Journal of Mathematical Physics

RESEARCH ARTICLE | FEBRUARY 07 2023

Large deviations for quadratic functionals of stable Gauss–Markov chains and entropy production

Marco Zamparo ; Massimiliano Semeraro

J. Math. Phys. 64, 023302 (2023)
https://doi.org/10.1063/5.0096315

Journal of Mathematical Physics ARTICLE scitation.org/journal/jmp

Large deviations for quadratic functionals of stable Gauss–Markov chains and entropy production

Cite as: J. Math. Phys. 64, 023302 (2023); doi: 10.1063/5.0096315
Submitted: 17 April 2022 • Accepted: 16 January 2023 •
Published Online: 7 February 2023

Marco Zamparo[a)] and Massimiliano Semeraro[b)]

AFFILIATIONS

Dipartimento di Fisica, Università degli Studi di Bari and INFN, Sezione di Bari, via Amendola 173, 70126 Bari, Italy

[a)] Author to whom correspondence should be addressed: marco.zamparo@uniba.it
[b)] E-mail: massimiliano.semeraro@uniba.it

ABSTRACT

In this paper, we establish a large deviation principle for the entropy production rate of possible non-stationary, centered stable Gauss–Markov chains, verifying the Gallavotti–Cohen symmetry. We reach this goal by developing a large deviation theory for quasi-Toeplitz quadratic functionals of multivariate centered stable Gauss–Markov chains, which differ from a perfect Toeplitz form by the addition of quadratic boundary terms.

I. INTRODUCTION

Let on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ be given a sequence $\{X_n\}_{n\geq 1}$ of random variables taking values in a Polish space $\mathcal{X}$. For any integer $N \geq 1$, let $\mu_N^+ := \mathbb{P}[(X_1, \ldots, X_N) \in \cdot]$ and $\mu_N^- := \mathbb{P}[(X_N, \ldots, X_1) \in \cdot]$ be the probability measures on the Borel σ-field $\mathscr{B}(\mathcal{X}^N)$ induced by the direct process and the reverse process, respectively. The *entropy production rate* up to time N is the real random variable on $(\Omega, \mathscr{F}, \mathbb{P})$ defined by

$$e_N := \begin{cases} \dfrac{1}{N} \ln \left[\dfrac{d\mu_N^+}{d\mu_N^-}(X_1, \ldots, X_N) \right] & \text{if } \mu_N^+ \ll \mu_N^-, \\ +\infty & \text{otherwise.} \end{cases}$$

The entropy production rate turns out to be a natural measure of irreversibility since $e_N = 0$ for all $N \geq 1$ if and only if the sequence $\{X_n\}_{n\geq 1}$ is reversible, namely, if and only if $(X_N, \ldots, X_1)$ is distributed as $(X_1, \ldots, X_N)$ for every N. The use of the entropy production rate to quantify the irreversibility of a stochastic process was proposed by Kurchan[1] and in more generality by Lebowitz, Spohn, and Maes,[2,3] who extended the seminal work by Gallavotti and Cohen[4] in the context of deterministic dynamical systems. Since then, the entropy production rate has become a basic topic in non-equilibrium statistical physics.[5–13] The entropy production rate came out with a supposed symmetry associated with its large fluctuations, which, in fact, was discovered by Gallavotti and Cohen[4] prompted by results of computer simulations.[14] They dubbed this symmetry a "fluctuation theorem." The appropriate formalism for describing the large fluctuations of the entropy production rate is large deviation theory.[15,16] The entropy production rate e_N is said to satisfy a *large deviation principle* with the *rate function* I if there exists a function I with compact level sets such that for each Borel set $\mathcal{B} \subseteq \mathbb{R}$,

$$-\inf_{w\in\mathcal{B}^o}\{I(w)\} \leq \liminf_{N\uparrow\infty} \frac{1}{N} \ln \mathbb{P}[e_N \in \mathcal{B}] \leq \limsup_{N\uparrow\infty} \frac{1}{N} \ln \mathbb{P}[e_N \in \mathcal{B}] \leq -\inf_{w\in\bar{\mathcal{B}}}\{I(w)\},$$

where $\mathcal{B}^o$ and $\bar{\mathcal{B}}$ are the interior and the closure of $\mathcal{B}$, respectively. The "fluctuation theorem" refers to a property of the function I. The rate function I is said to satisfy the *Gallavotti–Cohen symmetry* if for all $w \in \mathbb{R}$,

$$I(-w) = I(w) + w.$$

It has been pointed out that the Gallavotti–Cohen symmetry is an intrinsic property of I, which is met whenever e_N satisfies a large deviation principle.[8–10]

In this paper, we investigate the large fluctuations of the entropy production rate and the Gallavotti–Cohen symmetry for a possible non-stationary, centered stable Gauss–Markov chain $\{X_n\}_{n\geq 1}$ valued in $\mathcal{X} := \mathbb{R}^d$ with any dimension $d \geq 1$. Thus, we assume that there exists a *drift matrix* $S \in \mathbb{R}^{d\times d}$ with spectral radius $\rho(S) < 1$ such that

$$X_{n+1} = SX_n + G_n \tag{1.1}$$

for all $n \geq 1$, with $\{G_n\}_{n\geq 1}$ being a sequence of i.i.d. standard Gaussian random vectors valued in $\mathbb{R}^d$ and independent of X_1. We suppose that X_1 is a Gaussian random vector with mean zero and general positive-definite covariance matrix Σ_o. The process $\{X_n\}_{n\geq 1}$ is stationary if and only if $\Sigma_o = \Sigma_s := \sum_{k\geq 0} S^k (S^\top)^k$, and it is reversible if and only if S is symmetric and $\Sigma_o = \Sigma_s$.[17] Stability corresponds to the hypothesis $\rho(S) < 1$, which implies that Σ_s actually exists.

The large deviation principle for the entropy production rate and the Gallavotti–Cohen symmetry have been rigorously established for finite-state Markov chains with discrete and continuous time.[8,18] The same has been done for multivariate stationary Ornstein–Uhlenbeck processes with normal drift matrix,[19–21] i.e., with a drift matrix that commutes with its adjoint, and for a model of heat conduction through a chain of anharmonic oscillators coupled to two reservoirs at different temperatures.[22] The mathematical tool underlying these works is the Gärtner–Ellis theorem,[15,16] and the lack of a large deviation principle for the entropy production rate of more general stochastic processes is due to non-satisfiability of the hypotheses of that theorem. An attempt to overcome the limitations of the Gärtner–Ellis theorem has been done for stationary diffusion processes, for which a large deviation principle for the entropy production rate has been obtained in the limit of vanishing noise by resorting to the classical Freidlin–Wentzell theory.[23] The autoregressive model (1.1) we consider basically is the discrete-time version of a d-dimensional centered Ornstein–Uhlenbeck process. The main contribution of our work stems from the fact that we do not assume that the drift matrix S is normal or that the chain is stationary. This generality prevents the use of the Gärtner–Ellis theorem to get at a large deviation principle for the entropy production rate e_N. The way we go around this key point is to regard e_N as a quadratic functional and to establish a large deviation principle for the class of quadratic functionals to which the entropy production rate belongs via a time-dependent change of probability measure. We need such a general principle to also tackle a problem of large fluctuations in an active matter model.[24] The following lemma provides the explicit expression of e_N for the model (1.1) as a quadratic form. The simple proof is reported in Appendix A. We denote by $\langle\cdot,\cdot\rangle$ the standard inner product of $\mathbb{R}^d$.

Lemma 1.1. Let $\{X_n\}_{n\geq 1}$ be a d-dimensional centered Gauss–Markov chain with drift matrix S and initial positive-definite covariance Σ_o. For each $N \geq 1$,

$$Ne_N = \frac{1}{2}\big\langle X_1, (I - \Sigma_o^{-1} - S^\top S)X_1\big\rangle + \frac{1}{2}\big\langle X_N, (\Sigma_o^{-1} + S^\top S - I)X_N\big\rangle + \sum_{n=2}^{N}\big\langle X_n, (S - S^\top)X_{n-1}\big\rangle.$$

Lemma 1.1 shows that the entropy production Ne_N of the process $\{X_n\}_{n\geq 1}$ is a particular instance of a quasi-Toeplitz quadratic functional W_N having the form

$$W_N := \frac{1}{2}\langle X_1, LX_1\rangle + \frac{1}{2}\sum_{n=1}^{N}\langle X_n, UX_n\rangle + \sum_{n=2}^{N}\langle X_n, VX_{n-1}\rangle + \frac{1}{2}\langle X_N, RX_N\rangle, \tag{1.2}$$

with L, U, V, and R being four matrices in $\mathbb{R}^{d\times d}$ with L, U, and R symmetric. In fact, W_N turns out to be Ne_N when $L := I - \Sigma_o^{-1} - S^\top S$, $U := 0$, $V := S - S^\top$, and $R := \Sigma_o^{-1} + S^\top S - I$. The circumstance to be stressed is that the problem of the entropy production rate leads to perturb a perfect Toeplitz structure, defined by the matrices U and V, through the addition of the quadratic boundary terms $\langle X_1, LX_1\rangle$ and $\langle X_N, RX_N\rangle$ in such a way that the coefficient matrix of W_N differs from a block tridiagonal Toeplitz matrix by the first and last diagonal blocks. This circumstance required new large deviation principles for quadratic forms of Gauss–Markov chains to be developed and to be added to the literature on large deviations for Gaussian processes. Similar to e_N, we say that W_N/N satisfies a *large deviation principle* with the *rate function* I if there exists a function I with compact level sets such that for each Borel set $\mathcal{B} \subseteq \mathbb{R}$,

$$-\inf_{w\in\mathcal{B}^o}\{I(w)\} \leq \liminf_{N\uparrow\infty}\frac{1}{N}\ln\mathbb{P}\left[\frac{W_N}{N}\in\mathcal{B}\right] \leq \limsup_{N\uparrow\infty}\frac{1}{N}\ln\mathbb{P}\left[\frac{W_N}{N}\in\mathcal{B}\right] \leq -\inf_{w\in\bar{\mathcal{B}}}\{I(w)\}.$$

Large deviation principles for Gaussian processes have been an active field of research since the pioneering works by Donsker and Varadhan[25] and Bryc and Dembo[26] on the large fluctuations of empirical measures for stationary Gaussian processes. The focus soon moved to large deviations of quadratic functionals,[27,28] which, in general, cannot be tackled by a direct application of the Gärtner–Ellis theorem since steepness of the asymptotic cumulant generating function is not guaranteed. The asymptotic cumulant generating function does not contain apparently the whole information on the large deviation property of the process: there is a loss of information passing to the limit. For Toeplitz quadratic forms of stationary centered Gaussian sequences, large deviation principles are now well-established,[29–31] as well as some moderate deviation principles.[32] These results have been obtained by a sharp study of the spectrum of a product of two Toeplitz matrices. For stationary centered Gaussian sequences, large deviations have been also characterized for special Hermitian quadratic forms[29,33] and a sample path large deviation principle has been deduced for the squares of the process.[34] Despite this progress, there are no general results to deal with non-stationary Gaussian sequences and perturbations of Toeplitz quadratic functionals, which pose very specific problems.

The typical value v of W_N/N in the large N limit is described by the following law of large numbers, which is verified in Appendix B:

Proposition 1.1. Let $\{X_n\}_{n\geq1}$ be a d-dimensional centered Gauss–Markov chain with drift matrix S and stationary covariance Σ_s. Then,

$$\lim_{N\uparrow\infty}\frac{W_N}{N}=\frac{1}{2}\,\mathrm{tr}\Big[\Big(U+V^{\top}S+S^{\top}V\Big)\Sigma_s\Big]=:v\quad \mathbb{P}-a.s.$$

The typical value v does not depend on the initial condition and on the boundary perturbations. On the contrary, we shall see that non-stationarity and quadratic boundary terms, which are intimately related for Gauss–Markov chains, can affect deviations of quadratic functionals from the mean and shape rate functions. This is not surprising since squares of Gaussian random variables have an exponential, rather than super-exponential, tail probability. Researchers have already come across this issue. In fact, the maximum likelihood estimator and the Yule–Walker estimator for the drift parameter of a one-dimensional autoregressive stable process satisfy large deviation principles with different rate functions.[29] These two estimators are connected to quadratic functionals that differ exactly by a quadratic boundary term. A similar phenomenology holds for the continuous-time counterpart, i.e., the Ornstein–Uhlenbeck process.[35] Coming closer to statistical physics, quadratic boundary perturbations of the entropy production rate for Ornstein–Uhlenbeck processes with normal drift matrix have been considered to account for the heat dissipation rate of a network of thermally driven harmonic oscillators.[20]

The remainder of this paper is organized as follows: In Sec. I A, we present the main results of this work: a large deviation principle for quadratic functionals of type (1.2) in the context of multivariate centered stable Gauss–Markov chains and the Gallavotti–Cohen symmetry of the entropy production rate function. In Sec. I B, we apply the theory to the very special class of stable Gauss–Markov chains with normal drift matrix, making contact with previous results. Section II provides the proof of the large deviation principle for the quadratic functionals. Section III reports the proof of the Gallavotti–Cohen symmetry of the entropy production rate function.

A. Main results

From now on, we regard $\mathbb{R}^d$ as a subset of $\mathbb{C}^d$, and we denote by $\langle\cdot,\cdot\rangle$ the standard inner product of $\mathbb{C}^d$. We write $A>0$ to specify a positive-definite Hermitian matrix $A\in\mathbb{C}^{d\times d}$.

Fix matrices S, Σ_o, L, U, V, and R in $\mathbb{R}^{d\times d}$ with $\rho(S)<1$, $\Sigma_o>0$, and L, U, and R symmetric. According to (1.1) and (1.2), they define a centered stable Gauss–Markov chain $\{X_n\}_{n\geq1}$ and a quadratic functional W_N for each $N\geq1$. For every $\lambda\in\mathbb{R}$ and $\theta\in[0,2\pi]$, we make use of S, U, and V to construct the Hermitian matrix,

$$F_\lambda(\theta):=\Big(I-S^{\top}e^{\mathrm{i}\theta}\Big)\Big(I-Se^{-\mathrm{i}\theta}\Big)-\lambda\Big(U+Ve^{-\mathrm{i}\theta}+V^{\top}e^{\mathrm{i}\theta}\Big)\in\mathbb{C}^{d\times d},\tag{1.3}$$

and we set

$$f_\lambda:=\inf_{\theta\in[0,2\pi]}\inf_{\substack{z\in\mathbb{C}^d\\ z\neq0}}\left\{\frac{\langle z,F_\lambda(\theta)z\rangle}{\langle z,z\rangle}\right\}.\tag{1.4}$$

As f_λ bounds the spectrum of $F_\lambda(\theta)$ from below for all θ, if $f_\lambda>0$, then the functions that map $\theta\in[0,2\pi]$ in $\ln\det F_\lambda(\theta)$ and $F_\lambda^{-1}(\theta)$ are well-defined and continuous. Thus, for each $\lambda\in\mathbb{R}$ such that $f_\lambda>0$, we can introduce the integrals

$$\varphi(\lambda):=-\frac{1}{4\pi}\int_0^{2\pi}\ln\,\det\,F_\lambda(\theta)\,d\theta\tag{1.5}$$

and

$$\Phi_\lambda(n):=\frac{1}{2\pi}\int_0^{2\pi}F_\lambda^{-1}(\theta)e^{-\mathrm{i}n\theta}d\theta\tag{1.6}$$

with $n \in \mathbb{Z}$. We point out that the set of λ for which $f_\lambda > 0$ is an interval since f_λ is concave with respect to λ. It will turn out that $\varphi(\lambda)$ is the value at λ of the cumulant generating function of W_N in the large N limit: $\lim_{N\uparrow\infty}(1/N)\ln\mathbb{E}[e^{\lambda W_N}] = \varphi(\lambda)$. In order to specify the effective domain of the asymptotic cumulant generating function, we need the following technical lemma, whose proof is reported in Sec. II:

Lemma 1.2. Let $\lambda \in \mathbb{R}$ be such that $f_\lambda > 0$. The following conclusions hold:

1. $H_\lambda := I + (S + \lambda V)\Phi_\lambda(1) \in \mathbb{C}^{d\times d}$ *is invertible, and the matrix*

$$\mathcal{L}_\lambda := \Sigma_o^{-1} + S^\top S - \lambda(U + L) - (S^\top + \lambda V^\top)\Phi_\lambda(0)H_\lambda^{-1}(S + \lambda V) \in \mathbb{C}^{d\times d}$$

is Hermitian;

2. $K_\lambda := I + \Phi_\lambda(1)(S + \lambda V) \in \mathbb{C}^{d\times d}$ *is invertible, and the matrix*

$$\mathcal{R}_\lambda := I - \lambda(U + R) - (S + \lambda V)K_\lambda^{-1}\Phi_\lambda(0)(S^\top + \lambda V^\top) \in \mathbb{C}^{d\times d}$$

is Hermitian.

Lemma 1.2 states that the matrices $\mathcal{L}_\lambda$ and $\mathcal{R}_\lambda$ are well-defined and Hermitian when $\lambda \in \mathbb{R}$ satisfies $f_\lambda > 0$. It makes then sense to consider the extended real numbers,

$$\lambda_- := \inf\{\lambda \in \mathbb{R} \ : \ f_\lambda > 0,\ \mathcal{L}_\lambda > 0,\ \text{and}\ \mathcal{R}_\lambda > 0\} \tag{1.7}$$

and

$$\lambda_+ := \sup\{\lambda \in \mathbb{R} \ : \ f_\lambda > 0,\ \mathcal{L}_\lambda > 0,\ \text{and}\ \mathcal{R}_\lambda > 0\}. \tag{1.8}$$

We are now in the position to present the first main result of this paper, which establishes a large deviation principle for W_N/N and is proved in Sec. II via a time-dependent change of measure.

Theorem 1.1. *The following conclusions hold:*

1. $\lambda_- < 0 < \lambda_+$ *and the convex function I that maps $w \in \mathbb{R}$ in $I(w) := \sup_{\lambda\in(\lambda_-,\lambda_+)}\{w\lambda - \varphi(\lambda)\}$ has compact level sets;*
2. *the quadratic functional W_N/N associated with the stable Gauss–Markov chain $\{X_n\}_{n\ge1}$ satisfies a large deviation principle with the rate function I.*

Theorem 1.1 outperforms the Gärtner–Ellis theorem, which requires that the asymptotic cumulant generating function exists and defines an essentially smooth, lower semicontinuous function.[15,16] In Sec. II, we shall prove that $\lim_{N\uparrow\infty}(1/N)\ln\mathbb{E}[e^{\lambda W_N}] = \varphi(\lambda)$ if $\lambda \in (\lambda_-,\lambda_+)$ and $\lim_{N\uparrow\infty}(1/N)\ln\mathbb{E}[e^{\lambda W_N}] = +\infty$ if $\lambda \notin \overline{(\lambda_-,\lambda_+)}$. We shall also verify that the function φ that maps $\lambda \in (\lambda_-,\lambda_+)$ in $\varphi(\lambda)$ is convex and differentiable so that the limits $\lim_{\lambda\downarrow\lambda_-}\varphi(\lambda) =: \varphi_-$, $\lim_{\lambda\uparrow\lambda_+}\varphi(\lambda) =: \varphi_+$, $\lim_{\lambda\downarrow\lambda_-}\varphi'(\lambda) =: d_-$, and $\lim_{\lambda\uparrow\lambda_+}\varphi'(\lambda) =: d_+$ exist. If even the limit $\lim_{N\uparrow\infty}(1/N)\ln\mathbb{E}[e^{\lambda W_N}]$ existed for all $\lambda \in \mathbb{R}$ and defined a lower semicontinuous function as demanded by the Gärtner–Ellis theorem, what is generally missing to guarantee essentially smoothness of the asymptotic cumulant generating function is the steepness of φ, i.e., the property that $d_- = -\infty$ if $\lambda_- > -\infty$ and $d_+ = +\infty$ if $\lambda_+ < +\infty$. The lack of steepness produces affine stretches in the graph of the rate function. In fact, if $\lambda_- > -\infty$ and $d_- > -\infty$, then $I(w) = w\lambda_- - \varphi_-$ for all $w < d_-$. Note that φ_- is finite in this case since $\varphi(\lambda) \le \varphi(0) + \varphi'(\lambda)\lambda = \varphi'(\lambda)\lambda$ for all $\lambda \in (\lambda_-,\lambda_+)$ by convexity, which gives $\varphi_- \le d_-\lambda_-$ by sending λ to λ_-. Similarly, $I(w) = w\lambda_+ - \varphi_+$ for all $w > d_+$ with φ_+ finite if $\lambda_+ < +\infty$ and $d_+ < +\infty$.

Although the initial condition Σ_o and the quadratic boundary perturbations represented by L and R cannot affect the typical value of W_N/N in the large N limit, nor the value of $\varphi(\lambda)$, they enter the rate function I. In fact, according to Lemma 1.2, the matrices Σ_o, L, and R participate in determining the boundary points λ_- and λ_+ of the effective domain of the asymptotic cumulant generating function via $\mathcal{L}_\lambda$ and $\mathcal{R}_\lambda$. The matrices Σ_o, L, and R play a relevant role when the effective domain (λ_-,λ_+) turns out to be smaller than the primary domain of λ's for which $f_\lambda > 0$ and an irrelevant role otherwise. As the function φ is not steep whenever the effective domain is smaller than the primary one, if Σ_o, L, and R entail a reduction in domain size, then the graph of the rate function necessarily exhibits some affine stretch. The following example involving a quadratic functional of a one-dimensional stable Gauss–Markov chain, which allows explicit calculations, demonstrates the role of the initial condition and of the boundary perturbations:

Example 1.1. Fix $a \in \mathbb{R}$ such that $|a| < 1$ and consider the one-dimensional autoregressive model $X_{n+1} = aX_n + G_n$ for $n \ge 1$. The large fluctuations of the quadratic functional $W_N := (1/2)\sum_{n=1}^N X_n^2$ have been already characterized for the non-stationary case $X_1 := 0$[36] and for the *stationary centered case corresponding to* $\Sigma_o = \Sigma_s := (1-a^2)^{-1}$.[28] *We can use our theory to investigate centered non-stationary situations with general initial variance* $\Sigma_o > 0$. *In this example* $S := a$, $L := 0$, $U := 1$, $V := 0$, *and* $R := 0$. *For all* λ *and* θ, *we find*

$$F_\lambda(\theta) = 1 + a^2 - \lambda - 2a\cos\theta$$

so that $f_\lambda = 1 + a^2 - \lambda - 2|a|$. *If* $f_\lambda > 0$, *i.e.,* $\lambda < \lambda_s := (1-|a|)^2$, *then easy calculations based on Lemma 1.2 yield*

$$\varphi(\lambda) = -\frac{1}{2}\ln\frac{1 + a^2 - \lambda + \sqrt{(1+a^2-\lambda)^2 - (2a)^2}}{2},$$

$$\mathcal{L}_\lambda = \Sigma_o^{-1} + \frac{a^2 - 1 - \lambda + \sqrt{(1+a^2-\lambda)^2 - (2a)^2}}{2},$$

and

$$\mathcal{R}_\lambda = \frac{1 - a^2 - \lambda + \sqrt{(1+a^2-\lambda)^2 - (2a)^2}}{2} > 0.$$

The function φ *is steep over the primary domain* $(-\infty, \lambda_s)$. *According to (1.7) and (1.8), when* $\Sigma_o = \Sigma_s$, *we have* $\mathcal{L}_\lambda = \mathcal{R}_\lambda > 0$ *for all* $\lambda < \lambda_s$ *so that* $\lambda_- = -\infty$, $\lambda_+ = \lambda_s$, *and*

$$I(w) = I_s(w) := \begin{cases} +\infty & \text{if } w \le 0, \\ (1+a^2)w - \frac{1}{2}\sqrt{1+(4aw)^2} + \frac{1}{2}\ln\frac{1+\sqrt{1+(4aw)^2}}{2a} & \text{if } w > 0. \end{cases}$$

The rate function I_s *does not exhibit any affine stretch. In general, we find* $\lambda_- = -\infty$ *and* $\lambda_+ = \lambda_s$ *if* $\Sigma_o^{-1} \ge 1 - |a|$ *and* $\lambda_- = -\infty$ *and* $\lambda_+ = (\Sigma_o^{-1} - 1 + a^2)/(1-\Sigma_o) < \lambda_s$ *if* $\Sigma_o^{-1} < 1 - |a|$. *In the former case,* φ *is steep, whereas steepness is missing in the latter case where* $d_+ = \varphi'(\lambda_+) < +\infty$. *While* $I = I_s$, *when* $\Sigma_o^{-1} \ge 1 - |a|$, *in the case* $\Sigma_o^{-1} < 1 - |a|$, *the rate function reads*

$$I(w) = \begin{cases} I_s(w) & \text{if } w < d_+, \\ w\lambda_+ - \varphi_+ & \text{if } w \ge d_+. \end{cases}$$

Suppose now that $\Sigma_o = \Sigma_s$ *and consider the boundary perturbation* bX_N^2 *with* $b \in \mathbb{R}$ *so that* $W_N := (1/2)\sum_{n=1}^N X_n^2 + (b/2)X_N^2$. *In this case* $S := a$, $L := 0$, $U := 1$, $V := 0$, *and* $R := b$. *While* f_λ *and* $\varphi(\lambda)$ *are the same as before, for* $\lambda < \lambda_s$, *we have*

$$\mathcal{L}_\lambda = \frac{1 - a^2 - \lambda + \sqrt{(1+a^2-\lambda)^2 - (2a)^2}}{2} > 0$$

and

$$\mathcal{R}_\lambda = \frac{1 - a^2 - \lambda + \sqrt{(1+a^2-\lambda)^2 - (2a)^2}}{2} - b\lambda.$$

The boundary perturbation plays an irrelevant role if $-1 \le b \le |a|/(1-|a|)$, *to which* $\lambda_- = -\infty$, $\lambda_+ = \lambda_s$, *and* $I = I_s$ *correspond. If instead* $b < -1$ *or* $b > |a|/(1-|a|)$, *then the boundary perturbation comes into play. Specifically,* $b < -1$ *entails* $\lambda_- = 1/(b+1) - a^2/b$ *and* $\lambda_+ = \lambda_s$. *Steepness of the asymptotic cumulant generating function is missed in this case as* $d_- = \varphi'(\lambda_-) > -\infty$ *and*

$$I(w) = \begin{cases} w\lambda_- - \varphi_- & \text{if } w \le d_-, \\ I_s(w) & \text{if } w > d_-. \end{cases}$$

Similarly, $b > |a|/(1-|a|)$ *gives* $\lambda_- = -\infty$, $\lambda_+ = 1/(b+1) - a^2/b < \lambda_s$, $d_+ = \varphi'(\lambda_+) < +\infty$, *and*

$$I(w) = \begin{cases} I_s(w) & \text{if } w < d_+, \\ w\lambda_+ - \varphi_+ & \text{if } w \ge d_+. \end{cases}$$

As $W_N/N = e_N$ for all $N \geq 1$, when $L := I - \Sigma_o^{-1} - S^\top S$, $U := 0$, $V := S - S^\top$, and $R := \Sigma_o^{-1} + S^\top S - I$, Theorem 1.1 immediately shows that the entropy production rate e_N satisfies a large deviation principle. The Hermitian matrix $F_\lambda(\theta)$ corresponding to e_N reads for each $\lambda \in \mathbb{R}$ and $\theta \in [0, 2\pi]$,

$$\begin{aligned} F_\lambda(\theta) &= \left(I - S^\top e^{i\theta}\right)\left(I - Se^{-i\theta}\right) + 2i\lambda\left(S - S^\top\right) \sin\theta \\ &= I + S^\top S - \left(S + S^\top\right)\cos\theta + i(2\lambda + 1)\left(S - S^\top\right)\sin\theta. \end{aligned} \tag{1.9}$$

The second main result of this paper, whose proof is reported in Sec. III, confirms the Gallavotti–Cohen symmetry. This symmetry comes from the manifest relationship $F_{-\lambda-1}(\theta) = F_\lambda(2\pi - \theta)$.

Theorem 1.2. *The following conclusions hold:*

1. *the entropy production rate e_N of the stable Gauss–Markov chain $\{X_n\}_{n\geq 1}$ satisfies a large deviation principle with the convex rate function I;*
2. *$\lambda_- = -\lambda_+ - 1$ and $I(-w) = I(w) + w$ for all $w \in \mathbb{R}$.*

If the drift matrix S is symmetric and $\Sigma_o = \Sigma_s$, then the process $\{X_n\}_{n\geq 1}$ is reversible and $e_N = 0$ for all $N \geq 1$. The following example shows that there is entropy production when S is symmetric but $\{X_n\}_{n\geq 1}$ is not stationary:

Example 1.2. Assume that the drift matrix S is symmetric. We have $\Sigma_s = (I - S^2)^{-1}$, and formula (1.9) gives $F_\lambda(\theta) = (I - Se^{i\theta})(I - Se^{-i\theta})$ for every λ and θ. One can easily verify that $f_\lambda = [1 - \rho(S)]^2 > 0$ and $\varphi(\lambda) = 0$ for all $\lambda \in \mathbb{R}$. Starting from the identity $(I - Se^{\pm i\theta})^{-1} = \sum_{k\geq 0} S^k e^{\pm ik\theta}$ as $\rho(S) < 1$, one can then deduce that for all $\lambda \in \mathbb{R}$,

$$\mathcal{L}_\lambda = \mathcal{R}_{-\lambda-1} = (\lambda + 1)\Sigma_o^{-1} - \lambda\Sigma_s^{-1}.$$

Fix $\Sigma_o > 0$ different from Σ_s and set $\Delta := (\Sigma_s - \Sigma_o)(\Sigma_s + \Sigma_o)^{-1}$. We claim that the spectral radius $\rho(\Delta)$ of Δ is strictly positive and that

$$\lambda_\pm = \frac{1}{2}\left[-1 \pm \frac{1}{\rho(\Delta)}\right]. \tag{1.10}$$

The entropy production rate satisfies a large deviation principle with the rate function

$$I(w) = \begin{cases} w\lambda_- & \text{if } w < 0, \\ w\lambda_+ & \text{if } w \geq 0. \end{cases}$$

To prove (1.10), let $A \in \mathbb{R}^{d\times d}$ be an invertible matrix such that $(1/2)(\Sigma_o^{-1} + \Sigma_s^{-1}) = AA^\top$ and set $B := (1/2)A^{-1}(\Sigma_o^{-1} - \Sigma_s^{-1})(A^\top)^{-1}$. The matrix A exists since $\Sigma_s > 0$ and $\Sigma_o > 0$, and the spectral radius $\rho(B)$ of the symmetric matrix B is strictly positive since $\Sigma_o \neq \Sigma_s$. Similarity transformations show that $\rho(B) = \rho(\Delta)$. We have $\mathcal{L}_\lambda = A[I + (2\lambda + 1)B]A^\top > 0$ and $\mathcal{R}_\lambda = A[I - (2\lambda + 1)B]A^\top > 0$ if and only if $|2\lambda + 1|\rho(B) < 1$. Thus, $(2\lambda_\pm + 1)\rho(B) = \pm 1$.

Theorem 1.2 characterizes the entropy production rate of general stable, Gaussian autoregressive sequences. It might also be the basis for complementing the literature on multivariate Ornstein–Uhlenbeck processes[19–21] with a large deviation principle for the entropy production rate under non-normal drift matrices. According to Ref. 8, the entropy production rate of an Ornstein–Uhlenbeck process can be expressed in terms of a stochastic integral of the process, thanks to a continuous-time analogous of Lemma 1.1. As the Ornstein–Uhlenbeck process observed at equispaced times is a Gaussian discrete-time autoregressive model, after discretization of the stochastic integral, the entropy production rate of an Ornstein–Uhlenbeck process becomes the entropy production rate of a Gaussian autoregressive sequence. If the latter is an exponentially good approximation[15] of the former, then a large deviation principle for the entropy production rate of an Ornstein–Uhlenbeck process with non-normal drift matrix can be deduced by our general theory through a limit procedure. Work is in progress along this line.

B. Entropy production with a normal drift matrix

Analyzing the role of the conditions $\mathcal{L}_\lambda > 0$ and $\mathcal{R}_\lambda > 0$ in determining those $\lambda \in \mathbb{R}$ for which $\lim_{N\uparrow\infty}(1/N)\ln\mathbb{E}[e^{\lambda W_N}] = \varphi(\lambda)$ is a difficult task. We stress that the satisfiability of these conditions shapes the effective domain (λ_-, λ_+) of the asymptotic cumulant generating function of W_N. Now, our interest is in the entropy production $W_N := Ne_N$. Computer simulations suggest that in the stationary case $\Sigma_o = \Sigma_s$, the Hermitian matrices $\mathcal{L}_\lambda$ and $\mathcal{R}_\lambda$ associated with Ne_N are automatically positive-definite for the values of λ that satisfy the primary constraint $f_\lambda > 0$. If this is true in general, then we will conclude that $\lambda_- = \inf\{\lambda \in \mathbb{R} : f_\lambda > 0\}$ and $\lambda_+ = \sup\{\lambda \in \mathbb{R} : f_\lambda > 0\}$ when $\Sigma_o = \Sigma_s$. While we leave this general problem as an open question, we verify the conjecture $\lambda_- = \inf\{\lambda \in \mathbb{R} : f_\lambda > 0\}$ and $\lambda_+ = \sup\{\lambda \in \mathbb{R} : f_\lambda > 0\}$ for a stationary stable Gauss–Markov chain $\{X_n\}_{n\ge 1}$ with normal drift matrix S. Then, here we assume that $S^\top S = SS^\top$. This case is very special because it allows for explicit results. We recall that large deviation principles have been recently established for the entropy production rate of stationary stable Ornstein–Uhlenbeck processes with the normal drift matrix.[19–21] In particular, Budhiraja *et al.*[21] exhibited the rate function explicitly, posing the question of whether the same could have been done for the discrete-time autoregressive model. Our work gives an affirmative answer to their question, and indeed, we provide a large deviation principle for any drift matrix.

Dealing with a normal drift matrix in the problem of entropy production basically means dealing with a diagonal drift matrix. In fact, normality of S implies that there exists a unitary matrix $\Gamma \in \mathbb{C}^{d\times d}$ such that $\Gamma S\Gamma^{-1}$ and $\Gamma S^\top \Gamma^{-1} = (\Gamma S\Gamma^{-1})^\dagger$ are both diagonal. Let $\alpha_k + \mathrm{i}\beta_k$ be the kth element of the diagonal of $\Gamma S\Gamma^{-1}$, with α_k and β_k real numbers, and notice that the stability hypothesis $\rho(S) < 1$ requires that $\alpha_k^2 + \beta_k^2 < 1$ as $\alpha_k + \mathrm{i}\beta_k$ obviously is an eigenvalue of S. We suppose that $\beta_k \neq 0$ for some k in order to not to fall again in the class of symmetric drift matrices. According to (1.9), $\Gamma F_\lambda(\theta)\Gamma^{-1}$ is diagonal for all $\lambda \in \mathbb{R}$ and $\theta \in [0, 2\pi]$, and the kth element of the diagonal of $\Gamma F_\lambda(\theta)\Gamma^{-1}$ reads

$$1 + \alpha_k^2 + \beta_k^2 - 2\alpha_k \cos\theta - 2\beta_k(2\lambda + 1)\sin\theta = \left(1 + \alpha_k^2 + \beta_k^2\right)[1 - \varrho_k \cos(\theta - \vartheta_k)],$$

with

$$\varrho_k := 2\frac{\sqrt{\alpha_k^2 + (2\lambda+1)^2\beta_k^2}}{1 + \alpha_k^2 + \beta_k^2} \ge 0$$

and

$$\vartheta_k := \arctan\left(\frac{\beta_k + 2\lambda\beta_k}{\alpha_k}\right).$$

We omit to indicate the dependence of ϱ_k and ϑ_k on λ for simplicity. We have

$$f_\lambda = \inf_{\theta\in[0,2\pi]} \min_{1\le k\le d}\left\{\left(1 + \alpha_k^2 + \beta_k^2\right)[1 - \varrho_k \cos(\theta - \vartheta_k)]\right\} = \min_{1\le k\le d}\left\{\left(1 + \alpha_k^2 + \beta_k^2\right)(1 - \varrho_k)\right\}$$

so that the condition $f_\lambda > 0$ on λ becomes $\max_{1\le k\le d}\{\varrho_k\} < 1$. If $\max_{1\le k\le d}\{\varrho_k\} < 1$, then we find from (1.5) that

$$\begin{aligned}
\varphi(\lambda) &= -\frac{1}{4\pi}\sum_{k=1}^{d}\int_0^{2\pi} \ln\left\{\left(1 + \alpha_k^2 + \beta_k^2\right)[1 - \varrho_k \cos(\theta - \vartheta_k)]\right\} d\theta \\
&= -\frac{1}{4\pi}\sum_{k=1}^{d}\int_0^{2\pi} \ln\left(1 - \varrho_k \cos\theta\right) d\theta - \frac{1}{2}\sum_{k=1}^{d} \ln\left(1 + \alpha_k^2 + \beta_k^2\right) \\
&= -\frac{1}{2}\sum_{k=1}^{d} \ln\frac{1 + \sqrt{1 - \varrho_k^2}}{2} - \frac{1}{2}\sum_{k=1}^{d} \ln\left(1 + \alpha_k^2 + \beta_k^2\right).
\end{aligned}$$

For each $n \in \mathbb{Z}$, the matrix $\Gamma\Phi_\lambda(n)\Gamma^{-1}$ defined by (1.6) is diagonal with the kth diagonal element equal to

$$\begin{aligned}
\frac{1}{2\pi}\int_0^{2\pi} \frac{e^{-\mathrm{i}n\theta}d\theta}{(1 + \alpha_k^2 + \beta_k^2)[1 - \varrho_k \cos(\theta - \vartheta_k)]} &= \frac{e^{-\mathrm{i}n\vartheta_k}}{1 + \alpha_k^2 + \beta_k^2}\frac{1}{2\pi}\int_0^{2\pi} \frac{\cos(n\theta)d\theta}{1 - \varrho_k \cos\theta} \\
&= \frac{1}{1 + \alpha_k^2 + \beta_k^2}\frac{e^{-\mathrm{i}n\vartheta_k}}{\sqrt{1 - \varrho_k^2}}\left(\frac{1 - \sqrt{1 - \varrho_k^2}}{\varrho_k}\right)^{|n|}.
\end{aligned}$$

Under the constraint $\max_{1\le k\le d}\{\varrho_k\} < 1$, the matrices $\mathcal{L}_\lambda$ and $\mathcal{R}_\lambda$ associated by Lemma 1.2 with $L := I - \Sigma_o^{-1} - S^\top S$, $U := 0$, $V := S - S^\top$, and $R := \Sigma_o^{-1} + S^\top S - I$ can be written as

$$\mathcal{L}_\lambda = (\lambda + 1)(\Sigma_o^{-1} - \Sigma_s^{-1}) + \mathcal{M}_\lambda \tag{1.11}$$

and

$$\mathcal{R}_\lambda = \lambda(\Sigma_s^{-1} - \Sigma_o^{-1}) + \mathcal{M}_\lambda, \tag{1.12}$$

where $\Gamma\mathcal{M}_\lambda\Gamma^{-1} \in \mathbb{C}^{d\times d}$ is diagonal with the kth diagonal element given by

$$\frac{1-\alpha_k^2-\beta_k^2}{2} + \frac{1+\alpha_k^2+\beta_k^2}{2}\sqrt{1-\varrho_k^2} > 0. \tag{1.13}$$

To obtain (1.11) and (1.12), we have used the facts that $\Sigma_s = (I - SS^\top)^{-1}$ and that $\Gamma\Sigma_s^{-1}\Gamma^{-1}$ is diagonal with the kth diagonal entry equal to $1-\alpha_k^2-\beta_k^2$. Importantly, the Hermitian matrix $\mathcal{M}_\lambda$ is positive-definite as demonstrated by (1.13).

If the chain $\{X_n\}_{n\geq1}$ is stationary, i.e., if $\Sigma_o = \Sigma_s$, then $\mathcal{L}_\lambda = \mathcal{M}_\lambda$ and $\mathcal{R}_\lambda = \mathcal{M}_\lambda$ are automatically positive-definite when $\max_{1\leq k\leq d}\{\varrho_k\} < 1$, namely, when $f_\lambda > 0$. Thus, the conjecture $\lambda_- = \inf\{\lambda\in\mathbb{R} : f_\lambda > 0\}$ and $\lambda_+ = \sup\{\lambda\in\mathbb{R} : f_\lambda > 0\}$ for a stationary stable Gauss–Markov chain is true if the drift matrix is normal. Furthermore, in this case, λ_- and λ_+ are the smallest and the largest values of λ for which $\max_{1\leq k\leq d}\{\varrho_k\} = 1$, which are explicitly given by the formulas

$$\lambda_+ = \lambda_o := -\frac{1}{2} + \min_{1\leq k\leq d}\left\{\sqrt{\frac{(1+\alpha_k^2+\beta_k^2)^2 - 4\alpha_k^2}{16\beta_k^2}}\right\}$$

and

$$\lambda_- = -\lambda_o - 1.$$

Note that λ_o is finite since we are supposing that $\beta_k \neq 0$ for some k. With such λ_- and λ_+, the function φ turns out to be steep in (λ_-,λ_+). Thus, for each $w\in\mathbb{R}$, there exists a unique $\lambda\in(\lambda_-,\lambda_+)$ such that $w = \varphi'(\lambda)$ and, as a consequence,

$$I(w) = w\lambda - \varphi(\lambda).$$

Basically, this is the result found by Budhiraja *et al.*[21] for the continuous-time model.

To conclude, let us briefly discuss what happens when the chain $\{X_n\}_{n\geq1}$ is not stationary, i.e., when $\Sigma_o \neq \Sigma_s$. If $\mathcal{L}_\lambda \succ 0$ and $\mathcal{R}_\lambda \succ 0$ for all $\lambda\in(-\lambda_o-1,\lambda_o)$, then $\lambda_+ = \lambda_o$ and $\lambda_- = -\lambda_o - 1$, as before, and the function φ is steep in (λ_-,λ_+). We have $\mathcal{L}_\lambda \succ 0$ and $\mathcal{R}_\lambda \succ 0$ for all $\lambda\in(-\lambda_o-1,\lambda_o)$ if $\mathcal{L}_{\lambda_o} = \mathcal{R}_{-\lambda_o-1} \succ 0$ and $\mathcal{R}_{\lambda_o} = \mathcal{L}_{-\lambda_o-1} \succ 0$ as formulas (1.11) and (1.12) show that the functions that map λ in $\langle z, \mathcal{L}_\lambda z\rangle$ and $\langle z, \mathcal{R}_\lambda z\rangle$ are concave for any given $z\in\mathbb{C}^d$. If, on the contrary, there exists $\lambda\in(-\lambda_o-1,\lambda_o)$ such that $\mathcal{L}_\lambda \not\succ 0$ or $\mathcal{R}_\lambda \not\succ 0$, then $\lambda_+ < \lambda_o$ and $\lambda_- = -\lambda_+ - 1 > -\lambda_o - 1$. For example, this occurs for $\Sigma_o = \sigma I$ with a sufficiently small $\sigma > 0$. In this case, φ is not steep in (λ_-,λ_+), and the rate function at $w\in\mathbb{R}$ has the value

$$I(w) = \begin{cases} w\lambda_- - \varphi_- & \text{if } w \leq d_-, \\ w\lambda - \varphi(\lambda) & \text{if } d_- < w < d_+, \\ w\lambda_+ - \varphi_+ & \text{if } w \geq d_+, \end{cases}$$

where, regarding the case $d_- < w < d_+$, λ is the unique real number in (λ_-,λ_+) that satisfies $w = \varphi'(\lambda)$. Breaking stationarity can then involve affine stretches in the graph of the entropy production rate function.

II. PROOF OF LEMMA 1.2 AND THEOREM 1.1

In this section, we prove Theorem 1.1, which states the large deviation principle for the quadratic functional W_N defined by (1.2). The Proof of Theorem 1.1 is based on a time-dependent change of measure and requires at first to study the asymptotics of the cumulant generating function of W_N as N goes to infinity. In turn, this asks for investigation of Hermitian block tridiagonal quasi-Toeplitz matrices that differ from Hermitian block tridiagonal Toeplitz matrices by the first and last diagonal blocks. In Sec. II A, we introduce these matrices and characterize their positive definiteness property and their determinant. Section II B uses the theory of Sec. II A to compute the scaled cumulant generating function of W_N in the large N limit. The upper large deviation bound for closed sets is proved in Sec. II C. Finally, the lower large deviation bound for open sets is established in Sec. II D. Along the way, we shall also verify Lemma 1.2.

As we have already said, we regard $\mathbb{R}^d$ as a subset of $\mathbb{C}^d$. We denote by $\langle\cdot,\cdot\rangle$ the standard inner product of $\mathbb{C}^d$ and by $\|\cdot\|$ the induced norm. If $\zeta = (\zeta_1,\ldots,\zeta_N)$ and $z = (z_1,\ldots,z_N)$ are two vectors in $(\mathbb{C}^d)^N$, with N being a positive integer, we understand that $\langle\zeta,z\rangle := \sum_{n=1}^N\langle\zeta_n,z_n\rangle$ and $\|z\|^2 := \sum_{n=1}^N\langle z_n,z_n\rangle = \sum_{n=1}^N\|z_n\|^2$. For positive integers M and N, $\mathsf{BL}_{M,N}$ is the set of complex block matrices with $M\times N$ square blocks of size d. For any $\mathsf{A}\in\mathsf{BL}_{N,N}$, $\|\mathsf{A}\|$ is the operator norm of A induced by the norm of $(\mathbb{C}^d)^N$,

$$\|\mathsf{A}\| := \sup_{\substack{z\in(\mathbb{C}^d)^N \\ z\neq0}}\left\{\frac{\|\mathsf{A}z\|}{\|z\|}\right\}.$$

Given a Hermitian matrix $\mathsf{A} \in \mathsf{BL}_{N,N}$, we denote by $r(\mathsf{A})$ the infimum of the Rayleigh quotient of A, that is, the smallest eigenvalue of A,

$$r(\mathsf{A}) := \inf_{\substack{z \in (\mathbb{C}^d)^N \\ z \neq 0}} \left\{ \frac{\langle z, \mathsf{A}z \rangle}{\langle z, z \rangle} \right\}.$$

If a Hermitian matrix $\mathsf{A} \in \mathsf{BL}_{N,N}$ is positive-definite, we write $\mathsf{A} > 0$. We have $\mathsf{A} > 0$ if and only if $r(\mathsf{A}) > 0$.

A. On Hermitian block tridiagonal quasi-Toeplitz matrices

The coefficient matrix of the quadratic functional W_N is an element from a sequence of Hermitian block tridiagonal matrices in the following class:

Definition 2.1. A sequence of matrices $\{\mathsf{Q}_N\}_{N\geq 1}$, with $\mathsf{Q}_N \in \mathsf{BL}_{N+2,N+2}$ for each N, is a Hermitian block tridiagonal quasi-Toeplitz (HQT) matrix sequence if there exist four square matrices A, D, B, and E of size d, with A, D, and B Hermitian, such that for all $N \geq 1$,

$$\mathsf{Q}_N = \begin{pmatrix} A & E^\dagger & & & \\ E & D & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & D & E^\dagger \\ & & & E & B \end{pmatrix}.$$

In this section, we characterize asymptotic positive definiteness and asymptotic determinants of matrices from a HQT matrix sequence, postponing the most technical proofs in the Appendixes C–G. We stress that a HQT matrix sequence is bounded in the following sense, which is demonstrated in Appendix C:

Lemma 2.1. Let $\{\mathsf{Q}_N\}_{N\geq 1}$ be a HQT matrix sequence with A, D, B, and E as in the above definition. Then, for each $N \geq 1$,

$$\|\mathsf{Q}_N\| \leq \sqrt{2\|A\|^2 + 3\|D\|^2 + 2\|B\|^2 + 6\|E\|^2}.$$

In order to deal with a HQT matrix sequence $\{\mathsf{Q}_N\}_{N\geq 1}$, it is convenient to isolate the *bulk matrix* $\mathsf{T}_N \in \mathsf{BL}_{N,N}$ of Q_N defined by

$$\mathsf{T}_N := \begin{pmatrix} D & E^\dagger & & \\ E & \ddots & \ddots & \\ & \ddots & \ddots & E^\dagger \\ & & E & D \end{pmatrix}. \tag{2.1}$$

The bulk matrix T_N is a Hermitian block tridiagonal Toeplitz matrix, which allows Q_N to be written as

$$\mathsf{Q}_N = \begin{pmatrix} A & E^\dagger \mathsf{C}^\dagger & 0 \\ \mathsf{C}E & \mathsf{T}_N & \mathsf{R}^\dagger E^\dagger \\ 0 & E\mathsf{R} & B \end{pmatrix}, \tag{2.2}$$

where

$$\mathsf{C} := \begin{pmatrix} I \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathsf{BL}_{N,1} \tag{2.3}$$

and

$$\mathsf{R} := \begin{pmatrix} 0 & \cdots & 0 & I \end{pmatrix} \in \mathsf{BL}_{1,N}. \tag{2.4}$$

When T_N is invertible, we introduce the *boundary matrix* $\mathsf{S}_N \in \mathsf{BL}_{2,2}$ of Q_N defined by

$$\mathsf{S}_N := \begin{pmatrix} A - E^\dagger \mathsf{C}^\dagger \mathsf{T}_N^{-1} \mathsf{C} E & -E^\dagger \mathsf{C}^\dagger \mathsf{T}_N^{-1} \mathsf{R}^\dagger E^\dagger \\ -E \mathsf{R} \mathsf{T}_N^{-1} \mathsf{C} E & B - E \mathsf{R} \mathsf{T}_N^{-1} \mathsf{R}^\dagger E^\dagger \end{pmatrix}. \tag{2.5}$$

Manifestly, S_N is a Hermitian matrix. The following lemma relates the positive definiteness and the determinant of Q_N to those of the bulk matrix T_N and the boundary matrix S_N. The proof is reported in Appendix D.

Lemma 2.2. Let $\{\mathsf{Q}_N\}_{N\geq 1}$ be a HQT matrix sequence with bulk matrices T_N and boundary matrices S_N. The following conclusions hold for any $N \geq 1$:

1. *if $r(\mathsf{Q}_N) \geq q$ for some real number $q > 0$, then $r(\mathsf{T}_N) \geq q$ (which implies that T_N is invertible) and $r(\mathsf{S}_N) \geq q$;*
2. *if $\mathsf{T}_N > 0$ (which implies that T_N is invertible) and $\mathsf{S}_N > 0$, then $\mathsf{Q}_N > 0$ and*

$$\ln \det \mathsf{Q}_N = \ln \det \mathsf{T}_N + \ln \det \mathsf{S}_N.$$

We now examine the bulk matrices. For each $\theta \in [0, 2\pi]$, let $F(\theta) \in \mathbb{C}^{d\times d}$ be a Hermitian matrix defined by

$$F(\theta) := E e^{-\mathrm{i}\theta} + D + E^\dagger e^{\mathrm{i}\theta},$$

with D and E being the matrices that identify the bulk matrix T_N of Q_N. In the theory of block Toeplitz matrices,[38] the function F that maps θ in $F(\theta)$ is called the *symbol* of the matrices T_N. We shall equally call F the symbol of T_N or the symbol of Q_N. The blocks of T_N are related to the Fourier coefficients of the symbol F. In fact, for all $N \geq 1$, $\zeta = (\zeta_1, \ldots, \zeta_N) \in (\mathbb{C}^d)^N$, and $z = (z_1, \ldots, z_N) \in (\mathbb{C}^d)^N$, we have

$$\langle \zeta, \mathsf{T}_N z \rangle = \sum_{m=1}^N \sum_{n=1}^N \left\langle \zeta_m, \frac{1}{2\pi} \int_0^{2\pi} F(\theta) e^{\mathrm{i}(m-n)\theta} d\theta \, z_n \right\rangle = \frac{1}{2\pi} \int_0^{2\pi} \left\langle \sum_{n=1}^N \zeta_n e^{-\mathrm{i}n\theta}, F(\theta) \sum_{n=1}^N z_n e^{-\mathrm{i}n\theta} \right\rangle d\theta. \tag{2.6}$$

The following lemma describes the positive definiteness and the determinant of the bulk matrices T_N. The proof is provided in Appendix E. We stress that if $\inf_{\theta\in[0,2\pi]}\{r(F(\theta))\} > 0$, then the function that associates $\theta \in [0, 2\pi]$ with $\ln \det F(\theta)$ is well-defined and continuous.

Lemma 2.3. Let T_N be the bulk matrices of a HQT matrix sequence with symbol F. The following conclusions hold:

1. *if there exists a diverging sequence $\{N_k\}_{k\geq 0}$ of positive integers such that $r(\mathsf{T}_{N_k}) \geq t$ for all $k \geq 0$ with some $t \in \mathbb{R}$, then $r(\mathsf{T}_N) \geq t$ for all $N \geq 1$;*
2. *$r(\mathsf{T}_N) \geq t$ for all $N \geq 1$ with some $t \in \mathbb{R}$ if and only if $\inf_{\theta\in[0,2\pi]}\{r(F(\theta))\} \geq t$;*
3. *if $\inf_{\theta\in[0,2\pi]}\{r(F(\theta))\} > 0$, then*

$$\lim_{N\uparrow\infty} \frac{1}{N} \ln \det \mathsf{T}_N = \frac{1}{2\pi} \int_0^{2\pi} \ln \det F(\theta)\, d\theta.$$

The analysis of the boundary matrices S_N is based on the possibility to determine a limit boundary matrix when N is sent to infinity. This is done by the following lemma, which is proved in Appendix F. Let A, D, B, and E be as in Definition 2.1. Set for each $n \in \mathbb{Z}$,

$$\Phi(n) := \frac{1}{2\pi} \int_0^{2\pi} F^{-1}(\theta) e^{-\mathrm{i}n\theta} d\theta,$$

which is a well-defined matrix under the hypothesis $\inf_{\theta\in[0,2\pi]}\{r(F(\theta))\} > 0$.

Lemma 2.4. Let S_N *be the boundary matrices of a HQT matrix sequence with symbol F. Assume that* $\inf_{\theta\in[0,2\pi]}\{r(F(\theta))\} > 0$. *The following conclusions hold:*

1. $H := I - E\Phi(1) \in \mathbb{C}^{d\times d}$ *and* $K := I - \Phi(1)E \in \mathbb{C}^{d\times d}$ *are invertible;*
2. *the limit* $\lim_{N\uparrow\infty} \mathsf{S}_N =: \mathsf{S}_\infty$ *exists and* $\mathsf{S}_\infty = \begin{pmatrix} \mathcal{L} & 0 \\ 0 & \mathcal{R} \end{pmatrix}$ *with Hermitian matrices* $\mathcal{L}$ *and* $\mathcal{R}$ *defined, respectively, by*
$$\mathcal{L} := A - E^\dagger \Phi(0) H^{-1} E$$
and
$$\mathcal{R} := B - EK^{-1}\Phi(0)E^\dagger.$$

We call S_∞ the *limit boundary matrix* of the HQT matrix sequence $\{\mathsf{Q}_N\}_{N\geq 1}$. Putting the pieces together in the following proposition, we finally solve the positive definiteness and the determinants of the matrices Q_N in the large N limit.

Proposition 2.1. Let $\{\mathsf{Q}_N\}_{N\geq1}$ *be a HQT matrix sequence with symbol F and limit boundary matrix* $\mathsf{S}_\infty = \begin{pmatrix} \mathcal{L} & 0 \\ 0 & \mathcal{R} \end{pmatrix}$. *Assume that* $\inf_{\theta\in[0,2\pi]}\{r(F(\theta))\} > 0$, $\mathcal{L} > 0$, *and* $\mathcal{R} > 0$. *Then,* $\mathsf{Q}_N > 0$ *for all sufficiently large N and*

$$\lim_{N\uparrow\infty} \frac{1}{N} \ln \det \mathsf{Q}_N = \frac{1}{2\pi}\int_0^{2\pi} \ln \det F(\theta)\, d\theta.$$

Proof. According to Lemma 2.3, the hypothesis $\inf_{\theta\in[0,2\pi]}\{r(F(\theta))\} > 0$ gives $\mathsf{T}_N > 0$ for all $N \geq 1$ and

$$\lim_{N\uparrow\infty} \frac{1}{N} \ln \det \mathsf{T}_N = \frac{1}{2\pi}\int_0^{2\pi} \ln \det F(\theta)\, d\theta.$$

It also shows that $\lim_{N\uparrow\infty} \mathsf{S}_N = \mathsf{S}_\infty$ exists by Lemma 2.4. Since $\mathsf{S}_\infty > 0$ by hypothesis, the boundary matrices S_N are positive definite for all sufficiently large N. It follows by Lemma 2.2 that $\mathsf{Q}_N > 0$ for all sufficiently large N and that

$$\lim_{N\uparrow\infty} \frac{1}{N} \ln \det \mathsf{Q}_N = \lim_{N\uparrow\infty} \frac{1}{N} \ln \det \mathsf{T}_N + \lim_{N\uparrow\infty} \frac{1}{N} \ln \det \mathsf{S}_N = \frac{1}{2\pi}\int_0^{2\pi} \ln \det F(\theta) d\theta.$$

□

B. The cumulant generating function of W_N

Let us move to the stable Gauss–Markov chain $\{X_n\}_{n\geq1}$ and the quadratic functional W_N. According to (1.1), for each $N \geq 1$, the law of $(X_1,\ldots,X_{N+2})$ is the multivariate Gaussian distribution that at $x = (x_1,\ldots,x_{N+2}) \in (\mathbb{R}^d)^{N+2}$ has probability density,

$$\frac{e^{-\frac{1}{2}\langle x, \Sigma_N^{-1} x\rangle}}{\sqrt{(2\pi)^{(N+2)d} \det \Sigma_N}} = \frac{e^{-\frac{1}{2}\langle x_1, \Sigma_o^{-1} x_1\rangle}}{\sqrt{(2\pi)^d \det \Sigma_o}} \prod_{n=2}^{N+2} \frac{1}{\sqrt{(2\pi)^d}}\, e^{-\frac{1}{2}\|x_n - Sx_{n-1}\|^2}.$$

We see that the inverse Σ_N^{-1} of the covariance matrix Σ_N of $(X_1,\ldots,X_{N+2})$ is the real symmetric block tridiagonal matrix in $\mathsf{BL}_{N+2,N+2}$ given by

$$\Sigma_N^{-1} := \begin{pmatrix} \Sigma_o^{-1} + S^\top S & -S^\top & & & \\ -S & I + S^\top S & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & I + S^\top S & -S^\top \\ & & & -S & I \end{pmatrix},$$

and we have $\det \Sigma_N = \det \Sigma_o$. Together with Σ_N^{-1}, we introduce the real symmetric block tridiagonal matrix $\mathsf{M}_N \in \mathsf{BL}_{N+2,N+2}$ defined by

$$\mathsf{M}_N := \begin{pmatrix} U + L & V^\top & & & \\ V & U & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & U & V^\top \\ & & & V & U + R \end{pmatrix}.$$

The matrix M_N allows us to express the quadratic functional W_{N+2} as $(1/2)\langle X, \mathsf{M}_N X\rangle$ with $X := (X_1, \ldots, X_{N+2})$. The *cumulant generating function* of W_N is the function that maps $\lambda \in \mathbb{R}$ in $(1/N)\ln \mathbb{E}[e^{\lambda W_N}]$. We start with the following elementary result involving Gaussian integrals:

Lemma 2.5. For each $N \geq 1$ and $\lambda \in \mathbb{R}$,

$$\ln \mathbb{E}\Big[e^{\lambda W_{N+2}}\Big] = \begin{cases} -\frac{1}{2}\ln\det\Sigma_o - \frac{1}{2}\ln\det\left(\Sigma_N^{-1} - \lambda\mathsf{M}_N\right) & \text{if } \Sigma_N^{-1} - \lambda\mathsf{M}_N > 0, \\ +\infty & \text{otherwise.}\end{cases}$$

We aim to investigate the asymptotics of the cumulant generating function. According to Definition 2.1, the matrices $\mathsf{Q}_N := \Sigma_N^{-1} - \lambda\mathsf{M}_N \in \mathsf{BL}_{N+2,N+2}$ with some $\lambda \in \mathbb{R}$ form a HQT matrix sequence. Explicitly, we have $A := \Sigma_o^{-1} + S^\top S - \lambda(U+L)$, $D := I + S^\top S - \lambda U$, $B := I - \lambda(U+R)$, and $E := -S - \lambda V$. The symbol F_λ of this HQT matrix sequence reads

$$\begin{aligned} F_\lambda(\theta) &:= -(S+\lambda V)e^{-\mathrm{i}\theta} + I + S^\top S - \lambda U - \left(S^\top + \lambda V^\top\right)e^{\mathrm{i}\theta} \\ &= \left(I - S^\top e^{\mathrm{i}\theta}\right)\left(I - S e^{-\mathrm{i}\theta}\right) - \lambda\left(U + V e^{-\mathrm{i}\theta} + V^\top e^{\mathrm{i}\theta}\right)\end{aligned}$$

for every $\theta \in [0, 2\pi]$. It is exactly the matrix (1.3). According to (1.4), the real number f_λ is related to Rayleigh quotients of the symbol F_λ by $\inf_{\theta\in[0,2\pi]}\{r(F_\lambda(\theta))\} = f_\lambda$. Lemma 2.4 proves the technical Lemma 1.2, and the matrices $\mathcal{L}_\lambda$ and $\mathcal{R}_\lambda$ defined by Lemma 1.2 enter the limit boundary matrix of $\{\mathsf{Q}_N\}_{N\geq 1}$: $\mathsf{S}_\infty = \begin{pmatrix}\mathcal{L}_\lambda & 0 \\ 0 & \mathcal{R}_\lambda\end{pmatrix}$. By combining Lemma 2.5 with Proposition 2.1, we get that if $f_\lambda > 0$, $\mathcal{L}_\lambda > 0$, and $\mathcal{R}_\lambda > 0$, then

$$\lim_{N\uparrow\infty}\frac{1}{N}\ln\mathbb{E}\Big[e^{\lambda W_N}\Big] = -\frac{1}{4\pi}\int_0^{2\pi}\ln\det F_\lambda(\theta)\, d\theta =: \varphi(\lambda), \tag{2.7}$$

with $\varphi(\lambda)$ being the integral already defined in (1.5). We want to prove here that the set

$$\Lambda := \{\lambda \in \mathbb{R} \;:\; f_\lambda > 0,\ \mathcal{L}_\lambda > 0,\ \text{and}\, \mathcal{R}_\lambda > 0\} \tag{2.8}$$

is an interval. Formulas (1.7) and (1.8) states that $\lambda_- = \inf\{\Lambda\}$ and $\lambda_+ = \sup\{\Lambda\}$. To begin with, we need the following bound for $r(\Sigma_N^{-1})$, which is based on the hypothesis that the spectral radius $\rho(S)$ of S is smaller than 1 and is proved in Appendix G:

Lemma 2.6. There exists a real number $\sigma > 0$ such that $r(\Sigma_N^{-1}) \geq \sigma$ for all $N \geq 1$.

The following lemma shows that Λ is a convex set, and hence, it is an interval:

Lemma 2.7. The following limits exist and are finite:

$$\lim_{N\uparrow\infty}\inf_{\substack{z\in(\mathbb{C}^d)^{N+2}\\ z\neq 0}}\left\{\frac{\langle z, \mathsf{M}_N z\rangle}{\langle z, \Sigma_N^{-1} z\rangle}\right\} =: \xi_-$$

and

$$\lim_{N\uparrow\infty}\sup_{\substack{z\in(\mathbb{C}^d)^{N+2}\\ z\neq 0}}\left\{\frac{\langle z,\mathsf{M}_N z\rangle}{\langle z,\Sigma_N^{-1}z\rangle}\right\}=:\xi_+.$$

If $\lambda\in\Lambda$, *then* $\lambda\xi_-\leq 1$ *and* $\lambda\xi_+\leq 1$. *If* $\lambda\in\mathbb{R}$ *is such that* $\lambda\xi_-<1$ *and* $\lambda\xi_+<1$, *then* $\lambda\in\Lambda$.

Proof. Fix a real number λ_o and set

$$\liminf_{N\uparrow\infty}\sup_{\substack{z\in(\mathbb{C}^d)^{N+2}\\ z\neq 0}}\left\{\lambda_o\frac{\langle z,\mathsf{M}_N z\rangle}{\langle z,\Sigma_N^{-1}z\rangle}\right\}=:\xi_o.$$

The limit ξ_o is finite since $\|\mathsf{M}_N\|\leq C$ for every $N\geq 1$ with some constant $C<+\infty$ by Lemma 2.1 and $\langle z,\Sigma_N^{-1}z\rangle\geq\sigma\langle z,z\rangle$ and for all $N\geq 1$ and $z\in(\mathbb{C}^d)^{N+2}$ by Lemma 2.6. Let us show that

$$\limsup_{N\uparrow\infty}\sup_{\substack{z\in(\mathbb{C}^d)^{N+2}\\ z\neq 0}}\left\{\lambda_o\frac{\langle z,\mathsf{M}_N z\rangle}{\langle z,\Sigma_N^{-1}z\rangle}\right\}\leq\xi_o. \tag{2.9}$$

By choosing $\lambda_o=-1$ and $\lambda_o=1$, this proves that the following limits exist and are finite:

$$\lim_{N\uparrow\infty}\inf_{\substack{z\in(\mathbb{C}^d)^{N+2}\\ z\neq 0}}\left\{\frac{\langle z,\mathsf{M}_N z\rangle}{\langle z,\Sigma_N^{-1}z\rangle}\right\}=:\xi_-$$

and

$$\lim_{N\uparrow\infty}\sup_{\substack{z\in(\mathbb{C}^d)^{N+2}\\ z\neq 0}}\left\{\frac{\langle z,\mathsf{M}_N z\rangle}{\langle z,\Sigma_N^{-1}z\rangle}\right\}=:\xi_+.$$

Pick an arbitrary real number $\xi>\xi_o$ and $\epsilon>0$ such that $\xi_o+2\epsilon\leq\xi$. Consider the HQT matrix sequence $\{\mathsf{Q}_N\}_{N\geq 1}$ with $\mathsf{Q}_N:=\xi\Sigma_N^{-1}-\lambda_o\mathsf{M}_N\in\mathsf{BL}_{N+2,N+2}$ for all $N\geq 1$. By definition of ξ_o, there exists a diverging sequence $\{N_k\}_{k\geq 0}$ of positive integers with the property that for all k and $z\in(\mathbb{C}^d)^{N_k+2}$,

$$\lambda_o\langle z,\mathsf{M}_{N_k}z\rangle\leq(\xi-\epsilon)\langle z,\Sigma_{N_k}^{-1}z\rangle.$$

It follows that $r(\mathsf{Q}_{N_k})\geq\epsilon r(\Sigma_{N_k}^{-1})\geq\epsilon\sigma>0$ for any $k\geq 0$. Then, part 1 of Lemma 2.2 tells us that $r(\mathsf{T}_{N_k})\geq\epsilon\sigma$ and $r(\mathsf{S}_{N_k})\geq\epsilon\sigma$ for all k, with T_N being the bulk matrix of Q_N and S_N being its boundary matrix. As a consequence, parts 1 and 2 of Lemma 2.3 give $r(\mathsf{T}_N)\geq\epsilon\sigma$ for every $N\geq 1$ and $r(F(\theta))\geq\epsilon\sigma$ for all $\theta\in[0,2\pi]$, with F being the symbol of the Hermitian block Toeplitz matrices T_N. This way, Lemma 2.4 shows that $\lim_{N\uparrow\infty}\mathsf{S}_N=\mathsf{S}_\infty$ exists and is well-defined. Since $r(\mathsf{S}_{N_k})\geq\epsilon\sigma$ for all k, we have $r(\mathsf{S}_\infty)\geq\epsilon\sigma$ so that $r(\mathsf{S}_N)>0$ for all sufficiently large N. In conclusion, we find that both $\mathsf{T}_N>\mathsf{O}$ and $\mathsf{S}_N>\mathsf{O}$ for all sufficiently large N, and part 2 of Lemma 2.2 ensures us that $\mathsf{Q}_N>\mathsf{O}$ for all such N. This means that

$$\langle z,\mathsf{Q}_N z\rangle=\langle z,\left(\xi\Sigma_N^{-1}-\lambda_o\mathsf{M}_N\right)z\rangle>0 \tag{2.10}$$

for all sufficiently large N and $z\in(\mathbb{C}^d)^{N+2}$. It follows that

$$\limsup_{N\uparrow\infty}\sup_{\substack{z\in(\mathbb{C}^d)^{N+2}\\ z\neq 0}}\left\{\lambda_o\frac{\langle z,\mathsf{M}_N z\rangle}{\langle z,\Sigma_N^{-1}z\rangle}\right\}\leq\xi,$$

which demonstrates (2.9), thanks to the arbitrariness of $\xi>\xi_o$.

Let us demonstrate now the connection between the set Λ and the number ξ_- and ξ_+. Fix $\lambda\in\Lambda$ and consider the HQT matrix sequence $\{\mathsf{Q}_N\}_{N\geq 1}$ with matrices $\mathsf{Q}_N:=\Sigma_N^{-1}-\lambda\mathsf{M}_N$. We already know that this HQT matrix sequence has symbol F_λ and limit boundary matrix

$\mathsf{S}_\infty = \begin{pmatrix} \mathcal{L}_\lambda & 0 \\ 0 & \mathcal{R}_\lambda \end{pmatrix}$. Since $f_\lambda > 0$, $\mathcal{L}_\lambda > 0$, and $\mathcal{R}_\lambda > 0$ by hypothesis, we have $\mathsf{Q}_N > 0$ for all sufficiently large N according to Proposition 2.1. This shows that

$$1 - \lambda \frac{\langle z, \mathsf{M}_N z\rangle}{\langle z, \Sigma_N^{-1} z\rangle} > 0$$

for all sufficiently large N and $z \in (\mathbb{C}^d)^{N+2}$. Thus, by taking the infimum over z, we have

$$1 - \lambda \sup_{\substack{z\in(\mathbb{C}^d)^{N+2}\\ z\neq 0}} \left\{ \frac{\langle z, \mathsf{M}_N z\rangle}{\langle z, \Sigma_N^{-1} z\rangle} \right\} > 0$$

if $\lambda \geq 0$ and

$$1 - \lambda \inf_{\substack{z\in(\mathbb{C}^d)^{N+2}\\ z\neq 0}} \left\{ \frac{\langle z, \mathsf{M}_N z\rangle}{\langle z, \Sigma_N^{-1} z\rangle} \right\} > 0$$

if $\lambda < 0$. By sending N to infinity, we realize that $\lambda\xi_+ \leq 1$ if $\lambda \geq 0$, which also gives $\lambda\xi_- \leq 1$ as $\xi_- \leq \xi_+$, and that $\lambda\xi_- \leq 1$ if $\lambda < 0$, which also gives $\lambda\xi_+ \leq 1$.

Conversely, if $\lambda \in \mathbb{R}$ is such that $\lambda\xi_- < 1$ and $\lambda\xi_+ < 1$, then there exists $\epsilon > 0$ such that $1 - \lambda\xi_- \geq 2\epsilon$ and $1 - \lambda\xi_+ \geq 2\epsilon$. This yields that for all sufficiently large N and $z \in (\mathbb{C}^d)^{N+2}$,

$$\langle z, \Sigma_N^{-1} z\rangle - \lambda\langle z, \mathsf{M}_N z\rangle \geq \epsilon\langle z, \Sigma_N^{-1} z\rangle.$$

This way, $r(\mathsf{Q}_N) \geq \epsilon r(\Sigma_N^{-1}) \geq \epsilon\sigma$ for all sufficiently large N, where $\mathsf{Q}_N := \Sigma_N^{-1} - \lambda\mathsf{M}_N$, and Lemma 2.1 has been invoked. It follows from Lemma 2.2 that $r(\mathsf{T}_N) \geq \epsilon\sigma$ and $r(\mathsf{S}_N) \geq \epsilon\sigma$ for all sufficiently large N so that $\inf_{\theta\in[0,2\pi]}\{r(F_\lambda(\theta))\} \geq \epsilon\sigma$ by Lemma 2.3 and $r(S_\infty) \geq \epsilon\sigma$ by Lemma 2.4 with $\mathsf{S}_\infty = \begin{pmatrix} \mathcal{L}_\lambda & 0 \\ 0 & \mathcal{R}_\lambda \end{pmatrix}$. Thus, $f_\lambda > 0$, $\mathcal{L}_\lambda > 0$, and $\mathcal{R}_\lambda > 0$ so that $\lambda \in \Lambda$. □

The limit (2.7) together with the fact that Λ is an interval finally give the following important result. We stress that Λ contains an open neighborhood of the origin, as it is manifest by Lemma 2.7, so that $\lambda_- < 0 < \lambda_+$.

Proposition 2.2. For all $\lambda \in (\lambda_-, \lambda_+)$,

$$\lim_{N\uparrow\infty} \frac{1}{N} \ln \mathbb{E}\left[e^{\lambda W_N}\right] = -\frac{1}{4\pi}\int_0^{2\pi} \ln\det F_\lambda(\theta)\, d\theta.$$

The function φ that maps any $\lambda \in (\lambda_-, \lambda_+)$ in

$$\varphi(\lambda) := -\frac{1}{4\pi}\int_0^{2\pi} \ln\det F_\lambda(\theta)\, d\theta$$

is convex. A rapid way to demonstrate this fact is to observe that φ is the limit of a sequence of convex functions by Proposition 2.2. The function φ is also differentiable since the $F_\lambda(\theta)$ is differentiable with respect to λ for each $\theta \in [0, 2\pi]$. The asymptotic theory of sequences of convex functions (see Ref. 39, Theorem 24.5) gives that for all $\lambda \in (\lambda_-, \lambda_+)$,

$$\lim_{N\uparrow\infty} \frac{1}{N}\frac{d}{d\lambda} \ln \mathbb{E}\left[e^{\lambda W_N}\right] = \varphi'(\lambda). \tag{2.11}$$

Limit (2.11) will serve us to verify the lower large deviation bound. Other notable consequences of convexity are that the limits $\lim_{\lambda\downarrow\lambda_-}\varphi(\lambda) =: \varphi_-$ and $\lim_{\lambda\uparrow\lambda_+}\varphi(\lambda) =: \varphi_+$ exist (see Ref. 39, Theorem 7.5) and that φ' is non-decreasing in such a way that also the limits $\lim_{\lambda\downarrow\lambda_-}\varphi'(\lambda) =: d_-$ and $\lim_{\lambda\uparrow\lambda_+}\varphi'(\lambda) =: d_+$ exist.

Remark 2.1. In Sec. I A, we have claimed that $\lim_{N\uparrow\infty}(1/N)\ln\mathbb{E}[e^{\lambda W_N}] = +\infty$ if $\lambda \notin \overline{(\lambda_-, \lambda_+)} = \bar{\Lambda}$. Although we do not need this limit to prove a large deviation principle, we can verify it as follows. Assume, for instance, that $\lambda_+ < +\infty$ and pick $\lambda > \lambda_+ > 0$. It must be $\lambda\xi_+ > 1$ since, on the contrary, $\lambda\xi_- \leq \lambda\xi_+ \leq 1$ and $\lambda \in \bar{\Lambda}$ as a consequence according to Lemma 2.7. Fix $N \geq 1$, and let $\mathsf{A} \in \mathsf{BL}_{N+2,N+2}$ be a real invertible matrix such *that* $\Sigma_N = \mathsf{A}\mathsf{A}^\top$*, which exists because* $\Sigma_N > 0$*. Denoting by* $m_1, \ldots, m_{(N+2)d}$ *the eigenvalues of the real symmetric matrix* $\mathsf{A}^\top \mathsf{M}_N \mathsf{A} \in \mathsf{BL}_{N+2,N+2}$*, we have*

$$\xi_N := \max\{m_1, \ldots, m_{(N+2)d}\} = \sup_{\substack{z\in(\mathbb{C}^d)^{N+2} \\ z\neq 0}} \left\{\frac{\langle z, \mathsf{A}^\top \mathsf{M}_N \mathsf{A} z\rangle}{\langle z, z\rangle}\right\} = \sup_{\substack{z\in(\mathbb{C}^d)^{N+2} \\ z\neq 0}} \left\{\frac{\langle \mathsf{A}z, \mathsf{M}_N \mathsf{A} z\rangle}{\langle z, z\rangle}\right\}$$
$$= \sup_{\substack{z\in(\mathbb{C}^d)^{N+2} \\ z\neq 0}} \left\{\frac{\langle z, \mathsf{M}_N z\rangle}{\langle \mathsf{A}^{-1}z, \mathsf{A}^{-1} z\rangle}\right\} = \sup_{\substack{z\in(\mathbb{C}^d)^{N+2} \\ z\neq 0}} \left\{\frac{\langle z, \mathsf{M}_N z\rangle}{\langle z, \Sigma_N^{-1} z\rangle}\right\}.$$

Thus, Lemma 2.7 tells us that the number $\lambda\xi_N$ *approaches* $\lambda\xi_+ > 1$ *at large N so that* $1 - \lambda\xi_N \leq 0$ *if N exceeds a threshold value* N_o*. This shows that the matrix* $\Sigma_N^{-1} - \lambda \mathsf{M}_N = (\mathsf{A}^{-1})^\top (\mathsf{I} - \lambda \mathsf{A}^\top \mathsf{M}_N \mathsf{A})\mathsf{A}^{-1}$ *is not positive-definite for* $N > N_o$*. Lemma 2.5 concludes the proof.*

C. The upper large deviation bound

In this section, we prove the upper large deviation bound for closed sets. We start with some standard results from the theory of large deviations that we shall use to prove both the upper large deviation bound and the lower large deviation bound. For each $\eta \in (\lambda_-, \lambda_+)$ and $N \geq 1$, let $\mathbb{P}_{\eta,N}$ be the probability measure on $(\Omega, \mathscr{F})$ defined by the exponential change of measure,

$$\frac{d\mathbb{P}_{\eta,N}}{d\mathbb{P}} := \frac{e^{\eta W_N}}{\mathbb{E}[e^{\eta W_N}]}. \tag{2.12}$$

Let φ_η be the function that maps any $\lambda \in \mathbb{R}$ in

$$\varphi_\eta(\lambda) := \limsup_{N\uparrow\infty} \frac{1}{N} \ln \mathbb{E}_{\eta,N}\left[e^{\lambda W_N}\right]. \tag{2.13}$$

Since $0 \in (\lambda_-, \lambda_+)$, we have $\mathbb{P}_{0,N} = \mathbb{P}$ for all $N \geq 1$ and $\varphi_0(\lambda) = \varphi(\lambda)$ for all $\lambda \in (\lambda_-, \lambda_+)$, with φ being the convex differentiable function introduced at the end of Sec. II B. Moreover, if $\lambda \in \mathbb{R}$ is such that $\lambda + \eta \in (\lambda_-, \lambda_+)$, then

$$\varphi_\eta(\lambda) = \limsup_{N\uparrow\infty} \frac{1}{N} \ln \frac{\mathbb{E}[e^{(\lambda+\eta)W_N}]}{\mathbb{E}[e^{\eta W_N}]} = \varphi(\lambda + \eta) - \varphi(\eta) < +\infty.$$

It follows that the function φ_η is finite and differentiable in an open neighborhood of the origin with $\varphi_\eta'(0) = \varphi'(\eta)$. The following lemma states an upper large deviation bound with respect to the measure $\mathbb{P}_{\eta,N}$. We recall that the Fenchel–Legendre transform I_η of φ_η is the convex function that associates $w \in \mathbb{R}$ with

$$I_\eta(w) := \sup_{\lambda\in\mathbb{R}} \{w\lambda - \varphi_\eta(\lambda)\}.$$

Lemma 2.8. Fix $\eta \in (\lambda_-, \lambda_+)$*. The following conclusions hold:*

1. *the Fenchel–Legendre transform* I_η *of* φ_η *has compact level sets;*
2. *for each closed set* $\mathcal{F} \subseteq \mathbb{R}$,
$$\limsup_{N\uparrow\infty} \frac{1}{N} \ln \mathbb{P}_{\eta,N}\left[\frac{W_N}{N} \in \mathcal{F}\right] \leq -\inf_{w\in\mathcal{F}} \{I_\eta(w)\};$$
3. *for each* $\epsilon > 0$*, there exists* $\kappa > 0$ *such that for all sufficiently large N*,
$$\mathbb{P}_{\eta,N}\left[\left|\frac{W_N}{N} - \varphi'(\eta)\right| \geq \epsilon\right] \leq e^{-\kappa N}.$$

Proof. I_η is lower semicontinuous as any Fenchel–Legendre transform (see Ref. 39, Theorem 12.2). Due to lower semicontinuity, the level sets of I_η are closed. In order to prove part 1, it remains to verify that they are bounded. As the function φ_η is finite in an open neighborhood of the origin, there exists $\delta > 0$ such that $\varphi_\eta(\delta) < +\infty$ and $\varphi_\eta(-\delta) < +\infty$. If $I_\eta(w) \leq a$, for given real numbers w and a, then $w\delta - \varphi_\eta(\delta) \leq I_\eta(w) \leq a$ and $-w\delta - \varphi_\eta(-\delta) \leq I_\eta(w) \leq a$ by definition, that is, $-[a + \varphi_\eta(-\delta)]/\delta \leq w \leq [a + \varphi_\eta(\delta)]/\delta$.

Part 2 is a standard result from large deviation theory (see Ref. 15, Theorem 2.3.6 and Exercise 2.3.25). In a nutshell, the upper large deviation bound for compact sets is a manipulation of the Chernoff bound and holds without any assumption on the function φ_η. Extension to all closed sets is made possible by finiteness of φ_η in an open neighborhood of the origin, which entails exponential tightness.

As far as part 3 is concerned, in light of part 2, it suffices to demonstrate that

$$\inf_{v\notin(w-\epsilon,w+\epsilon)}\{I_\eta(v)\}>0 \tag{2.14}$$

for each $\epsilon>0$, with $w:=\varphi'(\eta)$. To begin with, let us observe that $I_\eta(v)>0$ if $v\neq w$. On the contrary, if $I_\eta(v)=0$, then for all λ in a neighborhood of the origin, we would have $\varphi(\lambda+\eta)-\varphi(\eta)=\varphi_\eta(\lambda)\geq v\lambda$ by definition of $I_\eta(v)$. This would imply $v=\varphi'(\eta)=:w$, which contradicts the hypothesis $v\neq w$. We can now verify (2.14). Pick $\epsilon>0$ and notice that the set $\mathcal{A}:=\{v\in\mathbb{R}:I_\eta(v)\leq 1\}$ is compact by part 1. If $(w-\epsilon,w+\epsilon)^c\cap\mathcal{A}=\emptyset$, then $\inf_{v\notin(w-\epsilon,w+\epsilon)}\{I_\eta(v)\}\geq 1$. If $(w-\epsilon,w+\epsilon)^c\cap\mathcal{A}\neq\emptyset$, then there exists $v_\star\in(w-\epsilon,w+\epsilon)^c\cap\mathcal{A}$ such that $I_\eta(v)\geq I_\eta(v_\star)$ for all $v\in(w-\epsilon,w+\epsilon)^c\cap\mathcal{A}$ and, hence, for all $v\notin(w-\epsilon,w+\epsilon)$, as I_η is a lower semicontinuous function and $(w-\epsilon,w+\epsilon)^c\cap\mathcal{A}$ is a compact set. On the other hand, we have $I_\eta(v_\star)>0$ since $v_\star\neq w$. □

Lemma 2.8 gives the following upper large deviation bound for the quadratic functionals W_N of the stable Gauss–Markov processes $\{X_n\}_{n\geq 1}$:

Proposition 2.3. The following conclusions hold:

1. *the convex function I that maps $w\in\mathbb{R}$ in $I(w):=\sup_{\lambda\in(\lambda_-,\lambda_+)}\{w\lambda-\varphi(\lambda)\}$ has compact level sets;*
2. *for each closed set $\mathcal{F}\subseteq\mathbb{R}$,*

$$\limsup_{N\uparrow\infty}\frac{1}{N}\ln\mathbb{P}\left[\frac{W_N}{N}\in\mathcal{F}\right]\leq-\inf_{w\in\mathcal{F}}\{I(w)\}.$$

Proof. I is the Fenchel–Legendre transform of the function that associates $\lambda\in(\lambda_-,\lambda_+)$ with $\varphi(\lambda)$ and $\lambda\notin(\lambda_-,\lambda_+)$ with $+\infty$. Then, part 1 is proved in the same way of part 1 of Lemma 2.8. Part 2 follows from part 2 of Lemma 2.8 with $\eta=0$ as $I_0(w):=\sup_{\lambda\in\mathbb{R}}\{w\lambda-\varphi_0(\lambda)\}\geq\sup_{\lambda\in(\lambda_-,\lambda_+)}\{w\lambda-\varphi(\lambda)\}=:I(w)$ for every $w\in\mathbb{R}$. □

D. The lower large deviation bound

In this section, we prove the lower large deviation bound for open sets, namely, that for each open set $\mathcal{G}\subseteq\mathbb{R}$,

$$\liminf_{N\uparrow\infty}\frac{1}{N}\ln\mathbb{P}\left[\frac{W_N}{N}\in\mathcal{G}\right]\geq-\inf_{w\in\mathcal{G}}\{I(w)\},$$

where I is the function that maps $w\in\mathbb{R}$ in $I(w):=\sup_{\lambda\in(\lambda_-,\lambda_+)}\{w\lambda-\varphi(\lambda)\}$. This is tantamount to state that for all $w\in\mathbb{R}$ and $\delta>0$,

$$\liminf_{N\uparrow\infty}\frac{1}{N}\ln\mathbb{P}\left[\frac{W_N}{N}\in(w-\delta,w+\delta)\right]\geq-I(w). \tag{2.15}$$

We start with the following lower bound based on Lemma 2.8:

Lemma 2.9. Fix $w\in\mathbb{R}$ and assume that there exists $\eta\in(\lambda_-,\lambda_+)$ such that $w=\varphi'(\eta)$. Then, for every $\delta>0$,

$$\liminf_{N\uparrow\infty}\frac{1}{N}\ln\mathbb{P}\left[\frac{W_N}{N}\in(w-\delta,w+\delta)\right]\geq\varphi(\eta)-w\eta.$$

Proof. Let $\mathbb{P}_{\eta,N}$ and φ_η be the probability measure (2.12) and the function (2.13), respectively. Fix $\delta>0$ and pick $\epsilon\in(0,\delta)$. The fact that $\eta W_N-Nw\eta-N\epsilon|\eta|\leq 0$ if $W_N/N\in(w-\epsilon,w+\epsilon)$ gives for each $N\geq 1$,

$$\begin{aligned}\mathbb{P}\left[\frac{W_N}{N}\in(w-\delta,w+\delta)\right]&\geq e^{-Nw\eta-N\epsilon|\eta|}\,\mathbb{E}\left[e^{\eta W_N}\mathbb{1}_{\left\{\frac{W_N}{N}\in(w-\epsilon,w+\epsilon)\right\}}\right]\\&=e^{-Nw\eta-N\epsilon|\eta|}\,\mathbb{E}\left[e^{\eta W_N}\right]\mathbb{P}_{\eta,N}\left[\left|\frac{W_N}{N}-w\right|<\epsilon\right],\end{aligned}$$

and part 3 of Lemma 2.8 shows that

$$\lim_{N\uparrow\infty} \mathbb{P}_{\eta,N}\bigg[\bigg|\frac{W_N}{N} - w\bigg| < \epsilon\bigg] = 1.$$

Thus, by invoking Proposition 2.2, we obtain

$$\liminf_{N\uparrow\infty} \frac{1}{N} \ln \mathbb{P}\bigg[\frac{W_N}{N} \in (w-\delta, w+\delta)\bigg] \geq \varphi(\eta) - w\eta - \epsilon|\eta|.$$

The lemma follows from here by sending ϵ to 0. □

Lemma 2.9 allows us to demonstrate the lower large deviation bound (2.15) for w in the closure $\overline{(d_-,d_+)}$ of (d_-,d_+), where $d_- := \lim_{\lambda\downarrow\lambda_-}\varphi'(\lambda)$ and $d_+ := \lim_{\lambda\uparrow\lambda_+}\varphi'(\lambda)$ as in Sec. II B. Note that convexity and differentiability of φ yield $\varphi(\lambda) \geq \varphi(\eta) + \varphi'(\eta)\,(\lambda-\eta)$ for every λ and η in (λ_-,λ_+) so that $I(w) = w\eta - \varphi(\eta)$ if $w = \varphi'(\eta)$ for some $\eta \in (\lambda_-,\lambda_+)$. Since $d_- \leq \varphi'(0) \leq d_+$ as φ' is non-decreasing, if $d_- = d_+$, then $\overline{(d_-,d_+)}$ contains only $\varphi'(0)$ and bound (2.15) directly follows from Lemma 2.9 with $\eta = 0$. If $d_- < d_+$ and $w \in (d_-,d_+)$, then there exists $\eta \in (\lambda_-,\lambda_+)$ such that $w = \varphi'(\eta)$, and bound (2.15) follows again from Lemma 2.9 with such η. If $d_- < d_+ < +\infty$, then $\overline{(d_-,d_+)}$ contains d_+, and we tackle the case $w = d_+$ as follows: Fix $\delta > 0$. There exist $v \in (d_-,d_+)$ arbitrarily close to w and $\epsilon > 0$ such that $(v-\epsilon, v+\epsilon) \subseteq (w-\delta, w+\delta)$. This way, since (2.15) holds for v, we find

$$\liminf_{N\uparrow\infty} \frac{1}{N} \ln \mathbb{P}\bigg[\frac{W_N}{N} \in (w-\delta, w+\delta)\bigg] \geq \liminf_{N\uparrow\infty} \frac{1}{N} \ln \mathbb{P}\bigg[\frac{W_N}{N} \in (v-\epsilon, v+\epsilon)\bigg] \geq -I(v).$$

From here, we get bound (2.15) for $w = d_+$ by sending v to w and by observing that $\lim_{v\uparrow w} I(v) = I(w)$ by convexity and lower semicontinuity of I (see Ref. 39, Corollary 7.5.1). Similar arguments can be used to solve the case $-\infty < d_- < d_+$ and $w = d_-$.

In order to complete the proof of the lower large deviation bound (2.15), it remains to address the case $d_+ < +\infty$ and $w > d_+$, as well as the case $d_- > -\infty$ and $w < d_-$. They are similar so that we discuss in detail the former only, omitting the proof of the latter. Assume that $d_+ < +\infty$ and fix $w > d_+$. We claim that the case $\lambda_+ = +\infty$ is trivial so that we also suppose $\lambda_+ < +\infty$. In fact, convexity and differentiability of φ combined with $\varphi(0) = 0$ give $\varphi(\lambda) \leq \lambda\varphi'(\lambda)$ for all $\lambda \in (\lambda_-,\lambda_+)$. It follows that $I(w) \geq w\lambda - \varphi(\lambda) \geq \lambda[w - \varphi'(\lambda)]$ for all $\lambda \in (\lambda_-,\lambda_+)$. Thus, if $\lambda_+ = +\infty$, then we realize that $I(w) = +\infty$ by sending λ to λ_+, as $\lim_{\lambda\uparrow\lambda_+}\varphi'(\lambda) = d_+ < w$, and the lower bound (2.15) is trivial. Observe that if $\lambda_+ < +\infty$ and $d_+ < +\infty$, then $\lim_{\lambda\uparrow\lambda_+}\varphi(\lambda) =: \varphi_+ < +\infty$ as $\varphi(\lambda) \leq \lambda\varphi'(\lambda)$ for all $\lambda \in (\lambda_-,\lambda_+)$. Since the function that associates $\lambda \in (\lambda_-,\lambda_+)$ with $w\lambda - \varphi(\lambda)$ is increasing under the hypothesis $w > d_+$, we have $I(w) := \sup_{\lambda\in(\lambda_-,\lambda_+)}\{w\lambda - \varphi(\lambda)\} = w\lambda_+ - \varphi_+$.

The idea to prove (2.15) for $w > d_+$ and $\lambda_+ < +\infty$ is to make a change of measure like in Lemma 2.9, but this time the parameter η must depend on the time N. Let us introduce such parameter. Pick $N \geq 1$. Since the covariance matrix $\Sigma_N \in \mathsf{BL}_{N+2,N+2}$ is symmetric positive-definite, there exists a real invertible matrix $\mathsf{A} \in \mathsf{BL}_{N+2,N+2}$ such that $\Sigma_N = \mathsf{A}\mathsf{A}^\top$. Like in Remark 2.1, let $m_1,\ldots,m_{(N+2)d}$ be the eigenvalues of the real symmetric matrix $\mathsf{A}^\top\mathsf{M}_N\mathsf{A} \in \mathsf{BL}_{N+2,N+2}$ and observe that

$$\begin{aligned}\xi_N := \max\{m_1,\ldots,m_{(N+2)d}\} &= \sup_{\substack{z\in(\mathbb{C}^d)^{N+2}\\ z\neq 0}}\bigg\{\frac{\langle z, \mathsf{A}^\top\mathsf{M}_N\mathsf{A}z\rangle}{\langle z,z\rangle}\bigg\}\\ &= \sup_{\substack{z\in(\mathbb{C}^d)^{N+2}\\ z\neq 0}}\bigg\{\frac{\langle z, \mathsf{M}_N z\rangle}{\langle z, (\mathsf{A}^{-1})^\top\mathsf{A}^{-1}z\rangle}\bigg\} = \sup_{\substack{z\in(\mathbb{C}^d)^{N+2}\\ z\neq 0}}\bigg\{\frac{\langle z, \mathsf{M}_N z\rangle}{\langle z, \Sigma_N^{-1}z\rangle}\bigg\}.\end{aligned}$$

Similarly,

$$\min\{m_1,\ldots,m_{(N+2)d}\} = \inf_{\substack{z\in(\mathbb{C}^d)^{N+2}\\ z\neq 0}}\bigg\{\frac{\langle z, \mathsf{M}_N z\rangle}{\langle z, \Sigma_N^{-1}z\rangle}\bigg\}.$$

We have $\lim_{N\uparrow\infty}\xi_N = \xi_+$ by Lemma 2.7, whereas $\min\{m_1,\ldots,m_{(N+2)d}\}$ approaches ξ_- at large N. Lemma 2.7 also gives $\xi_+ > 0$ since $\lambda_+ < +\infty$ by hypothesis. Indeed, $\xi_+ \leq 0$ would entail that the set Λ contains all positive real numbers. Lemma 2.5 and the fact that $\det\Sigma_N = \det\Sigma_o$ show that if $\mathsf{I} - \lambda\mathsf{A}^\top\mathsf{M}_N\mathsf{A} > 0$, namely, if $1 - \lambda m_l > 0$ for $l = 1,\ldots,(N+2)d$, then

$$\ln\mathbb{E}\Big[e^{\lambda W_{N+2}}\Big] = -\frac{1}{2}\ln\det\big(\mathsf{I} - \lambda\mathsf{A}^\top\mathsf{M}_N\mathsf{A}\big) = -\frac{1}{2}\sum_{l=1}^{(N+2)d}\ln(1-\lambda m_l). \tag{2.16}$$

We claim that for all sufficiently large N, there exists $\eta_N \in (0, \xi_N)$ such that

$$\frac{1}{2(N+2)} \sum_{l=1}^{(N+2)d} \frac{m_l}{1 - \eta_N m_l} = w. \tag{2.17}$$

Note that $\xi_N > 0$ for all sufficiently large N as $\lim_{N\uparrow\infty} \xi_N = \xi_+ > 0$. In fact, identity (2.16) in combination with (2.11) yields that $[2(N+2)]^{-1}\sum_{l=1}^{(N+2)d} m_l$ approaches $\varphi'(0) \le d_+ < w$ when N is sent to infinity. Thus, for all sufficiently large N, the continuous function that maps $\lambda \in [0, \xi_N)$ in $[2(N+2)]^{-1}\sum_{l=1}^{(N+2)d} m_l(1 - \lambda m_l)^{-1}$ increases from a value smaller than w at $\lambda = 0$ to $+\infty$ at $\lambda = \xi_N$ so that there exists a unique η_N satisfying (2.17). We must have $\lim_{N\uparrow\infty} \eta_N = \lambda_+$. On the contrary, there would exist $\epsilon > 0$ and a diverging sequence $\{N_k\}_{k\ge 0}$ of positive integers such that $\eta_{N_k} < \lambda_+ - \epsilon$ for all $k \ge 0$. Then, for every k,

$$w = \frac{1}{2(N_k+2)} \sum_{l=1}^{(N_k+2)d} \frac{m_l}{1 - \eta_{N_k} m_l} \le \frac{1}{2(N_k+2)} \sum_{l=1}^{(N_k+2)d} \frac{m_l}{1 - (\lambda_+ - \epsilon) m_l}.$$

By sending k to infinity and by combining (2.16) with (2.11), from here, we would get $w \le \varphi'(\lambda_+ - \epsilon) \le d_+$, which contradicts the assumption $w > d_+$. Another property of η_N is that

$$\liminf_{N\uparrow\infty} \frac{1}{N+2} \ln \mathbb{E}\Big[e^{\eta_N W_{N+2}}\Big] \ge \varphi_+. \tag{2.18}$$

In order to verify this bound, fix $\lambda \in (0, \lambda_+)$ and bear in mind that $\eta_N \ge \lambda$ for all sufficiently large N as $\lim_{N\uparrow\infty} \eta_N = \lambda_+$ so that

$$-\ln(1 - \eta_N m_l) \ge -\ln(1 - \lambda m_l) + (\eta_N - \lambda) \min\{0, m_1, \ldots, m_{(N+2)d}\}$$

for every l and sufficiently large N. Then, for all sufficiently large N, we have

$$\begin{aligned}
\frac{1}{N+2} \ln \mathbb{E}\Big[e^{\eta_N W_{N+2}}\Big] &= -\frac{1}{2(N+2)} \sum_{l=1}^{(N+2)d} \ln(1 - \eta_N m_l) \\
&\ge -\frac{1}{2(N+2)} \sum_{l=1}^{(N+2)d} \ln(1 - \lambda m_l) + (\eta_N - \lambda) d \min\{0, m_1, \ldots, m_{(N+2)d}\} \\
&= \frac{1}{N+2} \ln \mathbb{E}\Big[e^{\lambda W_{N+2}}\Big] + (\eta_N - \lambda) d \min\{0, m_1, \ldots, m_{(N+2)d}\}.
\end{aligned}$$

By sending N to infinity and by recalling that $\min\{m_1, \ldots, m_{(N+2)d}\}$ approaches ξ_- in this limit, Proposition 2.2 shows that

$$\liminf_{N\uparrow\infty} \frac{1}{N+2} \ln \mathbb{E}\Big[e^{\eta_N W_{N+2}}\Big] \ge \varphi(\lambda) + (\lambda_+ - \lambda) d \min\{0, \xi_-\},$$

which demonstrates (2.18) once λ is sent λ_+.

We now move to bound (2.15) and put η_N into context. Fix $\delta > 0$ and pick $\epsilon \in (0, \delta)$. For all sufficiently large N, η_N is positive as $\lim_{N\uparrow\infty} \eta_N = \lambda_+$, and we have

$$\begin{aligned}
\mathbb{P}\Big[\frac{W_{N+2}}{N+2} \in (w - \delta, w + \delta)\Big] &\ge e^{-(N+2)(w+\epsilon)\eta_N} \, \mathbb{E}\Big[e^{\eta_N W_{N+2}} \mathbb{1}_{\{\frac{W_{N+2}}{N+2} \in (w-\epsilon, w+\epsilon)\}}\Big] \\
&= e^{-(N+2)(w+\epsilon)\eta_N} \, \mathbb{E}\Big[e^{\eta_N W_{N+2}}\Big] \, \mathbb{P}_{\eta_N, N+2}\Big[\Big|\frac{W_{N+2}}{N+2} - w\Big| < \epsilon\Big],
\end{aligned}$$

where $\mathbb{P}_{\eta_N, N+2}$ is the probability measure (2.12) associated with η_N. This bound, together with (2.18), yields

$$\begin{aligned}
\liminf_{N\uparrow\infty} \frac{1}{N} \ln \mathbb{P}\Big[\frac{W_N}{N} \in (w - \delta, w + \delta)\Big] &\ge \varphi_+ - w\lambda_+ - \epsilon\lambda_+ \\
&\quad + \liminf_{N\uparrow\infty} \frac{1}{N+2} \ln \mathbb{P}_{\eta_N, N+2}\Big[\Big|\frac{W_{N+2}}{N+2} - w\Big| < \epsilon\Big].
\end{aligned}$$

This way, as $w\lambda_+ - \varphi_+ = I(w)$, we get at the lower large deviation bound (2.15) from here if we can prove that

$$\lim_{\epsilon\downarrow 0}\liminf_{N\uparrow\infty}\frac{1}{N+2}\ln \mathbb{P}_{\eta_N,N+2}\bigg[\bigg|\frac{W_{N+2}}{N+2}-w\bigg|<\epsilon\bigg]=0. \tag{2.19}$$

Verifying (2.19) is our last task. To this aim, we resort to the following result, which was introduced by Bryc and Dembo (see Ref. 28, Lemma 2) to deal with a similar problem:

Lemma 2.10. If $\{Z_l\}_{l\geq 1}$ is a sequence of i.i.d. random variables with mean zero, finite second moment, and positive probability density function at 0 with respect to a probability measure P, then for each $\epsilon > 0$, there exists $p > 0$ such that the following property holds:

$$P\bigg[\bigg|\sum_{l\geq 1}a_l Z_l\bigg|<\epsilon\bigg]\geq p$$

for any numerical sequence $\{a_l\}_{l\geq 1}$ such that $\sum_{l\geq 1}|a_l|\leq 1$.

Let $\{Y_l\}_{l\geq 1}$ be a sequence of independent standard Gaussian random variables with respect to a probability measure P. Lemma 2.10 ensures that for each $\epsilon > 0$, there exists $p > 0$ with the property that

$$P\bigg[\bigg|\sum_{l\geq 1}a_l\big(Y_l^2-1\big)\bigg|<\frac{\epsilon}{1+|w|+3d|\xi_-|}\bigg]\geq p \tag{2.20}$$

for any numerical sequence $\{a_l\}_{l\geq 1}$ such that $\sum_{l\geq 1}|a_l|\leq 1$, with ξ_- being the number introduced by Lemma 2.7. We make use of property (2.20) to prove (2.19). Since the real symmetric matrix $\mathsf{I}-\eta_N\mathsf{A}^\top\mathsf{M}_N\mathsf{A}\in\mathsf{BL}_{N+2,N+2}$ has positive eigenvalues $1-\eta_N m_1,\ldots,1-\eta_N m_{(N+2)d}$, if we build a diagonal matrix D with $\sqrt{1-\eta_N m_1},\ldots,\sqrt{1-\eta_N m_{(N+2)d}}$ on the diagonal, then $\mathsf{I}-\eta_N\mathsf{A}^\top\mathsf{M}_N\mathsf{A}=\mathsf{O}^\top\mathsf{D}^2\mathsf{O}$ with an orthogonal matrix $\mathsf{O}\in\mathsf{BL}_{N+2,N+2}$. This way, if we write $W_{N+2}=(1/2)\langle X,\mathsf{M}_N X\rangle$ with $X:=(X_1,\ldots,X_{N+2})$, then standard manipulations of Gaussian integrals yield for all $k\in(\mathbb{R}^d)^{N+2}$,

$$\begin{aligned}\mathbb{E}_{\eta_N,N+2}\Big[e^{i\langle k,\mathsf{DOA}^{-1}X\rangle}\Big]&=\frac{\mathbb{E}\big[e^{i\langle(\mathsf{A}^\top)^{-1}\mathsf{O}^\top\mathsf{D}k,X\rangle+\frac{1}{2}\eta_N\langle X,\mathsf{M}_N X\rangle}\big]}{\mathbb{E}\big[e^{\frac{1}{2}\eta_N\langle X,\mathsf{M}_N X\rangle}\big]}\\&=e^{-\frac{1}{2}\langle(\mathsf{A}^\top)^{-1}\mathsf{O}^\top\mathsf{D}k,(\Sigma_N^{-1}-\eta_N\mathsf{M}_N)^{-1}(\mathsf{A}^\top)^{-1}\mathsf{O}^\top\mathsf{D}k\rangle}\\&=e^{-\frac{1}{2}\langle\mathsf{O}^\top\mathsf{D}k,(\mathsf{I}-\eta_N\mathsf{A}^\top\mathsf{M}_N\mathsf{A})^{-1}\mathsf{O}^\top\mathsf{D}k\rangle}=e^{-\frac{1}{2}\langle k,k\rangle}.\end{aligned}$$

This formula states that the characteristic function of the random vector $Y:=\mathsf{DOA}^{-1}X$ with respect to the probability measure $\mathbb{P}_{\eta_N,N+2}$ is the characteristic function of $(N+2)d$ independent standard Gaussian random variables. Thus, the components $Y_1,\ldots,Y_{(N+2)d}$ of Y are independent standard Gaussian random variables with respect to the probability measure $\mathbb{P}_{\eta_N,N+2}$. It follows from (2.20) that for each $\epsilon > 0$, there exists $p > 0$ with the property that

$$\mathbb{P}_{\eta_N,N+2}\Bigg[\Bigg|\sum_{l=1}^{(N+2)d}a_l\big(Y_l^2-1\big)\Bigg|<\frac{\epsilon}{1+|w|+3d|\xi_-|}\Bigg]\geq p \tag{2.21}$$

for all $N\geq 1$ and real numbers $a_1,\ldots,a_{(N+2)d}$ such that $\sum_{l=1}^{(N+2)d}|a_l|\leq 1$. Let us observe now that

$$W_{N+2}=\frac{1}{2}\langle X,\mathsf{M}_N X\rangle=\frac{1}{2}\langle\mathsf{O}^\top\mathsf{D}^{-1}Y,\mathsf{A}^\top\mathsf{M}_N\mathsf{A}\mathsf{O}^\top\mathsf{D}^{-1}Y\rangle=\frac{1}{2}\sum_{l=1}^{(N+2)d}\frac{m_l}{1-\eta_N m_l}Y_l^2.$$

This identity combined with (2.17) shows that for all $\epsilon > 0$ and sufficiently large N,

$$\mathbb{P}_{\eta_N,N+2}\bigg[\bigg|\frac{W_{N+2}}{N+2}-w\bigg|<\epsilon\bigg]=\mathbb{P}_{\eta_N,N+2}\Bigg[\Bigg|\sum_{l=1}^{(N+2)d}a_l\big(Y_l^2-1\big)\Bigg|<\frac{\epsilon}{1+|w|+3d|\xi_-|}\Bigg], \tag{2.22}$$

where for $l=1,\ldots,(N+2)d$, we have set

$$a_l:=\frac{1}{2(N+2)(1+|w|+3d|\xi_-|)}\frac{m_l}{1-\eta_N m_l}.$$

We have $\sum_{l=1}^{(N+2)d}|a_l| \leq 1$ for all sufficiently large N. In fact, since

$$\frac{|m_l|}{1-\eta_N m_l} \leq \frac{m_l}{1-\eta_N m_l} - 2\min\{0, m_1, \ldots, m_{(N+2)d}\}$$

for every l, by invoking (2.17) and by recalling that $\min\{m_1, \ldots, m_{(N+2)d}\}$ approaches ξ_- at large N, for all sufficiently large N, we find

$$\begin{aligned}(1+|w|+3d|\xi_-|)\sum_{l=1}^{(N+2)d}|a_l| &= \frac{1}{2(N+2)}\sum_{l=1}^{(N+2)d}\frac{|m_l|}{1-\eta_N m_l} \\ &\leq \frac{1}{2(N+2)}\sum_{l=1}^{(N+2)d}\frac{m_l}{1-\eta_N m_l} - 2d\min\{0, m_1, \ldots, m_{(N+2)d}\} \\ &\leq w + 2d\big|\min\{m_1, \ldots, m_{(N+2)d}\}\big| \leq |w| + 3d|\xi_-|.\end{aligned}$$

In conclusion, by comparing (2.22) with (2.21), we realize that for each $\epsilon > 0$, there exists $p > 0$ such that

$$\mathbb{P}_{\eta_N, N+2}\left[\left|\frac{W_{N+2}}{N+2} - w\right| < \epsilon\right] \geq p$$

for all sufficiently large N. This bound proves (2.19).

III. PROOF OF THEOREM 1.2

We know that the entropy production Ne_N is the quadratic functional W_N corresponding to the matrices $L := I - \Sigma_o^{-1} - S^\top S$, $U := 0$, $V := S - S^\top$, and $R := \Sigma_o^{-1} + S^\top S - I$. Part 1 of the theorem then follows from Theorem 1.1 with the rate function I that maps $w \in \mathbb{R}$ in $\sup_{\lambda\in(\lambda_-,\lambda_+)}\{w\lambda - \varphi(\lambda)\}$. It remains to verify the Gallavotti–Cohen symmetry stated by part 2.

Formula (1.9) shows that the Hermitian matrices $F_\lambda(\theta)$ associated with the entropy production satisfy $F_{-\lambda-1}(\theta) = F_\lambda(2\pi - \theta)$ for all $\lambda \in \mathbb{R}$ and $\theta \in [0, 2\pi]$. According to (1.4) and (1.5), this identity immediately gives $f_{-\lambda-1} = f_\lambda$ for any λ and $\varphi(-\lambda-1) = \varphi(\lambda)$ for any λ such that $f_{-\lambda-1} = f_\lambda > 0$. We shall show in a moment that $\lambda_- = -\lambda_+ - 1$. It follows that for every $w \in \mathbb{R}$,

$$\begin{aligned}I(-w) - w &= \sup_{\lambda\in(\lambda_-,\lambda_+)}\{w(-\lambda-1) - \varphi(\lambda)\} \\ &= \sup_{\lambda\in(-\lambda_+-1,-\lambda_--1)}\{w\lambda - \varphi(-\lambda-1)\} = \sup_{\lambda\in(\lambda_-,\lambda_+)}\{w\lambda - \varphi(\lambda)\} = I(w),\end{aligned}$$

which demonstrates the Gallavotti–Cohen symmetry for I.

Let us verify that $\lambda_- = -\lambda_+ - 1$. Recalling that $\lambda_- = \inf\{\Lambda\}$ and $\lambda_+ = \sup\{\Lambda\}$, with Λ being the set defined by (2.8), it suffices to prove that $-\lambda - 1 \in \Lambda$ whenever $\lambda \in \Lambda$. Fix $\lambda \in \Lambda$. Then, $f_{-\lambda-1} = f_\lambda > 0$, which implies that the matrices $\Phi_{-\lambda-1}(n)$ given by (1.6) and $H_{-\lambda-1}$, $K_{-\lambda-1}$, $\mathcal{L}_{-\lambda-1}$, and $\mathcal{R}_{-\lambda-1}$ introduced by Lemma 1.2 are well-defined. The identity $F_{-\lambda-1}(\theta) = F_\lambda(2\pi-\theta)$ shows that $\Phi_{-\lambda-1}(n) = \Phi_\lambda(-n) = \Phi_\lambda^\dagger(n)$ for all $n \in \mathbb{Z}$. The latter entails that $H_{-\lambda-1}$ and K_λ are related by the law

$$H_{-\lambda-1} = I + [(\lambda+1)S^\top - \lambda S]\Phi_{-\lambda-1}(1) = I + [(\lambda+1)S^\top - \lambda S]\Phi_\lambda^\dagger(1) = K_\lambda^\dagger.$$

This law induces a relationship between the matrices $\mathcal{L}_{-\lambda-1}$ and $\mathcal{R}_\lambda$. In fact,

$$\begin{aligned}\mathcal{L}_{-\lambda-1} &= (\lambda+1)I - \lambda(\Sigma_o^{-1} + S^\top S) - [(\lambda+1)S - \lambda S^\top]\Phi_{-\lambda-1}(0)H_{-\lambda-1}^{-1}[(\lambda+1)S^\top - \lambda S] \\ &= (\lambda+1)I - \lambda(\Sigma_o^{-1} + S^\top S) - [(\lambda+1)S - \lambda S^\top]\Phi_\lambda(0)(K_\lambda^{-1})^\dagger[(\lambda+1)S^\top - \lambda S],\end{aligned}$$

which, by taking adjoint on both the sides and by bearing in mind that $\mathcal{L}_\lambda$ is Hermitian, yields

$$\mathcal{L}_{-\lambda-1} = (\lambda+1)I - \lambda(\Sigma_o^{-1} + S^\top S) - [(\lambda+1)S - \lambda S^\top]K_\lambda^{-1}\Phi_\lambda(0)[(\lambda+1)S^\top - \lambda S] = \mathcal{R}_\lambda.$$

Since $\mathcal{R}_\lambda > 0$ by hypothesis, we obtain $\mathcal{L}_{-\lambda-1} > 0$. By similar arguments, we find that $K_{\lambda-1} = H_\lambda^\dagger$ and $\mathcal{R}_{-\lambda-1} = \mathcal{L}_\lambda > 0$. In conclusion, $f_{-\lambda-1} > 0$, $\mathcal{L}_{-\lambda-1} > 0$, and $\mathcal{R}_{-\lambda-1} > 0$ so that $-\lambda - 1 \in \Lambda$.

ACKNOWLEDGMENTS

The authors are grateful to Giuseppe Gonnella for suggesting the problem of large deviations for the entropy production rate and quadratic functionals of autoregressive models.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Marco Zamparo: Conceptualization (lead); Formal analysis (lead); Investigation (lead); Writing – original draft (lead); Writing – review & editing (lead). **Massimiliano Semeraro**: Formal analysis (supporting); Investigation (supporting); Writing – original draft (supporting); Writing – review & editing (supporting).

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

APPENDIX A: PROOF OF LEMMA 1.1

Fix $N \geq 1$. Let $\mu_N^+ := \mathbb{P}[(X_1,\ldots,X_N) \in \cdot]$ and $\mu_N^- := \mathbb{P}[(X_N,\ldots,X_1) \in \cdot]$ be the probability measures on the Borel sets of $(\mathbb{R}^d)^N$ induced by the Gauss–Markov chain $\{X_n\}_{n\geq 1}$. According to (1.1), μ_N^+ and μ_N^- are the multivariate Gaussian distributions that have densities

$$\frac{d\mu_N^+}{d\ell}(x_1,\ldots,x_N) := \frac{e^{-\frac{1}{2}\langle x_1,\Sigma_o^{-1}x_1\rangle}}{\sqrt{(2\pi)^d \det \Sigma_o}} \prod_{n=2}^N \frac{1}{\sqrt{(2\pi)^d}}\, e^{-\frac{1}{2}\|x_n - Sx_{n-1}\|^2}$$

and

$$\frac{d\mu_N^-}{d\ell}(x_1,\ldots,x_N) := \frac{d\mu_N^+}{d\ell}(x_N,\ldots,x_1) = \frac{e^{-\frac{1}{2}\langle x_N,\Sigma_o^{-1}x_N\rangle}}{\sqrt{(2\pi)^d \det \Sigma_o}} \prod_{n=2}^N \frac{1}{\sqrt{(2\pi)^d}}\, e^{-\frac{1}{2}\|x_{n-1} - Sx_n\|^2}$$

with respect to the Lebesgue measure ℓ. Thus, $\mu_N^+ \ll \mu_N^- \ll \ell$, and standard results about measure theory[40] give for all $(x_1,\ldots,x_N) \in (\mathbb{R}^d)^N$,

$$\begin{aligned}\ln\left[\frac{d\mu_N^+}{d\mu_N^-}(x_1,\ldots,x_N)\right] &= \ln\left[\frac{d\mu_N^+}{d\ell}(x_1,\ldots,x_N)\Big/\frac{d\mu_N^-}{d\ell}(x_1,\ldots,x_N)\right] \\ &= \frac{1}{2}\langle x_1,(I-\Sigma_o^{-1}-S^\top S)x_1\rangle + \frac{1}{2}\langle x_N,(\Sigma_o^{-1}+S^\top S - I)x_N\rangle + \sum_{n=2}^N \langle x_n,(S-S^\top)x_{n-1}\rangle.\end{aligned}$$

APPENDIX B: PROOF OF PROPOSITION 1.1

Clearly, we have

$$\lim_{N\uparrow\infty} \frac{\langle X_1, LX_1\rangle}{N} = 0 \quad \mathbb{P}-\text{ a.s.} \tag{B1}$$

The Markov sequence $\{X_n\}_{n\geq 1}$ is a positive Harris recurrent chain (see Ref. 37, Proposition 12.5.1). Its invariant distribution is a Gaussian distribution with mean zero and covariance matrix Σ_s. Thus, the law of large numbers for Markov chains (see Ref. 37, Theorem 17.1.7) gives

$$\lim_{N\uparrow\infty} \frac{1}{N}\sum_{n=1}^N \langle X_n, AX_n\rangle = \frac{1}{\sqrt{(2\pi)^d\ \det\ \Sigma_s}} \int_{\mathbb{R}^d} \langle x, Ax\rangle\, e^{-\frac{1}{2}\langle x,\Sigma_s^{-1}x\rangle}\, dx = \mathrm{tr}[A\Sigma_s] \quad \mathbb{P}-\text{ a.s.}$$

for each matrix $A \in \mathbb{R}^{d\times d}$. In turn, this limit implies

$$\lim_{N\uparrow\infty} \frac{\langle X_N, AX_N\rangle}{N} = 0 \quad \mathbb{P}-\text{a.s.}$$

This way, we get

$$\lim_{N\uparrow\infty} \frac{1}{N}\sum_{n=1}^{N} \langle X_n, UX_n\rangle = \mathrm{tr}[U\Sigma_s] \quad \mathbb{P}-\text{a.s.} \tag{B2}$$

and

$$\lim_{N\uparrow\infty} \frac{\langle X_N, RX_N\rangle}{N} = 0 \quad \mathbb{P}-\text{a.s.} \tag{B3}$$

In addition, the Markov sequence $\{Y_n\}_{n\geq 1}$ with $Y_1 := \begin{pmatrix} 0 \\ X_1 \end{pmatrix}$ and $Y_n := \begin{pmatrix} X_{n-1} \\ X_n \end{pmatrix}$ for $n \geq 2$ is a positive Harris recurrent chain (see Ref. 37, Proposition 12.5.1). In fact, it satisfies

$$Y_{n+1} = \begin{pmatrix} 0 & I \\ 0 & S \end{pmatrix} Y_n + \begin{pmatrix} 0 \\ I \end{pmatrix} G_n$$

for $n \geq 1$, with the spectral radius of $\begin{pmatrix} 0 & I \\ 0 & S \end{pmatrix}$ being $\rho(S) < 1$ and the matrices $\begin{pmatrix} 0 & I \\ 0 & S \end{pmatrix}$ and $\begin{pmatrix} 0 \\ I \end{pmatrix}$ forming a controllable pair. The invariant distribution of the chain $\{Y_n\}_{n\geq 1}$ is a Gaussian distribution with mean zero and covariance matrix $\begin{pmatrix} \Sigma_s & \Sigma_s S \\ S^\top \Sigma_s & \Sigma_s \end{pmatrix}$. Then, we can appeal to the law of large numbers for Markov chains (see Ref. 37, Theorem 17.1.7) once more to obtain

$$\lim_{N\uparrow\infty} \frac{1}{N}\sum_{n=2}^{N} \langle X_n, VX_{n-1}\rangle = \frac{1}{2}\mathrm{tr}\big[\big(V^\top S + S^\top V\big)\Sigma_s\big] \quad \mathbb{P}-\text{a.s.} \tag{B4}$$

Limits (B1–B4) prove the proposition.

APPENDIX C: PROOF OF LEMMA 2.1

For all $N \geq 1$ and $z = (z_1, \ldots, z_{N+2}) \in (\mathbb{C}^d)^{N+2}$, we have

$$\begin{aligned}
\|\mathsf{Q}_N z\|^2 &= \left\|Az_1 + E^\dagger z_2\right\|^2 + \sum_{n=2}^{N+1} \left\|Ez_{n-1} + Dz_n + E^\dagger z_{n+1}\right\|^2 + \|Ez_{N+1} + Bz_{N+2}\|^2 \\
&\leq (\|A\|\|z_1\| + \|E\|\|z_2\|)^2 + \sum_{n=2}^{N+1} (\|E\|\|z_{n-1}\| + \|D\|\|z_n\| + \|E\|\|z_{n+1}\|)^2 + (\|E\|\|z_{N+1}\| + \|B\|\|z_{N+2}\|)^2 \\
&\leq 2\|A\|^2\|z_1\|^2 + 2\|E\|^2\|z_2\|^2 + 3\sum_{n=2}^{N+1} \left(\|E\|^2\|z_{n-1}\|^2 + \|D\|^2\|z_n\|^2 + \|E\|^2\|z_{n+1}\|^2\right) + 2\|E\|^2\|z_{N+1}\|^2 + 2\|B\|^2\|z_{N+2}\|^2 \\
&\leq \left(2\|A\|^2 + 3\|D\|^2 + 2\|B\|^2 + 6\|E\|^2\right)\|z\|^2.
\end{aligned}$$

APPENDIX D: PROOF OF LEMMA 2.2

Fix $N \geq 1$. Assume that there exists $q > 0$ such that $\langle z, \mathsf{Q}_N z\rangle \geq q\langle z, z\rangle$ for all $z \in (\mathbb{C}^d)^{N+2}$. Bearing in mind (2.2) and by writing z as $(a, t_1, \ldots, t_N, b)$ with $s := (a, b) \in (\mathbb{C}^d)^2$ and $t := (t_1, \ldots, t_N) \in (\mathbb{C}^d)^N$, this condition reads

$$\langle a, Aa\rangle + 2\langle t, \mathsf{C}Ea\rangle + \langle t, \mathsf{T}_N t\rangle + 2\langle b, E\mathsf{R}t\rangle + \langle b, Bb\rangle \geq q\langle a, a\rangle + q\langle t, t\rangle + q\langle b, b\rangle.$$

This way, by setting $a := 0$ and $b := 0$, we find $\langle t, \mathsf{T}_N t\rangle \geq q\langle t,t\rangle$ for any $t \in (\mathbb{C}^d)^N$. This shows, in particular, that T_N is invertible. By setting $t := -\mathsf{T}_N^{-1}\mathsf{C}Ea - \mathsf{T}_N^{-1}\mathsf{R}^\dagger E^\dagger b$, we obtain

$$\langle a, \Big(A - E^\dagger \mathsf{C}^\dagger \mathsf{T}_N^{-1}\mathsf{C}E\Big)a\rangle - 2\langle a, E^\dagger \mathsf{C}^\dagger \mathsf{T}_N^{-1}\mathsf{R}^\dagger E^\dagger b\rangle + \langle b, \Big(B - E\mathsf{R}\mathsf{T}_N^{-1}\mathsf{R}^\dagger E^\dagger\Big)b\rangle \geq q\langle a,a\rangle + q\langle b,b\rangle,$$

that is, $\langle s, \mathsf{S}_N s\rangle \geq q\langle s,s\rangle$ for all $s \in (\mathbb{C}^d)^2$. Part 1 is thus verified.

As far as part 2 is concerned, if T_N is invertible, then we can write down the identity

$$\mathsf{Q}_N = \mathsf{L}^\dagger \begin{pmatrix} A - E^\dagger \mathsf{C}^\dagger \mathsf{T}_N^{-1}\mathsf{C}E & \mathsf{0} & -E^\dagger \mathsf{C}^\dagger \mathsf{T}_N^{-1}\mathsf{R}^\dagger E^\dagger \\ \mathsf{0} & \mathsf{T}_N & \mathsf{0} \\ -E\mathsf{R}\mathsf{T}_N^{-1}\mathsf{C}E & \mathsf{0} & B - E\mathsf{R}\mathsf{T}_N^{-1}\mathsf{R}^\dagger E^\dagger \end{pmatrix}\mathsf{L}, \tag{D1}$$

with

$$\mathsf{L} := \begin{pmatrix} I & \mathsf{0} & 0 \\ \mathsf{T}_N^{-1}\mathsf{C}E & \mathsf{I} & \mathsf{T}_N^{-1}\mathsf{R}^\dagger E^\dagger \\ 0 & \mathsf{0} & I \end{pmatrix} \in \mathsf{BL}_{N+2,N+2}.$$

Since $\det \mathsf{L} = 1$, it follows by permutations of rows and columns that

$$\begin{aligned}\det \mathsf{Q}_N &= \det \begin{pmatrix} A - E^\dagger \mathsf{C}^\dagger \mathsf{T}_N^{-1}\mathsf{C}E & \mathsf{0} & -E^\dagger \mathsf{C}^\dagger \mathsf{T}_N^{-1}\mathsf{R}^\dagger E^\dagger \\ \mathsf{0} & \mathsf{T}_N & \mathsf{0} \\ -E\mathsf{R}\mathsf{T}_N^{-1}\mathsf{C}E & \mathsf{0} & B - E\mathsf{R}\mathsf{T}_N^{-1}\mathsf{R}^\dagger E^\dagger \end{pmatrix} \\ &= \det \begin{pmatrix} \mathsf{T}_N & \mathsf{0} & \mathsf{0} \\ \mathsf{0} & A - E^\dagger \mathsf{C}^\dagger \mathsf{T}_N^{-1}\mathsf{C}E & -E^\dagger \mathsf{C}^\dagger \mathsf{T}_N^{-1}\mathsf{R}^\dagger E^\dagger \\ \mathsf{0} & -E\mathsf{R}\mathsf{T}_N^{-1}\mathsf{C}E & B - E\mathsf{R}\mathsf{T}_N^{-1}\mathsf{R}^\dagger E^\dagger \end{pmatrix} = \det \mathsf{T}_N \cdot \det \mathsf{S}_N.\end{aligned}$$

Moreover, given $z \in (\mathbb{C}^d)^{N+2}$, by writing $Lz = (a, t_1, \ldots, t_N, b)$ with $s := (a,b) \in (\mathbb{C}^d)^2$ and $t := (t_1, \ldots, t_N) \in (\mathbb{C}^d)^N$, we realize from (D1) that $\langle z, \mathsf{Q}_N z\rangle = \langle t, \mathsf{T}_N t\rangle + \langle s, \mathsf{S}_N s\rangle$. Thus, if $\mathsf{T}_N > \mathsf{0}$, $\mathsf{S}_N > \mathsf{0}$, and $z \neq 0$, then we have $\langle z, \mathsf{Q}_N z\rangle > 0$ since L is invertible.

APPENDIX E: PROOF OF LEMMA 2.3

Part 1 is immediate since $r(\mathsf{T}_N)$ is non-increasing with respect to N. In fact, given any $z = (z_1, \ldots, z_N) \in (\mathbb{C}^d)^N$, by setting $\zeta := (z_1, \ldots, z_N, 0) \in (\mathbb{C}^d)^{N+1}$, we see that $r(\mathsf{T}_{N+1})\langle z,z\rangle = r(\mathsf{T}_{N+1})\langle \zeta,\zeta\rangle \leq \langle \zeta, \mathsf{T}_{N+1}\zeta\rangle = \langle z, \mathsf{T}_N z\rangle$. Part 3 is nothing but the Szegö theorem for the determinant of Hermitian block Toeplitz matrices (see Ref. 38, Theorem 7). Let us focus on part 2. Assume that $r(\mathsf{T}_N) \geq t$ for all $N \geq 1$ and pick a positive continuous function φ with period 2π and a vector $u \in \mathbb{C}^d$. Due to the assumed properties of φ, there exists a sequence $\{p_N\}_{N\geq 0}$ of trigonometric polynomials that converges uniformly to $\sqrt{\varphi}$, with p_N having degree N (see Ref. 40, Theorem 4.25). Write $p_N(\theta)$ as $\sum_{n=-N}^{N} c_{N,n} e^{-in\theta}$ for each N and θ. Since $r(\mathsf{T}_{2N+1}) \geq t$, by setting $\zeta_n = z_n := c_{N,N-n+1}e^{\mathrm{i}(N+1)}u$ for $n = 1, \ldots, 2N+1$ in (2.6), we obtain

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi} \langle u, F(\theta)u\rangle\, p_N^2(\theta)\, d\theta &= \frac{1}{2\pi}\int_0^{2\pi} \Big\langle \sum_{n=1}^{2N+1} z_n e^{-in\theta}, F(\theta)\sum_{n=1}^{2N+1} z_n e^{-in\theta}\Big\rangle\, d\theta \\ &= \langle z, \mathsf{T}_{2N+1}z\rangle \geq t\sum_{n=1}^{2N+1}\langle z_n, z_n\rangle = t\langle u,u\rangle\, \frac{1}{2\pi}\int_0^{2\pi} p_N^2(\theta)\, d\theta\end{aligned}$$

for all $N \geq 0$. By sending N to infinity, we get

$$\frac{1}{2\pi}\int_0^{2\pi}\langle u, F(\theta)u\rangle\, \varphi(\theta)\, d\theta \geq t\langle u,u\rangle\, \frac{1}{2\pi}\int_0^{2\pi}\varphi(\theta)\, d\theta.$$

The arbitrariness of φ and u shows that $\langle z, F(\theta)z\rangle \geq t\langle z,z\rangle$ for all $\theta \in [0, 2\pi]$ and $z \in \mathbb{C}^d$.

Conversely, if $r(F(\theta)) \geq t$ for every $\theta \in [0, 2\pi]$, then by invoking (2.6) again, we can write for all $N \geq 1$ and $z = (z_1, \dots, z_N) \in (\mathbb{C}^d)^N$,

$$\langle z, \mathsf{T}_N z\rangle \geq \frac{t}{2\pi}\int_0^{2\pi}\left\langle \sum_{n=1}^N z_n e^{-\mathrm{i}n\theta}, \sum_{n=1}^N z_n e^{-\mathrm{i}n\theta}\right\rangle d\theta = t\sum_{n=1}^N \langle z_n, z_n\rangle.$$

APPENDIX F: PROOF OF LEMMA 2.4

Suppose for a moment that the matrix H is invertible. Then, the matrix K is proved to be invertible by contradiction. In fact, if K is not invertible, then there exists a vector $u \in \mathbb{C}^d$ different from 0 such that $Ku = [I - \Phi(1)E]u = 0$. We must have $Eu \neq 0$, otherwise $u = 0$. Since $HE = EK$, we get $HEu = 0$ with $Eu \neq 0$, which contradicts the assumption that H is invertible.

Let us demonstrate now that the matrix H is invertible. This will prove part 1 of the lemma. We proceed by contradiction. Suppose that there exists a vector $u \in \mathbb{C}^d$ different from 0 such that $Hu = 0$. Pick an arbitrary integer $N \geq 3$, and for $n = 1, \dots, N$, consider the vectors

$$z_n := \Phi(1-n)u = \frac{1}{2\pi}\int_0^{2\pi} F^{-1}(\theta)e^{\mathrm{i}(n-1)\theta}d\theta\, u.$$

We have $z_1 \neq 0$ since $\Phi(0)$ is invertible. We claim that

$$\mathsf{T}_N\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_{N-1} \\ z_N\end{pmatrix} = \begin{pmatrix} Dz_1 + E^\dagger z_2 \\ Ez_1 + Dz_2 + E^\dagger z_3 \\ \vdots \\ Ez_{N-2} + Dz_{N-1} + E^\dagger z_N \\ Ez_{N-1} + Dz_N\end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -E^\dagger\Phi(-N)u\end{pmatrix}. \tag{F1}$$

Indeed, for $n = 2, \dots, N-1$, we have

$$\begin{aligned} Ez_{n-1} + Dz_n + E^\dagger z_{n+1} &= \frac{1}{2\pi}\int_0^{2\pi}\left[Ee^{-\mathrm{i}\theta} + D + E^\dagger e^{\mathrm{i}\theta}\right]F^{-1}(\theta)e^{\mathrm{i}(n-1)\theta}d\theta\, u \\ &= \frac{1}{2\pi}\int_0^{2\pi} e^{\mathrm{i}(n-1)\theta}d\theta u = 0 \end{aligned}$$

and

$$\begin{aligned} Dz_1 + E^\dagger z_2 &= \frac{1}{2\pi}\int_0^{2\pi}\left[D + E^\dagger e^{\mathrm{i}\theta}\right]F^{-1}(\theta)\, d\theta\, u \\ &= \frac{1}{2\pi}\int_0^{2\pi}\left[F(\theta) - Ee^{-\mathrm{i}\theta}\right]F^{-1}(\theta)\, d\theta\, u = [I - E\Phi(1)]u = Hu = 0. \end{aligned}$$

Finally, we see that

$$\begin{aligned} Ez_{N-1} + Dz_N &= \frac{1}{2\pi}\int_0^{2\pi}\left[Ee^{-\mathrm{i}\theta} + D\right]F^{-1}(\theta)e^{\mathrm{i}(N-1)\theta}d\theta\, u \\ &= \frac{1}{2\pi}\int_0^{2\pi}\left[F(\theta) - E^\dagger e^{\mathrm{i}\theta}\right]F^{-1}(\theta)e^{\mathrm{i}(N-1)\theta}d\theta\, u = -E^\dagger\Phi(-N)u. \end{aligned}$$

Due to (F1), it follows from (2.6) with $\zeta = z := (z_1, \dots, z_N) \in (\mathbb{C}^d)^N$ and the hypothesis $t := \inf_{\theta\in[0,2\pi]}\{r(F(\theta))\} > 0$ that

$$-\left\langle E^\dagger\Phi(1-N)u, E^\dagger\Phi(-N)u\right\rangle = \langle z, \mathsf{T}_N z\rangle \geq t\sum_{n=1}^N\langle z_n, z_n\rangle \geq t\langle z_1, z_1\rangle.$$

This bound is absurd since $z_1 \neq 0$ and $\lim_{N\uparrow\infty}\Phi(-N) = 0$ by the Riemann–Lebesgue lemma.

Let us move to part 2. As $\inf_{\theta\in[0,2\pi]}\{r(F(\theta))\} > 0$, T_N is invertible by Lemma 2.3, and we can set

$$\mathsf{T}_N^{-1}\mathsf{C} =: \begin{pmatrix} C_1 \\ \vdots \\ C_N\end{pmatrix} \tag{F2}$$

and

$$\mathsf{R}\mathsf{T}_N^{-1} =: \begin{pmatrix} R_1 & \cdots & R_N \end{pmatrix}. \tag{F3}$$

The matrices C and R were defined in (2.3) and (2.4), respectively. We have $\mathsf{C}^\dagger\mathsf{T}_N^{-1}\mathsf{C} = C_1$ and $\mathsf{R}\mathsf{T}_N^{-1}\mathsf{R}^\dagger = R_N$, which, on the one hand, show that C_1 and R_N are Hermitian and, on the other hand, allow us to write

$$\mathsf{S}_N = \begin{pmatrix} A - E^\dagger C_1 E & -E^\dagger R_1^\dagger E^\dagger \\ -EC_N E & B - ER_N E^\dagger \end{pmatrix}.$$

Let us verify that C_1 approaches the matrix $\Phi(0)H^{-1}$ and R_N approaches the matrix $K^{-1}\Phi(0)$ when N is sent to infinity, whereas C_N and R_1 approach 0. These facts prove part 2 of the lemma.

To begin with, we observe that since $H^\dagger$ and K are non-singular and $\lim_{N\uparrow\infty}\Phi(\pm N) = 0$ by the Riemann–Lebesgue lemma, the matrix

$$\mathsf{Z} := \begin{pmatrix} H^\dagger & -\Phi(N)E \\ -\Phi(-N)E^\dagger & K \end{pmatrix} \in \mathsf{BL}_{2,2} \tag{F4}$$

is invertible if $N > N_o$, with $N_o \geq 2$ being a sufficiently large integer. Pick $N > N_o$. By multiplying (F2) by T_N on the left and (F3) by T_N on the right, we explicitly have

$$\begin{cases} DC_1 + E^\dagger C_2 = I & \text{for } n = 1, \\ EC_{n-1} + DC_n + E^\dagger C_{n+1} = 0 & \text{for } n = 2,\ldots,N-1, \\ EC_{N-1} + DC_N = 0 & \text{for } n = N \end{cases} \tag{F5}$$

and

$$\begin{cases} DR_1^\dagger + E^\dagger R_2^\dagger = 0 & \text{for } n = 1, \\ ER_{n-1}^\dagger + DR_n^\dagger + E^\dagger R_{n+1}^\dagger = 0 & \text{for } n = 2,\ldots,N-1, \\ ER_{N-1}^\dagger + DR_N^\dagger = I & \text{for } n = N. \end{cases} \tag{F6}$$

By multiplying the nth equation in (F5) by $e^{-\mathrm{i}n\theta}$ and then by carrying out the sum over n, we get

$$F(\theta)\sum_{n=1}^N C_n e^{-\mathrm{i}n\theta} = Ie^{-\mathrm{i}\theta} + E^\dagger C_1 + EC_N e^{-\mathrm{i}(N+1)\theta},$$

which gives for $n = 1,\ldots,N$,

$$\begin{aligned} C_n &= \frac{1}{2\pi}\int_0^{2\pi} F^{-1}(\theta)\Big[Ie^{\mathrm{i}(n-1)\theta} + E^\dagger C_1 e^{\mathrm{i}n\theta} + EC_N e^{\mathrm{i}(n-N-1)\theta}\Big]d\theta \\ &= \Phi(1-n) + \Phi(-n)E^\dagger C_1 + \Phi(N-n+1)EC_N. \end{aligned} \tag{F7}$$

Similarly, (F6) shows that for $n = 1,\ldots,N$,

$$R_n^\dagger = \Phi(N-n) + \Phi(-n)E^\dagger R_1^\dagger + \Phi(N-n+1)ER_N^\dagger. \tag{F8}$$

At this point, by setting $n := 1$ and $n := N$ in (F7) and by recalling that $C_1 = C_1^\dagger$, we realize that

$$\mathsf{Z}\begin{pmatrix} C_1^\dagger \\ C_N \end{pmatrix} = \begin{pmatrix} \Phi(0) \\ \Phi(1-N) \end{pmatrix},$$

with Z being the matrix defined in (F4). It follows that

$$\begin{pmatrix} C_1^\dagger \\ C_N \end{pmatrix} = \mathsf{Z}^{-1}\begin{pmatrix} \Phi(0) \\ \phi(1-N) \end{pmatrix}$$

as Z is invertible for $N > N_o$. Similarly, (F8) for $n := 1$ and $n := N$ and the fact that $R_N^\dagger = R_N$ yield

$$\begin{pmatrix} R_1^\dagger \\ R_N \end{pmatrix} = Z^{-1}\begin{pmatrix} \Phi(N-1) \\ \Phi(0) \end{pmatrix}.$$

This way, the Riemann–Lebesgue lemma entails that C_1 approaches $\Phi(0)H^{-1}$ and R_N approaches $K^{-1}\Phi(0)$ when N is sent to infinity, whereas C_N and R_1 approach 0.

APPENDIX G: PROOF OF LEMMA 2.6

As the spectral radius $\rho(S)$ of S is smaller than 1 by hypothesis, Gelfand's formula for spectral radii gives $\lim_{n\uparrow\infty}\|S^n\|^{\frac{1}{n}} = \rho(S) < 1$. Then, there exist $s \in (0,1)$ and a positive constant c such that $\|S^n\| \le cs^n$ for all $n \ge 0$. Let us show that the lemma holds with $\sigma := [1 \wedge r(\Sigma_o^{-1})](1-s)^2c^{-2}$, which is positive since, obviously, $r(\Sigma_o^{-1}) > 0$. Fix $N \ge 1$ and, to begin with, observe that

$$\begin{aligned} r(\Sigma_N^{-1}) &= \inf_{\substack{z\in(\mathbb{C}^d)^{N+2} \\ z\neq 0}}\left\{\frac{\langle z_1, \Sigma_o^{-1}z_1\rangle + \sum_{n=2}^{N+2}\|z_n - Sz_{n-1}\|^2}{\sum_{n=1}^{N+2}\langle z_n, z_n\rangle}\right\} \\ &\ge 1 \wedge r(\Sigma_o^{-1}) \inf_{\substack{z\in(\mathbb{C}^d)^{N+2} \\ z\neq 0}}\left\{\frac{\|z_1\|^2 + \sum_{n=2}^{N+2}\|z_n - Sz_{n-1}\|^2}{\sum_{n=1}^{N+2}\langle z_n, z_n\rangle}\right\}. \end{aligned}$$

Since for any $z = (z_1, \ldots, z_{N+2}) \in (\mathbb{C}^d)^{N+2}$, there exists $\zeta = (\zeta_1, \ldots, \zeta_{N+2}) \in (\mathbb{C}^d)^{N+2}$ such that $z_n = \sum_{k=1}^n S^{n-k}\zeta_k$ for each n, this bound yields

$$r(\Sigma_N^{-1}) \ge 1 \wedge r(\Sigma_o^{-1}) \inf_{\substack{\zeta\in(\mathbb{C}^d)^{N+2} \\ \zeta\neq 0}}\left\{\frac{\sum_{n=1}^{N+2}\|\zeta_n\|^2}{\sum_{n=1}^{N+2}\sum_{h=1}^n\sum_{k=1}^n\langle S^{n-h}\zeta_h, S^{n-k}\zeta_k\rangle}\right\}.$$

At this point, it suffices to invoke the Cauchy–Schwarz inequality to conclude that for every $\zeta = (\zeta_1, \ldots, \zeta_{N+2}) \in (\mathbb{C}^d)^{N+2}$,

$$\begin{aligned} \sum_{n=1}^{N+2}\sum_{h=1}^n\sum_{k=1}^n\langle S^{n-h}\zeta_h, S^{n-k}\zeta_k\rangle &\le \sum_{n=1}^{N+2}\sum_{h=1}^n\sum_{k=1}^n\|S^{n-h}\zeta_h\|\|S^{n-k}\zeta_k\| \\ &\le c^2\sum_{n=1}^{N+2}\sum_{h=1}^n\sum_{k=1}^n s^{2n-h-k}\|\zeta_h\|\|\zeta_k\| \\ &\le \frac{c^2}{2}\sum_{n=1}^{N+2}\sum_{h=1}^n\sum_{k=1}^n s^{2n-h-k}\big(\|\zeta_h\|^2 + \|\zeta_k\|^2\big) \\ &\le \frac{c^2}{(1-s)^2}\sum_{k=1}^{N+2}\|\zeta_k\|^2. \end{aligned}$$

REFERENCES

[1] J. Kurchan, "Fluctuation theorem for stochastic dynamics," J. Phys. A: Math. Gen. **31**, 3719–3729 (1998).
[2] J. L. Lebowitz and H. Spohn, "A Gallavotti-Cohen-type symmetry in the large deviation functional for stochastic dynamics," J. Stat. Phys. **95**, 333–365 (1999).
[3] C. Maes, "The fluctuation theorem as a Gibbs property," J. Stat. Phys. **95**, 367–392 (1999).
[4] G. Gallavotti and E. G. D. Cohen, "Dynamical ensembles in stationary states," J. Stat. Phys. **80**, 931–970 (1995).
[5] D. J. Searles and D. J. Evans, "Fluctuation theorem for stochastic systems," Phys. Rev. E **60**, 159–164 (1999).
[6] C. Maes, F. Redig, and A. V. Moffaert, "On the definition of entropy production, via examples," J. Math. Phys. **41**, 1528–1554 (2000).
[7] C. Maes and K. Netočný, "Time-reversal and entropy," J. Stat. Phys. **110**, 269–310 (2003).
[8] D. Q. Jiang, M. Qian, and M. P. Qian, *Mathematical Theory of Nonequilibrium Steady States* (Springer, Berlin, 2004).
[9] J. C. Reid, E. M. Sevick, and D. J. Evans, "A unified description of two theorems in non-equilibrium statistical mechanics: The fluctuation theorem and the work relation," Europhys. Lett. **72**, 726–732 (2005).
[10] H. Ge and D. Q. Jiang, "The transient fluctuation theorem of sample entropy production for general stochastic processes," J. Phys. A: Math. Theor. **40**, 713–723 (2007).
[11] T. Tomé and M. J. de Oliveira, "Entropy production in nonequilibrium systems at stationary states," Phys. Rev. Lett. **108**, 020601 (2012).
[12] G. T. Landi, T. Tomé, and M. J. de Oliveira, "Entropy production in linear Langevin systems," J. Phys. A: Math. Theor. **46**, 395001 (2013).
[13] R. Wang and L. Xu, "Asymptotics of the entropy production rate for d-dimensional Ornstein–Uhlenbeck processes," J. Stat. Phys. **160**, 1336–1353 (2015).
[14] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, "Probability of second law violation in steady flows," Phys. Rev. Lett. **71**, 2401–2404 (1993).
[15] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed. (Springer, New York, 1998).

[16]F. den Hollander, *Large Deviations* (American Mathematical Society, Providence, 2000).
[17]H. Ōsawa, "Reversibility of first-order autoregressive processes," Stoch. Process. Appl. **28**, 61–69 (1988).
[18]D.-Q. Jiang, M. Qian, and F.-X. Zhang, "Entropy production fluctuations of finite Markov chains," J. Math. Phys. **44**, 4176 (2003).
[19]V. Jakšić, C. A. Pillet, and A. Shirikyan, "Entropic fluctuations in Gaussian dynamical systems," Rep. Math. Phys. **77**, 335–376 (2016).
[20]V. Jakšić, C. A. Pillet, and A. Shirikyan, "Entropic fluctuations in thermally driven harmonic networks," J. Stat. Phys. **166**, 926–1015 (2017).
[21]A. Budhiraja, Y. Chen, and L. Xu, "Large deviations of the entropy production rate for a class of Gaussian processes," J. Math. Phys. **62**, 052702 (2021).
[22]L. Rey-Bellet and L. E. Thomas, "Fluctuations of the entropy production in anharmonic chains," Ann. Henri Poincaré **3**, 483–502 (2002).
[23]L. Bertini and G. Di Gesù, "Small noise asymptotic of the Gallavotti-Cohen functional for diffusion processes," ALEA, Lat. Am. J. Probab. Math. Stat. **12**, 743–763 (2015).
[24]G. Gonnella, M. Semeraro, A. Suma, and M. Zamparo, "Work fluctuations for a harmonically confined active Ornstein–Uhlenbeck particle," (unpublished).
[25]M. D. Donsker and S. R. S. Varadhan, "Large deviations for stationary Gaussian processes," Commun. Math. Phys. **97**, 187–210 (1985).
[26]W. Bryc and A. Dembo, "On large deviations of empirical measures for stationary Gaussian processes," Stoch. Process. Appl. **58**, 23–34 (1995).
[27]G. R. Benitz and J. A. Bucklew, "Large deviation rate calculations for nonlinear detectors in Gaussian noise," IEEE Trans. Inf. Theory **36**, 358–371 (1990).
[28]W. Bryc and A. Dembo, "Large deviations for quadratic functionals of Gaussian processes," J. Theor. Probab. **10**, 307–332 (1997).
[29]B. Bercu, F. Gamboa, and A. Rouault, "Large deviations for quadratic forms of stationary Gaussian processes," Stoch. Process. Appl. **71**, 75–90 (1997).
[30]F. Gamboa, A. Rouault, and M. Zani, "A functional large deviations principle for quadratic forms of Gaussian stationary processes," Stat. Probab. Lett. **43**, 299–308 (1999).
[31]S. Ihara, "Large deviation theorems for Gaussian processes and their applications in information theory," Acta Appl. Math. **63**, 165–174 (2000).
[32]Y. Kakizawa, "Moderate deviations for quadratic forms in Gaussian stationary processes," J. Multivar. Anal. **98**, 992–1017 (2007).
[33]B. Bercu, F. Gamboa, and M. Lavielle, "Sharp large deviations for Gaussian quadratic forms with applications," ESAIM Probab. Stat. **4**, 1–24 (2000).
[34]M. Zani, "Sample path large deviations for squares of stationary Gaussian processes," Theory Probab. Appl. **57**, 347–357 (2013).
[35]D. Florens-Landais and H. Pham, "Large deviations in estimation of an Ornstein-Uhlenbeck model," J. Appl. Probab. **36**, 60–77 (1999).
[36]W. Bryc and W. Smolenski, "On the large deviation principle for a quadratic functional of the autoregressive process," Stat. Probab. Lett. **17**, 281–285 (1993).
[37]S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability* (Springer, London, 1993).
[38]J. Gutiérrez-Gutiérrez and P. M. Crespo, "Asymptotically equivalent sequences of matrices and Hermitian block Toeplitz matrices with continuous symbols: Applications to MIMO systems," IEEE Trans. Inf. Theory **54**, 5671–5680 (2008).
[39]R. T. Rockafellar, *Convex Analysis* (Princeton University Press, Princeton, 1970).
[40]W. Rudin, *Real and Complex Analysis*, 3rd ed. (McGraw-Hill, New York, 1987).