Conditional Large Deviation Principle for Finite State Gibbs Random Fields

Zhiyi Chi[1]
April, 2002

Abstract. Let $X = \{X_t\}_{t\in\mathbb{Z}^d} \sim P$ and $Y = \{Y_t\}_{t\in\mathbb{Z}^d} \sim Q$ be two independent stationary random fields with finite state spaces. Suppose Y is a Gibbs field with summable potential. Given a random realization x of X, the conditional large deviation principle (LDP) associated with $(x_t, Y_t)_{t\in\mathbb{Z}^d}$ are established at 3 levels, for empirical means, marginals, and field. In general, the rate function is a random variable. However, when X is ergodic, the rate function is deterministic and a variational characterization of the conditional LDP in terms of the specific relative entropy with respect to $P\times Q$ is established. We then prove a factorization formula which characterizes the LDP associated with $P\times Q$ in terms of the LDP associated with P and the "quasi-quenched" LDP associated with $P\times Q$ relative to P. Finally, we show that if Q is a stationary Gibbs field with summable potential, then for any stationary P, the "quasi-quenched" LDP associated with $P \times Q$ relative to P exists and is equal to the expected value of the quenched LDP. As a consequence of the results, if the LDP holds for P, then the LDP holds for $P \times Q$ as well.

[1]Department of Statistics, The University of Chicago. Address: 5734 University Avenue, Chicago, IL 60637. E-mail: chi@galton.uchicago.edu.

[2]*MSC* 2000 *subject classifications*: Primary 60F10, secondary 60G60.

[3]*Key words and phrases*: conditional large deviation principle, Gibbs random field, factorization of LDP.

1 Introduction

In this article, we consider the conditional LDP for Gibbs random fields and related issues.

1.1 Conditional LDP

Let $X = \{X_t\}_{t\in\mathbb{Z}^d} \sim P$ and $Y = \{Y_t\}_{t\in\mathbb{Z}^d} \sim Q$ be two independent stationary random fields, with respective finite state spaces S_X and S_Y. Given a random realization x of X, there are three versions of LDP for $(x, Y) = (x_t, Y_t)_{t\in\mathbb{Z}^d}$, one for the empirical means, one for the empirical marginal fields, and one for the empirical field ("the process level LDP"), all of which henceforth are referred to as the conditional LDP of (X, Y), given $X = x$. The conditional LDP is also often termed "quenched LDP" because it is established for a fixed, though random, value of X. Several cases of conditional LDP were established by Comets [6], when both X and Y are i.i.d. In recent years, the topic has been actively studied in the context of lossy data compression using random code books (cf. [9, 17, 4, 10] and references therein). Results in this regard are referred to as generalized Asymptotic Equipartition Property (AEP). Except for [10], where the generalized AEP for random fields is considered, the focus of the studies has been on the conditional LDP for empirical means for processes on $\mathbb{Z}$ or $\mathbb{R}$ satisfying various mixing conditions.

We shall consider the conditional LDP for (X, Y) given $X = x$ when X is a stationary random field and Y is a Gibbs random field defined on $\mathbb{Z}^d$. For the topic of the (unconditional) LDP for Gibbs random fields, we refer the reader to [13]. First, fix some general notations. Let S be an arbitrary finite set, $D \subset \mathbb{Z}^d$, and $n, \nu \in \mathbb{N}$. Then

$S^{\mathbb{Z}^d}$ = the compact space of functions $\omega : \mathbb{Z}^d \to S$ equipped with product topology;
θ_t = the shift operator on $S^{\mathbb{Z}^d}$, $\theta_t x \mapsto \tilde{x}$, $\tilde{x}_s = x_{s+t}$, $\forall x \in S^{\mathbb{Z}^d}$, $\forall t \in \mathbb{Z}^d$;
$\mathcal{M}(S^D)$ = the compact space of probability measures on S^D topologized by weak convergence;
$\mathcal{M}_s(S^{\mathbb{Z}^d})$ = the closed subspace of $\mathcal{M}(S^{\mathbb{Z}^d})$ consisting of P such that $P = P \circ \theta_t^{-1}$, $\forall t \in \mathbb{Z}^d$;
z_D = the restriction of z on D, $\forall z \in S^{\mathbb{Z}^d}$;
C_n = the block $\{(t_1, \ldots, t_d) \in \mathbb{Z}^d : \ 0 \le t_i < n\}$;
$B(S^D, \mathbb{R}^\nu) = \{f : f \text{ is a bounded function } S^D \to \mathbb{R}^\nu\}$;
$\|f\| = \sup_{x\in S^D} |f(x)|$, $\forall f \in B(S^D, \mathbb{R}^\nu)$;
$B_r = \{u \in \mathbb{R}^\nu : |u| < r\}$, $r > 0$.

For the particular problems considered here, we also introduce the following notations

$\Sigma = S_X \times S_Y$, $\Omega_X = S_X^{\mathbb{Z}^d}$, $\Omega_Y = S_Y^{\mathbb{Z}^d}$,
$\mathcal{B}_X$ (resp. $\mathcal{B}_Y$) = the product topology of Ω_X (resp. Ω_Y);
$\mathcal{G}(U)$ = the class of all Gibbs measures with respect to (**wrt**) a stationary summable interaction potential U on $S_Y^{\mathbb{Z}^d}$;
$\mathcal{I}$ = the σ-algebra of shift invariant sets of Ω_X.

Suppose $X \sim P \in \mathcal{M}_s(\Omega_X)$ and $Y \sim Q \in \mathcal{G}(U) \cap \mathcal{M}_s(\Omega_Y)$ are independent. Since $(\Omega_X, \mathcal{B}_X)$ is a Borel space, there is a regular conditional distribution for X given $\mathcal{I}$, denoted $P(\,\cdot\,|\mathcal{I})(x)$, $x \in \Omega_X$, which is stationary and ergodic (section 4.3, [1]). Our result on the conditional LDP for the empirical means associated with (X, Y), given $X = x$, is as follows.

Theorem 1 (Conditional LDP for empirical means) For P-almost all $x \in \Omega_X$, for any $n_0, \nu \in \mathbb{N}$, and $f \in B(\Sigma^D, \mathbb{R}^\nu)$, with $D = C_{n_0}$, the empirical means

$$\frac{1}{n^d} \sum_{t \in C_n} f((\theta_t x, \theta_t Y)_D), \quad n \in \mathbb{N} \tag{1.1}$$

satisfy the LDP with a good rate function $I(\,\cdot\,; x) = \Lambda_x^*(\,\cdot\,)$ parameterized by x, where

$$\Lambda_x^*(u) = \sup_{\lambda \in \mathbb{R}^\nu} \{\langle \lambda, u\rangle - \Lambda_x(\lambda)\} \tag{1.2}$$

$$\Lambda_x(\lambda) = \lim_{n\to\infty} \Lambda_{n,x}(\lambda), \quad \Lambda_{n,x}(\lambda) = \Lambda_{C_n,x}(\lambda) \tag{1.3}$$

and for any finite $V \subset \mathbb{Z}^d$,

$$\Lambda_{V,x}(\lambda) = \frac{1}{|V|} E_P\left[\log E_Q \exp\left\{\sum_{t\in V} \langle \lambda, f((\theta_t\,\cdot\,, \theta_t Y)_D)\rangle\right\} \Big| \mathcal{I}\right](x). \tag{1.4}$$

Remark: When P is ergodic, the above conditional LDP has a deterministic good rate function. □

Given a probability space Ω, for any $\omega \in \Omega$, denote by δ_ω the probability measure on Ω, such that $\delta_\omega(B) = \mathbf{1}_B(\omega)$, $B \subset \Omega$ measurable. Given $X = x$, the empirical marginal measures are random elements of $\mathcal{M}(\Sigma^D)$ given by

$$\hat{P}_{x,C_n,D} = \frac{1}{n^d} \sum_{t\in C_n} \delta_{(\theta_t x, \theta_t Y)_D} \tag{1.5}$$

and the empirical fields are random elements of $\mathcal{M}(\Sigma^{\mathbb{Z}^d})$ given by

$$\hat{P}_{x,C_n} = \frac{1}{n^d} \sum_{t\in C_n} \delta_{(\theta_t x, \theta_t Y)}. \tag{1.6}$$

On the conditional LDP for the empirical marginals and fields, we have

Theorem 2 (Conditional LDP for empirical marginals and fields) For P-almost all $x \in \Omega_X$, the following hold.

(a) For any finite $D \subset \mathbb{Z}^d$, the empirical measures (1.5) satisfy the LDP in $\mathcal{M}(\Sigma^D)$ wrt the weak topology, with a convex good rate function parameterized by x,

$$I_D(\pi; x) = \sup_{f \in B(\Sigma^D, \mathbb{R})} \{E_\pi f - \Lambda_x(f)\} \tag{1.7}$$

where

$$\Lambda_x(f) = \lim_{n\to\infty} \frac{1}{n^d} E_P\left[\log E_Q \exp\left\{\sum_{t\in C_n} f((\theta_t\,\cdot\,,\, \theta_t Y)_D)\right\} \Big| \mathcal{I}\right](x). \tag{1.8}$$

(b) The empirical fields (1.6) satisfy the LDP in $\mathcal{M}(\Sigma^{\mathbb{Z}^d})$ wrt the weak topology, and with a convex good rate function parameterized by x,

$$I(\pi; x) = \sup\{I_{C_n}(\pi_{C_n}; x):\ n \geq 1\}. \tag{1.9}$$

Furthermore,

$$I(\pi; x) = \infty \quad \text{if } \pi \notin \mathcal{M}_s(\Sigma^{\mathbb{Z}^d}).$$

Remark: When P and Q are i.i.d., (1.7) is implied by Theorem IV.1 of [6]. □

Theorem 1 is established by adapting the asymptotic value method [2] and an idea of [13] to divide a block into smaller blocks, with modifications to handle the randomness introduced by X. By the separability of the space

$$B_c = \bigcup_{\nu=1}^{\infty} \bigcup_{D\subset\mathbb{Z}^d:\ |D|<\infty} B(\Sigma^D, \mathbb{R}^\nu),$$

it is seen that almost surely, given $X = x$, the LDP for the empirical means holds simultaneously for all $f \in B_c$. Then different levels of conditional LDP in Theorem 2 are obtained by lifting the LDP for the empirical means using Dawson-Gärtner projective limit theorem [8].

1.2 Variational characterization of conditional LDP

From the previous subsection, it is seen that when X is stationary, the conditional LDP in general is random. For example, the process level LDP for (X, Y) conditioning on X has a $(\mathcal{B} \times \mathcal{I})$-measurable rate function $I(\pi; x)$, with $\mathcal{B}$ the Borel σ-algebra induced by the weak topology of $\mathcal{M}(\Sigma^{\mathbb{Z}^d})$ (Proposition 1). When the stationary random field X is ergodic, the rate function is deterministic and, under certain circumstances, can be characterized as a constrained maximum function.

In data compression, a variational characterization of the rate function is established as follows [9, 17, 4, 10]. Suppose X and Y both are stationary and ergodic and Y satisfies some strong mixing condition. Then given $f \in B(\Sigma, \mathbb{R})$, for any $u < Ef(X, Y)$, almost surely, conditioning on $X = x$,

$$\lim_{n\to\infty} -\frac{1}{n^d} \log \Pr\left\{ \frac{1}{n^d} \sum_{t\in C_n} f(x_t, Y_t) \le u \right\} = R(P, Q, u) = \lim_{n\to\infty} \frac{1}{n^d} \inf_{V_n\in M_n(u)} H(V_n \| \Pi_n). \quad (1.10)$$

where $\Pi_n = (P \times Q)_{C_n}$ is the marginal of $P \times Q$ on Σ^{C_n},

$$H(\mu\|\nu) = \sum_{z\in\Sigma^{C_n}} \mu(z) \log \frac{\mu(z)}{\nu(z)}, \quad (1.11)$$

is the relative entropy of μ wrt ν, $\mu, \nu \in \mathcal{M}(\Sigma^{C_n})$, and

$$M_n(u) = \left\{ V_n \in \mathcal{M}(\Sigma^{C_n}) : S_X^{C_n}\text{-marginal of } V_n = P_{C_n},\ \text{and } \frac{1}{n^d} E_{V_n} \sum_{i\in C_n} f(x_i, Y_i) \le u \right\}.$$

It would be interesting to see if the rate functions of the conditional LDP have similar variational characterizations, in particular, ones in terms of the specific relative entropy wrt $P \times Q$, at least for the case where X is ergodic. Recall that given $\mu, \nu \in \mathcal{M}(\Sigma^{\mathbb{Z}^d})$, and $V \subset \mathbb{Z}^d$, letting $H_V(\mu\|\nu) = H_V(\mu_V\|\nu_V)$, the specific relative entropy of ν wrt μ is given by

$$h(\mu\|\nu) = \lim_{n\to\infty} \frac{1}{n^d} H_{C_n}(\mu\|\nu) \quad (1.12)$$

provided the limit exists.

In [6], for $X \sim P$ and $Y \sim Q$ both i.i.d., it was shown that the empirical means (1.1) satisfy the LDP wrt the deterministic convex rate function $J(u) = \inf h(\pi\|P \times Q)$, where the infimum is taken over all stationary random fields π on $\Sigma^{\mathbb{Z}^d}$ with Ω_X-marginal equal to P and $E_\pi f(X_1, Y_1) = u$. This constrained variational characterization was established first for the empirical fields, then for the empirical means by the contraction principle (Theorem III.1, [6]).

We shall establish variational characterizations for the conditional LDP when both P and Q are Gibbs. Comparing to the i.i.d. case in [6], it is now more difficult to directly establish the LDP for the empirical fields. Instead, we shall first deal with the LDP for the empirical means as in equation (1.10). Then we shall obtain a variational characterization for the LDP for the empirical fields by projective limit. The results by this approach are summarized as follows.

Theorem 3 (Variational characterization of the LDP of empirical means) Suppose that, for some summable interaction potential U^P, $P \in \mathcal{G}(U^P) \cap \mathcal{M}_s(\Omega_X)$ and is ergodic. In addition, suppose $Q \in \mathcal{G}(U) \cap \mathcal{M}_s(\Omega_Y)$. Define functions $\Lambda(\lambda)$, $\Lambda_n(\lambda)$, $\Lambda^*(\lambda)$ by (1.2)–(1.4), with x being omitted from the subscript, as all these functions are deterministic. Define

$$\Lambda_n^*(u) = \sup_{\lambda \in \mathbb{R}} \{\lambda u - \Lambda_n(\lambda)\}.$$

For $\pi \in \mathcal{M}(\Sigma^{\mathbb{Z}^d})$, denote by π_X the marginal distribution of Ω_X. Define

$$u_{\min}^{(\infty)} = \inf\{u : \sup_{n\ge 1} \Lambda_n^*(u) < \infty\} \qquad u_{\max}^{(\infty)} = \sup\{u : \sup_{n\ge 1} \Lambda_n^*(u) < \infty\} \tag{1.13}$$

Then for any u,

$$\Lambda^*(u) = J(u) = \lim_{\epsilon \to 0} \inf\{h(\pi \| P \times Q) : \ \pi \in \mathcal{M}_s(\Sigma^{\mathbb{Z}^d}), \ \pi_X = P, \ |E_\pi f - u| < \epsilon\}. \tag{1.14}$$

Moreover, for $u \ne u_{\min}^{(\infty)}, u_{\max}^{(\infty)}$,

$$\Lambda^*(u) = \tilde{J}(u) = \inf\{h(\pi \| P \times Q) : \ \pi \in \mathcal{M}_s(\Sigma^{\mathbb{Z}^d}), \ \pi_X = P, \ E_\pi f = u\}. \qquad \square \tag{1.15}$$

Remark: Since $P \times Q$ is a stationary Gibbs random field on $\Sigma^{\mathbb{Z}^d}$, $h(\,\cdot\,\| P \times Q)$ exists and is lower semi-continuous on $\mathcal{M}_s(\Sigma^{\mathbb{Z}^d})$ (cf. Eqs (2.15-16), [13]).

Corollary 1 (Variational characterization of empirical marginal fields) The rate function $I_D(\pi)$, $\pi \in \mathcal{M}(\Sigma^{\mathbb{Z}^d})$ in (1.7) has the following variational characterization

$$I_D(\pi) = \lim_{\epsilon \to 0} \sup_{\substack{f \in B(\Sigma^D, \mathbb{R}) \\ \|f\|=1}} \inf\{h(\gamma \| P \times Q) : \ \gamma \in \mathcal{M}_s(\Sigma^{\mathbb{Z}^d}), \ \gamma_X = P, \ |E_\gamma f - E_\pi f| < \epsilon\}. \tag{1.16}$$

1.3 The "quasi-quenched" LDP and a factorization formula

The results in the previous subsection expresses different levels of the conditional LDP associated with X and Y in terms of the product of the two random fields. Conversely, it may be asked whether the LDP for the product of two random fields can be expressed in terms of the conditional LDP. The main result we obtain in this regard is that if one of the two random fields satisfies the process level LDP, and the other one satisfies the so called "quasi-quenched" process level LDP defined in a moment (Definition 1), then their product satisfies the process level LDP.

First, given $x \in \Omega_X$, $y \in \Omega_Y$, define the empirical fields $\hat{P}_{x,C_n} \in \mathcal{M}(\Omega_X)$ and $\hat{P}_{x,y,C_n} \in \mathcal{M}(\Sigma^{\mathbb{Z}^d})$ by

$$\hat{P}_{x,C_n}(\,\cdot\,) = \frac{1}{n^d} \sum_{t \in C_n} \delta_{\theta_t x}(\,\cdot\,), \qquad \hat{P}_{x,y,C_n}(\,\cdot\,) = \frac{1}{n^d} \sum_{t \in C_n} \delta_{\theta_t x, \theta_t y}(\,\cdot\,). \tag{1.17}$$

Definition 1 ("Quasi-quenched LDP") Suppose $\mu \in \mathcal{M}(\Omega_X)$ and $\nu \in \mathcal{M}(\Omega_Y)$. If there is a rate function I on $\mathcal{M}(\Sigma^{\mathbb{Z}^d})$, such that for any $\{x^{(n)},\ n \in \mathbb{N}\} \subset \Omega_X$ satisfying $\hat{P}_{x^{(n)},C_n} \to \mu$ in weak topology, the following two bounds hold,

$$\liminf_{n\to\infty} \frac{1}{n^d} \log \nu\{\hat{P}_{x^{(n)},Y,C_n} \in G\} \ge -\inf_{\pi\in G} I(\pi), \quad \forall G \subset \mathcal{M}(\Sigma^{\mathbb{Z}^d}) \text{ open}$$
$$\limsup_{n\to\infty} \frac{1}{n^d} \log \nu\{\hat{P}_{x^{(n)},Y,C_n} \in F\} \le -\inf_{\pi\in F} I(\pi), \quad \forall F \subset \mathcal{M}(\Sigma^{\mathbb{Z}^d}) \text{ closed}$$

then (μ, ν) are said to satisfy the quasi-quenched process level LDP relative to μ. The rate function I will be denoted by $I_{\nu|\mu}$.

Remark: Because $\mathcal{M}(\Sigma^{\mathbb{Z}^d})$ is compact, I is necessarily a good rate function. □

The term "quasi-quenched LDP" is used to make distinction with the conditional LDP as well as the annealed LDP. The conditional LDP is for the empirical fields $\hat{P}_{x,Y,C_n}$, $n \ge 1$, with $x \in \Omega_X$ a single random realization of X, and the annealed LDP is for the empirical fields $\hat{P}_{X,Y,C_n}$, i.e., the process level LDP for $P \times Q$ [7]. In contrast, the quasi-quenched LDP involves a sequence of elements $x^{(n)} \in \Omega_X$. In general, if X is not ergodic, then $x^{(n)}$ are not necessarily identical or random realizations of X. For instance, consider $\mathbb{Z}^d = \mathbb{Z}$. If for each $n \in \mathbb{N}$, the first n elements of $x^{(n)}$ are made up of $\lfloor n/2 \rfloor$ consecutive 0's followed by $\lceil n/2 \rceil$ consecutive 1's, then the empirical measures given by $x^{(n)}$ converges to $\pi = \frac{1}{2}(\delta_{\mathbf{0}} + \delta_{\mathbf{1}})$, with $\mathbf{c} = (c, c, \ldots)$, which is stationary but not ergodic. Since no random realizations of X are involved in the quasi-quenched LDP, the corresponding rate function is deterministic, which is different from the conditional LDP.

Theorem 4 (Factorization of the LDP for product fields) Suppose $P \in \mathcal{M}(\Omega_X)$ satisfies the process level LDP with a good rate function I_P, i.e.

$$\liminf_{n\to\infty} \frac{1}{n^d} \log\ P\{\hat{P}_{X,C_n} \in G\} \ge -\inf_{\nu\in G} I_P(\nu), \quad G \in \mathcal{M}(\Omega_X) \text{ open}$$
$$\limsup_{n\to\infty} \frac{1}{n^d} \log\ P\{\hat{P}_{X,C_n} \in F\} \le -\inf_{\nu\in F} I_P(\nu), \quad F \in \mathcal{M}(\Omega_X) \text{ closed}$$

Suppose $Q \in \mathcal{M}(\Omega_Y)$ such that for any $\mu \in \mathcal{M}(\Omega_X)$ with $I_P(\mu) < \infty$, (μ, Q) satisfies the quasi-quenched process level LDP relative to μ with a good rate function $I_{Q|\mu}$. Then $P \times Q$ satisfies the process level LDP with a good rate function

$$I_{P\times Q}(\pi) = \inf_{\mu\in\mathcal{M}(\Omega_X)} \{I_{Q|\mu}(\pi) + I_P(\mu)\}, \tag{1.18}$$

for $\pi \in \mathcal{M}(\Sigma^{\mathbb{Z}^d})$. □

The idea of the proof for Theorem 4 is to condition on the empirical field induced by the random realization of X. Roughly speaking, letting $N(\,\cdot\,)$ stand for an infinitesimal neighborhood in weak topology of a measure, one may derive

$$\begin{aligned}
-I_{P\times Q}(\pi) &\approx \lim_{n\to\infty} \frac{1}{n^d} \log \Pr\{\hat{P}_{X,Y,C_n} \in N(\pi)\} \\
&\approx \lim_{n\to\infty} \frac{1}{n^d} \log \int \Pr\{\hat{P}_{X,Y,C_n} \in U \,|\, \hat{P}_{X,C_n} \in N(\mu)\} \Pr\{\hat{P}_{X,C_n} \in N(\mu)\} \\
&\approx \lim_{n\to\infty} \text{"Largest"} \frac{1}{n^d} \Big[\log \Pr\{\hat{P}_{X,Y,C_n} \in N(\pi) \,|\, \hat{P}_{X,C_n} \in N(\mu)\} + \log \Pr\{\hat{P}_{X,C_n} \in N(\mu)\}\Big] \\
&\approx -\inf\nolimits_{\mu\in\mathcal{M}(\Omega_X)} \{I_{Q|\mu}(\pi) + I_P(\mu)\}.
\end{aligned}$$

Since factorization naturally arises from conditioning, similar formulas have appeared in a few places in different contexts [3, 7]. In [3], factorization derives from a sequence of probability transition kernels satisfying the LDP continuity condition which is different from the quasi-quenched LDP condition in Definition 1. The form of the factorization (1.18) is similar to Eqs. (7) and (9) in [7] (also see Eq. (9) in [14]).

1.4 The quasi-quenched LDP for Gibbs fields

Comparing to the conditional LDP for the empirical fields, the quasi-quenched LDP imposes more restrictive conditions on Q. However, if Q is a stationary Gibbs random field with summable potential, then for any $P \in \mathcal{M}_s(\Omega_X)$, (P, Q) satisfies the quasi-quenched LDP relative to P. Furthermore, the associated rate function is deterministic, whose dual function is the expected value of the dual function associated with the conditional LDP. Indeed, we have

Theorem 5 Suppose $X \sim P \in \mathcal{M}_s(\Omega_X)$ and $Y \sim Q \in \mathcal{G}(U) \cap \mathcal{M}_s(\Omega_Y)$. Define

$$\Lambda_V(\lambda) = E_P \Lambda_{V,X}(\lambda), \quad \Lambda_n(\lambda) = E_P \Lambda_{n,X}(\lambda), \quad \Lambda(\lambda) = \lim_{n\to\infty} \Lambda_n(\lambda), \tag{1.19}$$

where $\Lambda_{V,X}(\lambda)$, $\Lambda_{n,X}(\lambda)$ are defined in (1.3) and (1.4). Because $\Lambda_{n,X}(\lambda)$ is bounded, equation (1.3) and dominated convergence imply that the limit in (1.19) exists. Define

$$\Lambda^*(u) = \sup_{\lambda\in\mathbb{R}^\nu} \{\langle \lambda, u\rangle - \Lambda(\lambda)\}. \tag{1.20}$$

Suppose $\{x^{(n)}\}_{n\geq 1}$ is a sequence of elements in Ω_X such that

$$\hat{P}_{x^{(n)},C_n} \to P, \qquad n\to\infty \tag{1.21}$$

in the weak topology. Then the empirical means of

$$\frac{1}{n^d} \sum_{t\in C_n} f((\theta_t x^{(n)}, \theta_t Y)_D),$$

satisfy the LDP with a good rate function $\Lambda^*(u)$. □

Consequently, by similar argument to that for Theorem 2 as well as by Theorem 4, there is

Corollary 2 Suppose $Q \in \mathcal{G}(U) \cap \mathcal{M}_s(\Omega_Y)$. For any $P \in \mathcal{M}_s(\Omega_X)$, (P, Q) satisfies the quasi-quenched process level LDP relative to P, with a good rate function

$$I_{Q|P}(\pi) = \sup\{I_{C_n,P}(\pi_{C_n}) : n \geq 1\}, \quad \pi \in \mathcal{M}(\Sigma^{\mathbb{Z}^d}), \tag{1.22}$$

where for any $\pi \in \mathcal{M}(\Sigma^D)$, $D \in \mathbb{Z}^d$ finite,

$$I_{D,P}(\pi) = \sup_{f\in B(\Sigma^D,\mathbb{R})} \{E_\pi f - \Lambda_P(f)\} \tag{1.23}$$

with

$$\Lambda_P(f) = \lim_{n\to\infty} \Lambda_{P,n}(f), \quad \Lambda_{P,n}(f) = \frac{1}{n^d} E_P \left[\log E_Q \exp\left\{\sum_{t\in C_n} f((\theta_t X,\, \theta_t Y)_D)\right\}\right]. \tag{1.24}$$

By Corollary 2, we get

Corollary 3 Suppose $Q \in \mathcal{G}(U) \cap \mathcal{M}_s(\Omega_Y)$. For any $P \in \mathcal{M}_s(\Omega_X)$ satisfying the process level LDP with a rate function I_P having the property that $I_P(\mu) = \infty$ for $\mu \notin \mathcal{M}_s(\Omega_X)$, $P \times Q$ satisfies the process level LDP with the rate function (1.18). □

The upper bound of the quasi-quenched LDP is proved by essentially the same argument for the upper bound of the conditional LDP. However, the argument for the conditional LDP to show that the lower and upper bounds arise from the same rate function does not work for the quasi-quenched LDP. A different approach is required to the lower bound of the quasi-quenched LDP. We shall adapt a method in [4], which relies more specifically on the properties of Gibbs random fields.

The rest of the article proceeds as follows. In section 2, we fix notations and collect preliminary results on Gibbs random fields. Theorem 1 and Theorem 2 are proved in section 3. Theorem 3 and Corollary 1 are shown in section 4. Theorem 4 is proved in section 5. Finally, Theorem 5 is proved in section 6.

2 Preliminaries

Given $t = (t_1, \ldots, t_d)$, $s = (s_1, \ldots, s_d) \in \mathbb{Z}^d$, $V \subset \mathbb{Z}^d$, $p \in \mathbb{N}$, define

Distance between t and s: $|t - s| = \max_{i=1,\ldots d}\{|t_i - s_i|\}$
Distance between t and V: $d(t, V) = \min_{\tau\in V}\{|t - \tau|\}$
Translation of V by t: $t + V = \{t + s : \ s \in V\}$
For any $s \in \mathbb{Z}^d$, refer to $s + C_n$ as an n-block
Outer boundary area of V: $\partial_p V = \{t \in V^c : d(t, V) \le p\}$
Inner boundary area of V: $d_p V = \{t \in V : d(t, V^c) \le p\}$
For $p = 1$, $d_p V$ is referred to as the boundary of V and denoted by ∂V
p-neighborhood of t: $N_p(t) = \{s \in \mathbb{Z}^d : \|s - t\| \le p\}$

For convenience, define shift operators in the following more general way. Given $t \in \mathbb{Z}^d$, and $V \subset \mathbb{Z}^d$, the shift operator θ_t wrt V is a map $S^V \to S^{-t+V}$, such that for any $x \in S^V$, $y = \theta_t x \in S^{-t+V}$, with $y_s = x_{s+t}$, $s \in -t + V$. When there is no confusion, V will not be specified.

For $V \subset \mathbb{Z}^d$, denote by $\mathcal{F}_V$ the σ-field on Ω generated by the projections $x \mapsto x(t)$, $t \in V$. For $\mu \in \mathcal{M}(\Omega)$ and $V \subset \mathbb{Z}^d$ with $|V| < \infty$, let μ_V be the marginal distribution of μ on S^V, and $\mu_V(\,\cdot\,|y)$ the marginal distribution of $\mu_V(\,\cdot\,|\mathcal{F}_{V^c})(y)$ on S^V, with the latter being the regular conditional distribution on $\mathcal{F}_{\mathbb{Z}^d}$ given $\mathcal{F}_{V^c}$ at $y \in \Omega$.

An *interaction potential* is a collection of maps $U_V : S^V \to \mathbb{R}$, $|V| < \infty$. It is *stationary* if

$$U_V(x_V) = U_{t+V}((\theta_t^{-1}x)_{t+V}), \quad x \in \Omega, \quad t \in \mathbb{Z}^d \tag{2.1}$$

U is called summable, if

$$\|U\| = \sum_{0\in V} \|U_V\| < \infty, \tag{2.2}$$

Define

$$\gamma_p = \sum_{0\in A,\ A\not\subset N_p(0)} \|U_A\|. \tag{2.3}$$

Then $\gamma_p \to 0$ as $p \to \infty$. A measure $Q \in \mathcal{M}(\Omega)$ is called a Gibbs measure wrt U if $Q_V(\,\cdot\,|y)$, $y \in \Omega$, can be chosen as

$$Q_V(x_V|y) = (Z_V(y))^{-1} \exp\{-E_V(x_V|y)\}, \text{ with } E_V(x_V|y) = \sum_{A\cap V\neq\emptyset} U_A(\xi_A), \quad x \in \Omega \tag{2.4}$$

where $Z_V(y)$ is a normalization constant and $\xi = \xi(x, y, A) \in \Omega$ is given by

$$\xi(t) = \mathbf{1}_V(t)x(t) + \mathbf{1}_{V^c}(t)y(t), \quad t \in \mathbb{Z}^d. \tag{2.5}$$

It is known that U is stationary and summable, then $\mathcal{G}(U) \cap \mathcal{M}_s(\Omega) \neq \emptyset$ (Theorem 4.3, [15]). The following standard lemma will be used in the proof of the results.

Lemma 1 For finite $V \subset \mathbb{Z}^d$ and $p \in \mathbb{N}$, define

$$K_{p,V} = 4\gamma_p|V| + 2^{d+1}p^d|\partial V|\|U\|. \tag{2.6}$$

Then for finite $W \subset \mathbb{Z}^d$ disjoint with V, and $y \in \Omega$,

$$e^{-K_{p,V}}Q_V(x_V) \le Q_V(x_V|y_W) \le e^{K_{p,V}}Q_V(x_V). \tag{2.7}$$

Proof. Given $x \in \Omega$, for y, $y' \in \Omega$, define ξ and ξ' by (2.5) correspondingly. Since $U_A(\xi_A) = U_A(\xi'_A)$ for $A \subset V$,

$$\begin{aligned}
|E_V(x_V|y) - E_V(x_V|y')| &= \left|\sum_{A\cap\partial V\neq\emptyset}\Big(U_A(\xi_A) - U_A(\xi'_A)\Big)\right| \\
&\le 2\sum_{A\cap\partial V\neq\emptyset}\|U_A\| \le \sum_{x\in V}\sum_{\substack{0\in A\\ A\not\subset N_p(0)}}\|U_{x+A}\| + \sum_{x\in d_pV\cup\partial_pV}\sum_{0\in A}\|U_{x+A}\| \\
&\overset{(a)}{\le} 2\gamma_p|V| + (2p)^d|\partial V|\|U\| = \frac{1}{2}K_{p,V},
\end{aligned}$$

where (a) is by stationarity of U. Therefore,

$$e^{-\frac{1}{2}K_{p,V}}e^{-E_V(x_V|y')} \le e^{-E_V(x_V|y)} \le e^{\frac{1}{2}K_{p,V}}e^{-E_V(x_V|y')}.$$

Take sum over all x_V to get

$$\begin{aligned}
& e^{-\frac{1}{2}K_{p,V}}Z_V(y') \le Z_V(y) \le e^{\frac{1}{2}K_{p,V}}Z_V(y'), \\
\Longrightarrow\quad & e^{-K_{p,V}}Q_V(x_V|y') \le Q_V(x_V|y) \le e^{K_{p,V}}Q_V(x_V|y').
\end{aligned}$$

Integrate over y' wrt $\mathrm{d}Q$ to get

$$e^{-K_{p,V}}Q_V(x_V|y) \le Q_V(x_V) \le e^{K_{p,V}}Q_V(x_V|y). \tag{2.8}$$

By conditioning,

$$Q_{V\cup W}(x_V, y_W) = \int Q_{V\cup W}(x_V, y_W|\eta)\,Q(\mathrm{d}\eta) = \int Q_V(x_V|\xi)Q_W(y_W|\eta)\,Q(\mathrm{d}\eta)$$

with $\xi(t) = \mathbf{1}_{W^c}(t)\eta(t) + \mathbf{1}_W(t)Y(t)$, $t \in \mathbb{Z}^d$. By the first inequality of (2.8),

$$Q_{V\cup W}(x_V, y_W) \ge \int e^{-K_{p,V}}Q_V(x_V)Q_W(y_W|\eta)\,Q(\mathrm{d}\eta) = e^{-K_{p,V}}Q_V(x_V)Q_W(y_W)$$

leading to the first inequality of (2.7). The second inequality is similarly proved. □

3 Conditional LDP for a Gibbs measure

Theorem 1 follows from several lemmas. The first one is a conditional-integral version of Bryc's asymptotic value result (Theorem 7.1, [2]). However, it only gives a lower bound of the conditional LDP.

Recall that $D = C_{n_0}$. Henceforth, for $V, W \subset \mathbb{Z}^d$, with $|V| < \infty$ and $\bigcup_{t\in V}(t+D) \subset W$, denote

$$\bar{f}_V(x,y) = \frac{1}{|V|}\sum_{t\in V} f((\theta_t x, \theta_t y)_D), \quad \bar{f}_n(x,y) = \bar{f}_{C_n}(x,y), \quad (x,y) \in \Sigma^W, \tag{3.1}$$

Lemma 2 *P-almost surely, for all $g : \mathbb{R}^\nu \to \mathbb{R}$ continuous concave,*

$$L(g;X) = \lim_{n\to\infty} \frac{1}{n^d} E_P\left[\log E_Q \exp\left\{n^d g(\bar{f}_n(X,Y))\right\} \,\Big|\, \mathcal{I}\right]$$

exists, and

$$\liminf_{n\to\infty} \frac{1}{n^d}\log E_Q \exp\left\{n^d g(\bar{f}_n(X,Y))\right\} = L(g;X). \tag{3.2}$$

The next result implies an upper bound for the conditional LDP, by using a conditional-integral version of the log-moment generating functions.

Lemma 3 *P-almost surely, for all $\lambda \in \mathbb{R}^\nu$, the limit in (1.3) exists, and*

$$\limsup_{n\to\infty} \frac{1}{n^d}\log E_Q \exp\left\{n^d \langle\lambda, \bar{f}_n(X,Y)\rangle\right\} = \Lambda_X(\lambda). \tag{3.3}$$

Lemma 4 *For $x \in \Omega_X$ such that (3.2) and (3.3) hold, let*

$$I(u;x) = \sup\{g(u) - L(g;x) : \ g \in C(\mathbb{R}^\nu) \text{ concave}\}.$$

For the other $x \in \Omega_X$, define $I(u;x)$ arbitrarily. Then P-almost surely, the realization of X is one such that (3.2) and (3.3) hold, and $I(\,\cdot\,;X) = \Lambda^*(\,\cdot\,;X)$, where $\Lambda^*(\,\cdot\,;X)$ is defined by equations (1.2) to (1.4). □

Assuming the lemmas for the moment, the proof for Theorem 1 precedes as follows.

Proof of Theorem 1: First of all, since f is bounded, the probability distributions of $\bar{f}_n(x,Y)$ consist an exponentially tight family. Suppose x is a random realization of X such that both (3.2) and (3.3) hold. Then for $G \subset \mathbb{R}^\nu$ open, it is not hard to show

$$\liminf_{n\to\infty} \frac{1}{n^d}\log \Pr\left\{\bar{f}_n(x,Y) \in G\right\} \geq -\inf_{u\in G} I(u;x), \tag{3.4}$$

following argument similar to the one for Lemma 4.4.6 [11]. More specifically, given $u \in G$, define $h : \mathbb{R}^\nu \to \mathbb{R}$ to be a continuous concave function such that $h(u) = 1$ and $h(u) \leq 0$ for $u \notin G$. For $m > 0$, define $h_m = m(h-1)$. Then

$$E_Q \exp\left\{n^d h_m(\bar{f}_n(x,Y))\right\} \leq e^{-mn^d} + \Pr\left\{\bar{f}_n(x,Y) \in G\right\}$$

On the other hand, h_m is continuous and concave with $h_m(u) = 0$. Then by (3.2), it is seen

$$\max\{\liminf_{n\to\infty} n^{-d}\log \Pr\left\{\bar{f}_n(x,Y) \in G\right\}, -m\} \geq L(h_m;x) \geq -I(u;x).$$

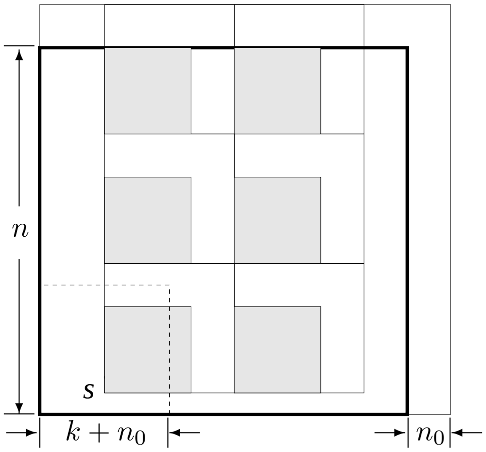

Figure 1. Spatial relationship of the blocks for calculation of the empirical means. The largest block is C_{n+n_0}, the one with thick boundaries is C_n, and the one with dotted boundaries is C_{k+n_0}. s is a point in C_{k+n_0}. Each shaded block is $s+t(k+n_0)+C_k$ for some $t \in \mathbb{Z}^d$ and belongs to J_s. Each small block that immediately contains $s+t(k+n_0)+C_k$ is $s+t(k+n_0)+C_{k+n_0}$, which belongs to I_s. The union of J_s over $s \in C_{k+n_0}$ consists of all $t + C_k \subset C_n$, $t \in \mathbb{Z}^d$. For the particular s, W_s is the union of the shaded blocks, and A_s the union of the blocks that immediately contain the shaded ones.

Letting $m \to \infty$ then finishes the proof for (3.4).

On the other hand, by Theorem 4.5.3 by [11], for any compact set $F \subset \mathbb{R}^\nu$,

$$\liminf_{n\to\infty} \frac{1}{n^d} \log \Pr\left\{\bar{f}_n(x,Y) \in F\right\} \le -\inf_{u\in F} \Lambda^*(u;x). \tag{3.5}$$

By the exponential tightness of the laws of $\bar{f}_n(x,Y)$, Lemma 3 leads to an upper bound of the LDP. Finally, the LDP is implied by Lemma 4. □

Proof of Lemma 2: Denote $\Gamma = \{u \in \mathbb{R}^\nu : |u| \le \|f\|\}$, and Γ° its inner part. For $V \subset \mathbb{Z}^d$, $|V| < \infty$, and $n \in \mathbb{N}$, define

$$\Lambda_{V,x} = \frac{1}{|V|} \log E_Q\left[e^{|V| g(\bar{f}_V(x,Y))}\right], \quad \Lambda_{n,x} = \Lambda_{C_n,x}. \tag{3.6}$$

Note that $g(\bar{f}_n(x,Y))$ only depends on $g|_\Gamma$. First, we show that if $g \in C(\Gamma)$ concave, then $\Lambda_{n,X}$ converges P-a.s. Given $\epsilon > 0$, fix integers $1 \ll p \ll k \ll n$. For each $s \in C_{k+n_0}$, let I_s be the collection of all the disjoint $(k+n_0)$-blocks $s + (k+n_0)t + C_{k+n_0}$ contained in C_{n+n_0}, $t \in \mathbb{Z}^d$, and J_s the collection of all the disjoint k-blocks $s+(k+n_0)t+C_k$ contained in C_n. Note the one-to-one correspondence between J_s and I_s,

$$V = t + C_k \in J_s \Longleftrightarrow V_e = t + C_{k+n_0} \in I_s \tag{3.7}$$

$$V_e = \textstyle\bigcup_{t\in V}(t+D). \tag{3.8}$$

Let $A_s = \bigcup_{V\in I_s} V$, $W_s = \bigcup_{V\in J_s} V$ (cf. Fig. 1). The union of J_s over $s \in C_{k+n_0}$ contains all $t + C_k \subset C_n$, $t \in \mathbb{Z}^d$, i.e.

$$\textstyle\bigcup_{s\in C_{k+n_0}} J_s = \{s + C_k : s \in T_{n,k}\}, \quad T_{n,k} = \{s \in \mathbb{Z}^d : s + C_k \subset C_n\}. \tag{3.9}$$

Then, letting $\kappa = k^d(k+n_0)^{-d}$, by $|d_{k+n_0}C_{n+n_0}| \le (k+n_0)^d|\partial C_{n+n_0}|$,

$$\begin{aligned}
& W_s \subset C_n, \quad C_{n+n_0} \setminus d_{k+n_0}C_{n+n_0} \subset A_s \subset C_{n+n_0} && (3.10)\\
\Longrightarrow\ & |W_s| = \sum_{V\in J_s} |V| = \kappa \sum_{V\in I_s} |V| = \kappa|A_s| \ge \kappa(|C_{n+n_0}| - (k+n_0)^d|\partial C_{n+n_0}|) \\
\Longrightarrow\ & |C_n \setminus W_s| < \epsilon n^d \\
\Longrightarrow\ & |\bar{f}_n(x,y) - \bar{f}_{W_s}(x,y)| \le 2\epsilon\|f\|, \qquad (x,y) \in \Sigma^{\mathbb{Z}^d}. && (3.11)
\end{aligned}$$

Because g is concave,

$$K_g = \sup\{|g(u)|,\ (1+\|f\|)\left|\frac{g(u)-g(v)}{u-v}\right|,\ u, v \in \Gamma^\circ,\ u \neq v\} < \infty. \tag{3.12}$$

Thus, by (3.11), $|W_s| = |J_s| k^d$, and the concavity of g,

$$g(\bar f_n(x,y)) \geq -2K_g\epsilon + g(\bar f_{W_s}(x,y)) \geq \frac{1}{|J_s|}\sum_{V\in J_s} g(\bar f_V(x,y))$$

$$\Longrightarrow \quad \Lambda_{n,X} \geq -2K_g\epsilon + \frac{1}{n^d}\log E_Q \exp\left\{\frac{n^d}{|J_s|}\sum_{V\in J_s} g(\bar f_V(X,Y))\right\}.$$

Given $V \in J_s$, it is easy to see that the value of $\bar f_V(X,Y)$ only depends on X_s, Y_s, $s \in \bigcup_{t\in V}(t+D) = V_e \subset A_s$. Also, by (2.7), (3.7), and successive conditioning, it is easy to verify

$$Q_{A_s}(y_{A_s}) \geq e^{-|J_s|K_{p,C_k}} \prod_{V\in I_s} Q_V(y_V).$$

Then,

$$\begin{aligned}
E_Q \exp\left\{\frac{n^d}{|J_s|}\sum_{V\in J_s} g(\bar f_V(X,Y))\right\} &= \sum_{y_{A_s}} \exp\left\{\frac{n^d}{|J_s|}\sum_{V\in J_s} g(\bar f_V(X_{A_s}, y_{A_s}))\right\} Q_{A_s}(y_{A_s}) \\
&\geq \sum_{y_{V_e}\in S_Y^{V_e}: V\in J_s} \exp\left\{\frac{n^d}{|J_s|}\sum_{V\in J_s} g(\bar f_V(X_{V_e}, y_{V_e}))\right\} e^{-|J_s|K_{p,C_k}} \prod_{V\in J_s} Q_{V_e}(y_{V_e}) \\
&= e^{-|J_s|K_{p,C_k}} \prod_{V\in J_s}\Big(\sum_{y_{V_e}\in S_Y^{V_e}} \exp\frac{n^d}{|J_s|} g(\bar f_V(X_{V_e}, y_{V_e}))\, Q_{V_e}(y_{V_e})\Big) \\
&= e^{-|J_s|K_{p,C_k}} \prod_{V\in J_s} E_Q \exp\left\{\frac{n^d}{|J_s|} g(\bar f_V(X,Y))\right\}.
\end{aligned}$$

Therefore, by the definition of K_{p,C_k}, for $1 \ll p \ll k \ll n$,

$$\begin{aligned}
\Lambda_{n,X} &\geq -2K_g\epsilon - \frac{|J_s|K_{p,C_k}}{n^d} + \frac{1}{n^d}\sum_{V\in J_s}\log E_Q \exp\left\{-\Big(\frac{n^d}{|J_s|} - k^d\Big) K_g + k^d\, g(\bar f_V(X,Y))\right\} \\
&\geq -2K_g\epsilon - \epsilon - \Big(1 - \frac{k^d|J_s|}{n^d}\Big) K_g + \frac{1}{n^d}\sum_{V\in J_s}\log E_Q\left[e^{k^d\, g(\bar f_V(X,Y))}\right] \\
&\geq -(3K_g+1)\epsilon + \frac{1}{n^d}\sum_{V\in J_s}\log E_Q\left[e^{k^d\, g(\bar f_V(X,Y))}\right].
\end{aligned}$$

The above inequality holds for each $s \in C_{k+n_0}$. By (3.9) and the fact that $J_s \cap J_t = \emptyset$, s, $t \in T_{n,k}$, $s \neq t$, averaging over $s \in C_{k+n_0}$ leads to

$$\begin{aligned}
\Lambda_{n,X} &\geq -(3K_g+1)\epsilon + \frac{1}{n^d}\frac{1}{(k+n_0)^d}\sum_{s\in T_{n,k}}\log E_Q\left[e^{k^d\, g(\bar f_{s+C_k}(X,Y))}\right] \\
&\overset{(a)}{=} -(3K_g+1)\epsilon + \frac{1}{n^d}\frac{1}{(k+n_0)^d}\sum_{s\in T_{n,k}}\log E_Q\left[e^{k^d\, g(\bar f_k(\theta_s X,Y))}\right],
\end{aligned}$$

where (a) is follows from the stationarity of Q. Thus, by (3.6),

$$\Lambda_{n,X} \geq -(3K_g+1)\epsilon + \frac{k^d}{(k+n_0)^d}\frac{1}{n^d}\sum_{s\in T_{n,k}} \Lambda_{k,\theta_s X}. \tag{3.13}$$

Take expectation wrt $P(\,\cdot\,|\mathcal{I})$ on both sides of (3.13). Because $P(\,\cdot\,|\mathcal{I})$ is stationary,

$$E_P\left[\Lambda_{n,X}|\mathcal{I}\right] \geq -(3K_g+1)\epsilon + \frac{k^d}{(k+n_0)^d}\frac{|T_{n,k}|}{n^d} E_P\left[\Lambda_{k,X}|\mathcal{I}\right]. \tag{3.14}$$

Fix k and take $\liminf_{n\to\infty}$ on both sides. Then $|T_{n,k}|n^{-d}\to 1$ and

$$\liminf_{n\to\infty} E_P[\Lambda_{n,X}|\mathcal{I}] \geq -(3K_g+1)\epsilon + \frac{k^d}{(k+n_0)^d} E_P[\Lambda_{k,X}|\mathcal{I}].$$

Let $k\to\infty$ on both sides. Then $k^d(k+n_0)^{-d}\to 1$ and

$$\liminf_{n\to\infty} E_P[\Lambda_{n,X}|\mathcal{I}] \geq -(3K_g+1)\epsilon + \limsup_{k\to\infty} E_P[\Lambda_{k,X}|\mathcal{I}].$$

Because ϵ is arbitrary, $E_P[\Lambda_{n,X}|\mathcal{I}]$ converges, with the limit denoted by $L(g;X)$.

Take $\liminf_{n\to\infty}$ on both sides of (3.13). Because X is stationary and $\mathcal{I}$ is the σ-algebra of shift invariant subsets of Ω_X, by the ergodic theorem (Theorem 6.21, [1]), P-almost surely,

$$\liminf_{n\to\infty} \Lambda_{n,X} \geq -(3K_g+1)\epsilon + \frac{k^d}{(k+n_0)^d} E_P[\Lambda_{k,X}|\mathcal{I}]. \tag{3.15}$$

Let $k\to\infty$ on the right hand side. Since ϵ is arbitrary, $\liminf_{n\to\infty}\Lambda_{n,X} \geq L(g;X)$. Since $\Lambda_{n,X}$ is bounded, Fatou's lemma gives $E_P[\liminf_{n\to\infty}\Lambda_{n,X}|\mathcal{I}] \leq L(g;X)$. By (3.6),

$$\liminf_{n\to\infty} \frac{1}{n^d}\log E_Q\left[e^{n^d g(\bar{f}_n(X,Y))}\right] = L(g;X), \quad P\text{-a.s.} \tag{3.16}$$

Thus the P-almost sure convergence is verified for each $g\in C(\Gamma)$ concave. The set of concave functions in $C(\Gamma)$ is separable. Moreover, given $x\in\Omega_X$,

$$\frac{1}{n^d}\log E_Q\left[e^{n^d g(\bar{f}_n(X,Y))}\right], \quad \frac{1}{n^d}E_P\left[\log E_Q\left[e^{n^d g(\bar{f}_n(\,\cdot\,,Y))}\right]\,\middle|\,\mathcal{I}\right](x), \quad n\geq 1$$

as a family of functions on $C(\Gamma)$ are equi-continuous. It is then not hard to show that P-almost surely, (3.16) holds for all $g\in C(\Gamma)$ concave. □

Proof of Lemma 3: Define $g(\,\cdot\,)$ such that it is equal to $\langle\lambda,\,\cdot\,\rangle$ on $\{u\in\mathbb{R}^\nu : |u|\leq \|f\|\}$. Then it can be shown that P-almost surely, for all $\lambda\in\mathbb{R}^\nu$,

$$\limsup_{n\to\infty}\frac{1}{n^d}\log E_Q\left[e^{|V|\langle\lambda,\bar{f}_V(X,Y)\rangle}\right] = \lim_{n\to\infty}\frac{1}{n^d}E_P\left[\log E_Q\left[e^{|V|\langle\lambda,\bar{f}_V(X,Y)\rangle}\right]\,\middle|\,\mathcal{I}\right] = \Lambda_X(\lambda).$$

Proof for this follows similar argument to Lemma 2, except using the second inequality of (2.7) instead of the first one, and the convexity of $\langle\lambda,\,\cdot\,\rangle$ instead of its concavity. □

To prove Lemma 4, we need one more auxiliary result.

Lemma 5 For any sequence $\{n_k\} \subset \mathbb{N}$, there is a sub-sequence $\{m_k\}$, such that for P-almost all $x \in \Omega_X$, $\bar{f}_{m_k}(x, Y)$ satisfy the LDP with a good rate function $I(u; x)$.

Proof. Again, let $\Gamma = \{u \in \mathbb{R}^\nu : |u| \le \|f\|\}$. First, given $g \in C(\Gamma)$ concave, let $F_n(g; X) = \Lambda_{n,X}$ given by (3.6). Let $G_n(X) = E_P[F_n(g; X)|\mathcal{I}]$ and $D_n = F_n - G_n$. Then $\lim_{k\to\infty} E_P[D_{n_k}] = 0$ and by Lemma 2, $\liminf_{k\to\infty} D_{n_k} = 0$. Since $D_{n_k} \xrightarrow{P} 0$ (cf. Lemma 2, [5]), there is a sub-sequence m_k, such that $D_{m_k} \to 0$, P-a.s., giving $F_{m_k}(g; X) \to L(g; X)$, P-a.s. Furthermore, $\{g \in C(\Gamma) : g$ concave$\}$ is separable, and given $x \in \Omega_X$, $\{F_n(g; x),\ n \ge 1\} \cup \{L(g; x)\}$ as a family of functions in g is equi-continuous. Based on the above two facts, by the diagonal argument, it is seen that from n_k one can find a sub-sequence m_k, such that P-almost surely, for all $g \in C(\Gamma)$ concave, $F_{m_k}(g; x) \to L(g; x)$ (cf. Theorem 3, [5]). Then by Bryc's inverse Varadhan lemma, for P-almost all $x \in \Omega_X$, $\bar{f}_{m_k}(x, Y)$ satisfy the LDP with a good rate function $I(u; x)$. □

Proof of Lemma 4: By Lemma 5, $I(u; x)$ is a good rate function of an LDP. Therefore, in order to demonstrate $I(u; x) = \Lambda^*(u; x)$, it is enough to show I is convex (Theorem 4.5.10, [11]).

Let $e = (1, 0, \dots, 0) \in \mathbb{Z}^d$. Denote

$$C'_n = (n + n_0)e + C_n, \quad V_n = C_n \cup C'_n.$$

Then repeating the same argument that leads to (3.4) and (3.5), it can be shown that P-almost surely, equations (3.4) and (3.5) hold simultaneously for C_n, C'_n, and V_n, and for all $g \in C(\Gamma)$ concave. Therefore, by Lemma 5, there is a sub-sequence $\{n_k\} \subset \mathbb{N}$, such that for P-almost all $x \in \Omega_X$, $\bar{f}_{C_{n_k}}(x, Y)$, $\bar{f}_{C'_{n_k}}(x, Y)$, and $\bar{f}_{V_{n_k}}(x, Y)$ satisfy the LDP with a good rate function $I(u; x)$.

Given $x \in \Omega_X$ and $V \subset \mathbb{Z}^d$ finite, denote by L_V the probability measure on $\mathbb{R}^\nu$ induced by $\bar{f}_V(x, Y)$. Given $r > 0$, fix a concave function g such that $g(0) = 0$, $g(u) < -1$, $u \notin B_r$. Given $u_1, u_2 \in \mathbb{R}^\nu$, let $\bar{u} = (u_1 + u_2)/2$. Denote $G_i = u_i + B_r$, $i = 1, 2$, and $G = \bar{u} + B_r$. For any $m > 0$,

$$J = \int e^{m|V_{n_k}|g(u-\bar{u})} L_{V_{n_k}}(\mathrm{d}u) \le e^{-m|V_{n_k}|} + L_{V_{n_k}}(G). \tag{3.17}$$

On the other hand, letting $g_i(u) = mg(u - u_i)$, $i = 1, 2$, by the concavity of g,

$$J = E_Q\left[e^{m|V_{n_k}|g(\bar{f}_{V_{n_k}}(x,Y)) - \bar{u})}\right] \ge E_Q\left[e^{|C_{n_k}|g_1(\bar{f}_{V_{n_k}}(x,Y))} e^{|C'_{n_k}|g_2(\bar{f}_{C'_{n_k}}(x,Y))}\right]$$

The two exponentials in the integral on the right hand side only depend on the values of Y on $C_{n_k+n_0}$ and $(n_k + n_0)e + C_{n_k+n_0}$, respectively. Therefore, given $p > 0$, (3.17) and (2.7) lead to

$$\begin{aligned} \frac{1}{|V_{n_k}|} \log\{e^{-m|V_{n_k}|} + L_{V_{n_k}}(G)\} \quad \ge \quad & \frac{1}{|V_{n_k}|} K_{p, C_{n_k+n_0}} + \frac{1}{2|C_{n_k}|} \log \int e^{|C_{n_k}|g_1(u)} L_{C_{n_k}}(\mathrm{d}u) \\ & + \frac{1}{2|C'_{n_k}|} \log \int e^{|C'_{n_k}|g_2(u)} L_{C'_{n_k}}(\mathrm{d}u). \end{aligned}$$

Let $k \to \infty$ and then $p \to \infty$ to get

$$\max\{\liminf_{k\to\infty} \frac{1}{|V_{n_k}|} \log L_{V_{n_k}}(G), -m\} \ge \frac{1}{2}(L(g_1; x) + L(g_2; x))$$

Since $L(g_i; x) = -(g_i(u_i) - L(g_i; x)) \ge -I(u_i)$, $i = 1, 2$, letting $m \to \infty$ leads to

$$\liminf_{k\to\infty} \frac{1}{|V_{n_k}|} \log L_{V_{n_k}}(\bar{u} + B_r) \ge -\frac{1}{2}(I(u_1) + I(u_2)).$$

Let $r \to 0$ to get $-I(\bar{u}) \ge -\frac{1}{2}(I(u_1) + I(u_2))$, proving the convexity of I. □

Proof of Theorem 2:

(a) Fix $\nu \in \mathbb{N}$. For $x \in \Omega_X$, $f \in B(\Sigma^D, \mathbb{R}^\nu)$, $g \in C(\mathbb{R}^\nu)$ concave, and finite $V \subset \mathbb{Z}^d$, write

$$\Lambda_{n,x}(f,g) = \frac{1}{n^d} \log E_Q \exp\left\{n^d g(\bar{f}_n(x,Y))\right\}, \quad M_{n,x}(f,g) = E_P[\Lambda_{n,\cdot}(f,g)|\mathcal{I}](x).$$

As a matter of fact, $\Lambda_{n,x}(f,g)$ is identical to $\Lambda_{n,x}$ in (3.6). The more complex notation is used to emphasize the dependence on both f and g.

From Lemma 2 and Lemma 3, for each $f \in B(\Sigma^D, \mathbb{R}^\nu)$, $g \in C(\mathbb{R}^\nu)$, there is $L(f,g,X)$ such that

$$\lim_{n\to\infty} M_{n,X}(f,g) = \liminf_{n\to\infty} \Lambda_{n,X}(f,g) = L(f,g,X), \quad P\text{-a.s.} \tag{3.18}$$

In addition, if $g(u) = \langle \lambda, u\rangle$ for some $\lambda \in \mathbb{R}^\nu$, then there is also

$$\limsup_{n\to\infty} \Lambda_{n,X}(f,g) = L(f,g,X). \tag{3.19}$$

Given $\nu, K \in \mathbb{N}$, let

$$F_K = \{f \in B(\Sigma^D, \mathbb{R}^\nu) : \|f\| \le K\}, \quad A_K = \{g \in C(\bar{B}_K), g \text{ concave}\}.$$

Given $(f,g) \in F_K \times A_K$, (3.18) holds P-a.s. On the other hand, given $x \in \Omega_X$, $M_{n,x}(f,g)$, $\Lambda_{n,x}(f,g)$, $n \ge 1$ consist a family of functions on $F_K \times A_K$. Under the sup norm, the family is equi-continuous and $F_K \times A_K$ is separable. Therefore, with probability 1, (3.18) holds simultaneously for all $(f,g) \in F_K \times A_K$. Consequently, by $B(\Sigma^D, \mathbb{R}^\nu) = \bigcup_K F_K$, with probability 1, (3.18) holds simultaneously for all $\nu \in \mathbb{N}$, $g \in C(\mathbb{R}^\nu)$, and $f \in B(\Sigma^D, \mathbb{R}^\nu)$. Similarly, with probability 1, (3.19) holds for all $f \in B(\Sigma^D, \mathbb{R}^\nu)$ and $\lambda \in \mathbb{R}^\nu$. Now by the same argument for Theorem 1, for P-almost all $x \in \Omega_X$, for $f \in B(\Sigma^D, \mathbb{R}^\nu)$, the empirical means $\bar{f}_n(x,Y)$ satisfy the LDP with a convex good rate function. Then the LDP for $\hat{P}_{x,C_n,D}$ follows.

(b) For P-almost all $\hat{x} \in \Omega_X$, for each $n \ge 1$, we can establish LDP for the empirical measure $\hat{P}_{C_n,\hat{x},Y}$. Thus the LDP in (b) follows from Dawson-Gärtner's theorem on the LDP for a project limit (Theorem 3.3, [8]). Then $I(\pi;\hat{x}) = \infty$ for $\pi \notin \mathcal{M}_s(\Sigma^{\mathbb{Z}^d})$, following argument similar to (Eq. 5.4.15, [12]). Indeed, for some finite $A \subset \mathbb{Z}^d$, $f \in B(\Sigma^A, \mathbb{R})$, and $l \in \mathbb{Z}^d$,

$$\int f((\theta_t x, \theta_t y)_A)\, \mathrm{d}\pi(x,y) \ge \int f((x,y)_A)\, \mathrm{d}\pi(x,y) + 1.$$

Choose $D = C_{n_0} \supset A \cup (l + A)$ and define $h(x,y) = f((\theta_l x, \theta_l y)_A) - f((x,y)_A)$. Then for any $M > 0$, $\int Mf\, \mathrm{d}\pi \ge M$. On the other hand, for any $C_n \subset D$,

$$|\bar{h}_n(x,y)| = |\bar{f}_n(\theta_l x, \theta_l y) - \bar{f}_n(x,y)| \le n^{-d}\|f\|\, |C_n \,\Delta\, (l + C_n)|,$$

where $C_n \,\Delta\, (l + C_n)$ stands for $C_n \cup (l + C_n) - C_n \cap (l + C_n)$. Thus

$$\left|\frac{1}{n^d} E_P\left[\log E_Q\left[e^{Mn^d \bar{h}_n(X,y)}\right] \,\Big|\, \mathcal{I}\right](\hat{x})\right| \le \frac{2\|f\| \cdot |C_n \,\Delta\, (l + C_n)|}{n^d} \to 0$$

leading to $L(Mh;\hat{x}) = 0$. Therefore, by $I(\pi;\hat{x}) \ge \int Mh\, \mathrm{d}\pi - L(Mh;\hat{x}) \ge M$, $I(\pi;\hat{x}) = \infty$. □

Proposition 1 Both $\Lambda_x(\lambda)$ and $\Lambda_x^*(u$ are $(\mathcal{B}(\mathbb{R}^\nu) \times \mathcal{I})$-measurable, where $\mathcal{B}(\mathbb{R}^\nu)$ is the Borel σ-algebra of $\mathbb{R}^\nu$. Furthermore, letting $\mathcal{B}$ be the Borel σ-algebra generated by the weak topology of $\mathcal{M}(\Sigma^{\mathbb{Z}^d})$, $I(\pi;x)$ is $(\mathcal{B} \times \mathcal{I})$-measurable.

Proof. From (1.3) and (1.4), for P-almost all $x \in \Omega_X$, $\Lambda_x(\lambda)$ is continuous, while given λ, $\Lambda_x(\lambda)$ as a function in x is $\mathcal{I}$-measurable. For $n \ge 1$, let $\mathcal{J}_n$ be the collection of all binary cells

$$I_{k,n} = \{(\lambda_1, \ldots, \lambda_\nu) \in \mathbb{R}^\nu : k_i 2^{-n} \le \lambda_i < (k_i+1)2^{-n},\ i = 1, \ldots, \nu\},\ k = (k_1, \ldots, k_\nu) \in \mathbb{Z}^\nu.$$

Then define $F_n(\lambda; x) = \sum_{k \in \mathbb{Z}^\nu} \Lambda_x(k2^{-n}) \mathbf{1}_{I_{k,n}}(\lambda)$. For each k, $\Lambda_x(k2^{-n})$ is $\mathcal{I}$-measurable, and $\mathbf{1}_{I_{k,n}}(\lambda)$ is $\mathcal{B}(\mathbb{R}^\nu)$ is measurable. Therefore, F_n is $(\mathcal{B}(\mathbb{R}^\nu) \times \mathcal{I})$-measurable. Because for each x, $\Lambda_x(\lambda)$ is continuous, $F_n \to \Lambda$ point wise, as $n \to \infty$, leading to $\Lambda_x(\lambda)$ being $(\mathcal{B}(\mathbb{R}^\nu) \times \mathcal{I})$-measurable.

By the continuity of $\Lambda_x(\lambda)$, given x, it is easy to see that for each $0 \le a < \infty$,

$$\{(u, x) : \Lambda_x^*(u) \le a\} = \bigcap_{\lambda \in Q} \{(u, x) : \langle \lambda, u \rangle - \Lambda_x(\lambda) \le a\} \in \mathcal{B}(\mathbb{R}^\nu) \times \mathcal{I},$$

where Q is the set of rational points in $\mathbb{R}^\nu$.

Given $n \ge 1$ and $f \in B(\Sigma^{C_n}, \mathbb{R})$, $E_\pi f - \Lambda_x(f)$ as a function in (π, x) is $(\mathcal{B} \times \mathcal{I})$-measurable. Since for each (π, x), $E_\pi f - \Lambda_x(f)$ is continuous in f, and $B(\Sigma^{C_n}, \mathbb{R})$ is separable, from (1.9), it is seen $I_{C_n}(\pi; x)$ is $(\mathcal{B} \times \mathcal{I})$-measurable, and hence so is $I(\pi; x)$ by (1.9). □

4 A variational characterization of the deterministic rate functions

Proof of Theorem 3: Fix $\epsilon > 0$ and $\pi \in \mathcal{M}_s(\Sigma^{\mathbb{Z}^d})$, such that $\pi_X = P$ and $|E_\pi f - u| < \epsilon$. Given $n \ge 1$, $x \in S_X^{C_{n+n_0}}$, and $\lambda \in \mathbb{R}$, since $\lambda n^d \bar{f}_n(x, y)$ is bounded, by Lemma 3.2.13 in [12],

$$\begin{aligned} &H_{C_{n+n_0}}(\pi_{C_{n+n_0}}(\,\cdot\,|x) \,\|\, Q_{C_{n+n_0}}) \\ &\quad \ge \lambda n^d \int \bar{f}_n(x, y)\, \mathrm{d}\pi_{C_{n+n_0}}(y|x) - \log \int \exp\left\{\lambda n^d \bar{f}_n(x, y)\right\} \mathrm{d}Q_{C_{n+n_0}}(y) \end{aligned}$$

Integrate both sides over x, then divide them by n^d . Since π is stationary and $\pi_X = P$, then

$$\frac{1}{n^d} H_{C_{n+n_0}}(\pi_{C_{n+n_0}} \| (P \times Q)_{C_{n+n_0}}) \ge \lambda\, E_\pi f - \Lambda_{C_n}(\lambda)$$

Since λ is arbitrary and $|E_\pi f - u| < \epsilon$,

$$\frac{1}{n^d} H_{C_{n+n_0}}(\pi_{C_{n+n_0}} \| (P \times Q)_{C_{n+n_0}}) \ge \sup_\lambda \{\lambda\, E_\pi f - \Lambda_{C_n}(\lambda)\} \ge \inf_{|u'-u|<\epsilon} \Lambda_{C_n}^*(u'). \tag{4.1}$$

By Theorem 5 of [16],

$$\Lambda^*(u) = \lim_{\epsilon \to 0} \limsup_{n \to \infty} \inf_{|u'-u|<\epsilon} \Lambda_{C_n}^*(u') = \lim_{\epsilon \to 0} \liminf_{n \to \infty} \inf_{|u'-u|<\epsilon} \Lambda_{C_n}^*(u'). \tag{4.2}$$

Let $n \to \infty$ and then $\epsilon \to 0$. Since π is arbitrary, by (4.1) and (4.2), $J(u) \ge \Lambda^*(u)$. Clearly $\tilde{J}(u) \ge J(u)$ and hence $\tilde{J}(u) \ge \Lambda^*(u)$.

To prove $\Lambda^*(u) \ge J(u)$, first assume $u \in (u_{\min}^{(\infty)}, u_{\max}^{(\infty)})$. In this case, it suffices to show $\Lambda^*(u) \ge \tilde{J}(u)$, which also implies (1.15). Define

$$u_{\min}^{(n)} = E_P[\operatorname*{ess\,inf}_Y\ \bar{f}_n(X, Y)] \qquad u_{\max}^{(n)} = E_P[\operatorname*{ess\,sup}_Y\ \bar{f}_n(X, Y)], \tag{4.3}$$

where $\bar{f}$ is defined as in (3.1). Then following the argument for (47) and (48) in [10],

$$u_{\min}^{(\infty)} = \sup_n u_{\min}^{(n)} = \lim_{n \to \infty} u_{\min}^{(n)} \qquad u_{\max}^{(\infty)} = \inf_n u_{\max}^{(n)} = \lim_{n \to \infty} u_{\max}^{(n)}. \tag{4.4}$$

Therefore, for all $n \ge 1$, the supremum of $\lambda u - \Lambda_{C_n}(\lambda)$ is achieved. Let

$$\lambda_n^* = \arg\sup_{\lambda \in \mathbb{R}} \{\lambda u - \Lambda_{C_n}(\lambda)\} \tag{4.5}$$

and define a probability measure $\hat{\pi}_n$ on Σ^{n+n_0} by

$$\hat{\pi}_n(x,y) = \frac{\exp\{\lambda_n^* n^d \bar{f}_n(x,y)\} \times Q_{C_{n+n_0}}(y)}{\sum_{y'} \exp\{\lambda_n^* n^d \bar{f}_n(x,y)\} \times Q_{C_{n+n_0}}(y')} P_{C_{n+n_0}}(x), \quad x \in S_X^{C_{n+n_0}},\ y \in S_Y^{C_{n+n_0}} \tag{4.6}$$

Then it is not hard to show (1) $S_X^{C_{n+n_0}}$-marginal of $\hat{\pi}_n$ is $P_{C_{n+n_0}}$, (2) $E_{\hat{\pi}_n} \bar{f}_n(X,Y) = u$, as seen from $\Lambda'_{C_n}(\lambda_n^*) = u$, and (3) $n^{-d} H_{C_{n+n_0}}(\hat{\pi}_n \,\|\, (P \times Q)_{C_{n+n_0}}) = \Lambda^*_{C_n}(u)$ (cf. [9, 17]).

Define $\tilde{\pi}_n \in \mathcal{M}(\Sigma^{\mathbb{Z}^d})$ as the product of independent copies of $\hat{\pi}_n$ on disjoint $(n+n_0)$-blocks $t(n+n_0) + C_{n+n_0}$, i.e.

$$\tilde{\pi}_n = \prod_{t \in \mathbb{Z}^d} \hat{\pi}_n \circ \theta^{-1}_{t(n+n_0)}$$

and $\pi_n \in \mathcal{M}_s(\Sigma^{\mathbb{Z}^d})$ by

$$\pi_n = \frac{1}{(n+n_0)^d} \sum_{t \in C_{n+n_0}} \tilde{\pi}_n \circ \theta_t^{-1}. \tag{4.7}$$

Since $\mathcal{M}_s$ is compact, there is $\pi \in \mathcal{M}_s(\Sigma^{\mathbb{Z}^d})$ and a subsequence π_{n_k}, such that $\pi_{n_k} \to \pi$ in weak topology. Next show

(A) $\pi_X = P$,
(B) $E_\pi f = u$, and
(C) $\Lambda^*(u) \ge h(\pi \| P \times Q)$.

For statement (A), given $m \ge 1$, let π_{X,C_m} be the marginal of π_X on $S_X^{C_m}$. Then

$$\pi_{X,C_m} = \lim_{k\to\infty} \frac{1}{(n_k+n_0)^d} \sum_{t \in C_{n_k+n_0}} (\tilde{\pi}_{n_k} \circ \theta_t^{-1})_{X,C_m}$$

When $t \in C_{n_k+n_0} \setminus d_m C_{n_k+n_0}$,

$$(\tilde{\pi}_{n_k} \circ \theta_t^{-1})_{X,C_m} = (\hat{\pi}_{n_k} \circ \theta_t^{-1})_{X,C_m} = ((\hat{\pi}_{n_k})_X \circ \theta_t^{-1})_{C_m} = (P_{C_{n_k+n_0}} \circ \theta_t^{-1})_{C_m} = P_{C_m}$$

where the last equality is by the stationarity of P and $t + C_m \subset C_{n_k+n_0}$. On the other hand, $|d_m C_{n_k+n_0}| = o(|C_{n_k+n_0}|)$. Therefore, letting $k \to \infty$, there is $\pi_{X,C_m} = P_{C_m}$, and thus $\pi_X = P$.

To prove statement (B), i.e. $E_\pi f = u$,

$$\begin{aligned}
E_{\pi_n} f(X_D, Y_D) &= \frac{1}{(n+n_0)^d} \sum_{t \in C_{n+n_0}} E_{\tilde{\pi}_n \circ \theta_t^{-1}} f(X_D, Y_D) \\
&= \frac{1}{(n+n_0)^d} \sum_{t \in C_n} E_{\tilde{\pi}_n} f(X_{t+D}, Y_{t+D}) + \frac{1}{(n+n_0)^d} \sum_{i \in C_{n+n_0} \setminus C_n} E_{\tilde{\pi}_n} f(X_{t+D}, Y_{t+D}) \\
&= \frac{1}{(n+n_0)^d} E_{\tilde{\pi}_n} \bar{f}_n(X,Y) + \frac{1}{(n+n_0)^d} \sum_{t \in C_{n+n_0} \setminus C_n} E_{\tilde{\pi}_n} f(X_{t+D}, Y_{t+D})
\end{aligned}$$

By the definition of $\tilde{\pi}_n$, $E_{\tilde{\pi}_n} \bar{f}_n(X,Y) = E_{\hat{\pi}_n} \bar{f}_n(X,Y) = u$. Let $n \to \infty$ along n_k. Then the first summand on the right hand side converges to u. By the boundedness of f and $|C_{n+n_0} \setminus C_n| = o(n^d)$, the second summand is $o(1)$. Therefore $E_\pi f = u$.

For statement (C), by the lower semi-continuity of $h(\,\cdot\,\|P\times Q)$, in order to show $h(\pi\|P\times Q) \le \Lambda^*(u)$, it is enough to show

$$\limsup_{n\to\infty} h(\pi_n\|P\times Q) \le \Lambda^*(u).$$

Given $p \ge 1$, fix $m \gg n \gg p$. Denote by π_{n,C_m} the marginal of π_n on Σ^{C_m}. By the convexity of $H_{C_m}(\,\cdot\,\|(P\times Q)_{C_m})$ and the stationarity of $P\times Q$,

$$\begin{aligned} H_{C_m}(\pi_{n,C_m}\|(P\times Q)_{C_m}) &\le \frac{1}{(n+n_0)^d} \sum_{t\in C_{n+n_0}} H_{C_m}((\tilde{\pi}_n\circ\theta_t^{-1})_{C_m}\|(P\times Q)_{C_m}) \\ &= \frac{1}{(n+n_0)^d} \sum_{t\in C_{n+n_0}} H_{t+C_m}(\tilde{\pi}_{n,t+C_m}\|(P\times Q)_{t+C_m}). \end{aligned} \tag{4.8}$$

Let $\mathcal{C}$ be the collection of all $s(n+n_0)+C_{n+n_0}$, $s\in\mathbb{Z}^d$. Given $t\in C_{n+n_0}$, define

$$\begin{aligned} J &= \{V\in\mathcal{C} : V\subset t+C_m\} \\ I &= \{V:\ V = W\cap(t+C_m)\ne\emptyset,\ \text{with } W\cap(t+C_m)^c\ne\emptyset,\ W\in\mathcal{C}\} \end{aligned}$$

Then $\tilde{\pi}_{n,t+C_m} = \prod_{V\in J\cup I}\tilde{\pi}_{n,V}$. On the other hand, since $P\times Q$ is a stationary Gibbs field with summable interaction potential, by Lemma 1,

$$(P\times Q)_{t+C_m} \ge \prod_{V\in J\cup I} e^{-K_{p,V}}(P\times Q)_V,$$

with $K_{p,V}$ defined in terms of the potential of the Gibbs field $P\times Q$. Therefore,

$$H_{t+C_m}(\tilde{\pi}_{n,t+C_m}\|(P\times Q)_{t+C_m}) \le \sum_{V\in J\cup I} K_{p,V} + \sum_{V\in J\cup I} H_V(\tilde{\pi}_{n,V}\|(P\times Q)_V).$$

For $V\in J$, by the construction of $\tilde{\pi}_n$, and the stationary of $P\times Q$,

$$H_V(\tilde{\pi}_{n,V}\|(P\times Q)_V) = H_{C_{n+n_0}}(\hat{\pi}_n\|(P\times Q)_{C_{n+n_0}}) = n^d\Lambda^*_{C_n}(u)$$

In addition, $K_{p,V} = K_{p,C_{n+n_0}}$, $V\in J$. On the other hand, for some constant $M>0$ which only depends on n but not m, for all $V\in I$, $K_{p,V} + H_V(\tilde{\pi}_{n,V}\|(P\times Q)_V)\le M$. Therefore,

$$H_{t+C_m}(\tilde{\pi}_{n,t+C_m}\|(P\times Q)_{t+C_m}) \le |J|K_{p,C_{n+n_0}} + n^d|J|\Lambda^*_{C_n}(u) + |I|M.$$

Divide both sides by $|C_m|$ and let $m\to\infty$. Then $|C_m|^{-1}|J|\to\frac{1}{(n+n_0)^d}$ and $|C_m|^{-1}|I|\to 0$.

By (4.8),

$$h(\pi_n\|P\times Q) \le \frac{K_{p,C_{n+n_0}}}{(n+n_0)^d} + \frac{n^d}{(n+n_0)^d}\Lambda^*_n(u). \tag{4.9}$$

Let $n\to\infty$. By Theorem 27 of [10], $\Lambda^*_n(u)\to\Lambda^*(u)$. Then by (2.6),

$$\limsup_{n\to\infty} h(\pi_n\|P\times Q) \le 4\gamma_p + \Lambda^*(u).$$

Since p is arbitrary, letting $p \to \infty$ finishes the proof.

When $u < u_{\min}^{(\infty)}$ or $u > u_{\max}^{(\infty)}$, it is easy to see $\Lambda^*(u) = \infty$, giving $\Lambda^*(u) \ge \tilde{J}(u) \ge J(u)$. It remains to prove (1.14) for $u = u_{\min}^{(\infty)}$ or $u_{\max}^{(\infty)}$. It is easy to check that

$$u_{\min}^{(n)} \le u_{\min}^{(nk)} \le u_{\min}^{(\infty)} \le u_{\max}^{(\infty)} \le u_{\max}^{(nk)} \le u_{\max}^{(n)}, \quad n, k \ge 1.$$

If for some n, $u_{\min}^{(n)} = u_{\max}^{(n)}$. Then $u = u_{\min}^{(n)}$, implying $E_{P\times Q}[\bar{f}_n(X,Y)] = u$, and hence

$$E_{P\times Q}[f(X_D, Y_D)] = u.$$

Therefore, letting $\pi = P \times Q$ in (1.14), $J(u) = 0$ and hence $\Lambda^*(u) \ge J(u)$.

Suppose $u_{\min}^{(n)} < u_{\max}^{(n)}$ for all $n \ge 1$. Let $u = u_{\min}^{(\infty)}$. Given $\epsilon > 0$, choose n large enough so that $u_{\min}^{(\infty)} - \epsilon < u_{\min}^{(n)} \le u_{\min}^{(\infty)} \le u_{\max}^{(\infty)} \le u_{\max}^{(n)}$. Then $A = (u_{\min}^{(n)}, u_{\max}^{(n)}) \cap (u_{\min}^{(\infty)} - \epsilon, u_{\min}^{(\infty)} + \epsilon) \ne \emptyset$. Choose $u' \in A$ and repeat the proof from (4.5) to (4.9). Then it is seen

$$h(\pi_n \| P \times Q) \le \frac{K_{p,C_{n+n_0}}}{(n+n_0)^d} + \Lambda_n^*(u').$$

Since

$$\Lambda_n^*(u_{\min}^{(n)}) = \lim_{u \downarrow u_{\min}^{(n)}} \Lambda_n^*(u), \quad \Lambda_n^*(u_{\max}^{(n)}) = \lim_{u \uparrow u_{\max}^{(n)}} \Lambda_n^*(u), \text{ and}$$
$$\Lambda_n^*(u) = \infty, \quad \text{for } u \notin [u_{\min}^{(n)}, u_{\max}^{(n)}],$$

then by the arbitrariness of u',

$$\inf\left\{h(\pi \| P\times Q):\ \pi \in \mathcal{M}_s(\Sigma^{\mathbb{Z}^d}),\ \pi_X = P,\ |E_\pi f - u| < \epsilon\right\} \le \frac{K_{p,C_{n+n_0}}}{(n+n_0)^d} + \inf_{|u'-u|<\epsilon} \Lambda_n^*(u').$$

Now let $n \to \infty$ and then $\epsilon \to 0$. By (4.2), $J(u) \le 4\gamma_p + \Lambda^*(u)$. Let $p \to \infty$ to finish the proof. The case $u = u_{\max}^{(\infty)}$ is similarly proved. □

Proof of Corollary 1: Denote $B_1 = \{f \in B(\Sigma^D, \mathbb{R}) : \|f\| = 1\}$. For any $f \ne 0$, $g = \frac{f}{\|f\|} \in B_1$. Therefore, by (1.7)

$$I_D(\pi) = \sup_{f\in B_1} \sup_{\lambda\in\mathbb{R}} \{\lambda\, E_\pi f - \Lambda(\lambda f)\}$$

with $\Lambda(f)$ defined by (1.8). By Theorem 3, it is seen

$$I_D(\pi) = \sup_{f\in B_1} \lim_{\epsilon\to 0} \inf\{h(\gamma \| P\times Q):\ \gamma \in \mathcal{M}_s(\Sigma^{\mathbb{Z}^d}),\ \gamma_X = P,\ |E_\gamma f - E_\pi f| < \epsilon\}.$$

It is easy to show the limit and the supremum are exchangeable, hence proving (1.16). □

5 Factorization formula

Proof of Theorem 4: First prove the lower bound of the LDP. Let $G \subset \mathcal{M}(\Sigma^{\mathbb{Z}^d})$ be an open subset. Fix $\pi \in G$ and $\mu \in \mathcal{M}(\Omega_X)$. Then there is a sequence of open subsets $U_n \subset \mathcal{M}(\Omega_X)$ with $U_n \downarrow \mu$, such that

$$\liminf_{n\to\infty} \frac{1}{n^d} \log P\{\hat{P}_{X,C_n} \in U_n\} \ge -I_P(\mu). \tag{5.1}$$

To see this, choose open sets $V_k \downarrow \mu$, and let $N_k \uparrow \infty$, such that for all $n \geq N_k$,

$$\frac{1}{n^d} \log\, P\{\hat{P}_{X,C_n} \in V_k\} \geq -I_P(\mu) - \frac{1}{k}.$$

Let $U_n = V_1$ for all $n < N_2$ and $U_n = V_k$ for $N_k \leq n < N_{k+1}$, $k > 1$. Then U_n satisfy (5.1).

Now show

$$\liminf_{n\to\infty} \frac{1}{n^d} \log \inf_{x:\hat{P}_{x,C_n}\in U_n} Q\{\hat{P}_{x,Y,C_n} \in G\} \geq -I_{Q|\mu}(\pi). \tag{5.2}$$

Indeed, assume (5.2) is not true. Then there are $\epsilon > 0$ and $x^{(n)} \in U_n$, $n \geq 1$, such that

$$\liminf_{n\to\infty} \frac{1}{n^d} \log\, Q\{\hat{P}_{x^{(n)},Y,C_n} \in G\} \leq -I_{Q|\mu}(\pi) - \epsilon.$$

Since $\hat{P}_{x^{(n)},C_n} \to \mu$, this contradicts the assumption that (μ, Q) satisfies the quasi-quenched process level LDP wrt μ with rate function $I_{Q|\mu}$.

By conditioning,

$$\begin{aligned}\frac{1}{n^d} \log \Pr\{\hat{P}_{X,Y,C_n} \in G\} &\geq \frac{1}{n^d} \log \Pr\{\hat{P}_{X,Y,C_n} \in G,\ \hat{P}_{X,C_n} \in U_n\} \\ &\geq \frac{1}{n^d} \log \inf_{x:\hat{P}_{x,C_n}\in U_n} Q\{\hat{P}_{x,Y,C_n} \in G\} + \frac{1}{n^d} \log P\{\hat{P}_{X,C_n} \in U_n\}\end{aligned}$$

Take $\liminf_{n\to\infty}$ on both sides and use (5.1) and (5.2) to get

$$\liminf_{n\to\infty} \frac{1}{n^d} \log \Pr\{\hat{P}_{X,Y,C_n} \in G\} \geq -I_{Q|\mu}(\pi) - I_P(\mu),$$

which leads to the lower bound.

To prove the upper bound, let F be a closed subset of $\mathcal{M}(\Sigma^{\mathbb{Z}^d})$. Fix $\epsilon, \delta > 0$. For each real valued function f, denote

$$f^\delta = \min\{f - \delta, 1/\delta\}.$$

Extend the definition of $I_{Q|\mu}$ so that $I_{Q|\mu}(\pi) \equiv 0$ if $I_P(\mu) = \infty$.

Fix $\mu \in \mathcal{M}(\Omega_X)$ and $\pi \in \mathcal{M}(\Sigma^{\mathbb{Z}^d})$. If $I_P(\mu) < \infty$, then by the assumption of Theorem 4, (μ, Q) satisfies the quasi-quenched LDP relative to μ. Because $\mathcal{M}(\Sigma^{\mathbb{Z}^d})$ is Hausdorff and regular (cf. Theorem D.8, [11]), π has an open neighborhood $G_{\mu,\pi} \subset \mathcal{M}(\Sigma^{\mathbb{Z}^d})$, such that

$$\inf_{\gamma\in\bar{G}_{\mu,\pi}} I_{Q|\mu}(\gamma) \geq I^\delta_{Q|\mu}(\pi). \tag{5.3}$$

(cf. (4.1.3), [11]). If $I_P(\mu) = \infty$, then by the extended definition of $I_{Q|\mu}$, such an open neighborhood obviously exists. On the other hand, by argument similar to that for (5.2), μ has an open neighborhood $U_{\mu,\pi} \subset \mathcal{M}(\Omega_X)$, such that

$$\limsup_{n\to\infty} \frac{1}{n^d} \log \sup_{x:\hat{P}_{x,C_n}\in U_{\mu,\pi}} Q\{\hat{P}_{x,Y,C_n} \in \bar{G}_{\mu,\pi}\} \leq -\inf_{\gamma\in\bar{G}_{\mu,\pi}} I_{Q|\mu}(\gamma) + \epsilon. \tag{5.4}$$

Furthermore, shrinking $U_{\mu,\pi}$ if necessary while keeping $\mu \in U_{\mu,\pi}$, $U_{\mu,\pi}$ satisfies

$$\inf_{\nu\in\bar{U}_{\mu,\pi}} I_P(\nu) \geq I^\delta_P(\mu). \tag{5.5}$$

Since $\mathcal{M}(\Omega_X)$ and $\mathcal{M}(\Sigma^{\mathbb{Z}^d})$ are compact, their closed subsets are compact as well. In particular, F is compact, and hence so is $\mathcal{M}(\Omega_X) \times F$. Therefore, there are (μ_i, π_i), $\mu_i \in F$, $\pi_i \in \mathcal{M}(\Omega_X)$, $i = 1, \ldots, N$, such that

$$\mathcal{M}(\Omega_X) \times F \subset \bigcup_{i=1}^N (U_{\mu_i,\pi_i} \times G_{\mu_i,\pi_i}).$$

Then, because X and Y are independent,

$$\begin{aligned}
\Pr\{\hat{P}_{X,Y,C_n} \in F\} &\le \sum_{i=1}^N \Pr\{\hat{P}_{X,C_n} \in U_{\mu_i,\pi_i}, \hat{P}_{X,Y,C_n} \in \bar{G}_{\mu_i,\pi_i}\} \\
&\le \sum_{i=1}^N \Big[\Pr\{\hat{P}_{X,C_n} \in \bar{U}_{\mu_i,\pi_i}\} \times \sup_{x:\hat{P}_{x,C_n} \in U_{\mu_i,\pi_i}} \Pr\{\hat{P}_{x,Y,C_n} \in \bar{G}_{\mu_i,\pi_i}\}\Big].
\end{aligned}$$

Because P satisfies the LDP with rate function I_P,

$$\limsup_{n\to\infty} \frac{1}{n^d} \log \Pr\{\hat{P}_{X,C_n} \in \bar{U}_{\mu_i,\pi_i}\} \le -\inf_{\nu \in \bar{U}_{\mu_i,\pi_i}} I_P(\nu).$$

By (5.3)–(5.5),

$$\begin{aligned}
\limsup_{n\to\infty} \frac{1}{n^d} \log \Pr\{\hat{P}_{X,Y,C_n} \in F\} &\le \max_i \{-I^\delta_{Q|\mu_i}(\pi_i) - I^\delta_P(\mu_i)\} + \epsilon \\
&\le -\inf_{\pi\in F} \inf_{\mu \in \mathcal{M}(\Omega_X)} \{I^\delta_{Q|\mu}(\pi) + I^\delta_P(\mu)\} + \epsilon = -B_\delta + \epsilon.
\end{aligned}$$

Since ϵ is arbitrary, in order to finish the proof, it is enough to show

$$\lim_{\delta \downarrow 0} B_\delta = A, \qquad A = \inf_{\pi \in F} \inf_{\mu\in\mathcal{M}(\Omega_X)} \{I_{Q|\mu}(\pi) + I_P(\mu)\}.$$

First, it is easy to see $B_\delta \le A$ and is non-increasing in $\delta > 0$. Therefore, $\limsup B_\delta \le A$. On the other hand, because rate functions are non-negative,

$$\begin{aligned}
I^\delta_{Q|\mu}(\pi) + I^\delta_P(\mu) &= \min\{I_{Q|\mu}(\pi) - \delta, \frac{1}{\delta}\} + \min\{I_P(\mu) - \delta, \frac{1}{\delta}\} \\
&\ge \min\{I_{Q|\mu}(\pi) + I_P(\mu) - 2\delta, \frac{1}{2\delta}\} - \delta \ge \min\{A - 2\delta, \frac{1}{2\delta}\} - \delta.
\end{aligned}$$

Take infimum on both sides over $\mu \in F$ and $\pi \in \mathcal{M}(\Sigma^{\mathbb{Z}^d})$, then let $\delta \downarrow 0$ to get $\liminf B_\delta \ge A$. □

Proposition 2 Given $\pi \in \mathcal{M}(\mathbb{Z}^d)$, if $I_{Q|\mu}(\pi)$ is lower semi-continuous in μ, then there is $\mu^* \in \mathcal{M}(\Omega_X)$, such that

$$I_{P\times Q}(\pi) = I_{Q|\mu^*}(\pi) + I_P(\mu^*). \tag{5.6}$$

Proof. If $I_{P\times Q}(\pi) = \infty$, then by (1.18), (5.6) holds for any $\mu \in \mathcal{M}(\Omega_X)$. If $I_{P\times Q}(\pi) < \infty$, choose $\mu_n \in \mathcal{M}(\Omega_X)$, so that $I_{Q|\mu_n}(\pi) + I_P(\mu_n) \downarrow I_{P\times Q}(\pi)$. Since $\mathcal{M}(\Omega_X)$ is compact and complete, there is a convergent subsequence $\mu_{n_i} \to \mu^* \in \mathcal{M}(\Omega_X)$ in weak topology. Because both $I_{Q|\mu}(\pi)$ and $I_P(\mu)$ are lower semi-continuous,

$$I_{Q|\mu^*}(\pi) + I_P(\mu^*) \le \liminf_{n\to\infty} (I_{Q|\mu_n}(\pi) + I_P(\mu_n)) = I_{P\times Q}(\pi)$$

which, combined with (1.18), yields (5.6). □

6 The quasi-quenched LDP for Gibbs random fields

Most part of this section is mainly the proof of Theorem 5. However, we first give a result in regard to Proposition 2 based on Corollary 2.

Corollary 4 If $Q \in \mathcal{G}(U) \cap \mathcal{M}_s(\Omega_Y)$, then for any $\pi \in \mathcal{M}(\Sigma^{\mathbb{Z}^d})$, the infimum in (1.18) is achieved.

Proof. By Proposition 2, it is enough to show $I_{Q|\mu}(\pi)$ is lower semi-continuous in μ. Given π, from (1.22) and (1.23), it is seen that $I_{Q|P}(\pi)$ as a function in P is the supremum of functions $E_\pi f - \Lambda_P(f) = E_\pi f - \lim_n \Lambda_{P,n}(f)$, over $f \in B(\Sigma^{C_h}, \mathbb{R})$, $h \geq 1$ where $\Lambda_P(f)$ and $\Lambda_{P,n}(f)$ are defined in (1.24). Since the supremum of semi-continuous functions is again lower semi-continuous, it suffices to show, for each such f, $\Lambda_P(f)$ is a continuous function in P.

Note that for any $P \in \mathcal{M}(\Omega_X)$ and $n \geq 1$, $|\Lambda_{P,n}(f)| \leq \|f\|$. Given $\epsilon > 0$, repeat the argument that leads to (3.14). It is seen that, for $h \ll k \ll n$, for *all* $P \in \mathcal{M}(\Omega_X)$,

$$\Lambda_{P,n}(f) \geq -(3\|f\| + 1)\epsilon - \epsilon|\Lambda_{P,k}(f)| + \Lambda_{P,k}(f) \geq -(4\|f\| + 1)\epsilon + \Lambda_{P,k}(f).$$

Likewise $\Lambda_{P,n}(f) \leq (4\|f\| + 1)\epsilon + \Lambda_{P,k}(f)$. Thus

$$\begin{aligned} & |\Lambda_{P,n}(f) - \Lambda_{P,k}(f)| \leq (4\|f\| + 1)\epsilon, \ \ h \ll k \ll n \\ \Longrightarrow \quad & |\Lambda_{P,n}(f) - \Lambda_{P,m}(f)| \leq 2(4\|f\| + 1)\epsilon, \ \ n, m \text{ large enough.} \end{aligned}$$

Therefore, $\Lambda_{P,n}(f)$ converges to $\Lambda_P(f)$ uniformly on $\mathcal{M}_s(\Omega_X)$. For each $n \in \mathbb{N}$, $\Lambda_{P,n}(f)$ is continuous in P, therefore, $\Lambda_P(f)$ is continuous in P as well. □

Proof of Theorem 5: To prove the upper bound, for $x \in \Omega_X$ and finite $V \subset \mathbb{Z}^d$, denote

$$\Lambda_{V,x}(\lambda) = \frac{1}{|V|} \log E_Q \left[e^{|V|\langle \lambda, \bar{f}(x,Y,V)\rangle} \right], \quad \Lambda_{n,x}(\lambda) = \Lambda_{C_n,x}(\lambda).$$

Then, by Theorem 4.5.3 of [11] and the boundedness of f, it is enough to show

$$\limsup_{n\to\infty} \Lambda_{n,x^{(n)}}(\lambda) \leq \Lambda(\lambda) \tag{6.1}$$

Let $g(u) = \langle \lambda, u \rangle$. We follow closely the argument that leads to (3.13) with two exceptions. First, instead of the first inequality of (2.7), the second inequality is used. Second, instead of the concavity of a linear function, its convexity is used. Again, denote $T_{n,k} = \{s : s + C_k \subset C_n\}$. For $1 \ll k \ll n$,

$$\Lambda_{n,x^{(n)}}(\lambda) X \leq (3K_g + 1)\epsilon + \frac{k^d}{(k+n_0)^d} \frac{1}{n^d} \sum_{s \in T_{n,k}} \Lambda_{k,\theta_s x^{(n)}}(\lambda).$$

Because $g(u) = \langle \lambda, u \rangle$, by (3.12),

$$\begin{aligned} & K_g \leq (1 + \|f\|)|\lambda|, \\ \Longrightarrow \quad & \Lambda_{n,x^{(n)}}(\lambda) \leq 3\epsilon(1 + \|f\|)|\lambda| + \epsilon + \frac{k^d}{(k+n_0)^d} \frac{1}{n^d} \sum_{s \in T_{n,k}} \Lambda_{k,\theta_s x^{(n)}}(\lambda). \end{aligned}$$

Also,

$$\begin{aligned}
&\qquad |\Lambda_{k,\theta_s x^{(n)}}(\lambda)| \le \|f\||\lambda|, \quad \text{and} \quad |T_{n,k}| = (1+o(1))n^d, \quad \frac{n}{k}, k\to\infty \\
\Longrightarrow &\qquad \Lambda_{n,x^{(n)}}(\lambda) \le 3\epsilon(2+\|f\|)|\lambda| + \epsilon + \frac{k^d}{(k+n_0)^d}\frac{1}{n^d}\sum_{s\in C_n}\Lambda_{k,\theta_s x^{(n)}}(\lambda) \\
&\qquad = 3\epsilon(2+\|f\|)|\lambda| + \epsilon + \frac{k^d}{(k+n_0)^d}\int \Lambda_{k,\zeta}(\lambda)\,\mathrm{d}\hat{P}_{x^{(n)},C_n}(\zeta)
\end{aligned}$$

Notice that $\Lambda_{k,\zeta}(\lambda)$ is a function only depending on the values of ζ_t, $t \in C_{k+n_0}$. Therefore, (1.21) implies the integral on the right end converges to $E_P\Lambda_{k,X}(\lambda)$ and hence

$$\limsup_{n\to\infty}\Lambda_{n,x^{(n)}}(\lambda) \le 3\epsilon(2+\|f\|)|\lambda| + \epsilon + \frac{k^d}{(k+n_0)^d}E_P\Lambda_{k,X}(\lambda).$$

Let $k\to\infty$. Because ϵ is arbitrary, $\limsup_n \Lambda_{n,x^{(n)}}(\lambda) \le \Lambda(\lambda)$.

To prove the lower bound, it is enough to show that for any open set $G \subset \mathbb{R}^\nu$ and $u \in G$,

$$\liminf_{n\to\infty}\frac{1}{n^d}\log\Pr\{\bar{f}_n(x^{(n)},Y)\in G\} \ge -\Lambda^*(u). \tag{6.2}$$

Fix $1 \ll p \ll k \ll n$ and $r > 0$, such that $B_{2r}(u) \subset G$. Denote I the collection of disjoint $(k+n_0)$-blocks $(k+n_0)t + C_{k+n_0}$, $t \in \mathbb{Z}^d$, that are contained in C_{n+n_0}. For each $V = s + C_{k+n_0} \in I$, let $k(V) = s + C_k$ and J be the collection of $k(V)$, $V \in I$. Define $W = \bigcup_{V\in J} V$. Then by (3.10), for $x \in \Omega_X$, $y \in \Omega_Y$,

$$\begin{aligned}
&\Big|\,\bar{f}_n(x,y) - \frac{1}{|J|}\sum_{V\in J}\bar{f}_V(x,y)\,\Big| \\
&= \Big|\,\frac{1}{n^d}\sum_{t\in C_n} f((\theta_t x,\theta_t y)_D) - \frac{1}{|J|}\frac{1}{k^d}\sum_{V\in J}\sum_{t\in V} f((\theta_t x,\theta_t y)_D)\,\Big| \\
&= \Big|\,\frac{1}{n^d}\sum_{t\in C_n} f((\theta_t x,\theta_t y)_D) - \frac{1}{|W|}\sum_{t\in W} f((\theta_t x,\theta_t y)_D)\,\Big| \\
&\le \frac{1}{n^d}\Big|\sum_{t\in C_n\setminus W} f((\theta_t x,\theta_t y)_D)\Big| + \Big(\frac{1}{|W|}-\frac{1}{n^d}\Big)\sum_{t\in W}|f((\theta_t x,\theta_t y)_D)| \\
&\le \frac{2|C_n\setminus W|\times\|f\|}{n^d} \ \le\ r.
\end{aligned}$$

Together with

$$\frac{1}{|J|}\sum_{V\in J}\bar{f}_V(x,Y) = \frac{1}{|I|}\sum_{V\in I}\bar{f}_{k(V)}(x,Y),$$

this implies

$$Q\left\{\bar{f}_n(x,Y)\in B_{2r}(u)\right\} \ge Q\left\{\frac{1}{|I|}\sum_{V\in I}\bar{f}_{k(V)}(x,Y)\in B_r(u)\right\}$$

Define random variable ξ_V, $V \in I$, such that

$$\xi_V \in S_Y^V, \qquad \xi_V \text{ i.i.d. } \sim Q_{C_{n+n_0}}. \tag{6.3}$$

Then by (2.7), for $y \in \Omega_Y$,

$$\Pr\{Y_V = y_V,\ V \in I\} \ge e^{-|I|K_{p,C_{k+n_0}}} \prod_{V\in I} \Pr\{\xi_V = y_V\}$$

$$\Longrightarrow \quad Q\{\bar f_n(x, Y) \in B_{2r}(u)\} \ge e^{-|I|K_{p,C_{k+n_0}}} \Pr\left\{ \frac{1}{|I|} \sum_{V\in I} \bar f_{k(V)}(x_V, \xi_V) \in B_r(u) \right\}. \tag{6.4}$$

For $n \ge 1$, let

$$\zeta_{n,V} = k^d\, \bar f_{k(V)}(x_V^{(n)}, \xi_V),\ V \in I, \qquad T_n = \frac{1}{k^d |I|} \sum_{V\in I} \zeta_{n,V}, \tag{6.5}$$

and μ_n the joint distribution of $\zeta_{n,V}$, $V \in I$. Then

$$\frac{1}{n^d} \log Q\{\bar f_n(x^{(n)}, Y) \in B_{2r}(u)\} \ge \frac{-|I|K_{p,C_{k+n_0}}}{n^d} + \frac{1}{n^d} \log \mu_n\{T_n \in B_r(u)\}. \tag{6.6}$$

To estimate $n^{-d} \log \mu_n\{T_n \in B_r(u)\}$, consider the log-moment generating function of T_n. For $\lambda \in \mathbb{R}^\nu$, let

$$\Lambda_{n,k}(\lambda) \quad = \quad \frac{1}{k^d |I|} \log E_{\mu_n} \left[e^{k^d |I| \langle \lambda, T_n\rangle} \right]. \tag{6.7}$$

First consider the case where $\Lambda_{n,k}(\lambda)$ satisfies the following condition,

(A) $\langle \lambda, u\rangle - \Lambda_{n,k}(\lambda)$ as a function in λ achieves

$$\Lambda^*_{n,k}(u) = \sup_{\lambda \in \mathbb{R}^\nu} \{\langle \lambda, u\rangle - \Lambda_{n,k}(\lambda)\}$$

at $\eta_{n,k}$, and $|\eta_{n,k}|$ is bounded for $1 \ll k \ll n$.

Lemma 6 Under condition (A),

$$\lim_{k\to\infty} \lim_{n\to\infty} \Lambda^*_{n,k}(u) = \Lambda^*(u).$$

Assume the lemma to be true for now. Define

$$\mathrm{d}\tilde\mu_n(\zeta_{n,V},\ V \in I) = \frac{e^{k^d\, |I|\, \langle \eta_{n,k},\, T_n\rangle}}{e^{k^d\, |I|\, \Lambda_{n,k}(\eta_{n,k})}} \,\mathrm{d}\mu_n(\zeta_{n,V},\ V \in I). \tag{6.8}$$

Then for any $\delta \in (0, r)$,

$$\begin{aligned}
\frac{1}{n^d} \log \mu_n\{T_n \in B_r(u)\} &\ge \frac{1}{n^d} \log E_{\tilde\mu_n} \left[\mathbf{1}_{\{T_n \in B_\delta(u)\}} \frac{e^{k^d\, |I|\, \Lambda_{n,k}(\eta_{n,k})}}{e^{k^d\, |I|\, \langle \eta_{n,k},\, T_n\rangle}} \right] \\
&= \frac{1}{n^d} \log E_{\tilde\mu_n} \left[\mathbf{1}_{\{T_n \in B_\delta(u)\}} \frac{e^{k^d\, |I|\, \Lambda_{n,k}(\eta_{n,k})}}{e^{k^d\, |I|\, \langle \eta_{n,k},\, u\rangle}} e^{k^d\, |I|\, \langle \eta_{n,k},\, u - T_n\rangle} \right] \\
&\ge -\frac{k^d |I|}{n^d} \Lambda^*_{n,k}(u) - \frac{k^d |I|}{n^d} |\eta_{n,k}| \delta + \frac{1}{n^d} \log \tilde\mu_n\{T_n \in B_\delta(u)\}.
\end{aligned} \tag{6.9}$$

Because $\langle\lambda, u\rangle - \Lambda_{n,k}(\lambda)$ is differentiable, by (A), $u = \nabla\Lambda_{n,k,}(\eta_{n,k})$. On the other hand, by (6.7) and (6.8),

$$\nabla\Lambda_{n,k,}(\eta_{n,k}) = \frac{E_{\mu_n}[T_n e^{k^d|I|\langle\eta_{n,k},T_n\rangle}]}{E_{\mu_n}[e^{k^d|I|\langle\eta_{n,k},T_n\rangle}]} = \frac{E_{\mu_n}[T_n e^{k^d|I|\langle\eta_{n,k},T_n\rangle}]}{e^{k^d|I|\Lambda_{n,k,}(\eta_{n,k})}} = E_{\tilde\mu_n}[T_n].$$

Therefore,

$$E_{\tilde\mu_n}\left[\frac{1}{|I|}\sum_{V\in I}\frac{1}{k^d}\zeta_{n,V}\right] = E_{\tilde\mu_n}[T_n] = u$$

Under μ_n, $\zeta_{n,V}$, $V \in I$ are independent. By (6.8), $k^{-d}\zeta_{n,V}$ are independent under $\tilde\mu_n$. In addition, $k^{-d}\zeta_{n,V}$ are bounded by $\|f\|$. Let $n\to\infty$ to get $|I|\to\infty$. By the weak law of large numbers, $\tilde\mu_n\{T_n \in B_\delta(u)\}\to 1$, as $n\to\infty$. Let $n\to\infty$ followed by $k\to\infty$ to get $n^{-d}k^d|I|\to 1$. By Lemma 6 and (6.9),

$$\liminf_{n\to\infty}\frac{1}{n^d}\log\mu_n\{T_n \in B_r(u)\} \ge -\Lambda^*(u) - \delta\limsup_{k\to\infty}\limsup_{n\to\infty}|\eta_{n,k}|. \tag{6.10}$$

Together with (6.6) and (2.6), this implies

$$\liminf_{n\to\infty}\frac{1}{n^d}\log Q\{\bar f_n(x,Y)\in B_{2r}(u)\} \quad\ge\quad -4\gamma_p - \Lambda^*(u) - \delta\,\limsup_{k\to\infty}\limsup_{n\to\infty}|\eta_{n,k}|.$$

By assumption (A), $\limsup_k\limsup_n|\eta_{n,k}|$ is finite. Since δ is arbitrary, letting $p\to\infty$ finishes the proof of (6.2).

To prove (6.2) without (A), μ_n is regularized as follows. Given $M > 0$, let $W_t \in \mathbb{R}^\nu$, $t \in \mathbb{Z}^d$ be i.i.d. random variables taking finite values, such that

$$\log E\left[e^{\langle\lambda,W_t\rangle}\right] \ge \begin{cases} \dfrac{1}{4M}|\lambda|^2 & \text{if } |\lambda| \le 4(\|f\|+|u|+1)M \\ (\|f\|+|u|+1)|\lambda| & \text{if } |\lambda| > 4(\|f\|+|u|+1)M \end{cases} \tag{6.11}$$

Such random variables can be obtained by appropriate quantization of $\mathbf{1}_{\{N_t\le R\}}N_t/\sqrt{M}$ with suitable R, where N_t are i.i.d. standard normal random variables. Suppose the support of W_t is S_2. Then it is seen (Y,W) is a stationary Gibbs field with summable interaction potentials on $(S_Y\times S_2)^{\mathbb{Z}^d}$. For $x\in\Omega_X$, $(y,w)\in(S_Y\times S_2)^{\mathbb{Z}^d}$, define

$$F(x,(y,w)) = f(x,y) + w(\mathbf{0}),$$

where $w(\mathbf{0})$ is the value of w at the origin $\mathbf{0}$. Modify ξ_V in (6.3) so that they are independent of W. Then, defining $\bar F_n$ by (3.1)

$$\begin{aligned}&\Pr\left\{\bar F_n(x,(Y,W))\in B_{2r}(u)\right\}\\ &\ge e^{-|I|K_{p,C_{k+n_0}}}\Pr\left\{\frac{1}{|I|}\sum_{V\in I}\bar f_{k(V)}(x_V,\xi_V) + \frac{1}{|I|}\sum_{V\in I}\frac{1}{k^d}\sum_{t\in k(V)}W_t \in B_r(u)\right\}.\end{aligned}$$

Comparing (6.4)–(6.7), this suggest the following modification be used,

$$\zeta'_{n,V} = \zeta_{n,V} + \sum_{t\in k(V)}W_t,\quad V\in I,\quad \mu'_n = \text{the joint distribution of } \zeta'_{n,V}, V\in I,$$

$$T'_n = T_n + \frac{1}{k^d|I|}\sum_{V\in I}\sum_{t\in k(V)}W_t = T_n + \frac{1}{k^d|I|}\sum_{V\in J}\sum_{t\in V}W_t$$

$$\Lambda'_{n,k}(\lambda) = \Lambda_{n,k}(\lambda) + \log E\left[e^{\langle\lambda,W_{\mathbf{0}}\rangle}\right],$$

By (6.11), $\Lambda'_{n,k}(\lambda) \ge \Lambda'_{n,k}(\lambda)$, and hence $\Lambda'^*_{n,k}(u) \le \Lambda^*_{n,k}(u)$. Since $|\Lambda_{n,k}(\lambda)| \le \|f\||\lambda|$, then for $|\lambda| > 4(\|f\| + |u| + 1)$,

$$\langle\lambda, u\rangle - \Lambda'_{n,k}(\lambda) = \langle\lambda, u\rangle - \Lambda_{n,k}(\lambda) - \log E\left[e^{\langle\lambda, W_{\mathbf{0}}\rangle}\right] \le (|u| + \|f\|)|\lambda| - (\|f\| + |u| + 1)|\lambda| < 0.$$

Since $\Lambda'^*_{n,k}(u)$ is non-negative, this implies that it must be achieved within $|\lambda| \le 4(\|f\|+|u|+1)$ and hence $\langle\lambda, u\rangle - \Lambda'_{n,k}(\lambda)$ satisfies (A). Then for $r > 0$, the preceding argument leads to

$$\frac{1}{n^d}\log \mu'_n\{T'_n \in B_{r/2}(u)\} \ge -\Lambda'^*(u) \ge -\Lambda^*(u)$$

On the other hand, by the definition of T'_n,

$$\mu_n\{T_n \in B_r(u)\} + \Pr\left\{\frac{1}{k^d|I|}\Big|\sum_{V\in J}\sum_{t\in V} W_t\Big| \ge \frac{r}{2}\right\} \ge \mu'_n\{T'_n \in B_{r/2}(u)\}$$

By LDP for W_t,

$$\begin{aligned}
&\limsup_{k\to\infty}\limsup_{n\to\infty}\frac{1}{k^d|J|}\log\Pr\left\{\frac{1}{k^d|I|}\Big|\sum_{V\in J}\sum_{t\in V} W_t\Big| \ge \frac{r}{2}\right\}\\
\le\;& -\inf_{|u|>r/2}\sup_{\lambda\in\mathbb{R}^\nu}\left\{\langle\lambda, u\rangle - \log E[e^{\langle\lambda, W_{\mathbf{0}}\rangle}]\right\}\\
\le\;& -\inf_{|u|>r/2}\sup_{|\lambda|\le 4|u|M}\left\{\langle\lambda, u\rangle - \frac{|\lambda|^2}{4M}\right\} \le -\frac{Mr^2}{4}.
\end{aligned}$$

Therefore,

$$\max\left\{\liminf_{n\to\infty}\log\frac{1}{n^d}\mu_n\{T_n \in B_r(u)\},\ -\frac{Mr^2}{4}\right\} \ge \liminf_{n\to\infty}\frac{1}{n^d}\log\mu'_n\{T'_n \in B_{r/2}(u)\} \ge \Lambda^*(u)).$$

Letting $n\to\infty$ followed by $M\to\infty$ then finishes the proof. □

Proof of Lemma 6: The proof is based on the following result.

Proposition 3 Whether (A) is satisfied or not,

$$\lim_{k\to\infty}\lim_{n\to\infty}\Lambda_{n,k}(\lambda) = \Lambda(\lambda). \tag{6.12}$$

Assume Proposition 3 to be true for now. When (A) is satisfied, we can fix $M > 0$, such that $|\eta_{n,k}| < M$. Because $\Lambda_{n,k}$ and Λ are convex, (6.12) implies that the convergence is uniform on for $|\lambda| \le M$. Therefore, given $\epsilon > 0$, when $n, k \gg 1$, $\Lambda_{n,k}(\eta_{n,k}) > \Lambda(\eta_{n,k}) - \epsilon$, implying

$$\Lambda^*_{n,k}(u) = \langle\eta_{n,k}, u\rangle - \Lambda_{n,k}(\eta_{n,k}) < \langle\eta_{n,k}, u\rangle - \Lambda(\eta_{n,k}) + \epsilon \le \Lambda^*(u) + \epsilon.$$

Thus $\limsup_k \limsup_n \Lambda^*_{n,k}(u) \le \Lambda^*(u)$.

On the other hand, choose $\lambda_i \in \mathbb{R}^\nu$, such that if $\Lambda^*(u) < \infty$, then $\langle\lambda_i, u\rangle - \Lambda(\lambda_i) \ge \Lambda^*(u) - 1/i$, and if $\Lambda^*(u) = \infty$, $\langle\lambda_i, u\rangle - \Lambda(\lambda_i) \ge i$. Then,

$$\liminf_{k\to\infty}\liminf_{n\to\infty}\Lambda^*_{n,k}(u) \ge \langle\lambda_i, u\rangle - \lim_{k\to\infty}\lim_{n\to\infty}\Lambda_{n,k}(\lambda_i) = \langle\lambda_i, u\rangle - \Lambda(\lambda_i),$$

leading to $\liminf_k \liminf_n \Lambda^*_{n,k}(u) \ge \Lambda^*(u)$. □

Proof of Proposition 3: The idea of the proof is as follows. By (6.7),

$$\Lambda_{n,k}(\lambda) = \frac{1}{k^d|I|}\sum_{V\in I}\log\int e^{k^d\langle\lambda,\,\bar f_{k(V)}(x_V^{(n)},\xi)\rangle}\,\mathrm{d}Q_V(\xi), \tag{6.13}$$
$$I = \text{the set of disjoint blocks } (k+n_0)t + C_{k+n_0} \text{ in } C_{n+n_0},\ t\in\mathbb{Z}^d.$$

Since I does not contain all the $(k+n_0)$-blocks in C_{n+n_0}, the assumption $\hat P_{x^{(n)},C_n}\to P$ can not apply directly to the sum on the right hand side of (6.13). To solve this problem, given $1\ll h\ll k$, each $(k+n_0)$-block in I is divided into h-blocks and $\Lambda_{n,k}$ is approximated by the sum of the log-moment generating functions of $\bar f_V(x^{(n)},\xi)$, across all h-blocks V contained in the blocks in I. When n and k are large enough, the collection of such h-blocks approximate that of all h-blocks in C_{n+n_0}, with $o(n^d)$ difference in cardinality. Then we can apply the assumption $\hat P_{x^{(n)},C_n}\to P$ to the approximating sum to get the convergence.

Fix $\epsilon>0$ and $n_0\ll p\ll h\ll k\ll n$. Define $T_{k,h}$ and $T_{n,h}$ in the same way as $T_{n,k}$ in (3.9). Let $A_{n,k}=\{s: s+C_{k+n_0}\in I\}$. Then, given $\lambda\in\mathbb{R}^\nu$, by (6.13) and the stationarity of Q,

$$\Lambda_{n,k}(\lambda) = \frac{1}{|I|}\sum_{s\in A_{n,k}}\Lambda_{k,\theta_s x^{(n)}}(\lambda), \quad \Lambda_{k,\theta_s x^{(n)}}(\lambda) \triangleq \frac{1}{k^d}\log E_Q\left[e^{k^d\langle\lambda,\bar f_k(\theta_s x^{(n)},Y)\rangle}\right] \tag{6.14}$$

In fact, $\Lambda_{k,\theta_s x^{(n)}}(\lambda)$ is identical to $\Lambda_{k,\theta_s x^{(n)}}$ defined in (3.6), with $g(\cdot)=\langle\lambda,\cdot\rangle$. The more complex notation is used to stress the dependence on λ. By (3.12), $K_g\le(1+\|f\|)|\lambda|$. Then (3.13) yields

$$\Lambda_{k,\theta_s x^{(n)}}(\lambda) \ge -\epsilon - 3(1+\|f\|)|\lambda|\epsilon + \frac{|C_h|}{k^d|C_{h+n_0}|}\sum_{t\in T_{k,h}}\Lambda_{h,\theta_{t+s}x^{(n)}}(\lambda). \tag{6.15}$$

Combine (6.14) and (6.15) to get

$$\Lambda_{n,k}(\lambda) \ge -\epsilon - 3(1+\|f\|)|\lambda|\epsilon + \frac{|C_h|}{k^d|C_{h+n_0}||I|}\sum_{s\in A_{n,k}}\sum_{t\in T_{k,h}}\Lambda_{h,\theta_{s+t}x^{(n)}}(\lambda). \tag{6.16}$$

The map $(s,t)\mapsto s+t$ is one-to-one from $A_{n,k}\times T_{k,h}$ onto $W=\{s+t: s\in A_{n,k}, t\in T_{k,h}\}\subset T_{n,h}$. Indeed, for $s,s'\in A_{n,k}$, $t,t'\in T_{k,h}$, if $s+t=s'+t'$, then $s+t+C_h=s'+t'+C_h$. Since $t+C_h$, $t'+C_h\subset C_k$, then $s+C_k$ and $s'+C_k$ intersect. Both belong to I, which consists of disjoint blocks. It is seen $s=s'$, implying $t+C_h$ and $t'+C_h$ intersect, and hence $t=t'$. By (6.16),

$$\Lambda_{n,k}(\lambda) \ge -\epsilon - 3(1+\|f\|)|\lambda|\epsilon + \frac{|C_h|}{k^d|C_{h+n_0}||I|}\sum_{t\in W}\Lambda_{h,\theta_t x^{(n)}}(\lambda). \tag{6.17}$$

Recall $1\ll h\ll k\ll n$. It is easy to see

$$\begin{gathered}
|W| = |A_{n,k}||T_{k,h}| \ge |T_{n,h}|(1-\epsilon) \implies |T_{n,h}\setminus W| \le \epsilon|T_{n,h}| \\
\frac{|C_h|}{k^d|C_{h+n_0}||I|} = \frac{1}{n^d}(1+o(1)), \quad |T_{n,h}| = (1+o(1))|C_n|, \quad |\Lambda_{h,x^{(n)}}(\lambda)| \le \|f\||\lambda| \\
\implies \quad \Lambda_{n,k}(\lambda) \ge -M\epsilon + \frac{1}{n^d}\sum_{t\in C_n}\Lambda_{h,\theta_t x^{(n)}}(\lambda) = -M\epsilon + \int \Lambda_{h,\xi}(\lambda)\,\hat P_{x^{(n)},C_n}(\mathrm{d}\xi),
\end{gathered}$$

where M is a constant only depending on f and λ. Let $n,k,h\to\infty$ in sequel. As ϵ is arbitrary, $\liminf_k\liminf_n\Lambda_{n,k}(\lambda)\ge\Lambda(\lambda)$. Similarly, $\limsup_k\limsup_n\Lambda_{n,k}(\lambda)\le\Lambda(\lambda)$. □

Acknowledgment

I am grateful to Ioannis Kontoyiannis for suggesting the topic of conditional LDP for Gibbs random fields and referring me to the paper of Föllmer and Orey [13]. I would also like to thank both him and Amir Dembo for valuable comments on the manuscript.

References

[1] Leo Breiman. *Probability.* Addison-Wesley, Inc., Reading, Massachusetts, 1968.

[2] Włodzimierz Bryc. On the large deviation principles for stationary weakly dependent random fields. *Ann. Probab.*, 20(2):1004–1030, 1992.

[3] Narasinga R Chaganty. Large deviations for joint distributions and statistical applications. *Sankhyā* **A***: The Indian J. Statist.*, 59(2):147–166, 1997.

[4] Zhiyi Chi. The first order asymptotic of waiting times with distortion between stationary processes. *IEEE Trans. Infor. Theory*, 47(1):338–347, Jan 2001.

[5] Zhiyi Chi. Stochastic sub-additivity approach to conditional large deviation principle. *Ann. Probab.*, 29(3):1303–1328, 2001.

[6] F. Comets. Large deviation estimates for a conditional probability distribution. Applications to random interaction Gibbs measures. *Probab. Theory Related Fields*, 80:407–432, 1989.

[7] F. Comets, Nina Gantert, and Ofer Zeitouni. Quenched, annealed and functional large deviations for one dimensional random walk in random environment. *Probab. Theory Related Fields*, 118:65–114, 2000.

[8] D. Dawson and J. Gärtner. Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics*, 20:247–308, 1987.

[9] Amir Dembo and Ioannis Kontoyiannis. The asymptotics of waiting times between stationary processes, allowing distortion. *Ann. Appl. Probab.*, 9:413–429, May 1999.

[10] Amir Dembo and Ioannis Kontoyiannis. Source coding, large deviations, and approximate pattern matching. *IEEE Trans. Infor. Theory*, 48(6):(in press), 2002.

[11] Amir Dembo and Ofer Zeitouni. *Large Deviations Techniques And Applications.* Springer-Verlag, New York, 1998. Second edition.

[12] J. D. Deuschel and D. W. Stroock. *Large Deviations.* Academic Press, Boston, MA, 1989.

[13] Hans Föllmer and Stenven Orey. Large deviations for the empirical field of a Gibbs measure. *Ann. Probab.*, 16(3):961–977, 1988.

[14] Nina Gantert and Ofer Zeitouni. Large deviations for one-dimensional radnom walk in a random environment – a survey. In P. Revesz and B. Toth, editors, *Proceedings of the Conference on Random Walks, Budapest, 1998*, pages 127–165, Budapest, Hungary, 1999. Bolyai Society Mathematical Studies 9.

[15] C. Preston. *Random Fields*, volume 534 of *Lecture Notes in Math.* Springer, New York, 1976.

[16] R. A. Wijsman. Convergence of sequences of convex sets, cones, and functions. *Bull. Amer. Math. Soc.*, 70:186–188, 1964.

[17] En-hui Yang and Zhen Zhang. On the redundancy of lossy source coding with abstract alphabets. *IEEE Trans. Infor. Theory*, 45(4):1092–1110, May 1999.

Zhiyi Chi
Department of Statistics
The University of Chicago
5734 University Avenue
Chicago, IL, 60637