

A UNIFIED APPROACH TO MONTE CARLO LIKELIHOOD ESTIMATION OF GIBBS RANDOM FIELD IMAGES

Gerasimos Potamianos and John Goutsias

Department of Electrical and Computer Engineering
Image Analysis and Communications Laboratory
The Johns Hopkins University
Baltimore, MD 21218, USA

Abstract

We present a unified analysis of Monte Carlo algorithms for estimating the likelihood function of a fully, or partially observed Gibbs random field image. We show that such an estimation reduces to estimating the partition functions of suitably chosen Gibbs random fields. We then proceed to study the computational complexity, with respect to image size, of Monte Carlo partition function estimation algorithms. We show that such algorithms can be classified into two categories: Algorithms [A], which are characterized by an exponential complexity, and Algorithms [B], which are characterized by a polynomial complexity, Turing reducible to the problem of sampling from the Gibbs probability measure. We compare these algorithms, by using both theoretical results and simulation experiments, and we determine the most efficient Algorithm [A] and the most efficient Algorithm [B]. We propose a generalized version of the most efficient Algorithm [B], suitable for statistical inference of a fully, as well as partially, observed Gibbs random field image.

1. Introduction

Gibbs random fields (GRFs) constitute a popular class of statistical models for images [1]. However, a number of theoretical and computational problems are associated with these models, the most prominent one being the exact calculation of the associated *partition function*. Given a *full* (i.e., noiseless), or *partial* (i.e., degraded by noise) observation of a GRF image, it is not possible to compute its likelihood function without knowledge of certain partition functions [2]. This clearly imposes restrictions to maximum likelihood estimation and hypothesis testing, needed in applications such as unsupervised image segmentation, model order determination, and texture classification [1].

Since exact computation of the partition and likelihood functions is impossible in general, approximate methods have to be used in order to obtain estimates of these quantities. The most widely applicable and reliable ones are stochastic simulations using Monte Carlo techniques. A respectable number of such Monte Carlo partition function estimation methods have been recently proposed [3]-[7]. However, their properties and relative merits need investigation. A first attempt towards a unified analysis of some of these methods appears in [2], where we have studied Monte Carlo algorithms which draw samples from a single probability mass function. We refer to such algorithms, as *Algorithms [A]*. We review these methods in Section 3, following Section 2, where we introduce all necessary notation and background.

This paper is a natural continuation of our previous work in [2]. In Section 4, we study the Monte Carlo partition function estimation algorithms proposed in [4]-[7]. These

algorithms draw samples from multiple Gibbs probability measures. In Section 5, we use both our theoretical analysis and a number of simulation experiments to determine the most efficient Algorithm [A], and the most efficient Algorithm [B]. By combining the advantages of both algorithms, we develop a new efficient scheme for partition function estimation, in Section 6. This scheme seems very appropriate for optimal statistical inference for GRF images. Finally, in Section 7, we summarize our results and provide some concluding remarks.

2. Partition Function and Likelihood

Let us consider a collection of $M \times N$ sites, associated with the rectangular lattice (image) $\Lambda = \{(i, j) : 1 \leq i \leq M, 1 \leq j \leq N\}$. A discrete-valued random variable H_{ij} is assigned at each site $(i, j) \in \Lambda$, taking values h_{ij} from a finite state-space E , which contains $R \geq 2$ distinct values. The resulting random field $\mathbf{H} = \{H_{ij} : 1 \leq i \leq M, 1 \leq j \leq N\}$ can take any one of the R^{MN} possible realizations $\mathbf{h} = \{h_{ij} : 1 \leq i \leq M, 1 \leq j \leq N\} \in E^{MN}$, with probability mass function $Pr[\mathbf{H} = \mathbf{h}]$. We restrict \mathbf{H} to be a GRF, whose probability mass function is given by the Gibbs probability measure [1], [3],

$$\pi(\mathbf{h}) = \frac{1}{Z} A(\mathbf{h}) > 0, \quad (1a)$$

for all $\mathbf{h} \in E^{MN}$, where

$$A(\mathbf{h}) = \exp\left\{-\frac{1}{T} U(\mathbf{h})\right\}, \quad \text{and} \quad Z = \sum_{\text{states } \mathbf{h}} A(\mathbf{h}). \quad (1b)$$

In (1), Z is a normalizing constant known as the *partition function*, T is a positive parameter known as the *temperature*, $U(\bullet)$ is the *energy function*, that incorporates the local interactions among neighboring sites, whereas the summation in (1b) is carried over all R^{MN} states.¹ Calculating Z is, therefore, prohibitive, even for moderate lattice sizes.

Without loss of generality, we assume second-order GRFs [3]. In this case,

$$A(\mathbf{h}) = e^{-\frac{1}{T} U(\mathbf{h})} = \prod_{i=1}^M \prod_{j=1}^N \sigma_{ij}(h_{ij}, h_{i-1,j}, h_{i,j-1}, h_{i,j-1}), \quad (2a)$$

and (see also (1))

$$Z = \sum_{\text{states } \mathbf{h}} \prod_{i=1}^M \prod_{j=1}^N \sigma_{ij}(h_{ij}, h_{i-1,j}, h_{i,j-1}, h_{i,j-1}) = \sum_{\text{states } \mathbf{h}} A(\mathbf{h}), \quad (2b)$$

where $\sigma_{ij}(x, y, z, \omega)$, $(x, y, z, \omega) \in E^4$, $(i, j) \in \Lambda$, is the *local transfer function* (LTF) of the GRF \mathbf{H} . The LTF depends on the temperature T , and is positive and finite for $0 < T \leq +\infty$. We assume a bounded LTF, i.e.,

$$\sigma_{\inf} = \inf_{M, N \rightarrow +\infty} \sigma_{\min}, \quad \text{and} \quad \sigma_{\sup} = \sup_{M, N \rightarrow +\infty} \sigma_{\max}, \quad (3a)$$

¹ Generally speaking, most of the quantities used in this paper depend on the sizes M and N of lattice Λ , as well as on temperature T . This dependence is often suppressed in order to simplify notation.

exist, are positive, and finite, for all $0 < T \leq +\infty$, where

$$\sigma_{\min} = \min \{ \sigma_{ij}(x, y, z, \omega) : x, y, z, \omega \in E ; (i, j) \in \Lambda \}, \quad (3b)$$

and

$$\sigma_{\max} = \max \{ \sigma_{ij}(x, y, z, \omega) : x, y, z, \omega \in E ; (i, j) \in \Lambda \}. \quad (3c)$$

In practice, we usually assume homogeneous LTFs, i.e., $\sigma_{ij}(x, y, z, \omega) = \sigma(x, y, z, \omega)$, for all $(i, j) \in \Lambda$, with necessary modifications at the *boundary* sites [1], [3].

A *fully observed* GRF image is a realization \mathbf{h}_o of a GRF \mathbf{H} , drawn with $Pr[\mathbf{H} = \mathbf{h}_o] = \pi(\mathbf{h}_o)$. Its likelihood function is given by (see (1) and (2))

$$L(\mathbf{h}_o) = \frac{1}{Z} \prod_{(x,y,z,\omega) \in E^4} \prod_{i=1}^M \prod_{j=1}^N \sigma_{ij}(x, y, z, \omega)^{v_{\mathbf{h}_o, ij}(x, y, z, \omega)}, \quad (4)$$

where $v_{\mathbf{h}_o, ij}(x, y, z, \omega)$, $(x, y, z, \omega) \in E^4$, $(i, j) \in \Lambda$, are the *sufficient statistics* of the Gibbs probability measure (2) [8]. Each such statistic can be calculated from the observation \mathbf{h}_o , and equals to one, whenever $(h_{ij}, h_{i-1, j}, h_{i-1, j-1}, h_{i, j-1}) = (x, y, z, \omega)$, or to zero, otherwise. Equation (4) simplifies in the case of a homogeneous LTF into

$$L(\mathbf{h}_o) = \frac{1}{Z} \prod_{(x,y,z,\omega) \in E^4} \sigma(x, y, z, \omega)^{v_{\mathbf{h}_o}(x, y, z, \omega)}, \quad (5a)$$

where

$$v_{\mathbf{h}_o}(x, y, z, \omega) = \sum_{i=1}^M \sum_{j=1}^N v_{\mathbf{h}_o, ij}(x, y, z, \omega), \quad (5b)$$

denotes the number of elementary squares of the form $(h_{ij}, h_{i-1, j}, h_{i-1, j-1}, h_{i, j-1}) = (x, y, z, \omega)$ in \mathbf{h}_o .

Assume now that the GRF \mathbf{H} is not observed, but is transformed into an observable random field $\mathbf{Y} = \{y_{ij} : 1 \leq i \leq M, 1 \leq j \leq N; y_{ij} \in E'\}$, defined on the same lattice Λ , where E' is a finite set containing $R' \geq 2$ distinct values. The transformation is described by means of $Pr[\mathbf{Y} = \mathbf{y} | \mathbf{H} = \mathbf{h}] > 0$, for all $\mathbf{y} \in E'^{MN}$, $\mathbf{h} \in E^{MN}$. A *partially observed* GRF image is then a realization \mathbf{y}_o of \mathbf{Y} , with likelihood function

$$\bar{L}(\mathbf{y}_o) = \frac{1}{\sum_{\text{states } \mathbf{h}} A(\mathbf{h})} \sum_{\text{states } \mathbf{h}} A(\mathbf{h}) Pr[\mathbf{Y} = \mathbf{y}_o | \mathbf{H} = \mathbf{h}] = \frac{\bar{Z}(\mathbf{y}_o)}{Z}. \quad (6)$$

Under natural assumptions on the degradation process [9], the pair (\mathbf{H}, \mathbf{Y}) is a GRF. Here, we assume that

$$Pr[\mathbf{Y} = \mathbf{y} | \mathbf{H} = \mathbf{h}] = \prod_{i=1}^M \prod_{j=1}^N Pr(y_{ij} | y_{i-1, j}, y_{i-1, j-1}, y_{i, j-1}, h_{ij}, h_{i-1, j}, h_{i-1, j-1}, h_{i, j-1}), \quad (7)$$

in which case, $\bar{Z}(\mathbf{y}_o)$ is the partition function of the second-order GRF $\{\mathbf{H} | \mathbf{Y}\}$, with a non-homogeneous LTF given by (see (2a), (6), and (7))

$$\begin{aligned} \bar{\sigma}_{ij}(x, y, z, \omega) \\ = \sigma_{ij}(x, y, z, \omega) Pr(y_{ij} | y_{i-1, j}, y_{i-1, j-1}, y_{i, j-1}, h_{ij}, h_{i-1, j}, h_{i-1, j-1}, h_{i, j-1}). \end{aligned}$$

The new LTF is clearly bounded, assuming that the conditional probabilities in (7) are bounded, and, therefore, (3) holds. Clearly, both partition functions Z and $\bar{Z}(\mathbf{y}_o)$ are of the form (2b). Therefore, it suffices to analyze Monte Carlo partition function algorithms for sums of the form (2b).

Monte Carlo algorithms are stochastic in nature [3]-[7]; therefore, and after a finite number of iterations, they can only produce a partition function estimate \hat{Z} , within an accuracy ε from Z , and confidence $1 - \xi$ [6], i.e.,

$$Pr[-\varepsilon/MN \leq f(\hat{Z}) - f(Z) \leq \varepsilon/MN] \geq 1 - \xi, \quad (8)$$

where $f(Z) = 1/MN \times \ln Z$, and $0 < \varepsilon, \xi < 1$. We consider $f(Z)$ here, rather than Z itself, because $\ln \sigma_{\inf} \leq f(Z) - \ln R \leq \ln \sigma_{\sup}$ (see also (2), (3)), whereas Z need not be bounded as $M, N \rightarrow +\infty$. In addition, and for a homogeneous LTF, $\lim_{M, N \rightarrow +\infty} f(Z)$ exists, is finite at all positive temperatures, and is an

analytic function of the LTF, except at *critical temperatures* [10].

As a final remark, the study of partition function estimation algorithms, requires the study of Monte Carlo likelihood function estimation algorithms. Indeed (see (4), (6)), the likelihood function of a fully, or partially, observed GRF image is a ratio of two partition functions. In (4), the numerator partition function is trivially computed from the observed data. Clearly, estimation of the two partition functions such that (8) holds, will suffice for

$$Pr[-2\varepsilon/MN \leq f(\hat{L}) - f(L) \leq 2\varepsilon/MN] \geq 1 - 2\xi$$

to be satisfied (similarly for L). Therefore, we concentrate here on the study of Monte Carlo GRF partition function estimation algorithms.

3. Algorithms [A] for GRF Partition Function Estimation

Algorithms [A] are based on drawing samples from a single probability mass function. They have been studied in [2], and for the case of a homogeneous LTF. Under this assumption, their computational complexity, such that (8) holds, is exponential with respect to image size.

A-1. Markov Mesh Based Sampling Algorithms [3].

Let us consider a special case of a second order GRF, namely a second order *Markov mesh* [11]. Its probability mass function is given by

$$P(\mathbf{h}) = \prod_{i=1}^M \prod_{j=1}^N \tau_{ij}(h_{ij}, h_{i-1, j}, h_{i-1, j-1}, h_{i, j-1}), \quad \text{for all } \mathbf{h} \in E^{MN}, \quad (9a)$$

where the LTF $\tau_{ij}(\bullet, \bullet, \bullet, \bullet)$ is positive, and satisfies

$$\sum_{u \in E} \tau_{ij}(u, y, z, \omega) = 1, \quad \text{for every } (y, z, \omega) \in E^3, \quad (9b)$$

and $(i, j) \in \Lambda$. The partition function in this case equals to one. Furthermore, samples of this random field can be drawn exactly in $O(RMN)$ time [11].

We now write

$$Z = \sum_{\text{states } \mathbf{h}} \left[\frac{A(\mathbf{h})}{P(\mathbf{h})} \right] P(\mathbf{h}) = \sum_{\text{states } \mathbf{h}} Q(\mathbf{h}) P(\mathbf{h}) = \mathbf{E}_P[Q(\mathbf{H})], \quad (10a)$$

where $\mathbf{E}_P[\bullet]$ denotes expectation, and (see (2a) and (9))

$$Q(\mathbf{h}) = \frac{A(\mathbf{h})}{P(\mathbf{h})} = \prod_{i=1}^M \prod_{j=1}^N \frac{\sigma_{ij}(h_{ij}, h_{i-1, j}, h_{i-1, j-1}, h_{i, j-1})}{\tau_{ij}(h_{ij}, h_{i-1, j}, h_{i-1, j-1}, h_{i, j-1})}, \quad (10b)$$

for all states $\mathbf{h} \in E^{MN}$. In this case,

$$Z_P(K) = \frac{1}{K} \sum_{k=1}^K Q(\mathbf{H}_k), \quad (11)$$

is an *unbiased* and *consistent Monte Carlo estimator* of Z [3], [8]. In (11), $\{\mathbf{H}_k, k = 1, 2, \dots, K\}$ is a collection of i.i.d. Markov meshes, statistically equivalent to \mathbf{H} .

The main focus of our work in [2], [3] is to choose the appropriate probability mass function $P(\mathbf{h})$, that satisfies (9), and results in small estimation variance; i.e., minimizing

$$D_{IS}(\pi, P) = \frac{1}{MN} \ln \sum_{\text{states } \mathbf{h}} \left[\frac{\pi(\mathbf{h})}{P(\mathbf{h})} \right]^2 P(\mathbf{h}) = \frac{1}{MN} \ln \left[1 + \frac{\mathbf{Var}_P[Q^2(\mathbf{H})]}{Z^2} \right], \quad (12)$$

with respect to P . This quantity characterizes the convergence properties of estimator (11) (see Theorem 1). The problem under consideration is, therefore, equivalent to approximating a general GRF by a Markov mesh [11].

Given the LTF $\sigma_{ij}(\bullet, \bullet, \bullet, \bullet)$ of the original GRF, we have considered in [2] and [3] three such approximating Markov meshes, namely $P^{iid}(\mathbf{h})$, $P^*(\mathbf{h})$, and $P^{**}(\mathbf{h})$, with LTFs

$$\tau_{ij}^{iid}(x, y, z, \omega) = 1/R, \quad (13a)$$

$$\tau_{ij}^*(x, y, z, \omega) = \frac{\left[\frac{\partial \ln Z}{\partial \sigma_{ij}(x, y, z, \omega)} \right] \sigma_{ij}(x, y, z, \omega)}{\sum_{u \in E} \left[\frac{\partial \ln Z}{\partial \sigma_{ij}(u, y, z, \omega)} \right] \sigma_{ij}(u, y, z, \omega)}, \quad (13b)$$

and

$$\tau_{ij}^{**}(x, y, z, \omega) = \frac{\sigma_{ij}(x, y, z, \omega)}{\sum_{u \in E} \sigma_{ij}(u, y, z, \omega)}, \quad (13c)$$

for all $(x, y, z, \omega) \in E^4$, and all $(i, j) \in \Lambda$, respectively. We now have the following theorem:

Theorem 1: For any Monte Carlo estimator $Z_P(K)$ of the partition function Z , given by (11), there exists an integer

$$K_{\min} = \left\lceil \frac{1}{\xi \varepsilon^2} \frac{\mathbf{Var}_P[Q(\mathbf{H})]}{Z^2} \right\rceil,$$

where $\lceil x \rceil$ denotes the smallest integer $\geq x$, such that²

$$\Pr \left[1 - \varepsilon \leq \frac{Z(K)}{Z} \leq 1 + \varepsilon \right] \geq 1 - \xi, \quad \text{for every } K \geq K_{\min}, \quad (14)$$

with $0 < \varepsilon, \xi < 1$. Furthermore, assuming a homogeneous LTF,

$$0 \leq \lim_{M, N \rightarrow +\infty} \frac{1}{MN} \ln K_{\min} = \lim_{M, N \rightarrow +\infty} D_{IS}(\pi, P) = D_P(T) < +\infty, \quad (15)$$

is satisfied at any temperature $0 < T \leq +\infty$; i.e., the method has exponential computational complexity with respect to image size. In particular,

$$D_{P^{iid}}(T) = \lim_{M, N \rightarrow +\infty} D_{IS}(\pi, P^{iid}) = \ln R + \lim_{M, N \rightarrow +\infty} \frac{1}{MN} \ln \frac{Z_{T/2}}{Z_T^2}, \quad (16)$$

where Z_T denotes the partition function of a GRF at temperature T , $0 \leq D_{P^{iid}}(T) \leq \ln R$, $D_{P^{iid}}(+\infty) = 0$, and $D_{P^{iid}}(0) = \ln R$, provided that $\ln(\eta_{\min}) = o(MN)$, as $M, N \rightarrow +\infty$, where η_{\min} is the number of minimum energy states in the GRF model.

Limits (15) and (16) do not necessarily exist when non-homogeneous LTFs are employed. Theorem 1 shows that in the case of a homogeneous LTF, the estimation of Z , through (10), (11), and by using samples drawn from $P^{iid}(\mathbf{h})$, fails at “low” temperatures, becoming computationally equivalent to a brute-force calculation of (2b), as the temperature approaches zero. In addition, and from a variety of simulation experiments, we have determined that

$$0 \approx D_{P^*}(T) \leq D_{P^{**}}(T) \ll D_{P^{iid}}(T), \quad (17)$$

i.e., choice (13b) is typically the best in practice, especially at low temperatures (i.e., GRFs with strong interactions). Extensive simulation experiments indicate that this is also the case, when a non-homogeneous LTF is of interest.

A-2. The First Ogata-Tanemura Method [4].

This algorithm is based on drawing samples directly from the Gibbs probability measure under consideration, which are then used to form a certain Gibbs average. This is a non-trivial procedure, since, in general, such samples can only be approximately obtained [2], [13]. Here, we choose to discuss computational complexity issues, related to this algorithm, *Turing reducible* to the problem of obtaining Gibbs samples. Therefore, we consider having independent such samples, readily available to us at a unit cost.

The First Ogata-Tanemura estimation method is based on the identity (see also (1))

$$\frac{R^{MN}}{Z} = \sum_{\text{states } \mathbf{h}} A^{-1}(\mathbf{h}) \pi(\mathbf{h}) = \mathbf{E}_{\pi}[A^{-1}(\mathbf{H})],$$

² Conditions (8) and (14) are equivalent [12].

provided that $A(\mathbf{h}) > 0$, for all states \mathbf{h} . Clearly,

$$Z_{OT_1}(K) = \frac{R^{MN}}{\frac{1}{K} \sum_{k=1}^K A^{-1}(\mathbf{H}_k)} = \frac{R^{MN}}{A_{inv}(K)}, \quad (18)$$

is a Monte Carlo estimator of Z . In (18), $\{\mathbf{H}_k, k = 1, 2, \dots, K\}$ is a collection of i.i.d. GRFs, statistically equivalent to \mathbf{H} . Estimator $A_{inv}(K)$ is *asymptotically normal* (as $K \rightarrow +\infty$) [8], and as a result of the *delta method* [8], estimator (18) is an *asymptotically unbiased* and *consistent* estimator of Z . We now have the following theorem:

Theorem 2: At any temperature T , such that

$$T \leq \frac{1}{(t+3)\ln R} \frac{U_{\max} - U_{\min}}{MN}, \quad t \geq 0, \quad (19a)$$

we have that

$$\frac{\lim_{K \rightarrow +\infty} K \mathbf{Var}[Z_{OT_1}(K)]}{K \mathbf{Var}[Z_{P^{iid}}(K)]} \geq R^{tMN}, \quad (19b)$$

where $U_{\max} = \max\{U(\mathbf{h}), \mathbf{h} \in E^{MN}\}$, and $U_{\min} = \min\{U(\mathbf{h}), \mathbf{h} \in E^{MN}\}$. Furthermore, and under the assumption of a homogeneous LTF, the sufficient number, K_{\min} , of independent Gibbs samples for estimator (18) to satisfy (14), enjoys the limit

$$\lim_{M, N \rightarrow +\infty} \frac{1}{MN} \ln K_{\min} = D_{OT_1}(T), \quad (20a)$$

with

$$0 \leq D_{OT_1}(T) = -2 \ln R + \lim_{M, N \rightarrow +\infty} \frac{1}{MN} \ln [Z(\sigma)Z(\sigma^{-1})]. \quad (20b)$$

In (20), $Z(\sigma)$ denotes the partition function of a GRF with LTF σ , $D_{OT_1}(+\infty) = 0$, and, $D_{OT_1}(0) = +\infty$, provided that $U_{\max} > U_{\min}$, for all sufficiently large M, N .

From (19), we conclude that *Algorithm [A-2]* becomes computationally worse than *Algorithm [A-1]*, when samples from $P^{iid}(\mathbf{h})$ are used, and for the case of GRFs with strong interactions (low T). In the case of a homogeneous LTF, (20) shows that *Algorithm [A-2]* is of an exponential computational complexity (Turing reducible to the problem of sampling from the Gibbs probability measure), becoming computationally worse than the brute force summation of (1b), as the temperature approaches zero. Clearly, this method is inappropriate for partition function calculations.

4. Algorithms [B] for GRF Partition Function Estimation

Algorithms [B] are based on drawing samples from multiple Gibbs probability measures. They have been proposed in [5]-[7]. Here, we present them in a unified manner, and for the case of a general GRF. In order to simplify notation and presentation, we assume that independent samples of the Gibbs probability measure are available to us at a unit cost. In practice, the Gibbs Sampler Algorithm [13] is used to estimate the required Gibbs averages. This case is treated in [12].

B-1. The Second Ogata-Tanemura Method [5].

From (1b), it is clear that the partition function can be trivially calculated at $T = +\infty$, since, in this case³

$$Z_{+\infty} = \lim_{T \rightarrow +\infty} Z_T = R^{MN}. \quad (21)$$

Now, observe that [6]

³ In the following, we need to explicitly show dependence of Z , $A(\mathbf{h})$, and $\pi(\mathbf{h})$, on temperature T . We shall therefore write Z_T , $A_T(\mathbf{h})$, and $\pi_T(\mathbf{h})$ for Z , $A(\mathbf{h})$, and $\pi(\mathbf{h})$, respectively.

$$\frac{Z_{T_1}}{Z_{T_0}} = \exp \left[\int_{T_0}^{T_1} \frac{\partial \ln Z_T}{\partial T} dT \right] = \exp \left[\int_{T_0}^{T_1} \mathbf{E}_{\pi_T} [U(\mathbf{H})] \frac{dT}{T^2} \right]. \quad (22)$$

By letting $T_0 = +\infty$, using (21), and changing variables from T to $\beta = 1/T$ in (22), we obtain

$$f(Z_{T_1}) = \frac{1}{MN} \ln Z_{T_1} = \ln R - \int_0^{1/T_1} I\left(\frac{1}{\beta}\right) d\beta, \quad (23a)$$

where

$$I\left(\frac{1}{\beta}\right) = I(T) = \mathbf{E}_{\pi_T} \left[\frac{U(\mathbf{H})}{MN} \right]. \quad (23b)$$

A Monte Carlo estimator for (23b), can now be obtained at any temperature T , by means of

$$I(T; K) = \frac{1}{MN} \frac{1}{K} \sum_{k=1}^K U(\mathbf{H}_k), \quad (24)$$

where $\{\mathbf{H}_k, k=1, 2, \dots, K\}$ is a collection of i.i.d. GRFs, with $Pr[\mathbf{H}_k = \mathbf{h}] = \pi_T(\mathbf{h})$. In practice, these random fields are generated by the Gibbs Sampler Algorithm, and are, therefore, correlated, and only asymptotically distributed according to the Gibbs probability measure [12], [13].

Let us now consider $\eta_s + 1$ equally spaced points $\beta_{(i)} = 1/T_{(i)}$ in the interval $[0, 1/T_1]$, defined by

$$\beta_{(i)} = \frac{1/T_1}{\eta_s} i \quad \text{or} \quad T_{(i)} = \frac{\eta_s}{i} T_1, \quad \text{for } i = 0, 1, \dots, \eta_s, \quad (25)$$

where, $\eta_s \bmod 2^{s-1} = 0$, and s is a positive integer. A simple and efficient way for calculating the integral in (23a) is by means of a *Romberg numerical integration procedure* of order $2s$, as

$$\int_0^{1/T_1} I\left(\frac{1}{\beta}\right) d\beta \approx RO(s, \eta_s), \quad (26a)$$

where

$$RO(1, \eta_1) = \frac{1}{\eta_1 T_1} \left[\frac{I(+\infty)}{2} + \sum_{i=1}^{\eta_1-1} I(T_{(i)}) + \frac{I(T_1)}{2} \right], \quad (26b)$$

and

$$RO(s+1, \eta_{s+1}) = \frac{1}{2^{2s}-1} \left[2^{2s} RO(s, \eta_{s+1}) - RO(s, \frac{\eta_{s+1}}{2}) \right], \quad (26c)$$

for $s \geq 1$. Equation (26b) is the well known *trapezoidal integration rule*. A better approximation of the integral in (23a) can be achieved by using the *Simpson's integration rule* ($s=2$) [12]. Generally,

$$RO(s, \eta_s) = \frac{1}{\eta_s T_1} \sum_{i=0}^{\eta_s} d_{s, \eta_s, i} I(T_{(i)}), \quad (27)$$

where $d_{s, \eta_s, i}$ are positive constants which satisfy $\sum_{i=0}^{\eta_s} d_{s, \eta_s, i} = \eta_s$, and can be computed by using (26).

A Monte Carlo estimator for $f(Z_{T_1})$, by means of a *Romberg numerical integration procedure* of order $2s$, will now be given by (see (23a), (24), (26a), (27))

$$F_{s, OT_2}(T_1) = \ln R - \frac{1}{\eta_s T_1} \sum_{i=0}^{\eta_s} d_{s, \eta_s, i} I(T_{(i)}; K_{s,i}). \quad (28)$$

This estimator can be shown to be *asymptotically unbiased* and *consistent*, as $\eta_s \rightarrow +\infty$, and $K_{s,i} \rightarrow +\infty$, for all $i = 0, 1, \dots, \eta_s$. We now have the following theorem:

Theorem 3: *To calculate $f(Z_{T_1})$, by means of the Monte Carlo estimation scheme (24), (25), and (28), such that (8) is satisfied, it suffices to:*

1. Use $\eta_s + 1$ equally spaced points $\beta_{(i)}$, given by (25), where

$$\eta_s = \left\lceil \left[f_{2s} / \varepsilon \right] \frac{1}{2s} \left[MN \max \{ |\ln \sigma_{\max}|, |\ln \sigma_{\min}| \} \right]^{1 + \frac{1}{2s}} \right\rceil_{2^{s-1}} \\ = O((MN)^{1 + \frac{1}{2s}}), \quad (29a)$$

with f_{2s} being positive constants.⁴ In (29a), $[x]_l = \min\{y$

⁴ For example, $f_2 = 1$, $f_4 = 5/3$, $f_6 = 1784/45$, and $f_8 = 7.394 \times 10^3$ [12].

positive integer : $y \bmod l = 0$ and $y \geq x$ }, with x nonnegative real, and l integer.

2. Use $K_{s,i}$ i.i.d. samples, drawn from $\pi_{T_{(i)}}(\mathbf{h})$, where

$$K_{s,i} = \left\lceil \frac{4\alpha_{s,i}}{\xi \varepsilon^2} \left[\frac{MN}{T_1} \right]^2 \eta_s \mathbf{Var}_{\pi_{T_{(i)}}} \left[\frac{U(\mathbf{H})}{MN} \right] \right\rceil = O((MN)^{3 + \frac{1}{2s}}), \quad (29b)$$

for every $i = 0, 1, \dots, \eta_s$. In (29a), σ_{\min} and σ_{\max} are given by (3b,c), where the LTF is considered at temperature T_1 , whereas, $a_{s,i} = 1$, for $i = 1, 2, \dots, \eta_s - 1$, $a_{s,0} = a_{s,\eta_s} = 2$.

The computational cost of satisfying (8), by means of (24), (25), and (28), is clearly $\sum_{i=0}^{\eta_s} K_{s,i}$ Monte Carlo iterations, which, and according to Theorem 3, is of the order of $O((MN)^{4+1/s})$ (i.e., polynomial with respect to image size), Turing reducible to the problem of sampling from the Gibbs probability measure. This complexity can become arbitrarily close to $O((MN)^4)$, as $s \rightarrow +\infty$.

B-2. The Jerrum-Sinclair Method [6], [7].

This partition function estimation method has been proposed by Jerrum and Sinclair, for special GRFs [6], and has been extended to the case of general GRFs by Geyer and Thompson [7]. It is based on the identity (see also (1))

$$\frac{Z_{T_1}}{Z_{T_0}} = \sum_{\text{states } \mathbf{h}} \left[\frac{A_{T_1}(\mathbf{h})}{A_{T_0}(\mathbf{h})} \right] \pi_{T_0}(\mathbf{h}) = \mathbf{E}_{\pi_{T_0}} \left[\frac{A_{T_1}(\mathbf{H})}{A_{T_0}(\mathbf{H})} \right]. \quad (30)$$

If Z_{T_0} is known, a Monte Carlo estimator for Z_{T_1} will be

$$Z_{T_1}(K) = Z_{T_0} \frac{1}{K} \sum_{k=1}^K \frac{A_{T_1}(\mathbf{H}_k)}{A_{T_0}(\mathbf{H}_k)}, \quad (31)$$

where, similarly to (24), we consider $\{\mathbf{H}_k, k=1, 2, \dots, K\}$ to be a collection of i.i.d. GRFs with $Pr[\mathbf{H}_k = \mathbf{h}] = \pi_{T_0}(\mathbf{h})$.

Estimator (31) provides good estimates of Z_{T_1} only if the Gibbs probability measures π_{T_0} and π_{T_1} , or, equivalently, temperatures T_0 and T_1 , are ‘‘close’’ enough to each other. Indeed, let $T_0 > T_1$, and define $T_d = (1/T_1 - 1/T_0)^{-1}$. By virtue of the *Central Limit Theorem* [8], the convergence rate of (31) to Z_{T_1} is determined by $\mathbf{Var}_{\pi_{T_0}}[g(\mathbf{H})] / \mathbf{E}_{\pi_{T_0}}^2[g(\mathbf{H})]$, where $g(\bullet) = A_{T_1}(\bullet) / A_{T_0}(\bullet)$. Note that [12]

$$\frac{\mathbf{Var}_{\pi_{T_0}}[g(\mathbf{H})]}{\mathbf{E}_{\pi_{T_0}}^2[g(\mathbf{H})]} < \frac{\max_{\text{states } \mathbf{h}} g(\mathbf{h})}{\min_{\text{states } \mathbf{h}} g(\mathbf{h})} \leq \left[\frac{\sigma_{\max}}{\sigma_{\min}} \right] \frac{MN}{T_d}, \quad (32)$$

where σ_{\max} and σ_{\min} are given by (3b,c), with the LTF considered at temperature $T=1$. Since we are interested in the computational complexity of the method, as $M, N \rightarrow +\infty$, we would like to bound the ratio $\mathbf{Var}_{\pi_{T_0}}[g(\mathbf{H})] / \mathbf{E}_{\pi_{T_0}}^2[g(\mathbf{H})]$ by

an upper bound which is not exponential, with respect to image size MN . If $T_d \geq MN$, then an upper bound in (32) will be $\sigma_{\max} / \sigma_{\min}$ which does not depend on MN . In practice, and if $T_d < MN$, we can introduce $\eta + 1$ temperatures $T_{(i)} \in [T_1, T_0]$, $i = 0, 1, \dots, \eta$, such that $T_{(0)} = T_0$, $T_{(\eta)} = T_1$, and

$$0 < \frac{1}{T_{(i+1)}} - \frac{1}{T_{(i)}} \leq \frac{1}{MN}, \quad \text{for every } i = 0, 1, \dots, \eta - 1. \quad (33)$$

Clearly, by letting $T_0 = +\infty$, the choice

$$\eta = \left\lceil \frac{MN}{T_1} \right\rceil, \quad T_{(0)} = +\infty, \quad T_{(\eta)} = T_1, \quad \text{and} \quad T_{(i)} = \frac{MN}{i}, \quad (34)$$

for $i = 1, 2, \dots, \eta - 1$, satisfies (33). Then, by using (21), (31), and (34), we have that

$$f(Z_{T_1}) = \frac{1}{MN} \ln Z_{T_1} = \ln R + \frac{1}{MN} \sum_{i=0}^{\eta-1} \ln \mathbf{E}_{\pi_{T_{(i)}}} \left[\frac{A_{T_{(i+1)}}(\mathbf{H})}{A_{T_{(i)}}(\mathbf{H})} \right]. \quad (35)$$

A Monte Carlo estimator for $f(Z_{T_1})$ will now be given by

$$F_{JS}(T_1) = \ln R + \frac{1}{MN} \sum_{i=0}^{\eta-1} \ln \left[\frac{1}{K_i} \sum_{k=1}^{K_i} \frac{A_{T_{(i+1)}}(\mathbf{H}_k)}{A_{T_{(i)}}(\mathbf{H}_k)} \right], \quad (36)$$

where, at each temperature $T_{(i)}$, $\{\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_{K_i}\}$ is a collection of statistically independent GRFs with identical probability mass functions $P_r[\mathbf{H}_k = \mathbf{h}] = \pi_{T_{(i)}}(\mathbf{h})$. Estimator (36) is an *asymptotically unbiased* and *consistent* estimator for $f(Z_{T_1})$, as $K_i \rightarrow +\infty$ [12]. We now have the following theorem:

Theorem 4: To calculate $f(Z_{T_1})$, by means of the Monte Carlo estimation scheme (34), (36), such that (8) is satisfied, it suffices to:

1. Use $\eta = \lceil MN/T_1 \rceil$ temperatures $T_{(i)}$, $i = 0, 1, \dots, \eta-1$, given by (34); and,
2. Use K_i i.i.d. samples, drawn from $\pi_{T_{(i)}}(\mathbf{h})$, where

$$K_i = \left\lceil \frac{1}{\xi} \eta (1 - e^{-\epsilon/\eta})^{-2} \frac{\sigma_{\max}}{\sigma_{\min}} \right\rceil = O((MN)^3), \quad (37)$$

for every $i = 0, 1, \dots, \eta-1$. In (37), σ_{\max} and σ_{\min} are given by (3b,c), with the LTF evaluated at temperature one.

According to Theorem 4, the computational cost of (36) is of a polynomial order $O((MN)^4)$, Turing reducible to sampling from the Gibbs probability measure. Therefore, and for sufficiently large M, N , *Algorithm [B-2]* will be computationally more preferable than *Algorithm [B-1]*.

We would like to conclude this section with three important remarks. First, Theorems 3 and 4 can be slightly modified, in order to apply to the case where averages of the forms (24) and (31) are estimated, by using independent samples drawn from probability mass functions which are sufficiently “close” to the required Gibbs probability measures [2]. Such samples can be obtained by using independent runs of the Gibbs Sampler Algorithm [12], [13]. Second, Theorem 3 is the first result of this type regarding *Algorithm [B-1]*, whereas Theorem 4 generalizes similar results obtained by Jerrum and Sinclair in [6]. Finally, assumption (3a) is necessary to derive the computational complexity results in (29), (37).

5. Simulation Results and Comparisons

The algorithms in Class [A] and Class [B] are, clearly, different in nature. Therefore, in this Section, we compare them within their own class, and determine the most efficient Algorithm [A] and [B].

A. Comparison of Algorithms [A].

Theorems 1 and 2 provide us with a means of comparing *Algorithms [A-1]* with *Algorithm [A-2]*. Clearly, and in order to study their relative efficiency, we need estimates of (see (12), (20))

$$D_P(M, N, T) = \frac{1}{MN} \ln \mathbf{E}_P[Q^2(\mathbf{h})] - \frac{2}{MN} \ln Z(\sigma) \quad (38a)$$

(where P stands for P^{iid} , P^* , or P^{**}), and

$$D_{OT_1}(M, N, T) = \frac{1}{MN} \ln Z(\sigma) + \frac{1}{MN} \ln Z(\sigma^{-1}) - 2 \ln R, \quad (38b)$$

respectively. Consistent Monte Carlo estimators for both quantities can be obtained by using, say, *Algorithm [B-2]*, with obvious adjustments. To simplify notation, we drop the dependence of these estimates on the image size.

We have compared such estimated complexity coefficients $D_{P^{iid}}(T)$, $D_{P^*}(T)$, $D_{P^{**}}(T)$, and $D_{OT_1}(T)$, to each other, as well as to $D_{BF}(T) = \ln R$ (complexity coefficient for performing the brute-force summation (2b) [12]), for a variety of GRFs. Such simulations verify Theorems 1, 2, and (17). See, for

example, Figure 1, where we depict such complexity coefficients, for the 32×32 site binary GRF1, with homogeneous LTF given in Table I, and $T \in [0.5, 2.0]$.

Clearly, *Algorithm [A-1]*, which is based on sampling from $P^*(\mathbf{h})$, is the superior one. Sampling from $P^{**}(\mathbf{h})$ can be advantageous for GRF models with weak interactions, because, in this case, $D_{P^*}(T) \approx D_{P^{**}}(T)$, and (13c) is much simpler to calculate than (13b).

B. Comparison of Algorithms [B].

Theorems 3 and 4 provide us with a means of comparing *Algorithm [B-1]* with *Algorithm [B-2]*. From (29a) and (34), it is clear that $\eta_s / \eta = O((MN)^{1/2s})$. In addition, by using (29), (34), and (37), we obtain

$$\frac{CPU_s}{CPU} = \frac{\sum_{i=0}^{\eta_s} K_{s,i}}{\sum_{i=0}^{\eta-1} K_i} \leq \frac{\xi \epsilon^2}{2} \frac{T_1^4}{(MN)^4} \frac{\sigma_{\min}}{\sigma_{\max}} \sum_{i=0}^{\eta_s} K_{s,i} = O((MN)^{1/s}), \quad (39)$$

where, CPU_s and CPU are the required computer times for (8) to hold, when using *Algorithm [B-1]* with Romberg integration of order $2s$, or *Algorithm [B-2]*, respectively. In (39), we assume that samples from the underlying Gibbs probability measure are available at a unit cost, and that computation of the functionals $U(\mathbf{H}_k)$ and $A(\mathbf{H}_k)$, in (24) and (36) respectively, is negligible. Clearly, *Algorithm [B-2]* is faster than *Algorithm [B-1]*, for sufficiently large images.

We have found that, although the values of η_s and η , obtained in Theorems 3 and 4 guarantee (8), they are unnecessarily large, as compared to the values that may be used in practice. We would, therefore, like to obtain the minimum possible values for η_s and η , as well as the corresponding computer times, that will still yield partition function estimators satisfying (8). We denote such values by $\bar{\eta}_s$, $\bar{\eta}$, and \bar{CPU}_s , \bar{CPU} , respectively. We can determine them exactly, only for the case of exactly solvable GRF models, i.e., models with analytically known $I(T)$ and Z_T , as

$$\bar{\eta}_s = \min \left\{ v : v \geq 1, v \bmod 2^{s-1} = 0, \text{ and } \left| \int_0^{1/T_1} I\left(\frac{1}{\beta}\right) d\beta - RO(s, v) \right| \leq \frac{\epsilon}{2MN} \right\}, \quad (40a)$$

and

$$\bar{\eta} = \arg \left\{ \min_{1 \leq v \leq \eta} \sum_{i=0}^{v-1} \left[\frac{2v^3}{\xi \epsilon^2} \left(\frac{Z_{T_{(i)}} Z_{T'_{(i)}}}{Z_{T'_{(i+1)}}} - 1 \right) \right] \right\}, \quad (40b)$$

where

$$T'_{(i)} = \left(\frac{2}{T_{(i+1)}} - \frac{1}{T_{(i)}} \right)^{-1}, \quad \text{for } i = 0, 1, 2, \dots, v-1. \quad (40c)$$

In addition, $\bar{CPU}_s = \sum_{i=0}^{\bar{\eta}_s} K_{s,i}$, and $\bar{CPU} = \sum_{i=0}^{\bar{\eta}-1} K_i$, where the summands are given by (29b) and (37). More details can be found in [12]. Figures 2-4 depict $\bar{\eta}_1$, $\bar{\eta}_2$, $\bar{\eta}_3$, $\bar{\eta}$, \bar{CPU}_2 , and \bar{CPU} , for the case of the exactly solvable 128×128 site GRF2 model with LTF given on Table I. From these figures, as well as from a variety of similar simulations in [12], we conclude that *Algorithm [B-2]* is the best Algorithm [B], whereas Simpson’s integration rule constitutes the most appropriate implementation of *Algorithm [B-1]*.

6. A Unifying GRF Partition Function Estimation Algorithm

In this Section, we seek connections between the best Algorithms in Classes [A] and [B], namely, *Algorithm [A-1]* and *Algorithm [B-2]*. We study the *Generalized Algorithm [B-2]* [7], suitable for addressing statistical inference tasks for GRF images. This algorithm reduces to *Algorithms [A-1]*, or *[B-2]*, in special cases.

Notice first that, in (34), we consider a number of Gibbs probability measures, $\pi_{T_i}(\mathbf{h})$, at equally spaced inverse temperatures $\beta_i = 1/T_i \in [0, 1/T_1]$. The value $\beta = 0$ corresponds to the i.i.d. Markov mesh $P^{iid}(\mathbf{h})$, with LTF τ^{iid} given by (13a), whereas the value $\beta = 1/T_1$ corresponds to the GRF whose partition function we wish to estimate, with LTF, let's say, $\sigma_{(1)}$. The LTF of the Gibbs probability measure considered in (35), lies on a one-parameter curve in the LTF space, connecting $\sigma_{(1)}$ and τ^{iid} (see Figure 5).

There is clearly no reason to restrict such curves to include τ^{iid} . Indeed, given any two LTFs

$$\sigma_{(1)} = \{ \sigma_{ij}^{(1)}(x, y, z, \omega) : x, y, z, \omega \in E ; (i, j) \in \Lambda \}, \quad (41a)$$

and

$$\sigma_{(0)} = \{ \sigma_{ij}^{(0)}(x, y, z, \omega) : x, y, z, \omega \in E ; (i, j) \in \Lambda \}, \quad (41b)$$

we can estimate the ratio of their corresponding partition functions, namely $Z(\sigma_{(1)})/Z(\sigma_{(0)})$, as

$$f(\hat{Z}(\sigma_{(1)})) - f(\hat{Z}(\sigma_{(0)})) = \frac{1}{MN} \sum_{l=0}^{\eta-1} \ln \left[\frac{1}{K_l} \sum_{k=1}^{K_l} \frac{A_{\sigma_{l+1}}(\mathbf{H}_k)}{A_{\sigma_l}(\mathbf{H}_k)} \right], \quad (42a)$$

where

$$\sigma_l = \left\{ \left[\sigma_{ij}^{(1)}(x, y, z, \omega) \right]^{\frac{l}{n}} \times \left[\sigma_{ij}^{(0)}(x, y, z, \omega) \right]^{1-\frac{l}{n}} : x, y, z, \omega \in E ; (i, j) \in \Lambda \right\}, \quad (42b)$$

for $l = 0, 1, \dots, n$. In (42a), and at each LTF σ_l , $\{\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_{K_l}\}$ is a collection of statistically independent GRFs with identical probability mass functions $Pr[\mathbf{H}_k = \mathbf{h}] = \pi_{\sigma_l}(\mathbf{h})$. We refer to (42) as the *Generalized Algorithm [B-2]*. Notice that the LTF σ_l , $l = 0, 1, \dots, n$, lies on a one-parameter curve connecting $\sigma_{(1)}$ and $\sigma_{(0)}$ (see Figure 5). Naturally, (42) reduces to (36), when $\sigma_{(0)} = \tau^{iid}$, i.e., to *Algorithm [B-2]*. Clearly, (42) gives *asymptotically unbiased and consistent* estimators of the ratio $Z(\sigma_{(1)})/Z(\sigma_{(0)})$. Similarly to Theorem 4, we now have the following theorem:

Theorem 5: *To calculate the difference $f(Z(\sigma_{(1)})) - f(Z(\sigma_{(0)}))$ by means of the Monte Carlo estimation scheme (42), such that the resulting estimate satisfies (8), it suffices to:*

1. Use $\eta = \lceil MN \rceil$ in (42); and
2. Use K_l i.i.d. samples, drawn from $\pi_{\sigma_l}(\mathbf{h})$, where

$$K_l = \left\lceil \frac{1}{\xi} \eta (1 - e^{-\xi/\eta})^{-2} d(\sigma_{(1)}, \sigma_{(0)}) \right\rceil = O((MN)^3),$$

for every $l = 0, 1, \dots, \eta-1$, and

$$d(\sigma_{(1)}, \sigma_{(0)}) = \max_{\substack{(i,j) \in \Lambda \\ (x,y,z,\omega) \in E^4}} \left[\frac{\sigma_{ij}^{(1)}(x,y,z,\omega)}{\sigma_{ij}^{(0)}(x,y,z,\omega)} \right] \max_{\substack{(i,j) \in \Lambda \\ (x,y,z,\omega) \in E^4}} \left[\frac{\sigma_{ij}^{(0)}(x,y,z,\omega)}{\sigma_{ij}^{(1)}(x,y,z,\omega)} \right].$$

We would now like to emphasize the following two points of practical importance:

First, quite often, $d(\sigma_{(1)}, \tau_{(1)}^*) < d(\sigma_{(1)}, \tau^{iid})$, where $\tau_{(1)}^*$ is the LTF of a Markov mesh which approximates a GRF with LTF $\sigma_{(1)}$, as in (13b). Therefore, and in order to estimate the partition function of a GRF with LTF $\sigma_{(1)}$, it is advantageous, instead of using (11), to first estimate $\tau_{(1)}^*$, and then run the *Generalized Algorithm [B-2]* along the curve defined by $\tau_{(1)}^*$ and $\sigma_{(1)}$ (see (42) and Figure 5). If, in (42), we use $n=1$, the *Generalized Algorithm [B-2]* reduces to the best *Algorithm [A-1]*.

Second, if we assume that $\sigma_{(1)}$ and $\sigma_{(0)}$ are sufficiently "close" to each other, then $d(\sigma_{(1)}, \sigma_{(0)}) < d(\sigma_{(0)}, \tau^{iid}) + d(\sigma_{(1)}, \tau^{iid})$. Therefore, using (42), as compared to using (36) twice, results in computational savings, when the problem is to estimate partition function or likelihood function ratios.

The second case arises in hypothesis testing problems. On the other hand, and in maximum likelihood estimation problems, we often need likelihood function estimates in a neighborhood of a specific LTF. Clearly, in both cases, the *Generalized Algorithm [B-2]* provides us with a flexible stochastic simulation scheme. Therefore, (42) is very suitable for statistical inference problems, based on likelihood calculations.

7. Conclusions

We have presented a unified analysis of stochastic simulation algorithms for partition function and, consequently, likelihood function estimation of general GRF images. We have focused our attention on determining their computational complexity, with respect to image size. By using Theorems 1-4, we have been able to classify these algorithms into two categories: *Algorithms [A]*, with exponential complexity, and *Algorithms [B]*, with polynomial complexity, Turing reducible to sampling from the Gibbs probability measure. We have determined that *Algorithms [A-1]* and *[B-2]* are the most efficient ones, within their own class. We have then introduced the *Generalized Algorithm [B-2]*, which contains both *Algorithm [A-1]* and *Algorithm [B-2]* as special cases, and discussed its suitability in GRF statistical inference problems.

Acknowledgements

This work was supported by the office of Naval Research, Mathematical Sciences Division, under ONR Grant N00014-90-J-1345.

References

- [1] R.C. Dubes and A.K. Jain, "Random field models in image analysis," *J. of Applied Statistics*, vol. 16, pp. 131-164, 1989.
- [2] G. Potamianos and J. Goutsias, "An analysis of Monte Carlo methods for likelihood estimation of Gibbsian images," *Proceedings of the ICASSP*, vol. 5, pp. 519-522, Minneapolis, Minnesota, 1993.
- [3] G. Potamianos and J. Goutsias, "Partition function estimation of Gibbs random field images using Monte Carlo simulations," *IEEE Transactions on Information Theory*, vol. 39, pp. 1322-1332, 1993.
- [4] Y. Ogata and M. Tanemura, "Estimation of interaction potentials of spatial point patterns through the maximum likelihood procedure," *Annals of the Institute of Statistical Mathematics*, vol. 33, Part B, pp. 315-338, 1981.
- [5] Y. Ogata, "A Monte Carlo method for an objective Bayesian procedure," *Annals of the Institute of Statistical Mathematics*, vol. 42, pp. 403-433, 1990.
- [6] M. Jerrum and A. Sinclair, "Polynomial-time approximation algorithms for the Ising model," *SIAM Journal on Computing*, vol. 22, pp. 1087-1116, 1993.
- [7] C.J. Geyer and E.A. Thompson, "Constrained Monte Carlo maximum likelihood for dependent data," (with Discussion), *Journal of the Royal Statistical Society, Series B*, vol. 54, pp. 657-699, 1992.
- [8] Bickel, P. J. and Doksum, K. A., *Mathematical Statistics*. Holden-Day, Oakland, California, 1977.
- [9] Comets, F. and Gidas, B., "Parameter estimation for Gibbs distributions from partially observed data," *The Annals of Applied Probability*, vol. 2, pp. 142-170, 1992.
- [10] R.S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*. New York City, NY: Springer-Verlag, 1985.

- [11] J. Goutsias, "Unilateral approximation of Gibbs random field images," *CVGIP: Graphical Models and Image Processing*, vol. 53, pp. 240-257, 1991.
- [12] G. Potamianos and J. Goutsias, "Analysis of stochastic simulation algorithms for partition function estimation of Gibbs images," *Technical Report JHU/ECE 93-05*, ECE Dept., The Johns Hopkins University, Baltimore, MD, 1993.
- [13] J.K. Goutsias, "A theoretical analysis of Monte Carlo algorithms for the simulation of Gibbs random field images," *IEEE Transactions on Information Theory*, vol. 37, pp. 1618-1628, 1991.

Table I: LTF of the two GRFs considered in Section 5.

LTF	GRF1	GRF2
$T \times \ln \sigma_{ij}(0,0,0,0)$	0.000000	0.881376
$T \times \ln \sigma_{ij}(0,0,0,1)$	0.000000	0.000000
$T \times \ln \sigma_{ij}(0,0,1,0)$	0.000000	0.881376
$T \times \ln \sigma_{ij}(0,0,1,1)$	0.000000	0.000000
$T \times \ln \sigma_{ij}(0,1,0,0)$	0.000000	0.000000
$T \times \ln \sigma_{ij}(0,1,0,1)$	0.601010	-0.881376
$T \times \ln \sigma_{ij}(0,1,1,0)$	0.000000	0.000000
$T \times \ln \sigma_{ij}(0,1,1,1)$	0.601010	-0.881376
$T \times \ln \sigma_{ij}(1,0,0,0)$	0.000000	-0.881376
$T \times \ln \sigma_{ij}(1,0,0,1)$	-1.262626	0.000000
$T \times \ln \sigma_{ij}(1,0,1,0)$	0.601010	-0.881376
$T \times \ln \sigma_{ij}(1,0,1,1)$	-0.661616	0.000000
$T \times \ln \sigma_{ij}(1,1,0,0)$	-1.262626	0.000000
$T \times \ln \sigma_{ij}(1,1,0,1)$	-1.924242	0.881376
$T \times \ln \sigma_{ij}(1,1,1,0)$	-0.661616	0.000000
$T \times \ln \sigma_{ij}(1,1,1,1)$	-1.323232	0.881376

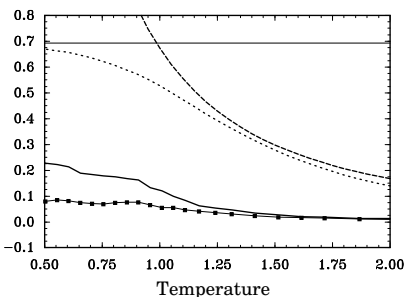


Figure 1: Monte Carlo estimates of the complexity coefficients associated with *Algorithms [A]*, and for the GRF1 model of Table I, considered on a 32×32 site lattice.

— $D_{BF}(T)$ $D_{p^{iid}}(T)$ - - - $D_{p^*}(T)$
 - · - $D_{p^{**}}(T)$ - - - - $D_{OT_1}(T)$

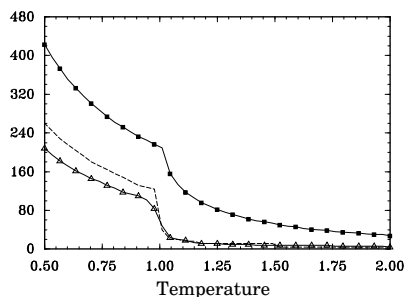


Figure 2: Values of $\bar{\eta}_1$, $\bar{\eta}_2$, and $\bar{\eta}_3$, required for estimating $f(Z_T)$ of the GRF2 model of Table I, considered on a 128×128 site lattice, by means of *Algorithm [B-1]*. In this case $\epsilon = 0.1$.

— $\bar{\eta}_1$ - - - $\bar{\eta}_2$ $\bar{\eta}_3$

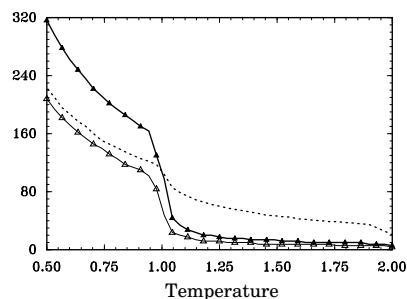


Figure 3: Values of $\bar{\eta}_2$ and $\bar{\eta}$, required for estimating $f(Z_T)$ of the GRF2 model of Table I, considered on a 128×128 site lattice, by means of *Algorithm [B-2]*.

— $\bar{\eta}_2$, for $\epsilon = 0.1$;
 - - $\bar{\eta}_2$, for $\epsilon = 0.01$;
 $\bar{\eta}$, for all $0 < \epsilon < 1$.

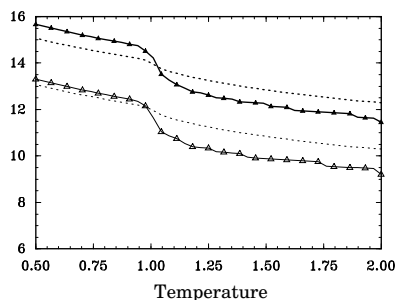


Figure 4: Values of \overline{CPU}_2 and \overline{CPU} associated with estimating $f(Z_T)$ of the GRF2 model of Table I, considered on a 128×128 site lattice, and for $\epsilon = 0.1$, and $\epsilon = 0.01$. In this case, $\xi = 0.05$.

— $\log_{10} \overline{CPU}_2$, for $\epsilon = 0.1$;
 $\log_{10} \overline{CPU}$, for $\epsilon = 0.1$;
 - - $\log_{10} \overline{CPU}_2$, for $\epsilon = 0.01$;
 - · - $\log_{10} \overline{CPU}$, for $\epsilon = 0.01$.

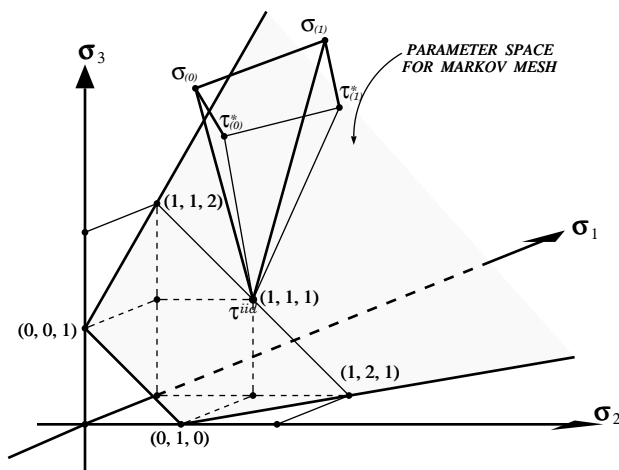


Figure 5: Parameter space for a simple GRF model with homogeneous LTF $\sigma = [1, \sigma_1, \sigma_2, \sigma_3]$. The parameter space for the Markov mesh LTF is a plane.