T, -q_0^*) = \mathcal{J}(\mathrm{BR}, a; T, q_0^*)
$$

(see Section 2 for details). For the remainder of Section 7 we will write BR to denote both of these strategies; it will be clear from context which strategy we are referring to.

We now define the strategy LaS. For the remainder of Section 7, we write q and u to denote $q^{\mathrm{L}a\mathrm{S}}$ and $u^{\mathrm{L}a\mathrm{S}}$.

Testing Epoch.

We let τ_C denote the first time $t \in (0,T)$ for which $q(t) = (1+\varepsilon)q_0$ or $q(t) = -q_0^*$, if such a time exists. If no such time exists we set $\tau_C = T$. The Testing Epoch ends when we reach time τ_C. During the Testing Epoch we set $u = 0$. If $\tau_C = T$, then the game is over when we reach time τ_C. If $\tau_C < T$, then at time τ_C we enter the Control Epoch.

CONTROL EPOCH.

In a moment, we will define a stopping time τ_D. With probability 1 we have $\tau_D \in (\tau_C, T]$. If the Control Epoch occurs, then it lasts from time τ_C until time τ_D.

In the event that we enter the Control Epoch at some time $\tau_C \in (0, T)$ for which $q(\tau_C) = -q_0^*$, we set $\tau_D = T$ so that the Control Epoch begins at time τ_C and position $-q_0^*$ and lasts until the game ends at time T. In this case, during the Control Epoch we execute the strategy BR.

In the event that we enter the Control Epoch at some time $\tau_C \in (0, T)$ for which $q(\tau_C) = (1 + \varepsilon) q_0$, we define a random variable

$$\bar{a} = \frac{\log(1 + \varepsilon)}{\tau_C}.$$

We then set

$$u(t) = -2\bar{a} q(t)$$

during the Control Epoch. In this case, we define the stopping time τ_D to be equal to the first time $t \in (\tau_C, T)$ for which $|q(t)| = q_0^*$ if such a time exists and equal to T if no such time exists.

In either case, if $\tau_D = T$, then the Control Epoch ends along with the game at time T. If $\tau_D < T$, then at time τ_D we enter the Disaster Mitigation Epoch.

DISASTER MITIGATION EPOCH.

If we enter the Disaster Mitigation Epoch at some time $\tau_D \in (\tau_C, T)$, then we have $|q(\tau_D)| = q_0^*$. We then execute the strategy BR from time τ_D until the game ends at time T.

This concludes the definition of the strategy LaS.

We define three random variables:

$$\mathrm{COST}_T = \int_0^{\tau_C} q^2(t)\, dt, \quad \mathrm{COST}_C = \int_{\tau_C}^{\tau_D} (q^2(t) + u^2(t))\, dt,$$

and

$$\mathrm{COST}_D = \int_{\tau_D}^{T} (q^2(t) + u^2(t))\, dt.$$

These are, respectively, the costs incurred during the Testing Epoch, the Control Epoch, and the Disaster Mitigation Epoch for a given realization of the noise. Note that they all depend on a.

Let $X \subset (0, \infty)$ be an arbitrary subset. We will be interested in events of the form "$\tau_C < T$ and $q(\tau_C) = (1 + \varepsilon) q_0$ and $\bar{a} \in X$". We will abbreviate this by writing just "$\bar{a} \in X$"; note that $\bar{a}$ is undefined if either $\tau_C = T$ or if $q(\tau_C) = -q_0^*$.

By Lemma 3.4, provided A is sufficiently large depending on T we have

$$c_{T,q_0} \le \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \ge A. \tag{7.2}$$

We will make use of this fact throughout this section.

7.1. The cost during the Testing Epoch

By Lemma 4.3, provided ε is sufficiently small depending on T we have

$$\mathrm{E}[\mathrm{Cost}_T] < C_{T,q_0} \cdot \varepsilon^{1/4} \quad \text{for any } a \in \mathbb{R} \tag{7.3}$$

(recall that q_0^* is determined by T, q_0). Combining this with (7.2) gives

$$\mathrm{E}[\mathrm{Cost}_T] < C_{T,q_0} \cdot \varepsilon^{1/4} \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \geq A. \tag{7.4}$$

7.2. The cost during the Control Epoch

Our goal in this section is to control the expected value of the random variable Cost_C, defined above.

Observe that

$$\mathrm{E}[\mathrm{Cost}_C] = \mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{\bar{a}>0}] + \mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{q(\tau_C)=-q_0^*}]. \tag{7.5}$$

In the event that we enter the Control Epoch at some time $\tau_C \in (0, T)$ for which $q(\tau_C) = -q_0^*$, we play the strategy BR starting from position $-q_0^*$ and time τ_C until the game ends at time T. By (7.1), we therefore have

$$\mathrm{E}[\mathrm{Cost}_C \,|\, (q(\tau_C) = -q_0^*)] \leq C_{T,q_0} \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \in \mathbb{R}. \tag{7.6}$$

Provided A is sufficiently large depending on ε, and ε is sufficiently small depending on T, Lemma 4.3 implies

$$\mathrm{Prob}\,(q(\tau_C) = -q_0^*) \leq C_{q_0} \cdot \varepsilon^{1/4} \quad \text{for any } a \geq A.$$

Therefore,

$$\mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{q(\tau_C)=-q_0^*}] \leq C_{T,q_0} \cdot \varepsilon^{1/4} \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \geq A. \tag{7.7}$$

By Corollary 3.8,

$$\mathcal{J}_0(a; T, q_0) \leq C_{T,q_0} \cdot A \quad \text{for any } a \leq A.$$

Combining this with (7.6) gives

$$\mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{q(\tau_C)=-q_0^*}] \leq C_{T,q_0} \cdot A \quad \text{for any } a \leq A. \tag{7.8}$$

We will make use of the following two observations throughout the remainder of this section. First, we note that

$$\mathrm{E}[\mathrm{Cost}_C \,|\, \bar{a} = \tilde{a}] \leq C_{T,q_0} (\tilde{a}^2 + 1) \quad \text{for any } \tilde{a} > 0, a \in \mathbb{R}. \tag{7.9}$$

This follows by observing that, in the event that we enter the Control Epoch, we have $|q| \leq q_0^*$ and $u = -2\bar{a}q$ during the Control Epoch with probability 1. Second, we define

$$A^* = \max\{A, |a|\}.$$

Taking ε sufficiently small depending on T, Lemma 4.5 implies that

$$\mathrm{Prob}\,(\bar{a} > 10^n A^*) \leq C \exp(-c_\varepsilon q_0^2 10^n A^*) \quad \text{for any } n \geq 1, a \in \mathbb{R}. \tag{7.10}$$

We now introduce a parameter $1 \gg \delta > 0$ that will be chosen below to be sufficiently small depending on ε. Note that since δ is determined by ε, we are allowed to take A sufficiently large depending on δ.

Assume that $a \geq A$. Observe that

$$\begin{aligned}\mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{\bar{a}>0}] = \mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{|a-\bar{a}|<\delta a}] + \mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{\bar{a}<(1-\delta)a}] \\ + \mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{\bar{a}>(1+\delta)a}].\end{aligned} \tag{7.11}$$

In the event that $\bar{a} = \tilde{a}$ for some $\tilde{a} > 0$, then during the Control Epoch we play the constant gain strategy $\mathrm{CG}(2\tilde{a})$ (see Section 3, specifically the discussion before Corollary 3.7, for the definition of the strategy CG). Therefore, by Corollary 3.7,

$$\mathrm{E}[\mathrm{Cost}_C \,|\, (|a-\bar{a}| < \delta a)] < (1 + 4a^2(1+\delta)^2)\, \frac{q_0^2 + T}{2a(1 - C\delta)}.$$

Taking A sufficiently large and δ sufficiently small, both depending on ε, we have

$$\mathrm{E}[\mathrm{Cost}_C \,|\, (|a-\bar{a}| < \delta a)] < 2a(1+\varepsilon)\,(q_0^2 + T).$$

By Lemma 3.4, provided A is sufficiently large depending on ε and T, we have

$$\mathcal{J}_0(a; T, q_0) \geq 2a(1-\varepsilon)\,(q_0^2 + T).$$

We deduce that

$$\mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{|a-\bar{a}|<\delta a}] < (1 + C\varepsilon) \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \geq A. \tag{7.12}$$

We continue to assume that $a \geq A$. By (7.9), we have

$$\mathrm{E}[\mathrm{Cost}_C \,|\, \bar{a} < (1-\delta)a] \leq C_{T,q_0} \cdot a^2.$$

Provided A is sufficiently large depending on ε, and ε is sufficiently small depending on T, Lemma 4.6 implies that

$$\mathrm{Prob}\,(|\bar{a} - a| \geq \delta a) \leq C \exp(-c_{\varepsilon,\delta}\, q_0^2\, a). \tag{7.13}$$

Therefore,

$$\mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{\bar{a}<(1-\delta)a}] < C_{T,q_0} \cdot a^2 \cdot \exp(-c_{\varepsilon,\delta}\, q_0^2\, a).$$

Taking A sufficiently large depending on ε, δ, q_0 and T gives

$$\mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{\bar{a}<(1-\delta)a}] < \varepsilon \quad \text{for any } a \geq A. \tag{7.14}$$

We continue to assume that $a \geq A$. By (7.9), we have

$$\mathrm{E}[\mathrm{Cost}_C \,|\, (1+\delta)a < \bar{a} < 10a] < C_{T,q_0} \cdot a^2, \tag{7.15}$$

$$\mathrm{E}[\mathrm{Cost}_C \,|\, 10^n a < \bar{a} < 10^{n+1} a] < C_{T,q_0} \cdot 10^{2n}\, a^2 \quad \text{for any } n \geq 1. \tag{7.16}$$

Since

$$\mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{\bar{a}>(1+\delta)a}] = \mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{(1+\delta)a<\bar{a}<10a}] + \sum_{n=1}^{\infty} \mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{10^n a<\bar{a}<10^{n+1}a}],$$

equations (7.10) (our assumption that $a \geq A$ implies that $A^* = a$), (7.13), (7.15) and (7.16) imply that

$$\mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{\bar{a} > (1+\delta)a}] < C_{T,q_0} \cdot a^2 \exp(-c_{\varepsilon,\delta}\, q_0^2\, a) + \sum_{n=1}^{\infty} C_{T,q_0} \cdot 10^{2n} a^2 \exp(-c_\varepsilon\, q_0^2\, 10^n a).$$

Taking A sufficiently large depending on ε, δ, T and q_0 gives

$$\mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{\bar{a} > (1+\delta)a}] < \varepsilon \quad \text{for any } a \geq A. \tag{7.17}$$

Combining (7.11), (7.12), (7.14) and (7.17) gives

$$\mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{\bar{a} > 0}] < (1 + C\varepsilon) \cdot \mathcal{J}_0(a; T, q_0) + C'\varepsilon \quad \text{for any } a \geq A;$$

the inequality (7.2) then implies that

$$\mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{\bar{a} > 0}] < (1 + C_{T,q_0} \cdot \varepsilon) \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \geq A. \tag{7.18}$$

We now suppose that $|a| < A$. Note that

$$\mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{\bar{a} > 0}] = \mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{\bar{a} < 10A}] + \sum_{n=1}^{\infty} \mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{10^n A < \bar{a} < 10^{n+1} A}]. \tag{7.19}$$

Equation (7.9) implies

$$\mathrm{E}[\mathrm{Cost}_C \,|\, \bar{a} < 10A] < C_{T,q_0} A^2, \tag{7.20}$$

$$\mathrm{E}[\mathrm{Cost}_C \,|\, 10^n A < \bar{a} < 10^{n+1} A] \leq C_{T,q_0}\, 10^{2n} A^2 \quad \text{for any } n \geq 1. \tag{7.21}$$

We combine (7.19)– (7.21) with (7.10) (our assumption $|a| < A$ implies that $A^* = A$) to deduce that

$$\mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{\bar{a} > 0}] \leq C_{T,q_0} A^2 + \sum_{n=1}^{\infty} C_{T,q_0}\, 10^{2n} A^2 \exp(-c_\varepsilon\, q_0^2\, 10^n A). \tag{7.22}$$

Taking A sufficiently large depending on ε and q_0 gives

$$\mathrm{E}[\mathrm{Cost}_C \cdot \mathbb{1}_{\bar{a} > 0}] < C_{T,q_0} \cdot A^2 \quad \text{for any } |a| < A. \tag{7.23}$$

We now assume that $a < -A$. Note that

$$\mathrm{Prob}\,(\bar{a} < 10|a|) = \mathrm{Prob}\,(0 < \bar{a} < 10|a|) \leq \mathrm{Prob}\,(\bar{a} > 0).$$

Applying Lemma 4.4, we get

$$\mathrm{Prob}\,(\bar{a} < 10|a|) \leq C_{\varepsilon,T,q_0} \cdot \exp(-c_\varepsilon\, q_0^2 |a|). \tag{7.24}$$

Note that

$$\mathrm{E}[\textsc{Cost}_C \cdot \mathbb{1}_{\bar{a}>0}] = \mathrm{E}[\textsc{Cost}_C \cdot \mathbb{1}_{\bar{a}<10|a|}] + \sum_{n=1}^{\infty} \mathrm{E}[\textsc{Cost}_C \cdot \mathbb{1}_{10^n|a|<\bar{a}<10^{n+1}|a|}].$$

We combine this with (7.9), (7.10) and (7.24) to get

$$\mathrm{E}[\textsc{Cost}_C \cdot \mathbb{1}_{\bar{a}>0}] \le C_{\varepsilon,T,q_0}|a|^2 \exp(-c_\varepsilon q_0^2 |a|) + C_{T,q_0} \sum_{n=1}^{\infty} 10^{2n} |a|^2 \exp(-c_\varepsilon q_0^2 10^n |a|).$$

Taking A sufficiently large depending on ε, T and q_0 gives

$$\mathrm{E}[\textsc{Cost}_C \cdot \mathbb{1}_{\bar{a}>0}] < \varepsilon \quad \text{for any } a < -A. \tag{7.25}$$

We combine (7.5), (7.7) and (7.18) to get that

$$\mathrm{E}[\textsc{Cost}_C] < (1 + C_{T,q_0} \cdot \varepsilon^{1/4})\, \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \ge A, \tag{7.26}$$

and we combine (7.5), (7.8), (7.23) and (7.25) to get

$$\mathrm{E}[\textsc{Cost}_C] < C_{T,q_0} \cdot A^2 \quad \text{for any } a \le A. \tag{7.27}$$

7.3. The cost during the disaster mitigation epoch

Our goal in this section is to control the expected value of the random variable $\textsc{Cost}_D$.

Let $\mathcal{D}$ denote the event that we reach the Disaster Mitigation Epoch, i.e., $\mathcal{D}$ is the event that $\tau_C < T$, $q(\tau_C) = (1+\varepsilon)q_0$, and $\tau_D < T$. For a subset $X \subseteq (0,\infty)$, we let $\mathcal{D}(X)$ denote the event that $\mathcal{D}$ occurs and $\bar{a} \in X$. Note that $\mathcal{D} = \mathcal{D}((0,\infty))$.

We remark that

$$\mathrm{E}[\textsc{Cost}_D] = \mathrm{E}[\textsc{Cost}_D \cdot \mathbb{1}_{\mathcal{D}}].$$

Suppose that we enter the Disaster Mitigation Epoch at some time $\tau_D \in (0,T)$, and suppose that $\bar{a} = \tilde{a}$ for some $\tilde{a} > 0$. Starting from position $|q| = q_0^*$ and time τ_D until time T, we execute the strategy BR. Therefore, by (7.1), we have

$$\mathrm{E}[\textsc{Cost}_D \,|\, \mathcal{D} \text{ AND } \bar{a} = \tilde{a}] \le C_{T,q_0} \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \in \mathbb{R}, \tilde{a} > 0. \tag{7.28}$$

We remark that the right-hand side of (7.28) is independent of $\tilde{a}$, and so we have

$$\mathrm{E}[\textsc{Cost}_D \,|\, \mathcal{D}(X)] \le C_{T,q_0} \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \in \mathbb{R}, X \subset (0,\infty).$$

Combining this with Corollary 3.8 gives

$$\mathrm{E}[\textsc{Cost}_D \,|\, \mathcal{D}(X)] \le C_{T,q_0} \cdot \max\{a, 1\} \quad \text{for any } a \in \mathbb{R}, X \subset (0,\infty). \tag{7.29}$$

In particular, taking $X = (0,\infty)$ implies that

$$\mathrm{E}[\textsc{Cost}_D] \le C_{T,q_0} \cdot A \quad \text{for any } a \le A. \tag{7.30}$$

We now show that the probability of reaching the Disaster Mitigation Epoch is small when $a \geq A$.

Assume $a \geq A$. Observe that

$$\mathrm{Prob}\,(\mathcal{D}((0, 3a/4))) \leq \mathrm{Prob}\,(0 < \bar{a} < 3a/4).$$

Provided A is sufficiently large depending on ε, and ε is sufficiently small depending on T, Lemma 4.6 implies that

$$\mathrm{Prob}\,(\mathcal{D}((0, 3a/4))) \leq C \exp(-c_\varepsilon q_0^2 a). \tag{7.31}$$

Now suppose that the event $\bar{a} = \tilde{a}$ occurs for some $\tilde{a} > 3a/4$. In this case, the probability that we enter the Disaster Mitigation Epoch is less than or equal to

$$\mathrm{Prob}\,(\exists t \in [0, T] : |\bar{q}(t)| \geq q_0^*),$$

where $\bar{q}\colon [0, T] \to \mathbb{R}$ is governed by

$$d\bar{q} = (a - 2\tilde{a})\bar{q}\,dt + dW_t, \quad \bar{q}(0) = (1 + \varepsilon)q_0.$$

Since $a - 2\tilde{a} < 0$, we apply Lemma 4.4 to get

$$\mathrm{Prob}\,(\exists t \in [0, T] : |\bar{q}(t)| \geq q_0^*) \leq C_{T,q_0} \cdot \exp(-c_{T,q_0} a).$$

This holds for any $\tilde{a} > 3a/4$; we deduce that

$$\mathrm{Prob}\,(\mathcal{D}((3a/4, \infty))) \leq C_{T,q_0} \cdot \exp(-c_{T,q_0} a). \tag{7.32}$$

Combining (7.29), (7.31) and (7.32) gives

$$\begin{aligned} \mathrm{E}[\textsc{Cost}_D] &= \mathrm{E}[\textsc{Cost}_D \cdot \mathbb{1}_{\mathcal{D}((0,3a/4))}] + \mathrm{E}[\textsc{Cost}_D \cdot \mathbb{1}_{\mathcal{D}((3a/4,\infty))}] \\ &\leq C_{T,q_0} \cdot a \cdot \exp(-c_\varepsilon q_0^2 a) + C_{T,q_0} \cdot a \cdot \exp(-c_{T,q_0} a) \end{aligned}$$

for any $a \geq A$. Taking A sufficiently large depending on ε, T and q_0 gives

$$\mathrm{E}[\textsc{Cost}_D] < \varepsilon \quad \text{for any } a \geq A.$$

Combining this with (7.2) gives

$$\mathrm{E}[\textsc{Cost}_D] < \varepsilon \cdot C_{T,q_0} \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \geq A. \tag{7.33}$$

7.4. Proof of Theorem 2.4

Note that

$$\mathcal{J}(\mathrm{LaS}, a; T, q_0) = \mathrm{E}[\textsc{Cost}_T] + \mathrm{E}[\textsc{Cost}_C] + \mathrm{E}[\textsc{Cost}_D]. \tag{7.34}$$

By (7.4), (7.26) and (7.33), we have

$$\mathcal{J}(\mathrm{LaS}, a; T, q_0) < (1 + C_{T,q_0} \cdot \varepsilon^{1/4}) \cdot \mathcal{J}_0(a; T, q_0) \quad \text{for any } a \geq A.$$

By (7.3), (7.27) and (7.30), we have

$$\mathcal{J}(\mathrm{L}a\mathrm{S}, a; T, q_0) < C_{T,q_0} \cdot A^2 \quad \text{for any } a \le A.$$

This proves Theorem 2.4.

Acknowledgments. We thank Amir Ali Ahmadi, Brittany Hamfeldt, Elad Hazan, Robert Kohn, Sam Otto, Allan Sly, and Melanie Weber for helpful conversations. We are grateful to the Air Force Office of Scientific Research, and Frederick Leve, for their generous support. The second named author thanks the Joachim Herz foundation for support. We thank the anonymous referees for helpful feedback.

Funding. This work was supported by AFOSR grant FA9550-19-1-0005 and by the Joachim Herz Foundation.

References

[1] Abbasi-Yadkori, Y. and Szepesvári, C.: Regret bounds for the adaptive control of linear quadratic systems. In *Proceedings of the 24th annual conference on learning theory (Budapest, 2011)*, pp. 1–26. Proc. Mach. Learn. Res. 19, PMLR, 2011.

[2] Abeille, M. and Lazaric, A.: Thompson sampling for linear-quadratic control problems. In *Proceedings of the 20th international conference on artificial intelligence and statistics (Fort Lauderdale, FL, 2017)*, pp. 1246–1254. Proc. Mach. Learn. Res. 54, PMLR, 2017.

[3] Åström, K. J.: *Introduction to stochastic control theory*. Mathematics in Science and Engineering 70, Academic Press, New York-London, 1970. Zbl 0226.93027 MR 0270799

[4] Bertsekas, D. P.: *Dynamic programming and optimal control. Vol. I*. Fourth edition. Athena Scientific, Belmont, MA, 2017. Zbl 1375.90299 MR 3644954

[5] Bertsekas, D. P.: *Dynamic programming and optimal control. Vol. II*. Fourth edition. Athena Scientific, Belmont, MA, 2012. Zbl 1298.90001 MR 3642732

[6] Brazy, D.: *Cockpit Voice Recorder DCA09MA026, Group chairman's factual report of investigation*. National Transportation Safety Board Docket no. SA-532, Exhibit no. 12, 2009.

[7] Carruth, J., Eggl, M. F., Fefferman, C. and Rowley, C. W.: Optimal agnostic control of unknown linear dynamics in a bounded parameter range. *Rev. Mat. Iberoam.* **41** (2025), no. 2, 651–744.

[8] Carruth, J., Eggl, M. F., Fefferman, C., Rowley, C. W. and Weber, M.: Controlling unknown linear dynamics with bounded multiplicative regret. *Rev. Mat. Iberoam.* **38** (2022), no. 7, 2185–2216. Zbl 1510.93373 MR 4526313

[9] Cesa-Bianchi, N. and Lugosi, G.: *Prediction, learning, and games*. Cambridge University Press, Cambridge, 2006. Zbl 1114.91001 MR 2409394

[10] Chen, X. and Hazan, E.: Black-box control for linear dynamical systems. In *Proceedings of 34th conference on learning theory (Boulder, CO, 2021)*, pp. 1114–1143. Proc. Mach. Learn. Res. 134, PMLR, 2021.

[11] Chen, X., Minasyan, E., Lee, J. D. and Hazan, E.: Regret guarantees for online deep control. In *Proceedings of the 5th annual learning for dynamics and control conference (Philadelphia, PA, 2023)*, 1032–1045. Proc. Mach. Learn. Res. 211, PMLR, 2023.

[12] Cohen, A., Koren, T. and Mansour, Y.: Learning linear-quadratic regulators efficiently with only $\sqrt{T}$ regret. In *Proceedings of the 36th international conference on machine learning (Long Beach, Ca, 2019)*, pp. 1300–1309. Proc. Mach. Learn. Res. 97, PMLR, 2019.

[13] Dean, S., Mania, H., Matni, N., Recht, B. and Tu, S.: Regret bounds for robust adaptive control of the linear quadratic regulator. In *NIPS'18: Proceedings of the 32th international conference on neural information processing systems*, 4192–4201. Curran Associates, NY, 2024.

[14] Dean, S., Tu, S., Matni, N. and Recht, B.: Safely learning to control the constrained linear quadratic regulator. In *2019 American control Conference (ACC)*, pp. 5582–5588. IEEE, 2019.

[15] Duchi, J., Hazan, E. and Singer, Y.: Adaptive subgradient methods for online learning and stochastic optimization. *J. Mach. Learn. Res.* **12** (2011), 2121–2159. Zbl 1280.68164 MR 2825422

[16] Faury, L., Russac, Y., Abeille, M. and Calauzènes, C.: Optimal regret bounds for generalized linear bandits under parameter drift. In *Proceedings in algorithmic learning theory*, 37 pp. Proc. Mach. Learn. Res. 132, PMLR, 2021.

[17] Fefferman, C., Guillén Pegueroles, B., Rowley, C. W. and Weber, M.: Optimal control with learning on the fly: a toy problem. *Rev. Mat. Iberoam.* **38** (2022), no. 1, 175–187. Zbl 1485.93628 MR 4382468

[18] Feller, W.: *An introduction to probability theory and its applications. Vol. II*. Second edition. John Wiley & Sons, New York-London-Sydney, 1971. Zbl 0219.60003 MR 0270403

[19] Furieri, L., Zheng, Y. and Kamgarpour, M.: Learning the globally optimal distributed LQ regulator. In *Proceedings of the 2nd conference on Learning for dynamics and control*, pp. 287–297. Proc. Mach. Learn. Res. 120, PMLR, 2020.

[20] Goel, G. and Hassibi, B.: Competitive control. *IEEE Trans. Automat. Control* **68** (2023), no. 9, 5162–5173. Zbl 07746851 MR 4639976

[21] Gurevich, D., Goswami, D., Fefferman, C. L. and Rowley, C. W.: Optimal control with learning on the fly: System with unknown drift. In *Proceedings of the 4th annual learning for dynamics and control conference*, pp. 870–880. Proc. Mach. Learn. Res. 168, PMLR, 2022.

[22] Hazan, E. and Singh, K.: Introduction to online nonstochastic control. Preprint, arXiv: 2211.09619v3, 2024.

[23] Jedra, Y. and Proutiere, A.: Minimal expected regret in linear quadratic control. In *Proceedings of the 25th international conference on artificial intelligence and statistics*, pp. 10234–10321. Proc. Mach. Learn. Res. 151, PMLR, 2022.

[24] Kargin, T., Lale, S., Azizzadenesheli, K., Anandkumar, A. and Hassibi, B.: Thompson sampling achieves $\tilde{O}(\sqrt{T})$ regret in linear quadratic control. In *Proceedings of 35th conference on learning theory*, pp. 3235-3284. Proc. Mach. Learn. Res. 178, PMLR, 2022.

[25] Kumar, R., Dean, S. and Kleinberg, R.: Online convex optimization with unbounded memory. In *NIPS'23: Proceedings of the 37th international conference on neural information processing systems*, article no. 1141, 26229–26270. Curran Associates, NY, 2024.

[26] Malik, D., Pananjady, A., Bhatia, K., Khamaru, K., Bartlett, P. L. and Wainwright, M. J.: Derivative-free methods for policy optimization: Guarantees for linear quadratic systems. In *Proceedings of the 22th international conference on artificial intelligence and statistics*, 2916–2925. Proc. Mach. Learn. Res. 89, PMLR, 2019.

[27] Mania, H., Tu, S. and Recht, B.: Certainty equivalence is efficient for linear quadratic control. In *NIPS'19: Proceedings of the 33th international conference on neural information processing systems*, article no. 911, 10154–10164. Curran Associates, NY, 2019.

[28] Martin, A., Furieri, L., Dörfler, F., Lygeros, J. and Ferrari-Trecate, G.: Safe control with minimal regret. In *Proceedings of the 4th annual learning for dynamics and control conference (Stanford, CA, 2022)*, pp. 726–738. Proc. Mach. Learn. Res. 168, PMLR, 2022.

[29] Minasyan, E., Gradu, P., Simchowitz, M. and Hazan, E.: Online control of unknown time-varying dynamical systems. In *NIPS'21: Proceedings of the 35th international conference on neural information processing systems*, article no. 1219, 15934–15945. Curran Associates, NY, 2024.

[30] Powell, W. B.: *Approximate dynamic programming. Solving the curses of dimensionality*. Wiley Ser. Probab. Stat., Wiley-Interscience John Wiley & Sons, Hoboken, NJ, 2007. Zbl 1156.90021 MR 2347698

[31] Robbins, H.: Some aspects of the sequential design of experiments. *Bull. Amer. Math. Soc.* **58** (1952), 527–535. Zbl 0049.37009 MR 0050246

[32] Simchowitz, M., Mania, H., Tu, S., Jordan, M. I. and Recht, B.: Learning without mixing: Towards a sharp analysis of linear system identification. In *Proceedings of the 31st conference on learning theory*, pp. 439–473. Proc. Mach. Learn. Res. 75, PMLR, 2018.

[33] Vermorel, J. and Mohri, M.: Multi-armed bandit algorithms and empirical evaluation. In *Machine Learning: ECML 2005*, 437–448. Lecture Notes in Computer Science 3720, Springer, Berlin, Heidelberg, 2005.

[34] Wagenmaker, A. and Jamieson, K.: Active learning for identification of linear dynamical systems. In *Proceedings of 33th conference on learning theory*, pp. 3487–3582. Proc. Mach. Learn. Res. 125, PMLR, 2020.

[35] Wei, C.-Y. and Luo, H.: Non-stationary reinforcement learning without prior knowledge: An optimal black-box approach. In *Proceedings of the 34th conference on learning theory (Boulder, CO, 2021)*, pp. 34300–4354. Proc. Mach. Learn. Res. 134, PMLR, 2021.

Received September 19, 2023.

Jacob Carruth
Department of Mathematics, Princeton University
Fine Hall, Washington Road, Princeton, NJ 08544, USA;
jcarruth@math.princeton.edu

Maximilian F. Eggl
Institute of Experimental Epileptology and Cognition, University of Bonn
Venusberg-Campus 1, 53127 Bonn, Germany;
meggl@uni-bonn.de

Charles Fefferman
Department of Mathematics, Princeton University
Fine Hall, Washington Road, Princeton, NJ 08544, USA;
cf@math.princeton.edu

Clarence W. Rowley
Department of Mechanical and Aerospace Engineering, Princeton University
Engineering Quadrangle, Princeton, NJ 08544, USA;
cwrowley@princeton.edu