Large-Time and Small-Ball Asymptotics for Quadratic Functionals of Gaussian Diffusions

S. V. LotFotsky*
Department of Mathematics, USC
Los Angeles, CA 90089 USA

M. Moers
Department of Mathematics, USC
Los Angeles, CA 90089 USA

November 10, 2015

Abstract

Using asymptotic analysis of the Laplace transform, we establish almost sure divergence of certain integrals and derive logarithmic asymptotic of small ball probabilities for quadratic forms of Gaussian diffusion processes. The large time behavior of the quadratic forms exhibits little dependence on the drift and diffusion matrices or the initial conditions, and, if the noise driving the equation is not degenerate, then similar universality also holds for small ball probabilities. On the other hand, degenerate noise leads to a variety of different asymptotics of small ball probabilities, including unexpected influence of the initial conditions.

AMS Subject Classification 60G15, 60J60.

Keywords: algebraic Riccati equation, logarithmic asymptotic, Ornstein-Uhlenbeck process, small ball constant, small ball rate.

1 Introduction

Given a Gaussian process $\boldsymbol{y} = \boldsymbol{y}(s),\ 0 < s < t$, with values in $\mathbb{R}^{\mathrm{d}}$ and a non-negative-definite symmetric matrix $Q \in \mathbb{R}^{\mathrm{d}\times\mathrm{d}}$, how large and how small can the random variable

$$\int_0^t \boldsymbol{y}^\top(s) Q \boldsymbol{y}(s) ds$$

be? For example,

[Q1] Does the integral

$$\int_0^\infty \boldsymbol{y}^\top(s) Q \boldsymbol{y}(s) ds$$

diverge with probability one? [While the expected value of the integral is easy to study, the almost-sure divergence is non-trivial unless $\boldsymbol{y}$ is ergodic.]

*Corresponding author. E-mail: lototsky@usc.edu, tel. (+1) 213 740 2389, fax: (+1) 213 740 2424.
Published in Asymptotic Analysis, Vol. 95 (2015) pp. 345–374, DOI 10.3233/ASY-151328 IOS Press

[Q2] What is the asymptotic of

$$\mathbb{P}\left(\int_0^t \boldsymbol{y}^\top(s)Q\boldsymbol{y}(s)ds \le \varepsilon\right)$$

as $0 < \varepsilon \to 0$?

Questions of this type arise in the analysis of various statistical estimators [12, Chpater 17], and in the study of Gaussian measures on Hilbert spaces [10], and, while the scalar case ($\mathrm{d} = 1$) has been getting a lot of the attention, much less is known in multi-dimensional setting.

The objective of this paper is to investigate questions [Q1] and [Q2] for a particular class of multi-dimensional Gaussian processes, namely, Gaussian diffusions. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ be a stochastic basis with an m-dimensional standard Brownian motion $\boldsymbol{w}$, and let $A \in \mathbb{R}^{\mathrm{d}\times\mathrm{d}}$ and $B \in \mathbb{R}^{\mathrm{d}\times m}$ be constant non-random matrices; $m \le \mathrm{d}$. Let $\boldsymbol{y} = \boldsymbol{y}(t)$ be the solution of

$$d\boldsymbol{y}(t) = A\boldsymbol{y}(t)dt + Bd\boldsymbol{w}(t),\ t > 0, \tag{1.1}$$

with initial condition $\boldsymbol{y}(0)$ that is independent of $\boldsymbol{w}$ and is a Gaussian vector with mean $\boldsymbol{m}$ and covariance $K \ge 0$. We refer to $\boldsymbol{y}$ as a `multi-dimensional Gaussian diffusion`; a popular alternative name is a multi-dimensional Ornstein-Uhlenbeck process.

Throughout the paper, a column vector is denoted by a lower-case bold letter (Greek or Latin), e.g. $\boldsymbol{\mu}$ or $\boldsymbol{y}$, whereas an upper-case regular Latin letter, e.g. A, means a matrix; the identity matrix is I. For a matrix A, $A^\top$ means transposition. The same notation $^\top$ will also be used for column vectors to produce a row vector. The notation $A \ge 0$ means that A is a symmetric non-negative-definite matrix: $A = A^\top$ and $\boldsymbol{x}^\top A\boldsymbol{x} \ge 0$; for such matrices, $A^{1/2}$ denotes the non-negative-definite symmetric square root of A. The trace of a square matrix A is $\mathrm{Tr}(A)$ and the determinant is $\det(A)$. The Euclidean norm of a vector and the induced matrix norm are both denoted by $\|\cdot\|$. The notation $\dot{}$, as in $\dot{\mathfrak{A}}$, means the derivative of the function $\mathfrak{A} = \mathfrak{A}(t)$ [scalar, vector, or matrix] with respect to t. Zero matrix and zero number are both 0; zero vector is $\mathbf{0}$.

One way to address both [Q1] and [Q2] is to investigate the Laplace transform function

$$\Psi(Q;t) = \mathbb{E}\exp\left(-\int_0^t \boldsymbol{y}^\top(s)Q\boldsymbol{y}(s)ds\right) \tag{1.2}$$

and then to apply a suitable Tauberian theorem. This approach requires asymptotic analysis of the function Ψ in various regimes, which, in turn, requires a *workable* closed-form expression for Ψ. Fortunately, the paper [8] and the book [9] provide all the necessary tools to carry out the asymptotic analysis of (1.2) when $\boldsymbol{y}$ satisfies (1.1) and the noise is non-degenerate in the sense that the matrix $BB^\top$ has rank d. The resulting answers to [Q1] and [Q2] turn out rather universal in the sense that there is minimal dependence on the drift matrix A and the initial condition $\boldsymbol{y}(0)$. In particular,

[A1] The integral

$$\int_0^\infty \boldsymbol{y}^\top(s)Q\boldsymbol{y}(s)ds$$

diverges with probability one (as long as it is not identically zero, which only happens when $\mathrm{Tr}(Q) = 0$). In fact, this divergence takes place under a much weaker condition than non-degeneracy of $BB^\top$, namely, when the pair (A, B) is controllable; see Theorem 4.3 below. The proof relies on the large-time asymptotic of $\Psi(Q;t)$, that is, analysis of $\Psi(Q;t)$ as $t \to \infty$.

[A2] If $B = I$ and the covariance matrix K of the initial condition is non-singular, then

$$\lim_{\varepsilon\to 0+} \varepsilon \ln \mathbb{P}\left(\int_0^t \boldsymbol{y}^\top(s)Q\boldsymbol{y}(s)ds \le \varepsilon\right) = \lim_{\varepsilon\to 0+} \varepsilon \ln \mathbb{P}\left(\int_0^t \boldsymbol{w}^\top(s)Q\boldsymbol{w}(s)ds \le \varepsilon\right) = -\frac{t^2}{8}\Big(\mathrm{Tr}\big(Q^{1/2}\big)\Big)^2. \tag{1.3}$$

Theorem 4.5 below provides the general result, which covers $B \ne I$ and a singular matrix K. The proof relies on high frequency asymptotic of $\Psi(Q;t)$, that is, analysis of $\Psi(\lambda Q;t)$ as $\lambda \to \infty$.

The paper is organized as follows. Section 2 presents the necessary background on the Laplace transform and small ball probabilities. Section 3 summarizes the main properties of the solution of (1.1), including the formula for $\Psi(Q;t)$. Section 4 contains the main contributions of the paper, related to items [A1] and [A2] above. Section 5 demonstrates how degenerate matrix $BB^\top$ can dramatically change (1.3).

2 Background

The main challenge in answering question [Q1] is often finding a rigorous proof of an "obvious" result. Question [Q2] presents a somewhat different challenge: getting useful information from the general answer. Indeed, after diagonalizing the matrix Q and expanding the process $\boldsymbol{y}$ in the eigenfunctions of its covariance operator, and under an additional assumption that $\mathbb{E}\boldsymbol{y}(t) = \boldsymbol{0}$, $t \ge 0$, we get

$$\int_0^t \boldsymbol{y}^\top(s)Q\boldsymbol{y}(s)ds = \sum_{k\ge 1} \lambda_k \xi_k^2$$

for some $\lambda_k > 0$ and iid standard normal ξ_k. Then, as shown in [16],

$$\mathbb{P}\left(\int_0^t \boldsymbol{y}^\top(s)Q\boldsymbol{y}(s)ds \le \varepsilon^2\right) \sim \left(4\pi \sum_{n\ge 1}\left(\frac{\lambda_n\gamma(\varepsilon)}{1+2\lambda_n\gamma(\varepsilon)}\right)^2\right)^{-1/2} \times \exp\left(\varepsilon^2\gamma(\varepsilon) - \frac{1}{2}\sum_{n\ge 1}\ln\big(1+2\lambda_n\gamma(\varepsilon)\big)\right), \tag{2.1}$$

in the sense that, as $\varepsilon \to 0$, the ratio of the expressions on the left and right sides of (2.1) approaches 1. The function $\gamma = \gamma(\varepsilon)$ is defined implicitly by the relation

$$\varepsilon^2 = \sum_{n\ge 1} \frac{\lambda_n}{1+2\lambda_n\gamma(\varepsilon)},$$

and this implicit dependence on ε is the main drawback of (2.1) in concrete applications.

Sometimes (2.1) can lead to an explicit asymptotic of the probability on the left-hand side (see [10, Section 6.1] and references therein), but when $\boldsymbol{y}$ is a multi-dimensional Gaussian diffusion, a completely different approach, based on asymptotic analysis of the function $\Psi(Q;t)$ from (1.2), appears to be a better option. In particular, this approach makes it possible to handle both questions [Q1] and [Q2].

In fact, the large-time asymptotic of $\Psi(Q;t)$ provides an immediate answer to [Q1].

Proposition 2.1 *If* $\lim_{t\to+\infty}\Psi(Q;t)=0$, *then* $\lim_{t\to+\infty}\int_0^t \boldsymbol{y}^\top(s)Q\boldsymbol{y}(s)ds=+\infty$ *with probability one.*

Proof. If $\lim_{t\to+\infty}\Psi(Q;t)=0$ but

$$\mathbb{P}\left(\lim_{t\to+\infty}\int_0^t \boldsymbol{y}^\top(s)Q\boldsymbol{y}(s)ds\leq c<\infty\right)=\delta>0,$$

then (1.2) implies $\Psi(Q;t)\geq \delta\, e^{-c}>0$ for all $t>0$, a contradiction.

□

The high frequency asymptotic of $\Psi(Q;t)$ provides an answer to [Q2] via an exponential Tauberian theorem (Theorem 2.2 below). This theorem is a modification of [10, Theorem 3.5] (which, in turn, is a modification of [1, Theorem 4.12.9]).

Theorem 2.2 *Let ξ be a non-negative random variable. Then*

$$\lim_{\lambda\to+\infty}\lambda^{-\gamma}\ln\left(\mathbb{E}e^{-\lambda\xi}\right)=-\alpha \quad \text{for some} \quad \alpha>0,\ 0<\gamma<1, \tag{2.2}$$

holds if and only if

$$\lim_{\varepsilon\to0}\varepsilon^{\gamma/(1-\gamma)}\ln\mathbb{P}(\xi\leq\varepsilon)=-\big((1-\gamma)\alpha\big)^{1/(1-\gamma)}\left(\frac{\gamma}{1-\gamma}\right)^{\gamma/(1-\gamma)}. \tag{2.3}$$

Technically, (2.3) is only logarithmic asymptotic of the probability and is not as strong as (2.1), but, for many applications, the logarithmic asymptotic is good enough, and, by providing an explicit dependence on ε, it is also much more useful.

We write (2.3) as

$$\ln\mathbb{P}(\xi\leq\varepsilon)\sim-\left((1-\gamma)\alpha\right)^{1/(1-\gamma)}\left(\frac{\gamma}{1-\gamma}\right)^{\gamma/(1-\gamma)}\varepsilon^{-\gamma/(1-\gamma)},\ \varepsilon\to0, \tag{2.4}$$

and say that the random variable ξ has the `small ball rate`

$$\varpi=\frac{\gamma}{1-\gamma}$$

and the `small ball constant`

$$\mathfrak{C}=\left((1-\gamma)\alpha\right)^{1/(1-\gamma)}\left(\frac{\gamma}{1-\gamma}\right)^{\gamma/(1-\gamma)}.$$

The two extreme cases of (2.3), corresponding to $\gamma=1$ (infinite small ball rate) and $\gamma=0$ (zero small ball rate), are a straightforward exercise in elementary probability.

Proposition 2.3 *Let ξ be a non-negative random variable. Then*

1. $\lim_{\lambda\to+\infty}\mathbb{E}e^{-\lambda\xi}=p_0>0$ *is equivalent to* $\mathbb{P}(\xi=0)=p_0$.
2. $\lim_{\lambda\to+\infty}\frac{\ln\mathbb{E}e^{-\lambda\xi}}{\lambda}=-\varepsilon_0<0$ *is equivalent to* $\varepsilon_0=\inf\big\{\varepsilon>0:\mathbb{P}(\xi\geq\varepsilon)=1\big\}>0$.

3 Multi-Dimensional Gaussian Diffusions

The main object of study in this paper is the $\mathbb{R}^{\mathrm{d}}$-valued process $\boldsymbol{y} = \boldsymbol{y}(t)$ defined by (1.1). Here is a summary of the basic properties of $\boldsymbol{y}$.

Proposition 3.1 *The solution of* (1.1) *is a Gaussian process*

$$\boldsymbol{y}(t) = e^{tA}\boldsymbol{y}(0) + \int_0^t e^{(t-s)A} B\, d\boldsymbol{w}(s) \tag{3.1}$$

with mean

$$\boldsymbol{\mu}(t) = e^{tA}\,\boldsymbol{m} \tag{3.2}$$

and covariance matrix

$$R(t) = e^{tA}Ke^{tA^\top} + \int_0^t e^{sA}BB^\top e^{sA^\top}\, ds. \tag{3.3}$$

The matrix $R(t)$ has the following properties:

1. *It is the solution of the initial value problem*

$$\dot{R}(t) = AR(t) + R(t)A^\top + BB^\top,\ t > 0,\quad R(0) = K; \tag{3.4}$$

2. *When $K = 0$, it is non-degenerate for every $t > 0$ if and only if the pair (A, B) is controllable: the $\mathrm{d} \times m\mathrm{d}$ matrix*

$$[B\ AB\ \cdots\ A^{\mathrm{d}-1}B]$$

has rank d*;*

3. *If $K = 0$, the pair (A, B) is controllable, and the eigenvalues of the matrix A have non-positive real parts, then, for every $t_\circ > 0$, there exists a positive numbers $c_\circ$ such that, for all $t > t_\circ > 0$,*

$$c_\circ^{\mathrm{d}} \le \det\big(R(t)\big) \le \|B^\top\|^{2\mathrm{d}}\, t^{\mathrm{d}}. \tag{3.5}$$

Moreover, if the pair (A, B) is controllable and all eigenvalues of A have strictly negative real parts, then equation (1.1) *is ergodic, and the stationary distribution is Gaussian with mean zero and non-singular covariance matrix R_∞ that is the unique solution of*

$$AR_\infty + R_\infty A^\top + BB^\top = 0. \tag{3.6}$$

Proof. Direct computations show that the solution of (1.1) is (3.1), from which (3.2) and (3.3) immediately follow. The solution of (3.4) is unique and is indeed (3.3). The equivalence between non-degeneracy of $R(t)$, $t > 0$, and controllability of (A, B) is well-known (e.g. [9, Corollary 4.3.2]).

To establish (3.5), note that, for every unit vector $\boldsymbol{u} \in \mathbb{R}^{\mathrm{d}}$,

$$\boldsymbol{u}^\top R(t)\boldsymbol{u} = \int_0^t \|B^\top e^{sA^\top}\boldsymbol{u}\|^2 ds.$$

In particular, the smallest eigenvalue of $R(t)$ is a non-decreasing function of t and, if all the eigenvalues of A have non-positive real parts, then the largest eigenvalue of $R(t)$ is bounded above by $\|B^\top\|^2 t$. Then (3.5) holds with

$$c_\circ = \min_{\|\boldsymbol{u}\|=1} \boldsymbol{u}^\top R(t_\circ)\boldsymbol{u}.$$

Finally, controllability of (A, B) and stability of A imply ergodicity of (1.1) [3, Theorem 9.1.1]. Existence, uniqueness, and non-degeneracy of the solution of (3.6) follow from [9, Theorem 5.3.1]. This solution of (3.6) is

$$R_\infty = \int_0^{+\infty} e^{sA} BB^\top e^{sA^\top} ds$$

(cf. [9, Formula (5.3.3)]). In particular,

$$R_\infty = \lim_{t\to+\infty} R(t)$$

is the covariance matrix of the stationary distribution for the solutions of (1.1).

This completes the proof of Proposition 3.1.

□

Given a symmetric non-negative definite matrix $Q \in \mathbb{R}^{\mathrm{d}\times\mathrm{d}}$, consider the function $\Psi(Q;t)$ from (1.2). The key to computing the function Ψ is the algebraic Riccati equation

$$C^\top BB^\top C - C^\top A - A^\top C = 2Q \tag{3.7}$$

for the unknown matrix $C \in \mathbb{R}^{\mathrm{d}\times\mathrm{d}}$.

Theorem 3.2 *Assume that equation* (3.7) *has a symmetric solution* $C = C^\top$ *and define the process* $\boldsymbol{y}_*$ *by*

$$d\boldsymbol{y}_* = (A - BB^\top C)\boldsymbol{y}_* dt + B d\boldsymbol{w}(t),\ \ \boldsymbol{y}_*(0) = \boldsymbol{y}(0).$$

Then

$$\Psi(Q;t) = e^{-(t/2)\mathrm{Tr}(BB^\top C)}\, \mathbb{E}\exp\left(\frac{1}{2}\Big(\boldsymbol{y}_*^\top(t)C\boldsymbol{y}_*(t) - \boldsymbol{y}_*^\top(0)C\boldsymbol{y}_*(0)\Big)\right). \tag{3.8}$$

In particular,

1. If $\boldsymbol{y}(0) = \boldsymbol{m}$ *is non-random, then*

$$\Psi(Q;t) = \frac{\exp\left(-\frac{1}{2}\left(t\,\mathrm{Tr}(BB^\top C) + \boldsymbol{m}^\top C\boldsymbol{m} - \boldsymbol{\mu}_*^\top(t)\big(I - CR_*(t)\big)^{-1}C\boldsymbol{\mu}_*(t)\right)\right)}{\sqrt{\det\big(I - CR_*(t)\big)}}, \tag{3.9}$$

where

$$\boldsymbol{\mu}_*(t) = e^{t(A-BB^\top C)}\boldsymbol{m},$$
$$R_*(t) = \int_0^t e^{s(A-BB^\top C)} BB^\top e^{s(A-BB^\top C)^\top} ds. \tag{3.10}$$

2. If $\boldsymbol{y}(0)$ is a Gaussian random vector with mean $\boldsymbol{m}$ and covariance K, then

$$\Psi(Q;t)=\frac{\exp\left(-\frac{1}{2}\left(t\,\mathrm{Tr}(BB^\top C)-\widetilde{\boldsymbol{\mu}}_*^\top(t)\big(I-\widetilde{C}\widetilde{R}_*(t)\big)^{-1}\widetilde{C}\widetilde{\boldsymbol{\mu}}_*(t)\right)\right)}{\sqrt{\det\big(I-\widetilde{C}\widetilde{R}_*(t)\big)}}, \tag{3.11}$$

where $\widetilde{C}$ and $\widetilde{R}_*(t)$ are 2d*-by-*2d *block matrices*

$$\widetilde{C}=\begin{pmatrix} -C & 0\\ 0 & C\end{pmatrix},\ \widetilde{R}_*(t)=\begin{pmatrix} K & Ke^{t(A-BB^\top C)^\top}\\ e^{t(A-BB^\top C)}K & e^{t(A-BB^\top C)}Ke^{t(A^\top-CBB^\top)}+R_*(t)\end{pmatrix}, \tag{3.12}$$

and

$$\widetilde{\boldsymbol{\mu}}_*(t)=\begin{pmatrix}\boldsymbol{m}\\ \boldsymbol{\mu}_*(t)\end{pmatrix}\in\mathbb{R}^{2\mathrm{d}}. \tag{3.13}$$

Proof. Theorem 3.2 was essentially proved in [8], because the more recent results from [9] about solvability of (3.7) made it possible to remove the additional restriction (stability of A) used in [8]. There are two main steps in the proof: (a) getting (3.8) via a change of measure, which an interested reader can easily do using a Girsanov-type formula, such as [11, Formula (7.138)]; (b) evaluating the right-hand side of (3.8) using the equality

$$\mathbb{E}e^{\boldsymbol{\xi}^\top G\boldsymbol{\xi}}=\big(\det(I-2GR)\big)^{-1/2}\exp\left(\boldsymbol{\mu}^\top(I-2GR)^{-1}G\boldsymbol{\mu}\right), \tag{3.14}$$

where $\boldsymbol{\xi}$ is a Gaussian random vector with mean $\boldsymbol{\mu}$ and covariance matrix R, and G is a symmetric matrix.

□

Corollary 3.3 *Non-zero initial conditions do not increase the value of Ψ:*

$$\Psi(t;Q;\boldsymbol{m},K)\le\Psi(t,Q;\boldsymbol{0},0). \tag{3.15}$$

Proof. We fix t,Q, C. Because

$$\boldsymbol{y}_*(t)=e^{t(A-BB^\top C)}\boldsymbol{y}(0)+\int_0^t e^{s(A-BB^\top C)}Bd\boldsymbol{w}(s),$$

the Gaussian random vectors

$$\boldsymbol{\zeta}=\boldsymbol{y}_*(t)-e^{t(A-BB^\top C)}\boldsymbol{y}(0)$$

and

$$\boldsymbol{\eta}=e^{t(A-BB^\top C)}\boldsymbol{y}(0)$$

are independent and $\mathbb{E}\boldsymbol{\zeta}=\boldsymbol{0}$. By (3.8),

$$\Psi(Q;t;\boldsymbol{m},K)=e^{-(t/2)\mathrm{Tr}(BB^\top C)}\,\mathbb{E}\exp\left(\frac{1}{2}\left((\boldsymbol{\zeta}+\boldsymbol{\eta})^\top C(\boldsymbol{\zeta}+\boldsymbol{\eta})-\boldsymbol{\eta}^\top C_t\boldsymbol{\eta}\right)\right),$$

where

$$C_t=e^{-t(A-BB^\top C)^\top}Ce^{-t(A-BB^\top C)}.$$

We now compute the expectation by conditioning on $\boldsymbol{\eta}$ and using (3.14) twice. The result is

$$\Psi(Q;t;\boldsymbol{m},K)=\Psi(Q;t;\mathbf{0},0)\Phi(t,Q,K)e^{\boldsymbol{m}^\top N\boldsymbol{m}}$$

with some function Φ and a matrix N; the matrix N depends on K and t, but the function Φ does not depend on $\boldsymbol{m}$. While this expression is not as explicit as (3.11), it does establish (3.15). Indeed, by definition,

$$\Psi(t;Q;\boldsymbol{m},K)\le 1, \tag{3.16}$$

so if the matrix N is not non-positive definite, then it would be possible to violate (3.16) with a suitable scaling of $\boldsymbol{m}$. By considering $\boldsymbol{m}=\mathbf{0}$, a similar argument shows that we also must have $\Phi(t,Q,K)\le 1$.

□

Let us summarize the basic facts about the symmetric algebraic Riccati equation

$$XDX-XA-A^\top X=Q, \tag{3.17}$$

with known matrices A,D,Q in $\mathbb{R}^{\mathrm{d}\times\mathrm{d}}$ such that $D\ge 0$ and $Q\ge 0$.

A standard assumption about (3.17) is that the pair (A,D) is either controllable or stabilizable; it is also often assumed that the pair (Q,A) is either observable or detectable. Below are *practical definitions* of these four concepts:

- (A,D) is `controllable` if the rank of the $\mathrm{d}\times 2\mathrm{d}$ matrix $[zI-A\ D]$ is equal to d for all complex numbers z [9, Theorem 4.3.3];
- (A,D) is `stabilizable` if the rank of the $\mathrm{d}\times 2\mathrm{d}$ matrix $[zI-A\ D]$ is equal to d for all complex numbers z with non-negative real part [9, Theorem 4.5.6(a)];
- (Q,A) is `observable` if $(A^\top,Q^\top)$ is controllable [9, Proposition 4.2.2];
- (Q,A) is `detectable` if $(A^\top,Q^\top)$ is stabilizable [the original definition];

In particular,

1. If D is invertible, then the pair (A,D) is controllable for every matrix A;
2. If A is stable (all eigenvalues of A have negative real parts), then the pair (A,D) is stabilizable for every matrix D `and` the pair (A,Q) is detectable for every matrix Q;
3. The pair (A,B) is controllable if and only if the pair $(A,BB^\top)$ is controllable [9, Corollary 4.1.3]; recall that the definition of controllability of (A,B) if $B\in\mathbb{R}^{\mathrm{d}\times m}$, $m\le \mathrm{d}$, is in Proposition 3.1.
4. For every matrix $C\in\mathbb{R}^{\mathrm{d}\times\mathrm{d}}$, the pair $(A+BB^\top C,BB^\top)$ is controllable if and only if the pair $(A,BB^\top)$ is controllable [9, Lemma 4.4.1].

The following is the summary of the main results from [9] about equation (3.17).

Proposition 3.4 *Consider equation* (3.17) *with* $D\ge 0$ *and* $Q\ge 0$.

1. *If the pair* (Q,A) *is observable, then every symmetric solution of* (3.17) *is non-singular [9, Problem 7.11.15].*

2. *If the pair (A, D) is stabilizable, then there exists a symmetric solution X_+ of* (3.17)*, such that $X_+ \geq 0$, the eigenvalues of $A - DX_+$ have non-positive real parts, and $X_+ - X \geq 0$ for every symmetric solution X of* (3.17)*; X_+ is called the* `maximal symmetric solution` *of* (3.17) *[9, Theorems 9.1.1 and 9.1.2].*
3. *If (A, D) is stabilizable and (Q, A) is detectable, then the eigenvalues of $A - DX_+$ have strictly negative real parts [9, Theorem 9.1.2].*
4. *If (A, D) is stabilizable, X_+ is the* `maximal symmetric solution` *of* (3.17)*, and $X_1 = X_1^\top$ satisfies*
$$X_1 D X_1 - X_1 A - A^\top X_1 = Q_1,$$
where $Q_1 \geq 0$ and $Q - Q_1 \geq 0$, then $X_+ - X_1 \geq 0$ [9, Corollary 9.1.6].

Equation (3.7) can have more than one symmetric solution, and Theorem 3.2 indicates that any such solution can be used to compute $\Psi(Q;t)$. If the pair $(A, BB^\top)$ is controllable, then it is often convenient to take $C = C_+$, so that (3.8) becomes

$$\Psi(Q;t) = e^{-(t/2)\mathrm{Tr}(CBB^\top)}\,\mathbb{E}\exp\left(\frac{1}{2}\Big(\boldsymbol{y}_+^\top(t)C_+\boldsymbol{y}_+(t) - \boldsymbol{y}^\top(0)C_+\boldsymbol{y}(0)\Big)\right), \tag{3.18}$$

where

$$d\boldsymbol{y}_+(t) = (A - BB^\top C_+)\boldsymbol{y}_+(t)dt + Bd\boldsymbol{w}(t),\ t > 0,\ \boldsymbol{y}_+(0) = \boldsymbol{y}(0). \tag{3.19}$$

Indeed, Propositions 3.1 and 3.4 suggest, and the following result confirms, that representation (3.18) can have advantages over the more general (3.8).

Proposition 3.5 *Assume that the pair $(A, BB^\top)$ is controllable, let C_+ be the maximal symmetric solution of* (3.7)*, and consider the process $\boldsymbol{y}_+ = \boldsymbol{y}_+(t)$ from* (3.19)*. Define*

$$\boldsymbol{\mu}_+(t) = \mathbb{E}\boldsymbol{y}_+(t),\ R_+(t) = \mathbb{E}\Big(\big(\boldsymbol{y}_+(t) - \boldsymbol{\mu}_+(t)\big)\big(\boldsymbol{y}_+(t) - \boldsymbol{\mu}_+(t)\big)^\top\Big).$$

Then

$$\limsup_{t\to+\infty} \|\boldsymbol{\mu}_+(t)\| < \infty.$$

If, in addition, the pair (A, Q) is detectable, then the process $\boldsymbol{y}_+$ is ergodic; the stationary distribution of $\boldsymbol{y}_+$ is Gaussian with mean zero and covariance matrix

$$R_+ = \lim_{t\to+\infty} R_+(t) = \int_0^\infty e^{t(A-BB^\top C_+)}BB^\top e^{t(A-BB^\top C_+)^\top}dt. \tag{3.20}$$

Proof. Everything follows from Propositions 3.1 and 3.4. In particular, to claim ergodicity, note that detectability of (A, Q) implies stability of $A - BB^\top C_+$, and, by [9, Lemma 4.4.1], controllability of $(A, BB^\top)$ implies controllability of $(A - BB^\top C_+, B)$. □

We will also need some basic facts about the Riccati differential equation

$$\dot{X}(t) + X(t)QX(t) = AX(t) + X(t)A^\top + D,\ t > 0,\quad X(0) = 0, \tag{3.21}$$

with constant square matrices A, D, Q.

Proposition 3.6 *Consider equation* (3.21) *under the assumptions that* $Q \geq 0$, $D \geq 0$, *and the pair* (A, D) *is controllable.*

1. *There exists a unique symmetric solution* $X = X(t)$ *of* (3.21), *and* $X(t) > 0$ *for all* $t > 0$ *[12, Lemma 16.3].*
2. *If* $X_1 = X_1(t)$ *is a symmetric solution of*

$$\dot{X}_1(t) + X_1(t)Q_1X_1(t) = AX_1(t) + X_1(t)A^\top + D, \; X_1(0) = 0,$$

with $Q_1 - Q \geq 0$, *then, for all* $t > 0$, $X(t) - X_1(t) \geq 0$ *and, as a consequence,* $\det\big(X(t)\big) \geq \det\big(X_1(t)\big)$ *[14, Theorem 1]* and *[13, Exercise 13.5.1].*

Formula (3.11) provides an expression for the function Ψ *in quadratures*: there are no *differential* Riccati equations to solve, as, for example in [17, Corollary 1] or [6, Section 4.1]. Still, further simplification of (3.11) or even of (3.9) is, in general, not possible, often because of complications related to evaluation of (3.10) when the matrices A, B, and C do not commute.

Below are two examples when the right-hand side of (3.9) can be simplified further. Both examples can be considered multi-dimensional analogues of the one-dimensional Ornstein-Uhlenbeck process

$$dx(t) = -ax(t)dt + \sigma dw(t), \; t > 0, \quad x(0) = 0, \tag{3.22}$$

for which it is known [12, Lemma 17.3] that

$$\mathbb{E}e^{-\lambda \int_0^t x^2(s)ds} = \frac{1}{\sqrt{e^{-at}\left(\frac{a\sinh\left(t\sqrt{a^2+2\sigma^2\lambda}\right)}{\sqrt{a^2+2\sigma^2\lambda}} + \cosh\left(t\sqrt{a^2+2\sigma^2\lambda}\right)\right)}}. \tag{3.23}$$

Proposition 3.7 *Assume that* $\boldsymbol{y}(0) = \boldsymbol{0}$, $A = A^\top$, $B = \sigma I$, $\sigma \neq 0$, *and the pair* (Q, A) *is observable. Define the matrix*

$$\Lambda = (A^2 + 2\sigma^2 Q)^{1/2}$$

as the symmetric non-negative-definite square root of $A^2 + 2\sigma^2 Q$. *Then*

1. *The matrix* Λ *is invertible;*
2. *The function* $\Psi(Q;t)$ *has the following representation:*

$$\Psi(Q;t) = \frac{e^{-\mathrm{Tr}(A)\,t/2}}{\sqrt{\det\big(\cosh(t\Lambda) - A\Lambda^{-1}\sinh(t\Lambda)\big)}}. \tag{3.24}$$

Proof. With $A = A^\top$, equation (3.7) becomes $C^2 - AC - CA = 2\sigma^2 Q$ or

$$(C - A)^2 = \Lambda^2. \tag{3.25}$$

Then

$$C_+ = A + \Lambda$$

is the maximal symmetric solution of (3.25), and, because $A - C_+ = -\Lambda$, the matrices Λ and $I - A\Lambda^{-1}$ are non-degenerate by Proposition 3.4.

Next,

$$\begin{aligned}I-CR(t)&=I-\frac{1}{2}(A+\Lambda)\Lambda^{-1}\big(I-e^{-2t\Lambda}\big)\\&=\frac{1}{2}\left(I+e^{-2t\Lambda}\right)-\frac{1}{2}A\Lambda^{-1}\big(I-e^{-2t\Lambda}\big)=\big(\cosh(t\Lambda)-A\Lambda^{-1}\sinh(t\Lambda)\big)e^{-t\Lambda},\end{aligned}$$

and (3.24) follows because $\det\left(e^{-t\Lambda}\right)=e^{-t\,\mathrm{Tr}(\Lambda)}$.

$\square$

The second example shows that, even when system (1.1) is not diagonal, the random variable $\int_0^t \|\boldsymbol{y}(s)\|^2 ds$ can have the same distribution as

$$\sum_{k=1}^{\mathrm{d}}\int_0^t x_k^2(s)ds,$$

where $x_1,\dots,x_{\mathrm{d}}$ are iid one-dimensional Ornstein-Uhlenbeck processes of the type (3.22); cf. [12, Lemma 17.5] when $\mathrm{d}=2$.

Proposition 3.8 *Assume that $\boldsymbol{y}(0)=\mathbf{0}$, $A=-aI+A_1$, $B=\sigma I$, and $Q=\lambda I$, where $a\in\mathbb{R}$, $\lambda>0$, $\sigma\neq 0$, and A_1 is a skew-symmetric matrix: $A_1=-A_1^\top$. Then*

$$\Psi(Q;t)=\left(\Psi_{OU}(\lambda;t)\right)^{\mathrm{d}}, \tag{3.26}$$

where

$$\Psi_{OU}(\lambda;t)=\left(\frac{e^{at}}{\cosh(\varrho t)+(a/\varrho)\sinh(\varrho t)}\right)^{1/2}$$

is the right-hand side of (3.23), with

$$\varrho=(a^2+2\sigma^2\lambda)^{1/2}.$$

Proof. Equation (3.7) becomes

$$C^2+(aI+A_1)C+C(aI-A_1)=2\sigma^2\lambda I. \tag{3.27}$$

Substituting $C=\gamma I$, $\gamma\in\mathbb{R}$, in (3.27) and using $CA_1=A_1C$ yields

$$\gamma^2+2a\gamma=2\sigma^2\lambda \text{ or } \gamma=-a\pm\varrho.$$

With $C=(-a+\varrho)I$, equality (3.10) becomes

$$R_*(t)=\int_0^t e^{-2\varrho s\,I}\,ds=\frac{1-e^{-2\varrho t}}{2\varrho}\,I,$$

and (3.26) follows from (3.11).

$\square$

In the case of non-zero initial condition, an explicit and manageable expression for Ψ exists when $\mathrm{d}=1$: for the one-dimensional OU process (3.22) with initial condition having mean x_0 and variance σ_0^2, direct computations lead to

$$\mathbb{E}\exp\left(-\lambda\int_0^t x^2(s)ds\right)=\frac{\Psi_{OU}(\lambda;t)}{\sqrt{1+2\sigma_0^2\psi}}\exp\left(-\frac{\psi x_0^2}{1+2\sigma_0^2\psi}\right), \tag{3.28}$$

where

$$\psi=\frac{\varrho-a}{2\sigma^2}\left(1-\frac{e^{-\varrho t}}{\cosh(\varrho t)+(a/\varrho)\sinh(\varrho t)}\right).$$

4 Asymptotic Analysis of the Laplace Transform

As a first application of Theorem 3.2, let us investigate the asymptotic of the function $\Psi(Q;t)$ as $t \to +\infty$.

Theorem 4.1 *Assume that the pair* $(A, BB^\top)$ *is controllable and* $\mathrm{Tr}(Q) > 0$. *Let* C_0 *be the maximal symmetric solution of*

$$CBB^\top C - CA - A^\top C = 0.$$

Denote by C_+ *the maximal symmetric solution of* (3.7). *Then* $\mathrm{Tr}(C_+ - C_0) > 0$ *and*

$$\limsup_{t\to+\infty} \frac{\ln \Psi(Q;t)}{t} \le -\frac{\mathrm{Tr}(C_+ - C_0)}{2}. \tag{4.1}$$

Proof. By (3.15), it is enough to consider only zero initial condition $\boldsymbol{y}(0) = \boldsymbol{0}$. We continue to use the notation

$$D = BB^\top.$$

Then (3.18) becomes

$$\Psi(Q;t) = \frac{\exp\big(-\frac{1}{2}\,(t\,\mathrm{Tr}(DC_+))\big)}{\sqrt{\det\big(I - C_+R_+(t)\big)}}. \tag{4.2}$$

By Proposition 3.4, $C_+ - C_0 \ge 0$, and therefore $\mathrm{Tr}(DC_+ - DC_0) \ge 0$. If $\mathrm{Tr}(DC_+ - DC_0) = 0$, then $DC_+ = DC_0$, so that $\boldsymbol{y}_+ = \boldsymbol{y}_\circ$, where $\boldsymbol{y}_\circ$ is the solution of (3.19) with C_0 instead of C_+. Equality (3.18) then implies $\Psi(Q;t) = 1$, contradicting the assumption that $\mathrm{Tr}(Q) > 0$. In other words, $\mathrm{Tr}(Q) > 0$ implies

$$\delta = \mathrm{Tr}(DC_+ - DC_0) > 0. \tag{4.3}$$

By Proposition 3.1 and [9, Lemma 4.4.1], $R_+(t) > 0$ for every $t > 0$ and

$$\dot{R}_+(t) = (A - DC_+)R_+(t) + R_+(t)(A^\top - C_+D) + D.$$

Similarly, the matrix

$$R_0(t) = \int_0^t e^{s(A-DC_0)}De^{s(A^\top - C_0D)}ds \tag{4.4}$$

is positive-definite for every $t > 0$.

Define the matrices

$$S(t) = \big(R_+(t)\big)^{-1},\ V(t) = S(t) - C_+,\ \text{and}\ \ U(t) = \begin{cases} \big(V(t)\big)^{-1}, & \text{if } t > 0; \\ 0, & \text{if } t = 0. \end{cases} \tag{4.5}$$

Then (4.2) becomes

$$\Psi^2(Q;t) = \det\big(e^{-tC_+}U(t)\big) \times \det\big(S(t)\big), \tag{4.6}$$

or, using (4.3) and $\det\big(e^X\big) = e^{\mathrm{Tr}(X)}$,

$$\Psi^2(Q;t) = e^{-\delta\, t}\det\big(S(t)\big) \times \det\big(e^{-tC_0}U(t)\big).$$

By (3.5) and Proposition 3.4(2),

$$\lim_{t\to\infty} \frac{\ln\Big(\det\big(S(t)\big)\Big)}{t} = 0,$$

and, to establish (4.1), it remains to verify that

$$\limsup_{t\to\infty} \frac{\ln\Big(\det\big(e^{-tC_0}U(t)\big)\Big)}{t} \le 0. \tag{4.7}$$

Let us derive the differential equation satisfied by $U(t)$. For every invertible matrix $X = X(t)$, the inverse matrix $Y(t) = \big(X(t)\big)^{-1}$ satisfies

$$\dot{Y}(t) = -Y(t)\dot{X}(t)Y(t), \tag{4.8}$$

which follows after applying the product rule to $X(t)Y(t) = I$. Applying (4.8) with $X(t) = R_+(t)$ we get

$$\dot{S}(t) = -S(A - DC_+) - (A^\top - C_+D)S - SDS. \tag{4.9}$$

After combining (4.5) and (4.9),

$$\dot{V}(t) = \dot{S}(t) = -(V + C_+)(A - DC_+) - (A^\top - C_+D)(V + C_+) - (V + C_+)D(V + C_+);$$

simplifying the result using (3.7),

$$\dot{V}(t) = -VA - A^\top V - VDV + 2Q. \tag{4.10}$$

One more application of (4.8), this time with $X(t) = V(t)$, together with (4.10), gives the differential equation satisfied by U:

$$\dot{U}(t) + 2U(t)QU(t) = AU(t) + U(t)A^\top + D. \tag{4.11}$$

By Proposition 3.6, $U(t) > 0$ for all $t > 0$ and

$$\det\big(U(t)\big) \le \det\big(U_0(t)\big),$$

where $U_0(t)$ satisfies

$$\dot{U}_0(t) = AU_0(t) + U_0(t)A^\top + D.$$

Next, let us apply (4.6) when $Q = 0$:

$$\Psi^2(0;t) = \det\big(S_0(t)\big) \times \det\big(e^{-tC_0}U_0(t)\big),$$

where $S_0(t) = R_0^{-1}(t)$ and $R_0(t)$ is from (4.4). Since $\Psi(0;t) = 1$ for all $t \ge 0$,

$$\det\big(R_0(t)\big) = \det\big(e^{-tC_0}U_0(t)\big) \ge \det\big(e^{-tC_0}U(t)\big).$$

By (3.5) and Proposition 3.4(2),

$$\lim_{t\to\infty} \frac{\ln\Big(\det\big(R_0(t)\big)\Big)}{t} = 0,$$

which implies (4.7).

This completes the proof of Theorem 4.3.

□

Combining (4.1) with Proposition 2.1, we get an answer to question [Q1] originally posed in the introduction.

Proposition 4.2 *If $Q \geq 0$ with $\mathrm{Tr}(Q) > 0$, $\boldsymbol{y} = \boldsymbol{y}(t)$ is the solution of* (1.1)*, and the pair (A, B) is controllable, then*

$$\int_0^{+\infty} \boldsymbol{y}^\top(t) Q \boldsymbol{y}(t) dt = +\infty \tag{4.12}$$

with probability one.

Proof. Indeed, (4.1) implies that

$$\lim_{t \to +\infty} \Psi(Q; t) = 0$$

as long as Q is not identically zero.

□

While intuitively obvious, (4.12) is surprisingly difficult to prove: even the one-dimensional case relies on the Laplace transform [12, Section 17.3]. Note also that (4.12) is, in general, not true without the controllability assumption: consider

$$A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \ B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \ Q = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

so that, with $\boldsymbol{y}(0) = (y_1(0) \ \ y_2(0))^\top$,

$$\boldsymbol{y}^\top(t) Q \boldsymbol{y}(t) = y_1^2(0) e^{-2t}$$

and the left-hand side of (4.12) is bounded above by $y_1^2(0)$. Further analysis of this example shows that, without the controllability assumption, the integral in (4.12) can either converge or diverge, depending on the initial condition $\boldsymbol{y}(0)$ and the matrix Q.

A more precise asymptotic of $\Psi(Q; t)$ as $t \to \infty$ exists under additional assumptions, and, even though it does not add anything as far as answering question [Q1], the result can, for example, provide an explicit solution to some optimization problems (cf. [4, Equation (1.10)]).

Theorem 4.3 *Assume that the pair (A, B) is controllable, the pair (Q, A) is observable, and the initial condition $\boldsymbol{y}(0)$ is a Gaussian vector with mean $\boldsymbol{m}$ and covariance matrix K. Let C_+ be the maximal symmetric solution of* (3.7)*. Then*

1. *The matrix R_+ from* (3.20) *is well-defined and non-singular;*
2. *The matrix $I - C_+ R_+$ is non-singular and, as $t \to +\infty$,*

$$\Psi(Q; t) \sim \frac{e^{-(t/2)\mathrm{Tr}(BB^\top C_+)}}{\sqrt{\det(I - C_+ R_+)}} \times \frac{e^{-(1/2)\boldsymbol{m}^\top (I + C_+ K)^{-1} C_+ \boldsymbol{m}}}{\sqrt{\det(I + C_+ K)}}. \tag{4.13}$$

Proof. Define the Gaussian process $\boldsymbol{y}_+ = \boldsymbol{y}_+(t)$ by (3.19). Proposition 3.5 implies that the process $\boldsymbol{y}_+$ has a unique invariant measure, which is Gaussian with mean zero and non-singular covariance matrix R_+. Passing to the limit as $t \to +\infty$ in (3.12) and (3.13), we find

$$\lim_{t \to +\infty} \big(I - \widetilde{C} \widetilde{R}_*(t)\big) = \begin{pmatrix} I + CK & 0 \\ 0 & I - CR_+ \end{pmatrix}, \ \lim_{t \to +\infty} \widetilde{\boldsymbol{\mu}}_*(t) = \begin{pmatrix} \boldsymbol{m} \\ 0 \end{pmatrix}.$$

It remains to verify that the matrix $I - C_+R_+$, or, equivalently, $\big(R_+\big)^{-1} - C_+$, is non-singular; then relation (4.13) will follow from (3.11).

To show that the matrix $\big(R_+\big)^{-1} - C_+$ is non-singular, note that, by [12, Theorem 16.2], the matrix $V = V(t)$ solving equation (4.10) has a non-singular limit as $t \to \infty$ that does not depend on the initial condition; by construction, this limit coincides with $\big(R_+\big)^{-1} - C_+$.

□

There is an alternative form of (4.13) using the *minimal symmetric solution* C_- of (3.7). Indeed, consider the equation

$$XDX + XA + A^\top X = 2Q. \tag{4.14}$$

Then C is solution of (3.7) if and only if $X = -C$ is a solution of (4.14). In particular, $X_- = -C_+$ is the minimal symmetric solution of (4.14). Applying [9, Theorem 7.5.1] we conclude that $X_+ = X_- + \big(R_+\big)^{-1} = \big(R_+\big)^{-1} - C_+$ is the *maximal* symmetric solution of (4.14). Note that direct computations using (4.9), with $\lim_{t\to+\infty} S(t) = R_+^{-1}$, confirm that $\big(R_+\big)^{-1} - C_+$ is a symmetric solution of (4.14), but an additional argument is still necessary to claim that it is indeed the maximal solution. By construction, $X_+ - X_- = C_+ - C_-$, which leads to an equivalent form of (4.13):

$$\Psi(Q;t) \sim e^{-(t/2)\mathrm{Tr}(BB^\top C_+)} \left(\frac{\det(C_+ - C_-)}{(-1)^{\mathrm{d}} \det(C_-)}\right)^{1/2} \times \frac{e^{-(1/2)\boldsymbol{m}^\top (I+C_+K)^{-1}C_+\boldsymbol{m}}}{\sqrt{\det(I + C_+K)}}. \tag{4.15}$$

On the one hand, the assumption about observability of (Q, A) cannot, in general, be omitted: take $\boldsymbol{y}(0) = \boldsymbol{0}$, $A = 0$, $B = I$, and

$$Q = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{pmatrix}.$$

In this case, the right-hand sides of both (4.15) and (4.13) are not defined because the matrices $C_\pm = \pm 2Q$ are singular and the matrix R_+ does not exist.

On the other hand, (4.15) can hold without the observability assumption: take $\boldsymbol{y}(0) = \boldsymbol{0}$,

$$A = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix},\ B = I,\ Q = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{pmatrix}.$$

so that $C_\pm = A \pm I$. In this case, $\det C_+ = 0$ but $\det C_- = 2$, $\det(C_+ - C_-) = 4$, $-\mathrm{Tr}(C_+)/2 = -1/2$, and the right-hand side of (4.15) gives the correct asymptotic $\Psi(Q;t) \sim \sqrt{2}e^{-t/2}$. Incidentally, note that $A - C_+ = -I$ is stable. Further analysis of this example shows that, if the matrices Q and A are both diagonal, then the observability condition can be replaced by a weaker condition of detectability, which, in this case, means that, for every zero entry on the diagonal of Q, the corresponding diagonal entry of A must be negative.

Next, we study the high frequency asymptotic of the function $\Psi(Q;t)$, that is,

$$\lim_{\lambda\to+\infty} \Psi(\lambda Q;t)$$

for fixed $t > 0$ and fixed $Q \geq 0$. As paper [2] demonstrates, if the matrix $BB^\top$ is not invertible, the high-frequency asymptotic can depend on the matrix Q in a rather complicated way even with zero initial conditions. In the next section, we will see that, when the noise is degenerate, the non-zero initial conditions can also change the asymptotic in a profound way. A truly universal result exists only when the matrix $BB^\top$ is invertible, which we will assume through the rest of this section.

If the matrix $D = BB^\top$ is invertible, then a linear transformation reduces the problem to the case $B = I$. Indeed, define

$$\bar{A} = D^{-1/2} A D^{1/2},\ \bar{Q} = D^{1/2} Q D^{1/2}.$$

The process $\boldsymbol{y}$ has the same distribution as the solution of

$$d\tilde{\boldsymbol{y}}(t) = A\tilde{\boldsymbol{y}}(t)dt + D^{1/2} d\boldsymbol{v}(t),\ t > 0,\ \tilde{\boldsymbol{y}}(0) = \boldsymbol{y}(0),$$

where $\boldsymbol{v}$ is a d-dimensional standard Brownian motion, and then

$$\Psi(Q; t) = \mathbb{E} \exp\left(- \int_0^t \bar{\boldsymbol{y}}^\top(s) \bar{Q} \bar{\boldsymbol{y}}(s) ds \right),$$

where $\bar{\boldsymbol{y}} = D^{-1/2} \tilde{\boldsymbol{y}}$ is the solution of

$$d\bar{\boldsymbol{y}}(t) = \bar{A}\bar{\boldsymbol{y}}(t)dt + d\boldsymbol{v}(t),\ t > 0,\ \bar{\boldsymbol{y}}(0) = D^{-1/2}\boldsymbol{y}(0).$$

Denote by C_λ the maximal symmetric solution of

$$C^2 - \bar{A}^\top C - C\bar{A} = 2\lambda \bar{Q},$$

and define

$$\bar{\boldsymbol{m}} = (BB^\top)^{-1/2}\boldsymbol{m}, \quad \boldsymbol{\mu}_\lambda(t) = e^{t(\bar{A}-C_\lambda)}\bar{\boldsymbol{m}}, \qquad R_\lambda(t) = \int_0^t e^{s(\bar{A}-C_\lambda)} e^{s(\bar{A}^\top - C_\lambda)} ds.$$

By Theorem 3.2,

$$\Psi(\lambda Q; t) = \frac{\exp\left(-\frac{1}{2}\left(t \operatorname{Tr}(C_\lambda) - \widetilde{\boldsymbol{\mu}}_\lambda^\top(t) S_\lambda \widetilde{\boldsymbol{\mu}}_\lambda(t)\right)\right)}{\sqrt{\det\left(I - \widetilde{C}_\lambda \widetilde{R}_\lambda(t)\right)}}, \tag{4.16}$$

where

$$S_\lambda = \left(I - \widetilde{C}_\lambda \widetilde{R}_\lambda(t)\right)^{-1} \widetilde{C}_\lambda,$$

$\widetilde{C}_\lambda$ and $\widetilde{R}_\lambda(t)$ are 2d-by-2d block matrices

$$\widetilde{C}_\lambda = \begin{pmatrix} -C_\lambda & 0 \\ 0 & C_\lambda \end{pmatrix},\ \widetilde{R}_\lambda(t) = \begin{pmatrix} K & K e^{t(\bar{A}^\top - C_\lambda)} \\ e^{t(\bar{A}-C_\lambda)} K & e^{t(\bar{A}-C_\lambda)} K e^{t(\bar{A}^\top - C_\lambda)} + R_\lambda(t) \end{pmatrix},$$

and $\widetilde{\boldsymbol{\mu}}_\lambda(t)$ is a vector in $\mathbb{R}^{2\mathrm{d}}$:

$$\widetilde{\boldsymbol{\mu}}_\lambda(t) = \begin{pmatrix} \boldsymbol{m} \\ \boldsymbol{\mu}_\lambda(t) \end{pmatrix}.$$

We will also need the matrix

$$Q_K = \lim_{\lambda \to +\infty} (I + \sqrt{2\lambda}\, \bar{Q}^{1/2} K)^{-1} \bar{Q}^{1/2}. \tag{4.17}$$

The limit in (4.17) exists by a monotonicity argument, and there are two obvious particular cases:

1. If $K=0$, then $Q_K = \bar{Q}^{1/2}$;
2. If $\bar{Q}^{1/2}K$ is invertible, then $Q_K = 0$.

Theorem 4.4 *Assume that the matrix $BB^\top$ is invertible. Denote by $\bar{Q}^{1/2}$ the symmetric non-negative square root of $\bar{Q}$.*

Then

$$\lim_{\lambda\to+\infty} \frac{\ln \Psi(\lambda Q; t)}{\lambda^{1/2}} = -2^{-1/2}\Big(t\,\mathrm{Tr}\big(\bar{Q}^{1/2}\big) + \bar{\boldsymbol{m}}^\top Q_K \bar{\boldsymbol{m}}\Big). \tag{4.18}$$

Proof. In view of (4.16), we need to verify the following:

$$\lim_{\lambda\to+\infty} \frac{(t/2)\mathrm{Tr}(C_\lambda)}{\sqrt{\lambda}} = 2^{-1/2} t\,\mathrm{Tr}\big(\bar{Q}^{1/2}\big), \tag{4.19}$$

$$\liminf_{\lambda\to\infty} \frac{\ln\Big(\det\big(I - \widetilde{C}_\lambda \widetilde{R}_\lambda(t)\big)\Big)}{\sqrt{\lambda}} = 0, \tag{4.20}$$

$$\lim_{\lambda\to+\infty} \frac{\widetilde{\boldsymbol{\mu}}_\lambda^\top(t) S_\lambda \widetilde{\boldsymbol{\mu}}_\lambda(t)}{2\sqrt{\lambda}} = 2^{-1/2}\bar{\boldsymbol{m}}^\top Q_K \bar{\boldsymbol{m}}. \tag{4.21}$$

If $\hat{C}_\lambda = C_\lambda/\sqrt{\lambda}$, then

$$\hat{C}_\lambda^2 - \frac{\bar{A}^\top}{\sqrt{\lambda}}\hat{C}_\lambda - \hat{C}_\lambda \frac{\bar{A}^\top}{\sqrt{\lambda}} = 2\bar{Q}.$$

By [9, Theorem 11.2.1],

$$\lim_{\lambda\to+\infty} \hat{C}_\lambda = \big(2\bar{Q}\big)^{1/2}, \tag{4.22}$$

which implies (4.19). Then

$$\lim_{\lambda\to+\infty} \lambda^{-1/2}(I + C_\lambda K)^{-1} C_\lambda = \sqrt{2}\, Q_K \tag{4.23}$$

and

$$\lim_{\lambda\to+\infty} \lambda^n \|C_\lambda e^{t(\bar{A}-C_\lambda)}\| = 0,\ n>0. \tag{4.24}$$

To establish (4.20) note that (4.22) and (4.24) imply

$$\liminf_{\lambda\to\infty} \frac{\ln\Big(\det\big(I - \widetilde{C}_\lambda \widetilde{R}_\lambda(t)\big)\Big)}{\sqrt{\lambda}} = \liminf_{\lambda\to\infty} \frac{\ln\Big(\det\big(I - C_\lambda R_\lambda(t)\big)\Big)}{\sqrt{\lambda}}.$$

Accordingly, define the matrix

$$U_\lambda(t) = \begin{cases} \Big(\big(R_\lambda(t)\big)^{-1} - C_\lambda\Big)^{-1}, & \text{if } t>0; \\ 0, & \text{if } t=0. \end{cases}$$

Then

$$1 \le \frac{1}{\det\Big(\big(I - C_\lambda R_\lambda(t)\big)\Big)} = \frac{\det\big(U_\lambda(t)\big)}{\det\big(R_\lambda(t)\big)}, \tag{4.25}$$

and equality (4.20) will follow from

$$\liminf_{\lambda\to\infty} \frac{\ln\Big(\det\big(R_\lambda(t)\big)\Big)}{\sqrt{\lambda}} = 0 \ \text{ and} \tag{4.26}$$

$$\limsup_{\lambda\to\infty} \frac{\ln\Big(\det\big(U_\lambda(t)\big)\Big)}{\sqrt{\lambda}} = 0. \tag{4.27}$$

To verify (4.26), denote by $\kappa_0 > 0$ the largest eigenvalue of the matrix $\big(2\bar{Q}\big)^{1/2}$. By (4.22) and continuous dependence of eigenvalues on the elements of the matrix (e.g. [15, Theorem 5.2]), the real part of every eigenvalue of $\bar{A}^\top - C_\lambda$ will be greater than or equal to $-2\sqrt{\lambda}\,\kappa_0$ for all sufficiently large λ. Proposition 3.1 then implies existence of a positive number δ such that, for all $\lambda > 0$,

$$\det\big(R_\lambda(t)\big) \geq \delta\,\lambda^{-\mathrm{d}/2},$$

from which (4.26) follows.

To verify (4.27), note that, by (4.11),

$$\dot{U}_\lambda(t) + 2\lambda U_\lambda(t)\bar{Q}U_\lambda(t) = \bar{A}U_\lambda(t) + U_\lambda(t)\bar{A}^\top + I,\ U_\lambda(0) = 0.$$

By (4.25),

$$\det\big(U_\lambda(t)\big) \geq \det\big(R_\lambda(t)\big),$$

whereas, by Proposition 3.6,

$$\det\big(U_\lambda(t)\big) \leq \nu_0^{\mathrm{d}},\quad \lambda > 0, \tag{4.28}$$

where ν_0 is the largest eigenvalue of the matrix $U_0(t)$, and then (4.27) follows from (4.28).

Finally, (4.21) follows from (4.23) and (4.24).

This completes the proof of Theorem 4.4.

$\square$

With the help of Theorem 4.4, we can now answer question [Q2] posed in Introduction.

Theorem 4.5 *Assume that, in equation* (1.1), *the matrix* $BB^\top$ *is invertible and the initial condition* $\boldsymbol{y}(0)$ *is independent of* $\boldsymbol{w}$ *and is a Gaussian vector with mean* $\boldsymbol{m}$ *and covariance* K*. Then, for every* $Q \geq 0$,

$$\ln \mathbb{P}\left(\int_0^t \boldsymbol{y}^\top(s)Q\boldsymbol{y}(s)ds \leq \varepsilon\right) \sim -\frac{1}{8}\Big(t\,\mathrm{Tr}\big(\bar{Q}^{1/2}\big) + \bar{\boldsymbol{m}}^\top Q_K\bar{\boldsymbol{m}}\Big)^2 \varepsilon^{-1}, \tag{4.29}$$

where $\bar{Q} = (BB^\top)^{1/2}Q(BB^\top)^{1/2}$, $\bar{\boldsymbol{m}} = (BB^\top)^{-1/2}\boldsymbol{m}$, $\bar{Q}^{1/2}$ *is the symmetric non-negative square root of* $\bar{Q}$, *and* Q_K *is the matrix from* (4.17).

Proof. With (4.18) in mind, we use (2.4), taking $\gamma = 1/2$ and

$$\alpha = 2^{-1/2}\Big(t\,\mathrm{Tr}\big(\bar{Q}^{1/2}\big) + \bar{\boldsymbol{m}}^\top Q_K\bar{\boldsymbol{m}}\Big).$$

$\square$

Theorem 4.5 shows that if the matrix $BB^\top$ is invertible, then the random variable

$$\xi = \int_0^t \boldsymbol{y}^\top(s) Q \boldsymbol{y}(s) ds$$

has small ball rate $\varpi = 1$ for every initial condition $\boldsymbol{y}(0)$, every drift matrix A, and every non-zero matrix $Q \geq 0$. The small ball constant

$$\mathfrak{C} = \frac{1}{8}\Big(t \operatorname{Tr}(\bar{Q}^{1/2}) + \bar{\boldsymbol{m}}^\top Q_K \bar{\boldsymbol{m}}\Big)^2$$

depends on both $\boldsymbol{m}$ and Q, but does not depend on the matrix A. If $A = 0$ and $\boldsymbol{y}(0) = \mathbf{0}$, then $\boldsymbol{y}$ is a Brownian motion with covariance matrix $BB^\top$ so that $\int_0^t \boldsymbol{y}^\top(s) Q \boldsymbol{y}(s) ds$ is equal in distribution to $\sum_{k=1}^{\mathrm{d}} \bar{\sigma}_k^2 \int_0^t w_k^2(s) ds$, where $w_1, \ldots, w_{\mathrm{d}}$ are independent one-dimensional standard Brownian motions and $\bar{\sigma}_1^2, \ldots, \bar{\sigma}_{\mathrm{d}}^2$ are the eigenvalues of the matrix $\bar{Q}$. The corresponding small deviations bound can then be derived, for example, from [10, Corollary 3.1], and this bound coincides with (4.29). In other words, as $\varepsilon \to 0$,

$$\begin{aligned}
\ln \mathbb{P}\left(\int_0^t \boldsymbol{y}^\top(s) Q \boldsymbol{y}(s) ds \leq \varepsilon\right) &\sim \ln \mathbb{P}\left(\sum_{k=1}^{\mathrm{d}} \int_0^t \bar{\sigma}_k^2 w_k^2(s) ds \leq \varepsilon\right) \\
&\sim -\frac{t^2}{8}\left(\sum_{k=1}^{\mathrm{d}} \bar{\sigma}_k\right)^2 \varepsilon^{-1}.
\end{aligned}$$

It is somewhat remarkable that the variance of the initial condition $\boldsymbol{y}(0)$ can affect the small ball constant for $\int_0^t \boldsymbol{y}^\top(s) Q \boldsymbol{y}(s) ds$. Indeed, consider the one-dimensional OU process (3.22) with initial condition having mean x_0 and variance σ_0^2. Equality (3.28) and the asymptotic relation (2.4) imply

$$\ln \mathbb{P}\left(\int_0^t x^2(s) ds \leq \varepsilon\right) \sim \begin{cases} -\dfrac{\sigma^2 t^2}{8}\,\varepsilon^{-1}, & \text{if } \sigma_0^2 > 0 \\[2ex] -\dfrac{(\sigma^2 t + x_0^2)^2}{8\sigma^2}\,\varepsilon^{-1}, & \text{if } \sigma_0^2 = 0. \end{cases}$$

More generally, if both K and Q are invertible, then $Q_K = 0$ and the small ball constant does not depend on the initial condition:

$$\ln \mathbb{P}\left(\int_0^t \boldsymbol{y}^\top(s) Q \boldsymbol{y}(s) ds \leq \varepsilon\right) \sim -\frac{t^2}{8}\Big(\operatorname{Tr}(\bar{Q}^{1/2})\Big)^2 \varepsilon^{-1}.$$

On the other hand, consider $A = 0$,

$$B = Q = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad K = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

so that

$$(I + C_\lambda K)^{-1} = \begin{pmatrix} 1/(1+\sqrt{2\lambda}) & 0 \\ 0 & 1 \end{pmatrix} \text{ and } Q_K = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

With $\boldsymbol{m} = (m_1\ m_2)^\top$, we get

$$\ln \mathbb{P}\left(\int_0^t \boldsymbol{y}^\top(s) Q \boldsymbol{y}(s) ds \leq \varepsilon\right) \sim -\frac{1}{8}\Big(2t + m_2^2\Big)^2 \varepsilon^{-1}.$$

5 Degenerate Noise: a Two-Dimensional Example

The objective of this section is to demonstrate how degenerate noise can destroy universality of the conclusion of Theorem 4.4. We will consider a two-dimensional example, corresponding to (1.1) with

$$A = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ \sigma \end{pmatrix}. \tag{5.1}$$

In other words, $\boldsymbol{y}^\top = (x \ \ \dot{x})$, and $x = x(t)$ satisfies

$$\ddot{x}(t) + a\dot{x}(t) + bx(t) = \sigma \dot{w}(t), \ t > 0. \tag{5.2}$$

Let

$$Q = \left(\begin{array}{cc} y & r \\ r & z \end{array} \right), \qquad y \geq 0, \ z \geq 0, \ yz \geq r^2,$$

be a symmetric non-negative definite matrix; occasionally, we will write $Q = (y \ r; r \ z)$.

We will now compute

$$\Psi(Q;t) = \mathbb{E} \exp \left(- \int_0^t \boldsymbol{y}^\top(s) Q \boldsymbol{y}(s) ds \right)$$

using Theorem 3.2.

To this end, define the numbers p and q by

$$p = \sqrt{a^2 + 2\sigma^2 z + 2q - 2b}, \quad q = \sqrt{b^2 + 2\sigma^2 y}.$$

To eliminate the trivial cases, we will always assume that $p > 0$ and $q > 0$.

Next, define the number $\nu = \sqrt{|(p^2/4) - q|}$ and the functions

$$\varphi(t) = \begin{cases} \frac{\sin \nu t}{\nu} \, e^{-pt/2}, & \text{if } p^2 - 4q < 0, \\ \frac{\sinh \nu t}{\nu} \, e^{-pt/2}, & \text{if } p^2 - 4q > 0, \\ te^{-pt/2}, & \text{if } p^2 - 4q = 0, \end{cases} \tag{5.3}$$

and

$$\mathfrak{a}(t) = \frac{1}{2pq} \big(1 - \dot{\varphi}^2(t) - (p^2 + q)\varphi^2(t) - p\dot{\varphi}(t)\varphi(t) \big), \tag{5.4a}$$

$$\mathfrak{b}(t) = \frac{\varphi^2(t)}{2}, \tag{5.4b}$$

$$\mathfrak{c}(t) = \frac{1}{2p} \big(1 - \dot{\varphi}^2(t) - q\varphi^2(t) \big). \tag{5.4c}$$

Direct computations show that the maximal symmetric solution of (3.7) with matrices A and B from (5.1) is

$$C = \sigma^{-2} \begin{pmatrix} pq - 2\sigma^2 \lambda r - ab & q - b \\ q - b & p - a \end{pmatrix},$$

and the matrix $R_*(t)$ from (3.10) is

$$R_*(t) = \sigma^2 \begin{pmatrix} \mathfrak{a}(t) & \mathfrak{b}(t) \\ \mathfrak{b}(t) & \mathfrak{c}(t) \end{pmatrix}.$$

If $\boldsymbol{y}(0) = \mathbf{0}$, that is, $x(0) = \dot{x}(0) = 0$ in (5.2), then (3.9) implies

$$\Psi(Q; t) = e^{-(p-a)t/2} \left[\det(I - CR_*(t))\right]^{-1/2}. \tag{5.5}$$

Sometimes, formula (5.5) can be simplified further. For example, consider the integrated Brownian motion $x(t) = \int_0^t w(s)ds$, corresponding to $a = b = 0$, $r = z = 0$, and $\sigma = 1$. Then (5.5) becomes

$$\Psi(t) = \frac{2}{\sqrt{\cos\left(2^{3/4} t y^{1/4}\right) + \cosh\left(2^{3/4} t y^{1/4}\right) + 2}},$$

a well-known result: cf. [7, Section 4.2.1] or [5, Theorem 3.1].

Here are two new examples.

1. Random harmonic oscillator $\ddot{x}(t) + bx(t) = \dot{w}(t)$, $b > 0$, corresponding to $a = 0$, $z = 0$, and $\sigma = 1$:

$$\Psi(t) = \frac{2\sqrt[4]{b^2+2y}}{\sqrt{(\sqrt{b^2+2y}-b)\cos(\sqrt{2}t\sqrt{\sqrt{b^2+2y}+b})+(\sqrt{b^2+2y}+b)\cosh(\sqrt{2}t\sqrt{\sqrt{b^2+2y}-b})+2\sqrt{b^2+2y}}};$$

2. Joint integrated Brownian motion and Brownian motion:

$$\Psi(t) = \begin{cases} \dfrac{2^{3/4}\sqrt{2y-z^2}}{\sqrt{(-\sqrt{y}z+\sqrt{2}y-\sqrt{2}z^2)\cos(t\sqrt{2\sqrt{2}\sqrt{y}-2z})+(\sqrt{y}z+\sqrt{2}y-\sqrt{2}z^2)\cosh(t\sqrt{2\sqrt{2}\sqrt{y}+2z})+2\sqrt{2}y}}, & z < \sqrt{2y}, \\[2ex] \dfrac{2^{3/4}\sqrt{z^2-2y}}{\sqrt{(-\sqrt{y}z+\sqrt{2}y-\sqrt{2}z^2)\cosh(t\sqrt{2z-2\sqrt{2}\sqrt{y}})+(\sqrt{y}z+\sqrt{2}y-\sqrt{2}z^2)\cosh(t\sqrt{2\sqrt{2}\sqrt{y}+2z})+2\sqrt{2}y}}, & z > \sqrt{2y}, \\[2ex] \dfrac{2}{\sqrt{\sqrt{2}t^2\sqrt{y}+\frac{3}{2}\cosh\left(2^{5/4}ty^{1/4}\right)+\frac{5}{2}}} = \dfrac{2}{\sqrt{t^2z+\frac{3}{2}\cosh\left(2t\sqrt{z}\right)+\frac{5}{2}}}, & z = \sqrt{2y}. \end{cases}$$

Even when the initial conditions are zero, the hight-frequency asymptotic of the function Ψ can depend on the matrix Q in a non-trivial way.

Theorem 5.1 *Assume that $x(0) = \dot{x}(0) = 0$. If $z = 0$ and $y > 0$, then*

$$\lim_{\lambda \to +\infty} \frac{\ln \Psi(\lambda Q; t)}{\lambda^{1/4}} = -2^{-1/4}\, \sigma^{1/2}\, y^{1/4}\, t. \tag{5.6}$$

If $z > 0$, then

$$\lim_{\lambda \to +\infty} \frac{\ln \Psi(\lambda Q; t)}{\lambda^{1/2}} = -2^{-1/2}\, \sigma\, z^{1/2}\, t. \tag{5.7}$$

Proof. When using (5.5), we now keep in mind that the matrix Q is re-scaled by the factor λ, so that y, r, z must be replaced with $\lambda y, \lambda r, \lambda z$. Throughout the proof we use the $O(\cdot)$ and $o(\cdot)$ notations for asymptotic comparison as $\lambda \to +\infty$, and write $f \sim g$ if $\lim_{\lambda \to +\infty} f/g = 1$.

Both (5.6) and (5.7) will follow from (5.5) once we show that

$$0 < \liminf_{\lambda \to +\infty} \det\left(I - CR_*(t)\right) \le \limsup_{\lambda \to +\infty} \det\left(I - CR_*(t)\right) < \infty. \tag{5.8}$$

The key question becomes the asymptotic, as $\lambda \to +\infty$, of the function φ from (5.3) for fixed $t > 0$. There are three cases to consider.

Case 1: $y > 0,\ r = z = 0$. As $\lambda \to +\infty$, we have $p^2 \sim 2q \sim 2(2\sigma^2\lambda y)^{1/2}$, so that $p^2 - 4q < 0$ for large λ and then, by (5.3),

$$\lim_{\lambda\to+\infty} \lambda^n \varphi(t) = \lim_{\lambda\to+\infty} \lambda^n \dot{\varphi}(t) = 0 \quad \text{for all } n > 0.$$

Then $\lim_{\lambda\to+\infty} \lambda^n \mathfrak{b}(t) = 0$ for all $n > 0$,

$$\mathfrak{a}(t) \sim \frac{1}{2pq},\ \mathfrak{c}(t) \sim \frac{1}{2p},\ \big(I - CR_*(t)\big) \sim \frac{1}{2}\left(\begin{array}{cc} 1 & -\frac{1}{p} \\ -\frac{q}{p} & 1 \end{array}\right),$$

so that

$$\det\big(I - CR_*(t)\big) \sim \frac{1}{4}\left(1 - \frac{q}{p^2}\right) \sim \frac{1}{8},$$

and (5.8) follows.

Case 2: $y = r = 0,\ z > 0$. Now q does not depend on λ and, as $\lambda \to +\infty$, we have $p^2 \sim 2\sigma^2\lambda z$, so that $p^2 - 4q > 0$ for large λ, $2\nu = \sqrt{p^2 - 4q} \sim p$, and

$$p - 2\nu = p - \sqrt{p^2 - 4q} = \frac{4q}{p + \sqrt{p^2 - 4q}} \sim \frac{2q}{p}.$$

As a result, (5.3) and (5.4a)–(5.4c) lead to

$$e^{(2\nu-p)t/2} \sim 1,\ \varphi(t) = \frac{\sinh(\nu t)}{\nu} e^{-pt/2} \sim \frac{1}{p},\ \dot{\varphi}(t) = \left(\cosh(\nu t) - \frac{p}{2\nu}\sinh(\nu t)\right) e^{-pt/2} \sim -\frac{q}{p^2}.$$

Then

$$I - CR_*(t) = \left(\begin{array}{cc} 1 & 0 \\ 0 & 1/2 \end{array}\right) + o(1),\ \ \det\big(I - CR_*(t)\big) = \frac{1}{2} + o(1),$$

and (5.8) follows.

Case 3: $y > 0,\ z > 0,\ r^2 \le yz$. We have $p^2 \sim 2\sigma^2\lambda z,\ q \sim \sqrt{2\sigma^2\lambda y}$ so that $p^2 - 4q > 0$ for large λ, $2\nu = \sqrt{p^2 - 4q} \sim p$, and

$$p - \sqrt{p^2 - 4q} = \frac{4q}{p + \sqrt{p^2 - 4q}} \sim 2\sqrt{\frac{y}{z}}.$$

With notations

$$\theta = \sqrt{\frac{y}{z}},\quad h = e^{-\theta t},\quad \beta = 1 - \frac{r}{\sqrt{yz}}, \tag{5.9}$$

equalities (5.3) and (5.4a)–(5.4c) lead to

$$\frac{q}{p} \sim \beta,\ e^{(2\nu-p)t/2} \sim h,$$
$$\varphi(t) = \frac{\sinh(\nu t)}{\nu} e^{-pt/2} \sim \frac{h}{p},\ \dot{\varphi}(t) = \left(\cosh(\nu t) - \frac{p}{2\nu}\sinh(\nu t)\right) e^{-pt/2} \sim -\frac{qh}{p^2} \sim -\frac{\theta h}{p},$$
$$\mathfrak{a}(t) \sim \frac{1 - h^2}{2pq},\ \mathfrak{b}(t) \sim \frac{h^2}{2p^2},\ \mathfrak{c}(t) \sim \frac{1}{2p}.$$

Then

$$I - CR_*(t) = \begin{pmatrix} 1 - \frac{\beta(1-h^2)}{2} + o(1) & O(1/p) \\ O(1) & \frac{1}{2} + o(1) \end{pmatrix},$$

$$\det\big(I - CR_*(t)\big) \sim \frac{2 - \beta(1-h^2)}{4},$$

and (5.8) follows.

This completes the proof of Theorem 5.1.

□

Theorems 2.2 and 5.1 lead to the logarithmic asymptotic of probability of small deviations for the corresponding quadratic functionals. The result shows that the small ball rate depends on whether $z = 0$ (cf. (5.10)) or $z > 0$ (cf. (5.11)). Equivalently, the random variables $\int_0^t x^2(s)ds$ and $\int_0^t \dot{x}^2(s)ds$ have different small ball rates.

Theorem 5.2 *Let $x = x(t)$ be the solution of equation* (5.2)*. For $y > 0$,*

$$\ln \mathbb{P}\left(y \int_0^t x^2(s)ds \le \varepsilon\right) \sim -\frac{3\sigma^{2/3}y^{1/3}t^{4/3}}{8}\,\varepsilon^{-1/3}. \tag{5.10}$$

If $\boldsymbol{y}(t) = \big(x(t),\ \dot{x}(t)\big)^\top$, $Q = (y\ r; r\ z)$, $y \ge 0,\ z > 0,\ r^2 \le yz$, then

$$\ln \mathbb{P}\left(\int_0^t \boldsymbol{y}^\top(s)Q\boldsymbol{y}(s)ds \le \varepsilon\right) \sim -\frac{\sigma^2 z t^2}{8}\,\varepsilon^{-1}. \tag{5.11}$$

Proof. To establish (5.10), note that (5.6) is (2.2) with

$$\gamma = 1/4 \ \text{ and } \ \alpha = 2^{-1/4}\sigma^{1/2}y^{1/4}t.$$

Then (5.10) follows from (2.4).

To establish (5.11), note that (5.7) is (2.2) with

$$\gamma = 1/2 \ \text{ and } \ \alpha = 2^{-1/2}\sigma z^{1/2}t.$$

Then (5.11) follows from (2.4).

□

Let us now consider the effects of non-zero initial conditions. As we saw in the previous section, if the matrix $BB^\top$ is non-singular, then the initial conditions can increase the small ball constant $\mathfrak{C}$, but do not change the small ball rate ϖ (cf. Theorem 4.5). For equation (5.2), the initial conditions can change the small ball rate as well. The most dramatic change takes place when $yz > r^2 \ge 0$ and the initial condition is non-random ($K = 0$) with $x(0) = m_1 \ne 0$. In this case, the quadratic form becomes uniformly bounded from below: there exists $\varepsilon_0 > 0$ such that

$$\mathbb{P}\left(\int_0^t \boldsymbol{y}^\top(s)Q\boldsymbol{y}(s)ds \le \varepsilon_0\right) = 0. \tag{5.12}$$

Informally, the small ball rate ϖ becomes infinite. While this might appear surprising at first, a simple application of the Cauchy-Schwartz inequality shows that the result is to be expected. For example, take

$$x(t) = m_1 + \int_0^t w(s)ds,\ \dot{x}(t) = w(t),\ y = z = 1,\ r = 0, \tag{5.13}$$

so that

$$\int_0^t \boldsymbol{y}^\top(s)Q\boldsymbol{y}(s)ds = \int_0^t \left(\left(m_1 + \int_0^s w(\tau)d\tau\right)^2 + w^2(s)\right) ds.$$

Using

$$2m_1 \int_0^s w(\tau)d\tau \geq -\frac{m_1^2}{2} - 2\left(\int_0^s w(\tau)d\tau\right)^2,$$

we get

$$\boldsymbol{y}^\top(s)Q\boldsymbol{y}(s) \geq \frac{m_1^2}{2} - \left(\int_0^s w(\tau)d\tau\right)^2 + w^2(s).$$

Next, by Cauchy-Schwartz,

$$\left(\int_0^s w(\tau)d\tau\right)^2 \leq s\int_0^s w^2(\tau)d\tau,$$

so that

$$\int_0^t \boldsymbol{y}^\top(s)Q\boldsymbol{y}(s)ds \geq \frac{m_1^2 t}{2} + \left(1 - \frac{t^2}{2}\right)\int_0^t w^2(s)ds,$$

and, as long as $t \leq \sqrt{2}$, we get (5.12) with $\varepsilon_0 = m_1^2 t/2$.

Asymptotic analysis of $\Psi(\lambda Q; t)$ as $\lambda \to +\infty$, together with Proposition 2.3, leads to a sharper bound in a more general setting; cf. (5.14) below. In particular, for (5.13), we actually have (5.12) for every $t > 0$ with $\varepsilon_0 = m_1^2 \tanh(t)$.

Theorem 5.3 *Assume $\sqrt{yz} > |r| \geq 0$, $K = 0$, and $m_1 \neq 0$. Then, using the notations from* (5.9),

$$\mathbb{P}\left(\int_0^t \boldsymbol{y}^\top(s)Q\boldsymbol{y}(s)ds \leq \varepsilon\right) = 0 \text{ for } \varepsilon < \sqrt{yz}\,\frac{\beta(2-\beta)(1-h^2)}{2-\beta(1-h^2)}\,m_1^2. \tag{5.14}$$

Proof. By Proposition 2.3, we need to show that

$$\lim_{\lambda\to+\infty} \frac{\ln \Psi(\lambda Q; t)}{\lambda} = -\sqrt{yz}\,\frac{\beta(2-\beta)(1-h^2)}{2-\beta(1-h^2)}\,m_1^2. \tag{5.15}$$

To simplify the presentation, and keeping in mind that the matrix C depends on λ, we will write

$$P_\lambda = e^{t(A-BB^\top C)},\ \boldsymbol{\mu}_\lambda = P_\lambda \boldsymbol{m},\ \boldsymbol{u}_\theta = \begin{pmatrix} 1 \\ -\theta \end{pmatrix},$$

and then (3.11) becomes

$$2\ln\Psi(\lambda Q; t) = -pt - \ln\Big(\det\big(\widetilde{S}_\lambda\big)\Big) + \widetilde{\boldsymbol{\mu}}_\lambda^\top \widetilde{S}_\lambda \widetilde{\boldsymbol{\mu}}_\lambda,$$

where

$$\widetilde{S}_\lambda = (I - \widetilde{C}_\lambda\widetilde{R}_\lambda)^{-1}\widetilde{C}_\lambda,\ \ \widetilde{C}_\lambda = \begin{pmatrix} -C & 0 \\ 0 & C \end{pmatrix},$$

$$\widetilde{R}_\lambda = \begin{pmatrix} K & KP_\lambda^\top \\ P_\lambda K & P_\lambda K P_\lambda^\top + R_* \end{pmatrix},\ \ R_*(t) = \sigma^2 \begin{pmatrix} \mathfrak{a}(t) & \mathfrak{b}(t) \\ \mathfrak{b}(t) & \mathfrak{c}(t) \end{pmatrix},\ \ \widetilde{\boldsymbol{\mu}}_\lambda = \begin{pmatrix} \boldsymbol{m} \\ \boldsymbol{\mu}_\lambda \end{pmatrix}.$$

By construction, $\det\left(\widetilde{S}_\lambda\right) = O(\lambda^n)$ for some $n \ge 0$, and therefore

$$\ln \Psi(\lambda Q; t) \sim -\frac{pt}{2} + \frac{1}{2}\widetilde{\boldsymbol{\mu}}^\top \widetilde{S}_\lambda \widetilde{\boldsymbol{\mu}}. \tag{5.16}$$

If $K = 0$, then, by (5.16),

$$\ln \Psi(\lambda Q; t) \sim 2^{-1/2}\sigma z^{1/2} t\, \lambda^{1/2} - \frac{1}{2}\boldsymbol{m}^\top C \boldsymbol{m} + \frac{1}{2}\boldsymbol{\mu}_\lambda^\top (I - CR_*)^{-1} C \boldsymbol{\mu}_\lambda. \tag{5.17}$$

If $\sqrt{yz} > |r|$, then

$$C = \begin{pmatrix} 2\sqrt{yz}\,\beta\,\lambda + O(\lambda^{1/2}) & \sqrt{2y}\sigma^{-1}\lambda^{1/2} + O(1) \\ \sqrt{2y}\sigma^{-1}\lambda^{1/2} + O(1) & \sqrt{2z}\sigma^{-1}\lambda^{1/2} + O(1) \end{pmatrix},$$

$$R_* = \begin{pmatrix} (1-h^2)4^{1/2}(yz)^{-1/2}\lambda^{-1} + O(\lambda^{-3/2}) & h^2(4z)^{-1}\lambda^{-1} + O(\lambda^{-3/2}) \\ h^2(4z)^{-1}\lambda^{-1} + O(\lambda^{-3/2}) & \sigma 2^{-3/2} z^{-1/2}\,\lambda^{-1/2} + O(\lambda^{-1}) \end{pmatrix},$$

$$C^{-1} = \begin{pmatrix} (2\beta)^{-1}(yz)^{-1/2}\,\lambda^{-1} + O(\lambda^{-3/2}) & -(2\beta z)^{-1}\,\lambda^{-1} + O(\lambda^{-3/2}) \\ -(2\beta z)^{-1}\,\lambda^{-1} + O(\lambda^{-3/2}) & \sigma(2z)^{-1/2}\,\lambda^{-1/2} + O(\lambda^{-1}) \end{pmatrix},$$

$$(C^{-1} - R_*)^{-1} = \frac{8\sqrt{2y}\,z\,\beta}{\sigma\big(2 - \beta(1-h^2)\big)}$$

$$\times \begin{pmatrix} \sigma^2 2^{-3/2} z^{-1/2}\,\lambda + O(\lambda^{1/2}) & (2+\beta h^2)(4z\beta)^{-1}\,\lambda^{1/2} + O(1) \\ (2+\beta h^2)(4z\beta)^{-1}\,\lambda^{1/2} + O(1) & \big(2 - \beta(1-h^2)\big)(4\beta)^{-1}(yz)^{-1/2}\,\lambda^{1/2} + O(1) \end{pmatrix},$$

$$\boldsymbol{\mu}_\lambda = m_1 h \boldsymbol{u}_\theta + O(\lambda^{-1/2}).$$

If $m_1 \neq 0$, then, by (5.17),

$$\ln \Psi(\lambda Q; t) \sim -\sqrt{yz}\,\beta m_1^2 \left(1 - \frac{2h^2}{2 - \beta(1-h^2)}\right)\lambda = -\sqrt{yz}\, m_1^2\, \frac{\beta(2-\beta)(1-h^2)}{2-\beta(1-h^2)}\,\lambda,$$

and (5.15) follows. Note that the right-hand side of (5.15) vanishes when $\beta = 0$ or $\beta = 2$, that is, exactly when the matrix Q is singular.

□

Further inspection of (5.16) suggests the following qualitative summary of the effects of non-zero initial conditions on small ball probabilities:

- If $\det K > 0$, then the small ball asymptotic at logarithmic scale does not depend on the initial conditions, that is, (5.11) holds when $z > 0$ and (5.10) holds when $z = 0,\ y > 0$.
- If $\det(K) = 0$, then the initial conditions can increase the small ball rate ϖ in the following two cases: (a) $y > 0,\ z = 0$ and (b) $yz > r^2 \ge 0$.
- In all other cases, the initial conditions do not change the small ball rate ϖ, but can increase the small ball constant $\mathfrak{C}$.

At this point, the amount of computations necessary to derive the corresponding quantitative results requires writing a separate paper.

References

[1] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, Cambridge University Press, Cambridge, 1987.

[2] X. Chen and W. V. Li. Quadratic functionals and small ball probabilities for the m-fold integrated Brownian motion. *Ann. Probab.*, **31** (2003), 1052–1077.

[3] G. Da Prato and J. Zabczyk. *Ergodicity for infinite-dimensional systems*, Cambridge University Press, Cambridge, 1996.

[4] M. D Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time. IV. *Comm. Pure Appl. Math.*, **36** (1983), 183–212.

[5] D. Khoshnevisan and Z. Shi. Chung's law for integrated Brownian motion. *Trans. Amer. Math. Soc.*, **350** (1998), 4253–4264.

[6] M. L. Kleptsyna and A. Le Breton. Optimal linear filtering of general multidimensional Gaussian processes and its application to Laplace transforms of quadratic functionals. *J. Appl. Math. Stochastic Anal.*, **14** (2001), 215–226.

[7] M. L. Kleptsyna and A. Le Breton. A Cameron-Martin type formula for general Gaussian processes—a filtering approach. *Stoch. Stoch. Rep.*, **72** (2002), 229–250.

[8] K. Koncz. On the parameter estimation of diffusional type processes with constant coefficients (elementary Gaussian processes). *Anal. Math.*, **13** (1987), 75–91.

[9] P. Lancaster and L. Rodman. *Algebraic Riccati equations.* Oxford University Press, New York, 1995.

[10] W. V. Li and Q.-M. Shao. Gaussian processes: inequalities, small ball probabilities and applications, in: *Stochastic processes: theory and methods*, D. N. Shanbhag and C. R. Rao, eds, Handbook of Statistics, Vol. 19, North-Holland, Amsterdam, 2001, pp. 533–597.

[11] R. S. Liptser and A. N. Shiryaev. *Statistics of random processes, I: General theory*, Springer, Berlin, 2001.

[12] R. S. Liptser and A. N. Shiryaev. *Statistics of random processes, II: Applications*, Springer, Berlin, 2001.

[13] L. Mirsky. *An introduction to linear algebra.* Dover Publications, Inc., New York, 1990.

[14] H. L. Royden. Comparison theorems for the matrix Riccati equation. *Comm. Pure Appl. Math.*, **41** (1988), 739–746.

[15] D. Serre. *Matrices: Theory and applications*, Springer, New York, 2010.

[16] G. N. Sytaya. On some asymptotic representation of the Gaussian measure in a Hilbert space. *Theory of Stochastic Processes*, **2** (1974), 93–104. Russian.

[17] A. I. Yashin. An extension of the Cameron-Martin result. *J. Appl. Probab.*, **30** (1993), 247–251.