JOURNAL OF MULTIVARIATE ANALYSIS 51, 414–431 (1994)

Mixed Limit Theorems for Pattern Analysis

ULF GRENANDER*

Brown University

AND

JAYARAM SETHURAMAN†

Florida State University

Limit theorems are derived for probability measures of random configurations over graphs which are used as prior distributions in pattern theory. For one-dimensional graphs, these limits can be viewed as distributions of certain stochastic processes, while in higher dimensions the limits will in some cases have to be interpreted as belonging to Schwartz distributions. Such limit distributions are easy to use in pattern analysis, and greatly reduce the computing effort required in comparison with stochastic relaxation methods. © 1994 Academic Press, Inc.

1. INTRODUCTION

Consider a *graph* $(\mathcal{S}, \sigma)$ where $\mathcal{S}$ is a collection of *sites* and σ is the collection of *edges* or *segments* in $\mathcal{S}$. An example of such a graph is the one-dimensional graph with sites $\mathbf{Z}_n = \{0, 1, ..., n-1\}$ with pairs of successive sites like $(i, i+1)$ as edges, wherein we identify n with 0. A *configuration* $\mathbf{x} = (x_i, i \in \mathcal{S})$ is a point in $\mathcal{X}^{\mathcal{S}}$, that is, a map from $\mathcal{S}$ into $\mathcal{X}$. The space $\mathcal{X}$ is usually called the *generator space*. This paper is concerned with distributions of random configurations. Such distributions appear as both prior and posterior distributions in Bayesian pattern analysis, where they play an important role.

Consider the simple case when $\mathcal{S}$ is the finite set $\mathbf{Z}_n = \{0, 1, ..., n-1\}$ and $\mathcal{X} = \mathcal{R}$ the real line. A reasonably general distribution for random configurations can be described as follows, based on two non-negative

Received March 10, 1993; revised February 18, 1994.

AMS 1980 subject classifications: primary 62E20; secondary 62N05, 62G05.

Key words and phrases: pattern analysis, mixed limit theorems, Bayesian image analysis.

* Research supported by Army Research Office Grant ARO GC-R-312279.

† Research supported by Army Research Office Grant DAAL03-90-G-0103, DAAH04-93-G-0201.

0047-259X/94 $6.00

functions, the *acceptor function* $A: \mathscr{X} \times \mathscr{X} \to \mathscr{R}$ and the *weight function* $Q: \mathscr{X} \to \mathscr{R}$. A distribution of random configurations, often used as a prior distribution, is defined as follows through its density function, with respect to some fixed measure, (often Lebesgue or counting measure),

$$\frac{1}{Z} \prod_{\sigma} A(y_{i_1}, y_{i_1}) \prod_{i=0}^{n-1} Q(y_i)$$

where the σ under the first product sign means that the product is taken over all pairs (i_1, i_2) which are connected by a segment σ in the graph. The constant Z that normalizes this density function to have mass one is the so-called partition function.

Such prior measures are known in pattern theory as *regularly controlled probabilities* and in physics as *Gibbs distributions*, and have also appeared in other applications.

The weight function Q describes how frequent the various values in the generators space are expected to be. Note that Q need not, in general, be a marginal density.

The acceptor function A expresses the coupling between y_{i_1} and y_{i_2} if i_1 and i_2 are joined by a segment in σ. If A decreases when y_{i_1} and y_{i_2} become more distant from each other the coupling is attractive; in the opposite case the coupling is repellent.

When $\mathscr{X}$ is a vector space one often writes $A((y_{i_1} - y_{i_2})/\varepsilon)$ instead of $A(y_{i_1}, y_{i_2})$, where ε is a *coupling parameter*. Suppose that $A(y)$ is decreasing in $\|y\|$. A small ε means strong coupling, and the values at neighboring sites will tend to be more alike. A large ε means weak coupling and neighboring sites will tend to be more alike.

Probability measures of this type depend upon Q and A in a way that is not very direct. In applications, for example to image processing, one of the first tasks, after choosing σ, Q, and A, is to synthesize the patterns, i.e., simulate the prior measure in order to get some insight in how the assumptions fit the situation one tries to model.

Such simulations can always be done, at least in principle, by stochastic relaxation which means the following. Using the well known fact that the prior above describes a Markov process $(Y_i, i = 0, 1, ..., n-1)$, one can use a variation of the Metropolis algorithm; see Metropolis *et al.* (1953). For any given site i, consider the conditional distribution of Y_i given the rest of the configuration $(Y_0, Y_1, ..., Y_{n-1})$. The Markovian nature of the process tells us that the conditional distribution depends only on the neighbors of Y_i. Simulate the conditional distribution, and go to another site. Repeat the procedure by sweeping across the whole configuration may times. It is then known that the random configuration converges in law to the prior from which we started.

The advantage of stochastic relaxation is its general applicability. Unfortunately it requires massive computation and may not be feasible when the following conditions apply:

(i) the size n of the graph is big,

(ii) the cardinality of the generator space $\mathscr{X}$ is large or infinite,

(iii) the couplings are strong, i.e., ε is small, or

(iv) the conditional distributions are not of standard (easily simulated form).

To deal with situations where these conditions hold, attempts have been made to prove limit theorems when $n \to \infty$, or $\varepsilon \to 0$, or the so-called *mixed case* when n and ε tend to their respective limits simultaneously. The first result, see Chow and Grenander (1985), assumed σ to be a linear chain, $\mathscr{X} = \mathscr{R}$, and used an acceptor function like a rectangular window

$$A(y) = \begin{cases} 1, & \text{if } |y| \leqslant 1, \\ 0, & \text{if } |y| > 1. \end{cases} \tag{1.1}$$

It showed, under conditions that will not be stated here, that $Y_i/\sqrt{\varepsilon}$ tends in distribution to a non-degenerate Gaussian distribution.

Attempts have been made, so far without success, to extend this to more general graphs, in particular to lattice type connectors that are used in image processing. The method in Chow and Grenander (1985) employed singular perturbation methods for integral operators. Here we shall use a more probabilistic method that extends to all the lattice connectors that we have examined. It uses the acceptor function $A(x) = \exp(-x^2)$ and some of its simple extensions. The choice of acceptor function is probably not essential as has been shown in Chow and Grenander (1985).

We shall prove not only that marginal distributions converge but that the entire joint distribution converges to the joint distribution of a particular stochastic process. In some cases, however, it will be necessary to interpret the sample "functions" of the limiting stochastic process as Schwartz distributions.

This paper will only deal with pattern synthesis, but it is known that the understanding of the limiting behavior of the priors will have consequences for analysis, for example, recognition of patterns, and facilitate the computing needed.

A reader who is not familiar with the background of the paper may consult Grenander (1976, pp. 63–92), Grenander (1981, pp. 194–317), and Geman and Geman (1984).

2. Convergence of Gaussian Markov Random Fields

Let the joint distribution of $\mathbf{Y}_n = (Y_0, Y_1, ..., Y_{n-1})$ in $\mathscr{R}^n$ be given by a pdf (w.r.t. Lebesgue measure) which is proportional to

$$\exp\left\{-\frac{1}{2\varepsilon^2}\sum_{i=0}^{n-1}(y_{i+1}-y_i)^2 - \frac{q}{2}\sum_{i=0}^{n-1} y_i^2\right\},$$

where $y_0 = y_n$. We will consider the normalized configuration $\mathbf{X}_n$ defined by $\mathbf{X}_n = (Y_0/\sqrt{\varepsilon}, ..., Y_{n-1}/\sqrt{\varepsilon})$. The joint distribution of $\mathbf{X}_n$ will be denoted by P_n and its pdf $p_n(\mathbf{x})$ is given by

$$p_n(\mathbf{x}) = \frac{1}{Z_n(q, \varepsilon)} \exp\left\{-\frac{1}{2\varepsilon}\sum_{i=0}^{n-1}(x_{i+1}-x_i)^2 - \frac{q\varepsilon}{2}\sum_{i=0}^{n-1} x_i^2\right\}. \tag{2.1}$$

An explicit expression for the partition function $Z_n(q, \varepsilon)$ is given by

$$Z_n(q, \varepsilon) = \int \exp\left\{-\frac{1}{2\varepsilon}\sum_{i=0}^{n-1}(x_{i+1}-x_i)^2 - \frac{q\varepsilon}{2}\sum_{i=0}^{n-1} x_i^2\right\} d\mathbf{x}. \tag{2.2}$$

Let M be the circle of unit circumference. Define a map $g_n(\mathbf{x})$ from the space of configurations $\mathbf{x} = (x_0, ..., x_{n-1})$ into $C(M)$, the space of continuous functions on M, as

$$g_n(\mathbf{x}, t) = x_{[nt]} + n\left(t - \frac{[nt]}{n}\right)(x_{[nt]+1} - x_{[nt]}), \tag{2.3}$$

where, as before, we identify x_n with x_0. Consider the process $\{X_n(t) = g_n(\mathbf{X}_n, t),\ t \in M\}$. Let $\mathscr{P}_n =^{\text{def}} P_n g_n^{-1}$ be the distribution of $\{X_n(t), t \in M\}$ in $C(M)$.

Let $\{X(t), t \in M\}$ be a stationary Gaussian Markov process with zero means and with covariance function $R(s, t)$ given by

$$R(s, t) = \frac{\cosh((s - t - 1/2)\sqrt{q})}{2\sqrt{q}\sinh(\sqrt{q}/2)}. \tag{2.4}$$

We will denote the distribution of this process by $\mathscr{P}$.

Theorem 1 below will show that $\{\mathbf{X}_n(t),\ t \in M\}$ will converge, under some conditions, to the process $\{X(t), t \in M\}$, i.e., $P_n g_n^{-1} \to \mathscr{P}$ weakly.

Throughout this paper we will be taking limits as $\varepsilon \to 0$, $n \to \infty$ such that $n\varepsilon \to 1$. Such limits are known as mixed limits. For simplicity, we will just say $n \to \infty$ to mean this kind of mixed limit, except in the formal statements of theorems.

THEOREM 1. *Let $\varepsilon \to 0$ and $n \to \infty$ such that $n\varepsilon \to 1$. Then*

$$\{X_n(t), t \in M\} \xrightarrow{d} \{X(t), t \in M\}.$$

Proof. We will first show that

$$R_n(s, t) \to R(s, t) \qquad \text{uniformly in } s, t, \tag{2.5}$$

where $R_n(s, t)$ is the covariance function of $\{X_n(t), t \in M\}$ and $R(s, t)$ is the covariance function of $\{X(t), t \in M\}$. This will establish that the finite dimensional distributions converge.

The pdf of $\mathbf{X}_n$ given in (2.1) can be rewritten as

$$\frac{1}{Z_n(q, \varepsilon)} \exp\left\{-\frac{1}{2}\mathbf{x}'A\mathbf{x}\right\},$$

where

$$A = \begin{pmatrix} \frac{2}{\varepsilon} + q\varepsilon & -\frac{1}{\varepsilon} & 0 & \cdots & -\frac{1}{\varepsilon} \\ -\frac{1}{\varepsilon} & \frac{2}{\varepsilon} + q\varepsilon & -\frac{1}{\varepsilon} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -\frac{1}{\varepsilon} & 0 & 0 & \cdots & \frac{2}{\varepsilon} + q\varepsilon \end{pmatrix}$$

and

$$Z_n(q, \varepsilon) = (2\pi)^{n/2} \, |A|^{-1/2}. \tag{2.6}$$

To evaluate limiting values of $Z_n(q, \varepsilon)$ and the limit distribution of $\mathscr{P}_n$ we wil need the following standard results on the inverses of some matrices; see, e.g., Davis (1979).

Let

$$D_\rho \stackrel{\text{def}}{=} \begin{pmatrix} 1 & -\rho & 0 & \cdots 0 \\ 0 & 1 & -\rho & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -\rho & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Its inverse D_ρ^{-1} is given by

$$D_\rho^{-1} = \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho^{n-1} & 1 & \rho & \cdots & \rho^{n-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \rho & \rho^2 & \rho^3 & \cdots & 1 \end{pmatrix} \cdot \frac{1}{(1-\rho^n)},$$

and further more the matrix $B_\rho =^{\text{def}} D'_\rho D_\rho$ is given by

$$B_\rho = \begin{pmatrix} 1+\rho^2 & -\rho & 0 & \cdots & -\rho \\ -\rho & 1+\rho^2 & -\rho & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -\rho & 0 & 0 & \cdots & 1+\rho^2 \end{pmatrix}.$$

Finally $B_\rho^{-1} = D_\rho^{-1}(D'_\rho)^{-1}$ is given by $1/((1-\rho^2)(1-\rho^n))\,E_\rho$ where the (r, s)th element of E_ρ is given by

$$E_\rho(r, s) = \rho^{s-r} + \rho^{n-s+r}$$

for $0 \leqslant r,\ s \leqslant n$.

Using the definitions of B_ρ above, we may write

$$A = \alpha B_\rho,$$

where

$$\alpha(1+\rho^2) = \frac{2}{\varepsilon} + q\varepsilon \tag{2.7}$$

$$\alpha\rho = \frac{1}{\varepsilon}. \tag{2.8}$$

Thus

$$C \stackrel{\text{def}}{=} A^{-1} = \frac{1}{\alpha} B_\rho^{-1} = \frac{1}{\alpha(1-\rho^2)(1-\rho^n)} E_\rho,$$

where E_ρ is as defined above. Thus, if $0 \leqslant s,\ t \leqslant 1$, then the $([nt], [ns])$th element of C is

$$\frac{1}{\alpha(1-\rho^2)(1-\rho^n)} [\rho^{[ns]-[nt]} + \rho^{n-[ns]+[nt]}]. \tag{2.9}$$

Solving for α and ρ from (2.7) and (2.8), we get

$$\rho = 1 - \varepsilon\sqrt{q} + O(\varepsilon^2) \sim 1$$

and

$$\alpha \sim \frac{1}{\varepsilon}(1 + \varepsilon\sqrt{q}) \sim \frac{1}{\varepsilon}.$$

Thus the limits of (2.9) as $n \to \infty$ becomes

$$\lim \frac{\varepsilon[(1-\varepsilon\sqrt{q}+O(\varepsilon^2))^{[ns]-[nt]}+(1-\varepsilon\sqrt{q}+O(\varepsilon^2))^{n-[ns]+[nt]}]}{(1-(1-\varepsilon\sqrt{q}+O(\varepsilon^2))^2)(1-(1-\varepsilon\sqrt{q}+O(\varepsilon^2))^n)}$$
$$=\frac{\exp(-(s-t)\sqrt{q})+\exp(-(1-s+t)\sqrt{q})}{2\sqrt{q}\,(1-\exp(\sqrt{q}))}$$
$$=\frac{\cosh((s-t-1/2)\sqrt{q})}{2\sqrt{q}\sinh(\sqrt{q}/2)}$$
$$=R(s,t).$$

This shows that $R_n([ns]/n, [nt]/n) \to R(s,t)$ uniformly in s, t. To show that $R_n(s,t) \to R(s,t)$, we note that

$$\begin{aligned}|X_n(s)-X_n([ns]/n)| &= |n(s-[ns]/n)|\,|X_{[ns]+1}-X_{[ns]}|\\ &\leqslant |X_{[ns]+1}-X_{[ns]}| \overset{d}{=} |X_0-X_1|\end{aligned}$$

and hence

$$\begin{aligned}E\,|X_n(s)-X_n([ns]/n)|^2 &\leqslant 2[C_{00}-C_{01}]\\ &=\frac{2[1+\rho^n-\rho-\rho^{n-1}]}{\alpha(1-\rho^2)(1-\rho^n)}\\ &=\frac{2(1-\rho^{n-1})}{\alpha(1+\rho)(1-\rho^n)}\\ &\sim\frac{2\varepsilon(1-(1-\varepsilon\sqrt{q})^{n-1})}{(1+1-\varepsilon\sqrt{q})(1-(1-\varepsilon\sqrt{q})^n)}\\ &\sim\varepsilon\\ &\to 0\end{aligned}$$

uniformly in s as $n \to \infty$. This proves that

$$R_n(s,t) \to R(s,t)$$

uniformly for all s, t and completes the proof of (2.5).

We next show that the distributions of $\{X_n(t), t \in M\}$ are tight. This will complete the proof of the theorem.

Since the distributions are Gaussian, it follows from (2.5) that

$$\begin{aligned}
E[|X_n(t)-X_n(s)|^4] &= 3(E[|X_n(t)-X_n(s)|^2])^2 \\
&= 3[R_n(t,t)-2R_n(s,t)+R_n(s,s)]^2 \\
&\to 3[R(t,t)-2R(s,t)+R(s,s)]^2 \qquad \text{uniformly in } (s,t) \\
&= 3\,\frac{[\cosh(\sqrt{q}/2)-\cosh((s-t-1/2)\sqrt{q})]^2}{q[\sinh(\sqrt{q}/2)]^2} \\
&\leqslant 3\left(\frac{\sinh(\sqrt{q})}{\sinh(\sqrt{q}/2)}\right)^2 (s-t)^2 \qquad \text{if} \quad |s-t|\leqslant 1/2
\end{aligned}$$

by using the fact that $\sinh(x)$, the derivative of $\cosh(x)$, is increasing in x. This shows that there is a $K<\infty$ and an n_0 such that

$$E[|X_n(t)-X_n(s)|^4] \leqslant K(t-s)^2$$

for $|t-s|<1/2$ and for $n \geqslant n_0$. From Thm. 12.3 on p. 95 of Bilingsley (1968), it follows that the distributions of $\{X_n(t), t \in M\}$ are tight.

This completes the proof of Theorem 1. ▮

Since the matrix A appearing in (2.1) is a circulant, its eigen values are $(2/\varepsilon)+q\varepsilon-(2/\varepsilon)\cos(2\pi r/n)$, $r=0, 1, ..., n-1$. The following corollary is an immediate consequence from this observation and (2.6).

COROLLARY 1.

$$Z_n(q,\varepsilon) = (2\pi)^{n/2}\left[\prod_0^{n-1}\left(\frac{2}{\varepsilon}+q\varepsilon-\frac{2}{\varepsilon}\cos\left(\frac{2\pi r}{n}\right)\right)\right]^{-1/2}.$$

One can use the following representation to define the process $\{X(t), t \in M\}$ in terms of a Wiener process $\{W(t), t \in M\}$ which we state below as a theorem.

THEOREM 2. *The process* $\{X(t), t \in M\}$ *can be expressed as*

$$X(t) = \frac{e^{\sqrt{q}/2}}{2\sinh(\sqrt{q}/2)} \int_0^1 e^{-\sqrt{q}[(t-s) \bmod 1]}\, dW(s), \tag{2.10}$$

where $\{W(t), t \in M\}$ *is standard Wiener process in* $C(M)$.

Proof. Consider the stationary Gaussian process $\{Z_\alpha(t), t \in M\}$ defined by

$$Z_\alpha(t) = \int_0^1 \exp(-\alpha[(t-s) \bmod 1])\, dW(s)$$

and denote its convariance function by $S_\alpha(r, s)$. By a direct evaluation we see that

$$\begin{aligned}
S_\alpha(0, t) &= \int_0^1 \exp(-\alpha[1 - s + (t - s) \bmod 1])\, ds \\
&= \int_0^t \exp(-\alpha[1 + t] + 2\alpha s)\, ds + \int_t^1 \exp(-\alpha[2 + t] + 2\alpha s)\, ds \\
&= \frac{1}{2\alpha} \{e^{-\alpha(1+t)}(e^{2\alpha t} - 1) + e^{-\alpha(2+t)}(e^{2\alpha} - e^{2\alpha t})\} \\
&= \frac{e^{-\alpha}}{2\alpha} \{e^{\alpha t} - e^{-\alpha t} + e^{\alpha - \alpha t} - e^{-\alpha + \alpha t}\} \\
&= \frac{e^{-\alpha}}{2\alpha} \{e^{\alpha t} + e^{\alpha - \alpha t}\}\{1 - e^{-\alpha}\} \\
&= \frac{e^{-\alpha}}{2\alpha} \{e^{\alpha t - \alpha/2} + e^{\alpha t + \alpha/2}\}\{e^{\alpha/2} - e^{-\alpha/2}\} \\
&= \frac{2e^{-\alpha}}{\alpha} \sinh\left(\frac{\alpha}{2}\right) \cosh\left(\left(t - \frac{1}{2}\right)\alpha\right).
\end{aligned}$$

Since

$$X(t) = \frac{e^{\sqrt{q}/2}}{2 \sinh(\sqrt{q}/2)} Z_{\sqrt{q}}(t),$$

we obtain

$$\begin{aligned}
\operatorname{Cov}(X(0), X(t)) &= \frac{e^q}{4(\sinh(\sqrt{q}/2)^2} S(0, t) \\
&= \frac{e^q}{4(\sinh(\sqrt{q}/2)^2} \frac{2e^{-\sqrt{q}}}{\sqrt{q}} \cosh\left(\left(t - \frac{1}{2}\right)\sqrt{q}\right) \sinh\left(\frac{\sqrt{q}}{2}\right) \\
&= \frac{1}{2\sqrt{q}} \frac{\cosh((t - 1/2)\sqrt{q})}{\sinh(\sqrt{q}/2)} \\
&= R(0, t)
\end{aligned}$$

This completes the proof of Theorem 2. ∎

3. Convergence of Some Non-Gaussian Random Fields

We will now consider a more general joint distribution for $\mathbf{Y}_n = (Y_0, Y_1, ..., Y_{n-1})$ than the simple Gaussian distribution used in Section 2. Let the joint probability distribution of $\mathbf{Y}_n$ in $\mathscr{R}^n$ be given by its pdf (w.r.t. Lebesgue measure) which is proportional to

$$\prod_{i=0}^{n-1} A\left(\frac{(y_{i+1} - y_i)}{\varepsilon}\right) \prod_{i=0}^{n-1} Q(y_i),$$

where

$$A(x) = \exp(-\tfrac{1}{2}a(x)), \qquad a(x) = x^2, \tag{3.1}$$

$$Q(x) = \exp(-\tfrac{1}{2}[qx^2 + b(x)]), \qquad \text{and} \tag{3.2}$$

$$b(x) = x^2 c(x), \quad c(x) \to 0 \qquad \text{as} \quad x \to 0,$$

and

$$-\delta \leqslant c(x) \leqslant x^2 \qquad \text{for some} \quad \delta \in (0, q). \tag{3.3}$$

As before we will consider the normalized configuration $\mathbf{X}_n = (Y_0/\sqrt{\varepsilon}, ..., Y_{n-1}/\sqrt{\varepsilon})$. Then the joint probability distribution Q_n of $\mathbf{X}_n$ in $\mathscr{R}_n$ has pdf $q_n(\mathbf{x})$ w.r.t. Lebesgue measure) given by

$$q_n(\mathbf{x}) = \frac{1}{Z_n(A, Q, \varepsilon)} \prod_{i=0}^{n-1} A\left(\frac{(x_{i+1} - x_i)}{\sqrt{\varepsilon}}\right) \prod_{i=0}^{n-1} Q(\sqrt{\varepsilon}\, x_i). \tag{3.4}$$

The following theorem, which generalizes Theorem 1, gives the weak convergence of the distributions of the process $\{g_n(\mathbf{X}_n, t), t \in M\}$, where g_n is as defined in (2.3).

THEOREM 3. *Let* $\mathbf{X}_n$ *have distribution given by* (3.4) *and let the functions* $A(x)$, $Q(x)$ *and* $b(x)$ *satisfy conditions* (3.1), (3.2), *and* (3.3). *Then, as* $n \to \infty$, $\varepsilon \to 0$ *such that* $n\varepsilon \to 1$, *the distribution of* $g(\mathbf{X}_n, \cdot)$ *converges weakly to the distribution* $\mathscr{P}$ *of* $X(\cdot)$ *given in Theorem* 1.

Proof. Let P_n be the probability measure defined in (2.1). The proof of this theorem will compare the measures Q_n and P_n and use the arguments of contiguity due to Le Cam (1960).

From Theorem 1, the distribution of X_0 under P_n is normal with mean 0 and variance $C_{00} = (1 + \rho^n)/\alpha(1 - \rho^2)(1 - \rho^n)$, and this variance converges to

$$\frac{\cosh(\sqrt{q}/2)}{2\sqrt{q}\sinh(\sqrt{q}/2)}.$$

Since $x_0^2 \, |c(\sqrt{\varepsilon}\, x_0)| \leqslant \delta x_0^2 + \varepsilon x_0^4$ from (3.3), it follows that $X_0^2 \, |c(\sqrt{\varepsilon}\, X_0)|$ is uniformly integrable under P_n. Once again, since $x_0^2 \, |c(\sqrt{\varepsilon}\, x_0)| \to 0$ as $n \to \infty$ for each x_0 from (3.3), it follows that

$$E_{P_n}(X_0^2 \, |c(\sqrt{\varepsilon}\, X_0)|) \to 0 \qquad \text{as} \quad n \to \infty. \tag{3.5}$$

Define

$$D_n = \frac{1}{2} \sum_{i=0}^{n-1} b(\sqrt{\varepsilon}\, x_i) = \frac{\varepsilon}{2} \sum_{i=0}^{n-1} X_i^2 c(\sqrt{\varepsilon}\, X_i). \tag{3.6}$$

Note that

$$|D_n| \leqslant \frac{\varepsilon}{2} \sum_0^{n-1} X_i^2 \, |c(\sqrt{\varepsilon}\, X_i)| \tag{3.7}$$

and

$$\exp(-D_n) \leqslant \exp\left(\frac{\delta\varepsilon}{2} \sum_0^{n-1} X_i^2\right). \tag{3.8}$$

From (3.7) and (3.5), we obtain

$$E_{P_n} |D_n| \leqslant \frac{\varepsilon}{2} \sum_0^{n-1} E_{P_n}(X_i^2 \, |c(\sqrt{\varepsilon}\, X_i)|) = \frac{n\varepsilon}{2} E_{P_n}(X_0^2 \, |c(\sqrt{\varepsilon}\, X_0)|) \to 0.$$

Thus $D_n \to 0$ in probability under P_n as $n \to \infty$.

Let $\theta > 1$ be such that $\delta\theta < q$. Such a θ exists in view of (3.3). By examining the identity for the partition function in (2.2) and using the bound in (3.8), we obtain

$$E_{P_n}(\exp(-\theta D_n)) \leqslant E_{P_n}\left(\exp\left(\frac{\theta\delta\varepsilon}{2} \sum_0^{n-1} X_i^2\right)\right) = \frac{Z_n(q - \theta\delta, \varepsilon)}{Z_n(q, \varepsilon)}.$$

Using the closed form expression for $Z_n(q, \varepsilon)$ in Corollary 1, we obtain

$$\begin{aligned}
0 &\leqslant \log Z_n(q - \theta\delta, \varepsilon) - \log Z_n(q, \varepsilon) \\
&= -\frac{1}{2} \sum_0^{n-1} \log\left(\frac{(2/\varepsilon) + (q - \theta\delta)\varepsilon - (2/\varepsilon)\cos(2\pi r/n)}{(2/\varepsilon) + q\varepsilon - (2/\varepsilon)\cos(2\pi r/n)}\right) \\
&= -\frac{1}{2} \sum_0^{n-1} \log\left(1 - \frac{\theta\delta\varepsilon}{(q\varepsilon + (2/\varepsilon)(1 - \cos(2\pi r/n)))}\right) \\
&\sim -\frac{1}{2} \sum_0^{n-1} \log\left(1 - \frac{\theta\delta}{(q + 2n^2(1 - \cos(2\pi r/n)))}\right) \\
&= -\frac{1}{2} \sum_{|r-0| \leqslant n\alpha} [\ldots] + \sum_{|r-0| > n\alpha} [\ldots] \qquad \text{(say)},
\end{aligned}$$

where the index r is taken mod n and α is a positive number which will be chosen later.

For any α and for $|r-0|>n\alpha$,

$$n^2\left(1-\cos\left(\frac{2\pi r}{n}\right)\right)\geqslant n^2(-\cos(2\pi\alpha))$$

and thus

$$\sum_{|r-0|>n\alpha}[\ldots]\leqslant -\frac{1}{2}n\log\left(1-\frac{\theta\delta}{(q+n^2(1-\cos(2\pi\alpha)))}\right)\leqslant\frac{C_1}{n},$$

where C_1 is a finite constant.

There is a sufficiently small $\alpha>0$ such that, for $|r-0|\leqslant n\alpha$ and large n,

$$n^2\left(1-\cos\left(\frac{2\pi r}{n}\right)\right)\geqslant\pi^2r^2$$

and

$$\sum_{|r-0|\leqslant n\alpha}[\ldots]\leqslant\frac{1}{2}\sum_{|r-0|\leqslant n\alpha}\log\left(1-\frac{\theta\delta}{q+\pi^2r^2}\right)\leqslant C_2\sum_1^{\infty}\frac{1}{r^2}$$

for some finite constant C_2.

This shows that $\exp(-D_n)$ is uniformly integrable under P_n. We had already shown that $D_n\to 0$ in probability. Thus

$$E_{P_n}(\exp(-D_n))\to 1. \tag{3.9}$$

Comparing the distributions in (2.1) and (3.4), we obtain the likelihood ratio

$$L_n(\mathbf{x})\overset{\text{def}}{=}\frac{dQ_n}{dP_n}(\mathbf{x})=\frac{Z_n(q,\varepsilon)}{Z_n(A,Q,\varepsilon)}\exp(-D_n). \tag{3.10}$$

Since $E_{P_n}(L_n)\equiv 1$, we also have

$$E_{P_n}(\exp(-D_n))=\frac{Z_n(A,Q,\varepsilon)}{Z_n(q,\varepsilon)}. \tag{3.11}$$

Comparing (3.9) with (3.11), we obtain

$$\frac{Z_n(A,Q,\varepsilon)}{Z_n(q,\varepsilon)}\to 1.$$

This shows that $L_n\to 1$ in probability under P_n. Hence $\{Q_n\}$ is contiguous with respect to $\{P_n\}$. Thus $Q_n g_n^{-1}$ and $P_n g_n^{-1}$ have the same limiting distributions, namely, $\mathscr{P}$. This completes the proof of Theorem 3. ∎

4. Gaussian Random Fields on d-Dimensional Graphs

We will now extend the results of Sections 2 and 3 to graphs which are d-dimensional lattices. Consider the graph $\mathscr{S} = \mathbf{Z}_n^d$ where $\mathbf{Z}_n = \{0, 1, ..., n-1\}$. A point in $\mathscr{S}$ is of the form $\mathbf{i} = (i_1, i_2, ..., i_d)$ with $i_r \in \mathbf{Z}_n$, $r = 1, ..., d$. Let $\mathbf{0} = (0, ..., 0)$ and let N_0 be a neighborhood of $\mathbf{0}$ which is symmetric, that is, $N_0 = M_0 \cup (-M_0)$, where M_0 is a subset of $\mathscr{S} \backslash \mathbf{0}$. Several straight forward but interesting generalizations of the results of the first two sections of this paper can be made as follows. Let $q > 0$ and $\mathbf{c} = (c_\mathbf{j}, \mathbf{j} \in N_0)$ be constants. These constants can depend on n, and when necessary, we will indicate the dependence by using symbols like $\mathbf{c}(n)$, $c_\mathbf{j}(n)$, etc. The vector $\mathbf{c}$ specifies the strength of the coupling between the site $\mathbf{0}$ and its neighbors in M_0. Consider a probability distribution for random configurations $\mathbf{Y}(n) = (Y_\mathbf{j}(n), \mathbf{j} \in \mathscr{S})$ given by a pdf (with respect to Lebesgue measure) which is proportional to

$$\exp\left(-\frac{1}{2}\left[q \sum_\mathbf{j} y_\mathbf{j}^2 + \frac{1}{\varepsilon^2} \sum_{\mathbf{j} \in \mathscr{S}, \mathbf{j} \in N_0} (y_\mathbf{j} - y_{\mathbf{j}+\mathbf{j}'})^2 c_\mathbf{j}'\right]\right).$$

We will consider the normalized configuration, $X_\mathbf{j}(n) = \varepsilon^p Y_\mathbf{j}(n)$, where the power p will be chosen suitably later in (4.7). The joint distribution of $\mathbf{X}(n)$ is given by the pdf

$$\begin{aligned} p_n(\mathbf{x}) &= Z_n(q, \mathbf{c}, p, \varepsilon) \exp\left(-\frac{1}{2}\left[a(n) \sum_\mathbf{j} x_\mathbf{j}^2 + \sum_{\mathbf{j} \in \mathscr{S}, \mathbf{j}' \in N_0} (x_\mathbf{j} - x_{\mathbf{j}+\mathbf{j}'})^2 b_{\mathbf{j}'}(n)\right]\right) \\ &= Z_n(q, \mathbf{c}, p, \varepsilon) \exp\left(-\frac{1}{2} \mathbf{x}' Q \mathbf{x}\right), \end{aligned}$$

where

$$a(n) = q\varepsilon^{-2p}, \tag{4.1}$$

$$b_\mathbf{j}(n) = c_\mathbf{j} \varepsilon^{-2-2p}, \tag{4.2}$$

$$Q(\mathbf{i}, \mathbf{j}) = Q(0, \mathbf{j} - \mathbf{i}) \qquad \text{and}$$

$$Q(0, \mathbf{j}) = [a(n) + 2B(n)]\, I(\mathbf{0} = \mathbf{j}) - 2b_\mathbf{j} I(\mathbf{j} \in N_0), \tag{4.3}$$

where $I(\cdot)$ is the usual indicator function. For $\mathbf{j}, \mathbf{k} \in \mathscr{S}$ define the inner product $\langle \mathbf{k}, \mathbf{j} \rangle = \sum_{1 \leqslant r \leqslant d} j_r k_r$. The matrix Q is a circulant matrix and its eigen values are given by

$$\begin{aligned} \lambda_\mathbf{k}(n) &= \sum_\mathbf{j} Q(\mathbf{0}, \mathbf{j}) \exp\left(\frac{2\pi i \langle \mathbf{k}, \mathbf{j} \rangle}{n}\right) \\ &= a(n) + 2 \sum_{\mathbf{j} \in N_0} b_\mathbf{j}(n) - \sum_{\mathbf{j} \in N_0} b_\mathbf{j}(n) \exp\left(\frac{2\pi i \langle \mathbf{k}, \mathbf{j} \rangle}{n}\right) \end{aligned}$$

$$= a(n) + 2 \sum_{\mathbf{j} \in N_0} b_{\mathbf{j}}(n) - 2 \sum_{\mathbf{j} \in M_0} b_{\mathbf{j}}(n) \cos\left(\frac{2\pi\langle \mathbf{k}, \mathbf{j}\rangle}{n}\right)$$

$$= a(n) + 2 \sum_{\mathbf{j} \in M_0} b_{\mathbf{j}}(n)\left(1 - \cos\left(\frac{2\pi\langle \mathbf{k}, \mathbf{j}\rangle}{n}\right)\right) \tag{4.4}$$

since $N_0 = M_0 \cup (-M_0)$.

Let $n^* = n^d$. From the Karhunen–Loève representation, we have

$$X_{\mathbf{j}}(n) = \frac{1}{\sqrt{n^*}} \sum_{\mathbf{k} \in \mathscr{S}} \exp\left(\frac{2\pi i\langle \mathbf{k}, \mathbf{j}\rangle}{n}\right) \frac{1}{\sqrt{\lambda_{\mathbf{k}}(n)}} Z_{\mathbf{k}}, \tag{4.5}$$

where $\{Z_{\mathbf{k}}, \mathbf{k}\} \in \mathscr{S}$ are i.i.d. standard normal random variables. Thus the covariance function of $\{X_{\mathbf{j}}(n), \mathbf{j} \in \mathscr{S}\}$ is given by

$$r_{\mathbf{j}, \mathbf{j}'} = \frac{1}{n^*} \sum_{\mathbf{k} \in \mathscr{S}} \exp\left(\frac{2\pi i\langle \mathbf{k}, \mathbf{j} - \mathbf{j}'\rangle}{n}\right) \frac{1}{\lambda_{\mathbf{k}}(n)}.$$

We will use the test function

$$\phi_{\mathbf{k}}(\mathbf{t}) = \exp(-2\pi i\langle \mathbf{k}, \mathbf{t}\rangle), \qquad t \in M, \quad \mathbf{k} \in \mathscr{S}$$

to interpolate $\mathbf{X}(n)$ to obtain a random function $X(n, \mathbf{t})$ on $C(M^d)$ as follows:

$$X(n, \mathbf{t}) = \frac{1}{\sqrt{n^*}} \sum_{\mathbf{k} \in \mathscr{S}} \phi_{\mathbf{k}}(\mathbf{t}) \frac{1}{\sqrt{\lambda_{\mathbf{k}}(n)}} Z_{\mathbf{k}}. \tag{4.6}$$

The random function $X(n, \mathbf{t})$ can also be defined directly in terms of $\mathbf{X}(n)$, by combining (4.5) and (4.6). The covariance function of $\{X(n, \mathbf{t}), \mathbf{t} \in M^d\}$ is given by

$$r_n(\mathbf{t}, \mathbf{t}') = \frac{1}{n^*} \sum_{\mathbf{k} \in \mathscr{S}} \exp(2\pi i\langle \mathbf{k}, \mathbf{t} - \mathbf{t}'\rangle) \frac{1}{\lambda_{\mathbf{k}}(n)}.$$

Thus the finite dimensional distributions of $\{X(n, \mathbf{t}), \mathbf{t} \in M^d\}$ converge if and only if

$$n^* \lambda_{\mathbf{k}} \to \beta_{\mathbf{k}} \qquad \text{where} \quad 0 < \beta \leqslant \infty.$$

The tightness of the distributions of $\{X(n, \mathbf{t})\}$, $\mathbf{t} \in M^d\}$ can be established in the same way as in Section 2. The limiting distribution of $\{X(n, \mathbf{t})\}$, $\mathbf{t} \in M^d\}$ is thus Gaussian and will be in $C(M^d)$ if and only if $\sum_{\mathbf{k}} (1/\beta_{\mathbf{k}}) < \infty$. When this sum is infinite, the limiting distribution will have to be interpreted as a distribution of the space of Schwartz distributions.

As before consider the mixed limits as $n \to \infty$, $\varepsilon \to 0$ so that $n\varepsilon \to 1$. We will therefore let $n = 1/\varepsilon$. The appropriate choices of p will be

$$p = -\frac{d}{2}. \tag{4.7}$$

Suppose that

$$\lim 2n^2 \sum_{\mathbf{j} \in M_0} c_{\mathbf{j}} \left(1 - \cos\left(\frac{2\pi\langle \mathbf{k}, \mathbf{j} \rangle}{n}\right)\right) = \alpha_{\mathbf{k}}. \tag{4.8}$$

By reworking the expression for $\lambda_{\mathbf{k}}(n)$ in (4.4), we find

$$\begin{aligned} n^*\lambda_{\mathbf{k}}(n) &= n^* \left\{a(n) + 2 \sum_{\mathbf{j} \in M_0} b_{\mathbf{j}}(n) \left(1 - \cos\left(\frac{2\pi\langle \mathbf{k}, \mathbf{j} \rangle}{n}\right)\right)\right\} \\ &= \varepsilon^{-2p} n^* \left\{q + \frac{1}{\varepsilon^2 n^2} 2n^2 \sum_{\mathbf{j} \in M_0} c_{\mathbf{j}} \left(1 - \cos\left(\frac{2\pi < \mathbf{k}, \mathbf{j} \rangle}{n}\right)\right)\right\} \\ &\to \beta_{\mathbf{k}} \\ &= q + \alpha_{\mathbf{k}}. \end{aligned}$$

Thus we have a limiting Gaussian distribution for the process $\{X_n(\mathbf{t}), \mathbf{t} \in M^d\}$ if $\sum_{\mathbf{k}} (1/\beta_{\mathbf{k}}) < \infty$, which is equivalent to

$$\sum_{\mathbf{k}} \frac{1}{\alpha_{\mathbf{k}}} < \infty, \tag{4.9}$$

where $\alpha_{\mathbf{k}}$ are given by (4.8). When (4.9) is not satisfied, we will have to interpret the limiting distribution to be in the space of Schwartz distributions on M^d.

We will now look at several examples and verify conditions (4.8) and (4.9).

EXAMPLE 1. Suppose that the constant $c_{\mathbf{j}}(n) \equiv c_{\mathbf{j}}$ do not depend on n. In this case, (4.8) reduces to

$$\lim 2n^2 \sum_{\mathbf{j} \in M_0} c_{\mathbf{j}} \left(1 - \cos\left(\frac{2\pi < \mathbf{k}, \mathbf{j} \rangle}{n}\right)\right) = 4\pi^2 \sum_{\mathbf{j} \in M_0} c_{\mathbf{j}} \langle \mathbf{k}, \mathbf{j} \rangle^2 = \alpha_{\mathbf{k}}.$$

Suppose further that $d = 1$, $M_0 = \{1\}$ and $c_1 = 1$. This is the example considered in Section 2. Then $\alpha_k = 4\pi^2 k^2$ and (4.9) is satisfied. In this case we get a genuine Gaussian process on $C(M)$ and re-establish another version of Theorem 1.

EXAMPLE 2. Let $d=1$, $M_0=\{1,2\}$, $c_1=4n^2$, $c_2=-n^2$. Then (4.8) reduces to

$$\begin{aligned}
&2n^2\left[4n^2\left(1-\cos\left(\frac{2\pi k}{n}\right)\right)-n^2\left(1-\cos\left(\frac{4\pi k}{n}\right)\right)\right]\\
&\quad=2n^2\left[4n^2\left(1-\cos\left(\frac{2\pi k}{n}\right)\right)-2n^2\left(1-\cos\left(\frac{2\pi k}{n}\right)^2\right)\right]\\
&\quad=4n^4\left(1-\cos\left(\frac{2\pi k}{n}\right)\right)^2\\
&\quad\to 16\pi^4k^4.
\end{aligned}$$

In this case $\alpha_k=16\pi^4k^4$ and (4.9) is satisfied. Once again, in this case we get a genuine Gaussian process on $C(M)$.

EXAMPLE 3. Let $d=2$, $M_{0,0}=\{(0,1),(1,0)\}$, $c_{0,1}=c_{1,0}=1$. Then (4.8) reduces to

$$2n^2\left[\left(1-\cos\left(\frac{2\pi k_1}{n}\right)+\left(1-\cos\left(\frac{2\pi k_2}{n}\right)\right]\to 4\pi^2(k_1^2+k_2^2).\right.\right.$$

Thus $\alpha_k=4\pi^2(k_1^2+k_2^2)$, $\sum(1/\alpha_k)=\infty$ and (4.9) is not satisfied. In this case we do not have a limiting Gaussian process in $C(M^2)$ but on the space of Schwartz distributions on M^2.

EXAMPLE 4. Let $d=2$, $M_{0,0}=\{(0,1), (1,0), (1,1), (1,-1)\}$, $c_{0,1}=c_{1,0}=c_{1,1}=c_{1,-1}=1$. Then (4.8) reduces to

$$\begin{aligned}
&2n^2\left[\left(1-\cos\left(\frac{2\pi k_1}{n}\right)\right)+\left(1-\cos\left(\frac{2\pi k_2}{n}\right)\right)\right.\\
&\qquad\left.+\left(1-\cos\left(\frac{2\pi(k_1+k_2)}{n}\right)\right)+\left(1-\cos\left(\frac{2\pi(k_1-k_2)}{n}\right)\right)\right]\\
&\quad\to 4\pi^2(k_1^2+k_2^2+(k_1+k_2)^2+(k_1-k_2)^2)\\
&\quad=12\pi^2(k_1^2+k_2^2).
\end{aligned}$$

Thus $\alpha_k=12\pi^2(k_1^2+k_2^2)$, $\sum(1/\alpha_k)=\infty$ and (4.9) is satisfied. In this case also we do not have a limiting Gaussian process in $C(M^2)$ but on the space of Schwartz distributions on M^2.

EXAMPLE 5. Let $d=2$, $M_{0,0}=\{(0,1),(1,0),(1,1),(1,-1)\}$, $c_{0,1}=c_{1,0}=n^2$, $c_{1,1}=c_{1,-1}=-n^2/2$. Then (4.8) reduces to

$$
\begin{aligned}
2n^2 \Bigg[n^2 \left(1-\cos\left(\frac{2\pi k_1}{n}\right)\right) &+ n^2\left(1-\cos\left(\frac{2\pi k_2}{n}\right)\right) \\
&-\frac{n^2}{2}\left(1-\cos\left(\frac{2\pi(k_1+k_2)}{n}\right)\right)+\frac{n^2}{2}\left(1-\cos\left(\frac{2\pi(k_1-k_2)}{n}\right)\right)\Bigg] \\
&=2n^4\left[\left(1-\cos\left(\frac{2\pi k_1}{n}\right)\right)\left(1-\cos\left(\frac{2\pi k_2}{n}\right)\right)\right] \\
&\to 8\pi^4 k_1^2 k_2^2.
\end{aligned}
$$

Thus $\alpha_k=8\pi^4k_1^2k_2^2$ and (4.9) is satisfied. Thus in this case, we do have a limiting Gaussian process in $C(M^2)$.

5. REMARKS

One can use the ideas of contiguity and uniform integrability, as done Section 3, to study more general Markov random fields and obtain similar limiting Gaussian processes.

Kurien and Sethuraman (1993b) examined other examples of Gaussian random fields like the above with different couplings, and have shown that, under mixed limits, the limiting Gaussian process can exhibit a phase transition based upon a constant of proportionality, namely, *temperature*, that multiplies the parameters q and the coupling constants $\mathbf{c}$.

In another direction, Chow (1990) and Kurien and Sethuraman (1993a) have extended our model in Theorem 1 to more general Markov random fields and obtained the limit distributions. In Kurien and Sethuraman (1993a) this limiting process is a stable process.

REFERENCES

1. BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.
2. CHOW, Y., AND GRENANDER, U. (1985). A singular perturbation problem. *J. Integral Equations* **9** 63–73.
3. CHOW, Y. (1990). A limit theorem for pattern synthesis in image processing *J. Multivariate Anal.* **34** 75–94.
4. DAVIS, P. J. (1979). *Circulant Matrices.* Wiley, New York.
5. GEMAN, S., AND GEMAN, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.* **6** 721–741.
6. GRENANDER, U. (1976). Pattern synthesis. *Lectures in Pattern Theory*, Vol. I. Springer-Verlag, Berlin/New York.

7. GRENANDER, U. (1981). Regular structures. *Lectures in Pattern Theory*, Vol. III. Springer-verlag, Berlin/New York.
8. KURIEN, T. V., AND SETHURAMAN, J. (1993a). A mixed limit theorem for stable random fields. *J. Multivariate Anal.* **47** 152–162.
9. KURIEN, T. V., AND SETHURAMAN, J. (1993b). Singularities in Gaussian random fields. *J. Theor. Probab. Appl.* **6** 89–99.
10. LE CAM, L. (1960). Locally asymptotically normal families of distributions, *Univ. of California Publ. Statist.* **3** 37–98.
11. METROPOLIS, N., ROSENBLUTH, M. N., ROSENBLUTH, A. W., TELLER, A. H., AND TELLER, E. (1953). Equations of state calculations by fast computing machines, *J. Chem. Phys.* **21** 1087–1091.