Rev. Mat. Iberoam. (Online first)
DOI 10.4171/RMI/1510

Published by EMS Press

Optimal agnostic control of unknown linear dynamics in a bounded parameter range

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

Abstract. Here and in a follow-on 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 is bounded, i.e., that $|a| \le a_{\mathrm{MAX}}$, and we study two variants of the control problem. In the first variant, Bayesian control, we are given a prior probability distribution for a and we seek a strategy that minimizes the expected value of a given cost function. Assuming that we can solve a certain PDE (the Hamilton–Jacobi–Bellman equation), we produce optimal strategies for Bayesian control. In the second variant, agnostic control, we assume nothing about a and we seek a strategy that minimizes a quantity called the regret. We produce a prior probability distribution $d\,\mathrm{Prior}(a)$ supported on a finite subset of $[-a_{\mathrm{MAX}}, a_{\mathrm{MAX}}]$ so that the agnostic control problem reduces to the Bayesian control problem for the prior $d\,\mathrm{Prior}(a)$.

Contents

1. Introduction

Here and in [7, 8, 17], we explore a new flavor of adaptive control theory, which we call “agnostic control”. Our introduction borrows heavily from that of the follow-on paper [7].

Mathematics Subject Classification 2020: 49N30.
Keywords: optimal control, adaptive control, 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 follow-on 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.) Here, we suppose that the unknown a is confined to a known interval $[-a_{\mathrm{MAX}}, a_{\mathrm{MAX}}]$ and we do not assume that we are given a Bayesian prior probability distribution for it. In [7], we extend our results to deal with the case in which a may be any real number.

Modulo an arbitrarily small increase in regret, we reduce the problem, here and in [7], 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, the Bellman equation. We prove here 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,[1] which seem to behave as expected.)

Let us now explain the above in more detail.

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; these random variables depend on our strategy σ, and we often write $q^\sigma(t)$ and $u^\sigma(t)$ to make that dependence explicit.

[1] 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).

Over a time horizon $T > 0$, we incur a COST, given[2] by

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

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, but we do not know the parameter a.

We would like to keep our cost as low as possible. We write $\text{ECOST}(\sigma, a)$ to denote the expected value of the COST (1.2) for the given a in (1.1).

In this paper, we study two variants of the above control problem, which we call *Bayesian control* and *agnostic control*.

For Bayesian control, we are given a prior probability distribution $d\text{Prior}(a)$ for the unknown a in (1.1). We assume that $d\text{Prior}$ is supported in an interval $[-a_{\text{MAX}}, a_{\text{MAX}}]$. Our task is to pick the strategy σ to minimize

$$\text{ECOST}(\sigma, d\text{Prior}) = \int_{-a_{\text{MAX}}}^{a_{\text{MAX}}} \text{ECOST}(\sigma, a)\, d\text{Prior}(a). \tag{1.3}$$

For agnostic control, we are given that a belongs to a known interval $[-a_{\text{MAX}}, a_{\text{MAX}}]$, but we are not given a prior probability distribution $d\text{Prior}(a)$, so we cannot define an expected cost by (1.3). Instead, our goal will be to minimize *worst-case regret*, defined by comparing the performance of our strategy with that of an opponent who knows the value of a and plays optimally. Let $\sigma_{\text{opt}}(a)$ be the optimal strategy for known a. Thus $\text{ECOST}(\sigma, a)$ is minimized over all σ by taking $\sigma = \sigma_{\text{opt}}(a)$.[3] We will introduce several variants of the notion of regret.

To a given strategy σ, we associate the following functions on $[-a_{\text{MAX}}, a_{\text{MAX}}]$:

- *Additive regret*, defined as

$$\text{AReg}(\sigma, a) = \text{ECOST}(\sigma, a) - \text{ECOST}(\sigma_{\text{opt}}(a), a) \geq 0.$$

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

$$\text{MReg}(\sigma, a) = \frac{\text{ECOST}(\sigma, a)}{\text{ECOST}(\sigma_{\text{opt}}(a), a)} \geq 1.$$

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

$$\text{HReg}_\gamma(\sigma, a) = \frac{\text{ECOST}(\sigma, a)}{\text{ECOST}(\sigma_{\text{opt}}(a), a) + \gamma}.$$

See [7] for a discussion of the regimes in which these three notions provide useful information.

Writing $\text{REGRET}(\sigma, a)$ to denote any one of the above three functions on the interval $[-a_{\text{MAX}}, +a_{\text{MAX}}]$, we define the *worst-case regret*:

$$\text{REGRET}^*(\sigma) = \sup\{\text{REGRET}(\sigma, a) : a \in [-a_{\text{MAX}}, a_{\text{MAX}}]\}. \tag{1.4}$$

We seek a strategy σ having the least possible worst-case regret.

[2] By rescaling, we can consider seemingly different cost functions of the form $\int_0^T (q^2 + \lambda u^2)$ for $\lambda > 0$.
[3] See standard textbooks (e.g., [3]) for the computation of $\sigma_{\text{opt}}(a)$ and its expected cost.

Thus, we have posed two problems: for Bayesian control, find the strategy that minimizes expected cost; for agnostic control, find a strategy that minimizes worst-case regret.

To prepare to present our results, we next discuss a relevant PDE, the *Bellman equation*.

We will see that the problem of Bayesian control is intimately connected to the following PDE for an unknown function $S(q, t, \zeta_1, \zeta_2)$ of four variables:

$$\begin{aligned} 0 = \partial_t S + (\bar{a}(\zeta_1, \zeta_2) q + u_{\mathrm{opt}})\, \partial_q S + \bar{a}(\zeta_1, \zeta_2) q^2\, \partial_{\zeta_1} S + q^2\, \partial_{\zeta_2} S + \frac{1}{2}\, \partial_q^2 S \\ + q\, \partial^2_{q\zeta_1} S + \frac{1}{2} q^2\, \partial^2_{\zeta_1} S + (q^2 + u^2_{\mathrm{opt}}), \end{aligned} \tag{1.5}$$

where

$$u_{\mathrm{opt}} = -\frac{1}{2}\, \partial_q S, \tag{1.6}$$

with terminal condition

$$S|_{t=T} = 0. \tag{1.7}$$

Here, $\bar{a}(\zeta_1, \zeta_2)$ is a known, smooth function of two variables.

We have succeeded in finding numerical solutions of (1.5)–(1.7), but we lack rigorous proofs of existence and smoothness of solutions. Accordingly, we impose a *PDE assumption* to the effect that (1.5)–(1.7) admit a solution S satisfying plausible estimates (see Section 4.3). Our numerics suggest that the PDE assumption is correct. *Our results below are conditional on the PDE assumption.*

We are ready to state our main results. We begin with Bayesian control. For a function $\bar{a}(\zeta_1, \zeta_2)$ given in terms of $d\mathrm{Prior}$ by an elementary formula, we define a function $u_{\mathrm{opt}}(q, t, \zeta_1, \zeta_2)$ as in (1.5)–(1.7), and then specify a strategy $\sigma = \sigma_{\mathrm{Bayes}}(d\mathrm{Prior})$ by setting

$$u^\sigma(t) = u_{\mathrm{opt}}(q^\sigma(t), t, \zeta_1(t), \zeta_2(t)), \quad \text{with} \tag{1.8}$$

$$\zeta_1(t) = \int_0^t q^\sigma(s)\,[dq^\sigma(s) - u^\sigma(s)\, ds] \quad \text{and} \quad \zeta_2(t) = \int_0^t (q^\sigma(s))^2\, ds. \tag{1.9}$$

Note that $\zeta_1(t)$ and $\zeta_2(t)$ are determined by past history up through time t, hence so is $u^\sigma(t)$ in (1.8). As explained in [7], heuristic reasoning suggests that $\sigma_{\mathrm{Bayes}}(d\mathrm{Prior})$ is the optimal strategy for Bayesian control with prior belief $d\mathrm{Prior}$. Our rigorous result confirms this intuition. Recall that q_0 is our starting position.

Theorem 1.1. *Fix a probability distribution* $d\mathrm{Prior}$ *on* $[-a_{\mathrm{MAX}}, a_{\mathrm{MAX}}]$*, and let* S*,* u_{opt} *and* $\sigma = \sigma_{\mathrm{Bayes}}(d\mathrm{Prior})$ *be as above. Then the following hold.*

(A) $\mathrm{E{\small COST}}(\sigma, d\mathrm{Prior}) = S(q_0, 0, 0, 0)$.

(B) *Let* σ' *be any other strategy. Then*

$$\mathrm{E{\small COST}}(\sigma', d\mathrm{Prior}) \geq \mathrm{E{\small COST}}(\sigma, d\mathrm{Prior}).$$

For a class of "tame strategies" σ', we can sharpen (B) above to a quantitative result. A *tame strategy* σ' satisfies the estimate

$$|u^{\sigma'}(t)| \leq \hat{C}\,[|q^{\sigma'}(t)| + 1] \quad (\text{all } t \in [0, T])$$

with probability 1, for a constant $\hat{C}$, called a *tame constant for* σ'. The quantitative version of (B) is as follows.

Theorem 1.2 (Quantitative uniqueness). *Let* $d\text{Prior}$ *and* $\sigma = \sigma_{\text{Bayes}}(d\text{Prior})$ *be as in Theorem 1.1. Given* $\varepsilon > 0$ *and given a constant* $\hat{C}$, *there exists* $\delta > 0$ *for which the following holds.*

Let σ' *be a tame strategy with tame constant* $\hat{C}$.

If $\text{ECOST}(\sigma', d\text{Prior}) \le \text{ECOST}(\sigma, d\text{Prior}) + \delta$, *then the expected value of*

$$\int_0^T \{|q^\sigma(t) - q^{\sigma'}(t)|^2 + |u^\sigma(t) - u^{\sigma'}(t)|^2\}\, dt$$

is less than ε.

Quantitative uniqueness will play a crucial rôle in our analysis of agnostic control.

Our main result for agnostic control is as follows.

Theorem 1.3. *Fix* $[-a_{\text{MAX}}, a_{\text{MAX}}]$, q_0, T *(and* γ, *if we use hybrid regret). Then there exist a probability measure* $d\text{Prior}$, *a finite subset* $E \subset [-a_{\text{MAX}}, a_{\text{MAX}}]$, *and a strategy* σ, *for which the following hold.*

(I) σ *is the optimal Bayesian strategy for the prior probability distribution* $d\text{Prior}$.

(II) $d\text{Prior}$ *is supported in the finite set* E.

(III) E *is precisely the set of points* $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$ *at which the function*

$$[-a_{\text{MAX}}, a_{\text{MAX}}] \ni a \mapsto \text{REGRET}(\sigma, a)$$

achieves its maximum.

(IV) $\text{REGRET}^*(\sigma) \le \text{REGRET}^*(\sigma')$ *for any other strategy* σ'.

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 $d\text{Prior}$, even though in fact we know nothing about a except that it lies in $[-a_{\text{MAX}}, a_{\text{MAX}}]$.

Let $d\text{Prior}$, σ and E be as in (I), (II) and (III) of Theorem 1.3. Since σ is the optimal Bayesian strategy for $d\text{Prior}$ (by (I)), and since $d\text{Prior}$ is supported on the finite set E (by (II)), we have for any other strategy σ' that

$$\text{ECOST}(\sigma, a_0) \le \text{ECOST}(\sigma', a_0) \quad \text{for some } a_0 \in E.$$

In particular, we have

$$\text{REGRET}(\sigma, a_0) \le \text{REGRET}(\sigma', a_0) \quad \text{for some } a_0 \in E.$$

Combining this with (III), we see that for any $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$, we have

$$\text{REGRET}(\sigma, a) \le \text{REGRET}(\sigma', a_0).$$

Therefore (I), (II) and (III) of Theorem 1.3 easily imply (IV). The hard part of Theorem 1.3 is the assertion that there exist $d\text{Prior}$, E and σ satisfying (I), (II) and (III).

Theorem 1.3 lets us search for optimal agnostic strategies: We first guess a finite set E and a probability measure $d\text{Prior}$ concentrated on E. By solving the Bellman equation, we produce the optimal Bayesian strategy $\sigma = \sigma_{\text{Bayes}}(d\text{Prior})$, which allows us to compute the function $[-a_{\text{MAX}}, a_{\text{MAX}}] \ni a \mapsto \text{REGRET}(\sigma, a)$. If the maximum of that function occurs precisely at the points of E, then σ is the desired optimal agnostic strategy. Otherwise, we modify our guess $(E, d\text{Prior})$. We have carried this out numerically for several $[-a_{\text{MAX}}, a_{\text{MAX}}]$, q_0 and T.

This concludes our introductory discussion of agnostic control for bounded a (i.e., for $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$).

We briefly touch on another variant of the control problem (1.1): *agnostic control for unbounded* a.

Suppose we assume absolutely nothing about our unknown a; it might be any real number. For any strategy σ, we define $\text{REGRET}^*(\sigma)$ as in (1.4), except that now the sup is taken over all $a \in \mathbb{R}$. Our task is to pick σ to minimize $\text{REGRET}^*(\sigma)$.

Our companion paper [7] analyzes this problem by comparing optimal agnostic control for arbitrary a with the case in which a is confined to a large interval $[-a_{\text{MAX}}(\varepsilon), a_{\text{MAX}}(\varepsilon)]$, depending on a small parameter $\varepsilon > 0$. Roughly speaking, [7] shows that any strategy for a confined to $[-a_{\text{MAX}}(\varepsilon), +a_{\text{MAX}}(\varepsilon)]$ may be modified to produce a strategy for arbitrary $a \in \mathbb{R}$, with an increase in worst-case hybrid regret of at most ε. (See [7] for precise statements.)

Recap. Let us summarize what we have achieved. Suppose our goal is to minimize worst-case hybrid regret in the setting in which 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 $[-a_{\text{MAX}}, a_{\text{MAX}}]$. We then look for a probability measure $d\text{Prior}$ living on a finite set $E \subset [-a_{\text{MAX}}, a_{\text{MAX}}]$, such that the regret of the optimal Bayesian strategy for $d\text{Prior}$ 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, smoothness, and size of solutions to the Bellman equation. In numerical simulations, we have produced evidence for our PDE assumptions, and we have produced optimal agnostic strategies for cases in which the unknown a is confined to an interval.

Ideas from the proofs. We mention one significant technical point regarding the proofs of Theorems 1.1, 1.2 and 1.3: we need a rigorous definition of a strategy. Certainly the phrase "a rule for determining $u(t)$ from past history" is not precise.

We want to allow $u(t)$ to depend discontinuously on past history $(q(s))_{s\in[0,t]}$. For instance, we should be allowed to set

$$u(t) = \begin{cases} -q(t) & \text{if } |q(t)| > 1, \\ 0 & \text{otherwise.} \end{cases}$$

On the other hand, we had better make sure that we can produce solutions of our stochastic ODE

$$dq = (aq + u)dt + dW.$$

We proceed as follows.

At first we fix a partition

$$0 = t_0 < t_1 < \cdots < t_N = T \tag{1.10}$$

of the time interval $[0, T]$. We restrict ourselves to strategies σ in which the control $u(t)$ is constant in each interval $[t_\nu, t_{\nu+1})$, and in which, for each ν, $u(t_\nu)$ is determined by $q(t_1), \ldots, q(t_\nu)$, together with "coin flips" $\vec{\xi} = (\xi_1, \xi_2, \cdots) \in \{0, 1\}^{\mathbb{N}}$. For all ν, we assume that $u(t_\nu)$ is a Borel measurable function of $(q(t_1), \ldots, q(t_\nu), \vec{\xi})$, and that

$$|u(t_\nu)| \leq C_{\text{TAME}}\,[|q(t_\nu)| + 1].$$

A strategy as above is called a *tame strategy associated to the partition* (1.10), with a *tame constant* C_{TAME}. For such strategies, it is easy to define the solutions $q^\sigma(t)$, $u^\sigma(t)$ of our stochastic ODE (1.1).

Most of our work lies in controlling and optimizing tame strategies associated to a sufficiently fine partition. In particular, we will prove approximate versions of Theorems 1.1 and 1.2 in the setting of such strategies (see Lemmas 4.6 and 4.9, respectively).

We will then define a tame strategy (not associated to any partition) by considering a sequence $\pi_1, \pi_2, \ldots$ of ever-finer partitions of $[0, T]$. To each partition π_n we associate a tame strategy σ_n with a tame constant C_{TAME} independent of n. If the resulting $q^{\sigma_n}(t)$ and $u^{\sigma_n}(t)$ tend to limits, in an appropriate sense, as $n \to \infty$, then we declare these limits $q(t)$ and $u(t)$ to arise from a *tame strategy* σ with a *tame constant* C_{TAME}.

Finally, we drop the restriction to tame strategies and consider general strategies. To do so, we consider a sequence $(\sigma_n)_{n=0,1,2,\ldots}$ of tame strategies, *not* assumed to have a tame constant independent of n. If the relevant $q^{\sigma_n}(t)$ and $u^{\sigma_n}(t)$ converge, in a suitable sense, as $n \to \infty$, then we say that the limits $q(t)$ and $u(t)$ arise from a strategy σ.

It is not hard to pass from tame strategies associated to partitions of $[0, T]$ to general tame strategies, and then to pass from such tame strategies to general strategies. The work in proving Theorems 1.1 and 1.2 lies in our close study of tame strategies associated to fine partitions.

We provide only a few comments on the proof of Theorem 1.3. The main work lies in proving an analogue of Theorem 1.3 in which the unknown a is confined to a finite set $A \subset [-a_{\text{MAX}}, a_{\text{MAX}}]$, rather than to the whole of $[-a_{\text{MAX}}, a_{\text{MAX}}]$. We apply that analogue to a sequence $A_1, A_2, \ldots$ of fine nets in $[-a_{\text{MAX}}, a_{\text{MAX}}]$, e.g., $A_n = [-a_{\text{MAX}}, a_{\text{MAX}}] \cap 2^{-n}\mathbb{Z}$, and pass to the limit as $n \to \infty$ using a weak compactness argument. To establish the result for finite A, we proceed by induction on the number of elements of A. Details may be found in Section 6.

Future directions. Our work suggests several unsolved problems, among which we mention:

- Prove (or disprove) the PDE assumption.
- Consider problems in which the particle lives in $\mathbb{R}^N$, not just in $\mathbb{R}^1$; and in which the dynamics of the particle depend on more than one unknown parameter. Can that be done without rendering the relevant numerics hopelessly impractical?
- Even for the model problem considered in this paper, improve the numerics to let us produce optimal agnostic strategies for a larger range of a_{MAX} and T than we can deal with today.

We speculate briefly on a particular model problem in which we do not know a priori what our control does.

Consider a particle governed by the stochastic ODE

$$dq(t) = au(t)\,dt + dW(t), \quad q(0) = 0. \tag{1.11}$$

As usual, $q(t)$ denotes position, $u(t)$ is our control, $W(t)$ is Brownian motion, and we incur a cost

$$\int_0^T \{(q(t))^2 + (u(t))^2\}\,dt.$$

In the simplest case, suppose we know a priori that $a = 1$ or $a = -1$, each with probability $1/2$. We write $\text{E}\textsc{cost}(\sigma)$ to denote the expected cost incurred by executing a strategy σ, and we set

$$\text{E}\textsc{cost}^* = \inf\{\text{E}\textsc{cost}(\sigma) : \text{all strategies } \sigma\}. \tag{1.12}$$

For this simple model problem, we conjecture that the inf in (1.12) is not achieved by any strategy σ, because heuristic reasoning suggests that there is a regime in which we would like to set $u = \pm\infty$ to gain instant information about a.

Clearly, there is much to be done before we can claim to understand agnostic control theory.

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 obtained 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 [7, 8, 17], and in 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 in [7], 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. The game

We will deal with random variables

(2.1) $a_{\text{TRUE}} \in [-a_{\text{MAX}}, +a_{\text{MAX}}]$,

(2.2) "Coin flips" $\xi_1, \xi_2, \ldots \in \{0, 1\}$ (we write $\vec{\xi}$ for $(\xi_1, \xi_2, \ldots)$), and

(2.3) Brownian motion $W(t)$, starting at $W(0) = 0$.

The random variable a_{TRUE} is deterministic and known in Section 3, unknown but subject to a known prior in Sections 4 and 5, and unknown without a known prior in Section 6.

The ξ_ν, the real number a_{TRUE}, and the Brownian motion are mutually independent. The variable a_{TRUE} has a prior probability distribution given by the measure $d\text{Prior}(a)$ in Sections 4 and 5; and each ξ_ν is equal to 0 with probability $1/2$, and to 1 with probability $1/2$.

When a_{TRUE} has a prior probability distribution, we write $\text{Prob}[\mathcal{E}]$ to denote the probability of an event $\mathcal{E}$ with respect to the above probability space, and we write $\text{E}[X]$ for the expected value of X with respect to that probability space.

Given $a \in [-a_{\mathrm{MAX}}, +a_{\mathrm{MAX}}]$, we write $\mathrm{Prob}_a[\mathcal{E}]$ for the probability of event $\mathcal{E}$ conditioned on the event $a_{\mathrm{TRUE}} = a$, and we write $\mathrm{E}_a[X]$ for the expectation of X conditioned on $a_{\mathrm{TRUE}} = a$; $\mathrm{Prob}_a[\cdots]$ and $\mathrm{E}_a[\cdots]$ make sense without an assumed prior for a.

Similarly, given $a \in [-a_{\mathrm{MAX}}, +a_{\mathrm{MAX}}]$ and $\vec{\eta} = (\eta_1, \eta_2, \dots) \in \{0,1\}^{\mathbb{N}}$, we shall write $\mathrm{Prob}_{a,\vec{\eta}}[\mathcal{E}]$ and $\mathrm{E}_{a,\vec{\eta}}[X]$ to denote the probability and expectation, respectively, conditioned on the event $a_{\mathrm{TRUE}} = a$ and $\xi_\nu = \eta_\nu$ for all ν.

Also, we write $\mathrm{E}_{\vec{\eta}}[X]$ and $\mathrm{Prob}_{\vec{\eta}}[\mathcal{E}]$ to denote expectation and probability, respectively, conditioned on $\vec{\xi} = \vec{\eta}$.

The above conditional expectations make sense even if, for instance, a is not in the support of $d\mathrm{Prior}$.

Fix a *terminal time* $T > 0$ and a partition

$$0 = t_0 < t_1 < \cdots < t_N = T \quad \text{of } [0, T].$$

Fix a *starting position* $q_0 \in \mathbb{R}$. A *tame rule* at time t_ν is a Borel measurable function $\sigma_{t_\nu} : \mathbb{R}^\nu \times \{0,1\}^{\mathbb{N}} \to \mathbb{R}$, satisfying the estimate

$$|\sigma_{t_\nu}(q_1, \dots, q_\nu, \vec{\xi})| \le C_{\mathrm{TAME}}[|q_\nu| + 1] \tag{2.4}$$

for all $(q_1, \dots, q_\nu, \vec{\xi}) \in \mathbb{R}^\nu \times \{0,1\}^{\mathbb{N}}$. If $\nu = 0$, then σ_{t_ν} is simply a function on $\{0,1\}^{\mathbb{N}}$. (We use the product topology on $\mathbb{R}^\nu \times \{0,1\}^{\mathbb{N}}$ to define Borel measurability. We require Borel measurability to avoid technicalities. In particular, the composition of Borel measurable functions is Borel measurable, whereas the composition of Lebesgue measurable functions need not be Lebesgue measurable.)

A *tame strategy* is an array $\sigma = (\sigma_{t_\nu})_{\nu=0,1,\dots,N-1}$, where, for each ν, σ_{t_ν} is a tame rule at time t_ν with the same C_{TAME} serving in (2.4) for all the t_ν. We call C_{TAME} a *tame constant* for the strategy σ. Until further notice, we say simply "strategy" in place of "tame strategy". If the σ_{t_ν} do not depend on the coin flips $\vec{\xi}$, we call σ a *deterministic strategy*. We will often write σ_ν in place of σ_{t_ν}.

Given a strategy $\sigma = (\sigma_{t_\nu})_{\nu=0,1,\dots,N-1}$, we define random variables $q^\sigma(t)$ for $t \in [0, T]$ and $u^\sigma(t)$ for $t \in [0, T)$, as follows.

By induction on ν, we define $q^\sigma(t)$ for $t \in [0, t_\nu]$ and $u^\sigma(t)$ for $t \in [0, t_\nu)$.

In the base case $\nu = 0$, we set $q^\sigma(t) = q_0$ for $t \in [0, t_\nu] = \{0\}$. Since $[0, t_\nu) = [0, 0)$ is empty, there is no need to define u^σ in the base case.

For the induction step, we fix $\nu \ge 0$, and assume that we have defined $q^\sigma(t)$ for $t \in [0, t_\nu]$ and $u^\sigma(t)$ for $t \in [0, t_\nu)$. We extend the definition of $q^\sigma(t)$ to $t \in [0, t_{\nu+1}]$, and that of $u^\sigma(t)$ to $t \in [0, t_{\nu+1})$, as follows:

- for $t \in [t_\nu, t_{\nu+1})$, we set

$$u^\sigma(t) = \sigma_{t_\nu}(q^\sigma(t_1), \dots, q^\sigma(t_\nu), \vec{\xi});$$

- for $t \in [t_\nu, t_{\nu+1}]$, we define $q^\sigma(t)$ as the solution of the stochastic ODE

$$dq^\sigma(t) = (a_{\mathrm{TRUE}}\, q^\sigma(t) + u^\sigma(t))\, dt + dW(t),$$

  with the initial value $q^\sigma(t_\nu)$ already given by our induction hypothesis.

This completes our induction on ν, so we have defined the random variables $q^\sigma(t), u^\sigma(t)$.

In addition to $q^\sigma(t)$ and $u^\sigma(t)$, we define random variables $\zeta_1^\sigma(t_\nu)$ and $\zeta_2^\sigma(t_\nu)$ $(0 \leq \nu < N)$ by the following induction.

$$\begin{aligned}
&\zeta_1^\sigma(t_0) = \zeta_2^\sigma(t_0) = 0 \quad \text{(recall, } t_0 = 0\text{);}\\
&\zeta_1^\sigma(t_{\nu+1}) = \zeta_1^\sigma(t_\nu) + q^\sigma(t_\nu) \cdot [\Delta q_\nu^\sigma - u^\sigma(t_\nu)\,\Delta t_\nu],\\
&\qquad \text{where } \Delta q_\nu^\sigma = q^\sigma(t_{\nu+1}) - q^\sigma(t_\nu) \text{ and } \Delta t_\nu = t_{\nu+1} - t_\nu;\\
&\zeta_2^\sigma(t_{\nu+1}) = \zeta_2^\sigma(t_\nu) + (q^\sigma(t_\nu))^2 \Delta t_\nu.
\end{aligned}$$

Thus,

$$\zeta_1^\sigma(t_\nu) = \sum_{0 \leq \mu < \nu} q^\sigma(t_\mu)\,(\Delta q_\mu^\sigma - u^\sigma(t_\mu)\,\Delta t_\mu) \quad \text{and} \quad \zeta_2^\sigma(t_\nu) = \sum_{0 \leq \mu < \nu} (q^\sigma(t_\mu))^2\,\Delta t_\mu.$$

We will try to pick our strategy σ to make the expected value of

$$\int_0^T [(q^\sigma(t))^2 + (u^\sigma(t))^2]\,dt$$

as small as possible.

3. Tame strategies associated to partitions

3.1. Setup

In this section, we take a_{TRUE} to be fixed, $a_{\text{TRUE}} = a$, and we suppose that our strategy makes no use of coin flips.

We fix a partition

$$0 = t_0 < t_1 < \cdots < t_N = T \tag{3.1}$$

of a time interval $[0, T]$.

We fix a (deterministic) strategy σ for the game with starting position q_0. We assume that our strategy is tame, i.e.,

$$|u^\sigma(t_\nu)| \leq C_{\text{TAME}}\,[|q^\sigma(t_\nu)| + 1] \tag{3.2}$$

for a constant C_{TAME}.

We write c, C, C', etc., to denote constants determined by

- C_{TAME} in (3.2),
- an upper bound for the time horizon T,
- an upper bound for a_{TRUE},
- an upper bound for $|q_0|$.

These symbols may denote different constants in different occurrences.

We define

$$\Delta t_\nu := t_{\nu+1} - t_\nu \quad \text{for all } \nu \ (0 \leq \nu < N),$$

and we assume that

$$(\Delta t_{\text{MAX}}) := \max_\nu \Delta t_\nu \text{ is less than a small enough constant } c. \tag{3.3}$$

We write $X = O(Y)$ to denote the estimate $|X| \leq CY$.

We write q_ν to denote $q^\sigma(t_\nu)$, u_ν to denote $u^\sigma(t_\nu)$, and Δq_ν to denote $q_{\nu+1} - q_\nu$.

Note that

$$u_\nu = \sigma_\nu(q_1, \dots, q_\nu), \tag{3.4}$$

where σ_ν is given by the strategy σ for decisions at time t_ν.

Thanks to (3.2), we have

$$|u_\nu| \leq C[|q_\nu| + 1]. \tag{3.5}$$

Recall that the q_ν evolve as follows.

The random variable q_0 is given, and u_0 is specified by the strategy σ. The random variables $q_{\nu+1}$ and $u_{\nu+1}$ are then determined from q_ν and u_ν as follows.

- We solve the stochastic ODE

$$dq(t) = (aq(t) + u_\nu)\,dt + dW(t) \quad \text{for } t \in [t_\nu, t_{\nu+1}], \tag{3.6}$$

  with initial condition $q(t_\nu) = q_\nu$. Here, a is the (given) value of a_{TRUE}, and $W(t)$ denotes Brownian motion at time t.
- We set $q_{\nu+1} = q(t_{\nu+1})$.
- We set $u_{\nu+1} = \sigma_{\nu+1}(q_1, \dots, q_{\nu+1})$ (compare with (3.4)).

Thus, the q_ν, u_ν are random variables defined by induction on ν.

Solving the ODE (3.6) using an integrating factor, we find that

$$q(t) - q(t_\nu) = (aq_\nu + u_\nu)\left[\frac{e^{a(t-t_\nu)} - 1}{a}\right] + \int_{t_\nu}^{t} e^{a(t-s)}\,dW(s) \tag{3.7}$$

for $t \in [t_\nu, t_{\nu+1}]$. In particular,

$$\Delta q_\nu = q_{\nu+1} - q_\nu = (aq_\nu + u_\nu)\,\Delta t_\nu^* + \Delta W_\nu, \tag{3.8}$$

where

$$\Delta t_\nu^* = \left[\frac{e^{a\Delta t_\nu} - 1}{a}\right] \tag{3.9}$$

and

$$\Delta W_\nu = \int_{t_\nu}^{t_{\nu+1}} e^{a(t_{\nu+1}-s)}\,dW(s). \tag{3.10}$$

(If $a = 0$, we interpret the above fractions in square brackets as $(t - t_\nu)$ in (3.7), and Δt_ν in (3.9).) We warn the reader that $\Delta W_\nu \neq W(t_{\nu+1}) - W(t_\nu)$.

Note that ΔW_ν is a normal random variable with mean 0 and variance

$$\Delta\tilde{t}_\nu = \left[\frac{e^{2a\Delta t_\nu} - 1}{2a}\right] \tag{3.11}$$

(again, equal to Δt_ν if $a = 0$).

Note that

$$\Delta\tilde{t}_\nu, \Delta t_\nu^* = \Delta t_\nu + O((\Delta t_\nu)^2). \tag{3.12}$$

Note also that $\Delta\tilde{t}_\nu$, Δt_ν^* and ΔW_ν depend on a.

We introduce the sigma algebras $\mathcal{F}_\nu$, defined as the algebra of events determined by the ΔW_μ for $0 \le \mu < \nu$. Note that q_ν, u_ν and ΔW_μ ($\mu < \nu$) are $\mathcal{F}_\nu$-measurable (i.e., they are deterministic once we condition on $\mathcal{F}_\nu$), while the ΔW_μ for $\mu \ge \nu$ are independent of $\mathcal{F}_\nu$.

Remark 3.1. Thanks to equation (3.8), the sigma algebra $\mathcal{F}_\nu$ may be equivalently defined to consist of all events determined by $q_1, \dots, q_\nu$. This equivalence holds because a_{TRUE} has been fixed ($a_{\text{TRUE}} = a$).

3.2. Estimates for probabilities of outliers

We suppose

$$Q > C \quad \text{for a large enough } C, \tag{3.13}$$

and we estimate the probability that $\max_\nu |q_\nu| > Q$. To do so, we set

$$u_\nu^1 = u_\nu / q_\nu \text{ and } u_\nu^0 = 0 \text{ if } |q_\nu| > 1;$$
$$u_\nu^1 = 0 \text{ and } u_\nu^0 = u_\nu \text{ otherwise.}$$

Thus,

$$u_\nu = u_\nu^1 q_\nu + u_\nu^0 \tag{3.14}$$

and

$$|u_\nu^1|, |u_\nu^0| \le C. \tag{3.15}$$

Also, u_ν^1 and u_ν^0 are $\mathcal{F}_\nu$-measurable.

Thanks to (3.14), we can rewrite (3.8) in the form

$$\Delta q_\nu = (a + u_\nu^1)\, q_\nu (\Delta t_\nu^*) + u_\nu^0 (\Delta t_\nu^*) + \Delta W_\nu,$$

or equivalently,

$$[e^{-a t_{\nu+1}} q_{\nu+1}] = (1 + e^{-a \Delta t_\nu}\, \Delta t_\nu^* u_\nu^1)\, [e^{-a t_\nu} q_\nu] + e^{-a t_{\nu+1}}\, \Delta t_\nu^* u_\nu^0 + e^{-a t_{\nu+1}} \Delta W_\nu.$$

(See (3.9).)

Setting

$$M_\nu = \prod_{0 \le \mu < \nu} (1 + e^{-a \Delta t_\mu}\, \Delta t_\mu^* u_\mu^1)^{-1}, \tag{3.16}$$

$$m_\nu = e^{-a t_{\nu+1}}\, \Delta t_\nu^* u_\nu^0, \quad \text{and} \tag{3.17}$$

$$q_\nu^* = M_\nu\, e^{-a t_\nu}\, q_\nu, \tag{3.18}$$

we see that

$$q_{\nu+1}^* = q_\nu^* + M_{\nu+1}\, m_\nu + M_{\nu+1}\, e^{-a t_{\nu+1}}\, \Delta W_\nu, \tag{3.19}$$

and that

$$q_0^* = q_0. \tag{3.20}$$

Since

$$|e^{-a\Delta t_\mu} u^i_\mu| \le C \ (i = 0, 1) \quad \text{and} \quad \sum_\mu (\Delta t^*_\mu) \le C,$$

we have

$$c < M_\nu < C \quad \text{and} \quad |m_\nu| < C \Delta t_\nu. \tag{3.21}$$

Moreover, $M_{\nu+1}$ is $\mathcal{F}_\nu$-measurable.

Let $\lambda \in \mathbb{R}$, to be fixed below. From (3.19) and (3.21), we have

$$\exp(\lambda q^*_{\nu+1}) \le \{\exp(\lambda q^*_\nu) \cdot \exp(C|\lambda|\Delta t_\nu)\} \exp(\{M_{\nu+1} e^{-at_{\nu+1}}\}\lambda \Delta W_\nu). \tag{3.22}$$

Here, the quantities in curly brackets are $\mathcal{F}_\nu$-measurable, while ΔW_ν is independent of $\mathcal{F}_\nu$.

Recalling that ΔW_ν is a normal random variable, with mean 0 and variance $O(\Delta t_\nu)$, we deduce from (3.21) and (3.22) that

$$\mathrm{E}[\exp(\lambda q^*_{\nu+1}) \,|\, \mathcal{F}_\nu] \le \exp(\lambda q^*_\nu) \, \exp(C|\lambda|\Delta t_\nu) \, \exp(C\lambda^2 \Delta t_\nu).$$

Thus, the random variables

$$Z_\nu = \exp(-C[|\lambda| + \lambda^2] t_\nu) \, \exp(\lambda q^*_\nu) \quad (0 \le \nu \le N) \tag{3.23}$$

form a supermartingale, with

$$Z_0 = \exp(\lambda q_0) \le \exp(C|\lambda|).$$

Consequently, for any $Q > 0$ we have

$$\mathrm{Prob}\big[\max_\nu Z_\nu > \exp(|\lambda| Q)\big] \le \exp(|\lambda|(C - Q)).$$

By definition (3.23), this means that

$$\mathrm{Prob}[\lambda q^*_\nu - C[|\lambda| + \lambda^2] t_\nu > |\lambda| Q \text{ for some } \nu] \le \exp(|\lambda|(C - Q)). \tag{3.24}$$

Taking Q greater that $2C$ in (3.24) (see (3.13)), and picking $\lambda = \pm Q$, we learn from (3.24) that

$$\mathrm{Prob}[|q^*_\nu| > CQ \text{ for some } \nu] \le C \exp(-cQ^2).$$

Recalling (3.18) and (3.21), we conclude that

$$\mathrm{Prob}\big[\max_\nu |q_\nu| > Q\big] \le C \exp(-cQ^2) \tag{3.25}$$

if Q satisfies (3.13).

Thus, we have succeeded in estimating the probability that $\max_\nu |q_\nu|$ is large.

Immediately from (3.5) and (3.25), we have also

$$\mathrm{Prob}\big[\max_\nu |u_\nu| > Q\big] \le C \exp(-cQ^2) \tag{3.26}$$

if Q satisfies (3.13).

We now turn our attention to

$$\zeta_1(t_\nu) = \sum_{0 \le \mu < \nu} q_\mu \, [\Delta q_\mu - u_\mu \, \Delta t_\mu] \tag{3.27}$$

and

$$\zeta_2(t_\nu) = \sum_{0 \le \mu < \nu} q_\mu^2 \, \Delta t_\mu. \tag{3.28}$$

Note that $\zeta_1(t_\nu)$ can be rewritten as

$$\begin{aligned} \zeta_1(t_\nu) &= \sum_{0 \le \mu < \nu} \left\{ \frac{1}{2} \, (q_{\mu+1}^2 - q_\mu^2) - \frac{1}{2} \, (\Delta q_\mu)^2 \right\} - \sum_{0 \le \mu < \nu} u_\mu \, q_\mu \, \Delta t_\mu \\ &= \frac{1}{2} \, q_\nu^2 - \frac{1}{2} \, q_0^2 - \frac{1}{2} \sum_{\mu=0}^{\nu-1} (\Delta q_\mu)^2 - \sum_{0 \le \mu < \nu} u_\mu \, q_\mu \, \Delta t_\mu. \end{aligned} \tag{3.29}$$

Now suppose that

$$\max_\nu |q_\nu|, \max_\nu |u_\nu| \le CQ, \quad \text{with } Q \text{ as in (3.13).} \tag{3.30}$$

Then from (3.28) and (3.29), we have

$$|\zeta_2(t_\nu)| \le CQ^2 \quad (\text{all } \nu), \tag{3.31}$$

and

$$|\zeta_1(t_\nu)| \le CQ^2 + \left| \sum_{0 \le \mu < \nu} \{ (\Delta q_\mu)^2 - \Delta \tilde{t}_\mu \} \right| \quad (\text{all } \nu), \tag{3.32}$$

since also

$$\sum_{\mu=0}^{N} (\Delta \tilde{t}_\mu) \le C \sum_{\mu=0}^{N} \Delta t_\mu \le C'.$$

We will show that

$$\operatorname{Prob}\left[\max_\nu \left| \sum_{0 \le \mu < \nu} \{ (\Delta q_\mu)^2 - \Delta \tilde{t}_\mu \} \right| > CQ^2 (\Delta t_{\text{MAX}})^{1/2} \right] \le C \exp(-c\,Q^2). \tag{3.33}$$

In view of (3.25), (3.26), and (3.33), estimates (3.31) and (3.32) imply the inequalities

$$\operatorname{Prob}\left[\max_\nu |\zeta_1(t_\nu)| > CQ^2 \right] \le C \exp(-c\,Q^2), \tag{3.34}$$

$$\operatorname{Prob}\left[\max_\nu |\zeta_2(t_\nu)| > CQ^2 \right] \le C \exp(-c\,Q^2), \tag{3.35}$$

for Q as in (3.13). Thus, to prove (3.34) and (3.35), it remains only to prove (3.33). Estimate (3.33) will have further applications in a later section.

We now prove (3.33).

From (3.8) and (3.12), we have

$$
\begin{aligned}
\sum_{\mu<\nu}\{(\Delta q_\mu)^2 - \Delta\tilde{t}_\mu\} &= \sum_{\mu<\nu}(aq_\mu + u_\mu)^2\,(\Delta t^*_\mu)^2 \\
&\quad + 2\sum_{\mu<\nu}(aq_\mu + u_\mu)\,(\Delta t^*_\mu)\,\Delta W_\mu + \sum_{\mu<\nu}\{(\Delta W_\mu)^2 - \Delta\tilde{t}_\mu\} \\
&\equiv \text{TERM}\,1(\nu) + \text{TERM}\,2(\nu) + \text{TERM}\,3(\nu),
\end{aligned}
\tag{3.36}
$$

with

$$
0 \le \text{TERM}\,1(\nu) \le C\,\max_\mu\{|q_\mu| + |u_\mu|\}^2 \cdot (\Delta t_{\text{MAX}}), \quad \text{all } \nu. \tag{3.37}
$$

To estimate TERM $2(\nu)$, we fix $Q > C$ as in (3.13) and study the random variables

$$
\begin{aligned}
Y_\nu &= \exp(-\hat{C}\,\lambda^2\,Q^2\,(\Delta t_{\text{MAX}})^2\,t_\nu) \\
&\quad \cdot \exp\Big(\lambda\sum_{\mu<\nu}(aq_\mu + u_\mu)\cdot \mathbb{1}_{|aq_\mu+u_\mu|<CQ}\,(\Delta t^*_\mu)\,\Delta W_\mu\Big)
\end{aligned}
\tag{3.38}
$$

for a large enough constant $\hat{C}$, and for $\lambda \in \mathbb{R}$ to be picked below. Since $(aq_\mu + u_\mu)\cdot \mathbb{1}_{|aq_\mu+u_\mu|<CQ}$ is $\mathcal{F}_\nu$-measurable for $\mu \le \nu$, we have

$$
\begin{aligned}
\mathrm{E}[Y_{\nu+1}\,|\,\mathcal{F}_\nu] &= Y_\nu \cdot \exp(-\hat{C}\,\lambda^2\,Q^2\,(\Delta t_{\text{MAX}})^2\,\Delta t_\nu) \\
&\quad \cdot \mathrm{E}\big[\exp([\lambda\,(aq_\nu + u_\nu)\cdot \mathbb{1}_{|aq_\nu+u_\nu|<CQ}\cdot(\Delta t^*_\nu)]\,\Delta W_\nu \,|\, \mathcal{F}_\nu\big] \\
&\le Y_\nu \cdot \exp(-\hat{C}\lambda^2\,Q^2\,(\Delta t_{\text{MAX}})^2\,\Delta t_\nu) \\
&\quad \cdot \exp\big(C\,[\lambda(aq_\nu + u_\nu)\cdot \mathbb{1}_{|aq_\nu+u_\nu|<CQ}\cdot(\Delta t^*_\nu)]^2\,\Delta\tilde{t}_\nu\big).
\end{aligned}
\tag{3.39}
$$

(Here, we use the fact that ΔW_ν is independent of $\mathcal{F}_\nu$, and normal with mean 0 and variance $\Delta\tilde{t}_\nu$.)

If we take $\hat{C}$ large enough, then the product of the exponentials on the right in (3.39) is less than 1. Thus,

$$
\mathrm{E}[Y_{\nu+1}\,|\,\mathcal{F}_\nu] \le Y_\nu, \quad \text{with } Y_0 \equiv 1.
$$

So the Y_ν form a supermartingale. Consequently,

$$
\mathrm{Prob}(\exists\,\nu \text{ such that } Y_\nu > \exp(c\,Q^2)) \le \exp(-c\,Q^2). \tag{3.40}
$$

We now pick $\lambda = [\mathrm{sgn}]\tilde{c}\,[(\Delta t_{\text{MAX}})]^{-1}$, with sgn $= \pm 1$ and $\tilde{c} > 0$ a small enough constant. Combining (3.38) and (3.40) then yields the estimate

$$
\begin{aligned}
\mathrm{Prob}\Big(\exists\,\nu \text{ such that } [(\Delta t_{\text{MAX}})]^{-1}\Big|\sum_{\mu<\nu}(aq_\mu + u_\mu)\cdot \mathbb{1}_{|aq_\mu+u_\mu|<CQ}(\Delta t^*_\mu)\Delta W_\mu\Big| > CQ^2\Big) \\
\le \exp(-c\,Q^2).
\end{aligned}
$$

Consequently,

$$
\begin{aligned}
\mathrm{Prob}\Big(\exists\,\nu \text{ such that } \Big|\sum_{\mu<\nu}(aq_\mu + u_\mu)(\Delta t^*_\mu)\Delta W_\mu\Big| > CQ^2(\Delta t_{\text{MAX}})\Big) \\
\le \exp(-c\,Q^2) + \mathrm{Prob}(\exists\,\mu \text{ such that } |aq_\mu + u_\mu| > CQ).
\end{aligned}
$$

Recalling (3.36), we have

$$\text{Prob}\big(\max_\nu |\text{TERM}\,2(\nu)| > CQ^2(\Delta t_{\text{MAX}})\big) \leq \exp(-cQ^2) + \text{Prob}\big(\max_\mu\{|q_\mu| + |u_\mu|\} > cQ\big). \tag{3.41}$$

We turn our attention to TERM 3(ν). Let

$$\mathcal{Z}_\nu = \exp(-\hat{C}\lambda^2(\Delta t_{\text{MAX}})t_\nu)\cdot\exp\Big(\lambda\sum_{\mu<\nu}\{(\Delta W_\mu)^2 - \Delta\tilde{t}_\mu\}\Big) \tag{3.42}$$

for a large enough constant $\hat{C}$, and for $\lambda \in \mathbb{R}$ to be picked below.

Then $\mathcal{Z}_0 \equiv 1$, and

$$\text{E}[\mathcal{Z}_{\nu+1}\,|\,\mathcal{F}_\nu] = \mathcal{Z}_\nu \exp(-\hat{C}\lambda^2(\Delta t_{\text{MAX}})\Delta t_\nu)\cdot\text{E}[\exp(\lambda\{(\Delta W_\nu)^2 - \Delta\tilde{t}_\nu\})]. \tag{3.43}$$

For $|\lambda| < c(\Delta t_{\text{MAX}})^{-1}$, we have

$$\begin{aligned}\text{E}[\exp(\lambda\{(\Delta W_\nu)^2 - \Delta\tilde{t}_\nu\})] &= \exp(-\lambda\Delta\tilde{t}_\nu)\cdot\frac{1}{\sqrt{2\pi\Delta\tilde{t}_\nu}}\int_{-\infty}^{\infty} e^{\lambda x^2}e^{-x^2/2\Delta\tilde{t}_\nu}\,dx\\ &= \exp(-\lambda\Delta\tilde{t}_\nu)\cdot(1-2\lambda(\Delta\tilde{t}_\nu))^{-1/2}\\ &\leq \exp(C\lambda^2(\Delta t_\nu)^2) \leq \exp(C\lambda^2(\Delta t_{\text{MAX}})\Delta t_\nu).\end{aligned} \tag{3.44}$$

Substituting (3.44) into (3.43), and taking $\hat{C}$ large enough, we find that

$$\text{E}[\mathcal{Z}_{\nu+1}\,|\,\mathcal{F}_\nu] \leq \mathcal{Z}_\nu,$$

i.e., the $\mathcal{Z}_\nu$ form a supermartingale. This holds for $|\lambda| < c(\Delta t_{\text{MAX}})^{-1}$. Since $\mathcal{Z}_0 \equiv 1$, it follows that

$$\text{Prob}(\exists\,\nu \text{ such that } \mathcal{Z}_\nu > \exp(cQ^2)) \leq \exp(-cQ^2).$$

Taking $\lambda = (\Delta t_{\text{MAX}})^{-1/2}\cdot[\text{sgn}]$, with sgn $= \pm 1$, we conclude that

$$\text{Prob}\Big(\exists\,\nu \text{ such that } \exp\Big(-\hat{C}t_\nu + (\Delta t_{\text{MAX}})^{-1/2}\Big|\sum_{\mu<\nu}\{(\Delta W_\mu)^2 - \Delta\tilde{t}_\mu\}\Big|\Big) > \exp(cQ^2)\Big) \leq \exp(-cQ^2),$$

so that

$$\text{Prob}\Big(\max_\nu\Big|\sum_{\mu<\nu}\{(\Delta W_\mu)^2 - \Delta\tilde{t}_\mu\}\Big| > CQ^2(\Delta t_{\text{MAX}})^{1/2}\Big) \leq \exp(-cQ^2).$$

Recalling (3.36), we conclude that

$$\text{Prob}\big(\max_\nu |\text{TERM}\,3(\nu)| > CQ^2(\Delta t_{\text{MAX}})^{1/2}\big) \leq \exp(-cQ^2). \tag{3.45}$$

Estimates (3.37), (3.41) and (3.45) control TERM 1(ν), TERM 2(ν) and TERM 3(ν). Substituting these estimates into (3.36), we learn that

$$\mathrm{Prob}\Big(\max_{\nu}\Big|\sum_{\mu<\nu}\{(\Delta q_\mu)^2-\Delta\tilde{t}_\mu\}\Big| > C'Q^2(\Delta t_{\mathrm{MAX}})^{1/2}\Big) \\ \leq C\exp(-cQ^2)+C\,\mathrm{Prob}\big(\max_{\mu}\{|q_\mu|+|u_\mu|\}>cQ\big).$$

Finally, recalling (3.25) and (3.26), we see that

$$\mathrm{Prob}\Big(\max_{\nu}\Big|\sum_{\mu<\nu}\{(\Delta q_\mu)^2-\Delta\tilde{t}_\mu\}\Big| > C''Q^2(\Delta t_{\mathrm{MAX}})^{1/2}\Big) \leq C\exp(-cQ^2),$$

completing the proof of (3.33).

Next, we estimate

$$\Delta q_\nu^\sigma = q^\sigma(t_{\nu+1})-q^\sigma(t_\nu),\quad \Delta\zeta_{1,\nu}^\sigma=\zeta_1^\sigma(t_{\nu+1})-\zeta_1^\sigma(t_\nu),\quad \Delta\zeta_{2,\nu}^\sigma=\zeta_2^\sigma(t_{\nu+1})-\zeta_2^\sigma(t_\nu).$$

Recall that

$$\Delta q_\nu^\sigma=(aq_\nu^\sigma+u_\nu^\sigma)\Delta t_\nu^*+\Delta W_\nu,\quad \Delta\zeta_{1,\nu}^\sigma=q_\nu^\sigma(\Delta q_\nu^\sigma-u_\nu^\sigma\Delta t_\nu),\quad \Delta\zeta_{2,\nu}^\sigma=(q_\nu^\sigma)^2\,\Delta t_\nu.$$

Let $Q\geq C$ for large enough C, and suppose $|q_\nu^\sigma|\leq Q$. Then also $|u_\nu^\sigma|\leq C[|q_\nu^\sigma|+1]\leq C'Q$, so

$$\begin{aligned}&|\Delta q_\nu^\sigma|\leq CQ(\Delta t_\nu)+|\Delta W_\nu|,\\ &|\Delta\zeta_{1,\nu}^\sigma|\leq Q(|\Delta q_\nu^\sigma|+CQ\Delta t_\nu)\leq C'Q^2(\Delta t_\nu)+Q|\Delta W_\nu|,\ \text{and}\\ &|\Delta\zeta_{2,\nu}^\sigma|\leq Q^2(\Delta t_\nu),\end{aligned}$$

hence for $p\geq 1$, we have

$$(|\Delta q_\nu^\sigma|+|\Delta\zeta_{1,\nu}^\sigma|+|\Delta\zeta_{2,\nu}^\sigma|)^p\leq C_p\,Q^{2p}(\Delta t_\nu)^p+C_p\,Q^p|\Delta W_\nu|^p.$$

Recall that ΔW_ν is independent of $\mathcal{F}_\nu$ and normal, with mean 0 and variance at most $C\Delta t_\nu$. It follows that, for any $p\geq 1$, we have

$$\mathrm{E}[(|\Delta q_\nu^\sigma|+|\Delta\zeta_{1,\nu}^\sigma|+|\Delta\zeta_{2,\nu}^\sigma|)^p\,|\,\mathcal{F}_\nu]\leq C_p\,Q^{2p}(\Delta t_\nu)^p+C_p\,Q^p(\Delta t_\nu)^{p/2} \tag{3.46}$$

whenever $|q_\nu^\sigma|\leq Q$. (Recall, q_ν^σ is deterministic once we condition on $\mathcal{F}_\nu$.) In particular, (3.46) implies that

$$\mathrm{Prob}[|\Delta q_\nu^\sigma|+|\Delta\zeta_{1,\nu}^\sigma|+|\Delta\zeta_{2,\nu}^\sigma|>(\Delta t_\nu)^{2/5}\,|\,\mathcal{F}_\nu]\leq C(\Delta t_\nu)^{1000} \tag{3.47}$$

if $|q_\nu^\sigma|\leq Q$ and $C\leq Q\leq(\Delta t_\nu)^{-1/1000}$.

Together with (3.46) for $p=2,4$ and Cauchy–Schwarz, (3.47) implies the estimate

$$\mathrm{E}\big[(|\Delta q_\nu^\sigma|+|\Delta\zeta_{1,\nu}^\sigma|+|\Delta\zeta_{2,\nu}^\sigma|)^p\cdot\mathbb{1}_{|\Delta q_\nu^\sigma|+|\Delta\zeta_{1,\nu}^\sigma|+|\Delta\zeta_{2,\nu}^\sigma|>(\Delta t_\nu)^{2/5}}\,|\,\mathcal{F}_\nu\big] \\ \leq C(\Delta t_\nu)^{100}\quad\text{for } p=1,2,$$

provided $|q_\nu^\sigma|\leq Q$ and $C'\leq Q\leq(\Delta t_\nu)^{-1/1000}$ for large enough C'.

Next, we estimate $|q^\sigma(t) - q^\sigma_\nu| = |q^\sigma(t) - q^\sigma(t_\nu)|$ for $t \in [t_\nu, t_{\nu+1}]$.
Recall that

$$q^\sigma(t) - q^\sigma_\nu = (aq^\sigma_\nu + u^\sigma_\nu) \cdot \Big[\frac{e^{a(t-t_\nu)} - 1}{a}\Big] + \int_{t_\nu}^t e^{a(t-s)}\, dW(s) \quad \text{for } t \in [t_\nu, t_{\nu+1}].$$

If $|q^\sigma_\nu| \le Q$ with $Q \ge C$ (for large enough C), then also $|u^\sigma_\nu| \le CQ$, hence

$$|q^\sigma(t) - q^\sigma_\nu| \le CQ(\Delta t_\nu) + \Big|\int_{t_\nu}^t e^{a(t-s)}\, dW(s)\Big| \quad \text{for } t \in [t_\nu, t_{\nu+1}].$$

Applying the reflection principle [18] to the Gaussian process

$$W^\#(t) = \int_{t_\nu}^t e^{-as}\, dW(s) \quad (t \ge t_\nu),$$

we see that

$$\text{Prob}\Big[\max_{t\in[t_\nu,t_{\nu+1}]} \Big|\int_{t_\nu}^t e^{a(t-s)}\, dW(s)\Big| > CQ(\Delta t_\nu)^{1/2}\Big] \le C\exp(-cQ^2).$$

Since $\int_{t_\nu}^t e^{a(t-s)}\, dW(s)$ $(t \ge t_\nu)$ is independent of $\mathcal{F}_\nu$, it now follows that

$$\text{Prob}\Big[\max_{t\in[t_\nu,t_{\nu+1}]} |q^\sigma(t) - q^\sigma_\nu| > C'Q(\Delta t_\nu)^{1/2} \Big| \mathcal{F}_\nu\Big] \le C\exp(-cQ^2)$$

provided $|q^\sigma_\nu| \le Q$ and $C \le Q$ for large enough C. Taking

$$Q = \frac{(\Delta t_\nu)^{2/5}}{C'(\Delta t_\nu)^{1/2}},$$

we find that

$$\text{Prob}\Big[\max_{t\in[t_\nu,t_{\nu+1}]} |q^\sigma(t) - q^\sigma_\nu| > (\Delta t_\nu)^{2/5} \Big| \mathcal{F}_\nu\Big] \le C(\Delta t_\nu)^{1000} \tag{3.48}$$

provided $|q^\sigma_\nu| \le c \cdot (\Delta t_\nu)^{-1/10}$.

Let us summarize the results of the above discussion.

Lemma 3.2 (Lemma on rare events). *We condition on $a_{\text{TRUE}} = a$ and $\vec{\xi} = \vec{\eta}$. Fix a strategy σ. For constants c and C depending only on upper bounds for $|q_0|$, a_{MAX}, C_{TAME}, and T, the following holds.*

Suppose $\Delta t_{\text{MAX}} \equiv \max_\nu(t_{\nu+1} - t_\nu) < c$. Then, for $Q > C$, the following hold with probability $> 1 - \exp(-cQ^2)$:

- $|q^\sigma(t_\nu)|, |u^\sigma(t_\nu)| \le Q$ *for all ν.*
- $|\zeta^\sigma_1(t_\nu)|, |\zeta^\sigma_2(t_\nu)| \le Q^2$ *for all ν.*
- $\big|\sum_{0\le\mu<\nu}(q^\sigma(t_{\mu+1}) - q^\sigma(t_\mu))^2 - t_\nu\big| \le Q^2(\Delta t_{\text{MAX}})^{1/2}$ *for all ν.*

Moreover, suppose we fix ν and condition on $\mathcal{F}_\nu$, the sigma algebra of events determined by the $q^\sigma(t_\mu)$ $(0 \le \mu \le \nu)$. Suppose that $|q_\nu| < (\Delta t_\nu)^{-1/1000}$.

Then

$$\mathrm{E}\big[(|\Delta q^\sigma_\nu| + |\Delta\zeta^\sigma_{1,\nu}| + |\Delta\zeta^\sigma_{2,\nu}|)^p \cdot \mathbb{1}_{|\Delta q^\sigma_\nu|+|\Delta\zeta^\sigma_{1,\nu}|+|\Delta\zeta^\sigma_{2,\nu}|>(\Delta t_\nu)^{2/5}} \,|\, \mathcal{F}_\nu\big] \le C(\Delta t_\nu)^{100}$$

for $p = 1, 2$, *and*

$$\mathrm{Prob}\big[|\Delta q^\sigma_\nu| + |\Delta\zeta^\sigma_{1,\nu}| + |\Delta\zeta^\sigma_{2,\nu}| > (\Delta t_\nu)^{2/5} \,|\, \mathcal{F}_\nu\big] \le C(\Delta t_\nu)^{1000}.$$

Also, we have

$$\mathrm{Prob}\Big[\max_{t\in[t_\nu,t_{\nu+1}]} |q^\sigma(t) - q^\sigma_\nu| > (\Delta t_\nu)^{2/5} \,|\, \mathcal{F}_\nu\Big] \le C(\Delta t_\nu)^{1000}$$

provided $|q^\sigma_\nu| \le c \cdot (\Delta t_\nu)^{-1/10}$. *Finally, for* $Q \ge C$, *we have*

$$\mathrm{Prob}\Big[\max_{t\in[t_\nu,t_{\nu+1}]} |q^\sigma(t) - q^\sigma_\nu| > C'Q(\Delta t_\nu)^{1/2} \,|\, \mathcal{F}_\nu\Big] \le C\exp(-cQ^2) \quad \textit{if } |q_\nu| \le Q.$$

Proof. To deduce the third bullet point from (3.33), we note that

$$\sum_{0\le\mu<\nu} (\Delta\tilde{t}_\mu) = \sum_{0\le\mu<\nu} [(\Delta t_\mu) + O((\Delta t_\mu)^2)] = t_\nu \cdot (1 + O(\Delta t_{\mathrm{MAX}})).$$

The remaining assertions of the lemma have already been proved as stated. ■

3.3. The probability density

We continue to adopt the assumptions and notation of Section 3.2. Our goal is to derive, for fixed $\bar{N} \le N$, an approximate formula for the joint probability density of $(q_1, \dots, q_{\bar{N}}) = (q^\sigma(t_1), \dots, q^\sigma(t_{\bar{N}}))$. Let us denote this joint probability density by $\Phi(\bar{q}_1, \dots, \bar{q}_{\bar{N}})$. Thus,

$$\mathrm{Prob}((q^\sigma(t_1), \dots, q^\sigma(t_{\bar{N}})) \in E) = \int_E \Phi(\bar{q}_1, \dots, \bar{q}_{\bar{N}})\, d\bar{q}_1 \cdots d\bar{q}_{\bar{N}} \tag{3.49}$$

for measurable sets $E \subset \mathbb{R}^{\bar{N}}$.

By formula (3.8),

$$\Delta q_\nu = (aq_\nu + u_\nu)\,\Delta t^*_\nu + \Delta W_\nu$$

with ΔW_ν mutually independent and normal, with mean 0 and variance $\Delta\tilde{t}_\nu$; consequently, the joint probability Φ is given by

$$\Phi(\bar{q}_1, \dots, \bar{q}_{\bar{N}}) = \prod_{\nu=0}^{\bar{N}-1} \phi_\nu \tag{3.50}$$

with

$$\phi_\nu = \frac{1}{\sqrt{2\pi\Delta\tilde{t}_\nu}} \exp\Big(-\frac{1}{2\Delta\tilde{t}_\nu}[\Delta\bar{q}_\nu - (a\bar{q}_\nu + \bar{u}_\nu)\,\Delta t^*_\nu]^2\Big). \tag{3.51}$$

Here, $\Delta\bar{q}_\nu \equiv \bar{q}_{\nu+1} - \bar{q}_\nu$, $\bar{q}_0 \equiv q_0$, and $\bar{u}_\nu$ denotes the control exercised by the strategy σ at time t_ν given that $q^\sigma(t_\mu) = \bar{q}_\mu$ for $0 \le \mu \le \nu$ and $\vec{\xi} = \vec{\eta}$. Note that $\bar{u}_\nu$ is determined by $\bar{q}_1, \dots, \bar{q}_\nu$ (and q_0).

We make the following assumptions on $(\bar{q}_1, \ldots, \bar{q}_{\bar{N}})$:

$$\max_\nu (|\bar{q}_\nu| + |\bar{u}_\nu|) \le Q, \tag{3.52}$$

$$\Big| \sum_\nu (\Delta \bar{q}_\nu)^2 - t_{\bar{N}} \Big| \le Q^2 (\Delta t_{\text{MAX}})^{1/4}, \tag{3.53}$$

with $Q \ge C$ given.

Thanks to Lemma 3.2, (3.52) and (3.53) are very likely true for $(\bar{q}_1, \ldots, \bar{q}_{\bar{N}}) = (q^\sigma(t_1), \ldots, q^\sigma(t_{\bar{N}}))$. Under assumptions (3.52) and (3.53), we will simplify the expressions (3.50) and (3.51).

First of all, since $\Delta t^*_\nu = \Delta t_\nu + O((\Delta t_\nu)^2)$ we have

$$[\Delta \bar{q}_\nu - (a\bar{q}_\nu + \bar{u}_\nu)\Delta t^*_\nu] = [\Delta \bar{q}_\nu - (a\bar{q}_\nu + \bar{u}_\nu)\Delta t_\nu] + \text{ERR}_\nu,$$

with

$$\text{ERR}_\nu = O(Q(\Delta t_\nu)^2)$$

thanks to (3.52). Hence,

$$\begin{aligned}
&[\Delta \bar{q}_\nu - (a\bar{q}_\nu + \bar{u}_\nu)\Delta t^*_\nu]^2 \\
&\quad = [\Delta \bar{q}_\nu - (a\bar{q}_\nu + \bar{u}_\nu)^2 \Delta t_\nu]^2 + \text{ERR}_\nu^2 + 2\,\text{ERR}_\nu [\Delta \bar{q}_\nu - (a\bar{q}_\nu + \bar{u}_\nu)\Delta t_\nu] \\
&\quad = [\Delta \bar{q}_\nu - (a\bar{q}_\nu + \bar{u}_\nu)\Delta t_\nu]^2 + O(Q^2(\Delta t_\nu)^3) + 2\,\text{ERR}_\nu(\Delta \bar{q}_\nu).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\sum_\nu \frac{1}{2\Delta \tilde{t}_\nu} [\Delta \bar{q}_\nu - (a\bar{q}_\nu + \bar{u}_\nu)\Delta t^*_\nu]^2 \\
&= \sum_\nu \frac{1}{2\Delta \tilde{t}_\nu} [\Delta \bar{q}_\nu - (a\bar{q}_\nu + \bar{u}_\nu)\Delta t_\nu]^2 + \sum_\nu O(Q^2(\Delta t_\nu)^2) + \sum_\nu \frac{\text{ERR}_\nu}{\Delta \tilde{t}_\nu}(\Delta \bar{q}_\nu).
\end{aligned} \tag{3.54}$$

The last sum has absolute value at most

$$C \sum_\nu (\Delta t_\nu)^{-1/2} \left\{ \frac{\text{ERR}_\nu}{\Delta \tilde{t}_\nu} \right\}^2 + C \sum_\nu (\Delta t_\nu)^{1/2} (\Delta \bar{q}_\nu)^2. \tag{3.55}$$

The expression (3.55) is, in turn, at most

$$\begin{aligned}
&\sum_\nu O(Q^2(\Delta t_\nu)^{3/2}) + C(\Delta t_{\text{MAX}})^{1/2} \sum_\nu (\Delta \bar{q}_\nu)^2 \\
&\quad = O(Q^2(\Delta t_{\text{MAX}})^{1/2}) + C(\Delta t_{\text{MAX}})^{1/2} \Big| \sum_\nu (\Delta \bar{q}_\nu)^2 - t_{\bar{N}} \Big| + C t_{\bar{N}} (\Delta t_{\text{MAX}})^{1/2} \\
&\quad = O(Q^2(\Delta t_{\text{MAX}})^{1/2}),
\end{aligned}$$

thanks to (3.53). Therefore, (3.54) implies that

$$\begin{aligned}
&\sum_\nu \frac{1}{2\Delta \tilde{t}_\nu} [\Delta \bar{q}_\nu - (a\bar{q}_\nu + \bar{u}_\nu)\Delta t^*_\nu]^2 \\
&\qquad = \sum_\nu \frac{1}{2\Delta \tilde{t}_\nu} [\Delta \bar{q}_\nu - (a\bar{q}_\nu + \bar{u}_\nu)\Delta t_\nu]^2 + O(Q^2(\Delta t_{\text{MAX}})^{1/2}).
\end{aligned} \tag{3.56}$$

We want to replace $\Delta\tilde{t}_\nu$ by Δt_ν on the right in (3.56). To do so, note that

$$\frac{1}{2\Delta\tilde{t}_\nu} - \frac{1}{2\Delta t_\nu} = -\frac{a}{2} + O(\Delta t_\nu),$$

thanks to (3.11). Consequently,

$$\begin{aligned}
&\sum_\nu \left(\frac{1}{2\Delta\tilde{t}_\nu} - \frac{1}{2\Delta t_\nu}\right)[\Delta\bar{q}_\nu - (a\bar{q}_\nu + \bar{u}_\nu)\Delta t_\nu]^2 \\
&= \sum_\nu [a + O(\Delta t_\nu)]\,(a\bar{q}_\nu + \bar{u}_\nu)\Delta t_\nu(\Delta\bar{q}_\nu) + \sum_\nu \Big[-\frac{a}{2} + O(\Delta t_\nu)\Big](\Delta\bar{q}_\nu)^2 \\
&\quad + \sum_\nu \Big[-\frac{a}{2} + O(\Delta t_\nu)\Big](a\bar{q}_\nu + \bar{u}_\nu)^2(\Delta t_\nu)^2 \\
&\equiv \text{TERM } \alpha + \text{TERM } \beta + \text{TERM } \gamma
\end{aligned} \tag{3.57}$$

Now

$$\text{TERM } \beta = -\frac{a}{2}\sum_\nu(\Delta\bar{q}_\nu)^2 + O(\Delta t_{\text{MAX}})\sum_\nu(\Delta\bar{q}_\nu)^2 = -\frac{a}{2}\,t_{\bar{N}} + O(Q^2(\Delta t_{\text{MAX}})^{1/4})$$

by (3.53), while

$$\text{TERM } \gamma = \sum_\nu O(Q^2(\Delta t_\nu)^2) = O(Q^2(\Delta t_{\text{MAX}}))$$

by (3.52). To estimate TERM α, we apply (3.52) and (3.53) to write

$$\begin{aligned}
|\text{TERM } \alpha| &\le C\sum_\nu(a\bar{q}_\nu + \bar{u}_\nu)^2 2(\Delta t_\nu)^{3/2} + C\sum_\nu(\Delta t_\nu)^{1/2}(\Delta\bar{q}_\nu)^2 \\
&\le O(Q^2(\Delta t_{\text{MAX}})^{1/2}) + C(\Delta t_{\text{MAX}})^{1/2}\sum_\nu(\Delta\bar{q}_\nu)^2 \\
&= O(Q^2(\Delta t_{\text{MAX}})^{1/2}) + C(\Delta t_{\text{MAX}})^{1/2}\Big[\sum_\nu(\Delta\bar{q}_\nu)^2 - t_{\bar{N}}\Big] = O(Q^2(\Delta t_{\text{MAX}})^{1/2}).
\end{aligned}$$

Combining our estimates for Terms α, β and γ, and recalling (3.57), we learn that

$$\sum_\nu \left(\frac{1}{2\Delta\tilde{t}_\nu} - \frac{1}{2\Delta t_\nu}\right)[\Delta\bar{q}_\nu - (a\bar{q}_\nu + \bar{u}_\nu)\,\Delta t_\nu]^2 = -\frac{a}{2}\,t_{\bar{N}} + O(Q^2(\Delta t_{\text{MAX}})^{1/4}).$$

Consequently, (3.56) implies that

$$\begin{aligned}
&\sum_\nu \frac{1}{2\Delta\tilde{t}_\nu}[\Delta\bar{q}_\nu - (a\bar{q}_\nu + \bar{u}_\nu)\,\Delta t^*_\nu]^2 \\
&\quad = -\frac{a}{2}\,t_{\bar{N}} + \sum_\nu \frac{1}{2\Delta t_\nu}[\Delta\bar{q}_\nu - (a\bar{q}_\nu + \bar{u}_\nu)\,\Delta t_\nu]^2 + O(Q^2(\Delta t_{\text{MAX}})^{1/4}).
\end{aligned} \tag{3.58}$$

Again applying (3.11), we see that

$$\frac{1}{\sqrt{2\pi\Delta\tilde{t}_\nu}} = \frac{1}{\sqrt{2\pi\Delta t_\nu}}\left(1 - \frac{1}{2}a\Delta t_\nu + O(\Delta t_\nu)^2\right) = \frac{1}{\sqrt{2\pi\Delta t_\nu}}\exp\left(-\frac{1}{2}a\Delta t_\nu + O(\Delta t_\nu)^2\right),$$

so that

$$\begin{aligned}\prod_{\nu=0}^{\bar{N}-1} \frac{1}{\sqrt{2\pi \Delta \tilde{t}_\nu}} &= \Big(\prod_{\nu=0}^{\bar{N}-1} \frac{1}{\sqrt{2\pi \Delta t_\nu}} \Big) \exp\Big(-\frac{a}{2} \sum_\nu \{(\Delta t_\nu) + O(\Delta t_\nu)^2\} \Big) \\ &= \Big(\prod_{\nu=0}^{\bar{N}-1} \frac{1}{\sqrt{2\pi \Delta t_\nu}} \Big) \exp\Big(-\frac{a}{2} t_{\bar{N}} + O(\Delta t_{\text{MAX}}) \Big).\end{aligned} \tag{3.59}$$

Putting (3.58) and (3.59) into (3.50) and (3.51), we find that

$$\begin{aligned}\Phi(\bar{q}_1, \ldots, \bar{q}_{\bar{N}}) = \prod_{\nu=0}^{\bar{N}-1} \Big\{ \frac{1}{\sqrt{2\pi \Delta t_\nu}} \exp\Big(-\frac{1}{2\Delta t_\nu} [\Delta \bar{q}_\nu - (a\bar{q}_\nu + \bar{u}_\nu)\Delta t_\nu]^2 \Big) \Big\} \\ \cdot (1 + O(Q^2 (\Delta t_{\text{MAX}})^{1/4})).\end{aligned} \tag{3.60}$$

In particular, the factors $\exp(\frac{a}{2} t_{\bar{N}})$ arising from (3.58) and (3.59) cancel.

We record our result (3.60) as a lemma.

Lemma 3.3 (Lemma on the probability distribution). *We condition on $a_{\text{TRUE}} = a$ and $\vec{\xi} = \vec{\eta}$.*

Then, for constants c and C determined by q_0, a_{MAX}, C_{TAME}, and an upper bound for T, the following holds.

Suppose $\Delta t_{\text{MAX}} = \max_\nu (t_{\nu+1} - t_\nu) < c$. Fix $\bar{N} \le N$. Let $\Phi(\bar{q}_1, \ldots, \bar{q}_{\bar{N}})$ be the joint probability density for $(q^\sigma(t_1), \ldots, q^\sigma(t_{\bar{N}}))$.

Let $Q > C$, and suppose $(\bar{q}_1, \ldots, \bar{q}_{\bar{N}})$ satisfies

$$\max_\nu (|\bar{q}_\nu| + |\bar{u}_\nu|) \le Q \quad \text{and} \quad \Big| \sum_\nu (\bar{q}_{\nu+1} - \bar{q}_\nu)^2 - t_{\bar{N}} \Big| \le Q^2 (\Delta t_{\text{MAX}})^{1/4},$$

where $\bar{u}_\nu$ is the control exercised by the strategy σ at time t_ν when

$$(q^\sigma(t_1), \ldots, q^\sigma(t_\nu)) = (\bar{q}_1, \ldots, \bar{q}_\nu) \quad \text{and} \quad \vec{\xi} = \vec{\eta}.$$

Then

$$\begin{aligned}\Phi(\bar{q}_1, \ldots, \bar{q}_{\bar{N}}) = \prod_{\nu=0}^{\bar{N}-1} \Big\{ \frac{1}{\sqrt{2\pi \Delta t_\nu}} \exp\Big(-\frac{1}{2\Delta t_\nu} [\Delta \bar{q}_\nu - (a\bar{q}_\nu + \bar{u}_\nu)\Delta t_\nu]^2 \Big) \Big\} \\ \cdot (1 + O(Q^2 (\Delta t_{\text{MAX}})^{1/4})).\end{aligned}$$

Here, $\Delta \bar{q}_\nu = \bar{q}_{\nu+1} - \bar{q}_\nu$ (with $\bar{q}_0 = q_0$), $\Delta t_\nu = t_{\nu+1} - t_\nu$, and $O(Q^2 (\Delta t_{\text{MAX}})^{1/4})$ denotes a quantity whose absolute value is at most $CQ^2 (\Delta t_{\text{MAX}})^{1/4}$.

3.4. Analytic continuation

In this section, we prepare to make an analytic continuation of the function mapping $a \in [-a_{\text{MAX}}, +a_{\text{MAX}}]$ to the expected cost incurred by a strategy σ assuming that $a_{\text{MAX}} = a$. We set up notation.

We fix a tame strategy $\sigma = (\sigma_\nu)_{0 \le \nu < N}$ and coin flips $\vec{\eta}$; we write $\bar{u}_\nu$ to denote the control exercised by the strategy σ at time t_ν assuming that $q^\sigma(t_\mu) = \bar{q}_\mu$ for $1 \le \mu \le \nu$ and $\vec{\xi} = \vec{\eta}$ (i.e., $\bar{u}_\nu = u^\sigma(t_\nu) = \sigma_\nu(\bar{q}_1, \ldots, \bar{q}_\nu, \vec{\eta})$).

We define functions

$$\psi_\nu(\Delta\bar{q}, \bar{q}, \bar{u}, a) := \frac{1}{\sqrt{2\pi\Delta t_\nu}} \exp\Big(- \frac{[\Delta\bar{q} - (a\bar{q} + \bar{u})\Delta t_\nu]^2}{2\Delta t_\nu}\Big) \tag{3.61}$$

and

$$\Psi(\bar{q}_1, \dots, \bar{q}_N, a) = \prod_{\nu=0}^{N-1} \psi_\nu(\Delta\bar{q}_\nu, \bar{q}_\nu, \bar{u}_\nu(\bar{q}_1, \dots, \bar{q}_\nu), a). \tag{3.62}$$

In (3.61), $\Delta\bar{q}$, $\bar{q}$ and $\bar{u}$ are real variables, while $a \in \mathbb{C}$. In (3.62), $\Delta\bar{q}_\nu := \bar{q}_{\nu+1} - \bar{q}_\nu$, with $\bar{q}_0 := q_0$. Again, in (3.62), $a \in \mathbb{C}$.

We introduce a set

$$\begin{aligned} E = \Big\{(\bar{q}_1, \dots, \bar{q}_N) \in \mathbb{R}^N : \max_\nu |\bar{q}_\nu| \le (\Delta t_{\text{MAX}})^{-1/16} \\ \text{and } |\sum_\nu (\bar{q}_{\nu+1} - \bar{q}_\nu)^2 - T| \le (\Delta t_{\text{MAX}})^{1/8}\Big\}. \end{aligned} \tag{3.63}$$

We denote by

$$\Phi(\bar{q}_1, \dots, \bar{q}_N, a) \quad (a \in [-a_{\text{MAX}}, +a_{\text{MAX}}])$$

the probability density for $(q^\sigma(t_1), \dots, q^\sigma(t_N))$ assuming that $a_{\text{TRUE}} = a$ and $\vec{\xi} = \vec{\eta}$.

According to Lemmas 3.2 and 3.3, we have

$$\Phi(\bar{q}_1, \dots, \bar{q}_N, a) = \Psi(\bar{q}_1, \dots, \bar{q}_N, a) \cdot (1 + \text{ERR}(\bar{q}_1, \dots, \bar{q}_N, a)) \tag{3.64}$$

for $(\bar{q}_1, \dots, \bar{q}_N) \in E$ and $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$, with

$$|\text{ERR}(\bar{q}_1, \dots, \bar{q}_N, a)| \le (\Delta t_{\text{MAX}})^{1/8}, \tag{3.65}$$

$$\mathrm{E}_{a,\vec{\eta}}[\mathbb{1}_{(q^\sigma(t_1),\dots,q^\sigma(t_N)) \notin E}] \le C \exp(-c(\Delta t_{\text{MAX}})^{-1/8}). \tag{3.66}$$

Since $|u^\sigma| \le C[|q^\sigma| + 1]$ for a tame rule σ, we have, for $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$,

$$\sum_\nu \{(u_\nu^\sigma)^2 + (q_\nu^\sigma)^2\}\Delta t_\nu \le C \max_\nu |q_\nu|^2 + C. \tag{3.67}$$

By Lemma 3.2, we have

$$\mathrm{E}_{a,\vec{\eta}}[\{\max_\nu |q^\sigma(t_\nu)|^2 + |u^\sigma(t_\nu)|^2\}^2] \le C. \tag{3.68}$$

We study the function

$$[-a_{\text{MAX}}, +a_{\text{MAX}}] \ni a \mapsto \mathrm{E}_{a,\vec{\eta}}\Big[\sum_{\nu=0}^{N-1} \{(q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2\}\Delta t_\nu\Big];$$

we denote this function by $\text{ECOST}(a)$.

The above definitions and estimates yield:

$$\text{ECOST}(a) = \mathrm{E}_{a,\vec{\eta}}\Big[\sum_{\nu=0}^{N-1}\{(q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2\}\,\Delta t_\nu \cdot \mathbb{1}_{(q^\sigma(t_1),\dots,q^\sigma(t_N))\in E}\Big] + \text{ERROR}\,1(a), \tag{3.69}$$

with

$$\begin{aligned} |\text{ERROR}\,1(a)| &\le \mathrm{E}_{a,\vec{\eta}}[C\,\max_\nu\{|q^\sigma(t_\nu)|^2 + |u^\sigma(t_\nu)|^2\}\cdot \mathbb{1}_{(q^\sigma(t_1),\dots,q^\sigma(t_\nu))\notin E}] \\ &\le \Big(\mathrm{E}_{a,\vec{\eta}}[C\,\max\{|q^\sigma(t_\nu)|^2 + |u^\sigma(t_\nu)|^2\}^2]\Big)^{1/2} \\ &\quad \cdot \big(\mathrm{Prob}_{a,\vec{\eta}}((q^\sigma(t_1),\dots,q^\sigma(t_N)) \notin E)\big)^{1/2} \\ &\le C' \exp(-c'(\Delta t_{\mathrm{MAX}})^{-1/8}). \end{aligned} \tag{3.70}$$

Moreover,

$$\begin{aligned} &\mathrm{E}_{a,\vec{\eta}}\Big[\sum_\nu\{(q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2\}\,\Delta t_\nu \cdot \mathbb{1}_{(q^\sigma(t_1),\dots,q^\sigma(t_\nu))\in E}\Big] \\ &= \int_{(\bar{q}_1,\dots,\bar{q}_N)\in E}\Big(\sum_\nu\{\bar{q}_\nu^2 + \bar{u}_\nu^2\}\,\Delta t_\nu\Big)\Phi(\bar{q}_1,\dots,\bar{q}_N,a)\,d\bar{q}_1\cdots d\bar{q}_N \\ &= \int_{(\bar{q}_1,\dots,\bar{q}_N)\in E}\Big\{\Big(\sum_\nu\{\bar{q}_\nu^2 + \bar{u}_\nu^2\}\,\Delta t_\nu\Big)\Psi(\bar{q}_1,\dots,\bar{q}_N,a) \\ &\quad \cdot (1 + \text{ERR}(\bar{q}_1,\dots,\bar{q}_N,a))\,d\bar{q}_1\cdots d\bar{q}_N\Big\} \\ &= (1+\text{ERROR}\,2(a))\int_{(\bar{q}_1,\dots,\bar{q}_N)\in E}\Big\{\Big(\sum_\nu\{\bar{q}_\nu^2+\bar{u}_\nu^2\}(\Delta t_\nu)\Big)\Psi(\bar{q}_1,\dots,\bar{q}_N,a)\,d\bar{q}_1\cdots d\bar{q}_N\Big\}, \end{aligned}$$

with

$$|\text{ERROR}\,2(a)| \le C(\Delta t_{\mathrm{MAX}})^{1/8}, \tag{3.71}$$

thanks to (3.64) and (3.65). Together with (3.69) and (3.70), this yields

$$\begin{aligned} \text{ECOST}(a) &= \text{ERROR}\,1(a) \\ &\quad + (1 + \text{ERROR}\,2(a))\cdot\int_E\Big\{\Big(\sum_\nu\{\bar{q}_\nu^2 + \bar{u}_\nu^2\}\,\Delta t_\nu\Big)\Psi(\bar{q}_1,\dots,\bar{q}_N,a)\Big\}\,d\bar{q}_1\cdots d\bar{q}_N, \end{aligned}$$

with ERROR $1(a)$ and ERROR $2(a)$ controlled by (3.70) and (3.71). Since also

$$0 \le \text{ECOST}(a) \le C$$

by Lemma 3.2, it follows that

$$\begin{aligned} \text{ECOST}(a) = \int_{(\bar{q}_1,\dots,\bar{q}_N)\in E}&\Big\{\Big(\sum_\nu\{\bar{q}_\nu^2 + \bar{u}_\nu^2\}\,\Delta t_\nu\Big) \\ &\cdot\Psi(\bar{q}_1,\dots,\bar{q}_N,a)\Big\}\,d\bar{q}_1\cdots d\bar{q}_N + \text{ERROR}\,3(a), \end{aligned} \tag{3.72}$$

with

$$|\text{ERROR}\,3(a)| \le C(\Delta t_{\text{MAX}})^{1/8}. \tag{3.73}$$

Equation (3.72) and estimate (3.73) hold for $a \in [-a_{\text{MAX}}, +a_{\text{MAX}}]$. We make an analytic continuation of the integral in (3.72) from $a \in [-a_{\text{MAX}}, +a_{\text{MAX}}]$ to $a = a_R + i a_I$ in the rectangle

$$\mathcal{R} = \{a_R + i a_I : a_R \in [-a_{\text{MAX}}, a_{\text{MAX}}], a_I \in [-\delta, \delta]\},$$

for a small $\delta > 0$ to be picked below.

We write $I(a)$ to denote the integral in (3.72). Thus,

$$I(a) = \int_E \Big(\sum_\nu \{\bar{q}_\nu^2 + \bar{u}_\nu^2\}\,\Delta t_\nu\Big)\Psi(\bar{q}_1, \ldots, \bar{q}_N, a)\; d\bar{q}_1 \cdots d\bar{q}_N \tag{3.74}$$

for $a = a_R + i a_I \in \mathcal{R}$, and

$$|\text{ECOST}(a) - I(a)| \le C(\Delta t_{\text{MAX}})^{1/8} \quad \text{for } a \in [-a_{\text{MAX}}, +a_{\text{MAX}}]. \tag{3.75}$$

A glance at (3.61) and (3.62) shows that the integrand in (3.74) has the form

$$B(\bar{q}_1, \ldots, \bar{q}_N)\exp(a^2 G_2(\bar{q}_1, \ldots, \bar{q}_N) + a G_1(\bar{q}_1, \ldots, \bar{q}_N) + G_0(\bar{q}_1, \ldots, \bar{q}_N)),$$

where B, G_0, G_1 and G_2 are bounded measurable functions of $(\bar{q}_1, \ldots, \bar{q}_N)$ on E. Moreover, the region of integration, E, is bounded; see (3.63). Therefore, $I(a)$ is an analytic function on $\mathcal{R}$.

Next, we estimate

$$\int_E \Big(\sum_\nu \{\bar{q}_\nu^2 + \bar{u}_\nu^2\}\Delta t_\nu\Big)\,|\Psi(\bar{q}_1, \ldots, \bar{q}_N, a_R + i a_I)|\; d\bar{q}_1 \cdots d\bar{q}_N \tag{3.76}$$

for $a_R + i a_I \in \mathcal{R}$. From (3.61), we have

$$|\psi_\nu(\Delta\bar{q}, \bar{q}, \bar{u}, a_R + i a_I)| = \exp\Big(\frac{a_I^2}{2}\,\bar{q}^2\,\Delta t_\nu\Big)\,\psi_\nu(\Delta\bar{q}, \bar{q}, \bar{u}, a_R),$$

hence (3.62) implies that

$$|\Psi(\bar{q}_1, \ldots, \bar{q}_N, a_R + i a_I)| = \exp\Big(\sum_\nu \frac{a_I^2}{2}\,\bar{q}_\nu^2\,\Delta t_\nu\Big)\,\Psi(\bar{q}_1, \ldots, \bar{q}_N, a_R).$$

Moreover, for $(\bar{q}_1, \ldots, \bar{q}_N) \in E$ and $a_R \in [-a_{\text{MAX}}, +a_{\text{MAX}}]$, (3.64) and (3.65) yield

$$\Psi(\bar{q}_1, \ldots, \bar{q}_N, a_R) \le 2\,\Phi(\bar{q}_1, \ldots, \bar{q}_N, a_R).$$

Consequently, for $a_R \in [-a_{\text{MAX}}, a_{\text{MAX}}]$, the integrand in (3.76) is at most

$$2\Big(\sum_\nu \{\bar{q}_\nu^2 + \bar{u}_\nu^2\}\,\Delta t_\nu\Big)\;\exp\Big(\frac{a_I^2}{2}\sum_\nu \bar{q}_\nu^2\,\Delta t_\nu\Big)\,\Phi(\bar{q}_1, \ldots, \bar{q}_N, a_R).$$

Since $\Phi(\bar{q}_1, \ldots, \bar{q}_N, a_R)$ is the probability density for $(q^\sigma(t_1), \ldots, q^\sigma(t_N))$ assuming $a_{\text{TRUE}} = a_R$ and $\vec{\xi} = \vec{\eta}$, it follows that the integral (3.76) is at most

$$2\,\mathrm{E}_{a_R,\vec{\eta}}\Big[\mathbb{1}_{(q^\sigma(t_1),\ldots,q^\sigma(t_N))\in E} \cdot \Big(\sum_\nu \{q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2\}\,\Delta t_\nu\Big) \cdot \exp\Big(\frac{a_I^2}{2}\sum_\nu (q^\sigma(t_\nu))^2\,\Delta t_\nu\Big)\Big] \quad \text{for } a \in \mathcal{R}.$$

Recall that $|u^\sigma(t_\nu)| \le C[|q^\sigma(t_\nu)| + 1]$ and that $|a_I| \le \delta$ for $a = a_R + i a_I \in \mathcal{R}$.

Consequently, for $a_R + i a_I \in \mathcal{R}$, we have

$$\Big(\sum_\nu \{(q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2\}\,\Delta t_\nu\Big)\exp\Big(\frac{a_I^2}{2}\sum_\nu (q^\sigma(t_\nu))^2\,\Delta t_\nu\Big) \le C_\delta \cdot \exp\big(C\delta^2 \max_\nu |q^\sigma(t_\nu)|^2\big).$$

So the integral (3.76) is at most

$$C_\delta\,\mathrm{E}_{a_R,\vec{\eta}}\big[\exp\big(C\delta^2 \max_\nu |q^\sigma(t_\nu)|^2\big)\big] \quad \text{for } a \in \mathcal{R}. \tag{3.77}$$

On the other hand, Lemma 3.2 gives

$$\mathrm{E}_{a_R,\vec{\eta}}[\exp(c \max_\nu |q^\sigma(t_\nu)|^2)] \le C \tag{3.78}$$

for $a_R \in [-a_{\text{MAX}}, a_{\text{MAX}}]$ and $c > 0$ small enough.

We now fix $\delta = \hat{c}$ small enough that $C\delta^2 < c$ with C and c as in (3.77) and (3.78). We conclude that the integral (3.76) is less than a large constant C, independent of $a \in \mathcal{R}$.

So we have shown that $I(a)$ is analytic and bounded for

$$a \in (-a_{\text{MAX}}, +a_{\text{MAX}}) \times (-\delta, \delta).$$

Recalling that we have taken $\delta = \hat{c}$ and that (3.75) holds, we obtain the following result.

Lemma 3.4 (Analytic continuation lemma). *Let σ be a tame strategy, and let $\vec{\eta} \in \{0,1\}^{\mathbb{N}}$. Then there exists an analytic function $I_{\vec{\eta}}(a)$ on the rectangle*

$$\mathcal{R} = \{a_R + i a_I : a_R \in (-a_{\text{MAX}}, a_{\text{MAX}}), a_I \in (-\hat{c}, \hat{c})\}$$

such that

$$|I_{\vec{\eta}}(a)| \le C \quad \text{on } \mathcal{R}$$

and

$$\Big|\mathrm{E}_{a,\vec{\eta}}\Big[\sum_{\nu=0}^{N-1}\{(q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2\}\,\Delta t_\nu\Big] - I_{\vec{\eta}}(a)\Big| \le C(\Delta t_{\text{MAX}})^{1/8}$$

for $a \in (-a_{\text{MAX}}, +a_{\text{MAX}})$. Explicitly,

$$I_{\vec{\eta}}(a) = \int_E \Big(\sum_{\nu=0}^{N-1}\{q_\nu^2 + u_\nu^2\}\,\Delta t_\nu\Big) \cdot \prod_{\nu=0}^{N-1}\Big\{\frac{1}{\sqrt{2\pi\Delta t_\nu}}\exp\Big(-\frac{[q_{\nu+1} - q_\nu - (aq_\nu + u_\nu)\Delta t_\nu]^2}{2\Delta t_\nu}\Big)\Big\}\,dq_1 \cdots dq_N,$$

where

$$E = \Big\{(\bar{q}_1, \dots, \bar{q}_N) \in \mathbb{R}^N : \max_\nu |\bar{q}_\nu| \le (\Delta t_{\text{MAX}})^{-1/16} \text{ and } \Big| \sum_\nu (\bar{q}_{\nu+1} - \bar{q}_\nu)^2 - T \Big| \le (\Delta t_{\text{MAX}})^{1/8} \Big\}. \tag{3.79}$$

and u_ν denotes the value assigned by σ to the control at time t_ν assuming that $q^\sigma(t_\mu) = q_\mu$ for $0 \le \mu \le \nu$ and $\vec{\xi} = \vec{\eta}$.

3.5. Moments of increments

We retain the assumptions and notation of Section 3.2. Recall that $\mathcal{F}_\nu$ is the sigma algebra of events determined by $q^\sigma(t_\mu)$ $(0 \le \mu \le \nu)$, and that

$$\Delta t_\nu = t_{\nu+1} - t_\nu, \qquad \Delta q_\nu = q^\sigma(t_{\nu+1}) - q^\sigma(t_\nu),$$
$$\Delta\zeta_{1,\nu} = \zeta_1^\sigma(t_{\nu+1}) - \zeta_1^\sigma(t_\nu), \quad \Delta\zeta_{2,\nu} = \zeta_2^\sigma(t_{\nu+1}) - \zeta_2^\sigma(t_\nu).$$

We suppose that $a_{\text{TRUE}} = a$ and $\vec{\xi} = \vec{\eta}$.

Recall also that

$$\Delta q_\nu = (aq_\nu + u_\nu)\Delta t_\nu^* + \Delta W_\nu, \tag{3.80}$$

$$\Delta\zeta_{1,\nu} = q_\nu(\Delta q_\nu - u_\nu \Delta t_\nu) = aq_\nu^2 \Delta t_\nu^* + q_\nu u_\nu(\Delta t_\nu^* - \Delta t_\nu) + q_\nu \Delta W_\nu, \tag{3.81}$$

$$\Delta\zeta_{2,\nu} = q_\nu^2 \Delta t_\nu. \tag{3.82}$$

We condition on $\mathcal{F}_\nu$; thus, q_ν and u_ν are deterministic, while ΔW_ν is normal, with mean 0 and variance $\Delta\tilde{t}_\nu$. We suppose that

$$|q_\nu|, |u_\nu| \le Q, \quad \text{with } Q \ge C \text{ given.} \tag{3.83}$$

Then

$$\Delta q_\nu = O(Q)\Delta t_\nu + \Delta W_\nu, \quad \Delta\zeta_{1,\nu} = O(Q^2)\Delta t_\nu + q_\nu \Delta W_\nu, \quad \Delta\zeta_{2,\nu} = O(Q^2)\Delta t_\nu,$$

so that

$$(\Delta q_\nu)^2 = O(Q^2)(\Delta t_\nu)^2 + 2\,O(Q)\,\Delta t_\nu\,\Delta W_\nu + (\Delta W_\nu)^2, \tag{3.84}$$

$$(\Delta\zeta_{1,\nu})(\Delta q_\nu) = O(Q^3)\,(\Delta t_\nu)^2 + O(Q^2)\,\Delta t_\nu\,\Delta W_\nu + q_\nu(\Delta W_\nu)^2, \tag{3.85}$$

$$(\Delta\zeta_{2,\nu})(\Delta q_\nu) = O(Q^3)\,(\Delta t_\nu)^2 + O(Q^2)\,\Delta t_\nu\,\Delta W_\nu, \tag{3.86}$$

$$(\Delta\zeta_{1,\nu})^2 = O(Q^4)\,(\Delta t_\nu)^2 + O(Q^3)\,(\Delta t_\nu)\,(\Delta W_\nu) + q_\nu^2(\Delta W_\nu)^2, \tag{3.87}$$

$$(\Delta\zeta_{2,\nu})(\Delta\zeta_{1,\nu}) = O(Q^4)\,(\Delta t_\nu)^2 + O(Q^3)\,(\Delta t_\nu)\,(\Delta W_\nu) \tag{3.88}$$

$$(\Delta\zeta_{2,\nu})^2 = O(Q^4)\,(\Delta t_\nu)^2. \tag{3.89}$$

Moreover, all the quantities $O(Q^{\text{power}})$ above are deterministic once we condition on $\mathcal{F}_\nu$.

Therefore, (3.80)–(3.82) and (3.84)–(3.89) yield the following:

(3.90) $$\mathrm{E}[\Delta q_\nu \,|\, \mathcal{F}_\nu] = (aq_\nu + u_\nu)\Delta t_\nu + O(Q(\Delta t_\nu)^2),$$

(3.91) $$\mathrm{E}[\Delta \zeta_{1,\nu} \,|\, \mathcal{F}_\nu] = aq_\nu^2(\Delta t_\nu) + O(Q^2(\Delta t_\nu)^2),$$

(3.92) $$\mathrm{E}[\Delta \zeta_{2,\nu} \,|\, \mathcal{F}_\nu] = q_\nu^2 \Delta t_\nu,$$

(3.93) $$\mathrm{E}[(\Delta \zeta_{1,\nu})^2 \,|\, \mathcal{F}_\nu] = q_\nu^2 \Delta t_\nu + O(Q^4(\Delta t_\nu)^2),$$

(3.94) $$\mathrm{E}[(\Delta \zeta_{1,\nu})(\Delta \zeta_{2,\nu}) \,|\, \mathcal{F}_\nu] = O(Q^4(\Delta t_\nu)^2), ,$$

(3.95) $$\mathrm{E}[(\Delta \zeta_{2,\nu})^2 \,|\, \mathcal{F}_\nu] = O(Q^4(\Delta t_\nu)^2),$$

(3.96) $$\mathrm{E}[(\Delta q_\nu)^2 \,|\, \mathcal{F}_\nu] = \Delta t_\nu + O(Q^2(\Delta t_\nu)^2),$$

(3.97) $$\mathrm{E}[(\Delta q_\nu)(\Delta \zeta_{1,\nu}) \,|\, \mathcal{F}_\nu] = q_\nu \Delta t_\nu + O(Q^3(\Delta t_\nu)^2),$$

(3.98) $$\mathrm{E}[(\Delta q_\nu)(\Delta \zeta_{2,\nu}) \,|\, \mathcal{F}_\nu] = O(Q^3(\Delta t_\nu)^2).$$

(Here we have used the fact that Δt_ν^* and $\Delta \tilde{t}_\nu$ are $\Delta t_\nu + O((\Delta t_\nu)^2)$.)

We define an event

(3.99) $$\text{TAME}_\nu = \{|\Delta q_\nu| \le 2(\Delta t_\nu)^{2/5}, |\Delta \zeta_{1,\nu}| \le 2(\Delta t_\nu)^{2/5}, |\Delta \zeta_{2,\nu}| \le 2(\Delta t_\nu)^{2/5}\}.$$

From Lemma 3.2, we obtain the estimate

$$\mathrm{E}[(\Delta q_\nu)^{\alpha_0}(\Delta \zeta_{1,\nu})^{\alpha_1}(\Delta \zeta_{2,\nu})^{\alpha_2} \cdot \mathbb{1}_{\text{NOT TAME}_\nu} \,|\, \mathcal{F}_\nu] = O((\Delta t_\nu)^{100})$$

for integers $\alpha_0, \alpha_1, \alpha_2 \ge 0$ with $\alpha_0 + \alpha_1 + \alpha_2 \le 2$. Also from Lemma 3.2, we recall that

$$\mathrm{Prob}[\text{NOT TAME}_\nu \,|\, \mathcal{F}_\nu] \le C \cdot (\Delta t_\nu)^{1000}.$$

Consequently, (3.90)–(3.98) imply the conclusions of the following lemma.

Lemma 3.5 (Lemma on moments of increments). *We fix $a_{\text{TRUE}} = a$ and $\vec{\xi} = \vec{\eta}$. We suppose that*

$$Q \ge C \quad \text{and} \quad (\Delta t_{\text{MAX}}) < Q^{-1000}.$$

Define the event

$$\text{TAME}(\nu) = \{|\Delta q_\nu^\sigma| \le 2(\Delta t_\nu)^{2/5}, |\Delta \zeta_{1,\nu}^\sigma| \le 2(\Delta t_\nu)^{2/5}, |\Delta \zeta_{2,\nu}^\sigma| \le 2(\Delta t_\nu)^{2/5}\}.$$

Let $\mathcal{F}_\nu$ be the sigma algebra of events determined by $q^\sigma(t_\mu)$ $(0 \le \mu \le \nu)$.

Fix ν, and suppose that

$$|q^\sigma(t_\nu)|, |u^\sigma(t_\nu)| \le Q.$$

Then the following hold:

$$\mathrm{E}[(\Delta q_\nu^\sigma)\,\mathbb{1}_{\text{TAME}(\nu)} \,|\, \mathcal{F}_\nu] = (aq_\nu^\sigma + u_\nu^\sigma)(\Delta t_\nu) + \text{ERR}\,1,$$
$$\mathrm{E}[(\Delta \zeta_{1,\nu}^\sigma)\,\mathbb{1}_{\text{TAME}(\nu)} \,|\, \mathcal{F}_\nu] = a(q_\nu^\sigma)^2\,(\Delta t_\nu) + \text{ERR}\,2,$$
$$\mathrm{E}[(\Delta \zeta_{2,\nu}^\sigma)\,\mathbb{1}_{\text{TAME}(\nu)} \,|\, \mathcal{F}_\nu] = (q_\nu^\sigma)^2\,(\Delta t_\nu) + \text{ERR}\,3$$
$$\mathrm{E}[(\Delta q_\nu^\sigma)^2\,\mathbb{1}_{\text{TAME}(\nu)} \,|\, \mathcal{F}_\nu] = (\Delta t_\nu) + \text{ERR}\,4,$$
$$\mathrm{E}[(\Delta q_\nu^\sigma)(\Delta \zeta_{1,\nu}^\sigma)\,\mathbb{1}_{\text{TAME}(\nu)} \,|\, \mathcal{F}_\nu] = q_\nu^\sigma(\Delta t_\nu) + \text{ERR}\,5,$$

$$\begin{aligned}
&\mathrm{E}[(\Delta q^\sigma_\nu)\,(\Delta\zeta^\sigma_{2,\nu})\,\mathbb{1}_{\text{TAME}(\nu)}\,|\,\mathcal{F}_\nu] = \text{ERR}\,6,\\
&\mathrm{E}[(\Delta\zeta^\sigma_{1,\nu})^2\,\mathbb{1}_{\text{TAME}(\nu)}\,|\,\mathcal{F}_\nu] = (q^\sigma_\nu)^2(\Delta t_\nu) + \text{ERR}\ 7,\\
&\mathrm{E}[(\Delta\zeta^\sigma_{1,\nu})\,(\Delta\zeta^\sigma_{2,\nu})\,\mathbb{1}_{\text{TAME}(\nu)}\,|\,\mathcal{F}_\nu] = \text{ERR}\,8,\\
&\mathrm{E}[(\Delta\zeta^\sigma_{2,\nu})^2\,\mathbb{1}_{\text{TAME}(\nu)}\,|\,\mathcal{F}_\nu] = \text{ERR}\,9,
\end{aligned}$$

where

$$|\text{ERR}\,1|,\ldots,|\text{ERR}\,9| \le C'Q^4(\Delta t_\nu)^2.$$

Also, under the above assumptions, we have

$$\text{Prob}[\text{NOT TAME}(\nu)\,|\,\mathcal{F}_\nu] \le (\Delta t_\nu)^{20}.$$

Here, of course,

$$\begin{aligned}
q^\sigma_\nu &= q^\sigma(t_\nu), & u^\sigma_\nu &= u^\sigma(t_\nu),\\
\Delta t_\nu &= t_{\nu+1} - t_\nu, & \Delta q^\sigma_\nu &= q^\sigma(t_{\nu+1}) - q^\sigma(t_\nu),\\
\Delta\zeta^\sigma_{1,\nu} &= \zeta^\sigma_1(t_{\nu+1}) - \zeta^\sigma_1(t_\nu), & \Delta\zeta^\sigma_{2,\nu} &= \zeta^\sigma_2(t_{\nu+1}) - \zeta^\sigma_2(t_\nu).
\end{aligned}$$

The constants C and C' are determined by q_0, a_{MAX}, C_{TAME} and an upper bound for T.

3.6. Stability under change of assumption

Let $f(\bar q_1,\ldots,\bar q_N)$ be a nonnegative function on $\mathbb{R}^N$. For $\vec\eta \in \{0,1\}^{\mathbb{N}}$ and for $a_1, a_2 \in [-a_{\text{MAX}}, +a_{\text{MAX}}]$, we compare

$$\mathrm{E}_{a_1,\vec\eta}[f(q^\sigma(t_1),\ldots,q^\sigma(t_N))]$$

with

$$\mathrm{E}_{a_2,\vec\eta}[f(q^\sigma(t_1),\ldots,q^\sigma(t_N))]$$

for a tame strategy σ.

To do so, let $\Phi(\bar q_1,\ldots,\bar q_N, a)$ denote the probability density of

$$(q^\sigma(t_1),\ldots,q^\sigma(t_N))$$

assuming that $\vec\xi = \vec\eta$ and $a_{\text{TRUE}} = a$.

According to Lemma 3.3, the following holds for $a = a_1, a_2$. Let $Q > C$ and $Q \le (\Delta t_{\text{MAX}})^{-1/1000}$, and suppose that

$$\begin{aligned}
&\max_\nu(|\bar q_\nu| + |\bar u_\nu|) \le Q,\\
&\Big|\sum_{0\le\nu<N}(\bar q_{\nu+1} - \bar q_\nu)^2 - T\Big| \le Q^2(\Delta t_{\text{MAX}})^{1/4},
\end{aligned} \tag{3.100}$$

where $\bar u_\nu$ denotes the value of $u^\sigma(t_\nu)$ assuming $q^\sigma(t_\mu) = \bar q_\mu$ for $\mu \le \nu$. (Recall that the $\bar u_\nu$ do not depend on a.) Then

$$\begin{aligned}
&\Phi(\bar q_1,\ldots,\bar q_N, a)\\
&\quad = \prod_{0\le\nu<N}\left\{\frac{1}{\sqrt{2\pi\Delta t_\nu}}\exp\Big(-\frac{1}{2\Delta t_\nu}\,[(\bar q_{\nu+1} - \bar q_\nu) - (a\bar q_\nu + \bar u_\nu)\,\Delta t_\nu]^2\Big)\right\}\\
&\quad\cdot(1 + \text{ERR}(\bar q_1,\ldots,\bar q_N, a)
\end{aligned} \tag{3.101}$$

with

$$|\mathrm{ERR}(\bar{q}_1, \ldots, \bar{q}_N, a)| \leq C Q^2 (\Delta t_{\mathrm{MAX}})^{1/4} \tag{3.102}$$

and $\bar{q}_0 \equiv q_0$.

Applying (3.101) and (3.102) with $a = a_1$ and with $a = a_2$, we see that if (3.100) holds, then

$$\begin{aligned}\Phi(\bar{q}_1, \ldots, \bar{q}_N, a_2) = \Phi(\bar{q}_1, \ldots, \bar{q}_N, a_1) & \\ \cdot \exp\Big(-\frac{1}{2}\,[a_2^2 - a_1^2]\,\bar{\xi}_2^{\sigma}(T) + [a_2 - a_1]\,\bar{\xi}_1^{\sigma}(T)\Big) & \\ \cdot (1 + \mathrm{ERR}(\bar{q}_1, \ldots, \bar{q}_N, a_1, a_2)), &\end{aligned} \tag{3.103}$$

with

$$\bar{\xi}_1^{\sigma}(T) := \sum_{0 \leq \nu < N} \bar{q}_\nu \left([\bar{q}_{\nu+1} - \bar{q}_\nu] - \bar{u}_\nu\, \Delta t_\nu\right), \quad \bar{\xi}_2^{\sigma}(T) := \sum_{0 \leq \nu < N} \bar{q}_\nu^2\, \Delta t_\nu, \tag{3.104}$$

and

$$|\mathrm{ERR}(\bar{q}_1, \ldots, \bar{q}_N, a_1, a_2)| \leq C Q^2 (\Delta t_{\mathrm{MAX}})^{1/4}. \tag{3.105}$$

If (3.100) holds, then

$$|\bar{\xi}_1^{\sigma}(T)|, |\bar{\xi}_2^{\sigma}(T)| \leq C Q^2 \tag{3.106}$$

(see (3.29)).

Letting $\mathcal{E}$ denote the event that $(q^\sigma(t_1), \ldots, q^\sigma(t_N))$ satisfies (3.100), we conclude from (3.103), (3.105) and (3.106) that

$$\begin{aligned}&\mathrm{E}_{a_2,\vec{\eta}}\,[f(q^\sigma(t_1), \ldots, q^\sigma(t_N)) \cdot \mathbb{1}_{\mathcal{E}}] \\ &\leq \mathrm{E}_{a_1,\vec{\eta}}\,[f(q^\sigma(t_1), \ldots, q^\sigma(t_N)) \cdot \mathbb{1}_{\mathcal{E}}] \cdot \exp(C Q^2 |a_2 - a_1|)\,(1 + \mathrm{ERR}_f(a_1, a_2)),\end{aligned} \tag{3.107}$$

with

$$|\mathrm{ERR}_f(a_1, a_2)| \leq C Q^2 (\Delta t_{\mathrm{MAX}})^{1/4}. \tag{3.108}$$

On the other hand, Lemma 3.2 shows that the complement of $\mathcal{E}$, denoted ${}^c\mathcal{E}$, satisfies

$$\mathrm{Prob}_{a_2,\vec{\eta}}\,[{}^c\mathcal{E}] \leq \exp(-c\,Q^2),$$

and therefore, by Cauchy–Schwarz, we have

$$\begin{aligned}&\mathrm{E}_{a_2,\vec{\eta}}\,[f(q^\sigma(t_1), \ldots, q^\sigma(t_N)) \cdot \mathbb{1}_{{}^c\mathcal{E}}] \\ &\qquad\leq \exp(-c' Q^2) \left(\mathrm{E}_{a_2,\vec{\eta}}\,[f^2(q^\sigma(t_1), \ldots, q^\sigma(t_N))]\right)^{1/2}.\end{aligned} \tag{3.109}$$

Combining (3.108) and (3.109), we obtain the following result.

Lemma 3.6 (Lemma on change of assumption). *Suppose $Q \geq C$ for a large enough constant C. Let $a_1, a_2 \in [-a_{\text{MAX}}, +a_{\text{MAX}}]$, let $\vec{\eta} \in \{0,1\}^{\mathbb{N}}$, and let $f(\bar{q}_1, \ldots, \bar{q}_N)$ be a nonnegative function on $\mathbb{R}^{\mathbb{N}}$. Let σ be a tame strategy. Assume that $Q \leq (\Delta t_{\text{MAX}})^{-1/1000}$. Then*

$$\begin{aligned} &\mathrm{E}_{a_2,\vec{\eta}}[f(q^\sigma(t_1), \ldots, q^\sigma(t_N))] \\ &\quad \leq \exp(CQ^2|a_2 - a_1|)(1 + CQ^2(\Delta t_{\text{MAX}})^{1/4})\,\mathrm{E}_{a_1,\vec{\eta}}[f(q^\sigma(t_1), \ldots, q^\sigma(t_N))] \\ &\qquad + \exp(-cQ^2)\Big(\mathrm{E}_{a_2,\vec{\eta}}[f^2(q^\sigma(t_1), \ldots, q^\sigma(t_N))]\Big)^{1/2}. \end{aligned}$$

3.7. Disasters due to undercontrol

Our tame strategies σ are defined to guarantee that

$$|u^\sigma(t_\nu)| \leq C[|q^\sigma(t_\nu)| + 1], \tag{3.110}$$

which is reasonable, given our assumption that $a_{\text{TRUE}} \in [-a_{\text{MAX}}, a_{\text{MAX}}]$. However, if $a_{\text{TRUE}} \gg a_{\text{MAX}}$, then we expect (3.110) to undercontrol, leading to exponentially large expected cost.

The following lemma confirms that intuition.

Lemma 3.7 (Lemma on undercontrol). *Let σ be a deterministic strategy satisfying (3.110), and suppose $a_{\text{TRUE}} = a$, where a exceeds a large enough constant C_*. Write $\mathrm{E}_a[\ldots]$ for the corresponding expectation. Assume that Δt_{MAX} is less than a small enough positive number determined by a and T. Then*

$$\mathrm{E}_a\Big[\sum_{0 \leq \nu < N} \{q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2\}\,\Delta t_\nu\Big] \geq cT^2 \exp(caT).$$

Proof. We write q_ν for $q^\sigma(t_\nu)$, Δq_ν for $q_{\nu+1} - q_\nu$, u_ν for $u^\sigma(t_\nu)$, and Δt_ν for $t_{\nu+1} - t_\nu$. We let $\mathcal{F}_\nu$ denote the sigma algebra of events determined by $q_1, \ldots, q_\nu$. Thus, q_ν and u_ν are deterministic once we condition on $\mathcal{F}_\nu$.

Recall that

$$\Delta q_\nu = (aq_\nu + u_\nu)(\Delta t_\nu^*) + \Delta W_\nu, \tag{3.111}$$

where ΔW_ν is normal with mean 0 and variance

$$\Delta \tilde{t}_\nu = \frac{\exp(2a\Delta t_\nu) - 1}{2a};$$

moreover, ΔW_ν is independent of $\mathcal{F}_\nu$. Here,

$$\Delta t_\nu^* = \frac{\exp(a\Delta t_\nu) - 1}{a}.$$

Since $\Delta t_\nu < \Delta t_{\text{MAX}}$ is less than a small enough positive number determined by a and T, we have

$$|\Delta\tilde{t}_\nu - \Delta t_\nu|, |\Delta t_\nu^* - \Delta t_\nu| < 10^{-3}\Delta t_\nu. \tag{3.112}$$

From (3.111) we have

$$\begin{aligned} q_{\nu+1}^2 &= q_\nu^2 + 2q_\nu[(aq_\nu + u_\nu)(\Delta t_\nu^*) + \Delta W_\nu] + (aq_\nu + u_\nu)^2(\Delta t_\nu^*)^2 \\ &\quad + 2(aq_\nu + u_\nu)(\Delta t_\nu^*)\Delta W_\nu + (\Delta W_\nu)^2 \\ &\geq q_\nu^2(1 + 2a\Delta t_\nu^*) + 2q_\nu u_\nu(\Delta t_\nu^*) + [2q_\nu + 2(aq_\nu + u_\nu)(\Delta t_\nu^*)]\Delta W_\nu + (\Delta W_\nu)^2. \end{aligned}$$

Consequently,

$$\mathrm{E}_a[q_{\nu+1}^2 \,|\, \mathcal{F}_\nu] \geq q_\nu^2(1 + 2a\Delta t_\nu^*) + 2q_\nu u_\nu(\Delta t_\nu^*) + (\Delta\tilde{t}_\nu). \tag{3.113}$$

For $\delta > 0$ to be picked in a moment, we have

$$|2q_\nu u_\nu| \leq \delta^{-2} q_\nu^2 + \delta^2 u_\nu^2 \leq C\delta^{-2} q_\nu^2 + C\delta^2,$$

thanks to (3.110).

Putting this inequality into (3.113), we find that

$$\mathrm{E}_a[q_{\nu+1}^2 \,|\, \mathcal{F}_\nu] \geq q_\nu^2(1 + [2a - C\delta^{-2}]\Delta t_\nu^*) + (\Delta\tilde{t}_\nu - C\delta^2\Delta t_\nu^*). \tag{3.114}$$

We take δ to be a small enough constant c such that

$$\Delta\tilde{t}_\nu - C\delta^2\Delta t_\nu^* \geq \frac{1}{2}\Delta t_\nu;$$

see (3.112). Since a exceeds a large enough constant C, we then have

$$[2a - C\delta^{-2}] > a,$$

so from (3.114) we obtain

$$\mathrm{E}_a[q_{\nu+1}^2 \,|\, \mathcal{F}_\nu] \geq q_\nu^2(1 + a\Delta t_\nu^*) + \frac{1}{2}\Delta t_\nu.$$

Again applying (3.112), and recalling that $\Delta t_\nu < \Delta t_{\mathrm{MAX}}$ is less than a small enough positive number determined by a and T, we conclude that

$$\mathrm{E}_a[q_{\nu+1}^2 \,|\, \mathcal{F}_\nu] \geq \exp\left(\frac{1}{2}a\Delta t_\nu\right)q_\nu^2 + \frac{1}{2}\Delta t_\nu,$$

and therefore

$$\mathrm{E}_a[q_{\nu+1}^2] \geq \exp\left(\frac{1}{2}a\Delta t_\nu\right)\mathrm{E}_a[q_\nu^2] + \frac{1}{2}\Delta t_\nu. \tag{3.115}$$

Since (3.115) implies that

$$\mathrm{E}_a[q_{\nu+1}^2] \geq \mathrm{E}_a[q_\nu^2] + \frac{1}{2}\Delta t_\nu \text{ for each } \nu,$$

we conclude that

$$\mathrm{E}_a[q_\nu^2] \geq \frac{1}{2}t_\nu \quad \text{for each } \nu.$$

We pick ν_0 so that

$$\frac{1}{2}T < t_{\nu_0} < \frac{2}{3}T.$$

(Our smallness assumption on Δt_{MAX} implies that such a ν_0 exists.) Then

$$\mathrm{E}_a[q_{\nu_0}^2] \geq \frac{1}{2}t_{\nu_0} > \frac{1}{4}T.$$

Returning to (3.115), we have

$$\mathrm{E}_a[q_{\nu+1}^2] \geq \exp\Big(\frac{1}{2}a\Delta t_\nu\Big)\mathrm{E}_a[q_\nu^2] \quad \text{for each } \nu,$$

hence for $\nu \geq \nu_0$ we have

$$\mathrm{E}_a[q_\nu^2] \geq \exp\Big(\frac{1}{2}a[t_\nu - t_{\nu_0}]\Big)\,\mathrm{E}[q_{\nu_0}^2] \geq \frac{1}{4}T\exp\Big(\frac{1}{2}a[t_\nu - t_{\nu_0}]\Big).$$

In particular,

$$\mathrm{E}_a[q_\nu^2] \geq \frac{1}{4}T\exp(caT) \quad \text{for } t_\nu \in \Big[\frac{3}{4}T, T\Big].$$

Consequently,

$$\begin{aligned}\mathrm{E}_a\Big[\sum_{0\leq\nu<N}\{q_\nu^2 + u_\nu^2\}\,\Delta t_\nu\Big] &\geq \mathrm{E}_a\Big[\sum_{t_\nu\in[\frac{3}{4}T,T]} q_\nu^2\,\Delta t_\nu\Big] \geq \frac{1}{4}T\exp(caT)\cdot\sum_{t_\nu\in[\frac{3}{4}T,T]}\Delta t_\nu \\ &\geq cT^2\exp(caT),\end{aligned}$$

since each $\Delta t_\nu < \Delta t_{\mathrm{MAX}}$ is less than a small enough positive number determined by a and T. The proof of the lemma is complete. ∎

3.8. Costing by integrals

Let σ be a deterministic tame strategy. We condition on

$$a_{\mathrm{TRUE}} = a \in [-a_{\mathrm{MAX}}, +a_{\mathrm{MAX}}],$$

and write $\mathrm{Prob}_a[\cdot]$ and $E_a[\cdot]$ to denote the corresponding probability and expectation.

We want to compare

$$\mathrm{COST}(\sigma) = \int_0^T \{(q^\sigma(t))^2 + (u^\sigma(t))^2\}\,dt$$

with

$$\mathrm{COST}_D(\sigma) = \sum_{0\leq\nu<N}\{(q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2\}\,\Delta t_\nu$$

("D" for "discrete").

Lemma 3.8 (Lemma on costing by integrals). *For any $m \geq 1$, we have*

$$\mathrm{E}_a[|\mathrm{COST}(\sigma) - \mathrm{COST}_D(\sigma)|^m] \leq C_m(\Delta t_{\mathrm{MAX}})^{m/2}.$$

Proof. Recall that $u^\sigma(t) = u^\sigma(t_\nu)$ for $t \in [t_\nu, t_{\nu+1}]$. Hence,

$$\begin{aligned}\mathrm{COST}(\sigma) - \mathrm{COST}_D(\sigma) &= \sum_{0\leq\nu<N}\int_{t_\nu}^{t_{\nu+1}}\{(q^\sigma(t))^2 - (q^\sigma(t_\nu))^2\}\,dt \\ &= \sum_{0\leq\nu<N}\int_{t_\nu}^{t_{\nu+1}}\big\{[q^\sigma(t) - q^\sigma(t_\nu)]^2 + 2q^\sigma(t_\nu)[q^\sigma(t) - q^\sigma(t_\nu)]\big\}\,dt.\end{aligned}$$

Setting

$$\textsc{Osc}(\nu) = \max_{t \in [t_\nu, t_{\nu+1}]} |q^\sigma(t) - q^\sigma(t_\nu)|,$$

we therefore have

$$|\textsc{Cost}(\sigma) - \textsc{Cost}_D(\sigma)| \leq \sum_{0 \leq \nu < N} \{(\textsc{Osc}(\nu))^2 + 2|q^\sigma(t_\nu)|\,(\textsc{Osc}(\nu))\}\,\Delta t_\nu,$$

so that by Hölder's inequality,

$$|\textsc{Cost}(\sigma) - \textsc{Cost}_D(\sigma)|^m \leq C_m \sum_{0 \leq \nu < N} \{(\textsc{Osc}(\nu))^{2m} + |q^\sigma(t_\nu)|^m\,(\textsc{Osc}(\nu))^m\}\,\Delta t_\nu.$$

Consequently,

$$\begin{aligned} &\mathrm{E}_a[|\textsc{Cost}(\sigma) - \textsc{Cost}_D(\sigma)|^m] \\ &\qquad \leq C_m \sum_{0 \leq \nu < N} \mathrm{E}_a[(\textsc{Osc}(\nu)^{2m}]\,(\Delta t_\nu) \\ &\qquad + C_m \sum_{0 \leq \nu < N} (\mathrm{E}_a[|q^\sigma(t_\nu)|^{2m}])^{1/2}\,(\mathrm{E}_a[(\textsc{Osc}(\nu))^{2m}])^{1/2}\Delta t_\nu. \end{aligned} \tag{3.116}$$

We now estimate the right-hand side of (3.116).

Fix a large enough constant C_*, and let $\mathcal{F}_\nu$ denote the sigma algebra of events determined by $q^\sigma(t_\mu)$ for $\mu = 1, \ldots, \nu$.

Lemma 3.2 shows that

$$\mathrm{E}_a[|q^\sigma(t_\nu)|^{2m}] \leq C_m \tag{3.117}$$

and that, given $Q_2 \geq Q_1 \geq C_*$, we have

$$\mathrm{Prob}_a[\textsc{Osc}(\nu) > Q_2(\Delta t_\nu)^{1/2}\,|\,\mathcal{F}_\nu] \leq C\exp(-cQ_2^2) \quad \text{if } |q^\sigma(t_\nu)| \leq Q_1. \tag{3.118}$$

From (3.118), we see that

$$\mathrm{E}_a[(\textsc{Osc}(\nu))^{2m}\,|\,\mathcal{F}_\nu] \leq C_m Q_1^{2m}(\Delta t_\nu)^m \quad \text{if } |q^\sigma(t_\nu)| \leq Q_1.$$

In particular,

$$\mathrm{E}_a[(\textsc{Osc}(\nu))^{2m} \cdot \mathbb{1}_{|q^\sigma(t_\nu)| \leq C_*}] \leq C_m(\Delta t_\nu)^m \tag{3.119}$$

and, for $k \geq 0$,

$$\begin{aligned} &\mathrm{E}_a[(\textsc{Osc}(\nu))^{2m} \cdot \mathbb{1}_{|q^\sigma(t_\nu)| \in [C_* 2^k, C_* 2^{(k+1)})}] \\ &\qquad \leq C_m \cdot (C_* 2^{(k+1)})^{2m}(\Delta t_\nu)^m \cdot \mathrm{Prob}_a[|q^\sigma(t_\nu)| > C_* 2^k]. \end{aligned} \tag{3.120}$$

Another application of Lemma 3.2 gives

$$\mathrm{Prob}_a[|q^\sigma(t_\nu)| > C_* 2^k] \leq C\exp(-c2^{2k}),$$

so that (3.120) implies that

$$\mathrm{E}_a[(\mathrm{OSC}(\nu))^{2m} \cdot \mathbb{1}_{|q^\sigma(t_\nu)| \in [C_* 2^k, C_* 2^{(k+1)})}] \le C_m \cdot (2^{2mk})(\Delta t_\nu)^m \cdot \exp(-c 2^{2k}).$$

Summing over $k \ge 0$, and combining the result with (3.119), we learn that

$$\mathrm{E}_a[(\mathrm{OSC}(\nu))^{2m}] \le C_m (\Delta t_\nu)^m. \tag{3.121}$$

From (3.116), (3.117), and (3.121) we conclude that

$$\mathrm{E}_a[|\mathrm{COST}(\sigma) - \mathrm{COST}_D(\sigma)|^m] \le C_m \sum_{0 \le \nu < N} (\Delta t_\nu)^{m/2+1} \le C_m' (\Delta t_{\mathrm{MAX}})^{m/2},$$

which is the conclusion of the lemma. ■

3.9. Continuous vs discrete

Let σ be a tame strategy, possibly depending on coin flips $\vec{\xi}$. In this section, we compare the following random functions of time:

$$\begin{aligned}
q_C^\sigma(t) &= q^\sigma(t) \quad \text{for } t \in [0, T],\\
q_D^\sigma(t) &= q^\sigma(t_\nu) \quad \text{for } t \in [t_\nu, t_{\nu+1}),\ \text{each } \nu < N,\\
\zeta_{1,C}^\sigma(t) &= \frac{1}{2}(q^\sigma(t))^2 - \frac{1}{2} q_0^2 - \frac{1}{2} t - \int_0^t u^\sigma(s) q^\sigma(s)\, ds \quad \text{for } t \in [0, T],\\
\zeta_{1,D}^\sigma(t) &= \frac{1}{2}(q^\sigma(t_\nu))^2 - \frac{1}{2} q_0^2 - \frac{1}{2} \sum_{0 \le \mu < \nu} (\Delta q^\sigma(t_\mu))^2 - \sum_{0 \le \mu < \nu} u^\sigma(t_\mu) q^\sigma(t_\mu)\, \Delta t_\mu\\
&\qquad \text{for } t \in [t_\nu, t_{\nu+1}),\ \text{each } \nu < N,\\
\zeta_{2,C}^\sigma(t) &= \int_0^t (q^\sigma(s))^2\, ds \quad \text{for } t \in [0, T],\\
\zeta_{2,D}^\sigma(t) &= \sum_{0 \le \mu < \nu} (q^\sigma(t_\mu))^2\, \Delta t_\mu \quad \text{for } t \in [t_\nu, t_{\nu+1}),\ \text{each } \nu < N.
\end{aligned}$$

We recall that $u^\sigma(t)$ is constant on $[t_\nu, t_{\nu+1})$ for each ν, so there is no need to introduce analogous quantities for u^σ.

We establish the following result.

Lemma 3.9 (Lemma on continuous variants). *Let Q be greater than a large enough constant C. Then there exists an event $\mathrm{BAD}(\sigma, Q)$ with the following properties.*

- *For each $a \in [-a_{\mathrm{MAX}}, +a_{\mathrm{MAX}}]$ and $\vec{\eta} \in \{0,1\}^{\mathbb{N}}$, we have*

$$\mathrm{Prob}_{a,\vec{\eta}}[\mathrm{BAD}(\sigma, Q)] \le C \exp(-c Q^2).$$

- *If $\mathrm{BAD}(\sigma, Q)$ does not occur, then*

$$\begin{aligned}
\max_{t \in [0,T)} |q_C^\sigma(t) - q_D^\sigma(t)| &\le C Q (\Delta t_{\mathrm{MAX}})^{1/4}\\
\max_{t \in [0,T)} |\zeta_{1,C}^\sigma(t) - \zeta_{1,D}^\sigma(t)| &\le C Q^2 (\Delta t_{\mathrm{MAX}})^{1/4},\\
\max_{t \in [0,T)} |\zeta_{2,C}^\sigma(t) - \zeta_{2,D}^\sigma(t)| &\le C Q^2 (\Delta t_{\mathrm{MAX}})^{1/4}.
\end{aligned}$$

Proof. From Lemma 3.2, the following have probability at most $C\exp(-cQ^2)$ when conditioned on $a_{\text{TRUE}} = a$, $\vec{\xi} = \vec{\eta}$ (any $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$, $\vec{\eta} \in \{0,1\}^{\mathbb{N}}$):

$$\max_{\nu}\left|\sum_{0\le\mu<\nu}(\Delta q_\mu^\sigma)^2 - t_\nu\right| > Q^2(\Delta t_{\text{MAX}})^{1/2}, \tag{3.122}$$

$$\max_{\mu}|q_\mu^\sigma| > Q, \tag{3.123}$$

$$\max_{\mu}|u_\mu^\sigma| > Q. \tag{3.124}$$

Moreover, with

$$\text{Osc}(\nu) := \max_{t\in[t_\nu,t_{\nu+1}]}|q^\sigma(t) - q^\sigma(t_\nu)|$$

and $\mathcal{F}_\nu^\sigma$ defined as the sigma algebra of events determined by the $q^\sigma(t_\mu)$ for $\mu \le \nu$, Lemma 3.2 gives also the following, for each fixed ν:

$$\text{Prob}_{a,\vec{\eta}}[\text{Osc}(\nu) > CQ(\Delta t_\nu)^{1/2}\,|\,\mathcal{F}_\nu] \le C\exp(-cQ^2) \quad \text{if } |q_\nu| \le Q.$$

This implies that

$$\text{Prob}_{a,\vec{\eta}}[\text{Osc}(\nu) > CQ(\Delta t_\nu)^{1/2} \quad \text{and} \quad |q_\nu| \le Q] \le C\exp(-cQ^2).$$

Since also

$$\text{Prob}_{a,\vec{\eta}}[|q_\nu| > Q] \le C\exp(-cQ^2),$$

it follows that

$$\text{Prob}_{a,\vec{\eta}}[\text{Osc}(\nu) > CQ(\Delta t_\nu)^{1/2}] \le C\exp(-cQ^2)$$

for each fixed ν, and for all $Q \ge C$.

Taking $Q(\Delta t_\nu)^{-1/4}$ in place of Q here, we find that

$$\text{Prob}_{a,\vec{\eta}}[\text{Osc}(\nu) > CQ(\Delta t_\nu)^{1/4}] \le C\exp(-cQ^2(\Delta t_\nu)^{-1/2}) \le C'\exp(-cQ^2)(\Delta t_\nu).$$

Summing over ν, we find that the event

$$\text{Osc}(\nu) > CQ(\Delta t_\nu)^{1/4} \quad \text{for some } \nu, \tag{3.125}$$

conditioned on $a_{\text{TRUE}} = a$ and $\vec{\xi} = \vec{\eta}$, has probability at most $C\exp(-cQ^2)$.

We now define $\text{BAD}(Q,\sigma)$ to be the event that at least one of the conditions (3.122)–(3.125) holds. Then, as claimed,

$$\text{Prob}_{a,\vec{\eta}}[\text{BAD}(Q,\sigma)] \le C\exp(-cQ^2)$$

for any $a \in [-a_{\text{MAX}}, +a_{\text{MAX}}]$ and any $\vec{\eta} \in \{0,1\}^{\mathbb{N}}$.

We now suppose that $\text{BAD}(Q,\sigma)$ does not occur, and compare $q_C^\sigma(t)$ with $q_D^\sigma(t)$, $\zeta_{1,C}^\sigma(t)$ with $\zeta_{1,D}^\sigma(t)$, and $\zeta_{2,C}^\sigma(t)$ with $\zeta_{2,D}^\sigma(t)$.

Since $\text{BAD}(Q,\sigma)$ does not occur, we have

$$\max_{\nu}|q^\sigma(t_\nu)|, \max_{\nu}|u^\sigma(t_\nu)| \le Q, \tag{3.126}$$

$$\max_{\nu}\left|\sum_{0\le\mu<\nu}(\Delta q_\nu^\sigma)^2 - t_\nu\right| \le Q^2(\Delta t_{\text{MAX}})^{1/2}, \tag{3.127}$$

$$\max_{\nu}\text{Osc}(\nu) \le CQ(\Delta t_{\text{MAX}})^{1/4}. \tag{3.128}$$

For any ν, and any $t \in [t_\nu, t_{\nu+1})$, we have

$$|q^\sigma_C(t) - q^\sigma_D(t)| = |q^\sigma(t) - q^\sigma(t_\nu)| \le \mathrm{OSC}(\nu) \le CQ(\Delta t_{\mathrm{MAX}})^{1/4}.$$

Thus,

$$\max_{t\in[0,T]} |q^\sigma_C(t) - q^\sigma_D(t)| \le CQ(\Delta t_{\mathrm{MAX}})^{1/4},$$

as claimed. Next, for any ν and any $t \in [t_\nu, t_{\nu+1})$, we have

$$\begin{aligned} \zeta^\sigma_{1,C}(t) - \zeta^\sigma_{1,D}(t) &= \frac{1}{2}[(q^\sigma(t))^2 - (q^\sigma(t_\nu))^2] + \frac{1}{2}\Big[\sum_{0\le\mu<\nu} (\Delta q^\sigma_\mu)^2 - t_\nu\Big] \\ &\quad - \frac{1}{2}(t - t_\nu) - \sum_{0\le\mu<\nu} \int_{t_\mu}^{t_{\mu+1}} \{u^\sigma(t_\mu)[q^\sigma(s) - q^\sigma(t_\mu)]\}\, ds \\ &\quad - \int_{t_\nu}^{t} u^\sigma(t_\nu) q^\sigma(s)\, ds \\ &\equiv \mathrm{TERM}\,1 + \mathrm{TERM}\,2 - \mathrm{TERM}\,3 - \mathrm{TERM}\,4 - \mathrm{TERM}\,5. \end{aligned}$$

(Here, we have used the fact that $u^\sigma(s) = u^\sigma(t_\mu)$ for $s \in [t_\mu, t_{\mu+1})$.) We note that

$$\max_{t\in[0,T]} |q^\sigma(t)| \le \max_{0\le\nu<N} \{|q^\sigma(t_\nu)| + \mathrm{OSC}(\nu)\} \le CQ, \tag{3.129}$$

by (3.126) and (3.128). Hence, for $t \in [t_\nu, t_{\nu+1})$, we have

$$|\mathrm{TERM}\,1| \le \max_{\tilde{t}\in[0,T]} |q^\sigma(\tilde{t})| \cdot |q^\sigma(t) - q^\sigma(t_\nu)| \le CQ\,\mathrm{OSC}(\nu) \le CQ^2(\Delta t_{\mathrm{MAX}})^{1/4}.$$

So

$$|\mathrm{TERM}\,1| \le CQ^2(\Delta t_{\mathrm{MAX}})^{1/4} \quad \text{for all } t \in [0, T).$$

Next, (3.127) tells us that

$$|\mathrm{TERM}\,2| \le Q^2(\Delta t_{\mathrm{MAX}})^{1/2} \quad \text{for all } t \in [0, T).$$

Clearly

$$|\mathrm{TERM}\,3| \le (\Delta t_{\mathrm{MAX}}).$$

Furthermore,

$$\begin{aligned} |\mathrm{TERM}\,4| &\le \sum_{0\le\mu<N} |u^\sigma(t_\mu)| \int_{t_\mu}^{t_{\mu+1}} |q^\sigma(s) - q^\sigma(t_\mu)|\, ds \\ &\le CQ \sum_{0\le\mu<N} \mathrm{OSC}(\mu)\Delta t_\mu \le CQ \cdot \max_\mu \mathrm{OSC}(\mu) \le CQ^2(\Delta t_{\mathrm{MAX}})^{1/4} \end{aligned}$$

thanks to (3.126) and (3.128). Finally,

$$|\mathrm{TERM}\,5| \le \max_\nu |u^\sigma(t_\nu)| \cdot \max_{\tilde{t}\in[0,T]} |q^\sigma(\tilde{t})| \cdot (\Delta t_{\mathrm{MAX}}) \le CQ^2(\Delta t_{\mathrm{MAX}}),$$

by (3.126) and (3.129). Combining our estimates for TERMS 1–5, we find that

$$|\zeta^{\sigma}_{1,C}(t) - \zeta^{\sigma}_{1,D}(t)| \le CQ^2(\Delta t_{\text{MAX}})^{1/4},$$

as claimed. We pass to $\zeta^{\sigma}_{2,C}$ and $\zeta^{\sigma}_{2,D}$. For $t \in [t_\nu, t_{\nu+1})$, we have

$$\begin{aligned}
|\zeta^{\sigma}_{2,C}(t) - \zeta^{\sigma}_{2,D}(t)| &= \Big| \sum_{0\le\mu<\nu} \int_{t_\mu}^{t_{\mu+1}} \{(q^\sigma(s))^2 - (q^\sigma(t_\mu))^2\}\, ds + \int_{t_\nu}^{t} (q^\sigma(s))^2\, ds \Big| \\
&\le \sum_{0\le\mu<\nu} C\Big(\max_{\tilde t\in[0,T]} |q^\sigma(\tilde t)|\Big) \int_{t_\mu}^{t_{\mu+1}} |q^\sigma(s) - q^\sigma(t_\mu)|\, ds + \max_{\tilde t\in[0,T]} |q^\sigma(\tilde t)|^2 \cdot \Delta t_{\text{MAX}} \\
&\le CQ \sum_{0\le\mu<\nu} \text{OSC}(\mu)\Delta t_\mu + CQ^2\Delta t_{\text{MAX}} \le C'Q \max_{0\le\mu<N} \text{OSC}(\mu) + CQ^2 \Delta t_{\text{MAX}} \\
&\le C''Q^2(\Delta t_{\text{MAX}})^{1/4}, \quad \text{thanks to (3.128) and (3.129).}
\end{aligned}$$

The proof of the lemma is complete. ∎

3.10. Refining a partition

Let σ be a tame strategy $\sigma = (\sigma_{t_\nu})_{0\le\nu<N}$ associated to a partition

$$0 = t_0 < t_1 < \cdots < t_N = T. \tag{3.130}$$

Suppose

$$0 = \hat t_0 < \hat t_1 < \cdots < \hat t_{\hat N} = T \tag{3.131}$$

is a refinement of the partition (3.130).

We would like to associate a tame strategy $\hat\sigma = (\hat\sigma_{\hat t_\mu})_{0\le\mu<\hat N}$ to the partition (3.131) in such a way that

$$\begin{aligned} q^{\hat\sigma}(t) &= q^\sigma(t) \quad \text{for all } t\in[0,T], \text{ and} \\ u^{\hat\sigma}(t) &= u^\sigma(t) \quad \text{for all } t\in[0,T). \end{aligned} \tag{3.132}$$

That would tell us that refining the partition (3.130) allows additional tame strategies, but does not rule out any tame strategies σ.

Unfortunately, no such $\hat\sigma$ exists. The problem is that, in order to be a tame strategy, $\hat\sigma$ must satisfy

$$|u^{\hat\sigma}(\hat t_\mu)| \le C_{\text{TAME}}[|q^{\hat\sigma}(\hat t_\mu)| + 1] \quad \text{with probability 1}. \tag{3.133}$$

It may happen that

$$|u^\sigma(\hat t_\mu)| = |u^\sigma(t_\nu)| \gg C_{\text{TAME}} \cdot [|q^\sigma(\hat t_\mu)| + 1]$$

for some $\hat t_\mu \in (t_\nu, t_{\nu+1})$, in which case (3.133) contradicts (3.132). Accordingly, we modify (3.132), as follows.

For each μ $(0 \le \mu < \hat N)$, define $\bar\nu(\mu)$ to be the index ν for which $t_\nu \le \hat t_\mu < t_{\nu+1}$.

Define a stopping time τ by setting

$$\tau = \begin{cases} \text{least } \hat{t}_\mu \text{ such that } |u^\sigma(t_{\bar{\nu}(\mu)})| > 2C_{\text{TAME}}[|q^\sigma(\hat{t}_\mu)| + 1], & \text{if such a } \hat{t}_\mu \text{ exists,} \\ T, & \text{otherwise.} \end{cases}$$

Then define random processes $\hat{q}(t)$, $\hat{u}(t)$ ($t \in [0, T]$) by setting

$$\begin{cases} \hat{q}(t) = q^\sigma(t) \text{ and } \hat{u}(t) = u^\sigma(t), & \text{for } 0 \le t < \tau, \\ \hat{u}(t) = 0, & \text{for } \tau \le t \le T, \end{cases}$$

and on $[\tau, T]$ defining $\hat{q}$ by

$$\begin{cases} d\hat{q}(t) = (a_{\text{TRUE}} \cdot \hat{q}(t))\, dt + dW(t), \\ \text{with initial condition } \hat{q}(\tau) = q^\sigma(\tau). \end{cases}$$

Note that

$$|\hat{u}(\hat{t}_\mu)| \le 2C_{\text{TAME}}[|\hat{q}(\hat{t}_\mu)| + 1] \quad \text{for all } \hat{t}_\mu. \tag{3.134}$$

It is a tedious exercise (left to the reader) to exhibit a tame strategy $\hat{\sigma} = (\hat{\sigma}_{\hat{t}_\mu})_{0 \le \mu < \hat{N}}$ associated to the partition (3.131), such that $\hat{q}(t) = q^{\hat{\sigma}}(t)$ for all $t \in [0, T]$ and $\hat{u}(t) = u^{\hat{\sigma}}(t)$ for all $t \in [0, T)$.

Whereas our original tame strategy σ satisfies

$$|\sigma_{t_\nu}(q_1, \ldots, q_\nu, \vec{\xi})| \le C_{\text{TAME}}[|q_\nu| + 1], \tag{3.135}$$

the strategy $\hat{\sigma}$ satisfies instead

$$|\hat{\sigma}_{\hat{t}_\mu}(q_1, \ldots, q_\mu, \vec{\xi})| \le 2C_{\text{TAME}}[|q_\mu| + 1]; \tag{3.136}$$

compare with (3.134).

In place of (3.132), we will show that $q^{\hat{\sigma}}(t)$ and $u^{\hat{\sigma}}(t)$ are likely very close to $q^\sigma(t)$ and $u^\sigma(t)$, respectively. To see this, we fix $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$ and $\vec{\eta} \in \{0, 1\}^{\mathbb{N}}$, and condition on $a_{\text{TRUE}} = a$ and $\vec{\xi} = \vec{\eta}$.

We define random variables

$$\text{OSC}(\nu) = \max_{t \in [t_\nu, t_{\nu+1}]} |q^\sigma(t) - q^\sigma(t_\nu)| \tag{3.137}$$

and an event

$$\text{DISASTER}: \ \text{OSC}(\nu) \ge 1 \ \text{ for some } \nu. \tag{3.138}$$

In Section 3.8, we proved that

$$\mathrm{E}_{a,\vec{\eta}}[(\text{OSC}(\nu))^m] \le C_m (\Delta t_\nu)^{m/2} \quad \text{for all } m \ge 1. \tag{3.139}$$

Hence,

$$\text{Prob}_{a,\vec{\eta}}[\text{OSC}(\nu) \ge 1] \le C_{\bar{m}} (\Delta t_\nu)^{\bar{m}} \quad \text{for all } \bar{m} \ge 1,$$

and consequently

$$\text{Prob}_{a,\vec{\eta}}[\text{DISASTER}] \leq \sum_{0\leq\nu<N} C_{\bar{m}}(\Delta t_\nu)^{\bar{m}} \leq C'_{\bar{m}-1}(\Delta t_{\text{MAX}})^{\bar{m}-1} \tag{3.140}$$

for any $\bar{m} \geq 1$.

Next, we prepare to estimate

$$\int_0^T \{|q^{\hat{\sigma}}(t) - q^\sigma(t)|^m + |u^{\hat{\sigma}}(t) - u^\sigma(t)|^m\}\, dt.$$

Let

$$q_0^\sigma(t) = q^\sigma(t_\nu) \quad \text{for } t \in [t_\nu, t_{\nu+1}),\ 0 \leq \nu < N.$$

Then

$$\int_{t_\nu}^{t_{\nu+1}} |q_0^\sigma(t) - q^\sigma(t)|^{2m}\, dt \leq (\text{OSC}(\nu))^{2m}\Delta t_\nu,$$

hence

$$\begin{aligned} \text{E}_{a,\vec{\eta}}\Big[\int_0^T |q_0^\sigma(t) - q^\sigma(t)|^{2m}\, dt\Big] &\leq \sum_\nu \text{E}_{a,\vec{\eta}}[(\text{OSC}(\nu))^{2m}]\Delta t_\nu \\ &\leq C_m(\Delta t_{\text{MAX}})^m, \end{aligned} \tag{3.141}$$

thanks to (3.139). Similarly,

$$\text{E}_{a,\vec{\eta}}\Big[\int_0^T |q_0^{\hat{\sigma}}(t) - q^{\hat{\sigma}}(t)|^{2m}\, dt\Big] \leq C_m(\Delta t_{\text{MAX}})^m, \tag{3.142}$$

where $q_0^{\hat{\sigma}}(t) = q^{\hat{\sigma}}(\hat{t}_\mu)$ for $t \in [\hat{t}_\mu, \hat{t}_{\mu+1})$, $0 \leq \mu < \hat{N}$. We have also

$$\text{E}_{a,\vec{\eta}}\Big[\int_0^T |q_0^\sigma(t)|^{2m}\, dt\Big] = \text{E}_{a,\vec{\eta}}\Big[\sum_{0\leq\nu<N} |q^\sigma(t_\nu)|^{2m}\Delta t_\nu\Big] \leq C_m, \tag{3.143}$$

thanks to Lemma 3.2. Similarly,

$$\text{E}_{a,\vec{\eta}}\Big[\int_0^T |q_0^{\hat{\sigma}}(t)|^{2m}\, dt\Big] \leq C_m. \tag{3.144}$$

From (3.141) and (3.143), we obtain

$$\text{E}_{a,\vec{\eta}}\Big[\int_0^T |q^\sigma(t)|^{2m}\, dt\Big] \leq C_m \quad \text{for } m \geq 1. \tag{3.145}$$

Similarly, from (3.142) and (3.144), we have

$$\text{E}_{a,\vec{\eta}}\Big[\int_0^T |q^{\hat{\sigma}}(t)|^{2m}\, dt\Big] \leq C_m \quad \text{for } m \geq 1. \tag{3.146}$$

Turning to u^σ and $u^{\hat{\sigma}}$, we recall that $u^\sigma(t) = u^\sigma(t_\nu)$ for $t \in [t_\nu, t_{\nu+1})$, $0 \leq \nu < N$; hence,

$$\text{E}_{a,\vec{\eta}}\Big[\int_0^T |u^\sigma(t)|^{2m}\, dt\Big] = \text{E}_{a,\vec{\eta}}\Big[\sum_{0\leq\nu<N} |u^\sigma(t_\nu)|^{2m}\Delta t_\nu\Big] \leq C_m \quad (m \geq 1), \tag{3.147}$$

by Lemma 3.2. Similarly,

$$\mathrm{E}_{a,\vec{\eta}}\Big[\int_0^T |u^{\hat{\sigma}}(t)|^{2m}\,dt\Big] \le C_m \quad (m \ge 1). \tag{3.148}$$

From (3.145)–(3.147), we see that

$$\mathrm{E}_{a,\vec{\eta}}\Big[\int_0^T \{|q^{\hat{\sigma}}(t) - q^{\sigma}(t)| + |u^{\hat{\sigma}}(t) - u^{\sigma}(t)|\}^{2m}\,dt\Big] \le C_m \quad (m \ge 1). \tag{3.149}$$

Moreover, unless DISASTER occurs, we have $\tau = T$, hence

$$q^{\hat{\sigma}}(t) = q^{\sigma}(t) \text{ and } u^{\hat{\sigma}}(t) = u^{\sigma}(t) \quad \text{for all } t \in [0, T]. \tag{3.150}$$

Indeed, if DISASTER does not occur, then for $0 \le \nu < N$ and $t \in [t_\nu, t_{\nu+1})$ we have

$$|q^{\sigma}(t) - q^{\sigma}(t_\nu)| \le 1,$$

hence $[|q^{\sigma}(t)| + 1]$ and $[|q^{\sigma}(t_\nu)| + 1]$ differ by at most a factor of 2. Since $|u^{\sigma}(t_\nu)| \le C_{\mathrm{TAME}}[|q^{\sigma}(t_\nu)| + 1]$, it follows that $|u^{\sigma}(t_\nu)| \le 2C_{\mathrm{TAME}}[|q^{\sigma}(t)| + 1]$ for $t \in [t_\nu, t_{\nu+1})$, $0 \le \nu < N$. In particular,

$$|u^{\sigma}(t_{\bar{\nu}(\mu)})| \le 2C_{\mathrm{TAME}}[|q^{\sigma}(t_\mu)| + 1] \quad \text{for all } \mu\ (0 \le \mu < \hat{N}). \tag{3.151}$$

Comparing (3.151) with the definition of τ, we see that, as claimed, $\tau = T$ unless DISASTER occurs.

Thus (3.150) holds unless DISASTER occurs.

From (3.149), (3.150) and (3.140), we now have

$$\begin{aligned} &\mathrm{E}_{a,\vec{\eta}}\Big[\int_0^T \{|q^{\hat{\sigma}}(t) - q^{\sigma}(t)| + |u^{\hat{\sigma}}(t) - u^{\sigma}(t)|\}^{m}\,dt\Big] \\ &\qquad = \mathrm{E}_{a,\vec{\eta}}\Big[\int_0^T \{|q^{\hat{\sigma}}(t) - q^{\sigma}(t)| + |u^{\hat{\sigma}}(t) - u^{\sigma}(t)|\}^{m}\,\mathbb{1}_{\mathrm{DISASTER}}\,dt\Big] \\ &\qquad \le \Big(\mathrm{E}_{a,\vec{\eta}}\Big[\int_0^T \{|q^{\hat{\sigma}}(t) - q^{\sigma}(t)| + |u^{\hat{\sigma}}(t) - u^{\sigma}(t)|\}^{2m}\,dt\Big]\Big)^{1/2} \\ &\qquad\quad \cdot (\mathrm{Prob}_{a,\vec{\eta}}[\mathrm{DISASTER}])^{1/2} \\ &\qquad \le C_{m,\bar{m}} \cdot (\Delta t_{\mathrm{MAX}})^{(\bar{m}-1)/2} \quad \text{for any } m, \bar{m} \ge 1. \end{aligned} \tag{3.152}$$

We record this result as a lemma.

Lemma 3.10 (Refinement lemma). *Let σ be a tame strategy associated to a partition*

$$0 = t_0 < t_1 < \cdots < t_N = T, \tag{A}$$

and let

$$0 = \hat{t}_0 < \hat{t}_1 < \cdots < \hat{t}_{\hat{N}} = T \tag{B}$$

be a refinement of the partition (A).

Then there exists a tame strategy $\hat{\sigma}$ associated to the partition (B)*, such that*

$$\mathrm{E}_{a,\vec{\eta}}\Big[\int_0^T \{|q^{\hat{\sigma}}(t) - q^{\sigma}(t)| + |u^{\hat{\sigma}}(t) - u^{\sigma}(t)|\}^m \, dt\Big] \leq C_{m,\bar{m}}(\Delta t_{\mathrm{MAX}})^{\bar{m}}$$

for all $m, \bar{m} \geq 1$ and all $a \in [-a_{\mathrm{MAX}}, +a_{\mathrm{MAX}}]$, $\vec{\eta} \in \{0,1\}^{\mathbb{N}}$.

The strategy $\hat{\sigma}$ satisfies the same estimates as we assumed for σ (see 3.110*), except that C_{TAME} is replaced by $2C_{\mathrm{TAME}}$.*

4. Bayesian strategies associated to partitions

4.1. Setup

In this section, we take a_{TRUE} to be governed by a known prior probability distribution $d\mathrm{Prior}(a)$, concentrated on an interval $[-a_{\mathrm{MAX}}, +a_{\mathrm{MAX}}]$.

We fix a partition

$$0 = t_0 < t_1 < \cdots < t_N = T \tag{4.1}$$

of the time interval $[0, T]$.

We fix a deterministic strategy σ for the game starting at position q_0.

We assume that our strategy σ is tame, i.e.,

$$|u^{\sigma}(t_\nu)| \leq C^{\sigma}_{\mathrm{TAME}}\,[|q^{\sigma}(t_\nu)| + 1] \tag{4.2}$$

for a constant $C^{\sigma}_{\mathrm{TAME}}$. We call $C^{\sigma}_{\mathrm{TAME}}$ the *tame constant* for σ.

We write

$$\begin{aligned} q^{\sigma}_{\nu} &= q^{\sigma}(t_\nu), & \Delta q^{\sigma}_{\nu} &= q^{\sigma}_{\nu+1} - q^{\sigma}_{\nu}, \\ \zeta^{\sigma}_{1,\nu} &= \zeta^{\sigma}_{1}(t_\nu), & \Delta \zeta^{\sigma}_{1,\nu} &= \zeta^{\sigma}_{1,\nu+1} - \zeta^{\sigma}_{1,\nu}, \\ \zeta^{\sigma}_{2,\nu} &= \zeta^{\sigma}_{2}(t_\nu), & \Delta \zeta^{\sigma}_{2,\nu} &= \zeta^{\sigma}_{2,\nu+1} - \zeta^{\sigma}_{2,\nu}, \\ u^{\sigma}_{\nu} &= u^{\sigma}(t_\nu). \end{aligned}$$

Until further notice, c, C, C', etc., will denote constants determined by $C^{\sigma}_{\mathrm{TAME}}$ in (4.2) together with a_{MAX} and upper bounds for T and $|q_0|$. The symbols c, C, C', etc., may denote different constants in different occurrences. We assume that

$$\Delta t_{\mathrm{MAX}} \equiv \max_{0 \leq \nu < N} (t_{\nu+1} - t_\nu) < c \quad \text{for a small enough constant } c. \tag{4.3}$$

We write $X = O(Y)$ to indicate that $|X| \leq CY$. We write $\mathcal{F}^{\sigma}_{\nu}$ to denote the sigma algebra of events determined by $q^{\sigma}(t_\mu)$ for $\mu = 0, 1, \ldots, \nu$. Note that $\mathcal{F}^{\sigma}_{\nu}$ depends on σ, because a_{TRUE} is not deterministic.

We write Prob[...] to denote probability, and we write E[...] to denote expectation.

If we condition on $a_{\mathrm{TRUE}} = a$, then we write $\mathrm{Prob}_a[\ldots]$ and $\mathrm{E}_a[\ldots]$ to denote the corresponding probability and expectation. Thus, for instance, $\mathrm{E}_a[X \,|\, \mathcal{F}^{\sigma}_{\nu}]$ denotes the expected value of X conditioned on $\mathcal{F}^{\sigma}_{\nu}$, given that $a_{\mathrm{TRUE}} = a$.

For any event $\mathcal{E}$, we have

$$\operatorname{Prob}[\mathcal{E}] = \int_{-a_{\text{MAX}}}^{a_{\text{MAX}}} \operatorname{Prob}_a[\mathcal{E}]\, d\operatorname{Prior}(a) \tag{4.4}$$

$$\operatorname{Prob}[\mathcal{E} \mid \mathcal{F}_\nu^\sigma] = \int_{-a_{\text{MAX}}}^{a_{\text{MAX}}} \operatorname{Prob}_a[\mathcal{E} \mid \mathcal{F}_\nu^\sigma]\, d\operatorname{Post}(a \mid \mathcal{F}_\nu^\sigma), \tag{4.5}$$

where $d\operatorname{Post}(a \mid \mathcal{F}_\nu^\sigma)$ is the posterior probability distribution for a_{TRUE} conditioned on $\mathcal{F}_\nu^\sigma$.

Similarly, for any random variable X, we have

$$\mathrm{E}[X] = \int_{-a_{\text{MAX}}}^{a_{\text{MAX}}} \mathrm{E}_a[X]\, d\operatorname{Prior}(a) \tag{4.6}$$

$$\mathrm{E}[X \mid \mathcal{F}_\nu^\sigma] = \int_{-a_{\text{MAX}}}^{a_{\text{MAX}}} \mathrm{E}_a[X \mid \mathcal{F}_\nu^\sigma]\, d\operatorname{Post}(a \mid \mathcal{F}_\nu^\sigma). \tag{4.7}$$

Thanks to (4.4)–(4.7), the following results are immediate from Lemma 3.2.

Lemma 4.1 (Bayesian lemma on rare events). *Suppose $Q > C$ for a large enough C. Then the following hold with probability $> 1 - \exp(-cQ^2)$:*

- $|q^\sigma(t_\nu)|, |u^\sigma(t_\nu)| \le Q$, *for all* ν.
- $|\zeta_1^\sigma(t_\nu)|, |\zeta_2^\sigma(t_\nu)| \le Q^2$, *for all* ν.
- $|\sum_{0\le\mu<\nu}(q^\sigma(t_{\mu+1}) - q^\sigma(t_\mu))^2 - t_\nu| \le Q^2(\Delta t_{\text{MAX}})^{1/2}$, *for all* ν.

Moreover, suppose we fix ν and condition on $\mathcal{F}_\nu^\sigma$, the sigma algebra of events determined by $q^\sigma(t_\mu)$ for $0 \le \mu \le \nu$. Suppose $|q^\sigma(t_\nu)| \le Q$, where $C \le Q \le (\Delta t_\nu)^{-1/1000}$ (for large enough C).

Then for $p = 1, 2$, we have

$$\mathrm{E}[(|\Delta q_\nu^\sigma| + |\Delta\zeta_{1,\nu}^\sigma| + |\Delta\zeta_{2,\nu}^\sigma|)^p \cdot \mathbb{1}_{|\Delta q_\nu^\sigma| + |\Delta\zeta_{1,\nu}^\sigma| + |\Delta\zeta_{2,\nu}^\sigma| > (\Delta t_\nu)^{2/5}} \mid \mathcal{F}_\nu^\sigma] \le C \cdot (\Delta t_\nu)^{100};$$

also

$$\operatorname{Prob}\Big[\max_{t\in[t_\nu, t_{\nu+1}]} |q^\sigma(t) - q_\nu^\sigma| > (\Delta t_\nu)^{2/5} \mid \mathcal{F}_\nu^\sigma\Big] \le C \cdot (\Delta t_\nu)^{1000}$$

and

$$\operatorname{Prob}[|\Delta q_\nu^\sigma| + |\Delta\zeta_{1,\nu}^\sigma| + |\Delta\zeta_{2,\nu}^\sigma| > (\Delta t_\nu)^{2/5} \mid \mathcal{F}_\nu^\sigma] \le C \cdot (\Delta t_\nu)^{1000}.$$

4.2. Posterior probabilities and expectations

Fix $Q > C$ for large enough C. For each ν, let $\text{OK}(\nu)$ denote the set of all $(\bar q_1, \dots, \bar q_\nu) \in \mathbb{R}^\nu$ such that

$$\max_{\mu\le\nu} |\bar q_\mu| + |\bar u_\mu| \le Q \tag{4.8}$$

and

$$\Big|\sum_{\mu<\nu} (\Delta\bar q_\mu)^2 - t_\nu\Big| \le Q^2(\Delta t_{\text{MAX}})^{1/4}, \tag{4.9}$$

where $\Delta\bar q_\mu = \bar q_{\mu+1} - \bar q_\mu$, $\bar q_0 \equiv q_0$, and $\bar u_\mu$ is defined to be the value assigned to $u^\sigma(t_\mu)$ provided $q^\sigma(t_\gamma) = \bar q_\gamma$ for $\gamma \le \mu$.

According to the Bayesian lemma on rare events, we have

$$\text{Prob}\,[\text{NOT OK}(\nu)] \leq \exp(-cQ^2). \tag{4.10}$$

Now, thanks to (4.3) and Lemma 3.3, the following holds for $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$ and $(\bar{q}_1, \ldots, \bar{q}_\nu) \in \text{OK}(\nu)$:

$$\begin{aligned}\text{Prob}[a_{\text{TRUE}} \in [a, a+da] \text{ and } q^\sigma(t_\mu) \in [\bar{q}_\mu, \bar{q}_\mu + d\bar{q}_\mu] \text{ for } \mu \leq \nu]\\ = (1 + O(Q^2(\Delta t_{\text{MAX}})^{1/4})) \cdot (\star) \cdot d\bar{q}_1 \cdots d\bar{q}_\nu,\end{aligned}$$

where

$$(\star) = \Big[d\text{Prior}(a) \cdot \prod_{0 \leq \mu < \nu} \Big\{\frac{1}{\sqrt{2\pi\Delta t_\mu}} \exp\Big(-\frac{1}{2\Delta t_\mu}(\Delta\bar{q}_\mu - (a\bar{q}_\mu + \bar{u}_\mu)\Delta t_\mu)^2\Big)\Big\}\Big].$$

Note that

$$(\star) = d\text{Prior}(a) \cdot \exp\Big(-\frac{1}{2}\bar{\zeta}_{2,\nu}\, a^2 + \bar{\zeta}_{1,\nu}\, a\Big) \cdot \{\text{Factor independent of } a\},$$

where

$$\bar{\zeta}_{2,\nu} = \bar{\zeta}_{2,\nu}(\bar{q}_1, \ldots, \bar{q}_\nu) = \sum_{0 \leq \mu < \nu} \bar{q}_\mu^2\, \Delta t_\mu, \tag{4.11}$$

$$\bar{\zeta}_{1,\nu} = \bar{\zeta}_{1,\nu}(\bar{q}_1, \ldots, \bar{q}_\nu) = \sum_{0 \leq \mu < \nu} \bar{q}_\mu\,(\Delta\bar{q}_\mu - \bar{u}_\mu\,\Delta t_\mu). \tag{4.12}$$

Hence, for

$$(\bar{q}_1, \ldots, \bar{q}_\nu) \in \text{OK}(\nu),$$

the posterior probability distribution for a_{TRUE}, given that $q^\sigma(t_\mu) = \bar{q}_\mu$ for $\mu = 1, \ldots, \nu$, is given by

$$\begin{aligned} d\text{Post}(a|\bar{q}_1, \ldots, \bar{q}_\nu)\\ = (1 + O(Q^2(\Delta t_{\text{MAX}})^{1/4})) \cdot \frac{d\text{Prior}(a) \cdot \exp\big(-\frac{1}{2}\bar{\zeta}_{2,\nu}a^2 + \bar{\zeta}_{1,\nu}a\big)}{Z}\end{aligned} \tag{4.13}$$

for a normalizing constant Z. Since

$$\int_{-a_{\text{MAX}}}^{a_{\text{MAX}}} d\text{Post}(a|\bar{q}_1, \ldots, \bar{q}_\nu) = 1,$$

equation (4.13) holds with

$$Z = \int_{-a_{\text{MAX}}}^{a_{\text{MAX}}} d\text{Prior}(a) \cdot \exp\Big(-\frac{1}{2}\bar{\zeta}_{2,\nu}\, a^2 + \bar{\zeta}_{1,\nu}\, a\Big). \tag{4.14}$$

We have thus proven the following result.

Lemma 4.2 (Lemma on posterior probabilities). *Suppose $Q \geq C$ for a large enough C. Fix ν, and suppose $(q^\sigma(t_1), \dots, q^\sigma(t_\nu))$ satisfy*

$$\max_{\mu \leq \nu} |q^\sigma(t_\mu)| + |u^\sigma(t_\mu)| \leq Q, \tag{4.15}$$

$$\Big| \sum_{\mu<\nu} (\Delta q^\sigma(t_\mu))^2 - t_\nu \Big| \leq Q^2 (\Delta t_{\text{MAX}})^{1/4}. \tag{4.16}$$

Then the posterior probability distribution for a_{TRUE} conditioned on $\mathcal{F}_\nu^\sigma$ is given by

$$\begin{aligned} d\,\mathrm{Post}(a \,|\, \mathcal{F}_\nu^\sigma) &= (1 + O(Q^2(\Delta t_{\text{MAX}})^{1/4})) \\ &\quad \cdot \frac{\exp\Big(-\frac{1}{2}\zeta_2^\sigma(t_\nu) a^2 + \zeta_1^\sigma(t_\nu) a\Big) \cdot d\,\mathrm{Prior}(a)}{Z}, \end{aligned} \tag{4.17}$$

with

$$Z = \int_{-a_{\text{MAX}}}^{a_{\text{MAX}}} \exp\Big(-\frac{1}{2}\zeta_2^\sigma(t_\nu) b^2 + \zeta_1^\sigma(t_\nu) b\Big)\, d\,\mathrm{Prior}(b). \tag{4.18}$$

Corollary 4.3. *Under the assumption of the above lemma, we have*

$$\mathrm{E}[a_{\text{TRUE}} \,|\, \mathcal{F}_\nu^\sigma] = O(Q^2(\Delta t_{\text{MAX}})^{1/4}) + \bar{a}(\zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu)), \tag{4.19}$$

where

$$\bar{a}(\bar{\zeta}_1, \bar{\zeta}_2) \equiv \frac{\int_{-a_{\text{MAX}}}^{a_{\text{MAX}}} a \exp\big(-\frac{a^2}{2}\bar{\zeta}_2 + a\bar{\zeta}_1\big)\, d\,\mathrm{Prior}(a)}{\int_{-a_{\text{MAX}}}^{a_{\text{MAX}}} \exp\big(-\frac{a^2}{2}\bar{\zeta}_2 + a\bar{\zeta}_1\big)\, d\,\mathrm{Prior}(a)} \tag{4.20}$$

for $(\bar{\zeta}_1, \bar{\zeta}_2) \in \mathbb{R}^2$.

Thanks to the above corollary, formula (4.7), and Lemma 3.5, we now have the following results.

Lemma 4.4 (Lemma on posterior expectations). *Suppose $Q \geq C$ for large enough C, and assume that $\Delta t_{\text{MAX}} \leq Q^{-1000}$.*

Define the event

$$\text{TAME}(\nu) = \{|\Delta q_\nu^\sigma| \leq 2(\Delta t_\nu)^{2/5}, |\Delta \zeta_{1,\nu}^\sigma| \leq 2(\Delta t_\nu)^{2/5}, |\Delta \zeta_{2,\nu}^\sigma| \leq 2(\Delta t_\nu)^{2/5}\}.$$

Fix ν, and suppose we have

$$\max_{\mu \leq \nu} |q^\sigma(t_\mu)| + |u^\sigma(t_\mu)| \leq Q$$

and

$$\Big| \sum_{0 \leq \mu < \nu} (\Delta q^\sigma(t_\mu))^2 - t_\nu \Big| \leq Q^2 (\Delta t_{\text{MAX}})^{1/4}.$$

Then the following hold:

$$\begin{aligned}
&\mathrm{E}[(\Delta q^\sigma_\nu)\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\,|\,\mathcal{F}^\sigma_\nu] = [\bar a(\zeta^\sigma_1(t_\nu),\zeta^\sigma_2(t_\nu))\,q^\sigma_\nu + u^\sigma_\nu](\Delta t_\nu) + \mathrm{ERR}\,1,\\
&\mathrm{E}[(\Delta \zeta^\sigma_{1,\nu})\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\,|\,\mathcal{F}^\sigma_\nu] = \bar a(\zeta^\sigma_{1,\nu}(t_\nu),\zeta^\sigma_{2,\nu}(t_\nu))\cdot (q^\sigma_\nu)^2(\Delta t_\nu) + \mathrm{ERR}\,2,\\
&\mathrm{E}[(\Delta \zeta^\sigma_{2,\nu})\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\,|\,\mathcal{F}^\sigma_\nu] = (q^\sigma_\nu)^2(\Delta t_\nu) + \mathrm{ERR}\,3,\\
&\mathrm{E}[(\Delta q^\sigma_\nu)^2\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\,|\,\mathcal{F}^\sigma_\nu] = (\Delta t_\nu) + \mathrm{ERR}\,4,\\
&\mathrm{E}[(\Delta q^\sigma_\nu)\cdot(\Delta \zeta^\sigma_{1,\nu})\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\,|\,\mathcal{F}^\sigma_\nu] = q^\sigma_\nu(\Delta t_\nu) + \mathrm{ERR}\,5,\\
&\mathrm{E}[(\Delta q^\sigma_\nu)\cdot(\Delta \zeta^\sigma_{2,\nu})\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\,|\,\mathcal{F}^\sigma_\nu] = \mathrm{ERR}\,6,\\
&\mathrm{E}[(\Delta \zeta^\sigma_{1,\nu})^2\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\,|\,\mathcal{F}^\sigma_\nu] = (q^\sigma_\nu)^2(\Delta t_\nu) + \mathrm{ERR}\,7,\\
&\mathrm{E}[(\Delta \zeta^\sigma_{1,\nu})(\Delta \zeta^\sigma_{2,\nu})\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\,|\,\mathcal{F}_\nu] = \mathrm{ERR}\,8,\\
&\mathrm{E}[(\Delta \zeta^\sigma_{2,\nu})^2\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\,|\,\mathcal{F}^\sigma_\nu] = \mathrm{ERR}\,9,
\end{aligned}$$

where

$$|\mathrm{ERR}\,1|,\ldots,|\mathrm{ERR}\,9| \le C' Q^4 (\Delta t_{\mathrm{MAX}})^{1/4}\Delta t_\nu.$$

Moreover, under the above assumptions on $(q^\sigma(t_1),\ldots,q^\sigma(t_\nu))$, *we have*

$$\mathrm{Prob}[\text{NOT TAME}(\nu)\,|\,\mathcal{F}^\sigma_\nu] \le (\Delta t_\nu)^{20}.$$

Here,

$$\bar a(\zeta_1,\zeta_2) := \frac{\int_{-a_{\mathrm{MAX}}}^{a_{\mathrm{MAX}}} a \exp\big(-\frac{a^2}{2}\zeta_2 + a\zeta_1\big)\,d\mathrm{Prior}(a)}{\int_{-a_{\mathrm{MAX}}}^{a_{\mathrm{MAX}}} \exp\big(-\frac{a^2}{2}\zeta_2 + a\zeta_1\big)\,d\mathrm{Prior}(a)}.$$

4.3. The PDE assumption

For our given probability distribution $d\mathrm{Prior}(a)$, we fix the function $\bar a(\zeta_1,\zeta_2)$ given in the preceding section, and we introduce the following PDE for an unknown function $S(q,t,\zeta_1,\zeta_2)$ defined on $\mathbb{R}\times[0,T]\times\mathbb{R}\times[0,\infty)$:

$$\begin{aligned}
0 = \partial_t S + (\bar a(\zeta_1,\zeta_2)q + u_{\mathrm{opt}})\,\partial_q S + \bar a\,(\zeta_1,\zeta_2)q^2\,\partial_{\zeta_1} S + q^2\,\partial_{\zeta_2} S + \frac{1}{2}\,\partial_q^2 S\\
+ q\,\partial_{q\zeta_1} S + \frac{1}{2}\,q^2\,\partial^2_{\zeta_1} S + (q^2 + u^2_{\mathrm{opt}}),
\end{aligned} \tag{4.21}$$

where

$$u_{\mathrm{opt}} = -\frac{1}{2}\,\partial_q S. \tag{4.22}$$

We *assume* the existence of a solution of the above PDE, satisfying the following additional conditions.

The TERMINAL CONDITION:

$$S(q,t,\zeta_1,\zeta_2) = 0 \quad \text{at } t = T. \tag{4.23}$$

POSITIVITY:

$$S(q,t,\zeta_1,\zeta_2) \ge 0. \tag{4.24}$$

ESTIMATES: We assume $S \in C^{2,1}$. For $|\alpha| \le 3$, we have almost everywhere that

$$|\partial^\alpha S(q,t,\zeta_1,\zeta_2)| \le K \cdot [1 + |q| + |\zeta_1| + |\zeta_2|]^{m_0} \tag{4.25}$$

for constants $K, m_0 \ge 1$. Moreover,

$$|u_{\text{opt}}(q,t,\zeta_1,\zeta_2)| \le C^{\text{opt}}_{\text{TAME}}\,[|q| + 1]. \tag{4.26}$$

Note that (4.25) holds everywhere, not just almost everywhere, when $|\alpha| \le 2$.

We call K, m_0, $C^{\text{opt}}_{\text{TAME}}$, a_{MAX}, and our upper bounds for T and $|q_0|$ the BOILERPLATE CONSTANTS. We now broaden our definition of constants c, C, C', etc., to allow them to depend on the BOILERPLATE CONSTANTS. As usual, these symbols may denote different constants in different occurrences.

We strengthen our large Q assumption. More precisely, we assume from now on that $Q \ge C$ for a large enough constant C. Since the meaning of the constant C has changed, the above is stronger than our previous large Q assumption.

We assume that

$$\Delta t_{\text{MAX}} \le Q^{-1000}.$$

4.4. The allegedly optimal strategy

Let $u_{\text{opt}} = u_{\text{opt}}(q,t,\zeta_1,\zeta_2)$ be as in Section 4.3. We define a strategy $\tilde{\sigma}$ based on the function u_{opt}.

Given ν ($1 \le \nu \le N$) and given real numbers $q_1, \dots, q_\nu$, we define numbers u_μ, $\zeta_{1,\mu}$ and $\zeta_{2,\mu}$ by induction on $\mu = 0, 1, \dots, \nu$ so that $\zeta_{1,0} = \zeta_{2,0} = 0$, and, for each μ,

$$\begin{aligned} u_\mu &= u_{\text{opt}}(q_\mu, t_\mu, \zeta_{1,\mu}, \zeta_{2,\mu}), \\ \zeta_{1,\mu} &= \sum_{0 \le \gamma < \mu} q_\gamma\,([q_{\gamma+1} - q_\gamma] - u_\gamma \Delta t_\gamma), \\ \zeta_{2,\mu} &= \sum_{0 \le \gamma < \mu} q_\gamma^2\,\Delta t_\gamma, \end{aligned}$$

We then set $\tilde{\sigma}_\nu(q_1, \dots, q_\nu)$ equal to the above u_μ for $\mu = \nu$.

We define our ALLEGEDLY OPTIMAL STRATEGY $\tilde{\sigma}$ to be the collection of tame rules

$$\tilde{\sigma} = (\tilde{\sigma}_\nu)_{\nu = 0,1,\dots,N-1}.$$

Thanks to our PDE assumption (see (4.26)), each $\tilde{\sigma}_\nu$ is indeed a tame rule with tame constant $C^{\text{opt}}_{\text{TAME}}$, hence $\tilde{\sigma}$ is a strategy.

4.5. Performance of competing strategies

We have just defined the allegedly optimal strategy $\tilde{\sigma}$, that is based on the functions $u_{\text{opt}}(q,t,\zeta_1,\zeta_2)$, $S(q,t,\zeta_1,\zeta_2)$, and $\bar{a}(\zeta_1,\zeta_2)$.

In this section, we compare the performance of $\tilde{\sigma}$ with that of an arbitrary competing (deterministic, tame) strategy σ, also defined with respect to the given partition $0 = t_0 < t_1 < \cdots < t_N = T$. We assume that

$$(\Delta t_{\text{MAX}}) \le Q^{-2000\,m_0}. \tag{4.27}$$

(See Section 4.3 for m_0.) Thanks to our assumption 4.26, the strategy $\tilde{\sigma}$ satisfies

$$|u^{\tilde{\sigma}}(t_\nu)| \le C^{\text{opt}}_{\text{TAME}}\,[|q^{\tilde{\sigma}}(t_\nu)| + 1],$$

i.e., $\tilde{\sigma}$ is tame with constant $C^{\text{opt}}_{\text{TAME}}$. We assume that the strategy σ is tame with constant C^{σ}_{TAME}, i.e., we assume that

$$|u^{\sigma}(t_\nu)| \le C^{\sigma}_{\text{TAME}}\,[|q^{\sigma}(t_\nu)| + 1].$$

In this section, we write c, C, C', etc., to denote constants determined by C^{σ}_{TAME} and the BOILERPLATE CONSTANTS (one of which is $C^{\text{opt}}_{\text{TAME}}$). We strengthen our large Q assumption by supposing that $Q > C$ for a large enough C. Since the meaning of constants C has changed, this indeed strengthens our previous large Q assumption.

We define

$$\mathcal{U} = \{(q, \zeta_1, \zeta_2) \in \mathbb{R}^3 : |q|, |\zeta_1|, |\zeta_2| \le Q\}, \tag{4.28}$$

$$\mathcal{U}^+ = \{(q, \zeta_1, \zeta_2) \in \mathbb{R}^3 : |q|, |\zeta_1|, |\zeta_2| \le 2Q\}. \tag{4.29}$$

For each μ ($0 \le \mu \le N$) we introduce the event

(4.30)

$$\text{OK}_\mu = \Big\{(q^\sigma(t_\mu), \zeta_1^\sigma(t_\mu), \zeta_2^\sigma(t_\mu)) \in \mathcal{U}, \Big|\sum_{0\le\gamma<\mu} (\Delta q^\sigma(t_\gamma))^2 - t_\mu\Big| \le Q^2(\Delta t_{\text{MAX}})^{1/4}\Big\}.$$

We define the event

$$\text{DISASTER} = \{\text{OK}_\mu \text{ fails for some } \mu \le N\} \tag{4.31}$$

and the stopping time

$$\tau = \begin{cases} t_\mu \text{ for the least } \mu \text{ for which } \text{OK}_\mu \text{ fails}, & \text{if there is such a } \mu, \\ T = t_N & \text{otherwise.} \end{cases} \tag{4.32}$$

Note that τ is indeed a stopping time with respect to the $\mathcal{F}^\sigma_\nu$, i.e., the event $\tau > t_\nu$ is $\mathcal{F}^\sigma_\nu$-measurable, for each ν.

We define a *cost-to-go* by setting

$$\text{CTG}^\sigma(t_\nu) = \sum_{t_\nu \le t_\mu < \tau} [(q^\sigma(t_\mu))^2 + (u^\sigma(t_\mu))^2]\,\Delta t_\mu. \tag{4.33}$$

To measure the difference between the strategies σ and $\tilde{\sigma}$, we introduce the random variable

$$\text{DISCREP}^\sigma(t_\mu) = u^\sigma(t_\mu) - u_{\text{opt}}(q^\sigma(t_\mu), t_\mu, \zeta_1^\sigma(t_\mu), \zeta_2^\sigma(t_\mu)). \tag{4.34}$$

Note that

$$u_{\text{opt}}(q^\sigma(t_\mu), t_\mu, \zeta_1^\sigma(t_\mu), \zeta_2^\sigma(t_\mu))$$

is not the same as

$$u^{\tilde{\sigma}}(t_\mu) = u_{\text{opt}}(q^{\tilde{\sigma}}(t_\mu), t_\mu, \zeta_1^{\tilde{\sigma}}(t_\mu), \zeta_2^{\tilde{\sigma}}(t_\mu)).$$

In the next lemma, we compare the cost-to-go of σ with that of $\tilde{\sigma}$. We will be conditioning on $\mathcal{F}^\sigma_\nu$, so the quantities $q^\sigma(t_\nu)$, $u^\sigma(t_\nu)$, $\zeta_1^\sigma(t_\nu)$ and $\zeta_2^\sigma(t_\nu)$ may be regarded as deterministic.

Lemma 4.5 (Main lemma on competing strategies). *Fix* ν $(0 \le \nu \le N)$. *Suppose*

$$(\star) \qquad (q^\sigma(t_\nu), \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu)) \in \mathcal{U}.$$

Then

$$\text{(A)} \qquad \begin{aligned} &\mathrm{E}[\mathrm{CTG}^\sigma(t_\nu) \,|\, \mathcal{F}_\nu^\sigma] + \hat{C}\, Q^{2m_0}\, \mathrm{Prob}[\mathrm{DISASTER} \,|\, \mathcal{F}_\nu^\sigma] \\ &\quad \ge S(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu)) + \mathrm{E}\Big[\sum_{t_\nu \le t_\mu < \tau} (\mathrm{DISCREP}^\sigma(t_\mu))^2 \Delta t_\mu \,|\, \mathcal{F}_\nu^\sigma \Big] \\ &\qquad - Q^{2m_0}(T - t_\nu) \cdot (\Delta t_{\mathrm{MAX}})^{1/20}. \end{aligned}$$

If $\sigma = \tilde{\sigma}$, *then*

$$\text{(B)} \qquad \begin{aligned} \mathrm{E}[\mathrm{CTG}^\sigma(t_\nu) \,|\, \mathcal{F}_\nu^\sigma] &\le S(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu)) \\ &\quad + \hat{C}\, Q^{2m_0}\, \mathrm{Prob}[\mathrm{DISASTER} \,|\, \mathcal{F}_\nu^\sigma] + Q^{2m_0}(T - t_\nu) \cdot (\Delta t_{\mathrm{MAX}})^{1/20}. \end{aligned}$$

We fix the large constant $\hat{C}$ throughout this section.

Proof. We proceed by downward induction on ν.

In the base case, $\nu = N$. Since $\tau \le T = t_N$, we have $\mathrm{CTG}^\sigma(t_\nu) = 0$. Also,

$$S(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu)) = 0$$

by the terminal condition for our PDE. Moreover, the sum

$$\sum_{t_\nu \le t_\mu < \tau} (\mathrm{DISCREP}^\sigma(t_\mu))^2 \Delta t_\mu$$

is empty, and $T - t_\nu = 0$. Therefore, (A) asserts that

$$\hat{C}\, Q^{2m_0}\, \mathrm{Prob}[\mathrm{DISASTER} \,|\, \mathcal{F}_\nu^\sigma] \ge 0,$$

while (B) asserts that if $\sigma = \tilde{\sigma}$, then

$$0 \le \hat{C}\, Q^{2m_0}\, \mathrm{Prob}[\mathrm{DISASTER} \,|\, \mathcal{F}_\nu^\sigma].$$

These two (equivalent) inequalities are obviously correct, so our lemma holds in the base case $\nu = N$.

For the induction step, we fix $\nu < N$, and assume that (A) and (B) hold with $(\nu + 1)$ in place of ν. We will deduce (A) and (B) for the given ν. To do so, we first dispose of a trivial case.

Suppose for a moment that OK_μ *fails for some* $\mu \le \nu$. Then DISASTER occurs, and $\tau \le t_\nu$; consequently, $\mathrm{CTG}^\sigma(t_\nu) = 0$, and

$$\sum_{t_\nu \le t_\mu < \tau} (\mathrm{DISCREP}^\sigma(t_\mu))^2 \Delta t_\mu = 0.$$

Therefore, (A) asserts that

$$\hat{C}\, Q^{2m_0} \ge S(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu)) - Q^{2m_0}(T - t_\nu) \cdot (\Delta t_{\mathrm{MAX}})^{1/2},$$

while (B) asserts that if $\sigma = \tilde{\sigma}$ then

$$0 \leq S(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu)) + \hat{C}\, Q^{2m_0} + Q^{2m_0} \cdot (T - t_\nu) \cdot (\Delta t_{\mathrm{MAX}})^{1/20}.$$

These inequalities are immediate from our assumptions on the PDE solution S, together with our hypothesis

$$(q^\sigma(t_\nu), \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu)) \in \mathcal{U}.$$

Thus, our induction step is complete in the trivial case in which OK_μ fails for some $\mu \leq \nu$.

From now on, we assume that

$$\mathrm{OK}_\mu \text{ holds for all } \mu \leq \nu. \tag{4.35}$$

Thus,

$$\tau \geq t_{\nu+1} \tag{4.36}$$

For the moment, we condition on $\mathcal{F}_{\nu+1}^\sigma$, and distinguish two cases:

Case I. $(q^\sigma(t_{\nu+1}), \zeta_1^\sigma(t_{\nu+1}), \zeta_2^\sigma(t_{\nu+1})) \in \mathcal{U}$,
Case II. $(q^\sigma(t_{\nu+1}), \zeta_1^\sigma(t_{\nu+1}), \zeta_2^\sigma(t_{\nu+1})) \notin \mathcal{U}$.

In *Case* I, our inductive hypothesis tells us the following:

$$\begin{aligned} &\mathrm{E}[\mathrm{CTG}^\sigma(t_{\nu+1}) \,|\, \mathcal{F}_{\nu+1}^\sigma] + \hat{C}\, Q^{2m_0}\, \mathrm{Prob}[\mathrm{DISASTER} \,|\, \mathcal{F}_{\nu+1}^\sigma] \\ &\geq S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta_1^\sigma(t_{\nu+1}), \zeta_2^\sigma(t_{\nu+1})) + \mathrm{E}\Big[\sum_{t_{\nu+1} \leq t_\mu < \tau} (\mathrm{DISCREP}^\sigma(t_\mu))^2 \Delta t_\mu \,|\, \mathcal{F}_{\nu+1}^\sigma \Big] \\ &\quad - Q^{2m_0}(T - t_{\nu+1}) \cdot (\Delta t_{\mathrm{MAX}})^{1/20}; \end{aligned}$$

and if $\sigma = \tilde{\sigma}$ then

$$\begin{aligned} &\mathrm{E}[\mathrm{CTG}^\sigma(t_{\nu+1}) \,|\, \mathcal{F}_{\nu+1}^\sigma] \\ &\qquad \leq S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta_1^\sigma(t_{\nu+1}), \zeta_2^\sigma(t_{\nu+1})) + \hat{C}\, Q^{2m_0}\, \mathrm{Prob}[\mathrm{DISASTER} \,|\, \mathcal{F}_{\nu+1}^\sigma] \\ &\qquad\quad + Q^{2m_0}(T - t_{\nu+1}) \cdot (\Delta t_{\mathrm{MAX}})^{1/20}. \end{aligned}$$

Since

$$\mathrm{CTG}^\sigma(t_\nu) = \mathrm{CTG}^\sigma(t_{\nu+1}) + [(q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2]\, \Delta t_\nu,$$

thanks to (4.36), the above inequalities yield at once that

$$\begin{aligned} &\mathrm{E}[\mathrm{CTG}^\sigma(t_\nu) \,|\, \mathcal{F}_{\nu+1}^\sigma] + \hat{C}\, Q^{2m_0}\, \mathrm{Prob}[\mathrm{DISASTER} \,|\, \mathcal{F}_{\nu+1}^\sigma] \\ &\qquad \geq S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta_1^\sigma(t_{\nu+1}), \zeta_2^\sigma(t_{\nu+1})) + [(q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2]\, \Delta t_\nu \\ &\qquad\quad + \mathrm{E}\Big[\sum_{t_{\nu+1} \leq t_\mu < \tau} (\mathrm{DISCREP}^\sigma(t_\mu))^2 \Delta t_\mu \,|\, \mathcal{F}_{\nu+1}^\sigma \Big] \\ &\qquad\quad - Q^{2m_0} \cdot (T - t_{\nu+1}) \cdot (\Delta t_{\mathrm{MAX}})^{1/20}; \end{aligned} \tag{AI}$$

and if $\sigma = \tilde{\sigma}$, then

$$\text{(BI)} \qquad \begin{aligned} &\mathrm{E}[\textsc{Ctg}^{\sigma}(t_\nu) \,|\, \mathcal{F}^{\sigma}_{\nu+1}] \\ &\quad \le S(q^{\sigma}(t_{\nu+1}), t_{\nu+1}, \zeta_1^{\sigma}(t_{\nu+1}), \zeta_2^{\sigma}(t_{\nu+1})) + [(q^{\sigma}(t_\nu))^2 + (u^{\sigma}(t_\nu))^2]\,\Delta t_\nu \\ &\qquad + \hat{C}\, Q^{2m_0}\, \mathrm{Prob}[\textsc{Disaster} \,|\, \mathcal{F}^{\sigma}_{\nu+1}] + Q^{2m_0}(T - t_{\nu+1}) \cdot (\Delta t_{\text{MAX}})^{1/20}. \end{aligned}$$

Estimates (AI) and (BI) hold in Case I.

We turn to *Case* II. In that case, $\textsc{ok}_{\nu+1}$ fails, $\textsc{Disaster}$ occurs, and $\tau = t_{\nu+1}$, hence

$$\textsc{Ctg}^{\sigma}(t_\nu) = [(q^{\sigma}(t_\nu))^2 + (u^{\sigma}(t_\nu))^2]\,\Delta t_\nu, \quad \mathrm{Prob}[\textsc{Disaster} \,|\, \mathcal{F}^{\sigma}_{\nu+1}] = 1, \quad \text{and}$$
$$\sum_{t_{\nu+1} \le t_\mu < \tau} (\textsc{Discrep}^{\sigma}(t_\mu))^2\, \Delta t_\mu = 0.$$

So we have the following estimates:

$$\text{(AII)} \qquad \begin{aligned} &\mathrm{E}[\textsc{Ctg}^{\sigma}(t_\nu) \,|\, \mathcal{F}^{\sigma}_{\nu+1}] + \hat{C}\, Q^{2m_0}\, \mathrm{Prob}[\textsc{Disaster} \,|\, \mathcal{F}^{\sigma}_{\nu+1}] \\ &\qquad \ge \hat{C}\, Q^{2m_0} + \mathrm{E}\Big[\sum_{t_{\nu+1} \le t_\mu < \tau} (\textsc{Discrep}^{\sigma}(t_\mu))^2 \Delta t_\mu \,|\, \mathcal{F}^{\sigma}_{\nu+1} \Big] \\ &\qquad\quad - Q^{2m_0}(T - t_{\nu+1}) \cdot (\Delta t_{\text{MAX}})^{1/20}; \end{aligned}$$

and if $\sigma = \tilde{\sigma}$ then

$$\text{(BII)} \qquad \begin{aligned} &\mathrm{E}[\textsc{Ctg}^{\sigma}(t_\nu) \,|\, \mathcal{F}^{\sigma}_{\nu+1}] \\ &\quad \le [(q^{\sigma}(t_\nu))^2 + (u^{\sigma}(t_\nu))^2]\,\Delta t_\nu + \hat{C}\, Q^{2m_0}\, \mathrm{Prob}[\textsc{Disaster} \,|\, \mathcal{F}^{\sigma}_{\nu+1}] \\ &\qquad + Q^{2m_0} \cdot (T - t_{\nu+1}) \cdot (\Delta t_{\text{MAX}})^{1/20}. \end{aligned}$$

Estimates (AII) and (BII) hold in Case II.

Combining estimates (AI), (BI), (AII) and (BII), we obtain the following inequalities, which hold in both Cases I and II:

$$\begin{aligned} &\mathrm{E}[\textsc{Ctg}^{\sigma}(t_\nu) \,|\, \mathcal{F}^{\sigma}_{\nu+1}] + \hat{C}\, Q^{2m_0}\, \mathrm{Prob}[\textsc{Disaster} \,|\, \mathcal{F}^{\sigma}_{\nu+1}] \\ &\quad \ge S(q^{\sigma}(t_{\nu+1}), t_{\nu+1}, \zeta_1^{\sigma}(t_{\nu+1}), \zeta_2^{\sigma}(t_{\nu+1})) \cdot \mathbb{1}_{\text{Case I}} \\ &\qquad + C\, \hat{Q}^{2m_0} \mathbb{1}_{\text{Case II}} + [(q^{\sigma}(t_\nu))^2 + (u^{\sigma}(t_\nu))^2]\,\Delta t_\nu \cdot \mathbb{1}_{\text{Case I}} \\ &\qquad + \mathrm{E}\Big[\sum_{t_{\nu+1} \le t_\mu < \tau} (\textsc{Discrep}^{\sigma}(t_\mu))^2 \Delta t_\mu \,|\, \mathcal{F}^{\sigma}_{\nu+1} \Big] - Q^{2m_0} \cdot (T - t_{\nu+1}) \cdot (\Delta t_{\text{MAX}})^{1/20}; \end{aligned}$$

and if $\sigma = \tilde{\sigma}$, then

$$\begin{aligned} \mathrm{E}[\textsc{Ctg}^{\sigma}(t_\nu) \,|\, \mathcal{F}^{\sigma}_{\nu+1}] \le\; & S(q^{\sigma}(t_{\nu+1}), t_{\nu+1}, \zeta_1^{\sigma}(t_{\nu+1}), \zeta_2^{\sigma}(t_{\nu+1})) \cdot \mathbb{1}_{\text{Case I}} \\ & + [(q^{\sigma}(t_\nu))^2 + (u^{\sigma}(t_\nu))^2]\,\Delta t_\nu + \hat{C}\, Q^{2m_0}\, \mathrm{Prob}[\textsc{Disaster} \,|\, \mathcal{F}^{\sigma}_{\nu+1}] \\ & + Q^{2m_0} \cdot (T - t_{\nu+1}) \cdot (\Delta t_{\text{MAX}})^{1/20}. \end{aligned}$$

We now cease conditioning on $\mathcal{F}^\sigma_{\nu+1}$ and instead condition on $\mathcal{F}^\sigma_\nu$. From the two preceding inequalities we obtain the following estimates, valid whenever (4.35) holds:

$$\text{(A*)}\quad \begin{aligned} &\mathrm{E}[\mathrm{CTG}^\sigma(t_\nu)\,|\,\mathcal{F}^\sigma_\nu] + \hat{C}\,Q^{2m_0}\,\mathrm{Prob}[\mathrm{DISASTER}\,|\,\mathcal{F}^\sigma_\nu] \\ &\quad \geq \mathrm{E}[S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta^\sigma_1(t_{\nu+1}), \zeta^\sigma_2(t_{\nu+1}))\cdot \mathbb{1}_{\mathrm{CASE\ I}}\,|\,\mathcal{F}^\sigma_\nu] \\ &\qquad + \mathrm{E}[\hat{C}\,Q^{2m_0}\mathbb{1}_{\mathrm{CASE\ II}}\,|\,\mathcal{F}^\sigma_\nu] + [(q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2]\,\Delta t_\nu\cdot \mathrm{E}[\mathbb{1}_{\mathrm{CASE\ I}}\,|\,\mathcal{F}^\sigma_\nu] \\ &\qquad + \mathrm{E}\Big[\sum_{t_{\nu+1}\leq t_\mu<\tau}(\mathrm{DISCREP}^\sigma(t_\mu))^2\,\Delta t_\mu\,|\,\mathcal{F}^\sigma_\nu\Big] \\ &\qquad - Q^{2m_0}\cdot(T - t_{\nu+1})\cdot(\Delta t_{\mathrm{MAX}})^{1/20}; \end{aligned}$$

and if $\sigma = \tilde{\sigma}$, then

$$\text{(B*)}\quad \begin{aligned} \mathrm{E}[\mathrm{CTG}^\sigma(t_\nu)\,|\,\mathcal{F}^\sigma_\nu] \leq{}& \mathrm{E}[S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta^\sigma_1(t_{\nu+1}), \zeta^\sigma_2(t_{\nu+1}))]\cdot \mathbb{1}_{\mathrm{CASE\ I}}\,|\,\mathcal{F}^\sigma_\nu] \\ &+ [(q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2]\,\Delta t_\nu + \hat{C}\,Q^{2m_0}\,\mathrm{Prob}[\mathrm{DISASTER}\,|\,\mathcal{F}^\sigma_\nu] \\ &+ Q^{2m_0}\cdot(T - t_{\nu+1})\cdot(\Delta t_{\mathrm{MAX}})^{1/20}. \end{aligned}$$

Next, from Lemma 4.4, recall the event

$$\mathrm{TAME}(\nu) = \{|\Delta q^\sigma(t_\nu)|, |\Delta\zeta^\sigma_1(t_\nu)|, |\Delta\zeta^\sigma_2(t_\nu)| \leq 2(\Delta t_\nu)^{2/5}\}, \tag{4.37}$$

and the estimate

$$\mathrm{Prob}[\mathrm{NOT\ TAME}(\nu)\,|\,\mathcal{F}^\sigma_\nu] \leq (\Delta t_\nu)^{20}. \tag{4.38}$$

(Note that Lemma 4.4 applies, thanks to (4.35).) If $\mathrm{TAME}(\nu)$ occurs, then, since

$$(q^\sigma(t_\nu), \zeta^\sigma_1(t_\nu), \zeta^\sigma_2(t_\nu)) \in \mathcal{U}$$

(by the hypothesis of the present lemma), we have

$$(q^\sigma(t_{\nu+1}), \zeta^\sigma_1(t_{\nu+1}), \zeta^\sigma_2(t_{\nu+1})) \in \mathcal{U}^+,$$

hence

$$0 \leq S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta^\sigma_1(t_{\nu+1}), \zeta^\sigma_2(t_{\nu+1})) \leq CQ^{m_0}.$$

Therefore, if we take $\hat{C}$ large enough, then

$$\begin{aligned} S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta^\sigma_1(t_{\nu+1}), \zeta^\sigma_2(t_{\nu+1}))\cdot \mathbb{1}_{\mathrm{CASE\ I}} + \frac{1}{2}\,\hat{C}\,Q^{2m_0}\cdot \mathbb{1}_{\mathrm{CASE\ II}} \\ \geq S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta^\sigma_1(t_{\nu+1}), \zeta^\sigma_2(t_{\nu+1}))\cdot \mathbb{1}_{\mathrm{TAME}(\nu)} \end{aligned} \tag{4.39}$$

and

$$\frac{1}{2}\,\hat{C}\,Q^{2m_0}\cdot \mathbb{1}_{\mathrm{CASE\ II}} \geq [(q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2]\,\Delta t_\nu\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\cdot \mathbb{1}_{\mathrm{CASE\ II}}. \tag{4.40}$$

Putting (4.39) and (4.40) into (A*), we learn that

$$\begin{aligned}
&\mathrm{E}[\textsc{Ctg}^\sigma(t_\nu)\,|\,\mathcal{F}_\nu^\sigma] + \hat{C}\,Q^{2m_0}\mathrm{Prob}[\textsc{Disaster}\,|\,\mathcal{F}_\nu^\sigma] \\
&\qquad \geq \mathrm{E}[S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta_1^\sigma(t_{\nu+1}), \zeta_2^\sigma(t_{\nu+1}))\cdot \mathbb{1}_{\textsc{tame}(\nu)}\,|\,\mathcal{F}_\nu^\sigma] \\
&\qquad\quad + [(q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2]\,\Delta t_\nu\, \mathrm{Prob}[\textsc{tame}(\nu)\,|\,\mathcal{F}_\nu^\sigma] \\
&\qquad\quad + \mathrm{E}\Big[\sum_{t_{\nu+1}\leq t_\mu<\tau} (\textsc{Discrep}^\sigma(t_\mu))^2\,\Delta t_\mu\,\Big|\,\mathcal{F}_\nu^\sigma\Big] \\
&\qquad\quad - Q^{2m_0}\cdot(T - t_{\nu+1})\cdot(\Delta t_{\textsc{max}})^{1/20}.
\end{aligned} \tag{A\#}$$

To obtain an analogous result from (B*), we note that

$$\begin{aligned}
&S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta_1^\sigma(t_{\nu+1}), \zeta_2^\sigma(t_{\nu+1}))\cdot\mathbb{1}_{\textsc{Case I}} \\
&\quad \leq S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta_1^\sigma(t_{\nu+1}), \zeta_2^\sigma(t_{\nu+1}))\cdot\mathbb{1}_{(q^\sigma(t_{\nu+1}),\zeta_1^\sigma(t_{\nu+1}),\zeta_2^\sigma(t_{\nu+1}))\in\mathcal{U}^+} \\
&\quad \leq S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta_1^\sigma(t_{\nu+1}), \zeta_2^\sigma(t_{\nu+1}))\cdot\mathbb{1}_{\textsc{tame}(\nu)} + CQ^{2m_0}\cdot\mathbb{1}_{\textsc{not tame}(\nu)},
\end{aligned}$$

and therefore,

$$\begin{aligned}
&\mathrm{E}[S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta_1^\sigma(t_{\nu+1}), \zeta_2^\sigma(t_{\nu+1}))\cdot\mathbb{1}_{\textsc{Case I}}\,|\,\mathcal{F}_\nu^\sigma] \\
&\quad \leq \mathrm{E}[S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta_1^\sigma(t_{\nu+1}), \zeta_2^\sigma(t_{\nu+1}))\cdot\mathbb{1}_{\textsc{tame}(\nu)}\,|\,\mathcal{F}_\nu^\sigma] + CQ^{2m_0}(\Delta t_\nu)^{20},
\end{aligned} \tag{4.41}$$

thanks to (4.38).

Putting (4.41) into (B*), we obtain for $\sigma = \tilde{\sigma}$ the inequality

$$\begin{aligned}
\mathrm{E}[\textsc{Ctg}^\sigma(t_\nu)\,|\,\mathcal{F}_\nu^\sigma] &\leq \mathrm{E}[S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta_1^\sigma(t_{\nu+1}), \zeta_2^\sigma(t_{\nu+1}))\cdot\mathbb{1}_{\textsc{tame}(\nu)}\,|\,\mathcal{F}_\nu^\sigma] \\
&\quad + CQ^{2m_0}(\Delta t_\nu)^{20} + [(q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2]\,\Delta t_\nu \\
&\quad + \hat{C}\,Q^{2m_0}\,\mathrm{Prob}[\textsc{Disaster}\,|\,\mathcal{F}_\nu^\sigma] \\
&\quad + Q^{2m_0}(T - t_{\nu+1})\cdot(\Delta t_{\textsc{max}})^{1/20}.
\end{aligned} \tag{B\#}$$

Estimates (A#) (for general σ) and (B#) (for $\sigma = \tilde{\sigma}$) hold whenever (4.35) is satisfied.

We next study the quantity

$$\mathrm{E}[S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta_1^\sigma(t_{\nu+1}), \zeta_2^\sigma(t_{\nu+1}))\cdot\mathbb{1}_{\textsc{tame}(\nu)}\,|\,\mathcal{F}_\nu^\sigma],$$

which appears in (A#) and (B#).

Thanks to conditions (4.35), (4.30) and (4.37), the points $(q^\sigma(t_\nu), \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu))$ and $(q^\sigma(t_{\nu+1}), \zeta_1^\sigma(t_{\nu+1}), \zeta_2^\sigma(t_{\nu+1}))$ both lie in $\mathcal{U}^+$ provided $\textsc{tame}(\nu)$ holds; see (4.29). Recall that the third derivatives of S are assumed to be bounded a.e. by CQ^{m_0} on $\mathcal{U}^+$. Therefore, if $\textsc{tame}(\nu)$ holds, then

$$\begin{aligned}
&\Big|S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta_1^\sigma(t_{\nu+1}), \zeta_2^\sigma(t_{\nu+1})) \\
&\quad - \sum_{p_1+p_2+p_3+p_4\leq 2}\Big\{\frac{(\partial_t^{p_1}\partial_q^{p_2}\partial_{\zeta_1}^{p_3}\partial_{\zeta_2}^{p_4}S)}{p_1!p_2!p_3!p_4!}(\Delta t_\nu)^{p_1}(\Delta q^\sigma(t_\nu))^{p_2}(\Delta\zeta_1^\sigma(t_\nu))^{p_3}(\Delta\zeta_2^\sigma(t_\nu))^{p_4}\Big\}\Big| \\
&\quad \leq CQ^{m_0}\max\{(\Delta t_\nu), |\Delta q^\sigma(t_\nu)|, |\Delta\zeta_1^\sigma(t_\nu)|, |\Delta\zeta_2^\sigma(t_\nu)|\}^3 \leq C'Q^{m_0}[(\Delta t_\nu)^{2/5}]^3,
\end{aligned} \tag{4.42}$$

where the partials of S are evaluated at $(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu))$, and we have made use of (4.37).

Moreover, the summands in (4.42) satisfy

$$|(\partial_t^{p_1} \cdots \partial_{\zeta_2}^{p_4} S)\,(\Delta t_\nu)^{p_1}\,(\Delta q^\sigma(t_\nu))^{p_2}\,(\Delta \zeta_1^\sigma(t_\nu))^{p_3}\,(\Delta \zeta_2^\sigma(t_\nu))^{p_4}| \leq C Q^{m_0}\,(\Delta t_\nu)^{p_1+\frac{2}{5}[p_2+p_3+p_4]}$$

whenever (4.35) and TAME(ν) hold.

Consequently, (4.42) implies that

$$\begin{aligned}
&\Bigg| S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta_1^\sigma(t_{\nu+1}), \zeta_2^\sigma(t_{\nu+1})) \cdot \mathbb{1}_{\mathrm{TAME}(\nu)} \\
&\quad - \sum_{\substack{p_1+p_2+p_3+p_4\leq 2 \\ p_1+\frac{2}{5}(p_2+p_3+p_4)\leq 1}} \Bigg\{ \frac{(\partial_t^{p_1}\partial_q^{p_2}\partial_{\zeta_1}^{p_3}\partial_{\zeta_2}^{p_4} S)}{p_1!p_2!p_3!p_4!} (\Delta t_\nu)^{p_1} (\Delta q^\sigma(t_\nu))^{p_2} \\
&\qquad \cdot (\Delta \zeta_1^\sigma(t_\nu))^{p_3} (\Delta \zeta_2^\sigma(t_\nu))^{p_4} \cdot \mathbb{1}_{\mathrm{TAME}(\nu)} \Bigg\} \Bigg| \\
&\quad \leq C Q^{m_0} (\Delta t_\nu)^{6/5}.
\end{aligned} \tag{4.43}$$

The summands entering into (4.43) arise from S, $\partial_t S$, $\partial_q S$, $\partial_{\zeta_1} S$, $\partial_{\zeta_2} S$, $\partial_q^2 S$, $\partial_{q\zeta_1} S$, $\partial_{q\zeta_2} S$, $\partial_{\zeta_1}^2 S$, $\partial_{\zeta_1\zeta_2} S$ and $\partial_{\zeta_2}^2 S$. We conclude that

$$\left| S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta_1^\sigma(t_{\nu+1}), \zeta_2^\sigma(t_{\nu+1})) \cdot \mathbb{1}_{\mathrm{TAME}(\nu)} - \sum_{i=0}^{10} \mathrm{TERM}\, i \right| \leq C Q^{m_0}\,(\Delta t_\nu)^{6/5} \tag{4.44}$$

whenever (4.35) holds, where

$$\mathrm{TERM}\,0 = S \cdot \mathbb{1}_{\mathrm{TAME}(\nu)}, \tag{4.45}$$

$$\mathrm{TERM}\,1 = (\partial_t S) \cdot (\Delta t_\nu) \cdot \mathbb{1}_{\mathrm{TAME}(\nu)}, \tag{4.46}$$

$$\mathrm{TERM}\,2 = (\partial_q S) \cdot (\Delta q^\sigma(t_\nu)) \cdot \mathbb{1}_{\mathrm{TAME}(\nu)}, \tag{4.47}$$

$$\mathrm{TERM}\,3 = (\partial_{\zeta_1} S) \cdot (\Delta \zeta_1^\sigma(t_\nu)) \cdot \mathbb{1}_{\mathrm{TAME}(\nu)}, \tag{4.48}$$

$$\mathrm{TERM}\,4 = (\partial_{\zeta_2} S) \cdot (\Delta \zeta_2^\sigma(t_\nu)) \cdot \mathbb{1}_{\mathrm{TAME}(\nu)}, \tag{4.49}$$

$$\mathrm{TERM}\,5 = \frac{1}{2}\,(\partial_q^2 S) \cdot (\Delta q^\sigma(t_\nu))^2 \cdot \mathbb{1}_{\mathrm{TAME}(\nu)}, \tag{4.50}$$

$$\mathrm{TERM}\,6 = (\partial_{q\zeta_1} S) \cdot (\Delta q^\sigma(t_\nu))\,(\Delta \zeta_1^\sigma(t_\nu)) \cdot \mathbb{1}_{\mathrm{TAME}(\nu)}, \tag{4.51}$$

$$\mathrm{TERM}\,7 = (\partial_{q\zeta_2} S) \cdot (\Delta q^\sigma(t_\nu))\,(\Delta \zeta_2^\sigma(t_\nu)) \cdot \mathbb{1}_{\mathrm{TAME}(\nu)}, \tag{4.52}$$

$$\mathrm{TERM}\,8 = \frac{1}{2}\,(\partial_{\zeta_1}^2 S) \cdot (\Delta \zeta_1^\sigma(t_\nu))^2 \cdot \mathbb{1}_{\mathrm{TAME}(\nu)}, \tag{4.53}$$

$$\mathrm{TERM}\,9 = (\partial_{\zeta_1\zeta_2} S) \cdot (\Delta \zeta_1^\sigma(t_\nu))(\Delta \zeta_2^\sigma(t_\nu)) \cdot \mathbb{1}_{\mathrm{TAME}(\nu)}, \tag{4.54}$$

$$\mathrm{TERM}\,10 = \frac{1}{2}\,(\partial_{\zeta_2}^2 S) \cdot (\Delta \zeta_2^\sigma(t_\nu))^2 \cdot \mathbb{1}_{\mathrm{TAME}(\nu)}. \tag{4.55}$$

Here, again, S and its partial derivatives are evaluated at

$$(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu));$$

hence these $(\partial^\alpha S)$ are deterministic once we condition on $\mathcal{F}_\nu^\sigma$. Consequently,

$$\begin{aligned}
&\mathrm{E}[\mathrm{TERM}\,0\,|\,\mathcal{F}_\nu^\sigma] = S\cdot \mathrm{Prob}[\mathrm{TAME}(\nu)\,|\,\mathcal{F}_\nu^\sigma],\\
&\mathrm{E}[\mathrm{TERM}\,1\,|\,\mathcal{F}_\nu^\sigma] = (\partial_t S)\cdot(\Delta t_\nu)\cdot \mathrm{Prob}[\mathrm{TAME}(\nu)\,|\,\mathcal{F}_\nu^\sigma],\\
&\mathrm{E}[\mathrm{TERM}\,2|\mathcal{F}_\nu^\sigma] = (\partial_q S)\cdot \mathrm{E}[\Delta q^\sigma(t_\nu)\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\,|\,\mathcal{F}_\nu^\sigma],\\
&\mathrm{E}[\mathrm{TERM}\,3\,|\,\mathcal{F}_\nu^\sigma] = (\partial_{\zeta_1} S)\cdot \mathrm{E}[\Delta \zeta_1^\sigma(t_\nu)\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\,|\,\mathcal{F}_\nu^\sigma],\\
&\mathrm{E}[\mathrm{TERM}\,4\,|\,\mathcal{F}_\nu^\sigma] = (\partial_{\zeta_2} S)\cdot \mathrm{E}[\Delta \zeta_2^\sigma(t_\nu)\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\,|\,\mathcal{F}_\nu^\sigma],\\
&\mathrm{E}[\mathrm{TERM}\,5\,|\,\mathcal{F}_\nu^\sigma] = \frac{1}{2}\,(\partial_q^2 S)\cdot \mathrm{E}[(\Delta q^\sigma(t_\nu))^2\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\,|\,\mathcal{F}_\nu^\sigma],\\
&\mathrm{E}[\mathrm{TERM}\,6\,|\,\mathcal{F}_\nu^\sigma] = (\partial_{q\zeta_1} S)\cdot \mathrm{E}[(\Delta q^\sigma(t_\nu))(\Delta \zeta_1^\sigma(t_\nu))\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\,|\,\mathcal{F}_\nu^\sigma],\\
&\mathrm{E}[\mathrm{TERM}\,7\,|\,\mathcal{F}_\nu^\sigma] = (\partial_{q\zeta_2} S)\cdot \mathrm{E}[(\Delta q^\sigma(t_\nu))(\Delta \zeta_2^\sigma(t_\nu))\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\,|\,\mathcal{F}_\nu^\sigma],\\
&\mathrm{E}[\mathrm{TERM}\,8\,|\,\mathcal{F}_\nu^\sigma] = \frac{1}{2}\,(\partial_{\zeta_1}^2 S)\cdot \mathrm{E}[(\Delta \zeta_1^\sigma(t_\nu))^2\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\,|\,\mathcal{F}_\nu^\sigma],\\
&\mathrm{E}[\mathrm{TERM}\,9\,|\,\mathcal{F}_\nu^\sigma] = (\partial_{\zeta_1\zeta_2} S)\cdot \mathrm{E}[(\Delta \zeta_1^\sigma(t_\nu))(\Delta \zeta_2^\sigma(t_\nu))\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\,|\,\mathcal{F}_\nu^\sigma],\\
&\mathrm{E}[\mathrm{TERM}\,10\,|\,\mathcal{F}_\nu^\sigma] = \frac{1}{2}\,(\partial_{\zeta_2}^2 S)\cdot \mathrm{E}[(\Delta \zeta_2^\sigma(t_\nu))^2\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\,|\,\mathcal{F}_\nu^\sigma].
\end{aligned}$$

The expectations on the right-hand sides have been computed modulo a small error in Lemma 4.4. Applying that lemma, recalling that $|\partial^\alpha S| \le CQ^{m_0}$ for $|\alpha| \le 2$, and substituting the results into (4.44), we find that

$$\begin{aligned}
&\mathrm{E}[S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta_1^\sigma(t_{\nu+1}), \zeta_2^\sigma(t_{\nu+1}))\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\,|\,\mathcal{F}_\nu^\sigma]\\
&\quad = S + (\Delta t_\nu)\{\partial_t S + [\bar a\, q^\sigma(t_\nu) + u^\sigma(t_\nu)]\,\partial_q S + \bar a\cdot (q^\sigma(t_\nu))^2\,\partial_{\zeta_1} S\\
&\quad + (q^\sigma(t_\nu))^2\partial_{\zeta_2} S + \frac{1}{2}\,\partial_q^2 S + q^\sigma(t_\nu)\partial_{q\zeta_1} S + \frac{1}{2}\,(q^\sigma(t_\nu))^2\partial_{\zeta_1}^2 S\} + \mathrm{ERROR},
\end{aligned} \tag{4.56}$$

where S and its derivatives are evaluated at the point $(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu))$, $\bar a$ denotes $\bar a(\zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu))$, and

$$|\mathrm{ERROR}| \le CQ^{m_0}(\Delta t_\nu)^{6/5} + CQ^{m_0+4}(\Delta t_{\mathrm{MAX}})^{1/4}\Delta t_\nu \le CQ^{m_0}(\Delta t_{\mathrm{MAX}})^{1/5}\Delta t_\nu. \tag{4.57}$$

(We have used our assumption (4.27).) Since S satisfies our PDE (4.21), equation (4.56) may be rewritten in the equivalent form

$$\begin{aligned}
&\mathrm{E}[S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta_1^\sigma(t_{\nu+1}), \zeta_2^\sigma(t_{\nu+1}))\cdot \mathbb{1}_{\mathrm{TAME}(\nu)}\,|\,\mathcal{F}_\nu^\sigma]\\
&\quad = S(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu))\\
&\qquad + (\Delta t_\nu)\cdot \big\{(\partial_q S)\cdot [u^\sigma(t_\nu) - u_{\mathrm{opt}}(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu))]\\
&\qquad - [(q^\sigma(t_\nu))^2 + (u_{\mathrm{opt}}(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu)))^2]\big\}\\
&\qquad + \mathrm{ERROR},
\end{aligned} \tag{4.58}$$

with

$$|\mathrm{ERROR}| \le CQ^{m_0}(\Delta t_{\mathrm{MAX}})^{1/5}\,\Delta t_\nu. \tag{4.59}$$

Recalling that $\partial_q S = -2u_{\rm opt}$ (see (4.22)), we see that the expression in curly brackets in (4.58) is equal to

$$\{-2u_{\rm opt}\cdot[u^\sigma - u_{\rm opt}] - (q^\sigma)^2 - u_{\rm opt}^2\} = \{(u^\sigma - u_{\rm opt})^2 - (q^\sigma)^2 - (u^\sigma)^2\},$$

where we have written u^σ for $u^\sigma(t_\nu)$, and $u_{\rm opt}$ for $u_{\rm opt}(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu))$. Moreover,

$$u^\sigma - u_{\rm opt} = u^\sigma(t_\nu) - u_{\rm opt}(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu)) \equiv \textsc{Discrep}^\sigma(t_\nu)$$

(see (4.34)). Therefore, (4.58) and (4.59) are equivalent to

$$\begin{aligned}(4.60)\qquad &\mathrm{E}[S(q^\sigma(t_{\nu+1}), t_{\nu+1}, \zeta_1^\sigma(t_{\nu+1}), \zeta_2^\sigma(t_{\nu+1}))\cdot \mathbb{1}_{\textsc{tame}(\nu)}\,|\,\mathcal{F}_\nu^\sigma] \\ &\quad = S(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu)) \\ &\quad\quad + (\Delta t_\nu)\{(\textsc{Discrep}^\sigma(t_\nu))^2 - (q^\sigma(t_\nu))^2 - (u^\sigma(t_\nu))^2\} + \textsc{error},\end{aligned}$$

with

$$(4.61)\qquad |\textsc{error}| \le CQ^{m_0}(\Delta t_{\textsc{max}})^{1/5}\Delta t_\nu.$$

Here, (4.60) and (4.61) are valid wherever (4.35) holds.

We now substitute (4.60), (4.61) into (A#) and (B#), to obtain the following results, valid whenever (4.35) holds:

$$\begin{aligned}(\mathrm{A\#\#})\qquad &\mathrm{E}[\textsc{Ctg}^\sigma(t_\nu)\,|\,\mathcal{F}_\nu^\sigma] + \hat{C}\,Q^{2m_0}\,\mathrm{Prob}[\textsc{Disaster}\,|\,\mathcal{F}_\nu^\sigma] \\ &\quad \ge \{S(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu)) + (\Delta t_\nu)(\textsc{Discrep}^\sigma(t_\nu))^2 \\ &\quad\quad - [(q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2]\,\Delta t_\nu - CQ^{m_0}(\Delta t_{\textsc{max}})^{1/5}\Delta t_\nu\} \\ &\quad\quad + [(q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2]\,\Delta t_\nu\cdot \mathrm{Prob}[\textsc{tame}(\nu)\,|\,\mathcal{F}_\nu^\sigma] \\ &\quad\quad + \mathrm{E}\Big[\sum_{t_{\nu+1}\le t_\mu<\tau}(\textsc{Discrep}^\sigma(t_\mu))^2\Delta t_\mu\,|\,\mathcal{F}_\nu^\sigma\Big] \\ &\quad\quad - Q^{2m_0}(T - t_{\nu+1})\cdot(\Delta t_{\textsc{max}})^{1/20};\end{aligned}$$

and if $\sigma = \tilde{\sigma}$, then (since $\textsc{Discrep}^\sigma(t_\nu) = 0$) we have

$$\begin{aligned}(\mathrm{B\#\#})\qquad &\mathrm{E}[\textsc{Ctg}^\sigma(t_\nu)\,|\,\mathcal{F}_\nu^\sigma] \le \{S(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu)) \\ &\quad + CQ^{m_0}(\Delta t_{\textsc{max}})^{1/5}\Delta t_\nu\} + CQ^{2m_0}(\Delta t_\nu)^{20} \\ &\quad + \hat{C}\,Q^{2m_0}\,\mathrm{Prob}[\textsc{Disaster}\,|\,\mathcal{F}_\nu^\sigma] + Q^{2m_0}(T - t_{\nu+1})\cdot(\Delta t_{\textsc{max}})^{1/20}.\end{aligned}$$

In (A##), we note that

$$[(q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2]\,\mathrm{Prob}[\textsc{not tame}\,|\,\mathcal{F}_\nu^\sigma] \le CQ^2(\Delta t_\nu)^{20};$$

see hypothesis ($\star$) of this lemma. Therefore, in (A##), the terms

$$[(q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2]\,\Delta t_\nu\cdot\mathrm{Prob}[\textsc{tame}(\nu)] \quad\text{and}\quad -[(q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2]\,\Delta t_\nu$$

nearly cancel: they produce a term dominated by $CQ^2(\Delta t_\nu)^{20}$. Moreover, thanks to (4.36), we have

$$(\Delta t_\nu)(\text{DISCREP}^\sigma(t_\nu))^2 + \mathrm{E}\Big[\sum_{t_{\nu+1}\le t_\mu<\tau} (\text{DISCREP}^\sigma(t_\mu))^2 \Delta t_\mu \,\Big|\, \mathcal{F}^\sigma_\nu\Big] = \mathrm{E}\Big[\sum_{t_\nu\le t_\mu<\tau} (\text{DISCREP}^\sigma(t_\mu))^2 \Delta t_\mu \,\Big|\, \mathcal{F}^\sigma_\nu\Big].$$

(We have also used the fact that $\text{DISCREP}^\sigma(t_\nu)$ is deterministic when conditioned on $\mathcal{F}^\sigma_\nu$.)

Finally, our assumptions that $\Delta t_{\text{MAX}} \le Q^{-2000m_0}$ and $Q > C$ (C large enough) imply that

$$\begin{aligned} CQ^2(\Delta t_\nu)^{20} &+ CQ^{m_0}(\Delta t_{\text{MAX}})^{1/5}\Delta t_\nu + Q^{2m_0}(T - t_{\nu+1})\cdot(\Delta t_{\text{MAX}})^{1/20} \\ &\le Q^{2m_0}(\Delta t_{\text{MAX}})^{1/20}\cdot(t_{\nu+1} - t_\nu) + Q^{2m_0}(T - t_{\nu+1})\cdot(\Delta t_{\text{MAX}})^{1/20} \\ &= Q^{2m_0}(T - t_\nu)\cdot(\Delta t_{\text{MAX}})^{1/20}, \end{aligned}$$

and also that the sum of terms

$$CQ^{m_0}(\Delta t_{\text{MAX}})^{1/5}(\Delta t_\nu) + CQ^{2m_0}(\Delta t_\nu)^{20} + Q^{2m_0}(T - t_{\nu+1})(\Delta t_{\text{MAX}})^{1/20}$$

is at most

$$Q^{2m_0}(\Delta t_{\text{MAX}})^{1/20}\Delta t_\nu + Q^{2m_0}(T - t_{\nu+1})\cdot(\Delta t_{\text{MAX}})^{1/20} = Q^{2m_0}(T - t_\nu)\cdot(\Delta t_{\text{MAX}})^{1/20}.$$

In view of the above remarks, (A##) and (B##) imply the following results, valid whenever (4.35) holds:

$$\begin{aligned} \mathrm{E}[\text{CTG}^\sigma(t_\nu)\,|\,\mathcal{F}^\sigma_\nu] &+ \hat{C}\,Q^{2m_0}\,\text{Prob}[\text{DISASTER}\,|\,\mathcal{F}^\sigma_\nu] \\ &\ge S(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu)) + \mathrm{E}\Big[\sum_{t_\nu\le t_\mu<\tau} (\text{DISCREP}^\sigma(t_\mu))^2\Delta t_\mu \,\Big|\, \mathcal{F}^\sigma_\nu\Big] \\ &\quad - Q^{2m_0}\cdot(T - t_\nu)\cdot(\Delta t_{\text{MAX}})^{1/20}; \end{aligned}$$

and if $\sigma = \tilde{\sigma}$, then

$$\begin{aligned} \mathrm{E}[\text{CTG}^\sigma(t_\nu)\,|\,\mathcal{F}^\sigma_\nu] \le S(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu)) &+ \hat{C}\,Q^{2m_0}\,\text{Prob}[\text{DISASTER}\,|\,\mathcal{F}^\sigma_\nu] \\ &+ Q^{2m_0}\cdot(T - t_\nu)\cdot(\Delta t_{\text{MAX}})^{1/20}. \end{aligned}$$

These are precisely our desired conclusions (A) and (B). Our downward induction on ν is complete, thus proving Lemma 4.5. ■

We now draw conclusions from Lemma 4.5. Setting $\nu = 0$, we obtain the following results, comparing the allegedly optimal strategy $\tilde{\sigma}$ to the competing strategy σ.

$$\begin{aligned} &\mathrm{E}\Big[\sum_{0\le t_\mu<\tau} \{(q^\sigma(t_\mu))^2 + (u^\sigma(t_\mu))^2\}\Delta t_\mu\Big] + \hat{C}\,Q^{2m_0}\,\text{Prob}[\text{DISASTER}] \\ &\qquad \ge S(q_0, 0, 0, 0) + \mathrm{E}\Big[\sum_{0\le t_\mu<\tau} (\text{DISCREP}^\sigma(t_\mu))^2\Delta t_\mu\Big] - Q^{2m_0}(\Delta t_{\text{MAX}})^{1/20}\,T; \end{aligned} \tag{4.62}$$

and if $\sigma = \tilde{\sigma}$, then

$$\mathrm{E}\Big[\sum_{0\le t_\mu<\tau}\{(q^\sigma(t_\mu))^2+(u^\sigma(t_\mu))^2\}\,\Delta t_\mu\Big] \le S(q_0,0,0,0)+\hat{C}\,Q^{2m_0}\,\mathrm{Prob}[\textsc{Disaster}]+Q^{2m_0}(\Delta t_{\mathrm{MAX}})^{1/20}\,T. \tag{4.63}$$

From (4.28), (4.30), (4.31) and Lemma 4.1, we have

$$\mathrm{Prob}[\textsc{Disaster}] \le C\exp(-cQ). \tag{4.64}$$

Let us investigate what happens if DISASTER occurs.
Since

$$|u^\sigma(t_\mu)| \le C[|q^\sigma(t_\mu)|+1],$$

we have

$$\sum_{0\le\mu<N}\{(q^\sigma(t_\mu))^2+(u^\sigma(t_\mu))^2\}\,\Delta t_\mu \le C\big(\max_\mu |q^\sigma(t_\mu)|\big)^2+C.$$

Also,

$$\begin{aligned}|\textsc{Discrep}^\sigma(t_\mu)| &= |u^\sigma(t_\mu)-u_{\mathrm{opt}}(q^\sigma(t_\mu),t_\mu,\zeta_1^\sigma(t_\mu),\zeta_2^\sigma(t_\mu))|\\ &\le |u^\sigma(t_\mu)|+C[|q^\sigma(t_\mu)|+1]\end{aligned}$$

by our PDE assumption (see (4.26)). Hence,

$$\sum_{0\le\mu<N}(\textsc{Discrep}^\sigma(t_\mu))^2\,\Delta t_\mu \le C\big(\max_\mu |q^\sigma(t_\mu)|\big)^2+C.$$

The above remarks and Lemma 4.1 yield the estimates

$$\mathrm{E}\Big[\Big(\sum_{0\le\mu<N}\{(q^\sigma(t_\mu))^2+(u^\sigma(t_\mu))^2\}\,\Delta t_\mu\Big)^2\Big] \le C$$

and

$$\mathrm{E}\Big[\Big(\sum_{0\le\mu<N}(\textsc{Discrep}^\sigma(t_\mu))^2\,\Delta t_\mu\Big)^2\Big] \le C.$$

Consequently, Cauchy–Schwarz and (4.64) imply that

$$\mathrm{E}\Big[\Big(\sum_{0\le\mu<N}\{(q^\sigma(t_\mu))^2+(u^\sigma(t_\mu))^2\}\,\Delta t_\mu\Big)\cdot\mathbb{1}_{\textsc{Disaster}}\Big] \le C\cdot(\mathrm{Prob}[\textsc{Disaster}])^{1/2} \le C'\exp(-c'Q) \tag{4.65}$$

and

$$\mathrm{E}\Big[\Big(\sum_{0\le\mu<N}(\textsc{Discrep}^\sigma(t_\mu))^2\Delta t_\mu\Big)\cdot\mathbb{1}_{\textsc{Disaster}}\Big] \le C\cdot(\mathrm{Prob}[\textsc{Disaster}])^{1/2} \le C'\exp(-cQ). \tag{4.66}$$

So we have controlled the consequences of DISASTER.

On the other hand, if DISASTER does not occur, then $\tau = T$; see (4.31) and (4.32). Therefore,

$$\mathrm{E}\Big[\Big(\sum_{0\le\mu<N}\{(q^\sigma(t_\mu))^2+(u^\sigma(t_\mu))^2\}\,\Delta t_\mu\Big)\cdot\mathbb{1}_{\text{NON-DISASTER}}\Big] \le \mathrm{E}\Big[\sum_{0\le t_\mu<\tau}\{(q^\sigma(t_\mu))^2+(u^\sigma(t_\mu))^2\}\,\Delta t_\mu\Big] \tag{4.67}$$

and

$$\mathrm{E}\Big[\Big(\sum_{0\le\mu<N}(\text{DISCREP}^\sigma(t_\mu))^2\Delta t_\mu\Big)\cdot\mathbb{1}_{\text{NON-DISASTER}}\Big] \le \mathrm{E}\Big[\sum_{0\le t_\mu<\tau}(\text{DISCREP}^\sigma(t_\mu))^2\Delta t_\mu\Big]. \tag{4.68}$$

Also, obviously,

$$\mathrm{E}\Big[\sum_{0\le\mu<N}\{(q^\sigma(t_\mu))^2+(u^\sigma(t_\mu))^2\}\,\Delta t_\mu\Big] \ge \mathrm{E}\Big[\sum_{0\le t_\mu<\tau}\{(q^\sigma(t_\mu))^2+(u^\sigma(t_\mu))^2\}\,\Delta t_\mu\Big]. \tag{4.69}$$

Substituting (4.69), (4.68) and (4.64) into (4.62), we find that

$$\mathrm{E}\Big[\sum_{0\le\mu<N}\{(q^\sigma(t_\mu))^2+(u^\sigma(t_\mu))^2\}\,\Delta t_\mu\Big]+CQ^{2m_0}\exp(-cQ)$$
$$\ge S(q_0,0,0,0)+\mathrm{E}\Big[\Big(\sum_{0\le\mu<N}(\text{DISCREP}^\sigma(t_\mu))^2\Big)\cdot\mathbb{1}_{\text{NON-DISASTER}}\Big]-Q^{2m_0}(\Delta t_{\text{MAX}})^{1/20}T.$$

Together with (4.66), this in turn yields:

$$\begin{aligned}\mathrm{E}\Big[\sum_{0\le\mu<N}\{(q^\sigma(t_\mu))^2+(u^\sigma(t_\mu))^2\}\,\Delta t_\mu\Big] &\ge S(q_0,0,0,0)+\mathrm{E}\Big[\sum_{0\le\mu<N}(\text{DISCREP}^\sigma(t_\mu))^2\Big)\Big]\\ &\quad - CQ^{2m_0}\exp(-cQ)-Q^{2m_0}(\Delta t_{\text{MAX}})^{1/20}T.\end{aligned} \tag{4.70}$$

Similarly, suppose $\sigma = \tilde{\sigma}$. Then, substituting (4.67) and (4.64) into (4.63), we find that

$$\mathrm{E}\Big[\Big(\sum_{0\le\mu<N}\{(q^\sigma(t_\mu))^2+(u^\sigma(t_\mu))^2\}\,\Delta t_\mu\Big)\cdot\mathbb{1}_{\text{NON-DISASTER}}\Big] \le S(q_0,0,0,0)+CQ^{2m_0}\exp(-cQ)+Q^{2m_0}(\Delta t_{\text{MAX}})^{1/20}T.$$

Together with (4.65), this implies that

$$\mathrm{E}\Big[\sum_{0\le\mu<N}\{(q^\sigma(t_\mu))^2+(u^\sigma(t_\mu))^2\}\Delta t_\mu\Big] \le S(q_0,0,0,0)+CQ^{2m_0}\exp(-cQ)+Q^{2m_0}(\Delta t_{\text{MAX}})^{1/20}T. \tag{4.71}$$

We have proven (4.70) for $Q \geq C$ and $\Delta t_{\text{MAX}} \leq Q^{-2000m_0}$; see (4.27). Similarly, (4.71) holds for $Q \geq C$, $\Delta t_{\text{MAX}} \leq Q^{-2000m_0}$, $\sigma = \tilde{\sigma}$.

Now suppose $\varepsilon > 0$ is given. We pick $Q \geq C$ so large that

$$CQ^{2m_0} \exp(-cQ) < \varepsilon/2$$

in (4.70) and (4.71), and also so large that

$$Q^{2m_0}(Q^{-2000m_0})^{1/20}\, T < \varepsilon/2.$$

Then, if $\Delta t_{\text{MAX}} < Q^{-2000m_0}$, we obtain from (4.70) and (4.71) the estimates

$$\mathrm{E}\Big[\sum_{0\leq\mu<N} \{(q^\sigma(t_\mu))^2 + (u^\sigma(t_\mu))^2\}\Delta t_\mu\Big] \geq S(q_0,0,0,0) + \mathrm{E}\Big[\sum_{0\leq\mu<N} (\textsc{Discrep}^\sigma(t_\mu))^2 \Delta t_\mu\Big] - \varepsilon,$$

and if $\sigma = \tilde{\sigma}$, then

$$\mathrm{E}\Big[\sum_{0\leq\mu<N} \{(q^\sigma(t_\mu))^2 + (u^\sigma(t_\mu))^2\}\Delta t_\mu\Big] \leq S(q_0,0,0,0) + \varepsilon.$$

Thus, we have proven the following result, *modulo* our PDE assumption.

Lemma 4.6 (First Bayesian main lemma). *Let $\tilde{\sigma}$ be the* ALLEGEDLY OPTIMAL STRATEGY *for the partition $0 = t_0 < t_1 < \cdots < t_N = T$, and let $\varepsilon > 0$ be given.*

Let σ be another tame deterministic strategy for the same partition. Assume that σ satisfies

$$|u^\sigma(t_\nu)| \leq C^\sigma_{\text{TAME}}[|q^\sigma(t_\nu)| + 1].$$

Suppose that

$$\Delta t_{\text{MAX}} = \max_\nu (t_{\nu+1} - t_\nu)$$

is less than a small enough $\delta > 0$, determined by ε, C^σ_{TAME}, and the BOILERPLATE CONSTANTS. *Then*

$$\mathrm{E}\Big[\sum_{0\leq\nu<N} \{(q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2\}\Delta t_\nu\Big] \geq S(q_0,0,0,0) + \mathrm{E}\Big[\sum_{0\leq\nu<N} (\textsc{Discrep}^\sigma(t_\nu))^2 \Delta t_\nu\Big] - \varepsilon$$

and

$$\mathrm{E}\Big[\sum_{0\leq\nu<N} \{(q^{\tilde{\sigma}}(t_\nu))^2 + (u^{\tilde{\sigma}}(t_\nu))^2\}\Delta t_\nu\Big] \leq S(q_0,0,0,0) + \varepsilon,$$

where

$$\textsc{Discrep}^\sigma(t_\nu) = u^\sigma(t_\nu) - u_{\text{opt}}(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu)).$$

Corollary 4.7. *Let σ and $\tilde{\sigma}$ be as above. If*

$$\mathrm{E}\Big[\sum_{0\le\nu<N}\{(q^{\sigma}(t_\nu))^2+(u^{\sigma}(t_\nu))^2\}\,\Delta t_\nu\Big]\le \mathrm{E}\Big[\sum_{0\le\nu<N}\{(q^{\tilde{\sigma}}(t_\nu))^2+(u^{\tilde{\sigma}}(t_\nu))^2\}\,\Delta t_\nu\Big]+\varepsilon,$$

then

$$\mathrm{E}\Big[\sum_{0\le\nu<N}(\mathrm{DISCREP}^{\sigma}(t_\nu))^2\,\Delta t_\nu\Big]\le 3\varepsilon.$$

4.6. Stability of the allegedly optimal strategy

We begin by setting up the notation for this section.

- $\tilde{\sigma}$ denotes the ALLEGEDLY OPTIMAL STRATEGY.
- σ denotes some other tame deterministic strategy based on the same partition $0=t_0<t_1<\cdots<t_N=T$ as $\tilde{\sigma}$.
- q^{σ}_{ν} denotes $q^{\sigma}(t_\nu)$, and Δq^{σ}_{ν} denotes $q^{\sigma}(t_{\nu+1})-q^{\sigma}(t_\nu)$.
- $q^{\tilde{\sigma}}_{\nu}$ denotes $q^{\tilde{\sigma}}(t_\nu)$, and $\Delta q^{\tilde{\sigma}}_{\nu}$ denotes $q^{\tilde{\sigma}}(t_{\nu+1})-q^{\tilde{\sigma}}(t_\nu)$.
- $\zeta^{\sigma}_{1,\nu}$ denotes $\zeta^{\sigma}_{1}(t_\nu)$, and $\Delta\zeta^{\sigma}_{1,\nu}$ denotes $\zeta^{\sigma}_{1}(t_{\nu+1})-\zeta^{\sigma}_{1}(t_\nu)$.
- $\zeta^{\tilde{\sigma}}_{1,\nu}$ denotes $\zeta^{\tilde{\sigma}}_{1,\nu}(t_\nu)$, and $\Delta\zeta^{\tilde{\sigma}}_{1,\nu}$ denotes $\zeta^{\tilde{\sigma}}_{1}(t_{\nu+1})-\zeta^{\tilde{\sigma}}_{1}(t_\nu)$.
- $\zeta^{\sigma}_{2,\nu}$ denotes $\zeta^{\sigma}_{2}(t_\nu)$, and $\Delta\zeta^{\sigma}_{2,\nu}$ denotes $\zeta^{\sigma}_{2}(t_{\nu+1})-\zeta^{\sigma}_{2}(t_\nu)$.
- $\zeta^{\tilde{\sigma}}_{2,\nu}$ denotes $\zeta^{\tilde{\sigma}}_{2,\nu}(t_\nu)$, and $\Delta\zeta^{\tilde{\sigma}}_{2,\nu}$ denotes $\zeta^{\tilde{\sigma}}_{2}(t_{\nu+1})-\zeta^{\tilde{\sigma}}_{2}(t_\nu)$.
- u^{σ}_{ν} denotes $u^{\sigma}(t_\nu)$, and $u^{\tilde{\sigma}}_{\nu}$ denotes $u^{\tilde{\sigma}}(t_\nu)$.

We recall that

$$|u^{\sigma}(t_\nu)|\le C^{\sigma}_{\mathrm{TAME}}[|q^{\sigma}(t_\nu)|+1]\quad\text{and}\quad|u^{\tilde{\sigma}}(t_\nu)|\le C^{\mathrm{opt}}_{\mathrm{TAME}}[|q^{\tilde{\sigma}}(t_\nu)|+1].$$

In this section, c, C, C', etc., denote constants determined by $C^{\sigma}_{\mathrm{TAME}}$ and the BOILERPLATE CONSTANTS (one of which is $C^{\mathrm{opt}}_{\mathrm{TAME}}$). As usual, these symbols may denote different constants in different occurrences.

We fix $a\in[-a_{\mathrm{MAX}},+a_{\mathrm{MAX}}]$, and condition on $a_{\mathrm{TRUE}}=a$. We write $\mathrm{Prob}_a[\cdots]$ and $\mathrm{E}_a[\cdots]$ to denote the corresponding probability and expectation, respectively. We write $\mathcal{F}_\nu$ to denote the sigma algebra of events determined by the Brownian motion $(W(t))_{t\in[0,t_\nu]}$ (and by $a_{\mathrm{TRUE}}=a$).

Recall that

(4.72) $$\Delta q^{\sigma}_{\nu}=(aq^{\sigma}_{\nu}+u^{\sigma}_{\nu})(\Delta t^{*}_{\nu})+\Delta W_\nu,$$

(4.73) $$\Delta q^{\tilde{\sigma}}_{\nu}=(aq^{\tilde{\sigma}}_{\nu}+u^{\tilde{\sigma}}_{\nu})(\Delta t^{*}_{\nu})+\Delta W_\nu,$$

(4.74) $$\Delta\zeta^{\sigma}_{1,\nu}=q^{\sigma}_{\nu}(\Delta q^{\sigma}_{\nu}-u^{\sigma}_{\nu}\Delta t_\nu),$$

(4.75) $$\Delta\zeta^{\tilde{\sigma}}_{1,\nu}=q^{\tilde{\sigma}}_{\nu}(\Delta q^{\tilde{\sigma}}_{\nu}-u^{\tilde{\sigma}}_{\nu}\Delta t_\nu),$$

(4.76) $$\Delta\zeta^{\sigma}_{2,\nu}=(q^{\sigma}_{\nu})^2\,\Delta t_\nu,$$

(4.77) $$\Delta\zeta^{\tilde{\sigma}}_{2,\nu}=(q^{\tilde{\sigma}}_{\nu})^2\,\Delta t_\nu,$$

and that

$$q_0^\sigma = q_0^{\tilde\sigma} = q_0, \qquad \zeta_{1,0}^\sigma = \zeta_{1,0}^{\tilde\sigma} = \zeta_{2,0}^\sigma = \zeta_{2,0}^{\tilde\sigma} = 0. \tag{4.78}$$

In (4.72) and (4.73), $\Delta t_\nu^* = \Delta t_\nu + O((\Delta t_\nu)^2)$, and ΔW_ν is a normal random variable with mean 0 and variance $O(\Delta t_\nu)$, independent of $\mathcal{F}_\nu$.

The quantities $q_\nu^\sigma, q_\nu^{\tilde\sigma}, u_\nu^\sigma, u_\nu^{\tilde\sigma}, \zeta_{1,\nu}^\sigma, \zeta_{1,\nu}^{\tilde\sigma}, \zeta_{2,\nu}^\sigma$ and $\zeta_{2,\nu}^{\tilde\sigma}$ are deterministic once we condition on $\mathcal{F}_\nu^\sigma$.

As in Section 4.5, we define

$$\textsc{Discrep}^\sigma(t_\nu) = u^\sigma(t_\nu) - u_{\text{opt}}(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu)). \tag{4.79}$$

Our goal is to show that if

$$\mathrm{E}_a\Big[\sum_{0\le\nu<N} (\textsc{Discrep}^\sigma(t_\nu))^2\,\Delta t_\nu\Big] \text{ is small;} \tag{4.80}$$

then also

$$\mathrm{E}_a\Big[\sum_{0\le\nu<N} \{|q^\sigma(t_\nu) - q^{\tilde\sigma}(t_\nu)|^2 + |u^\sigma(t_\nu) - u^{\tilde\sigma}(t_\nu)|^2\}\,\Delta t_\nu\Big] \text{ is small.} \tag{4.81}$$

Here, (4.80) asserts that σ does something close to what $\tilde\sigma$ would do in the circumstances encountered by σ. On the other hand, (4.81) asserts that σ and $\tilde\sigma$ produce nearly equal outcomes. To show that (4.80) implies (4.81), we introduce the vector

$$\mathcal{X}_\nu = \begin{pmatrix} q_\nu^\sigma - q_\nu^{\tilde\sigma} \\ \zeta_{1,\nu}^\sigma - \zeta_{1,\nu}^{\tilde\sigma} \\ \zeta_{2,\nu}^\sigma - \zeta_{2,\nu}^{\tilde\sigma} \end{pmatrix} \equiv \begin{pmatrix} \mathcal{X}_{\nu,1} \\ \mathcal{X}_{\nu,2} \\ \mathcal{X}_{\nu,3} \end{pmatrix} \in \mathbb{R}^3. \tag{4.82}$$

Thanks to (4.78), we have

$$\mathcal{X}_0 = 0. \tag{4.83}$$

For a large enough C, we introduce a positive number Q satisfying

$$Q \ge C. \tag{4.84}$$

Under the assumption

$$|q_\nu^\sigma|, |q_\nu^{\tilde\sigma}|, |\zeta_{1,\nu}^\sigma|, |\zeta_{1,\nu}^{\tilde\sigma}|, |\zeta_{2,\nu}^\sigma|, |\zeta_{2,\nu}^{\tilde\sigma}| \le Q, \tag{4.85}$$

we will estimate

$$\Delta\mathcal{X}_\nu \equiv \begin{pmatrix} \Delta\mathcal{X}_{\nu,1} \\ \Delta\mathcal{X}_{\nu,2} \\ \Delta\mathcal{X}_{\nu,3} \end{pmatrix} \equiv \mathcal{X}_{\nu+1} - \mathcal{X}_\nu.$$

Until further notice, we fix ν and assume (4.85).

We write $G_1, G_2, \ldots$ to denote random variables satisfying

(4.86) G_i is deterministic once we condition on $\mathcal{F}_\nu$, and

(4.87) $|G_i| \le CQ^{m_0}$, with m_0 as in our PDE assumption (see (4.25)).

To estimate $\Delta \mathcal{X}_\nu$, we first apply our PDE assumption (see (4.25)) to show that

$$\begin{aligned} &u_{\text{opt}}(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu)) - u^{\tilde{\sigma}}(t_\nu) \\ &\qquad = u_{\text{opt}}(q^\sigma(t_\nu), t_\nu, \zeta_1^\sigma(t_\nu), \zeta_2^\sigma(t_\nu)) - u_{\text{opt}}(q^{\tilde{\sigma}}(t_\nu), t_\nu, \zeta_1^{\tilde{\sigma}}(t_\nu), \zeta_2^{\tilde{\sigma}}(t_\nu)) \\ &\qquad = G_1 [q_\nu^\sigma - q_\nu^{\tilde{\sigma}}] + G_2 [\zeta_{1,\nu}^\sigma - \zeta_{1,\nu}^{\tilde{\sigma}}] + G_3 [\zeta_{2,\nu}^\sigma - \zeta_{2,\nu}^{\tilde{\sigma}}] \end{aligned}$$

with G_1, G_2 and G_3 as in (4.86) and (4.87).

Consequently, (4.79) implies that

$$u^\sigma(t_\nu) - u^{\tilde{\sigma}}(t_\nu) = \text{DISCREP}^\sigma(t_\nu) + G_1 \mathcal{X}_{\nu,1} + G_2 \mathcal{X}_{\nu,2} + G_3 \mathcal{X}_{\nu,3}. \tag{4.88}$$

We subtract (4.73) from (4.72) and apply (4.88). Thus,

$$\begin{aligned} \Delta \mathcal{X}_{\nu,1} &= a \mathcal{X}_{\nu,1}(\Delta t_\nu^*) + \text{DISCREP}^\sigma(t_\nu)(\Delta t_\nu^*) \\ &\qquad + G_1 \mathcal{X}_{\nu,1}(\Delta t_\nu^*) + G_2 \mathcal{X}_{\nu,2}(\Delta t_\nu^*) + G_3 \mathcal{X}_{\nu,3}(\Delta t_\nu^*), \end{aligned}$$

which implies that

$$\begin{aligned} \Delta \mathcal{X}_{\nu,1} &= G_4 \mathcal{X}_{\nu,1}(\Delta t_\nu) + G_5 \mathcal{X}_{\nu,2}(\Delta t_\nu) \\ &\qquad + G_6 \mathcal{X}_{\nu,3}(\Delta t_\nu) + \text{DISCREP}^\sigma(t_\nu) \cdot (\Delta t_\nu^*) \end{aligned} \tag{4.89}$$

with G_4, G_5 and G_6 satisfying (4.86) and (4.87). Next, we deduce from (4.72) and (4.74) that

$$\begin{aligned} \Delta \zeta_{1,\nu}^\sigma &= q_\nu^\sigma \cdot \big([(a q_\nu^\sigma + u_\nu^\sigma)(\Delta t_\nu^*) + \Delta W_\nu] - u_\nu^\sigma \Delta t_\nu\big) \\ &= a (q_\nu^\sigma)^2 (\Delta t_\nu^*) + q_\nu^\sigma \Delta W_\nu + q_\nu^\sigma \cdot u_\nu^\sigma \cdot (\Delta t_\nu^* - \Delta t_\nu), \end{aligned}$$

and similarly,

$$\Delta \zeta_{1,\nu}^{\tilde{\sigma}} = a (q_\nu^{\tilde{\sigma}})^2 (\Delta t_\nu^*) + q_\nu^{\tilde{\sigma}} \Delta W_\nu + q_\nu^{\tilde{\sigma}} \cdot u_\nu^{\tilde{\sigma}} \cdot (\Delta t_\nu^* - \Delta t_\nu).$$

Subtracting, and recalling our assumptions (4.85), we find that

$$\Delta \mathcal{X}_{\nu,2} = G_7 \mathcal{X}_{\nu,1}(\Delta t_\nu) + \mathcal{X}_{\nu,1} \Delta W_\nu + G_8 (\Delta t_\nu)^2, \tag{4.90}$$

with G_7 and G_8 as in (4.86) and (4.87). (Here, we use also the estimate

$$|u^\sigma(t_\nu)| \le C [|q^\sigma(t_\nu)| + 1],$$

as well as the corresponding estimate for $u^{\tilde{\sigma}}$ and $q^{\tilde{\sigma}}$.)

Again recalling (4.85), we see from (4.76) and (4.77) that

$$\Delta \mathcal{X}_{\nu,3} = G_9 \mathcal{X}_{\nu,1}(\Delta t_\nu), \tag{4.91}$$

with G_9 as in (4.86) and (4.87).

Equations (4.89), (4.90) and (4.91) tell us that

$$\Delta \mathcal{X}_\nu = G \mathcal{X}_\nu(\Delta t_\nu) + H \mathcal{X}_\nu(\Delta W_\nu) + F_\nu, \tag{4.92}$$

where

- the entries of the matrix G satisfy (4.86) and (4.87),

- H is the constant matrix
$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
- and the vector F_ν satisfies
$$|F_\nu| \leq C|\textsc{Discrep}^\sigma(t_\nu)| \cdot (\Delta t_\nu) + CQ^{m_0}(\Delta t_\nu)^2.$$

We now estimate

$$\begin{aligned} \mathrm{E}_a[|\mathcal{X}_{\nu+1}|^2 - |\mathcal{X}_\nu|^2 \,|\, \mathcal{F}_\nu] &= 2\mathrm{E}_a[(\Delta\mathcal{X}_\nu)\cdot\mathcal{X}_\nu \,|\, \mathcal{F}_\nu] + \mathrm{E}_a[|\Delta\mathcal{X}_\nu|^2 \,|\, \mathcal{F}_\nu] \\ &\equiv 2\cdot\textsc{term}\,1 + \textsc{term}\,2. \end{aligned} \tag{4.93}$$

Recall that G and $\mathcal{X}_\nu$ are deterministic once we condition on $\mathcal{F}_\nu$, while ΔW_ν is independent of $\mathcal{F}_\nu$, with mean 0 and variance $\leq C(\Delta t_\nu)$. Hence, from (4.92) we have the following estimates:

$$\begin{aligned} \textsc{term}\,1 &= \mathrm{E}_a[\mathcal{X}_\nu \cdot (\Delta\mathcal{X}_\nu) \,|\, \mathcal{F}_\nu] \\ &\leq CQ^{m_0}(\Delta t_\nu)|\mathcal{X}_\nu|^2 + C|\textsc{Discrep}^\sigma(t_\nu)|\cdot|\mathcal{X}_\nu|(\Delta t_\nu) + C|\mathcal{X}_\nu|\,Q^{m_0}(\Delta t_\nu)^2, \\ \textsc{term}\,2 &= \mathrm{E}_a[|\Delta\mathcal{X}_\nu|^2 \,|\, \mathcal{F}_\nu] \\ &\leq C|G\mathcal{X}_\nu(\Delta t_\nu)|^2 + C|H\mathcal{X}_\nu|^2\,\mathrm{E}_a[|\Delta W_\nu|^2] \\ &\quad + C(\textsc{Discrep}^\sigma(t_\nu))^2(\Delta t_\nu)^2 + CQ^{2m_0}(\Delta t_\nu)^4. \end{aligned}$$

Our assumption (4.85) implies that $|\mathcal{X}_\nu| \leq CQ$, hence the above estimates imply that

$$\textsc{term}\,1 \leq CQ^{m_0}(\Delta t_\nu)|\mathcal{X}_\nu|^2 + C(\textsc{Discrep}^\sigma(t_\nu))^2(\Delta t_\nu) + CQ^{m_0+1}(\Delta t_\nu)^2 \tag{4.94}$$

and

$$\begin{aligned} \textsc{term}\,2 \leq\ & CQ^{2m_0+2}(\Delta t_\nu)^2 + C|\mathcal{X}_\nu|^2(\Delta t_\nu) \\ & + C(\textsc{Discrep}^\sigma(t_\nu))^2(\Delta t_\nu)^2 + CQ^{2m_0}(\Delta t_\nu)^4. \end{aligned} \tag{4.95}$$

Putting (4.94) and (4.95) into (4.93), we learn that

$$\begin{aligned} \mathrm{E}_a[|\mathcal{X}_{\nu+1}|^2 \,|\, \mathcal{F}_\nu] \leq\ & (1 + CQ^{m_0}(\Delta t_\nu))|\mathcal{X}_\nu|^2 + C(\textsc{Discrep}^\sigma(t_\nu))^2\,\Delta t_\nu \\ & + CQ^{2m_0+2}(\Delta t_\nu)^2. \end{aligned} \tag{4.96}$$

We have proven (4.96) under the assumption (4.85). We now drop assumption (4.85), and let $\mathcal{E}_\nu$ denote the event

$$\{|q^\sigma_\mu|, |q^{\tilde\sigma}_\mu|, |\zeta^\sigma_{1,\mu}|, |\zeta^{\tilde\sigma}_{1,\mu}|, |\zeta^\sigma_{2,\mu}|, |\zeta^{\tilde\sigma}_{2,\mu}| \leq Q \ \text{ for all } \mu \leq \nu\}.$$

Note that $\mathbb{1}_{\mathcal{E}_{\nu+1}} \leq \mathbb{1}_{\mathcal{E}_\nu}$, and that (4.85) holds whenever $\mathcal{E}_\nu$ occurs. Moreover, $\mathcal{E}_\nu$ is deterministic once we condition on $\mathcal{F}_\nu$. Therefore, from (4.96) we deduce that

$$\begin{aligned} &\mathrm{E}_a[|\mathcal{X}_{\nu+1}|^2\mathbb{1}_{\mathcal{E}_{\nu+1}} \,|\, \mathcal{F}_\nu] \leq \mathrm{E}_a[|\mathcal{X}_{\nu+1}|^2 \cdot \mathbb{1}_{\mathcal{E}_\nu} \,|\, \mathcal{F}_\nu] \\ &\quad \leq (1 + CQ^{m_0}(\Delta t_\nu))|\mathcal{X}_\nu|^2\mathbb{1}_{\mathcal{E}_\nu} + C(\textsc{Discrep}^\sigma(t_\nu))^2\,\Delta t_\nu + CQ^{2m_0+2}(\Delta t_\nu)^2. \end{aligned} \tag{4.97}$$

We now cease conditioning on $\mathscr{F}_\nu$, and condition merely on $a_{\text{TRUE}} = a$. From (4.97) we learn that

$$\begin{aligned} \mathrm{E}_a[|\mathcal{X}_{\nu+1}|^2 \mathbb{1}_{\mathcal{E}_{\nu+1}}] \leq (1 + CQ^{m_0}(\Delta t_\nu))\, \mathrm{E}_a[|\mathcal{X}_\nu|^2 \mathbb{1}_{\mathcal{E}_\nu}] \\ + C\mathrm{E}_a[(\mathrm{DISCREP}^\sigma(t_\nu))^2](\Delta t_\nu) + CQ^{2m_0+2}(\Delta t_\nu)^2. \end{aligned} \tag{4.98}$$

Recall that $\mathcal{X}_0 = 0$.

We impose the smallness assumption

$$(\Delta t_{\text{MAX}})^{1/2} \leq Q^{-(2m_0+2)}. \tag{4.99}$$

Then (4.98) implies that

$$\mathrm{E}_a[|\mathcal{X}_\nu|^2 \cdot \mathbb{1}_{\mathcal{E}_\nu}] \leq C \exp(CQ^{m_0} t_\nu) \left\{ \mathrm{E}_a\Big[\sum_{0 \leq \mu < \nu} (\mathrm{DISCREP}^\sigma(t_\mu))^2 \Delta t_\mu \Big] + C(\Delta t_{\text{MAX}})^{1/2} \right\}$$

for $0 \leq \nu \leq N$. Since $\mathbb{1}_{\mathcal{E}_N} \leq \mathbb{1}_{\mathcal{E}_\nu}$, it follows that

$$\begin{aligned} &\mathrm{E}_a[|\mathcal{X}_\nu|^2 \mathbb{1}_{\mathcal{E}_N}] \\ &\quad \leq C \exp(CQ^{m_0}) \left\{ \mathrm{E}_a\Big[\sum_{0 \leq \mu < N} (\mathrm{DISCREP}^\sigma(t_\mu))^2 \Delta t_\mu \Big] + (\Delta t_{\text{MAX}})^{1/2} \right\}. \end{aligned} \tag{4.100}$$

In particular, since $q_\nu^\sigma - q_\nu^{\tilde{\sigma}}$ is the first component of $\mathcal{X}_\nu$, we have

$$\begin{aligned} &\mathrm{E}_a[|q_\nu^\sigma - q_\nu^{\tilde{\sigma}}|^2 \cdot \mathbb{1}_{\mathcal{E}_N}] \\ &\quad \leq C \exp(CQ^{m_0}) \left\{ \mathrm{E}_a\Big[\sum_{0 \leq \mu < N} (\mathrm{DISCREP}^\sigma(t_\mu))^2 \Delta t_\mu \Big] + (\Delta t_{\text{MAX}})^{1/2} \right\}. \end{aligned} \tag{4.101}$$

Moreover, (4.88) and (4.100) together yield

$$\begin{aligned} &\mathrm{E}_a[|u_\nu^\sigma - u_\nu^{\tilde{\sigma}}|^2 \cdot \mathbb{1}_{\mathcal{E}_N}] \\ &\quad \leq CQ^{2m_0} \exp(CQ^{m_0}) \left\{ \mathrm{E}_a\Big[\sum_{0 \leq \mu < N} (\mathrm{DISCREP}^\sigma(t_\mu))^2 \Delta t_\mu \Big] + (\Delta t_{\text{MAX}})^{1/2} \right\} \\ &\qquad + C\mathrm{E}_a[(\mathrm{DISCREP}^\sigma(t_\nu))^2]. \end{aligned} \tag{4.102}$$

Summing (4.101) and (4.102) against Δt_ν, we find that

$$\begin{aligned} &\mathrm{E}_a\Big[\sum_{0 \leq \nu < N} \{|q_\nu^\sigma - q_\nu^{\tilde{\sigma}}|^2 + |u_\nu^\sigma - u_\nu^{\tilde{\sigma}}|^2\} \Delta t_\nu \cdot \mathbb{1}_{\mathcal{E}_N} \Big] \\ &\quad \leq CQ^{2m_0} \exp(CQ^{m_0}) \left\{ \mathrm{E}_a\Big[\sum_{0 \leq \mu < N} (\mathrm{DISCREP}^\sigma(t_\mu))^2 \Delta t_\mu \Big] + (\Delta t_{\text{MAX}})^{1/2} \right\}. \end{aligned} \tag{4.103}$$

We now turn to the case in which $\mathcal{E}_N$ does not occur. Recall that $\mathcal{E}_N$ fails precisely when, for some ν, we have

$$\max\{|q_\nu^\sigma|, |q_\nu^{\tilde{\sigma}}|, |\zeta_{1,\nu}^\sigma|, |\zeta_{1,\nu}^{\tilde{\sigma}}|, |\zeta_{2,\nu}^\sigma|, |\zeta_{2,\nu}^{\tilde{\sigma}}|\} > Q.$$

Thanks to Lemma 3.2, applied to the strategies σ and $\tilde{\sigma}$, we have

$$\mathrm{Prob}_a[\mathcal{E}_N \text{ fails}] \leq C \exp(-cQ) \tag{4.104}$$

and also that

$$\mathrm{E}_a[(\max_\nu\{|q_\nu^\sigma| + |q_\nu^{\tilde\sigma}| + |u_\nu^\sigma| + |u_\nu^{\tilde\sigma}|\})^4] \leq C,$$

hence

$$\mathrm{E}_a\Big[\Big(\sum_{0\leq\nu<N} \{|q_\nu^\sigma - q_\nu^{\tilde\sigma}|^2 + |u_\nu^\sigma - u_\nu^{\tilde\sigma}|^2\}\,\Delta t_\nu\Big)^2\Big] \leq C. \tag{4.105}$$

From (4.104), (4.105), and Cauchy–Schwarz, we obtain the estimate

$$\mathrm{E}_a\Big[\sum_{0\leq\nu<N} \{|q_\nu^\sigma - q_\nu^{\tilde\sigma}|^2 + |u_\nu^\sigma - u_\nu^{\tilde\sigma}|^2\}\,\Delta t_\nu \cdot \mathbb{1}_{\mathcal{E}_N \text{ fails}}\Big] \leq C\exp(-cQ). \tag{4.106}$$

Finally, combining (4.103) and (4.106), we find that

$$\begin{aligned}
&\mathrm{E}_a\Big[\sum_{0\leq\nu<N} \{|q_\nu^\sigma - q_\nu^{\tilde\sigma}|^2 + |u_\nu^\sigma + u_\nu^{\tilde\sigma}|^2\}\,\Delta t_\nu\Big] \\
&\leq CQ^{2m_0}\exp(CQ^{m_0})\Big\{\mathrm{E}_a\Big[\sum_{0\leq\mu<N} (\mathrm{DISCREP}^\sigma(t_\mu))^2\Delta t_\mu\Big] + (\Delta t_{\mathrm{MAX}})^{1/2}\Big\} \\
&\quad + C\exp(-cQ).
\end{aligned} \tag{4.107}$$

We have proven (4.107) under the assumption (4.99).

Now let $\varepsilon > 0$ be given.

We take Q in (4.107) large enough so that $C\exp(-cQ) < \varepsilon/3$.

Having picked Q, we strengthen our smallness assumption (4.99) by demanding that

$$CQ^{2m_0}\exp(CQ^{m_0})\cdot(\Delta t_{\mathrm{MAX}})^{1/2} \leq \frac{\varepsilon}{3}.$$

If also

$$\mathrm{E}_a\Big[\sum_{0\leq\mu<N} (\mathrm{DISCREP}^\sigma(t_\mu))^2\,\Delta t_\mu\Big] \leq \frac{\varepsilon}{3}\,[CQ^{2m_0}\exp(CQ^{m_0})]^{-1},$$

then (4.107) implies that

$$\mathrm{E}_a\Big[\sum_{0\leq\nu<N} \{|q_\nu^\sigma - q_\nu^{\tilde\sigma}|^2 + |u_\nu^\sigma - u_\nu^{\tilde\sigma}|^2\}\,\Delta t_\nu\Big] < \varepsilon.$$

Thus, we have proven the following result.

Lemma 4.8 (Stability lemma). *Let $\varepsilon > 0$, let $\tilde\sigma$ be the* ALLEGEDLY OPTIMAL STRATEGY *for the partition $0 = t_0 < t_1 < \cdots < t_N = T$, and let σ be another tame deterministic strategy for that same partition. Assume that σ satisfies*

$$|u^\sigma(t_\nu)| \leq C^\sigma_{\mathrm{TAME}}[|q^\sigma(t_\nu)| + 1].$$

Fix $a \in [-a_{\mathrm{MAX}}, +a_{\mathrm{MAX}}]$. Suppose that

$$\Delta t_{\mathrm{MAX}} = \max_\nu(t_{\nu+1} - t_\nu) < \delta$$

and

$$\mathrm{E}_a\Big[\sum_{0\le\nu<N}(u^\sigma(t_\nu)-u_{\mathrm{opt}}(q^\sigma(t_\nu),t_\nu,\zeta_1^\sigma(t_\nu),\zeta_2^\sigma(t_\nu))^2\,\Delta t_\nu\Big]<\delta$$

for a small enough $\delta>0$ *determined by* ε, C^σ_{TAME}, *and the* BOILERPLATE CONSTANTS. *Then*

$$\mathrm{E}_a\Big[\sum_{0\le\nu<N}\big\{|u^\sigma(t_\nu)-u^{\tilde\sigma}(t_\nu)|^2+|q^\sigma(t_\nu)-q^{\tilde\sigma}(t_\nu)|^2\big\}\,\Delta t_\nu\Big]<\varepsilon.$$

4.7. The second Bayesian main lemma

Lemma 4.9 (The second Bayesian main lemma). *Let* $\varepsilon>0$, *let* $\tilde\sigma$ *be the* ALLEGEDLY OPTIMAL STRATEGY *for a partition of* $[0,T]$, *and let* σ *be another deterministic tame strategy for that same partition. Assume that* σ *satisfies*

$$|u^\sigma(t_\nu)|\le C^\sigma_{\text{TAME}}[|q^\sigma(t_\nu)|+1].$$

Suppose that

$$\Delta t_{\text{MAX}}<\delta$$

and that

$$\begin{aligned}\mathrm{E}\Big[\sum_{0\le\nu<N}\{(q^\sigma(t_\nu))^2+(u^\sigma(t_\nu))^2\}\,\Delta t_\nu\Big]\\ \le \mathrm{E}\Big[\sum_{0\le\nu<N}\{(q^{\tilde\sigma}(t_\nu))^2+(u^{\tilde\sigma}(t_\nu))^2\}\,\Delta t_\nu\Big]+\delta,\end{aligned}\tag{4.108}$$

for a small enough $\delta>0$ *determined by* ε, C^σ_{TAME}, *and the* BOILERPLATE CONSTANTS.

Then, for every $a\in[-a_{\text{MAX}},+a_{\text{MAX}}]$, *we have*

$$\mathrm{E}_a\Big[\sum_{0\le\nu<N}\{|q^\sigma(t_\nu)-q^{\tilde\sigma}(t_\nu)|^2+|u^\sigma(t_\nu)-u^{\tilde\sigma}(t_\nu)|^2\}\,\Delta t_\nu\Big]<\varepsilon.$$

Proof. First, we notice that we can replace the BOILERPLATE CONSTANT $C^{\text{opt}}_{\text{TAME}}$ by the constant $\max\{C^\sigma_{\text{TAME}},C^{\text{opt}}_{\text{TAME}}\}$.

Observe that

$$\text{DISCREP}^\sigma(t_\nu):=u^\sigma(t_\nu)-u_{\mathrm{opt}}(q^\sigma(t_\nu),t_\nu,\zeta_1^\sigma(t_\nu),\zeta_2^\sigma(t_\nu))$$

satisfies

$$|\text{DISCREP}^\sigma(t_\nu)|\le|u^\sigma(t_\nu)|+C[|q^\sigma(t_\nu)|+1],$$

hence

$$\sum_{0\le\nu<N}(\text{DISCREP}^\sigma(t_\nu))^2\,\Delta t_\nu\le C'\max_{0\le\nu<N}\{|q^\sigma(t_\nu)|^2+|u^\sigma(t_\nu)|^2+1\}.$$

Lemma 3.2 therefore yields the estimate

$$\mathrm{E}_a\Big[\Big\{\sum_{0\le\nu<N}(\text{DISCREP}^\sigma(t_\nu))^2\,\Delta t_\nu\Big\}^p\Big]\le C_p\tag{4.109}$$

for any $p\ge1$ and any $a\in[-a_{\text{MAX}},a_{\text{MAX}}]$.

Now let $\varepsilon > 0$ be given. We pick a small enough $\varepsilon_1 > 0$, depending on ε, then we pick a large enough $Q > 1$, depending on ε_1, next we pick a small enough $\varepsilon_2 > 0$, depending on Q, and finally we pick a small enough $\delta > 0$ depending on ε_2. We then argue as follows.

Suppose that $\Delta t_{\text{MAX}} < \delta$, and suppose (4.108) holds. Since $\delta > 0$ has been picked small enough, depending on ε_2, Corollary 4.7 tells us that

$$\mathrm{E}\Big[\sum_{0\le\nu<N} (\textsc{Discrep}^{\sigma}(t_\nu))^2\,\Delta t_\nu\Big] < 3\varepsilon_2. \tag{4.110}$$

For any random variable X, the expected value $\mathrm{E}[X]$ is an average of $\mathrm{E}_a[X]$ over $a \in [-a_{\text{MAX}}, +a_{\text{MAX}}]$ with respect to our given prior probability distribution for a_{TRUE}. Therefore, (4.110) implies that for some $a_1 \in [-a_{\text{MAX}}, a_{\text{MAX}}]$, we have

$$\mathrm{E}_{a_1}\Big[\sum_{0\le\nu<N} (\textsc{Discrep}^{\sigma}(t_\nu))^2\,\Delta t_\nu\Big] < 4\varepsilon_2.$$

Consequently, for any $a \in [-a_{\text{MAX}}, +a_{\text{MAX}}]$, Lemma 3.6 and estimate (4.109) with $p = 2$ give

$$\mathrm{E}_{a}\Big[\sum_{0\le\nu<N} (\textsc{Discrep}^{\sigma}(t_\nu))^2\,\Delta t_\nu\Big] \le \exp(CQ^2)\cdot 4\varepsilon_2 + C\exp(-cQ^2). \tag{4.111}$$

Since ε_2 has been picked small enough depending on Q, while Q has been picked large enough depending on ε_1, the right-hand side of (4.111) is less than $C'\exp(-cQ^2) < \varepsilon_1$. Thus, for any $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$, we have

$$\mathrm{E}_{a}\Big[\sum_{0\le\nu<N} (\textsc{Discrep}^{\sigma}(t_\nu))^2\,\Delta t_\nu\Big] < \varepsilon_1. \tag{4.112}$$

Since ε_1 has been picked small enough depending on ε, estimate (4.112) and Lemma 4.8 imply that

$$\mathrm{E}_{a}\Big[\sum_{0\le\nu<N} \{|u^{\sigma}(t_\nu) - u^{\tilde\sigma}(t_\nu)|^2 + |q^{\sigma}(t_\nu) - q^{\tilde\sigma}(t_\nu)|^2\}\,\Delta t_\nu\Big] < \varepsilon$$

for all $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$, completing the proof of the lemma. ■

Corollary 4.10. *Under the assumptions of Lemma 4.9 we have*

$$\Big|\mathrm{E}_{a}\Big[\sum_{0\le\nu<N} \{(q^{\sigma}(t_\nu))^2 + (u^{\sigma}(t_\nu))^2\}\,\Delta t_\nu\Big] - \mathrm{E}_{a}\Big[\sum_{0\le\nu<N} \{(q^{\tilde\sigma}(t_\nu))^2 + (u^{\tilde\sigma}(t_\nu))^2\}\,\Delta t_\nu\Big]\Big| \le C\varepsilon^{1/2}$$

for each $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$.

Proof. The corollary follows from Lemma 4.9, together with Minkowski's inequality and the estimate

$$\mathrm{E}_{a}\Big[\sum_{0\le\nu<N} \{(q^{\tilde\sigma}(t_\nu))^2 + (u^{\tilde\sigma}(t_\nu))^2\}\,\Delta t_\nu\Big] \le C,$$

which in turn follows from Lemma 3.2. ■

4.8. Allowing for dependence on coin flips

Let $\tilde{\sigma}$ be the ALLEGEDLY OPTIMAL STRATEGY associated to a partition of $[0, T]$, and let σ be a tame strategy associated to the same partition. In this section, we allow σ to depend on the coin flips $\vec{\xi}$. For fixed $\vec{\eta} \in \{0, 1\}^{\mathbb{N}}$, we write $\sigma_{\vec{\eta}}$ for the strategy prescribed by σ in case $\vec{\xi} = \vec{\eta}$. We write $\mathrm{Prob}_B[\dots]$ and $\mathrm{E}_B[\dots]$ ("B" for "Bernoulli") to denote probability and expectation with respect to the natural (product) probability measure on $\{0, 1\}^{\mathbb{N}}$, in which each ξ_ν is equal to 0 with probability $1/2$.

Our goal here is to extend Lemma 4.9 to the case of the $\vec{\xi}$-dependent strategy σ.

To do so, we denote

$$\textsc{Cost}_D(\sigma) = \sum_{0 \le \nu < N} \{(q^\sigma(t_\nu))^2 + (u^\sigma(t_\nu))^2\} \Delta t_\nu,$$

and similarly for $\textsc{Cost}_D(\tilde{\sigma})$ and $\textsc{Cost}_D(\sigma_{\vec{\eta}})$.

Because σ is tame, we have

$$|u^\sigma(t_\nu)| \le C^\sigma_{\text{TAME}} \cdot [|q^\sigma(t_\nu)| + 1] \tag{4.113}$$

for a constant C^σ_{TAME}.

Let $\varepsilon > 0$ be given, and let $\delta > 0$ be less than a small enough positive number determined by ε, C^σ_{TAME}, and the BOILERPLATE CONSTANTS.

Suppose that $\Delta t_{\text{MAX}} < \delta$, and that

$$\mathrm{E}[\textsc{Cost}_D(\sigma)] \le \mathrm{E}[\textsc{Cost}_D(\tilde{\sigma})] + \delta. \tag{4.114}$$

We want to show that

$$\mathrm{E}_a\Big[\sum_{0 \le \nu < N} \{|q^\sigma(t_\nu) - q^{\tilde{\sigma}}(t_\nu)|^2 + |u^\sigma(t_\nu) - u^{\tilde{\sigma}}(t_\nu)|^2\} \Delta t_\nu\Big] < \varepsilon$$

for all $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$. To see this, we first pick $\hat{\delta}$ small enough, determined by ε, C^σ_{TAME}, and the BOILERPLATE CONSTANTS; then we pick δ small enough, determined by $\hat{\delta}$, C^σ_{TAME}, and the BOILERPLATE CONSTANTS.

Lemma 4.6, with $\hat{\delta}^2$ in place of ε, shows that

$$\mathrm{E}[\textsc{Cost}_D(\sigma_{\vec{\eta}})] \ge \mathrm{E}[\textsc{Cost}_D(\tilde{\sigma})] - \hat{\delta}^2 \quad \text{for all } \vec{\eta} \in \{0, 1\}^{\mathbb{N}}. \tag{4.115}$$

On the other hand, (4.114) shows that

$$\mathrm{E}_B[\{\mathrm{E}[\textsc{Cost}_D(\sigma_{\vec{\eta}})] - \mathrm{E}[\textsc{Cost}_D(\tilde{\sigma})] + \hat{\delta}^2\}] \le \delta + \hat{\delta}^2 < 2\hat{\delta}^2, \tag{4.116}$$

since δ is less than a small enough constant depending on $\hat{\delta}$. The quantity in curly brackets in (4.116) is nonnegative, thanks to (4.115). Therefore, if we set

$$\textsc{Goodflips} = \big\{\vec{\eta} \in \{0, 1\}^{\mathbb{N}} : \mathrm{E}[\textsc{Cost}_D(\sigma_{\vec{\eta}})] \le \mathrm{E}[\textsc{Cost}_D(\tilde{\sigma})] + \hat{\delta}\big\}, \tag{4.117}$$

$$\textsc{Badflips} = \big\{\vec{\eta} \in \{0, 1\}^{\mathbb{N}} : \mathrm{E}[\textsc{Cost}_D(\sigma_{\vec{\eta}})] > \mathrm{E}[\textsc{Cost}_D(\tilde{\sigma})] + \hat{\delta}\big\}, \tag{4.118}$$

then

$$\mathrm{Prob}_B[\textsc{Badflips}] < 10\hat{\delta}. \tag{4.119}$$

Moreover, since $\sigma_{\vec{\eta}}$ and $\tilde{\sigma}$ are tame for each $\vec{\eta}$, Lemma 3.2 implies that for each $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$, we have

$$\mathrm{E}_a\Big[\Big\{\sum_{0\le\nu<N}[|q^{\sigma_{\vec{\eta}}}(t_\nu)-q^{\tilde{\sigma}}(t_\nu)|^2+|u^{\sigma_{\vec{\eta}}}(t_\nu)-u^{\tilde{\sigma}}(t_\nu)|^2]\,\Delta t_\nu\Big\}^2\Big]\le C,$$

where we allow constants C to depend on C^{σ}_{TAME} in (4.113). Consequently,

$$\mathrm{E}_a\Big[\Big\{\sum_{0\le\nu<N}[|q^{\sigma}(t_\nu)-q^{\tilde{\sigma}}(t_\nu)|^2+|u^{\sigma}(t_\nu)-u^{\tilde{\sigma}}(t_\nu)|^2]\,\Delta t_\nu\Big\}^2\Big]\le C.$$

Together with (4.119) and Cauchy–Schwarz, this implies that

$$\mathrm{E}_a\Big[\Big\{\sum_{0\le\nu<N}[|q^{\sigma}(t_\nu)-q^{\tilde{\sigma}}(t_\nu)|^2+|u^{\sigma}(t_\nu)-u^{\tilde{\sigma}}(t_\nu)|^2]\,\Delta t_\nu\Big\}\cdot\mathbb{1}_{\text{BADFLIPS}}(\vec{\xi})\Big]\le C\hat{\delta}^{1/2}. \tag{4.120}$$

On the other hand, if $\vec{\eta}\in$ GOODFLIPS, then Lemma 4.9 tells us that

$$\mathrm{E}_a\Big[\sum_{0\le\nu<N}\{|q^{\sigma_{\vec{\eta}}}(t_\nu)-q^{\tilde{\sigma}}(t_\nu)|^2+|u^{\sigma_{\vec{\eta}}}(t_\nu)-u^{\tilde{\sigma}}(t_\nu)|^2\}\,\Delta t_\nu\Big]<\frac{\varepsilon}{2}.$$

Consequently,

$$\mathrm{E}_a\Big[\Big\{\sum_{0\le\nu<N}[|q^{\sigma}(t_\nu)-q^{\tilde{\sigma}}(t_\nu)|^2+|u^{\sigma}(t_\nu)-u^{\tilde{\sigma}}(t_\nu)|^2]\,\Delta t_\nu\Big\}\cdot\mathbb{1}_{\text{GOODFLIPS}}(\vec{\xi})\Big]\le\frac{\varepsilon}{2}. \tag{4.121}$$

We learn from (4.120) and (4.121) that

$$\mathrm{E}_a\Big[\sum_{0\le\nu<N}\{|q^{\sigma}(t_\nu)-q^{\tilde{\sigma}}(t_\nu)|^2+|u^{\sigma}(t_\nu)-u^{\tilde{\sigma}}(t_\nu)|^2\}\,\Delta t_\nu\Big]<\varepsilon. \tag{4.122}$$

Thus, we have shown that (4.113), (4.114) and $\Delta t_{\text{MAX}}<\delta$ imply (4.122).

That is, Lemma 4.9 holds without the assumption that σ is deterministic. From now on, when we apply that lemma, we need not check that σ is deterministic.

4.9. Reformulating the main Bayesian results

Let σ be a tame strategy. Recall, from Section 3.9, that $q^{\sigma}_{C}(t)=q^{\sigma}(t)$ and $q^{\sigma}_{D}(t)=q^{\sigma}(t_\nu)$ for $t\in[t_\nu,t_{\nu+1})$, $0\le\nu<N$. Also, $u^{\sigma}(t)=u^{\sigma}(t_\nu)$ for $t\in[t_\nu,t_{\nu+1})$, $0\le\nu<N$. Therefore,

$$\sum_{0\le\nu<N}\{|u^{\sigma}(t_\nu)-u^{\tilde{\sigma}}(t_\nu)|^2+|q^{\sigma}(t_\nu)-q^{\tilde{\sigma}}(t_\nu)|^2\}\,\Delta t_\nu = \int_0^T\{|u^{\sigma}(t)-u^{\tilde{\sigma}}(t)|^2+|q^{\sigma}_D(t)-q^{\tilde{\sigma}}_D(t)|^2\}\,dt,$$

where $\tilde{\sigma}$ is the ALLEGEDLY OPTIMAL STRATEGY for some Bayesian prior. Similarly,

$$\sum_{0\le\nu<N}\{|u^{\sigma}(t_\nu)|^2+|q^{\sigma}(t_\nu)|^2\}\,\Delta t_\nu=\int_0^T\{|u^{\sigma}(t)|^2+|q^{\sigma}_D(t)|^2\}\,dt,$$

and the analogous formula holds for $\tilde{\sigma}$.

From Section 3.9, we have

$$\mathrm{E}_a\left[\max_{t\in[0,T]} |q^\sigma(t) - q^\sigma_D(t)|^m\right] \le C_m(\Delta t_{\mathrm{MAX}})^{m/4} \quad \text{for any } m \ge 1$$

and any $a \in [-a_{\mathrm{MAX}}, a_{\mathrm{MAX}}]$; the analogous estimate holds for $\tilde{\sigma}$ in place of σ.

In view of the above remarks, we can reformulate our main previous results, replacing $q^\sigma_D(t)$ by $q^\sigma(t)$. We define

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

We combine the above discussion with Lemmas 3.4, 3.7, 4.6, and 4.9 to deduce the following.

Theorem 4.11 (Main theorem on Bayesian strategies). *Let σ be a tame strategy, satisfying*

$$|u^\sigma(t_\nu)| < C^\sigma_{\mathrm{TAME}}[|q^\sigma(t_\nu)| + 1].$$

Fix a Bayesian prior $d\mathrm{Prob}(a)$ on $[-a_{\mathrm{MAX}}, +a_{\mathrm{MAX}}]$, and suppose our PDE assumption holds for the PDE arising from that prior. Let $\tilde{\sigma}$ denote the ALLEGEDLY OPTIMAL STRATEGY *for the same partition of $[0, T]$ used to define σ.*

Then given $\varepsilon > 0$, there exists $\delta > 0$ determined by ε, together with the BOILERPLATE CONSTANTS *and the constant C^σ_{TAME}, such that the following holds. Suppose $\Delta t_{\mathrm{MAX}} < \delta$. Then:*

(1) *There holds*

$$|\mathrm{E}[\mathrm{Cost}(\tilde{\sigma})] - S(q_0, 0, 0, 0)| < \varepsilon,$$

where S is our PDE solution.

(2) *If*

$$\mathrm{E}[\mathrm{Cost}(\sigma)] < \mathrm{E}[\mathrm{Cost}(\tilde{\sigma})] + \delta,$$

then for any $a \in [-a_{\mathrm{MAX}}, +a_{\mathrm{MAX}}]$, we have

$$\mathrm{E}_a\left[\int_0^T \{|u^\sigma(t) - u^{\tilde{\sigma}}(t)|^2 + |q^\sigma(t) - q^{\tilde{\sigma}}(t)|^2\}\, dt\right] < \varepsilon.$$

(3) *There exists an analytic function $I(a)$, defined on $\mathcal{R} \equiv (-a_{\mathrm{MAX}}, +a_{\mathrm{MAX}}) \times (-\hat{c}, \hat{c})$ for some $\hat{c}$, such that $|I(a)| \le C$ on $\mathcal{R}$, and*

$$|I(a) - \mathrm{E}_a[\mathrm{Cost}(\sigma)]| < \varepsilon \quad \textit{for } a \in (-a_{\mathrm{MAX}}, +a_{\mathrm{MAX}}).$$

(4) *If a exceeds a large enough constant C, then*

$$\mathrm{E}_a[\mathrm{Cost}(\sigma)] > cT^2 \exp(caT).$$

Note that strictly speaking, we proved (3) and (4) for deterministic strategies, so they hold for σ once we condition on $\vec{\xi} = \vec{\eta}$ for a fixed $\vec{\eta} \in \{0, 1\}^{\mathbb{N}}$. Integrating over $\vec{\eta}$, we obtain (3) and (4) as stated.

4.10. Comparing the allegedly optimal strategies for two partitions

Let π be a partition $0 = t_0 < t_1 < \cdots < t_N = T$, and let π' be a refinement of π, given by $0 = t'_0 < t'_1 < \cdots < t'_N = T$. Let $\tilde{\sigma}$ and $\tilde{\sigma}'$ be the corresponding ALLEGEDLY OPTIMAL STRATEGIES. We set $\Delta t_{\text{MAX}} = \max_\nu (t_{\nu+1} - t_\nu)$ and $\Delta t'_{\text{MAX}} = \max_\nu (t'_{\nu+1} - t'_\nu) \le \Delta t_{\text{MAX}}$. We will prove the following result.

Lemma 4.12. *Given $\varepsilon > 0$, there exists $\delta > 0$ such that if $\Delta t_{\text{MAX}} < \delta$, then*

$$\mathrm{E}_a\Big[\int_0^T \{|q^{\tilde{\sigma}}(t) - q^{\tilde{\sigma}'}(t)|^2 + |u^{\tilde{\sigma}}(t) - u^{\tilde{\sigma}'}(t)|^2\}\, dt\Big] < \varepsilon$$

for all $a \in [-a_{\text{MAX}}, +a_{\text{MAX}}]$.

Proof. Given $\varepsilon > 0$, we pick $\hat{\delta} > 0$ small enough, and then pick $\delta > 0$ small enough, depending on $\hat{\delta}$.

Suppose $\Delta t_{\text{MAX}} < \delta$; then also $\Delta t'_{\text{MAX}} < \delta$. Theorem 4.11 gives

$$\Big|\mathrm{E}\Big[\int_0^T \{|q^{\tilde{\sigma}}(t)|^2 + |u^{\tilde{\sigma}}(t)|^2\}\, dt\Big] - S(q_0, 0, 0, 0)\Big| < \hat{\delta}, \tag{4.123}$$

$$\Big|\mathrm{E}\Big[\int_0^T \{|q^{\tilde{\sigma}'}(t)|^2 + |u^{\tilde{\sigma}'}(t)|^2\}\, dt\Big] - S(q_0, 0, 0, 0)\Big| < \hat{\delta}. \tag{4.124}$$

In particular,

$$\mathrm{E}\Big[\int_0^T \{|q^{\tilde{\sigma}}(t)|^2 + |u^{\tilde{\sigma}}(t)|^2\}\, dt\Big] < C, \tag{4.125}$$

$$\mathrm{E}\Big[\int_0^T \{|q^{\tilde{\sigma}'}(t)|^2 + |u^{\tilde{\sigma}'}(t)|^2\}\, dt\Big] < C. \tag{4.126}$$

Lemma 3.10 gives a tame strategy $\hat{\sigma}$ associated to the partition π' for which

$$\mathrm{E}_a\Big[\int_0^T \{|q^{\tilde{\sigma}}(t) - q^{\hat{\sigma}}(t)|^2 + |u^{\tilde{\sigma}}(t) - u^{\hat{\sigma}}(t)|^2\}\, dt\Big] \le \hat{\delta}^2 \tag{4.127}$$

for every $a \in [-a_{\text{MAX}}, +a_{\text{MAX}}]$, hence

$$\mathrm{E}\Big[\int_0^T \{|q^{\tilde{\sigma}}(t) - q^{\hat{\sigma}}(t)|^2 + |u^{\tilde{\sigma}}(t) - u^{\hat{\sigma}}(t)|^2\}\, dt\Big] \le \hat{\delta}^2.$$

Together with (4.125), this implies that

$$\mathrm{E}\Big[\int_0^T \{|q^{\hat{\sigma}}(t)|^2 + |u^{\hat{\sigma}}(t)|^2\}\, dt\Big] \le \mathrm{E}\Big[\int_0^T \{|q^{\tilde{\sigma}}(t)|^2 + |u^{\tilde{\sigma}}(t)|^2\}\, dt\Big] + C\hat{\delta}.$$

Thanks to (4.123) and (4.124), this in turn implies that

$$\mathrm{E}\Big[\int_0^T \{|q^{\hat{\sigma}}(t)|^2 + |u^{\hat{\sigma}}(t)|^2\}\, dt\Big] \le \mathrm{E}\Big[\int_0^T \{|q^{\tilde{\sigma}'}(t)|^2 + |u^{\tilde{\sigma}'}(t)|^2\}\, dt\Big] + C\hat{\delta}. \tag{4.128}$$

Recall that $\tilde{\sigma}'$ is the ALLEGEDLY OPTIMAL STRATEGY associated to the partition π', while $\hat{\sigma}$ is another tame strategy associated to π'.

Consequently, by virtue of Theorem 4.11, (4.128) implies that

$$\mathrm{E}_a\left[\int_0^T \{|q^{\hat{\sigma}}(t) - q^{\hat{\sigma}'}(t)|^2 + |u^{\hat{\sigma}}(t) - u^{\hat{\sigma}'}(t)|^2\}\,dt\right] < \frac{\varepsilon}{2} \tag{4.129}$$

for every $a \in [-a_{\mathrm{MAX}}, +a_{\mathrm{MAX}}]$.

From (4.127) and (4.129), we have

$$\mathrm{E}_a\left[\int_0^T \{|q^{\tilde{\sigma}}(t) - q^{\tilde{\sigma}'}(t)|^2 + |u^{\tilde{\sigma}}(t) - u^{\tilde{\sigma}'}(t)|^2\}\,dt\right] < \varepsilon,$$

completing the proof of the lemma. ■

Corollary 4.13. *Given $\varepsilon > 0$ there exists $\delta > 0$ for which the following holds. Let*

$$\pi : 0 = t_0 = t_0 < t_1 < \cdots < t_N = T \quad \text{and} \quad \pi' : 0 = t'_0 < t'_1 < \cdots < t'_{N'} = T$$

be partitions, let

$$\Delta t_{\mathrm{MAX}} = \max_\nu (t_{\nu+1} - t_\nu) \quad \text{and} \quad \Delta t'_{\mathrm{MAX}} = \max_\nu (t'_{\nu+1} - t'_\nu),$$

and let $\tilde{\sigma}$ and $\tilde{\sigma}'$ be the ALLEGEDLY OPTIMAL STRATEGIES *associated to π and π', respectively.*

If $\Delta t_{\mathrm{MAX}}, \Delta t'_{\mathrm{MAX}} < \delta$, then

$$\mathrm{E}_a\left[\int_0^T \{|q^{\tilde{\sigma}}(t) - q^{\tilde{\sigma}'}(t)|^2 + |u^{\tilde{\sigma}}(t) - u^{\tilde{\sigma}'}(t)|^2\}\,dt\right] < \varepsilon$$

for every $a \in [-a_{\mathrm{MAX}}, +a_{\mathrm{MAX}}]$.

Proof. Compare both $\tilde{\sigma}$ and $\tilde{\sigma}'$ to the ALLEGEDLY OPTIMAL STRATEGY arising from a common refinement of π and π'. ■

5. Decisions in continuous time

5.1. Tame strategies with decisions in continuous time

Suppose that for each $n = 1, 2, 3, \ldots$, we are given a tame strategy σ^n associated to a partition π^n of the time interval $[0, T]$. Say π^n is given by

$$0 = t_0^n < t_1^n < \cdots < t_{N(n)}^n = T, \tag{5.1}$$

and σ^n is given by the collection of tame rules

$$\sigma^n = (\sigma^n_{t^n_\nu})_{\nu=0,1,\ldots,N(n)-1},$$

where each $\sigma^n_{t^n_\nu}$ is a function of ν real variables $\bar{q}_1, \ldots, \bar{q}_\nu$ and the coin flips $\vec{\xi}$.

We assume that

$$\Delta t^n_{\text{MAX}} := \max_\nu (t^n_{\nu+1} - t^n_\nu) \to 0 \quad \text{as } n \to \infty, \tag{5.2}$$

and that

$$|\sigma^n_{t^n_\nu}(\bar{q}_1, \dots, \bar{q}_\nu, \vec{\xi})| \le C^{\vec{\sigma}}_{\text{TAME}}[|\bar{q}_\nu| + 1] \quad \text{for all } \bar{q}_1, \dots, \bar{q}_\nu \tag{5.3}$$

for all n, ν, with $C^{\vec{\sigma}}_{\text{TAME}}$ independent of $n, \nu, \bar{q}_1, \dots, \bar{q}_\nu$.

Each σ^n gives rise to control trajectories and particle trajectories $u^{\sigma^n, a}(t)$ and $q^{\sigma^n, a}(t)$, respectively, for $t \in [0, T]$, once we specify that $a_{\text{TRUE}} = a$. Here, a is an arbitrary real number, not necessarily belonging to the interval $[-a_{\text{MAX}}, +a_{\text{MAX}}]$. We define the *expected cost* of σ^n by

$$\text{ECOST}(\sigma^n, a) = \mathrm{E}_a\Big[\int_0^T \{(u^{\sigma^n, a}(t))^2 + (q^{\sigma^n, a}(t))^2\}\, dt\Big].$$

We say that $(\sigma^n)_{n\ge1}$ is a *Cauchy sequence of uniformly tame strategies* if (5.2) and (5.3) hold, and

$$\lim_{n,m\to\infty} \mathrm{E}_a\Big[\int_0^T \{|u^{\sigma^n, a}(t) - u^{\sigma^m, a}(t)|^2 + |q^{\sigma^n, a}(t) - q^{\sigma^m, a}(t)|^2\}\, dt\Big] = 0, \tag{5.4}$$

uniformly for a in any bounded subset of $\mathbb{R}$.

If $(\sigma^n)_{n\ge1}$ and $(\hat{\sigma}^n)_{n\ge1}$ are two Cauchy sequences of uniformly tame strategies, then we call those sequences *equivalent* if we have

$$\lim_{n\to\infty} \mathrm{E}_a\Big[\int_0^T \{|u^{\sigma^n, a}(t) - u^{\hat{\sigma}^n, a}(t)|^2 + |q^{\sigma^n, a}(t) - q^{\hat{\sigma}^n, a}(t)|^2\}\, dt\Big] = 0 \tag{5.5}$$

for each $a \in \mathbb{R}$.

If $(\sigma^n)_{n\ge1}$ is a Cauchy sequence of uniformly tame strategies, then for each $a \in \mathbb{R}$, there exist random functions $u^a(t)$ and $q^a(t)$ such that

$$\lim_{n\to\infty} \mathrm{E}_a\Big[\int_0^T \{|u^{\sigma^n, a}(t) - u^a(t)|^2 + |q^{\sigma^n, a}(t) - q^a(t)|^2\}\, dt\Big] = 0, \tag{5.6}$$

uniformly for a in any bounded subset of $\mathbb{R}$.

Moreover, if two Cauchy sequences $(\sigma^n)_{n\ge1}$ and $(\hat{\sigma}^n)_{n\ge1}$ are equivalent, then for each $a \in \mathbb{R}$, the u^a and q^a defined by those sequences are equal, for a.e. $t \in [0, T]$, almost surely with respect to Prob_a.

We define a *tame strategy* (for decisions in continuous time) to be an equivalence class of Cauchy sequences of uniformly tame strategies, with respect to the equivalence relation (5.5). We denote a tame strategy by $\vec{\sigma} = [[(\sigma^n)_{n\ge1}]]$, and we say that $C^{\vec{\sigma}}_{\text{TAME}}$ in (5.3) is a *tame constant* for $\vec{\sigma}$. If $\vec{\sigma} = [[(\sigma^n)_{n\ge1}]]$ is a tame strategy, then we write $q^{\vec{\sigma}, a}(t)$ and $u^{\vec{\sigma}, a}(t)$ to denote the functions $q^a(t)$ and $u^a(t)$ in (5.6).

If $\vec{\sigma} = [[(\sigma^n)_{n\ge1}]]$ is a tame strategy, then we define

$$\text{ECOST}(\vec{\sigma}, a) = \lim_{n\to\infty} \text{ECOST}(\sigma^n, a).$$

This quantity is well defined, since the limit exists and two equivalent Cauchy sequences produce the same expected cost. Immediately from our results on tame strategies associated to partitions of $[0, T]$, we have the following results, for any tame strategy $\vec{\sigma} = [[(\sigma^n)_{n\geq 1}]]$.

Note that

$$\mathrm{E}\textsc{cost}(\vec{\sigma}, a) = \mathrm{E}_a\Big[\int_0^T \{(q^{\vec{\sigma},a}(t))^2 + (u^{\vec{\sigma},a}(t))^2\}\,dt\Big]$$

for any tame strategy $\vec{\sigma}$ and any $a \in \mathbb{R}$. For large enough $a > 0$, we have

$$\mathrm{E}\textsc{cost}(\vec{\sigma}, a) \geq cT^2 \exp(caT). \tag{5.7}$$

Moreover, we will see that the function

(5.8) $[-a_{\mathrm{MAX}}, +a_{\mathrm{MAX}}] \ni a \mapsto \mathrm{E}\textsc{cost}(\vec{\sigma}, a)$ continues to a bounded analytic function on $(-a_{\mathrm{MAX}}, +a_{\mathrm{MAX}}) + i(-\hat{c}, \hat{c})$ for some $\hat{c} > 0$ determined by the BOILERPLATE CONSTANTS and the constant $C^{\vec{\sigma}}_{\mathrm{TAME}}$.

Moreover, we may replace a_{MAX} by any $\hat{a}_{\mathrm{MAX}} > a_{\mathrm{MAX}}$, and the assumptions of the preceding sections are still valid; the constants determined by the BOILERPLATE CONSTANTS will now depend on $\hat{a}_{\mathrm{MAX}}$. In particular, (5.8) immediately implies that the function

$$[-\hat{a}_{\mathrm{MAX}}, +\hat{a}_{\mathrm{MAX}}] \ni a \mapsto \mathrm{E}\textsc{cost}(\vec{\sigma}, a)$$

continues to a bounded analytic function on

$$(-\hat{a}_{\mathrm{MAX}}, +\hat{a}_{\mathrm{MAX}}) + i(-\hat{c}(\hat{a}_{\mathrm{MAX}}), +\hat{c}(\hat{a}_{\mathrm{MAX}}));$$

the bound for that analytic function depends on $\hat{a}_{\mathrm{MAX}}$. This implies that the function

$$\mathbb{R} \ni a \mapsto \mathrm{E}\textsc{cost}(\vec{\sigma}, a)$$

continues to an analytic function on a neighborhood of the real axis in $\mathbb{C}$. In other words,

$$\mathrm{E}\textsc{cost}(\vec{\sigma}, a) \text{ is a real-analytic function of } a \in \mathbb{R}. \tag{5.9}$$

Let us check our assertion (5.8). Recall our previous result on analytic continuation: given $\varepsilon > 0$, there exists $\delta > 0$ such that whenever $\Delta t^n_{\mathrm{MAX}} < \delta$, we have

$$\mathrm{E}\textsc{cost}(\sigma^n, a) = I^n(a) + \textsc{error}^n(a) \tag{5.10}$$

on $(-a_{\mathrm{MAX}}, +a_{\mathrm{MAX}})$, where

$$|\textsc{error}^n(a)| < \varepsilon \quad \text{for } a \in (-a_{\mathrm{MAX}}, a_{\mathrm{MAX}}), \tag{5.11}$$

and

$$I^n(a) \text{ is analytic on } \mathcal{R} \equiv (-a_{\mathrm{MAX}}, +a_{\mathrm{MAX}}) + i(-\hat{c}, \hat{c}), \tag{5.12}$$

with $|I^n(a)| \leq C$ everywhere on that rectangle. In particular, (5.10), (5.11) and (5.12) hold for all large enough n, since $\Delta t^n_{\mathrm{MAX}} \to 0$ as $n \to \infty$.

Since the $I^n(a)$ are uniformly bounded analytic functions on the rectangle $\mathcal{R}$, we may pick out a subsequence $I^{n_j}(a)$ that converges to a bounded analytic function $I^\infty(a)$ uniformly on compact subsets of $\mathcal{R}$. Applying (5.10) and (5.11) to σ^{n_j}, and passing to the limit as $j \to \infty$, we find that $\text{ECOST}(\vec{\sigma}, a) = I^\infty(a)$ for all $a \in (-a_{\text{MAX}}, a_{\text{MAX}})$, completing the proof of (5.8).

Next, we construct an ALLEGEDLY OPTIMAL STRATEGY with *decisions in continuous time*.

Fix a prior probability distribution $d\text{Prob}(a)$ on $[-a_{\text{MAX}}, +a_{\text{MAX}}]$. Given a tame strategy $\vec{\sigma} = [[(\sigma_n)_{n\geq 1}]]$, we define

$$\text{ECOST}(\sigma^n) = \int_{-a_{\text{MAX}}}^{a_{\text{MAX}}} \text{ECOST}(\sigma^n, a)\, d\text{Prob}(a),$$
$$\text{ECOST}(\vec{\sigma}) = \int_{-a_{\text{MAX}}}^{a_{\text{MAX}}} \text{ECOST}(\vec{\sigma}, a)\, d\text{Prob}(a).$$

Let π^n be a sequence of partitions of $[0, T]$ (as in (5.1)), with $\Delta t^n_{\text{MAX}} \to 0$ as $n \to \infty$.

For each n, let $\tilde{\sigma}^n$ denote the allegedly optimal strategy associated to the partition π^n. Corollary 4.13 tells us that $(\tilde{\sigma}^n)_{n\geq 1}$ satisfies condition (5.4). Moreover, we have assumed condition (5.2), and our PDE assumption (see (4.26)) tells us that (5.3) holds. Thus, the $(\tilde{\sigma}^n)_{n\geq 1}$ form a Cauchy sequence of uniformly tame strategies.

We write $\vec{\sigma}_{\text{opt}} = [[(\tilde{\sigma}^n)_{n\geq 1}]]$ to denote the resulting continuous tame strategy. Note that $\vec{\sigma}_{\text{opt}}$ is independent of the sequence of partitions used to define it. We will show that it is optimal for Bayesian control.

Let $\vec{\sigma} = [[(\sigma^n)_{n\geq 1}]]$ be a tame strategy with tame constant $C^{\vec{\sigma}}_{\text{TAME}}$. Then

$$\text{ECOST}(\vec{\sigma}) = \int_{-a_{\text{MAX}}}^{a_{\text{MAX}}} \text{ECOST}(\vec{\sigma}, a)\, d\text{Prob}(a) = \lim_{n\to\infty} \int_{-a_{\text{MAX}}}^{a_{\text{MAX}}} \text{ECOST}(\sigma^n, a)\, d\text{Prob}(a)$$
$$= \lim_{n\to\infty} \text{ECOST}(\sigma^n).$$

Here, the interchange of limit and integral is justified by the uniform convergence in a that we assumed in our definition of Cauchy sequences.

Let $\varepsilon > 0$ be given. For n large enough, condition (5.2) for the σ^n allows us to apply Theorem 4.11; we conclude that

$$\text{ECOST}(\sigma^n) \geq \text{ECOST}(\tilde{\sigma}^n) - \varepsilon$$

for n large enough. Here, $\tilde{\sigma}^n$ denotes the ALLEGEDLY OPTIMAL STRATEGY associated to the partition relevant to σ^n.

Passing to the limit as $n \to \infty$, we see that

$$\text{ECOST}([[(\sigma^n)_{n\geq 1}]]) \geq \text{ECOST}(\vec{\sigma}_{\text{opt}}).$$

Thus, indeed, $\vec{\sigma}_{\text{opt}}$ is the optimal tame Bayesian strategy; any competing tame Bayesian strategy has an expected cost at least that of $\vec{\sigma}_{\text{opt}}$.

Next, we compare the cost of $\vec{\sigma}_{\text{opt}}$ with that of another tame strategy $\vec{\sigma} = [[(\sigma_n)_{n\geq 1}]]$ conditioned on $a_{\text{TRUE}} = a$ for a given $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$.

We will prove the following assertion:

(5.13) Given $\varepsilon > 0$, there exists δ, depending on the BOILERPLATE CONSTANTS and the constant $C^{\vec{\sigma}}_{\text{TAME}}$, such that if $\text{ECOST}(\vec{\sigma}) \le \text{ECOST}(\vec{\sigma}_{\text{opt}}) + \delta$, then, for every $a \in [-a_{\text{MAX}}, +a_{\text{MAX}}]$, we have $|\text{ECOST}(\vec{\sigma}, a) - \text{ECOST}(\vec{\sigma}_{\text{opt}}, a)| < \varepsilon$.

To prove (5.13), we recall that $\vec{\sigma}_{\text{opt}} = [[(\tilde{\sigma}_n)_{n\ge 1}]]$, with $\tilde{\sigma}_n$ the allegedly optimal strategy associated to the partition associated to σ_n. Then

$$\text{ECOST}(\vec{\sigma}) = \lim_{n\to\infty} \text{ECOST}(\sigma_n), \tag{5.14}$$

$$\text{ECOST}(\vec{\sigma}_{\text{opt}}) = \lim_{n\to\infty} \text{ECOST}(\tilde{\sigma}_n), \tag{5.15}$$

$$\text{ECOST}(\vec{\sigma}, a) = \lim_{n\to\infty} \text{ECOST}(\sigma_n, a), \tag{5.16}$$

$$\text{ECOST}(\vec{\sigma}_{\text{opt}}, a) = \lim_{n\to\infty} \text{ECOST}(\tilde{\sigma}_n, a). \tag{5.17}$$

If $\text{ECOST}(\vec{\sigma}) \le \text{ECOST}(\vec{\sigma}_{\text{opt}}) + \delta$, then by (5.14) and (5.15), we have

$$\text{ECOST}(\sigma_n) \le \text{ECOST}(\tilde{\sigma}_n) + 2\delta \quad \text{for large enough } n.$$

Also, for large enough n, the partition of $[0, T]$ associated to σ_n, $\tilde{\sigma}_n$ has mesh less than 2δ. It therefore follows from Theorem 4.11 that

$$|\text{ECOST}(\sigma_n, a) - \text{ECOST}(\tilde{\sigma}_n, a)| \le \frac{\varepsilon}{2}$$

for all $a \in [-a_{\text{MAX}}, +a_{\text{MAX}}]$ and all large enough n.

From (5.16) and (5.17), we now see that

$$|\text{ECOST}(\vec{\sigma}, a) - \text{ECOST}(\vec{\sigma}_{\text{opt}}, a)| \le \varepsilon$$

for all $a \in [-a_{\text{MAX}}, +a_{\text{MAX}}]$, completing the proof of (5.13).

5.2. Not-necessarily-tame strategies

In this section, we drop the restriction to tame strategies.

Let $\vec{\sigma}_1, \vec{\sigma}_2, \vec{\sigma}_3, \ldots$ be tame strategies in the sense of Section 5.1. We do *not* assume the $\vec{\sigma}_n$ have a tame constant independent of n.

We say that the sequence $(\vec{\sigma}_n)_{n\ge 1}$ is *Cauchy* if

$$\lim_{m,n\to\infty} \mathrm{E}_a \Big[\int_0^T \big\{ |q^{\vec{\sigma}_n, a}(t) - q^{\vec{\sigma}_m, a}(t)|^2 + |u^{\vec{\sigma}_n, a}(t) - u^{\vec{\sigma}_m, a}(t)|^2 \big\}\, dt \Big] = 0,$$

uniformly for a in any bounded subset of $\mathbb{R}$. Two Cauchy sequences $(\vec{\sigma}_n)_{n\ge 1}$ and $(\vec{\sigma}^{\#}_n)_{n\ge 1}$ will be called *equivalent* if

$$\lim_{n\to\infty} \mathrm{E}_a \Big[\int_0^T \big\{ |q^{\vec{\sigma}_n, a}(t) - q^{\vec{\sigma}^{\#}_n, a}(t)|^2 + |u^{\vec{\sigma}_n, a}(t) - u^{\vec{\sigma}^{\#}_n, a}(t)|^2 \big\}\, dt \Big] = 0$$

for each $a \in \mathbb{R}$.

To a Cauchy sequence $(\vec{\sigma}_n)_{n\geq 1}$ as above, we associate the trajectories $q^a(t)$ and $u^a(t)$, for which we have

$$\lim_{n\to\infty} \mathrm{E}_a\left[\int_0^T \left\{|q^{\vec{\sigma}_n,a}(t) - q^a(t)|^2 + |u^{\vec{\sigma}_n,a}(t) - u^a(t)|^2\right\} dt\right] = 0.$$

Two equivalent Cauchy sequences yield the same q^a and u^a. We define a *strategy* to be an equivalence class of Cauchy sequences under the above equivalence relation. We denote strategies by $\vec{\vec{\sigma}} = [[(\vec{\sigma}_n)_{n\geq 1}]]$, and we write $q^{\vec{\vec{\sigma}},a}(t)$ and $u^{\vec{\vec{\sigma}},a}(t)$, respectively, to denote the above functions $q^a(t)$ and $u^a(t)$. We define

$$\textsc{Ecost}(\vec{\vec{\sigma}}, a) = \lim_{n\to\infty} \textsc{Ecost}(\sigma_n, a) = \mathrm{E}_a\left[\int_0^T \left\{(q^{\vec{\vec{\sigma}},a}(t))^2 + (u^{\vec{\vec{\sigma}},a}(t))^2\right\} dt\right].$$

These limits converge uniformly for a in any bounded subset of $\mathbb{R}$. Note that every tame strategy $\vec{\sigma}$ may also be regarded as a strategy as defined just above, namely the equivalence class associated with the constant sequence $\vec{\sigma}, \vec{\sigma}, \vec{\sigma}, \ldots$.

Suppose we are given a prior probability distribution $d\mathrm{Prior}$ on $[-a_{\mathrm{MAX}}, a_{\mathrm{MAX}}]$. We define

$$\textsc{Ecost}(\vec{\vec{\sigma}}, d\mathrm{Prior}) = \int_{-a_{\mathrm{MAX}}}^{a_{\mathrm{MAX}}} \textsc{Ecost}(\vec{\vec{\sigma}}, a)\, d\mathrm{Prior}(a).$$

If $\vec{\vec{\sigma}} = [[(\vec{\sigma}_n)_{n\geq 1}]]$, then

$$\begin{aligned}\textsc{Ecost}(\vec{\vec{\sigma}}, d\mathrm{Prior}) &= \int_{-a_{\mathrm{MAX}}}^{a_{\mathrm{MAX}}} \lim_{n\to\infty} \textsc{Ecost}(\vec{\sigma}_n, a)\, d\mathrm{Prior}(a)\\ &= \lim_{n\to\infty} \int_{-a_{\mathrm{MAX}}}^{a_{\mathrm{MAX}}} \textsc{Ecost}(\vec{\sigma}_n, a)\, d\mathrm{Prior}(a) = \lim_{n\to\infty} \textsc{Ecost}(\vec{\sigma}_n, d\mathrm{Prior});\end{aligned}$$

the interchange of limit and integral is justified by the uniform convergence noted above.

Now let $\vec{\sigma}_{\mathrm{Bayes}}(d\mathrm{Prior})$ be the optimal Bayesian strategy for $d\mathrm{Prior}$, given in Section 5.1. For any tame strategy $\vec{\sigma}$, we have seen that

$$\textsc{Ecost}(\vec{\sigma}, d\mathrm{Prior}) \geq \textsc{Ecost}(\vec{\sigma}_{\mathrm{Bayes}}(d\mathrm{Prior}), d\mathrm{Prior}).$$

In particular, if $\vec{\vec{\sigma}} = [[(\vec{\sigma}_n)_{n\geq 1}]]$, then

$$\textsc{Ecost}(\vec{\sigma}_n, d\mathrm{Prior}) \geq \textsc{Ecost}(\vec{\sigma}_{\mathrm{Bayes}}(d\mathrm{Prior}), d\mathrm{Prior}),$$

hence

$$\textsc{Ecost}(\vec{\vec{\sigma}}, d\mathrm{Prior}) = \lim_{n\to\infty} \textsc{Ecost}(\vec{\sigma}_n, d\mathrm{Prior}) \geq \textsc{Ecost}(\vec{\sigma}_{\mathrm{Bayes}}(d\mathrm{Prior}), d\mathrm{Prior}).$$

So we see that the strategy $\vec{\sigma}_{\mathrm{Bayes}}(d\mathrm{Prior})$ has expected cost less than or equal to that of any competing strategy $\vec{\vec{\sigma}}$.

From now on, we drop the arrows from our notation. When we mention a strategy, we will make clear whether it is a general strategy, a tame strategy, or a tame strategy associated to a partition of $[0, T]$.

Combining the results of this section with Theorem 4.11, we deduce Theorems 1.1 and 1.2 from the introduction.

6. Agnostic control

Throughout this section, the random variable $a_{\text{TRUE}} \in [-a_{\text{MAX}}, a_{\text{MAX}}]$ is unknown and we *do not* assume that we have a prior belief about a_{TRUE}.

6.1. Mixed strategies

Let

$$\phi : (\xi_1, \xi_2, \xi_3, \dots) \mapsto (\xi_1, \xi_3, \xi_5, \dots)$$

be the map from $\{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$ that erases every other bit. Fix a partition

$$0 = t_0 < t_1 < \cdots < t_N = T, \tag{6.1}$$

and let $\sigma = (\sigma_{t_\nu})_{0 \le \nu < N}$ be a tame strategy associated to the partition (6.1). Since σ is tame, we have

$$|u^\sigma(t_\nu)| \le C^\sigma_{\text{TAME}} \cdot [|q^\sigma(t_\nu)| + 1] \quad \text{for each } \nu. \tag{6.2}$$

We can pass from σ to the morally equivalent strategy

$$\sigma^{\#} = (\sigma^{\#}_{t_\nu})_{0 \le \nu < N}$$

by setting

$$\sigma^{\#}_{t_\nu}(q_1, \dots, q_\nu, \vec{\xi}) = \sigma_{t_\nu}(q_1, \dots, q_\nu, \phi(\vec{\xi})).$$

Thus, the strategy $\sigma^{\#}$ does precisely what σ does, except that whereas σ makes use of the bits $\xi_1, \xi_2, \xi_3, \dots$, $\sigma^{\#}$ makes use only of the bits $\xi_1, \xi_3, \xi_5, \dots$.

Now let $\sigma^0 = (\sigma^0_{t_\nu})_{0 \le \nu < N}$ and $\sigma^1 = (\sigma^1_{t_\nu})_{0 \le \nu < N}$ be two tame strategies, both associated to the partition (6.1), and let $\theta \in [0,1]$ be given. We define a *mixed strategy* σ^θ as follows. First, we pass from σ^0 and σ^1 to the strategies $\sigma^{0\#}$ and $\sigma^{1\#}$ as above. These strategies make use of the bits $\xi_1, \xi_3, \xi_5, \dots$, but ignore the bits $\xi_2, \xi_4, \xi_6, \dots$ We regard $\xi_2, \xi_4, \xi_6, \dots$ as the binary digits of a random variable Y taking values in $[0,1]$. If $Y \le \theta$, then we play the strategy $\sigma^{1\#}$ at all times t_ν. If instead $Y > \theta$, then we play the strategy $\sigma^{0\#}$ at all times t_ν.

Evidently,

$$\textsc{Ecost}(\sigma^{0\#}, a) = \textsc{Ecost}(\sigma^0, a) \quad \text{and} \quad \textsc{Ecost}(\sigma^{1\#}, a) = \textsc{Ecost}(\sigma^1, a)$$

for all $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$; and

$$\textsc{Ecost}(\sigma^\theta, a) = \theta\, \textsc{Ecost}(\sigma^1, a) + (1-\theta)\, \textsc{Ecost}(\sigma^0, a) \tag{6.3}$$

for all $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$, since $Y \le \theta$ with probability θ. Note that σ^θ is a tame strategy, with

$$C^{\sigma^\theta}_{\text{TAME}} \le \max\left\{C^{\sigma^0}_{\text{TAME}}, C^{\sigma^1}_{\text{TAME}}\right\}. \tag{6.4}$$

We have defined the intermediate strategy σ^θ when σ^0 and σ^1 are tame strategies associated to the same partition (6.1) of $[0,T]$.

We next extend our definition to tame strategies with decisions in continuous time. Fix a sequence $\pi_1, \pi_2, \dots$ of partitions of $[0,T]$, with $\text{mesh}(\pi_i) \to 0$ as $i \to \infty$. Let $\theta \in [0,1]$ be given. Let $\sigma^0_1, \sigma^0_2, \sigma^0_3, \dots$ and $\sigma^1_1, \sigma^1_2, \sigma^1_3, \dots$ be tame strategies, where, for each i, the strategies σ^0_i and σ^1_i are associated to the partition π_i of $[0,T]$.

Suppose that $(\sigma_i^0)_{i=1,2,\dots}$ and $(\sigma_i^1)_{i=1,2,\dots}$ are Cauchy sequences, in the sense of Section 5.1. Thus, $\vec{\sigma}^0 = [[(\sigma_i^0)_{i\geq 1}]]$ and $\vec{\sigma}^1 = [[(\sigma_i^1)_{i\geq 1}]]$ are tame strategies in the sense of that section. For each i, we pass from σ_i^0 and σ_i^1 to the mixed strategy σ_i^θ associated to the partition π_i of $[0, T]$. Then $\sigma_1^\theta, \sigma_2^\theta, \dots$ is again a Cauchy sequence in the sense of Section 5.1. We write $\vec{\sigma}^\theta$ to denote the tame strategy $[[(\sigma_i^\theta)_{i\geq 1}]]$. Then we have

$$\textsc{Ecost}(\vec{\sigma}^\theta, a) = \theta\,\textsc{Ecost}(\vec{\sigma}^1, a) + (1-\theta)\,\textsc{Ecost}(\vec{\sigma}^0, a)$$

for all $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$, and

$$C_{\text{TAME}}^{\sigma_i^\theta} \leq \max\left\{C_{\text{TAME}}^{\sigma_i^0}, C_{\text{TAME}}^{\sigma_i^1}\right\},$$

as follows easily from (6.3) and (6.4).

Note that we have restricted attention to tame strategies $[[(\sigma_i^0)_{i\geq 1}]]$ and $[[(\sigma_i^1)_{i\geq 1}]]$ in which, for each i, σ_i^0 and σ_i^1 are associated to the same partition of $[0, T]$. It would be natural to dispense with this restriction, but for our purposes that will not be necessary.

6.2. Efficient strategies are Bayesian

In this section, we deal with tame strategies in the sense of Section 5.1 of Section 5.

Suppose we are given a class of strategies, which we call the LEGAL STRATEGIES.

Assume that given two LEGAL STRATEGIES σ^0 and σ^1, and given $\theta \in [0, 1]$, there exists a LEGAL STRATEGY σ^θ for which we have

$$\textsc{Ecost}(\sigma^\theta, a) = (1-\theta)\,\textsc{Ecost}(\sigma^0, a) + \theta\,\textsc{Ecost}(\sigma^1, a) \tag{6.5}$$

for all $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$.

For example, suppose we fix a constant $\hat{C}$ and a sequence of partitions $(\pi_i)_{i\geq 1}$ of the interval $[0, T]$, with $\text{mesh}(\pi_i) \to 0$ as $i \to \infty$.

Then the class of all tame strategies $[[(\sigma_i)_{i\geq 1}]]$ with σ_i associated to π_i and $C_{\text{TAME}}^{\sigma_i} \leq \hat{C}$ satisfies (6.5), thanks to our discussion of mixed strategies in Section 6.1.

Fix a finite set $A \subset [-a_{\text{MAX}}, a_{\text{MAX}}]$, and let $\varepsilon \geq 0$ be given. (Note that we allow $\varepsilon = 0$.)

A LEGAL STRATEGY σ will be said to be *efficient with tolerance* ε if there does not exist another LEGAL STRATEGY σ' such that

$$\textsc{Ecost}(\sigma', a) < \textsc{Ecost}(\sigma, a) - \varepsilon \tag{6.6}$$

for all $a \in A$. This notion depends on the set A and the class of LEGAL STRATEGIES.

In this section, we use a simple convexity argument to prove the following result.

Lemma 6.1 (Efficient strategies are Bayesian). *Fix A, ε and a class of* LEGAL STRATEGIES *as above, and let $\hat{\sigma}$ be a* LEGAL STRATEGY. *Suppose $\hat{\sigma}$ is efficient with tolerance ε. Then there exists a prior probability distribution $(p(a))_{a\in A}$ such that for all other* LEGAL STRATEGIES *σ', we have*

$$\sum_{a\in A} p(a)\,\textsc{Ecost}(\hat{\sigma}, a) \leq \sum_{a\in A} p(a)\,\textsc{Ecost}(\sigma', a) + \varepsilon. \tag{6.7}$$

Proof. For any strategy σ, define the *cost vector* to be the vector $(\textsc{Ecost}(\sigma, a))_{a\in A} \in \mathbb{R}^A$. Thanks to (6.5), the set $\mathcal{K}$ of all cost vectors of legal strategies is convex. Define another convex set $\mathcal{K}_- \subset \mathbb{R}^A$ to consist of all vectors $(v_a)_{a\in A}$ such that $v_a < \textsc{Ecost}(\hat{\sigma}, a) - \varepsilon$ for all $a \in A$. Because $\hat{\sigma}$ is efficient with tolerance ε, the convex sets $\mathcal{K}$ and $\mathcal{K}_-$ are disjoint. Hence there exists a nonzero linear functional $\lambda: \mathbb{R}^A \to \mathbb{R}$ such that $\lambda(v) \leq \lambda(v^*)$ whenever $v \in \mathcal{K}_-$ and $v^* \in \mathcal{K}$. The functional λ has the form

$$\lambda((v_a)_{a\in A}) = \sum_{a\in A} p(a)\, v_a,$$

with at least one nonzero coefficient $p(a_0)$. By definition of $\mathcal{K}_-$, $\mathcal{K}$, and λ, the following holds:

(6.8) Let σ' be a LEGAL STRATEGY, and let $(v_a)_{a\in A}$ satisfy $v_a < \textsc{Ecost}(\hat{\sigma}, a) - \varepsilon$ for all $a \in A$. Then $\sum_{a\in A} p(a) v_a \leq \sum_{a\in A} p(a)\,\textsc{Ecost}(\sigma', a)$.

We claim that the $p(a)$ are all nonnegative. Indeed, suppose $p(\hat{a}) < 0$ for some $\hat{a} \in A$. We take $\sigma' = \hat{\sigma}$, $v_a = \textsc{Ecost}(\hat{\sigma}, a) - \varepsilon - 1$ for $a \in A \backslash \{\hat{a}\}$, and $v_{\hat{a}} = -\mathcal{V}$ for some large positive $\mathcal{V}$. If $\mathcal{V}$ is large enough, then the above σ' and $(v_a)_{a\in A}$ violate (6.8). So, as claimed, the $p(a)$ are all nonnegative.

Since also the $p(a)$ are not all zero, we may multiply the $p(a)$ by a positive normalizing constant to preserve (6.8) and achieve also

$$\sum_{a\in A} p(a) = 1. \tag{6.9}$$

Thus, $(p(a))_{a\in A}$ is a probability distribution.

Now let $\delta > 0$, and let $v_a = \textsc{Ecost}(\hat{\sigma}, a) - \varepsilon - \delta$ for $a \in A$. Thanks to (6.8) and (6.9), we have

$$\sum_{a\in A} p(a)\,\textsc{Ecost}(\hat{\sigma}, a) - \varepsilon - \delta \leq \sum_{a\in A} p(a)\,\textsc{Ecost}(\sigma', a)$$

for every legal strategy σ'. Since $\delta > 0$ may be taken arbitrarily small, inequality (6.7) follows, completing the proof of the lemma. ■

6.3. Regret

We fix continuous functions $\rho_0, \rho_1: \mathbb{R} \to \mathbb{R}$. We suppose that

$$|\rho_0(a)| \leq \tilde{C} \ \text{ and } \ 0 < \tilde{c} < \rho_1(a) < \tilde{C} \quad \text{for } a \in [-a_{\text{MAX}}, a_{\text{MAX}}].$$

For any strategy σ and any $a \in \mathbb{R}$, we define

$$\textsc{Regret}(\sigma, a) = \rho_0(a) + \rho_1(a) \cdot \textsc{Ecost}(\sigma, a).$$

We add the above $\tilde{c}$ and $\tilde{C}$ to our list of BOILERPLATE CONSTANTS. As usual, c, C, C', etc., denote constants depending only on the BOILERPLATE CONSTANTS. These symbols may denote different constants in different occurrences. The above notion of regret includes as special cases our earlier notions of additive, multiplicative, and hybrid regret.

6.4. The main lemma on agnostic control

Suppose our PDE assumption holds (with the same constants K, m_0 and $C^{\text{opt}}_{\text{TAME}}$) for every prior probability distribution on a given interval $[-a_{\text{MAX}}, a_{\text{MAX}}]$. Under that assumption (see Section 4.3), we prove the following result.

Lemma 6.2. *Let $\varepsilon > 0$, and let $A \subset [-a_{\text{MAX}}, a_{\text{MAX}}]$ be finite. Then there exist a subset $A_0 \subset A$, a probability measure μ, and a strategy $\tilde{\sigma}$, with the following properties.*

(1) *The measure μ is concentrated on A_0.*

(2) *$\tilde{\sigma}$ is the optimal Bayesian strategy for the prior μ.*

(3) *For $a \in A$ and $a_0 \in A_0$, we have*

$$\text{REGRET}(\tilde{\sigma}, a) \leq \text{REGRET}(\tilde{\sigma}, a_0) + \varepsilon.$$

In particular,

(4) $|\text{REGRET}(\tilde{\sigma}, a_0) - \text{REGRET}(\tilde{\sigma}, a_0')| \leq \varepsilon$ *for* $a_0, a_0' \in A_0$.

For the proof of the above lemma, we first fix a class of strategies that we call OK. For $i = 1, 2, 3, \ldots$, let σ_i be a tame strategy arising from the partition $[0, T] \cap 2^{-i} T\,\mathbb{Z}$ of the time interval $[0, T]$. Suppose that the σ_i form a Cauchy sequence in the sense of Section 5.1, and that the tame constants $C_{\text{TAME}}^{\sigma_i}$ are all less than or equal to the constant $C_{\text{TAME}}^{\text{opt}}$ in our PDE assumption (see Section 4.3, inequality (4.26)). Then the strategy $\vec{\sigma} = [[(\sigma_i)_{i=1,2,\ldots}]]$ will be called *OK*.

We make two crucial observations regarding OK strategies:

(1) For any prior μ on $[-a_{\text{MAX}}, a_{\text{MAX}}]$, the optimal Bayesian strategy $\tilde{\sigma}$ is OK.

(2) If σ and σ' are OK strategies, then so is the mixed strategy that plays strategy σ with probability θ and strategy σ' with probability $(1-\theta)$ (for $0 \leq \theta \leq 1$).

If A is any finite subset of $[-a_{\text{MAX}}, a_{\text{MAX}}]$ and σ is any strategy, we write $\text{MR}(\sigma, A)$ to denote the quantity $\max\{\text{REGRET}(\sigma, a) : a \in A\}$. For any strategy σ and any prior μ on a finite set A, we write

$$\text{ECOST}(\sigma, \mu) = \sum_{a \in A} \text{ECOST}(\sigma, a)\,\mu(a).$$

We now begin the proof of Lemma 6.2.

Proof. We proceed by induction on $\#A$, the number of elements of A.

In the base case, $\#A = 1$, i.e., $A = \{a_0\}$ for some $a_0 \in [-a_{\text{MAX}}, a_{\text{MAX}}]$. We take $A_0 = A$, $\mu =$ point mass at a_0, $\tilde{\sigma} =$ optimal known-a strategy for $a = a_0$. The conclusions of the lemma are obvious.

For the induction step, we fix $k \geq 2$ and assume the:

INDUCTION HYPOTHESIS. Our lemma holds whenever $\#A < k$.

We fix A with $\#A = k$, and prove the lemma for A.

Let $\varepsilon > 0$ be given. We pick $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_7 > 0$, with $\varepsilon_0 = \varepsilon$; and with ε_{i+1} small enough, depending on $\varepsilon_0, \ldots, \varepsilon_i$ and the BOILERPLATE CONSTANTS.

Let $\text{MR}_* = \inf\{\text{MR}(\sigma, A) : \sigma \text{ any OK strategy}\}$ and let σ_* be an OK strategy such that

$$\text{MR}(\sigma_*, A) \leq \text{MR}_* + \varepsilon_7.$$

For any other OK strategy σ', we have

$$\text{MR}(\sigma_*, A) \leq \text{MR}(\sigma', A) + \varepsilon_7. \tag{6.10}$$

If some OK strategy σ' satisfied

$$\textsc{Ecost}(\sigma', a) \leq \textsc{Ecost}(\sigma_*, a) - C\varepsilon_7$$

for all $a \in A$ and a large enough constant C, then σ' would violate (6.10). Therefore, σ_* is $C\varepsilon_7$-efficient on A for the class of OK strategies. Thanks to observation (2) and Lemma 6.1, there exists a probability measure μ on A such that

$$\textsc{Ecost}(\sigma_*, \mu) \leq \textsc{Ecost}(\sigma', \mu) + \varepsilon_6 \tag{6.11}$$

for any OK strategy σ'. In particular, (6.11) holds for the optimal Bayesian strategy for μ, denoted $\tilde{\sigma}$. (Here we use observation (1).) It therefore follows from Theorem 4.11 that

$$|\textsc{Ecost}(\sigma_*, a) - \textsc{Ecost}(\tilde{\sigma}, a)| \leq \varepsilon_5 \quad \text{for all } a \in A.$$

Together with (6.10), this shows that

$$\textsc{mr}(\tilde{\sigma}, A) \leq \textsc{mr}(\sigma', A) + C\varepsilon_5 \tag{6.12}$$

for all OK strategies σ'. It may happen that

$$\textsc{Regret}(\tilde{\sigma}, a) \geq \textsc{mr}(\tilde{\sigma}, A) - \varepsilon_3 \quad \text{for all } a \in A. \tag{6.13}$$

In that case, the conclusions of our lemma hold for $\tilde{\sigma}$, μ and $A_0 = A$. Hence, we may assume that (6.13) is false. Let

$$A_0 = \{a \in A : \textsc{mr}(\tilde{\sigma}, A) - \varepsilon_3 \leq \textsc{Regret}(\tilde{\sigma}, a) \leq \textsc{mr}(\tilde{\sigma}, a)\}. \tag{6.14}$$

Thus,

$$\textsc{mr}(\tilde{\sigma}, A) - \varepsilon_3 \leq \textsc{Regret}(\tilde{\sigma}, a) \leq \textsc{mr}(\tilde{\sigma}, A) \quad \text{for } a \in A_0 \tag{6.15}$$

$$\textsc{Regret}(\tilde{\sigma}, a) < \textsc{mr}(\tilde{\sigma}, A) - \varepsilon_3 \quad \text{for } a \in A \backslash A_0. \tag{6.16}$$

Since (6.13) is false, we have $\#A_0 < \#A$, so our INDUCTIVE HYPOTHESIS applies, i.e., our lemma holds for A_0.

Thus, there exist a subset $A_{00} \subset A_0$, a probability measure μ_0, and a strategy $\tilde{\sigma}_0$, with the following properties.

(6.17) μ_0 is concentrated on A_{00}.

(6.18) $\tilde{\sigma}_0$ is the optimal Bayesian strategy for the prior μ_0.

(6.19) $\textsc{Regret}(\tilde{\sigma}_0, a) \leq \textsc{Regret}(\tilde{\sigma}_0, a_0) + \varepsilon_7 \quad \text{for } a \in A_0, a_0 \in A_{00}$.

In particular,

$$|\textsc{Regret}(\tilde{\sigma}_0, a_0) - \textsc{Regret}(\tilde{\sigma}_0, a_0')| \leq \varepsilon_7 \quad \text{for } a_0, a_0' \in A_{00}. \tag{6.20}$$

From (6.19) and (6.20), we see that

$$\textsc{mr}(\tilde{\sigma}_0, A_0) - \varepsilon_6 \leq \textsc{Regret}(\tilde{\sigma}_0, a_0) \leq \textsc{mr}(\tilde{\sigma}_0, A_0) \quad \text{for } a_0 \in A_{00}. \tag{6.21}$$

Our plan is to prove that the conclusions of Lemma 6.2 for A hold for the set A_{00}, the measure μ_0, and the strategy $\tilde{\sigma}_0$; that will complete our induction on $\#A$ and prove Lemma 6.2. To carry out our plan, we first prove that

$$|\mathrm{MR}(\tilde{\sigma}, A) - \mathrm{MR}(\tilde{\sigma}_0, A_0)| \leq \varepsilon_2. \tag{6.22}$$

To see (6.22), we recall that

$$\mathrm{REGRET}(\sigma, a) = \rho_0(a) + \rho_1(a) \cdot \mathrm{ECOST}(\sigma, a)$$

with $c < \rho_1(a) < C$.

For $a_0 \in A_{00} \subset A_0$, estimates (6.15) and (6.21) therefore imply the inequalities

$$\left[\frac{\mathrm{MR}(\tilde{\sigma}, A) - \rho_0(a)}{\rho_1(a)}\right] - C\varepsilon_3 \leq \mathrm{ECOST}(\tilde{\sigma}, a) \leq \left[\frac{\mathrm{MR}(\tilde{\sigma}, A) - \rho_0(a)}{\rho_1(a)}\right], \tag{6.23}$$

$$\left[\frac{\mathrm{MR}(\tilde{\sigma}_0, A_0) - \rho_0(a)}{\rho_1(a)}\right] - C\varepsilon_6 \leq \mathrm{ECOST}(\tilde{\sigma}_0, a) \leq \left[\frac{\mathrm{MR}(\tilde{\sigma}_0, A_0) - \rho_0(a)}{\rho_1(a)}\right]. \tag{6.24}$$

Let

$$H_1 = \sum_{a \in A_{00}} \frac{\mu_0(a)}{\rho_1(a)} \quad \text{and} \quad H_0 = \sum_{a \in A_{00}} \frac{\mu_0(a)\rho_0(a)}{\rho_1(a)}. \tag{6.25}$$

Since μ_0 is a probability measure concentrated on A_{00}, and since $c < \rho_1(a) < C$, we have

$$c' < H_1 < C'. \tag{6.26}$$

Multiplying (6.23) and (6.24) by $\mu_0(a)$, and summing over $a \in A_{00}$, we obtain the inequalities

$$H_1\,\mathrm{MR}(\tilde{\sigma}, A) - H_0 - C\varepsilon_3 \leq \mathrm{ECOST}(\tilde{\sigma}, \mu_0) \leq H_1\,\mathrm{MR}(\tilde{\sigma}, A) - H_0, \tag{6.27}$$

$$H_1\,\mathrm{MR}(\tilde{\sigma}_0, A_0) - H_0 - C\varepsilon_6 \leq \mathrm{ECOST}(\tilde{\sigma}_0, \mu_0) \leq H_1\,\mathrm{MR}(\tilde{\sigma}_0, A_0) - H_0. \tag{6.28}$$

Moreover, since $\tilde{\sigma}_0$ is the optimal Bayesian strategy for the prior μ_0, we have

$$\mathrm{ECOST}(\tilde{\sigma}_0, \mu_0) \leq \mathrm{ECOST}(\tilde{\sigma}, \mu_0). \tag{6.29}$$

From (6.27), (6.28) and (6.29), we see that

$$H_1\,\mathrm{MR}(\tilde{\sigma}_0, A_0) - H_0 - C\varepsilon_6 \leq \mathrm{ECOST}(\tilde{\sigma}_0, \mu_0) \leq \mathrm{ECOST}(\tilde{\sigma}, \mu_0) \leq H_1\,\mathrm{MR}(\tilde{\sigma}, A) - H_0.$$

Thus,

$$H_1\mathrm{MR}(\tilde{\sigma}_0, A_0) \leq H_1\mathrm{MR}(\tilde{\sigma}, A) + C\varepsilon_6.$$

Thanks to (6.26), this tells us that

$$\mathrm{MR}(\tilde{\sigma}_0, A_0) \leq \mathrm{MR}(\tilde{\sigma}, A) + C\varepsilon_6.$$

So we have proven half of (6.22). In particular, (6.22) holds unless we have

$$\mathrm{MR}(\tilde{\sigma}_0, A_0) \leq \mathrm{MR}(\tilde{\sigma}, A) - \varepsilon_2. \tag{6.30}$$

To complete the proof of (6.22), we assume (6.30) and derive a contradiction as follows. Observation (1) tells us that both strategies $\tilde{\sigma}$ and $\tilde{\sigma}_0$ are OK. We form a mixed strategy σ_{MIX} by playing the strategy $\tilde{\sigma}$ with probability $(1 - \varepsilon_4)$ and the strategy $\tilde{\sigma}_0$ with probability ε_4. Observation (2) tells us that σ_{MIX} is an OK strategy. We will see that (6.30) implies that σ_{MIX} outperforms $\tilde{\sigma}$, contradicting (6.12). To see this, we first recall that since $\tilde{\sigma}$ and $\tilde{\sigma}_0$ are tame strategies, we have $\text{COST}(\tilde{\sigma}, a)$, $\text{COST}(\tilde{\sigma}_0, a) \leq C$ for all $a \in A$, hence

$$\text{(6.31)} \qquad \text{REGRET}(\tilde{\sigma}_0, a) \leq \text{REGRET}(\tilde{\sigma}, a) + C \quad \text{for any } a \in A.$$

Now suppose $a \in A_0$. Then (6.30) yields the inequalities

$$\begin{aligned} \text{REGRET}(\sigma_{\text{MIX}}, a) &= (1-\varepsilon_4)\,\text{REGRET}(\tilde{\sigma}, a) + \varepsilon_4\,\text{REGRET}(\tilde{\sigma}_0, a) \\ &\leq (1-\varepsilon_4)\,\text{MR}(\tilde{\sigma}, A) + \varepsilon_4\,\text{MR}(\tilde{\sigma}_0, A_0) \\ &\leq (1-\varepsilon_4)\,\text{MR}(\tilde{\sigma}, A) + \varepsilon_4\,[\text{MR}(\tilde{\sigma}, A) - \varepsilon_2] = \text{MR}(\tilde{\sigma}, A) - \varepsilon_4\varepsilon_2. \end{aligned}$$

On the other hand, for $a \in A \setminus A_0$, inequalities (6.16) and (6.31) imply that

$$\begin{aligned} \text{REGRET}(\sigma_{\text{MIX}}, a) &= (1-\varepsilon_4)\,\text{REGRET}(\tilde{\sigma}, a) + \varepsilon_4\,\text{REGRET}(\tilde{\sigma}_0, a) \\ &\leq (1-\varepsilon_4)\,[\text{MR}(\tilde{\sigma}, A) - \varepsilon_3] + \varepsilon_4\,[\text{MR}(\tilde{\sigma}, A) + C] \\ &= \text{MR}(\tilde{\sigma}, A) + C\varepsilon_4 - (1-\varepsilon_4)\varepsilon_3 \leq \text{MR}(\tilde{\sigma}, A) - \frac{1}{2}\varepsilon_3 \end{aligned}$$

(since $\varepsilon_4 \ll \varepsilon_3 \ll 1$). Thus, for all $a \in A$, we have

$$\text{REGRET}(\sigma_{\text{MIX}}, a) \leq \text{MR}(\tilde{\sigma}, A) - \min\left\{\frac{1}{2}\varepsilon_3, \varepsilon_4\varepsilon_2\right\} = \text{MR}(\tilde{\sigma}, A) - \varepsilon_2\varepsilon_4.$$

In other words,

$$\text{MR}(\sigma_{\text{MIX}}, A) \leq \text{MR}(\tilde{\sigma}, A) - \varepsilon_4\varepsilon_2.$$

As promised, this contradicts (6.12), completing the proof of (6.22).

Returning to (6.27) and (6.28), we now see that

$$\text{(6.32)} \qquad |\text{ECOST}(\tilde{\sigma}, \mu_0) - \text{ECOST}(\tilde{\sigma}_0, \mu_0)| \leq C\varepsilon_3,$$

thanks to (6.22) and (6.26).

Since $\tilde{\sigma}_0$ is the optimal Bayesian strategy for μ_0, and since $\tilde{\sigma}$ is tame with tame constant at most C, (6.32) and Theorem 4.11 together imply that

$$\text{(6.33)} \qquad |\text{ECOST}(\tilde{\sigma}, a) - \text{ECOST}(\tilde{\sigma}_0, a)| \leq \varepsilon_1 \quad \text{for all } a \in A.$$

We are ready to show that A_{00}, μ_0 and $\tilde{\sigma}_0$ satisfy the conclusions of Lemma 6.2 for A. Indeed, we know that μ_0 is a probability measure concentrated on A_{00}, and that $\tilde{\sigma}_0$ is the optimal Bayesian strategy for the prior μ_0. It remains only to show that

$$\text{(6.34)} \qquad \text{REGRET}(\tilde{\sigma}_0, a) \leq \text{REGRET}(\tilde{\sigma}_0, a_0) + \varepsilon \quad \text{for any } a_0 \in A_{00}, a \in A.$$

However, (6.33) yields

$$\text{(6.35)} \qquad \text{REGRET}(\tilde{\sigma}_0, a) \leq \text{REGRET}(\tilde{\sigma}, a) + C\varepsilon_1 \leq \text{MR}(\tilde{\sigma}, A) + C\varepsilon_1 \quad \text{for } a \in A,$$

while (6.21) and (6.22) yield

$$\text{(6.36)} \qquad \text{REGRET}(\tilde{\sigma}_0, a_0) \geq \text{MR}(\tilde{\sigma}_0, A_0) - \varepsilon_6 \geq \text{MR}(\tilde{\sigma}, A) - C\varepsilon_2$$

for $a_0 \in A_{00}$.

The desired estimate (6.34) is immediate from (6.35) and (6.36). Thus, as promised, the conclusions of our lemma hold, with A_{00}, μ_0 and $\tilde{\sigma}_0$ in place of A_0, μ and $\tilde{\sigma}$. Our induction on $\#A$ is complete, and Lemma 6.2 is proven. ∎

6.5. An interval of allowed parameters

The previous section produced nearly optimal agnostic strategies when the parameter a is known to belong to a finite set. In this section, we pass to the case in which a is known merely to belong to a given interval $[-a_{\text{MAX}}, +a_{\text{MAX}}]$. We continue to suppose our PDE assumption (from Section 4.3) holds for every prior probability distribution on the interval $[-a_{\text{MAX}}, a_{\text{MAX}}]$. Under this assumption, we will prove the following result.

Theorem 6.3. *There exists a Bayesian prior probability measure μ_∞ supported on a subset $A_\infty \subset [-a_{\text{MAX}}, +a_{\text{MAX}}]$, for which the optimal Bayesian strategy σ_∞ satisfies*

(A) *the function $a \mapsto \text{REGRET}(\sigma_\infty, a)$ is constant on A_∞, and*

(B) *the function $[-a_{\text{MAX}}, a_{\text{MAX}}] \ni a \mapsto \text{REGRET}(\sigma_\infty, a)$ is maximized on A_∞.*

Proof. Let $A_1, A_2, A_3, \ldots$ be a sequence of sets of the form

(6.37) $\quad A_N = [-a_{\text{MAX}}, a_{\text{MAX}}] \cap 2^{-m_N}\mathbb{Z}$, with $m_N \to \infty$ as $N \to \infty$; and let

(6.38) $\quad \varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots$ be a sequence of positive numbers tending to zero.

Applying Lemma 6.2 to each A_N, we obtain a probability measure μ_N, concentrated on a subset $A_N^0 \subset A_N$, such that the optimal Bayesian strategy σ_N for the prior μ_N satisfies

$$\text{MR}_N - \varepsilon_N \le \text{REGRET}(\sigma_N, a^0) \le \text{MR}_N \quad \text{for all } a^0 \in A_N^0, \text{ where} \tag{6.39}$$

$$\text{MR}_N = \max\{\text{REGRET}(\sigma_N, a) : a \in A_N\}. \tag{6.40}$$

Passing to a subsequence, we may assume that the μ_N converge weakly to a probability measure μ_∞ on $[-a_{\text{MAX}}, +a_{\text{MAX}}]$. Again passing to a subsequence, we may assume that the MR_N converge to a limit MR_∞ as $N \to \infty$. (Here we use the fact that the MR_N are bounded, thanks to Lemma 3.2.)

Let σ_∞ be the optimal Bayesian strategy for the prior μ_∞. After again passing to a subsequence, we will show that

$$\text{REGRET}(\sigma_N, a) \to \text{REGRET}(\sigma_\infty, a) \quad \text{as } N \to \infty, \tag{6.41}$$

uniformly for $a \in [-a_{\text{MAX}}, +a_{\text{MAX}}]$. The proof of (6.41) is the main step in our argument.

Let us recall how $\text{REGRET}(\sigma_N, a)$ and $\text{REGRET}(\sigma_\infty, a)$ are defined. Starting from the prior μ_N, we form the functions

$$\bar{a}_N(\zeta_1, \zeta_2) = \frac{\int_{-a_{\text{MAX}}}^{a_{\text{MAX}}} a \exp\left(-\frac{a^2}{2}\zeta_2 + a\zeta_1\right) d\mu_N(a)}{\int_{-a_{\text{MAX}}}^{a_{\text{MAX}}} \exp\left(-\frac{a^2}{2}\zeta_2 + a\zeta_1\right) d\mu_N(a)},$$

and similarly define $\bar{a}_\infty(\zeta_1, \zeta_2)$.

Using $\bar{a}_N$ in place of $\bar{a}$ in (4.21), we then obtain a PDE solution $S_N(q, t, \zeta_1, \zeta_2) \in C^{2,1}_{\text{loc}}(\mathbb{R} \times [0, T] \times \mathbb{R} \times [0, \infty))$, satisfying the conditions given in Section 4.3. Thanks to the estimates on $\partial^\alpha S_N$ ($|\alpha| \le 3$) in that section, we may again pass to a subsequence, and assume that

$$\partial^\alpha S_N \to \partial^\alpha S_\infty \quad \text{as } N \to \infty \text{ for } |\alpha| \le 2, \tag{6.42}$$

uniformly on compact subsets of $\mathbb{R} \times [0, T] \times \mathbb{R} \times [0, \infty)$. Here, $S_\infty \in C^{2,1}_{\text{loc}}$ satisfies all the estimates given in Section 4.3. Since the μ_N converge weakly to μ_∞, it follows that $\bar{a}_N(\zeta_1, \zeta_2) \to \bar{a}_\infty(\zeta_1, \zeta_2)$ as $N \to \infty$ for each $(\zeta_1, \zeta_2) \in \mathbb{R} \times [0, \infty)$. Together with (6.42) and the PDE satisfied by the S_N, this proves that S_∞ satisfies the PDE (4.21) for the Bayesian prior μ_∞.

Now define

$$u_N(q, t, \zeta_1, \zeta_2) = -\frac{1}{2}\, \partial_q S_N(q, t, \zeta_1, \zeta_2) \quad \text{and} \quad u_\infty(q, t, \zeta_1, \zeta_2) = -\frac{1}{2}\, \partial_q S_\infty(q, t, \zeta_1, \zeta_2)$$

for $(q, t, \zeta_1, \zeta_2) \in \mathbb{R} \times [0, T] \times \mathbb{R} \times [0, \infty)$. From (6.42) we have

$$u_N \to u_\infty \quad \text{as } N \to \infty, \tag{6.43}$$

uniformly on compact subsets of $\mathbb{R} \times [0, T] \times \mathbb{R} \times [0, \infty)$.

For each $k \ge 1$, we introduce the partition π_k of $[0, T]$ given by

$$0 = t_0^k < t_1^k < \cdots < t_k^k = T, \quad \text{with } t_\nu^k = \frac{\nu}{k}\, T. \tag{6.44}$$

Let $\sigma(N, k)$ be the ALLEGEDLY OPTIMAL STRATEGY for the Bayesian prior μ_N and the partition π_k. Thus,

$$u^{\sigma(N,k)}(t_\nu^k) = u_N(q^{\sigma(N,k)}(t_\nu^k), t_\nu^k, \zeta_1^{\sigma(N,k)}(t_\nu^k), \zeta_2^{\sigma(N,k)}(t_\nu^k)) \quad \text{for } 0 \le \nu < k. \tag{6.45}$$

Similarly, let $\sigma(\infty, k)$ be the ALLEGEDLY OPTIMAL STRATEGY for the Bayesian prior μ_∞ and the partition π_k. Thus,

$$u^{\sigma(\infty,k)}(t_\nu^k) = u_\infty(q^{\sigma(\infty,k)}(t_\nu^k), t_\nu^k, \zeta_1^{\sigma(\infty,k)}(t_\nu^k), \zeta_2^{\sigma(\infty,k)}(t_\nu^k)) \quad \text{for } 0 \le \nu < k. \tag{6.46}$$

By definition,

$$\text{ECOST}(\sigma_N, a) = \lim_{k\to\infty} \text{ECOST}(\sigma(N, k), a), \tag{6.47}$$

$$\text{ECOST}(\sigma_\infty, a) = \lim_{k\to\infty} \text{ECOST}(\sigma(\infty, k), a), \tag{6.48}$$

for $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$. Finally,

$$\text{REGRET}(\sigma_N, a) = \rho_0(a) + \rho_1(a)\text{ECOST}(\sigma_N, a), \tag{6.49}$$

$$\text{REGRET}(\sigma_\infty, a) = \rho_0(a) + \rho_1(a)\text{ECOST}(\sigma_\infty, a), \tag{6.50}$$

for $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$. This concludes our review of the definition of $\text{REGRET}(\sigma_N, a)$ and $\text{REGRET}(\sigma_\infty, a)$.

For large enough N and k, we will apply Lemma 4.8 of Section 4.6 to the Bayesian prior μ_∞, the ALLEGEDLY OPTIMAL STRATEGY $\sigma(\infty, k)$, and the alternative strategy

$\sigma(N,k)$. Thus, u_∞ will play the rôle of u_{opt} in Section 4.6, while $u^{\sigma(N,k)}$, given by (6.45), will play the rôle of u^σ in Section 4.6. The rôle of the quantity DISCREP^σ in Section 4.6 will therefore be played by

$$\begin{aligned}\text{(6.51)}\qquad \text{DISCREP}(N,k,t_\nu^k) &= u_N(q^{\sigma(N,k)}(t_\nu^k), t_\nu^k, \zeta_1^{\sigma(N,k)}(t_\nu^k), \zeta_2^{\sigma(N,k)}(t_\nu^k)) \\ &\quad - u_\infty(q^{\sigma(N,k)}(t_\nu^k), t_\nu^k, \zeta_1^{\sigma(N,k)}(t_\nu^k), \zeta_2^{\sigma(N,k)}(t_\nu^k)).\end{aligned}$$

To apply Lemma 4.8, we must estimate

$$\text{(6.52)}\qquad \mathrm{E}_a\Big[\sum_{0\le\nu<k} |\text{DISCREP}(N,k,t_\nu^k)|^2 \Delta t_\nu^k\Big] \quad \text{for } a \in [-a_{\text{MAX}}, a_{\text{MAX}}],$$

with $\Delta t_\nu^k = t_{\nu+1}^k - t_\nu^k = T/k$ (see (6.44)). To estimate the quantity in (6.52), we recall that

$$|u_N(q,t,\zeta_1,\zeta_2)|, |u_\infty(q,t,\zeta_1,\zeta_2)| \le C[|q|+1],$$

and, consequently,

$$\text{(6.53)}\qquad |\text{DISCREP}(N,k,t_\nu^k)| \le C[|q^{\sigma(N,k)}(t_\nu^*)| + 1].$$

For $Q \ge C$, define the events

$$\text{BAD}(N,k,Q) := \Big\{ \max_{0\le\nu<k} \{|q^{\sigma(N,k)}(t_\nu^k)| + |\zeta_1^{\sigma(N,k)}(t_\nu^k)| + |\zeta_2^{\sigma(N,k)}(t_\nu^k)|\} > Q \Big\},$$

$$\text{GOOD}(N,k,Q) := \Big\{ \max_{0\le\nu<k} \{|q^{\sigma(N,k)}(t_\nu^k)| + |\zeta_1^{\sigma(N,k)}(t_\nu^k)| + |\zeta_2^{\sigma(N,k)}(t_\nu^k)|\} \le Q \Big\}.$$

We take $k \ge C$ for a large C, so that our partition of $[0,T]$ is fine enough to allow us to apply Lemma 3.2. Lemma 3.2 then tells us that

$$\mathrm{E}_a\big[\max_{0\le\nu<k} \{|q^{\sigma(N,k)}(t_\nu^k)| + 1\}^2 \cdot \mathbb{1}_{\text{BAD}(N,k,Q)}\big] \le CQ^{-1}$$

for $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$. Hence, by (6.53),

$$\text{(6.54)}\qquad \mathrm{E}_a\Big[\sum_{0\le\nu<k} (\text{DISCREP}(N,k,t_\nu^k))^2\, \Delta t_\nu^k \cdot \mathbb{1}_{\text{BAD}(N,k,Q)}\Big] \le CQ^{-1}$$

for $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$. On the other hand, let $\delta > 0$ be given, and suppose

$$N \ge N_{\min}(\delta, Q) \quad \text{for a large enough } N_{\min}(\delta, Q).$$

Then, by comparing (6.43) with (6.51), we see that

$$|\text{DISCREP}(N,k,t_\nu^k)| \le \delta \quad \text{for all } \nu,$$

provided $\text{GOOD}(N,k,Q)$ occurs. Therefore,

$$\mathrm{E}_a\Big[\sum_{0\le\nu<k} (\text{DISCREP}(N,k,t_\nu^k))^2\, \Delta t_\nu^k \cdot \mathbb{1}_{\text{GOOD}(N,k,Q)}\Big] \le C\delta$$

for all $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$. Together with (6.54), this implies that

$$\mathrm{E}_a\Big[\sum_{0\le\nu<k}(\textsc{Discrep}(N,k,t_\nu^k))^2\,\Delta t_\nu^k\Big]\le C\delta + CQ^{-1}$$

for all $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$ provided $N \ge N_{\min}(\delta, Q)$ and $k \ge C$. Taking $Q = \delta^{-1}$, we have

$$\mathrm{E}_a\Big[\sum_{0\le\nu<k}(\textsc{Discrep}(N,k,t_\nu^k))^2\,\Delta t_\nu^k\Big]\le C\delta \tag{6.55}$$

for $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$, $N \ge N'_{\min}(\delta)$, and $k \ge C$. Also, recalling (6.44), we see that

$$\Delta t^k_{\text{MAX}} \equiv \max_{0\le\nu<k}(t^k_{\nu+1}-t^k_\nu)\le \frac{C}{k} < \delta, \quad \text{provided } k > \frac{C}{\delta}. \tag{6.56}$$

Now let $\varepsilon > 0$ be given, and let δ be small enough, depending on ε. Our results (6.55) and (6.56) are the hypotheses of Lemma 4.8, with μ_∞ in place of dPrior, and with $\sigma(N,k)$ in place of σ. Applying that lemma, we learn that

$$|\textsc{Ecost}(\sigma(N,k),a) - \textsc{Ecost}(\sigma(\infty,k),a)| \le \varepsilon,$$

all $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$, for $k \ge k_{\min}(\varepsilon)$ and $N \ge N''_{\min}(\varepsilon)$. Passing to the limit as $k \to \infty$ for fixed N, and recalling (6.47) and (6.48), we see that

$$|\textsc{Ecost}(\sigma_N,a) - \textsc{Ecost}(\sigma_\infty,a)| \le \varepsilon$$

for $N \ge N''(\varepsilon)$ and for all $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$. Since $\varepsilon > 0$ is arbitrary, we conclude that

$$\textsc{Ecost}(\sigma_N,a) \to \textsc{Ecost}(\sigma_\infty,a) \quad \text{as } N \to \infty,$$

uniformly for $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$. Thanks to (6.49) and (6.50), this in turn implies that

$$\textsc{Regret}(\sigma_N,a) \to \textsc{Regret}(\sigma_\infty,a) \quad \text{as } N \to \infty,$$

uniformly for $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$. So, at last, we have proven (6.41).

Notice that the functions $\textsc{Regret}(\sigma_N, a)$ and $\textsc{Regret}(\sigma_\infty, a)$ are continuous on $[-a_{\text{MAX}}, a_{\text{MAX}}]$. Thanks to (6.41), they have a common modulus of continuity, i.e.,

$$|\textsc{Regret}(\sigma_N,a_1) - \textsc{Regret}(\sigma_N,a_2)| \le \omega(|a_1-a_2|) \tag{6.57}$$

for $a_1, a_2 \in [-a_{\text{MAX}}, a_{\text{MAX}}]$ and for all $N \ge 1$, for a function $\omega(t)$ satisfying

$$\omega(t) \to 0 \quad \text{as } t \to 0^+. \tag{6.58}$$

We have defined the probability measure μ_∞ and its optimal Bayesian strategy σ_∞. To complete the proof of our theorem, we must define a set $A_\infty \subset [-a_{\text{MAX}}, a_{\text{MAX}}]$ and prove that

- μ_∞ is supported on A_∞,
- $\textsc{Regret}(\sigma_\infty, a)$ is constant on A_∞, and
- $\textsc{Regret}(\sigma_\infty, a)$ is maximized on A_∞ over all $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$.

We define A_∞ to consist of all $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$ such that for all $\eta > 0$ and all $N_* \ge 1$ there exists $a^0 \in A^0_N \cap (a-\eta, a+\eta)$ for some $N > N_*$. (Recall A^0_N from the defining conditions for the μ_N.)

Let us check that μ_∞ is supported in A_∞. Thus, let $a \in [-a_{\text{MAX}}, a_{\text{MAX}}] \backslash A_\infty$. Then for some open interval $I \ni a$ and some $N_* \geq 1$, we have $A_N^0 \cap I = \emptyset$ for $N > N_*$. Since μ_N is supported in A_N^0, we have $\mu_N(I) = 0$. Since the probability measures μ_N converge weakly to μ_∞ as $N \to \infty$, it follows that $\mu_\infty(I) = 0$. So $a \notin \text{support}(\mu_\infty)$, completing the proof that μ_∞ is supported in A_∞.

Next, suppose $a^0 \in A_\infty$. Then there exist sequences $N_\nu \to \infty$ and $a_\nu \to a^0$ as $\nu \to \infty$, with $a_\nu \in A_{N_\nu}^0$. From (6.39) we have

$$\text{MR}_{N_\nu} - \varepsilon_{N_\nu} \leq \text{REGRET}(\sigma_N, a_\nu) \leq \text{MR}_{N_\nu},$$

hence, thanks to (6.57),

$$\text{MR}_{N_\nu} - \varepsilon_{N_\nu} - \omega(|a_\nu - a^0|) \leq \text{REGRET}(\sigma_N, a^0) \leq \text{MR}_{N_\nu} + \omega(|a_\nu - a^0|). \tag{6.59}$$

As $\nu \to \infty$, we have $\text{MR}_{N_\nu} \to \text{MR}_\infty$, $\varepsilon_{N_\nu} \to 0$, and $\omega(|a_\nu - a^0|) \to 0$ thanks to (6.58). Therefore, (6.59) implies that

$$\lim_{N\to\infty} \text{REGRET}(\sigma_N, a^0) = \text{MR}_\infty.$$

Recalling (6.41), we see that

$$\text{REGRET}(\sigma_\infty, a^0) = \text{MR}_\infty \quad \text{for all } a^0 \in A_\infty. \tag{6.60}$$

On the other hand, let $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$. From (6.37) we obtain a sequence $a_N \in A_N$ $(N \geq 1)$ such that $a_N \to a$ as $N \to \infty$. Thanks to (6.40), we have

$$\text{REGRET}(\sigma_N, a_N) \leq \text{MR}_N \quad \text{for each } N,$$

hence

$$\text{REGRET}(\sigma_N, a) \leq \text{MR}_N + \omega(|a_N - a|), \tag{6.61}$$

by (6.57). As $N \to \infty$, we have $\text{MR}_N \to \text{MR}_\infty$ and $\omega(|a_N - a|) \to 0$ by (6.58). Therefore, (6.61) and (6.41) yield the inequality

$$\text{REGRET}(\sigma_\infty, a) \leq \text{MR}_\infty \quad \text{for all } a \in [-a_{\text{MAX}}, a_{\text{MAX}}]. \tag{6.62}$$

From (6.60) and (6.62), we see that $\text{REGRET}(\sigma_\infty, a)$ is constant on A_∞, and that the maximum of $\text{REGRET}(\sigma_\infty, a)$ over all $a \in [-a_{\text{MAX}}, a_{\text{MAX}}]$ is achieved on A_∞. The proof of our theorem is complete. ∎

Under additional assumptions on the functions $\rho_0(a)$ and $\rho_1(a)$ in Section 6.3, we can easily deduce that the set A_∞ in the above theorem is finite. Indeed, suppose ρ_0 and ρ_1 are real-analytic on $\mathbb{R}$, and suppose that for all $\varepsilon > 0$ we have

$$\rho_0(t) \geq -\exp(\varepsilon t) \text{ and } \rho_1(t) \geq \exp(-\varepsilon t) \quad \text{for large positive } t. \tag{6.63}$$

Recall that the function $a \mapsto \text{ECOST}(\sigma_\infty, a)$ is real-analytic on $\mathbb{R}$ and grows exponentially as $a \to \infty$. (See Theorem 4.11.)

Under our assumptions on ρ_0 and ρ_1, it follows that the function

$$a \mapsto \text{REGRET}(\sigma_\infty, a) = \rho_0(a) + \rho_1(a)\,\text{ECOST}(\sigma_\infty, a)$$

is again real-analytic on $\mathbb{R}$ and exponentially large as $a \to \infty$.

In particular, $[-a_{\text{MAX}}, a_{\text{MAX}}] \ni a \mapsto \text{REGRET}(\sigma_\infty, a)$ is a nonconstant real-analytic function. Since $\text{REGRET}(\sigma_\infty, a)$ is constant on A_∞, it follows that A_∞ is finite, as claimed.

Combining Theorem 6.3 with the fact that A_∞ is finite establishes parts (I), (II) and (III) of Theorem 1.3 in the introduction.

Acknowledgments. We thank Amir Ali Ahmadi, Brittany Hamfeldt, Elad Hazan, Robert Kohn, Sam Otto, Allan Sly, and Melanie Weber for helpful conversations, and we thank the anonymous referee for helpful feedback. We are grateful to the Air Force Office of Scientific Research, especially 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.: Controlling unknown linear dynamics with almost optimal regret. To appear in *Rev. Mat. Iberoam.* (2024), published online first.

[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