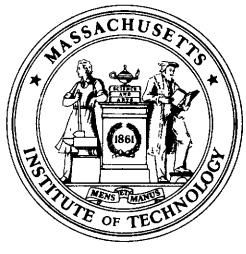
From Random Fields to Networks

Ibrahim M. Elfadel

RLE Technical Report No. 579

June 1993



The RESEARCH LABORATORY *of* ELECTRONICS
MASSACHUSETTS INSTITUTE of TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS 02139-4307

From Random Fields to Networks

Ibrahim M. Elfadel

RLE Technical Report No. 579

June 1993

Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139-4307

This work was supported in part by the National Science Foundation under Grant MIP 91-17724.

From Random Fields to Networks

by

Ibrahim M. Elfadel

Submitted to the Department of Mechanical Engineering
on February 1, 1993, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Gibbs random fields (GRF's) are probabilistic models inspired by statistical mechanics and used to model images in robot vision and image processing. In this thesis, we bring the analytical methods of statistical mechanics to bear on these models. Specifically, we address and solve the following fundamental problems:

1. Mean-field estimation of a constrained GRF model: The configuration space of a GRF model is often constrained to produce "interesting" patterns. We develop mean-field equations for estimating the means of these constrained GRF's. The novel feature of these equations is that the finiteness of graylevels is incorporated in a "hard" way in the equations.

2. Correlation-field estimation of a GRF model: GRF correlation functions are generally hard to compute analytically and expensive to compute numerically. We use the mean-field equations developed above to propose a new procedure for *estimating* these correlation functions. Our procedure, which is valid for both unconstrained and constrained models, is applied to the quadratic interaction model, and a new closed-form approximation for its correlation function in terms of the model parameters is derived.

3. Network representation: We show how the mean-field equations of the GRF model can be mapped onto the fixed-point equations of an analog winner-take-all (WTA) network. The building block of this network is the *generalized sigmoid mapping*, a natural generalization of the sigmoidal function used in artificial neural networks. This mapping turns out to have a very simple VLSI circuit implementation with desirable circuit-theoretic properties.

4. Solution algorithms: Iterated-map methods and ordinary differential equations (ODE's) are proposed to solve the network fixed-point equations. In the former, we show, using Lyapunov stability theory for discrete systems, that the worst that could happen during synchronous iteration is an oscillation of period 2. In the latter, we show that the ODE's are the gradient descent equations of energy functions that can be derived from the mean-field approximation. One of our gradient descent algorithms can be interpreted as the generalization to analog WTA networks of Hopfield's well-known algorithm for his analog network.

5. Temperature dependence: The GRF temperature parameter reflects the thermodynamic roots of the model. Using eigenstructure analysis, we study the temperature effect on

the stability of the WTA network fixed points. In particular, we derive new closed-form formulas for the *critical temperature* of a large class of models used in grayscale texture synthesis. The stability study is used to gain insight into the phase transition behavior of GRF's.

The implications of these results for image modeling, optimization, and analog hardware implementation of image processing and optimization algorithms are discussed.

Thesis Supervisor: John L. Wyatt, Jr.

Professor of Electrical Engineering and Computer Science

Acknowledgments

First and foremost, I would like to acknowledge the continuous encouragement and support I have received from my thesis supervisor, Professor John L. Wyatt, Jr. It has been for me an honor and a privilege to work with John both as his research and teaching assistant. His emphasis on clarity, depth, and rigor has certainly shaped my own approach to academic research. I would also like to express my gratitude for his close and interested reading of the thesis document.

I am thankful to my thesis committee members: Professor Berthold Horn for asking the right questions and Professor Alan Yuille for helping me answer them. I am also grateful to Alan for the time he gave me and for the many illuminating discussions we had.

Special thanks go to the chairman of my committee, Professor Derek Rowell, for his help in getting me better organized and more goal-oriented. Mission accomplished. Witness this document!

Many MIT faculty members contributed both academically and financially to my graduate education while at MIT. In particular, I would like to acknowledge the financial support that I received, in the form of research assistantships, from Professors David Hardt, Sanjoy Mitter, and Ronald Rivest, and in the form of a teaching assistantship from Professor Munther Dahleh.

I would also like to acknowledge and thank Professor Jacob White for his readiness to help and for his advice on both academic and non-academic matters.

Doctor Robert Armstrong, my officemate for the past three years, was always there to patiently answer my ill-posed questions about computing and text formatting with the well-thought replies of an outstanding expert. Bob also ended up proofreading parts of this thesis. All the remaining bugs and typos are solely my responsibility. A zillion thanks, Bob!

While working on my own dissertation, I had the unique privilege of collaborating with two excellent colleagues: Professor Rosalind Picard and Professor Andrew Lumsdaine. It is a pleasure to thank them for sharing research problems and ideas with me.

To past and present members of the VLSI-CAD group at RLE: thanks for your friendship and for sharing your thoughts, stories, jokes, and math questions with me.

To past and present members of the MIT Vision Chip Project: thanks for being so patient with this "theory stuff."

To my friends and colleagues in Syria, France, Japan, and the US: thanks for your cheers along the long road.

To my relatives in Syria who have been calling me "Doctor" since I was in junior high: sorry you had to wait so long!

Without my family's financial sacrifices, emotional support, and unconditional love, this thesis would have remained no more than a possibility. To them I say:"I love you all."

وَمَا تَوْفِيقِي إِلَّا بِٱللَّهِ

The work described in this thesis was supported in part by the National Science Foundation and the Defense Advanced Research Projects Agency (DARPA) under Grant No. MIP-88-14612, and by the National Science Foundation under Grant No. MIP-91-17724.

to

my mother

my sister

my brothers

and the blessed memory of my father

Contents

1

Introduction

. . . وَمَنْ يُرِيدُ عِلْمَ كُلِّ شَيْءٍ ، فَيَنْبَغِي لِأَهْلِهِ أَنْ يُدَاوُوْهُ !

أبو عثمان الجاحظ

The use of Gibbs random fields (GRFs) for modeling images is now well established in the image processing and computer vision literature. These random field models arise very often in the context of Bayesian formulations as our *a priori* assumptions about the noise-free image in image processing or the imaged 3D scene in computer vision [31, 58, 84]. They can also be interpreted as the statistical incarnations of regularizers under the regularization paradigm of early vision theory [76].

When considered for their own sake, these GRFs become probabilistic descriptions of grayscale patterns. For instance, early study of these models in the context of texture modeling concentrated on how they can be used to fit natural textures (wood, cork, textile, etc) [14]. Later, they were used for texture segmentation, e.g., [16, 30, 79] and more recently [37].

Another major area of application where these models are used is the development of integrated vision systems for object recognition. In [21], GRF modeling is used as a paradigm for integrating different visual cues like shading, stereo, and texture in order to recognize objects. A similar use of GRFs is the one described in [12] where they are used in the context of fusing the data obtained from different sensors to accomplish estimation and recognition tasks.

The scope of this thesis is much more limited. Here, we concentrate on some basic questions that are typically asked by the typical engineer when handed a random signal: "What is its mean? What is its covariance?" The mean and the covariance of a random signal are usually the quantities of engineering interest, especially in signal processing application. GRFs are remarkable in that these quantities are generally unknown because the normalizing constant of a Gibbs distribution, the famous partition function, is generally hard to compute. Most of the reserach work concerning this question has concentrated on the Gaussian case [6, 11, 48, 61] for which analytical results can be obtained.

The other question relevant to engineering applications is the dependence of the random signal on its parameters. Here also, GRFs are in a class by themselves in that their parameter estimation [1] problem remains largely open despite a sustained research effort by statisticians, engineers, and applied mathematicians [6, 14, 72, 91].

The usual way for attacking these questions in the engineering literature has been through the extensive simulations of specific models in specific contexts. In this thesis, we follow a different path – more analytical and more generic. It is more analytical for we bring the mathematical methods of statistical physics to bear on some of the above questions. And it is more generic for we show how the use of such methods can give answers for a whole class of models rather than specific, isolated ones.

The underlying theme of this dissertation is the following.

> *The mean-field model of statistical physics, when applied to a Gibbs distribution with discrete gray levels, can yield both moment estimates and parameter estimates.*

This theme is illustrated in the estimation of the correlation field (Chapter 4) and the estimation of the *critical temperatures* delimiting visually different Gibbs samples in the parameter space (Chapter 6).

As was mentioned above, GRFs fit naturally within the Bayesian framework of image processing and computer vision. In this framework, the *a priori* GRF model is usually competing against the sensor noise model. The impact of too much trust in the *a priori* model is sometimes heuristically expressed in such catch sentences as "too much prior is far more dangerous than none at all" [45]. Although this thesis does not deal directly with appliactions, we suggest that its framework and results could help quantify such statements.

[1] The word *parameter* is to be understood in a generic way. In the case of GRF, its meaning encompasses lattice size and geometry, the neighborhood size and geometry, the local potentials, the nature of global constraints, etc.

1.1 Notation and Definitions

In this section, we give the basic notation and definitions that will be used throughout this thesis. We are concerned with a *finite* [2] set $\mathcal{S}$ which will represent a two-dimensional image. A neighborhood system for this set is a family $\mathbf{N} = \{\mathcal{N}_s \subset \mathcal{S}, s \in \mathcal{S}\}$ of subsets satisfying the following two conditions:

1. $s \notin \mathcal{N}_s$ and

2. $s \in \mathcal{N}_r$ if and only if $r \in \mathcal{N}_s$, $\forall s, r \in \mathcal{S}$

The elements of $\mathcal{S}$ are the image *pixels*. Since the models considered in this thesis have their roots in statistical physics, we will sometimes use the word *site* to designate an element of $\mathcal{S}$. The pixels $r \in \mathcal{N}_s$ are the neighbors of s. The first condition above means that the site s is *not* its own neighbor. The second condition introduces a symmetry property in the definition of a neighborhood: s is a neighbor of r if and only if r is a neighbor of s. Note that when the set $\mathcal{S}$ is a square grid, the symmetry condition is equivalent to the absence of a preferred direction in the image. This is consistent with the expectation that an image model should be noncausal.

The pair $(\mathcal{S}, \mathbf{N})$ of the set and its neighborhood system is an undirected graph whose vertices are the pixels and whose edges are formed by pairs of neighbors. A subset $C \subset \mathcal{S}$ is called a *clique* if every two distinct pixels in C are neighbors. The concept of clique is important for defining Gibbs potentials and relating Gibbs distributions to Markov random fields.

To each pixel $s \in \mathcal{S}$, we assign a random variable X_s taking its values x_s in a set G_s usually called the spin, or pixel, state space. [3] The family $(X_s)_{s \in \mathcal{S}}$ is a *random field.* The product set

$$\Omega = \prod_{s \in \mathcal{S}} G_s$$

is the *configuration space* of the random field. All the random fields considered in this thesis are such that $G_s = G$, $\forall s \in \mathcal{S}$. Note that the nature of the field depends crucially on the topological nature of the set G. Some of the most commonly used G sets with examples of their associated random fields are listed below:

[2] See the Appendix for the complications that might occur when infinite lattices are considered.

[3] We follow the common custom of denoting a random variable with an upper-case letter and denoting its value with a lower-case letter.

1. $G = \Re$. The pixels of the well known Gaussian field have the real line for state space.

2. G is a compact interval $\subset \Re$. This is the pixel state space of the rotor model [32]. Typically, $G = [0, 2\pi]$. Note also that analog grayscale images have a compact pixel state space.

3. $G = \mathbf{Z}$. The so-called discrete Gaussian model [32] is defined in this manner.

4. G is a finite subset of $\mathbf{Z}$. For instance, the autobinomial model studied in [14] falls under this category. The set $G = \{0, 1, \ldots, 255\}$ is the pixel state space of all grayscale digitized images having an 8-bit pixel representation.

5. G is a binary set typically taken to be $\{-1, 1\}$. This is a special case of the previous one. A famous example in this category is the Ising model of magnetism. This is also a generic model for all binary images.

Models differ also in the nature of the underlying set $\mathcal{S}$. The most common cases are those for which $\mathcal{S}$ is a subset of $\mathbf{Z}$, which corresponds to one-dimensional random processes, or a subset of $\mathbf{Z}^2$, which corresponds to two-dimensional random fields. As was mentioned above, we mainly consider the case of a *finite* set $\mathcal{S}$. However in Chapter 2, we deal with Gaussian processes defined on *all* $\mathbf{Z}$. In the finite case, we make the assumption that the pixels of the set $\mathcal{S}$ satisfy a periodic boundary condition. This means that the graph $(\mathcal{S}, \mathbf{N})$ lies on a circle in the one-dimensional case, and on a torus in the two-dimensional case.

Furthermore, for a given underlying pixel set $\mathcal{S}$, models could differ in the geometry and order of the neighborhood system $\mathbf{N}$. Consider the case when $\mathcal{S}$ is a square grid. Figure 1-1 shows the commonly used neighborhood systems for this case. These neighborhoods are distinguished by their orders. The four nearest neighbors define a first-order neighborhood system, the eight nearest neighbors define a second-order neighborhood system, and so forth. Note that other neighborhood geometries are possible including ones in which the pixel neighbors are disconnected from the pixel itself. In Figure 1-2, we show the cliques associated with the second-order neighborhood model. The set of cliques include single sites, pairs of sites, triplets of sites, and quadruplets of sites.

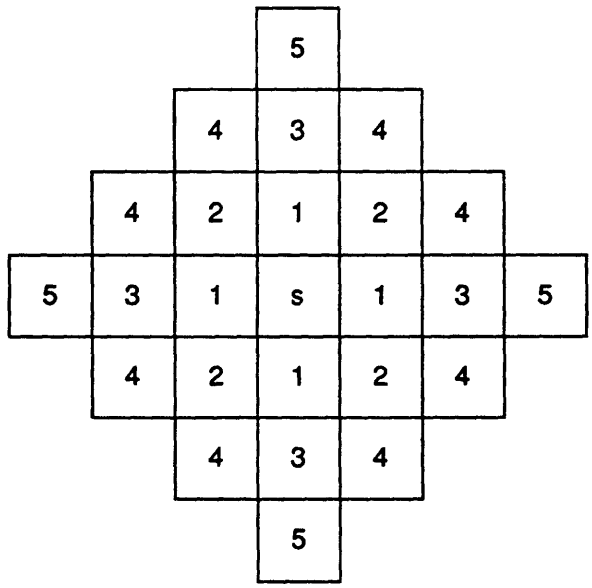


FIGURE 1-1: Neighborhood orders. The nearest-neighbor pixels denoted by 1 form a neighborhood of order 1. A second order neighborhood is formed by the eight nearest neighbors and so forth.

1.2 Markov Random Fields and Gibbs Distributions

Now we are in a position to introduce basic definitions that will be fundamental for this thesis. They are now part of the standard lore in the literature on Markov and Gibbs random fields [19, 31]. From the previous section, we have the following setup: a graph $(\mathcal{S}, \mathbf{N})$ and a family $\mathbf{X} = \{X_s, s \in \mathcal{S}\}$ of random variables indexed by the set $\mathcal{S}$ taking their values in a configuration space Ω. In order to avoid using the language of abstract integration theory, we will assume that the underlying pixel state space is discrete and finite, e.g., $G = \{0, 1, \ldots, n-1\}$. Let now P be a probability measure defined on Ω.

Definition 1 *The random field* $\mathbf{X}$ *is said to be a Markov random field (MRF) if the following two conditions are satisfied:*

1. $P(\mathbf{X} = \mathbf{x}) > 0, \ \forall \mathbf{x} \in \Omega$;
2. $P(X_s = x | X_r = x_r, \forall r \neq s) = P(X_s = x | X_r = x_r, \forall r \in \mathcal{N}_s), \ \forall s \in \mathcal{S}, \mathbf{x} \in \Omega$.

We should stress that both conditions are essential for the definition of an MRF. Note that when constraints are imposed on the possible configurations, the first condition might fail. More will be said on the constrained case in Chapter 2. The functions given in

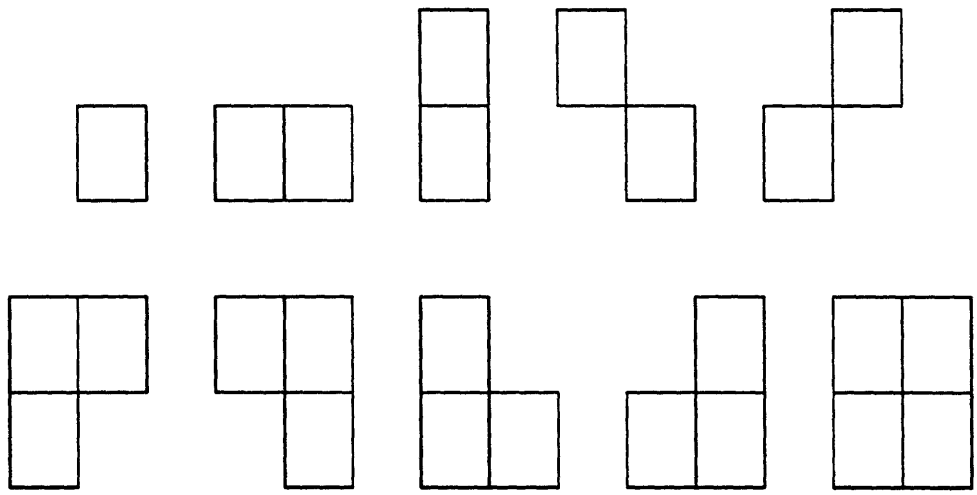


FIGURE 1-2: Cliques of a second-order neighborhood. Note the presence of cliques containing three sites and a clique containing four sites.

the left-hand side of the second condition, known as the Markov property, are called the *local characteristics* of the MRF. A compact notation for each one of them is $P_s(x_s|\mathbf{x}_{(s)})$, where $\mathbf{x}_{(s)}$ denotes the configuration obtained from $\mathbf{x}$ after the removal of the pixel s. A fundamental property that will be proved in Appendix A is that when the set $\mathcal{S}$ is finite, the probability distribution P is uniquely determined by its local characteristics. When P satisfies the Markov property, the local characteristics can be computed more easily, just by looking at the neighbors of each pixel instead of looking at the the rest of the configuration $\mathbf{x}_s$. Note that this computational locality will be violated when constraints are imposed on the the configuration set Ω.

Having defined Markov random fields, we now introduce their kin, Gibbs random fields. Again the concepts that will follow are standard, and our presentation is taken from [30].

Definition 2 *A* potential *is a collection* $V = \{V_A, A \subset \mathcal{S},\ V_A : \Omega \rightarrow \Re\}$ *of functions such that* $V_\emptyset = 0$ *(*$\emptyset$ *is the empty set), and* $V_A(\mathbf{x}) = V_A(\mathbf{x}')$ *if* $x_s = x'_s$, $\forall s \in A$.

Note that the above notion of potential is valid for any set $\mathcal{S}$. When this latter set is equipped with a neighborhood system $\mathbf{N}$ that makes the pair $(\mathcal{S}, \mathbf{N})$ a graph with a set of cliques $\mathcal{C}$, we have the following

Definition 3 *A* neighborhood-compatible potential, *or an* $\mathbf{N}$*-potential, is a potential such that* $V_A = 0, \forall A \notin \mathcal{C}$.

This definition means that interactions between pixels belonging to different cliques are not allowed. A Gibbs distribution is a representation of a *positive* probability distri-

bution motivated by the study of thermodynamic equilibrium in statistical physics. It is defined as follows.

Definition 4 *A Gibbs distribution with respect to the neighborhood system* **N** *is a measure of the form*

$$\Pi(\mathbf{x}) = Z^{-1} \exp[-\frac{1}{T}E(\mathbf{x})], \quad Z = \sum_{\mathbf{x}\in\Omega} \exp[-\frac{1}{T}E(\mathbf{x})], \tag{1.1}$$

where $E(\mathbf{x})$, the energy function, is given by

$$E(\mathbf{x}) = -\sum_{C\in\mathcal{C}} V_C(\mathbf{x}) \tag{1.2}$$

and "T" is the temperature.

The positive normalizing constant, Z, in (1.1) is known in statistical mechanics as the partition function. [4] Except in very few cases, this function is very hard to compute both analytically and numerically. Let us now give some comments regarding the above definition.

1. The measure defined by the Gibbs distribution is always positive, i.e., $\Pi(\mathbf{x}) > 0$.

2. Only cliques are involved in the summation defining the Gibbs energy. In other words, the potential V is a neighborhood-compatible potential in the sense of Definition 3. This fact is crucial for relating Gibbs distributions to MRFs.

3. The temperature parameter reflects the thermodynamical roots of the Gibbs distribution. The states of a system S of energy E, in equilibrium with an infinitely large heat reservoir at temperature T, are distributed according to the Gibbs distribution given above. This is the *canonical postulate* of statistical mechanics [85].

4. When the configuration set Ω is infinite, the condition $Z < +\infty$ must be imposed. This situation arises when the set $\mathcal{S}$ is infinite or when the pixel state space is infinite.

---

[4] The origin of this name for the normalizing constant is the following. Let Ω_e be the subset of the configuration space Ω defined by $\Omega_e = \{\mathbf{x} \in \Omega \text{such that} E(\mathbf{x}) = e\}$. Then we can write

$$\begin{aligned} Z &= \sum_e \sum_{\mathbf{x}\in\Omega_e} \exp[-\frac{1}{T}E(\mathbf{x})] \\ &= \sum_e |\Omega_e| \exp[-\frac{e}{T}]. \end{aligned}$$

The family $(\Omega_e)_e$ is a partition of the configuration space Ω. Moreover Z can be computed as an exponentially weighted sum of the partition set cardinalities.

5. The above definition does not involve constraints on the configuration space. We will see in Chapter 2 how constraints can be incorporated in the Gibbs distribution.

6. Finally, one can easily check that as the temperature parameter approaches 0, the mean of the Gibbs distribution approaches its mode. This observation is fundamental to applications like maximum *a posteriori* estimation.

It is easy to check that every positive measure Π can be written as

$$\Pi(\mathbf{x}) = Z^{-1} \exp[-E(\mathbf{x})].$$

(Just *define* $E(\mathbf{x}) = \log \frac{\Pi(0)}{\Pi(\mathbf{x})}$, and put $Z^{-1} = \Pi(0)$.) A more delicate problem is to find the *potential* V such that

$$E(\mathbf{x}) = -\sum_{A \subset \mathcal{S}} V_A(\mathbf{x}).$$

Note that in the above formula, the summation is taken over *all* the subsets of $\mathcal{S}$. In other words, Π is still not a Gibbs distribution in the sense of Definition 4. This brings us to the second delicate problem, namely, find sufficient conditions for the energy expression to be a summation over cliques only.

The answers to the above two questions are not obvious. They constitute the proof of the "if" part of the following famous theorem

Theorem 1 *[Hammersley-Clifford (1972)] Let* $\mathbf{N}$ *be a neighborhood system defined on a* finite *set* $\mathcal{S}$. *Then* Π *is a Gibbs distribution with respect to the neighborhood system* $\mathbf{N}$ *if and only if* Π *is a Markov random field with respect to* $\mathbf{N}$. *Moreover, for any clique* C *and any site* $s \in C$, *the* $\mathbf{N}$*-potential is given by*

$$V_C(\mathbf{x}) = \sum_{B \subset C} (-1)^{|C-B|} \ln \Pi_s(x_s | \mathbf{x}^B_{(s)}), \tag{1.3}$$

where $\mathbf{x}^B = (x^B_q)_{q \in \mathcal{S}}$ *denotes the following configuration*

$$x^B_q = \begin{cases} x_q, & q \in B \\ 0, & q \notin B \end{cases}$$

and $\Pi_s(x_s | \mathbf{x}^B_{(s)})$ *is the conditional probability at* s *given the configuration* $\mathbf{x}^B_{(s)}$.

The "only if" part is easy to prove. For the "if" part, we refer the reader to the excellent course notes of D. Geman [30]. We would like however to make the following remarks:

1. Theorem 1 will fail if the set $\mathcal{S}$ is not finite. More specifically, the local characteristics of an MRF are not sufficient, in the infinite case, to guarantee the existence of a unique Gibbs distribution over the lattice. The existence and uniqueness of Gibbs distributions for infinite sets are the fundamental questions of the mathematical theory of phase transitions. See Appendix A.

2. It will also fail if constraints are imposed on the configuration Ω. The effect of these constraints is often a change in the neighborhood system and therefore the cliques over which the energy summation should be taken. See Chapter 2.

3. Formula 1.3 gives the relationship between the local characteristics of the MRF and the neighborhood-compatible potential of the Gibbs distribution. The proof of this relationship requires the use of the Moebius inversion formula [77], well-known in combinatorics.

The practical value of the Hammersley-Clifford theorem is that it gives us a systematic way for constructing Markov random fields. This is usually done by choosing the family of clique potentials corresponding to the neighborhood system defined on $\mathcal{S}$. The "if" part of the theorem is also of practical use. Indeed, if we are given a set of local characteristics, then we can use (1.3) to find whether these local characteristics are consistent, i.e., define a unique joint probability distribution of which they are the local conditional probability distributions.

A generic example that we will deal with in this thesis is that of a system with at most quadratic interactions. In this case, only single-site and two-site cliques have non-zero potentials. It follows that the Gibbs energy of this system can be written as

$$E(\mathbf{x}) = \sum_{s \in \mathcal{S}} V_s(x_s) + \sum_{\{s,r\} \in \mathcal{C}_2} V_{\{s,r\}}(x_s, x_r), \tag{1.4}$$

where $\mathcal{C}_2$ is the set of two-site cliques of the graph $(\mathcal{S}, \mathbf{N})$. A typical case is when the set $\mathcal{S}$ is a square lattice of size $N \times N$ with four nearest-neighbor neighborhoods and periodic boundary conditions. Then $\mathcal{S} = \{s = (i,j) | 1 \leq i,j \leq N\}$, and we have

$$E(\mathbf{x}) = \sum_{i,j=1}^{N} V_{ij}(x_{ij}) + \sum_{i,j=1}^{N} [V_{ij;i+1,j}(x_{ij}, x_{i+1,j}) + V_{ij;i,j+1}(x_{ij}, x_{i,j+1})] \tag{1.5}$$

The first term in (1.5) can be considered the energy contribution of an external field applied to the lattice, while the second term is the energy contribution of the interactions between the lattice pixels. This situation is very common in the area of image processing and computer vision, where the external field is an input data term obtained through

an imaging sensor, while the internal field describes the interactions in the *a priori* image model. Tasks like image segmentation, motion estimation, and halftoning, can be modeled within the MRF framework using the above energy function with the external and internal potentials chosen appropriately.

It is important to note that the summation in (1.4) and (1.5) is over cliques. An equivalent expression can be obtained by considering potential functions indexed by ordered pairs (s,r) such that $V_{sr} = V_{rs} = \frac{1}{2}V_{\{s,r\}}$. Then we can write (1.4) as

$$E(\mathbf{x}) = \sum_{s\in\mathcal{S}} V_s(x_s) + \sum_{s\in\mathcal{S}}\sum_{r\in\mathcal{N}_s} V_{sr}(x_s, x_r), \tag{1.6}$$

where the second summation is now taken over all the neighbors of all the pixels.

The models in which only two-site interactions are considered were called automodels by Besag [6], and they include famous models in statistical physics like the binary Ising model and the multilevel Potts model as well as models originally introduced by Besag himself like the autobinomial model.

Once an energy function with potentials compatible with the neighborhood structure is defined, we can define the positive Gibbs distribution as in Definition 4.

We end this section by giving one more notation. Let P be a probability distribution on Ω, and let $s_1, s_2, \ldots, s_m$ be m sites in $\mathcal{S}$. Then the m-th order moment of the random variables $X_{s_1}, X_{s_2}, \ldots, X_{s_m}$ with respect to P is given by

$$R(s_1, s_2, \ldots, s_m) = \sum_{\mathbf{x}\in\Omega} x_{s_1} x_{s_2} \ldots x_{s_m} P(\mathbf{x}). \tag{1.7}$$

The first moment is simply the mean $E[X_s]$ of the random variable X_s, while the second-order moment is the correlation $E[X_s X_r]$ between the pixel values at sites s and r. Note that for the Gibbs distribution, the computation of any of these moments requires the analytical or numerical computation of the partition function – a generally difficult task.

1.3 Contributions

As was mentioned at the beginning of this chapter, many fundamental questions pertaining to the properties of Gibbs random fields still need to be addressed. In this thesis we will deal mainly with the first, i.e., the question related to the moments (mean, correlation, etc.) of the Gibbs random field, and only tangentially with the second, i.e, the difficult question of parameter estimation.

Our two primary goals are the following. First, we want to provide approximations for the mean and the correlation function for a large class of Gibbs distributions, namely those having a finite number of graylevels. Second, we want to show how analog reciprocal networks can be derived to compute the approximate mean function.

The class of Gibbs models we are dealing with in this thesis falls somewhere in between the binary Ising model of statistical physics and the Gaussian model of signal processing. It is well known that the mean and correlation functions are not sufficient to characterize a non Gaussian random signal. However, in engineering applications, the knowledge of these functions can be very useful to answer important signal estimation questions. The usual approach to getting these functions in the case of Gibbs random fields has been the use of Monte Carlo simulations. Using the analytical machinery of statistical physics, we show the following:

1. An estimate of the mean signal for a large class of unconstrained Gibbs random fields with a finite number of graylevels can be obtained as the global minimum of a cost function – the *effective energy.*

2. An estimate of the correlation function can be obtained from the mean estimate using a simple *exact* relationship between the mean of a Gibbs distribution and its covariance.

3. We define iterated-map dynamical systems to find the approximate mean of a Gibbs distribution. We conduct an in-depth study of the long term-behavior of such maps.

4. We show that these iterated-map systems are the discrete-time analog counterparts of discrete-time, discrete-signal dynamical systems that seek for the *modes* of the Gibbs distribution. These latter systems can be construed as a generalization of the binary Hopfield network [35].

5. We define continuous-time, gradient-descent dynamical systems to find the approximate mean of a Gibbs distribution. One of these gradient-descent dynamical systems can be construed as a generalization of the analog Hopfield network [36].

6. Finally, we conduct a stability analysis of one of our gradient-descent systems and use it to derive closed-form estimates for the critical temperatures of a gray-level GRF.

In summary, starting from a probabilistic model, a Gibbs *random field*, we have shown that a deterministic *network* can be designed to find an estimate of its mean field. [5] This

[5]Hence, the title of the thesis.

mean-field estimate can in turn be used to find an estimate of the correlation field.

The results described above can also be interpreted from an entirely different viewpoint – that of the design of *analog* networks for solving *discrete* optimization problems. To any cost function defined on a finite set, one can assign a Gibbs probability distribution. Computing the mode of this probability distribution is equivalent to globally minimizing the cost function. As it was mentioned earlier in this Chapter, one can compute the mode by first computing the mean as a function of temperature and then finding the limit of the mean as the temperature approaches 0. The paradigm used to find a deterministic network to estimate the mean field can thus be used to solve the discrete optimization problem using analog systems. The so called mean-field approximation can therefore be construed as a *digital-to-analog* theory converter.

1.4 Overview

This thesis is organized as follows. Chapter 2 is devoted to surveying what is known in the literature about the computation of the correlation function when the Gibbs random field is either Gaussian or Ising. We found it useful to introduce the notion of a Gaussian reciprocal process in the one-dimensional case, since it is this notion, and not that of the Gauss-Markov process, that can be generalized to the multidimensional case. For both the one-dimensional and two-dimensional case, the correlation function satisfies a finite-difference equation. The correlation function of one-dimensional Ising model is derived rigorously and all details are given, but for the two-dimensional model, we only give the major result obtained by Onsager in 1942 and stress the fundamental difference between the one-dimensional case and the two-dimensional case, and that is the existence of a finite, non-zero limit for the two-dimensional correlation function as the distance between pixels approaches infinity. This is the essence of the phenomenon of phase transition. At the end of Chapter 2, we prove the "grayscale" version of the *linear response theorem* which allows the derivation of the correlation function of a Gibbs distribution from its mean function.

Thus the problem of approximating the correlation function is reduced to finding an approximation of the mean function. Chapter 3 is precisely devoted to the question of approximating the mean for unconstrained *and* constrained Gibbs random fields with a finite number of gray levels. We try in our treatment to make explicit the different assumptions needed to carry out the algebraic manipulations. We also point out where we believe more work is required to make some of the heuristics more rigorous. The highlights of the third chapter are:

1. In the unconstrained case, a simple derivation of the mean-field approximation that does not require the use of an integral representation of the partition function.

2. In the constrained case, we develop a set of mean-field equations using a probability decomposition approach. Here also, an integral representation of the partition function is not required.

Once the mean estimates are derived in Chapter 3, we devote Chapter 4 to show how these mean estimates can be used to obtain correlation estimates. After developing a general procedure for getting these correlation estimates we turn to the quadratic interaction model and derive a new closed-form approximation for its covariance function.

The mean estimates are generally solutions of fixed-point equations. In Chapter 5, we propose iterated-map methods to find these fixed points. First it is shown that the mean-field iterated maps are the analog counterparts of discrete-state iterated maps that try to compute the fixed points of a winner-take-all (WTA) network. The discrete iterated-maps are shown to possess Lyapunov functions that can be related to the Gibbs energy, while the analog iterated-maps are shown to possess Lyapunov functions that can be related to the effective energy. We also derive a number of useful results concerning the convergence and oscillations of both the discrete-state and analog iterated maps. This chapter has also an in-depth study of the *generalized sigmoid mapping* which plays, in the context of WTA networks, a role similar to that of the sigmoid function in the Hopfield neural network. In particular, it is shown that this mapping has a simple, exact VLSI circuit implementation with desirable circuit-theoretic properties.

One problem with the analog iterated-map dynamics is that it can have an oscillation of period 2. In Chapter 6, we show the existence of gradient-descent dynamical systems whose equilibrium points are the fixed points of the analog WTA network. The results that we have described so far do not depend on the temperature parameter. In the context of the iterated-map and gradient descent dynamical system, the temperature plays the role of a bifurcation parameter. This role is investigated in this chapter, and new closed-form estimates for the *critical temperatures* of a large class of Gibbs models used in texture synthesis are derived.

In Chapter 7, we summarize our findings and give indications about possible research directions in image modeling, neural networks, and analog optimization.

2

Correlation Functions of Gaussian and Binary Gibbs Distributions

Dangereuses images qui sont là, comme partout, les substitus d'un savoir!

Gaston Bachelard

This chapter deals with the important issue of determining the correlation function of a Gibbs distribution. Gibbs random fields and their counterparts, Markov random fields, have been used extensively for the past decade as a priori image models for image processing and computer vision tasks [31, 84]. The image or spatial scene in these models is the realization of a random field defined through local dependence characteristics. In the Gibbs case these characteristics are defined by neighborhood-based energy interactions, while in the Markov case they are given by neighborhood-based conditional probabilities. When the neighborhood system is given and *no global constraints are imposed on the image*, the Markov and the Gibbs representations are equivalent. This is the famous Hammersley-Clifford theorem [77].

Despite the extensive use of these models, a number of their structural and statistical properties that are important for image modeling and processing remain unknown. Among the structural properties, we can mention the dependence of the model realizations on the functional form of the local interactions, their neighborhood size and their parameters, whether these interactions are expressed through local conditional probabilities as in the Markov representation or the interaction energies as in the Gibbs rep-

resentation. Understanding this dependence is important for model realization, model identification, and model parameter estimation, and it has been the object of steady research since the seminal paper of Besag [6] and the texture analysis and synthesis study of Cross and Jain [14].

Among the statistical properties of these models, the moments of the joint probability distribution of the family of random variables defining the random field are the most important. Already in 1980, Hassner and Sklansky [34] have mentioned the difficulty of analytically computing the statistical moments of Markov random field models as one of their major shortcomings. This fact sets these models apart from other more tractable models like the simultaneous autoregressive models (SAR) [6] or the mosaic models [1] for which it is possible to compute the correlation function between pixels located at an arbitrary distance on the image grid.

This chapter is precisely devoted to investigating this difficult question. Of course, there is no pretension that the answer is contained herein, for we believe there is no general, exact answer. But we have tried to show that for some specific models, it is possible to obtain the correlation function of the joint probability distribution. We have also hinted at the possibility of computing *approximate* correlation functions for other models – a hint that will be fully developed in Chapter 4.

There are two specific Markov random field models whose correlation functions are, to a large extent, known *even* in the two-dimensional case. The first is a model in which the pixel state space, i.e., the pixel graylevel, belongs to the real line, and where the local, neighborhood-based conditional dependence of the color of a given pixel on its neighbors is given by a Gaussian probability distribution. The second model is the famous Ising model of statistical physics that can be used as an *a priori* model for binary images (only two colors are allowed for each pixel.) For each of these models we start by introducing the one-dimensional version and compute its correlation function. Then we treat the two-dimensional case and show how the correlation function can be obtained. We hope that this method of exposition will make clear the similarities and the differences between the one-dimensional and the two-dimensional versions of each model. The case of the Ising model, which is the simplest one can think of for implementing local dependence between pixels, shows that there is indeed a big gap between the correlation structure of the one-dimensional case and the two-dimensional case.

It is worthwhile to note that although the Gaussian Gibbs model seems to be analytically tractable, its problems have not been exhausted yet *even* in the one-dimensional case! And it is still the subject of active research [49, 61].

Section 2.1 of this chapter will be devoted to a quick survey of the contexts in which

the correlation function of the image model is needed. These contexts, which are very common in image modeling and processing, justify by themselves the importance of considering the correlation question for Markov and Gibbs image models.

2.1 Need for Correlation Functions

2.1.1 Image Coding

It is well known from transform coding theory [40] that the Karhunen-Loeve (KL) transform is the optimum image transform coder with respect to all decorrelating transforms as well as all unitary transforms. The criterion of optimality is that of minimum average mean-square distortion [1] between the original data and the reproduced data. This minimum mean-square error criterion leads in a natural way to considering the correlation matrix of the image model. As it turns out, the KL transform diagonalizes this correlation matrix and packs the image energy into few leading eigenvalues that are coded and transmitted.

The information provided by the correlation matrix of the image model is also used for determining the distortion rate as well as for studying the effect of the block size on coder performance [40]. In these analyses, it is usually assumed that the correlation function exhibits an exponential decay *to zero* with the distance between pixels. This kind of assumption is of course inspired by that of the one-dimensional causal Gauss-Markov process.

2.1.2 Retinal Information Processing

In their attempts at formulating a coherent theory for early vision processing by mammals, some vision scientists [27, 3] have recently proposed an information-theoretic framework for deriving the spatiotemporal transfer function of retinal signal processing. This approach, which is based on the earlier work of [4], postulates that the goal of retinal processing is to produce the most "efficient" code for transmission down the optic nerve, considered a noisy channel with a given capacity. A crucial point in the application of such a theory to deriving the retinal transfer function is of course that of the statistical nature of the information source. Atick [3] assumed that for mammals this source should be the

[1] The word **average** refers to a *sample* average over the size of the block, while the word **mean** refers to an *ensemble* average with respect to the probability distribution of the *a priori* image model.

ensemble of images of natural scenes equipped with a Gaussian probability distribution whose *correlation matrix* is the one found by Field [27].

Markov and Gibbs random fields have been used as *a priori* models for a variety of scenes including natural ones [34], but there has been no attempt at finding whether they can produce correlation functions similar to the ones found in [27].

2.1.3 Pattern Analysis

Natural textures are generally random, and therefore they lend themselves easily to a statistical description. In fact, two of the six natural images that were used in [27] to determine experimentally the correlation function of natural scenes were of natural textures. Among the statistical measures used to analyze and distinguish textures are second-order statistics. A famous conjecture due to Julesz stated that humans are unable to distinguish between textures having the same second-order statistics [25]. Although there are now many counterexamples to this conjecture, second-order statistics like the auto-correlation function [40] are still useful tools for computing the tonal coarseness and dispersion of textures and as a first step in a texture discrimination algorithm.

2.1.4 Pattern Synthesis

Many of the image modeling procedures are based on the idea of white noise filtering. The problem is the following: given a two-dimensional spectral density function (SDF), $S(z_1, z_2)$, or equivalently a correlation function, find a two-dimensional stable linear filter $A(z_1, z_2)$ that realizes such an SDF, i.e., such that

$$S(z_1, z_2) = \sigma_n^2 A(z_1, z_2) A(z_1^{-1}, z_2^{-1}),$$

where σ_n^2 is the input white noise variance. As in the one-dimensional case this is a problem of spectral factorization. Note that in the two-dimensional case there is no causality imposed on the linear filter. Relaxing the causality requirement however does not make the problem any easier to solve since in the two-dimensional case, a given two-dimensional polynomial might not have a decomposition into a product of lower order polynomials.

The usual engineering approach is to find a stable linear filter such that the above equation is satisfied approximately rather than exactly. This approximate representation of SDF is called the minimum variance representation and is studied in [39].

2.2 One-Dimesional Case: Gaussian Processes

In this section, we start looking at some of the widely used image models from the viewpoint of their correlation functions. Among these models, Gauss-Markov random fields occupy a special place. The introduction of these fields into the image processing literature came almost a decade [89] before systematic investigations of Markov random fields in image modeling and processing were undertaken [34, 14, 31].

2.2.1 Causal Case: Gauss-Markov Processes

In this section, we show how to get the correlation function of a zero-mean, homogeneous, first-order, Gauss-Markov process. Let $X_n, n \in \mathbf{Z}$, be such a process. Using the language of random fields that we introduced in Chapter 1, the lattice [2] $\mathcal{S}$ here is identified with $\mathbf{Z}$ and the spin state space is identified with $\Re$. The Gaussian property means that every finite collection $\{X_{n_1}, X_{n_2}, \ldots, X_{n_p}\}$ of random variables has a Gaussian probability distribution, $\forall p \in \mathbf{N}, n_1, n_2, \ldots, n_p \in \mathbf{Z}$. For every random variable, the expectation $E[X_n] = 0$. The first order Markov property means that

$$p_{n+1}(x_{n+1}|x_n, x_{n-1}, \ldots, x_{n-k}) = p_{n+1}(x_{n+1}|x_n),\ \forall k, n \in \mathbf{Z},$$

where $p_{n+1}(x_{n+1}|x_n, \ldots, x_{n-k})$ is the conditional probability density of X_{n+1} given the past random variables $X_n, \ldots, X_{n-k}$. Using an image that is now part of the standard folklore in probability, we can interpret the above condition as saying that in a first order Markov process the future X_{n+1} conditioned on the past $\{X_{n-1}, \ldots, X_{n-k}\}$ and the present $\{X_n\}$ depends only on the present. We call this property the one-sided Markov property. The reason will become clear in the sequel. The homogeneity property means that the conditional probability density $p_n(\cdot|\cdot)$ is independent of n. We denote the common conditional density by $p(\cdot|\cdot)$. It is worthwhile to note that a Markov process could be homogeneous but not stationary [65].

For this random process, we are interested in computing the correlation function $E[X_m X_n]$ for any pair (X_m, X_n). This computation is rather difficult in the general case. But because of the Gauss-Markov nature of our process it is possible to write down a "state-space" representation that makes the computation quite easy. This can be done as follows. From the Markov condition, we get for the mean of X_n conditioned on the past $\{X_{n-l}, l \geq 1\}$,

$$E[X_n|X_{n-1}, \ldots, X_{n-k}] = E[X_n|X_{n-1}].\ \forall k \geq 1.$$

[2] Unlike the rest of the thesis, the lattice of the Gaussian case will be assumed infinite.

Moreover, because the X_{n-1} and X_n are jointly Gaussian, the conditional mean of X_n given X_{n-1} is also the best linear estimate of X_n based on X_{n-1}, i.e., $E[X_n|X_{n-1}]$ is proportional to X_{n-1}, and there exists α_n such that $E[X_n|X_{n-1}] = \alpha_n X_{n-1}$. To compute α_n, we use the standard conditional estimation property that the estimation error $Z_n = X_n - \alpha_n X_{n-1}$ is uncorrelated with the data. In other words,

$$E[Z_n X_{n-1}] = E[(X_n - \alpha_n X_{n-1})X_{n-1}] = 0,$$

which gives

$$\alpha_n = \frac{E[X_{n-1}X_n]}{E[X_{n-1}^2]}.$$

Introducing the standard deviations σ_n and correlation coefficients $\rho_{m,n}$, we can write

$$X_n = \frac{\sigma_n}{\sigma_{n-1}}\rho_{n-1,n}X_{n-1} + Z_n.$$

Because of the Markov property, not only is the error Z_n uncorrelated with X_{n-1} but also it is uncorrelated with all the $X_{n-k}, k \geq 1$. Therefore by multiplying the above equation by $X_m, m < n$, and taking the expectations of both sides, we get

$$\rho_{m,n} = \rho_{m,n-1}\rho_{n-1,n}.$$

Applying the same equality to $\rho_{m,n-1}$ and repeating, we get

$$\rho_{m,n} = \prod_{k=m}^{n-1} \rho_{k,k+1}.$$

If we now make the additional assumption that the process is *stationary*, i.e., for the Gaussian process, both the first and second order statistics are independent of n, then we obtain

$$\rho_{m,n} = \rho^{n-m}, \ m < n.$$

It follows that for the stationary, zero-mean, first-order, Markov-Gauss process, the correlation function is given by

$$E[X_m X_n] = \sigma^2 \rho^{|n-m|}, \ \forall m, n \in \mathbf{Z},$$

where $0 < |\rho| \leq 1$.

It is clear from the above expression that the correlation function of the stationary Gauss-Markov process exhibits an exponential decay versus the distance between the sites (unless of course $|\rho| = 1$.)

One important property of the error sequence in the 1D Gauss-Markov process is that it is uncorrelated, i.e., $E[Z_m Z_n] = 0$, and since the error sequence is Gaussian, the errors

are independent of each other. Moreover, in the stationary case, $E[Z_n^2] = \sigma^2(1-\rho^2)$. Therefore, the error sequence is white Gaussian of variance equal to $\sigma^2(1-\rho^2)$. It will be seen that this is no longer the case for the *noncausal* Gauss-Markov model: the error sequence *is* correlated.

2.2.2 Noncausal Case: Gaussian Reciprocal Processes

As was mentioned in Chapter 1, images are inherently noncausal, in the sense that there is no preferred direction for scanning the image. To use the poetic imagery of Markov chains, we can say that in an image, there is no past, no present and no future. However, each closed region in the image has an inside, a boundary and an outside, and its around these geometric concepts that a 2D model for conditional dependence should be constructed. In fact, the geometric ideas of inside, boundary, and outside can even be applied to the real line $\Re$ or to the set of integers $\mathbf{Z}$. Consider a zero-mean, Gaussian process $X_n, n \in \mathbf{Z}$. Then this process is said to be *reciprocal* [50] or *quasi-Markov* [10] if the conditional densities satisfy the following condition

$$p_n(x_n|x_{n-k},\ldots,x_{n-1},x_{n+1},\ldots,x_{n+l}) = p_n(x_n|x_{n-1},x_{n+1}),\ \forall n \in \mathbf{Z},\ k,l \geq 1. \qquad (2.1)$$

This condition means that the conditional probability *inside* the interval $]n-1,n+1[\subset \mathbf{Z}$ given the configurations *outside* the interval are equal to the conditional probability *inside* given only the configurations at the *boundary*. The "inside" set here is just one point $\{n\}$, while the "boundary" has two points $\{n-1,n+1\}$. Here also we will assume that the process is homogeneous so that the conditional density is independent of n. Note that the above definition is similar to a two-sided Markov property in which we look at the configurations on both sides of the considered site. There are two important questions that one should ask here. The first is about the relationship between the standard Markov property, which is a one-sided property, and the two-sided Markov property of the reciprocal Gaussian process. The second question is whether it is possible to characterize all the correlation functions of the stationary reciprocal Gaussian processes the same way we are able to characterize the correlation functions of the stationary Gauss-Markov processes.

Correlation Structure

The answer to the first question is that the one-sided Markov property implies the two-sided Markov property. The converse is however not true, i.e., there are reciprocal, stationary Gaussian processes that are not Gauss-Markov. Let us first state the following

Proposition 1 *If $X_n, n \in \mathbf{Z}$, is a Markov process then it is reciprocal.*

Before giving the proof, let us note that the theorem does not make the assumption that the process is Gaussian. Indeed, the proof, as will be shown, uses only the Markovian assumption. To simplify notation, let us denote by $\mathbf{l}$ and $\mathbf{r}$ the two families of variables $\{x_{n-k}, \ldots, x_{n-2}\}$ and $\{x_{n+2}, \ldots, x_{n+l}\}$, respectively. To prove Proposition 1, we need Lemmas 1 and 3 below. Lemma 1 gives an equivalent definition for a reciprocal process while Lemma 3 is an interesting, albeit little-known, conditional independence property for Markov chains.

Lemma 1 *The process $X_n, n \in \mathbf{Z}$, is reciprocal if and only if*

$$p(\mathbf{l}, x_n, \mathbf{r}|x_{n-1}, x_{n+1}) = p(x_n|x_{n-1}, x_{n+1})p(\mathbf{l}, \mathbf{r}|x_{n-1}, x_{n+1}). \tag{2.2}$$

Proof: Assume that condition (2.2) is satisfied. Then we have

$$\begin{aligned} p(x_n|\mathbf{l}, x_{n-1}, x_{n+1}, \mathbf{r}) &= \frac{p(\mathbf{l}, x_{n-1}, x_n, x_{n+1}, \mathbf{r})}{p(\mathbf{l}, x_{n-1}, x_{n+1}, \mathbf{r})} \\ &= \frac{p(\mathbf{l}, x_n, \mathbf{r}|x_{n-1}, x_{n+1})}{p(\mathbf{l}, \mathbf{r}|x_{n-1}, x_{n+1})} \text{ (definition of conditional densities)} \\ &= \frac{p(x_n|x_{n-1}, x_{n+1})p(\mathbf{l}, \mathbf{r}|x_{n-1}, x_{n+1})}{p(\mathbf{l}, \mathbf{r}|x_{n-1}, x_{n+1})} \text{ (by condition 2.2).} \end{aligned} \tag{2.3}$$

It follows that the process is reciprocal. Conversely assume that the process is reciprocal. Then

$$\begin{aligned} p(\mathbf{l}, x_n, \mathbf{r}|x_{n-1}, x_{n+1}) &= \frac{p(\mathbf{l}, x_{n-1}, x_n, x_{n+1}, \mathbf{r})}{p(x_{n-1}, x_{n+1})} \\ &= \frac{p(x_n|\mathbf{l}, x_{n-1}, x_{n+1}, \mathbf{r})p(\mathbf{l}, x_{n-1}, x_{n+1}, \mathbf{r})}{p(x_{n-1}, x_{n+1})} \\ &= p(x_n|x_{n-1}, x_{n+1})p(\mathbf{l}, \mathbf{r}|x_{n-1}, x_{n+1}), \end{aligned} \tag{2.4}$$

where the last equality resulted from the definitions of a reciprocal process and conditional densities. ■

In plain English, the above lemma is saying that the process is reciprocal if given the conditions on the boundaries the inside and outside are independent. This condition should be contrasted against the following known property ([65], p. 386) for one-sided Markov processes. The property is easy to prove and is stated in terms of a continuous-state Markov process.

Lemma 2 *Let $X_n, n \in \mathbf{Z}$, be a one-sided, continuous-state Markov process. If $k < m < n$, then*

$$p(x_k, x_n|x_m) = p(x_k|x_m)p(x_n|x_m). \tag{2.5}$$

This lemma states that given the present, the past and the future of a one-sided Markov process are independent. Note how the concepts of past, present and future of Markov processes were replaced by the concepts of inside, boundary, and outside for reciprocal processes. These latter concepts make the reciprocal process a better candidate for generalization to multidimensional cases. The next lemma is also concerned with the conditional independence of past and future in Markov processes.

Lemma 3 *Let $X_n, n \in \mathbf{Z}$, be a Markov process. Then we have*

$$p(\mathbf{l}, \mathbf{r}|x_{n-1}, x_{n+1}) = p(\mathbf{l}|x_{n-1})p(\mathbf{r}|x_{n+1}). \tag{2.6}$$

Proof:

$$\begin{aligned} p(\mathbf{l}, \mathbf{r}|x_{n-1}, x_{n+1}) &= \frac{p(\mathbf{l}, x_{n-1}, x_{n+1}, \mathbf{r})}{p(x_{n-1}, x_{n+1})} \\ &= \frac{p(\mathbf{l}, x_{n+1}, \mathbf{r}|x_{n-1})}{p(x_{n+1}|x_{n-1})} \\ &= \frac{p(\mathbf{l}|x_{n-1})p(x_{n+1}, \mathbf{r}|x_{n-1})}{p(x_{n+1}|x_{n-1})} \\ &= p(\mathbf{l}|x_{n-1})p(\mathbf{r}|x_{n+1}, x_{n-1}) \\ &= p(\mathbf{l}|x_{n-1})p(\mathbf{r}|x_{n+1}) \end{aligned}$$

The third equality results from Lemma 2 and the last one in nothing but the Markov property. ■

With Lemmas 1 and 3 in hand, we can now prove Proposition 1.

Proof: We want to prove that if X_n is Markov process then the condition given by Lemma 1 is satisfied. Thus we start with the left-hand side of (2.2) and prove, using the Markov condition, that (2.2) is satisfied.

$$\begin{aligned} p(\mathbf{l}, x_n, \mathbf{r}|x_{n-1}, x_{n+1}) &= \frac{p(\mathbf{l}, x_{n-1}, x_n, x_{n+1}, \mathbf{r})}{p(x_{n-1}, x_{n+1})} \\ &= \frac{p(\mathbf{l}, x_{n-1}, x_n, \mathbf{r}|x_{n+1})}{p(x_{n-1}|x_{n+1})} \\ &= \frac{p(\mathbf{l}, x_{n-1}, x_n|x_{n+1})}{p(x_{n-1}|x_{n+1})} p(\mathbf{r}|x_{n+1}), \end{aligned}$$

where the last inequality results from the one-sided Markov property. Now consider the first term in the expression above. Then we have

$$\begin{aligned} \frac{p(\mathbf{l}, x_{n-1}, x_n|x_{n+1})}{p(x_{n-1}|x_{n+1})} &= \frac{p(\mathbf{l}, x_{n-1}, x_n, x_{n+1})}{p(x_{n-1}, x_{n+1})} \\ &= \frac{p(\mathbf{l}, x_n, x_{n+1}|x_{n-1})}{p(x_{n+1}|x_{n-1})} \end{aligned}$$

$$
\begin{aligned}
&= \frac{p(\mathbf{l}|x_{n-1})p(x_n, x_{n+1}|x_{n-1})}{p(x_{n+1}|x_{n-1})} \\
&= p(\mathbf{l}|x_{n-1})p(x_n|x_{n-1}, x_{n+1}),
\end{aligned}
$$

where the third equality results from Lemma 2 and the last one follows from

$$
\frac{p(x_n, x_{n+1}|x_{n-1})}{p(x_{n+1}|x_{n-1})} = \frac{p(x_n, x_{n+1}, x_{n-1})}{p(x_{n+1}, x_{n-1})} = p(x_n|x_{n+1}, x_{n-1}).
$$

Assembling, we get

$$
p(\mathbf{l}, x_n, \mathbf{r}|x_{n-1}, x_{n+1}) = p(\mathbf{l}|x_{n-1})p(\mathbf{r}|x_{n+1})p(x_n|x_{n-1}, x_{n+1}).
$$

The desired result is obtained by applying Lemma 3. ∎

The converse of Proposition 1 is not true. In other words, there are reciprocal Gaussian processes that are not Gauss-Markov. This fact is in sharp contrast to the case where the state space of the Markov chain is *finite.* Then one can prove using the theory of Markov random fields that there is in fact an equivalence between a one-sided Markov process and a double-sided Markov process [30]. To see why the converse of Proposition 1 is not true, we need to study the correlation structure of Gaussian reciprocal process and show that there are Gaussian reciprocal processes having correlation functions that are *not* of exponential decay. We shall proceed as for the Gauss-Markov case and exploit the optimal estimation properties of Gaussian processes. We will consider zero-mean processes and denote by $R(m,n) = E[X_m X_n]$ the correlation function of the process X_n. When this process is stationary $R(m,n) = R(n-m)$.

Proposition 2 *Let $X_n, n \in \mathbf{Z}$, be a zero-mean, stationary Gaussian reciprocal process, such that $R(0) = 1$. Denote by $a = R(1)$ and $b = 1 + R(2)$. Then the correlation function satisfies the difference equation*

$$
aR(l+2) - bR(l+1) + aR(l) = 0, \ \forall l \in \{\mathbf{Z} - (-1)\} \tag{2.7}
$$

Proof: First we will treat the case $|R(2)| \neq 1$. For $l > 2$, consider the random variables $X_{n-l}, X_{n-1}, X_n, X_{n+1}$. Then by the reciprocal property,

$$
p(x_n|x_{n-l}, x_{n-1}, x_{n+1}) = p(x_n|x_{n-1}, x_{n+1}).
$$

It follows that

$$
E[X_n|X_{n-l}, X_{n-1}, X_{n+1}] = E[X_n|X_{n-1}, X_{n+1}].
$$

Since the process is Gaussian, the conditional mean is a linear combination of the data, i.e., $\exists\ u, v \in \Re$ such that

$$
E[X_n|X_{n-1}, X_{n+1}] = uX_{n-1} + vX_{n+1}.
$$

Moreover the "estimation error", or the residual, given by

$$Z_n = X_n - E[X_n|X_{n-1}, X_{n+1}] = X_n - uX_{n-1} - vX_{n+1}$$

is uncorrelated with the data. In other words,

$$E[Z_nX_{n-1}] = E[Z_nX_{n+1}] = E[Z_nX_{n-l}] = 0, \ \forall l \neq 0.$$

The above three equations along with the expression of Z_n give

$$\begin{aligned} R(1) - uR(0) - vR(2) &= 0 \\ R(1) - uR(2) - vR(0) &= 0 \\ R(l) - uR(l-1) - vR(l+1) &= 0, \ \forall l \neq 0. \end{aligned}$$

Solving the first two equations for u and v gives

$$u = v = \frac{R(1)}{R(0) + R(2)} = \frac{a}{b},$$

which upon substitution in the last equation gives

$$aR(l+1) - bR(l) + aR(l-1) = 0, \ \forall l \neq 0.$$

The difference equation (2.7) is obtained by replacing l with $l+1$.

Now consider the case $|R(2)| = 1$. Since $R(0) = [R(0)]^2 = 1$, the correlation matrix of the Gaussian vector $(X_0, X_2)^T$ has a determinant equal to $1 - [R(2)]^2 = 0$, i.e., we can write $X_2 = cX_0$. Using the reciprocal property, we get that the error

$$Z_1 = X_1 - E[X_1|X_{-1}, X_0, X_2] = X_1 - E[X_1|X_0]$$

is uncorrelated with X_0, X_{-1} and X_2. The last equation is a result of the linear dependence between X_0 and X_2. Also, $E[X_1|X_0] = \kappa X_0$. Hence

$$E[Z_1X_0] = E[Z_1X_{-1}] = E[Z_1X_2] = 0.$$

We can then write for the first two equations

$$R(1) = \kappa \quad \text{and} \quad ; \ R(2) = \kappa R(1).$$

It follows that $R(2) = [R(1)]^2$, which means that $R(2) = 1$ and that $R(1) = \pm 1$. To obtain the correlation $R(n)$ it is sufficient to notice that the error Z_1 is uncorrelated with X_{-n+1} which gives the recurrent equation

$$R(n) = R(1)R(n-1).$$

Therefore

$$R(n) = [R(1)]^n,$$

and the correlation function in this case is either $R(n) = 1$ or $R(n) = (-1)^n$. ■

Remarks: Let us pause here for a while to give the following remarks:

1. The case when $|R(2)| = 1$ leads to a correlation function that is either 1 or $(-1)^n$. The latter situation is similar to the "uninteresting" $|\rho| = 1$ for the Gauss-Markov case.

2. In the case when $|R(2)| \neq 1$, we get that $u = v$. The parameters u and v can be understood as bonding parameters in the sense of Markov random field models. Their equality is a necessary outcome of the stationarity assumptions. This is in line with the observation made in [6] that the bonding parameters in a homogeneous Markov random field are symmetric.

3. Since every Gauss-Markov process is also a Gaussian reciprocal process, the correlation function of a zero-mean, stationary Gauss-Markov process such that $E[X_n^2] = 1$ always satisfies the difference equation (2.7). To check this, note that for a Gauss-Markov process we have $\rho = a, R(2) = \rho^2 = a^2$. Substituting into (2.7), we get the identity

$$a^3 - (1 + a^2)a + a = 0$$

as can be easily checked.

What makes the class of reciprocal processes larger than the class of Gaussian processes is that the difference equation (2.7) admits solutions other that the ones that lead to the exponentially decaying correlation of the Gauss-Markov process. Specifically, we have

Proposition 3 *If X_n, $n \in \mathbf{Z}$, is a zero-mean, stationary, reciprocal Gaussian process such that $E[X_n^2] = 1$, then its correlation function is one of the following*

1. $R(n) = \lambda^n$, $0 \leq |\lambda| \leq 1$;
2. $R(n) = \cos(\theta n)$, $\theta \in [0, 2\pi]$.

Proof: We start with the difference equation (2.7) and set $R(l) = z^l$. Then z must satisfy the quadratic equation

$$az^2 - bz + a = 0. \tag{2.8}$$

We have of course three cases:

1. $b^2 = 4a^2$. Then (2.8) has one solution $z_- = z_+ = \pm 1$. We set $\lambda = \pm 1$. Clearly the correlation function has Form 1 above.

2. $b^2 > 4a^2$. Equation (2.8) has two *real* solutions

$$z_{\pm} = \frac{1}{2a}(b \pm \sqrt{b^2 - 4a^2})$$

with $z_+ > 1$ and $z_- < 1$. Therefore

$$R(l) = Az_-^l + Bz_+^l.$$

Since $|R(l)| \leq 1$ and $R(0) = 1$, we have $A = 1$ and $B = 0$, i.e., posing $z_- = \lambda$ with $|\lambda < 1$, $R(l) = \lambda^l$, which has Form 1.

3. $b^2 < 4a^2$. In this case, (2.8) has two *complex* conjugate solutions ξ and $\bar{\xi}$ such that $|\xi| = 1$. We can then write

$$R(l) = A\exp(il\theta) + B\exp(-il\theta),$$

with $\tan(\theta) = \sqrt{\frac{4a^2}{b^2} - 1}$. Since $R(l)$ is real, we get

$$R(l) = C\cos(l\theta) + D\sin(l\theta).$$

Since $R(0) = 1$, we get $C = 1$. Moreover for the condition $|R(l) \leq 1$, $\forall l$, to be satisfied, we must have $D = 0$. It follows that $R(l) = \cos(l\theta)$, i.e, Form 2.

■

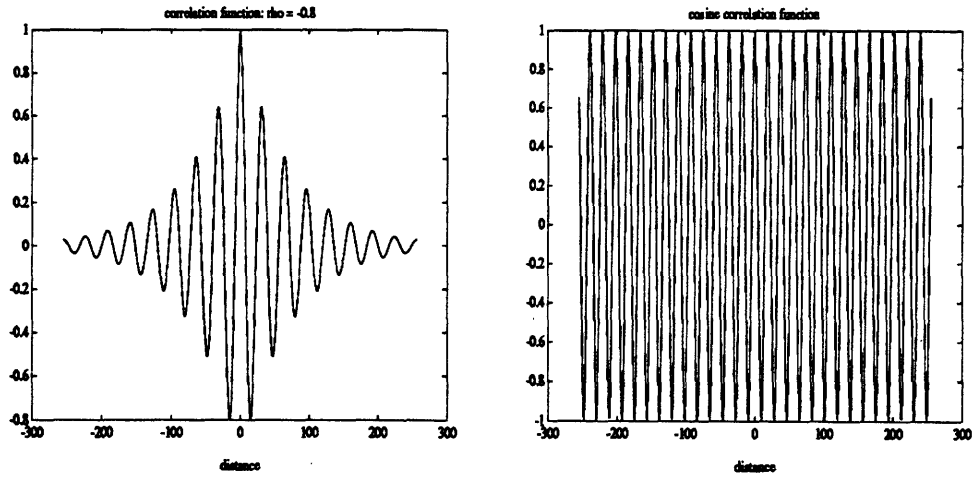
Note that if we impose the condition $|R(l)| < 1, l > 0$, then $R(n) = \lambda^n$, $0 < |\lambda| < 1$, or $R(n) = \cos(\theta n)$ with $\frac{\theta}{\pi}$ irrational.

The interest of the above proposition comes from the fact that the two-sided Markov property of reciprocal processes leads to a much richer class of correlation functions. Indeed, the almost periodic nature of $R(n) = cos(n\theta)$ with $\frac{\theta}{\pi}$ irrational contrasts sharply with the exponentially decaying correlation of Gauss-Markov processes as can be seen in Figure 2-1.

Noise Process

Beside this fundamental difference in the correlation functions, there is another difference related to the behavior of the error process $Z_n, n \in \mathbf{Z}$. While this process is independent, identically distributed, in the stationary Gauss-Markov case, it is correlated in the stationary, reciprocal, Gaussian case. In other words the autoregressive equations

$$X_n = uX_{n-1} + vX_{n+1} + Z_n,\ n \in \mathbf{Z}$$



(a) (b)

FIGURE 2-1: (a) Exponentially decaying correlation function. (b) Cosine correlation function for the Gaussian reciprocal process.

for a reciprocal Gaussian process are not driven by white noise. Note in particular the noncausal nature of these equations. It is possible to compute the correlation structure of the input sequence, $Z_n, n \in \mathbf{Z}$, as function of the autoregressive coefficients u, v. Indeed, when $|R(2)| \neq 1$, we can write the above equation as

$$Z_n = X_n - uX_{n-1} - uX_{n+1},\ n \in \mathbf{Z},$$

where we have used the fact that $v = u$ (see Proposition 3). Now Z_{n+1} is uncorrelated with X_{n-1}, X_n and X_{n+2}. Therefore

$$E[Z_n Z_{n+1}] = -uE[X_{n+1}Z_{n+1}] = -uE[Z_{n+1}^2].$$

Similarly

$$E[Z_n Z_{n-1}] = -uE[X_{n-1}Z_{n-1}] = -uE[Z_{n-1}^2].$$

Using the correlation properties between the error sequence Z_n and X_n we can also prove that

$$E[Z_n Z_{n+l}] = 0,\ \forall l \in \mathbf{Z},\ |l| > 1.$$

The above equations mean that unlike the Gauss-Markov case, the error sequence of the reciprocal process exhibits local (nearest-neighbor) correlations that are determined by the autoregressive parameter u and the variance of the noise $E[Z_n^2]$. An easy computation shows that

$$E[Z_n^2] = E[Z_n X_n] = 1 - 2uR(1).$$

2.2.3 Noncausal Autoregressive Process

This local correlation behavior of the noise sequence makes the reciprocal process inherently different from the general noncausal, autoregressive (AR) model [40], in which we have

$$X_n = \alpha_- X_{n-1} + \alpha_+ X_{n+1} + Z_n, \tag{2.9}$$

where Z_n is a sequence of zero-mean, identical, independently distributed random variables. In this class of models the autoregressive parameters α_+ and α_- can be different, i.e, the model is not realizable using a homogeneous, reciprocal process. The following proposition concerning the correlation function of a process described by (2.9) gives the reason.

Proposition 4 *The correlation function of the noncausal, autoregressive model (2.9) satisfies a 4th order difference equation.*

Proof: Denote by $R(l) = E[X_n X_{n+l}]$ the correlation function. From (2.9), we have

$$Z_n = X_n - \alpha_- X_{n-1} - \alpha_+ X_{n+1},$$

and using the fact the error sequence is zero-mean and that Z_n and $Z_{n+l}, l \neq 0$ are independent therefore uncorrelated, we get

$$\begin{aligned} E[Z_n Z_{n+l}] &= E\left[(X_n - \alpha_- X_{n-1} - \alpha_+ X_{n+1})(X_{n+l} - \alpha_- X_{n+l-1} - \alpha_+ X_{n+l+1})\right] \\ &= (1 + \alpha_-^2 + \alpha_+^2)R(l) - (\alpha_- + \alpha_+)R(l-1) - \\ &\quad (\alpha_- + \alpha_+)R(l+1) + \alpha_-\alpha_+ R(l-2) + \alpha_-\alpha_+ R(l+2) \\ &= 0. \end{aligned}$$

Replacing l with $l+2$ and rearranging we get for $R(l)$ the difference equation

$$\begin{aligned} \alpha_-\alpha_+ R(l+4) &- (\alpha_- + \alpha_+)R(l+3) + (1 + \alpha_-^2 + \alpha_+^2)R(l+2) \\ &- -(\alpha_- + \alpha_+)R(l+1) + \alpha_-\alpha_+ R(l). \end{aligned}$$

■

Note that the coefficients of the above equation are symmetric with respect to α_- and α_+. Let us denote by $\mu = 1 + \alpha_-^2 + \alpha_+^2, \nu = \alpha_- + \alpha_+$ and $\zeta = \alpha_-\alpha_+$. Then the difference equation above becomes

$$\zeta R(l+4) - \nu R(l+3) + \mu R(l+2) - \nu R(l+1) + \zeta R(l) = 0. \tag{2.10}$$

As for the reciprocal process, it is possible to compute the different forms for the correlation function $R(l)$. Indeed, The characteristic equation of the 4th order difference equation (2.10) is given by

$$\zeta z^4 - \nu z^3 + \mu z^2 - \nu z + \zeta = 0. \tag{2.11}$$

Note that the roots of this fourth-order equation are invariant under the transformation $z \rightarrow \frac{1}{z}$. It follows that if z is a root of (2.10) then $x = z + \frac{1}{z}$ is a root of the *quadratic* equation

$$4\zeta x^2 - 3\mu x + 2\nu - 3\zeta = 0. \tag{2.12}$$

For each root x of (2.12), there corresponds two roots of (2.12) given by

$$z = \frac{1}{2}(z \pm \sqrt{x^2 - 4}).$$

The different possibilities for the correlation function can be determined according to the discriminant of (2.12), the sign of $\frac{2\nu - 3\zeta}{\zeta}$, and the condition $|R(l)| \leq R(0)$, $l \in \mathbf{Z}$.

It is clear that the correlation structure of the noncausal, autoregressive model is richer than that of the Gaussian reciprocal process. Note also that in the former case, the parameters of the model need not be symmetric, i.e., we could have $\alpha_- \neq \alpha_+$. Moreover, the noise process of the model need not be Gaussian. In fact the above reasoning about the correlation function remains valid even if the independent random variables Z_n have different distributions.

The noncausal autoregressive models are often encountered when one discretizes boundary-value differential equations and drive the resulting difference equations with zero-mean white noise. For instance, the 1D Poisson equation would correspond to the case where $\alpha_- = \alpha_+ = \frac{1}{2}$. This aspect of the noncausal autoregressive models and its applications to noncausal smoothing problems is treated in [48].

Finally, it is worthwhile to mention that this richness in the correlation structure of the noncausal, autoregressive model is obtained *at the expense of a more complicated conditional probability structure.*

2.2.4 Conditional Probability of the Autoregressive Model

In order to explain this remark, [3] we will compute the conditional probability for the nearest-neighbor autoregressive model given in (2.9). A first look at (2.9) might lead us

[3]This section was formulated as an exam problem for 6.432, the core MIT course on stochastic processes, in Spring 1992. John Wyatt was lecturer.

to believe that the noncausal process defined by this equation is a (first-order) reciprocal process, since only the two nearest neighbors are involved in the definition of every random variable. This is actually wrong. The process is in fact *second-order* reciprocal, in the sense that the conditional probability of the variable X_n given *all* the remaining variables is equal to its conditional probability given the variables of its nearest-neighbors *and* next-nearest neighbors. First let us make the following simplifying assumptions:

1. Assume that the noise process is a zero mean, independent, identically distributed Gaussian process of unit variance.

2. Assume periodic boundary conditions, i.e, $X_{N+1} = X_1$. In other words, the process is defined on a circle instead of being defined on a line.

Our objective is to compute the conditional probabilities

$$p(x_n|x_m, 1 \leq m \leq N,\ m \neq n),\ 1 \leq n \leq N.$$

In order to do this, we introduce the following notations. We let $\mathbf{X}$ and $\mathbf{Z}$ denote the random vectors $(X_1, \ldots, X_N)$ and $(Z_1, \ldots, Z_N)$, respectively. Also, we denote by $\mathbf{x}_i$ the N-dimensional vector $(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_N)$ obtained from $\mathbf{x}$ by replacing the i-th component with 0.

With these notations, Equation (2.9) can be written in a linear system form as

$$\mathbf{BX} = \mathbf{Z}, \tag{2.13}$$

where the matrix $\mathbf{B}$ is given by

$$\mathbf{B} = \begin{pmatrix} 1 & -\alpha_+ & 0 & \ldots & 0 & -\alpha_- \\ -\alpha_- & 1 & -\alpha_+ & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & -\alpha_- & 1 & -\alpha_+ \\ -\alpha_+ & 0 & \ddots & 0 & -\alpha_- & 1 \end{pmatrix}$$

and is assumed invertible.

To find the conditional probabilities, we need first to compute the joint probability distribution $p(\mathbf{x})$.

Lemma 4 *The random vector* $\mathbf{X}$ *is Gaussian of covariance matrix equal to* $(\mathbf{B}^T\mathbf{B})^{-1}$, *so that we have*

$$p(\mathbf{x}) = \frac{|\mathbf{B}|}{(2\pi)^{N/2}} \exp\left[-\frac{1}{2}\mathbf{x}^T\mathbf{B}^T\mathbf{B}\mathbf{x}\right].$$

Proof: Since $\mathbf{B}$ is invertible, we can write for any N-dimensional vector $\mathbf{r}$, $\mathbf{r}^T\mathbf{X} = \mathbf{r}^T\mathbf{B}^{-1}\mathbf{Z}$, which is a Gaussian variable by the fact that $\mathbf{Z}$ is a Gaussian vector. It follows that $\mathbf{X}$ itself is a Gaussian vector. Moreover, the correlation of the Gaussian vector $\mathbf{X}$ is given by

$$E[\mathbf{X}\mathbf{X}^T] = (\mathbf{B}^T\mathbf{B})^{-1},$$

since $E[\mathbf{Z}\mathbf{Z}^T]$ is the identity matrix. The expression of $p(\mathbf{x})$ results from the fact that $\mathbf{X}$ is also zero-mean. ■

The following formula satisfied by the conditional probabilities will also be useful

Lemma 5 *Let $p(\mathbf{x})$ be a probability density such that $\forall \mathbf{x} \in \Re^N, p(\mathbf{x}) > 0$. Denote by*

$$Q(\mathbf{x}) = \ln\left[\frac{p(\mathbf{x})}{p(\mathbf{0})}\right].$$

Then

$$\frac{p(x_n|x_m, m \neq n)}{p(0|x_m, m \neq n)} = \frac{p(\mathbf{x})}{p(\mathbf{x}_n)} = \exp(Q(\mathbf{x}) - Q(\mathbf{x}_n)),$$

where 0 is the value taken by the variable at the n-th site.

Proof: To prove the above inequality, we use the definition of $Q(\mathbf{x})$ and the fact that

$$\begin{aligned} p(\mathbf{x}) &= p(x_n|x_m, m \neq n)p(x_1, \ldots, x_{n-1}, x_{n+1}, \ldots, x_N) \\ p(\mathbf{x}_n) &= p(0|x_m, m \neq n)p(x_1, \ldots, x_{n-1}, x_{n+1}, \ldots, x_N). \end{aligned}$$

■

Now we are ready to obtain the conditional probabilities $p(x_n|x_m, m \neq n)$. Applying the above lemma to the Gaussian distribution obtained in Lemma (4), we get

$$Q(\mathbf{x}) - Q(\mathbf{x}_n) = \frac{1}{2}\left[\|\mathbf{B}\mathbf{x}_n\|^2 - \|\mathbf{B}\mathbf{x}\|^2\right].$$

In the above expression, all the terms that do not contain x_n or its nearest neighbors x_{n-1} and x_{n+1} will cancel out, and we are left with the terms corresponding to $n-1, n$ and $n+1$. After some simple algebra we get

$$Q(\mathbf{x}) - Q(\mathbf{x}_n) = -(1 + \alpha_-^2 + \alpha_+^2)x_n^2 + 2(\alpha_- + \alpha_+)x_n(x_{n-1} + x_{n+2}) - 2\alpha_-\alpha_+ x_n(x_{n-1} + x_{n+2}).$$

In other words, we have for the conditional probability the expression

$$\begin{aligned} \frac{p(x_n|x_m, m \neq n)}{p(0|x_m, m \neq n)} &= \exp\{-x_n[(1 + \alpha_-^2 + \alpha_+^2)x_n \\ &+ 2(\alpha_- + \alpha_+)(x_{n-1} + x_{n+2}) - 2\alpha_-\alpha_+(x_{n-1} + x_{n+2})]\}. \end{aligned}$$

Note that the denominator is a real number whose value can be determined using the fact that the conditional probability integrates to one over the random variable values at one site. There is a number of interesting remarks that can be made.

1. The exponent in the conditional probability is invariant with respect to the transformation $\alpha_- \rightarrow \alpha_+$ and $\alpha_+ \rightarrow \alpha_-$. Writing

$$\begin{aligned} \sigma^2 &= (1 + \alpha_-^2 + \alpha_+^2)^{-1} \\ a &= \sigma^2(\alpha_- + \alpha_+) \\ b &= \sigma^2 \alpha_- \alpha_+ \end{aligned}$$

we get

$$\frac{p(x_n|x_m, m \neq n)}{p(0|x_m, m \neq n)} = \exp\left\{-\frac{x_n}{\sigma^2}\left[x_n + 2a(x_{n-1} + x_{n+2}) - 2b(x_{n-1} + x_{n+2})\right]\right\}.$$

It is clear from the above expression that there is symmetry in the parameters bonding site n to its neighbors and its nearest neighbors.

2. Although we started with an equation in which x_n *seemed* to depend only on x_{n-1} and x_{n+1}, the conditional probability depends not only on these sites but also on the next-nearest neighbors x_{n-2} and x_{n+2}. In other words, we can write

$$p(x_n|x_m, m \neq n) = p(x_n|x_m, m = n-2, n-1, n+1, n+2).$$

3. From the above expression, one can deduce that X_n, $1 \leq n \leq N$ is a second order one-dimensional Markov random field. The neighborhood structure of this field is the one defined not only by the nearest-neighbor but also by the next-nearest neighbor.

4. In order to get the best estimate of X_n given the remaining variables, we need to compute the expectation of X_n conditioned on the remaining variables. Completing the square in the exponent of the conditional probability, we get

$$\frac{p(x_n|x_m, m \neq n)}{p(0|x_m, m \neq n)} = \exp\left\{-\frac{1}{\sigma^2}\left[(x_n - m_n)^2 + m_n^2\right]\right\},$$

where

$$m_n = E[X_n|X_k, k \neq n] = -a(x_{n-1} + x_{n+1}) + d(x_{n-2} + x_{n+2}).$$

The question remains as to what it takes so that the noncausal autoregressive equation (2.9) defines a one-dimensional Markov field with nearest-neighbor neighborhood structure. We have the following proposition.

Proposition 5 *The periodic, noncausal, autoregressive process defined by equation (2.9) defines a first-order Markov random field if and only if the following two conditions are satisfied*

(i) The matrix $\mathbf{B}$ *is symmetric* ($\alpha_- = \alpha_+$) *positive, definite.*

(ii) The noise process $\mathbf{Z}$ *is correlated with a correlation matrix* $\mathbf{B}^{-1}$.

Proof: The proof is easy and is left as an exercise to keep the reader alert. ∎

The above observations are not new. In fact they were made by Besag in his 1974 paper [6]. The elementary one-dimensional treatment here should help the reader appreciate the tradeoff that exists between the MRF model order and the correlation structure of the noise sequence. Moreover, it should make more apparent the difference between models based on conditional probabilities and noise-driven models.

2.3 Two-Dimensional Case: Gauss-Markov Random Fields

We have seen in Section 2.2.2, that in the 1D case, the Gaussian noncausal models exhibit a richer correlation structure than the causal ones. Whether the correlation structure has exponential decay or not depends on the parameters of the autoregressive equation representing the model. We have also seen that within the Markov context, it is the (noncausal) reciprocal process that is the better candidate for generalization to the multidimensional case.

In order to avoid dealing with the mathematically tricky infinite lattice (see Appendix A) or the cumbersome boundary conditions of the rectangular lattice, we will assume throughout that the lattice is toroidal.[4] In order to avoid complex notations, we restrict ourselves to the nearest-neighbor neighborhood structure on a toroidal lattice $\mathcal{S}$ of size $|\mathcal{S}| = N^2$. [5]

If $A \subset \mathcal{S}$, we denote by $\mathbf{X}_A$ the random vector describing the configurations of lattice sites in A. The boundary of A is a subset of $\mathcal{S}$, denoted ∂A, and defined by

$$\partial A = \{s \in \mathcal{S} - A |\ \exists r \in A, s \in \mathcal{N}_r\}.$$

The boundary site configurations are described by the random vector $\mathbf{x}_{\partial A}$.

The Gauss-Markov random field is introduced using conditional probabilities. Later in this section we show how it can be written as a Gibbs distribution with appropriately

[4]Only very recently did researchers start considering finite lattices with non-toroidal boundary conditions for the Gaussian case [48, 61].

[5]The neighborhood assumption is not as restrictive as it may sound. Indeed, if we extend the pixel state space by considering pixel variables in $\Re^d$, $d > 1$, then higher order neighborhood models can be treated as nearest-neighbor models as well [48].

chosen local energy interactions. The following definition develops the concept of inside-outside mentioned in Section 2.2.2 for the reciprocal process.

Definition 5 *A Gaussian random field $X_s, s \in \mathcal{S}$ is said to have the Markov property if for any subset $A \subset \mathcal{S}$ we have*

$$p_A(\mathbf{x}_A|\mathbf{x}_{\mathcal{S}\setminus A}) = p_A(\mathbf{x}_A|\mathbf{x}_{\partial A}). \tag{2.14}$$

In other words, the probability density for the A configurations conditioned on the configurations of the "environment" is equal to the probability density conditioned on the configurations of the A boundary. The conditional densities depend on the set A, hence the index A in Equation (2.14). As for the reciprocal process, one can prove that given the boundary configuration $x_{\partial A}$, the "inside" configuration x_A and the "outside" configuration are independent. When the set A is reduced to one site s, we can write Equation (2.14) as

$$p_s(x_s|\mathbf{x}_{\mathcal{S}\setminus\{s\}}) = p_s(\mathbf{x}_s|x_r, r \in \mathcal{N}_s),$$

which is the second condition given in the MRF definition 1. [6]

2.3.1 Correlation Function

The random field will be assumed homogeneous, i.e., the conditional densities are independent of the site location s. As for the reciprocal process, we will show that the correlation function satisfies, on the square lattice, a difference equation that we will indicate how to solve to obtain the correlation structure of the Gauss-Markov field. For this we use a vector notation to identify sites on the lattice. Choosing some site o as an origin, the site s will be denoted by a vector $\mathbf{s}$ whose origin is located at o. We let $\mathbf{h} = (1,0)$ and $\mathbf{v} = (0,1)$ be the horizontal and vertical unit vectors on the lattice. The four nearest-neighbor of a site $\mathbf{s}$ will be $\mathbf{s}+\mathbf{h}, \mathbf{s}+\mathbf{v}, \mathbf{s}-\mathbf{h}$, and $\mathbf{s}-\mathbf{v}$.

Proposition 6 *The correlation function of a two-dimensional zero-mean, homogeneous, nearest-neighbor Gauss-Markov field satisfies the two-dimensional difference equation*

$$R(\mathbf{l}) - a_e R(\mathbf{l}-\mathbf{h}) - a_n R(\mathbf{l}-\mathbf{v}) - a_w R(\mathbf{l}-\mathbf{h}) - a_s R(\mathbf{l}-\mathbf{v}) = d\delta(\mathbf{l}), \tag{2.15}$$

for $a_e, a_n, a_w, w_s, d \in \Re$.

[6] We do not discuss the important problem of knowing whether this condition, which is implied by (2.14), is sufficient for (2.14) to hold [10].

Proof: Since the field is homogeneous and Gaussian, the conditional expectation of $X_{\mathbf{s}}$ given a configuration of the nearest-neighbors, $X_{\mathbf{r}}, \mathbf{r} \in \mathcal{N}_{\mathbf{s}}$, is a linear combination of the $X_{\mathbf{r}}$'s, i.e., there exist $a_e, a_n, a_w, a_s \in \Re$, the indices designating the east, north, west, and south neighbors, such that

$$X_{\mathbf{s}} = a_e X_{\mathbf{s}+\mathbf{h}} + a_n X_{\mathbf{s}+\mathbf{v}} + a_w X_{\mathbf{s}-\mathbf{h}} + a_s X_{\mathbf{s}-\mathbf{v}} + Z_{\mathbf{s}}, \tag{2.16}$$

where $Z_{\mathbf{s}}$, $\mathbf{s} \in \mathcal{S}$, is an error, or residual, field. The random variable $Z_{\mathbf{s}}$ is uncorrelated with the field variables $X_{\mathbf{r}}$ where $\mathbf{r}$ is a neighbor of $\mathbf{s}$. Moreover because of the Markov condition, it is uncorrelated with *all* the field variables other than $X_{\mathbf{s}}$. Let $\mathbf{s}+\mathbf{l}$ denote some other site. Then since $E[X_{\mathbf{s}+v\mathbf{l}} Z_{\mathbf{s}}] = 0, \mathbf{l} \neq 0$, and posing $E[X_{\mathbf{s}} Z_{\mathbf{s}}] = d, \mathbf{l} \neq 0$, we get

$$\begin{aligned} E[X_{\mathbf{s}+\mathbf{l}} X_{\mathbf{s}}] &= a_e E[X_{\mathbf{s}+\mathbf{l}} X_{\mathbf{s}+\mathbf{h}}] + a_n E[X_{\mathbf{s}+\mathbf{l}} X_{\mathbf{s}+\mathbf{v}}] \\ &+ a_w E[X_{\mathbf{s}+\mathbf{l}} X_{\mathbf{s}-\mathbf{h}}] + a_s E[X_{\mathbf{s}+\mathbf{l}} X_{\mathbf{s}-\mathbf{v}}] + d\delta(\mathbf{l}) \end{aligned} \tag{2.17}$$

which, because of the homogeneity assumption, gives (2.15). ∎

The coefficients a_e, a_w, a_n, a_s can be determined from the correlation values by writing (2.15) for $\mathbf{l} = \mathbf{h} \pm \mathbf{v}$ and using the fact that $R(\mathbf{l}) = R(-\mathbf{l})$. This results in the following linear system of equations

$$\begin{bmatrix} a & b & c & d \\ b & a & d & c \\ b & e & c & f \\ e & b & f & c \end{bmatrix} \cdot \begin{bmatrix} a_e \\ a_n \\ a_w \\ a_s \end{bmatrix} = \begin{bmatrix} g \\ h \\ g \\ h \end{bmatrix}, \tag{2.18}$$

where the coefficients of the 4×4 matrix and the right-hand side vector are determined by the values of the correlation function R. It is important to note that these necessary equations impose a fundamental structural constraint on the field parameters, a_e, a_w, a_n, a_s, which is $a_e = a_w$ and $a_n = a_s$, i.e., in the nearest-neighbor lattice, the eastern parameter is equal to the western one, and the northern parameter is equal to the southern one. In order to see this, notice that the above equations are invariant under the mapping north $\rightarrow$ south, east $\rightarrow$ west. This result for the parameter symmetry is the counterpart in the two-dimensional case to the equality of the two coefficients u and v for the 1D reciprocal case.

The other point of analogy between the one-dimensional reciprocal case and the two-dimensional Gauss-Markov case is in the structure of the noise field $Z_{\mathbf{s}}$. We have seen that for the reciprocal Gaussian process the noise process Z_n is not white but correlated. This fact gets inherited by the two-dimensional Gauss-Markov field. The correlation

function of the two-dimensional noise field can be computed as follows. From (2.16) we get

$$Z_{\mathbf{s}} = X_{\mathbf{s}} - a_e X_{\mathbf{s}+\mathbf{h}} - a_n X_{\mathbf{s}+\mathbf{v}} - a_e X_{\mathbf{s}-\mathbf{h}} - a_n X_{\mathbf{s}-\mathbf{v}}. \tag{2.19}$$

Multiplying the above equation with $Z_{\mathbf{s}\pm\mathbf{h}}$ and using the uncorrelation property between the error and the data, we get

$$E[Z_{\mathbf{s}\pm\mathbf{h}} Z_{\mathbf{s}}] = -a_e E[Z_{\mathbf{s}\pm\mathbf{h}} X_{\mathbf{s}\pm\mathbf{h}}] = -a_e E[Z^2_{\mathbf{s}\pm\mathbf{h}}].$$

Similarly,

$$E[Z_{\mathbf{s}\pm\mathbf{v}} Z_{\mathbf{s}}] = -a_n E[Z_{\mathbf{s}\pm\mathbf{v}} X_{\mathbf{s}\pm\mathbf{v}}] = -a_n E[Z^2_{\mathbf{s}\pm\mathbf{v}}].$$

Note that we have

$$E[Z^2_{\mathbf{s}\pm\mathbf{v}}] = E[Z^2_{\mathbf{s}\pm\mathbf{h}}] = E[Z^2_{\mathbf{s}}] = d.$$

For $\mathbf{l} \neq \pm\mathbf{h}$ or $\mathbf{l} \pm \mathbf{v}$, we have

$$E[Z_{\mathbf{l}} Z_{\mathbf{s}}] = 0.$$

The most important aspect of the noise correlation function is that it is not white and that its structure is determined by that of the Gaussian field itself.

2.3.2 Power Spectrum

Solving the two-dimensional difference equation to obtain the fundamental forms of the correlation function is more difficult than in the 1D case.[7] However, because of linearity and stationarity, the two-dimensional power spectrum, which is the Fourier transform of the correlation function, can be computed. Indeed we have

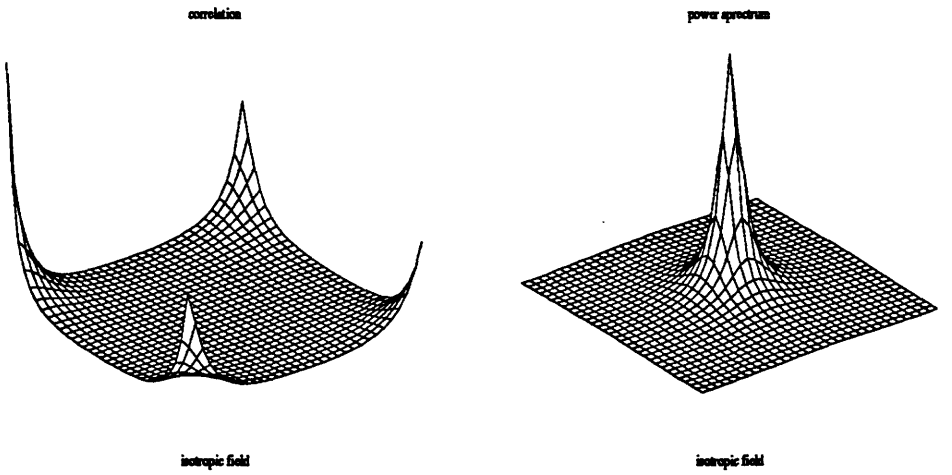
$$R(\mathbf{l}) - a_e(R(\mathbf{l}+\mathbf{h}) + R(\mathbf{l}-\mathbf{h})) - a_n(R(\mathbf{l}+\mathbf{v}) + R(\mathbf{l}-\mathbf{v})) = d\delta(\mathbf{l}).$$

Taking the two-dimensional Fourier transform of both sides, and denoting by $S(\lambda_1, \lambda_2)$ the power spectrum of the random field we get

$$S(\lambda_1, \lambda_2) = d(1 - 2a_e \cos\lambda_1 - 2a_n \cos\lambda_2)^{-1}.$$

Since the power spectrum is a non-negative function, the above expression imposes yet another constraint on the parameters of the Gaussian field. Indeed for $S(\lambda_1, \lambda_2)$ to be non-negative, it is sufficient that $|a_e| + |a_n| < \frac{1}{2}$. Note that this fact is not apparent from the system of linear equations (2.18) used to compute these parameters.

[7] In fact the solution is possible in the case when we impose toroidal boundary conditions [11].



(a) (b)

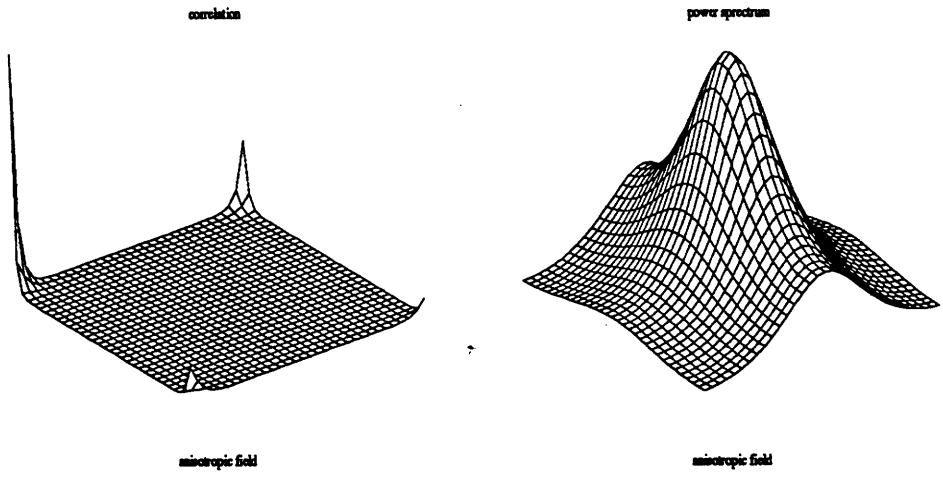
FIGURE 2-2: Correlation function (a) and power spectrum (b) of an isotropic Gauss-Markov random field. The origin of the correlation plot is at the leftmost corner.

Figure 2-2 shows the plots of the correlation function and the power spectrum of an isotropic Gauss-Markov field in which $a_e = a_n = .24$. The power spectrum, plotted over the grid $[-\pi, +\pi] \times [-\pi, +\pi]$, indicates that most of the signal energy is concentrated around the origin. The correlation plot, in which the origin is at the leftmost corner, indicates that that there is a decay in the correlation as distance increases. The raised corners are due to the high curvature areas in the power spectrum.

Similarly, Figure 2-3 shows the plots of the correlation function and the power spectrum of an anisotropic Gauss-Markov field in which $a_e = .24$, $a_n = .10$. The power spectrum, plotted over the grid $[-\pi, +\pi] \times [-\pi, +\pi]$, has a symmetric distribution about the $\lambda_2 = 0$ axis. Note how the smoothness of the power spectrum affects the height of the correlation function at the three corners other than the origin.

Important Remark: The above computation of the power spectrum used the implicit assumption that $d > 0$. What if $d = 0$? Can we define a Gaussian MRF? The answer is yes, we can! In fact the correlation functions in this case correspond to the cosine correlation function in the reciprocal case. The reader can easily verify in the latter that the cosine correlation function does correpond to a process with zero noise sequence *if the parameter* $\frac{\theta}{\pi}$ *is assumed irrational.*

When the field is defined on an infinite lattice, the degenerate zero-noise case can be eliminated by assuming that the field is regular [49]. At least in the one-dimesional case, i.e, $\mathcal{S} = \mathbf{Z}$, this additional assumption has far-reaching and somewhat disappointing implications. Indeed, Levy [49] has proved that a regular, reciprocal, stationary Gaussian



(a) (b)

FIGURE 2-3: Correlation function (a) and power spectrum (b) of an anisotropic Gauss-Markov random field. The origin of the correlation plot is at the leftmost corner.

field is necessarily Markov! In other words, noncausality is not adding anything new!

2.3.3 Gauss-Gibbs Distribution: Hammersley-Clifford Revisited

We would like to end this section by mentioning how one can derive a Gibbs probability distribution for the two-dimensional Gauss-Markov random field. One can see this as a (yet another) proof of the famous Hammersley-Clifford theorem [6] about the equivalence between Markov random fields and Gibbs random fields.

To understand this paragraph, let us review the different representations of a Gauss-Markov random field. This field was first defined using Markovian conditional probabilities (Markov representation). Next, an autoregressive model driven by colored noise was introduced based on the properties of Gaussian estimation (autoregressive representation). Now, we would like to introduce the third representation, namely, the one corresponding to Definition 4 (Gibbs representation).

In order to obtain the Gibbs probability distribution, one needs to define the *clique* potentials. In the nearest-neighbor lattice, a clique is one of the following three subsets

$$\{\mathbf{s}\}, \{\mathbf{s}, \mathbf{s}+\mathbf{h}\}, \{\mathbf{s}, \mathbf{s}+\mathbf{v}\}.$$

Denoting by σ^2 the variance of the noise process Z_s, the clique potentials are defined by

$$V(x_{\mathbf{s}}) = \frac{x_{\mathbf{s}}^2}{\sigma^2}$$

$$V(x_{\mathbf{s}}, x_{\mathbf{s+h}}) = -\frac{a_e x_{\mathbf{s}} x_{\mathbf{s+h}}}{\sigma^2}$$
$$V(x_{\mathbf{s}}, x_{\mathbf{s+v}}) = -\frac{a_n x_{\mathbf{s}} x_{\mathbf{s+v}}}{\sigma^2}$$

The energy of the Gibbs distribution is then defined as the sum over all cliques of the clique potentials given above, i.e.,

$$E[\mathbf{x}] = \frac{1}{\sigma^2} \sum_{\mathbf{s} \in \mathcal{S}} (x_{\mathbf{s}}^2 - a_e x_{\mathbf{s}} x_{\mathbf{s+h}} - a_n x_{\mathbf{s}} x_{\mathbf{s+h}}),$$

so that the Gibbs distribution is given as usual by [8]

$$p(\mathbf{x}) = \frac{1}{Z} \exp[-\frac{1}{T} E(\mathbf{x})].$$

The existence of the above distribution depends on the normalizing constant Z being non-zero. We will now give a necessary and sufficient condition for a non-vanishing normalizing constant. It is not difficult to verify that the energy $E(\mathbf{x})$ can also be written as

$$E(\mathbf{x}) = \frac{1}{\sigma^2} \mathbf{x}^\top \mathbf{H} \mathbf{x},$$

where $\mathbf{H}$ is, for the periodic boundary condition, a block circulant matrix [40]. Using the rules of Gaussian integration, it is not difficult to see that

$$\int_{\Re^{|\mathcal{S}|}} \exp\left[-\frac{1}{T\sigma^2} \mathbf{x}^\top \mathbf{H} \mathbf{x}\right] d\mathbf{x} = \frac{|det(\mathbf{H})|^{1/2}}{(\pi \sigma^2 T)^{|\mathcal{S}|/2}}.$$

It follows that Z is non-zero, and therefore the Gibbs distribution exists, if and only if the determinant of the matrix $\mathbf{H}$ is non-zero. Note that this imposes a constraint on the model parameters a_e and a_n that will insure that the matrix $\mathbf{H}$ is positive definite. This constraint is the same as the one obtained from the positivity of the power spectrum [61], i.e., $|a_e| + |a_n| < \frac{1}{2}$.

2.4 Ising Models

In the previous section, we have dealt with a lattice system in which the pixel state space is the whole real line. We have seen that in the Gaussian case, it is possible to obtain either a closed-form expression for the correlation function or a recurrence equation whose solution can ultimately give this function. Perhaps surprisingly, getting the correlation function for Gibbs random fields with finite pixel state space is much harder. We

[8] This expression for the Gibbs distribution can also be computed directly using the formula linking the probability density with the set of its conditional probabilities. See Appendix B.

should recognize though that finite discrete problems do not present the same analytical possibilities as infinite continuous ones. A case in point is combinatorial optimization which is generally much harder to deal with than for instance convex optimization.

In this section, we discuss how to obtain the correlation function of the simplest Gibbs distribution on a square lattice having a finite, discrete state space and a nearest-neighbor interaction, namely, the so-called Ising model. As for the Gaussian case, we deal first with the one-dimensional case, the Ising chain with a periodic boundary condition, and then we consider the two-dimensional case. The method used to derive the correlation function in both cases is called the transfer matrix method. It reduces the calculation of the partition function to the computation of the eigenvalues of a certain matrix. This method is rather general and has been used for solving exactly a number of models, other than the 2D, zero external field Ising model, in statistical physics [5]. It is however radically different from the method used for the Gaussian case which is based on a linear representation of the Gauss-Markov random field. Such a representation is lacking when the pixel has a finite state space.

2.4.1 One-Dimensional Ising Model

We consider a linear Ising chain of length N, in which each pixel $k \in \{1, \ldots, N\}$ has a pixel value $x_k \in \{-1, +1\}$. We also assume that the chain is periodic, i.e., $x_{N+1} = x_0$. The energy function of the chain is

$$E(\mathbf{x}) = -J \sum_{n=1}^{N} x_n x_{n+1} = -J E_1(\mathbf{x}) \tag{2.20}$$

When $J > 0$ the chain is called attractive, or ferromagnetic, because the state that minimizes the energy is the one in which neighboring pixels will be alike, i.e., both up (+1) or both down (-1). The chain is repulsive, or antiferrormagnetic, when $J < 0$. In this case, the energy is minimized when neighboring sites have different states. The periodic Ising chain is not just a simplification of the Ising lattice to the 1D case. It is also a valid model for some biological systems. For instance, the very complex hemoglobin molecule, the oxygen carrier in the red blood cells, can be modeled by an Ising ring of four sites [85].

The Gibbs probability distribution of the models is given by

$$P(\mathbf{x}) = Z_N^{-1} \exp\left[-\beta E_1(\mathbf{x})\right], \tag{2.21}$$

where $\beta = \frac{J}{T}$, and the subscript N is a reminder that the partition function is computed for N sites. This partition function can be computed exactly as is shown in the following

Lemma 6 *The partition function of a 1D periodic Ising chain having N sites is given by*

$$Z_N = \lambda_1^N + \lambda_2^N$$

where λ_i, $i = 1, 2$ are the eigenvalues of the 2×2 matrix

$$\mathbf{L} = \begin{bmatrix} L(+1,+1) & L(+1,-1) \\ L(-1,+1) & L(-1,-1) \end{bmatrix},$$

with $L(x, x') = \exp(-\beta x x')$.

Proof: We have

$$\begin{aligned} Z_N &= \sum_{\mathbf{x}} e^{-\beta E_1(\mathbf{x})} \\ &= \sum_{\mathbf{x}} \prod_{n=1}^{N} e^{-\beta x_n x_{n+1}} \\ &= \sum_{\mathbf{x}} \prod_{n=1}^{N} L(x_n, x_{n+1}) \\ &= \sum_{x_1=\pm 1} \cdots \sum_{x_N=\pm 1} \prod_{n=1}^{N} L(x_n, x_{n+1}) \\ &= \sum_{x_1=\pm 1} L^N(x_1, x_{N+1}) \\ &= L^N(1,1) + L^N(-1,-1) \end{aligned}$$

where $L^N(x_n, x_{n+1})$ denotes the (x_n, x_{n+1}) coefficient of the matrix $\mathbf{L}$ raised to the power N. Note that in the last equation, we used the periodic boundary condition $x_{N+1} = x_1$. It follows that we can write

$$Z_N = \text{Trace}(\mathbf{L}^N) = \lambda_1^N + \lambda_2^N,$$

where $\lambda_i, i = 1, 2$ are the eigenvalues of the matrix $\mathbf{L}$. ∎

The matrix $\mathbf{L}$ is called the spin-to-spin *transfer matrix* of the Ising chain. We will see in the next paragraph that the two-dimensional case requires the use of a column-to-column transfer matrix. Since $\mathbf{L}$ is symmetric, its eigenvalues are real and can be computed explicitly. Indeed, we have $\lambda_1 = 2\cosh\beta$ and $\lambda_2 = 2\sinh\beta$. The orthonormal eigenvectors corresponding to these eigenvalues are $\mathbf{v}_1 = \frac{1}{\sqrt{2}}(1,1)^T$ and $\mathbf{v}_2 = \frac{1}{\sqrt{2}}(1,-1)^T$, respectively.

Computation of the correlation function for the 1D Ising chain: Let $k, l \in \{1, \ldots N\}$ be two sites on the chain. Then the correlation between the pixels x_k and x_l is given by

$$R(k,l) = Z_N^{-1} \sum_{\mathbf{x}} x_k x_l \exp(-\beta E_1(\mathbf{x})). \tag{2.22}$$

Using the lemma above and again the transfer matrix $\mathbf{L}$ and its eigenstructure, we can prove the following

Proposition 7 *The correlation function between any two pixels $1 \leq k, l \leq l$ on a periodic Ising chain of length N is given by*

$$R(k,l) = \frac{(\tanh\beta)^{|l-k|} + (\tanh\beta)^{N-|k-l|}}{1 + (\tanh\beta)^N}. \tag{2.23}$$

Proof: Assume that $k < l$. Then we can write for the summation over the chain configurations

$$\begin{aligned} R_{kl} &= Z_N^{-1} \sum_{\mathbf{x}} x_k x_l \exp\left[-\beta E(\mathbf{x})\right] \\ &= Z_N^{-1} \sum_{\mathbf{x}} \prod_{n=1}^{k-1} L(x_n, x_{n+1}) x_k \prod_{n=k}^{l-1} L(x_n, x_{n+1}) x_l \prod_{n=l}^{N} L(x_n, x_{n+1}) \\ &= Z_N^{-1} \sum_{x_k, x_l = \pm 1} x_k L^{N-l+k}(x_k, x_l) L^{l-k}(x_k, x_l). \end{aligned}$$

In order to compute the last summation, note that the transfer matrix $\mathbf{L}$ raised to any power s is given by

$$\mathbf{L}^s = \lambda_1^s \mathbf{V}_1 + \lambda_2^s \mathbf{V}_2,$$

where $\mathbf{V}_i = \mathbf{v}_i \mathbf{v}_i^T, i = 1, 2$, is the rank one matrix corresponding to the eigenvector $\mathbf{v}_i$. It follows that

$$\begin{aligned} R(k,l) &= Z_N^{-1} \sum_{x_k, x_l = \pm 1} x_k x_l \left[\lambda_1^{N-l+k} V_1(x_k, x_l) + \lambda_2^{N-l+k} V_2(x_k, x_l)\right] \\ &\quad \left[\lambda_1^{l-k} V_1(x_l, x_k) + \lambda_2^{l-k} V_2(x_l, x_k)\right] \\ &= Z_N^{-1} \left[\lambda_1^{N-l+k} \lambda_2^{l-k} + \lambda_1^{l-k} \lambda_2^{N+k-l}\right]. \end{aligned}$$

Using the expression of Z_N as given by Lemma (6) and the explicit expression of the eigenvalues we get

$$\begin{aligned} R(k,l) &= \frac{(\lambda_2/\lambda_1)^{l-k} + (\lambda_2/\lambda_1)^{N-l+k}}{1 + (\lambda_2/\lambda_1)^N} \\ &= \frac{(\tanh\beta)^{l-k} + (\tanh\beta)^{N-l+k}}{1 + (\tanh\beta)^N}. \end{aligned}$$

The above expression is valid for $l > k$. Exchanging the role of k and l, we get the expression given by (2.23). ∎

The above formula for the correlation function of a periodic Ising chain of length N calls the following remarks:

1. The formula is valid for both attractive $J > 0$ and repulsive $J < 0$ interactions.

2. The correlation between pixels k and l depends only on the difference $k - l$, i.e., $R(k, l) = R_{k-l}$. This is a direct result of the shift invariance property of the Ising ring. Note also that the sequence $R(n)$ is even, i.e., $R(n) = R(-n)$.

3. For $\beta \neq 0$ (finite temperature) $|\tanh \beta| < 1$. Therefore as the number of pixels becomes very large, i.e., $N \to \infty$, we get

$$R(k - l) = (\tanh \beta)^{|k-l|}, k, l \in \mathbf{Z},$$

   and the even correlation function is $R(n) = (\tanh \beta)^{|n|}, n \in \mathbf{Z}$.

4. Comparing the correlation function of the shift-invariant Ising ring in the infinite length limit with that of the stationary Gauss-Markov process shows that they are identical if we impose that the variance and the nearest-neighbor correlation coefficient of the Gauss-Markov process are $\sigma_n = 1$ and $\rho = \tanh \beta$, respectively.

5. For the Ising ring, the parameter β which is the ratio of the bonding strength J and the temperature T controls the amount of correlation between pixels. For finite, non-zero temperature and finite, non-zero bonding strength, we have $-1 < \tanh \beta < 1$, so that the correlation sequence R_n decays exponentially *to zero* as $|n| \to \infty$. This situation is similar to that of a Gauss-Markov process with $|\rho| < 1$.

6. This exponential decay to zero means that the Ising lattice does not exhibit any long-range order. In particular, the oscillatory behavior that the correlation function of a 1D Gaussian reciprocal process exhibits does not exist for the Ising ring. However, when the temperature is very low or the bonding strength is very large, we have perfect order. This means that in the attractive case, all the pixels have value +1 or all the pixels have value -1, while in the repulsive case the pixels are alternatively +1 or -1.

In the next paragraph, we will see that the exponential decrease of the correlation function with pixel distance is also valid for the 2D Ising lattice. However, there is a fundamental difference with the 1D case which is that the limit as the distance goes to infinity might not be zero if the parameter β is above certain critical value β_c.

2.4.2 Two-Dimensional Ising Model

In this section, we show how the transfer matrix method can be used to compute both the partition function and the correlation function of the zero external field, isotropic Ising lattice. The exact computation of the partition function was first done by the Norwegian physicist Onsager [63] and constituted a major advance in statistical physics. The exposition given here about the two-dimensional Ising model belongs to the standard lore of the theory of critical phenomena [38]. Our treatment follows the old but now classic survey paper of Newell and Montroll [62].

Again our objective is to compute the correlation function

$$R(s,r) = Z_N^{-1} \sum_{\mathbf{x}} x_s x_r \mathrm{e}^{-\beta E(\mathbf{x})}, \tag{2.24}$$

measuring the second-order dependence between two sites on the lattice designated by the indices s and r. But now we have a square lattice $\mathcal{S}$ of size $N \times N$. Also the interaction between lattice pixels is nearest-neighbor, and the energy of the lattice can be written as

$$E(\mathbf{x}) = - \sum_{1 \leq p,q \leq N^2} J_{pq} x_p x_q,$$

where $J_{pq} = J$ if and only if $p \in \mathcal{N}_q$. The lattice pixels take the values $+1$ or -1. As in the 1D case, the parameter β is the ratio of the bonding strength J and the temperature T. From the expression of the energy, it is clear that when $J > 0$ the energy of the lattice is minimized when all the pixels have the same value. All the lattice is black or all the lattice is white. This is the ferromagnetic case. When $J < 0$, the energy is minimized when the lattice has a black-and-white checkerboard pattern. This is the antiferromagnetic case.

Before delving into the technical details, it might be worthwhile to describe informally the behavior of the correlation function. Under the assumption of a toroidal, ferromagnetic lattice with the number of rows and number of columns becoming infinite, the correlation function of two sites on the same row or two sites on the same column will

a) exponentially decrease to *zero* with the distance between the two sites if the parameter β is less than or equal to some critical value β_c that can be computed exactly;

b) exponentially decrease to a *non-zero* value with distance between the two sites if β is strictly greater than β_c.

The above behavior reveals a fundamental difference between the 1D Ising chain and the two-dimensional Ising lattice. While the 1D Ising chain does not exhibit any long-range order for any finite, non-zero β, the long-range behavior of the two-dimensional Ising lattice depends on the value of the parameter β. The long-range correlation for the two-dimensional Ising model at β values above the critical one should remind us of the long-term dependence that the increments of fractional Brownian motions [55] with parameter $H > 1/2$ exhibit. Let us for now stress the difference between the 1D and the two-dimensional Ising lattice by quoting from [62]

> One must be an optimist to expect a one-dimensional system to behave like its two- or three-dimensional analog. If one is hopeful that techniques used to solve a one-dimensional problem will help solve a two- or three-dimensional, [9] his optimism quickly subsides when he tries to apply them to the Ising lattice.

Computation of the correlation function for the two-dimensional Ising model: In this paragraph, we will give enough details to show how the transfer matrix method, properly generalized to handle two-dimensional lattices, will help obtain an *asymptotic* (i.e., when the number of lattice sites becomes infinite) value for the correlation function between any pair of sites on the same row. A similar result can be had for any pair of sites on the same column. As in the 1D case the idea is to write the correlation in terms of the eigenstructure of a transfer matrix.

Specifically, let m and n be the number of rows and columns respectively.[10] Let $s = (i, j)$ and $r = (i, j + k)$ be two lattice sites on the same row, and let their pixel values be denoted by $x_{i,j}$ and $x_{i,j+k}$ respectively. For these pixels, the expression for the correlation function is

$$R(i, j; i, j + k) = Z_{M,N}^{-1} \sum_{\mathbf{x}} x_{i,j} x_{i,j+k} \mathrm{e}^{-E(\mathbf{x})/T}. \tag{2.25}$$

Using the fact that the lattice is periodic in columns, we have $x_{p,n+1} = x_{p,1}, \forall p \in \{1, \ldots, m\}$, the energy of the model can be separated into the contributions due to the horizontal bonds and those due to the vertical bonds. Therefore one can write,

$$E(\mathbf{x}) = -J \sum_{p=1}^{m-1} \sum_{q=1}^{n} x_{p,q} x_{p+1,q} - J \sum_{p=1}^{m} \sum_{q=1}^{n} x_{p,q} x_{p,q+1} \tag{2.26}$$

[9] Almost fourty years after the paper of Montroll and Newell, the 3D Ising model is still unsolved.

[10] Although these numbers are the same for our square lattice, we distinguish between them here because their infinite limits will be taken at two different times.

The interaction energy can be separated into an intra-column interaction, i.e, inside each column, and inter-column interaction, i.e, between every pair of neighboring columns. If we now denote a column configuration by $\mathbf{c}_q = (x_{1,q}, \ldots, x_{m,q})$ we can express the intra-column interaction by

$$\xi_1(\mathbf{c}_q) = -J \sum_{p=1}^{m-1} x_{p,q} x_{p+1,q}$$

and the inter-column interaction by

$$\xi_2(\mathbf{c}_q, \mathbf{c}_{q+1}) = -J \sum_{p=1}^{m} x_{p,q} x_{p+1,q}$$

where we assume that $\mathbf{c}_{n+1} = \mathbf{c}_1$. Using the expression of ξ_1 and ξ_2 we can write the total energy as

$$E(\mathbf{x}) = \sum_{q=1}^{n} \xi_1(\mathbf{c}_q) + \xi_2(\mathbf{c}_q, \mathbf{c}_{q+1}).$$

With this expression of the energy the partition function can be written as

$$\begin{aligned} Z_{m,n} &= \sum_{\mathbf{c}_q, 1 \le q \le n} \prod_{q=1}^{n} L(\mathbf{c}_q, \mathbf{c}_{q+1}) \\ &= \sum_{\mathbf{c}_1} L^n(\mathbf{c}_1, \mathbf{c}_1) \end{aligned}$$

where

$$\begin{aligned} L(\mathbf{c}, \mathbf{c}') &= \exp(-\beta \xi_1(\mathbf{c}) - \beta \xi_2(\mathbf{c}, \mathbf{c}')) \\ &= \exp\left[\beta \sum_{p=1}^{m-1} x_p x_{p+1}\right] \exp\left[\beta \sum_{p=1}^{m} x_p x'_p\right], \end{aligned}$$

where in the last equation we have dropped the column index q and kept only the row index. As for the 1D case, the symbol $L^n(\mathbf{c}, \mathbf{c})$ denotes the $(\mathbf{c}, \mathbf{c})$ component of the $2^m \times 2^m$ matrix $\mathbf{L}$, whose elements are defined above, raised to the power n. It follows that the partition function is given by

$$Z_{m,n} = \text{Trace}(\mathbf{L}^n) = \sum_{q=1}^{2^m} \lambda_q^n,$$

where λ_q is the q-th eigenvalue of the matrix $\mathbf{L}$. As in the 1D case, the computation of the partition function is reduced to finding the eigenstructure of the matrix $\mathbf{L}$. Notice however that the problem exhibits a curse of dimensionality. While for the 1D case the matrix $\mathbf{L}$ was 2×2, in the two-dimensional case it is $2^m \times 2^m$. It follows that the complexity of computing the partition function grows exponentially with the column (or

row) size of the lattice. Note that in the thermodynamic limit, i.e., when $m, n \to \infty$, only the knowledge of the largest eigenvalue is necessary. Indeed, for finite m the matrix $\mathbf{L}$ has strictly positive elements, and it follows by the Perron-Frobenius theorem that the eigenvalue of maximum modulus is strictly non-degenerate. Let λ_1 be this eigenvalue. Then when $n \to \infty$ with m fixed, the partition function will be of the order of λ_1^n. The fundamental contribution of Onsager was to have explicitly computed the largest eigenvalue of the matrix $\mathbf{L}$ for any finite m. His expression for λ_1 was [85]

$$\lambda_1 = (2\sinh 2\beta)^{m/2} \exp\left[\frac{1}{2}(\gamma_1 + \gamma_2 + \ldots + \gamma_{2m-1})\right],$$

where

$$\cosh \gamma_k = \cosh 2\beta \coth 2\beta - \cos(\frac{\pi k}{m}).$$

Having evaluated the denominator of the correlation function expression, let us now concentrate on the numerator. Decomposing the exponential into a product, we can write

$$\sum_{\mathbf{x}} x_{i,j} x_{i,j+k} e^{-\beta E(\mathbf{x})} = \sum_{\mathbf{c},\mathbf{c}'} x_i L^k(\mathbf{c}, \mathbf{c}') x_i' L^{n-r}(\mathbf{c}', \mathbf{c}),$$

where $\mathbf{c} = (x_1, \ldots, x_m)$ and $\mathbf{c}' = (x_1', \ldots, x_m')$. Using the spectral decomposition of the matrix $\mathbf{L}$, we can write

$$\mathbf{L}^s = \sum_{q=1}^{2^m} \lambda_q^s \mathbf{v}_q \mathbf{v}_q^T$$

where $\mathbf{v}_q$ is the eigenvector corresponding to the eigenvalue λ_q. For the $(\mathbf{c}, \mathbf{c}')$ component of the matrix $\mathbf{L}^s$, we have

$$\mathbf{L}^s(\mathbf{c}, \mathbf{c}') = \sum_{q=1}^{2^m} \lambda_q^s(\mathbf{c}) \mathbf{v}_q \mathbf{v}_q^T(\mathbf{c}')$$

It follows that the numerator of the correlation function can be written as

$$\sum_{p,q=1}^{2^m} \lambda_p^{n-k} \lambda_q^k \sum_{\mathbf{c},\mathbf{c}'} x_k x_k' \mathbf{v}_p(\mathbf{c}) \mathbf{v}_p(\mathbf{c}') \mathbf{v}_q(\mathbf{c}) \mathbf{v}_q(\mathbf{c}').$$

Note that the last summation can be written compactly as the square of the scalar product of $\mathbf{v}_p$ and $x_k \mathbf{v}_q$, i.e., we can write the above sum as

$$\sum_{p,q=1}^{2^m} \lambda_p^{n-k} \lambda_q^k (\mathbf{v}_p, x_k \mathbf{v}_q)^2$$

Using the expression of the partition function, the correlation function for any pair of pixels on a given row can be written as

$$R_{m,n}(i, j; i, j+k) = \frac{\sum_{p,q=1}^{2^m} \lambda_p^{n-k} \lambda_q^k (\mathbf{v}_p, x_k \mathbf{v}_q)^2}{\sum_{p=1}^{2^m} \lambda_p^n}. \tag{2.27}$$

Fixing m and k and letting n go to infinity, the above expression for the correlation function becomes

$$R_m(i,j;i,j+k) = \sum_{p=1}^{2^m} (\lambda_j/\lambda_1)^k (\mathbf{v}_1, x_k \mathbf{v}_q)^2, \tag{2.28}$$

where we have dropped the dependence of the correlation function on the number of columns n. The first term in the above summation is zero [85]. It follows that as the distance k between the two pixels become large, the correlation function is of the order of

$$(\lambda_2 \lambda_1)^k (\mathbf{v}_1, x_k \mathbf{v}_2)^2.$$

It follows that the presence of long-range order, i.e., low-frequency content for the lattice configurations depend on whether the eigenvalue of maximum modulus is degenerate or not. It has already been mentioned that for a finite m this eigenvalue is non- degenerate, and therefore $|\lambda_2/\lambda_1| < 1$. Hence, for m finite and as $k \to \infty$,

$$\lim_{k\to\infty} R_m(i,j;i,j+k) = 0.$$

In other words, a cylindrical Ising lattice having an infinite number of columns and a finite number of rows cannot exhibit long-range order no matter what the value of the parameter β is. However as the number of rows becomes infinite, there is a *possibility* for the occurrence of long-range order. Whether this will actually happen or not depends on the value of the parameter β. Here, we merely state the result [85]

$$\lim_{k\to\infty} \lim_{m\to\infty} R_m(i,j;i,j+k) = \begin{cases} \lim_{m\to\infty} (\mathbf{v}_1, x_k \mathbf{v}_2)^2 > 0 & \text{for } \beta > \beta_c \\ 0 & \text{for } \beta < \beta_c. \end{cases} \tag{2.29}$$

2.4.3 Other Exactly Solvable Models

Baxter's book [5] contains many examples other than the 2-D Ising lattice that are exactly solvable, i.e., for which the Z partition function and therefore the thermodynamic properties of the lattice can be determined in closed form. The main tool in Baxter's approach is the transfer matrix method. Among the models treated by Baxter using this method are the ice-type models, the square lattice eight-vertex models, the Kagomé lattice eight-vertex models, etc. These models remain an untapped source of information about the representation of 2D patterns.

2.4.4 Effect of Constraints

Very often, one needs to impose constraints on the type of configurations that can be produced by the Gibbs distribution model. For instance, a binary alloy can be modeled

with an Ising lattice in which the ratio between the "up" spins and the "down" spins is kept constant [62]. In the area of texture modeling, Cross and Jain [14] have argued that the use of a histogram constraint improves the grayscale textural quality of a texture sampled from a Gibbs distribution by insuring the presence of all graylevels at any instant of the sampling process. It is intuitively clear however that imposing global constraints on the configuration space could drastically change the Markov dependence of the random field. This could happen even in the 1D case, as the following simple example shows.

Example 1 *Consider an Ising ring with four spins* $\{s_1, s_2, s_3, s_4\}$ *located at the four cardinal points of the circle. Remember that the Ising lattice is a periodic, nearest-neighbor lattice having the following cliques*

$$\{1,2\}, \{2,3\}, \{3,4\}, \{4,1\}.$$

The Gibbs energy of the system is

$$E(\mathbf{s}) = s_1 s_2 + s_2 s_3 + s_3 s_4 + s_4 s_1.$$

In the absence of constraints, the conditional probabilities of the different pixels satisfy the Markov dependence property

$$p(s_i | s_j, j \neq i) = p(s_i | s_{i-1}, s_{i+1}),\ i = 1,2,3,4,$$

where by periodicity, we have identified s_0 *with* s_4 *and* s_5 *with* s_1. *When we impose the global constraint that the number of* $+1$ *spins be equal to the number of* -1 *spins, the above expression for the conditional probability is no longer valid. Indeed, from Table 1, we get*

$$Z = \sum_{\mathbf{s}} e^{-E(\mathbf{s})/T} = 4 + 2e^{4/T}.$$

Moreover, we have

$$p(s_1 = -1 | s_2 = -1, s_4 = +1) = \frac{1}{2}$$
$$p(s_1 = +1 | s_2 = -1, s_4 = +1) = \frac{1}{2}.$$

But

$$p(s_1 = -1 | s_2 = -1, s_3 = +1, s_4 = +1) = 1$$
$$p(s_1 = +1 | s_2 = -1, s_3 = +1, s_4 = +1) = 0.$$

It follows that this constrained Ising ring does not have the Markov property, although we can define for it a Gibbs probability distribution.

s_1	s_2	s_3	s_4	$E(\mathbf{s})$
-1	-1	+1	+1	0
-1	+1	+1	-1	0
+1	+1	-1	-1	0
+1	-1	-1	+1	0
-1	+1	-1	+1	-4
+1	-1	+1	-1	-4

Table 2-1: Energy table for a four-pixel periodic Ising ring with uniform histogram constraint.

It is clear from the above example that the nearest-neighbor neighborhood structure is no longer appropriate for expressing the local dependence in the lattice. A larger neighborhood, here the whole lattice, is needed to express the dependence of a pixel value on the values of other pixels.

This can also be seen by looking at the energy function of the constrained system. Indeed, the usual way of expressing constraints on the configuration space of the lattice is through penalty functions added to the energy of the lattice. As an example, let us take a ferromagnetic binary alloy in which the two species (white pixels and black pixels) are mixed in equal amounts. Then one way for expressing this constraint is through the condition [11]

$$\sum_{s\in\mathcal{S}} x_s = 0,\ x_s \in \{-1,+1\}, \forall s \in \mathcal{S}.$$

It is clear that the above sum is zero if and only if the number of white pixels (spin value $+1$) is equal to the number of black pixels (spin value -1). In order to express penalty incurred by deviating from the global constraint, the above sum is squared, scaled with a "large" positive number μ and then added to the energy of the lattice, so that the latter now becomes

$$E_\mu(\mathbf{x}) = -\frac{1}{2}\sum_{s\in\mathcal{S}}\sum_{r\in\mathcal{N}_s} x_s x_r + \frac{\mu}{\beta}\left(\sum_{s\in\mathcal{S}} x_s\right)^2. \tag{2.30}$$

It is important to note that when the penalty function is expanded, quadratic interactions will appear between distant sites. In other words, the interaction energy is no longer restricted to pairs of nearest-neighbor sites. Another way of stating the same thing is that the nearest-neighbor Markov property is lost as a result of imposing the global constraints.

[11] We assume that the number of lattice sites is even.

With this expression of the lattice energy, the Gibbs distribution of the system becomes

$$p_\mu(\mathbf{x}) = Z(\mu)^{-1} \exp(-\beta E_\mu(\underline{x})),$$

where we have indicated the dependence of the distribution on the weight μ of the penalty function

$$J(\mathbf{x}) = \left(\sum_{s \in \mathcal{S}} x_s \right)^2 .$$

The penalty function J operates as an energy barrier to prevent the configurations that do not satisfy the global constraints from occurring. To see this more clearly, let us find the limit Gibbs distribution as the weight $\mu \to \infty$. We have

$$\begin{aligned} p_\mu(\mathbf{x}) &= \frac{\exp(-\beta E(\underline{x}) - \mu J(\mathbf{x}))}{\sum_\xi \exp(-\beta E(\xi) - \mu J(\xi))} \\ &= \frac{\exp(-\beta E(\underline{x}))}{\sum_\xi \exp(-\beta E(\xi)) \exp(-\mu(J(\xi) - J(\mathbf{x})))} \end{aligned}$$

Let $\mathbf{x}$ be a given configuration whose penalty function value is $J(x) > 0$. When μ is large, the terms of the partition function in which $J(\xi) < J(\mathbf{x})$ become negligible and only the terms in which $J(\xi) \leq J(\mathbf{x})$ are counted. In other words, the summation for the partition function becomes

$$\sum_{\xi : J(\xi) < J(\mathbf{x})} \exp(-\beta E(\xi)) \exp(-\mu(J(\xi) - J(\mathbf{x}))),$$

i.e., only the configurations having a penalty less than the barrier $J(\mathbf{x})$ are counted in the computation of the partition function. In the extreme case when $\mathbf{x}$ satisfies the global constraint $J(\mathbf{x}) = 0$ and when the penalty weight μ approaches ∞, only the configurations ξ such that $J(\xi) = 0$ are summed over. The probability distribution in this limit becomes

$$p(\mathbf{x}) = \frac{\mathbf{1}_{(J=0)}(\mathbf{x}) \exp(-\beta E(\underline{x}))}{\sum_{\xi : J(\xi)=0} \exp(-\beta E(\xi))},$$

where $\mathbf{1}_{(J=0)}$ is the characteristic function of the set $\{\mathbf{x} | J(\mathbf{x}) = 0\}$.

2.5 General Case: Mean-Field Approximation

It is clear from the previous sections, that computing *exactly* the correlation function of a Gibbs random fields is a rather difficult task. In most cases the *exact* computation of the partition function is a necessary but usually very difficult step that requires the use of special techniques with which the engineering community is not very familiar.

Even a binary model as simple as the Ising lattice requires the use of rather complicated algebraic methods like the transfer matrix method.[12] It is therefore legitimate to ask the question whether it is possible to compute the correlation function *approximately.* Because of its importance in the theory of critical phenomena, physicists have dealt with the problem of approximating the partition function for a long time [38]. Indeed, knowing the partition function makes possible the computation of all the important macroscopic thermodynamic quantities like free energy, magnetization and susceptibility. Usually, a critical phenomenon manifests itself as an abrupt change of these macroscopic quantities as function of the lattice temperature or the strength of an external field applied to the lattice. In fact it turns out that the latter physical situation is crucial in deriving an expression for the correlation function.

For any Gibbs system we have the following important lemma.[13]

Lemma 7 *Let* $P(\mathbf{x}) = \frac{1}{Z}\exp\left[-\frac{1}{T}E(\mathbf{x})\right]$ *be the probability distribution of an arbitrary Gibbs random field. Denote by* $\overline{x_p}$ *the mean at site* p *and by* $\overline{x_p x_q}$ *the correlation function between the two variables* x_p *and* x_q*. Let* $h_q x_q$ *be the value of a linear external field at site* q *(add* $-Th_q x_q$ *to the energy term.) Then we have*

$$\frac{\partial \overline{x_p}}{\partial h_q}(h_q = 0) = \overline{x_p x_q} - \overline{x_p}\,\overline{x_q}. \tag{2.31}$$

Proof: First write the expression of the mean as

$$\overline{x_p}(h_q) = \frac{1}{Z(h_q)}\sum_{\mathbf{x}} x_s \exp\left[-\frac{1}{T}(E(\mathbf{x}) - Th_q x_q)\right],$$

where

$$Z(h_q) = \sum_{\mathbf{x}} \exp\left[-\frac{1}{T}(E(\mathbf{x}) - Th_q x_q)\right].$$

Taking the partial derivative of the mean with respect to h_q, we get

$$\begin{aligned}\frac{\partial \overline{x_p}}{\partial h_q}(h_q) &= -\frac{1}{Z(h_q)^2}\frac{\partial Z(h_q)}{\partial h_q}\sum_{\mathbf{x}} x_p \exp\left[-\frac{1}{T}(E(\mathbf{x}) - Th_q x_q)\right] \\ &+ \frac{1}{Z(h_q)}\sum_{\mathbf{x}} x_p x_q \exp\left[-\frac{1}{T}(E(\mathbf{x}) - Th_q x_q)\right].\end{aligned}$$

It is not difficult to check that the ratio $\frac{1}{Z(h_q)}\frac{\partial Z(h_q)}{\partial h_q}$ is the mean $\overline{x_q}(h_q)$ at site q. By evaluating the right-hand side at zero, we get Equation (2.31) above. ∎

[12]For the Ising lattice, combinatorial methods can also be used to compute the partition function [85].

[13]This lemma is usually given for binary Ising models in statistical mechanics textbooks where it is known as the "linear response theorem" ([66], p. 33). As is clear from the simple proof, it is however valid under very general conditions.

Note that the right-hand side is the *covariance* function between sites p and q that we will denote by Γ_{pq}. The above formula for this function calls the following remarks. First, it is valid for any energy function $E(\mathbf{x})$, including one that would incorporate all the constraints of the system, e.g., the uniqueness and uniform histogram constraints that we will be dealing with in the next chapter. Second, it is valid for any temperature T whether it be below or above the critical temperature, see Chapter 6. Third, it is valid for any lattice structure, in particular, the cubic lattice structure that arises as binary matching elements [14] used to represent the site states when they belong to a finite discrete set, see Chapter 5. Fourth, it works *even* when the site random variables are continuous.[15] Finally, this formula can be generalized to compute higher order moments by exciting the lattice linearly at more than one site and computing the partial derivative of the mean with respect to the strength of the excitation inputs at the different sites.

The formula given by Lemma 7 above will be the basis for finding an approximation of the correlation function for *any* Gibbs distribution. The approximation is derived from the so-called mean-field theory in Statistical Mechanics. In the next chapter, we will introduce four different ways for finding a mean-field approximation for a Gibbs random field. We will attempt at clarifying the relationships between these different ways. Also, we will show how the mean-field equations, no matter how they are obtained, can be used to find approximations for the correlation function of the Gibbs system.

[14] See Chapter 3.

[15] The definition of Z in this case should be done with care to insure that the Z integral (rather than summation) is finite.

3

Mean-Field Approximations of Gibbs Distributions

When you follow two separate chains of thought, Watson, you will find some point of intersection which should approximate the truth.

Sherlock Holmes

A recent trend in computer vision has been to attempt to derive deterministic algorithms from stochastic-based, mainly Bayesian, formulations of early vision tasks. There are two main reasons for seeking deterministic algorithms. The first is fast processing speed as compared with Monte Carlo algorithms, and the second is analog hardware implementation. The well-known statistical physics paradigm of mean-field theory [2] provides an analytical framework for the derivation of such deterministic algorithms. Recent applications of the mean-field approach have been reported in [28] for surface reconstruction and [29, 8, 7] for image smoothing and segmentation. In [12], the mean-field aproximation is used as a tool for developing deterministic algorithms to solve the correspondence problem in stereo vision.

Our objective in this thesis is quite different though. We would like to show how the *mean-field* approximation can lead to a *correlation-field* approximation. This chapter is devoted to developing different mean-field theories for different classes of Gibbs models: binary, grayscale, unconstrained, or constrained. The actual computation of the

correlation-field, given in the next chapter, is based on Formula (2.31) proved at the end of the previous chapter.

In Chapter 6, we will see how we can use mean-field theory to get another type of approximation – that of the critical temperature indicating the presence of a "phase transition" between visually different patterns.

We will develop four different methods for deriving the mean-field equations. The first method is valid only for binary fields with quadratic interactions between pixel values. This type of interaction by itself does not make the random field Gaussian, because the pixel state space is discrete and finite. Moreover, there is no requirement that the interactions be only nearest-neighbors. When the pixel state space has more than two levels, with the interaction potentials remaining pairwise, we will show (second method) that it is still possible to use the results of the binary case. For both cases, we assume that no constraints are imposed on the configuration space. When we do impose constraints or when the interaction potential is not pairwise, we use two recent approaches developed by statistical physicists to formulate mean-field equations. The first approach is energy-based but contains one heuristic step that we will try to justify mathematically. The second approach is probabilistic, and constraints are imposed via classical Lagrange multiplier theory. But in both cases, the physical intuition is the same: on a given pixel, the other pixels act through their mean graylevel values. In other words, one can neglect the "details" of the field action on a given pixel, and this approximation becomes all the more valid when the noise variance, i.e., the temperature, of the field becomes small. The mean-field approximation has also an information-theoretic interpretation that was developed in [8].

3.1 Unconstrained Binary Quadratic Case

Since our interest is to deal with general lattice systems with discrete but finite state space, let us look again at a binary system in which we allow distant pixels to interact. In other words, we assume that we have a Gibbs system whose energy is given by

$$E(\mathbf{x}) = -\frac{1}{2} \sum_{s,r \in \mathcal{S}} J_{sr} x_s x_r, \tag{3.1}$$

where $x_s \in \{-1, +1\}$, $\forall s \in \mathcal{S}$. In this system, we don't assume that the interaction coefficients are nearest-neighbors. This allows for long-range interactions similar to the ones that exist in Hopfield networks [36] or the ones introduced through the imposition of global constraints on the lattice, Section 2.4.4. The stumbling block here is again the

computation of the partition function

$$Z = \sum_{\mathbf{x}} \exp\left[-\frac{1}{T}E(\mathbf{x})\right]. \tag{3.2}$$

We have seen in Chapter 2 that even in the homogeneous, isotropic, nearest-neighbor case of the Ising model, Z is very hard to compute exactly. We can however compute it *approximately*, using the fact that the interaction between pixel values is quadratic.

3.1.1 Effective Energy

First, we can write the energy function as

$$E(\mathbf{x}) = -\frac{1}{2}\mathbf{x}^T\mathbf{J}\mathbf{x}.$$

Moreover, by completing the square of the exponent, we can show that a Gaussian function satisfies the identity [1] [66, 80]

$$\int_{\Re^{|\mathcal{S}|}} d\mathbf{w} \exp\left[-\mathbf{w}^\top\mathbf{A}\mathbf{w} + \mathbf{b}^\top\mathbf{w}\right] = \frac{\pi^{|\mathcal{S}|/2}}{\sqrt{det(A)}} \exp\left(\frac{1}{4}\mathbf{b}^\top\mathbf{A}^{-1}\mathbf{b}\right), \tag{3.3}$$

where the integration element is given by

$$d\mathbf{w} = \prod_{s\in\mathcal{S}} dw_s.$$

For the integral in Equation (3.3) to exist, the matrix $\mathbf{A}$ must be positive definite. In the expression of the partition function we let the binary vector $\mathbf{x}$ play the role of the vector $\mathbf{b}$ of the above Gaussian identity. It follows that the partition function summation can be expressed as

$$\begin{aligned} Z &= \sum_{\mathbf{x}} \exp\left[\frac{1}{2T}\mathbf{x}^T\mathbf{J}\mathbf{x}\right] \\ &= C_1(det(\mathbf{J}), |\mathcal{S}|, T) \sum_{\mathbf{x}} \int_{\Re^{|\mathcal{S}|}} d\mathbf{w} \exp\left[-\mathbf{w}^T\left(\frac{2\mathbf{J}}{T}\right)^{-1}\mathbf{w} + \mathbf{x}^T\mathbf{w}\right], \end{aligned}$$

[1] For the integral of the right-hand side to exist, we must assume that the matrix $\mathbf{A}$ is positive definite. We make this assumption throughout Section 3.1

where [2] we have used the invertibility of the interaction coefficient matrix $\mathbf{J}$. Using easy induction on the size of the lattice, $|\mathcal{S}|$, one can prove the following identity

$$\sum_{(x_s)} \prod_{s\in\mathcal{S}} \exp(w_s x_s) = \prod_{s\in\mathcal{S}} \sum_{x_s=\pm 1} \exp(w_s x_s). \tag{3.4}$$

Also,

$$\sum_{x_s=\pm 1} \exp(x_s w_s) = 2\cosh w_s = 2\exp(\log\cosh w_s). \tag{3.5}$$

In the integrals used in the partition function, we introduce the change of variables

$$\mathbf{w} = \frac{1}{T}\mathbf{J}\mathbf{v}. \tag{3.6}$$

In terms of the new lattice variables v_s, $s \in \mathcal{S}$, we can write the partition function as [3] This equality results from (3.5) and from the change of variable (3.6) under the integration sign. Note that C_2 is, like C_1, independent of $\mathbf{w}$.

$$Z = C_2(det(\mathbf{J}), |\mathcal{S}|, T) \int_{\mathbb{R}^{|\mathcal{S}|}} d\mathbf{v} \exp\left[-\frac{1}{2T}\mathbf{v}^T\mathbf{J}\mathbf{v} + \sum_{s\in\mathcal{S}} \log\cosh\left(\frac{1}{T}\sum_{r\in\mathcal{S}} J_{sr}v_r\right)\right]. \tag{3.7}$$

The net result of all this algebra is the transformation of the partition function from a summation over the discrete variables, $\mathbf{x}$, into an integral over the continuous variables, $\mathbf{v}$. Such transformations can be done in the general case where the energy is not quadratic. More will be said on this in Section 3.3.1.

It is classical to call [28] the exponent of the exponential function in the Z integral the *effective* energy E_{eff}, so we have by definition

$$E_{eff}(\mathbf{v}) = \frac{1}{2}\mathbf{v}^T\mathbf{J}\mathbf{v} - T\sum_{s\in\mathcal{S}} \log\cosh\left(\frac{1}{T}\sum_{r\in\mathcal{S}} J_{sr}v_r\right). \tag{3.8}$$

Note how the second term in the expression of $E_{eff}(\mathbf{v})$ depends explicitly on temperature. We will show later that at the global minimum of E_{eff}, the effective energy has the same

[2] The constant $C_1(det(\mathbf{J}), |\mathcal{S}|, T)$ is independent of $\mathbf{w}$ and $\mathbf{x}$ and is explicitly given by the formula

$$C_1(det(\mathbf{J}), |\mathcal{S}|, T) = \sqrt{\left(\frac{T}{2\pi}\right)^{|\mathcal{S}|} / det(\mathbf{J})}.$$

[3] Here also, we have

$$C_2(det(\mathbf{J}), |\mathcal{S}|, T) = |det(\mathbf{J})| \left(\frac{2}{T}\right)^{|\mathcal{S}|} C_1(det(\mathbf{J}), |\mathcal{S}|, T).$$

thermodynamic structure as the Hemholtz free energy of the system. At this stage, it is worthwhile to mention that in general E_{eff} is *not* a convex function $\mathbf{v}$. However, in the limit of high temperatures, it is easy to show that the quadratic term will dominate in the expression of E_{eff}, and because $\mathbf{J}$ was assumed positive definite, one can conclude that the effective energy becomes convex as $T \to +\infty$.

3.1.2 Approximation of the Partition Function

The above integral expression of Z is exact but very complicated. The usual heuristic argument at this stage is to say that the integrand contribution to the integral comes chiefly from the region where the integrand has a global maximum, that is, the region where the effective energy has a global minimum, see Appendix B. This leads to the approximation

$$Z \approx C_3(det(\mathbf{H}^*), |\mathcal{S}|, T) \exp\left[-\frac{1}{T} E_{eff}(\mathbf{v}^*)\right], \tag{3.9}$$

where $\mathbf{v}^*$ is a point at which the effective energy reaches a global minimum,[4] and $\mathbf{H}^*$ is the Hessian of the effective energy computed at the global minimum $\mathbf{v}^*$. Note that this Hessian matrix is positive definite [5]. A necessary condition satisfied by $\mathbf{v}^*$ is of course that the gradient of E_{eff} vanish, i.e.,

$$\nabla E_{eff}(\mathbf{v}^*) = \mathbf{0}.$$

Differentiating $E_{eff}(\mathbf{v})$ with respect to v_q, $q \in \mathcal{S}$, we obtain

$$\sum_{s \in \mathcal{S}} J_{sq} \left[v_s - \tanh\left(\frac{1}{T} \sum_{r \in \mathcal{S}} J_{sr} v_r \right) \right] = 0,$$

which because of the invertibility of the matrix $\mathbf{J}$ gives

$$v_s = \tanh(\frac{1}{T} \sum_{r \in \mathcal{S}} J_{sr} v_r), \quad \forall s \in \mathcal{S}. \tag{3.10}$$

The global minimum point $\mathbf{v}^*$ belong to the set of solutions of the above system of nonlinear algebraic equations. This set might contain more than one solution because of the nonconvexity of the effective energy.

---

[4]The interest in this point will be explained in Subsection 3.1.4 where we will show that $\mathbf{v}^*$ can be considered an approximation of the true mean of $\mathbf{x}$ with respect to the Gibbs distribution.

[5]The explicit expression of $C_3(det(\mathbf{H}^*), |\mathcal{S}|, T)$ is given by

$$C_3(det(\mathbf{H}^*), |\mathcal{S}|, T) = C_2(|\mathcal{S}|, T, det(\mathbf{J}))(\pi T)^{|\mathcal{S}|/2} / \sqrt{det(\mathbf{H}^*)}.$$

The above system can be specialized to the case of the Ising model in which $J_{sr} = J$ if and only if r and q are a nearest-neighbor pair. Then we will have

$$v_s = \tanh(\beta \sum_{r \in \mathcal{N}_s} v_r), \quad \forall r \in \mathcal{S},$$

where again $\beta = J/T$. The parameter β controls the slope of the sigmoid function $x \rightarrow \tanh(\beta x)$. As $\beta \rightarrow \infty$, the sigmoid function approaches a sign function. This can happen when either the system temperature is too small or the interaction between pixels is too strong.

For the general case and in the limit of small temperatures, the iterative solution of the above equations gives

$$v_r^{(n+1)} = sign\left(\sum_{q \in \mathcal{S}} J_{rq} v_q^{(n)}\right),$$

which is identical to Hopfield's update rule for a neural network of binary neurons in which all the neuron thresholds are set to zero [35].

3.1.3 Effective Energy in the Presence of an External Field

Because of its importance in computing approximations for the moments of the Gibbs distribution, this paragraph is devoted to computing the effective energy and the partition function in the presence of an external field applied to the spins of the lattice. Although the algebra is lengthy and tedious, it is conceptually simple: it is based on a linear change of variables to profit from (3.3).

Let then $\mathbf{h} = (h_s, s \in \mathcal{S})$ be the external field applied to the lattice.[6] The energy of the system can be written as

$$E(\mathbf{x}) = -\frac{1}{2} \sum_{s,r \in \mathcal{S}} J_{sr} x_s x_r - \sum_{s \in \mathcal{S}} h_s x_s, \tag{3.11}$$

or in vector form as

$$E(\mathbf{x}) = -\frac{1}{2} \mathbf{x}^T \mathbf{J} \mathbf{x} - \mathbf{h}^T \mathbf{x}, \tag{3.12}$$

which can also be written as

$$E(\mathbf{x}) = -\frac{1}{2} (\mathbf{x} + \mathbf{J}^{-1}\mathbf{h})^T \mathbf{J} (\mathbf{x} + \mathbf{J}^{-1}\mathbf{h}) + \frac{1}{2} \mathbf{h}^T \mathbf{J}^{-1} \mathbf{h}. \tag{3.13}$$

[6] In the neural network literature, the h_s's can also represent a family of thresholds that bias the neuron outputs.

Now we make the change of variable

$$\mathbf{y} = \mathbf{x} + \mathbf{J}^{-1}\mathbf{h}$$

and write the partition function of the system whose energy is given by (3.13) as

$$Z(\mathbf{h}) = \exp\left[-\frac{1}{2T}\mathbf{h}^T\mathbf{J}\mathbf{h}\right]\sum_{\mathbf{y}}\exp\left[\frac{1}{2T}\mathbf{y}^T\mathbf{J}\mathbf{y}\right], \tag{3.14}$$

where the above notation is used to emphasize the dependence of the partition function on the external field $\mathbf{h}$. As in the previous paragraph we can apply (3.3) to transform the above expression into an integral

$$\begin{aligned}
Z(\mathbf{h}) \quad &= C_1(det(\mathbf{J}), |\mathcal{S}|, T) \quad \exp\left[\frac{-\mathbf{h}^T\mathbf{J}^{-1}\mathbf{h}}{2T}\right] \times \\
&\qquad \int_{\Re^{|S|}} d\mathbf{w} \sum_{\mathbf{y}} \exp\left[-\mathbf{w}^T\left(\frac{2\mathbf{J}}{T}\right)^{-1}\mathbf{w} + \mathbf{y}^T\mathbf{w}\right] \\
&= C_1(det(\mathbf{J}), |\mathcal{S}|, T) \quad \exp\left[\frac{-\mathbf{h}^T\mathbf{J}^{-1}\mathbf{h}}{2T}\right] \times \\
&\qquad \int_{\Re^{|S|}} d\mathbf{w} \exp\left[-\mathbf{w}^T\left(\frac{2\mathbf{J}}{T}\right)^{-1}\mathbf{w} + \mathbf{h}^T\mathbf{J}^{-1}\mathbf{w}\right] \sum_{\mathbf{x}} \exp\left[\mathbf{x}^T\mathbf{w}\right] \\
&= 2^{|\mathcal{S}|}C_1(det(\mathbf{J}), |\mathcal{S}|, T) \quad \exp\left[\frac{-\mathbf{h}^T\mathbf{J}^{-1}\mathbf{h}}{2T}\right] \int_{\Re^{|S|}} d\mathbf{w} \times \\
&\qquad \exp\left[-\mathbf{w}^T\left(\frac{2\mathbf{J}}{T}\right)^{-1}\mathbf{w} + \mathbf{h}^T\mathbf{J}^{-1}\mathbf{w} + \sum_{s\in\mathcal{S}} \log\cosh w_s\right].
\end{aligned}$$

Now we introduce the change of variable

$$\mathbf{w} = \frac{1}{T}(\mathbf{J}\mathbf{v} + \mathbf{h}),$$

and write the above partition function as

$$\begin{aligned}
Z(\mathbf{h}) \quad = \quad & C_2(det(\mathbf{J}), |\mathcal{S}|, T) \times \\
& \int_{\Re^{|S|}} d\mathbf{v} \exp\left(\frac{1}{T}\left[-\frac{1}{2}\mathbf{v}^T\mathbf{J}\mathbf{v} + T\sum_{s\in\mathcal{S}} \log\cosh\left[\frac{1}{T}\left(\sum_{r\in\mathcal{S}} J_{sr}v_r + h_s\right)\right]\right]\right)
\end{aligned} \tag{3.15}$$

Note that the constant C_2 is independent of the external field $\mathbf{h}$ and depends only on the determinant of $\mathbf{J}$, the temperature T, and the size of the lattice $|\mathcal{S}|$. This fact about C_2 will be important in the next section where we use Z to prove that the solutions of Equations (3.10) are approximations of the pixel means.

The effective energy of the lattice system under the external field is given by

$$E_{eff}(\mathbf{v},\mathbf{h}) = \frac{1}{2}\mathbf{v}^T\mathbf{J}\mathbf{v} - T\sum_{s\in\mathcal{S}} \log\cosh\left[\frac{1}{T}\left(\sum_{r\in\mathcal{S}} J_{sr}v_r + h_s\right)\right], \tag{3.16}$$

where the arguments of E_{eff} emphasize its dependence on both the internal field $\mathbf{v}$ and the external field $\mathbf{h}$. The difference with the $E_{eff}(\mathbf{v})$ of the zero field system is just in the bias applied, in the non-quadratic term, to the average internal field at each pixel. Note that we recover the zero external field case by setting $\mathbf{h}$ to zero in (3.16).

An approximation of the partition function can be obtained by replacing the integral in (3.15) by the maximum value of the integrand. This latter can be obtained by computing the effective energy at one of its minima. These can be obtained as solutions of the gradient equations

$$\nabla_{\mathbf{v}} E_{eff}(\mathbf{v}(\mathbf{h}),\mathbf{h}) = \mathbf{0}.$$

The notation $\mathbf{v}(\mathbf{h})$ is used to emphasize the dependence of the extreme point of the effective energy on the external field. Differentiating $E_{eff}(\mathbf{v},\mathbf{h})$ with respect to v_q, $q \in \mathcal{S}$, we obtain

$$\sum_{s\in\mathcal{S}} J_{sq}\left[v_s - \tanh\left[\frac{1}{T}\left(\sum_{r\in\mathcal{S}} J_{sr}v_r + h_s\right)\right]\right] = 0,$$

which because of the invertibility of the matrix $\mathbf{J}$ gives

$$v_s = \tanh\left[\frac{1}{T}\left(\sum_{r\in\mathcal{S}} J_{sr}v_r + h_s\right)\right], \quad \forall s \in \mathcal{S}. \tag{3.17}$$

Solving this system of nonlinear equations gives the point $\mathbf{v}(\mathbf{h})$. Note that because of the continuity of the solutions with respect to the external field we have

$$\lim_{\mathbf{h}\to\mathbf{0}} \mathbf{v}(\mathbf{h}) = \mathbf{v}^*.$$

As for the approximation of the partition function, it is given by [7]

$$Z(\mathbf{h}) \approx C_3(det(\mathbf{H}^*(\mathbf{h})), |\mathcal{S}|, T)(\exp\left[-\frac{1}{T}E_{eff}(\mathbf{v}(\mathbf{h}),\mathbf{h})\right]. \tag{3.18}$$

3.1.4 Approximation of the Mean-Field

Now we want to show that the vector $\mathbf{v}^*$ is an approximation for the mean vector $\overline{\mathbf{x}}$. Using the expression of the partition function in the presence of an external field affecting

[7] C_3 has the same functional form as C_2. Note however that C_3 depends on the external field $\mathbf{h}$. This is because the Hessian is computed at the global minimum $\mathbf{v}(\mathbf{h})$ which depends on $\mathbf{h}$.

the energy as described in (3.11), we can easily verify that the mean is exactly given by the formula

$$\overline{x_s} = T\left(\frac{\partial \log Z(\mathbf{h})}{\partial h_s}\right)_{\mathbf{h}=0}. \tag{3.19}$$

Using (3.15), we get

$$\log Z(\mathbf{h}) = \log C_2(det(\mathbf{J}), |\mathcal{S}|, T) + \log \int_{\Re^{|\mathcal{S}|}} d\mathbf{v} \exp\left[-\frac{1}{T} E_{eff}(\mathbf{v}, \mathbf{h})\right]$$

Differentiating the above expression with respect to each of the $h_s, s \in \mathcal{S}$, and using the fact that $C_2(det(\mathbf{J}), |\mathcal{S}|, T)$ is independent of $\mathbf{h}$, we obtain

$$\begin{aligned} \overline{x_s}(\mathbf{h}) &= T\left(\frac{\partial \log Z(\mathbf{h})}{\partial h_s}\right) \\ &= -\frac{\int_{\Re^{|\mathcal{S}|}} d\mathbf{v} \frac{\partial E_{eff}(\mathbf{v},\mathbf{h})}{\partial h_s} \exp\left[-\frac{1}{T} E_{eff}(\mathbf{v}, \mathbf{h})\right]}{\int_{\Re^{|\mathcal{S}|}} d\mathbf{v} \exp\left[-\frac{1}{T} E_{eff}(\mathbf{v}, \mathbf{h})\right]} \end{aligned}$$

Applying the saddle-point approximation method [8] to both the numerator and denominator of the right-hand side, we get [9]

$$\overline{x_s}(\mathbf{h}) \approx -\frac{\partial E_{eff}(\mathbf{v}^*(\mathbf{h}), \mathbf{h})}{\partial h_s} \tag{3.20}$$

But we have from (3.16)

$$-\frac{\partial E_{eff}(\mathbf{v}^*(\mathbf{h}), \mathbf{h})}{\partial h_s} = \tanh\left[\frac{1}{T}\left(\sum_{sr} J_{sr} v_r^*(\mathbf{h}) + h_s\right)\right] = v_s^*(\mathbf{h})$$

Computing the above equation and (3.20) at $\mathbf{h} = \mathbf{0}$, we get

$$\overline{x_s} = v_s^*,$$

which means that the point obtained by solving the set (3.10) of nonlinear equations is actually an approximation of the mean field value with respect to the Gibbs distribution. This situation is general, and it will be extended to grayscale image models, i.e, having more than two colors, and satisfying global constraints. It is worthwhile noting that the formula used to get the mean from the partition function can be generalized to get any moment. This is precisely what we have shown in the previous chapter 2 for the second-order moments.

[8] See Appendix B.

[9] Note that the scaling terms that result from the saddle-point method are the same and therefore cancel out.

3.1.5 Effective Energy and Free Energy

We have mentioned that the effective energy has the peculiarity of being explicitly dependent on temperature. This explicit dependence is found both in the expressions of Helmholtz's free energy and Gibbs's free energy. [10] We will show now that the expression of the effective energy *computed at a point*, $\mathbf{v}^*$, obtained as a solution of (3.10) is actually a mean-field approximation to the Helmholtz free energy F. Before proceeding with the algebraic proof, which is quite involved, let us first give an approximation of the entropy for the system of pixels whose states are defined by the mean-field equations given in (3.10). This approximation is based on the following

Lemma 8 *The entropy S of a system of binary $\in \{-1,+1\}$, independent, pixels with mean-values $\overline{\mathbf{x}} = (\overline{x}_s, s \in \mathcal{S})$, is given by*

$$S(\overline{\mathbf{x}}) = -\frac{1}{2}\sum_{s\in\mathcal{S}}\left[(1+\overline{x}_s)\log\frac{1}{2}(1+\overline{x}_s) + (1-\overline{x}_s)\log\frac{1}{2}(1-\overline{x}_s)\right]. \tag{3.21}$$

Proof: In order to define the entropy, we need to make the change of variable $\overline{x}_s \rightarrow \frac{1}{2}(1+\overline{x}_s)$, which maps $(-1,+1)$ onto $(0,1)$. The entropy of the source at site s can then be defined as [65]

$$S(\overline{x}_s) = -\left[\frac{1}{2}(1+\overline{x}_s)\log\frac{1}{2}(1+\overline{x}_s) + \frac{1}{2}(1-\overline{x}_s)\log\frac{1}{2}(1-\overline{x}_s)\right].$$

The entropy of all the lattice sources is the sum of the site entropies, i.e.,

$$S(\overline{\mathbf{x}}) = \sum_{s\in\mathcal{S}} S(\overline{x}_s) = -\frac{1}{2}\sum_{s\in\mathcal{S}}\left[(1+\overline{x}_s)\log\frac{1}{2}(1+\overline{x}_s) + (1-\overline{x}_s)\log\frac{1}{2}(1-\overline{x}_s)\right].$$

■

Now we give the main proposition of this paragraph

Proposition 8 *The effective energy $E_{eff}(\mathbf{v}^*)$ computed at the states given by (3.10) is a mean-field approximation of the Helmholtz free energy and is given by*

$$E_{eff}(\mathbf{v}^*) = E(\mathbf{v}^*) - TS(\mathbf{v}^*), \tag{3.22}$$

where $E(\mathbf{v}^*)$ is the Gibbs energy computed at $\mathbf{v}^*$ and $S(\mathbf{v}^*)$ is given by

$$S(\mathbf{v}^*) = -\frac{1}{2}\sum_{s\in\mathcal{S}}\left[(1+v_s^*)\log\frac{1}{2}(1+v_s^*) + (1-v_s^*)\log\frac{1}{2}(1-v_s^*)\right]$$

[10] The Helmholtz free energy, usually denoted by F, is given by $F = U - TS$, while the Gibbs free energy, G, is given by $G = H - TS = U + PV - TS$, where U, H, S, P, V, and T are respectively the system internal energy, enthalpy, entropy, pressure, volume, and temperature.

Proof: Let us denote by

$$w_s^* = \frac{1}{T}\sum_r J_{sr} v_r^*,$$

then the expression of the effective energy becomes

$$E_{eff}(\mathbf{v}^*) = \frac{T}{2}\mathbf{v}^{*T}\mathbf{w}^* - T\sum_{s\in\mathcal{S}} \log\cosh w_s^*. \tag{3.23}$$

From (3.10), we have

$$v_s^* = \tanh w_s^*,$$

or

$$w_s^* = \tanh^{-1} v_s^* = \frac{1}{2}\log\frac{1+v_s^*}{1-v_s^*},$$

which gives

$$\cosh w_s^* = \left(\frac{1+v_s^*}{1-v_s^*}\right)^{1/2} + \left(\frac{1-v_s^*}{1+v_s^*}\right)^{1/2} = \frac{2}{[(1-v_s^*)(1+v_s^*)]^{1/2}}$$

It follows that

$$\log\cosh w_s^* = -\frac{1}{2}[\log\frac{1}{2}(1-v_s^*) + \log\frac{1}{2}(1+v_s^*)]. \tag{3.24}$$

On the other hand, we can split the first term in the energy function as

$$\frac{T}{2}\mathbf{v}^{*T}\mathbf{u} = -\frac{1}{2}\mathbf{v}^{*T}\mathbf{J}\mathbf{v}^* + T\sum_{s\in\mathcal{S}} v_s^* w_s^*.$$

Therefore

$$E_{eff}(\mathbf{v}^*) = E(\mathbf{v}^*) + \frac{T}{2}\sum_{s\in\mathcal{S}} v_s^* \log\frac{1+v_s^*}{1-v_s^*} - T\sum_{s\in\mathcal{S}}\log\cosh w_s^*.$$

Substituting from (3.24) and expanding, we get after some algebra

$$E_{eff}(\mathbf{v}^*) = E(\mathbf{v}^*) + \frac{T}{2}\sum_{s\in\mathcal{S}}\left[(1+v_s^*)\log\frac{1}{2}(1+v_s^*) + (1-v_s^*)\log\frac{1}{2}(1-v_s^*)\right].$$

Using (3.21), the above equation becomes

$$E_{eff}(\mathbf{v}^*) = E(\mathbf{v}^*) - TS(\mathbf{v}^*).$$

■

Comparing the above expression of $E_{eff}(\mathbf{v}^*)$ with the expression

$$F = \overline{E} - TS$$

of Helmholtz free energy computed, we can conclude the following

1. The internal energy $\overline{E}$, which is the ensemble average of the Gibbs energy with respect to the Gibbs distribution, can be approximated with the Gibbs energy computed at the solution of the mean-field equations $E(\mathbf{v}^*)$. Note that this statement hides actually two approximations and not just one. The first is

$$\overline{E} \approx E(\overline{\mathbf{x}}),$$

and the second is

$$E(\overline{\mathbf{x}}) \approx E(\mathbf{v}^*).$$

2. Computed at the mean-field solution $\mathbf{v}^*$, the entropy $S(\mathbf{v}^*)$ is that of a system of binray, independent pixels with mean values $v_s^*, s \in \mathcal{S}$. In other words, the mean-field approximation amount to making an independence assumption on the pixel random variables.

3. Based on the above two remarks, we can approximate the Helmholtz free energy as

$$F \approx E_{eff}(\mathbf{v}^*) = E(\mathbf{v}^*) - TS(\mathbf{v}^*).$$

4. Finally, we find it appealing that the expression of the entropy derived from the effective energy is identical to the *information-theoretic* entropy of a set of binary independent sources with symbol probabilities $\frac{1}{2}(1 + v_s^*)$ and $\frac{1}{2}(1 - v_s^*)$.

3.1.6 Remarks

We end this subsection about the binary case with the following remarks.

1. The Gaussian formula (3.3) can be used only if the matrix $\mathbf{J}$ is positive definite. Note however that the expression of $E_{eff}(\mathbf{v})$ (3.8) does not require that $\mathbf{J}$ be positive definite or even invertible.

2. When the matrix $\mathbf{J}$ is positive definite, the mean-field approximation formulas are general, in the sense that no specific assumptions are made on the neighborhood structure, the homogeneity, or the isotropy of the field.

3. The validity of the mean-field approximation depends on both the lattice size and temperature. The higher the number of pixels and the smaller the temperature, the better the approximation. A heuristic argument of why this is so is given in Appendix B.

3.2 Unconstrained Multilevel Case

In the previous section, we dealt with an image model in which the pixel state space was binary. In this section, we treat the case of a Gibbs random field image model in which the energy function is given by

$$E(\mathbf{x}) = \sum_{s \in \mathcal{S}} \sum_{r \in \mathcal{N}_s} V(x_s, x_r), \tag{3.25}$$

where now $x_s \in G = \{0, 1, \ldots, n-1\}$, n being the number of graylevels. The above energy is derived from the general expression (1.6) where we have dropped the energy due to the external field and assumed that the pairwise interaction potentials are homogeneous, i.e., independent of the lattice site s, and isotropic, i.e., independent of the lattice direction from s to r. A major result of this section is the proof that the fixed point equations (3.10) of the binary case admit a natural generalization to the multilevel case. The crucial assumption that allows such a generalization to be made is the fact that only pairwise interactions are allowed among the different pixels. Another result is the definition of generalized sigmoid mappings that are the multidimensional counterparts of the single variable sigmoid function which is of widespread use in the neural network literature.

3.2.1 Effective Energy

In order to derive an approximation of the partition function Z, we need to compute the effective energy. We accomplish this by transforming the energy given in (3.25) into an expression similar to that of (3.1). This transformation is done via the use of binary matching variables [93].

Before doing this transformation, the following notation will be useful. Let $\mathcal{S}_g$ be the subset of $\mathcal{S}$ defined by $\{s \in \mathcal{S}, x_s = g\}$. They are the lattice subsets of constant graylevels, and they clearly form a partition of the lattice.

Let now $y_{sg} \in \{0, 1\}$ be a binary variable where the first index $s \in \mathcal{S}$ denotes a lattice site, while the second index $g \in G$ denotes a graylevel value. The variable v_{sg} takes one of the two values 0 or 1 according to the following rule: $v_{sg} = 1$, if $s \in \mathcal{S}_g$, and $v_{sg} = 0$, otherwise. The binary variable v_{sg} matches a site to a label, here a graylevel value. Since every site s has one and only one graylevel, the matching variables obey the following constraints

$$\sum_{g \in G} y_{sg} = 1, \quad \forall s \in \mathcal{S}. \tag{3.26}$$

In terms of these binary matching variables, the energy is given by

$$E(\mathbf{y}) = \sum_{g\in G}\sum_{g'\in G} V(g,g') \sum_{s\in\mathcal{S}}\sum_{r\in\mathcal{N}_s} y_{sg}y_{rg'}. \tag{3.27}$$

where $\mathbf{y}$ is now the vector of dimension $n|\mathcal{S}|$ of all binary matching variables $y_{sg}, s \in \mathcal{S}, g \in G$, and is defined as follows

$$\mathbf{y} = \left(\mathbf{y}_1^T, \ldots, \mathbf{y}_s^T, \ldots, \mathbf{y}_{|\mathcal{S}|}^T\right)^T,$$

where $\mathbf{y}_s = (y_{s0}, \ldots, y_{s(n-1)})^T$ is the vector of binary matching elements at site $s \in \mathcal{S}$. If T_n denotes the set of n-dimensional vectors

$$\mathbf{e}_a = (\delta_{1a}, \ldots, \delta_{na}), a \in G,$$

which form the canonical basis of $\Re^n$, then $\mathbf{y}_s \in T_n$, $\forall s \in \mathcal{S}$, and $\mathbf{y} \in (T_n)^{|\mathcal{S}|}$. The values $V(g,g')$ of the interaction potential defines an $n \times n$ matrix $\mathbf{V} = [V_{gg'}]_{g,g'\in G}$ that we call the *color interaction matrix*. We can retrieve an expression similar to (3.1) as follows. First define the matrix $\mathbf{L}$ by $L_{sr} = 1$ if $r \in \mathcal{N}_s$ and 0 otherwise.[11] Then the above energy function can be written as

$$E(\mathbf{y}) = -\frac{1}{2}\sum_{g,g'\in G}\sum_{s,r\in\mathcal{S}} (-2L_{sr}V_{gg'})\, y_{sg}y_{rg'}. \tag{3.28}$$

The above equation exhibits a clear decoupling between the *horizontal* connectivity between sites, and *vertical* connectivity between colors at each site. Note also that it is possible to start with an energy function defined as in (3.28) instead of (3.27), and to introduce lattice anisotropy by making the coefficients of the matrix $\mathbf{L}$ depend on the neighbor orientation. We will return to this point in Chapter 6.

We can go one step further and write the energy as

$$E(\mathbf{y}) = -\frac{1}{2}\mathbf{y}^T\mathbf{J}\mathbf{y} \tag{3.29}$$

with $-\frac{1}{2}\mathbf{J} = \mathbf{L} \otimes \mathbf{V}$, the outer, or Kronecker, product of the matrices $\mathbf{L}$ and $\mathbf{V}$ [40]. Explicitly, we have

$$J_{sr}^{gg'} = -2L_{sr}V_{gg'},\ \forall s,r \in \mathcal{S},\ \forall g,g' \in G.$$

We can now use (3.3) and write the partition function in terms of the integral [12]

[11] The matrix $\mathbf{L}$ is symmetric because of the assumed symmetry of the neighborhood system.

[12] The constant $K_1(det(\mathbf{J}), n|\mathcal{S}|, T)$ is independent of $\mathbf{w}$ and $\mathbf{y}$ and is explicitly given by the formula

$$K_1(det(\mathbf{J}), n|\mathcal{S}|, T) = \sqrt{\left(\frac{T}{2\pi}\right)^{n|\mathcal{S}|} / det(\mathbf{J})}.$$

$$Z = K_1(det(\mathbf{J}), n|\mathcal{S}|, T) \int_{\Re^{n|\mathcal{S}|}} d\mathbf{w} \exp\left[-\mathbf{w}^T \left(\frac{2\mathbf{J}}{T}\right)^{-1} \mathbf{w}\right] \sum_{\mathbf{y}} \exp\left[\mathbf{y}^T \mathbf{w}\right]. \tag{3.30}$$

The fundamental difference with the binary case of the previous subsection is in the computation of the summation term inside the integrand. This summation can be written as

$$\sum_{\mathbf{y}} \exp\left[\mathbf{y}^T \mathbf{w}\right] = \sum_{(\mathbf{y}_1 \in T_n)} \cdots \sum_{(\mathbf{y}_s \in T_n)} \cdots \sum_{(\mathbf{y}_{|\mathcal{S}|} \in T_n)} \prod_{s \in \mathcal{S}} \exp(\mathbf{y}_s^T \mathbf{w}_s). \tag{3.31}$$

Because of the uniqueness condition on the binary matching variable at each site, we can prove, using for instance induction on the number of sites, that the right-hand side of the above equation is equal to

$$\prod_{s \in \mathcal{S}} \sum_{g \in G} \exp(w_{sg}) = \exp\left[\sum_{s \in \mathcal{S}} \log\left(\sum_{g \in G} \exp(w_{sg})\right)\right]. \tag{3.32}$$

Substituting into (3.30), we get for the partition function the expression

$$Z = K_1(det(\mathbf{J}), n|\mathcal{S}|, T) \int_{\Re^{n|\mathcal{S}|}} d\mathbf{w} \exp\left[-\mathbf{w}^T \left(\frac{2\mathbf{J}}{T}\right)^{-1} \mathbf{w} + \sum_{s \in \mathcal{S}} \log\left(\sum_{g \in G} \exp(w_{sg})\right)\right]. \tag{3.33}$$

Changing variables according to

$$\mathbf{w} = \frac{1}{T} \mathbf{L} \otimes \mathbf{V}\mathbf{v}$$

we can write the effective energy of the graylevel system as

$$E_{eff}(\mathbf{v}) = \frac{1}{2} \mathbf{v}^T \mathbf{J} \mathbf{v} - T \sum_{s \in \mathcal{S}} \log\left[\sum_{g \in G} \exp\left(\frac{1}{T} \sum_{rg'} J_{sr}^{ab} v_{rg'}\right)\right]. \tag{3.34}$$

It is worthwhile to compare the above expression with the one for the binary case given in (3.8). For one thing, in both expressions, the quadratic interaction term is still governed by the matrix $\mathbf{J}$ of the original Gibbs system. For another, the arguments of the logarithmic terms are similar in that they are normalized summations of exponentials.

Again, we can approximate the Z partition function by replacing the integration with the maximum value of the integrand. These values are reached at the global minima of the effective energy $E_{eff}(\mathbf{v})$ which are among the solutions of

$$\nabla E_{eff}(\mathbf{v}^*) = \mathbf{0}.$$

Differentiating $E_{eff}(\mathbf{v})$ with respect to v_{sa}, $s \in \mathcal{S}$, and using the invertibiliy assumption on the matrix $\mathbf{J}$, we obtain

$$\begin{aligned} \mathbf{v}_s^* &= \mathbf{F}(\mathbf{w}_s^*), \ \forall s \in \mathcal{S} \\ \mathbf{w}_s^* &= \frac{1}{T}\sum_r \mathbf{J}_{sr}\mathbf{v}_r^* \end{aligned} \tag{3.35}$$

where $\mathbf{J}_{sr} = [J_{sr}^{ab}]_{(a,b)\in G^2}$ and $\mathbf{F} = (F_1, \ldots, F_n)$ is an infinitely differentiable mapping from $\Re^n$ onto the simplex

$$\mathcal{T}_n = \{\mathbf{z} \in \Re^n, \sum_{i=1}^n z_i = 1, \ z_i \geq 0, \ i = 1, \ldots, n\} \tag{3.36}$$

and is defined by

$$z_i = F_i(\mathbf{p}) = \frac{\exp(p_i)}{\sum_{k=1}^n \exp(p_k)}, \ i = 1, \ldots, n. \tag{3.37}$$

It is easy to check that $\mathbf{z}$ thus defined $\in \mathcal{T}_n$. Notice that the denominator in the definition of $\mathbf{F}$ is the same for all components. It is a normalizing constant that acts like a local partition function for the analog values of the binary matching variables. In the special case where $n = 2$ (the binary case), the mapping $\mathbf{F}$ reduces to the famous sigmoid function of Hopfield and backpropagation neural networks. This is why we call $\mathbf{F}$ the *generalized sigmoid mapping.* An in-depth study of the properties of the mapping $\mathbf{F}$ will be given in Chapter 5.

The similarities between the fixed point equations (3.10) of the binary case and the fixed point equations (3.35) of the n-dimensional case are striking. The scalar function tanh is now replaced by the n-dimesional mapping $\mathbf{F}$, and the variables at each site are now n-dimensional vectors rather than scalars. The argument of the tanh as well as that of $\mathbf{F}$ is still a linear combination of the effects that the other lattice sites have on the current pixel.

As in the binary case, an approximation of the partition function can be obtained as

$$Z \approx K_3(det(\mathbf{H}^*), n|\mathcal{S}|, T) \exp\left[-\frac{1}{T}E_{eff}(\mathbf{v}^*)\right], \tag{3.38}$$

where $\mathbf{H}^*$ is the Hessian of the effective energy computed at the global minimum $\mathbf{v}^*$ of the effective energy. Also, as in the binary case, we can show that this global minimum is indeed an approximation of the exact mean of the Gibbs system, i.e.,

$$\overline{\mathbf{y}_s} \approx \mathbf{v}_s^*, \ \forall s \in \mathcal{S}.$$

This fact, whose proof parallels that of the binary case, justifies calling (3.35) mean-field equations and the solution corresponding to the global minimum of the effective energy the mean-field approximation.

3.2.2 Effective Energy and Free Energy

In this paragraph, we relate the effective energy computed at the solutions of (3.35) to the Helmholtz free energy by explicitly computing the entropy at these solution points.

Before doing this computation, which is actually very similar to the one given for the binary case, we would like to mention that for each $s \in \mathcal{S}$, the solutions $v^*_{sg}, g \in G$, of the fixed point equations admit an interesting probabilistic interpretation. Indeed, from (3.35), it is clear that $v^*_{sg} > 0$ and that

$$\sum_{g \in G} v^*_{sg} = 1. \tag{3.39}$$

Therefore, v^*_{sg} can be considered the probability that pixel s on the lattice has color g. The pixel will be assigned the color having the highest probability. Using this probabilistic interpretation, the mean value of the graylevel x_s at pixel s can then be approximated as

$$\overline{x_s} \approx \sum_{g \in G} g v^*_{sg}.$$

The probabilistic nature of the analog matching variables will also appear in the entropy expression given in (3.41). First notice that because of (3.39) and the definition of w^*_{sg} in (3.35), the effective energy can be written as

$$E_{eff}(\mathbf{v}^*) = \frac{1}{2}\mathbf{v}^{*T}\mathbf{J}\mathbf{v}^* - T\sum_{s \in \mathcal{S}}\left(\sum_{g \in G} v^*_{sg}\right)\log\left(\sum_{g' \in G}\exp(w^*_{sg'})\right). \tag{3.40}$$

But

$$\log\sum_{g' \in G}\exp(w^*_{sg'}) = w_{sg} - \log v^*_{sg},\ \forall s \in \mathcal{S},\ \forall g \in G.$$

Substituting in (3.40), we get

$$\begin{aligned}
E_{eff}(\mathbf{v}^*) &= \frac{1}{2}\mathbf{v}^{*T}\mathbf{J}\mathbf{v}^* - T\sum_{s \in \mathcal{S}}\left(\sum_{g \in G} v^*_{sg}\right)\log\left(\sum_{g' \in G}\exp(w^*_{sg'})\right) \\
&= \frac{1}{2}\mathbf{v}^{*T}\mathbf{J}\mathbf{v}^* - T\sum_{s \in \mathcal{S}}\sum_{g \in G} v^*_{sg}(w^*_{sg} - \log v^*_{sg}) \\
&= \frac{1}{2}\mathbf{v}^{*T}\mathbf{J}\mathbf{v}^* - T\mathbf{v}^{*T}\mathbf{w}^* + T\sum_{s \in \mathcal{S}}\sum_{g \in G} v^*_{sg}\log v^*_{sg} \\
&= -\frac{1}{2}\mathbf{v}^{*T}\mathbf{J}\mathbf{v}^* + T\sum_{s \in \mathcal{S}}\sum_{g \in G} v^*_{sg}\log v^*_{sg} \qquad (3.41) \\
&= E(\mathbf{v}^*) - TS(\mathbf{v}^*)
\end{aligned}$$

The last equality above results from the fact that the entropy of a source s with n symbols having probabilities $v_{sg}^*, g \in G$, is given for each lattice site by

$$S_s(\mathbf{v}_s) = -\sum_{g \in G} v_{sg} \log v_{sg}$$

From (3.41) it follows that in the grayscale case, as in the binary case, the total entropy of the lattice is equal to the sum of the site entropies. This implies that in the limit of the mean-field approximation, the random variables (or vectors) assigned to the pixels are independent [65]. Later in this chapter, Section 3.3.2, we will show how starting from the assumption that the pixel random vectors are independent we can derive a mean-field approximation for the image model.

3.2.3 Generalization

The standing assumption on the energy function, in the previous subsection, was that the clique interaction potential between pixels V_{sr} is homogeneous, i.e., independent of the site s, and isotropic, i.e., independent of the site neighbors r. This assumption was made to introduce the tensor notation for the product between the lattice connectiveity matrix $\mathbf{L}$ and the color interaction matrix $\mathbf{V}$ – a notation that will come in very handy when we compute the *critical temperatures* for the mean-field phase transitions in Chapter 6. In case the interaction potential is not homogeneous or isotropic, we can still define a *tensor* $\mathbf{J}$ (rather than a matrix) by

$$J_{sr}^{ab} = L_{sr} V_{sr}(a, b), \ \forall s, r \in \mathcal{S}, \forall a, b \in G. \tag{3.42}$$

The $\mathbf{J}$ tensor subscripts s, r relate to the lattice action while the superscripts a, b relate to the interaction between graylevels. The tensor $\mathbf{J}$ satisfies the following symmetry relationship

$$J_{sr}^{ab} = J_{rs}^{ba} \tag{3.43}$$

As function of the binary matching elements the energy function is again quadratic, and is given by [13]

$$E(\mathbf{y}) = -\sum_{s,r \in \mathcal{S}} \sum_{a,b \in G} J_{sr}^{ab} y_{sa} y_{rb} = -\mathbf{y}^T \mathbf{J} \mathbf{y} \tag{3.44}$$

If the $\mathbf{J}$ tensor is positive definite we can apply the Gaussian trick (3.3) to derive both the effective energy and the mean-field equations. Here however, we will follow another

[13] Note that this energy function differs from that of (3.28) by a factor of $\frac{1}{2}$. This amounts to choosing $L_{sr} = -1$ if $r \in \mathcal{N}_s$.

path. We will show that if we are not directly interested in the effective energy of the system, or in other words, we do not care of which energy function the fixed-points are the extrema, then we can derive the mean-field equations directly from the partition function. One advantage of this calculation is that it will shed new light on the meaning of the mean-field equations and on the nature of the mean-field approximation of the partition function. Another advantage is that it leads to the mean-field equations even if the quadratic interaction tensor $\mathbf{J}$ is not positive definite.

We start with the following energy function

$$E(\mathbf{y},\mathbf{h}) = -\sum_{s,r\in\mathcal{S}}\sum_{a,b\in G} J_{sr}^{ab} y_{sa} y_{rb} - \sum_{s\in\mathcal{S}}\sum_{a\in G} h_{sa} y_{sa}, \tag{3.45}$$

where we impose the uniqueness constraint

$$\sum_{a\in G} y_{sa} = 1, \ \forall s \in \mathcal{S},$$

at each pixel s of the lattice. The above energy function contains a linear term to account for the excitation with an external field $\mathbf{h}$. The partition function of the lattice is given by

$$Z(\mathbf{h}) = \sum_{\mathbf{y}} \exp\left[-\frac{1}{T} E(\mathbf{y},\mathbf{h})\right], \tag{3.46}$$

where the summation is taken over all configurations satisfying the uniqueness constraint. Now consider a generic, binary matching variable y_{sa}. Then the effect of the rest of the binary matching variables[14] on y_{sa} is given by

$$f_{sa}(h_{sa}) = \sum_{r\in\mathcal{S}}\sum_{b\in G} J_{sr}^{ab} y_{rb} + h_{sa}, \tag{3.47}$$

so that the partition function can now be written as

$$Z(\mathbf{h}) = \sum_{\mathbf{y}} \exp\left[\frac{1}{T}\sum_{sa} y_{sa} f_{sa}\right],$$

which can also be written as

$$Z(\mathbf{h}) = \sum_{\mathbf{y}} \exp\left[\frac{1}{T}\sum_{s} \mathbf{y}_s^T \mathbf{f}_s\right],$$

where we have replaced the summation over the graylevels $a \in G$ by a dot product between the binary matching vector $\mathbf{y}_s$ and the field effect $\mathbf{f}_s$ at site s. Now again the

[14]Our derivation here assumes that the auto-connect matrices $\mathbf{J}_{ss}, s \in \mathcal{S}$, are all zero.

problem is how to compute $Z(\mathbf{h})$. An approximation $Z^*(\mathbf{h})$ of $Z(\mathbf{h})$ can be obtained if we replace the field effect $\mathbf{f}_s$ at each pixel s by an *estimate* of its mean

$$\mathbf{f}_s^*(\mathbf{h}) = \sum_{r \in \mathcal{S}} \mathbf{J}_{sr} \mathbf{v}_r^*(\mathbf{h}) + \mathbf{h}_s,$$

where $\mathbf{J}_{sr} = [J_{sr}^{ab}]_{a,b \in G}$ and $\mathbf{v}_r^*(\mathbf{h})$ is an estimate of the mean vector of $\mathbf{y}_r$ *in the presence* of the external field $\mathbf{h}$. Then

$$Z(\mathbf{h}) \approx Z^*(\mathbf{h}) = \sum_{\mathbf{y} \in (T_n)^{|\mathcal{S}|}} \exp\left[\frac{1}{T} \sum_s \mathbf{y}_s^T \mathbf{f}_s^*(\mathbf{h})\right].$$

Replacing $\mathbf{f}_s^*/T$ with $\mathbf{w}_s^*(\mathbf{h})$, the summation on the right-hand side of the above equation becomes exactly the one in (3.31), and we can use the result in (3.32) to write the partition function approximation as

$$Z^*(\mathbf{h}) = \prod_{s \in \mathcal{S}} \sum_{a \in G} \exp(w_{sa}^*(\mathbf{h})).$$

The problem of course is how to compute v_{sa}^*. This can be accomplished if we remember that the mean at any pixel can be obtained from the partition function by

$$\overline{y_{sa}} = T \left(\frac{\partial \log Z}{\partial h_{sa}}\right)_{\mathbf{h}=\mathbf{0}}$$

This relation defines a set of *consistency* equations for the mean estimates v_{sa}^* that are obtained from

$$v_{sa}^* = T \left(\frac{\partial \log \hat{Z}}{\partial h_{sa}}\right)_{\mathbf{h}=\mathbf{0}},$$

where we have denoted by v_{sa}^* the estimate $v_{sa}^*(\mathbf{0})$ at zero field. The above equation gives

$$v_{sa}^* = \frac{\exp(w_{sa}^*)}{\sum_{b \in G} \exp(w_{sb}^*)}.$$

Using the function $\mathbf{F}$ defined in (3.37), we get for the $\mathbf{v}_s^*$'s a set of fixed point equations similar to those of (3.35), i.e.,

$$\begin{aligned} \mathbf{v}_s^* &= \mathbf{F}(\mathbf{w}_s^*), \\ \mathbf{w}_s^* &= \frac{1}{T}\left(\sum_{r \in \mathcal{S}} \mathbf{J}_{sr} \mathbf{v}_r^*\right), \quad \forall s \in \mathcal{S}. \end{aligned} \tag{3.48}$$

Of course, when the tensor $\mathbf{J} = \mathbf{L} \otimes \mathbf{V}$, we get exactly the fixed point equations of (3.35).

Having shown a different way for getting estimates for the means v_{sa}, let us contrast this method with the one that led to (3.35).

1. The second method is more direct. However, it produces less. In particular, it does not tell us whether the fixed points of (3.48) are the extrema of some energy function. We know by the first method that when the $\mathbf{J}$ is positive definite, they are extrema of the effective energy (3.34).

2. The second method shows indeed that at the solutions of (3.35), the partition function is approximated by a product, over the lattice sites, of *local* functions, namely,

$$\sum_{a \in G} \exp\left[\frac{1}{T} \sum_{r \in S} \sum_{b \in G} J_{sr}^{ab} v_{sa}^{*}\right].$$

This would have been the case had we assumed that the Gibbs probability distribution $P(\mathbf{y})$ can be decomposed into a product, over the lattice sites, of single-site probability distributions, i.e., that the pixels states are independent random variables (or vectors).

3. It is not clear how global constraints on the patterns can be introduced in the direct method. We will see in the next section that there is an easy way for incorporating them in the effective-energy method to derive the mean-field equations.

3.3 Constrained Multilevel Case

3.3.1 Saddle Point Method

As it was mentioned at the end of the previous section, the methods based on the Gaussian integration formula (3.3) will fail if the matrix $\mathbf{J}$ is not positive definite. They also fail if the energy function itself cannot be transformed into a quadratic form. We now introduce a method that could handle these situations. We will see in particular that it will be possible to get the effective energy for *any* energy function $E(\mathbf{x})$, and that it will be possible to incorporate, in the formulation, global constraints on the different patterns. Here also, an approximation of the partition function can be had, and as usual, this approximation requires the knowledge of the minima of the effective energy function, minima that can be obtained from a set of fixed point equations. A very important spinoff of this method is that it provides us with a systematic procedure to transform discrete, nonlinear, constrained optimization problems into continuous nonlinear programs that can be solved using continuation methods [64]

Our derivation of these equations is inspired by, but slightly different from, the one in [68]. The basic differences are in the way we handle the constraints and apply the

heuristics of the saddle-point method. To prepare the reader for the long algebra that will follow, we will give the main steps of this method:

1. First, express the constraints on the matching vector using delta functions.
2. Transform the summation in the partition function into an integral.
3. Define the effective energy and use the saddle point method to approximate the partition function.
4. Get the mean field equations for the binary matching vector as extrema the effective energy.

The binary matching variables will be as usual denoted by the vector $\mathbf{y}$. The energy function of the lattice will now be assumed an *arbitrary* function $E(\mathbf{y})$ of the matching vector $\mathbf{y}$. In addition to the uniqueness constraints

$$\sum_{g \in G} y_{sg} = 1, \quad \forall s \in \mathcal{S}, \tag{3.49}$$

we will assume that the number of sites in each graylevel set $\mathcal{S}_g$ is fixed and that they are all equal to the same number

$$\sum_{s \in \mathcal{S}} y_{sg} = |\mathcal{S}_g| = \gamma, \quad \forall g \in G, \tag{3.50}$$

which is equivalent to saying that the pattern histogram is uniform. The partition function of this system is given by

$$Z = \sum_{\mathbf{y}} e^{-\frac{1}{T}E(\mathbf{y})}, \tag{3.51}$$

where the summation is taken over all configurations satisfying the uniqueness and uniform histogram constraints.

An early version of this section appeared in [23].

Constraint Representation

The introduction of the constraints as delta functions is based on the following trick. [15] Each term in the partition function summation can be written as

$$e^{-\frac{1}{T}E(\mathbf{y})} = \int_{\mathbf{v}} e^{-\frac{1}{T}E(\mathbf{v})} \delta(\mathbf{y} - \mathbf{v}) d\mathbf{v} \tag{3.52}$$

[15] This trick is pretty standard in statistical physics. See, for instance, Reif [78].

The dummy integration vector $\mathbf{v}$, expressing the domain of the matching vector $\mathbf{y}$, should satisfy the uniform histogram constraints (3.50). These can be included in the above integral using delta functions as follows

$$e^{-\frac{1}{T}E(\mathbf{y})} = \int_{\mathbf{v}} e^{-\frac{1}{T}E(\mathbf{v})}\delta(\mathbf{y}-\mathbf{v}) \prod_{g\in G} \delta(z_g)d\mathbf{v} \tag{3.53}$$

where

$$z_g = -\gamma + \sum_{s\in\mathcal{S}} v_{sg},\ \forall g \in G.$$

Note that because of the presence of the delta functions in the above integrand of (3.53), the value of the integral does not change if we include in the integrand the functions $\exp[\mathbf{w}^T(\mathbf{y}-\mathbf{v})]$ and $\exp(\mathbf{p}^T\mathbf{z})$, where $\mathbf{w}$ and $\mathbf{p}$ are arbitrary vectors in $\Re^{n|\mathcal{S}|}$ and $\Re^n$, respectively. With these additional functions, Equation (3.53) becomes

$$e^{-\frac{1}{T}E(\mathbf{y})} = \int_{\mathbf{v}} e^{-\frac{1}{T}E(\mathbf{v})}e^{\mathbf{w}^T(\mathbf{y}-\mathbf{v})}\delta(\mathbf{y}-\mathbf{v}) \prod_{g\in G} e^{p_g z_g}\delta(z_g)d\mathbf{v}. \tag{3.54}$$

The delta functions can also be expressed as integrals using the known fact that a delta function is the inverse Fourier transform of a function identically equal to unity [16]. In other words, let $\mathbf{u} \in \Re^N$, then

$$\delta(\mathbf{u}) = \int_{\omega} e^{i\omega^T\mathbf{u}}d\omega, \tag{3.55}$$

where $\omega^T\mathbf{u}$ is the scalar product of the frequency vector ω and the space vector $\mathbf{u}$, and $i^2 = -1$. Applied to the delta functions introduced above, equation (3.55) gives

$$\delta(\mathbf{v}-\mathbf{y}) = \int_{\mathbf{b}} e^{i\mathbf{b}^T(\mathbf{v}-\mathbf{y})}d\mathbf{b}, \tag{3.56}$$

and

$$\prod_{g\in G} \delta(z_g) = \int_{\mathbf{c}} e^{i\mathbf{c}^T\mathbf{z}}d\mathbf{c}, \tag{3.57}$$

Substituting (3.56) and (3.57) into (3.54), we get the triple integral

$$e^{-\frac{1}{T}E(\mathbf{y})} = \int_{\mathbf{v}}\int_{\mathbf{b}}\int_{\mathbf{c}} e^{-\frac{1}{T}E(\mathbf{v})}e^{(\mathbf{w}+i\mathbf{b})^T(\mathbf{y}-\mathbf{v})}e^{(\mathbf{p}+i\mathbf{c})^T\mathbf{z}}d\mathbf{v}d\mathbf{b}d\mathbf{c}. \tag{3.58}$$

Only the uniform histogram constraint was included in this integration. The uniqueness constraint (3.49) will be imposed during the actual computation of the partition function in the next subsection.

[16] Again, we are doing these algebraic manipulations the same way they are done in statistical physics textbooks, where (3.55) is usually called the integral representation of the delta function [78]. Note that we have dropped the $\frac{1}{2\pi}$ factor of the Fourier integral.

Partition Function

The partition function Z is obtained by summing (3.58) over all configurations $\mathbf{y}$. Permuting the summation and integration signs (this is legitimate, since we have only a finite number of terms in the summation), we get

$$Z = \int_{\mathbf{v}} \int_{\mathbf{b}} \int_{\mathbf{c}} e^{-\frac{1}{T}E(\mathbf{v})} e^{-(\mathbf{w}+i\mathbf{b})^T\mathbf{v}} e^{(\mathbf{p}+i\mathbf{c})^T\mathbf{z}} \left(\sum_{\mathbf{y}} e^{(\mathbf{w}+i\mathbf{b})^T\mathbf{y}} \right) d\mathbf{v}d\mathbf{b}d\mathbf{c}. \tag{3.59}$$

In order to compute the summation over $\mathbf{y}$ inside the integral, we need the following equality

$$\sum_{\mathbf{y}} e^{(\mathbf{w}+i\mathbf{b})^T\mathbf{y}} = \prod_{s\in\mathcal{S}} \sum_{g\in G} e^{w_{sg}+ib_{sg}}, \tag{3.60}$$

which can be easily proved using the uniqueness constraint on each $\mathbf{y}_s$. Note that the right-hand side of (3.60) can also be written as

$$\exp\left[\sum_{s\in\mathcal{S}} \log\left(\sum_{g\in G} e^{w_{sg}+ib_{sg}} \right) \right]. \tag{3.61}$$

Using (3.59) and (3.61), we can now write the partition function in integral form as

$$\begin{aligned} Z \;=\; & \int_{\mathbf{v}} \int_{\mathbf{b}} \int_{\mathbf{c}} d\mathbf{v}d\mathbf{b}d\mathbf{c} \times \\ & \exp\left[-\frac{1}{T}E(\mathbf{v}) - (\mathbf{w}+i\mathbf{b})^T\mathbf{v} + \sum_{s\in\mathcal{S}} \log\left(\sum_{g\in G} e^{w_{sg}+ib_{sg}} \right) + (\mathbf{p}+i\mathbf{c})^T\mathbf{z} \right] \end{aligned} \tag{3.62}$$

Effective Energy

The exponent of the exponential in the previous expression is analytic in the complex vectors $\mathbf{w}+i\mathbf{b}$ and $\mathbf{p}+i\mathbf{c}$. The integrations with respect to the vectors $\mathbf{b}$ and $\mathbf{c}$ are contour integrations in the complex vector spaces $\mathbf{C}^{n|\mathcal{S}|}$ and $\mathbf{C}^n$, respectively. Because of the analyticity of the integrand, the value of the integral will not be affected if we deform the integration contours. In particular, we can have these contours pass through the real vectors $\mathbf{v}^*$, $\mathbf{w}^*$ and $\mathbf{p}^*$ where the *effective energy*

$$E_{eff}(\mathbf{v},\mathbf{w},\mathbf{p}) = E(\mathbf{v}) + T\mathbf{w}^T\mathbf{v} - T\sum_{s\in\mathcal{S}} \log\left(\sum_{g\in G} e^{w_{sg}} \right) - T\mathbf{p}^T\mathbf{z} \tag{3.63}$$

reaches a global minimum. Therefore, the contour integrals are taken along $\mathbf{w}^* + \Re^{n|\mathcal{S}|}$ for $\mathbf{b}$ and $\mathbf{p}^* + \Re^n$ for $\mathbf{c}$. Because the oscillations introduced by the terms $i\mathbf{b}^T\mathbf{v}^*$ and $i\mathbf{c}^T\mathbf{p}^*$ at large values of $\mathbf{b}$ and $\mathbf{c}$ tend to cancel out contributions to the integral, the maximum contributions to the Z partition function comes from the neighborhoods of $\mathbf{b} = \mathbf{0}$ and

$\mathbf{c} = \mathbf{0}$. It follows [17] that the partition function can be approximated by computing the value of the effective energy at the global minimum

$$Z \approx \hat{Z} = \kappa \exp\left[-\frac{1}{T} E_{eff}(\mathbf{v}^*, \mathbf{w}^*, \mathbf{p}^*)\right]. \tag{3.64}$$

Looking back at the expression of the effective energy (3.63), we could make the following remarks:

1. The uniform histogram constraints are being accounted for through the term containing the vector $\mathbf{p}$. Note that this vector acts like a Lagrange vector multiplier for this set of constraints.

2. The expression of the effective energy contains, as a special case, the one derived for a quadratic energy function in Section 3.2. Indeed, for the quadratic case we have

   $$\mathbf{w} = \frac{1}{T}\mathbf{J}\mathbf{v},$$

   and therefore we can rewrite Equation (3.34) as

   $$E_{eff}(\mathbf{v}) = E(\mathbf{v}) + T\mathbf{w}^T\mathbf{v} - T\sum_{s\in\mathcal{S}} \log\left(\sum_{g\in G} e^{w_{sg}}\right),$$

   which is similar to (3.63).

The above two remarks suggest that although the algebraic steps that have led to (3.63) are rather hard to justify rigorously, they lead to results that agree with the ones obtained previously.

Mean-Field Equations

It was mentioned in the previous paragraph that the partition function can be approximated by (3.64). In order to compute the right-hand side of this formula, we need to compute the point$(\mathbf{v}^*, \mathbf{w}^*, \mathbf{p}^*)$, where E_{eff} has a saddle point. A necessary condition satisfied by this point is

$$\frac{\partial E_{eff}}{\partial \mathbf{v}} = \frac{\partial E_{eff}}{\partial \mathbf{w}} = \frac{\partial E_{eff}}{\partial \mathbf{p}} = \mathbf{0}. \tag{3.65}$$

[17] It should be clear that the above statements do not constitute a proof. They are just a plausibility argument to justify dropping the dependence on the imaginary parts of the constraint variables.

Computing these partial derivatives with respect to v_{ka}, w_{ka} and p_a, respectively, where $k \in \mathcal{S}$ and $a \in G$, the fixed point equations become

$$\frac{\partial E_{eff}}{\partial v_{ka}} = \frac{\partial E}{\partial v_{ka}} + Tw_{ka} - Tp_a = 0 \tag{3.66}$$

$$\frac{\partial E_{eff}}{\partial w_{ka}} = v_{ka} - \frac{e^{w_{ka}}}{\sum_{b \in G} e^{w_{kb}}} = 0 \tag{3.67}$$

$$\frac{\partial E_{eff}}{\partial p_a} = \gamma - \sum_{k \in \mathcal{S}} v_{ka} = 0 \tag{3.68}$$

Note that equation (3.67) implies that

$$\sum_{a \in G} v_{ka} = 1, \ \forall k \in \mathcal{S},$$

which is the uniqueness constraint as applied to the *analog* matching variables. At every pixel of the lattice the analog matching variables are constrained to lie in the simplex $\mathcal{T}_n$ defined in (3.36). This has the effect of reducing the dimensionality of the solution space at each site by one. Note also that the fraction term in (3.67) is exactly the same as the one defined in (3.37). Equation (3.68) is nothing but the uniform histogram constraints as applied to the mean-field values of the binary matching variables.

Using vector notations for each site, we can write the mean-field equations (3.66), (3.67), and (3.68) in an more compact form as follows:

$$\frac{\partial E_{eff}}{\partial \mathbf{v}_s} = \nabla_s E(\mathbf{v}) - T\mathbf{w}_s - T\mathbf{p} = \mathbf{0}$$

$$\frac{\partial E_{eff}}{\partial \mathbf{w}_s} = \mathbf{v}_s - \mathbf{F}(\mathbf{w}_s) = \mathbf{0}$$

$$\frac{\partial E_{eff}}{\partial \mathbf{p}} = \mathbf{\Gamma} - \sum_{s \in \mathcal{S}} \mathbf{v}_s = \mathbf{0},$$

where in the first equation, $\nabla_s E(\mathbf{v})$ is the gradient of the energy function with respect to the analog matching vector at site s, and in the last equation $\mathbf{\Gamma}$ is the n-dimensional vector $(\gamma, \ldots, \gamma)^T$. If we eliminate the vector $\mathbf{w}_s$ between the first and the second equation, we obtain the following system of algebraic equations where the unknowns are the analog matching variables $\mathbf{v}_s, s \in \mathcal{S}$, and the Lagrange vector multiplier $\mathbf{p}$ associated with the uniform histogram constraint:

$$\mathbf{v}_s = \mathbf{F}\left(\frac{1}{T}\nabla_s E(\mathbf{v}) - \mathbf{p}\right) \tag{3.69}$$

$$\sum_{s \in \mathcal{S}} \mathbf{v}_s = \mathbf{\Gamma}.$$

The vector $\mathbf{p}$ acts in (3.69) as a constant external field. When there is no uniform histogram constraint ($\mathbf{p} = 0$), (3.69) becomes the natural generalization of (3.69) to

models with arbitrary energy functions. It is easy to check that we get (3.35) back from (3.69) if we assume that $E(\mathbf{y})$ is quadratic and given by, for instance, (3.44).

Using the formula

$$\overline{\mathbf{y}_s} = T\left(\frac{\partial \log Z(\mathbf{h})}{\partial h_s}\right)_{\mathbf{h}=0},$$

it is easy to see that the solutions of (3.69) give, as in the quadratic case, approximations for the means $\overline{\mathbf{y}_s}, s \in \mathcal{S}$.

Free Energy and Entropy

At a solution of (3.69), the uniform histogram constraint is exactly satisfied so that

$$E_{eff}(\mathbf{v}^*, \mathbf{w}^*, \mathbf{p}^*) = E(\mathbf{v}^*) + T\mathbf{v}^{*T}\mathbf{w}^* - T\sum_{s\in\mathcal{S}} \log\left(\sum_{g\in G} e^{w^*_{sg}}\right).$$

Moreover the uniqueness constraint is also satisfied, therefore

$$E_{eff}(\mathbf{v}^*, \mathbf{w}^*, \mathbf{p}^*) = E(\mathbf{v}^*) + T\mathbf{v}^{*T}\mathbf{w}^* - T\sum_{s\in\mathcal{S}}\left(\sum_{g\in G} v^*_{sg}\right) \log\left(\sum_{g'\in G} e^{w^*_{sg'}}\right).$$

But since $\mathbf{v}^*_s = F(\mathbf{w}^*_s)$, $\forall s \in \mathcal{S}$, we get that

$$\log\left(\sum_{g\in G} e^{w^*_{sg}}\right) = w^*_{sg} - \log v^*_{sg},$$

Substituting into the effective energy, we obtain

$$E_{eff}(\mathbf{v}^*, \mathbf{w}^*, \mathbf{p}^*) = E(\mathbf{v}^*) + T\sum_{s\in\mathcal{S}}\sum_{g\in G} v^*_{sg} \log v^*_{sg} \tag{3.70}$$

Noting that the second term of the right-hand side is the negative entropy at a fixed point $\mathbf{v}^*$, the effective energy becomes simply

$$E_{eff}(\mathbf{v}^*, \mathbf{w}^*, \mathbf{p}^*) = E(\mathbf{v}^*) - TS(\mathbf{v}^*).$$

It follows that here also, the effective energy is identical to a mean-field approximation to the Helmholtz free energy of the system at the fixed point. Observe that the expression of the entropy is exactly the one for $|\mathcal{S}|$ independent sources, each one of them producing n symbols with probabilities specified by $v^*_{sg}, g \in G$..

3.3.2 Probability Factorization Method

In the previous section, we have presented a method for deriving a set of algebraic equations that can be used to compute approximations of the pixel means of a Gibbs distribution having a discrete, finite scale of graylevels. The main characteristics of this method are that it allows us to incorporate global constraints and leads to a natural parameterization of the analog matching variables as outputs of the generalized sigmoid mapping $\mathbf{F}$. We have also shown that at the solutions of the algebraic equations (3.69), the values of the effective energy approximate the free energy and that the entropy is the same as the one that would have been obtained had the pixel states been independent random variables.

In this section, we look at the problem from the other end. What if we *start* by making the approximation that pixel states are independent random variables or vectors? What does the expression of the free energy become? And how can we use it to find mean-field approximations similar to the ones given by equations (3.69)?

The free energy is defined as $F = \overline{E} - TS$ [78]. Here, $\overline{E}$ is the internal energy, i.e., the ensemble mean of the energy with respect to the Gibbs distribution, and S is the entropy, defined below, of the GRF model. Our objective is to obtain an approximation for both $\overline{E}$ and S.

Entropy Approximation

The exact expression [18] for the entropy is

$$S = - \sum_{\mathbf{y} \in (T_n)^{|\mathcal{S}|}} P(\mathbf{y}) \ln P(\mathbf{y}), \tag{3.71}$$

where $(T_n)^{|\mathcal{S}|}$ is the set of all lattice configurations [19] and $P(\mathbf{y})$ is the probability of a configuration $\mathbf{y} \in (T_n)^{|\mathcal{S}|}$. Note that $\mathbf{y}$ is exactly given by the binary matching vector $\mathbf{y}^T = (\mathbf{y}_1^T, \ldots, \mathbf{y}_{|\mathcal{S}|}^T)$. Therefore we can write that $P(\mathbf{y}) = P(\mathbf{y}_1, \ldots, \mathbf{y}_{|\mathcal{S}|})$. If we make the *approximation* that the binary matching vectors of any two different sites are independent random vectors, then the joint probability distribution of $\mathbf{y}$ can be *approximated* by a product of $|\mathcal{S}|$ independent probability distributions each of which dependent on one and

[18] In statistical physics, the Boltzman constant, k, is put in front of (3.71) to give entropy the units of energy per degree of temperature. In this thesis, units are chosen so as to make $k = 1$.

[19] The reader is reminded that T_n, the set of vertices of the simplex $\mathcal{T}_n$, is the configuration space of the binary matching vector $\mathbf{y}_s$. T_n has n configurations corresponding to the unit vectors $\mathbf{e}_a, a \in G$ that form the canonical basis of $\Re^n$.

only one binary matching vector, i.e.,

$$P(\mathbf{y}) = P(\mathbf{y}_1, \ldots, \mathbf{y}_{|\mathcal{S}|}) \approx \prod_{s \in \mathcal{S}} P(\mathbf{y}_s). \tag{3.72}$$

Given the fact that the entropy of two independent sources is the sum of their entropies [65], we can write

$$S \approx -\sum_{s \in \mathcal{S}} \sum_{\mathbf{y}_s \in T_n} P(\mathbf{y}_s) \ln P(\mathbf{y}_s) \tag{3.73}$$

Let now v_{sa} be the probability of the event $\mathbf{v}_s = \mathbf{e}_a$, then the mean vector at site s is given by

$$\mathbf{v}_s = \overline{\mathbf{y}}_s = \sum_{a \in G} v_{sa} \mathbf{e}_a, \tag{3.74}$$

with $\sum_{a \in G} v_{sa} = 1$, and the entropy of the random vector $\mathbf{v}_s$ is given in [65]

$$-\sum_{a \in G} v_{sa} \ln v_{sa}. \tag{3.75}$$

Substituting into (3.73), we get the following approximation for the entropy

$$S \approx -\sum_{s \in \mathcal{S}} \sum_{a \in G} v_{sa} \ln v_{sa}. \tag{3.76}$$

Internal Energy Approximation

An approximation of the internal energy $\overline{E}$ can be obtained using the same probability factorization idea. Indeed for a *quadratic* energy function, we have successively

$$\begin{aligned}
\overline{E} &= \sum_{\mathbf{y} \in \Omega} E(\mathbf{y}) P(\mathbf{y}) \\
&= \sum_{\mathbf{y} \in (T_n)^{|\mathcal{S}|}} E(\mathbf{y}_1, \ldots, \mathbf{y}_{|\mathcal{S}|}) P(\mathbf{y}_1, \ldots, \mathbf{y}_{|\mathcal{S}|}) \\
&\approx \sum_{\mathbf{y} \in (T_n)^{|\mathcal{S}|}} E(\mathbf{y}_1, \ldots, \mathbf{y}_{|\mathcal{S}|}) \prod_{q \in \mathcal{S}} P(\mathbf{y}_q) \\
&\approx \sum_{y_1 \in T_n} \cdots \sum_{y_{|\mathcal{S}|} \in T_n} \sum_{s,r \in \mathcal{S}} J_{sr} \mathbf{y}_s^T \mathbf{V} \mathbf{y}_r \prod_{q \in \mathcal{S}} P(\mathbf{y}_q) \\
&\approx \sum_{s,r \in \mathcal{S}} J_{sr} \mathbf{v}_s^T \mathbf{V} \mathbf{v}_r = E(\mathbf{v}).
\end{aligned} \tag{3.77}$$

In other words, the ensemble average of the energy can be approximated by the energy computed at the ensemble average of the vector $\mathbf{y}$. If the energy function is *not* quadratic, we can still make the double approximation $\overline{E(\mathbf{y})} \approx E(\overline{\mathbf{y}}) \approx E(\mathbf{v})$, but this is only a zero-order approximation that assumes that the fluctuations of the random vectors around their means are extremely small.

Free Energy Approximation

Combining (3.76) with (3.77), we get an approximation of the free energy

$$F = \overline{E} - TS \approx E(\mathbf{v}) + T \sum_{s \in \mathcal{S}} \sum_{a \in G} v_{sa} \log v_{sa}. \tag{3.78}$$

Observe that this approximation is the same as the one obtained from the saddle-point method (3.70). It follows that another way for getting the mean-field approximation would be to find the minima of

$$E(\mathbf{v}) + T \sum_{s \in \mathcal{S}} \sum_{a \in G} v_{sa} \log v_{sa},$$

subject to the uniqueness constraints and uniform histogram constraints. This leads to defining the effective energy, actually the Lagrangian, by

$$\begin{aligned} E_{eff}[\mathbf{v}, \mathbf{p}, \mathbf{q}] &= E(\mathbf{v}) + T \sum_{sa} v_{sa} \log v_{sa} \\ &+ \sum_{a \in G} p_a (\sum_{s} v_{sa} - \gamma) + \sum_{s \in \mathcal{S}} q_s (\sum_{a} v_{sa} - 1), \end{aligned} \tag{3.79}$$

where the $\mathbf{p} = (p_a, a \in G)$ and $\mathbf{q} = (q_s, s \in \mathcal{S})$ are the Lagrange vector multipliers corresponding to the uniform histogram and the uniqueness constraint, respectively.

Mean-Field Equations

The extrema of the effective energy function 3.79 are given by

$$\frac{\partial E_{eff}}{\partial v_{ka}} = \frac{\partial E}{\partial v_{ka}} + T(1 + \log v_{ka}) + q_k + p_a = 0 \tag{3.80}$$

$$\frac{\partial E_{eff}}{\partial q_s} = -1 + \sum_{a \in G} v_{sa} = 0 \tag{3.81}$$

$$\frac{\partial E_{eff}}{\partial p_a} = -\gamma + \sum_{k \in \mathcal{S}} v_{ka} = 0 \tag{3.82}$$

The second and the third sets of equations represent the uniqueness and histogram constraints. In Chapter 6, we will use these equations to derive estimates of the critical temperatures corresponding to mean-field phase transitions. For now, the reader should note that these equations are different from the ones obtained using the saddle-point method (3.69) . For one thing, the generalized sigmoid mapping does not appear in the new set of equations. This is because the uniqueness constraints were imposed through Lagrange multipliers rather than by directly computing the partition function.

3.4 On the Legendre Transformation

In classical mechanics, the Legendre transformation is used to derive the Hamiltonian formalism from the Lagrangian formalism. In classical thermodynamics, it is used, among many other instances, to pass from an internal energy representation to an enthalpy representation, or from Helmholtz's free energy to Gibbs's. Geometrically, there is nothing to the Legendre transformation more than the idea that a curve in the plane could be either represented by a locus of points or as an envelop of a family of lines, and that the two representations are equivalent. [20]

In this short section, we will use the Legendre transformation to show how we can obtain equivalent representations for the mean-field systems described by the effective energy expressions that we have derived in this chapter. We restrict ourselves to the system described by an arbitrary energy function $E(\mathbf{y})$, Section 3.3.1, but without imposing any global constraint. In this case the effective energy is given by

$$E_{eff}(\mathbf{v},\mathbf{w}) = E(\mathbf{v}) + T\mathbf{w}^T\mathbf{v} - T\sum_{s\in\mathcal{S}} \log\left(\sum_{g\in G} e^{w_{sg}}\right) \tag{3.83}$$

The Legendre transformation of the above energy function with respect to the *generalized coordinates* $\mathbf{w}$ is given by [21]

$$H_{eff}(\mathbf{v},\mathbf{w}) = E_{eff}(\mathbf{v},\mathbf{w}) - \mathbf{w}^T\nabla_{\mathbf{w}}E_{eff}(\mathbf{v},\mathbf{w}), \tag{3.84}$$

where the symbol $\nabla_{\mathbf{x}}$ denotes the gradient with respect to $\mathbf{x}$. Observe that at the critical points of E_{eff} the two functions have the same values. Simple algebra shows that

$$\mathbf{w}^T\nabla_{\mathbf{w}}E_{eff}(\mathbf{v},\mathbf{w}) = T\mathbf{w}^T\mathbf{v} - T\sum_{s\in\mathcal{S}}\sum_{g\in G} w_{sg}F_g(\mathbf{w}_s).$$

Substituting this equation along with (3.83) into (3.84), we get

$$H_{eff}(\mathbf{v},\mathbf{w}) = E(\mathbf{v}) - T\sum_{s\in\mathcal{S}}\sum_{g\in G} F_g(\mathbf{w}_s)\log\left(\sum_{g'\in G} e^{w_{sg'}}\right) - T\sum_{s\in\mathcal{S}}\sum_{g\in G} w_{sg}F_g(\mathbf{w}_s) \tag{3.85}$$

From the definition of the generalized sigmoid mapping $\mathbf{F}$, we have

$$\log\left(\sum_{g'\in G} e^{w_{sg'}}\right) = w_{sg} - \log F_g(\mathbf{w}_s).$$

[20] An insightful treatment that we particularly like is the one by Callen [9].

[21] Note that we continue to express H_{eff} in terms of the original coordinates $\mathbf{w}$.

Substituting back in (3.85), we get simply

$$H_{eff}(\mathbf{v}, \mathbf{w}) = E(\mathbf{v}) + T \sum_{s \in S} \sum_{g \in G} F_g(\mathbf{w}_s) \log F_g(\mathbf{w}_s). \tag{3.86}$$

This representation has the following appealing features

1. There is a nice separation between the contributions of the system states to H_{eff}. All the contributions of $\mathbf{v}$ are through the energy term $E(\mathbf{v})$, while all the contributions of $\mathbf{w}$ are through the entropy term.

2. The generalized sigmoid mappings allow us to give a compact representation of the system entropy, S, *away* from the fixed points.

$$S = S(\mathbf{w}) = -\sum_{s \in S} \sum_{g \in G} F_g(\mathbf{w}_s) \log F_g(\mathbf{w}_s).$$

The above representation also justifies the probabilistic interpretation of the generalized sigmoid mappings: $F_g(\mathbf{w}_s)$ is the probability that pixel s has color g. Moreover, at the fixed point, we have by (3.69)

$$v^*_{sg} = F_g(\mathbf{w}^*_s).$$

Substituting in (3.86), we recover the expression of the free energy at the fixed point, namely,

$$H_{eff}(\mathbf{v}^*, \mathbf{w}^*) = E_{eff}(\mathbf{v}^*, \mathbf{w}^*) = E(\mathbf{v}^*) + T \sum_{s \in S} \sum_{g \in G} v^*_{sg} \log v^*_{sg},$$

expression that was already given in (3.41) and (3.70).

4

Correlation-Field Approximations of Gibbs Distributions

When I hear you give your reasons, the thing always appears to me so ridiculously simple that I could easily do it myself, though at each successive instance of your reasoning I am baffled until you explain your process.

Dr. Watson

In this chapter, we will use the mean-field equations developed in Chapter 3 to answer the following important question: "What is the correlation function between the two graylevel values of a pair of pixels in a Gibbs random field image model?" Chapter 2 was devoted to investigating some remarkable examples, the Gauss-Markov random field and the Ising model, for which we can answer this question exactly.

Now we would like to argue that for the other Gibbs models in which the pixel state space is neither continuous, like the Gaussian model, nor binary, like the Ising model, we can obtain *estimates* for the correlation field from the very equations that give us estimates for the mean values of its pixels. Formula (2.31) that we proved in Chapter 2 will be at the core of our argument.

Before going into the different cases, it is probably worthwhile to re-examine the different approximations that we have made in order to arrive at the mean-field equations. [1]

[1]The equations we are referring to are the ones we obtained in Chapter 3, namely, (3.10), (3.17), (3.35), (3.48), and (3.69).

1. **Saddle point approximation:** This approximation was used to derive (3.10), (3.17), (3.48), and (3.69). It amounts to estimating the value of an integral by computing its integrand at a well-chosen point. We have mentioned that the validity of this approximation improves as we increase the lattice size.

2. **Stationary phase approximation:** We used this approximation in Section 3.3.1 to argue that the principal contribution to the integral form of the Z partition function (3.62) comes from points lying in the real vector space $\Re^{n|\mathcal{S}|}$ rather than the complex vector space $\mathbf{C}^{n|\mathcal{S}|}$.

3. **Probability factorization approximation:** We used this method to derive the effective energy from an assumption on the joint probability distribution of the Gibbs random field. One aspect of this approximation, in the case of a general energy function, is that it identifies the mean of a function, e.g., the energy, with the value of this function computed at the pixel means. This latter approximation was also used to relate the effective energy to the Helmholtz free energy.

The approximation method for the correlation field that we propose here adds yet another approximation. For it consists in identifying the derivative of the mean estimate with the derivative of the exact mean so as to be able to apply (2.31) and obtain the correlation function.

For any two pixels p, q on the lattice we denote by Γ_{pq} the correlation[2] between the random variables x_p and x_q, i.e,

$$\Gamma_{pq} = \overline{(x_p - \overline{x}_p)(x_q - \overline{x}_q)}, \tag{4.1}$$

where the expectation operator[3] is denoted by an overline.

According to Lemma 7, the correlation is given by

$$\Gamma_{pq} = \frac{\partial \overline{x_p}}{\partial h_q}(h_q = 0), \tag{4.2}$$

where h_q is the strength of an external field applied to the lattice at pixel q. Let us remind ourselves here that in (4.2), the external field is scaled with temperature before it is applied, i.e., it is the term $-Th_qx_q$ that is added to the energy function rather than the term $-h_qx_q$. When the latter is added to the energy, Formula 4.2 should be replaced with

$$\Gamma_{pq} = T\frac{\partial \overline{x_p}}{\partial h_q}(h_q = 0). \tag{4.3}$$

[2]In this chapter, we abuse terminolgy and call Γ_{pq} a correlation function. See ([66], p. 33).

[3]This symbol was chosen so as not to confuse the common expectation symbol E with the energy function.

The approximation formula for the correlation function is

$$\Gamma_{pq} \approx \hat{\Gamma}_{pq} = T\frac{\partial \hat{x}_p}{\partial h_q}(h_q = 0), \tag{4.4}$$

where $\hat{x}_p$ and $\hat{\Gamma}_{pq}$ denote *estimates* of the mean at pixel p and the correlation between pixels p and q, respectively. It is this approximation that we will use in the rest of this chapter.

The intuitive idea for using the above formula is quite simple. If we want to know the inherent correlation between x_p and x_q, i.e., how much x_p fluctuates around its mean because of fluctuations that x_q experiences around its own mean, we consider the mean behavior of the system. Then we excite it by a small perturbation Δh_q at site q, measure the change $\Delta \hat{x}_p$ in the mean value of x_p due to the perturbation, and compute the ratio $\frac{\Delta \hat{x}_p}{\Delta h_q}$. The change in $\hat{x}_p$ is due to the effects of the external perturbation Δh_q *and* the internal interactions between p and the rest of the lattice due to the fluctuations of x_q. Therefore to get the correlation between x_p and x_q due solely to the internal field interactions, we need to take the limit of the above ratio as $\Delta h_q \to 0$, which gives Formula 4.2.

The rest of this chapter will be devoted to investigating the use of the mean-field approximations of Chapter 3 to compute the correlation-field approximations according to Formula 4.4.

4.1 Unconstrained Binary Case

Our starting point is (3.17) that we reproduce here for ease of reference

$$v_p = \tanh\left[\frac{1}{T}\left(\sum_{q\in\mathcal{S}} J_{pq}v_q + h_p\right)\right], \quad \forall p \in \mathcal{S}. \tag{4.5}$$

The above equation was derived with an external field of energy $-\mathbf{h}^T\mathbf{x}$ added to the interaction field energy $E(\mathbf{x})$. Therefore to compute the correlation field in this case, (4.4) needs to be applied.

When the external field is zero, $\mathbf{h} = \mathbf{0}$, then it is clear that $\mathbf{v} = 0$ is always a solution of (4.5) at all temperatures. This corresponds to a zero mean value for all pixels. For an external field $\mathbf{h}$ applied to the lattice, the perturbation of the zero mean solution can be obtained from (4.5) using a first-order Taylor series expansion of the tanh function in the neighborhood of zero. This yields

$$v_p = \frac{1}{T}\left(\sum_{q\in\mathcal{S}} J_{pq}v_q + h_p\right), \quad \forall p \in \mathcal{S}. \tag{4.6}$$

If we denote by $\mathbf{I}_{\mathcal{S}}$ the identity matrix of order $|\mathcal{S}| \times |\mathcal{S}|$, then the above equation can be written in matrix form as

$$(T\mathbf{I}_{\mathcal{S}} - \mathbf{J})\mathbf{v} = \mathbf{h}. \tag{4.7}$$

Note that if T is high enough, the above matrix is guaranteed to be invertible. A sufficient condition for this to happen is that T be greater than the largest eigenvalue[4] of $\mathbf{J}$. Under this condition, the perturbation of the mean field $\mathbf{v}$ around the zero solution due to an external field $\mathbf{h}$ is given by

$$\mathbf{v} = (T\mathbf{I}_{\mathcal{S}} - \mathbf{J})^{-1}\mathbf{h}. \tag{4.8}$$

If we denote by $\mathbf{\Gamma}$ the $|\mathcal{S}| \times |\mathcal{S}|$ correlation matrix of the binary random field, and by $\hat{\mathbf{\Gamma}}$ its approximation then we get by Formula 4.4

$$\hat{\mathbf{\Gamma}} = (\mathbf{I}_{\mathcal{S}} - \frac{1}{T}\mathbf{J})^{-1}. \tag{4.9}$$

The following remarks are called for:

1. The matrix $\hat{\mathbf{\Gamma}}$ is symmetric due to the symmetry assumption on $\mathbf{J}$. Conversely, if $\hat{\mathbf{\Gamma}}$ is to be symmetric as an approximation of the symmetric correlation matrix Γ, then $\mathbf{J}$ must be symmetric.

2. What if T is small? If T is so small that the Hessian at the zero solution of (4.5) is not positive semidefinite, then the zero mean solution cannot be a local minimum of the effective energy. In other words, we need to Taylor-expand the right-hand side of (4.6) around a different point.

Using the above approximation of the correlation function, we can define a Gaussian distribution having

$$\hat{\mathbf{\Gamma}}^{-1} = \frac{1}{T}(TI_{\mathcal{S}} - \mathbf{J})$$

for information matrix. [5] Since the field has a zero mean, the Gaussian distribution is given by

$$p(\mathbf{v}) = \kappa_1 \exp\left[-\frac{1}{2T}\mathbf{v}^T(-\mathbf{J} + T\mathbf{I}_{\mathcal{S}})\mathbf{v}\right],$$

where κ_1 is a normalizing constant.

This distribution has the same first and second order statistics as the high temperature binary model and could be construed a Gaussian "approximation" of the binary field. The previous statement should be understood informally, since we have not given a precise definition to the concept of distance between probability distributions.

[4]The matrix $\mathbf{J}$ is assumed symmetric, and therefore all its eigenvalues are real.

[5]An information matrix is the inverse of a non-singular correlation matrix.

4.2 Unconstrained Multilevel Case

In this section we conduct an analysis similar to the one we have given for the binary case. The difficulty in the multilevel case is in the way the colors are represented in the mean-field solution. Indeed, the mean-field equations give the analog mean values, $v_{sg} \in (0,1), s \in \mathcal{S}, g \in G$, of the binary matching variables y_{sg} rather than the colors themselves. An interpretation of v_{sg} that is suggested by both the uniqueness constraint (3.26) and the expression of the entropy at the mean-field solution, e.g., (3.40), is that it represents the probability that site s gets the color g. In terms of the binary matching variables y_{sg}, one can always write

$$x_s = \sum_{g \in G} g y_{sg} = \mathbf{g}^T \mathbf{y}_s, \ \forall s \in \mathcal{S}, \tag{4.10}$$

where $\mathbf{g}$ is the vector of the n pixel graylevels $(0, 1, \ldots, n-1)$. It follows that the mean color is given by

$$\overline{x}_s = \sum_{g \in G} g \overline{y}_{sg} = \mathbf{g}^T \overline{\mathbf{y}}_s. \tag{4.11}$$

The other difficulty is how one should apply (4.3) given that the external field $\mathbf{h}$ should be applied on the two-dimensional lattice pixels $x_s, s \in \mathcal{S}$, and not the three-dimensional lattice of the binary matching variables $y_{sg}, s \in \mathcal{S}, g \in G$. This problem can be solved by decomposing the pixel color over the binary matching variables as in (4.10). Then the energy due to the external field h_s at s can be written as

$$-h_s x_s = -h_s \sum_{g \in G} g y_{sg} = -h_s \mathbf{g}^T \mathbf{y}_s. \tag{4.12}$$

The above equation means that applying an external field h_s on x_s is equivalent to applying an external field $h_s \mathbf{g}$ on $\mathbf{y}_s$. The external excitation at site s corresponding to color g is denoted $h_{sg} = g h_s$. Using (4.12), the fixed-point equations (3.48) become

$$\mathbf{v}_s = \mathbf{F}\left[\frac{1}{T}\left(\sum_{r \in \mathcal{S}} \mathbf{J}_{sr} \mathbf{v}_r + h_s \mathbf{g}\right)\right], \forall s \in \mathcal{S}, \tag{4.13}$$

where $\mathbf{J}_{sr}$ is the $n \times n$ matrix $[J_{sr}^{ab}]_{(a,b) \in G \times G}$.

4.2.1 A General Relationship

Using the above remarks, we can write down a general relationship between the correlation-field expressed in terms of the actual color means and the correlations based

on the means of the binary matching elements. The interest of this relationship is that it reduces the computation of

$$\Gamma_{pq} = T\frac{\partial \overline{x}_p}{\partial h_q}$$

to that of

$$\Gamma_{pq}^{ab} = T\frac{\partial \overline{y}_{pa}}{\partial h_{qb}}.$$

Proposition 9 *For any $p, q \in \mathcal{S}$, we have*

$$\Gamma_{pq} = \sum_{a,b \in G} ab\Gamma_{pq}^{ab}. \tag{4.14}$$

Proof: First notice that because of (4.11), we have

$$\frac{\partial \overline{x}_p}{\partial h_q} = \sum_{a \in G} a\frac{\partial \overline{y}_{pa}}{\partial h_q}.$$

Next by the chain rule, we have

$$\frac{\partial \overline{y}_{pa}}{\partial h_q} = \sum_{b \in G} \frac{\partial \overline{y}_{pa}}{\partial h_{qb}}\frac{\partial h_{qb}}{\partial h_q},$$

where according to (4.12),

$$\frac{\partial h_{qb}}{\partial h_q} = b.$$

Equation (4.14) follows immediately. ∎

The meaning of (4.14) should be clear. To pass from the color correlations to the matching variable correlations the latter ones need to be weighted by the color products ab and summed to get the color-based correlation coefficients. Proposition 9 can also be proved directly by going back to the definition of the means of x_s and y_{sg} as function of the Gibbs distribution. It is important to keep in mind that the correlation coefficients are not independent. Indeed, because of the uniqueness constraint on the binary matching elements, we have for their means

$$\sum_{a \in G} \overline{y_{pa}} = 1, \ \forall p \in \mathcal{S},$$

which upon partial differentiation with respect to h_{qb} gives

$$\sum_{a \in G} \Gamma_{pq}^{ab} = 0, \ \forall p, q \in \mathcal{S}, \ \forall b \in G.$$

Finally, let us note that Equation 4.14 can be written more compactly as

$$\Gamma_{pq} = \mathbf{g}^T \mathbf{\Gamma}_{pq} \mathbf{g},$$

where $\mathbf{g}$ is the color vector and $\mathbf{\Gamma}_{pq}$ is the $n \times n$ correlation matrix between the matching variables at sites p and q.

As a result of the above proposition, we can now concentrate on approximating the correlation coefficients Γ_{pq}^{ab}. But before doing this we need some preliminary properties satisfied by the general fixed point equations(4.13).

4.2.2 Preliminary Properties

In this paragraph, we assemble some of the properties of (4.13) because of their usefulness in the derivation of the correlation-field approximation.

At high temperatures, the generalized sigmoid mapping $\mathbf{F}$ maps $\Re^n$ onto the neighborhood of the point $\frac{1}{n}\mathbf{e}$, where $\mathbf{e} = (1, 1, \ldots, 1)^T$. Therefore in the limit of high temperatures, $\mathbf{v}_s = \frac{1}{n}\mathbf{e}, \forall s \in \mathcal{S}$, is always a solution. This remains the case *even* in the presence of an external field. In the binary case of the previous section, the zero mean-field was a valid solution at zero external field for all temperatures. A similar situation happens in the multilevel case if we assume that the sum

$$\sum_{r \in \mathcal{S}, b \in G} J_{sr}^{ab}$$

is independent of both s and a. Indeed for any $\alpha \in \Re, \mathbf{F}(\alpha\mathbf{e}) = \frac{1}{n}\mathbf{e}$, which means that $\mathbf{v}_s = \frac{1}{n}\mathbf{e}$ is always a solution at zero external field. We call this solution the *trivial* mean-field solution. Its probabilistic interpretation is that every pixel has an equal probability for being assigned any color.

Another fact about $\mathbf{F}$ is that its gradient can be given in closed form as function of $\mathbf{F}$ itself. We will show in Chapter 5 that the Jacobian $D\mathbf{F}(\mathbf{z})$ of $\mathbf{F}(\mathbf{z})$ is given by

$$D\mathbf{F}(\mathbf{z}) = diag(F_a(\mathbf{z})) - \mathbf{F}(\mathbf{z})\mathbf{F}(\mathbf{z})^T. \tag{4.15}$$

At the trivial mean-field solution, we have

$$\mathbf{M} \triangleq D\mathbf{F}(\frac{1}{n}\mathbf{e}) = \frac{1}{n}\mathbf{I} - \frac{1}{n^2}\mathbf{O} \tag{4.16}$$

where $\mathbf{I}$ is the $n \times n$ identity matrix, and $\mathbf{O} \triangleq \mathbf{e}\mathbf{e}^T$ is the $n \times n$ matrix having all its coefficients equal to 1. Note that since $\mathbf{M}\mathbf{e} = 0$, the matrix $\mathbf{M}$ is singular. It is also symmetric, therefore digonalizable with an orthogonal matrix. Let $\mathbf{e}' \in \mathbf{e}^\perp$, the $(n-1)$-dimensional linear space orthogonal to the vector $\mathbf{e}$. Then we have $\mathbf{M}\mathbf{e}' = \frac{1}{n}\mathbf{e}'$, which means that any vector in $\mathbf{e}^\perp$ is an eigenvector of $\mathbf{M}$ corresponding to the eigenvalue $\frac{1}{n}$. It follows that $\mathbf{M}$ has two eigenvalues: $\lambda_1 = \frac{1}{n}$ of multiplicity $(n-1)$, and

$\lambda_2 = 0$ of multiplicity 1. To construct the matrix $\mathbf{K}$ that diagonalizes $\mathbf{M}$, we choose the last column vector $\mathbf{k}_n = \mathbf{e}$, and the first $n-1$ column vectors $\mathbf{k}_1, \ldots, \mathbf{k}_{n-1}$ equal to an arbitrary orthonormal basis of the $(n-1)$-dimensional space $\mathbf{e}^\perp$. We remind the reader that for an orthogonal matrix, we have $\mathbf{K}^{-1} = \mathbf{K}^T$.

4.2.3 General Case: Orthonormal Transformation

We are now interested in computing approximations of the correlation coefficients Γ^{ab}_{pq}. As for the binary case, we Taylor-expand the right-hand side of (4.13) to the first order in the neighborhood of the trivial mean-field solution. Using the expression of the generalized sigmoid function gradient, we get

$$\frac{1}{n}\mathbf{e} + \mathbf{v}_s = \frac{1}{n}\mathbf{e} + \mathbf{M}\left[\frac{1}{T}\left(\sum_{r \in \mathcal{S}} \mathbf{J}_{sr}\mathbf{v}_r + h_s\mathbf{g}\right)\right]. \tag{4.17}$$

In the above equation, $\mathbf{v}_s$ is a perturbation of the analog matching variables around the trivial mean-field solution $\frac{1}{n}\mathbf{e}$, which cancels out from both sides. Since $\mathbf{e}^T\mathbf{M} = 0$, a necessary condition for $\mathbf{v}_s$ to be a valid perturbation is that

$$\mathbf{e}^T\mathbf{v}_s = \sum_{g \in G} v_{sg} = 0.$$

In other words, the perturbation $\mathbf{v}_s$ at every site s must be in $\mathbf{e}^\perp$, the vector space orthogonal to $\mathbf{e}$. This also means that the above linear equations in $\mathbf{v}_s$ are not independent. However they can be made independent using the linear transformation

$$\mathbf{u}_s = \mathbf{K}^T\mathbf{v}_s.$$

Observe that because the $\mathbf{e}^T$ is the last row in $\mathbf{K}^T$, the last component in $\mathbf{u}_s$ is 0. Multiplying (4.17) through by $\mathbf{K}^T$ we get

$$\mathbf{u}_s = \mathbf{K}^T\mathbf{M}\mathbf{K}\left[\frac{1}{T}\left(\sum_{r \in \mathcal{S}} \mathbf{K}^T\mathbf{J}_{sr}\mathbf{K}\mathbf{u}_r + h_s\mathbf{K}^T\mathbf{g}\right)\right], \ \forall s \in \mathcal{S}. \tag{4.18}$$

The matrix $\mathbf{K}^T\mathbf{M}\mathbf{K} = diag(\frac{1}{n}, \ldots, \frac{1}{n}, 0)$. Therefore for every vector $\mathbf{u}_s$ the last equation is identically satisfied and gives $0 = 0$. So we can eliminate $|\mathcal{S}|$ equations from the above linear system and solve for $\mathbf{u}'_s = (u_1, \ldots, u_{n-1})^T$ instead of solving for $\mathbf{u}_s$. Denoting by $\mathbf{J}'_{sr}$ the matrix obtained from $\mathbf{K}^T\mathbf{J}_{sr}\mathbf{K}$ by deleting the last row and column and by $\mathbf{g}'$ the vector obtained from $\mathbf{K}^T\mathbf{g}$ by deleting the last component, we get for the $\mathbf{u}'_s$'s the linear system

$$\mathbf{u}'_s = \frac{1}{n}\left[\frac{1}{T}\left(\sum_{r \in \mathcal{S}} \mathbf{J}'_{sr}\mathbf{u}'_r + h_s\mathbf{g}'\right)\right], \ \forall s \in \mathcal{S}. \tag{4.19}$$

Denoting by $\mathbf{u}'$ and $\mathbf{h}'$ the vectors obtained by stacking the the site vectors $\mathbf{u}'_s$ and $h_s\mathbf{g}'$, respectively, we get

$$(nT\mathbf{I} - \mathbf{J}')\mathbf{u}' = \mathbf{h}', \tag{4.20}$$

where $\mathbf{J}'$ is the block matrix $[\mathbf{J}'_{sr}]_{s,r\in\mathcal{S}}$. Formally, this linear equation for $\mathbf{u}'$ looks like the linear equation (4.7) given for the binary model. Some of the issues related to the use of (4.20) to compute the the correlation coefficients Γ^{ab}_{pq} are

1. We need to be able to invert the matrix $(nT\mathbf{I} - \mathbf{J}')$. This means that either n, the number of colors, or T, the temperature, needs to be high enough. A necessary and sufficient condition for $(nT\mathbf{I} - \mathbf{J}')$ to be invertible is that the product nT be greater than the largest eigenvalues of $\mathbf{J}'$.

2. The relationship between the correlations based on the variables $\mathbf{u}$ and those based on the variables $\mathbf{v}$ is given by the well known formula for the linear transformations of random vectors

$$\mathbf{\Gamma}_{pq}(\mathbf{v}) = \mathbf{K}\mathbf{\Gamma}_{pq}(\mathbf{u})\mathbf{K}^T \tag{4.21}$$

   where the boldface $\mathbf{\Gamma}_{pq}(\mathbf{v})$ matrix denotes the correlation matrix between the matching vectors $\mathbf{v}_p$ and $\mathbf{v}_q$, and similarly for $\mathbf{\Gamma}_{pq}(\mathbf{u})$.

3. How is the correlation matrix $\mathbf{\Gamma}_{pq}(\mathbf{u})$ is derived from the correlation matrix $\mathbf{\Gamma}_{pq}(\mathbf{u}')$? Here also we should go to first principles. We know that the last component of $\mathbf{u}_s$ is 0 for all $s \in \mathcal{S}$. Therefore, $\mathbf{\Gamma}_{pq}(\mathbf{u})$ is obtained from $\mathbf{\Gamma}_{pq}(\mathbf{u}')$ by adding one row and one column of zeros to the latter.

4. Is the final result dependent on the choice of the orthogonal transformation matrix $\mathbf{K}$, i.e., on the orthogonal basis of the linear subspace $\mathbf{e}^T$? The answer is no. One can prove, although the algebra is tedious, that the effects of the orthogonal transformation matrix $\mathbf{K}$ in (4.18) and (4.21) "cancel out."

Now we are ready to give a general procedure for obtaining an approximation $\hat{\mathbf{\Gamma}}$ of the correlation matrix $\mathbf{\Gamma}$ for the color random variables $\mathbf{x}$ distributed according to a general Gibbs distribution with pairwise interaction potentials.

1. Choose the matrix $\mathbf{K}$ and compute

$$\mathbf{\Gamma}(\mathbf{u}') = (nT\mathbf{I} - \mathbf{J}')^{-1}.$$

2. Construct the matrix $\mathbf{\Gamma}(\mathbf{u})$ by adding to each block, $\mathbf{\Gamma}_{pq}(\mathbf{u})$, a row and a column of zeros.

3. Compute $\mathbf{\Gamma}(\mathbf{v})$ according to (4.21).

4. Finally compute the color correlation coefficients Γ_{pq} using (4.14).

In this procedure, the computational burden comes from the computation of the inverse matrix. We will see in Chapter 6 that for some special structures of the model matrix $\mathbf{J}$ we can use FFT algorithms to invert the matrix $(nT\mathbf{I} - \mathbf{J}')$.

4.2.4 Special Case: The Autobinomial Model

An interesting case where we do not have to go through the rather cumbersome method developed for the general multilevel model is that of the autobinomial model, first introduced by Besag [6] and subsequently used in [14] for texture synthesis. This model is defined by the coefficients

$$J_{sr}^{ab} = -2L_{sr}V_{ab} = 2L_{sr}ab, \ \forall s, r \in \mathcal{S}, a, b \in G, \tag{4.22}$$

where L_{sr} is a neighborhood function that is nonzero if and only if r is a neighbor of s. Note that any anisotropy in the model can be incorporated in the definition of L_{sr}. For instance, when the neighborhood $\mathcal{N}_s$ is that of the four nearest neighbors on the square, toroidal grid, we can write

$$L_{sr} = \begin{cases} 0 & \text{if } r \notin \mathcal{N}_s \\ \beta_h & \text{if } r \text{ is a west or east neighbor of } s \\ \beta_v & \text{if } r \text{ is a north or south neighbor of } s \end{cases} \tag{4.23}$$

Using the graylevel vector $\mathbf{g}$, the color interaction matrix for this model can be written as

$$\mathbf{V} = -\mathbf{g}\mathbf{g}^T.$$

It follows that

$$\mathbf{J}_{sr} = 2L_{sr}\mathbf{g}\mathbf{g}^T,$$

and the model matrix $\mathbf{J}$ is given by

$$\mathbf{J} = 2\mathbf{L} \otimes \mathbf{g}\mathbf{g}^T. \tag{4.24}$$

In order to get an estimate on the perturbation around the mean color, we multiply (4.17) through by $\mathbf{g}^T$ to obtain

$$\hat{\mathbf{x}}_s = \mathbf{g}^T\mathbf{M}\left[\frac{1}{T}\left(\sum_{r\in\mathcal{S}} 2L_{sr}\mathbf{g}\mathbf{g}^T\mathbf{v}_r + h_s\mathbf{g}\right)\right]. \tag{4.25}$$

Easy algebra shows that

$$\begin{aligned} \mathbf{g}^T\mathbf{g} &= \frac{1}{6}(n-1)n(2n-1) \\ \mathbf{g}^T\mathbf{M}\mathbf{g} &= \frac{1}{n}\mathbf{g}^T\mathbf{g} - \frac{1}{n^2}\mathbf{g}^T\mathbf{O}\mathbf{g} = \frac{1}{12}(n^2-1) \end{aligned}$$

It follows that

$$\hat{\mathbf{x}}_s = \frac{n^2-1}{12T}\left(\sum_{r\in\mathcal{S}} 2L_{sr}\hat{\mathbf{x}}_r + h_s\right). \tag{4.26}$$

The perturbations $\hat{\mathbf{x}}$ on the mean colors in the autobinomial case satisfy the following system of linear equations

$$(\frac{12T}{n^2-1}\mathbf{I}_{\mathcal{S}} - 2\mathbf{L})\hat{\mathbf{x}} = \mathbf{h} \tag{4.27}$$

For a ratio $12T/(n^2-1)$ higher than the largest eigenvalue of $\mathbf{L}$, we can approximate the correlation function of the autobinomial model around the trivial mean-field solution by

$$\hat{\mathbf{\Gamma}} = (\frac{12}{n^2-1}\mathbf{I}_{\mathcal{S}} - \frac{2}{T}\mathbf{L})^{-1}, \tag{4.28}$$

where for this model, the trivial mean-field solution corresponds to a mean color value given by

$$\overline{x}_s = \frac{1}{n}\sum_{g\in G} g = \frac{n-1}{2}.$$

The natural question that arises here is "how is this multilevel case related to the binary case treated in the previous section?" Substituting n with 2 in (4.27) gives

$$(4T\mathbf{I}_{\mathcal{S}} - 2\mathbf{L})\hat{\mathbf{x}} = \mathbf{h}. \tag{4.29}$$

This equation differs from (4.8) by the presence of the factor 4 in front of the temperature. The reason is due to the difference between the pixel values in the two binary models. In (4.8), the pixel mean values are in the segment (-1,1), while in the present case they are in (0,1). We can pass from one to the other using the change of variable $x \rightarrow \frac{x+1}{2}$. This and the quadratic interaction result in the 4 factor of (4.29).

Also, as in the binary case, we can give a Gaussian "approximation" of the autobinomial model based on the mean $\overline{x}_s = (n-1)/2$ and the information matrix $\frac{1}{T}(\frac{12T}{n^2-1}\mathbf{I}_{\mathcal{S}} - 2\mathbf{L})$ by

$$p(\mathbf{x}) = \kappa_2 \exp\left[-\frac{1}{T}(\mathbf{x}-\overline{\mathbf{x}})^T\left(\frac{12T}{n^2-1}\mathbf{I}_{\mathcal{S}} - 2\mathbf{L}\right)(\mathbf{x}-\overline{\mathbf{x}})\right],$$

where κ_2 is a normalizing constant. For specific structures of the model matrix $\mathbf{L}$, much more can be said about the values of T for which the above approximation exists, see Chapter 6. An important observation can already be made: the original autobinomial model and the corresponding Gaussian model share the same neighborhood structure and bonding parameters.

4.3 Constrained Multilevel Case

So far, the only constraint that has been imposed on the binary matching variables is the *local* uniqueness constraint which specifies that one and only one binary matching variable is nonzero at each site. The natural question is whether the general procedure that we have developed in the previous section carries over to the case where *global* constraints are imposed on the binary matching variables. One type of such constraints is the uniform histogram constraint that was introduced in Section 3.3.2. From a pattern synthesis point of view, this constraint guarantees that all colors are present in equal amounts in the synthesized pattern. In graph optimization, it corresponds to an equipartitioning of the graph vertices.

It was noticed in Chapter 2 that Formula (2.31) is valid whether the configurations are constrained or not. Also, the general remarks given in the previous section about the relationships between the color correlations and the matching variable correlations remain valid for the constrained case.

Our objective in this section is therefore to show that the correlation estimates

$$\hat{\Gamma}^{bd}_{rm} = T\frac{\partial v_{rb}}{\partial h_{md}}(\mathbf{h} = \mathbf{0}) \tag{4.30}$$

can also be obtained using the information contained in the critical point equations of the effective energy. In the binary and the multilevel *unconstrained* cases, we used first-order Taylor series expansions of the fixed point equations to compute the correlation approximations as the coefficients of the inverse of a certain matrix. For the multilevel *constrained* case, we will show how the mean-field equations (3.80) derived in Section 3.3.2 can be used to obtain a system of linear equations that is satisfied by the correlation approximations. An important advantage of this method is to show the close relationship between these approximations and the Hessian of the system energy. This relationship is an instance of a deeper fact relating Green's functions to correlation functions in continuous field problems [66]. We gave the reader a glimpse at this fact in the description that followed (4.4).

Our starting point is (3.80) to which we add the effect due to an external field represented by the field strength $\mathbf{h} = (h_{md}, m \in \mathcal{S}, d \in G)$. The equation then becomes

$$\frac{\partial E}{\partial v_{sa}} + T(1 + \log v_{sa}) - h_{md}\delta_{ad}\delta_{sm} + q_s + p_a = 0,\ s \in \mathcal{S},\ a \in G, \tag{4.31}$$

where q_s and p_a are the Lagrange multipliers corresponding to the uniqueness and uniform histogram constraint. The Kronecker delta symbols express the fact that the effect of

the external field is linear in the binary matching variables. The solutions $\mathbf{v}, \mathbf{q}$ and $\mathbf{p}$ of this equation are functions of the external field $\mathbf{h} = (h_{md}, m \in \mathcal{S}, d \in G)$.

In order to obtain the equations satisfied by the correlation estimates, the above equation is differentiated with respect to h_{md} and the chain rule is applied to the energy partial derivatives

$$\sum_{rb} \frac{\partial^2 E}{\partial v_{sa} \partial v_{rb}} \frac{\partial v_{rb}}{\partial h_{md}} + \frac{T}{v_{sa}} \frac{\partial v_{sa}}{\partial h_{md}} - \delta_{ad}\delta_{sm} + \frac{\partial q_s}{\partial h_{md}} + \frac{\partial p_a}{\partial h_{md}} = 0, \; s \in \mathcal{S}, \; a \in G. \quad (4.32)$$

Letting $\mathbf{h}$ go to zero, denoting by $\mathbf{v}^* = (v^*_{sa}, s \in \mathcal{S}, a \in G)$ a solution of the mean-field equations, and substituting the partial derivatives from (4.30), we get the following system of linear equations for the correlation estimates

$$\sum_{rb} \frac{\partial^2 E(\mathbf{v}^*)}{\partial v_{sa} \partial v_{rb}} \hat{\Gamma}^{bd}_{rm} + \frac{T}{v^*_{sa}} \hat{\Gamma}^{ad}_{sm} - T\delta_{ad}\delta_{sm} + T\frac{\partial q_s}{\partial h_{md}} + T\frac{\partial p_a}{\partial h_{md}} = 0, \; s \in \mathcal{S}, \; a \in G. \quad (4.33)$$

The above equations can be written in terms of the Hessian coefficients

$$H^{ab}_{sr} = \frac{\partial^2 E(\mathbf{v}^*)}{\partial v_{sa} \partial v_{rb}} + \frac{T}{v^*_{sa}} \delta_{sr}\delta_{ab}$$

of the *unconstrained* effective energy computed at the mean-field solution $\mathbf{v}^*$ as

$$\sum_{rb} H^{ab}_{sr} \hat{\Gamma}^{bd}_{rm} - T\delta_{ad}\delta_{sm} + T\frac{\partial p_a}{\partial h_{md}} + T\frac{\partial q_s}{\partial h_{md}} = 0, \; s \in \mathcal{S}, \; a \in G. \quad (4.34)$$

The partial derivatives of p_a and q_s express the sensitivity of the Lagrange multipliers to the perturbation of the system energy with an external field. These sensitivities are unknown but can be obtained using the constraint equations (3.81) and (3.82). Indeed, if we differentiate the constraint equations with respect to $h_{md}, m \in \mathcal{S}, d \in G$, and compute the resulting partial derivatives at $\mathbf{h} = \mathbf{0}$, we get for the uniqueness and uniform histogram constraints

$$\sum_{b} \hat{\Gamma}^{bd}_{rm} = 0, \; \forall r, m \in \mathcal{S}, \; \forall d \in G, \quad (4.35)$$

and

$$\sum_{r} \hat{\Gamma}^{bd}_{rm} = 0, \; \forall b, d \in G, \forall m \in \mathcal{S}, \quad (4.36)$$

respectively. The number of unknowns $\hat{\Gamma}^{bd}_{rm}, \frac{\partial p_a}{\partial h_{md}}, \frac{\partial q_s}{\partial h_{md}}$ in (4.33), (4.35) and (4.36) is equal to the number of equations. Moreover the system is linear in the unknowns and, in theory, can be solved to give the correlation functions as well as the sensitivity of the Lagrange multipliers with respect to the external field.

The above procedure for computing the correlation estimates has a number of advantages. Among them are

1. It is valid for any energy function.

2. It is valid at any temperature T above the critical temperature (see Chapter 6).

3. The configuration constraints can be incorporated via Lagrange multipliers, and the sensitivities of the latter with respect to external fields can be computed along with the correlation estimates.

It is worthwhile to look at (4.34) when no constraints are imposed on the binary matching variables. Then the Lagrange multipliers are identically zero, and these equations become

$$\sum_{rb} H_{sr}^{ab} \hat{\Gamma}_{rm}^{bd} = T\delta_{ad}\delta_{sm}, \ \forall m, s \in \mathcal{S}, \ \forall a, d \in G, \tag{4.37}$$

which can be written in tensor form as

$$\mathbf{H}\hat{\mathbf{\Gamma}} = T\mathbf{I}_{\mathcal{S}} \otimes \mathbf{I}_G, \tag{4.38}$$

where $\mathbf{I}_{\mathcal{S}}$ and $\mathbf{I}_G$ are the $|\mathcal{S}| \times |\mathcal{S}|$ and $n \times n$ identity matrices, respectively. Under this form it is clear that the correlation tensor and the effective energy Hessian *computed at the global minimum* of the effective energy are inverse of each other. Note that as the temperature decreases to zero the global minima of the effective energy and those of the energy become identical. When the uniqueness and uniform histogram constraints are imposed, we can obtain the correlation estimates of the constrained case from those of the unconstrained case by projecting the latter on the constraint linear spaces defined by (4.35) and (4.36).

Having computed the correlation estimates of the binary matching variables, we can get the correlation estimates of the graylevels using (4.14)

4.4 Summary

It is probably time to stop and reflect about what we have accomplished. In Chapter 2, we raised the question of computing the correlation function of a Gibbs distribution – a probabilistic model of widespread use in image modeling and machine vision. We carefully considered some of the cases for which the correlation function is exactly known: the continuous Gaussian model and the binary Ising model. The question remained whether the correlation function can be computed approximately for other models. Using (2.31), also known as the linear response theorem in Statistical Physics [66], we have proposed a procedure for computing correlation-field approximations for a large class of discrete, multilevel Gibbs distributions. This procedure is new. So is the explicit formula giving

the correlation approximation for the autobinomial model. Moreover, we have considered the constrained graylevel case, and here also, we have shown that approximations of the correlation coefficients can be obtained as functions of the Hessian of the effective energy computed at the mean-field solution.

The mean-field and correlation field approximations can be used to define Gaussian distributions that have the same neighborhood and bonding parameter characteristics as the original discrete, graylevel distributions. We believe that the relationships between these Gaussian models and their discrete, graylevel parents need further study. Of particular importance is to understand when we can use the Gaussian model, which can be sampled relatively easily, instead of the discrete, graylevel model whose samples require time-consuming Monte Carlo methods to generate.

5

Mean-Field Theories and their Network Representations

If you have built castles in the air, your work need not be lost; that is where they should be. Now put the foundations under them.

Henry David Thoreau

5.1 Introduction

Even the casual reader should have noticed that up until now, our work has drawn heavily on the analytical methods of statistical mechanics. In this chapter, we shift our viewpoint from thermodynamics to dynamics. Computationally, the shift is from the study of cost functions and their extrema to that of algorithms and stability.

This chapter is written somewhat in the spirit of Hopfield's two seminal papers [35, 36] on the binary and analog neural networks that now bear his name. Our starting point is a dynamical system defined on a finite set. This set was the vertices of the N-dimensional hypercube in the case of the discrete Hopfield model, where N is the number of neurons. In our model, the discrete winner-take-all (WTA) model, the finite set is the cartesian product of N sets, each one of them being the set of vertices of the simplex $\mathcal{T}_n$. This simplex is a *natural* one, since it is the one associated with the uniqueness constraint

of the binary matching elements. We show that our dynamical system possesses an "energy" function that decreases along its trajectories. We then concentrate on the study of one mode of operation of the WTA system, the synchronous one, and give a sufficient condition that guarantees the absence of oscillations in the network. Our proof uses an apparently novel argument of interlacing Lyapunov sequences.

The gap between Hopfield's discrete [35] and analog [36] models can be bridged using mean-field theory. We alluded to this fact in Chapter 3. See also [58] and [92]. Interestingly, Hopfield himself was not aware of this connection until after the publication of his 1984 paper on the analog model [John Wyatt, private communication]. We will show, using a simple mean-field method that does not require the use of path integrals à la Peterson [68] or à la Simic [80], how an analog WTA dynamical system can be derived from the discrete one. There is an interesting parallelism in the mean-field derivation between the Hopfield model and the WTA model. The now famous sigmoid function is a direct result of the mean-field approximation. It operates as a smooth version of the threshold function of the discrete model. The mean-field approximation of the discrete WTA network yields a *generalized* sigmoid mapping that is a smoothed version of the discrete WTA mapping. We state and prove useful properties about the generalized sigmoid mapping, including its being the gradient map of a *convex* function. We then concentrate on the global dynamics of the synchronous, analog WTA network. Here also, we show the existence of a Lyapunov function that is nonincreasing along the system trajectories. Even better, we use this function and the properties of the generalized sigmoid mapping to prove that the only ω-limit sets of the analog WTA dynamical system are either fixed points or limit cycles of period 2. We strengthen this result and show that under a condition similar to that of the discrete WTA network, the *only* ω-limit sets are the fixed points.

Ultimately, we would like to be able to implement the analog WTA network in analog hardware. It came as a surprise to us that the generalized sigmoid mapping can be implemented *exactly* using MOSFET transistors operated in subthreshold CMOS technology. We give two such implementations and comment on some of their circuit-theoretic properties.

Finally, we extract what we believe is the essence of the Hopfield-type dynamics to extend the Hopfield and WTA models to neural networks in which the neuron states live in abstract spaces. We treat in some detail the Hilbert space case and show that results similar to those of the analog WTA model hold as well. We argue that this new layer of abstraction is exactly what is needed to extend the "neural network" paradigm to networks in which the neurons are the sites of phenomena described by, for instance,

elliptic partial differential equations.

We leave out many classical subjects. For instance, we do not discuss the important subjects of learning, memory and recall, and information capacity. With the formal framework developed here, it will be possible to investigate these subjects by "copying" the work that has already been done in the context of Hopfield networks.

5.2 Discrete Winner-Take-All Networks

Our motivation for adopting the formal model described in the next subsection for the discrete WTA network is threefold:

- First, the work described in Chapter 3 showed that whenever the mean-field approximation of a multilevel Gibbs random field is being derived, the mapping of the pixel state onto a vector of binary matching variables was a required step. This mapping results in the change of the pixel state space from the space of labels or graylevels to the set of the simplex $\mathcal{T}_n$ vertices (see (5.1)). This motivated our interest in defining a dynamical system directly on this set.

- A number of WTA mechanisms have been proposed in the computer vision literature to implement different visual tasks. For instance, Marroquin [58] used an algorithm similar to the synchronous dynamical system (5.3) to implement the Marr and Poggio stereo matching algorithm [57]. He made a number of interesting observations concerning network convergence and limit-cycling. Much more definite results will be stated in this chapter about both the discrete and analog WTA networks.

- The winner-take-all mechanism has also made its way into the analog VLSI implementation of early vision algorithms. Mahowald and Delbrück [54] designed an analog VLSI circuit to implement the cooperative stereo algorithm of Marr and Poggio (op. cit.). Their VLSI system implements the local vertical inhibition mechanism among different features using an analog WTA circuit due to Lazzaro [47]. In these and similar circuits, stability is always an issue [82]. But as a first step to sort out the stability question, it is essential to delineate the stability problems due to hardware implementation from those due to the algorithm itself. Hence, the main objective of our WTA framework is to investigate the stability and oscillation issues inherent to the WTA mechanism.

5.2.1 Notation and Framework

Let $\mathcal{S}$ be a set of sites. They could represent either pixels in an image, neurons in a neural network, or spin sites in a magnet model. The number of these sites will be denoted $|\mathcal{S}|$. To each site $s \in \mathcal{S}$, we assign a vertex $\mathbf{y}_s$ of the simplex

$$\mathcal{T}_n = \{\mathbf{z} \in \Re^n, z_a \geq 0, \ \forall a \in \mathcal{G}, \text{ and } \sum_{k \in G} z_k = 1\}, \tag{5.1}$$

where $\mathcal{G}$ is the finite set of integers $\{1, 2, \ldots, n\}$. Every vertex in the simplex encodes one of the state sites. This state could be either a graylevel or a label. We denote the set of vertices of $\mathcal{T}_n$ by T_n. The configuration set or the state space of all the sites is the set

$$\mathbf{\Omega} \triangleq (T_n)^{|\mathcal{S}|}. \tag{5.2}$$

An element $\mathbf{y} \in \mathbf{\Omega}$ is given by $\mathbf{y} = (\mathbf{y}_1^T, \ldots, \mathbf{y}_s^T, \ldots, \mathbf{y}_{|\mathcal{S}|}^T)^T$. Now on the state space $\mathbf{\Omega}$, we define the discrete dynamical system

$$\mathbf{y}_s(k+1) = \mathbf{W}(\mathbf{u}_s(k)) \tag{5.3}$$

$$\mathbf{u}_s(k) = \sum_{r \in \mathcal{S}} \mathbf{J}_{sr}\mathbf{y}_r(k), \ \forall s \in \mathcal{S}, \ \forall k \geq 0 \tag{5.4}$$

The matrices $\mathbf{J}_{sr}$ are $n \times n$ symmetric matrices, and the (a, b), $1 \leq a, b \leq n$, coefficient of each one of them will be denoted by J_{sr}^{ab}. They define connectivity strength between the states of any pair of sites. If sites s and r do not interact then $\mathbf{J}_{sr} = \mathbf{0}$. Throughout the chapter we make the assumption that $\mathbf{J}_{sr}$ is symmetric and that $\mathbf{J}_{sr} = \mathbf{J}_{rs}$. Thus the block matrix $\mathbf{J}$ is symmetric. The mapping $\mathbf{W}$ maps $\Re^n$ to the simplex $\mathcal{T}_n$ and is defined by

$$\mathbf{W}(\mathbf{z}) = (W_1(\mathbf{z}), \ldots, W_n(\mathbf{z}))^T$$

where for every $a \in G$,

$$W_a(\mathbf{z}) = \begin{cases} 1 & \text{if } z_a > z_b, \ \forall b \in \mathcal{G} \\ 0 & \text{otherwise} \end{cases} \tag{5.5}$$

In words, the mapping $\mathbf{W}$ assigns 1 to the component of $\mathbf{z}$ which is greater than all other components and 0 to all the other components. Therefore $\mathbf{W}$ defines a winner-take-all (WTA) mapping on $\Re^n$. The dynamic system (5.3) will be called the discrete winner-take-all (WTA) dynamical system. The above definition does not say what happens if there are more than one winner among the components of the vector $\mathbf{z}$. Let $A = \{a \in \mathcal{G} | z_a = \max(z_1, \ldots, z_n)\}$. The set A is the set of indices corresponding to the maximum

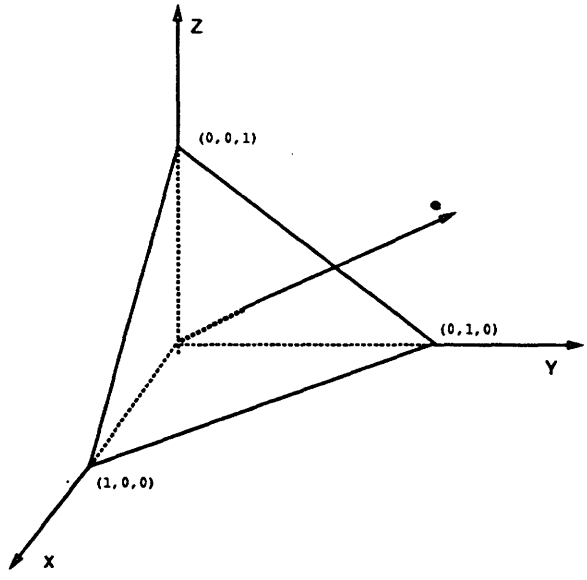


FIGURE 5-1: The 3D simplex. The vertices of this simplex represent the possible states at each site. The configuration space is the cartesian product of $\mathcal{S}$ copies of these vertices.

components of $\mathbf{z}$. The number of these indices will be denoted $|A|$. Then the definition of W_a can be extended to include the case where $|A| \geq 2$ as follows

$$W'_a(\mathbf{z}) = \begin{cases} \frac{1}{|A|} & \text{if } z_a \geq z_b,\ \forall b \in \mathcal{G} \\ 0 & \text{otherwise} \end{cases} \tag{5.6}$$

Observe that with this extended definition of $\mathbf{W}$ the image of $\Re^n$ might contain points that are not among the vertices of $\mathcal{T}_n$. If we want to restrict (5.3) to being a dynamic system on $\mathbf{\Omega}$, then we can choose

$$W_a(\mathbf{z}) = \begin{cases} 1 & \text{if } a = \min A = \min\{b \in \mathcal{G} | z_b = \max(z_1, \ldots, z_n)\} \\ 0 & \text{otherwise} \end{cases} \tag{5.7}$$

In the sequel, we adopt this latter definition of $\mathbf{W}$. In this case, the dynamic system is defined on $\mathbf{\Omega}$. In Figure 5-1, we show the state space of each pixel in the case when n the number of labels is 3. The discrete case evolves on the vertices of the simplex, while the analog state will evolve in its interior. If we imagine that at each site we have a column of binary variables representing the coordinates of each vertex, then the WTA network has a 3D architecture similar to the one shown in Figure 5-2.

The first part of the following lemma about the properties of the mapping $\mathbf{W}$ will be very useful in studying the WTA dynamical system.

Lemma 9 *The winner-take-all mapping* $\mathbf{W} : \Re^n \rightarrow T_n$ *satisfies the following properties:*

i) $\mathbf{z}_1^T(\mathbf{W}(\mathbf{z}_1) - \mathbf{W}(\mathbf{z}_2)) \geq 0$, $\forall \mathbf{z}_1, \mathbf{z}_2 \in \Re^n$.

ii) The function $\psi : \Re^n \rightarrow \Re$ *defined by* $\psi(\mathbf{z}) = \mathbf{z}^T\mathbf{W}(\mathbf{z})$ *is convex.*

Proof: i) Let $i = min\{j \in G | z_{1j} \geq z_{1m}, \forall m \in G\}$. Then by definition of $\mathbf{W}$, we have

$$\mathbf{z}_1^T\mathbf{W}(\mathbf{z}_1) = z_{1i}.$$

Moreover, $\mathbf{z}_1^T\mathbf{W}(\mathbf{z}_2) = z_{1m} \leq z_{1i}$. Therefore

$$\mathbf{z}_1^T(\mathbf{W}(\mathbf{z}_1) - \mathbf{W}(\mathbf{z}_2)) = z_{1i} - z_{1m} \geq 0.$$

2) To prove that ψ is convex, we need to prove that the inequality

$$\psi(t\mathbf{z}_1 + (1-t)\mathbf{z}_2) \leq t\psi(\mathbf{z}_1) + (1-t)\psi(\mathbf{z}_2)$$

is satisfied $\forall \mathbf{z}_1, \mathbf{z}_2 \in \Re^n$ and $\forall t \in [0,1]$. This is indeed the case since the function ψ is nothing but the function $\mathbf{z} = (z_1, \ldots, z_n) \rightarrow \max_i(z_i)$, which is a convex function. ■

There are two modes in which the state of (5.3) can be updated: synchronous and asynchronous.

1. Synchronous mode: In this operation mode, all the sites are free to change at each time step. In updating a site at time $k+1$, only the state of the system at time k is used. This update rule is similar to a Gauss-Jacobi iteration [64]. The way our dynamical system was defined in (5.3) corresponds to this operation mode.

2. Asynchronous mode: According to this operation mode, only one site s is updated at each time step k. As a result, the update of each site uses the most recent information about the state of the system. This update rule is similar to a Gauss-Seidel iteration [64]. At each time step k, we denote by s_k the site selected for updating. Then the operation mode becomes

$$\mathbf{y}_{s_{k+1}}(k+1) = \mathbf{W}(\mathbf{u}_s(k)) \tag{5.8}$$

$$\mathbf{u}_s(k) = \sum_{r \in \mathcal{S}} \mathbf{J}_{sr}\mathbf{y}_r(k) \tag{5.9}$$

$$\mathbf{y}_s(k+1) = \mathbf{y}_s(k), \text{ if } s \neq s_{k+1} \tag{5.10}$$

The initial condition of (5.3) and (5.8) will be denoted by $(\mathbf{y}_s(0), s \in \mathcal{S})$.

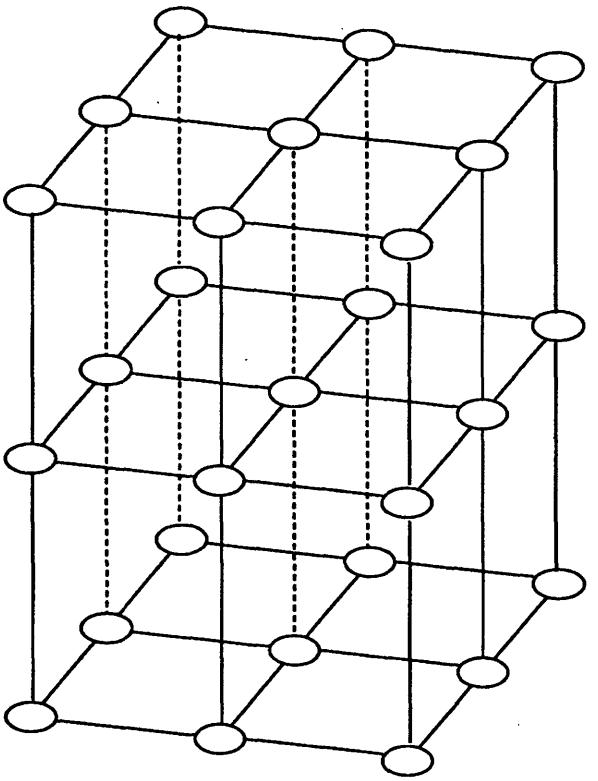


FIGURE 5-2: The winner-take-all network. Each column represents the state at a given site in the lattice. Each node in a given column is labeled either 0 or 1. Because of the uniqueness constraint, all the nodes are labeled 0 except one. So the sum along each column is 1. The horizontal layers represent the connectivity between the sites. In this particular case, we have nearest-neighbor connections [59].

5.2.2 A Lyapunov Function for the Discrete WTA Dynamics

Lyapunov functions are a standard tool for proving the stability of equilibrium points in the qualitative theory of ordinary differential equations [86]. They can also be defined and used for the study of the stability of fixed points in iterative processes and nonlinear difference equations [46]. However in all these contexts there is the underlying assumption that the state space is a continuum, i.e., $\Re^d$ or an open set of $\Re^d$. Moreover, in addition to the well-posedness conditions imposed on the dynamical system (Lipschitz continuity for the differential equation and continuity for the difference equation), the Lyapunov function itself is required to be continuous on the state space. The WTA iterative process that we have defined above has an underlying state space $\Omega = (T_n)^{|\mathcal{S}|}$ that is discrete and finite. Therefore the standard definitions of Lyapunov stability theory need to be adapted to the situation at hand. If $\mathbf{y} \in \Omega$, we denote by $\mathcal{K}\mathbf{y} \in \Omega$ the new state resulting from updating the WTA dynamics (either synchronous or asynchronous) to $\mathbf{y}$. Then we have the following definition

Definition 6 *A function $E : \Omega \rightarrow \Re$ is said to be a Lyapunov function for the discrete dynamical system*

$$\mathbf{y}(k+1) = \mathcal{K}\mathbf{y}(k),\ \mathbf{y}(0) \in \Omega$$

if the change

$$\Delta E \triangleq E(\mathcal{K}\mathbf{y}) - E(\mathbf{y}) \tag{5.11}$$

is nonpositive $\forall \mathbf{y} \in \Omega$.

The above definition implies that a Lyapunov function is always nonincreasing along the trajectories of the system. Moreover, since the state space is finite, any function is both bounded above and below. It follows that along any trajectory, a Lyapunov function always converges to a limit. The rest of this subsection will be devoted to finding Lyapunov functions for the WTA dynamical systems both in the asynchronous and synchronous modes.

Let E be the quadratic function defined on the state space Ω as follows

$$E(\mathbf{y}) = -\sum_{s,r \in \mathcal{S}} \mathbf{y}_s^T \mathbf{J}_{sr} \mathbf{y}_r. \tag{5.12}$$

For now we will call $E(\mathbf{y})$ the "energy" of the system configuration $\mathbf{y}^T = (\mathbf{y}_1, \ldots, \mathbf{y}_{|\mathcal{S}|})$. For the WTA dynamics defined in (5.3) or (5.8), denote by $E(k)$ the system energy at time k and by

$$\mathbf{d}_s(k) = \mathbf{y}_s(k+1) - \mathbf{y}_s(k) \tag{5.13}$$

the state transition vector at site s. Then we have

Proposition 10 *The one-step energy change in terms of the site state transition vectors is given by*

$$\Delta E_k \triangleq E_{k+1} - E_k = -2\sum_{s\in\mathcal{S}} \mathbf{d}_s^T(k)\mathbf{u}_s(k) - \sum_{s,r\in\mathcal{S}} \mathbf{d}_s^T(k)\mathbf{J}_{sr}\mathbf{d}_r(k). \tag{5.14}$$

Proof: To prove this equality, start from the right hand side and expand the second term to get the three-term expression

$$-\sum_{s,r\in\mathcal{S}} \mathbf{y}_s^T(k+1)\mathbf{J}_{sr}\mathbf{y}_r(k+1) - \sum_{s,r\in\mathcal{S}} \mathbf{y}_s^T(k)\mathbf{J}_{sr}\mathbf{y}_r(k) + 2\sum_{s,r\in\mathcal{S}} \mathbf{y}_s^T(k)\mathbf{J}_{sr}\mathbf{y}_r(k+1),$$

where we have used the symmetry of $\mathbf{J}_{sr}$ in the last term. Using the expression of $\mathbf{u}_s(k)$, the first term in the right-hand side of (5.14) gives

$$2\sum_{s,r\in\mathcal{S}} \mathbf{y}(k)_s^T\mathbf{J}_{sr}\mathbf{y}_r(k) - 2\sum_{s,r\in\mathcal{S}} \mathbf{y}_s^T(k)\mathbf{J}_{sr}\mathbf{y}_r(k+1).$$

Adding the last two expressions and using the definition of E we get the desired equality (5.14). ■

With this expression, we have the following

Proposition 11 *Under the assumption that $\mathbf{J}_{ss}$ is positive semidefinite, $\forall s \in \mathcal{S}$, the energy sequence $E_k, k \geq 0$ (5.12) is non-increasing along the trajectories of the asynchronous, discrete winner-take-all system (5.8). Therefore it is a Lyapunov function.*

Proof: Let s^* be the one and only site to be updated at time $k+1$. Then by Proposition 10, we have

$$\Delta E_k = -2\mathbf{d}_{s^*}^T(k)\mathbf{u}_{s^*}(k) - \mathbf{d}_{s^*}^T(k)\mathbf{J}_{s^*s^*}\mathbf{d}_{s^*}(k) \leq -\mathbf{d}_{s^*}^T(k)\mathbf{u}_{s^*}(k),$$

the term in $\mathbf{J}_{s^*s^*}$ being ≤ 0 by assumption. But we have

$$\mathbf{d}_{s^*}(k) = \mathbf{y}_{s^*}(k+1) - \mathbf{y}_{s^*}(k) = W(\mathbf{u}_{s^*}(k)) - W(\mathbf{u}_{s^*}(k-1)).$$

Applying the first part of Lemma 9, we get

$$\mathbf{d}_{s^*}^T(k)\mathbf{u}_{s^*}(k) \geq > 0.$$

It follows that $\Delta E_k \leq 0$ or $E_{k+1} \leq E_k$, which means that E is non-increasing along the trajectories of the asynchronous WTA dynamic system defined in (5.8) ■

Observe now that the function E is bounded below by

$$E_{\min} = -\sum_{s,r\in\mathcal{S}} \|\mathbf{J}_{sr}\|,$$

where $\|\mathbf{J}_{sr}\|$ is the matrix norm of $\mathbf{J}_{sr}$ induced by, for instance, the l_2 norm on $\Re^n$. The sequence $E_k, k \geq 0$, is non-increasing and bounded below. Therefore it converges to a limit E_{asyn}.

It is important to note that the above proposition remains valid even if the block matrices $\mathbf{J}_{ss}, s \in \mathcal{S}$, are all zero. When they are positive definite, the Lyapunov sequence becomes *strictly* decreasing. Hence, we have the following

Corollary 1 *Under the assumption that $\mathbf{J}_{ss}$ is positive definite, $\forall s \in \mathcal{S}$, the energy sequence $E_k, k \geq 0$ (5.12) is strictly decreasing along the trajectories of the asynchronous, discrete winner-take-all system (5.8). Therefore, it always converges to a fixed point.*

The above corollary allows us to rule out oscillations in asynchronous, discrete, WTA networks in which the autoconnection matrices are positive definite.

Note that the energy sequence along the trajectories of the *synchronous* WTA dynamic system (5.3) is not guaranteed to be non-increasing unless we assume that the block matrix $\mathbf{J} = [\mathbf{J}_{sr}]_{s,r \in \mathcal{S}}$ is positive semidefinite.

Proposition 12 *Suppose that the matrix $\mathbf{J} = [\mathbf{J}_{sr}]_{s,r \in \mathcal{S}}$ is positive semidefinite. Then the energy sequence $E_k, k \geq 0$, (5.12) is non-increasing along the trajectories of the synchronous, discrete winner-take-all system (5.3). Therefore it is a Lyapunov function.*

Proof: First, we define the $n|\mathcal{S}|$-dimensional vector $\mathbf{d} = \left(\mathbf{d}_1^T, \ldots, \mathbf{d}_{\mathcal{S}}^T\right)^T$. With this and the definition of the block matrix $\mathbf{J}$, the second term in the expression of ΔE_k (5.14) becomes

$$\Delta E_k = -2 \sum_{s \in \mathcal{S}} \mathbf{d}_s^T(k) \mathbf{u}_s(k) - \mathbf{d}^T(k) \mathbf{J} \mathbf{d}(k).$$

Let now

$$\Sigma(k) = \{s \in \mathcal{S} | \mathbf{y}_s(k+1) \neq \mathbf{y}_s(k)\}$$

be the subset of sites whose states have changed at the $(k+1)$st iteration. Then

$$\begin{aligned} \sum_{s \in \mathcal{S}} \mathbf{d}_s^T(k) \mathbf{u}_s(k) &= \sum_{s \in \Sigma(k)} \left(\mathbf{y}_s(k+1) - \mathbf{y}_s(k)\right)^T \mathbf{u}_s(k) \\ &= \sum_{s \in \Sigma(k)} \left(W(\mathbf{u}_s(k)) - W(\mathbf{u}_s(k-1))\right)^T \mathbf{u}_s(k) \\ &\geq 0, \end{aligned}$$

where the above inequality results from Lemma 9, part (i). From the above inequality and the fact that $\mathbf{J}$ is positive semidefinite, we get that $\Delta E_k \leq 0$ or $E_{k+1} \leq E_k$, i.e., the energy sequence $E_k, k \geq 0$, is non-increasing along the trajectories of the synchronous, discrete WTA dynamic system (5.3). ∎

By the above proposition and the lower-boundedness of the energy function, it results that the energy sequence $E_k, k \leq 0$, converges to a limit E_{syn}. Note that both E_{asyn} and E_{syn} depend on the initial condition of the WTA system.

The condition imposed on the matrix $\mathbf{J}$ in Proposition 12 is stronger than needed. Indeed for the energy sequence $E_k, k \geq 0$, to be non-increasing along the trajectories of (5.3), it is sufficient that $\mathbf{J}$ be non-negative on a set $\mathcal{D}$ of vectors defined as follows:

$$\mathcal{D} \triangleq \{\mathbf{d} \in \Re^{n|\mathcal{S}|} \text{ such that } \exists \mathbf{y}, \mathbf{y}' \in \Omega, \mathbf{y} \neq \mathbf{y}', \ \mathbf{d} = \mathbf{y} - \mathbf{y}'\}. \tag{5.15}$$

In words, a vector $\mathbf{d} = \left(\mathbf{d}_1^T, \ldots, \mathbf{d}_s^T, \ldots, \mathbf{d}_{|\mathcal{S}|}^T\right)^T$ is in $\mathcal{D}$ if and only if every $\mathbf{d}_s \in \Re^n$ can be written as a difference $\mathbf{y}_s - \mathbf{y}'_s$, where $\mathbf{y}_s$ and $\mathbf{y}'_s$ are two vertices of $\mathcal{T}_n$. Note that $\mathcal{D}$ does not contain the zero vector. Therefore there should exist at least one site $s \in \mathcal{S}$ such that $\mathbf{y}_s \neq \mathbf{y}'_s$.

An interesting situation arise when the diagonal blocks of $\mathbf{J}$ are all zero. This is usually the case when the the matrix $\mathbf{J}$ is induced by the coefficients of a Markov random field with pairwise interactions and no external field. Then in this case, only along the asynchronous WTA trajectories is the quadratic energy guranteed to be non-increasing. This is because of the following

Proposition 13 *Let* $\mathbf{P}$ *be an* $N \times N$ *symmetric matrix such that* $P_{ll} = 0$ *for all* $1 \leq l \leq N$. *Then* $\mathbf{P}$ *is positive semidefinite if and only if* $\mathbf{P}$ *is the zero matrix.*

Proof: The "if" part is trivial. For the "only if" part, notice that on the one hand,

$$trace(\mathbf{P}) = \sum_{l=1}^{N} P_{ll} = 0.$$

On the other hand, the eigenvalues $\lambda_l, 1 \leq l \leq N$, of $\mathbf{P}$ are all nonnegative real, and since the trace is an invariant for $\mathbf{P}$

$$trace(\mathbf{P}) = \sum_{l=1} \lambda_l = 0$$

It follows that $\lambda_l = 0$ for all $1 \leq l \leq N$. Therefore $\mathbf{P}$ is the zero matrix. ∎

5.2.3 Thresholds and External Inputs

It is perfectly possible to change the dynamics of the WTA so that thresholds and external inputs are included. For instance, the synchronous system 5.3 becomes

$$\mathbf{y}_s(k+1) = \mathbf{W}(\mathbf{u}_s(k)) \tag{5.16}$$

$$\mathbf{u}_s(k) = \sum_{r \in \mathcal{S}} \mathbf{J}_{sr}\mathbf{y}_r(k) + \mathbf{h}_s - \mathbf{t}_s, \ \forall s \in \mathcal{S}, \ \forall k \geq 0 \tag{5.17}$$

where $\mathbf{h}_s$ is now an external vector field applied at site s and $\mathbf{t}_s$ is a vector of thresholds. The computation of the winning components in the vector $\mathbf{u}_s$ must of course involve both the components of the external field $\mathbf{h}_s$ and the components of the threshold vector $\mathbf{t}_s$. It is not difficult to check that the energy function given below is bounded and non-increasing along the trajectories of (5.16)

$$E = -\frac{1}{2}\sum_{s,r\in\mathcal{S}} \mathbf{y}_s^T \mathbf{J}_{sr}\mathbf{y}_r - \sum_{s\in\mathcal{S}} \mathbf{y}_s^T \mathbf{h}_s + \sum_{s\in\mathcal{S}} \mathbf{y}_s^T \mathbf{t}_s \tag{5.18}$$

if the block matrix $\mathbf{J} = [\mathbf{J}_{sr}]$ is positive semidefinite. Similarly, when the matrices $\mathbf{J}_{ss}$ are all zeros, the above function defines non-increasing sequence along the trajectories of the asynchronous WTA system with external inputs and thresholds.

5.2.4 Global Dynamics of the Synchronous Discrete WTA Network

The convergence of the Lyapunov sequence E_k to a lower bound is an important result but it does not tell us very much about the global dynamics of the system trajectories. However because of the finiteness of the state space Ω, every orbit $(\mathbf{y}(k))_{k\geq 0}$ starting at an arbitrary $\mathbf{y}(0) \in \Omega$ will converge to a limit cycle. In other words, there will be $l \geq 0, p > 0$ such that the finite sequence $\mathbf{y}(l), \mathbf{y}(l+1), \ldots, \mathbf{y}(l+p-1)$ satisfies $\mathbf{y}(l+i) \neq \mathbf{y}(l+j), 0 \leq i < j \leq p-1$ and $\mathbf{y}(l+p) = \mathbf{y}(p)$. The integer p is the period of the cycle. Our objective in this paragraph is to prove that under a mild condition on the matrix $\mathbf{J}$ the only limit cycles of the synchronous WTA system are the fixed points. This will be done by introducing a new Lyapunov function inspired by the work of Marcus and Westervelt on the analog Hopfield network [56].

Let

$$\Phi(\mathbf{u}) \triangleq -\sum_{s\in\mathcal{S}} \mathbf{W}^T(\mathbf{u}_s)\mathbf{u}_s \tag{5.19}$$

here $\mathbf{W}$ is the winner-take-all mapping defined in (5.3). As usual, we are interested in computing the change

$$\Delta\Phi_k \triangleq \Phi_{k+1} - \Phi_k$$

along the trajectories of the WTA dynamical system, where $\Delta\Phi_k = \Phi(\mathbf{u}(k))$. Note that

$$\begin{aligned}
\Phi_k &= -\sum_{s\in\mathcal{S}} \mathbf{W}^T(\mathbf{u}_s(k))\mathbf{u}_s(k) \\
&= -\sum_{s\in\mathcal{S}} \mathbf{y}_s^T(k+1)\mathbf{u}_s(k) \\
&= -\sum_{s,r\in\mathcal{S}} \mathbf{y}_s^T(k+1)\mathbf{J}_{sr}\mathbf{y}_r(k)
\end{aligned} \tag{5.20}$$

which means that at each time step both the current state and the next state are involved in the computation of Φ_k. Contrast this with the Lyapunov function E, where only the current state is needed in the computation. We have the following

Proposition 14 *The function Φ defined in (5.19) is a concave Lyapunov function for the synchronous WTA dynamic system (5.3).*

Proof: Using the definition of Φ (5.19) and the first part of Lemma 9, we can see that Φ is concave as the negative of a sum of convex functions. Now from (5.20) we have

$$\Delta\Phi_k = -\sum_{s,r\in\mathcal{S}} (\mathbf{y}_s(k+2) - \mathbf{y}_s(k))^T \mathbf{J}_{sr}\mathbf{y}_r(k+1),$$

which can also be written as

$$\Delta\Phi_k = -\sum_{s\in\mathcal{S}} (\mathbf{W}(\mathbf{u}_s(k+1)) - \mathbf{W}(\mathbf{u}_s(k-1)))^T \mathbf{u}_s(k+1),$$

where we have used the update equations of the synchronous WTA system (5.3). Now using the first part of Lemma 9, we get that

$$(\mathbf{W}(\mathbf{u}_s(k+1)) - \mathbf{W}(\mathbf{u}_s(k-1)))^T \mathbf{u}_s(k+1) \geq 0,\ \forall s \in \mathcal{S},\ \forall k \geq 1.$$

Therefore, $\Delta\Phi_k \leq 0$ and the sequence $(\Phi_k)_{k\geq 0}$ is non-increasing along the trajectories of the synchronous WTA system. ∎

Here also for each initial condition, $\mathbf{y}(0) \in \Omega$, the sequence $(\Phi_k)_{k\geq 0}$ will converge to a limit that is a minimum for the orbit $(\mathbf{y}(t))_{t\geq 0}$.

The objective of the rest of this section is to prove that if the block matrix $\mathbf{J}$ satsfies a weak form of positive definiteness, then all the limit cycles of the synchronous WTA system are of period 1, i.e, they are all fixed points for the discrete system (5.3). The following lemma on the relationship between the two Lyapunov sequences $(E_k)_{k\geq 0}$ and $(\Phi_k)_{k\geq 0}$ along the trajectories of (5.3) is fundamental for the proof.

Lemma 10 *Assume the block matrix $\mathbf{J}$ positive semidefinite. Then the sequences $(E_k)_{k\geq 0}$ and $(\Phi_k)_{k\geq 0}$ are interlaced along the trajectories of System (5.3), i.e.,*

$$E_{k+1} \leq \Phi_k \leq E_k \leq \Phi_{k-1} \quad \forall k \geq 0.$$

Proof: First, let us write E_k in terms of the auxiliary variables $\mathbf{u}_s(k)$ and the winner-take-all mapping $\mathbf{W}$. From (5.12) and the definition of $\mathbf{u}_s$, we get

$$E_k = -\sum_{s\in\mathcal{S}} \mathbf{u}_s^T(k)\mathbf{W}\left(\mathbf{u}_s(k-1)\right).$$

Now

$$\Phi_k - E_k = -\sum_{s\in\mathcal{S}} \mathbf{u}_s^T(k)\left(\mathbf{W}(\mathbf{u}_s(k)) - \mathbf{W}(\mathbf{u}_s(k-1))\right).$$

Applying the first part of Lemma (9) to every term in the above summation, we get

$$\Phi_k - E_k \leq 0.$$

Similarly,

$$\begin{aligned}\Phi_k - E_{k+1} &= -\sum_{s\in\mathcal{S}} \left(\mathbf{u}_s(k) - \mathbf{u}_s(k+1)\right)^T \mathbf{W}(\mathbf{u}_s(k)) \\ &= -\sum_{s,r\in\mathcal{S}} \left(\mathbf{y}_s(k) - \mathbf{y}_s(k+1)\right)^T \mathbf{J}_{sr}\left(\mathbf{y}_r(k+1) - \mathbf{y}_r(k) + \mathbf{y}_r(k)\right),\end{aligned}$$

where we have used the definitions of the auxiliary variables $\mathbf{u}_s$ and the iteration equations (5.3). Moreover, we have added and subtracted the term $\mathbf{y}_r(k)$ so as to make use of the nonnegativity of the matrix $\mathbf{J}$ and write

$$-\sum_{s,r\in\mathcal{S}} \left(\mathbf{y}_s(k) - \mathbf{y}_s(k+1)\right)^T \mathbf{J}_{sr}\left(\mathbf{y}_r(k+1) - \mathbf{y}_r(k)\right) \geq 0.$$

Therefore

$$\begin{aligned}\Phi_k - E_{k+1} &\geq -\sum_{s,r\in\mathcal{S}} \left(\mathbf{y}_s(k) - \mathbf{y}_s(k+1)\right)^T \mathbf{J}_{sr}\mathbf{y}_r(k) \\ &= -\sum_{s\in\mathcal{S}} \mathbf{u}_s^T(k)\left(\mathbf{W}(\mathbf{u}_s(k-1) - \mathbf{W}(\mathbf{u}_s(k))\right)\end{aligned}$$

where we have used again (5.3). Applying Lemma 9 to every term of the right-hand side we get the desired inequality $\Phi_k \geq E_{k+1}$. ∎

We are in position to state a major theorem for the global dynamics of synchronous WTA networks, but we need first to introduce some new notation. For each site $s \in \mathcal{S}$, let H_s be the vector space in $\Re^n$ spanned by the edges of the simplex $\mathcal{T}_n$. This space is of dimension $n-1$. Denote by

$$H = \oplus_{s\in\mathcal{S}} H_s$$

The direct sum of these subspaces. Note that if $\mathbf{v}$ and $\mathbf{v}'$ are in Ω then $\mathbf{d} = \mathbf{v} - \mathbf{v}'$ is in H. We say that the matrix $\mathbf{J}$ is positive definite on H if $\mathbf{v}^T\mathbf{J}\mathbf{v} > 0$, for all $\mathbf{v} \in H, \mathbf{v} \neq 0$.

Theorem 2 *Suppose that the block matrix* $\mathbf{J}$ *is positive definite on* H. *Then all the trajectories of the synchronous, discrete winner-take-all network converge to fixed points.*

Stated differently, this theorem asserts that under a mild assumption on the matrix $\mathbf{J}$ describing the interaction between the different sites, the synchronous winner-take-all

network has no limit cycles of period ≥ 1. One can also show that if this condition is not satisfied then limit cycles of period 2 can occur.

Proof: We know that because of the finiteness of the state space Ω, every trajectory converge to a limit cycle of period, say p. Our objective is to prove that $p = 1$. So let $\mathbf{y}(1)$ be an initial point on the limit cycle. Then by periodicity $\mathbf{y}(1) = \mathbf{y}(p)$. Therefore, $E_1 = E_p$ and $\Phi_1 = \Phi_p$. By Propositions 12 and 14 both $(E_k)_{k\geq 0}$ and $(\Phi_k)_{k\geq 0}$ are non-increasing along the trajectories of (5.3). Using Lemma 10, we obtain

$$E_1 = \Phi_1 = E_2 = \Phi_2 = \ldots = E_p = \Phi_p,$$

i.e., both sequences are constant over the limit cycle, and their values are equal. Let $\mathbf{d}(k) = \mathbf{y}(k+1) - \mathbf{y}(k), 1 \leq k \leq (p-1)$ be the transition vector between any two successive points on the limit cycle. We want to prove that $\mathbf{d}(k) = 0$, i.e, that the *state vector* is also constant. Indeed by Proposition 10, it is easy to check that

$$\mathbf{d}^T(k)\mathbf{J}\mathbf{d}(k) = E_k - E_{k+1} + 2(\Phi_k - E_k).$$

But because the Lyapunov sequences are constant, we get

$$\mathbf{d}^T(k)\mathbf{J}\mathbf{d}(k) = 0.$$

which implies, since $\mathbf{d}_k \in H$ and since the block matrix $\mathbf{J}$ is positive definite on H, that $d_k, 1 \leq k \leq p-1$, itself is $\mathbf{0}$. In other words, the limit cycle is reduced to a single point. ■

We will see that this global property remains valid when we replace the winner-take-all mapping an analog approximation, which will be done later in this chapter. [1]

5.2.5 Relationship to the Binary Hopfield Network

In this short subsection, we show that the binary Hopfield network can be easily mapped onto a discrete WTA network whose dynamics is governed by one of the two systems (5.3) or (5.8).

We assume the reader familiar with Hopfield's 1982 paper. Denote the connectivity matrix of the Hopfield model by $\mathcal{T} = [T_{sr}]_{s,r\in\mathcal{S}}$. The neurons $V_s, s \in \mathcal{S}$, of the Hopfield model can have one of two values 0 or 1. Using the method of binary matching elements, we can map 0 and 1 respectively onto the vertices $(0,1)^T$ and $(1,0)^T$ of the 2D simplex

[1] Note that as for Proposition 12, it would have been enough to assume that the matrix $\mathbf{J}$ is positive on the set $\mathcal{D}$ defined in (5.15). We have preferred to give the stronger sufficient condition of positive definiteness on H because it is the one that is needed for the analog case.

$x_1 + x_2 = 1,\ x_1, x_2 \geq 0$. The 2D binary vector corresponding to V_s will be denoted $\mathbf{v}_s$, and as a result of the mapping we have $\mathbf{v}_s = (V_s, 1 - V_s)^T$. For each pair $(s,r) \in \mathcal{S} \times \mathcal{S}$ define the interaction *matrix* $\mathbf{J}_{sr}$ by

$$\mathbf{J}_{sr} = \begin{bmatrix} T_{sr} & 0 \\ 0 & 0 \end{bmatrix}.$$

Hopfield's dynamical system is defined as in [35] by [2]

$$\begin{aligned} V_s &\rightarrow 1, \text{ if } \sum_{r \in \mathcal{S}} T_{sr} V_r > 0 \\ V_s &\rightarrow 0, \text{ if } \sum_{r \in \mathcal{S}} T_{sr} V_r < 0. \end{aligned}$$

Now denote by $\mathbf{u}_s = (u_{s0}, u_{s1})^T$ the sum [3]

$$\sum_{r \in \mathcal{S}} \mathbf{J}_{sr} \mathbf{V}_r.$$

From the definition of $\mathbf{J}_{sr}$, it is easy to see that the Hopfield dynamics can be written as

$$\mathbf{v}_s \rightarrow (1,0)^T, \text{ if } u_{s1} > u_{s0}$$

$$\mathbf{v}_s \rightarrow (0,1)^T, \text{ if } u_{s0} > u_{s1}.$$

Therefore the update conditions for both the Hopfield model and the WTA model are equivalent. Using the definition of the WTA mapping $\mathbf{W}$, the above dynamics can simply be written as

$$\mathbf{v}_s(k+1) = \mathbf{W}(\mathbf{u}_s(k))$$

Another mapping of the binary Hopfield model onto a discrete WTA model is to use the following definition of interaction matrices

$$\mathbf{J}_{sr} = \begin{bmatrix} T_{sr} & -T_{sr} \\ -T_{sr} & T_{sr} \end{bmatrix}.$$

The reader will easily verify that the quadratic "energy" corresponding to these interaction matrices is the same as the one that would result from replacing the zero-one values of the neuron outputs with ± 1. The identification of the binary Hopfield network with a WTA network implies that all our conclusions about the global dynamics of the WTA network are also true for the Hopfield network.

[2] It is worthwhile to mention that we do not exclude the site s itself from the summation giving the total action of the network on the neuron at site s. Also, Hopfield considered only the asynchronous update mode, although he mentioned the possibility of using a synchronous one.

[3] Because of the definition of $\mathbf{J}_{sr}$, the first component of $\mathbf{u}_s$ is always equal to 0, hence the notation u_{s0}.

5.3 Analog Winner-Take-All Networks

In this section, we give the analog counterpart of the discrete WTA networks described in the previous section. The only heuristic step in our presentation is the use of the mean-field approximation to pass from the discrete WTA network to the analog one. Although an axiomatic presentation of the WTA dynamical system is possible, we have preferred this path for three reasons. The first is to indicate that there is a strong link between the discrete network and the analog one. The second is to show how the generalized sigmoid mapping, which is the fundamental building block in the analog WTA network, arises in a rather natural way when we use statistical physics to find analog counterparts to discrete systems defined on the finite state space Ω. The third reason is to point out the role of thermodynamics in studying the global dynamics of the analog WTA networks. This section is organized into several subsections. In Subsection 5.3.1, we use mean-field theory to derive the analog WTA network from the discrete one defined in (5.3). Subsection 5.3.2 is devoted to a detailed study of some of the properties of the generalized sigmoid mapping. These properties will play a central role in Subsection 5.3.3 where the dynamics of the synchronous WTA network is studied. Our major result here is that all the limit cycles have either a period of 2 or are fixed points. This result extends, to the analog WTA network, the one that Marcus and Westervelt [56] obtained for the analog Hopfield network. Moreover, if the connectivity matrix of the network satisfies a weak form of positive-definiteness that we will specify, then the only limit cycles are the fixed points. Analog hardware implementations of the generalized sigmoid mapping, the building block of the WTA network, are given in Section 5.4.

5.3.1 Derivation of the Analog WTA Network

To make this chapter self contained, we give a detailed derivation of the fixed-point equations of the analog WTA network using one of the methods presented in Chapter 3.

We have shown in the previous section that the quadratic energy function

$$E(\mathbf{y}) = -\sum_{s,r\in\mathcal{S}} \mathbf{y}_s^T \mathbf{J}_{sr} \mathbf{y}_r$$

is non-increasing along the trajectories of the synchronous discrete WTA network (5.3) if the connectivity matrix $\mathbf{J} = [\mathbf{J}_{sr}]_{(s,r)\in\mathcal{S}^2}$ is positive semidefinite. We have also shown that if the autoconnection matrices $\mathbf{J}_{ss}$ are positive semidefinite, then the same energy function is non-increasing along the trajectories of the asynchronous, discrete WTA network (5.8). The above energy function can be used to define a Gibbs distribution on the state space

Ω. The partition function (refer to Chapter 1) of this distribution is given by

$$Z = \sum_{\mathbf{y}\in\Omega} \exp\left[-\frac{1}{T}E(\mathbf{y})\right].$$

We are now in a situation similar to that of Section 3.2 where we have used the method of binary matching elements to transform the Gibbs distribution of a multilevel pixel state space to a Gibbs distribution in which the pixel state space is the set T_n of the simplex $\mathcal{T}_n$ vertices. The fastest way to derive the fixed-point equations of the analog WTA network is to proceed as in Subsection 3.2.3. This method has the advantage of not requiring that $\mathbf{J}$ be positive semidefinite on the block matrix $\mathbf{J}$. However for the algebra to work out correctly, we should assume that the auto-connect matrices $\mathbf{J}_{ss} = \mathbf{0}, \forall s \in \mathcal{S}$. This situation is typical of Markovian fields. Denote by $\mathbf{v}_s$ an approximation of the ensemble average $\overline{\mathbf{y}}_s$ of $\mathbf{y}_s, s \in \mathcal{S}$, with respect of the Gibbs distribution. Let

$$\mathbf{w}_s = \sum_{r\in\mathcal{S}} \mathbf{J}_{sr}\mathbf{v}_r + \mathbf{h}_s$$

be an approximation of the effect of the lattice on the state of site s. We have included in $\mathbf{w}_s$ the effect of an external field $\mathbf{h}_s, s \in \mathcal{S}$, so as to be able to get consistency relations for $\mathbf{v}_s$ at zero external field.

An approximation $\hat{Z}$ of the partition function can be obtained by replacing the actual lattice field effect at site s with the approximation $\mathbf{w}_s$ of its mean. Thus

$$\begin{aligned} Z \approx \hat{Z} &= \sum_{\mathbf{y}\in\Omega} \exp\left[\frac{1}{T}\mathbf{y}_s^T\mathbf{w}_s\right] \\ &= \sum_{\mathbf{y}_1\in T_n} \cdots \sum_{\mathbf{y}_S\in T_n} \exp\left[\frac{1}{T}\sum_{s\in\mathcal{S}}\mathbf{y}_s^T\mathbf{w}_s\right] \\ &= \sum_{\mathbf{y}_1\in T_n} \cdots \sum_{\mathbf{y}_S\in T_n} \prod_{s\in\mathcal{S}} \exp\left[\frac{1}{T}\mathbf{y}_s^T\mathbf{w}_s\right]. \end{aligned}$$

Since the vectors $\mathbf{y}_s$ can only be the vertices of the simplex $\mathcal{T}_n$, the above summation can be written as

$$\hat{Z} = \prod_{s\in\mathcal{S}} \sum_{g\in G} \exp(w_{sg}/T).$$

Using the formula

$$\mathbf{v}_s = T\left(\frac{\partial \log \hat{Z}}{\partial \mathbf{h}_s}\right)_{\mathbf{h}=0},$$

where

$$\log \hat{Z} = \sum_{s\in\mathcal{S}} \log \sum_{g\in G} \exp(w_{sg}/T), \tag{5.21}$$

we can obtain the following set of *consistency* relationships for the vectors $\mathbf{v}_s, s \in \mathcal{S}$

$$\mathbf{v}_s = \mathbf{F}(\frac{1}{T}\mathbf{w}_s), \ \forall s \in \mathcal{S}, \tag{5.22}$$

where the mapping $\mathbf{F}$ is the generalized sigmoid mapping that was introduced in Chapter 3. This mapping is a continuous one from $\Re^n$ into the simplex $\mathcal{T}_n$ whose j-th component is defined by

$$F_j(\mathbf{z}) = \frac{e^{z_j}}{\sum_{i=1}^n e^{z_i}}, \ \forall \mathbf{z} = (z_1, z_2, \ldots, z_n)^T \in \Re^n, \tag{5.23}$$

Often we will deal with the case where the argument of the generalized sigmoid mapping is scaled with the inverse temperature $\frac{1}{T}$. In this case, we will use the notation

$$\mathbf{F}_T(\mathbf{z}) = \mathbf{F}(\frac{1}{T}\mathbf{z}), \ \mathbf{z} \in \Re^n.$$

The consistency relationships given in (5.22) are nothing but the fixed-point equations of the discrete dynamical system

$$\begin{aligned} \mathbf{v}_s(k+1) &= \mathbf{F}_T(\mathbf{w}_s(k)) \\ \mathbf{w}_s(k) &= \sum_{s,r \in \mathcal{S}} \mathbf{J}_{sr}\mathbf{v}_r(k). \end{aligned} \tag{5.24}$$

Here also, we can define a synchronous mode of operation and an asynchronous one. In the sequel, we will concentrate on the synchronous system in which all sites are free to update their states at each time step $k+1$ and only information available at time step $k \geq 0$ is used in the update.

Because the components of $\mathbf{F}$ are nonnegative and sum up to 1, the state $\mathbf{v}_s(k) \in \mathcal{T}_n$. Therefore, the state space of the above discrete dynamical system is the compact, convex set $\mathbf{K} = (\mathcal{T}_n)^{|\mathcal{S}|}$. Formally this system can be written as

$$\mathbf{v}(k+1) = \mathcal{K}_T(\mathbf{v}(k)), \ \mathbf{v}(0) \in \mathbf{K}, \tag{5.25}$$

where $\mathcal{K}_T$ is the nonlinear mapping induced by $\mathbf{F}_T$ and the connectivity matrices $[\mathbf{J}_{sr}]_{s,r \in \mathcal{S}}$. A priori it is not clear that a fixed point for the dynamical system (5.24) exists at all. In fact there is at least one: it is the fixed point whose existence is guaranteed by Brouwer's theorem ([13], p. 149), that we state here for ease of reference.

Theorem 3 *If* $\mathbf{K}$ *is a nonempty compact convex subset of* $\Re^d$ *and* $\mathcal{K} : \mathbf{K} \to \mathbf{K}$ *is a continuous map, then there is a point* $\mathbf{v} \in \mathbf{K}$ *such that* $\mathbf{v} = \mathcal{K}(\mathbf{v})$.

The conditions of the above theorem are satisfied by our mapping $\mathcal{K}_T$ and the subset $\mathbf{K}$. The problem remains whether the trajectories of (5.24) converge to one of the fixed points of $\mathcal{K}$. A sufficient condition for this to happen will be given in Subsection 5.3.3. For now we need to justify the name "analog winner-take-all network" given to the system (5.24). This will become clear after we give some of the properties of the generalized sigmoid mapping in the next subsection.

5.3.2 The Generalized Sigmoid Mapping

The generalized sigmoid mapping $\mathbf{F}$ is as old as Statistical Mechanics. In fact it corresponds to the Gibbs probability distribution assigned to a thermodynamic system in equilibrium with a heat source at temperature $T = 1$ having n states with energies specified by the real numbers $-z_1, \ldots, -z_n$. Indeed in this case, the probability p_j that the system be in the j-th energy state is given by

$$p_j = \frac{e^{z_j}}{\sum_{i=1}^{n} e^{z_i}},$$

which is exactly the expression of the j-th component of the generalized sigmoid mapping as given in (5.23). This probabilistic interpretation of $\mathbf{F}$ is useful in its own right. Because the exponential is an increasing function, the mapping $\mathbf{F}$ assigns the highest probability to the component j such that $z_j \geq z_l, l \neq j$. In fact there is a close relationship between $\mathbf{F}$ and the winner-take-all mapping $\mathbf{W}'$ defined in (5.6) for we have the following proposition

Proposition 15 *The mapping F_T converges pointwise to $\mathbf{W}'$ over $\Re^n$ as $T \to 0$ and to $\frac{1}{n}\mathbf{e} = \frac{1}{n}(1, 1, \ldots, 1)^T \in \Re^n$ as $T \to \infty$.*

Proof: Easy proof. ∎

Now denote by $\mathcal{H}_{ij}, 1 \leq i \neq j \leq n$ the hyperplanes $z_i = z_j$ in $\Re^n$, and let $\mathcal{H}$ be the union of all these hyperplanes. Then it should be clear from the previous proof that the mapping $\mathbf{F}_T$ converges pointwise to the winner-take-all mapping $\mathbf{W}$ defined in (5.7) as $T \to 0$ over $\Re^n \backslash \mathcal{H}$. It follows that except for the zero- measure set $\mathcal{H}$, the dynamical system (5.3) with the finite discrete state space Ω is the *zero temperature* limit of the dynamical system (5.24) with the continuous state space $\mathcal{K}$. We will show in the next subsection that *even* at nonzero temperature, both systems have similar behavior when the matrix $\mathbf{J}$ is positive definite on H.

Although most of the following properties of $\mathbf{F}$ are easy to prove, we have decided to assemble them here since to the best of our knowledge, they do not appear elsewhere in the literature.

Proposition 16 *The generalized sigmoid mapping is invariant under translation along the vector* **e**.

Proof: Let $\alpha \in \Re$. Then easy algebra shows for all $\mathbf{z} \in \Re^n$ that

$$\mathbf{F}(\mathbf{z} + \alpha\mathbf{e}) = \mathbf{F}(\mathbf{z}).$$

■

The above proposition looks benign but will be used in the next subsection to finish our proof that the limit cycles of the dynamical system (5.24) are either fixed points or have a period of 2. When the consistency relationships (5.22) were derived, **F** was obtained through a differentiation process. Hence,

Proposition 17 *The generalized sigmoid mapping is a gradient map, i.e, there exists* $\mathcal{P} : \Re^n \rightarrow \Re$ *such that*

$$\mathbf{F} = \nabla\mathcal{P}.$$

Proof: Write

$$\mathcal{P}(\mathbf{z}) \triangleq \log \sum_{g=1}^{n} \exp(z_g), \forall \mathbf{z} \in \Re^n. \tag{5.26}$$

Then it is clear that $\mathbf{F}(\mathbf{z}) = \nabla\mathcal{P}(\mathbf{z})$. ■

The symbol $\mathcal{P}$ was chosen to indicate that it is a *potential* function. It is known that the derivative of the sigmoid function can be expressed in terms of the sigmoid function itself [36]. The generalized sigmoid mapping exhibit a similar property, for we have

Proposition 18 *The Jacobian* $D\mathbf{F}$ *of the generalized sigmoid mapping is a symmetric* $n \times n$ *matrix that satisfies*

$$D\mathbf{F}(\mathbf{z}) = diag\,(F_g(\mathbf{z})) - \mathbf{F}(\mathbf{z})\mathbf{F}(\mathbf{z})^T. \tag{5.27}$$

Moreover, it is always singular with the vector **e** *being the only eigenvector corresponding to the zero eigenvalue.*

Proof: Let $\mathbf{z} \in \Re^n$. Then we have by definition for $1 \leq j \leq n$

$$F_j(\mathbf{z}) = \frac{e^{z_j}}{\sum_{l=1}^{n} e^{z_l}}.$$

Let now $i \neq j$ be in $\{1, \ldots, n\}$. Then the partial derivative of F_j with respect to z_i is given by

$$\begin{aligned} \frac{\partial F_j(\mathbf{z})}{\partial z_i} &= -\frac{e^{z_j} e^{z_i}}{\left(\sum_{l=1}^{n} e^{z_l}\right)^2} \\ &= -F_j(\mathbf{z}) F_i(\mathbf{z}) \end{aligned}$$

Similarly, for $i = j$ we have

$$\begin{aligned}\frac{\partial F_j(\mathbf{z})}{\partial z_j} &= \frac{e^{z_j}}{\sum_{l=1}^n e^{z_l}} - \frac{e^{z_j}e^{z_i}}{\left(\sum_{l=1}^n e^{z_l}\right)^2} \\ &= F_j(\mathbf{z}) - F_j(\mathbf{z})F_j(\mathbf{z})\end{aligned}$$

Assembling the above expressions in a matrix form, we get (5.27). To simplify notation define $\mathbf{f} \triangleq \mathbf{F}(\mathbf{z})$, and $\mathbf{M} \triangleq D\mathbf{F}(\mathbf{z})$. Note that

$$f_j > 0,\ 1 \leq j \leq n,\ \sum_{j=1}^n f_j = 1,$$

and

$$\mathbf{M} = diag(f_j) - \mathbf{f}\mathbf{f}^T.$$

Then $\mathbf{M}\mathbf{e} = \mathbf{f} - \mathbf{f}\mathbf{f}^T\mathbf{e}$. But $\mathbf{f}^T\mathbf{e} = \sum_j f_j = 1$. Therefore $\mathbf{M}\mathbf{e} = \mathbf{0}$. This proves that $\mathbf{e}$ is an eigenvector associated to the eigenvalue 0. It also proves that the row sum of $\mathbf{M}$ is constant and equal to 0. Since $\mathbf{M}$ is symmetric this is also true for the column sum. Let us now prove that the null space of $\mathbf{M}$ is generated by $\mathbf{e}$. So let ζ be another vector in the null space of $\mathbf{M}$ then the i-th component of $\mathbf{M}\zeta$ satisfies

$$f_i\zeta_i - f_i\mathbf{f}^T\zeta = 0,\ 1 \leq i \leq n.$$

Since $f_i > 0$, we get that

$$\zeta_i = \mathbf{f}^T\zeta,\ 1 \leq i \leq n.$$

In other words, all the components of ζ are equal, i.e, $\zeta \in \text{span}(\mathbf{e})$. ■

The Jacobian of $\mathbf{F}$ is a symmetric matrix and therefore is diagonalizable using an orthogonal matrix. Note that the first column of this orthogonal matrix can be taken equal to $\mathbf{e}$. Note also that $\mathbf{e}^\perp$, the vector subspace orthogonal to $\mathbf{e}$, is invariant under $\mathbf{M}$ and that the restriction of $\mathbf{M}$ to this subspace is invertible.

Now we would like to prove the most important property of the generalized sigmoid mapping, and that is the convexity of its "potential" function $\mathcal{P}$. The convexity of $\mathcal{P}$ will play both a theoretical role when we study the dynamics of the analog WTA network and a practical role in that it will insure that any electrical n-port built to implement the generalized sigmoid function will be locally passive.

In order to make the proof as self contained as possible, we recall the basic definition of convex functions and an equivalent characterization of these functions in terms of arbitrary convex combinations of points.

Definition 7 *A function $\mathcal{F} : \Re^n \to R$ is said to be convex if, for all $\mathbf{z}_1, \mathbf{z}_2 \in \Re^n$ and $0 < \alpha < 1$*

$$\mathcal{F}(\alpha\mathbf{z}_1 + (1-\alpha)\mathbf{z}_2) \leq \alpha\mathcal{F}(\mathbf{z}_1) + (1-\alpha)\mathcal{F}(\mathbf{z}_2) \tag{5.28}$$

The following proposition gives an equivalent characterization of convexity, and our experience has confirmed its extreme usefulness in deriving meaningful inequalities.

Proposition 19 *A function $\mathcal{F} : \Re^n \rightarrow R$ is said to be convex if and only if, for all $\mathbf{z}_1, \mathbf{z}_2, \ldots \mathbf{z}_m \in \Re^n$ and all nonnegative numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$ with $\sum_{l=1}^m \alpha_l = 1$*

$$\mathcal{F}\left(\sum_{l=1}^m \alpha_l \mathbf{z}_l\right) \leq \sum_{l=1}^m \alpha_l \mathcal{F}(\mathbf{z}_l) \tag{5.29}$$

Proof: To show that (5.29) is sufficient take $m = 2$ which gives the convexity definition. To show that it is necessary, proceed by induction on m, the number of points. A detailed proof is given in ([64], p. 83). ■

When the function $\mathcal{F}$ is differentiable, its gradient can be used to characterize its convexity. The next proposition will be useful when we study the dynamics of the analog WTA system. Its proof can be found in ([64], p. 85).

Proposition 20 *Suppose that $\mathcal{F} : \Re^n \rightarrow R$ is a differentiable function over $\Re^n$, and denote its gradient map by $\mathcal{F}'$. Then $\mathcal{F}$ is convex if and only if, for all $\mathbf{z}_1, \mathbf{z}_2 \in \Re^n$*

$$\mathcal{F}(\mathbf{z}_2) - \mathcal{F}(\mathbf{z}_1) \geq \mathcal{F}'(\mathbf{z}_1)(\mathbf{z}_2 - \mathbf{z}_1). \tag{5.30}$$

When $\mathcal{F}$ is twice differentiable, we can use its Hessian $\mathcal{F}''$ to characterize its convexity. Again, see ([64], p. 87).

Proposition 21 *Suppose that $\mathcal{F} : \Re^n \rightarrow R$ is a twice differentiable function over $\Re^n$, with a Hessian matrix denoted by $\mathcal{F}''$. Then $\mathcal{F}$ is convex if and only if $\mathcal{F}''(\mathbf{z})$ is positive semidefinite for all $\mathbf{z} \in \Re^n$.*

The following lemma is based on the above characterization and will be instrumental in proving the convexity of $\mathcal{P}$.

Lemma 11 *Let $z_1, z_2, \ldots, z_m$ be a family of real numbers, and let $\alpha_1, \alpha_2, \ldots, \alpha_m$ a family of nonnegative real numbers with $\sum_{l=1}^m \alpha_l = 1$. Then we have*

$$\left(\sum_{l=1}^m \alpha_l z_l\right)^2 \leq \sum_{l=1}^m \alpha_l z_l^2 \tag{5.31}$$

Proof: The function $\mathcal{F}(z) = z^2$ is a convex function on $\Re$ by Proposition 21. Inequality 5.31 results easily from applying Proposition 19 to the square function $\mathcal{F}$. ■

Last in this subsection but not least, we state the following

Proposition 22 *The function $\mathcal{P}$, of which the generalized sigmoid mapping is a gradient map, is convex on $\Re^n$.*

Proof: To prove this, we use Proposition 21. Let $\mathbf{z}$ be an arbitrary point in $\Re^n$. The Hessian of $\mathcal{P}$ at $\mathbf{z}$ is nothing but the Jacobian of the generalized sigmoid function $\mathbf{F}, D\mathbf{F}(\mathbf{z})$ whose expression is given in Proposition 18. With the notation used in the proof of Proposition 18, this expression is

$$\mathbf{M} = diag(f_j) - \mathbf{f}\mathbf{f}^T.$$

We want to prove that this matrix is positive semidefinite, i.e.,

$$\zeta^T \mathbf{M} \zeta \geq 0, \; \forall \zeta \in \Re^n.$$

But we have

$$\zeta^T \mathbf{M} \zeta = \sum_{l=1}^{n} f_l \zeta_l^2 - \left(\sum_{l=1}^{n} f_l \zeta_l \right)^2.$$

Since the real variables $f_l, 1 \leq l \leq n$, are nonnegative and satisfy $\sum_{l=1}^{n} f_l = 1$, we can apply Lemma 11 to conclude that

$$\zeta^T \mathbf{M} \zeta \geq 0.$$

■

The convexity of $\mathcal{P}$ was also independently observed in [44] where it played an important role in the analysis of mean-field algorithms for solving the linear assignment problem.

Finally, it should be noted that if $\mathbf{F}$ is the gradient map of $\mathcal{P}$ then $\mathbf{F}_T$ is the gradient map of $T\mathcal{P}_T$ where $\mathcal{P}_T(.) = \mathcal{P}(\frac{1}{T}.)$.

These properties of the generalized sigmoid mapping will be used in the next subsection devoted to the study of the global dynamics of the analog, synchronous WTA network (5.24).

5.3.3 Global Dynamics of the Analog Synchronous WTA Network

In this subsection, we concentrate on studying the global dynamics of the synchronous, analog WTA network. The update equations of this network were derived from the discrete WTA network in Subsection 5.3.1. They are re-written below for ease of reference

$$\begin{aligned} \mathbf{v}_s(k+1) &= \mathbf{F}_T(\mathbf{w}_s(k)) \\ \mathbf{w}_s(k) &= \sum_{r \in \mathcal{S}} \mathbf{J}_{sr} \mathbf{v}_r(k), \end{aligned} \tag{5.32}$$

The state space of this process is the compact, convex set $\mathbf{K} = (\mathcal{T}_n)^{|\mathcal{S}|}$. We have also written this iteration in the more compact form

$$\mathbf{v}(k+1) = \mathcal{K}_T(\mathbf{v}(n)),\ \mathbf{v}(0) \in \mathbf{K} \tag{5.33}$$

where $\mathcal{K}_T$ is the nonlinear mapping induced on $\mathbf{K}$ by $\mathbf{F}_T$ and the connectivity matrices $[\mathbf{J}_{sr}]_{s,r\in\mathcal{S}}$. By Brouwer's theorem (see 3), this dynamical system has at least one equilibrium point. This fact is interesting but is too general to be useful for this particular system. Since all the trajectories are bounded, the long-term behavior of the System 5.33 is revealed by the structure of its ω-limit sets [46], Proposition 1.10. Our major result here is that the ω-limit sets of (5.33) do not contain limit cycles of period greater than 2. Therefore, as T is varied, the trajectories cannot bifurcate to chaos via period-doubling [60].

We recall here the basic definition needed for the rest of this subsection.

Definition 8 *A function $E : \mathbf{K} \rightarrow \Re$ is said to be a Lyapunov function for the discrete dynamical system*

$$\mathbf{v}(k+1) = \mathcal{K}\mathbf{v}(k),\ \mathbf{v}(0) \in \mathbf{K}$$

if it is continuous on $\mathbf{K}$ and if the change

$$\Delta E \triangleq E(\mathcal{K}\mathbf{v}) - E(\mathbf{v}) \tag{5.34}$$

is nonpositive $\forall \mathbf{v} \in \mathbf{K}$.

This definition differs from that given in Definition 6 in that the Lyapunov function is required to be continuous.

We have shown in Chapter 3 that the fixed-point equations (5.22) can also be obtained as the extrema of an effective energy that is usually derived by transforming the partition function sum into an integral. However, this effective energy is not guaranteed to be a Lyapunov function along the trajectories of (5.33). But because $\mathbf{F}$, the generalized sigmoid mapping, is the gradient map of a convex "potential", we can guarantee the following

Proposition 23 *With*

$$\mathbf{w}_s = \sum_{r\in\mathcal{S}} \mathbf{J}_{sr}\mathbf{v}_r,\ \ \forall s \in \mathcal{S}, \tag{5.35}$$

and $\mathcal{P}_T$ being the gradient map of $\mathbf{F}_T$, the function $E : \mathbf{J}(\mathbf{K}) \rightarrow \Re$ defined by

$$E(\mathbf{w}) = \sum_{s\in\mathcal{S}} \mathbf{w}_s^T \mathbf{F}_T(\mathbf{w}_s) - T\sum_{s\in\mathcal{S}} \left[\mathcal{P}_T(\mathbf{w}_s) + \mathcal{P}_T\left(\sum_{r\in\mathcal{S}} \mathbf{J}_{sr}\mathbf{F}_T(\mathbf{w}_r)\right)\right], \tag{5.36}$$

is a Lyapunov function along the trajectories of the analog synchronous WTA system (5.32).

Proof: First, it is clear that E is continuous on $\mathbf{K}$. Now, let us compute $E(\mathbf{w})$ along the trajectories of (5.32). At the k-th iterate $E_k = E(\mathbf{w}(k))$, we have

$$E_k = \sum_{s,r \in \mathcal{S}} \mathbf{v}_s^T(k) \mathbf{J}_{sr} \mathbf{v}_r(k+1) - T \sum_{s \in \mathcal{S}} \left(\mathcal{P}_T(\mathbf{w}_s(k)) + \mathcal{P}_T(\mathbf{w}_s(k+1)) \right).$$

Easy algebra shows that the change along the trajectories is successively given by

$$\begin{aligned}
\Delta E_k &= \sum_{s,r \in \mathcal{S}} \left(\mathbf{v}_s(k+2) - \mathbf{v}_s(k) \right)^T \mathbf{J}_{sr} \mathbf{v}_r(k+1) - T \sum_{s \in \mathcal{S}} \left[\mathcal{P}_T(\mathbf{w}_s(k+2)) - \mathcal{P}_T(\mathbf{w}_s(k)) \right] \\
&= \sum_{s \in \mathcal{S}} \left[\mathbf{v}_s^T(k+1)(\mathbf{w}_s(k+2) - \mathbf{w}_s(k)) - T \left(\mathcal{P}_T(\mathbf{w}_s(k+2)) - \mathcal{P}_T(\mathbf{w}_s(k)) \right) \right] \\
&= \sum_{s \in \mathcal{S}} \left[\mathbf{F}_T(\mathbf{w}_s(k))^T (\mathbf{w}_s(k+2) - \mathbf{w}_s(k)) - T \left(\mathcal{P}_T(\mathbf{w}_s(k+2)) - \mathcal{P}_T(\mathbf{w}_s(k)) \right) \right].
\end{aligned}$$

Now in order to prove that ΔE_k is nonpositive, it is sufficient to prove that every term inside the above summation is nonpositive. Since by Proposition 22, $T\mathcal{P}_T$ is convex with gradient map $\mathbf{F}_T$, we have according to Proposition 20 for all $s \in \mathcal{S}$

$$T \left(\mathcal{P}_T(\mathbf{w}_s(k+2)) - \mathcal{P}_T(\mathbf{w}_s(k)) \right) \geq \mathbf{F}_T(\mathbf{w}_s(k))^T (\mathbf{w}_s(k+2) - \mathbf{w}_s(k)).$$

Therefore $\Delta E_k \leq 0$, and the sequence E_k is nonincreasing along the trajectories of the (5.32). ∎

Note that since this Lyapunov function is continuous on the compact set $\mathbf{K}$, it reaches its minimum $E_{\min}$. Therefore, there is $\mathbf{v}(0) \in \mathbf{K}$ such that the sequence $E_k = E(\mathbf{v}(k))$ converges to $E_{\min}$. Moreover, for every initial condition $\mathbf{v}(0) \in \mathbf{K}$, there is a point $\mathbf{v}^*$ such that $E_k \geq E^* = E(\mathbf{v}^*)$ and $E_k \to E^*$.

The Lyapunov function proposed above is again inspired by the one that Marcus and Westervelt [56] used to study the convergence and stability properties of the analog Hopfield network. The most interesting aspect of this Lyapunov function is that "it looks one-step-forward," thus producing a Lyapunov sequence that depends both on $\mathbf{v}_s(k)$ and its update $\mathbf{v}_s(k+1)$. Using the effective energy (see, e.g., (3.40))

$$E_{eff}(\mathbf{v}) = \frac{1}{2} \mathbf{v}^T \mathbf{J} \mathbf{v} - T \sum_{s \in \mathcal{S}} \mathcal{P}_T \left(\sum_{r \in \mathcal{S}} \mathbf{J}_{sr} \mathbf{v}_r \right)$$

as a Lyapunov function candidate produces a sequence that depends only on the current state $\mathbf{v}_s(k)$, and the effective energy change along the trajectory could be either positive or negative.

We are now in a position to make precise conclusions about the ω-limit sets of the analog, synchronous WTA networks. First let us recall the definition of ω limit sets and state a discrete version of LaSalle's invariance principle ([46], p. 9) which basically tells us where in the state space to look for them when we have a Lyapunov function handy.

Definition 9 *A point* $\mathbf{p} \in \mathbf{K}$ *is said to be an* ω*-limit point of the dynamical system*

$$\mathbf{v}(k+1) = \mathcal{K}(\mathbf{v}(k)) \tag{5.37}$$

defined on $\mathbf{K}$ *if there is an initial condition* $\mathbf{v}(0)$ *and a sequence* $(k_i)_{i\geq 0}$ *such that the sequence* $\mathbf{v}(k_i)$ *converges to* $\mathbf{p}$. *The* ω*-limit set of the dynamical system is the set of all* ω*-limit points.*

Theorem 4 *[***LaSalle's Invariance Principle [46]***] If* E *is a Lyapunov function for (5.37) on* $\mathbf{K}$ *then every* ω*-limit set of (5.37) is contained in the (LaSalle) set*

$$\mathcal{L} = \{\mathbf{v} \in \mathbf{K} | \Delta E = E(\mathcal{K}\mathbf{v}) - E(\mathbf{v}) = 0\}.$$

Our statement of LaSalle's invariance principle differs somewhat from the one given in Proposition 2.6 in LaSalle's monograph (op. cit.) in that the trajectories of the analog WTA system are always bounded. Moreover, our statement is about the ω-limit sets of the system rather than the individual trajectories.

Now we can state our major result for this subsection.

Theorem 5 *The only* ω*-limit sets of the analog, synchronous, WTA dynamical system (5.32) are the limit cycles of period less than or equal to* 2.

This theorem is a non-obvious extension to the class of analog, synchronous, WTA networks of the result that Marcus and Westervelt obtained for the analog, synchronous, Hopfield network. The proof of this theorem uses the following lemma satisfied by the generalized sigmoid mapping $\mathbf{F}$ and its potential $\mathcal{P}$.

Lemma 12 *Let* $\mathbf{z}_0$ *be an arbitrary vector in* $\Re^n$. *Then* $\mathbf{z} \in \Re^n$ *satisfies the equation*

$$F_T(\mathbf{z}_0)^T(\mathbf{z} - \mathbf{z}_0) - T\left(\mathcal{P}_T(\mathbf{z}) - \mathcal{P}_T(\mathbf{z}_0)\right) = 0 \tag{5.38}$$

if and only if there is $\mu \in \Re$ *such that* $\mathbf{z} = \mathbf{z}_0 + \mu\mathbf{e}$.

Proof: "If" part. Assume

$$\exists \mu \in \Re, \text{ such that } \mathbf{z} - \mathbf{z}_0 = \mu\mathbf{e}.$$

Then because the components of the generalized sigmoid mapping sum up to one, we get

$$F_T(\mathbf{z}_0)^T(\mathbf{z} - \mathbf{z}_0) = F_T(\mathbf{z}_0)^T\mu\mathbf{e} = \mu.$$

On the other hand we can immediately see from the expression of $\mathcal{P}_T$ that

$$\mathcal{P}_T(\mathbf{z}_0 + \mu\mathbf{e}) = \frac{\mu}{T} + \mathcal{P}_T(\mathbf{z}_0).$$

Combining the last two equations, it results that (5.38) is satisfied.

"Only if" part. Assume that $\mathbf{z} \in \Re^n$ satisfies (5.38). Then we can always find $\mu \in \Re$ such that $\mathbf{z} - \mathbf{z}_0 = \mu\mathbf{e} + \mathbf{p}$, where $\mathbf{p}$ is a vector orthogonal to $\mathbf{e}$. Substituting into (5.38), we get the following equation in $\mathbf{p}$

$$F_T(\mathbf{z}_0)^T\mathbf{p} - T\left(\mathcal{P}_T(\mathbf{z}_0 + \mathbf{p}) - \mathcal{P}_T(\mathbf{z}_0)\right) = 0.$$

Using a second-order Taylor series expansion of $\mathcal{P}_T$ in the neighborhood of $\mathbf{z}_0$, we can write ([17], p. 193)

$$F_T(\mathbf{z}_0)^T\mathbf{p} - T\left(\mathcal{P}_T(\mathbf{z}_0 + \mathbf{p}) - \mathcal{P}_T(\mathbf{z}_0)\right) = -\int_0^1 dt(1-t)\mathbf{p}^T D\mathbf{F}_T(\mathbf{z}_0 + t\mathbf{p})\mathbf{p},$$

From the equation satisfied by $\mathbf{p}$, it follows that

$$\int_0^1 dt(1-t)\mathbf{p}^T D\mathbf{F}_T(\mathbf{z}_0 + t\mathbf{p})\mathbf{p} = 0.$$

Since $t \rightarrow D\mathbf{F}_T(\mathbf{z}_0 + t\mathbf{p})$ is continuous and positive semidefinite, we get that

$$\mathbf{p}^T D\mathbf{F}_T(\mathbf{z}_0 + t\mathbf{p})\mathbf{p} = 0, \ \forall t \in [0,1].$$

But we have seen in Proposition 18 that the null space of the Hessian of $\mathcal{P}_T$, i.e, the Jacobian of $\mathbf{F}_T$ is the one-dimensional subspace spanned by $\mathbf{e}$. Since $\mathbf{p}$ is orthogonal to $\mathbf{e}$, we conclude that $\mathbf{p} = \mathbf{0}$. It follows that $\mathbf{z} = \mathbf{z}_0 + \mu\mathbf{e}$. Q.E.D. ∎

Let us now prove Theorem 5.

Proof: According to LaSalle's invariance principle, it is sufficient to study the LaSalle set of the Lyapunov function given in Proposition 23. So let $\mathbf{v}(0) \in \mathbf{K}$, and compute $\Delta E_1 = E(\mathcal{K}\mathbf{v})(0)) - E(\mathbf{v}(0)) = E_1 - E_0$. Then we have

$$\Delta E_1 = \sum_{s \in \mathcal{S}} \left[\mathbf{F}_T(\mathbf{w}_s(0))^T(\mathbf{w}_s(2) - \mathbf{w}_s(0)) - T\left(\mathcal{P}_T(\mathbf{w}_s(2)) - \mathcal{P}_T(\mathbf{w}_s(0))\right)\right].$$

By convexity of $T\mathcal{P}_T$, the terms inside the summation are all nonpositive. Therefore setting ΔE_1 to zero implies that for all $s \in \mathcal{S}$

$$\mathbf{F}_T(\mathbf{w}_s(0))^T(\mathbf{w}_s(2) - \mathbf{w}_s(0)) - T\left(\mathcal{P}_T(\mathbf{w}_s(2)) - \mathcal{P}_T(\mathbf{w}_s(0))\right) = 0.$$

Applying Lemma 5.38, we conclude that for every $s \in \mathcal{S}$, $\exists \mu_s$ such that

$$\mathbf{w}_s(2) = \mathbf{w}_s(0) + \mu_s\mathbf{e}.$$

We know from Proposition 16 that

$$\mathbf{F}_T(\mathbf{z} + \alpha\mathbf{e}) = \mathbf{F}_T(\mathbf{z}).$$

Therefore, for all $s \in \mathcal{S}$,

$$\mathbf{v}_s(2) = \mathbf{v}_s(0),$$

which means that the LaSalle set $\mathcal{L}$ contains only periodic orbits of period 2 or less. ■

The above result is important because it proves that all the trajectories of the analog, synchronous WTA network converge to one of two attractors: a periodic orbit of period exactly 2 or a fixed point.

It is important to note that all these results about the global dynamics of the WTA network were obtained without making any assumption about the network connectivity matrix $\mathbf{J}$. Additional assumptions on $\mathbf{J}$ will of course buy us more. Indeed we have the following

Theorem 6 *Suppose the connectivity matrix of the WTA network positive definite* [4] *on H. Then the only ω-limit sets of the analog, synchronous WTA system (5.32) are the fixed points.*

Proof: According to Theorem 5, the only ω-limit cycles are the fixed points and the limit cycles of period 2. Suppose then that the ω-limit set is not a fixed point, and let $\mathbf{v}(0)$ and $\mathbf{v}(1)$ be its two points. Observe that we have by periodicity $\mathbf{v}(2) = \mathbf{v}(0)$. We want to prove that in fact $\mathbf{v}(1) = \mathbf{v}(0)$. Now $\mathbf{v}(1) - \mathbf{v}(0)$ is a vector in H. Since $\mathbf{J}$ is positive definite on H, it is sufficient to prove that the quantity

$$\beta = (\mathbf{v}(1) - \mathbf{v}(0))^T \mathbf{J} (\mathbf{v}(1) - \mathbf{v}(0))$$

is nonpositive. But

$$\begin{aligned} \beta &= \mathbf{v}^T(1)\mathbf{J}(\mathbf{v}(1) - \mathbf{v}(0)) - \mathbf{v}^T(0)\mathbf{J}(\mathbf{v}(1) - \mathbf{v}(0)) \\ &= \mathbf{v}^T(1)\mathbf{J}(\mathbf{v}(1) - \mathbf{v}(0)) + \mathbf{v}^T(2)\mathbf{J}(\mathbf{v}(2) - \mathbf{v}(1)), \end{aligned}$$

where in the last term, we have used the fact that $\mathbf{v}(2) = \mathbf{v}(0)$. Now

$$\begin{aligned} \mathbf{v}^T(1)\mathbf{J}(\mathbf{v}(1) - \mathbf{v}(0)) &= \sum_{s \in \mathcal{S}} \mathbf{F}_T(\mathbf{w}_s(0))^T(\mathbf{w}_s(1) - \mathbf{w}_s(0)) \\ &\leq T \sum_{s \in \mathcal{S}} [\mathcal{P}_T(\mathbf{w}_s(1)) - \mathcal{P}_T(\mathbf{w}_s(0))], \end{aligned}$$

where the last inequality is a result of the convexity of $\mathcal{P}$ and Proposition 20. Similarly, we have

$$\begin{aligned} \mathbf{v}^T(2)\mathbf{J}(\mathbf{v}(2) - \mathbf{v}(1)) &\leq T \sum_{s \in \mathcal{S}} [\mathcal{P}_T(\mathbf{w}_s(2)) - \mathcal{P}_T(\mathbf{w}_s(1))] \\ &\leq T \sum_{s \in \mathcal{S}} [\mathcal{P}_T(\mathbf{w}_s(0)) - \mathcal{P}_T(\mathbf{w}_s(1))], \end{aligned}$$

[4] Recall that we used the symbol H to denote the direct sum of $\oplus_{s \in \mathcal{S}} H_s$, where H_s is the $(n-1)$-dimensional subspace spanned by the edges of the simplex $\mathcal{T}_n$ at site s.

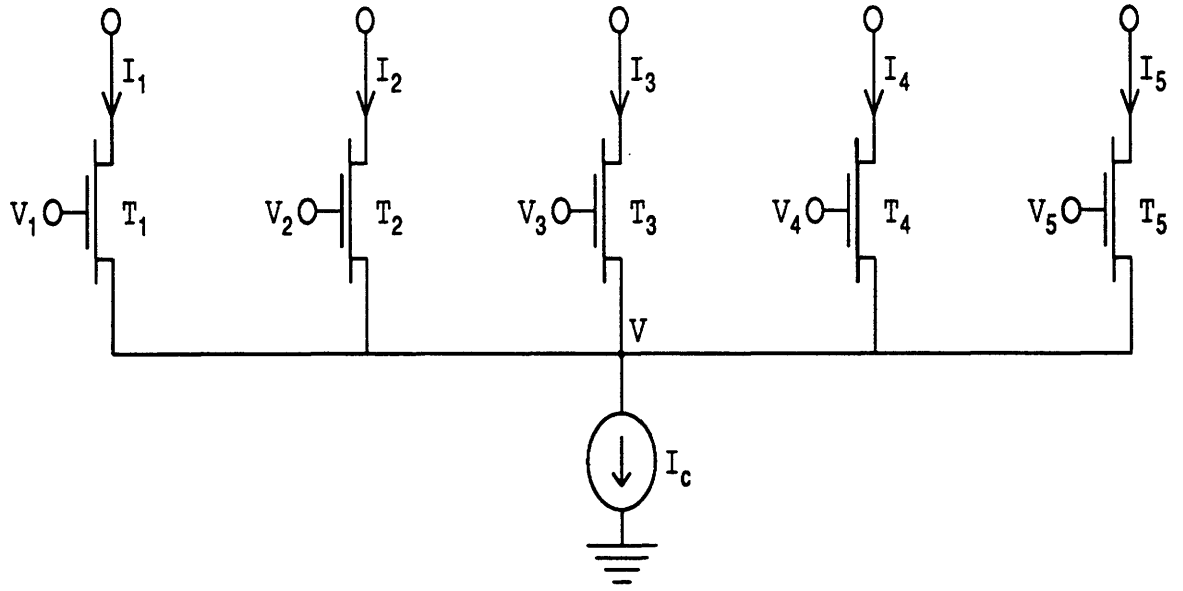


FIGURE 5-3: A circuit implementation of the generalized sigmoid mapping defined on $\Re^5$. This circuit is operated in subthreshold mode, takes the gates voltages as inputs and gives the drain currents as outputs. The control current source I_c could also be implemented as a transistor whose drain current is in saturation.

where again we have used periodicity: $\mathcal{P}_T(\mathbf{w}_s(2)) = \mathcal{P}_T(\mathbf{w}_s(0))$. Assembling all the inequalities, we conclude that

$$\begin{aligned} \beta &\leq T \sum_{s \in \mathcal{S}} [\mathcal{P}_T(\mathbf{w}_s(1)) - \mathcal{P}_T(\mathbf{w}_s(0)) + \mathcal{P}_T(\mathbf{w}_s(0)) - \mathcal{P}_T(\mathbf{w}_s(1))] \\ &\leq 0. \end{aligned}$$

Therefore $\mathbf{v}(1) = \mathbf{v}(0)$, i.e, the limit cycle is a fixed point. Q.E.D. ■

5.4 Circuit Implementation

5.4.1 Implementation of the Generalized Sigmoid Mapping

In this subsection, we propose a CMOS circuit implementation of the generalized sigmoid mapping. The circuit uses Carver Mead's analog computation paradigm [59] of using the physics of the silicon medium to perform useful computations. Thus the circuit will be operated in the subthreshold region where the drain-to-source saturation current is an increasing exponential function of the gate voltages. A schematic of the circuit is given in Figure 5-3. Our circuit is a natural generalization of the differential pair circuit used in the transconductance differential amplifier [59].

We would like to show that when the transistors are operated in the subthreshold region the drain currents $I_1, \ldots, I_n$ are the outputs of a generalized sigmoid mapping whose inputs are the gate voltages $V_1, \ldots, V_n$. One of the advantages of subthreshold operation is that the drain current saturates in few thermal potentials $V_0 = \frac{kT}{q}$ of drain-to-source voltage. Moreover, in this operation mode, the drain current is an exponential function of the gate-to-source voltage. For everyone of the transistors in Figure 5-3, we can write ([59], p.68),

$$I_m = I_0 \exp\left[(\kappa V_m - V)/V_0\right], \; 1 \leq m \leq n, \tag{5.39}$$

where I_0 and κ are process-dependent parameters. Applying KCL at the common source gives

$$\sum_{p=1}^{n} I_p = I_c. \tag{5.40}$$

Substituting for the drain currents, the KCL equation becomes

$$\sum_{p=1}^{n} I_0 \exp(\kappa V_p/V_0) \exp(-V/V_0) = I_c, \tag{5.41}$$

which gives

$$I_0 \exp(-V/V_0) = I_c \left[\sum_{p=1}^{n} \exp(\kappa V_p/V_0)\right]^{-1}. \tag{5.42}$$

Substituting back in (5.39), we get for the drain currents

$$I_m = I_c \frac{\exp(\kappa V_m/V_0)}{\sum_{m=1}^{n} \exp(\kappa V_p/V_0)}. \tag{5.43}$$

This circuit has the interesting properties of being unclocked and parallel. Moreover, the (scaled) uniqueness constraint is imposed naturally through the KCL equation and the control current source. An easy way of implementing this current source is by a transistor with a saturated drain current. From a complexity point of view, this circuit is most striking since it computes n exponentials, n ratios, and $n-1$ sums in one time constant!

Another interesting computational feature of this circuit is that it implements, in a natural way, the winner-take-all (WTA) maooing. Indeed, when one of the gate voltages, say V_m, is more positive by many V_0's than the other gate voltages, all the transistors other than T_m are turned off, and the drain current I_m is approximately equal to the control current I_c. A good, honest, analog WTA circuit should exhibit the property that when the inputs are equal the outputs are also equal. This is trivially satisfied by our circuit.

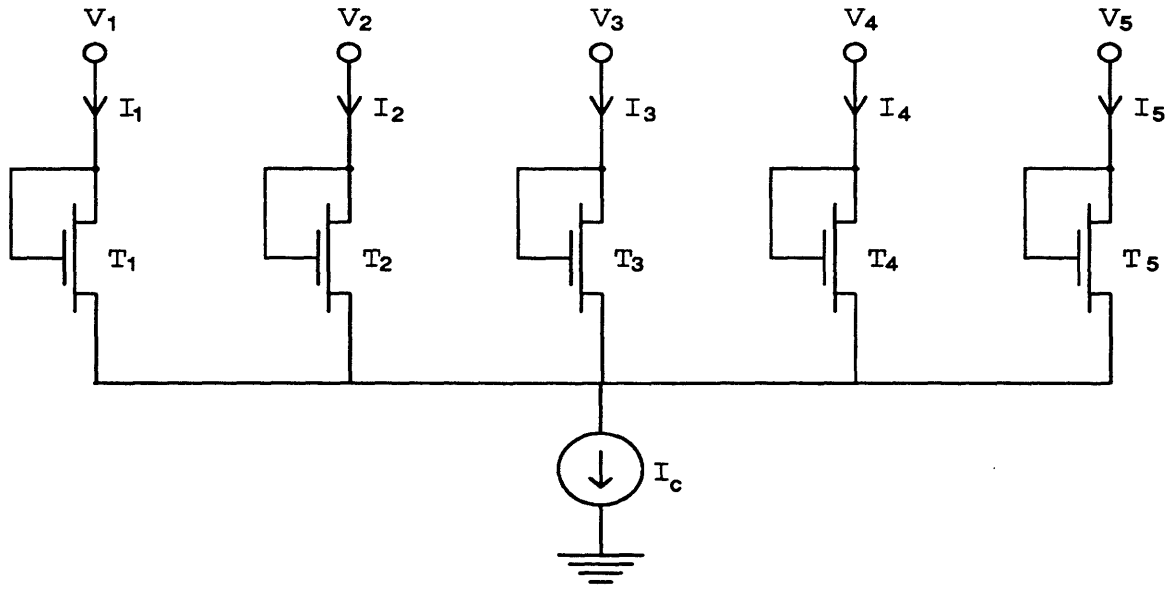


FIGURE 5-4: Modified circuit implementation of the generalized sigmoid mapping defined on $\Re^5$. In this circuit all the transistors are diode-connected, and all the drain currents are well in saturation region. Note that for every transistor, both the voltage input and the current output are on the same wire – the drain.

One disadvantage of the above circuit is that the input voltages and output currents have different wires. This makes the incorporation of the circuit as a building block in larger networks more difficult because of the additional wiring involved. One possible solution that will preserve the mapping between input voltages and output currents is to short the gate and the drain of each of the transistors and apply the input voltages to the drains. The modified circuit is shown in Figure 5-4.[5] Each transistor is said to be diode-connected ([59], p. 39). In normal operating conditions, the drain current has the same exponential dependence on the gate-to-source voltage as in the previous circuit. The resulting input/output mappings from voltages to currents is still a generalized sigmoid mapping. In the next paragraph, we concentrate on this modified circuit and show that it satisfies some desirable circuit-theoretic properties.

5.4.2 Circuit-Theoretic Properties

In this subsection, we give the circuit-theoretic properties of the generalized sigmoid mapping circuit shown in Figure 5-4. As it was mentioned in the previous section, the fundamental difference between the original circuit of Figure 5-3 and the modified circuit

[5] I would like to acknowledge helpful discussions with John Wyatt and to thank him for suggesting the diode-connected configuration shown in Figure 5-4.

of Figure 5-4 is that in the latter the input and output for each cell are on the same wire. This allows us to consider the modified circuit a multiport circuit element.

We let $\mathbf{v}$ and $\mathbf{i}$ be the n-dimensional vectors representing the input gate voltages and the output drain currents, respectively. We let $\mathcal{V}$ be the open set of $\Re^n$ of admissible input voltages. This set is determined by the conditions of normal operations, including the requirement that the transistors be operated in subthreshold mode.

The mapping $\mathbf{F} : \mathcal{V} \rightarrow \Re^n$ such that $\mathbf{v} = \mathbf{F}(\mathbf{i})$, where $\mathbf{F}$ is the generalized sigmoid mapping, represents a voltage-controlled, nonlinear, resistive n- port. We abuse notation and identify this n-port with its representation mapping $\mathbf{F}$. Its properties are given by the following proposition.[6]

Proposition 24 *The generalized sigmoid n-port $\mathbf{F}$ is reciprocal, locally passive and having a co-content function given by*

$$\Phi(\mathbf{v}) = \frac{1}{\kappa} I_c V_0 \ln \sum_{m=1}^{n} \exp(\kappa v_m / V_0) \tag{5.44}$$

Proof: A sufficient condition for reciprocity is that the voltage-controlled representation $\mathbf{F}$ be a gradient map. It can be easily checked that

$$\mathbf{F}(\mathbf{v}) = \nabla \Phi(\mathbf{v}).$$

The above equation also proves that Φ is indeed the co-content function of the nonlinear, resistive n-port $\mathbf{F}$.

To prove its local passivity, we need to show that the Jacobian of $\mathbf{F}, \mathbf{J_F}(\mathbf{v})$, is positive semidefinite for all $\mathbf{v} \in \mathcal{V}$. But this is an immediate result of the convexity of the co-content, i.e, the potential function of the generalized sigmoid mapping. ■

Thus, with this reciprocal, locally passive implementation of the generalized sigmoid mapping, we have added a new circuit element to the library of the circuit designer. Note that since this circuit element implements in an analog way the uniqueness constraint of the binary matching variables, it can be considered a *nonlinear* constraint box [33] that can be used in analog, reciprocal, locally passive networks to implement optimization algorithms.

[6]The concepts of reciprocity, passivity, content, and co-content are fundamental to nonlinear circuit theory. They are carefully developed in [90].

5.5 Generalization: Synchronous "Neural Networks" on Functional Spaces

In this short section, [7] we go beyond the usual practice of assigning binary or real numbers to sites in neural networks and assign functions from a functional space (here, we restrict ourselves to a Hilbert space.) Thus, every "neuron" becomes the site of a nonlinear, distributed parameter system communicating with the external world and with the neighboring sites. We concentrate on the parallel, synchronous, operating mode and give sufficient conditions for the state of the network to be the solution of an optimization problem.

Notation Let $\mathcal{H}$ be a real Hilbert space. Recall [13] that Riesz's representation theorem allows us to identify $\mathcal{H}$ with its dual space. Let $\mathcal{S}$ be a finite set of N sites. The *state* of each site $s \in \mathcal{S}$ is an element v_s in a convex closed bounded subset $K \subset \mathcal{H}$. The symbol $\mathcal{L}$ will denote the Banach algebra of bounded linear operators on $\mathcal{H}$. Let $\{L_{rs} \in \mathcal{L}, r, s \in \mathcal{S}\}$ be a symmetric (i.e, $L_{rs} = L_{sr}$) family of such operators. For each site s, we define

$$u_s \triangleq \sum_{r \in \mathcal{S}} L_{sr} v_r \ \in \mathcal{H}.$$

Note that u_s lies in a convex bounded subset of $\mathcal{H}$ that we denote by U_s.

A Synchronous Neural Network Let now $\{F_s, s \in \mathcal{S}\}$ be a family of nonlinear maps from $\mathcal{H}$ into $\mathcal{H}$. We assume that every F_s is a compact map on the closure of U_s, denoted $\overline{U_s}$, such that $F_s(\overline{U_s}) \subseteq K$. Define now the synchronous dynamical system on the product Hilbert space $\mathcal{H}^N$ by

$$v_s^{t+1} = F_s(u_s^t),\ v_s^0 \in K \tag{5.45}$$

$$u_s^t = \sum_{r \in \mathcal{S}} L_{sr} v_r^t. \tag{5.46}$$

We assume that the update is being done synchronously, that is, at each time instant $t+1$, all the sites are free to change their states using the information at time instant t. The synchronous mode of operation is the one to use if the different sites are to operate in parallel. This functional-analytic framework is a straightforward generalization of the discrete- or real-valued one currently encountered in the neural network literature. The

[7] This section was originally written as an independent research note and could therefore be read independently of the rest of this chapter.

dynamical system (5.45) can be written in a more compact form as

$$v^{t+1} = \mathbf{F}(u^t),\ v^0 \in K^N \tag{5.47}$$

$$u^t = \mathbf{L}v^t, \tag{5.48}$$

where $\mathbf{F} = (F_1, \ldots, F_N), v = (v_1, \ldots, v_N), u = (u_1, \ldots, u_N)$, and $\mathbf{L}$ is a symmetric $N \times N$ matrix with coefficients in $\mathcal{L}$.

Proposition 25 *The dynamical system (5.47) admits at least one equilibrium point, i.e.,* $\exists v \in K^N$ *such that* $v = \mathbf{F}(\mathbf{L}v)$.

Proof: The mapping $\mathbf{G} = \mathbf{F} \circ \mathbf{L}$ is compact on the closed bounded convex subset $K^N \subseteq \mathcal{H}^N$, and $G(K^N) \subseteq K^N$. The statement is then a straightforward application of Shauder's fixed point theorem [13]. ∎

This fixed point is not necessarily unique, and it is not clear whether the sequence $v^t, t \geq 0$ converges to any of the fixed points or any limit cycles. However, The study of the long-term behavior would be substantially facilitated if system (5.47) had a Lyapunov functional. We give below a sufficient condition for the existence of such a functional.

Theorem 7 *Suppose that every map* F_s *is the Gâteaux-derivative of a convex Gâteaux-differentiable functional* $\phi_s : \mathcal{H} \to \Re$. *Then there exists a continuous functional* $\Phi : \mathcal{H}^N \to \Re$ *such that the sequence* $\Phi_t \triangleq \Phi(v^t)$ *is nonincreasing along the trajectories of 5.47.*

Proof: Denote by $< . , . >$ the inner product in the Hilbert space $\mathcal{H}$. Define the functional

$$\Phi(u) = \sum_{s \in \mathcal{S}} < u_s, F_s(u_s) > - \sum_{s \in \mathcal{S}} (\phi_s(u_s) + \phi_s(F_s(u_s))). \tag{5.49}$$

It is clear that Φ is continuous on $\mathcal{H}$. Moreover, along the trajectories of 5.47 we have

$$\begin{aligned} \Delta\Phi &= \Phi(Fu) - \Phi(u) \\ &= \sum_{s \in \mathcal{S}} < F_s(u_s), F_s \circ F_s(u_s) - u_s > - \sum_{s \in \mathcal{S}} (\phi_s(F_s \circ F_s(u_s)) - \phi_s(u_s)). \end{aligned} \tag{5.50}$$

Since the functional ϕ_s is convex, we have the inequality

$$\phi_s(F_s \circ F_s(u_s)) - \phi_s(u_s) \geq < F_s(u_s), F_s \circ F_s(u_s) - u_s > .$$

Therefore $\Delta\Phi \leq 0$ and the sequence Φ_t is nonincreasing along the trajectories. ∎.

The above Lyapunov functional is *formally* the same as the one used in Subsection 5.3.3. The theorems stated for the global dynamics of the analog, synchronous

WTA network suggest that similar conclusions hold in this abstract setting if we assume that for every s, the mapping F_s and its "potential" ϕ_s satisfy the property

$$< F_s(u), w - u >= \phi_s(w) - \phi_s(u) \text{ iff } w = v. \tag{5.51}$$

This property can be construed as a solvability condition for an equation parameterized by u and with unknown w. A careful look at the properties of the generalized sigmoid mapping, specifically (5.38), indicates that a much weaker condition is actually sufficient. This condition is that the mappings F_s satisfy the following:

(P) For every $s \in \mathcal{S}$ and $u \in \mathcal{H}$, let $A_s(u)$ be the set of points $w \in \mathcal{H}$ such that (5.51) with parameter u and unknown w is satisfied. This set is nonempty because it contains u. Then we say that the mapping F_s satisfies the property P if it is constant over $A_s(u)$ and equal to $F_s(u)$.

Then we have the following

Proposition 26 *Suppose that every F_s satisfies the property P. Then any periodic orbit of the dynamical system (5.47) has a period of at most 2.*

Proof: Let $v^1, v^2, \ldots, v^p$ be a periodic orbit of period p. Then since the sequence $\Phi_t, 1 \leq t \leq p$ is nonincreasing around the orbit and since $\Phi_{p+1} = \Phi_1$ (by periodicity), we conclude that $\Phi_1 = \Phi_2 = \ldots = \Phi_p$, and therefore we have $\Delta\Phi = 0$ around the orbit. Since all the terms inside the summation in (5.50) are of the same sign, it follows that

$$\phi_s(F_s \circ F_s(u_s^1)) - \phi_s(u_s^1) - < F_s(u_s^1), F_s \circ F_s(u_s^1) - u_s^1 >= 0.$$

But by the assumption made on F_s and ϕ_s we get that

$$F_s^3(u_s^1)) = F_s(u_s^1)).$$

Or $v_s^4 = v_s^2$, which means that the period of the orbit is at most 2. ∎

It is important to note that the above conclusions are independent of the interaction operators $L_{rs}, r, s \in \mathcal{S}$. If we now make the additional assumption that the matrix operator $\mathbf{L}$ from $\mathcal{H}^N \rightarrow \mathcal{H}^N$ is positive definite, then using the fact that $\mathbf{L}$ defines a norm on $\mathcal{H}^N$, we can prove that the only periodic orbits of the dynamical system (5.47) are the fixed points.

Conclusion The above treatment did not include external inputs. These could be easily added to the system dynamics, and the Lyapunov functional can be easily changed to account for their presence. The conclusions about the global dynamics of the network remain however unchanged. This generalization is an attempt to look at neural networks from the perspective of large scale, hierarchical systems rather than the "small scale" perspective of biology and VLSI circuit implementation. Our model could include systems as varied as assemblies of flexible trusses in a large space structure or processors running the related waveform relaxation programs, where the "waveform" is the static load distribution for each truss. More theoretical work is needed to understand the practical implications of our assumptions on the nonlinear mappings at each site, like their compactness and monotonicity. Moreover, it is important to characterize the mappings that satisfy the additional condition stated in the preamble to Proposition 26.

5.6 Summary

In this chapter, we have studied, in some depth, the global dynamics of both the discrete and analog winner-take-all networks. The major results of this chapter are

1. For the discrete, synchronous WTA network, if the connectivity block matrix of the network satisfies a weakened form of positive definiteness, then the fixed points are the only limit cycles of the network.

2. The analog WTA network can be derived from the discrete WTA network using a very simple mean-field approximation method.

3. The dynamics of the analog WTA network is governed by the generalized sigmoid mapping that we have proven to be the gradient map of a *convex* potential.

4. The generalized sigmoid mapping admits a simple *exact* VLSI circuit implementation operating in subthreshold mode.

5. The worst that can happen in *any* synchronous, analog WTA network with a symmetric connectivity matrix is an oscillation of period 2.

6. If the connectivity block matrix of the analog WTA network satisfies a weakened form of positive definiteness, then the worst can be avoided and the fixed points become the *only* limit cycles of the network.

7. The neuron state does not have to be a number. It can be, for instance, a multidimensional signal of finite energy whose dynamics is governed by a partial differential equations. Even in this case, we can define "neural networks" whose dynamics is governed by a Lyapunov function similar to the one of the analog WTA network.

We have left out many interesting issues: learning, content addressable memories, pattern recognition, information capacity, etc. We feel that all these issues can be dealt with in the framework provided in this chapter. For instance, we can use the outer-product/Hebbian learning rule to define the connectivity matrix $\mathbf{J}$. We can use the WTA network to store and recall *gray-level* patterns the same way the Hopfield network is used to store and recall *binary* patterns. These stored patterns can be used as the basis for a pattern recognition task of grayscale images.

The passage from the discrete WTA model to the analog one resulted in the introduction of a new parameter: the temperature. The results that we have stated for the global dynamics of the analog WTA model do *not* depend on temperature. The "hot" question for the moment is then: What is the impact of temperature on the dynamics of the analog WTA model?

For an answer, the interested (and not yet tired) reader might like to have a peep at the next chapter.

6

Solution of the Mean-Field Equations

Singularity is almost invariably a clue.

Sherlock Holmes

6.1 Introduction

In this chapter, we address in some detail the temperature question posed at the end of the previous chapter. Our presentation will be in three parts. The first part will deal with the iterated-map synchronous WTA network with continuous states that we introduced in the previous chapter. We will show how the temperature can affect the nature of the network fixed point(s). In the second part, we will show how the fixed point sof the WTA network can be retrieved using continuous-time, gradient-descent methods. A nice result in this part is a generalization of the analog Hopfield dynamics [36] to the case of WTA networks. We will push the generalization further and construct Hopfield-type analog dynamics on "neural" networks in which the "neurons" have multidimensional, memoryless input/output mappings with convex potentials. Although we could have studied the temperature-dependence question in this part, we have preferred to defer it to the third part which deals with the more general case of *constrained*

WTA networks. Specifically, [1] we will deal with one of the constrained multilevel Gibbs models that were introduced in Chapter 3 and consider the fixed-point equations of the probability decomposition method, see Section 3.3.2. In this context, we will propose gradient descent algorithms rather than iterated-map algorithms for the solution of the fixed-point equations and treat in some details specific models that arise in the context of image and texture modeling. In particular, we will see how the imposition of the commonly used periodic boundary conditions will make the computation of the so-called *critical temperature* possible. In fact, we will be able to provide closed-form formulas for these critical temperatures as function of the model parameters, i.e, the connectivity block matrix $\mathbf{J} = [\mathbf{J}_{sr}]_{s,r\in\mathcal{S}}$. The approach we take is conceptually very simple but algebraically somewhat tedious. The basic idea is to study the eigenstructure of the iteration Jacobian in the synchronous analog WTA network or the eigenstructure of the effective energy Hessian in the case of the probability decomposition mean-field model.

Recently there has been some emphasis on the role of "phase transitions" in image and texture modeling using Gibbs random fields [30, 19, 69]. We will show how our analytical models can provide some insights into the phase transition phenomenon. Since we feel it is important to distinguish between the empirical notion of phase transition that we use here and the mathematical one, we give in Appendix A a short overview of the probabilistic theory of phase transitions for Gibbs systems as developed by the Russian mathematician Dobrushin in the late sixties and early seventies.

6.2 Unconstrained Multilevel Iterated Maps

Recall the equations that give the fixed-points or the mean-field solutions of the unconstrained multilevel Gibbs model (3.35)

$$\begin{aligned} \mathbf{v}_s &= \mathbf{F}(\frac{1}{T}\mathbf{w}_s) \\ \mathbf{w}_s &= \sum_{s\in\mathcal{S}} \mathbf{J}_{sr}\mathbf{v}_r. \end{aligned} \tag{6.1}$$

We know from the previous chapter that they are also the fixed-point equations of the analog WTA dynamical system without external inputs. We remind the reader that we used these equations in Chapter 4 to compute approximations of the correlation function of an unconstrained multilevel Gibbs random field. In this section, we study in more details the properties of the solutions.

[1] An early version of the third part will appear in the Journal of Mathematical Imaging and Vision as a paper co-authored with Alan Yuille under the title "Mean-Field Phase Transition and Correlation Function for Gibbs Random Fields."

6.2.1 High-Temperature Solutions

We know from Proposition 15 of the previous Chapter that as $T \to \infty$, the generalized sigmoid mapping converges to the constant mapping from $\Re^n \to \mathcal{T}_n$ that assigns to any $\mathbf{z}$ the vector $\frac{1}{n}\mathbf{e}$, where n is the number of labels or graylevels, and $\mathbf{e}$ is the vector of components $(1, 1, \ldots, 1)^T$ orthogonal to the simplex $\mathcal{T}_n$. We have therefore the following

Theorem 8 *For T high enough, the fixed-point equations have a unique fixed point $\mathbf{v}_s(T)$, $s \in \mathcal{S}$. Moreover*

$$\lim_{T \to \infty} \mathbf{v}_s(T) = \frac{1}{n}\mathbf{e}, \ \forall s \in \mathcal{S}$$

Note that because of Brouwer's fixed-point theorem, there is always at least one solution at *any* temperature. This proposition says that for high enough temperatures there is one and only one solution that converges to the trivial fixed-point $\mathbf{v}_s^{tr} = \frac{1}{n}\mathbf{e}$, $s \in \mathcal{S}$ as T approaches $+\infty$. First, let us first prove the convergence statement.

Proof: Assuming T high enough, the unique fixed-point $\mathbf{v}_s(T)$ is bounded for all $s \in \mathcal{S}$, and the argument of the generalized sigmoid mapping is given by

$$\frac{1}{T} \sum_{r \in \mathcal{S}} \mathbf{J}_{sr} \mathbf{v}_r(T)$$

and converges to the vector $\mathbf{0}$ as $T \to +\infty$, and therefore $\mathbf{v}_s(T) \to \mathbf{F}(\mathbf{0}) = \frac{1}{n}\mathbf{e}$ as $T \to +\infty$. ∎

The proof of uniqueness [2] is not as easy. It is based on finding a lower bound $T_{\min} \leq T$ such that the global mapping $\mathcal{K}_T$ (see (5.25)) is a contraction mapping on the state space $\mathbf{K} = (\mathcal{T}_n)^{|\mathcal{S}|}$. Specifically, we have the following.

Proposition 27 *The global mapping $\mathcal{K}_T$ of the WTA network (5.24) is a contraction mapping on $\mathbf{K}$ with respect to the ℓ_2 norm for*

$$T \geq T_{\min} = \rho(\mathbf{J}), \tag{6.2}$$

where $\rho(\mathbf{J})$ is the spectral radius of the matrix $\mathbf{J}$.

The proof of this proposition is based on the following lemma about the eigenvalues of the Jacobian matrix of $D\mathbf{F}(\mathbf{z})$ at an arbitrary point $\mathbf{z} \in \Re^n$.

Lemma 13 *The spectral radius $\rho(D\mathbf{F}(\mathbf{z}))$ satisfies*

$$\rho(D\mathbf{F}(\mathbf{z})) \leq \max_{1 \leq k \leq n} (F_k(\mathbf{z})), \ \forall \mathbf{z} \in \Re^n \tag{6.3}$$

[2] John Wyatt came up with the main ideas of this proof while reading an early draft of this chapter. I would like to thank him for allowing me to include it in the final thesis document.

Proof: Since the Jacobian of $D\mathbf{F}$ is symmetric positive semidefinite its spectral radius is equal to its largest eigenvalue. From Proposition 18, we have

$$D\mathbf{F}(\mathbf{z}) = diag\,(f_1, \ldots, f_n) - \mathbf{f}\mathbf{f}^T$$

where we have put $\mathbf{f} \triangleq \mathbf{F}(\mathbf{z})$ to simplify notation. Let (λ, ζ) be an (eigenvalue, eigenvector) pair of $D\mathbf{F}(\mathbf{z})$. Then we have

$$f_k\zeta_k - f_k\mathbf{f}^T\zeta = \lambda\zeta_k,\ \forall 1 \leq k \leq n.$$

We have the following two cases:

(i) $\mathbf{f}^T\zeta = 0$. Then one can show that this could happen if and only if $\mathbf{f} = \frac{1}{n}$, in which case $\lambda = \frac{1}{n}$ is an eigenvalue of multiplicity $n-1$, and (6.3) is satisfied.

(ii) $\mathbf{f}^T\zeta \neq 0$. Then the eigenvector ζ can be scaled so that $\mathbf{f}^T\zeta = 1$. It follows that

$$\zeta_k = \frac{f_k}{f_k - \lambda},\ 1 \leq k \leq n.$$

Now because of the normalization $\mathbf{f}^T\zeta = 1$, we conclude that the eigenvalues of the Jacobian matrix satisfies the algebraic equation

$$\sum_{k=1}^{n} \frac{f_k^2}{f_k - \lambda} = 1.$$

To conclude the proof of this lemma, just notice that if $\lambda \geq \max_{1\leq k\leq n} f_k$, all the terms of the above equation are negative, and therefore it cannot be satisfied. Hence,

$$\rho(D\mathbf{F}(\mathbf{z})) \leq \max_{1\leq k\leq n}(F_k(\mathbf{z})),\ \forall \mathbf{z} \in \Re^n.$$

■

Corollary 2 *The sepctral radius* $\rho(D\mathbf{F}(\mathbf{z})) < 1,\ \forall \mathbf{z} \in \Re^n$.

Proof: This results from the previous lemma and from the fact that $\max_{1\leq k\leq n}(F_k(\mathbf{z})) < 1,\ \forall \mathbf{z} \in \Re^n$. ■

Let us now go back to the proof of Proposition 27.

Proof: The mapping $\mathcal{K}_T$ is continuously differentiable on the interior set of $\mathbf{K}$. A sufficient condition for a continuously diffrentiable function to be contracting with resepct to the ℓ_2 norm is that the spectral radius of its Jacobian matrix be $\leq \gamma < 1$. Now the Jacobian of $\mathcal{K}_T$ can be obtained using the chain rule, and we have

$$D\mathcal{K}_T(\mathbf{v}) = [D_{sr}\mathcal{K}_T(\mathbf{v})]_{(s,r)\in\mathcal{S}^2}$$

with

$$D_{sr}\mathcal{K}_T(\mathbf{v}) = \frac{1}{T} D\mathbf{F}_T(\mathbf{w}_s)\mathbf{J}_{sr}.$$

In other words,

$$D\mathcal{K}_T(\mathbf{v}) = \frac{1}{T} diag(D\mathbf{F}_T(\mathbf{w}_1), \ldots D\mathbf{F}_T(\mathbf{w}_{|\mathcal{S}|}))\mathbf{J}.$$

Because of the multiplicative property of induced matrix norms,

$$\begin{aligned} \rho(D\mathcal{K}_T(\mathbf{v})) &\leq \frac{1}{T}\rho(diag(D\mathbf{F}_T(\mathbf{w}_1), \ldots DF_T(\mathbf{w}_{|\mathcal{S}|})))\rho(\mathbf{J}) \\ &\leq \frac{1}{T} \max_{1\leq s\leq|\mathcal{S}|} \rho(D\mathbf{F}_T(\mathbf{w}_s))\rho(\mathbf{J}) \\ &\leq \frac{1}{T}\rho(\mathbf{J}), \end{aligned}$$

where the last inequality results from Corollary 2. To conclude, just define $T_{\min} = \rho(\mathbf{J})$. Then if $T > T_{\min}$, the mapping $\mathcal{K}_T$ will be contracting. ∎

From this it follows that the synchronous iteration of the WTA network is the Picard iteration of a contracting map. It therefore converges to a unique fixed point.

Using the probabilistic interpretation of the generalized sigmoid mapping, the infinite-temperature solution corresponds to a configuration in which the labels have equal likelihood to appear at each site. The pattern of the network is random-looking or noisy – a state that fits the intuitive understanding of thermal agitation at high temperatures.

Because the above solution is approached as T becomes very high for *all* connectivity matrices $\mathbf{J}$, it will be called the *trivial* solution. There are however special cases where this trivial solution is a valid solution at *every* temperature. The case below is a typical one.

Proposition 28 *Assume that for every $s \in \mathcal{S}$ the matrix*

$$\mathbf{N}_s = \sum_{r\in\mathcal{S}} \mathbf{J}_{sr}$$

has $\mathbf{e}$ as an eigenvector corresponding to an eigenvalue λ_s. Then for any temperature $T > 0$, the trivial solution $\mathbf{v}_s^{tr}, s \in \mathcal{S}$, is a solution of (6.1).

Proof: When $\mathbf{v}_s = \frac{1}{n}\mathbf{e}$, the argument of $\mathbf{F}$ in (6.1) can be written as

$$\frac{1}{T}\mathbf{w}_s = \frac{1}{nT}\mathbf{N}_s\mathbf{e} = \frac{\lambda_s}{nT}\mathbf{e}.$$

Using Proposition (16) of the previous chapter, we conclude that $\mathbf{F}(\frac{\lambda_s}{nT}\mathbf{e}) = \frac{1}{n}\mathbf{e}$. It follows that the trivial solution is a fixed point of (6.1) for any temperature $T > 0$. ∎

It is clear that the point $\frac{1}{n}\mathbf{e}$ is geometrically a privileged point of the simplex $\mathcal{T}_n$. In fact, it is also a privileged solution for the fixed-point equations since every other solution

in the interior of $\mathcal{T}_n, int(\mathcal{T}_n)$, can be mapped to the trivial solution $\mathbf{v}_s^{ir}$ using a smooth (nonlinear!) transformation that maps every $\mathcal{T}_n$ onto itself.

Lemma 14 *Let* $\zeta = (\zeta_1, \ldots, \zeta_n)^T \in int(\mathcal{T}_n)$. *Then the transformation* $\mathbf{B}_\zeta : \Re^n \rightarrow \Re^n$ *defined by*

$$B_{\zeta_j}(\mathbf{z}) = \frac{\zeta_j^{-1} z_j}{\sum_{i=1}^n \zeta_i^{-1} z_i} \tag{6.4}$$

is a bijective mapping from $\mathcal{T}_n$ *onto itself such that* $\mathbf{B}_\zeta(\zeta) = \frac{1}{n}\mathbf{e}$.

Proof: First of all, note that $\mathbf{B}_\zeta$ is well defined on $\mathcal{T}_n$ since ζ has only nonzero components, $\mathbf{z} \in \mathcal{T}_n$ cannot have all its components zero, and they cannot be orthogonal. Therefore the denominator in (6.4) is always nonzero. Moreover, it is easy to see that $B_{\zeta_j}(\mathbf{z}) \geq 0,\ 1 \leq j \leq n$, and $\sum_{j=1}^n B_{\zeta_j}(\mathbf{z}) = 1$. Therefore $\mathbf{B}_\zeta(\mathcal{T}_n) \subseteq \mathcal{T}_n$. Also since the ζ_j's are nonzero $B_{\zeta_j}(\mathbf{z}_1) = B_{\zeta_j}(\mathbf{z}_2)$ implies that $z_{1j} = z_{2j},\ 1 \leq j \leq n$, i.e., $\mathbf{B}_\zeta$ is injective. Moreover, if $\mathbf{z}' \in \mathcal{T}_n$ then it is easy to check that the point $\mathbf{z}$ defined by

$$z_j = \frac{\zeta_j z_j'}{\sum_{i=1}^n \zeta_i z_i'}$$

is in $\mathcal{T}_n$ and satisfies $\mathbf{B}_\zeta(\mathbf{z}) = \mathbf{z}'$. This means that $\mathcal{B}$ is surjective on $\mathcal{T}_n$ and that the transformation

$$z_j \rightarrow \frac{\zeta_j z_j}{\sum_{i=1}^n \zeta_i z_i},\ 1 \leq j \leq n$$

is the inverse transformation. ■

We denote the inverse transformation of $\mathbf{B}_\zeta$ by $\mathbf{B}_\zeta^{-1}$. Now if $\mathbf{v}_s^* \in int(\mathcal{T}_n),\ \forall s \in \mathcal{S}$, is a solution of the fixed-point equations (6.1), then it can be mapped onto the trivial solution using the transformations $\mathbf{B}_{\mathbf{v}_s^*}, s \in \mathcal{S}$ applied on each of the equations. Note also that using the new variables

$$\mathbf{w}_s = \mathbf{B}_{\mathbf{v}_s^*}(\mathbf{v})$$

we can write a new set of fixed-point equations

$$\mathbf{w}_s = \mathbf{B}_{\mathbf{v}_s^*}\left\{\mathbf{F}\left[\frac{1}{T}\sum_{r\in\mathcal{S}} \mathbf{J}_{sr}\mathbf{B}_{\mathbf{v}_r^*}^{-1}\mathbf{w}_r\right]\right\}$$

having the trivial solution for a fixed point. This transformation can be made at any temperature. The dependence of these transformations on temperature is through the fixed point $\mathbf{v}_s^*, s \in \mathcal{S}$.

6.2.2 Stability of the High-Temperature Solution

The previous section discussed the existence and uniquness of a fixed point at high enough temperatures. Under some conditions on the connectivity matrices $\mathbf{J}$, the trivial solution obtained in the limit of infinite temperatures can be a (not necessarily unique) WTA fixed point at any temperature. We also pointed out to the possibility of using a state-space preserving nonlinear transformation to map *any* fixed point on to the trivial solution. In this section we would like to investigate the stability of the trivial solution under the synchronous analog WTA dynamical system given in (5.24) which we rewrite here for ease of reference

$$\begin{aligned} \mathbf{v}_s(k+1) &= \mathbf{F}\left(\frac{1}{T}\mathbf{w}_s(k)\right) \\ \mathbf{w}_s(k) &= \sum_{s,r\in\mathcal{S}} \mathbf{J}_{sr}\mathbf{v}_r(k). \end{aligned} \tag{6.5}$$

As function of temperature, the above system can be considered a one-parameter family of d-dimensional maps where $d = n^{|\mathcal{S}|}$ with a fixed point at $\mathbf{v}_s^{tr} = \frac{1}{n}\mathbf{e}$, $\forall s \in \mathcal{S}$. The stability of this equilibrium point as function of temperature can be studied by looking at eigenstructure of the linearized map at the equilibrium solution [87]. To get this linearized map, we need to compute the Jacobian $\mathbf{D}$ of the mapping defining the dynamical system (6.5). We have the following

Proposition 29 *The linearized map of $\mathcal{K}_T$ at a point that maps onto the trivial solution is given by the block matrix $\mathbf{D} = [\mathbf{D}_{sr}]_{s,r\in\mathcal{S}}$ where*

$$\mathbf{D}_{sr} = \frac{1}{T}\mathbf{M}\mathbf{J}_{sr} \tag{6.6}$$

is a symmetric matrix with

$$\mathbf{M} = \frac{1}{n}\mathbf{I} - \frac{1}{n^2}\mathbf{O}.$$

We encountered the matrices $\mathbf{O}$ and $\mathbf{M}$ for the first time in Chapter 4 (see (4.16)) where we had to a compute correlation-field approximation for the based on the mean-field approximation.

Proof: In order to get the (s,r) block of the Jacobian of the right-hand side in (6.5), we compute its gradient with respect to the vector $\mathbf{v}_r$. Applying the chain rule, we get

$$\frac{1}{T}D\mathbf{F}\left(\frac{1}{T}\mathbf{w}_s\right)\frac{\partial \mathbf{w}_s}{\partial \mathbf{v}_r}.$$

Using Proposition 18, the fact that we are at the trivial solution, and the linear dependence of $\mathbf{w}_s$ on $\mathbf{v}_r$, the above expression becomes

$$\frac{1}{T}\mathbf{M}\mathbf{J}_{sr},$$

which is the desired formula for $\mathbf{D}_{sr}$. This matrix is symmetric because both $\mathbf{M}$ and $\mathbf{J}_{sr}$ are symmetric. ∎

Let us make the following remarks:

1. Although $\mathbf{D}_{sr}$ depends on the matrix $\mathbf{J}_{sr}$, only the columns of the matrix $\mathbf{J}_{sr}$ that are parallel to the simplex $\mathcal{T}_n$ affect $\mathbf{D}_{sr}$. Indeed, we know that the vector $\mathbf{e}$ spans the one-dimensional null space of $\mathbf{M}$. Every column vector $\mathbf{j}$ of $\mathbf{J}_{sr}$ can be written as the direct sum of $\mathbf{j}_1 \in \mathcal{T}_n$ and $\mathbf{j}_2 \in \text{span}(\mathbf{e})$. Since $\mathbf{M}\mathbf{j}_2 = 0$, we conclude that only the parts of the column vectors parallel to $\mathcal{T}_n$ affect the matrix $\mathbf{D}_{sr}$.

2. The above conclusion is actually valid for *all* fixed points and not just for the trivial fixed point. The latter case has however the special feature that *any* vector $\mathbf{j}_2 \in \mathcal{T}_n$ is an eigenvector corresponding to the eigenvalue $\frac{1}{n}$. It follows that when the column vectors of $\mathbf{J}_{sr}$ are all orthogonal to $\mathbf{e}$, we can write

$$\mathbf{D}_{sr} = \frac{1}{nT}\mathbf{J}_{sr}. \tag{6.7}$$

   In other words, the linearized map at the trivial fixed point is directly proportional to the block connectivity matrix, and we have

$$\mathbf{D} = \frac{1}{nT}\mathbf{J}. \tag{6.8}$$

   It is easy to verify that under the above condition on the column vectors of $\mathbf{J}$, the null space of $\mathbf{J}$ and therefore $\mathbf{D}$ has dimension $|\mathcal{S}|$.

3. It follows from the above analysis that the eigenvalues of the linearized map $\mathbf{D}$ are proportional to the eigenvalues of the block connectivity matrix $\mathbf{J}$, and the coefficient of proportionality is $\frac{1}{nT}$.

4. Finally, note that the above remarks are valid whether the matrix $\mathbf{J}$ is positive semidefinite or not. In general, $\mathbf{J}$ can have positive as well as negative eigenvalues.

The local stability at the trivial solution depends on the eigenvalues of the matrix $\mathbf{D}$. If the magnitudes of these eigenvalues are smaller than 1 then the trivial solution is stable. It is clear from (6.8), that the temperature T and the number of colors or labels n play similar roles. A large number of labels, like a high temperature, will make the trivial fixed point more stable. In other words, the iteration will converge faster to the noisy pattern characteristic of thermal agitation. Now let $|\lambda|_{\max}(\mathbf{J})$ be the largest *magnitude* for the eigenvalues of $\mathbf{J}$, then we have the following

Proposition 30 *Suppose n, the number of colors, given and suppose that*

$$T > \frac{|\lambda|_{\max}(\mathbf{J})}{n}.$$

Then the trivial solution is stable.

Proof: The direct proof is easy. The result can also be deduced from the proof of Theorem 8 by noting that at the trivial solution the spectral radius of every diagonal block $D\mathbf{F}\left(\frac{1}{T}\mathbf{w}_s\right)$ is $\frac{1}{n}$. ∎

We will see in Subsection 6.4.1 that there are important special cases in which it is possible to compute the eigenstructure of the matrix $\mathbf{J}$ explicitly. We will do this computation in the context of the probability factorization mean-field equations for the constrained multilevel Gibbs model.

Since the block connectivity matrix is assumed symmetric, the Jacobian matrix $\mathbf{D}$ cannot have complex conjugate eigenvalues, and therefore a "Hopf" type bifurcation[3] from the trivial fixed point is not possible. It follows that there are only two types of bifurcation as the temperature decreases from high values where the trivial fixed point is stable to low values where it is unstable:

1. The temperature is such that the linearized map has eigenvalues equal to 1.

2. The temperature is such that the linearized map has eigenvalues equal to -1.

Note that the second situation can occur for positive temperatures only if the block connectivity matrix $\mathbf{J}$ is not positive semidefinite. The reader should recall (Theorem 5) that in this case, the analog synchronous WTA dynamics could have limit cycles of period 2. The work on understanding the bifurcation behavior of the analog WTA system in this case is still in progress. For the case when the connectivity matrix is positive semidefinite on vector space H (see Theorem 6), the only bifurcations that can occur are from a stable trivial fixed point to stable nontrivial fixed points.

6.3 Unconstrained Multilevel Gradient Descent

By far, the most common way to find solutions of the fixed-point equations uses the gradient descent method. This method is a natural one because these points are the extrema of the effective energy obtained from the saddle-point approximation. The

[3] A historically more correct name for iterated maps is "Naimark-Sacker" bifurcation. See Wiggins ([87], p. 374-381).

objective of this section is to show that there is actually more than one way to write the gradient-descent dynamical system of the unconstrained, multilevel, mean-field equations developed in Chapter 3.2.

First let us recall the effective energy given in (3.34)

$$E_{eff}(\mathbf{v}) = \frac{1}{2}\mathbf{v}^T\mathbf{J}\mathbf{v} - T\sum_{s\in\mathcal{S}}\log\left[\sum_{g\in G}\exp\left(\frac{1}{T}\sum_{r\in\mathcal{S},g'\in G}J_{sr}^{gg'}v_{rg'}\right)\right], \tag{6.9}$$

where E_{eff} is expressed in terms of the mean-field variables $\mathbf{v}$. Recall also that $J_{sr}^{gg'} = L_{sr}V(g,g')$ and $\mathbf{J}_{sr} = L_{sr}\mathbf{V}$. The gradient descent equations for these variables can be compactly written as

$$\mathbf{C}\frac{d\mathbf{v}}{dt} = -\nabla E_{eff}(\mathbf{v}) \tag{6.10}$$

where $\mathbf{C}$ is a diagonal matrix of positive "capacitances." Computing the right-hand side, we get for every $s \in \mathcal{S}$,

$$\mathbf{C}_s\frac{d\mathbf{v}_s}{dt} = -\left[\mathbf{w}_s - \sum_{r\in\mathcal{S}}\mathbf{J}_{sr}\mathbf{F}(\mathbf{w}_r/T)\right] \tag{6.11}$$

$$\mathbf{w}_s = \sum_{r\in\mathcal{S}}\mathbf{J}_{sr}\mathbf{v}_r, \tag{6.12}$$

where $\mathbf{F}$ is the generalized sigmoid mapping, and $\mathbf{C}_s$ is an invertible diagonal matrix of positive entries. We have $\mathbf{C} = diag(\mathbf{C}_1,\ldots,\mathbf{C}_{|\mathcal{S}|})$. These gradient descent equations are written in terms of the "output" variables $\mathbf{v}_s$ of the generalized sigmoid mapping. An alternative way for writing these equations is by using the "input" variables $\mathbf{w}_s$. Assuming that the block connectivity matrix $\mathbf{J}$ is invertible, we can write the effective energy function in terms of the "input" variables as

$$E_{eff}(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{J}^{-1}\mathbf{w} - T\sum_{s\in\mathcal{S}}\log\left(\sum_{g\in G}\exp[w_{sg}/T]\right), \tag{6.13}$$

Differentiating with respect to $\mathbf{w}_s$, we get the gradient-descent equations

$$\mathbf{C}_s\frac{d\mathbf{w}_s}{dt} = -\left[\sum_{r\in\mathcal{S}}\mathbf{J}_{sr}^{-1}\mathbf{w}_r - \mathbf{F}(\mathbf{w}_s/T)\right], \tag{6.14}$$

where $\mathbf{C}_s$ is again a diagonal matrix of positive entries corresponding for the site s.

Using the fact that

$$\mathbf{v}_s = \sum_{r\in\mathcal{S}}\mathbf{J}_{sr}^{-1}\mathbf{w}_r, \tag{6.15}$$

the above gradient descent equations can be written as

$$\mathbf{C}_s\frac{d\mathbf{w}_s}{dt} = -\mathbf{v}_s + \mathbf{F}(\mathbf{w}_s/T). \tag{6.16}$$

Note that under the assumption that the block connectivity matrix is invertible both sets (6.11) and (6.13) of gradient descent equations will give the *same* set of equilibrium points, i.e.,

$$\mathbf{v}_s = \mathbf{F}\left(\frac{1}{T}\sum_{r\in\mathcal{S}} \mathbf{J}_{sr}\mathbf{v}_r\right).$$

Note also that in both cases, it is the constraint

$$\mathbf{w}_s = \sum_{r\in\mathcal{S}} \mathbf{J}_{sr}\mathbf{v}_r$$

that is enforced between the mean-field variables $\mathbf{v}_s, s \in \mathcal{S}$, and the auxiliary variables $\mathbf{w}_s, s \in \mathcal{S}$ in the gradient descent dynamical systems.

Another possibility is of course to adopt the constraint

$$\mathbf{v}_s = \mathbf{F}(\mathbf{w}_s/T),\ \forall s \in \mathcal{S}, \tag{6.17}$$

and define the dynamical system

$$\mathbf{C}_s\frac{d\mathbf{w}_s}{dt} = -\mathbf{w}_s + \sum_{r\in\mathcal{S}} \mathbf{J}_{sr}\mathbf{v}_r. \tag{6.18}$$

The natural question is whether the above system minimizes some "energy" function. The answer is yes, but the surprising fact is that the Lyapunov function of this system is not the effective energy E_{eff} but its *Legendre transform*, H_{eff}, which was derived in Section 3.4. Its complete expression is given by

$$H_{eff}(\mathbf{v},\mathbf{w}) = E(\mathbf{v}) + T\sum_{s\in\mathcal{S}}\sum_{g\in G} F_g(\mathbf{w}_s/T)\log F_g(\mathbf{w}_s/T). \tag{6.19}$$

which is valid for an arbitrary energy function $E(\mathbf{v})$. Imposing the input-output constraints (6.17) has two important implications. First, the function H_{eff} becomes

$$H_{eff}(\mathbf{v}) = E(\mathbf{v}) + T\sum_{s\in\mathcal{S}}\sum_{g\in G} v_{sg}\log v_{sg}, \tag{6.20}$$

and, second, the time derivatives of the output variables v_{sg} sum up to zero, i.e,

$$\sum_{g\in G} \dot{v}_{sg} = 0,\ \forall s \in \mathcal{S}, \tag{6.21}$$

as a result of the fact that

$$\sum_{g\in G} v_{sg} = 1,\ \forall s \in \mathcal{S}.$$

Geometrically, the "velocity constraint" means that the velocity vector $\dot{\mathbf{v}}$ is parallel to the simplex $\mathcal{T}_n$. For the case when the energy function is quadratic,

$$E(\mathbf{v}) = -\frac{1}{2}\sum_{r,s\in\mathcal{S}} \mathbf{v}_s^T \mathbf{J}_{sr}\mathbf{v}_r,$$

we have

$$H_{eff}(\mathbf{v}) = -\frac{1}{2}\sum_{r,s\in\mathcal{S}} \mathbf{v}_s^T \mathbf{J}_{sr}\mathbf{v}_r + T\sum_{s\in\mathcal{S}}\sum_{g\in G} v_{sg}\log v_{sg}, \tag{6.22}$$

and the following theorem

Theorem 9 *The function H_{eff} of (6.22) is nonincreasing along the trajectories of the dynamical system (6.18).*

Proof: Compute the derivative of $H_{eff}(\mathbf{v})$ along the trajectories of (6.18). We have

$$\dot{H}_{eff}(\mathbf{v}) = -\sum_{s,r\in\mathcal{S}} \dot{\mathbf{v}}_s^T \mathbf{J}_{sr}\mathbf{v}_r + T\sum_{s\in\mathcal{S}}\sum_{g\in G}(1+\log v_{sg})\dot{v}_{sg}.$$

Now because of (6.21), the term

$$\sum_{s\in\mathcal{S}}\sum_{g\in G}\dot{v}_{sg}$$

sums up to zero. Moreover, we have

$$\log v_{sg} = \frac{w_{sg}}{T} - \log\left(\sum_{g'\in G}\exp[\frac{w_{sg'}}{T}]\right).$$

Again because of (6.21), the term

$$\sum_{s\in\mathcal{S}}\sum_{g\in G}\left(\log\sum_{g'\in G}\exp(\frac{w_{sg'}}{T})\right)\dot{v}_{sg}$$

sums up to zero, and the time derivative of H_{eff} reduces to

$$\dot{H}_{eff}(\mathbf{v}) = -\sum_{s,r\in\mathcal{S}} \dot{\mathbf{v}}_s^T \mathbf{J}_{sr}\mathbf{v}_r + \sum_{s\in\mathcal{S}}\sum_{g\in G} w_{sg}\dot{v}_{sg},$$

which can also be written as

$$\dot{H}_{eff}(\mathbf{v}) = -\sum_{s\in\mathcal{S}} \dot{\mathbf{v}}_s^T\left(\sum_{r\in\mathcal{S}}\mathbf{J}_{sr}\mathbf{v}_r - \mathbf{w}_s\right).$$

The inner term is nothing but $\mathbf{C}_s^{-1}\dot{\mathbf{w}}_s$. As for $\dot{\mathbf{v}}_s$, it is given by

$$\dot{\mathbf{v}}_s = \frac{1}{T}D\mathbf{F}(\mathbf{w}_s/T)\dot{\mathbf{w}}_s.$$

Moreover, we know that (see Proposition 22) the Jacobian of $\mathbf{F}$ is symmetric and positive semidefinite. Therefore,

$$\dot{\mathbf{w}}_s^T D\mathbf{F}(\mathbf{w}_s/T)\mathbf{C}_s^{-1}\dot{\mathbf{w}}_s \geq 0, \ \forall s \in \mathcal{S},$$

and it follows that

$$\dot{H}_{eff}(\mathbf{v}) \leq 0.$$

■

The above result can be strengthened to a *strict* inequality by noticing that the value of the generalized sigmoidal mapping at an arbitrary point depends only on the orthogonal projection of this point on the hyperplane $\mathbf{e}^T\mathbf{z}_s = 0$, see Proposition 16. Moreover, according to Proposition 18, the Jacobian of $\mathbf{F}$ is positive *definite* on this hyperplane. Abusing notations and denoting by $\mathbf{w}_s, s \in \mathcal{S}$, the projections of the auxiliary variables on the hyperplanes $\mathbf{e}^T\mathbf{z}_s = 0$, we conclude that $\dot{\mathbf{w}}_s$ lies in this hyperplane and that

$$\dot{\mathbf{w}}_s^T D\mathbf{F}(\mathbf{w}_s/T)\mathbf{C}_s^{-1}\dot{\mathbf{w}}_s > 0, \ \forall s \in \mathcal{S}.$$

Therefore

$$\dot{H}_{eff}(\mathbf{v}) < 0.$$

Corollary 3 *The function H_{eff} defined in (6.22) is a Lyapunov function of the system (6.18) that is strictly decreasing along its trajectories.*

Proof: The strict decrease of H_{eff} was shown in the paragraph preceding this corollary. In addition, H_{eff} is continuous over the compact subset $(\mathcal{T}_n)^{|\mathcal{S}|}$. Therefore, it is bounded below. H_{eff} is therefore a Lyapunov function of (6.18). ■

Three facts played an essential role in obtaining the result of theorem 9 and its strengthened version. The first is that the time derivative of each of the output variables $\dot{\mathbf{v}}_s$ is constrained to be in the hyperplane $\mathbf{e}^T\mathbf{z}_s = 0$. The second is the particular form of the generalized sigmoid mapping that allowed us to recover the auxiliary variable w_{sg} from the logarithm of the output variable v_{sg}. The third fact is that the potential function of the generalized sigmoid mapping is convex.

It will be interesting to see how essential these facts are to the final result. In other words, what mappings $\mathbf{F}$ will make system (6.18) have a function similar to H_{eff} that is nonincreaisng along its trajectories?

It turns out that the *only* requirement is that $\mathbf{F}$ be the gradient map of a convex potential. Indeed, we have the following important generalization of Theorem 9.

Theorem 10 *Let $\Phi_s, s \in \mathcal{S}$, be a family of twice-differentiable, convex functions defined on $\Re^n$ with gradient maps denoted by $\mathbf{F}_s, s \in \mathcal{S}$. Then the function*

$$H(\mathbf{v}) = -\frac{1}{2} \sum_{s,r \in \mathcal{S}} \mathbf{v}_s^T \mathbf{J}_{sr} \mathbf{v}_r + \sum_{s \in \mathcal{S}} \left[\mathbf{w}_s^T \mathbf{v}_s - \Phi_s(\mathbf{w}_s) \right] \tag{6.23}$$

where

$$\mathbf{v}_s = \mathbf{F}_s(\mathbf{w}_s)$$

is nonincreasing along the trajectories of the dynamical system (6.18).

Proof: Computing the time derivative of $H(\mathbf{v})$ along the trajectories of (6.18), we get

$$\dot{H}(\mathbf{v}) = -\sum_{s \in \mathcal{S}} \dot{\mathbf{w}}_s^T D\mathbf{F}_s(\mathbf{w}_s) \mathbf{C}_s^{-1} \dot{\mathbf{w}}_s.$$

Since the function Φ_s is convex and the diagonal matrix $\mathbf{C}_s$ is positive, the term

$$\dot{\mathbf{w}}_s^T D\mathbf{F}_s(\mathbf{w}_s) \mathbf{C}_s^{-1} \dot{\mathbf{w}}_s \geq 0,$$

which yields

$$\dot{H}(\mathbf{v}) \leq 0.$$

■

The above theorem calls the following remarks:

1. First, it is important to notice that the convex functions Φ_s, and therefore $\mathbf{F}_s$, are site-dependent. A practical case where a situation like this may arise is where the generalized sigmoid mappings are considered at different temperatures.

2. For the convex function Φ_s, the function

   $$\Phi_s^*(\mathbf{v}_s) = \max_{\mathbf{w}} \left[\mathbf{w}_s^T \mathbf{v}_s - \Phi_s(\mathbf{w}_s) \right]$$

   is also convex, and it is called the convex conjugate ([83], page 731), of Φ_s. The mapping that assigns to a (not necessarily smooth) convex function Φ its convex conjugate Φ^* as defined above is sometimes called the Legendre-Fenshel transform, *op. cit.*. Note that the value of $\mathbf{w}_s$ that maximizes the right-hand side is precisely obtained for all $\mathbf{w}_s$ such that $\mathbf{v}_s = \mathbf{F}_s(\mathbf{w}_s)$. This is exactly the condition at which the function H is computed. Therefore, we can write

   $$H(\mathbf{v}) = -\frac{1}{2} \sum_{s,r \in \mathcal{S}} \mathbf{v}_s^T \mathbf{J}_{sr} \mathbf{v}_r + \sum_{s \in \mathcal{S}} \Phi_s^*(\mathbf{v}_s). \tag{6.24}$$

3. Theorem 10 can be interpreted as a robustness theorem in the sense that as long as the mapping from the input variables $\mathbf{w}_s$ to the output variables $\mathbf{v}_s$ has a positive semidefinite Jacobian the system (6.18), will not oscillate. If in addition every mapping $\mathbf{F}_s$ has a convex compact image, then the continuous function $H(\mathbf{v})$ will be bounded below, thus becoming a Lyapunov function for (6.18).

4. It is not difficult to show [4] that for the generalized sigmoid mapping, we have

$$\Phi_s^*(\mathbf{v}_s) = T \sum_{g \in G} v_{sg} \ln v_{sg}.$$

Thus we recover the expression H_{eff} of Theorem 9 from (6.23), which was to be expected.

Finally, let us state some of the advantages of the gradient descent method in the input space $\mathbf{w} \in \Re^{n|\mathcal{S}|}$ as compared to the synchronous iteration method of the previous chapter or to the gradient descent in the output space $\mathbf{v} \in \Re^{n|\mathcal{S}|}$.

1. When the output space of the "neuron" is a compact subset of $\Re^n$ then the function H will be a Lyapunov function of the system, and the only attractors of will be the minima of H which are also the the fixed points of (6.1) or equilibrium points of (6.18). We know from the previous chapter that the synchronous iteration dynamics might have an oscillation of period 2, and that this oscillation can be ruled out by imposing additional constraints on the block connectivity matrix $\mathbf{J}$.

2. The Lyapunov function H is nonincreasing along the trajectories of (6.18) independently of the connectivity matrices $\mathbf{J}_{sr}$. The only requirement that needs to be imposed is a symmetry assumption on each of the blocks $\mathbf{J}_{sr}$ and on $\mathbf{J}$. Note that this was not the case for the gradient system (6.14) where we had to assume that the matrix $\mathbf{J}$ was invertible.

3. When we do descent in the output space, there is no guarantee that the outputs remain constrained to the simplex $\mathcal{T}_n$. This constraint is satisfied only at equilibrium. However, we will see in the next section that it is possible to change the gradient descent equations so as to enforce the output constraint during descent.

Under a condition similar to that of Proposition 28, we can show that the trivial fixed point $v_{sg} = \frac{1}{n}, g \in G, s \in \mathcal{S}$ is always a an equilibrium point for the dynamical system

[4] See, for instance, the proof given below in a circuit-theoretic context.

(6.18) at which $\mathbf{w}_s = \mathbf{0}$. We can also analyze the local stability of this equilibrium point by computing the Jacobian of the right-hand side of (6.18). Note that the condition for local stability is that all the eigenvalues of the Jacobian matrix be less than 0. The sufficient condition for stability that we obtain is similar to that of Proposition 30.

Circuit-theoretic remark: We have shown in Chapter 5 that the generalized sigmoid mapping has an exact circuit implementation using diode-connected MOSFET's operated in subthreshold mode, see Section 5.4. This implementation is a nonlinear, voltage-controlled, resistive n-port with a co-content function given by (Proposition 24)

$$\Phi(\mathbf{v}) = \frac{1}{\kappa} I_c V_0 \ln \sum_{m=1}^{n} \exp(\kappa v_m / V_0). \tag{6.25}$$

(It should be clear to the reader that in the context of this paragraph, the variables $v_m, 1 \leq m \leq n$, denote the input voltages of the resistive n-port.)

A natural question that arises from the remarks given above is the circuit-theoretic implications of applying the Legendre-Fenchel transform to the co-content function Φ. For the current vector $\mathbf{i}$, let

$$\Phi^*(\mathbf{i}) = \max_{\mathbf{v}} \left(\mathbf{i}^T \mathbf{v} - \Phi(\mathbf{v}) \right) \tag{6.26}$$

be the convex conjugate of Φ. Note that the function $\mathbf{v} \rightarrow \mathbf{i}^T \mathbf{v} - \Phi(\mathbf{v})$ is concave, and therefore $\Phi^*(\mathbf{i})$ is well defined but can be unbounded for some values of $\mathbf{i}$. However, if we restrict the voltages to be in some compact region of $\Re^n$, as is the usual case, then Φ^* will be bounded. The following proposition is the main result of this paragraph

Proposition 31 *The convex conjugate Φ^* defined in (6.26) is the content of a nonlinear, current-controlled, resistive n-port and is explicitly given by*

$$\Phi^*(\mathbf{i}) = \frac{I_c V_0}{\kappa} \sum_{m=1}^{n} \frac{i_m}{I_c} \ln \frac{i_m}{I_c}. \tag{6.27}$$

Proof: The function $\mathbf{v} \rightarrow \mathbf{i}^T \mathbf{v} - \Phi(\mathbf{v})$ it has maxima at the points where its gradient is $\mathbf{0}$, i.e,

$$\mathbf{i} = \mathbf{F}(\mathbf{v})$$

where $\mathbf{F}$ is the gradient map of Φ. From the explicit expression of $i_m, 1 \leq m \leq n$, see (5.43), we get that

$$\ln \frac{i_m}{I_c} = \frac{\kappa}{V_0} \left(v_m - \frac{1}{I_c} \Phi(\mathbf{v}) \right).$$

Therefore

$$\Phi^*(\mathbf{i}) = \sum_{m=1}^{n} \left[\frac{I_c V_0}{\kappa} \frac{i_m}{I_c} \ln \frac{i_m}{I_c} + \frac{i_m}{I_c} \Phi(\mathbf{v}) \right] - \Phi(\mathbf{v}).$$

Now notice that by KCL, we have

$$\sum_{m=1}^{n} \left[\frac{i_m}{I_c}\Phi(\mathbf{v})\right] - \Phi(\mathbf{v}) = 0.$$

Therefore

$$\Phi^*(\mathbf{i}) = \frac{I_c V_0}{\kappa} \sum_{m=1}^{n} \frac{i_m}{I_c} \ln \frac{i_m}{I_c},$$

which is the result claimed above. ∎

Implicit in the expression of Φ^* are the constraints that the currents must be non-negative and sum up to I_c. The expression of Φ^* is a strong reminder of the information-theoretic definition of entropy. We call this nonlinear, reciprocal, locally passive, resistive n-port the *entropic resistor*. The output voltage v_m is given as

$$v_m = \frac{\partial \Phi^*(\mathbf{i})}{\partial i_m} = \frac{V_0}{\kappa}\left(1 + \ln \frac{i_m}{I_c}\right). \tag{6.28}$$

The diode-connected circuit given in Figure 5-4 can be considered an implementation of the above v-i characteristics provided only $n-1$ current sources be used as inputs with KCL setting the n-th input current. It should be noted that biasing the voltages by the same amount does not affect the currents. This fact, which is apparent in (5.43) but not in (6.28), means that the latter equation does not provide the full inverse for the mapping from the voltage domain onto the current domain.

6.4 Constrained Multilevel Gradient Descent

In this section, we concentrate on the case where *global* constraints are imposed on the binary matching elements. We adopt the probability factorization point of view. We will show how in this case and under the practical assumptions usually made in image modeling, we can obtain closed form equations for the *temperatures* at which the trivial solution becomes unstable. We will also propose a gradient projection algorithm for finding the equilibrium points in terms of the *output* variables $\mathbf{v}_s, s \in \mathcal{S}$.

Recall that the intermediary step in the probability decomposition method was the approximation of the free energy F in terms of the mean-field variables. This approximation is given in (3.78). Notice that the right-hand side in this equation is similar [5] to H_{eff}, see (6.20). From the previous section, we know that H_{eff} is nonincreasing along the trajectories of the dynamical system (6.18), where the mean-field variables are

[5] The two expressions are not identical, because F is given at the mean-field values while H_{eff} is valid for all $\mathbf{v}_s, s \in \mathcal{S}$, satisfying the constraints (6.17).

obtained using the constraint equations (6.17). If the expression of the free energy F is used as the basis of a gradient descent update in the space of mean-field variables $\mathbf{v}_s, s \in \mathcal{S}$, then there is no guarantee that the velocity vector $\dot{\mathbf{v}}_s$ will remain parallel to the simplex $\mathcal{T}_n$ during the descent. This is why *two* sets of Lagrange multipliers are given in the expression of the probability decomposition effective energy (3.79)

$$\begin{aligned} E_{eff}[\mathbf{v}, \mathbf{p}, \mathbf{q}] &= E(\mathbf{v}) + T \sum_{s \in \mathcal{S}} \sum_{a \in G} v_{sa} \log v_{sa} \\ &\quad + \sum_{a \in G} p_a (\sum_{s \in \mathcal{S}} v_{sa} - \gamma) + \sum_{s \in \mathcal{S}} q_s (\sum_{a \in G} v_{sa} - 1). \end{aligned} \tag{6.29}$$

The first set $\mathbf{p} = (p_a, a \in G)$ corresponds to the *global* uniform histogram constraint, while the set $\mathbf{q} = (q_s, s \in \mathcal{S})$ corresponds to the uniqueness constraint at each pixel.

The mean-field solutions belong to the set of saddle points of the above expression of the effective energy. At these saddle points we have

$$\begin{aligned} \frac{\partial E_{eff}}{\partial v_{sa}} &= \frac{\partial E}{\partial v_{sa}} + T\{\log v_{sa} + 1\} + p_a + q_s = 0, \ \forall s \in \mathcal{S}, a \in G, \\ \frac{\partial E_{eff}}{\partial p_a} &= \sum_{s \in \mathcal{S}} v_{sa} - \gamma = 0, \ \forall a \in G, \\ \frac{\partial E_{eff}}{\partial q_s} &= \sum_{a \in G} v_{sa} - 1 = 0, \ \forall s \in \mathcal{S}. \end{aligned} \tag{6.30}$$

Among the first-order methods that we can use to solve these algebraic equations, we can mention the differential multiplier method [73], which uses a gradient *descent* for $E_{eff}[\mathbf{v}, \mathbf{p}, \mathbf{q}]$ with respect to $\mathbf{v}_s, \forall s \in \mathcal{S}$ and gradient *ascent* with respect to the Lagrange multipliers $\mathbf{p}$ and $\mathbf{q}$. A sufficient condition for this method to converge to a point where the constraints are satisfied *exactly* is that the Hessian of E_{eff} with respect to the mean-field variables be positive definite for all $\mathbf{v}, \mathbf{p}$ and $\mathbf{q}$, see Luenberger ([52], p. 427). Because the constraints are linear in the mean-field variables, this Hessian is identical to that of the "free energy" part of the effective energy. Note that the positive definite condition is satisfied in our case if the temperature T is high enough. For low temperatures, this method might fail in one of two ways. The first is common to all optimization methods where the objective function is non convex, and that is the convergence to a *local* rather than global minima. The second is peculiar to this method and is related to the problem constraints: the algorithm might converge to a point where the constraints are not satisfied.

An alternative first-order algorithm to the differential multiplier method is the gradient projection method also described in Luenberger ([52], p. 330). Here the descent

is also with respect to the mean-field variables but its taken along the negative of the gradient of the free energy projected on the hyperplanes of both the uniqueness and uniform histogram constraints. This projected gradient can be computed explicitly due to the simple form of the linear constraints, and a formula for the descent equation is given in the simulation subsection 6.4.3.

6.4.1 Bifurcation Temperature

In this subsection, we study the stability properties of the solutions of (6.30) in the particular case where the energy $E(\mathbf{v})$ is quadratic in the variables $v_{sa}, s \in \mathcal{S}, a \in G$, i.e.,

$$E(\mathbf{v}) = -\frac{1}{2} \sum_{s,r \in \mathcal{S}} \sum_{a,b \in G} J_{sr}^{ab} v_{sa} v_{rb},$$

where in this subsection, the coefficient of J_{sr}^{ab} is given by

$$J_{sr}^{ab} = -2 L_{sr} V_{ab}.$$

As usual, L_{sr} refers to the graph connectivity matrix, while V_{ab} refers to the color interaction coefficients. The expression of the energy becomes

$$E(\mathbf{v}) = \sum_{s,r \in \mathcal{S}} \sum_{a,b \in G} L_{sr} V_{ab} v_{sa} v_{rb}.$$

By adding the entropy term to the above expression, we obtain the free energy

$$F(\mathbf{v}) = \sum_{s,r \in \mathcal{S}} \sum_{a,b \in G} L_{sr} V_{ab} v_{sa} v_{rb} + T \sum_{s \in \mathcal{S}} \sum_{a \in G} v_{sa} \log v_{sa}.$$

Free Energy Hessian

In this section, we develop a perturbation analysis of the mean-field equations to estimate the location of transition temperatures. It is inspired by both the Gaussian perturbation analysis well known in statistical physics [53] and more recent work on the elastic net optimization approach to solving the traveling salesman problem [20].

The extrema of the effective energy obey the equations (6.30). Our first result follows directly:

Theorem 11 *If the interaction matrix L_{sr} is shift-invariant, i.e., $L_{sr} = A_{s-r}$ for some A, then $v_{sa} = 1/n$, $\forall s \in \mathcal{S}, a \in G$, is always an extremum of the effective energy.*

Proof: It is clear that $v_{sa} = 1/n, \forall s \in \mathcal{S}, a \in G$, satisfies the global constraints. Shift-invariance implies that $\sum_r L_{sr} = K$ is independent of the index s. Then setting $p_a = -2(K/n)\sum_b V_{ab} - T\{\log(1/n) + 1\}$ and $q_s = 0$ gives an extremum. ∎

We observe that shift-invariance is a very desirable property of the spatial interaction tensor. It is a homogeneity condition that ensures that interactions between different sites depend only on the relative positions of the sites. Henceforth we will assume that it is satisfied.

For high temperature the solution $v_{sa} = 1/n,\ \forall s \in \mathcal{S}, a \in G$, is the only solution of the mean-field equations. We call this point the *trivial solution* and will refer to it in the rest of this section as such. It corresponds to a situation where every site has the same value, which is the average of all the allowed colors values. All the structure is essentially averaged out. Such behavior is typical of the high temperature limit of most Gibbs distributions.

To determine the stability of the trivial solution, we must examine the Hessian of the effective energy evaluated at this solution. If the Hessian is positive definite then $v_{sa} = 1/n,\ \forall s \in \mathcal{S}, \forall a \in G$, is a stable solution. As the temperature decreases, however, a phase transition will occur. The Hessian will develop negative eigenvalues and the trivial solution will become unstable and a nonsymmetric, patterned, solution will develop. By computing the eigenvalues and eigenvectors of the Hessian, we can determine the critical temperature at which the phase transition occurs. Moreover, the form of these eigenvectors will approximately determine the form of the solutions near the critical temperature. We must take care to ensure that we only consider eigenvectors which are in directions allowed by the global constraints.

Differentiating (3.79) twice with respect to the variables $v_{sa}, s \in \mathcal{S}, a \in G$ gives as a generic coefficient for the energy Hessian at the trivial solution

$$H_{srab} = \frac{\partial^2 E}{\partial v_{sa} \partial v_{rb}} = 2L_{sr}V_{ab} + nT\delta_{sr}\delta_{ab}. \tag{6.31}$$

Note that because of the linearity in the constraints, the Lagrange multipliers do not appear in the expression of the Hessian. This Hessian can be written as the following sum of tensor products (also known in signal processing as Kronecker products [40])

$$\mathbf{H} = 2\mathbf{L} \otimes \mathbf{V} + nT\mathbf{I}_{\mathcal{S}} \otimes \mathbf{I}_G, \tag{6.32}$$

where $\mathbf{L}$ is the connectivity matrix of the image lattice, $\mathbf{V}$ is the color interaction matrix, and $\mathbf{I}_{\mathcal{S}}$ and $\mathbf{I}_G$ are the identity matrices of order $|\mathcal{S}|$ and n, respectively. In the following subsection, we give expressions for the eigenvalues and eigenvectors of the Hessian at $v_{sa} = 1/n,\ \forall s \in \mathcal{S}, a \in G$.

In the remaining of this subsection, we will assume that the lattice $|\mathcal{S}|$ is square of size $N \times N$.

Hessian Eigenstructure

The Hessian matrix given in (6.32) is of order $|\mathcal{S}|n$. Note that both $\mathbf{L}$ and $\mathbf{V}$ are symmetric therefore diagonalizable with real eigenvalues. The effect of the identity matrices of the second term in (6.32) on the eigenstructure of the Hessian is simply to shift the eigenvalues of $2\mathbf{L} \otimes \mathbf{V}$ by nT.

From the following simple linear algebra fact about tensor products [40], we can conclude that the Hessian matrix $\mathbf{H}$ is itself diagonalizable with real eigenvalues.

Fact: Let $\mathbf{x}$ be an eigenvector of $\mathbf{A}$ with an eigenvalue α and $\mathbf{y}$ an eigenvector of $\mathbf{B}$ with an eigenvalue β, then $x \otimes y$ is an eigenvector of $\mathbf{A} \otimes \mathbf{B}$ with an eigenvalue $\alpha\beta$.

In reality, much more can be said about the eigenstructure of $\mathbf{H}$, because our assumptions make the matrix $\mathbf{L}$ block circulant with all its blocks being circulant matrices [15]. Indeed, because of the shift-invariance assumption and the periodicity of the lattice, we can write

$$\mathbf{L} = Circulant(L_1, L_2, \ldots, L_N), \tag{6.33}$$

with

$$L_i = Circulant(c_{i1}, c_{i2}, \ldots, c_{iN}),\ \forall i = 1, \ldots, N, \tag{6.34}$$

where for an arbitrary N-tuple $\mathbf{A} = (a_1, a_2, \ldots, a_N)$, the notation $Circulant(a_1, a_2, \ldots, a_N)$ denotes the $N \times N$ matrix

$$\mathbf{A} = \begin{bmatrix} a_1 & a_2 & \ldots & \ldots & a_N \\ a_N & a_1 & a_2 & \ldots & a_{N-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_3 & a_4 & \ldots & \ldots & a_2 \\ a_2 & a_3 & \ldots & \ldots & a_1 \end{bmatrix}.$$

The actual values of the coefficients $c_{ij}, 1 \leq i, j \leq N$, depend on the order of the model, i.e., the neighborhood size, and its bonding parameters. The following theorem about the eigenvalues of $\mathbf{H}$ is however valid for any neighborhood size and any set of bonding parameters as long as the symmetry condition $L_{sr} = L_{rs}$ is enforced. Define $\xi_N = e^{\sqrt{-1}\frac{2\pi}{N}}$. Then we have the following

Theorem 12 *Let $\gamma_a, 1 \leq a \leq n$, be the eigenvalues of* $\mathbf{V}$*, and $\hat{c}_{pq}, 1 \leq p, q \leq N$ be the 2D discrete Fourier transform of $c_{ij}, 1 \leq i, j \leq N$ given by*

$$\hat{c}_{pq} = \sum_{i,j=1}^{N} c_{ij} \xi_N^{p(i-1)} \xi_N^{q(j-1)}. \tag{6.35}$$

Then the nN^2 eigenvalues of the Hessian matrix $\mathbf{H}$ *are given by*

$$\lambda_{pqa} = 2\hat{c}_{pq}\gamma_a + nT,\ 1 \leq p, q \leq N, 1 \leq a \leq n. \tag{6.36}$$

Proof: : The proof of this theorem is based on the tensor algebra fact mentioned above and the following lemma about block circulant matrices. ■

Lemma 15 *Let* $\mathbf{L}$ *be the block circulant matrix given by (6.33) and (6.34). Then the eigenvalues of* $\mathbf{L}$ *are exactly the 2D Fourier coefficients of the matrix $c_{ij}, 1 \leq i, j \leq N$, given by (6.35).*

Proof: For a proof, see [43]. ■

Reference [43] does not give the eigenvectors of the matrix $\mathbf{L}$, but it is not difficult to show the following

Lemma 16 *Let* $\mathbf{w}_p$ *be the vector* $\left(\xi_N^{p(k-1)}\right)_{1 \leq k \leq N}$*. Then the N^2 eigenvectors of the block circulant matrix* $\mathbf{L}$ *are given by the tensor products $\mathbf{w}_p \otimes \mathbf{w}_q,\ 1 \leq p, q \leq N$.*

Note that these eigenvectors are independent of the matrix $\mathbf{L}$ coefficients. Based on Lemma 16 and the Kronecker product fact stated above and denoting by $\mathbf{u}_a, 1 \leq a \leq n$ the eigenvectors of the matrix $\mathbf{V}$, we can now state

Theorem 13 *The nN^2 eigenvectors of the block circulant matrix* $\mathbf{L}$ *are given by the tensor products $\mathbf{w}_p \otimes \mathbf{w}_q \otimes \mathbf{u}_a,\ 1 \leq p, q \leq N, 1 \leq a \leq n$.*

The expression of the Hessian along with Theorems 12 and 13 means that we can separate the "spatial" or lattice effects from the "color" effects on the eigenstructure of the Hessian. Note that the uniqueness and uniform histogram constraints are translated into constraints on the acceptable eigenvectors along which we are allowed to perturb the mean-field values about the trivial solution. More specifically, we should choose, among the eigenvectors of $\mathbf{V}$ those that are orthogonal to $\mathbf{e}_G = (1, \ldots, 1) \in \Re^M$, and among the eigenvectors of $\mathbf{L}$, those that are orthogonal to $\mathbf{e}_S = (1, \ldots, 1) \in \Re^{N^2}$. For the former, the only eigenvector that is not orthogonal to $\mathbf{e}_S$ is the one corresponding to $p = q = N$. As for $\mathbf{V}$, the satisfaction of the constraints will depend on the specific coefficients V_{ab}.

To determine the critical temperatures we merely have to compute the temperature at which the first eigenvalue λ_{pqa} becomes negative, provided that its corresponding

eigenvector obeys the global constraints. From the form of λ_{pqa} given in Theorem 12, we see that at sufficiently high temperature all the eigenvalues will be positive and so $v_{sa} = 1/n,\ \forall s \in \mathcal{S}, a \in G$, is stable. As T decreases the details of the interactions $\{L_{sr}\}$ and $\{V_{ab}\}$ become important and negative eigenvalues may develop. In the next section, we consider two special cases and compute the phase transitions.

6.4.2 Special Cases

Two cases are of special interest because they have been widely used in the texture literature. The first is that of the Potts model [16], and the second is that of the autobinomial Markov random field model [14].

In both cases we assume a nearest neighbor interaction between the lattice sites having a horizontal bonding parameter β_h and a vertical bonding parameter β_v. The circulant blocks $L_i, 1 \leq i \leq N$ of the matrix $\mathbf{L}$ are given by

$$\begin{aligned} L_i &= 0,\ \forall i \neq 1, 2, N, \\ L_1 &= Circulant(0, \beta_h, 0, \ldots, 0, \beta_h), \\ L_2 &= Circulant(\beta_v, 0, \ldots, 0), \\ L_N &= Circulant(\beta_v, 0, \ldots, 0). \end{aligned}$$

A simple application of Lemma 15 will give the N^2 eigenvalues of the matrix $\mathbf{L}$

$$\hat{c}_{pq} = 2\left[\beta_v \cos(2\pi/N) + \beta_h \cos(2\pi q/N)\right]. \tag{6.37}$$

Note that the effects of an anisotropic lattice on the critical temperature can be easily studied using the above formula. However, to simplify the statements of our results, we will restrict our attention to the isotropic case for which $\beta_h = \beta_v = 1$. In this case and for any value of N, the eigenvalues are within the closed interval $(-4, +4)$. However, the eigenvector corresponding to the largest eigenvalue is the vector $\mathbf{e}_{\mathcal{S}}$ which is exactly perpendicular to the uniform histogram constraint hyperplane. Therefore, it has to be rejected as a valid perturbation.

The Potts Model

For this model the color interaction matrix is given by $V_{ab} = 1 - \delta_{ab}$. That is, there is a contribution to the texture energy if and only if the colors of two neighboring pixels are different. Note that in this case the color matrix is a circulant, symmetric matrix.

Therefore it is diagonalizable by a Fourier matrix, i.e., its ath eigenvector $\mathbf{u}_a$ has the components $\xi_n^{a(b-1)}, 1 \leq b \leq n$, and all the eigenvectors except one, namely $\mathbf{u}_n = \mathbf{e}_G$, satisfy the uniqueness constraint.

This color interaction matrix has only two eigenvalues: $(n-1)$ of multiplicity 1 and -1 of multiplicity $n-1$. The first corresponds to the eigenvector $\mathbf{e}_G$. Note that the second eigenvalue is independent of the number of colors.

The spatial interaction matrix $\mathbf{L}$ has eigenvalues

$$2\left[\cos(2\pi p/N) + \cos(2\pi q/N)\right].$$

The eigenvalue with $p = q = N$ has the eigenvector, $\mathbf{e}_S$, that does not satisfy the uniform histogram constraints. It follows that the allowed eigenvalues of the Hessian are

$$-4\left[\cos(2\pi p/N) + \cos(2\pi q/N)\right] + nT$$

for $(p,q) \neq (N,N)$. The trivial fixed point becomes unstable when the first of these eigenvalues reaches 0. It is straightforward to evaluate these eigenvalues to obtain the following

Theorem 14 *For the Potts spin model color interaction and the nearest-neighbor, isotropic, attractive, spatial interaction, the critical temperature is*

$$T_c = (4/n)\left[1 + \cos(2\pi/N)\right]$$

corresponding to the eigenvectors $\mathbf{v}_p \otimes \mathbf{v}_q \otimes \mathbf{u}_a$ *with* $(p,q) = (0,1) or (1,0)$ *and* $a \in \{1,\ldots,n-1\}$.

Observe that the critical temperature decreases inversely proportional to the number of colors but is relatively independent of the number of lattice sites for large N. Indeed as $N \to \infty$ the critical temperature tends to $8/n$.

The behavior of the system just below the critical temperature is given by a linear combination of the eigenvectors $\mathbf{v}_p \otimes \mathbf{v}_q \otimes \mathbf{u}_a$ with $(p,q) = (0,1)$ or $(1,0)$ and $a \in \{1,\ldots,n-1\}$.

The Autobinomial Model

For our second model we set $n = 2$ and consider the autobinomial Gibbs random field model with $(V_{11}, V_{12}, V_{21}, V_{22}) = (0,0,0,-1)$. This is closely related to the Ising spin model which, in the absence of global constraints, is exactly soluble with known critical temperature.

The eigenvectors of V_{ab} are $(1,0)$ and $(0,1)$ with corresponding eigenvalues 0 and 1 respectively. Neither of them satisfy the global constraints. Instead we must consider their combination $(1/\sqrt{2})(1,-1)$ which corresponds to projecting the unit eigenvector $(0,1)$ in the direction perpendicular to the constraints. This gives a contribution of (-1/2) to the quadratic expansion about $v_{sa} = 1/2$. Thus it effectively corresponds to an eigenvalue of $(-1/2)$.

Combining this with the results for the spatial interaction matrix gives "allowable" eigenvalues

$$-2\left[\cos(2p\pi/N) + \cos(2q\pi/N)\right] + 2T$$

for $(p,q) \neq (N,N)$. The following result is obtained

Theorem 15 *The critical temperature of the isotropic, attractive, nearest-neighbor, binary autobinomial model is given by* $T_c = 1 + \cos(2\pi/N)$.

Observe that the critical temperature is only weakly dependent on N for large N. Moreover, as $N \to \infty$ we get $T_c \to 2$.

Below the critical temperature the solution will be a linear combination of the two vectors with color part $(1,-1)$ and spatial part $\xi_N^{p(k-1)}\xi_N^{q(l-1)}$ for $(p,q) = (1,0)$ or $(0,1)$.

The result given in Theorem 15 can be generalized to an autobinomial model with n gray levels and a color interaction matrix $V_{ab} = -ab, 0 \leq a,b \leq n-1$. Indeed, using an argument similar to the one that precedes Theorem 15, we can show that

$$T_c = (\frac{n^2-1}{3})[1 + \cos(2\pi/N)].$$

This formula can alternatively be derived from Equation (4.27) which we used to obtain an approximation of the covarinace function of the autobinomial model.

The dependence of critical temperature on the square of the number of gray levels is in agreement with an empirical observation made in [69] concerning the regions of phase transition of the autobinomial model.

6.4.3 Simulations

In this section we compare our results for the Potts model and the autobinomial model with computer simulations using a novel projected gradient descent algorithm. This algorithm has the advantage of enforcing the constraints on the mean-field variables at every step of the update. Let

$$F(\mathbf{v}) = \sum_{s,r\in\mathcal{S}}\sum_{a,b\in G} L_{sr}V_{ab}v_{sa}v_{rb} + T\sum_{s\in\mathcal{S}}\sum_{a\in G} v_{sa}\log v_{sa}. \tag{6.38}$$

be the free energy. The components of of $DF(\mathbf{v})$ are given by the partial derivatives

$$\frac{\partial F}{\partial v_{sa}}.$$

We have the following

Proposition 32 *Denote by* $\nabla^P F(\mathbf{v}) = (\frac{\partial F}{\partial v_{sa}}^P)_{s,a}$ *the projection of the gradient of* $F(\mathbf{v})$ *on the uniqueness*

$$\sum_{a \in G} v_{sg} = 1, \ \forall s \in \mathcal{S},$$

and uniform histogram

$$\sum_{s \in \mathcal{S}} v_{sg} = 1, \ \forall a \in G,$$

constraint surfaces. Then we have

$$\frac{\partial F}{\partial v_{sa}}^P = \frac{\partial F}{\partial v_{sa}} - \frac{1}{N^2} \sum_r \frac{\partial F}{\partial v_{ra}} - \frac{1}{n} \sum_b \frac{\partial F}{\partial v_{sb}} + \frac{1}{nN^2} \sum_{rb} \frac{\partial F}{\partial v_{rb}}. \tag{6.39}$$

Proof: It is enough to check that

$$\sum_{s \in \mathcal{S}} \frac{\partial F}{\partial v_{sa}}^P = 0, \ \forall a \in G,$$

and that

$$\sum_{a \in G} \frac{\partial F}{\partial v_{sa}}^P = 0, \ \forall s \in \mathcal{S}.$$

For instance, for the first equality, if we sum up the right-hand side of (6.39) with respect to $s \in \mathcal{S}$, the first and the second terms will cancel each other, while the third and fourth terms will cancel each other. ∎

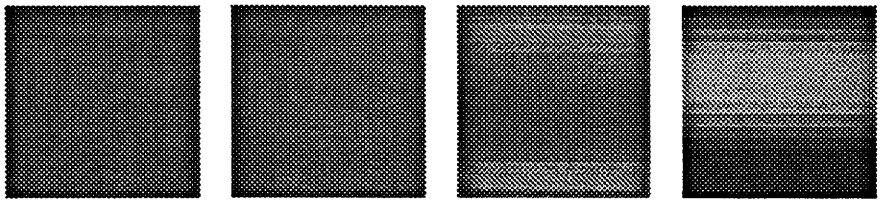
The *projected* gradient descent dynamics can now be written as

$$c\dot{v}_{sa} = -\frac{\partial F}{\partial v_{sa}}^P, \ \forall s \in \mathcal{S}, \ a \in G, \tag{6.40}$$

where c is a positive time constant. [6]

This gradient rule can be inserted in a continuation (or deterministic annealing) method that performs a projected-gradient descent while decreasing the temperature T. The intuition is that the global minima can be found at high temperature, where the free energy is convex, and can then be tracked down as the temperature decreases to the desired final temperature. This type of methods is not guaranteed to converge to the global optimum solution, but is empirically extremely successful [67].

[6] It is important that this time constant be the same for all variables v_{sa}, so that the velocities $\dot{v}_{sa}$ would be parallel to the constraint surfaces.



(a) (b) (c) (d)

FIGURE 6-1: States of the isotropic Potts mean-field model at different temperatures: (a) Initial condition; (b) T slightly higher than the critical temperature; (c) T slightly lower; (d) T lower. Note that for T slightly higher than the critical temperature, the mean-field states are attracted back to the trivial solution. The lattice size is 16×16.

The Potts Model

We consider the case with $n = 2$ colors and $N = 16$. From the previous section, Theorem 14, we find that the critical temperature is $T_c = 3.85$.

We performed simulations at temperatures $T = 3.9$, 3.75 and 3.5. The system was given the initial configuration $v_{sa} = (1/2) + n_{sa}$, $\forall s \in \mathcal{S}, a \in G$, where the $\{n_{sa}\}$ are chosen independently from the uniform distribution over the range $-0.005 \leq n_{sa} \leq 0.005$ and then projected to ensure that the $\{v_{sa}\}$ satisfy the global constraints.

To update the $\{v_{sa}\}$ we use the projected gradient, given by (6.39). In discrete time, the projected-gradient descent is given by

$$v_{sa}(t+1) = v_{sa}(t) - \zeta \frac{\partial F^P}{\partial v_{sa}}(\mathbf{v}(t)), \ \forall s \in \mathcal{S}, a \in G, \tag{6.41}$$

where the coefficient $\zeta = 1/(10 + D^P F \cdot D^P F)$. The term dependent on $D^P F$ in the denominator is used to stabilize the iterations where the gradient is large.

The experiments were run on a SUN4 with a stopping condition whenever $\left|D^P F\right| < 0.0005$. The typical number of iterations varied greatly with the temperature. For temperatures below the critical one, the number of iterations generally decreases with temperture, as we can see from Figures 6.4.3. 6.4.3 and 6.4.3(a).

The results confirm our theoretical predictions of the temperatures at which the trivial equilibrium point becomes unstable. The simulation at $T = 3.9$ rapidly converged to the state $v_{sa} = 1/2$, $\forall s \in \mathcal{S}, a \in G$. At $T = 3.75$ and $T = 3.5$ the state $v_{sa} = 1/2$, $\forall s \in \mathcal{S}, a \in G$ became unstable and the system converged to a new minimum. This is illustrated by Figures 6.4.3, 6.4.3 and 6.4.3(a).

In Figure 6.4.3, the grayscale images represent the mean-field values by pixel intensities defined as follows. Suppose the color values are specified to be $\mathbf{C} = \{C_a\}$. Then the mean-field values $\{v_{sa}\}$ will correspond to intensities $\{I_s\} = \{\sum_a v_{sa} C_a\}$. If the pattern

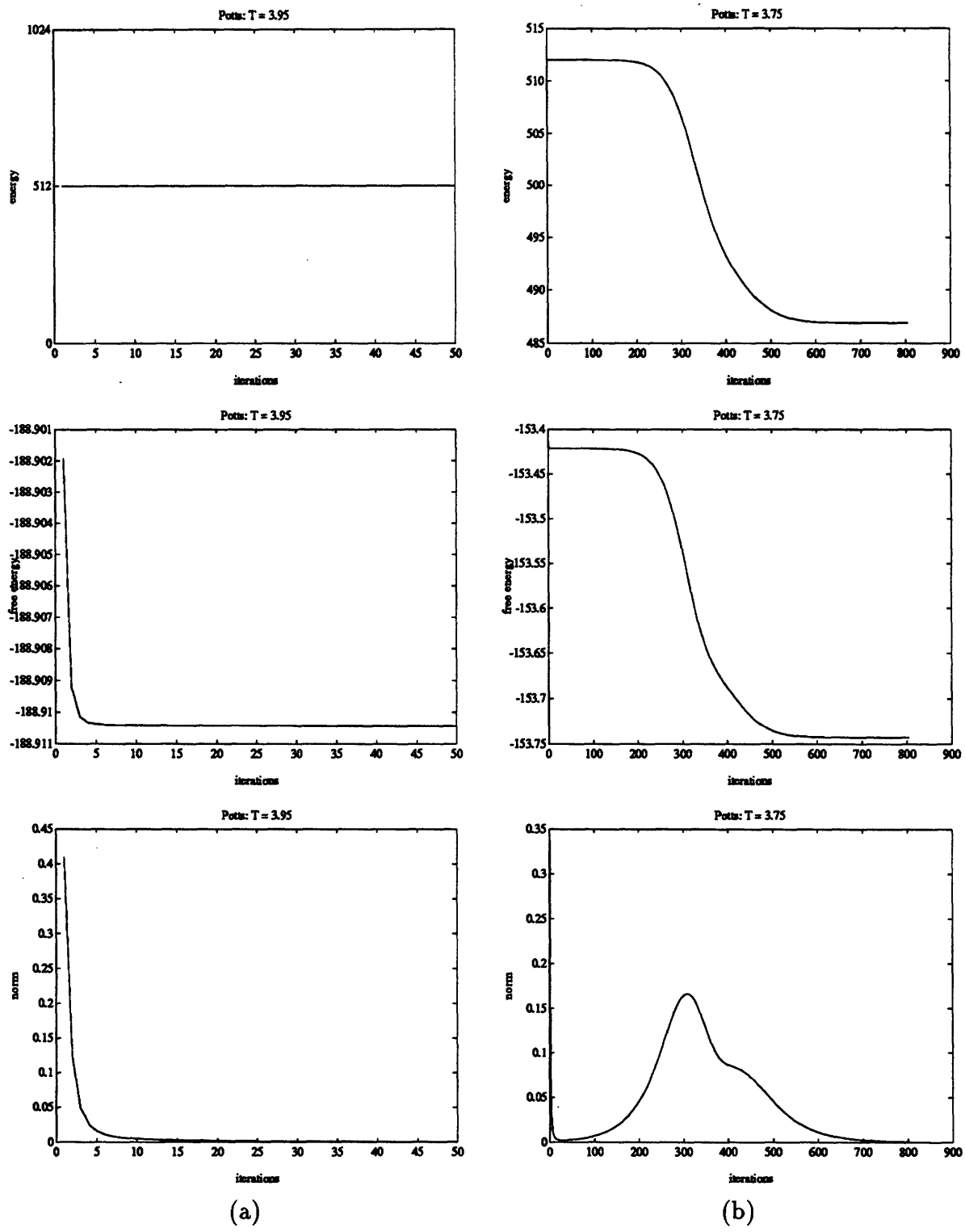


FIGURE 6-2: Potts model: Energy, free energy and gradient norm for temperatures slightly higher and slightly lower than the critical temperature. (a) $T = 3.9$; (b) $T = 3.75$.

is assumed binary with 0 corresponding to black and 255 corresponding to white, then $\mathbf{C} = \{0, 255\}$, and the intensities $\{I_s\} = \{255 w_{s1}\}$. Figure 6.4.3(a) gives the mean-field values of the initial conditions. Figures 6.4.3(b), 6.4.3(c), and 6.4.3(d) final states at $T = 3.9$, 3.75, 3.5 respectively. It is clear that the critical temperature lies between $T = 3.9$ and $T = 3.75$. Note that if we rounded off the $\{v_{sa}\}$ to the nearest integer we would obtain a purely binary black-white pattern. The patterns just above the critical temperature correspond to a random-looking binary texture, while those just below the critical temperature correspond to more ordered patterns. [7] It is in this sense that the critical temperature defines a phase transition.

In Figure 6-6(a), the theoretical prediction about the pattern below the critical temperature is illustrated. For a given column, the intensities are plotted versus the rows for both $T = 3.75$ and $T = 3.5$. The resulting curves are sine waves whose amplitudes measure the degree of saturation of the mean-field variables, i.e., how close they are to black or white. Note that this saturation increases as the inverse of temperature. The phases of these sine waves are random.

Figures 6.4.3(a) and (b), and 6.4.3(a) plot the energy, the free gradient and the norm of the free energy gradient, as a function of the number of iterations, at temperatures $T = 3.9$, $T = 3.75$ and $T = 3.5$ respectively. Figure 3(a) shows that at $T = 3.9$ the system rapidly converges to the solution $v_{sa} = 1/2$, $\forall s \in \mathcal{S}, a \in G$. Figures 6.4.3(b) and 6.4.3(a) (at $T = 3.75$ and $T = 3.5$) show that the system slowly escapes from the local minimum at $v_{sa} = 1/2$, $\forall s \in \mathcal{S}, a \in G$, and converges to a new solution. This convergence is significantly faster for $T = 3.5$.

The Autobinomial Model

Here also we consider the binary case $n = 2$, and again we choose $N = 16$. It follows from Theorem 15 that the critical temperature is $T_c = 1.92$.

We performed simulations using the same techniques described in the previous subsection. We ran the algorithm at $T = 2.1$, $T = 1.8$ and $T = 1.6$.

As predicted the trivial equilibrium state $v_{sa} = 1/2$, $\forall s \in \mathcal{S}, a \in G$, was stable for $T = 2.1$ and unstable for $T = 1.8$ and $T = 1.6$. See Figures 6.4.3, 6.4.3, 6.4.3(b) and 6-6(b).

In Figure 6.4.3, we show the mean values of the intensities as defined in the prvious paragraph. Figure 6.4.3(a) gives the mean intensities of the initial condition. Fig-

[7] In Figure 6.4.3, we give a pictorial representation of the mean-field values. The actual binary textures are obtained by thresholding the mean-field values about the value 1/2. When the mean-field values are close to but randomly distributed about 1/2 we get a random-looking pattern.

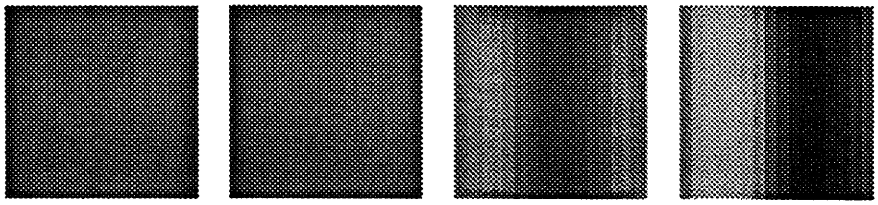


FIGURE 6-3: States of the isotropic autobinomial mean-field model at different temperatures: (a) Initial condition; (b) T slightly higher than the critical temperature; (c) T slightly lower; (d) T lower. Note that for T slightly higher than the critical temperature, the mean-field states are attracted back to the trivial solution. The lattice size is 16×16.

ures 6.4.3(b), 6.4.3(c), and 6.4.3(d) give the final states at $T = 2.1,\ 1.8,\ 1.6$, respectively. It is clear that the critical temperature lies between $T = 2.1$ and $T = 1.8$. As for the Potts model, the patterns just above the critical temperature correspond to a random-looking binary texture, while those just below the critical temperature correspond to more ordered patterns[2]. It is in this sense that the critical temperature defines a phase transition.

In Figure 6-6, the theoretical prediction about the pattern below the critical temperature is illustrated. For a given row, the intensities are plotted versus the columns for both $T = 1.8$ and $T = 1.6$. The resulting curves are sine waves whose amplitudes measure the degree of saturation of the mean-field variables, i.e., how close they are to black or white. Note that this saturation increases as the inverse of temperature. The phases of these sine waves are random.

Figures 6.4.3 and 6.4.3(b) , plot the energy, the free energy and the norm of the free energy gradient, as a function of the number of iterations, at temperatures $T = 2.1$, $T = 1.8$ and $T = 1.6$ respectively. Figure 6.4.3(a) shows that at $T = 2.1$ the system rapidly converges to the solution $v_{sa} = 1/2,\ \forall s \in \mathcal{S}, a \in G$. Figures 6.4.3(b) and 6.4.3(b) (at $T = 1.8$ and $T = 1.6$) show that the system slowly escapes from the local minimum at $v_{sa} = 1/2,\ \forall s \in \mathcal{S}, a \in G$, and converges to a new solution. This convergence is faster for $T = 1.6$.

6.4.4 Application: Phase Transition

Recent work on texture modeling using Gibbs random fields [22, 70] has shown that their descriptive "power" could be quite limited if the model parameters are not chosen appropriately. This is probably best illustrated with the role of the temperature parameter. Heuristically, we can say the following. When the temperature is high, the noise level in the Gibbs system is high, and the Gibbs distribution is close to a uniform

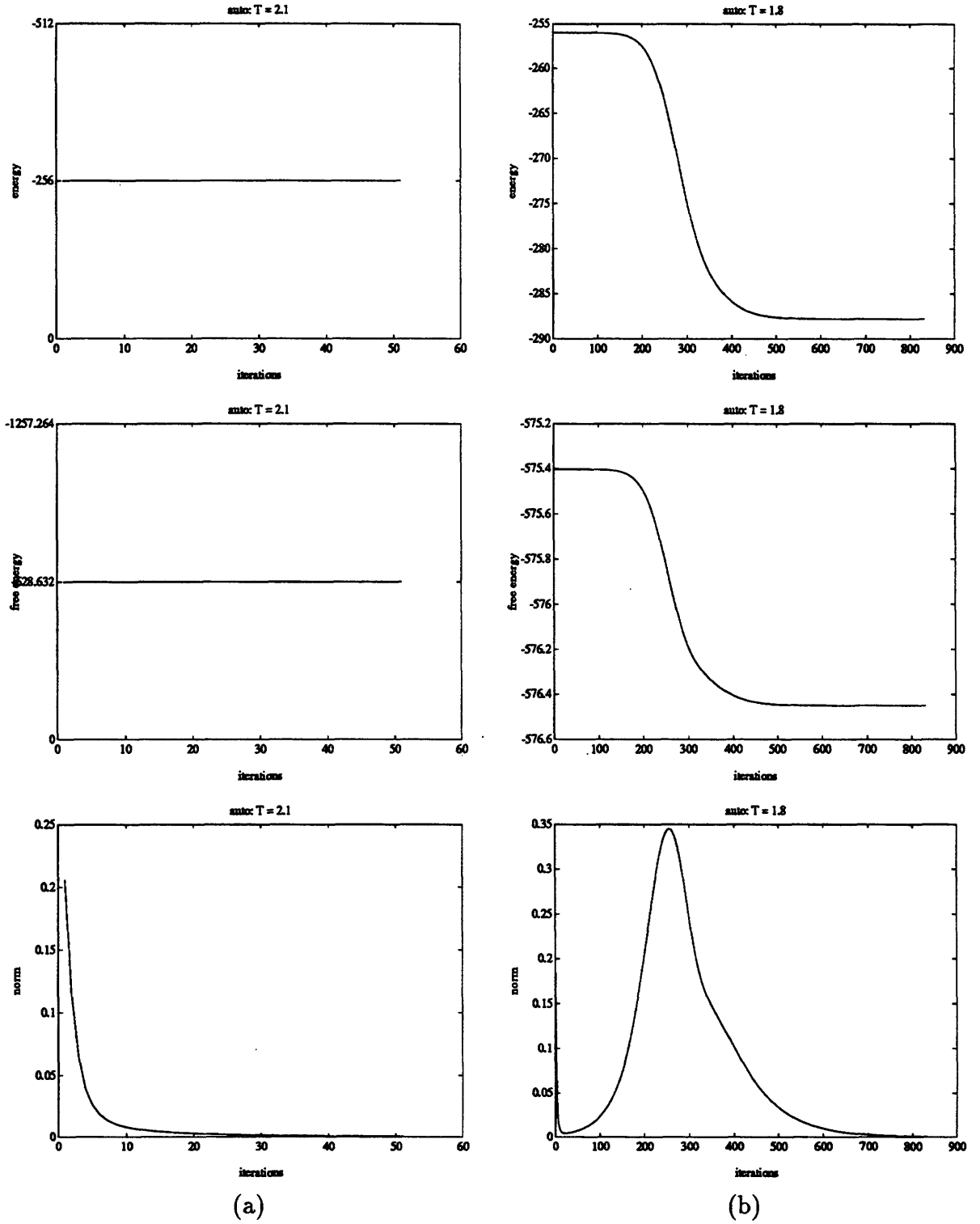


FIGURE 6-4: Autobinomial model: Energy, free energy and gradient norm for temperatures slightly higher and slightly lower than the critical temperature. (a) $T = 2.1$; (b) $T = 1.8$.

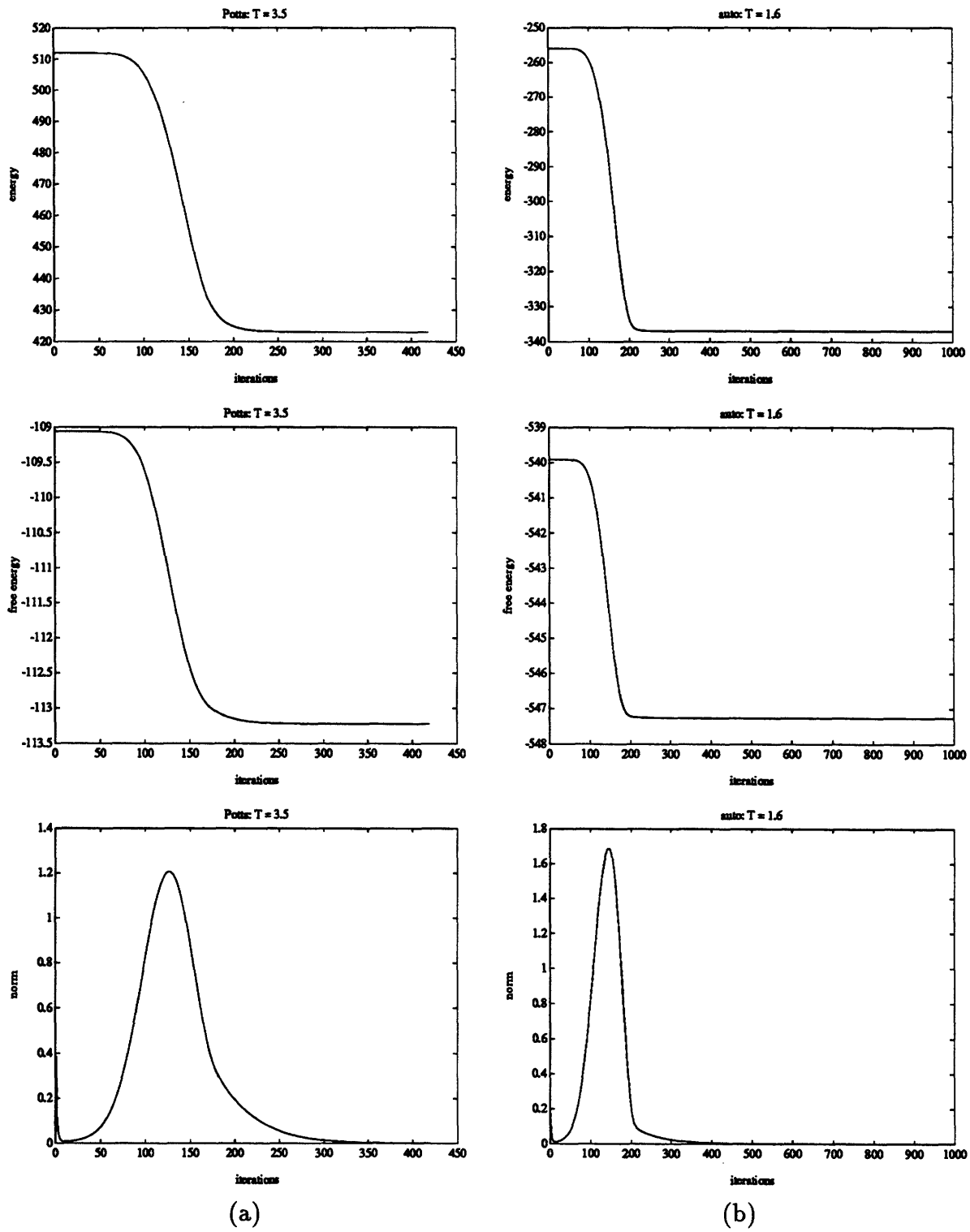
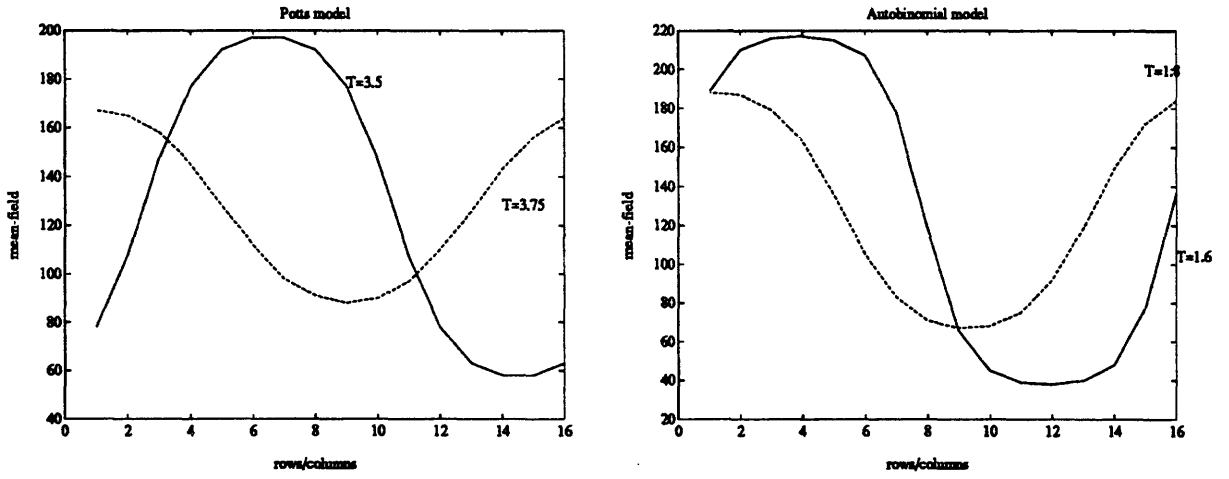


FIGURE 6-5: Energy, free energy and gradient norm for temperatures much lower than the critical temperatures. (a) Potts model: $T = 2.1$; (b) Autobinomial model $T = 1.8$. Note that the convergence to a nontrivial stable point is much faster than for temperatures slightly lower than the critical temperatures.



(a) (b)

FIGURE 6-6: The local minima for temperatures lower than the critical temperatures are approximated rather well by the eigenvectors of the energy Hessian matrix. The amplitude of the sine wave indicates the saturation of the mean-field values. It increases with inverse temperature. The phase is random. (a) Potts model. (b) Autobinomial model. The scale of the mean-feild values is 255 so as to match the grayscale figures.

distribution over the set of all possible pattern configurations. Sampling from this distribution is very likely to produce a noisy pattern without any textural properties. When the temperature of the Gibbs distribution is very low, the most probable patterns are the ground states of the Gibbs energy. In [70], it is shown that these ground state patterns have "unintertsing" textural properties. In other words, the patterns sampled from the Gibbs distribution have no textural characteristics at both ends of the temperature scale. Therefore we can argue that most of the interesting textural patterns are sampled when the temperature is "neither too low nor too high."

This intuition is borne out in the simulated annealing study of the Gibbs autobinomial model described in [71], where it has been shown experimentally that there are transition temperatures at which the pattern visual properties change noticeably. But the important problem of *predicting* the temperature range in which phase transitions are expected to occur has remained open. This temperature prediction is important for two main reasons. First, it allows to set up, *without* recourse to the time-consuming simulated annealing algorithm, upper and lower limits on the temperature scale within which "interesting" patterns are likely to get synthesized. Second, it helps in choosing the appropriate *a priori* model parameters when maximum *a posteriori* (MAP) estimates are sought for a given image processing or computer vision task.

The above two reasons have been pointed out in the recent literature on Gibbs random

fields and their siblings, Markov random fields. The importance of the first reason is mentioned in [19] where it is stated that [3]

> [...] a Markov random field exhibits a *phase transition* phenomenon when parameters are specified so that the short-term correlations among neighboring pixels develop into long-term correlations [...]. Models exhibiting *phase transition* are of limited value in image modeling because realizations tend to be dominated by one or two colors and do not exhibit fine structure or texture. [...] The presence of *phase transition* and the interpretation of model parameters have been ignored in some applications of random field models.

As for the second reason, the effect of phase transitions on the quality of MAP estimates, it is stated in [30] that

> [...] the efficacy (and visual appearance) of the MAP estimator degrades rapidly as the relative influence of the prior distribution [...] is increased with respect to the data component [...]. This phenomenon appears to be related to the existence of *phase transition* in the prior model: reconstructions may be dominated by the global properties of the prior, such as long-range order [...], which are characteristic of certain parameter ranges.

From a theoretical viewpoint, we should emphasize that the general problem of phase transitions for Gibbs random fields is a difficult one. In Appendix A, we give a quick survey of the mathematical aspects of this question and show how one can use the theoretical formulas already obtained in the literature to derive crude estimates of the temperature interval within which phase transitions are likely to occur. We also point out the fact that within this theoretical context, the mathematical problem of phase transitions for *constrained* Gibbs random fields remains unsolved.

In the mean-field context, we have shown how the change in temperature leads to a change in the equilibrium pattern from a noisy-looking pattern to a structured pattern. This type of behavior was observed in our deterministic simulations for both the Potts and autobinomial models. The main advantage of our mean-field solution is that it provides us with an *explicit* formula for the temperature at which this type of transition occurs. For instance, the critical temperature formula given for the Potts model in Theorem 14 shows that as the number of colors become higher, the critical temperature at which transition to a structured pattern occurs becomes lower. This decrease is like $\frac{1}{n}$ for the

[3] We have italicized the expression *phase transition* in these quotes.

Potts model. We can also show using the estimate of the correlation function given in (4.28) for the autobinomial model that the critical temperature of that model actually *increases* like $\frac{n^2-1}{3}$. Moreover, it is not difficult to incorporate in these formulas the anisotropy of the lattice. It is then possible to show using an eigenvector argument that if we have a nearest-neighbor, anisotropic, attractive lattice, then the structured pattern that will appear at the critical temperature will have the direction of the larger bonding parameter.

6.5 Summary

This chapter was devoted to an analysis of the effect of temperatures on the solutions of the mean-field equations of the grayscale Gibbs random field.

In the first part we studied the synchronous iterated map and showed that the center of mass of the simplex $\mathcal{T}_n$ plays a previliged role at each site. Using a nonlinear transformation [4] that preserves this simplex, we showed that all fixed points can be mapped onto the *trivial* fixed point at which the pixel mean-field matching vectors are represented by the simplex center of mass. The trivial fixed point is always a high-temperature solution of the fixed-point equations.

Next we studied the stablity of the trivial fixed point and showed that there is a lower bound on the temperature for which the trivial fixed point is stable. This lower bound depends on the number of colors and the eigenvalues of the block connectivity matrix.

An iterated map is but one method for finding the fixed points. Another method is to profit from the fact that the fixed-point equations can be obtained as the extremum equations of energy functions that can be derived from the mean-field approximation. We can then define gradient-descent dynamics whose attractors are the fixed points. In this context, we have obtained the following:

1. We have shown that the type of gradient descent to be used is decided by the relationship imposed between the mean-field variables $\mathbf{v}_s$ and the auxiliary variables $\mathbf{w}_s$.

2. If this relationship is one of linear dependence of $\mathbf{w}_s$ on $\mathbf{v}_r, r \in \mathcal{S}$, through the connectivity matrix $\mathbf{J}$ then we should use the effective energy E_{eff} to define the gradient descent equations in terms of the mean-field variables $\mathbf{v}_s$.

[4] This transformation is identical to the one used in interior-point methods for linear programming. See, for instance, Karmarkar's original paper [41].

3. If the relationship is imposed using the generalized sigmoid mapping, then H_{eff}, the Legendre transformation of E_{eff}, should be used to define gradient-descent equations in terms of the auxilary variables $\mathbf{w}_s$.

4. The latter method is a generalization of the Hopfield dynamics to the case of winner-take-all networks.

5. The realization that the generalized sigmoid mapping is convex allowed us to generalize both the Hopfield and WTA gradient descent to networks in which the neuron input-output mappings are n-dimensional but have convex potentials. This was done using the Legendre-Fenchel transform of the convex potentials.

6. When interpreted in circuit-theoretic terms, the Legendre-Fenshel transform allows us to obtain content functions from co-content functions. Thus the new gradient descent dynamics has an appealing circuit-theoretic interpretation in terms of content functions.

Finally, we studied the temperature dependence of the mean-field solutions in the context of the probability decomposition method. Using an eigenstructure analysis of the free energy Hessian, we have obtained, for the periodic lattice case, new closed form formulas for the critical temepratures at which the trivial equilibrium point becomes unstable. Using a novel projected-gradient decsent algorithm, we have verified these formulas for two Gibbs models used in texture synthesis: the Potts model and the auto-binomial model. We have also commented on the insights that these critical temperatures provide into the phenomenon of phase transition for Gibbs random fields.

7

Conclusions and Directions for Future Research

Just because some of us can read and write and do a little math,
that doesn't mean that we deserve to conquer the world.

Kurt Vonnegut

In this final chapter, we would like to give a survey of the contributions of this dissertation and point out some future research directions.

7.1 Conclusions

7.1.1 Correlation Functions of Gibbs Random Fields

The first major result of this thesis is the establishment of a close link between the mean-field approximation and the correlation-field approximation. The widely held belief that the mean-field approximation contains only information about the first-order statistics of the Gibbs distribution is unfounded. The crucial fact is that the mean-field approximation is more than an approximation of the Gibbs ensemble averages: it is actually an approximation of the Gibbs partition function which is the *generating function* of the Gibbs distribution. If we know the generating function of a probability

distribution then we have everything we need to compute its moments, in particular, the second-order moments.

There are two extreme cases in the realm of Gibbs distributions in which exact computations of the correlation functions are possible. These cases are the Ising model in which the pixel state space is binary and the Gaussian case in which the pixel state space is the real line. The class of Gibbs distributions with a finite number of graylevels, which contains all digital graylevel images, have correlation functions that are hard to derive analytically and expensive to compute numerically. Chapter 4 was devoted to explaining a new procedure for obtaining correlation-field approximations from the mean-field approximations. We applied this procedure to the autobinomial Gibbs distribution, a model of widespread use in texture synthesis, and derived a new formula for its correlation function that shows explicit dependence on the model temperature, number of graylevels, and lattice parameters and connectivity graph.

7.1.2 Winner-Take-All Networks

The important intermediary step in getting the correlation-field estimation is of course to get the mean field. The realization that the mean-field equations are actually the fixed points of a winner-take-all (WTA) network has led us, in Chapter 5, to initiate a systematic study of the dynamics of these networks both in the discrete and analog cases. Although one could define the analog network in an axiomatic fashion, i.e, independently from the discrete case, our presentation has the advantage of showing that there is a deep link between the two and that this link is provided by mean-field theory.

Among the new results, we should mention:

Iterated MAP Dynamics

1. For the discrete, synchronous WTA network, if the connectivity block matrix of the network satisfies a weakened form of positive definiteness, then the fixed points are the only limit cycles of the network.

2. The iterated map dynamics of the analog WTA network is governed by the generalized sigmoid mapping that we have proven to be the gradient map of a *convex* potential.

3. The worst that can happen in *any* synchronous, analog WTA network with a symmetric connectivity matrix is an oscillation of period 2.

4. If the connectivity block matrix of the analog WTA network satisfies the same weakened form of positive definiteness as in the discrete case, then the worst can be avoided and the fixed points become the *only* limit cycles of the network.

5. The neuron state does not have to be a number. It can be, for instance, a multidimensional signal of finite energy whose dynamics is governed by a partial differential equations. Even in this case, we can define "neural networks" whose dynamics is governed by Lyapunov functionals similar to those of the analog WTA network.

Gradient Descent Dynamics

In Chapter 6, we have also provided gradient-descent dynamics for WTA networks. These descent equations have the advantage over the iterated map dynamics of Chapter 5 of having only point attractors. We have discovered that the convexity of the sigmoid mapping potential is the crucial assumption that makes the effective energy functions derived from mean-field theory *bona fide* Lyapunov functions. This discovery allowed us to define gradient-descent equations for more general settings and to bring the role of the Legendre-Fenshel transform to the forefront. Here also we could have defined gradient dynamics in a functional space setting the same way we have done it for the iterated map dynamics.

Circuit-Theoretic Results

We have discovered that the generalized sigmoid mapping admits a simple *exact* VLSI circuit implementation operating in subthreshold mode.

We were also able to give a circuit-theoretic interpretation to the iterated-map and gradient-descent Lyapunov functions in terms of the co-content and content functions of the generalized sigmoid resistive element, respectively. We have proved that the content function of the generalized sigmoid mapping has the same functional form as information-theoretic entropy!

An aside discovery was that the content and co-content functions of a locally passive n-port resistive element are *dual* of each other in a precise sense, the duality relationship being given by the Legendre-Fenshel transform.

7.1.3 Stability and Critical Temperatures

Also in Chapter 6, we have studied the temperature dependence of the mean-field solutions in the context of the probability decomposition method. Using an eigenstructure

analysis of the free energy Hessian, we have obtained, for the periodic lattice case, new closed form formulas for the critical temperatures at which the trivial equilibrium point becomes unstable. Using a novel projected-gradient descent algorithm, we have verified these formulas for two Gibbs models used in texture synthesis: the Potts model and the autobinomial model. We have also commented on the insights that these critical temperatures provide into the phenomenon of phase transition for Gibbs random fields.

7.2 Directions for Future Research

7.2.1 Image Modeling and Processing

Mean-field theory has given us approximations of both the mean and the correlation of a graylevel Gibbs random field. As pointed out at the end of Chapter 4, they can be used to define a Gaussian joint probability distribution. This Gaussian distribution has the same underlying graph structure as the original discrete Gibbs distribution. It is also the maximum entropy distribution among all joint probability distributions having the same mean and correlation function [26]. Gaussian models have the advantage that they can be synthesized using fast frequency-domain techniques [89]. It will be worthwhile to compare patterns synthesized from these Gaussian models with the patterns obtained from their discrete parents using the time-consuming Monte-Carlo sampling procedures.

Discrete Gibbs random fields were also used to implement a labeling layer in hierechical models for segmenting textured images [16]. To the best of our knowledge, the mean-field counterparts of these hierechical models have not been investigated yet.

In this thesis, mean-field theory was used as an analytical framework for analyzing the grayscale Gibbs distributions which have been used for long as models for microtextures [14, 16]. A natural question is whether the mean-field dynamics can itself be used for pattern generation. We suggest that the useful region for investigating the pattern generation capabilities of the mean-field dynamics should be around the critical temperatures that were computed in closed form in Chapter 6. It might also be worthwhile to compare the deterministic mean-field dynamics with that of reaction-diffusion equations which are known to generate "interesting" patterns [88]. This comparison must be centered around the role of nonlinearities in both models. The mean-field nonlinearity (sigmoid, generalized sigmoid) is dissipative (it has a convex potential) while the reaction nonlinearity is not bound to be dissipative, and therefore, it is more likely to produce "richer" patterns ... at the expense of stability!

7.2.2 Neural Networks

In our analysis of the dynamics of winner-take-all (WTA) networks, we have assumed that the network connectivity matrix was a given of the problem. In the context of artificial neural networks, learning is the procedure whereby the network connectivity matrix is computed so that the network outputs (here, its stable configurations) have desirable features. In his original papers [35, 36], Hopfield suggested the use of a Hebbian/outer-product rule for defining the connectivity coefficients in his network. We can proceed in the same manner for the WTA network which thus becomes a content addressable memory (CAM) for *grayscale* patterns the same way the Hopfield network is a CAM for *binary* patterns. This line of research should open the way for a large number of interesting applications of the WTA network to optimization, image processing, and pattern recognition. Of course, the different problems that have popped up in the context of Hopfield networks will also pop up in this context: spurious minima, assymetric connections, dense connectivities, etc.

On a more speculative level, we suggest that the generalized sigmoid mapping can be considered a "grandmother" neuron [74] that fires along the axon of the winning label. Using this neuron as the output stage in a backpropagation network is worth investigating as a mechanism for implementing supervised competitive learning. Using the property of the generalized sigmoid mapping Jacobian, we can also write backpropagation learning equations for feedforward networks of these "grandmother" cells that are very similar to the backpropagation equations based on sigmoid functions. It is worthwhile noting that the lateral inhibition in the VLSI implementation of each of our "grandmother" cells has a $O(n)$ complexity rather than the usual $O(n^2)$. In the work of Poggio and Girosi [75], the role of the grandmother cell is played by the Gaussian radial basis function. This function seems to be biologically plausible [Poggio, private communication]. We ignore whether the simple VLSI implementation of the generalized sigmoid mapping is biologically plausible.

7.2.3 Constrained Optimization

The path that we have followed in this thesis starting with the correlation function of Gibbs random field image models with a finite number of graylevels is somewhat atypical. For in order to answer an inherently probabilistic question, we have adopted a deterministic method: mean-field theory. The reader should contrast this approach with the typical simulated annealing paradigm in which one usually starts with a deterministic question (find the minimum of a cost function) and uses the probabilistic framework as

an intermediary step.

Of course, the two method can be combined. Whenever we are looking for an analog solution of a discrete optimization problem, the discrete cost function can be considered the energy of a Gibbs distribution. Then the mean-field approximation can be used to get the analog algorithm. Many authors have already investigated such an approach [68, 80, 44].

We have shown how the application of such a method on a discrete cost functions with uniqueness constraints at each site results in an analog WTA network. From an optimization viewpoint, this method can be construed as a systematic way for coming up with interior-point algorithms for discrete *nonlinear* programming problems. Indeed, the analog counterpart of the discrete uniqueness constraint turned out, via the natural parameterization imposed by the generalized sigmoid mapping, to be the *interior* of the simplex $\mathcal{T}_n$. This simplex and its projective transformations are at the core of Karmarkar's interior-point method for *linear* programming. We believe that this analogy should be investigated further. Karmarkar's recent paper [42] about the continuous dynamics of his algorithm seems to be a good starting point.

A

Mathematical Theory of Phase Transitions

En ce qui concerne la connaissance théorique du réel ...
tout ce qui est facile à enseigner est inexact.

Gaston Bachelard

Abstract: In this Appendix, I try to explain (for myself, at the very least) the implications of the modern mathematical theory of phase transitions as far as the Gibbs systems used in image modeling are concerned. My objective is (to try) to make clear what probabilists mean by phase transition and to apply some of their results to the image models that are commonly used in image processing and computer vision.

A.1 Introduction

In probability theory, the problem of phase transitions arises in the following context. Given an *infinite* set of random variables and a family of conditional distributions for these random variables, find a *joint* probability distribution from which the family of conditional probability distributions can be derived. This infinite set of random variables is said to exhibit a phase transition when, to the family of conditional distributions,

correspond more than one joint probability distribution. In other words, phase transition is a failure in uniqueness. The situation at hand is not unlike that of Markov chains, where we are given a set of conditional probabilities, the transition probabilities, that might or might not admit a unique equilibrium distribution depending on the properties of the transition matrix. The fundamental difference between the situation considered in this note and that of Markov chains is that the set indexing the infinite family of random variables has no natural order and therefore we cannot use the language of semi groups, as for Markov chains, to study the "long term" behavior of the system.

This note is divided as follows. After giving the notation in the next section, we show, in Section 3, why the finite case does not pose any problems from the probabilistic point of view. Then in Section 4 we pass to the infinite case and deal with the existence problem for the joint distribution. Section 5 is devoted to the statement of a very general sufficient condition, due to Dobrushin, for the absence of a phase transition. Then in Section 6, we give a testable inequality, due to Simon, for Dobrushin's condition to be satisfied. This condition is applied to Gibbs systems with potential functions commonly used in image modeling. The consideration of an infinite family of random variables by probabilists at the outset is contrasted, in Section 7, with the idea of an infinite system as a thermodynamic limit – idea that is used in statistical physics to define phase transitions. Finally in Section 8, we give some thoughts and indicate an open research problem that might be taken up by someone more daring and capable than the author of this document.

A.2 Notation

Most of the time, we will work with a square regular lattice – the one usually used in image processing. In practice, this square lattice is a finite subset $\mathcal{S} \subset \mathbf{Z}^2$, but the mathematical theory of phase transitions is concerned with infinitely countable lattices. The problem of determining the size of the finite lattice that could capture those mathematical phase transitions as defined on the infinite lattice is, of course, of the realm of computational physics. This aspect of the problem, which is very important in applications, will not be dealt with here.

At each site $s \in \mathcal{S}$ of the infinite lattice, there is a pixel (or spin) whose state x_s is a random variable that can take its values from a set G. Typically this set is a discrete finite set or a compact interval of $\Re$. However both in image processing and statistical physics, models are used in which $G = \Re$, as for, e.g., Gaussian models, or $G = \mathbf{Z}$, as for the discrete Gaussian models. The configuration space is given by the set $\Omega = G^{\mathcal{S}}$.

This set is a countably infinite product and has at least the power of the continuum. Defining a probability measure on such sets should be done with care. The lattice will usually be equipped with a symmetric neighborhood system $\{\mathcal{N}_s, s \in \mathcal{S}\}$ that will define on the set $\mathcal{S}$ a graph structure. Although the neighborhood system is important for image processing applications, most of the mathematical questions pertaining to phase transitions can be formulated for an infinitely countable set $\mathcal{S}$ without any additional structure.

If $A \subseteq \mathcal{S}$ we denote by $x_A = (x_s, s \in A)$ the configuration of the sites belonging to the subset A. We also denote by $x_{(s)}$ the configuration of the subset $\mathcal{S}\backslash\{s\}$.

A.3 Finite-Lattice Case

Assume that we are given for the lattice $\mathcal{S}$ the set of conditional probability distributions

$$\{P_s(x_s|x_{(s)}), s \in \mathcal{S}\}.$$

Then it is natural to ask whether this set of conditional probability distributions, usually called *local characteristics* determine, in a unique manner, a joint probability distribution on the lattice. If the lattice is finite the answer is yes: a *unique* joint probability distribution can be recovered from the conditional probability distributions. So assume that $\mathcal{S}$ is finite and number its sites from 1 to $N = |\mathcal{S}|$ then we have the following

Proposition 33 [Besag [6]] *If the lattice is finite then the joint probability distribution is uniquely determined by its local characteristics.*

Proof: The proof uses the fact that the joint probability distribution for a finite lattice satisfies

$$\frac{P(\mathbf{x})}{P(\mathbf{y})} = \prod_{s=1}^{N} \frac{P(x_s|x_1, \ldots, x_{s-1}, y_{s+1}, \ldots, y_N)}{P(y_s|x_1, \ldots, x_{s-1}, y_{s+1}, \ldots, y_N)} \tag{A.1}$$

which can be easily proved by noting that

$$P(\mathbf{x}) = P(x_N|x_{(N)})P(x_{(N)}) = \frac{P(x_N|x_{(N)})}{P(y_N|x_{(N)})} P(x_1, \ldots, x_{N-1}, y_N) \tag{A.2}$$

and repeating the above formula until all the y_s's are exhausted. By fixing a reference configuration $\mathbf{y}$, the above formula proves the existence of the joint probability distribution. Uniqueness is proved by noticing that if P and P' are two joint distributions having the same local characteristics then $\frac{P(\mathbf{x})}{P(\mathbf{y})} = \frac{P'(\mathbf{x})}{P'(\mathbf{y})}$, which means that $P \equiv P'$. ∎

The above characterization of joint probability distributions for finite lattices calls the following remarks:

- The local characteristics need to be positive for Formula (A.1) to be valid for any two configurations $\mathbf{x}$ and $\mathbf{y}$.

- If some of the local characteristics take zero as a value, then this would mean that some of the lattice configurations are prohibited. In other words we have a constrained system. Let us denote by Ω' the subset of the configuration space Ω that is defined by the constraints. Then (A.1) can be rewritten by conditioning both sides on the event $\mathcal{A} = \{\mathbf{z} \in \Omega'\}$, i.e.,

$$\frac{P(\mathbf{x}|\mathcal{A})}{P(\mathbf{y}|\mathcal{A})} = \prod_{s=1}^{N} \frac{P(x_s|x_1,\ldots,x_{s-1},y_{s+1},\ldots,y_N,\mathcal{A})}{P(y_s|x_1,\ldots,x_{s-1},y_{s+1},\ldots,y_N,\mathcal{A})} \tag{A.3}$$

- Note that there is not (yet) any type of markovianity assumption on the system. Formula (A.1) is valid for any kind of probability distribution on a finite set. It should be stressed that the role of a local markovianity assumption on the local characteristics is to ease the computation of the latter by looking at a small subset of the configuration set.

- The set $\mathcal{S}$ is not assumed to have any specific structure. In particular, the graph structure introduced by the neighborhood system is not used.

- The question of the existence and uniqueness of the joint probability distribution is purely *probabilistic*. At this level, there is no need to define a Gibbs system describing the *physical* interactions between the site spins.

A.4 Infinite-Lattice Case

If the $|\mathcal{S}|$ is infinite then Formula (A.1) is no longer valid [30]. In order to describe the mathematical theory dealing with the questions of existence and uniqueness of a joint distribution in the infinite lattice case, we should first set up the right machinery for defining the conditional probabilities in a mathematically meaningful way. Although there are many presentations on how this can be done [77, 51, 32], we adopt here the original formulation of Dobrushin in which the state space G of each spin is a finite set. First we need to clarify what we mean by the probability distribution of the lattice. We denote by $F(\mathcal{S})$ the set of *finite* subsets of $\mathcal{S}$.

To each finite subset $\Lambda \in F(\mathcal{S})$ corresponds a probability distribution $P_\Lambda(x_s, s \in \Lambda)$ such that

$$Prob\{X_s = x_s, s \in \Lambda\} = P_\Lambda(x_s, s \in \Lambda). \tag{A.4}$$

The family $\mathbf{P} = \{P_\Lambda, \Lambda \in F(\mathcal{S})\}$ of all probability distributions of all the finite sets of the lattice must satisfy the following consistency conditions for any two finite subsets $\Lambda' \subset \Lambda \in F(\mathcal{S})$:

$$\sum_{s \in \Lambda \backslash \Lambda', x_s \in G} P_\Lambda(x_s, s \in \Lambda) = P_\Lambda(x_s, s \in \Lambda'), \tag{A.5}$$

which means that for two nested, finite sets the probability distribution of the smaller set can be obtained as a marginal distribution from the distribution of the larger set. The family $\mathbf{P}$ is called the distribution of the random field $\mathbf{X}$ defined on the infinite lattice.

The next step is to define what we mean by a conditional distribution on the infinite lattice. For this, we consider a finite set $\Lambda \in F(\mathcal{S})$ and a configuration $\mathbf{y} = \{y_s, s \in \mathcal{S}\backslash\Lambda\}$, and we associate to them the conditional probability distribution defined by

$$Prob\{X_s = x_s, s \in \Lambda | \mathbf{y}\} = Q_{\Lambda,\mathbf{y}}(x_s, s \in \Lambda). \tag{A.6}$$

As in the case of the family $\mathbf{P}$, we also require the family $\mathbf{Q} = \{Q_{\Lambda,\mathbf{y}}, \Lambda \in F(\mathcal{S}), \mathbf{y} \in G^{\mathcal{S}\backslash\Lambda}\}$ of all conditional distribution to satisfy a set of consistency conditions for finite, nested sets. In order to understand these conditions, one should start with the following lemma:

Lemma 17 *Let A, B, and C be three events. Then we have for the conditional probabilities*

$$Prob(A \cup B|C) = Prob(A|B \cup C)Prob(B|C) \tag{A.7}$$

Proof:

$$\begin{aligned} Prob(A \cup B|C)Prob(C) &= Prob(A \cup B \cup C) \\ &= Prob(A|B \cup C)Prob(B \cup C) \\ &= Prob(A|B \cup C)Prob(B|C)Prob(C). \end{aligned} \tag{A.8}$$

■

Let now $\Lambda' \subset \Lambda \in F(\mathcal{S})$ be two finite subsets and $\mathbf{y} \in G^{\mathcal{S}\backslash\Lambda}$ the lattice configuration *outside* the subset Λ. Define the events $A = \{X_s = x_s, s \in \Lambda'\}, B = \{X_s = x_s, s \in \Lambda\backslash\Lambda'\}$, and $C = \{X_s = y_s, s \in \mathcal{S}\backslash\Lambda\}$. Then we have

$$\begin{aligned} Prob(A \cup B|C) &= Q_{\Lambda,\mathbf{y}}(x_s, s \in \Lambda) \\ Prob(A|B \cup C) &= Q_{\Lambda',\tilde{\mathbf{y}}}(x_s, s \in \Lambda') \\ Prob(B|C) &= \sum_{x_s \in G, s \in \Lambda'} Q_{\Lambda,\mathbf{y}}(x_s, s \in \Lambda). \end{aligned} \tag{A.9}$$

where $\tilde{\mathbf{y}}$ is defined by $\tilde{y}_s = y_s, s \in \mathcal{S}\backslash\Lambda$ and $\tilde{y}_s = x_s, s \in \Lambda\backslash\Lambda'$. Applying Lemma 17 above we get the consistency conditions that must be satisfied by the family of conditional distributions $\mathbf{Q}$.

Given the above definitions, there are two important questions that must be dealt with. The first is: given a consistent family of conditional distributions $\mathbf{Q}$, does there exist a joint probability distribution such that the family $\mathbf{Q}$ would be its family of conditional distributions? The second is: what are the conditions that we should impose on the family $\mathbf{Q}$ so that the joint probability distribution is unique when it exists?

The existence question is not too difficult to solve. For the probability distribution to exist it is sufficient that the conditional probabilities be Markovian [1]. Before giving the precise definition of this notion, let us note the following

1. In the finite case, there was no need for any additional assumption on the conditional distributions for the joint probability to exist.

2. Another difference with the finite case is that in order to define the Markov property, we need to equip the set of sites $\mathcal{S}$ with a neighborhood system, i.e., with a graph structure in which the sites become the vertices of the graph.

To keep the definition as general as possible, we now assume that the lattice is a graph $(\mathcal{S}, E)$, where $\mathcal{S}$ is the set of vertices and E the set of edges. The *graph incidence function*, $c : \mathcal{S} \times \mathcal{S} \to \{0,1\}$, is defined by $c(s,r) = 1$, if there is an edge between the sites s and r, and $c(s,r) = 0$, otherwise. If $\Lambda \in \mathcal{S}$, the *boundary* of Λ, denoted by $\partial\Lambda$ is the subset of $\mathcal{S}$ defined by $\partial\Lambda = \{r \in \mathcal{S} | c(r,s) = 1, s \in \Lambda\}$. The boundary of a site s will be simply denoted ∂s. It will be assumed that $\partial s \in F(\mathcal{S}), \forall s \in \mathcal{S}$, so that if $\Lambda \in F(\mathcal{S})$ then $\partial\Lambda$ is also in $F(\mathcal{S})$.

Definition 10 *The family* $\mathbf{Q}$ *of conditional probability distribution is said to be* Markovian *if the following condition is satisfied*

$$Q_{\Lambda,\mathbf{y}}(x_s, s \in \Lambda) = Q_{\Lambda,\mathbf{y}_{\partial\Lambda}}(x_s, s \in \Lambda), \tag{A.10}$$

where $\mathbf{y}_{\partial\Lambda}$ *is the configuration of the lattice on the boundary of the set* Λ.

In other words, in order to determine the conditional probability distribution in a *finite* subset of the lattice given that the lattice configuration is known *outside* the subset, it is sufficient to condition on the lattice configuration at the *boundary* of the subset. Note that the definition of the Markov property is *global*, i.e., in terms of finite subsets of the lattice rather than *local* in terms of single spins. When the spin state space G is finite, the local and global definition are equivalent [18]. However, when $G = \Re$, as in the Gaussian case, then one can find random fields that are *locally* Markov but not Markov [10]. Now we can state Dobrushin's existence theorem:

[1] This assumption is actually stronger than the one needed. In fact, it is enough that the conditional probabilities be "asymptotically close" to Markovian.

Theorem 16 *Let (S, E) be an infinite graph where each vertex has a finite boundary, and let G be a finite set. Let* $\mathbf{Q}$ *be a Markovian family of consistent conditional distributions. Then there exists a joint distribution* $\mathbf{P}$ *having* $\mathbf{Q}$ *for conditional distributions.*

Proof: see [18]. ∎

For a given family $\mathbf{Q}$ of conditional distributions, we denote by $\mathcal{P}(\mathbf{Q})$ the set of joint distributions defined by $\mathbf{Q}$. In the case of a Markovian field, Theorem 16 says that $\mathcal{P}(\mathbf{Q})$ contains at least one element. The situation where $\mathcal{P}(\mathbf{Q})$ contains more than one element led Dobrushin to give the following definition [18].

Definition 11 *A phase transition is said to occur if* $\mathcal{P}(\mathbf{Q})$ *has two elements or more.*

To intuitively understand this definition, one should view the infinite graph as an infinite volume thermodynamic system (i.e., a system considered in the thermodynamic limit), in which the site interactions are described by the family $\mathbf{Q}$ of conditional distributions. The thermodynamic equilibrium of the system is described by a joint distribution $\mathbf{P}$. When more than one joint probability distribution exists for a given set of interactions, the system can exist in as many different thermodynamic equilibria as there are joint distributions. Each of these equilibria is called a phase. An important question in statistical physics is to know whether a system can exhibit a phase transition. In the following section, we give a sufficient condition for the absence of a phase transition also due to Dobrushin. We also indicate how this condition can be used to give the ranges of system parameters for which there can be no phase transition. Before closing this section, notice should be made of the fact that no mention has been made of Gibbs distributions in the formulation of the different definitions and problems. But it is clear that under the conditions of Theorem 16, and the condition that the conditional probabilities are positive, we can define, for each $\Lambda \in F(\mathcal{S})$, a Gibbs distribution. Moreover this Gibbs distribution will be, by the Hammersley-Clifford theorem, compatible with the graph structure of the lattice.

A.5 Dobrushin's Uniqueness Condition

Dobrushin's sufficient condition for the uniqueness of the joint probability distribution on the infinite lattice is formulated in terms of the *local* conditional distributions, i.e., the ones obtained when the finite subset Λ is restricted to one site. Again, we remind ourselves that this restriction is valid only because the spin state space G is finite. To intuitively understand the uniqueness condition, we could argue as follows. For a finite system, the joint probability condition is uniquely defined by the conditional distribution.

Therefore, one should try to make the infinite case as close as possible to the finite case by making sure that the local conditional distributions depend weakly on the configurations of distant sites and that no matter where we are on the lattice, the distance of the local conditional probabilities to the other ones is small. These notions will now be made rigorous.

Let $r, s \in \mathcal{S}$ be two different sites on the lattice, and let $Q_{r,\mathbf{y}}$ be the conditional probability distribution associated with the site r when the lattice configuration outside $\{r\}$ is $\mathbf{y}$. Let now $\mathbf{y}_{(s)}$ be any configuration for $\mathcal{S}\backslash\{r\}$ that agrees with $\mathbf{y}$ everywhere except at the site s. Consider now the *distance* between the two conditional distributions $Q_{r,\mathbf{y}}$ and $Q_{r,\mathbf{y}_{(s)}}$ defined by

$$d(Q_{r,\mathbf{y}}, Q_{r,\mathbf{y}_{(s)}}) = \frac{1}{2} \sum_{x \in G} |Q_{r,\mathbf{y}}(x) - Q_{r,\mathbf{y}_{(s)}}(x)|. \tag{A.11}$$

Note that this distance is zero when the two conditional distributions are identical. Note also that it depends on the value taken by the configuration $\mathbf{y}_{(s)}$ at the site s. We can eliminate this dependence by considering

$$d_{rs} = \sup_{\mathbf{y}, \mathbf{y}_s} d(Q_{r,\mathbf{y}}, Q_{r,\mathbf{y}_{(s)}}), \tag{A.12}$$

which measures how dependent the conditional distributions at site r are on the values taken by the field at site s. With these definitions, we can state Dobrushin's uniqueness theorem

Theorem 17 *If*

$$sup_{r \in \mathcal{S}} \sum_{s \neq r} d_{rs} < 1 \tag{A.13}$$

then $\mathcal{P}(\mathbf{Q})$ contains one and only one element.

An interesting interpretation of the above condition is that the infinite matrix $\mathbf{d} = [d_{rs}]$ (where we pose $d_{rr} = 0$) has an ℓ_∞-induced norm that is less than one. This matrix can therefore be considered a contracting operator on the space ℓ_∞ of bounded sequences. Note also that if the field is Markovian then $d_{rs} = 0, \forall s \notin \partial r$, which means that the matrix $\mathbf{d}$ is sparse with a banded structure. Then it becomes very easy to check Condition A.13 above.

A.6 Simon's Uniqueness Formula: Application to Gibbs Measures

In this section we present some applications of Dobrushin's uniqueness condition. All the applications use the formalism of Gibbs distributions. To introduce this formalism,

we will have to define the important notion of a *potential function.* We then present Simon's uniqueness formula that implies Dobrushin's uniqueness condition. Simon's formula is then applied to some Gibbs distribution commonly used in image processing and computer vision to derive upper bounds on the phase transition temperatures in terms of the parameters of the interaction potentials. These bounds are extremely useful when we want to sample the Gibbs distribution or when we run simulated annealing algorithms.

A.6.1 Potential Functions and Energy for the Infinite Case

The material presented in this section is pretty standard and could be found in, e.g., [32]. Our presentation however omits important mathematical details so as not to overload the unfamiliar reader.

Definition 12 *A function* $\Phi : F(\mathcal{S}) \times \Omega \rightarrow \Re$ *is called a potential if* $\forall \Lambda \in F(\mathcal{S})$ *and* $\forall \omega \in \Omega$ *the series*

$$E_{\Lambda}^{\Phi}(\omega) = \sum_{A \in F(\mathcal{S}), A \cap \Lambda \neq \emptyset} \Phi(A, \omega) \tag{A.14}$$

is summable.

The quantity $E_{\Lambda}^{\Phi}(\omega)$ is called the *total energy* of the subset of sites Λ for the configuration ω. In order to understand the above definition better, assume that Λ is a single site s. Then the total energy of s for the configuration ω is written

$$E_{s}^{\Phi}(\omega) = \sum_{s \in A \in F(\mathcal{S})} \Phi(A, \omega). \tag{A.15}$$

In other words, for a given configuration of the lattice sites, the contribution of the site s to the lattice energy is the sum of the (interaction) potentials computed at the configurations of the finite subsets containing s. This definition calls some remarks:

1. Assume that the set $\mathcal{S}$ is finite. Then in Definition A.14, we can take $\Lambda = \mathcal{S}$. Then the total energy of the lattice for the configuration ω is given by [77]

$$E_{s}^{\Phi}(\omega) = \sum_{A \subset \mathcal{S}} \Phi(A, \omega). \tag{A.16}$$

2. The summability condition is always satisfied for the finite case. Therefore any function defined on $F(\mathcal{S})$ with values in $\Re$ is a potential function. However, sometimes the conditions

   - $\Phi(\emptyset, \cdot) = 0$;

- $\forall A \in F(\mathcal{S}), \omega_A = \omega'_A \Rightarrow \Phi(A,\omega) = \Phi(A,\omega')$

are imposed, see [30].

3. Definition A.14 is valid for any set $\mathcal{S}$. When the set $\mathcal{S}$ has a graph structure as in image grids, we can be more specific about the values of the potential function over special subset of vertices, e.g., simplices in which $c(s,r) = 1$ for all pairs $(r,s), r \neq s$.

Examples

Let us now give some examples to illustrate the concept of a potential.

1. *Generic Case:* Very often the potential function is separable, i.e.,

$$\Phi(A,\omega) = J(A)\Psi(\omega_A),\ \forall A \in F(\mathcal{S}),\ \forall \omega \in \Omega \tag{A.17}$$

where $J(\emptyset) = 0$. Note that this potential function satisfies the two conditions mentioned above in the context of finite lattices.

2. *Pair Potentials:* In this case, we have $\Phi(A,\cdot) = 0$ as soon as $|A| > 2$. For the generic case, the pair potential is defined by

$$\Phi(A,\omega) = \begin{cases} J(r,s)\Psi(\omega_r,\omega_s) & \text{if } A = (r,s), r \neq s, \\ J(r,r)\psi(\omega_r) & \text{if } A = \{r\}, \\ 0 & \text{otherwise,} \end{cases} \tag{A.18}$$

where $J : \mathcal{S} \times \mathcal{S} \to \Re$ and $\Psi : G \times G \to \Re$ are symmetric. In this case, the formula for the total energy at site s for the configuration ω is given by

$$E_s^{\Phi}(\omega) = J(s,s)\psi(\omega_s) + \sum_{r \in (\mathcal{S} \setminus \{s\})} J(s,r)\Psi(\omega_s,\omega_r). \tag{A.19}$$

Note that under the summability condition, the above series, which might contain an infinite number of nonzero terms, always exists. The pair potential as defined here is valid for any set $\mathcal{S}$ whether it has a graph structure or not.

3. *Gibbs Potential:* Assume now that $\mathcal{S}$ is the set of vertices of a graph $(\mathcal{S}, E)$. A subset $\mathcal{C} \subset \mathcal{S}$ is called a *clique* [2] if the restriction of c to $\mathcal{C} \times \mathcal{C}$ is identically equal to one. In other words, any pair of vertices picked up from $\mathcal{C}$ are linked by an edge in E. We denote by $C(\mathcal{S})$ the set of cliques of the graph $(\mathcal{S}, E)$. Then a potential function

[2] In graph theory, the word *simplex* is usually used to designate a clique, while the word *clique* is reserved to simpleces of maximal number of vertices.

is a *Gibbs potential* if $\Phi(A,\cdot) = 0$, if $A \notin C(\mathcal{S})$. It should be noted that the Gibbs potential is defined only for finite cliques, i.e., for the set $FC(\mathcal{S}) = F(\mathcal{S}) \cap C(\mathcal{S})$. Let now s be a fixed lattice site, and let $C_s = \{\mathcal{C} \in FC(\mathcal{S}) | s \in \mathcal{C}\}$ be the set of finite cliques containing the site s. The energy contribution of site s for the configuration ω is given by

$$E_s^{\Phi}(\omega) = \sum_{\mathcal{C} \in C_s} J(\mathcal{C}) \Psi \omega_{\mathcal{C}}. \tag{A.20}$$

We note that in the finite case, we have $C(\mathcal{S}) \subset F(\mathcal{S})$.

4. *Multidimensional Ising Potential:* This is a Gibbs potential defined on the lattice $\mathbf{Z}^d, d \geq 2$, equipped with the nearest-neighbor neighborhood system. In this case, $FC(\mathcal{S}) = C(\mathcal{S}) = \mathcal{S} \cup \{\{r,s\} \subset \mathbf{Z}^d | c(s,r) = 1\}$, and the potential is a pair potential so that we can write for every $s \in \mathcal{S}$

$$E_s^{\Phi}(\omega) = J(s,s)\psi(\omega_s) + \sum_{\{s,r\} \in C(\mathcal{S})} J(s,r)\Psi(\omega_s, \omega_r). \tag{A.21}$$

A.6.2 Simon's Formula

In [81], Simon has shown that Dobrushin's uniqueness condition is implied by an easily testable inequality that involves the potential functions of the conditional Gibbs distributions. It was mentioned in Section 2 that under the Markovianity assumption it was possible to define a Gibbs distribution that would be compatible with the graph structure of the lattice. In fact, the Markov assumption is not really needed to *define* conditional Gibbs distributions even in the infinite dimensional case. This can be done as follows. Let Φ be a potential defined on the set $\mathcal{S}$. Let $\Lambda \in F(\mathcal{S})$ and $\mathbf{y}$ be the configuration of the sites $\in \mathcal{S} \backslash \Lambda$, the environment of Λ. Then the Gibbs conditional probability of the subset Λ for the environment configuration $\mathbf{y}$ is given by

$$Q_{\Lambda,\mathbf{y}}(\mathbf{x}) = \frac{1}{Z_{\Lambda}} \exp\left(-\frac{1}{T} E_{\Lambda}^{\Phi}(\mathbf{xy})\right). \tag{A.22}$$

where

$$Z_{\Lambda,\mathbf{y}} = \sum_{\mathbf{x} \in \Omega_{\Lambda}} \exp\left(-\frac{1}{T} E_{\Lambda}^{\Phi}(\mathbf{xy})\right)$$

is the *conditional* partition function associated with the finite subset Λ and the *boundary condition* $\mathbf{y}$. The symbol $\mathbf{xy}$ corresponds to the site configuration ω given in Definition (12). Note that the summation in the conditional partition function is only over the site configurations of the subset Λ. Note also that the above formula is implicitly assuming that the spin state space G is finite. In case G is $\Re$ as in the Gaussian model or a

compact interval in $\Re$ as in the Heisenberg model, the above definition of the conditional distributions should be changed to take into account the probability measure on the state space. Again, we should note that this definition is independent of the lattice structure but depends crucially on the potential function.

We are now in a position to state Simon's theorem. For a function f defined on Ω, $\|f\|$ will be the $\sup_{\omega\in\Omega}|f(\omega)|$. The potential function Φ is said to be *shift-invariant* or *homogeneous* if $\Phi(\Lambda+p,\tau_p\omega)=\Phi(\Lambda,\omega),\forall\Lambda\in F(\mathcal{S}),p\in\mathcal{S}$, where $\tau_p\omega=(\omega_{s-p})_{s\in\mathcal{S}}$ is the translate of the configuration $\omega\in\Omega$.

Theorem 18 *Let $\Phi: F(\mathcal{S})\times\Omega\to\Re$ be a homogeneous potential function defined on the set $\mathcal{S}$. If*

$$\frac{1}{T}\sum_{0\in\Lambda}(|\Lambda|-1)\|\Phi(\Lambda,\cdot)\|<1, \tag{A.23}$$

then the conditional Gibbs distributions satisfy Dobrushin's uniqueness condition.

Proof: see [81]. ■

The summation in Condition (A.23) is over all the finite subsets *containing* the site 0. Each term in the summation gives the maximum contribution over all configurations that this subset could contribute to the total energy. Note that because of translation invariance, the point at which Inequality A.23 is tested is irrelevant. Note also that this inequality is generally satisfied when the temperature T is high enough. It therefore gives an upper bound on the true critical temperature of the infinite system. This fact will be exploited in the next section to give upper bounds on the true critical temperatures for some commonly used Gibbs systems.

A.6.3 Applications

In this subsection, we give some applications of Simon's formula to the problem of existence of a phase transition for some of the potential function examples mentioned in the previous subsection. We will assume that $\mathcal{S}=\mathbf{Z}^2$ and that the potential function is a pairwise, homogeneous potential. Under these assumptions, the right hand side of Simon's inequality can be written as

$$\frac{1}{T}\sum_{0\neq s\in\mathcal{S}}\|\Phi(\{0,s\},\cdot)\|.$$

If in addition we assume that the potential Φ is factorizable, i.e.,

$$\Phi(\{0,s\},\omega)=J(0,s)\Psi(\omega_0,\omega_s)$$

as in the generic example above then Simon's inequality can be written as

$$\frac{1}{T}\sum_{0\neq s\in\mathcal{S}}|J(0,s)|\sup_{\omega_0,\omega_s\in G}|\Psi(\omega_0,\omega_s)|<1. \tag{A.24}$$

Before applying the above formula to actual systems by specifying the bond parameters $J_s = J(0,s)$ and the interaction function Ψ, we would like to make the following remarks:

1. The inequality does not distinguish between attractive and repulsive potentials – the sign of J_s does not enter into consideration.

2. The L_∞ norm of the function Ψ should exist. A sufficient condition for this is that Ψ be bounded on $G \times G$. It is clear that a Gaussian potential (i.e., pairwise, quadratic potential with $G = \Re$) does not satisfy this condition. However this condition is always satisfied whenever G is compact and Ψ continuous with respect to the natural topology of G.

3. Again, we stress that this is only a sufficient condition. It is unable, for instance, to tell us whether the Ising chain ($\mathcal{S} = \mathbf{Z}, G = \{-1,+1\}$) exhibits a phase transition below a certain temperature. In fact, we know by direct computation that it does not.

Now, we give some actual examples for applying Simon's inequality. As was mentioned, They are all for the case $\mathcal{S} = \mathbf{Z}^2$. We will also assume that the lattice has a homogeneous graph structure, i.e., each node on the lattice is connected to the same number of neighboring vertices. We denote this number by ν. This graph structure is introduced because it is the one that is prevalent in the image processing literature. The graph $\nu = 4$ (4 nearest neighbors) corresponds to a first order neighborhood model; $\nu = 8$ (eight nearest neighbors) corresponds to a second order neighborhood model and so on. Moreover, we consider only isotropic models.

1. *Potts model:* In this model, we have $G = \{1,\ldots,n\}, J_s = J, \forall s \in \mathcal{S}, \Psi(\omega_r,\omega_s) = \delta(\omega_r,\omega_s)$ where $\delta(\cdot,\cdot)$ is Kronecker's delta function. Simon's inequality becomes simply

$$\frac{\nu|J|}{T}<1, \tag{A.25}$$

or $T > \nu|J|$. In other words, if there is a phase transition it can only happen for a critical temperature $T_c \leq n|J|$. For the case where the bonding strength $|J| = 1$ and the $\nu = 4$, we get the upper bound $T_c \leq 4$. Note that this upper bound is independent of n, the number of colors. In Section 6.4.2, a mean-field estimate

of the critical temperature was given for the periodic lattice and was found to be $\hat{T}_c = 8/n$. When $n = 2$ we get the upper bound given by Simon's formula. As n increases, the mean-field estimate decreases consistently with Simon's upper bound. It is interesting to note that when n is very large, $\hat{T}_c \to 0$. This result is also consistent with the exact result, reported in [32], that for $n > 12$, there is no phase transition in the sense of Dobrushin.

2. *Autobinomial model:* Here also, we have $G = \{1, \ldots, n\}, J_s = J, \forall s \in \mathcal{S}, \Psi(\omega_r, \omega_s) = \omega_r, \omega_s$. The supremum of Ψ over $G \times G$ is n^2. There results that the Simon's inequality becomes

$$\frac{\nu|J|n^2}{T} < 1, \tag{A.26}$$

or $T > \nu|J|n^2$. For the nearest-neighbor, binary case $\nu = 4, n = 2$, and $|J| = 1$ we get for T_c the upper bound $T_c \leq 16$. This bound is actually very coarse. In Section 6.4.2, the mean-field estimate of the critical temperature for this model was found to be $\hat{T}_c = 2$. This is also consistent with Simon's upper bound.

3. *Heisenberg model:* Now we assume that $G = [0, 2\pi]$, a compact interval $\subseteq \Re, J_s = J, \forall s \in \mathcal{S}, \Psi(\omega_r, \omega_s) = \cos(\omega_r - \omega_s)$, which is upper bounded by 1. It follows that we get a result similar to that of the Potts model, i.e., $T_c \leq \nu|J|$.

A.7 Relationship to the Theory of Yang-Lee

In classical statistical physics, the partition function is used to define a phase transition. According to a generally accepted definition in this field, a phase transition occurs at a zero of the partition function considered as a function of the inverse temperature $\beta = \frac{1}{T}$. However for a finite lattice and when $\beta \in \Re$, the partition function is strictly positive and a zero cannot exist. But if β is treated as a complex variable, then a zero will exist. In the latter case however, the zero is a complex variable that is hard to relate to the physical temperature of the system. The situation can be remedied by considering the thermodynamic limit, i.e, what happens to the complex zeros of the partition function as the lattice size goes to infinity while the spin density remains constant. It is clear that as in the mathematical case, we might run into an existence problem for the thermodynamic limit. This existence depends, first, on how the thermodynamic limit is taken and, second, on the interaction potentials between spins [85]. If the thermodynamic limit exists, then the sequence of zeros of the partition functions of the finite lattices will

possess a clustering point on the real axis. This clustering point would then be used to define a phase-transition temperature.

A.8 Conclusions

In conclusion, we would like to state the following remarks:

1. Dobrushin's theory of Gibbs states does not exhaust all phase transitions. Indeed, Gruber (1976) has come up with an example in which the infinite system has a unique Gibbs distribution at all temperatures, i.e., there is no phase transition in the sense of Dobrushin, but where the free energy has a singularity, i.e., there is a phase transition in the sense of macroscopic thermodynamic. In other words, the existence of more than one Gibbs measure is the "worst" that can happen to the infinite system, but there are other "bad" things that can happen also.

2. The theory presented in this appendix does not deal with spin state space that is unbounded. In particular, it does not deal with the important case of Gaussian spins. Cassandro and his collaborators (1979) extended the theory to deal with the unbounded case and have come up with a generalization of Dobrushin's condition to this case. The case of Gibbs/Markov Gaussian fields is a research area in its own right that reduces in the 1D case to the study of the so-called reciprocal processes. It is of particular importance since the results obtained can be used to solve important engineering problems like multidimensional smoothing [50].

3. It is to be noted that the systems considered were all unconstrained. The introduction of constraints on the configurations introduces long range interactions that might lead to the violation of Dobrushin's uniqueness condition. It is therefore expected that in the constrained case, phase transitions could happen in regions that are otherwise prohibited. The problem of finding a sufficient condition for uniqueness in the constrained system similar to Dobrushin's is still apparently an open research question in the theory of Gibbs systems [D. W. Stroock, private communication].

B

Saddle-Point Approximation

Quite simple, my dear Watson.

Sherlock Holmes

In this appendix, we give more details on the saddle-point approximation that we used in Chapter 3 to replace the integral representation of the partition function with a simpler expression that depends on the point where the *effective* energy has a global minimum. This type of approximation is a multidimensional generalization of Laplace's method for finding the leading term in the asymptotic expansion as $x \to +\infty$ of integrals of the form

$$I(x) = \int_a^b e^{-x\phi(t)} dt, \tag{B.1}$$

where ϕ is a real continuous functions. The fundamental idea behind Lapalce's method is the following: If the real continuous function $\phi(t)$ has its global minimum at $c \in [a, b]$ then it is only the immediate neighborhood of $c \in [a, b]$ that contributes to the asymptotic expansion of $I(x)$. An excellent reference on this and related methods is the book by Erdelyi, *Asymptotic Expansions* [24].

There are two steps involved in the approximations:

1. Assuming $c \in$ the open interval (a, b), replace $I(x)$ with

$$I(x; \epsilon) = \int_{c-\epsilon}^{c+\epsilon} e^{-x\phi(t)} dt, \tag{B.2}$$

for ϵ sufficiently small.

2. Use the Taylor series expansions of ϕ in the neighborhood $(c-\epsilon, c+\epsilon)$ of c to find the asymptotic expansion of $I(x;\epsilon)$.

One can easily show that the quantity $I(x) - I(x;\epsilon)$ is exponentially small compared with $I(x)$ as $x \to +\infty$. This implies that the asymptotic expansions of $I(x)$ and $I(x;\epsilon)$ are identical as $x \to +\infty$. For more details, see the abovementioned reference.

The mechanics of the method is rather simple. Writing the second-order Taylor series expansion of $\phi(t)$ in the neighborhood of c, we get

$$\phi(t) = \phi(c) + \frac{1}{2}(t-c)^2\phi^{(2)}(c) + o\left((t-c)^2\right), \tag{B.3}$$

the first-order term being zero because $\phi'(c) = 0$ at the isolated global-minimum point c. We will assume from now on that the integration is over all the real line. Substituting into the integral and neglecting the contribution of the higher order terms of the Taylor series, we get

$$I(x) \approx e^{-x\phi(c)} \int_{-\infty}^{+\infty} e^{-\frac{x}{2}\phi^{(2)}(c)(t-c)^2}.$$

Since $\phi^{(2)}(c) > 0$, the above integral evaluates to

$$\sqrt{\frac{2\pi}{x\phi^{(2)}(c)}},$$

which gives

$$I(x) \approx \sqrt{\frac{2\pi}{x\phi^{(2)}}} e^{-x\phi(c)}.$$

When we are dealing with the multidimensional case as in the approximation of the partition function, the Gaussian integral evaluates to

$$\left(\frac{2\pi}{x}\right)^{|\mathcal{S}|/2} / \sqrt{det(\mathbf{H}^*)}$$

where $|\mathcal{S}|$ is the size of the lattice, and $\mathbf{H}^*$ is the Hessian of the multivariable function $\phi(\mathbf{v})$ computed at the global minimum point $\mathbf{v}^*$. The integral $I(x)$ can therefore be approximated in the multidimensional case by

$$I(x) \approx \left(\frac{2\pi}{x}\right)^{|\mathcal{S}|/2} / \sqrt{det(\mathbf{H}^*)} e^{-x\phi(\mathbf{v}^*)}. \tag{B.4}$$

This approximation is better, the larger the parameter x. In the case of the partition function approximation, this parameter depends on the lattice temperature, T, and the lattice size, $\mathcal{S}$. Indeed, one can write

$$\frac{1}{T}E_{eff}(\mathbf{v}) = \frac{|\mathcal{S}|}{T}(E_{eff}(\mathbf{v})/|\mathcal{S}|),$$

where the term between parentheses is the effective energy per pixel which can be considered approximately independent of the lattice size. This heuristic argument suggests that the mean-field approximation improves, the larger the lattice size.

In Section 3.1.4, we used both the above approximation and its generalization to the case where the integral to be expanded is of the form

$$I(x) = \int_a^b f(t)e^{-x\phi(t)}dt,$$

where f and ϕ are real continuous functions. In this case, one can show that the leading term in the expansion of $I(x)$ as $x \to +\infty$ is

$$\frac{\sqrt{2\pi}f(c)e^{-x\phi(c)}}{\sqrt{x}\phi(2)(c)}$$

where c is the point at which the function ϕ reaches a global minimum. For more details, see [24]. A similar leading term can also be obtained for the multivariable case.

As stated by Amit ([2], p. 137f), "Physics tends to lump together three methods - steepest descent, saddle-point and Laplace's method, unless distinctions are absolutely required." Indeed, the name "steepest-descent" is usually reserved for the case where the functions f and ϕ are complex analytic in the complex variable t and where $I(x)$ is a contour integration.

Bibliography

[1] N. Ahuja and B. Shachter. *Pattern Models.* Wiley, 1983.

[2] D. J. Amit. *Modeling Brain Function.* Cambridge University Press, 1989.

[3] J. J. Atick and A. N. Redlich. Quantitative tests of a theory of retinal processing: Contrast sensitivity curves. Tech. rep. iassns-hep-90/51, Institute for Advanced Study, October 1990.

[4] H. B. Barlow. The coding of sensory messages. In W. H. Thorpe and O. L. Zangwill, editors, *Current Problems in Animal Behavior.* Cambridge University Press, 1961.

[5] R. J. Baxter. *Exactly Solved Models in Statistical Mechanics.* Academic Press, Inc., London, 1982.

[6] J. Besag. Spatial interaction and the statistical analysis of lattice systems (with discussion). *Journal Royal Statistical Society, series B*, 36:192–236, 1974.

[7] G. L. Bilbro, W. E. Snyder, S. J. Garnier, and J. W. Gault. Mean field annealing: A formalism for constructing GNC-like algorithms. *IEEE Transactions on Neural Networks*, 3(1):131 – 138, 1992.

[8] G. L. Bilbro, W. E. Snyder, and R. C. Mann. Mean-field approximation minimizes relative entropy. *J. Opt. Soc. Am. A*, 8(2):290 – 294, 1991.

[9] H. B. Callen. *Thermodynamics.* John Wiley & Sons, 1960.

[10] S. C. Chay. On quasi-Markov random fields. *Journal of Multivariate Analysis*, 2:14 – 76, 1972.

[11] R. Chellappa. Two-dimensional discrete Gaussian Markov random field models for image processing. In L. N. Kanal and A. Rosenfeld, editors, *Progress in Pattern Recognition 2.* Elsevier Science Publishers B. V. (North Holland), 1985.

[12] J. J. Clark and A. L. Yuille. *Data Fusion for Sensory Informtaion Processing Systems.* Kluwer Academic Publishers, 1990.

[13] J. B. Conway. *A Course in Functional Analysis.* Springer-Verlag, second edition, 1990.

[14] G. R. Cross and A. K. Jain. Markov random field texture models. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 5(1):25–39, 1983.

[15] P. J. Davis. *Circulant Matrices.* Wiley, 1979.

[16] H. Derin and H. Elliott. Modeling and segmentation of noisy and textured images using Gibbs random fields. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 9(1):39–55, 1987.

[17] J. Dieudonné. *Fondements de l'Analyse Moderne.* Gauthiers-Villars, 1972.

[18] R. L. Dobrushin. Description of a random field by means of its conditional probabilities and conditions of its regularity. *Theory of Probability and its Applications*, 13:197–224, 1968.

[19] R. C. Dubes and A. K. Jain. Random field models in image analysis. *Journal of Applied Statistics*, 16(2):131 – 163, 1989.

[20] R. Durbin, R. Szeliski, and A. Yuille. An analysis of the elastic net approach to the traveling salesman problem. *Neural Computation*, 1:348–358, 1989.

[21] Jr. E. B. Gamble. *Integration of Early Visual Cues for Recognition.* PhD thesis, MIT, 1990.

[22] I. M. Elfadel and R. W. Picard. Miscibility matrices explain the behavior of grayscale textures generated by Gibbs random fields. *SPIE Proceedings, Intelligent Robots and Computer Vision IX*, 1381:524–535, 1990.

[23] I. M. Elfadel and A. L. Yuille. Mean-field theory for grayscale texture synthesis using Gibbs random fields. In *SPIE Proceedings, Stochastic and Neural Methods in Signal Processing, Image Processing and Computer Vision*, volume 1569, pages 248–259, San Diego, CA, 1991.

[24] A. Erdelyi. *Asymptotic Expansions.* Dover Publications, Inc., 1956.

[25] B. Julesz *et al.* Inability of humans to discriminate between textures that agree on second-order statistics-revisited. *Perception*, 2:391 – 405, 1973.

[26] A. Feinstein. *Foundations of Information Theory.* McGraw-Hill, 1958.

[27] D. J. Field. Relations between the statistics of natural images and the response of cortical cells. *Journal of the Optical Society of America A*, 4(12):2379 – 2394, 1987.

[28] D. Geiger and F. Girosi. Parallel and deterministic algorithms from MRF's: Surface reconstruction. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13(5):401–412, 1991.

[29] D. Geiger and A. Yuille. A common framework for image segmentation. *Int. J. Computer Vision*, 6:227 – 253, 1991.

[30] D. Geman. *Random Fields and Inverse Problems in Imaging.* Department of Mathematics and Statistics, University of Massachusetts, 1989. preprint.

[31] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 6(6):721–741, 1984.

[32] H. O. Georgii. *Gibbs Measures and Phase Transitions.* Walter de Gruyter, 1988.

[33] J. G. Harris. *Analog models for early vision.* PhD thesis, California Institute of Technology, Pasadena, CA, 1991. Dept. of Computation and Neural Systems.

[34] M. Hassner and J. Sklansky. The use of Markov random fields as models of texture. *Computer Graphics and Image Processing*, 12:357–370, 1980.

[35] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proc. Nat'l Acad. Sci., USA*, 79:2554–2558, 1982.

[36] J. J. Hopfield. Neurons with graded responses have collective computational properties like those of two-state neurons. *Proc. Nat'l Acad. Sci., USA*, 81:3088–3092, 1984.

[37] R. Hu and M. M. Fahmy. Texture segmentation based on a hierarchical Markov random field model. *Signal Processing*, 26:285 – 305, 1992.

[38] K. Huang. *Statistical Mechanics.* Wiley, 1963.

[39] A. K. Jain. Advances in mathematical models for image processing. *IEEE Proceedings*, 69(5):502 – 528, 1981.

[40] A. K. Jain. *Fundamentals of Digital Image Processing.* Prentice-Hall, 1989.

[41] N. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4:373 – 395, 1984.

[42] N. Karmarkar. Riemannian geometry underlying interior-point methods for linear programming. *Contemporary Mathematics*, 114:51 – 75, 1990.

[43] R. L. Kashyap. Univariate and multivariate random field models for images. *Computer Graphics and Image Processing*, 12:257 – 270, 1980.

[44] J. J. Kosowsky and A. L. Yuille. The invisible hand algorithm: Solving the assignment problem with statistical physics. TR # 91-1, Harvard Robotics Laboratory, 1991.

[45] D. S. Lalush. Simulation evaluation of Gibbs prior distributions for use in maximum *a posteriori* SPECT reconstruction. *IEEE Transactions on Medical Imaging*, 11(2):267 – 275, 1992.

[46] J. P. LaSalle. *The Stability and Control of Discrete Processes.* Springer-Verlag, 1986.

[47] J. Lazarro, S. Ryckebush, M. Mahowald, and C. Mead. Winner-take-all circuits of $O(n)$ complexity. In D. S. Touretsky, editor, *Advances in Neural Information Processing Systems I*, pages 703 – 711. Morgan Kaufman, 1989.

[48] B. C. Levy. Noncausal estimation of discrete-time Gauss-Markov random fields. In *Proc. Int. Symp. Math. Theory Networks Syst. (MTNS-89)*, pages 13–21, Cambridge, MA, 1990. Birkhauser.

[49] B. C. Levy. Regular and reciprocal multivariate stationary Gaussian processes over **z** are necessarily Markov. *Journal of Mathematical Systems, Estimation, and Control*, 2(2):133–154, 1992.

[50] B. C. Levy, P. Frezza, and A. J. Arthur. Modeling and estimation of discrete-time Gaussian reciprocal processes. *IEEE Trans. Automatic Control*, 35(9):1013–1023, 1990.

[51] T. M. Liggett. *Interacting Particle Systems.* Springer Verlag, 1985.

[52] D. G. Luenberger. *Linear and Nonlinear Programming.* Addison-Wesley, second edition, 1984.

[53] Shang-Keng Ma. *Modern Theory of Critical Phenomena.* Benjamin/Cummings, 1976.

[54] M. Mahowald and T. Delbrück. Cooperative stereo matching using static and dynamic image features. In Carver Mead and Mohammad Ismail, editors, *Analog VLSI Implementation of Neural Systems*, pages 213 – 233. Kluwer Academic Publishers, 1989.

[55] B. B. Mandlebrot and J. W. Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Review*, 10(4):422 – 437, 1968.

[56] C. M. Marcus and R. M. Westervelt. Dynamics of iterated map neural networks. *Phys. Rev. A*, 40(1):501 – 504, 1989.

[57] D. Marr and T. Poggio. Cooperative computation of stereo disparity. *Science*, 194:283 – 287, 1976.

[58] J. L. Marroquin. *Probabilistic Solution of Inverse Problems.* PhD thesis, MIT, 1985.

[59] Carver Mead. *Analog VLSI and Neural Systems.* Addison-Wesley, 1989.

[60] F. C. Moon. *Chaotic Vibrations.* Wiley, 1987.

[61] J. Moura and N. Balram. Recursive structure of noncausal Gauss-Markov fields. *IEEE Trans. Informatin Theory*, 38(2):334 –354, 1992.

[62] G. F. Newell and E. W. Montroll. On the theory of the Ising model of ferromagnetism. *Reviews of Modern Physics*, 25(2):353 – 389, 1953.

[63] L. Onsager. A two-dimensional model with an order-disorder transition. *Physical Review*, 65:117–118, 1944.

[64] J. M. Ortega and W. C. Rheinboldt. *Iterative Solution of Nonlinear Equations in Several Variables.* Academic Press, 1970.

[65] A. Papoulis. *Probability, Random Variables, and Stochastic Processes.* McGraw Hill, 1984.

[66] G. Parisi. *Statistical field theory.* Addison-Wesley, 1988.

[67] C. Peterson. Parallel distributed approaches to combinatorial optimization problems - benchmark studies on tsp. *Neural Computation*, 2(3):261 – 270, 1990.

[68] C. Peterson and B. Söderberg. A new method for mapping optimization problems onto neural networks. *International Journal of Neural Systems*, 1(1):3 – 22, 1989.

[69] R. W. Picard. *Texture Modeling: Temperature Effects on Markov/Gibbs Random Fields.* PhD thesis, MIT, 1991.

[70] R. W. Picard and I. M. Elfadel. On the structure of aura and co-occurrence matrices for the Gibbs texture model. *Journal of Mathematical Imaging and Vision*, 2:5–25, 1992.

[71] R. W. Picard and A. P. Pentland. Markov/Gibbs image modeling: Temperature and texture. In *Proc. SPIE Conf. on Intell. Robots and Comp. Vis.*, pages 15–26, Boston, MA, 1991.

[72] D. K. Pickard. Inference for discrete Markov fields: The simplest nontrivial case. *Journal of the American Statistical Association, Theory and Methods*, 82(397):90–96, 1987.

[73] John Platt. *Constraint Methods for Neural networks and Computer Graphics.* PhD thesis, California Institute of Technology, Pasadena, CA, 1989.

[74] T. Poggio. A theory of how the brain might work. In *Cold Spring Harbor Symposia on Quantitive Biology*, volume LV, pages 899–909, 1990.

[75] T. Poggio and F. Girosi. Networks for approximation and learning. *Proceedings of the IEEE*, 78(9):1481–1497, 1990.

[76] T. Poggio and V. Torre. Ill posed problems and regularization analysis in early vision. AI memo 773, MIT AI Laboratory, 1984.

[77] C. J. Preston. *Gibbs States on Countable Sets.* Cambridge University Press, 1974.

[78] F. Reif. *Fundamentals of Statistical and Thermal Physics.* McGraw-Hill, 1982.

[79] T. Simchony and R. Chellappa. Stochastic and deterministic algorithms for MAP texture segmentation. In *Proc. ICCASS*, volume 2, pages 1120 – 1123, New York, 1988.

[80] P. D. Simic. Statistical mechanics as the underlying theory of "elastic" and "neural" optimization. *Network*, 1:89 – 103, 1990.

[81] B. Simon. A remark on Dobrushin's uniqueness theorem. *Communications in Mathematical Physics*, 68:183 – 185, 1979.

[82] D. Standley and J. Wyatt. Stability criterion for lateral inhibition networks that is robust in the epresence of integrated circuit parasitics. *IEEE Trans. Circuits and Systems*, 36(5):675 – 681, 1989.

[83] G. Strang. *Introduction to Applied Mathematics.* Wellesley-Cambridge Press, 1986.

[84] R. Szeliski. *Bayesian Modelling of Uncertainty in Low-level Vision.* Kluwer Academic Publishers, 1989.

[85] C. J. Thompson. *Mathematical Statistical Mechanics.* Princeton University Press, 1972.

[86] M. Vidyasagar. *Nonlinear Systems Analysis.* Prentice-Hall, 1987.

[87] S. Wiggins. *Introduction to Applied Nonlinear Dynamical Systems and Chaos.* Springer-Verlag, 1990.

[88] Andrew Witkin and Michael Kass. Reaction-diffusion textures. *Computer Graphics*, 25(3), July 1991. Proc. Siggraph 91.

[89] J. W. Woods. Two-dimensional discrete markov fields. *IEEE Trans. Information Theory*, 18:232 – 240, 1972.

[90] J. L. Wyatt. *Lectures on Nonlinear Circuit Theory.* MIT course 6.335, 1989.

[91] L. Younes. Estimation and annealing for gibbsian fields. *Ann. Inst. Henri Poincaré*, 24:269–294, 1988.

[92] A. L. Yuille. Energy functions for early vision and analog networks. AI memo # 987, MIT, 1987.

[93] A. L. Yuille. Generalized deformable models, statistical physics, and matching problems. *Neural Computation*, 2:1–24, 1990.


وَقُلْ رَبِّ زِدْنِي عِلْمًا