Article

A Large Deviation Principle and an Expression of the Rate Function for a Discrete Stationary Gaussian Process

Olivier Faugeras and James MacLaurin *

INRIA Sophia Antipolis Mediterannee, 2004 Route Des Lucioles, Sophia Antipolis, France;
E-Mail: olivier.faugeras@inria.fr

* Author to whom correspondence should be addressed; E-Mail: james.maclaurin@inria.fr.

Abstract: We prove a large deviation principle for a stationary Gaussian process over $\mathbb{R}^b$, indexed by $\mathbb{Z}^d$ (for some positive integers d and b), with positive definite spectral density, and provide an expression of the corresponding rate function in terms of the mean of the process and its spectral density. This result is useful in applications where such an expression is needed.

Keywords: stationary Gaussian process; large deviation principle; spectral density

---

1. Introduction

In this paper, we prove a large deviation principle (LDP) for a spatially-stationary, indexed by $\mathbb{Z}^d$, Gaussian process over $\mathbb{R}^b$ and obtain an expression for the rate function. Our work in mathematical neuroscience involves the search for asymptotic descriptions of large ensembles of neurons [1–3]. Since there are many sources of noise in the brains of mammals [4], the mathematician interested in modeling certain aspects of brain functioning is often led to consider spatial Gaussian processes that (s)he uses to model these noise sources. This motivates us to use large deviation techniques. Being also interested in formulating predictions that can be experimentally tested in neuroscience laboratories, we strive to obtain analytical results, *i.e.*, effective results from which, for example, numerical simulations can be developed. This is why we determine a more tractable expression for the rate function in this article.

Our result concerns the large deviations of ergodic phenomena, the literature of which we now briefly survey. Donsker and Varadhan obtained a large deviation estimate for the law governing the empirical

process generated by a Markov process [5]. They then determined a large deviation principle for a $\mathbb{Z}$-indexed stationary Gaussian process, obtaining a particularly elegant expression for the rate function using spectral theory. Chiyonobu *et al.* [6–8] obtain a large deviation estimate for the empirical measure generated by ergodic processes satisfying certain mixing conditions. Baxter *et al.* [9] obtain a variety of results for the large deviations of ergodic phenomena, including one for the large deviations of $\mathbb{Z}$-indexed $\mathbb{R}^b$-valued stationary Gaussian processes. Steinberg *et al.* [10] have proven an LDP for a stationary $\mathbb{Z}^d$-indexed Gaussian process over $\mathbb{R}$, and [11] obtain an LDP for an $\mathbb{R}$-indexed, $\mathbb{R}$-valued stationary Gaussian process.

In the work we are developing [12], we need large deviation results for spatially-ergodic Ornstein–Uhlenbeck processes. This requires Theorem 1 of this paper.

In the first section, we make some preliminary definitions and state the chief theorem, Theorem 1, for zero mean processes. In the second section, we prove the theorem. In Appendix A, we state and prove several identities involving the relative entropy, which are necessary for the proof of Theorem 1. In Appendix B, we prove a general result for the large deviations of exponential approximations. We prove Corollary 2, which extends the result in Theorem 1 to non-zero mean processes in the Appendix C.

2. Preliminary Definitions

For some topological space Ω equipped with its Borelian σ-algebra $\mathcal{B}(\Omega)$, we denote the set of all probability measures by $\mathcal{M}(\Omega)$. We equip $\mathcal{M}(\Omega)$ with the topology of weak convergence. Our process is indexed by $\mathbb{Z}^d$. For $j \in \mathbb{Z}^d$, we write $j = (j(1), \ldots, j(d))$. For some positive integer $n > 0$, we let $V_n = \{j \in \mathbb{Z}^d : |j(\delta)| \leq n \text{ for all } 1 \leq \delta \leq d\}$. Let $\mathcal{T} = \mathbb{R}^b$, for some positive integer b. We equip $\mathcal{T}$ with the Euclidean topology and $\mathcal{T}^{\mathbb{Z}^d}$ with the cylindrical topology, and we denote the Borelian σ-algebra generated by this topology by $\mathcal{B}(\mathcal{T}^{\mathbb{Z}^d})$. For some $\mu \in \mathcal{M}(\mathcal{T}^{\mathbb{Z}^d})$ governing a process $(X^j)_{j \in \mathbb{Z}^d}$, we let $\mu^{V_n} \in \mathcal{M}(\mathcal{T}^{V_n})$ denote the marginal governing $(X^j)_{j \in V_n}$. For some $j \in \mathbb{Z}^d$, let the shift operator $S^j : \mathcal{T}^{\mathbb{Z}^d} \to \mathcal{T}^{\mathbb{Z}^d}$ be $S(\omega)^k = \omega^{j+k}$. We let $\mathcal{M}_s(\mathcal{T}^{\mathbb{Z}^d})$ be the set of all stationary probability measures μ on $(\mathcal{T}^{\mathbb{Z}^d}, \mathcal{B}(\mathcal{T}^{\mathbb{Z}^d}))$, such that, for all $j \in \mathbb{Z}^d$, $\mu \circ (S^j)^{-1} = \mu$. We use Herglotz's theorem to characterize the law $Q \in \mathcal{M}_s(\mathcal{T}^{\mathbb{Z}^d})$ governing a stationary process (X^j) in the following manner. We assume that $E[X^j] = 0$ and:

$$E[X_0{}^{\dagger}X_k] = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \exp(i\langle k, \theta \rangle)\tilde{G}(\theta)d\theta. \tag{1}$$

Our convention throughout this paper is to denote the transpose of X as ${}^{\dagger}X$ and the spectral density with a tilde. $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathbb{R}^d$. Here, $\tilde{G}(\theta)$ is a continuous function $[-\pi, \pi]^d \to \mathbb{C}^{b \times b}$, where we consider $[-\pi, \pi]^d$ to have the topology of a torus. In addition, $\tilde{G}(-\theta) = {}^{\dagger}\tilde{G}(\theta) = \bar{\tilde{G}}$ ($\bar{x}$ indicates the complex conjugate of x). We assume that for all $\theta \in [-\pi, \pi]^d$, $\det \tilde{G}(\theta) > \tilde{G}_{min}$ for some $\tilde{G}_{min} > 0$, from which it follows that for each θ, $\tilde{G}(\theta)$ is Hermitian positive definite. If $x \in \mathbb{C}^b$, then ${}^{\dagger}xX^j$ is a stationary sequence with spectral density ${}^{\dagger}x\tilde{G}\bar{x}$. We employ the operator norm over $\mathbb{C}^{b \times b}$. Let $p_n : \mathcal{T}^{\mathbb{Z}^d} \to \mathcal{T}^{\mathbb{Z}^d}$ be such that $p_n(\omega)^k = \omega^{k \mod V_n}$. Here, and throughout the paper, we take k

mod V_n to be the element $l \in V_n$, such that, for all $1 \leq \gamma \leq d$, $l(\gamma) = k(\gamma) \mod (2n+1)$. Define the process-level empirical measure $\hat{\mu}_n : \mathcal{T}^{\mathbb{Z}^d} \to \mathcal{M}_s\left(\mathcal{T}^{\mathbb{Z}^d}\right)$ as:

$$\hat{\mu}_n(\omega) = \frac{1}{(2n+1)^d} \sum_{k \in V_n} \delta_{S^k p_n(\omega)}. \tag{2}$$

Let $\Pi^n \in \mathcal{M}\left(\mathcal{M}_s(\mathcal{T}^{\mathbb{Z}^d})\right)$ be the image law of Q under $\hat{\mu}_n$. We note that in this definition, we need not have chosen V_n to have an odd side length $(2n+1)$: this choice is for notational convenience, and these results could easily be reproduced in the case that V_n has side length n.

In the context of mathematical neuroscience, (X^j) could correspond to a model of interacting neurons on a lattice ($d = 1, 2$ or 3), as in [12,13]. We note that the large deviation principle of this paper may be used to obtain an LDP for other processes using standard methods, such as the contraction principle or Lemma 13.

Definition 1. Let $(\Omega, \mathcal{H})$ be a measurable space and μ, ν probability measures.

$$I^{(2)}(\mu||\nu) = \sup_{f \in B} \left\{ \mathbb{E}^\mu[f] - \log \mathbb{E}^\nu[\exp(f)] \right\},$$

where B is the set of all bounded measurable functions. If Ω is Polishand $\mathcal{H} = \mathcal{B}(\Omega)$, then we only need to take the supremum over the set of all continuous bounded functions.

Let (Y^j) be a stationary Gaussian process on $\mathcal{T}$, such that $E[Y^j] = 0$, $E[Y^{j\dagger}Y^k] = 0$ and $E[Y^{j\,\dagger}Y^j] = \mathrm{Id}_b$. Each Y^j is governed by a law P, and we write the governing law in $\mathcal{M}_s(\mathcal{T}^{\mathbb{Z}^d})$ as $P^{\mathbb{Z}^d}$. It is clear that the governing law over V_n may be written as $P^{\otimes V_n}$ (that is the product measure of P, indexed over V_n).

Definition 2. Let $\mathcal{E}_2$ be the subset of $\mathcal{M}_s(\mathcal{T}^{\mathbb{Z}^d})$ defined by:

$$\mathcal{E}_2 = \{\mu \in \mathcal{M}_s(\mathcal{T}^{\mathbb{Z}^d}) \,|\, \mathbb{E}^\mu[\|\omega^0\|^2] < \infty\}.$$

We define the process-level entropy to be, for $\mu \in \mathcal{M}_s(\mathcal{T}^{\mathbb{Z}^d})$:

$$I^{(3)}(\mu) = \lim_{n \to \infty} \frac{1}{(2n+1)^d} I^{(2)}\left(\mu^{V_n}||P^{\otimes V_n}\right).$$

It is a consequence of Lemma 11 that if $\mu \notin \mathcal{E}_2$, then $I^{(3)}(\mu) = \infty$. For further discussion of this rate function and a proof that $I^{(3)}$ is well-defined, see [8].

Definition 3. A sequence of probability laws (Γ^n) on some topological space Ω equipped with its Borelian σ-algebra is said to satisfy a strong large deviation principle (LDP) with rate function $I : \Omega \to \mathbb{R}$ if I is lower semicontinuous, for all open sets O,

$$\varliminf_{n \to \infty} \frac{1}{(2n+1)^d} \log \Gamma^n(O) \geq -\inf_{x \in O} I(x)$$

and for all closed sets F:

$$\varlimsup_{n \to \infty} \frac{1}{(2n+1)^d} \log \Gamma^n(F) \leq -\inf_{x \in F} I(x).$$

If, furthermore, the set $\{x : I(x) \leq \alpha\}$ is compact for all $\alpha \geq 0$, we say that I is a good rate function.

Definition 4. For $\mu \in \mathcal{E}_2$, we denote the $\mathbb{C}^b \times \mathbb{C}^b$-valued spectral measure on $([-\pi, \pi]^d, \mathcal{B}([-\pi, \pi]^d))$ (which exists due to Herglotz's theorem) by $\tilde{\mu}$. We have:

$$\frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \exp(i\langle k, \theta \rangle) d\tilde{\mu}(\theta) = \mathbb{E}^\mu \left[\omega^{0\dagger} \omega^k \right].$$

For $\theta \in [-\pi, \pi]^d$, let $\tilde{H}(\theta) = \tilde{G}(\theta)^{-\frac{1}{2}}$ be the Hermitian positive definite square root of $\tilde{G}(\theta)^{-1}$ and:

$$\tilde{H}(\theta) = \sum_{j \in \mathbb{Z}^d} H^j \exp\left(-i\langle j, \theta \rangle\right). \tag{3}$$

The $b \times b$ matrices H^j are the coefficients of the absolutely convergent Fourier series (due to Wiener's theorem) of $\tilde{G}^{-1/2}$. Define $\beta : \mathcal{T}^{\mathbb{Z}^d} \to \mathcal{T}^{\mathbb{Z}^d}$ as follows:

$$(\beta(\omega))^k = \sum_{j \in \mathbb{Z}^d} {}^\dagger H^j \omega^{k-j}.$$

The theorem below is the chief result of this paper.

Theorem 1. (Π^n) *satisfies a strong LDP with good rate function* $I^{(3)}(\mu \circ \beta^{-1})$. *Here:*

$$I^{(3)}\left(\mu \circ \beta^{-1}\right) = \begin{cases} I^{(3)}(\mu) - \Gamma(\mu) & \text{if } \mu \in \mathcal{E}_2 \\ \infty & \text{otherwise,} \end{cases} \tag{4}$$

where:

$$\Gamma(\mu) = \begin{cases} \Gamma_1 + \Gamma_2(\mu) & \text{if } \mu \in \mathcal{E}_2 \\ 0 & \text{otherwise.} \end{cases}$$

Here:

$$\Gamma_1 = -\frac{1}{2(2\pi)^d} \int_{[-\pi,\pi]^d} \log \det \tilde{G}(\theta) d\theta$$

$$\Gamma_2(\mu) = \frac{1}{2(2\pi)^d} \int_{[-\pi,\pi]^d} \operatorname{tr}\left(\left(\mathrm{Id}_b - \tilde{G}(\theta)^{-1}\right) d\tilde{\mu}(\theta)\right).$$

Finally, the rate function uniquely vanishes at $\mu = Q$.

Corollary 2. *Suppose that the underlying process* Q *is defined as previously, except with mean* $E^Q[\omega^j] = c$ *for all* $j \in \mathbb{Z}^d$ *and some constant* $c \in \mathbb{R}^b$. *If we denote the image law of the empirical measure by* Π_c^n, *then* (Π_c^n) *satisfies a strong LDP with a good rate function (for* $\mu \in \mathcal{E}_2$):

$$I^{(3)}(\mu) - \Gamma(\mu) + \frac{1}{2} {}^\dagger c \tilde{G}(0)^{-1} c - {}^\dagger c \tilde{G}(0)^{-1} m^\mu. \tag{5}$$

Here, $m^\mu = E^{\mu(\omega)}[\omega^j]$ *for all* $j \in \mathbb{Z}^d$. *If* $\mu \notin \mathcal{E}_2$, *then the rate function is infinite. The rate function has a unique minimum, i.e.,* $I^{(3)}(\mu) - \Gamma(\mu) + \frac{1}{2} {}^\dagger c \tilde{G}(0)^{-1} c - {}^\dagger c \tilde{G}(0)^{-1} m^\mu = 0$ *if and only if* $\mu = Q$.

We prove this in Appendix C.

3. Proof of Theorem 1

In this proof, we essentially adapt the methods of [9,14]. We introduce the following metric over $\mathcal{T}^{\mathbb{Z}^d}$. For $j \in \mathbb{Z}^d$, let $\lambda^j = 3^{-d} \prod_{\delta=1}^{d} 2^{-|j(\delta)|}$. Define the metric d^λ as follows,

$$d^\lambda(x, y) = \sum_{j \in \mathbb{Z}^d} \lambda^j \min\left(\|x^j - y^j\|, 1\right),$$

where the above is the Euclidean norm. Let $d^{\lambda, \mathcal{M}}$ be the induced Prokhorov metric over $\mathcal{M}(\mathcal{T}^{\mathbb{Z}^d})$. These metrics are compatible with their respective topologies.

For $\theta \in [-\pi, \pi]^d$, let $\tilde{F}(\theta) = \tilde{G}(\theta)^{1/2}$ be the Hermitian positive square root of $\tilde{G}$ and:

$$\tilde{F}(\theta) = \sum_{j \in \mathbb{Z}^d} F^j \exp\left(-i\langle j, \theta\rangle\right). \tag{6}$$

The $b \times b$ matrices F^j are the coefficients of the absolutely convergent Fourier series of the positive square root. Define $\tau : \mathcal{T}^{\mathbb{Z}^d} \to \mathcal{T}^{\mathbb{Z}^d}$ and $\tau_{(M)} : \mathcal{T}^{\mathbb{Z}^d} \to \mathcal{T}^{\mathbb{Z}^d}$ as follows,

$$(\tau(\omega))^k = \sum_{j \in \mathbb{Z}^d} {}^\dagger F^j \omega^{k-j},$$

$$\left(\tau_{(n)}(\omega)\right)^k = \sum_{j \in V_n} {}^\dagger F^j \omega^{k-j} \prod_{\delta=1}^{d} \left(1 - \frac{|j(\delta)|}{2n+1}\right). \tag{7}$$

We note that $\tau = \beta^{-1}$ (on a suitable domain, where the series are convergent) and that $\tau_{(n)}$ is a continuous map, but τ is not continuous (in general).

We note (using Lemma 6) that $P^{\mathbb{Z}^d} \circ \tau^{-1}$ has spectral density $\tilde{G}$. We define:

$$\tilde{F}_{(n)}(\theta) = \sum_{j \in V_n} F^j \exp\left(-i\langle j, \theta\rangle\right) \prod_{\delta=1}^{d} \left(1 - \frac{|j(\delta)|}{2n+1}\right). \tag{8}$$

We write $\varepsilon(n) = \sup_{\theta \in [-\pi, \pi]^d} \left\|\tilde{F}(\theta) - \tilde{F}_{(n)}(\theta)\right\|^2$. By Fejer's theorem, $\varepsilon(n) \to 0$ as $n \to \infty$. We define $\tilde{G}_{(n)}(\theta) = \tilde{F}_{(n)}(\theta)^2$, noting that this is the spectral density of $P^{\mathbb{Z}^d} \circ \tau_{(n)}^{-1}$. Let $\Gamma_{(n)}(\mu) = \Gamma_{1,(n)} + \Gamma_{2,(n)}(\mu)$, where for $\mu \in \mathcal{E}_2$,

$$\Gamma_{1,(n)} = -\frac{1}{2(2\pi)^d} \int_{[-\pi,\pi]^d} \log \det \tilde{G}_{(n)}(\theta) d\theta, \tag{9}$$

$$\Gamma_{2,(n)}(\mu) = \frac{1}{2(2\pi)^d} \int_{[-\pi,\pi]^d} \mathrm{tr}\left(\left(\mathrm{Id_b} - \tilde{G}_{(n)}^{-1}(\theta)\right) d\tilde{\mu}(\theta)\right). \tag{10}$$

If $\mu \notin \mathcal{E}_2$, let $\Gamma_{1,(n)} = \Gamma_{2,(n)}(\mu) = 0$. Let $R^n \in \mathcal{M}(\mathcal{M}_s(\mathcal{T}^{\mathbb{Z}^d}))$ be the law governing $\hat{\mu}_n(Y)$, where we recall that the stationary process (Y^j) is defined just below Definition 1. Let $\Pi_{(m)}^n \in \mathcal{M}(\mathcal{M}_s(\mathcal{T}^{\mathbb{Z}^d}))$ be the law governing $\hat{\mu}_n(\tau_{(m)}(Y))$.

Lemma 3. *(R^n) satisfies a large deviation principle with good rate function $I^{(3)}(\mu)$. If $\mu \notin \mathcal{E}_2$, then $I^{(3)}(\mu) = \infty$.*

Proof. The first statement is proven in [8]. The last statement follows from Lemma 11 below. □

Lemma 4. $(\Pi^n_{(m)})$ *satisfies a strong LDP with good rate function given by, for* $\mu \in \mathcal{E}_2$:

$$I^{(3)}_{(m)}(\mu) = \inf_{\nu \in \mathcal{E}_2 : \mu = \nu \circ \tau^{-1}_{(m)}} I^{(3)}(\nu) = I^{(3)}(\mu) - \Gamma_{(m)}(\mu). \tag{11}$$

If $\mu \notin \mathcal{E}_2$, *then* $I^{(3)}_{(m)}(\mu) = \infty$.

Proof. The sequence of laws governing $\hat{\mu}_n(Y) \circ \tau^{-1}_{(m)}$ (as $n \to \infty$, with m fixed) satisfies a strong LDP with good rate function as a consequence of an application of the contraction principle to Lemma 3 (since $\tau_{(m)}$ is continuous). Now, through the same reasoning as in Lemma 2.1, Theorem 2.3 in [14], it follows from this that $(\Pi^n_{(m)})$ satisfies a strong LDP with the same rate function as that of $\hat{\mu}_n(Y) \circ \tau^{-1}_{(m)}$. The last identification in (11) follows from Lemmas 6 and 9 in Appendix A. We only need to take the infimum over $\mathcal{E}_2$, because by Lemma 3, $I^{(3)}(\nu)$ is infinite for $\nu \notin \mathcal{E}_2$. $\square$

Lemma 5. *If* $0 < \mathfrak{b} < \frac{1}{2\varepsilon(m)}$, *then for all* $n > 0$:

$$\frac{1}{(2n+1)^d} \log \mathbb{E}^{P^{\mathbb{Z}^d}}\left[\exp\left(\mathfrak{b} \sum_{k \in V_n} \| \tau_{(m)}(\omega)^k - \tau(\omega)^k \|^2 \right) \right] \leq -\frac{1}{2} \log\left(1 - 2\mathfrak{b}\varepsilon(m)\right).$$

The proof is almost identical to that in Lemma 2.4 in [14]. We are now ready to prove Theorem 1.

Proof. We apply Lemma 13 in the Appendix to the above result. We substitute $Y_{(m)} = \tau_{(m)}(\omega)$ and $W = \tau(\omega)$. Taking $m \to \infty$ in the equation in Lemma 5, we find that (38) is satisfied if we stipulate that $\kappa = 0$. After noting the LDP governing $\Pi^n_{(m)}$ in (11), we may thus conclude that (Π^n) satisfies a strong LDP with good rate function:

$$\lim_{\delta \to 0} \varliminf_{m \to \infty} \inf_{\gamma \in B^\delta(\mu)} I^{(3)}_{(m)}(\gamma), \tag{12}$$

where $B^\delta(\mu) = \{\gamma : d^{\lambda, \mathcal{M}}(\mu, \gamma) \leq \delta\}$.

It remains for us to identify (12) with the rate function in (4). We claim that for each $\delta > 0$,

$$\varliminf_{m \to \infty} \inf_{\gamma \in B^\delta(\mu)} \left(I^{(3)}_{(m)}(\gamma) \right) = \inf_{\gamma \in B^\delta(\mu)} (I^{(3)}(\gamma) - \Gamma(\gamma)). \tag{13}$$

To see this, we have from Lemma 11 that for all m and all γ and constants $\alpha_1 < 1$, $\alpha_3 > 1$ and $\alpha_2, \alpha_4 \in \mathbb{R}$, (note that if $\gamma \notin \mathcal{E}_2$ the inequalities below are immediate from the definitions):

$$I^{(3)}(\gamma) - \Gamma(\gamma) \geq (1 - \alpha_1) I^{(3)}(\gamma) - \alpha_2 \tag{14}$$

$$I^{(3)}_{(m)}(\gamma) \geq (1 - \alpha_1) I^{(3)}(\gamma) - \alpha_2 \tag{15}$$

$$I^{(3)}(\gamma) - \Gamma(\gamma) \leq (1 + \alpha_3) I^{(3)}(\gamma) - \alpha_4$$

$$I^{(3)}_{(m)}(\gamma) \leq (1 + \alpha_3) I^{(3)}(\gamma) - \alpha_4$$

We thus see that if $I^{(3)}(\gamma) = \infty$ for all $\gamma \in B^\delta(\mu)$, then (13) is identically infinite on both sides. Otherwise, it may be seen from (14) and (15) that it suffices to establish (13) in the case that $B^\delta(\mu) = B^\delta_l(\mu) := \{\gamma : d^{\lambda, \mathcal{M}}(\mu, \gamma) \leq \delta, I^{(3)}(\gamma) \leq l\}$ for some $l < \infty$. However, it follows from (29) and (34) that for all $\gamma \in B^\delta_l(\mu)$, there exist constants (α^m_5), which converge to zero as $m \to \infty$ and such that $\left| I^{(3)}_{(m)}(\gamma) - I^{(3)}(\gamma) + \Gamma(\gamma) \right| \leq \alpha^m_5$. We may thus conclude (13). The expression for the rate function in (4) now follows, since $I^{(3)}(\gamma) - \Gamma(\gamma)$ is lower semicontinuous, by Lemma 12.

For the second statement in the Theorem, if $I^{(3)}(\mu \circ \beta^{-1}) = 0$, then $I^{(2)}\left((\mu \circ \beta^{-1})^{V_n} || P^{\otimes V_n}\right) = 0$ for all n. However, since the relative entropy has a unique zero, this means that $(\mu \circ \beta^{-1})^{V_n} = P^{\otimes V_n}$ for all n. However, this means that $\mu \circ \beta^{-1} = P^{\mathbb{Z}^d}$, and therefore (using Lemma 6), $\mu = Q$. $\square$

Acknowledgments

Appendix

A. Properties of the Entropy

Let $\tilde{\mathcal{K}} : [-\pi, \pi]^d \to \mathbb{C}^{b \times b}$ possess an absolutely convergent Fourier series and be such that the eigenvalues of $\tilde{\mathcal{K}}(\theta)$ are strictly greater than zero for all θ. We require that $\tilde{\mathcal{K}}$ is the density of a stationary sequence, which means that we must also assume that for all θ:

$$\tilde{\mathcal{K}}(-\theta) = {}^\dagger\tilde{\mathcal{K}}(\theta) = \bar{\bar{\mathcal{K}}}(\theta).$$

This means, in particular, that $\tilde{\mathcal{K}}(\theta)$ is Hermitian. We write:

$$(\Delta(\omega))^k = \sum_{j \in \mathbb{Z}^d} {}^\dagger\mathcal{S}^j \omega^{k-j}, \text{ where} \tag{16}$$

$$\sum_{j \in \mathbb{Z}^d} \mathcal{S}^j \exp\left(-i\langle j, \theta\rangle\right) = \tilde{\mathcal{K}}(\theta)^{-\frac{1}{2}}. \tag{17}$$

Here, $\tilde{\mathcal{K}}^{-\frac{1}{2}}$ is understood to be the positive Hermitian square root of $\tilde{\mathcal{K}}^{-1}$. The Fourier series of $\tilde{\mathcal{K}}^{-\frac{1}{2}}$ is absolutely convergent as a consequence of Wiener's theorem. In this section, we determine a general expression for $I^{(3)}(\xi \circ \Delta^{-1})$. We are generalizing the result for $b = 1$ with $\mathbb{Z}$-indexing given in [14]. These results are necessary for the proofs in the previous section.

We similarly write that:

$$(\Upsilon(\omega))^k = \sum_{j \in \mathbb{Z}^d} {}^\dagger\mathcal{R}^j \omega^{k-j}, \text{ where} \tag{18}$$

$$\sum_{j \in \mathbb{Z}^d} \mathcal{R}^j \exp\left(-i\langle j, \theta\rangle\right) = \tilde{\mathcal{K}}(\theta)^{\frac{1}{2}}. \tag{19}$$

As previously, $\tilde{\mathcal{K}}^{\frac{1}{2}}$ is understood to be the positive definite Hermitian square root of $\tilde{\mathcal{K}}$. We note that $\mathcal{R}^{-j} = {}^\dagger\mathcal{R}^j$ and $\mathcal{S}^{-j} = {}^\dagger\mathcal{S}^j$.

Lemma 6. *For all $\xi \in \mathcal{E}_2$, $\xi \circ \Delta^{-1}$ and $\xi \circ \Upsilon^{-1}$ are in $\mathcal{E}_2$ and:*

$$\xi \circ \Delta^{-1} \circ \Upsilon^{-1} = \xi \circ \Upsilon^{-1} \circ \Delta^{-1} = \xi.$$

Proof. We make use of the following standard Lemma from [15], to which the reader is referred for the definition of an orthogonal stochastic measure. Let $(U^j) \in \mathbb{R}^b$ be a zero-mean stationary sequence governed by $\xi \in \mathcal{E}_2$. Then, there exists an orthogonal $\mathbb{R}^b$-valued stochastic measure $Z^\xi = Z^\xi(\Delta)$ $(\Delta \in \mathcal{B}([-\pi, \pi[^d))$, such that for every $j \in \mathbb{Z}^d$ (ξ a.s.):

$$U^j = \frac{1}{2\pi} \int_{[-\pi,\pi]^d} \exp(i\langle j, \theta \rangle) Z^\xi(d\theta). \tag{20}$$

Conversely, any orthogonal stochastic measure defines a zero-mean stationary sequence through (20). It may be inferred from this representation that:

$$Z^{\xi \circ \Delta^{-1}}(d\theta) = {}^\dagger \tilde{\mathcal{K}}(\theta)^{-\frac{1}{2}} Z^\xi(d\theta),$$
$$Z^{\xi \circ \Upsilon^{-1}}(d\theta) = {}^\dagger \tilde{\mathcal{K}}(\theta)^{\frac{1}{2}} Z^\xi(d\theta).$$

The proof that this is well-defined makes use of the fact that $\tilde{\mathcal{K}}^{\frac{1}{2}}$ and $\tilde{\mathcal{K}}^{-\frac{1}{2}}$ are uniformly continuous, since their Fourier series' each converge uniformly. This gives us the lemma. We note for future reference that, if ξ has spectral measure $d\tilde{\xi}(\theta)$, then the spectral density of $\xi \circ \Delta^{-1}$ is:

$$\tilde{\mathcal{K}}^{-\frac{1}{2}}(\theta) d\tilde{\xi}(\theta) \tilde{\mathcal{K}}^{-\frac{1}{2}}(\theta). \tag{21}$$

$\square$

It remains for us to determine a specific expression for $I^{(3)}(\xi \circ \Upsilon^{-1})$.

Definition 5. If $\xi \in \mathcal{E}_2$, we define:

$$\Gamma^\Delta(\xi) = \frac{1}{2} \left(\mathbb{E}^\xi \left[\|\omega^0\|^2 \right] - \mathbb{E}^{\xi \circ \Delta^{-1}} \left[\|\omega^0\|^2 \right] \right) - \frac{1}{2(2\pi)^d} \int_{[-\pi,\pi]^d} \log \det \left(\tilde{\mathcal{K}}(\theta) \right) d\theta. \tag{22}$$

Otherwise, we define $\Gamma^\Delta(\xi) = 0$.

Lemma 7. *If* $\xi \in \mathcal{E}_2$,

$$\Gamma^\Delta(\xi) = \frac{1}{2(2\pi)^d} \int_{[-\pi,\pi]^d} \mathrm{tr} \left(\left(\mathrm{Id}_b - \tilde{\mathcal{K}}(\theta)^{-1} \right) d\tilde{\xi}(\theta) \right) - \frac{1}{2(2\pi)^d} \int_{[-\pi,\pi]^d} \log \det \left(\tilde{\mathcal{K}}(\theta) \right) d\theta. \tag{23}$$

Proof. We see from (21) that $\xi \circ \Delta^{-1}$ has spectral measure $\tilde{\mathcal{K}}^{-\frac{1}{2}}(\theta) d\tilde{\xi}(\theta) \tilde{\mathcal{K}}^{-\frac{1}{2}}(\theta)$. We thus find that:

$$\begin{aligned}
\Gamma^\Delta(\xi) =& \frac{1}{2(2\pi)^d} \int_{[-\pi,\pi]^d} \mathrm{tr} \left(d\tilde{\xi}(\theta) \right) - \frac{1}{2(2\pi)^d} \int_{[-\pi,\pi]^d} \mathrm{tr} \left(\tilde{\mathcal{K}}^{-\frac{1}{2}}(\theta) d\tilde{\xi}(\theta) \tilde{\mathcal{K}}^{-\frac{1}{2}}(\theta) \right) \\
& - \frac{1}{2(2\pi)^d} \int_{[-\pi,\pi]^d} \log \det \left(\tilde{\mathcal{K}}(\theta) \right) d\theta. \\
=& \frac{1}{2(2\pi)^d} \int_{[-\pi,\pi]^d} \mathrm{tr} \left(d\tilde{\xi}(\theta) \right) - \frac{1}{2(2\pi)^d} \int_{[-\pi,\pi]^d} \mathrm{tr} \left(\tilde{\mathcal{K}}(\theta)^{-1} d\tilde{\xi}(\theta) \right) \\
& - \frac{1}{2(2\pi)^d} \int_{[-\pi,\pi]^d} \log \det \left(\tilde{\mathcal{K}}(\theta) \right) d\theta.
\end{aligned}$$

$\square$

Lemma 8. *For all $\xi \in \mathcal{E}_2$,*

$$I^{(3)}\left(\xi \circ \Delta^{-1}\right) \leq I^{(3)}\left(\xi\right) - \Gamma^{\Delta}(\xi).$$

Proof. We assume for now that there exists a q, such that $\mathcal{S}^j = 0$ for all $|j| \geq q$, denoting the corresponding map by Δ_q. Let $\mathcal{N}_q^m : \mathcal{T}^{V_m} \to \mathcal{T}^{V_m}$ be the following linear operator. For $j \in V_m$, let $(\mathcal{N}_q^m \omega)^j = \sum_{k \in V_m} \mathcal{S}^{(k-j) \bmod V_m} \omega^k$. Let $\grave{\xi}_{q,n} = \xi^{V_n} \circ (\mathcal{N}_q^n)^{-1}$. It follows from this assumption that the V_l marginals of $\grave{\xi}_{q,n}$ and $\xi \circ \Delta_q^{-1}$ are the same, as long as $l \leq n - q$. Thus:

$$I^{(2)}\left((\xi \circ \Delta_q^{-1})^{V_l} || P^{\otimes V_l}\right) = I^{(2)}\left((\grave{\xi}_{q,n})^{V_l} || P^{\otimes V_l}\right)$$
$$\leq I^{(2)}\left(\grave{\xi}_{q,n} || P^{\otimes V_n}\right). \tag{24}$$

This last inequality follows from a property of the relative entropy $I^{(2)}$, namely that it is nondecreasing as we take a 'finer' σ-algebra (it is a direct consequence of Lemma 2.3 in [5]. If ξ^{V_n} does not have a density for some n, then $I^{(3)}(\xi)$ is infinite and the Lemma is trivial. If otherwise, we may readily evaluate $I^{(2)}\left(\grave{\xi}_{q,n} || P^{\otimes V_n}\right)$ using a change of variable to find that:

$$I^{(2)}\left(\grave{\xi}_{q,n} || P^{\otimes V_n}\right) = I^{(2)}\left(\xi^{V_n} || P^{\otimes V_n}\right) + \frac{1}{2}\mathbb{E}^{\xi^{V_n}}\left[\|\mathcal{N}_q^n \omega\|^2 - \|\omega\|^2\right] + \frac{1}{2}\log\det\left(\mathcal{N}_q^n\right).$$

We divide (24) by $(2l+1)^d$, substitute the above result and, finally, take $l \to \infty$ (while fixing $n = l + q$) to find that:

$$I^{(3)}\left(\xi \circ \Delta_q^{-1}\right)$$
$$\leq I^{(3)}\left(\xi\right) + \frac{1}{2}\left(\mathbb{E}^{\xi \circ (\Delta)^{-1}}\left[\|\omega^0\|^2\right] - \mathbb{E}^{\xi}\left[\|\omega^0\|^2\right]\right) + \frac{1}{2(2\pi)^d}\int_{[-\pi,\pi]^d} \log\det\left(\tilde{\mathcal{K}}(\theta)\right) d\theta$$
$$= I^{(3)}\left(\xi\right) - \Gamma_q^{\Delta}(\xi).$$

Here, $\Gamma_q^{\Delta}(\xi)$ is equal to $\Gamma^{\Delta}(\xi)$, as defined above, subject to the above assumption that $\mathcal{S}^j = 0$ for $|j| > q$. On taking $q \to \infty$, it may be readily seen that $\Gamma_q^{\Delta} \to \Gamma^{\Delta}$ pointwise. Furthermore, the lower semicontinuity of $I^{(3)}$ dictates that:

$$I^{(3)}\left(\xi \circ \Delta^{-1}\right) \leq \varliminf_{q \to \infty} I^{(3)}\left(\xi \circ \Delta_q^{-1}\right),$$

which gives us the Lemma. $\square$

Lemma 9. *For all $\xi \in \mathcal{E}_2$, $I^{(3)}\left(\xi \circ \Delta^{-1}\right) = I^{(3)}\left(\xi\right) - \Gamma^{\Delta}\left(\xi\right)$.*

Proof. We find, similarly to the previous Lemma, that if $\gamma \in \mathcal{M}_s(\mathcal{T}^{\mathbb{Z}^d})$, then:

$$I^{(3)}\left(\gamma \circ \Upsilon^{-1}\right) \leq I^{(3)}\left(\gamma\right) + \frac{1}{2}\left[\mathbb{E}^{\gamma \circ \Upsilon^{-1}}\left[\|\omega^0\|^2\right] - \mathbb{E}^{\gamma}\left[\|\omega^0\|^2\right]\right] - \frac{1}{2(2\pi)^d}\int_{[-\pi,\pi]^d} \log\det\tilde{\mathcal{K}}(\theta)d\theta. \tag{25}$$

We substitute $\gamma = \xi \circ \Delta^{-1}$ into the above and, after noting Lemma 6, we find that:

$$I^{(3)}\left(\xi\right) \leq I^{(3)}\left(\xi \circ \Delta^{-1}\right) + \frac{1}{2}\left(\mathbb{E}^{\xi}\left[\|\omega^0\|^2\right] - \mathbb{E}^{(\xi \circ \Delta^{-1})}\left[\|\omega^0\|^2\right]\right) - \frac{1}{2(2\pi)^d}\int_{[-\pi,\pi]^d} \log\det\tilde{\mathcal{K}}(\theta)d\theta$$
$$= I^{(3)}\left(\xi \circ \Delta^{-1}\right) + \Gamma^{\Delta}(\xi). \tag{26}$$

The result now follows from the previous Lemma and (26). $\square$

We next prove some matrix identities, which are needed in the proof of Lemma 11.

Lemma 10. *If $A, B \in \mathbb{C}^{l \times l}$ (for some positive integer l) are both Hermitian, then:*

$$\mathrm{tr}(AB) \leq \mathrm{b}\|A\|\|B\| \leq \begin{cases} l\|A\|\mathrm{tr}(B) & \text{if} \quad B \text{ is positive} \\ l\mathrm{tr}(A)\|B\| & \text{if} \quad A \text{ is positive} \end{cases} \tag{27}$$

If A and B are positive and, in addition, A is invertible, then:

$$\mathrm{tr}\,(B) \leq l\|A^{-1}\|^2\mathrm{tr}(ABA). \tag{28}$$

Proof. The first part of (27) follows from von Neumann's trace inequality. Given two matrices A and B in $\mathbb{C}^{l \times l}$:

$$|\mathrm{tr}(AB)| \leq \sum_{t=1}^{l} \alpha_t \beta_t,$$

where the α_t's and β_t's are the singular values of A and B. In the case where A and B are Hermitian, the singular values are the magnitudes of the (real) eigenvalues. By Cauchy–Schwartz, $\sum_{t=1}^{l} \alpha_t \beta_t \leq \sqrt{\sum_{t=1}^{l} \alpha_t^2} \times \sqrt{\sum_{t=1}^{l} \beta_t^2} \leq l\|A\|\,\|B\|$. If B is positive, so are its eigenvalues and $\|B\| \leq \mathrm{tr}(B)$, hence (27). If A is invertible, $\mathrm{tr}(B) = \mathrm{tr}(A^{-2}ABA)$. If, moreover, A and B are both Hermitian positive, we obtain the second identity by applying (27) to the Hermitian matrix A^{-2} and the Hermitian positive matrix ABA. $\square$

Lemma 11. *For all $0 < a < \frac{1}{2}$,*

$$a\mathbb{E}^{\mu}[\|\omega^0\|^2] + \frac{b}{2}\log(1 - 2a) \leq I^{(3)}(\mu). \tag{29}$$

For all $\mu \in \mathcal{E}_2$, there exist constants $\alpha_1 < 1$, $\alpha_2 \in \mathbb{R}$, $\alpha_3 > 1$, $\alpha_4 \in \mathbb{R}$ and $\check{m} > 0$, such that for all $m > \check{m}$,

$$\Gamma_{(m)}(\mu) \leq \alpha_1 I^{(3)}(\mu) + \alpha_2, \tag{30}$$
$$\Gamma(\mu) \leq \alpha_1 I^{(3)}(\mu) + \alpha_2. \tag{31}$$
$$\Gamma_{(m)}(\mu) \geq -\alpha_3 I^{(3)}(\mu) + \alpha_4 \tag{32}$$
$$\Gamma(\mu) \geq -\alpha_3 I^{(3)}(\mu) + \alpha_4 \tag{33}$$

There exist constants $\alpha_3^m, \alpha_4^m > 0$ that converge to zero as $m \to \infty$, such that:

$$\left|\Gamma_{(m)}(\mu) - \Gamma(\mu)\right| \leq \alpha_3^m E^{\mu}[\|\omega^0\|^2] + \alpha_4^m. \tag{34}$$

Proof. It is a standard result that if $\mu \notin \mathcal{M}_s(\mathcal{T}^{\mathbb{Z}^d})$, then $I^{(3)}(\mu) = \infty$, for which the first result is evident. Let $\mu \in \mathcal{M}_s(\mathcal{T}^{\mathbb{Z}^d})$. For $w \in \mathcal{T}^{V_n}$, we let $f(\omega) = \sum_{k \in V_n} \|\omega^k\|^2$. The function $f_M(\omega) = f(\omega)1_{af(\omega) \leq M}$ is bounded, and hence, from Definition 1, we have:

$$a\int_{\mathcal{T}^{V_n}} f_M \, d\mu^{V_n} \leq \log \int_{\mathcal{T}^{V_n}} \exp(af_M) \, dP^{\otimes V_n} + I^{(2)}(\mu^{V_n}||P^{\otimes V_n}).$$

We obtain using an easy Gaussian computation that:

$$\log \int_{\mathcal{T}^{V_n}} \exp(af) \, dP^{\otimes V_n} = -\frac{(2n + 1)^d b}{2}\log(1 - 2a).$$

Upon taking $M \to \infty$ and applying the dominated convergence theorem, we obtain:

$$a\mathbb{E}^\mu[\|\omega^0\|^2] \leq -\frac{b}{2}\log(1-2a) + \frac{1}{(2n+1)^d}I^{(2)}(\mu^{V_n}\|P^{\otimes V_n}).$$

We have used the stationarity of μ. By taking the limit $n \to \infty$, we obtain the first inequality (29).

It follows from the definition that (30)–(33) are true if $\mu \notin \mathcal{E}_2$. Thus, we may assume that $\mu \in \mathcal{E}_2$. We choose $\check{m}$ to be such that the eigenvalues of $\tilde{F}_{(m)}$ (as defined in (8)) are strictly greater than zero for all $m > \check{m}$. It may be easily verified that:

$$\sum_{s=1}^{b} \tilde{\mu}_{ss}\left([-\pi,\pi]^d\right) = (2\pi)^d \mathbb{E}^\mu\left[\|\omega^0\|^2\right]. \tag{35}$$

We observe the following upper and lower bounds, which hold for all $m > \check{m}$ (and for Γ_1, too),

$$\Gamma_{1,(m)} \leq -\frac{1}{2}\inf_{m\geq\check{m},\theta\in[-\pi,\pi]^d}\log\det\tilde{G}_{(m)}(\theta) < \infty \tag{36}$$

$$\Gamma_{1,(m)} \geq -\frac{1}{2}\sup_{m\geq\check{m},\theta\in[-\pi,\pi]^d}\log\det\tilde{G}_{(m)}(\theta) > -\infty \tag{37}$$

We recall that, since $\tilde{G}_{(m)} = \tilde{F}_{(m)}^2$,

$$\Gamma_{2,(m)}(\mu) = \frac{1}{2(2\pi)^d}\int_{[-\pi,\pi]^d}\text{tr}\left(\mathrm{d}\tilde{\mu}(\theta)\right) - \frac{1}{2(2\pi)^d}\int_{[-\pi,\pi]^d}\text{tr}\left(\tilde{F}_{(m)}^{-2}(\theta)\mathrm{d}\tilde{\mu}(\theta)\right).$$

Note that $\text{tr}(\tilde{F}_{(m)}^{-2}(\theta)d\tilde{\mu}(\theta)) = \text{tr}(\tilde{F}_{(m)}^{-1}(\theta)(\tilde{F}_{(m)}^{-1}(\theta)d\tilde{\mu}(\theta)\tilde{F}_{(m)}^{-1}(\theta))\tilde{F}_{(m)}(\theta)) = \text{tr}(\tilde{F}_{(m)}^{-1}(\theta)d\tilde{\mu}(\theta)\tilde{F}_{(m)}^{-1}(\theta))$ and apply Lemma 10, (28) to this, obtaining:

$$\Gamma_{2,(m)}(\mu) \leq \frac{1}{2}\left(1-\alpha_1^*\right)\mathbb{E}^\mu\left[\|\omega^0\|^2\right],$$

where:

$$\alpha_1^* = \frac{1}{b}\inf_{\theta\in[-\pi,\pi]^d,m\geq\check{m}}\left(\|\tilde{F}_{(m)}(\theta)\|^{-2}\right) > 0.$$

If $\alpha_1^* \geq 1$, then (30) is clear, because $I^{(3)}(\mu) \geq 0$; the above inequality would imply that $\Gamma_{2,(m)}$ is negative and (36). Otherwise, we may use (29) and (36) to find that:

$$\Gamma_{(m)}(\mu) \leq \frac{1}{2a}(1-\alpha_1^*)I^{(3)}(\mu) - \frac{b(1-\alpha_1^*)}{4a}\log(1-2a) - \frac{1}{2}\inf_{m\geq\check{m},\theta\in[-\pi,\pi]^d}\log\det\tilde{G}_{(m)}.$$

We may substitute $a > \frac{1}{2}(1-\alpha_1^*)$, letting $\alpha_1 = \frac{1}{2a}(1-\alpha_1^*)$, into the above to obtain (30). The second inequality (31) follows by taking $m \to \infty$ in the first.

For the third inequality, we find using (27) that:

$$\Gamma_{2,(m)}(\mu) \geq \frac{1}{2}\left(1-\alpha_3^*\right)E^\mu[\|\omega^0\|^2],$$

where:

$$\alpha_3^* = b\inf_{\theta\in[-\pi,\pi]^d,m\geq\check{m}}\left(\|\tilde{F}_{(m)}(\theta)\|^2\right).$$

If $\alpha_3^* \leq 1$, then (32) is clear, because of the fact that $I^{(3)}(\mu) \geq 0$, the above inequality would mean that $\Gamma_{2,(m)}$ is positive and (37). Otherwise, we may use (29) and (37) to find that:

$$\Gamma_{2,(m)}(\mu) \geq \frac{1-\alpha_3^*}{2a} I^{(3)}(\mu) + \frac{b(1-\alpha_3^*)}{4a} \log(1-2a) - \frac{1}{2} \sup_{m \geq \check{m}, \theta \in [-\pi,\pi]^d} \log \det \tilde{G}_{(m)}(\theta).$$

This yields (32) on taking $a < \frac{\alpha_3^*-1}{2}$. Taking limits as $m \to \infty$ yields (33).

Now, making use of Lemma 10, (27), it may be seen that:

$$\left|\Gamma_{2,(m)}(\mu) - \Gamma_2(\mu)\right| = \frac{1}{2(2\pi)^d} \left|\int_{[-\pi,\pi]^d} \mathrm{tr}\left(\left(\tilde{G}^{-1}(\theta) - \tilde{G}_{(m)}^{-1}(\theta)\right) \mathrm{d}\tilde{\mu}(\theta)\right)\right|.$$
$$\leq \frac{b}{2(2\pi)^d} \mathcal{G}_m \int_{[-\pi,\pi]^d} \mathrm{tr}(\mathrm{d}\tilde{\mu}(\theta)).$$

Here, $\mathcal{G}_m = \sup_{\theta \in [-\pi,\pi]^d} \|\tilde{G}^{-1}(\theta) - \tilde{G}_{(m)}^{-1}(\theta)\|$. The convergence of $\Gamma_{1,(m)}$ to Γ_1 is clear. We thus obtain (34). $\square$

Lemma 12. $I^{(3)}(\mu) - \Gamma(\mu)$ *is lower-semicontinuous.*

Proof. Since $I^{(3)}(\mu) - \Gamma(\mu)$ is infinite for all $\mu \notin \mathcal{E}_2$, we only need to prove the Lemma for $\mu \in \mathcal{E}_2$. We need to prove that if $\mu^{(j)} \to \mu$, then $\varliminf_{j\to\infty} I^{(3)}(\mu^{(j)} \circ \beta^{-1}) \geq I^{(3)}(\mu \circ \beta^{-1})$. We may assume, without loss of generality, that:

$$\varliminf_{j\to\infty} I^{(3)}(\mu^{(j)} \circ \beta^{-1}) = \lim_{j\to\infty} I^{(3)}(\mu^{(j)} \circ \beta^{-1}) \in \mathbb{R} \cup \infty.$$

If $\mathbb{E}^{\mu^{(j)}}[\|v^0\|^2] \to \infty$, then by (29) in Lemma 11, $\lim_{j\to\infty} I^{(3)}(\mu^{(j)} \circ \beta^{-1}) = \infty$, satisfying the requirements of the Lemma.

Thus, we may assume that there exists a constant l, such that $\mathbb{E}^{\mu^{(j)}}[\|\omega^0\|^2] \leq l$ for all j. We therefore have that, for all m, because of (11) and (4),

$$\lim_{j\to\infty} I^{(3)}(\mu^{(j)} \circ \beta^{-1}) = \lim_{j\to\infty} \left(I_{(m)}^{(3)}\left(\mu^{(j)}\right) + \Gamma_{(m)}\left(\mu^{(j)}\right) - \Gamma\left(\mu^{(j)}\right)\right).$$

Making use of (32), we may thus conclude that:

$$\lim_{j\to\infty} I^{(3)}(\mu^{(j)} \circ \beta^{-1}) \geq \lim_{j\to\infty} \left(I_{(m)}^{(3)}(\mu^{(j)}) - \epsilon_{(m)}^*\right),$$

for some $\epsilon_{(M)}^*$, which goes to zero as $m \to \infty$. In addition, $\varliminf_{j\to\infty} I_{(m)}^{(3)}(\mu^{(j)}) \geq I_{(m)}^{(3)}(\mu)$ due to the lower semi-continuity of $I_{(m)}^{(3)}$. On taking $m \to \infty$, since $I_{(m)}^{(3)}(\mu) \to I^{(3)}(\mu \circ \beta^{-1})$, we find that:

$$\varliminf_{j\to\infty} I^{(3)}(\mu^{(j)} \circ \beta^{-1}) \geq I^{(3)}(\mu^{(j)} \circ \beta^{-1}).$$

$\square$

B. A Lemma on the Large Deviations of Stationary Random Variables

The following lemma is an adaptation of Theorem 4.9 in [9] to $\mathbb{Z}^d$. We state it in a general context. Let $\mathfrak{B}$ be a Banach Space with norm $\| \cdot \|$. For $j \in \mathbb{Z}^d$, let $\lambda_j = 3^{-d} \prod_{\delta=1}^d 2^{-|j(\delta)|}$. We note that $\sum_{k \in \mathbb{Z}^d} \lambda_k = 1$. Define the metric d^λ on $\mathfrak{B}^{\mathbb{Z}^d}$ by:

$$d^\lambda(x, y) = \sum_{j \in \mathbb{Z}^d} \lambda_j \min \left(\|x^j - y^j\|, 1 \right).$$

Let the induced Prokhorov metric on $\mathcal{M}(\mathfrak{B}^{\mathbb{Z}^d})$ be $d^{\lambda,\mathcal{M}}$. For $\omega \in \mathfrak{B}^{\mathbb{Z}^d}$ and $j \in \mathbb{Z}^d$, we define the shift operator as $S^j(\omega)^k = \omega^{j+k}$. Let $B^\delta(A) = \{x \in \mathfrak{B}^{\mathbb{Z}^d} : d^\lambda(x, y) \leq \delta \text{ for some } y \in A\}$ be the closed blowup and $B(\delta)$ be the closed blowup of $\{0\}$.

Suppose that for $m \in \mathbb{Z}_+$, $Y_{(m)}, W \in \mathfrak{B}^{\mathbb{Z}^d}$ are stationary random variables, governed by a probability law $\mathbb{P}$. We suppose that $\hat{\mu}_n(Y_{(m)})$ is governed by $\Pi^n_{(m)}$ and $\hat{\mu}^n(W)$ is governed by Π^n_W; these being the empirical process measures, defined analogously by (2). Suppose that for each m, $(\Pi^n_{(m)})$ satisfies an LDP with good rate function $J_{(m)}$. Suppose that $W = Y_{(m)} + Z_{(m)}$ for some series of stationary random variables $Z_{(m)}$ on $\mathfrak{B}^{\mathbb{Z}^d}$.

Lemma 13. *If there exists a constant $\kappa > 0$, such that for all $\mathfrak{b} > 0$:*

$$\varlimsup_{m \to \infty} \varlimsup_{n \to \infty} \frac{1}{(2n+1)^d} \log E \left[\exp \left(\mathfrak{b} \sum_{j \in V_n} \min \left(\|Z_{(m)}^j\|^2, 1 \right) \right) \right] < \kappa, \tag{38}$$

then (Π^n_W) satisfies an LDP with good rate function:

$$J(x) = \lim_{\epsilon \to 0} \varliminf_{m \to \infty} \inf_{y \in B^\epsilon(x)} J_{(m)}(y) = \lim_{\epsilon \to 0} \varlimsup_{m \to \infty} \inf_{y \in B^\epsilon(x)} J_{(m)}(y).$$

Proof. It suffices, thanks to Theorem 4.2.16, Exercise 4.2.29 in [16], to prove that for all $\epsilon > 0$,

$$\lim_{m \to \infty} \varlimsup_{n \to \infty} \frac{1}{(2n+1)^d} \log \mathbb{P} \left(d^{\lambda,\mathcal{M}}(\hat{\mu}^n(W), \hat{\mu}^n(Y_{(m)})) > \epsilon \right) = -\infty. \tag{39}$$

For $x \in \mathfrak{B}^{\mathbb{Z}^d}$, write $|x|_\lambda := d^\lambda(x, 0)$. Let $B \in \mathcal{B}(\mathfrak{B}^{\mathbb{Z}^d})$. Then, noting the definition of p_n just above (2),

$$\begin{aligned}
\hat{\mu}^n(W)(B) =& \frac{1}{(2n+1)^d} \sum_{j \in V_n} 1_B \left(S^j p_n(Y_{(m)}) + S^j p_n(Z_{(m)}) \right) \\
\leq& \frac{1}{(2n+1)^d} \sum_{j \in V_n} \left\{ 1_B \left(S^j p_n(Y_{(m)}) + S^j p_n(Z_{(m)}) \right) 1_{B(\epsilon)}(S^j p_n(Z_{(m)})) + 1_{B(\epsilon)^c} \left(S^j p_n(Z_{(m)}) \right) \right\} \\
\leq& \frac{1}{(2n+1)^d} \sum_{j \in V_n} 1_{B^\epsilon}(S^j p_n(Y_{(m)})) + \frac{1}{(2n+1)^d} \#\{j \in V_n : |S^j p_n(Z_{(m)})|_\lambda > \epsilon\} \\
\leq& \hat{\mu}^n(Y_{(m)})(B^\epsilon) + \frac{1}{(2n+1)^d} \# \left\{ j \in V_n : |S^j p_n(Z_{(m)})|_\lambda > \epsilon \right\}.
\end{aligned}$$

Thus:

$$
\mathbb{P}\left(d^{\lambda,\mathcal{M}}(\hat{\mu}^n(W), \hat{\mu}^n(Y_{(m)})) > \epsilon\right)
$$

$$
\leq \mathbb{P}\left(\frac{1}{(2n+1)^d}\#\left\{j \in V_n : |S^j p_n(Z_{(m)})|_\lambda > \epsilon\right\} > \epsilon\right)
$$

$$
\leq \mathbb{P}\left(\sum_{j \in V_n} |S^j p_n(Z_{(m)})|_\lambda^2 > (2n+1)^d \epsilon^3\right)
$$

$$
\leq \exp\left(-\mathfrak{b}(2n+1)^d \epsilon^3\right) E^{\mathbb{P}}\left[\exp\left(\mathfrak{b}\sum_{j \in V_n} |S^j p_n(Z_{(m)})|_\lambda^2\right)\right]
$$

$$
\leq \exp\left(-\mathfrak{b}(2n+1)^d \epsilon^3\right) E^{\mathbb{P}}\left[\exp\left(\mathfrak{b}\sum_{j \in V_n, k \in \mathbb{Z}^d} \lambda_k \min(\|p_n(Z_{(m)})^{j+k}\|^2, 1)\right)\right]
$$

for an arbitrary $\mathfrak{b} > 0$. Since $\sum_{k \in \mathbb{Z}^d} \lambda_k = 1$ and the exponential function is convex, by Jensen's inequality:

$$
\exp\left(-\mathfrak{b}(2n+1)^d \epsilon^3\right) E^{\mathbb{P}}\left[\exp\left(\mathfrak{b}\sum_{j \in V_n, k \in \mathbb{Z}^d} \lambda_k \min(\|p_n(Z_{(m)})^{j+k}\|^2, 1)\right)\right]
$$

$$
\leq \exp\left(-\mathfrak{b}(2n+1)^d \epsilon^3\right) E^{\mathbb{P}}\left[\sum_{k \in \mathbb{Z}^d} \lambda_k \exp\left(\mathfrak{b}\sum_{j \in V_n} \min(\|p_n(Z_{(m)})^{j+k}\|^2, 1)\right)\right]
$$

$$
= \exp\left(-\mathfrak{b}(2n+1)^d \epsilon^3\right) E^{\mathbb{P}}\left[\exp\left(\mathfrak{b}\sum_{j \in V_n} \min(\|Z_{(m)}^j\|^2, 1)\right)\right],
$$

by the stationarity of $Z_{(m)}$ and the fact that $\sum_{k \in \mathbb{Z}^d} \lambda_k = 1$. We may thus infer, using (38), that:

$$
\varlimsup_{m \to \infty} \varlimsup_{n \to \infty} \frac{1}{(2n+1)^d} \log \mathbb{P}\left(d^{\lambda,\mathcal{M}}(\hat{\mu}^n(W), \hat{\mu}^n(Y_{(m)})) > \epsilon\right) \leq -\mathfrak{b}\epsilon^3 + \kappa. \tag{40}
$$

Since $\mathfrak{b}$ is arbitrary, we may take $\mathfrak{b} \to \infty$ to obtain (39). $\square$

C. Proof of Corollary 2

We now prove Corollary 2.

Proof. Let $\phi : \mathcal{T} \to \mathcal{T}$ be $\phi(\omega) = \omega + c$ and $\phi^{\mathbb{Z}^d} : \mathcal{T}^{\mathbb{Z}^d} \to \mathcal{T}^{\mathbb{Z}^d}$ be $\phi^{\mathbb{Z}^d}(\omega)^j = \phi(\omega^j)$. Let $\Psi^{\mathbb{Z}^d} : \mathcal{M}_s(\mathcal{T}^{\mathbb{Z}^d}) \to \mathcal{M}_s(\mathcal{T}^{\mathbb{Z}^d})$ be $\Psi^{\mathbb{Z}^d}(\mu) = \mu \circ (\phi^{\mathbb{Z}^d})^{-1}$ and $\Psi : \mathcal{M}_s(\mathcal{T}) \to \mathcal{M}_s(\mathcal{T})$ be $\Psi(\nu) = \nu \circ \phi^{-1}$. It is easily checked that these maps are bicontinuous bijections for their respective topologies. Since $\Pi_c^n = \Pi^n \circ (\Psi^{\mathbb{Z}^d})^{-1}$, we have, by a contraction principle, Theorem 4.2.1 in [16], that (Π_c^n) satisfies a strong LDP with good rate function:

$$
I^{(3)}((\Psi^{\mathbb{Z}^d})^{-1}(\mu)) - \Gamma((\Psi^{\mathbb{Z}^d})^{-1}(\mu)). \tag{41}
$$

Clearly, $(\Psi^{\mathbb{Z}^d})^{-1}(\mu)$ is in $\mathcal{E}_2$ if and only if μ is in $\mathcal{E}_2$. Let $\nu = (\Psi^{\mathbb{Z}^d})^{-1}(\mu)$. It is well known that if $(\Psi^{\mathbb{Z}^d})^{-1}(\mu)^{V_n}$ is absolutely continuous relative to $P^{\otimes V_n}$, then the relative entropy may be written as:

$$I^{(2)}(((\Psi^{\mathbb{Z}^d})^{-1}(\mu))^{V_n}||P^{\otimes V_n}) = E^{((\Psi^{\mathbb{Z}^d})^{-1}(\mu))^{V_n}(x)}\left[\log \frac{d((\Psi^{\mathbb{Z}^d})^{-1}(\mu))^{V_n}}{dP^{\otimes V_n}}(x)\right]$$

$$= E^{\mu^{V_n}(x)}\left[\log \frac{d\mu^{V_n}}{d(\Psi(P))^{\otimes V_n}}(x)\right] = I^{(2)}(\mu^{V_n}||(\Psi(P))^{\otimes V_n}).$$

Otherwise, the relative entropy is infinite. Thus, if the relative entropy is finite, $(\Psi^{\mathbb{Z}^d})^{-1}(\mu)^{V_n}$ must possess a density, and this means that μ^{V_n} possesses a density, which we denote by $r(x): \mathcal{T}^{V_n} \to \mathcal{T}^{V_n}$. We note that the density of $(\Psi(P))^{\otimes V_n}$ is:

$$\rho^{V_n}(x) = (2\pi)^{-\frac{(2n+1)d_b}{2}} \exp\left(-\frac{1}{2}\sum_{j \in V_n} \|x^j - c\|^2\right). \tag{42}$$

Accordingly, we find that:

$$I^{(2)}\left(\mu^{V_n}||(\Psi(P))^{\otimes V_n}\right)$$
$$= E^{\mu^{V_n}(x)}\left[\log\left(\frac{r(x)}{\rho^{V_n}(x)}\right)\right]$$
$$= I^{(2)}\left(\mu^{V_n}||P^{\otimes V_n}\right) - (2n+1)^{d\dagger}m^\mu c + \frac{(2n+1)^d}{2}\|c\|^2.$$

We divide by $(2n+1)^d$ and take n to infinity to obtain:

$$I^{(3)}((\Psi^{\mathbb{Z}^d})^{-1}(\mu)) = I^{(3)}(\mu) - {}^\dagger m^\mu c + \frac{1}{2}\|c\|^2.$$

If μ^{V_n} does not possess a density for some n, then both sides of the above equation are infinite. It may be verified that the spectral density of ν is given by:

$$d\tilde{\nu}(\theta) = d\tilde{\mu}(\theta) + (2\pi)^d\delta(\theta)(c^\dagger c - m^{\mu\dagger}c - c^\dagger m^\mu).$$

On substituting this into the expression in Theorem 1, we find that:

$$\Gamma_2(\nu) = \Gamma_2(\mu) + \frac{1}{2}\mathrm{tr}\left(\left(\mathrm{Id_b} - \tilde{G}(0)^{-1}\right)c^\dagger c\right) - \frac{1}{2}\mathrm{tr}\left(\left(\mathrm{Id_b} - \tilde{G}(0)^{-1}\right)\left(m^{\mu\dagger}c + c^\dagger m^\mu\right)\right)$$
$$= \Gamma_2(\mu) + \frac{1}{2}{}^\dagger c\left(\mathrm{Id_b} - \tilde{G}(0)^{-1}\right)c - {}^\dagger c\left(\mathrm{Id_b} - \tilde{G}(0)^{-1}\right)m^\mu.$$

We have used the fact that $\tilde{G}(0)$ is symmetric. We thus obtain (5). This minimum of the rate function remains unique because of the bijectivity of $(\Psi^{\mathbb{Z}^d})$. $\square$

Acknowledgments

Author Contributions

Both authors contributed to all the article.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Baladron, J.; Fasoli, D.; Faugeras, O.; Touboul, J. Mean-field description and propagation of chaos in networks of Hodgkin-Huxley and Fitzhugh-Nagumo neurons. *J. Math. Neurosci.* **2012**, *2*, doi:10.1186/2190-8567-2-10.

2. Faugeras, O.; Maclaurin, J. Asymptotic description of neural networks with correlated synaptic weights; Rapport de recherche RR-8495; INRIA: Sophia Antipolis, France, March 2014.

3. Faugeras, O.; Touboul, J.; Cessac, B. A constructive mean field analysis of multi population neural networks with random synaptic weights and stochastic inputs. *Front. Comput. Neurosci.* **2009**, *3*, doi: 10.3389/neuro.10.001.2009.

4. Rolls, E.T.; Deco, G. *The Noisy Brain: Stochastic Dynamics as a Principle of Brain Function*; Oxford university press: Oxford, UK, 2010.

5. Donsker, M.D.; Varadhan, S.R.S. Asymptotic evaluation of certain markov process expectations for large time, iv. *Comm. Pure Appl. Math.* **1983**, *XXXVI*, 183–212.

6. Chiyonobu, T.; Kusuoka, S. The large deviation principle for hypermixing processes. *Probab. Theor. Relat. Field.* **1988**, *78*, 627–649.

7. Bryc, W.; Dembo, A. Large deviations and strong mixing. In *Annales de l'IHP Probabilités et statistiques*; Elsevier: Amsterdam, The Netherlands, 1996; Volume 32, pp. 549–569.

8. Deuschel, J.D.; Stroock, D.W.; Zessin, H. Microcanonical distributions for lattice gases. *Comm. Math. Phys.* **1991**, *139*. 83–101.

9. Baxter, J.R.; Jain, N.C. An approximation condition for large deviations and some applications. In *Convergence in Ergodic Theory and Probability*; Bergulson, V., Ed.; Ohio State University, Mathematical Research Institute: Columbus, OL, USA, 1993.

10. Steinberg, Y.; Zeitouni, O. On tests for normality. *IEEE Trans. Inf. Theory* **1992**, *38*. 1779–1787.

11. Bryc, W.; Dembo, A. On large deviations of empirical measures for stationary gaussian processes. *Stoch. Process. Appl.* **1995**, *58*, 23–34.

12. Faugeras, O.; MacLaurin, J. Asymptotic description of stochastic neural networks. I. Existence of a large deviation principle. *Comptes Rendus Mathematiques* **2014**, *352*, No. 10.

13. Faugeras, O.; MacLaurin, J. Large deviations of a spatially-ergodic neural network with learning. **2014**, arXiv:1404.0732.

14. Donsker, M.D.; Varadhan, S.R.S. Large deviations for stationary Gaussian processes. *Commun. Math. Phys.* **1985**, *97*, 187–210.

15. Shiryaev, A.N. *Probability*; Springer: Berlin, Germany, 1996.

16. Dembo, A.; Zeitouni, O. *Large Deviations Techniques*, 2nd ed.; Springer, Berlin, Germany, 1997.