Large deviations for quadratic functionals of Gaussian processes

Włodzimierz Bryc*
Department of Mathematics
University of Cincinnati
Cincinnati, OH 45 221
bryc@uc.edu

Amir Dembo†
Department of Mathematics
and Department of Statistics
Stanford University
Stanford, CA 94 305
amir@playfair.stanford.edu

April 9, 1993
Revised September 28, 1993

Abstract

The Large Deviation Principle is derived for several unbounded additive functionals of centered stationary Gaussian processes. For example, the rate function corresponding to $\frac{1}{T}\int_0^T X_t^2 dt$ is the Fenchel-Legendre transform of $L(y) = -\frac{1}{4\pi}\int_{-\infty}^{\infty} \log(1-4\pi y f(s))ds$, where X_t is a continuous time process with the bounded spectral density $f(s)$. The spectral density condition for quadratic functionals is weaker than known sufficient conditions for bounded continuous functionals. Similar results in the discrete-time version are obtained for the energy of multivariate Gaussian processes and for the sums of $p < 2$ powers.

1 Introduction

Let E be a separable Banach space. Throughout most of the paper $E = R$, except in Proposition 2.2, where $E = R^2$, and in Proposition 2.3, where $E = R^{d+1}$.

Suppose $\mathbf{S}_n, n > 0$, are E-valued random variables. We shall say that $\{n^{-1}\mathbf{S}_n\}$ satisfies the Large Deviation Principle (LDP), if there is a lower semicontinuous rate function $I : E \to [0, \infty]$, with compact level sets $I^{-1}([0, a])$ for all $a > 0$, and such that

$$\liminf_{n\to\infty} n^{-1} \log P(n^{-1}\mathbf{S}_n \in A) \geq -\inf_{x\in A} I(x)$$

for all open subsets $A \subset E$;

$$\limsup_{n\to\infty} n^{-1} \log P(n^{-1}\mathbf{S}_n \in A) \leq -\inf_{x\in A} I(x)$$

for all closed subsets $A \subset E$.

We shall work with the continuous indices n (which below are denoted by T rather than n) as well as with the discrete $n = 1, 2, \ldots$; in Section 2.3 we shall also consider other normalizations.

*Partially supported by C.P. Taft Memorial Fund
†Partially supported by NSF DMS92-09712 grant and by a US-ISRAEL BSF grant.
Key Words: large deviations, quadratic additive functionals, Gaussian processes.
AMS (1991) Subject Classification: 60F10

For a general stationary process X_j, the LDP for the empirical measures (i.e., in the discrete time setup corresponding to $\mathbf{S}_n = \sum_{j=1}^n \delta_{X_j}$) and the related question for bounded additive functionals (i.e., $\mathbf{S}_n = \sum_{j=1}^n F(X_j)$, with bounded $F(\cdot)$) have been studied by a number of authors under some restriction on dependence; see [10, Section 6.4] for a sample of results, and [10, Section 6.9 page 280] for relevant references. Gaussian processes were studied in [13], LDP for Gaussian fields is given in [21], see also [12] for an interesting case.

Large deviations for general unbounded additive functionals of Markov chains under minimal assumptions were studied e.g. in [19].

Quadratic forms in Gaussian random variables have been studied by various asymptotic methods e.g. in statistical and electrical engineering literature; for an early paper using the saddle point method to approximate the distribution for a fixed number of variables, see [18], see also [17]. There is also a number of papers on the Central Limit Theorem (CLT), see e.g. [1], [22] and the references therein. Several results directly pertinent to the LDP have appeared: [8] gives a version of the LDP restricted to certain sets and obtained using the Grenander-Szegö method as employed below (and also in [2] and [6]). Their results however deal with quadratic forms in implicit way and without explicit expressions for the rate function; [2] presents the heuristic reasoning that motivated and facilitated much of this paper; in [7], the LDP given as Corollary 2.1 below is stated under an additional technical assumption; in [5] explicit rate function is found for autoregressive AR(1) processes.

In this paper, the LDP is derived for several unbounded additive functionals of stationary centered Gaussian processes that posses spectral density. Of those, quadratic functionals received most attention - for electrical engineering motivation the reader is referred to [6]; motivation from control theory is presented in the introduction to [5]; statistical motivation can be read out from [8].

The following describes the contents of the paper. In Theorem 2.1 we show that $\frac{1}{T}\int_0^T X_t^2 dt$, where X_t is a continuous time process with the bounded spectral density $f(s)$, satisfies the LDP and the rate function is given by the Fenchel-Legendre transform of $L(y) = -\frac{1}{4\pi}\int_{-\infty}^{\infty} \log(1-4\pi y f(s))ds$. In Theorem 2.2 we show the corresponding multivariate discrete-time result. The LDP with normalization of $o(n)$ and the quadratic rate function (corresponding to more moderate deviations) is derived in Theorem 2.3 for unbounded spectral densities (see [10, Section 3.7] for such results in the context of Cramér's theorem). In Corollary 2.2 we analyze $\frac{1}{n}\sum_{j=1}^n |X_j|^p$ for $p<2$. Proposition 2.1 points out the relevance to the CLT. In Section 2.5 we incorporate a non-zero mean in the univariate version of Theorem 2.2, thus deriving the LDP for the empirical variance. In Section 2.6 the LDP is derived for the empirical autocorrelation vector of an i.i.d. process X_j and some counter intuitive results concerning the validity of this LDP when $\{X_j\}$ is an AR(1) process are presented. An approach to higher order expansions is sketched in Section 2.7.

2 Results

This section contains statements of our main results. The proofs are given in Section 3, except for those results that are marked as immediate consequences of other theorems.

2.1 Continuous time

Let $\{X_t\}$ be a real-valued, centered, separable stationary Gaussian process with the covariance $R(t) = E(X_0 X_t)$ and spectral density $f(s)$, i.e., $R(t) = \int_{-\infty}^{\infty} e^{its} f(s)ds$.

Denote $S_T = \int_0^T X_t^2 dt$, $M = \text{ess sup } f(s)$.

Theorem 2.1 *Suppose that $\{X_t\}_{t\geq 0}$ has bounded spectral density function $f(s) \in L_1(R, ds)$. Then $\{\frac{1}{T}S_T\}$ satisfies the LDP with the rate function*

$$I(x) = \sup_{-\infty<y<1/(4\pi M)} \{xy - L(y)\}, \tag{1}$$

where for $y < 1/(4\pi M)$

$$L(y) = -\frac{1}{4\pi}\int_{-\infty}^{\infty} \log(1 - 4\pi y f(s))ds. \tag{2}$$

As an application, suppose that X_t is the Ornstein-Uhlenbeck process, i.e., the stationary solution to $dX_t = -aX_t + \sqrt{a}dW_t$, $a > 0$. The spectral density is $f(s) = \frac{1}{\pi}\frac{a}{a^2+s^2}$ with $M = 1/(\pi a)$. Integrating expression (2) we get $L(y) = \frac{1}{2}a - \frac{1}{2}\sqrt{a^2 - 4ay}$, leading to $I(x) = \frac{a}{4}\left(\sqrt{x} - \frac{1}{\sqrt{x}}\right)^2$ for $x > 0$ and $I(x) = \infty$ otherwise.

2.2 Discrete time

The following result is the finite-dimensional discrete time version of Theorem 2.1.

Theorem 2.2 *Let* $\{\mathbf{X}_k\}_{k=1,2,\ldots}$ *be a centered, stationary Gaussian* R^d*-valued sequence with the spectral density* $\mathbf{F}(s) = [F_{i,j}(s)]$ *such that ess sup* $\|\mathbf{F}(s)\| < \infty$ *(where* $\|\mathbf{F}\|$ *denotes the operator norm associated with the matrix* $\mathbf{F}$*, c.f. (22) below). Then for every nonnegative definite symmetric real matrix* $\mathbf{W}$*,* $\{n^{-1}\sum_{j=1}^n \langle \mathbf{X}_j | \mathbf{W}\mathbf{X}_j\rangle\}$ *satisfies the LDP with the rate function*

$$I(x) = \sup_{-\infty < y < 1/(2M)} \{xy - L(y)\}, \tag{3}$$

where $M =$ *ess sup* $\|\mathbf{W}^{1/2}\mathbf{F}(s)\mathbf{W}^{1/2}\|$ *and for* $y < 1/(2M)$

$$L(y) = -\frac{1}{4\pi}\int_0^{2\pi} \log\det(I - 2y\mathbf{W}\mathbf{F}(s))ds. \tag{4}$$

Remark 2.1 *Clearly, Theorem 2.2 implies that the LDP holds also when* $\mathbf{W}$ *is a nonpositive definite symmetric real matrix. However, in Section 2.6 we give an example of* $\mathbf{W}$ *that is neither positive definite nor negative definite for which* $L(y) = \infty$ *even when all eigenvalues of* $2y\mathbf{W}\mathbf{F}(s)$ *are uniformly (in* s*) strictly less than 1.*

The following special case of Theorem 2.2 is of interest.

Corollary 2.1 *Let* $\{X_k\}_{k=1,2,\ldots}$ *be a real-valued, centered, stationary Gaussian process with bounded spectral density function* $f(s)$*. Then* $\{\frac{1}{n}\sum_{j=1}^n X_j^2\}$ *satisfies the LDP with the rate function of (3) where here* $M =$ *ess sup* $f(s)$ *and*

$$L(y) = -\frac{1}{4\pi}\int_0^{2\pi} \log(1 - 2yf(s))ds. \tag{5}$$

The next corollary follows from Corollary 2.1 if $p = 2$; it follows from [13] by an approximation argument if $p < 2$. The approximation argument does not apply to $p = 2$ case, because the cummulant generating function is unbounded in the latter case.

Corollary 2.2 *Suppose that* $\{X_k\}_{k=1,2,\ldots}$ *has continuous spectral density satisfying* $\int_0^{2\pi}\log f(s)ds > -\infty$*. If* $p \le 2$ *then* $\{\frac{1}{n}\sum_{j=1}^n |X_j|^p\}$ *satisfies the LDP.*

Remark 2.2 *Theorems 2.1 and 2.2 can be also extended to the multivariate index case (Gaussian random fields on* R^k *or* Z^k*). Indeed, [16, Chapter 8] develops the relevant abstract results.*

2.3 Unbounded spectral density

A suitably modified variant of the LDP holds true also when the spectral density is unbounded. Namely, taking $S_n = \sum_{j=1}^{n} X_j^2$ we shall show that for a certain sequence $m_n \to \infty$ random variables $\{m_n(\frac{1}{n}S_n - E(X_1^2))\}$ satisfy the upper and lower bounds with exponent m_n^2/n, i.e.,

$$\begin{aligned} -\inf_{x\in A^\circ} I(x) &\leq \liminf_{n\to\infty} \frac{m_n^2}{n} \log P(m_n(\frac{1}{n}S_n - E(X_1^2)) \in A) \\ &\leq \limsup_{n\to\infty} \frac{m_n^2}{n} \log P(m_n(\frac{1}{n}S_n - E(X_1^2)) \in A) \leq -\inf_{x\in \bar{A}} I(x), \end{aligned} \tag{6}$$

where A° and $\bar{A}$ denote the interior and the closure of a measurable set A respectively.

Theorem 2.3 *Suppose that real-valued, centered stationary Gaussian process $\{X_j\}_{j\geq 1}$ has spectral density function $f(s) \in L_q(ds)$, where $2 < q \leq \infty$. Let $\{m_n\}$be such that $n^{-1/q}m_n \to \infty$ (if $q = \infty$, assume $m_n \to \infty$), and $n^{-1/2}m_n \to 0$. Then $\{m_n(\frac{1}{n}S_n - E(X_1^2))\}$ satisfies the LDP (6) with the rate function*

$$I(x) = \frac{x^2}{2\sigma^2},$$

where

$$\sigma^2 = \frac{1}{\pi}\int_0^{2\pi} f^2(s)ds. \tag{7}$$

Remark 2.3 *With minor changes in the statement and in the proof, Theorem 2.3 holds true both in the multivariate setup of Theorem 2.2 and in the continuous time setup of Theorem 2.1 with the same $I(x)$, but with (7) replaced by*

$$\sigma^2 = \pi^{-1}\int_0^{2\pi} tr\ (\mathbf{F}(s))^2 ds \tag{8}$$

in the former case (taking $\mathbf{W} = \mathbf{I}$) and

$$\sigma^2 = 4\pi\int_{-\infty}^{\infty} f^2(s)ds \tag{9}$$

in the latter.

2.4 Normal convergence

Lemmas 3.3 and 3.6 from the proof of the LDP yield the following CLT. At least in the univariate discrete time setup this result is known, see [1, Theorem 2], [15, Theorem 2] for a direct proof (for non-normal convergence, see [22]). Related results are given in [3, Theorem 5] and the references therein, c.f. also [20, page 58, Theorem 3].

Proposition 2.1 *(i) If $\{X_t\}$ is a real-valued, centered, separable stationary Gaussian process with the spectral density $f(s) \in L_2(R, ds) \cap L_1(R, ds)$, then $\frac{1}{\sqrt{T}}\int_0^T (X_t^2 - E(X_0^2))dt$ is asymptotically normal $N(0, \sigma)$ as $T \to \infty$ with σ^2 given by (9).*

(ii) If $\{\mathbf{X}_k\}_{k=1,2,\ldots}$ is a centered, stationary Gaussian R^d-valued sequence with the spectral density $\mathbf{F}(s) = [F_{i,j}(s)]$, such that $tr\ (\mathbf{F}(s))^2$ is integrable, then

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\langle \mathbf{X}_i|\mathbf{X}_i\rangle - E(\langle \mathbf{X}_1|\mathbf{X}_1\rangle))$$

is asymptotically normal $N(0, \sigma)$ as $n \to \infty$ with σ^2 given by (8).

2.5 Non-centered processes and the LDP for the empirical variance

Many of the results presented above carry over to the case of non-centered stationary Gaussian processes by application of the contraction principle. For concreteness, consider the setup of Corollary 2.1, i.e. let $\{X_j\}$ be a real-valued centered stationary Gaussian process.

Proposition 2.2 *Suppose that spectral density $f(\cdot)$ is differentiable. Let $\mathbf{S}_n = [\sum_{j=1}^n X_j, \sum_{j=1}^n X_j^2]'$. Then $\{n^{-1}\mathbf{S}_n\}$ satisfies the LDP (in R^2) with the rate function*

$$J(x_1, x_2) = I(x_2 - x_1^2) + \frac{x_1^2}{2f(0)} , \tag{10}$$

where $I(\cdot)$ is the rate function given by (3) and (5), and if $f(0) = 0$ then $J(x_1, x_2) = \infty$ for $x_1 \neq 0$ while $J(0, x_2) = I(x_2)$.

Applying the contraction principle (see [10, Theorem 4.2.1]) with respect to the continuous function $g(x_1, x_2) = x_2 + 2x_1\mu + \mu^2 : R^2 \to R$, we see that for a non-centered process $Y_j = X_j + \mu$, the sequence $\{n^{-1}\sum_{j=1}^n Y_j^2\}$ satisfies the LDP (in R) with rate function

$$\tilde{J}(z) = \inf_{\{(x_1,x_2): z = g(x_1,x_2)\}} J(x_1, x_2) = \sup_{y < 1/(2M)} \{zy - \frac{\mu^2 y}{1 - 2yf(0)} - L(y)\} ,$$

where $M = \text{ess sup } f(s)$ and $L(y)$ given by (5), compare also [2, page 361]. Similarly, applying the contraction principle with respect to the continuous function $h(x_1, x_2) = x_2 - x_1^2$ results with the empirical variance of $\{X_j\}_{j=1}^n$ satisfying the LDP with the rate function $I(\cdot)$ given by (3) and (5) (i.e. the *same* rate as for $\{n^{-1}\sum_{j=1}^n X_j^2\}$).

2.6 The empirical autocorrelation vector

For $j \geq 0$, let $S_n^{(j)} = \sum_{k=1}^{n-j} X_k X_{k+j}$. Then $n^{-1}S_n^{(j)}$ is the j-th empirical autocorrelation based on the sample of size n. For fixed $d \geq 1$ let $\mathbf{S}_n = [S_n^{(0)}, \ldots, S_n^{(d)}] \in R^{d+1}$. If $f(\cdot)$ is the spectral density of $\{X_j\}$, denote

$$\mathbf{f}(s) = [f(s), f(s)\cos s, \ldots, f(s)\cos sd]' \in R^{d+1}.$$

Proposition 2.3 *Suppose that $\{X_k\}_{k=1,2,\ldots}$ are i.i.d. N(0,1) random variables. Then $\{\frac{1}{n}\mathbf{S}_n\}$ satisfies the LDP with the rate function*

$$I(\mathbf{x}) = \sup\{\langle \mathbf{x}|\mathbf{y}\rangle - L(\mathbf{y}) : \mathbf{y} \in D\},$$

where

$$D = \{\mathbf{y} \in R^{d+1} : \sup_{0 \leq s \leq 2\pi} \langle \mathbf{y}|\mathbf{f}(s)\rangle < 1/2\},$$

and for $\mathbf{y} \in D$

$$L(\mathbf{y}) = -\frac{1}{4\pi}\int_0^{2\pi} \log(1 - 2\langle \mathbf{y}|\mathbf{f}(s)\rangle)ds .$$

Remark 2.4 *The proof of Proposition 2.3 (with the same formula for the rate function) extends to any differentiable spectral density $f(s)$ provided that for all $\mathbf{y} \in D$*

$$\limsup_{n\to\infty} n^{-1} \log E(\exp(\langle \mathbf{y}|\mathbf{S}_n\rangle)) < \infty . \tag{11}$$

However, the example below shows that for $d = 1$ and for every AR(1) process with $0 < |a| < 1$, (11) is false for some $\mathbf{y} \in D$. Hence, in these cases even if $\{\frac{1}{n}\mathbf{S}_n\}$ satisfies the LDP, the rate function cannot be given by the expression as in Proposition 2.3.

Example 2.1 *Let X_k be an AR(1) process (with $\beta_0 = 1$, $\beta_1 = 0$ and $0 < |a| < 1$) corresponding to $r_i = E[X_0 X_i] = a^i/(1-a^2)$ for $i = 0, 1, \ldots$ and $f(s) = 1/(1 + a^2 - 2a\cos s)$. Therefore $\mathbf{y} = \lambda[1+a^2, -2a]' \in D$ for every $\lambda < 1/2$. Let $\mathbf{R}_n$ denote the covariance matrix of $\mathbf{X} = [X_1, \ldots, X_n]'$ and let $\mathbf{Y}_n$ be the $n \times n$ symmetric Toeplitz matrix corresponding to $y_0 = \lambda(1+a^2)$, $y_1 = -\lambda a$ and $y_i = 0$ for all $1 < i \le n-1$. Since $\mathbf{R}_n^{-1}[r_0, \ldots, r_{n-1}]' = [1, 0, \ldots, 0]'$, we have for $\lambda > (1-a^2)/2$ and all n large enough*

$$\langle [r_0, \ldots, r_{n-1}] | (\mathbf{R}_n^{-1} - 2\mathbf{Y}_n)[r_0, \ldots, r_{n-1}]' \rangle = r_0 - 2\lambda(1+a^2)\sum_{i=0}^{n-1} r_i^2 + 4\lambda a \sum_{i=0}^{n-2} r_i r_{i+1} < 0 ,$$

implying that $E(\exp(\lambda(1+a^2)S_n^{(0)} - 2\lambda a S_n^{(1)})) = \infty$ (see Lemma 3.1).

Note that the above expression is related to Theorem 2.2. Indeed,

$$\lambda(1+a^2)(S_n^{(0)} - \gamma X_n^2 - (1-\gamma)X_1^2) - 2\lambda a S_n^{(1)} = \sum_{j=1}^{n-1} \langle \mathbf{X}_j | \mathbf{W}_\gamma \mathbf{X}_j \rangle$$

where $\mathbf{X}_j = [X_j, X_{j+1}]' \in R^2$ and

$$\mathbf{W}_\gamma = \lambda \begin{bmatrix} \gamma(1+a^2) & -a \\ -a & (1-\gamma)(1+a^2) \end{bmatrix} .$$

Considering $\lambda \ge 0$, $\mathbf{W}_\gamma$ is nonnegative definite iff $\gamma \in [a^2/(1+a^2), 1/(1+a^2)]$. For this range of γ it follows by applying Lemma 3.6 to $\mathbf{Y}_j = \mathbf{W}_\gamma^{1/2}\mathbf{X}_j$ that for all $\lambda < 1/2$,

$$\lim_{n\to\infty} n^{-1} \log E(\exp(\lambda(1+a^2)(S_n^{(0)} - \gamma X_n^2 - (1-\gamma)X_1^2) - 2\lambda a S_n^{(1)})) = -\frac{1}{2}\log(1-2\lambda) . \qquad (12)$$

It can also be verified that for every $\gamma > 1/(1+a^2)$ the left side of (12) is infinite for some $\lambda \in (0, 1/2)$, while the eigenvalues of $\mathbf{W}_\gamma \mathbf{F}(s)$ (which are 0 and λ) are independent of γ.

Remark 2.5 *The example shows that the large deviations of the empirical autocorrelation vector are sensitive to boundary effects (the choice of γ), and that Theorem 2.2 does not extend to matrices $\mathbf{W}$ which are neither nonnegative definite nor nonpositive definite.*

2.7 Exact asymptotic

The following result comes essentially form [16, page 76]. Together with saddle point approximation, it can be used to find higher order asymptotic expansions for probabilities of "regular enough" sets in Corollary 2.1. We do not pursue this possibility here.

Corollary 2.3 *Suppose $\{X_k\}_{k\ge1}$ is a centered, real-valued stationary Gaussian sequence with bounded spectral density $f(s)$ and $M = \text{ess sup } f(s)$. Let $S_n = \sum_{k=1}^n X_k^2$ and $L(y)$ be defined by (5). Then for all $y < 1/(2M)$ the sequence $\{\exp(-nL(y))E(\exp(yS_n))\}$ is monotonically nonincreasing. If in addition $f(s)$ is differentiable and for some $\alpha > 0$ the function $f'(s)$ is uniformly Lipschitz continuous with exponent α then*

$$\lim_{n\to\infty} \exp(-nL(y))E(\exp(yS_n)) = \exp(L(y) - \frac{1}{2\pi}\int\int_{|z|\le1} |h_y'(z)|^2 d\sigma),$$

where

$$h_y(z) = \frac{1}{4\pi}\int_0^{2\pi} \log(1 - 2yf(s))\frac{1+ze^{-is}}{1-ze^{-is}}ds ,$$

and $\sigma(dz)$ is the surface measure on the unit disc in C.

3 Proofs

We shall need the following well known elementary result.

Lemma 3.1 *Suppose* $\mathbf{X} = [X_1, \ldots, X_n]'$ *is a real valued centered Gaussian vector with the covariance matrix* $\mathbf{R}$ *and let* $\mathbf{M}$ *be a symmetric real valued* $n \times n$*-matrix. Then with* $\lambda_1, \ldots, \lambda_n$ *the eigenvalues of the matrix* $\mathbf{MR}$

$$\log E \exp(z\langle \mathbf{X}|\mathbf{MX}\rangle) = -\frac{1}{2}\sum_{j=1}^{n} \log(1 - 2z\lambda_j)$$

for $z \in C$ *such that* $\max_j\{Re(z)\lambda_j\} < 1/2$*. Furthermore,* $\log E \exp(y\langle \mathbf{X}|\mathbf{MX}\rangle) = \infty$ *for* $y \in R$ *such that* $\max_j\{y\lambda_j\} \geq 1/2$*.*

With $\mathbf{X} = \mathbf{R}^{1/2}\mathbf{Z}$ and $\mathbf{Z}$ a standard multivariate normal, Lemma 3.1 follows by direct integration of the density of $\mathbf{Z}$.

Lemma 3.2 *If* $\{Y_j\}$ *are i.i.d. r.v. with mean zero, finite second moment and positive probability density function at 0, then for each* $\theta > 0$ *there is* $\delta > 0$ *such that*

$$\inf\{P(|\sum_{i=1}^{\infty} k_i Y_i| < \theta) : \sum_{i=1}^{\infty} |k_i| \leq 1\} \geq \delta.$$

Proof: Denote $\sigma^2 = E(Y^2)$ and fix the sequence $\{k_i\}$. Without loss of generality, we may assume that $|k_i| \geq |k_{i+1}|$ for all $i \geq 1$. Note that then the condition $\sum_j |k_j| \leq 1$ implies that $|k_j| \leq 1/j$ for all $j \geq 1$. Consequently, for every $r \geq 1$ by Chebyshev's inequality we have

$$P(|\sum_{i=r+1}^{\infty} k_i Y_i| < \theta) \geq 1 - \frac{\sigma^2}{\theta^2}\sum_{j=r+1}^{\infty}\frac{1}{j^2}. \tag{13}$$

Note that one can find $r_0 = r_0(\theta)$ such that the right hand side of (13) is strictly positive. Choose now such $r_0(\theta/2)$. By independence we have

$$P(|\sum_{i=1}^{\infty} k_i Y_i| < \theta) \geq P(|\sum_{i=1}^{r_0} k_i Y_i| < \theta/2) P(|\sum_{i=r_0+1}^{\infty} k_i Y_i| < \theta/2)$$

and, since $|k_i| \leq 1$, using (13) we get

$$\begin{aligned} P(|\sum_{i=1}^{\infty} k_i Y_i| < \theta) &\geq P(\max_{1 \leq i \leq r_0} |Y_i| < \theta/(2r_0)) P(|\sum_{i=r_0+1}^{\infty} k_i Y_i| < \theta/2) \\ &\geq P(|Y_1| < \theta/(2r_0))^{r_0}\left(1 - \frac{4\sigma^2}{\theta^2}\sum_{j=r_0+1}^{\infty}\frac{1}{j^2}\right) =: \delta. \end{aligned}$$

This ends the proof with $\delta > 0$ as defined above. ■

3.1 Proof of Theorem 2.1

For complex z with $Re(z) < \frac{1}{4\pi M}$, let $L_T(z) = \log E(\exp(zS_T))$.

The following Lemma was motivated by a heuristic argument in [2].

Lemma 3.3 *Under the assumptions of Theorem 2.1, for $Re(z) < \frac{1}{4\pi M}$ we have*

$$\lim_{T\to\infty} \frac{1}{T} L_T(z) = -\frac{1}{4\pi}\int_{-\infty}^{\infty} \log(1-4\pi z f(s))ds.$$

Proof: For $T > 0$, denote by $\lambda_j = \lambda_j(T)$ the eigenvalues of

$$\int_0^T R(t-s)g(s)ds = \lambda g(t) \in L_2([0,T]) \tag{14}$$

and let $e_j = e_j(t) \in L_2([0,T], dt)$ be the corresponding orthonormal eigenfunctions. Since by Mercer's theorem, $R(t-s) = \sum_j \lambda_j e_j(t) e_j(s)$ with positive and summable eigenvalues $\{\lambda_j\}$, we have the Karhunen-Loéve expansion $X_t = \sum_j \sqrt{\lambda_j}\gamma_j e_j(t)$, where γ_j are i.i.d. N(0,1). Note that

$$\sup_j \lambda_j = \sup_{g\in L_2, \|g\|=1} \int_0^T g(t)dt \int_0^T g(u)du \int_{-\infty}^{\infty} e^{i(t-u)s} f(s)ds.$$

Since for $T < \infty$ each square-integrable $g(\cdot)$ is integrable, we may switch the order of integration, which gives

$$\sup_j \lambda_j \le M \int_{-\infty}^{\infty} |\int_0^T g(t)e^{its}dt|^2 ds = 2\pi M, \tag{15}$$

where the last equality is by Plancherel's theorem. Therefore $Re(z) < 1/(4\pi M) \le 1/(2\lambda_j)$ and

$$\frac{1}{T} L_T(z) = 1/T \log E[\exp(zS_T)] = -1/(2T)\sum_{j=1}^{\infty} \log(1-2z\lambda_j) = -\frac{1}{2}\int_0^{2\pi M} \log(1-2zx)\mu_T(dx), \tag{16}$$

where $\mu_T(dx) := 1/T\sum_j \delta_{\lambda_j}(dx)$ denotes the distribution of the eigenvalues on $[0, 2\pi M]$. Fix z and choose $\delta > 0$ such that $2|z|\delta < 1$ and such that $\{s : 2\pi f(s) = \delta\}$ is of Lebesgue measure zero. By [16, page 139] for $k = 1, 2, \ldots$ we have

$$\lim_{T\to\infty} \int_0^{2\pi M} x^k \mu_T(dx) = (2\pi)^{k-1}\int_{-\infty}^{\infty} f^k(s)ds, \tag{17}$$

and also for every bounded continuous $F(\cdot)$

$$\lim_{T\to\infty} \int_\delta^{2\pi M} F(x)\mu_T(dx) = \frac{1}{2\pi}\int_{\{s:2\pi f(s)\ge\delta\}} F(2\pi f(s))ds. \tag{18}$$

Let $P_k(x)$ be the k-th Taylor polynomial for $x \mapsto \log(1-2zx)$. Notice that from (17) and (18), for each fixed k we get

$$\int_0^\delta P_k(x)\mu_T(dx) \to \frac{1}{2\pi}\int_{\{s:2\pi f(s)\le\delta\}} P_k(2\pi f(s))ds. \tag{19}$$

Clearly, for $0 \le x \le \delta$ we have

$$|P_k(x) - \log(1-2zx)| = |\sum_{j=k+1}^{\infty} (2zx)^j/j| < \frac{1}{k}\frac{(2x|z|)^{k+1}}{1-2|z|\delta} \le \frac{1}{k}\frac{2x|z|}{1-2|z|\delta}.$$

Given $\epsilon > 0$ choose $k > 2|z|(1-2|z|\delta)^{-1}\epsilon^{-1}$. Then by (19) choose $T_0 = T_0(k)$ such that for all $T > T_0$ we have

$$\Delta_1 := |\int_0^\delta P_k(x)\mu_T(dx) - \frac{1}{2\pi}\int_{\{s:2\pi f(s)\le\delta\}} P_k(2\pi f(s))ds| < \epsilon$$

and by (17) (with k=1)

$$\int_0^{2\pi M} x\mu_T(dx) < 2R(0).$$

Enlarging T_0 if necessary, by (18) we may also ensure

$$\Delta_2 := |\int_\delta^{2\pi M} \log(1-2zx)\mu_T(dx) - \frac{1}{2\pi}\int_{\{s:2\pi f(s)\geq\delta\}} \log(1-4\pi z f(s))ds| < \epsilon$$

for all $T > T_0$. Therefore for all $T > T_0$ we have

$$\begin{aligned} |\int_0^{2\pi M} \log(1-2zx)\mu_T(dx) - \frac{1}{2\pi}\int_{-\infty}^{\infty} \log(1-4\pi z f(s))ds| \quad &\leq \\ \Delta_1 + \Delta_2 + \epsilon\int_0^{2\pi M} x\mu_T(dx) + \epsilon\int_{-\infty}^{\infty} f(s)ds \quad &< \quad (2+3R(0))\epsilon. \blacksquare \end{aligned}$$

Remark 3.1 *By the induced convergence for analytic functions, from Lemma 3.3 it follows that*

$$T^{-1}\frac{d}{dy}L_T(y) \to \frac{d}{dy}L(y) = \int_{-\infty}^{\infty} \frac{f(s)}{1-4\pi y f(s)} ds$$

for all $y < \frac{1}{4\pi M}$ *(this can be also verified directly using [16, page 139]).*

Remark 3.2 *Let* $\lambda_1(T)$ *be the maximal eigenvalue of (14). Then* $\lambda_1(T) \leq 2\pi M$ *by (15), and therefore by [16, page 139] one has* $\lambda_1(T) \to 2\pi M$ *as* $T \to \infty$.

Proof of Theorem 2.1: By Remark 3.2 and Lemma 3.1 it follows that $L(y) = \lim_{T\to\infty} T^{-1}L_T(y)$ is infinite for $y > 1/(4\pi M)$, and by Lemma 3.3 $L(y)$ exists and given by (2) for all $y < 1/(4\pi M)$. Define $L(1/(4\pi M)) = \lim_{y\nearrow 1/(4\pi M)} L(y)$ (which by monotone convergence coincides with $L(1/(4\pi M))$ of (2)), and note that by the monotonicity of $L_T(y)$ with respect to y

$$\liminf_{T\to\infty,\, y_T\to 1/(4\pi M)} T^{-1}L_T(y_T) \geq L(1/(4\pi M)) \,. \tag{20}$$

If $L(1/(4\pi M)) = \infty$, then the result follows by the Gärtner-Ellis Theorem (see [10, Theorem 2.3.6]), for then (20) holds with equality, and $L(\cdot)$ is steep, i.e., $\lim_{y\nearrow 1/(4\pi M)} \frac{d}{dy}L(y) = \infty$. Since in general this is not the case (and it is not even clear that $T^{-1}L_T(1/(4\pi M))$ converges), we follow instead the strategy of parameter dependent change of measure, as outlined in [11]. Indeed, by the monotonicity of $L_T(\cdot)$ it follows that [11, (2.13) and (2.15)] hold. Excluding the trivial case of zero spectral density, since $L'(y) > 0$ is non-decreasing, there is $c > 0$ such that $L'(y) \to c$ as $y \nearrow 1/(4\pi M)$, and examining [11, Proposition 2.14] we see that the LDP with the rate function of (1) holds even for $L(1/(4\pi M)) < \infty$ as soon as $L(\cdot)$ is steep, i.e. $c = \infty$. Turning to deal with $L(\cdot)$ which is not steep, i.e. $c < \infty$, observe that then $I(\cdot)$ of (1) is continuous at $x = c$ and it is simple to check that for $x \geq c$

$$I(x) = \frac{x}{4\pi M} - L(\frac{1}{4\pi M}).$$

Thus, by [11, Proposition 2.14], suffices to show that for all $x > c$ and all $\epsilon > 0$ small enough

$$\liminf_{T\to\infty} \frac{1}{T}\log P(|T^{-1}S_T - x| < \epsilon) \geq -\frac{1}{4\pi M}(x+\epsilon) + L(\frac{1}{4\pi M}), \tag{21}$$

in order to complete the proof of the theorem. To this end, let $\lambda_1(T) \geq \lambda_2(T) \geq \ldots \geq \lambda_n(T) \geq \ldots$ be the eigenvalues of (14) and for $y < 1/(2\lambda_1)$ let

$$k_j(y,T) = \frac{\lambda_j}{T(1-2y\lambda_j)} \,.$$

Since $T^{-1}\frac{d}{dy}L_T(y) = \sum_j k_j(y,T)$ is monotone in y and approaches ∞ as y approaches $1/(2\lambda_1)$, there exists $y_T < 1/(2\lambda_1(T))$ such that $\sum_{j=1}^{\infty} k_j = x$ for $k_j = k_j(y_T, T)$. Moreover, for each fixed $y < 1/(4\pi M)$, by Remark 3.1 $\lim_T T^{-1}\frac{d}{dy}L_T(y) = \frac{d}{dy}L(y) \leq c < x$, while $\limsup_T y_T \leq 1/(4\pi M)$ by Remark 3.2; hence $y_T \to 1/(4\pi M)$. For y_T as above, define the measure Q_T via

$$\frac{dQ_T}{dP} = \exp(y_T S_T - L_T(y_T)),$$

and let V_T denote the r.v. $(T^{-1}S_T - x)$ under measure Q_T. Note that by (16) the Laplace transform of V_T is given by

$$E[e^{sV_T}] = \prod_{i=1}^{\infty} \exp(-sk_i)/\sqrt{1-sk_i},$$

where $k_i = k_i(y_T, T)$. Therefore V_T has the representation

$$V_T = \sum_{j=1}^{\infty} k_j(Z_j^2 - 1)$$

with Z_j i.i.d. normal N(0,1), and by Lemma 3.2 we deduce that $Q_T(|T^{-1}S_T - x| < \epsilon) \geq \delta$ for all $\epsilon > 0$ and some $\delta = \delta(\epsilon) > 0$ which is independent of T. Since $y_T \geq 0$ for all large T,

$$\begin{aligned} T^{-1}\log P(|T^{-1}S_T - x| < \epsilon) &= T^{-1}\log(\int \frac{dP}{dQ_T} 1_{|T^{-1}S_T - x|<\epsilon} dQ_T) \\ &\geq T^{-1}\log Q_T(|T^{-1}S_T - x| < \epsilon) - y_T(x+\epsilon) + T^{-1}L_T(y_T), \end{aligned}$$

and the lower bound (21) follows from (20). ■

3.2 Proof of Theorem 2.2

Throughout this proof we consider R^n, $n \geq 1$ as Hilbert subspaces of ℓ_2 with the inherited norms. For an $n \times n$-matrix $\mathbf{A}$, we consider the usual operator norm

$$||\mathbf{A}|| = \sup_{\mathbf{y} \in R^n \setminus \mathbf{0}} \frac{||\mathbf{A}\mathbf{y}||}{||\mathbf{y}||}, \tag{22}$$

and the Hilbert-Schmidt norm $|\mathbf{A}| = \sqrt{\text{tr}\ (\mathbf{A}\mathbf{A}')}$ (with the usual convention that $\mathbf{A}'$ is the conjugate transpose of the matrix $\mathbf{A}$). It is well known that $|\mathbf{ABC}| \leq ||\mathbf{A}|| \cdot |\mathbf{B}| \cdot ||\mathbf{C}||$, and that $||\mathbf{A}|| \leq |\mathbf{A}|$, see e.g. [14, Section XI.6]. The distribution of the eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$ of $\mathbf{A}$ is the discrete probability measure

$$\mu_n(dx) = n^{-1}\sum_{j=1}^{n} \delta_{\lambda_j}(dx)$$

(either on R or on C, depending on whether $\mathbf{A}$ is symmetric, or not). The following result is known.

Lemma 3.4 ([16, p 105]) *Suppose the $n \times n$ matrices $\mathbf{A}_n$ and $\mathbf{B}_n$ have the distribution of the eigenvalues μ_n and ν_n respectively and assume that*

$$\sup_n(||\mathbf{A}_n|| + ||\mathbf{B}_n||) < \infty, \tag{23}$$

and

$$\lim_{n\to\infty} n^{-1}|\mathbf{A}_n - \mathbf{B}_n|^2 = 0. \tag{24}$$

Then $\lim_{n\to\infty} |\int x^k \mu_n(dx) - \int x^k \nu_n(dx)| = 0$ *for every* $k = 1, 2, \ldots$

Let $\mathbf{R}_n = cov(\mathbf{X}_0, \mathbf{X}_n)$ be the $d \times d$-covariance matrices, and let μ_n be the distribution of the eigenvalues of the block-Toeplitz $nd \times nd$ matrix

$$\mathbf{A}_n = \begin{bmatrix} \mathbf{R}_0 & \mathbf{R}_1 & \dots & \mathbf{R}_{n-1} \\ \mathbf{R}_{-1} & \mathbf{R}_0 & \dots & \mathbf{R}_{n-2} \\ \vdots & & \ddots & \vdots \\ \mathbf{R}_{-(n-1)} & \mathbf{R}_{-(n-2)} & \dots & \mathbf{R}_0 \end{bmatrix}. \tag{25}$$

The asymptotic of μ_n follows by extending the argument of [16, page 113] as follows.

Lemma 3.5 *If $M = ess\ sup\|\mathbf{F}(s)\| < \infty$ then $\sup_n \|\mathbf{A}_n\| \leq M$. Moreover, for any $a < b$ such that $m(s : \lambda_j(s) = a) = m(s : \lambda_j(s) = b) = 0$ for $j = 1, \dots, d$,*

$$\lim_{n\to\infty} \mu_n([a, b]) = (2\pi d)^{-1} \sum_{j=1}^{d} m(s : a < \lambda_j(s) < b) \ , \tag{26}$$

where m is Lebesgue measure on $[0, 2\pi]$ and $\lambda_1(s) \geq \lambda_2(s) \geq \cdots \lambda_d(s) \geq 0$ are the eigenvalues of $\mathbf{F}(s)$ (recall that $\mathbf{F}(s), 0 \leq s \leq 2\pi$, are Hermitian, nonnegative definite matrices).

Proof: For $(n-1)/2 \geq A \geq 1$ let $\widehat{\mathbf{R}}_k = (1 - k/A)\mathbf{R}_k$ for $k = 0, \dots, A$ and $\widehat{\mathbf{R}}_k = 0$ for $k > A$, with $\widehat{\mathbf{R}}_{-k} = \widehat{\mathbf{R}}_k'$. Let $\mathbf{B}_{n,A}$ be the block-Toeplitz $nd \times nd$ matrix constructed as in (25) but with the blocks $\widehat{\mathbf{R}}_k$ instead of $\mathbf{R}_k$. Let $\mathbf{C}_{n,A}$ be the block-circulant matrix associated with $\mathbf{B}_{n,A}$, i.e. using the blocks $\widehat{\mathbf{R}}_{k \bmod n}$ in (25) instead of $\mathbf{R}_k$. Let $\mathbf{F}_A(s) = \sum_{k=-A}^{A} e^{-iks}\widehat{\mathbf{R}}_k$, with $\{\lambda_{j,k}\}_{j=1,\dots,d}$ denoting the eigenvalues of $\mathbf{F}_A(2\pi k/n)$, $k = 0, \dots, n-1$ and $\mathbf{v}_{j,k} \in R^d$ the corresponding eigenvectors. The usual argument for circulant matrices shows that for $j = 1, \dots, d, k = 0, \dots, n-1$ the nd-dimensional vectors

$$(\mathbf{v}_{j,k}, e^{2\pi ik/n}\mathbf{v}_{j,k}, \dots, e^{2\pi ik(n-1)/n}\mathbf{v}_{j,k})$$

are the linearly independent eigenvectors of $\mathbf{C}_{n,A}$ corresponding to the eigenvalues $\lambda_{j,k}$; therefore those are all the eigenvalues of $\mathbf{C}_{n,A}$. Consequently, $\|\mathbf{C}_{n,A}\| \leq \sup_s \|\mathbf{F}_A(s)\|$ and since

$$\mathbf{F}_A(s) = \frac{1}{2\pi} \int_0^{2\pi} \frac{\sin^2(A(s-t)/2)}{A \sin^2((s-t)/2)} \mathbf{F}(t) dt \ ,$$

clearly,

$$\sup_s \|\mathbf{F}_A(s)\| \leq \sup_s \|\mathbf{F}(s)\| \cdot \frac{1}{2\pi} \int_0^{2\pi} \frac{\sin^2(At/2)}{A \sin^2(t/2)} dt = M \ . \tag{27}$$

We turn now to prove that $\|\mathbf{A}_n\| \leq M$ and $\|B_{n,A}\| \leq M$. To this end, fix n, pick $\mathbf{x}_j \in R^d$ and write $\mathbf{x} = (\mathbf{x}_j)$ as a column vector. Then,

$$\begin{aligned} \langle \mathbf{x} | \mathbf{A}_n \mathbf{x} \rangle &= (2\pi)^{-1} \int_0^{2\pi} \langle \sum_{k=1}^{n} e^{-iks}\mathbf{x}_k | \mathbf{F}(s) \sum_{m=1}^{n} e^{ims}\mathbf{x}_m \rangle ds \leq (2\pi)^{-1} \int_0^{2\pi} \|\sum_{k=1}^{n} e^{-iks}\mathbf{x}_k\|^2 \|\mathbf{F}(s)\| ds \\ &\leq \sup_{0\leq s\leq 2\pi} \|\mathbf{F}(s)\| (\frac{1}{2\pi} \int_0^{2\pi} \|\sum_{k=1}^{n} e^{-iks}\mathbf{x}_k\|^2 ds) = M\|\mathbf{x}\|^2 . \end{aligned}$$

By a similar argument we have for $n > A$

$$\langle \mathbf{x} | \mathbf{B}_{n,A}\, \mathbf{x} \rangle = (2\pi)^{-1} \int_0^{2\pi} \langle \sum_{k=1}^{n} e^{-iks}\mathbf{x}_k | \mathbf{F}_A(s) \sum_{m=1}^{n} e^{ims}\mathbf{x}_m \rangle ds \leq \|\mathbf{x}\|^2 \sup_s \|\mathbf{F}_A(s)\| \leq M\|\mathbf{x}\|^2 .$$

This shows that matrices $\mathbf{A}_n$ and $\mathbf{B}_{n,A}$ and $\mathbf{C}_{n,A}$ satisfy (23) for every choice of $A \leq (n-1)/2$.

By applying Parseval's relation elementwise one has

$$\sum_{j=-\infty}^{\infty} |\mathbf{R}_j|^2 = (2\pi)^{-1} \int_0^{2\pi} |\mathbf{F}(s)|^2 ds \le d\,M^2 \ .$$

Since for every $n > A$ we have

$$n^{-1}|\mathbf{A}_n - \mathbf{B}_{n,A}|^2 \le 2\sum_{j=1}^{A} (j/A)^2 |\mathbf{R}_j|^2 + 2 \sum_{j=A+1}^{\infty} |\mathbf{R}_j|^2 \ ,$$

by Kronecker's Lemma it follows that $n^{-1}|\mathbf{A}_n - \mathbf{B}_{n,A}|^2$ can be made arbitrarily small (uniformly in $n > A$) by choosing A large enough. Therefore, by choosing first A large and then n large enough, we can make sure that (24) holds both for $|\mathbf{A}_n - \mathbf{B}_{n,A}|$ and for $|\mathbf{B}_{n,A} - \mathbf{C}_{n,A}|$ since

$$|\mathbf{B}_{n,A} - \mathbf{C}_{n,A}|^2 \le 2A \sum_{j=1}^{A} |\mathbf{R}_j|^2 \le A d\, M^2 \ .$$

Consequently, by Lemma 3.4 the asymptotic of μ_n is the same as the asymptotic of the distribution of the eigenvalues of $\mathbf{C}_{n,A}$ provided we let $n \to \infty$ first and then take $A \to \infty$.

Fix a positive integer ℓ. In view of the continuity of $\mathbf{F}_A(s)$ we have for any fixed $A \ge 1$

$$\lim_{n\to\infty} n^{-1} \sum_{k=1}^{n} \text{ tr } (\mathbf{F}_A(2\pi k/n)^\ell) = \frac{1}{2\pi} \int_0^{2\pi} \text{ tr } (\mathbf{F}_A(s)^\ell) ds \ .$$

Also

$$\begin{aligned} |(2\pi)^{-1} \int_0^{2\pi} \text{ tr } (\mathbf{F}_A(s)^\ell - \mathbf{F}(s)^\ell) ds|^2 &\le d(2\pi)^{-1} \int_0^{2\pi} |\mathbf{F}_A(s)^\ell - \mathbf{F}(s)^\ell|^2 ds \\ &\le d\ell^2 M^{2(\ell-1)} (2\pi)^{-1} \int_0^{2\pi} |\mathbf{F}_A(s) - \mathbf{F}(s)|^2 ds \ , \end{aligned}$$

and since,

$$(2\pi)^{-1} \int_0^{2\pi} |\mathbf{F}_A(s) - \mathbf{F}(s)|^2 ds = 2\sum_{j=1}^{A} (j/A)^2 |\mathbf{R}_j|^2 + 2 \sum_{j=A+1}^{\infty} |\mathbf{R}_j|^2 \ ,$$

we have for $A \to \infty$ that $\int_0^{2\pi}$ tr $(\mathbf{F}_A(s)^\ell - \mathbf{F}(s)^\ell) ds \to 0$, leading to

$$\lim_{A\to\infty} \lim_{n\to\infty} n^{-1} \sum_{k=1}^{n} \text{ tr } (\mathbf{F}_A(2\pi k/n)^\ell) = \frac{1}{2\pi} \int_0^{2\pi} \text{ tr } (\mathbf{F}(s)^\ell) ds \ .$$

With the above holding for every positive integer ℓ, the limit (26) follows by [16, page 105]. ■

Let $S_n = \sum_{j=1}^{n} \langle \mathbf{X}_j | \mathbf{X}_j \rangle$ and for complex z, let $L_n(z) = \log E(\exp(z S_n))$.

Lemma 3.6 *If* $\sup_s \|\mathbf{F}(s)\| = M < \infty$, *then the limit* $\lim_{n\to\infty} \frac{1}{n} L_n(z)$ *exists for every* z *in the half-plane* $Re\ z < \frac{1}{2M}$ *and*

$$\lim_{n\to\infty} \frac{1}{n} L_n(z) = -\frac{1}{4\pi} \int_0^{2\pi} \log\det(I - 2z\mathbf{F}(s)) ds \ . \qquad (28)$$

Remark 3.3 *For* $d = 1$ *this lemma is known, see [6, page 105], or [7, Example 3.1 a)].*

Proof: Clearly,

$$S_n = [\mathbf{X}_1, \ldots, \mathbf{X}_n][\mathbf{X}_1, \ldots, \mathbf{X}_n]'.$$

Therefore by Lemma 3.1, for $Re(z) < 1/(2\max_j \lambda_j)$

$$n^{-1}L_n(z) = -1/(2n)\sum_{j=1}^{nd} \log(1-2z\lambda_j),$$

where $\{\lambda_j\}$ are the eigenvalues of the symmetric nonnegative definite matrix $\mathbf{A}_n$.

Lemma 3.5 implies that $\max_j \lambda_j = \|\mathbf{A}_n\| \leq M$ for all n, and by (26) actually $\|\mathbf{A}_n\| \to M$ as $n \to \infty$. Consequently, (28) follows by applying (26) and observing that

$$n^{-1}L_n(z) = -\frac{d}{2}\int_0^M \log(1-2zx)\mu_n(dx) \, . \blacksquare$$

Remark 3.4 *By the induced convergence for analytic functions, from Lemma 3.6 it follows that for* $y < 1/(2M)$

$$n^{-1}\frac{d}{dy}L_n(y) \to \frac{d}{dy}L(y) = \frac{1}{2\pi}\sum_{j=1}^{d}\int_0^{2\pi} \frac{\lambda_j(s)}{1-2y\lambda_j(s)}ds,$$

where $\lambda_j(s), j = 1, \ldots, d$ *are the (nonnegative) eigenvalues of* $\mathbf{F}(s)$. *(This claim can also be verified directly from (26).)*

Proof of Theorem 2.2: For $\mathbf{W}$ an identity matrix, the proof repeats the reasoning from the proof of Theorem 2.1. Indeed, by Lemma 3.6, $n^{-1}L_n(y)$ converges to $L(y)$ of (4) for $y < 1/(2M)$, while by Lemmas 3.1 and 3.5, for $y > 1/(2M)$

$$L(y) = \lim_{n\to\infty} n^{-1}L_n(y) = \infty \, .$$

Excluding the trivial case of zero spectral density, notice that $L'(y) > 0$ is monotonically increasing for $y < 1/(2M)$, and let $c > 0$ be such that $L'(y) \to c$ as $y \nearrow 1/(2M)$. Define $L(1/(2M)) = \lim_{y\nearrow 1/(2M)} L(y)$. Since [11, (2.13) and (2.15)] hold by the monotonicity of $L_n(\cdot)$, if $L(y)$ is steep, i.e. $c = \infty$, then the LDP with the rate function $I(\cdot)$ of (3) and (4) follows by [11, Proposition 2.14] (even if $n^{-1}L_n(1/(2M))$ fails to converge). If $L(\cdot)$ is not steep then $I(x)$ is continuous at $x = c$ and $I(x) = \frac{x}{2M} - L(\frac{1}{2M})$ for all $x \geq c$. Letting $\{\lambda_j\}$ denote the nonnegative eigenvalues of the matrix $\mathbf{A}_n$, the n-dependent change of measure via $\frac{dQ_n}{dP} = \exp(y_n S_n - L_n(y_n))$ results with $n^{-1}S_n - x$ (under Q_n) having the representation $\sum_{j=1}^{nd} k_j(Z_j^2 - 1)$ with Z_j i.i.d. normal $N(0,1)$ and $k_j = \lambda_j/(n(1-2y_n\lambda_j))$, where $y_n < 1/(2\max_j \lambda_j)$ chosen such that $\sum_{j=1}^{nd} k_j = x$. Since $\max_j\{\lambda_j\} = \|\mathbf{A}_n\| \to M$ as $n \to \infty$ it follows by Remark 3.4, that $\lim_n y_n = 1/(2M)$ and the proof of the large deviations lower bound for $x > c$ is completed by applying Lemma 3.2 (note that $\liminf_n n^{-1}L_n(y_n) \geq L(1/(2M))$). For any $\mathbf{W}$ nonnegative definite symmetric real matrix, we have $\mathbf{W} = \mathbf{W}^{1/2}\mathbf{W}^{1/2}$ with $\mathbf{W}^{1/2}$ also nonnegative symmetric real matrix. Hence $\langle \mathbf{X}_j | \mathbf{W}\mathbf{X}_j \rangle = \langle \mathbf{Y}_j | \mathbf{Y}_j \rangle$ for $j = 1, 2, \ldots$, where $\mathbf{Y}_j = \mathbf{W}^{1/2}\mathbf{X}_j$ is a stationary process of bounded spectral density $\mathbf{W}^{1/2}\mathbf{F}(s)\mathbf{W}^{1/2}$. Therefore, the general case follows by applying the above proof to the process $\{\mathbf{Y}_j\}$. $\blacksquare$

Remark 3.5 *For* $d = 1$, *by Lemma 3.1 and [16, pages 38, 44],* $n^{-1}\log E(\exp((2M)^{-1}\sum_{j=1}^n X_j^2))$ *converges as* $n \to \infty$ *to* $L(1/(2M))$ *of (5). The validity of this result in the general context of Theorem 2.2 is not addressed here.*

3.3 Proof of Theorem 2.3

The proof is based on the Gärtner-Ellis Theorem (c.f. [10, Theorem 2.3.6 and Remark (a)]) used with the normalization $a_n = m_n^2/n \to 0$.

We shall need the following estimate for the maximal eigenvalue of the covariance matrices.

Lemma 3.7 *If $1 \leq q \leq \infty$ then there is $C < \infty$ such that for all $n > 1$ if $\mathbf{A}_n$ is the covariance matrix of $[X_1, \ldots, X_n]'$ then $\|\mathbf{A}_n\| \leq Cn^{\frac{1}{q}}$.*

Proof: Let $\mathbf{x} = [x_1, \ldots, x_n]'$ be such that $\|\mathbf{x}\| = 1$ and $\|\mathbf{A}_n\| = \langle \mathbf{x}|\mathbf{A}_n\mathbf{x}\rangle$. Then, denoting $1/p + 1/q = 1$, we have $\|\mathbf{A}_n\| = \frac{1}{2\pi}\int_0^{2\pi} f(s)|\sum x_j e^{ijs}|^2 ds \leq \|f\|_q (\frac{1}{2\pi}\int_0^{2\pi} |\sum x_j e^{ijs}|^{2p} ds)^{1/p} \leq C(\sum |x_j|)^{(2p-2)/p} \leq Cn^{1/q}$. ∎

Proof of Theorem 2.3: Denote $T_n = m_n(\frac{1}{n}S_n - EX_1^2)$ and as previously, let $\lambda_j = \lambda_j(n), 1 \leq j \leq n$, be the eigenvalues of the covariance of $X_1, \ldots, X_n$. Since by Lemma 3.7 and the choice of m_n $\max_j \lambda_j/m_n \to 0$, for every $y \in R$ and for all $n \geq n_0(y)$ we have

$$\log E\exp(nm_n^{-2}yT_n) = -ynm_n^{-1}EX_1^2 - \frac{1}{2}\sum_{j=1}^{n}\log(1 - 2y\lambda_j/m_n) .$$

Notice that by Taylor's Theorem for $|w| < 1$

$$\log(1-w) = -w - (1/2)w^2(1-tw)^{-2},$$

where $t = t(w) \in [0, 1]$. This is applied here to $w_j = 2y\lambda_j/m_n$ which by Lemma 3.7 satisfies $\sup_j |w_j| \to 0$ as $n \to \infty$, and hence, $|1 - t(w_j)w_j| \to 1$ uniformly in $1 \leq j \leq n$. This shows that the limit of

$$m_n^2 n^{-1}\log E\exp(nm_n^{-2}yT_n)$$

is the same as that of

$$ym_n(n^{-1}\sum_{j=1}^{n}\lambda_j - E(X_1^2)) + y^2 n^{-1}\sum_{j=1}^{n}\lambda_j^2$$

Clearly, $\sum_{j=1}^{n}\lambda_j = \operatorname{tr} \mathbf{A}_n = nE(X_1^2)$, and

$$n^{-1}\sum_{j=1}^{n}\lambda_j^2 = n^{-1}\operatorname{tr} \mathbf{A}_n^2 = \sum_{k=-(n-1)}^{n-1}(1 - |k|/n)r_k^2 = \sum_{k=-(n-1)}^{n-1} r_k^2 - 2\sum_{k=1}^{n-1}(k/n)r_k^2$$

Notice that by Parseval's identity $\sum_{k=-(n-1)}^{n-1} r_k^2 \to \sum_{k=-\infty}^{\infty} r_k^2 = \sigma^2/2$ as $n \to \infty$. On the other hand, by Kronecker's Lemma $\sum_{k=1}^{n-1}(k/n)r_k^2 \to 0$ as $n \to \infty$ leading to

$$\lim_{n\to\infty} m_n^2 n^{-1}\log E\exp(nm_n^{-2}yT_n) = \frac{1}{2}y^2\sigma^2.$$

This ends the proof by the Gärtner-Ellis Theorem. ∎

3.4 Proof of Proposition 2.1

For $f(s)$ or $\|\mathbf{F}(s)\|$ bounded, the CLT follows immediately from Lemmas 3.3 and 3.6 by a simple complex analysis argument given in [4, Proposition 1]. In general, for every $M < \infty$, we let $X_t = Y_t + Z_t$ in the continuous time setup and $\mathbf{X}_k = \mathbf{Y}_k + \mathbf{Z}_k$ in the discrete time setup; in the former case Y_t and Z_t are independent, real-valued, centered, separable stationary Gaussian processes with spectral densities $f_y(s) = \min(f(s), M)$ and $f_z(s) = f(s) - f_y(s)$, while in the latter $\mathbf{Y}_k$ and $\mathbf{Z}_k$

are independent, R^d-valued, centered, stationary Gaussian sequences, with the spectral densities $\mathbf{F}_y(s)$ and $\mathbf{F}_z(s)$ having the same eigenvectors as $\mathbf{F}(s)$ but with eigenvalues $\min(\lambda_j(s), M)$ and $\max(\lambda_j(s) - M, 0)$ respectively. Then, in the continuous time setup,

$$W_M := \frac{1}{\sqrt{T}} \int_0^T (X_t^2 - Y_t^2 - E(X_0^2 - Y_0^2))dt = \frac{1}{\sqrt{T}} \int_0^T (Z_t^2 - E(Z_0^2))dt + \frac{2}{\sqrt{T}} \int_0^T Y_t Z_t dt ,$$

has mean zero and variance bounded above by $\epsilon_M := 4\sigma(4\pi \int_{-\infty}^{\infty} f_z(s)^2 ds)^{1/2}$, while in the discrete time setup,

$$W_M := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\langle \mathbf{X}_i | \mathbf{X}_i \rangle - \langle \mathbf{Y}_i | \mathbf{Y}_i \rangle - E(\langle \mathbf{X}_0 | \mathbf{X}_0 \rangle - \langle \mathbf{Y}_0 | \mathbf{Y}_0 \rangle)) ,$$

has zero mean and variance bounded by $\epsilon_M := 4\sigma(\pi^{-1} \int_0^{2\pi} \mathrm{tr}(\mathbf{F}_z(s)^2) ds)^{1/2}$. Note that in both cases $\epsilon_M \to 0$ as $M \to \infty$, hence for every $\delta > 0$, by Chebyshev's inequality $P(|W_M| > \delta) < \epsilon_M/\delta^2 \to 0$ as $M \to \infty$ uniformly in T (n). Since $f_y(s)$ is bounded, $\frac{1}{\sqrt{T}} \int_0^T (Y_t^2 - E(Y_0^2))dt$ is asymptotically normal $N(0, \sigma_M)$ as $T \to \infty$, with $\sigma_M := (4\pi \int_{-\infty}^{\infty} f_y(s)^2 ds)^{1/2}$ monotonically increasing to σ as $M \to \infty$. Similarly, in the discrete time setup, $\|\mathbf{F}_y(s)\|$ is bounded and hence $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\langle \mathbf{Y}_i | \mathbf{Y}_i \rangle - E(\langle \mathbf{Y}_0 | \mathbf{Y}_0 \rangle))$ is asymptotically normal $N(0, \sigma_M)$ as $n \to \infty$, with $\sigma_M := (\pi^{-1} \int_0^{2\pi} \mathrm{tr}\ (\mathbf{F}_y(s))^2 ds)^{1/2} \nearrow \sigma$ as $M \to \infty$. The required CLT then follows by the continuity of the normal distribution function. ■

3.5 Proof of Proposition 2.2

For $\mathbf{y} = [y_1, y_2]$ define $L_n(\mathbf{y}) = \log E \exp(\langle \mathbf{y} | \mathbf{S}_n \rangle)$. Let $\mathbf{R}_n$ be the covariance matrix of $\mathbf{X} = [X_1, \ldots, X_n]'$ with $\lambda_1(n)$ denoting the maximal eigenvalue of $\mathbf{R}_n$, $\mathbf{I}_n$ denoting the identity matrix, and $\mathbf{e}_n = [1, 1, \ldots, 1]'$. By adapting the calculations of Lemma 3.1 we have for $y_2 < 1/(2\lambda_1(n))$

$$L_n(\mathbf{y}) = L_n([0, y_2]) + \frac{1}{2} y_1^2 \langle \mathbf{e}_n | \mathbf{R}_n^{1/2} (\mathbf{I}_n - 2y_2 \mathbf{R}_n)^{-1} \mathbf{R}_n^{1/2} \mathbf{e}_n \rangle$$

(and $L_n(\mathbf{y}) = \infty$ for all other values of $\mathbf{y}$).

Lemma 3.8 *If $y_2 < 1/(2M)$ then*

$$L(\mathbf{y}) = \lim_{n\to\infty} n^{-1} L_n(\mathbf{y}) = L(y_2) + \frac{y_1^2 f(0)}{2(1 - 2y_2 f(0))} ,$$

with $L(y)$ given by (5), and $n^{-1} L_n(\mathbf{y}) \to \infty$ when $y_2 > 1/(2M)$.

Proof: We have by [16, page 65] that $n^{-1} L_n([0, y_2]) \to L(y_2)$ for all $y_2 < 1/(2M)$ and $n^{-1} L_n(\mathbf{y}) \to \infty$ for all $y_2 > 1/(2M)$. Taking $y_2 < 1/(2M)$ we have by [16, pages 27, 53, 209] that

$$n^{-1} \langle \mathbf{e}_n | (\mathbf{I}_n - 2y_2 \mathbf{R}_n)^{-1} \mathbf{e}_n \rangle \to 1/(1 - 2y_2 f(0)) ,$$

and the proof is completed by noting that $2y_2 \mathbf{R}_n^{1/2} (\mathbf{I}_n - 2y_2 \mathbf{R}_n)^{-1} \mathbf{R}_n^{1/2} + \mathbf{I}_n = (\mathbf{I}_n - 2y_2 \mathbf{R}_n)^{-1}$. ■

Proof of Proposition 2.2: Defining $L([y_1, 1/(2M)]) = \lim_{y_2 \nearrow 1/(2M)} L(\mathbf{y})$ and $L(\mathbf{y}) = \infty$ for $y_2 > 1/(2M)$, it is easy to check that $J(x_1, x_2)$ of (10) is the Fenchel-Legendre transform of $L(\mathbf{y})$. Here again, it is easy to check that conditions [11, (2.13) and (2.15)] follow from the monotonicity of $L_n(\mathbf{y})$ with respect to y_2. Hence, suffices to show that $L(\mathbf{y})$ is steep, for then the LDP with rate function $J(\cdot)$ holds by [11, Proposition 2.14] (even if $n^{-1} L_n(\mathbf{y})$ fails to converge for $y_2 = 1/(2M)$). To that end, note that for $y_2 < 1/(2M)$

$$\frac{\partial L(\mathbf{y})}{\partial y_2} \geq \frac{1}{2\pi} \int_0^{2\pi} \frac{f(s)}{1 - 2y_2 f(s)} ds .$$

Hence, by the differentiability of $f(s)$ we have $\frac{\partial L(\mathbf{y})}{\partial y_2} \to \infty$ as $y_2 \nearrow 1/(2M)$ implying that $L(\mathbf{y})$ is steep (for more details, see the proof of Proposition 2.3). ■

3.6 Proof of Proposition 2.3

Let $L_n(\mathbf{y}) = \log E(\exp(\langle \mathbf{y}|\mathbf{S}_n\rangle))$ and let $\mathbf{Y}_n$ be the symmetric Toeplitz $n \times n$-matrix whose first row is $(y_0, \frac{1}{2}y_1, \ldots, \frac{1}{2}y_d, 0, \ldots, 0)$. Let $\mathbf{R}_n$ be the covariance matrix of $\mathbf{X} = [X_1, \ldots, X_n]'$. Since $\langle \mathbf{y}|\mathbf{S}_n\rangle = \mathbf{X}'\mathbf{Y}_n\mathbf{X}$, by Lemma 3.1 we have

$$L_n(\mathbf{y}) = -1/2 \sum_{j=1}^{n} \log(1 - 2\lambda_j(\mathbf{y})), \tag{29}$$

where $\lambda_j(\mathbf{y})$ are the eigenvalues of the matrix $\mathbf{M}_n = \mathbf{Y}_n\mathbf{R}_n$ and $\mathbf{y}$ is such that $\max_j\{\lambda_j(\mathbf{y})\} < 1/2$.

For i.i.d. X_j we have that $\mathbf{R}_n$ is the identity matrix, hence $\mathbf{M}_n = \mathbf{Y}_n$ is the symmetric Toeplitz matrix corresponding to the "signed" bounded spectral density $\langle \mathbf{y}|\mathbf{f}(s)\rangle$. In particular, by [16, page 65] for $\mathbf{y} \in D$

$$\frac{1}{n} L_n(\mathbf{y}) \to L(\mathbf{y}) = -\frac{1}{4\pi} \int_0^{2\pi} \log(1 - 2\langle \mathbf{y}|\mathbf{f}(s)\rangle) ds.$$

By [16, pages 38, 44] this relation holds also for $\mathbf{y} \in \partial D$, i.e. when $\sup_s \langle \mathbf{y}|\mathbf{f}(s)\rangle = \frac{1}{2}$, while $n^{-1}L_n(\mathbf{y}) \to \infty$ for all other values of $\mathbf{y}$.

Notice that if $\|\mathbf{y}\| < 1/(2(d+1))$ then $\mathbf{y} \in D$. Therefore, in order to establish the LDP, we need only to verify the steepness condition, i.e.,

$$\lim_{\mathbf{y}\to\mathbf{y}_0, \mathbf{y}\in D} \|L'(\mathbf{y})\| = \infty$$

for all $\mathbf{y}_0 \in \partial D$, see [10, Theorem 2.3.6]. To this end, fix $\mathbf{y}_0 \in \partial D$ and let $0 \le s_0 \le 2\pi$ be such that $\langle \mathbf{y}_0|\mathbf{f}(s_0)\rangle = 1/2$. It suffices to show that $|\langle \mathbf{y}_0|L'(\mathbf{y})\rangle| \to \infty$ as $\mathbf{y} \to \mathbf{y}_0, \mathbf{y} \in D$. Clearly,

$$\langle \mathbf{y}_0|L'(\mathbf{y})\rangle = \frac{1}{2\pi} \int_0^{2\pi} \frac{\langle \mathbf{y}_0|\mathbf{f}(s)\rangle}{1 - 2\langle \mathbf{y}|\mathbf{f}(s)\rangle} ds.$$

Let $I_+ = \{s : \langle \mathbf{y}_0|\mathbf{f}(s)\rangle \ge 0\}$, and $I_- = \{s : \langle \mathbf{y}_0|\mathbf{f}(s)\rangle < 0\}$. We have

$$\limsup_{\mathbf{y}\to\mathbf{y}_0} \left|\frac{1}{2\pi} \int_{I_-} \frac{\langle \mathbf{y}_0|\mathbf{f}(s)\rangle}{1 - 2\langle \mathbf{y}|\mathbf{f}(s)\rangle} ds\right| \le \|\mathbf{y}_0\|.$$

Since $\mathbf{f}(s)$ is differentiable, for each $\epsilon > 0$ there is $\delta > 0$ such that for $|s - s_0| < \delta$ we have $|\langle \mathbf{y}_0|\mathbf{f}(s)\rangle - \langle \mathbf{y}_0|\mathbf{f}(s_0)\rangle| < \epsilon\delta$ and $\langle \mathbf{y}_0|\mathbf{f}(s)\rangle \ge m > 0$; i.e. $(s_0 - \delta, s_0) \subset I_+$ (if $s_0 = 0$ replace $(s_0 - \delta, s_0)$ by $(s_0, s_0 + \delta)$). Then

$$\begin{aligned}
\int_{I_+} \frac{\langle \mathbf{y}_0|\mathbf{f}(s)\rangle}{1 - 2\langle \mathbf{y}|\mathbf{f}(s)\rangle} ds &\ge \int_{s_0-\delta}^{s_0} \frac{\langle \mathbf{y}_0|\mathbf{f}(s)\rangle}{1 - 2\langle \mathbf{y}|\mathbf{f}(s)\rangle} ds \\
&\ge m \int_{s_0-\delta}^{s_0} \frac{1}{2\langle \mathbf{y}_0|\mathbf{f}(s_0) - \mathbf{f}(s)\rangle + 2\langle \mathbf{y}_0 - \mathbf{y}|\mathbf{f}(s)\rangle} ds \ge m \frac{\delta}{2\epsilon\delta + 2\|\mathbf{y}_0 - \mathbf{y}\|}
\end{aligned}$$

Therefore $\liminf_{\mathbf{y}\to\mathbf{y}_0} \langle \mathbf{y}_0|L'(\mathbf{y})\rangle \ge m/(4\pi\epsilon) - \|\mathbf{y}_0\|$. Taking $\epsilon \to 0$, this ends the proof. ■

Acknowledgement The authors thank Ofer Zeitouni for improving their proof of Lemma 3.2.

References

[1] F. Avram, On Bilinear Forms in Gaussian Random Variables and Toeplitz Matrices, Probab. Th. Rel. Fields 79 (1988), 37–45.

[2] G. R. Benitz & J. A. Bucklew, Large Deviation Rate Calculations for Nonlinear Detectors in Gaussian Noise, IEEE Transactions on Information Theory 36 (1990), 358 – 371.

[3] R. C. Bradley, On the spectral density and asymptotic normality of weakly dependent random fields, Journal of Theoretical Probability 5 (1992), 355 –373.

[4] W. Bryc, A Remark On The Connection Between The Large Deviation Principle And The Central Limit Theorem, Stat. Probab. Letters, to appear.

[5] W. Bryc & W. Smolenski, On the Large Deviation Principle for a quadratic functional of the autoregressive process, Stat. Probab. Lett. 17 (1993), 281–285.

[6] J. A. Bucklew, Large deviations techniques in decision, simulation, and estimation, Wiley, 1990.

[7] J. A. Bucklew & J. S. Sadowsky, A Contribution to the Theory of Chernoff's Bounds, IEEE Transactions on Information Theory 39 (1993), 249–254.

[8] J. Coursol & D. Dacunha-Castelle, Sur la formule de Chernoff pour deux processus gaussiens stationnaires, C. R. Acad. Sci. Paris Ser. A-B, 288 (1979), A769–A770.

[9] A. de Acosta, Upper bounds for Large Deviations of Dependent Random Vectors, Z. Wahrscheinlichkeitsth. Verw. Gebiete 69 (1985), 551-565.

[10] A. Dembo & O. Zeitouni, Large Deviations Techniques and Applications, Jones and Bartlett, Boston, MA, 1993.

[11] A. Dembo & O. Zeitouni, Large deviations via parameter dependent change of measure and an application to the lower tail of Gaussian processes, to appear, proceedings of the Ascona Conference on Stochastic Analysis, 1993. (Editors: E. Bolthausen, M. Dozzi and F. Russo, Birkhauser).

[12] J. Deuschel, Invariance principle and empirical mean large deviations of the critical Ornstein-Uhlenbeck process, Ann. Probab. 17 (1989) 74–90.

[13] M. D. Donsker & S. R. S. Varadhan, Large deviations for Stationary Gaussian Processes, Commun. Math. Phys. 97 (1985), 187-210.

[14] N. Dunford & J. T. Schwartz, Linear operators, Interscience Publishers, New York, 1958-71.

[15] R. Fox & M. S. Taqqu, Central Limit Theorems for Quadratic Forms in Random variables having Long-Range Dependence, Probab. Th. Rel. Fields 74 (1987), 213–240.

[16] U. Grenander & G. Szegö, Toeplitz Forms and their Applications, University of California Press, 1958.

[17] T. Koopmans, Serial correlation and quadratic forms in normal variables, Ann. Math. Stat. 13 (1942), 14–33.

[18] J. E. Mazo & J. Salz, Probability of error for quadratic detectors, Bell System Techn. Journ. XLIV (1965), 2165-2186.

[19] P. Ney & E. Nummelin, Markov additive process I. Eigenvalue properties and limit theorems, Ann. Probab. 15 (1987), pp. 561-592; Markov additive process II , ibidem, pp. 593-609.

[20] M. Rosenblatt, Stationary Sequences and Random Fields, Birkhäuser, Boston, 1985.

[21] Y. Steinberg & O. Zeitouni, On tests for normality, IEEE Transactions on Information Theory 38 (1992), 1779–1787.

[22] N. Terrin & M. S. Taqqu, A Noncentral Limit Theorem for Quadratic Forms of Gaussian Stationary Sequences, Journal of Theoretical Probability 3 (1990), 449– 475.