Perfect Simulation of Spatial Processes

Publicações Matemáticas

Perfect Simulation of Spatial Processes

Nancy L. Garcia
UNICAMP

impa

23º Colóquio Brasileiro de Matemática

Impresso no Brasil / Printed in Brazil

Capa: Noni Geiger

23º Colóquio Brasileiro de Matemática

- Aspectos de Modelagem Matemática em Dinâmica dos Fluidos – André Nachbin
- Cálculo Variacional e Controle Ótimo – Antonio Leitão
- Computational Methods in the Local Theory of Curves – Abramo Hefez e Marcelo Escudeiro Hernandes
- Contornos, Conjuntos Convexos e Autômatos Celulares – André Toom
- Inteiros Quadráticos e o Grupo de Classes – Antonio J. Engler e Paulo Brumatti
- Introduction to Toric Varieties – Jean-Paul Brasselet
- Introdução aos Espaços de Escala (EDPs em Processamento de Imagens) – Ralph Teixeira
- Índices de Campos Holomorfos e Aplicações – Marcio G. Soares e Rogério S. Mol
- Notes on Morse Theory – Daniel V. Tausk, Francesco Mercuri e Paolo Piccione
- One Dimensional Dynamics: the Mathematical Tools – Edson de Faria e Welington de Melo
- Partial Regularity of Solutions of the 3-D Incompressible Navier-Stokes Equations – Hermano Frid e Mikhail Perepelitsa
- **Perfect Simulation of Spatial Processes – Nancy L. Garcia**
- Riemannian and Submanifold Geometry – Carlos Olmos
- Tópicos em Combinatória Contemporânea – Carlos Gustavo Moreira e Yoshiharu Kohayakawa
- Uma Introdução Sucinta a Algoritmos de Aproximação – Cristina G. Fernandes et al

Distribuição:
IMPA
Estrada Dona Castorina, 110
22460-320 Rio de Janeiro, RJ
e-mail: dic@impa.br
http://www.impa.br

ISBN: 85-244-0182-6

Preface

Simulating a spatial point process is not an easy task except for the Poisson point process. Usually Monte Carlo Markov Chain methods have been used to generate samples from spatial point processes, for a detailed review see Møller (1999). One of the most common approaches is to identify the point process as the invariant measure of a spatial birth and death chain and run the corresponding chain for a long time until the distribution of the chain is close to equilibrium, see Kelly and Ripley (1976), Preston (1977), Møller (1989), Baddeley and Møller (1989), Clifford and Nicholls (1994) for examples. The problem here is to assess how long the chain should run in order to achieve the desired approximation. In finite state Markov chains, this is related to mixing times and cut-off phenomena [see, Aldous and Fill (1999)]. For spatial point processes, the state space is uncountable and usually these techniques cannot be applied.

However, after the pioneer work of Propp and Wilson (1996) we can reach a much more ambitious goal: to simulate perfectly from the invariant distribution. In the last few years there have been an enormous amount of papers written on the subject ranging from suggestions of practical (and not so practical) methods of achieving such goal to applications of such methods to specific problems. This work presents a review of some of the schemes used to perfect sample from spatial processes. As examples, the area-interaction point process, Strauss process, penetrable sphere model, Peierls contours of the Ising model and continuous loss networks are studied under three proposed algorithms in the literature.

The outline of this book is as follows. Chapter 1 describes several point processes of interest beginning with the Poisson process which will be the basis for all of the other processes studied. Chapter 2 describes a graphical construction for the birth-and-death processes with desired invariant measure that can be used as a theoretical tool to prove probabilistic properties and also used as a basis for a simulation scheme. Chapter 3 describes three methods for simulating point processes, applying them to the processes described in Chapter 1.

At this point I would like to point out that this text do not intend to be an exhaustive account of the perfect simulation schemes applied to all possible spatial point processes. The objective of this work is to introduce the subject and to give a flavor of the techniques and some nice examples of applications. An interested reader should check periodically the site http://www.dimacs.rutgers.edu/~dbwilson/exact where one can find an enormous amount of information related to perfect simulation.

My knowledge about the subject was constructed by many interesting and enlightening discussions with several colleagues and coauthors. In particular, Tom Kurtz introduced me to the world of spatial point processes that can be obtained from Poisson point processes. To work with Pablo Ferrari and Roberto Fernández during the last few years have been a privilege and I have learned with them a great amount of things, one among many was perfect simulation. Of course, they are the ones to blame for any errors in these notes!

I would like to thank the organizers of the 23o. Colóquio Brasileiro de Matemática for the opportunity to teach this course and the Statistics Department of University of California - Berkeley for the hospitality, these notes are based on a series of invited lectures given there when I was a visiting scholar. My thanks to Nevena Marić who wrote the Matlab programs used for the pictures of the simulation procedure and to Laura Ramos for reading a preliminary version of the notes.

Contents

1 Point processes 1

1.1 A little bit of history 1

1.2 Poisson point processes 4

1.3 Marked Poisson processes 7

1.4 Interacting spatial processes 8

1.4.1 Area-interaction point processes 9

1.4.2 Strauss Processes 10

1.4.3 Penetrable spheres mixture model 11

1.4.4 Simulation procedures 11

1.5 Statistical mechanics models 12

1.5.1 Low-temperature Ising model 13

1.6 Loss networks with fixed routing 17

1.6.1 Continuous unbounded one-dimensional loss network . 18

2 Spatial birth-and-death processes 21

2.1 Spatial birth-and-death processes 21

2.1.1 Finite-volume construction 23

2.1.2 Infinite-volume construction 26

2.1.3 Time ergodicity 29

2.1.4 Time convergence and uniqueness 31

2.2 More general birth-and-death processes 33

2.2.1 Peierls contours of the Ising model 33

2.2.2 Loss networks 35

3 Perfect Simulation **37**
3.1 General definition 39
3.2 Coupling from the past 40
3.2.1 Discrete-time 41
3.2.2 Jump processes in continuous-time 46
3.2.3 Dominated CFTP 47
3.2.4 Point processes 48
3.2.5 User Impatience Bias. 50
3.3 Fill's interruptible algorithm 51
3.3.1 Application to attractive spin systems 54
3.3.2 Point processes 55
3.4 Backward-Forward Algorithm 58
3.4.1 Spatial point processes 58
3.4.2 Peierls contours of the ferromagnetic Ising model . . . 62
3.5 Concluding remarks 65

Chapter 1

Point processes

1.1 A little bit of history

The origins of the theory of point processes are certainly ancient, since way back in the past man have been counting stars in regions of the sky, recording natural events such as floodings, earthquakes and appearances of comets. In more recent times, point processes have been used in life tables, counting problems, particle physics, population processes, communication engineering, etc.

The theory of point processes, probably had its beginning with the study of intervals between occurrences of events. In particular, the renewal processes can be viewed as the study of sequences of intervals between successive changes of components subject to failures that are substituted for new ones as soon as a failure occurs. This subject was broadly studied in the 30's and reached its peak with the work of Feller, Smith among others in the period following the Second World War. However, its origin can be seen in early periods in demography problems, insurance and mortality tables. The first life table appeared in the work of John Graunt *Observations on the London Bills of Mortality* in 1662. This work is considered a mark in statistics history, notice that the exchange of letters between Pascal and Fermat occurred

in 1654 and was published in 1679. Several authors studied these problems related to what is known today as survival analysis, among them, A. de Moivre, Laplace, Euler, R.C. Ellis. In this century, at least three big areas of application were developed:

- queuing theory, more specifically telecommunications problems (Erlang (1909) derived the Poisson distribution for the number of calls in a fixed interval of time);
- Actuarial science using differential and integral equations for populational growth; and
- Reliability.

A basic approach for spatial point processes that can be generalized to higher dimensional spaces is the counting of number of events intervals or regions. An important mark is the work of Poisson (n.d.) that includes the derivation of the Poisson distribution as a binomial limit. The first discussions about counting processes would be Seidel (1876) and Abbé (1879) that deal with the occurrence of thunderstorms and the number of blood cells, respectively, apparently independent of the work of Poisson. Erlang (1909) derived the Poisson distribution for the number of calls in a fixed interval of time. Bateman (1910), a mathematical consultant in the work of Rutherford and Geiger in the counting of α particles, obtained the Poisson distribution as the solution of the system of differential equations:

$$\begin{aligned} p_n'(t) &= -\lambda p_n(t) + p_{n-1}(t), \qquad (n \geq 1) \\ p_0'(t) &= -\lambda p_0(t). \end{aligned}$$

These equations represent the formulation of the problem in terms of pure birth process and they were the first step in the quick development of the theory of birth and death processes in the following two decades (specially, McKendrick (1914), McKendrick (1926) and Yule (1924)). These works preceded the general formulations of birth and death processes as Markov processes in the 30's.

Non-Poissonian models appeared in ecology and other fields where the dispersion is bigger than the mean. Another important contribution in this field was made by Neyman (1939) when studying the distribution of beetle larvae in space considering the distance they appeared from the nest. Several articles developed this theme that also appeared in astronomy problems.

Another area of application of spatial point processes is particle systems and communication engineering. During and after the Second World War there was an explosive growth in theory and applications of stochastic processes. New applications were introduced and existing fields were explored in depth. Moreover, a great development in the theory was used in order to unify basic concepts. In queuing theory, C. Palm (1943) studied fluctuations in the intensity in traffic theory (general arrival flow to telecommunications systems). In his work there is a systematic description of the properties of renewal processes, for instance, the use of the Poisson process as the arrival process in a service station. His notion of regeneration point gave rise to several amazing applications. In Palm's terminology, the Poisson process is characterized as the only process where all points are regeneration points. Also, he conjectured and partially proved that the superposition of a large number of sparse renewal processes, in the limit, converges to a Poisson process. In his work, the term *point processes* was used for the first time. Khinchin (1955) gave a complete and more rigorous proof of Palm's results and extended them in several directions. This work initiated the development of this theory by Russian and east-European probabilists. During this time there was a great leap in the development of theoretical physics and the study of probability was a favorite in the Russian school.

The theory of weak convergence to measures in metric spaces by Prohorov (1956) and others preceed the study of general random measures that is the current view in the study of point processes. A point process models the random distribution of indistinguishable points in some space, for concreteness we take this space to be $\mathbb{R}^d$ or $\mathbb{Z}^d$. We identify a point process N

with the counting measure N given by assigning unit mass to each point, that is, $N(A)$ is the number of points in a set A. With this identification in mind, consider $\mathcal{N}(\mathbb{R}^d)$ be the set of counting measures on $\mathbb{R}^d$. The latter assumption implies that such a process N is determined by the probability distribution of the random variables $N(A)$ = number of points in $A \in \mathcal{B}(\mathbb{R}^d)$, the bounded subsets of $\mathbb{R}^d$. This section is a very short summary of a much more general general discussion given in Daley and Vere-Jones (1988). From now on, unless noted, we are going to consider only *orderly* processes, that is, processes that have at most one point per site.

1.2 Poisson point processes

The Poisson point process is one of the most popular models for counting problems. Besides being a good description of many natural phenomena, it is very simple from the computational point of view. Furthermore, or perhaps relatedly, it is used as a reference measure to define other types of processes. Its general definition is as follows.

Definition 1.2.1 *Let ν be a Radon measure on $\mathbb{R}^d$. A point process N_ν on $\mathbb{R}^d$ is a* Poisson process *with mean measure ν if its state space is $\mathcal{N} = \{N \in \{0,1\}^{\mathbb{R}^d} : N(x) = 1$ for only a countable number of $x \in \mathbb{R}^d\}$, and defining $N_\nu(A) = \int_A N_\nu(dx)$,*

(i) For any disjoint $A_1, A_2, \ldots, A_k \in \mathcal{B}(\mathbb{R}^d)$ the random variables $N_\nu(A_1)$, $N_\nu(A_2), \ldots, N_\nu(A_k)$ are independent, and

(ii) For each $A \in \mathcal{B}(\mathbb{R}^d)$ and $k \geq 0$

$$\mathbb{P}[N_\nu(A) = k] = \frac{e^{-\nu(A)}\nu(A)^k}{k!}. \tag{1.2.2}$$

We can think the process N_ν either as a random counting measure $N_\nu = \sum_{\delta_{\xi_i}}$ or as the random set of points $N_\nu = \{x \in \mathbb{R}^d : N_\nu(x) = 1\} = \{\xi_1, \xi_2, \ldots\}$.

A *κ-homogeneous Poisson process* is a process with $\nu = \kappa m_d$, where κ is a constant and m_d the Lebesgue measure on $\mathbb{R}^d$.

Algorithm 1.2.3 *The simulation of a κ-homogeneous Poisson process is simple:*

- *For each finite window W, generate $R \sim Poisson\Big(\kappa m_d(W)\Big)$;*
- *Given $R = r$ generate $U_1, \ldots, U_r$ independently distributed according to the the uniform distribution in W.*
- *Repeat independently for disjoint windows.*

Poisson processes have the "lack of memory" property. In particular, a Poisson process N conditional on n points of N being sited at $x_1, x_2, \ldots, x_n$ has the properties of $N + \delta_{\{x_1, x_2, \ldots, x_n\}}$. Thus, for all purposes the process "forgets" where it had the n points and behaves as it were N with the n points adjoined. The notion of conditioning in this case is not straightforward since the the event "having a point at x" has probability zero, either we may condition on a formal way in terms of Palm probabilities and Palm distributions (see Karr (1986), section 1.7 or Daley and Vere-Jones (1988), section 12.1) or we may make direct use of (i), for example Garcia (1995) showed

Proposition 1.2.4 *Let N be a Poisson process on $\mathbb{R}^d$ with mean measure μ. Let $c : \mathbb{R}^d \to [0, \infty)$, $\rho : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ and $\phi : [0, \infty) \to [0, \infty)$ be Borel measurable functions. Then*

$$\mathbb{E}[\int_{\mathbb{R}^d} c(z)\phi(\int_{\mathbb{R}^d} \rho(x, z)N(dx))N(dz)] =$$
$$= \int_{\mathbb{R}^d} c(z)\mathbb{E}[\phi(\rho(z, z) + \int_{\mathbb{R}^d - \{z\}} \rho(x, z)N(dx))]\mu(dz)$$

More general Poisson processes in which ν is absolutely continuous with respect to the Lebesgue measure in $\mathbb{R}^d$ with density w, can be simulated

using the *projection method* introduced by Kurtz (1989) and described by Garcia (1995). Consider the set

$$C_w = \left\{(x,s); x \in \mathbb{R}^d, s \in \mathbb{R}, 0 \leq s \leq w(x)\right\}, \tag{1.2.5}$$

and the Poisson process $N_{m_{d+1}}$ on $\mathbb{R}^{d+1}$ with Lebesgue mean measure m_{d+1}. Then the process N_w on $\mathbb{R}^d$ defined by

$$N_w(A) = N_{m_{d+1}}\Big(C_w \cap (A \times \mathbb{R})\Big) \tag{1.2.6}$$

is Poisson with mean ν. In words, it is enough to simulate $N_{m_{d+1}}$ as above, and then take the points that lie in C_w and project them onto $\mathbb{R}^d$. More generally, this scheme can be used for Poisson processes whose measure ν has the form

$$\nu(A) = m_{d+1}\Big(C \cap (A \times \mathbb{R})\Big) \tag{1.2.7}$$

for some $C \in \mathbb{R}^{d+1}$.

Doubly stochastic point processes or Cox processes If μ is a random Radon-measure on $\mathbb{R}^d$ independent of N_{m_d+1} such that

$$\mu(B) = \int_B w(x)\,dx.$$

then, denoting N_μ the process (1.2.6),

$$\mathbb{E}[N_\mu(A)] = \mu(A).$$

That is, conditional on μ, N_μ is a Poisson process with mean measure μ.

Finite total rate. For future purposes we consider the case $\nu(\mathbb{R}^d \times \mathbb{R}^+) < \infty$; we interpret the last coordinate as time. One can compute the distribution of the (not necessarily finite) time τ_1, the smaller time-coordinate of the points (if any) of the process. Indeed, calling N the point Poisson process with rate ν, for $0 \leq t \leq \infty$,

$$\mathbb{P}(\tau_1 > t) = \mathbb{P}(N(\mathbb{R}^d \times [0,t)) = 0) = \exp(-\nu(\mathbb{R}^d \times [0,t))). \tag{1.2.8}$$

In the case of one-dimensional processes ($d = 0$) the above reads

$$\mathbb{P}(\tau_1 > t) \;=\; \mathbb{P}(N([0,t)) = 0) \;=\; \exp(-\nu[0,t)) \,. \tag{1.2.9}$$

1.3 Marked Poisson processes

Sometimes it is convenient to allow each point of the process to have a mark belonging to a set $\mathcal{M}$. That is, a *marked point process* is a point process M on $\mathbb{R}^d \times \mathcal{M}$ such that the marginal process of locations $M(\cdot \times \mathcal{M})$ is a point process on $\mathbb{R}^d$.

Notice that not all point processes on a product space are marked point processes, for example a κ-homogeneous Poisson process on $\mathbb{R}^2$ cannot be represented as a marked point process on $\mathbb{R} \times \mathbb{R}$.

An important example is the *completely independent marked point process*. Let N be a marked point process on $\mathbb{R}^d \times \mathcal{M}$ with the property that the n random variables of the set

$$\{N(A_i \times B_i) : \text{ bounded } A_i \in \mathcal{B}_{\mathbb{R}^d}, B_i \in \mathcal{B}_{\mathcal{M}},\ i = 1, 2, \ldots, n\} \tag{1.3.1}$$

are mutually independent whenever A_i are disjoint. It is easy to see Daley and Vere-Jones (1988) that a marked point process with the complete independence property is fully specified by two components:
(i) a Poisson process of locations $N(\cdot \times \mathcal{M})$; and
(ii) a family of probability distributions $\{P(\cdot \mid x), x \in \mathbb{R}^d\}$ giving the distribution of the mark in $\mathcal{M}$.

A very important example of a completely independent marked point process is the so called Boolean model. Let N be a κ-homogeneous Poisson point process in $\mathbb{R}^d$, represent it by the location of its points as

$$N = \{\xi_1, \xi_2, \ldots\}. \tag{1.3.2}$$

Let $S_1, S_2, \ldots$ be a collection of independent $\mathcal{B}_{\mathbb{R}^d}$-valued random variables. That is, S_i is a random Borel set on $\mathbb{R}^d$ and construct the marked point process

$$M = \{(\xi_1, S_1), (\xi_2, S_2), \ldots\} \tag{1.3.3}$$

or represent it as a coverage process Hall (1988) on $\mathbb{R}^d$ given by

$$\mathcal{C} = \{\xi_i + S_i, i = 1, 2, \ldots\} \tag{1.3.4}$$

where $\xi + S = \{\xi + z; z \in S\}$. Boolean models have the property that the number of sets $C \in \mathcal{C}$ that cover a fixed point $x \in \mathbb{R}^d$ is a Poisson random variable with mean $\kappa\mathbb{E}(\text{vol}(S))$.

Poisson Cluster Processes When dealing with processes that can have more than one occurrence per site, sometimes we can work with a baseline orderly process and a marks that indicate how many occurrences at this site. For example, Neyman (1939) introduced a Poisson cluster process to model particle counts in entomology and bacteriology. This model was later applied to cosmology by Nyeman and Scott (n.d.) and it is also known as the Neyman-Scott process:

1. Locations are realizations of an inhomogeneous Poisson point process with mean measure μ;
2. To each location x associate independent and identically distributed random variables K_x according to a discrete probability distribution;
3. The distribution of the new K_x particles around the original location x are independent and identically distributed according to a d-dimensional density function f;
4. Superimpose the location of the new particles erasing the original locations.

Notice that this description gives an algorithm to simulate Neyman-Scott processes.

1.4 Interacting spatial processes

The independence property characterizes the Poisson process. Most of the applications deal with point processes having interaction between points. In

this work, we are going to consider a particular case, point processes with probability law (restricted to a finite box $\Lambda \subset \mathbb{R}^d$) which are absolutely continuous with respect to the probability law of a homogeneous Poisson point process. In fact, if we call μ_Λ^0 the law of the unit-homogeneous Poisson process, their distribution is characterized by the Radon-Nikodym derivative (or Gibbs measure) given by

$$\mu_\Lambda(dN) = \frac{1}{Z_\Lambda} e^{-H(N,\Lambda)} \mu_\Lambda^0(dN) \tag{1.4.1}$$

where $H(N,\Lambda)$ is the energy function, Z_Λ is a normalizing constant.

The Radon-Nikodym derivative can be thought as a measure of how much more likely are the configuration N in this process than in the Poisson process. That is, the set of possible configurations are the same for the interacting process and the Poisson process. However, their likelihood changes and the RN derivative measures this likelihood.

Questions of interest in the study of these processes are about the existence of limits of these measures as $\Lambda \to \mathbb{R}^d$. That is, is there a well-defined counting measure μ such that $\mu_\Lambda \to \mu$? In what sense? How rapidly? How to simulate from these infinite-volume measures?

1.4.1 Area-interaction point processes

In these processes, introduced by Baddeley and van Lieshout (1995), each point (=*germ*) has associated a *grain* formed by a copy of a fixed compact (and usually convex) set $G \subset \mathbb{R}^d$. The intersections of these grains determine a weight that corrects the otherwise Poissonian distribution of the germs. In this case, the Gibbs measure (1.4.1) is given by

$$\mu_\Lambda(dN) = \frac{\kappa^{N(\Lambda)}\, \phi^{-m_d(N \oplus G)}}{Z_\Lambda(\kappa,\phi)} \mu_\Lambda^0(dN)\,, \tag{1.4.2}$$

where κ and ϕ are positive parameters, $Z_\Lambda(\kappa,\phi)$ is a normalizing constant and $N \oplus G$ is the *coverage process* given by

$$N \oplus G := \bigcup_{x \in N} \{x + G\}\,. \tag{1.4.3}$$

Note that when N is Poisson process, this coverage process is a Boolean model. Hence the area-interacting process defined by (1.4.2) can be thought as a "weighted Boolean model" with weights depending exponentially on the area of the covered region.

The parameter ϕ controls the area-interaction between the points of N: the process is *attractive* if $\phi > 1$ and *repulsive* otherwise. If $\phi = 1$ the process is just the (unweighted) Boolean model with grain G and Poissonian rate κ. The case $\phi > 1$ is related to the *penetrable sphere model* introduced by Widow and Rowlinson (1970) and described in Section 1.4.3. The case of *area-exclusion* corresponds to a suitable limit $\phi \to 0$.

Baddeley and van Lieshout (1995) established basic existence and extension properties. For dimensions $d \geq 2$ and ϕ sufficiently small there is a phase transition (Lebowitz and Gallavotti, 1971; Ruelle, 1971), in the sense that limits of bounded-window distributions lead to several infinite-volume measures, depending on the boundary conditions chosen. Purely probabilistic literature focus rather on free boundary conditions, in which case there is a unique, well defined infinite-volume process.

1.4.2 Strauss Processes

A related process to the area-interaction point is the so-called Strauss process. In this case, the unit Poisson process is weighted according to an exponential of the number of pairs of points closer than a fixed threshold r. In this case, the Gibbs measure (1.4.1) is defined by

$$\mu_\Lambda(dN) = \frac{1}{Z_\Lambda} e^{\beta_1 N(\Lambda) + \beta_2 S(N,\Lambda)} \mu^0_\Lambda(dN) \tag{1.4.4}$$

where $S(N, \Lambda)$ is the number of unordered pairs such that $\|x_i - x_j\| < r$. The case $\beta_2 > 0$ was introduced by Strauss (1975) to model the clustering of Californian red wood seedlings around older stumps, however in this case (1.4.4) is not integrable, see Kelly and Ripley (1976).

1.4.3 Penetrable spheres mixture model

The penetrable sphere model was introduced by Widow and Rowlinson (1970) to study liquid-vapor phase transitions. It is a point process with two types of points, therefore it can be seen as a bi-dimensional point process (N, M) in the product space $\mathcal{N} \times \mathcal{N}$ which is absolutely continuous with respect to the product of two independent unit Poisson processes and Radon-Nikodym derivative given by

$$\tilde{\mu}_\Lambda(dN, dM) = \frac{1}{Z_\Lambda}\beta_1^{N(\Lambda)}\beta_2^{M(\Lambda)}1_{\{d(N,M)>R\}}(\mu_\Lambda^0 \times \mu_\Lambda^0)(dN, dM) \qquad (1.4.5)$$

where $d(N, M) = \min\{d(x, y); x \in N, y \in M\}$ is the shortest distance between a point of N and M. That is, in this model points of different type cannot be at a distance shorter than R.

Marginal and conditional distributions:

It is easy to see that the conditional distribution of N given M is a homogeneous Poisson process with intensity β_1 on $\Lambda \setminus (M \oplus G)$. Where G is a sphere of radius R. Similarly, the conditional distribution of M given N is a homogeneous Poisson process with intensity β_2 on $\Lambda \setminus (N \oplus G)$.

The marginal distribution of N is an area-interaction point process with $\kappa = \beta_1$ and $\phi = e^{\beta_2}$. Similarly, the marginal distribution of M is an area-interaction point process with $\kappa = \beta_2$ and $\phi = e^{\beta_1}$.

1.4.4 Simulation procedures

The measures defined by (1.4.2), (1.4.4) and (1.4.5) cannot be simulated so easily as in the Poissonian case. On the one hand, disjoint regions are no longer independent, due to the coverage, and, on the other hand, the normalizations Z_W are difficult to estimate. The usual approach is to obtain them as the invariant measure of a spatial birth-and-death process as discussed below (Section 2). Whereas for the measure defined by (1.4.5) we can use the fact

that the conditional distributions are homogeneous Poisson processes which are easy to simulate.

1.5 Statistical mechanics models

Spin systems, on a finite set $\Lambda \subset \mathbb{Z}^d$, model random configurations $\sigma \in \{-1,+1\}^\Lambda$. We can identify σ with a point process on Λ viewing

$$\sigma(B) = \sum_{i \in B} \mathbf{1}\{\sigma(i) = 1\}.$$

A Gibbs measure in this case, is the distribution of the system in equilibrium

$$\mu_\Lambda(\sigma) = \frac{1}{Z_\Lambda} e^{-H(\sigma,\Lambda)/kT} \tag{1.5.1}$$

where $H(\sigma,\Lambda)$ is the energy function, T is the absolute temperature and k is the Boltzmann constant. Usually, $\beta = 1/kT$ is used in (1.5.1).

We can partially order the set of configurations by declaring $\sigma \leq \tau$ whenever $\sigma(i) \leq \tau(i)$ for all $i \in \Lambda$ and we say that μ_Λ is *attractive* if the conditional probability of $\sigma(i) = +1$ is an increasing function of $\sigma(j)$ for $j \neq i$. Formally, fix $i \in V$ given $\sigma \in \{+1,-1\}^V$ and define $\sigma_+^{(i)}$ and $\sigma_-^{(i)}$ as

$$\sigma_\pm^{(i)}(j) = \begin{cases} \pm 1, & j = i; \\ \sigma(j), & j \neq i \end{cases}$$

We say that μ_Λ is *monotone* if

$$\frac{\mu_\Lambda(\sigma_-^{(i)})}{\mu_\Lambda(\sigma_+^{(i)})} \geq \frac{\mu_\Lambda(\tau_-^{(i)})}{\mu_\Lambda(\tau_+^{(i)})}$$

or, equivalently,

$$\mu_\Lambda(\sigma_-^{(i)})\mu_\Lambda(\tau_+^{(i)}) \geq \mu_\Lambda(\sigma_+^{(i)})\mu_\Lambda(\tau_-^{(i)})$$

for all configurations $\sigma \leq \tau$ and all $i \in V$.

There are several dynamics that have μ_Λ as their invariant measure, for monotone measures one of the most used is the *heat-bath algorithm*, which is a

procedure that visits all sites (deterministically or randomly that guarantees an infinite number of visits per site almost surely) and updates the value at site i according to the conditional probability for μ_Λ. For example, the uniformly random heat-bath algorithm can be update at time n as

$$\phi(\sigma, U_n, V_n) = \begin{cases} \sigma_-^{(V_n)}, & \text{if } U_n < \mu_\Lambda(\sigma_-^{(V_n)})/(\mu_\Lambda(\sigma_+^{(V_n)}) + \mu_\Lambda(\sigma_-^{(V_n)})), \\ \sigma_+^{(V_n)}, & \text{if } U_n \geq \mu_\Lambda(\sigma_-^{(V_n)})/(\mu_\Lambda(\sigma_+^{(V_n)}) + \mu_\Lambda(\sigma_-^{(V_n)})), \end{cases} \tag{1.5.2}$$

where $\{U_n, n \geq 1\}$ are independently and identically distributed $U(0,1)$ random variables, $\{V_n, n \geq 1\}$ are independently and identically distributed $U\{\Lambda\}$ and σ is the configuration at time $n-1$.

1.5.1 Low-temperature Ising model

The Ising model is a particular case of spin systems where

$$H(\sigma, \Lambda) = -\sum_{i<j\in\Lambda} \alpha_{ij}\sigma(i)\sigma(j) - \sum_{i\in\Lambda} B_i\sigma(i) \tag{1.5.3}$$

where B_i is the strength of the external field at site i and α_{ij} models the interaction strength between sites i and j. We are going to concentrate in the case where there is no external field ($B_i = 0$ for all i) and $\alpha_{ij} = 1$ if i and j are neighbors, that is $|i-j| = 1$ and $\alpha_{ij} = 0$ otherwise, for a more complete discussion see Liggett (1985). In this case, a very important question is about existence of phase transition.

It is immediate to see that when Λ is finite the Gibbs measure (state) given by

$$\mu_\Lambda(\sigma) = \frac{1}{Z_\Lambda} e^{-\beta \sum_{i,j\in\Lambda, |i-j|=1} \sigma(i)\sigma(j)} \tag{1.5.4}$$

where β is (up to constant) the inverse of the temperature, is well defined. However, when we want to consider the Ising model on a countable space S, we cannot apply (1.5.4) directly with Λ replaced by S. In this case, the Gibbs state is defined for configurations on finite subsets $\Lambda \subset S$ and then pass to

the limit. Formally, let $\Lambda \subset S$ be a finite set, $\Lambda^c = S \setminus \Lambda$ and $\xi \in \{-1,+1\}^{\Lambda^c}$. Let $\mu_{\Lambda,\xi}$ be the probability measure on $\{-1,+1\}^{\Lambda}$ given by

$$\mu_{\Lambda,\xi}(\sigma) = \frac{1}{Z_{\Lambda,\xi}} \exp\Big\{-\beta\{\sum_{i,j\in\Lambda,|i-j|=1} \sigma(i)\sigma(j) + \sum_{i\in\Lambda,j\in\Lambda^c,|i-j|=1} \sigma(i)\xi(j)\}\Big\} \tag{1.5.5}$$

where again $Z_{\Lambda,\xi}$ is a normalizing constant. In this case, $\mu_{\Lambda,\xi}$ is called the *G*ibbs state with boundary condition ξ (which can depend on Λ). Let

$$\mathcal{G} = \{\mu; \mu \text{ is a weak limit of } \mu_{\Lambda,\xi} \text{ for any } \Lambda \to S \text{ and boundary conditions } \xi\}. \tag{1.5.6}$$

Notice that definition (1.5.6) only makes sense if all μ_Λ are defined in the same probability space, but this can be accomplished by extending the $\mu_{\Lambda,\xi}$ to measures in the full space $\{-1,+1\}^S$ in the natural way, namely acting as the delta measure on ξ for events outside Λ.

Definition 1.5.7 *We say that the model exhibits a phase transition if $\mathcal{G}$ contains more than one element.*

Let $\mu_{\Lambda,+}$ and $\mu_{\Lambda,-}$ be defined by (1.5.5) with $\xi \equiv +1$ and $\xi \equiv -1$ respectively. It can be proved that

(i) $\mu_+ = \lim_{\Lambda \nearrow S} \mu_{\Lambda,+}$ and $\mu_- = \lim_{\Lambda \nearrow S} \mu_{\Lambda,-}$ exist;

(ii) phase transition occurs if, and only if, $\mu_+ \neq \mu_-$;

(iii) phase transition does not occur if, and only if, for all $x \in S$

$$\mu_+\{\sigma; \sigma(x) = +1\} = \mu_-\{\sigma; \sigma(x) = +1\}; \tag{1.5.8}$$

(iv) for all $x \in S$,

$$\mu_+\{\sigma; \sigma(x) = +1\} + \mu_-\{\sigma; \sigma(x) = +1\} = 1. \tag{1.5.9}$$

For the case, $S = \mathbb{Z}^d$ $d \geq 2$, there is phase transition for β sufficiently large. The proof of this affirmation can be found in Liggett (1985) and it uses

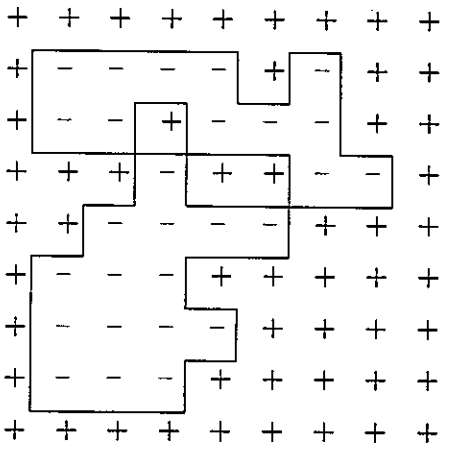

Figure 1.5.9: Peierls contours for 2-dimensional Ising model with positive boundary conditions in $\Lambda = \{1, \ldots, 9\}^2$

the well-known Peierls contours which allows to map the measures $\mu_{\Lambda,+}$ and $\mu_{\Lambda,-}$ of the ferromagnetic Ising model at low temperature into an ensemble of objects —the contours— interacting only by perimeter-exclusion. See, for instance, Section 5B of Dobrushin (1996), for a concise and rigorous account of this mapping. Contours are hyper-surfaces formed by a finite number of $(d-1)$-dimensional unit cubes —*links* for $d = 2$, *plaquettes* for higher dimensions— centered at points of $\mathbb{Z}^d$ and perpendicular to the edges of the dual lattice $\mathbb{Z}^d + (\frac{1}{2}, \cdots, \frac{1}{2})$. To formalize their definition, let us call two plaquettes *adjacent* if they share a $(d-2)$-dimensional face. A set of plaquettes, γ, is *connected* if for any two plaquettes in γ there exists a sequence of adjacent plaquettes in γ joining them. The set γ is *closed* if every $(d-2)$-dimensional face is covered by an even number of plaquettes in γ. *Contours* are connected and closed sets of plaquettes. For example, in two dimensions contours are closed polygonals. In this work we are going to study the construction and simulation of the measure μ_+ for values of β where there exists phase-transition.

Two contours γ and θ are said *compatible* and denoted by $\gamma \sim \theta$ if no plaquette of γ is adjacent to a plaquette of θ. In two dimensions, therefore, contours are compatible if and only if they do not share the endpoint of a link. In three dimensions two compatible contours can share vertices, but not sides of plaquettes. Ising spin configurations in a bounded region with "+" (or "−") boundary condition are in one-to-one correspondence with families of pairwise compatible contours.

The probability distribution for the set of contours in a bounded window Λ is absolutely continuous respect to the counting measure, and assigns to each configurations $\xi \in \{0,1\}^{\mathbf{G}(\Lambda)}$ ($\mathbf{G}(\Lambda)$ is the set of all possible contours inside Λ) probability weight given by

$$\mu_\Lambda(\xi) \;=\; \frac{1}{Z_\Lambda}\, e^{-\beta \sum_{\gamma:\xi(\gamma)=1} |\gamma|} \tag{1.5.10}$$

where $|\gamma|$ is the number of plaquettes comprising γ.[For simplicity we are absorbing in β a factor of 2]. Notice that

$$\mu_\Lambda(\xi) = \frac{1}{Z_\Lambda} \Big(\prod_{\gamma,\theta \in \xi} \mathbf{1}(\gamma \sim \theta) \Big)\, \mu^0_\Lambda(\xi) \tag{1.5.11}$$

(here Z_Λ is not necessarily the same as in (1.5.10)) where μ^0_Λ is the product of Poisson random variables with mean $e^{-\beta|\gamma|}$ for $\gamma \in \mathbf{G}(\Lambda)$. The state-space of μ_Λ is contained in $\{0,1\}^{\mathbf{G}}$, the one of μ^0_Λ is contained in $\mathbb{N}^{\mathbf{G}}$, where $\mathbf{G}$ is the set of all possible contours.

Contour ensembles can therefore be considered extreme cases of perimeter-interacting point processes (in a discretized space). They are extreme on two counts: (1) they involve perimeter-repulsion, i.e. a limit $\phi \to 0$, and (2) they correspond to an infinite (but countable) family of grains of arbitrarily large size.

An important issue for these contour ensembles is the extension of (1.5.10) to a well defined infinite-volume process. Traditionally this problem has been tackled via cluster expansions. Our alternative approach, besides yielding the perfect simulation scheme discussed in Section 3.4, allowed us to prove that

such an extension is possible and unique as long as

$$\alpha := \sup_{\gamma} \frac{1}{|\gamma|} \sum_{\theta \not\sim \gamma} |\theta| \, e^{-\beta|\theta|} < 1. \tag{1.5.12}$$

Fernández, Ferrari and Garcia (1999). This is a weaker condition than the one obtained by usual expansions (for instance in Lebowitz and Mazel (1998)). A condition of this sort is unavoidable because the contour description certainly can not remain valid at the Ising critical temperature, in fact at high enough dimension it breaks down at a temperature strictly below criticality Aizenman, Bricmont and Lebowitz (1987).

1.6 Loss networks with fixed routing

A loss network models, for instance, the occurrence of calls in a communication network. The network is formed by a countable family of links (e.g. $\mathbb{Z}^d$), and each link j comprises a number $C_j \in \mathbb{Z}_+$ of circuits joining its boundaries. A call is characterized by a route γ and a holding period. There is a countable family, G, of routes, each one defined by the numbers $A_{j\gamma}$ of circuits used from each link j. Calls requesting a route γ arrive as Poisson streams of rate $w(\gamma)$ and as γ varies it indexes independent Poisson streams. The call is lost if on any link j there are fewer than $A_{j\gamma}$ circuits free. Otherwise, the call is connected and simultaneously holds $A_{j\gamma}$ circuits from each link j for the holding period of the call. Holding periods are independent, and independent of earlier arrival times and holding periods. For a survey on loss networks, see Kelly (1991). The state space of a loss network is either $\{0,1\}^{\mathbf{G}}$ if no more than one call per route is allowed or $\mathbb{N}^{\mathbf{G}}$ when more than one call can use the same route at the same time. The notation $\xi_t(\gamma)$ indicates the number of calls occupying route γ at time t.

This type of loss networks is labelled *fixed-routing*, as opposed to the *alternative-routing* networks in which calls that are blocked on a route can search for another route. The latter are not addressed in the following. The main issues in the theory of loss networks are:

(i) To establish conditions granting the existence of the (infinite-volume) process, for finite time-intervals. This means, conditions precluding the occurrence of an "explosion" in a finite time.

(ii) To establish conditions for the existence of the process(es) for unbounded time-intervals, that is, conditions ensuring that the process has well defined limits for $t \to \pm\infty$.

(iii) To prove existence of invariant measures and to determine the regime in which there is uniqueness.

(iv) To analyze the properties of this(ese) measure(s), for instance mixing properties, finite-volume corrections and validity of the central limit theorem.

(v) To determine, or find bounds for, the speed of convergence to the invariant measure.

In Section 2 we will see the results of Fernández *et al.* (1998, 1999) which give an answer to all these questions. In fact, the technique of analysis employed there leads to a perfect simulation scheme (Section 3.4).

1.6.1 Continuous unbounded one-dimensional loss network

A natural generalization of the preceding setting is to consider *continuous* loss networks, that is, networks with routes in a continuous, as opposed to discrete, space. The simplest of these models is the loss network in $\mathbb{R}$ introduced by Kelly (1991). Callers of this network are arranged along an infinitely long cable and each call between two points s_1, $s_2 \in \mathbb{R}$ on the cable involves just the segment between them. The cable has the capacity to carry simultaneously up to C calls past any point along its length. Hence, a call attempt between s_1 and $s_2 \in \mathbb{R}$, $s_1 < s_2$, is lost if past any point of the interval $[s_1, s_2]$ the cable is already carrying C calls. Calls are attempted with

initial (leftmost) point following a homogeneous space-time Poisson process with rate density (i.e. time rate per unit length) κ, and (space) lengths given by a distribution π, independent of its leftmost point, with finite mean ρ. The holding time of a call has exponential distribution with mean one. The location of a call, its length and its duration are independent.

Ferrari and Garcia (1998) used a continuous (non-oriented) percolation argument to prove that this model has a unique invariant measure whenever π has finite third moment and the arrival rate κ is sufficiently small. The argument also shows that the process is ergodic, that is, converges to a unique invariant measure whatever the initial distribution.

Let $N = \{\xi_1, \xi_2, \ldots\}$ be a κ-homogeneous Poisson process on $\mathbb{R} \times [0, \infty)$, $H_1, H_2, \ldots$ be i.i.d. random variables with $\mathbb{E}[H_1] = 1$ and $W_1, W_2, \ldots$ i.i.d. random variables with common distribution π, the variables H's, W's and the Poisson process being independent. Consider the random rectangles

$$R_i = \{(x, y); \xi_{i1} \leq x \leq \xi_{i1} + W_i, \xi_{i2} \leq y \leq \xi_{i2} + H_i\},$$

then $\{R_i, i \geq 1\} = \{\xi_i + S_i, i \geq 1\}$ is a boolean model in the continuum $\mathbb{R} \times [0, \infty)$ Hall (1988, p.43) where $S_i = [0, W_i] \times [0, H_i]$ and it represents the independent process of attempted calls. Fix $(x, y) \in \mathbb{R} \times [0, \infty)$, then

$$\mathbb{P}((x, y) \text{ is not covered}) = \mathbb{P}(\text{ for all } i, (x, y) \notin R_i) = e^{-\kappa\mu^{-1}}.$$

Let $v(S_i)$ the content of the largest sphere contained in S_i and $V(S_i)$ the content of the smallest sphere containing S_i, then

$$v(S_i) = \pi(\frac{W_i}{2} \wedge \frac{H_i}{2})^2 \quad \text{and} \quad V(S_i) = \pi(\frac{W_i}{2} \vee \frac{H_i}{2})^2$$

since $\mathbb{E}[v(S_i)] > 0$, if we assume $\mathbb{E}[(\frac{W_i}{2} \vee \frac{H_i}{2})^3] < \infty$, there exists a critical value κ_c such that there is no continuum percolation Hall (1988, Theorem 4.11). That is, the number of rectangles in each clump is finite with probability 1. In this case, each $(s, t) \in \mathbb{R} \times [0, \infty)$ belongs to a finite number of random rectangles R_i and the loss network process can be constructed from the independent process by "erasing" the rectangles which lead to more

than C calls. When H corresponds to an exponential distribution it has all moments. Hence, for this argument it is sufficient to ask that the random variable W with distribution π have a third moment finite. A refinement of the above argument shows that this condition can be weakened to $2\kappa\rho < 1$. This construction is heart of the graphical representation of spatial birth and death processes defined in Section 2 that leads to the backward-forward perfect simulation scheme.

This invariant measure can be considered a generalized point process where the germs are the leftmost points of a call and the grains are randomly chosen segments. It can be generalized to any dimension, just substituting the segments by curves or arbitrary bounded sets.

Chapter 2

Spatial birth-and-death processes

The common feature linking all the spatial processes described in Section (1.2) is that *all these distributions can be realized as invariant measures of spatial interacting birth-and-death processes.*

2.1 Spatial birth-and-death processes

Non-spatial birth and death processes are continuous time Markov chains with $\{0, 1, 2, \ldots\}$ as state space and transition probabilities that are positive only to neighbors (see, for example,Feller (1968) for an overview). Preston (1977) introduced a birth and death process which takes into account the position of the individuals. He defines the process as a continuous time jump process with state space that contains all possible configurations of individuals. In our case, we are interested in point processes that are specified through a density (Radon Nikodym derivative) with respect to a unit Poisson

point process, that is,

$$\mu_\Lambda(dN) = \frac{1}{Z_\Lambda} e^{-H(N,\Lambda)} \mu_\Lambda^0(dN) \tag{2.1.1}$$

where $H(N,\Lambda)$ is the energy function, Z_Λ is a normalizing constant.

Therefore, the state space for this process is $\mathcal{S} = \{\mathbf{n}; H(\mathbf{n},\Lambda) < +\infty\}$, that is the set of configurations with positive density. Ripley (1977) showed that such measure μ_Λ is the invariant measure of a spatial birth and death process. We specify this process in terms of a nonnegative functions $\lambda : \mathbb{R}^d \times \mathcal{N}(\mathbb{R}^d) \to [0,\infty)$ and $\delta : \mathbb{R}^d \times \mathcal{N}(\mathbb{R}^d) \to [0,\infty)$. The meaning of λ is that if the point configuration at time t is $\mathbf{n} \in \mathcal{N}(\mathbb{R}^d)$, then the probability that a point is added to the configuration in a neighborhood of the point x having area ΔA in the next interval of length Δt is approximately $\lambda(x,\mathbf{n})\Delta A \Delta t$. The interpretation for δ is similar, except that, a point can only be deleted from the configuration if already present, that is, if the point configuration at time t is $\mathbf{n} \in \mathcal{N}(\mathbb{R}^d)$, then the probability that a point $x \in \mathbf{n}$ is deleted from the configuration in the next interval of length Δt is approximately $\delta(x,\mathbf{n})\Delta t$. In fact, there is more than one process that has the same invariant measure, we can choose λ and δ in such way that they satisfy the detailed balance condition:

$$\lambda(x,\mathbf{n}) e^{-H(\mathbf{n},\Lambda)} = \delta(x,\mathbf{n}) e^{-H(\mathbf{n}\cup\{x\},\Lambda)} \qquad \text{if } \mathbf{n}\cup\{x\} \in \mathcal{S}. \tag{2.1.2}$$

Notice that equation (2.1.2) says that any pair of birth and death rates such that

$$\frac{\lambda(x,\mathbf{n})}{\delta(x,\mathbf{n})} = \exp\{-H(\mathbf{n}\cup\{x\},\Lambda) + H(\mathbf{n},\Lambda)\} \tag{2.1.3}$$

will give raise to a process with invariant measure having the density given by (2.1.1). We can always take $\delta(x,\mathbf{n}) = 1$, that is, whenever a point is added to the configuration it lives an exponential amount of time independently of the configuration of the process.

The variable-birth and constant-death process has generator given by

$$Af(\eta) = \int (f(\eta+\delta_x) - f(\eta))\lambda(x,\eta)dx + \int (f(\eta-\delta_x) - f(\eta))\eta(dx) \tag{2.1.4}$$

for "suitable" functions f.

Consider first a finite range birth rate, that is, there exists a compact set G such that if $\eta_1 = \eta_2$ inside $x + G$ then $\lambda(x, \eta_1) = \lambda(x, \eta_2)$ satisfying

$$\bar{\lambda} = \sup_{x,\eta} \lambda(x, \eta) < \infty. \tag{2.1.5}$$

We are going to use a oriented percolation argument to show the existence of the (infinite-volume) process, for finite-time intervals. More restrictive conditions in $\bar{\lambda}$ will ensure ergodicity and exponential rate of convergence.

Graphical construction

In order to get a graphical construction for the process with generator (2.1.4), we begin with a $\bar{\lambda}$-homogeneous Poisson point process on $\mathbb{R}^d \times \mathbb{R}$. Denote it by $N = \{(\xi_1, T_1), (\xi_2, T_2), \ldots\}$. For each point (ξ_i, T_i), associate two independent marks $S_i \sim \exp(1)$ and $Z_i \sim U(0, 1)$.

We can see the marked point process $\mathrm{C} = \{(\xi_i, T_i, S_i, Z_i), i = 1, 2, \ldots\}$ as the graphical representation of a birth and death process with constant birth rate $\bar{\lambda}$ and constant death rate 1 (call this free process α) and Z_i will be used as the indicator of "allowed" births.

From now on, a marked point (ξ_i, T_i, S_i, Z_i) will be identified with a marked *cylinder* $((\xi_i + G) \times [T_i, T_i + S_i), Z_i)$ with *basis* ξ_i, *birth time* T_i, *lifetime* S_i and *flag* Z_i. Calling $C = (\xi, t, s, z)$, we use the notation

$$\mathrm{Basis}\,(C) = \xi, \quad \mathrm{Birth}\,(C) = t, \quad \mathrm{Life}\,(C) = [t, t + s], \quad \mathrm{Flag}\,(C) = z. \tag{2.1.6}$$

Define incompatibility between cylinders C and C' by

$$C' \not\sim C \quad \text{if and only if} \quad \mathrm{Basis}\,(C)+G \cap \mathrm{Basis}\,(C')+G \neq \emptyset \quad \text{and} \quad \mathrm{Life}\,(C) \cap \mathrm{Life}\,(C') \neq \emptyset, \tag{2.1.7}$$

otherwise $C' \sim C$ (compatible).

2.1.1 Finite-volume construction

The construction of the spatial birth and death process in a finite box Λ with an initial configuration $\eta_0 = \{\varphi_1, \varphi_2, \ldots\}$ using the Poisson processes

is straightforward. We use only the finite set $\{(\xi_i, T_i, S_i, Z_i) : \xi_i \in \Lambda, T_i > 0\}$. Let $\mathbf{C}^\Lambda = \{C \in \mathbf{C} : \text{Basis}\,(C) \cap \Lambda \neq \emptyset, \text{Birth}\,(C) > 0\}$. To each point φ_j present in the initial configuration η_0 we independently associate an exponential time $\tilde{S}_j$ and a cylinder $(\varphi_j, 0, \tilde{S}_j, 0)$. The collection of initial cylinders is called $\mathbf{C}_0^\Lambda$. We realize the dynamics η_t^Λ as a (deterministic) function of $\mathbf{C}^\Lambda$ and $\mathbf{C}_0^\Lambda$. Let

$$\mathbf{C}^\Lambda[0,t] = \{(\xi, s, l, z) \in \mathbf{C}^\Lambda \cup \mathbf{C}_0^\Lambda; 0 \leq s + l, s \leq t\}. \tag{2.1.8}$$

Consider $0 < t_1 < t_2 < \ldots < t_N$ as the birth and death marks T_i, S_i lying in the set $[0,t]$. By the properties of the Poisson process all of these times are distinct.

$$\{t_1, \ldots, t_N\} = [0,t] \cap \{s, s+l; (\xi, s, l, z) \in \mathbf{C}^\Lambda[0,t]\}. \tag{2.1.9}$$

We construct the process η_t^Λ inductively as follows:

FV.1. Suppose that η_u^Λ is already defined, and that $t_{i-1} \leq u < t_i$. We set

$$\eta_s^\Lambda = \eta_u^\Lambda, \quad \text{for all } u \leq s < t_i. \tag{2.1.10}$$

If $u \geq t_N$ then

$$\eta_s^\Lambda = \eta_u^\Lambda, \quad \text{for all } u \leq s. \tag{2.1.11}$$

FV.2. If t_i is a death time, that is, $t_i = s + l$ for some $(\xi, s, l, z) \in \mathbf{C}^\Lambda[0,t]$ then we delete the point ξ: we set

$$\eta_{t_i}^\Lambda = \eta_{t_i-}^\Lambda \setminus \{\xi\}. \tag{2.1.12}$$

Go back to **F.V.1**.

FV.3. If t_i is a birth time, that is, $t_i = s$ for some $(\xi, s, l, z) \in \mathbf{C}^\Lambda[0,t]$ then we do not add the point ξ if $z > \lambda(\xi_i, \eta_{t_i-}^\Lambda)$: we set

$$\eta_{t_i}^\Lambda = \eta_{t_i-}^\Lambda. \tag{2.1.13}$$

Otherwise the point ξ is added: we set

$$\eta_{t_i}^\Lambda = \eta_{t_i-}^\Lambda \cup \{\xi\}. \tag{2.1.14}$$

In either case, go back to **F.V.1**.

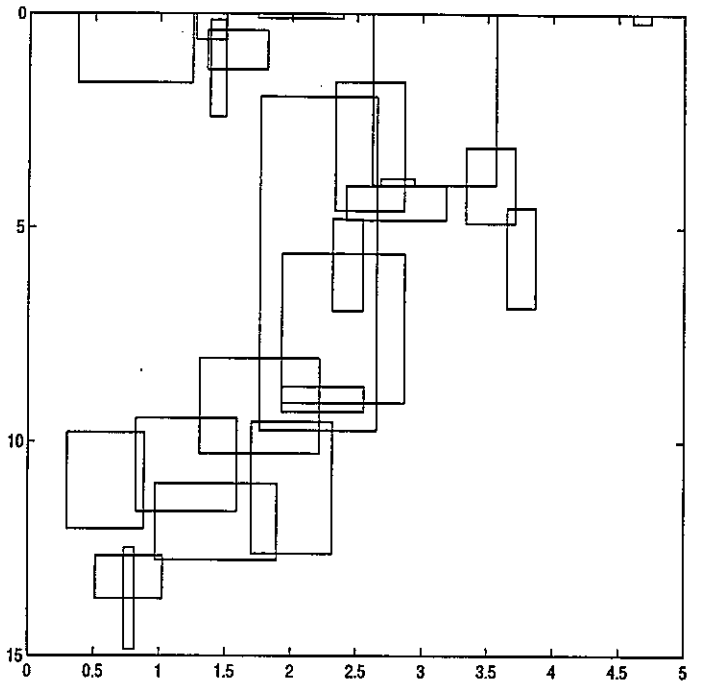

Figure 2.1.9: Independent process

It is tedious but easy to show that η_t^Λ has generator A^Λ defined as in (2.1.4) restricting the sums to the configurations contained in Λ. It is easy to find an invariant measure μ^Λ for this process (through the equation $\int A\mu^\Lambda(\eta) = 0$). Some regeneration argument should show that η_t^Λ converges in distribution to μ^Λ for any initial configuration η. This, in particular, implies that μ^Λ is the unique invariant measure for η_t^Λ.

Using the same Poisson marks for η_t^Λ and α_t (the process with constant birth rate $\bar{\lambda}$ and constant death rate 1), we have

$$\eta_t^\Lambda(A) \leq \alpha_t(A), \tag{2.1.15}$$

for all $A \subset \Lambda$ because in the process α_t all cylinders are kept. This implies

$$\mu^\Lambda\{\eta : \eta(A) = 0\} \geq \mathbb{P}\{\alpha : \alpha(A) = 0\}. \tag{2.1.16}$$

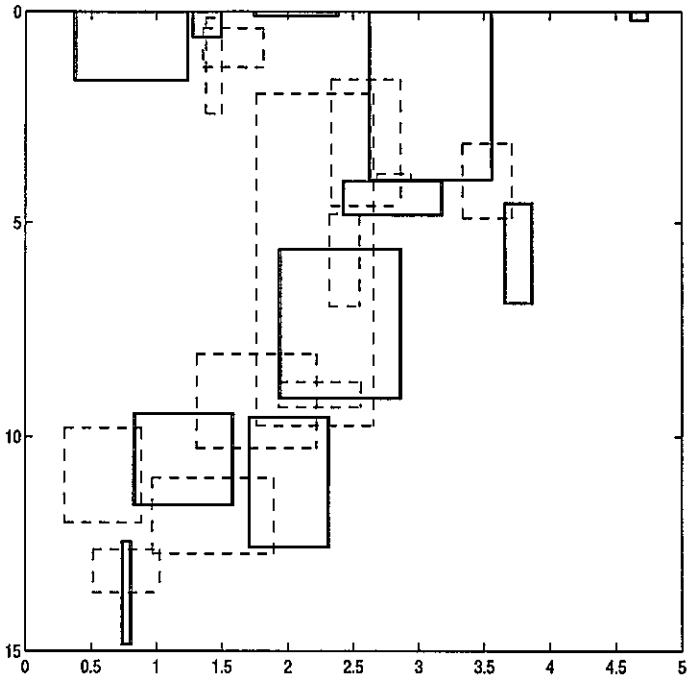

Figure 2.1.14: Thinned process

2.1.2 Infinite-volume construction

If we try to perform an analogous construction in infinite volume we are confronted with the problem that there is not a first mark. However, notice that the free process α_t always exist. The goal, is to find conditions under which the process α_t can be thinned by the finite-volume construction.

From Figure 2.1.14 it is easy to see that the presence of a cylinder C at time t depends only on the cylinders that were born before it, are alive at the time of birth of C and that intersects C. This is loosely speaking the definition of the *first generation of ancestors* of C. The second generation is formed by the ancestors of the first generations. Recursively we can construct the *clan* of ancestors. Therefore, for each cylinder of the free process present at time t, if we can go backwards in time up to time 0 and find only finitely

many cylinders we have to check then the process at time t is well-defined.

To be precise, define the total order $\prec$ in the set of cylinders induced by the birth times. That is $C \prec C'$ if and only if $\text{Birth}\,(C) \leq \text{Birth}\,(C')$.

For an arbitrary space-time point (x,t) define the set

$$\mathbf{A}_1^{x,t} = \{C \in \mathbf{C}\,;\, x \in \text{Basis}\,(C) + G,\ \text{Life}\,(C) \ni t\} \tag{2.1.17}$$

the set of cylinders containing the point (x,t).

For any cylinder C define the set of *ancestors* of C as the set

$$\mathbf{A}_1^C = \{C' \in \mathbf{C}\,;\, C' \prec C\,;\, C' \not\sim C\} \tag{2.1.18}$$

Notice that the definition of ancestor does not depend on the lifetime of C. Recursively for $n \geq 1$, the nth generation of ancestors are defined as

$$\mathbf{A}_n^{x,t} = \{C'' : C'' \in \mathbf{A}_1^{C'} \text{ for some } C' \in \mathbf{A}_{n-1}^{x,t}\}. \tag{2.1.19}$$

and for a given cylinder C,

$$\mathbf{A}_n^C = \{C'' : C'' \in \mathbf{A}_1^{C'} \text{ for some } C' \in \mathbf{A}_{n-1}^C\}. \tag{2.1.20}$$

We say that there is *backward oriented percolation* in $\mathbf{C}$ if there exists a space-time point (x,t) such that $\mathbf{A}_n^{x,t} \neq \emptyset$ for all n, that is, there exists a point with infinitely many generations of ancestors. Let the *clan* of the space-time point (x,t) be the union of its ancestors:

$$\mathbf{A}^{x,t} = \bigcup_{n \geq 1} \mathbf{A}_n^{x,t} \tag{2.1.21}$$

and $\mathbf{C}[0,t] = \{C \in \mathbf{C} : \text{Birth}\,(C) \in [0,t]\}$.

In the next theorem we give a sufficient condition for the existence of the infinite-volume process in any finite time interval in terms of backwards percolation.

Theorem 2.1.22 *If with probability one* $\mathbf{A}^{x,t} \cap \mathbf{C}[0,t]$ *is finite for any* $x, \in \mathbb{R}^d$ *and* $t \geq 0$, *then for any box* $\Lambda \subset \mathbb{R}^d$, *the process with generator* A^Λ *is well defined and has at least one invariant measure* μ^Λ.

Proof. We construct the process for $\Lambda = \mathbb{R}^d$. The construction for other Λ is analogous. The initial distribution is denoted $\eta_0 = \{\varphi_1, \varphi_2, \ldots\}$. For each $\varphi_j \in \eta_0$ let S_j be an independent exponentially distributed random time of mean 1. The time S_j represents the lifetime of the cylinder with basis φ_j, birth time 0 and flag 0. We call $\mathbf{C}(0)$ the set of cylinders $\{(\varphi_j, 0, S_j, 0); \varphi_j \in \eta_0\}$. Since the cylinders in $\mathbf{C}(0)$ have no ancestors in $\mathbf{C}[0,t]$, under the hypothesis of the theorem, every cylinder in $\mathbf{C}(0) \cup \mathbf{C}[0,t]$ has a finite number of ancestors in $\mathbf{C}[0,t]$.

It is easy to see that we can represent the constant birth and death process α_t as

$$\alpha_t = \{\text{Basis}\,(C); C \in \mathbf{C}(0) \cup \mathbf{C}[0,t], \text{Life}\,(C) \ni t\}. \tag{2.1.23}$$

We will construct η_t as a thinning of the constant birth and death process α_t, for this consider $C \in \mathbf{C}(0) \cup \mathbf{C}[0,t]$ such that Life $(C) \ni t$, then Basis $(C) \in \alpha_t$, to decide if Basis $(C) \in \eta_t$ we need to look at $\mathbf{C}(0) \cup \cup_{n \geq 1} \mathbf{A}_n^C$, which is finite by hypothesis. In this case, we can perform the mark-by-mark construction defined in the previous section can be accomplished in a finite number of steps.

It is possible to show that η_t has generator A given by (2.1.4). □

Invariant measure

The usual approach to find invariant measures for Markov processes is to construct the process beginning at $-\infty$ in a given configuration and cut the process at time 0. If, the process is independent of the initial configuration, the process at time 0 has invariant measure.

In the graphical construction above, the lack of percolation allows us to construct η_t as a thinning of α_t for times in the whole real line. Since the construction is time-translation invariant, the distribution of η_t will be invariant.

Theorem 2.1.24 *If with probability one there is no backwards oriented percolation in* $\mathbf{C}$*, then the process with generator* A *can be constructed in* $(-\infty, \infty)$

in such a way that the marginal distribution of η_t is invariant.

Definition 2.1.25 *The distribution of η_t is called μ.*

Remark: A consequence of Theorem 2.1.24 is that the spatial birth and death process in a finite box Λ can be constructed for all $t \in \mathbb{R}$.

2.1.3 Time ergodicity

The main theorem in this section shows that all that is needed for exponential convergence to a unique invariant measure is the absence of backwards and non-oriented percolation. The key to prove this fact is a domination by a branching process which will be sub-critical under the condition $\bar{\lambda} \leq 1/m^d(G)$.

Note that the collection of cylinders

$$\mathcal{C} = \{(\text{Basis}\,(C) + G) \times \text{Life}\,(C); C \in \mathbf{C}\} \tag{2.1.26}$$

is a boolean model Hall (1988) and for any $x \in \mathbb{R}^d$ and $t \geq 0$ we have that the number of hypercubes that cover (x,t) is Poisson distributed with mean $m^d(G)\bar{\lambda}$. In fact,

$$\begin{aligned} \mathbb{P}((x,t) \quad \text{not covered}) &= \mathbb{P}((x,t) \notin ((\text{Basis}\,(C) + G) \times \text{Life}\,(C) \quad \text{for any } C \in \mathbf{C}) \\ &= e^{-m^d(G)\bar{\lambda}}. \end{aligned} \tag{2.1.27}$$

Notice that the process of ancestors is not a Galton-Watson process, a cylinder in the first generation can also be in the second generation of a cylinder. However, we can define a Galton-Watson branching process $B_n \in \mathbb{N}$ such that the offspring distribution of a cylinder C has the same (marginal) law as the distribution of $\mathbf{A}_1^C$, but the branches behave independently. The key point is to fix a way to distribute common ancestors. Let Y_i^n be i.i.d.

non negative integer valued random variables with Poisson distribution with mean $m^d(G)\bar{\lambda}$. Define $B_0 = 1$ and

$$B_{n+1} = \sum_{i=1}^{B_n} Y_i^n \tag{2.1.28}$$

(with the convention $\sum_{i=1}^{0} Y_i^n = 0$). It is possible to couple the BO-cluster $\mathbf{A}^{x,t}$ and $(B_n)_{n\geq 0}$ in such a way that the number of ancestors in the nth generation of (x,t) is less than or equal to B_n. The total number of ancestors of (x,t) is bounded by

$$\|\mathbf{A}^{x,t}\| \leq \sum_{n\geq 0} B_n. \tag{2.1.29}$$

Therefore, there is no backward oriented percolation if the process is subcritical, that is,

$$\bar{\lambda} < (m^d(G))^{-1}. \tag{2.1.30}$$

Defining the time-length and the space-width of the family of cylinders $\mathbf{A}^{x,t}$ be respectively

$$\mathrm{TL}\,(\mathbf{A}^{x,t}) = t - \sup\{s : \mathrm{Life}\,(C) \ni s, \text{ for some } C \in \mathbf{A}^{x,t}\}, \tag{2.1.31}$$

$$\mathrm{SW}\,(\mathbf{A}^{x,t}) = |\cup_{C\in\mathbf{A}^{x,t}} \mathrm{Basis}\,(C) + G|, \tag{2.1.32}$$

we get

$$\mathrm{SW}\,(A^{x,0}) \leq m^d(G)B \tag{2.1.33}$$

$$\mathrm{TL}\,(A^{x,0}) \leq \sum_{i=1}^{B} \tilde{S}_i \tag{2.1.34}$$

where

$$B = \sum_{n\geq 0} B_n \tag{2.1.35}$$

and $\tilde{S}_i, i \geq 1$ are i.i.d. exponentially distributed random variables with mean 1.

Since

$$\mathbb{E}[B] = \frac{1}{1 - m^d(G)\bar{\lambda}} \tag{2.1.36}$$

we have

$$\mathbb{E}[\mathrm{SW}\,(A^{x,0})] \leq \frac{1}{m^d(G)^{-1} - \bar{\lambda}} \tag{2.1.37}$$

$$\mathbb{E}[\mathrm{TL}\,(A^{x,0})] \leq \frac{1}{1 - m^d(G)\bar{\lambda}}. \tag{2.1.38}$$

Moreover, the moment generating function of $\mathrm{TL}\,(A^{x,0})$ is given by

$$\mathbb{E}[a^{\mathrm{TL}\,(A^{x,0})}] = F[(1 - \log a)^{-1}] \tag{2.1.39}$$

where $F(b)$ is the generating function of Z and consequently,

$$\mathbb{P}[\mathrm{TL}\,(A^{x,0}) > bt] \leq F_B(\bar{b})e^{-bt}. \tag{2.1.40}$$

2.1.4 Time convergence and uniqueness

We say that two sets of cylinders $\mathbf{A}$ and $\mathbf{A}'$ are incompatible if there is a cylinder in $\mathbf{A}$ incompatible with a cylinder in $\mathbf{A}'$:

$$\mathbf{A} \not\sim \mathbf{A}' \text{ if and only if } C \not\sim C' \text{ for some } C \in \mathbf{A} \text{ and } C' \in \mathbf{A}'. \tag{2.1.41}$$

Theorem 2.1.42 *Assume that there is no backwards oriented percolation with probability one. Then,*

1. *Uniqueness. The measure μ is the unique invariant measure for the process η_t.*

2. *Time convergence. For any compact set A,*

$$\lim_{t\to\infty} \sup_A |\mathbb{E}\eta_t^{\eta}(A) - \mathbb{E}\eta(A)| = 0. \tag{2.1.43}$$

Furthermore,

$$\begin{aligned} &\sup_A |\mathbb{E}\eta(A) - \mathbb{E}\eta_t^{\eta}(A)| \\ &\quad \leq \mathbb{P}(\cup_{x\in A}\{\mathbf{A}^{x,t} \not\sim \mathbf{C}(0) \text{ or } \mathrm{TL}\,(\mathbf{A}^{x,t}) > t\}) \qquad (2.1.44) \\ &\quad \leq \Big(\mathbb{P}(\cup_{x\in A}\mathrm{TL}\,(\mathbf{A}^{x,0}) > bt) + e^{-(1-b)t}\,\mathbb{E}(\mathrm{SW}\,(\mathbf{A}^{x,0}))\Big) \qquad (2.1.45) \end{aligned}$$

for any $b \in (0,1)$.

3. *Space convergence. As* $\Lambda \to \mathbb{R}^d$, μ^Λ *converges weakly to* μ. *More precisely, if* A *is a finite set contained in* $\mathbb{R}^d$, *then*

$$|\mu(A) - \mu^\Lambda(A)| \leq \mathbb{P}\Big(\mathbf{A}(A) \neq \mathbf{A}^\Lambda(A))\Big). \tag{2.1.46}$$

Moreover, by construction we have that the invariant measure μ *is space-invariant. That is, we have spatial ergodicity of the stationary distribution.*

Proof. Existence of μ has been proven in Theorem 2.1.24.

In order to prove uniqueness of the invariant measure we use the same Poisson marks to construct simultaneously the stationary process η_t and a process starting at time zero with an arbitrary initial configuration η. The second process is called η_t^η, where $\eta_0^\eta = \eta$. The process η_t^η ignores the cylinders in $\mathbf{C}$ with birth times less than 0 and considers $\mathbf{C}(0) = \{(\varphi_j, 0, S_j, 0) : \varphi_j \in \eta\}$, the set of cylinders with basis given by the initial configuration η and birth time zero —the times S_j are exponentially distributed with mean 1 and independent of everything.

It is enough to prove that

$$\sup_A \mathbb{P}(|\eta_t(A) - \eta_t^\eta(A)| > 0) \to 0 \tag{2.1.47}$$

as $t \to \infty$.

Since we are using $\mathbf{C}$ to construct η_t and $\mathbf{C}[0,t] \cup \mathbf{C}(0)$ to construct η_t^η, it follows

$$|\eta_t^\eta(A) - \eta_t(A)| \leq \sum_{x \in A} \mathbf{1}\Big\{\Big(\mathbf{A}^{x,t} \not\sim \mathbf{C}(0) \text{ or } \mathrm{TL}\,(\mathbf{A}^{x,t}) > t\Big)\Big\} \tag{2.1.48}$$

Note that $\mathbf{A}^{x,t} \neq \emptyset$ for finitely many $x \in A$. The proof of the above results is done similarly as in Fernández, Ferrari and Garcia (1998). The estimates for the moments of $\mathrm{TL}\,(\mathbf{A}^{x,t})$ and $\mathrm{SW}\,(\mathbf{A}^{x,t})$ are given by (2.1.37), (2.1.38) and (2.1.39).

The arguments prove that the process converges, uniformly in the initial configuration, to the invariant measure μ. An immediate consequence is that μ is the unique invariant measure. Moreover, it is easy to see that the velocity of convergence is exponential.

2.2 More general birth-and-death processes

2.2.1 Peierls contours of the Ising model

In this case, we are considering a birth-and-death process where the individuals are contours in G, a birth rate depending on the configuration of contours already present and unit death rate. Specify this process in terms of a non-negative function $\lambda : \mathbf{G} \times \{0,1\}^{\mathbf{G}} \to [0,1]$ given by

$$\lambda(\gamma, \eta) = e^{-\beta|\gamma|} \mathbf{1}(\gamma \sim \eta). \tag{2.2.1}$$

The above process has generator given by

$$Af(\eta) = \sum_{\gamma \in \mathbf{G}} e^{-\beta|\gamma|} \mathbf{1}\{\eta^{+\gamma} \in \{0,1\}^{\mathbf{G}}\}[f(\eta^{+\gamma}) - f(\eta)] + \sum_{\gamma \in \mathbf{G}} \eta(\gamma)[f(\eta^{-\gamma}) - f(\eta)] \tag{2.2.2}$$

where f is a cylindrical function and

$$\eta^{\pm\gamma} = \eta \pm \delta_\gamma. \tag{2.2.3}$$

A construction similar to Section 2 shows that a sufficient condition for existence of the process is

$$\alpha := \sup_\gamma \frac{1}{|\gamma|} \sum_{\theta \not\sim \gamma} |\theta| \, e^{-\beta|\theta|} < \infty \tag{2.2.4}$$

where as ergodicity is obtained under the condition (1.5.12) ($\alpha < 1$). In this case, consider the origin of a contour to be the first point in lexicographic order and define

$$\lambda_\beta := \sum_{\gamma : \mathrm{ori}(\gamma) = 0} e^{-\beta|\gamma|}. \tag{2.2.5}$$

Begin with λ_β-independent Poisson streams $\{N_x; x \in \mathbb{Z}^d\}$. Denote

$$N_x = \{T_1(x), T_2(x), \ldots\} \tag{2.2.6}$$

and to each point $T_i(x)$ assign independent marks:

- $\gamma_i(x) = \gamma$ chosen from $\{\gamma : \mathrm{ori}(\gamma) = x\}$ with probability $\exp(-\beta|\gamma|)/\lambda_\beta$;
- $S_i(x) \sim \exp(1)$.

Consider the set of marked cylinders:

$$\begin{aligned} &\mathrm{C}\{(\gamma_i(x), T_i(x), S_i(x)); x \in \mathbb{Z}^d, i \geq 1\} \\ &\quad \equiv \{\gamma_i(x) \times [T_i(x), T_i(x) + S_i(x)]; x \in \mathbb{Z}^d, i \geq 1\} \end{aligned} \tag{2.2.7}$$

and as before for $C = (\gamma, t, s)$, we use the notation

$$\mathrm{Basis}\,(C) = \gamma, \quad \mathrm{Birth}\,(C) = t, \quad \mathrm{Life}\,(C) = [t, t+s]. \tag{2.2.8}$$

Define incompatibility between cylinders C and C' by

$$C' \not\sim C \quad \text{if and only if} \quad \mathrm{Basis}\,(C) \not\sim \mathrm{Basis}\,(C') \quad \text{and} \quad \mathrm{Life}\,(C) \cap \mathrm{Life}\,(C') \neq \emptyset, \tag{2.2.9}$$

otherwise $C' \sim C$ (compatible), where compatibility between contours was defined in Section 1.5.1.

The construction follows as Section 2.1.1 and 2.1.2, except that here we erase all incompatible cylinders (we do not need to check the flag). Conditions for lack of backward percolation and velocity of convergence are obtained through a domination by a multi-type branching process b_n, where the types are the contours and $b_n^\gamma(\theta)$ denotes the number of cylinders of basis θ in the nth generation of a cylinder C with basis γ. In this case,

$$\sum_\theta b_n^\gamma(\theta) \geq \|A_n^C\|. \tag{2.2.10}$$

It is easy to see that the mean number of descendents type θ from a mother type γ is given by

$$m(\gamma, \theta) = e^{-\beta|\theta|} 1(\gamma \not\sim \theta) \tag{2.2.11}$$

and

$$\sum_\theta m^n(\gamma, \theta) \leq \sum_\theta |\theta| m^n(\gamma, \theta) \leq |\gamma| \alpha^n \tag{2.2.12}$$

where α is defined by (1.5.12) and the process is sub-critical if $\alpha < 1$. Detailed calculations can be found in Fernández *et al.* (1998, 1999).

2.2.2 Loss networks

They can be associated to point processes with grains stochastically chosen (possibly with sizes forming an unbounded set). For instance, for the continuous loss network of Section 1.6.1, the germs are the leftmost points of calls and the grains are segments with random lengths. In this case, domination is done through a multi-type branching process with uncountable many types. Assume, in general, that the leftmost points of calls appear with rate $f(x)$ and that call lengths are given by a distribution π independent of x. We only require the latter to have a finite mean ρ. Consider a germ sitting at the origin, that is a call stretching from the origin to the right, born at time zero. Its ancestors correspond to cylinders with sufficient lifetime and with bases given by either calls starting at negative sites and passing through the origin, or calls of arbitrary length originating within the sites occupied by the initial call. Therefore, the sub-criticality parameter has two contributions:

$$\int_{-\infty}^{0} dx\, f(x)\, \pi(\{L > x\}) \;+\; \int_{0}^{\infty} \pi(dL) \int_{0}^{L} dx\, f(x)\,. \tag{2.2.13}$$

[We have omitted a factor 1 corresponding to the lifetime of the ancestors.] If the rate f is constant, say equal to κ, each contribution in (2.2.13) is equal to $\kappa\rho$. The finite-time process therefore exists as long as $\rho < \infty$ and the ergodic stationary process if

$$2\kappa\rho \;<\; 1\,. \tag{2.2.14}$$

Chapter 3

Perfect Simulation

Perfect simulations or *exact sampling* are labels for a recently developed set of techniques designed to produce output whose distribution is guaranteed to follow a given probability law. These techniques are particularly useful in relation with Markov Chain Monte Carlo, and their range of applicability is rapidly growing (see Green and Murdoch (1999), Section 1.3, and Mira, Møller and Roberts (1999) and Møller and Nicholls (1999) and references therein, or visit the site `http://dimacs.rutgers.edu/~dbwilson/exact`).

Several techniques have been suggested recently in the literature. The outbreak of these subject come with Propp and Wilson (1996) paper where they suggest a practical method of achieving a perfect sample of a Markov chain with finite state space. Their *Coupling from the Past* (CFTP) algorithm have been applied for infinite (or huge) state spaces require a *monotonicity property*: there must exist a "maximal" and a "minimal" states and a coupling such that the coalescence of trajectories starting from these two states imply the coalescence of all other trajectories ("monotone coupling"). Examples of processes with this property include Glauber dynamics of spin systems with the FKG property (Propp and Wilson, 1996) and attractive point processes (Kendall, 1997; Kendall, 1998). In fact, through a minor modification the algorithm is also applicable to repulsive point processes

(Kendall, 1998; Häggström and Nelander, 1998).

One of the biggest problem with CFTP technique is that it has the so called *impatient-user bias*. Fill (1998) introduced a technique, free of the impatient-user bias, based on the well-known rejection algorithm for generating independent random samples. Fill's algorithm (usually called *Interruptible Algorithm* (IA)) applies to Markov processes whose *time-reversed* process has a monotonicity property. Thus its range of applicability overlaps with that of the CFTP algorithm at reversible monotone processes like Glauber dynamics of attractive automata or ferromagnetic spin systems and attractive point processes (Fill, 1998; Thönnes, 1999).

Fernández, Ferrari and Garcia (2000) introduce yet a different perfect simulation scheme, *Backward-Forward Algorithm* (BFA), applicable, in principle, to any process that can be sampled from the invariant measure of a spatial birth and death process and is continuous with respect to a Poisson point process. This perfect simulation scheme allows to simulate the distribution of the *infinite-volume* invariant measure μ in a *finite region*, or window, Λ. For example, area-interaction point processes Baddeley and van Lieshout (1995), Strauss processes (Strauss, 1975), fixed-routing loss networks (Kelly, 1991) and Peierls contours of low-temperature Ising model (Fernández et al., 1998). Some of these processes have been subjected to other perfect simulation methods. For instance, attractive point processes can be simulated using dominated CFTP or IA methods, if the models can be sandwiched between a "maximal" and a "minimal" weighted Boolean models (Fill, 1998; Kendall, 1998; Thönnes, 1999). The CFTP algorithm can be applied to repulsive point processes as well (Kendall, 1998). Nevertheless, these treatments consider processes in a finite window with *fixed* boundary conditions. From the statistical mechanical point of view it is important to consider finite windows of an infinite-volume distribution. The only mention to this is by Kendall (1997), who points out a scheme valid when the underlying Boolean model does not exhibit (unoriented) percolation. In this case, the CFTP method can be extended by looking at $[-T, 0] \times [-K, K]^d$ for ever increasing T and K. The lack of percolation ensures that eventu-

ally the area-interaction process will not be affected by whatever boundary conditions are imposed outside $[-K, K]^d$. On the other hand, Häggström, van Lieshout and Møller (1999) (first appeared in 1996 as a research report) combine ideas from CFTP method with two component Gibbs sampler to deal with infinite area-interaction point processes. This paper together with Kendall (1998) (which also appeared as a technical report in 1996) are among the very first papers on perfect simulation of point processes.

BFA has distinctive features: there is no coupling involved, consequently the scheme is insensitive to the presence of monotonicity; it directly samples a finite window of the equilibrium measure in *infinite-volume* without further limit procedures; it relies on a graphical construction that has the added value of being a proven theoretical tool for the analysis of properties of the target measure (Fernández et al., 1998; Fernández et al., 1999) obtain mixing properties, finite-volume corrections and asymptotic (in temperature) distribution of "defects" of the low-temperature Ising translation-invariant extremal measures.

All of the three algorithms, dominated CFTP, IA and BFA algorithm are based on the construction of the underlying (marked) Poisson process where objects are born.

3.1 General definition

Let us start with an abstract definition embodying the three exact simulation algorithms, Coupling from the Past (CFTP), Fill's Interruptible (IA) and Backward-Forward Algorithm (BFA) described in the Introduction.

Definition 3.1.1 *A perfect simulation (or exact sampling) scheme for a probability space* $(\mathbf{X}, \Omega, \mathcal{F}, \mu)$ *consists in:*

(i) A process $\underline{V} = (V_t)_{t \geq 0}$,

(ii) *A* $\{\mathcal{G}_t\}$*-stopping-time,* $\tau = \tau(\underline{V})$*, where* $\mathcal{G}_t = \sigma(V_s, 0 \leq s \leq t)$*, such that*

$$\mathbb{P}(\tau(\underline{V}) < \infty) = 1\,, \tag{3.1.2}$$

(iii) *A random function* $\Phi_{\underline{V}} : \mathbb{R}_+ \to \mathbf{X}$ *such that*

$$\mathbb{P}(\Phi_{\underline{V}}(\tau) \in A) = \mu(A)\,. \tag{3.1.3}$$

The definition is completely general: $\underline{V}$ is some underlying process and $\mathbb{P}$ is a probability measure defined in a sufficiently large space encompassing $\mathbf{X}$ and the state-space of the process $\underline{V}$. In fact, in CFTP and our algorithm, the set up is such that

$$\mathbb{P}(\Phi_{\underline{V}}(t) \in A) = \mu(A) \quad , \quad \forall t \geq \tau\,. \tag{3.1.4}$$

This is not so in the IA algorithm. Property (3.1.4) stems from the fact that in CFTP and in our case, the algorithm "looks into the past", and the process $\underline{V}$ is related to *past history* or *ancestry* of what happens at a *fixed* time, say time zero. The IA algorithm, instead, is constructed on the basis of the forward evolution but incorporates a time-reversed trajectory for the acceptance-rejection procedure.

3.2 Coupling from the past

The *coupling from the past algorithm* (CFTP) was the first feasible algorithm for a perfect simulation scheme. It was introduced by Propp and Wilson (1996) and after this paper was made available, there was a sequence of articles applying the method to several different situations. We are going to describe the application to some situations ranging from the simplest (discrete-time finite state Markov chain, vertical CFTP) to more complicated problems (area-interaction point process and Strauss processes, dominated CFTP).

3.2.1 Discrete-time

Consider the problem of generating a random sample from a distribution μ on a finite set $\mathcal{S}$ which is the unique invariant measure of a discrete-time aperiodic, irreducible, positive recurrent Markov chain $\{X_t, t \in \mathbb{Z}\}$ with state space $\mathcal{S}$. Let P denote its transition matrix. Their approach can be described as follows: simulate the Markov chain, coupling all the paths beginning from all possible initial states, a predetermined amount of time (from $-T$ to 0), if all paths coalesce at time 0, the coalescent state X_0 has the desired distribution π. If the paths did not coalesce, restart the chain at $-T' < -T$, from all possible initial states, preappending new moves to the old ones. They show that if enough moves are preapended, eventually all the paths will coalesce and the resulting coalescent state X_0 is an unbiased sample from π.

In this case, the ingredients of the algorithm are:

- A discrete-time backwards process defined by a sequence $(U_i)_{i\leq 0}$ of independent random variables uniformly distributed in $[0,1]$. The forward process $\underline{V}$ is simply defined as its time-inversion: $V_i = U_{-i}$.

- An updating function $F : \mathbf{X} \times [0,1] \to \mathbf{X}$ such that the Markov chain constructed by setting $X_n = F(X_{n-1}, V_n)$ has μ as unique invariant measure.

The definition of τ and $\Phi_{\underline{V}}$ is based on iterations of F:

$$F_{[k,k']}(x, \underline{V}) = F(F_{[k,k'-1]}(x, \underline{V}), V_{k'}) \tag{3.2.1}$$

for $k' \geq k$, where $F_{[k,k]}(x, \underline{V}) = F(x, V_k)$. Notice that $F_{[k,k']}(x, \underline{V})$ depends only on $(V_k, \ldots, V_{k'})$. Now,

$$\tau = \min\left\{n : F_{[-n,0]}(x, \underline{V}) \text{ does not depend on } x\right\} \tag{3.2.2}$$

and

$$\Phi_{\underline{V}}(t) = F_{[-t,0]}(x, \underline{V}) = F_{[-\tau,0]}(x, \underline{V}) \tag{3.2.3}$$

for any $x \in \mathbf{X}$ and $t \geq \tau$. For $t < \tau$ the value of $\Phi_{\underline{V}}$ is arbitrary.

In words, the process is simulated from time $-t$ to time 0, using the *same* realization $u_{-t}, \ldots, u_{-1}$ of the random variables $(U_i)_{-t \leq i \leq -1}$ for *all* possible initial states $X_{-t} = x$. If all the resulting trajectories coalesce at or before time 0, the value of X_0 is taken to be a sample of μ. If not, the simulation is started some other time $t' > t$, using, for the period $[-t, 0]$, the previous realization $u_{-t}, \ldots, u_{-1}$ of the independent random variables. This (backwards) iteration is continued until all trajectories are seen to coalesce before time 0. The key points of this prescription are: (i) the use of the same random numbers to generate trajectories for different initial states (coupling), (ii) the keeping of a given realization of random numbers for a given period in all iterations, and (iii) the use of a *fixed* time —called time 0— to register the sample.

The efficiency of the algorithm depends on the choice of the function F. A badly designed coupling can lead to extremely large values of τ. As an example of this, consider a process with $\mathbf{X} = \{0, 1\}$ and with probability $1/2$ of jumping from any state to any other. Here $\mu(0) = \mu(1) = 1/2$. If one chooses $F(0, v) = 1 - F(1, v) = \mathbf{1}\{v < 1/2\}$, the resulting coupling time τ is infinite with probability one. The construction of "good" couplings requires the maximization of $\min_{x,y} \mathbb{P}(F(x, V_n) = F(y, V_n))$. This condition is strongly model-dependent. Every homogeneous (uniform) ergodic Markov process admits an F yielding a finite coalescence time, e.g. the Vaserstein coupling used by Dobrushin (1965) or the construction provided by Nummelin's splitting technique to get a constant $\epsilon > 0$ and a probability measure Q such that $\mathbb{P}(F_{[-1,0]}(x, \underline{V}) \in \cdot) \geq \epsilon Q(\cdot)$. See Foss and Tweedie (1998) for more details.

For completeness, let us see why the algorithm performs as stated. By definition of F,

$$\begin{aligned} \mu(x) &= \lim_{-T \to -\infty} \mathbb{P}\Big(F_{[-T,0]}(a, \underline{V}) = x\Big) \qquad (3.2.4) \\ &= \lim_{-T \to -\infty} \Big[\mathbb{P}\Big(F_{[-T,0]}(a, \underline{V}) = x\,,\, \tau \leq T\Big) + \mathbb{P}\Big(F_{[-T,0]}(a, \underline{V}) = x\,,\, \tau > T\Big)\Big] \end{aligned}$$

and by definition of τ,

$$\mathbb{P}\Big(F_{[-T,0]}(a,\underline{V}) = x\,,\, \tau \le T\Big) \;=\; \mathbb{P}\Big(\Phi_{\underline{V}}(t) = x\Big) \qquad \forall t \ge \tau\,. \tag{3.2.5}$$

On the other hand, if F is well chosen, τ is finite with probability one, hence

$$\mathbb{P}\Big(F_{[-T,0]}(a,\underline{V}) = x\,,\, \tau > T\Big) \;\le\; \mathbb{P}(\tau > T \mid X_{-T} = a) \;=\; \mathbb{P}(\tau > T) \underset{T\to\infty}{\longrightarrow} 0\,. \tag{3.2.6}$$

This shows property (3.1.3).

While it is true that trajectories also coalesce when looked *forward* in time, an algorithm based on this fact does not lead to a perfect simulation scheme. Indeed, if τ^* is the forward coalescing time, the analogous of (3.2.4) holds,

$$\begin{aligned}\mu(x) &= \lim_{T\to\infty} \mathbb{P}(F_{[0,T]}(a,\underline{V}) = x) \\ &= \lim_{T\to\infty} \Big[\mathbb{P}(F_{[0,T]}(a,\underline{V}) = x\,,\, \tau^* \le T) + \mathbb{P}(F_{[0,T]}(a,\underline{V}) = x\,,\, \tau^* > T)\Big]\,,\end{aligned} \tag{3.2.7}$$

but in general

$$\lim_{T\to\infty} \mathbb{P}\Big(F_{[0,T]}(a,\underline{V}) = x\,,\, \tau^* \le T\Big) \;\neq\; \mathbb{P}\Big(\Phi_{\underline{V}}(\tau^*) = x\Big) \tag{3.2.8}$$

(in fact, the limit may not even exist).

Example 3.2.9

Let's consider a very simple example, take the random walk on $\{0,1,2,3,4\}$ with transition probabilities

$$P(i,i+1) = P(i+1,i) = 1/2, \text{ for } i = 1,2,3;$$

$$P(0,0) = P(0,1) = P(4,4) = P(4,3) = 1/2.$$

$$\pi = (1/5, 1/5, 1/5, 1/5, 1/5, 1/5)$$

In this case we have a stochastic flow defined as: if $X_0 = i$ then $X_n = F_{0,n}(i)$

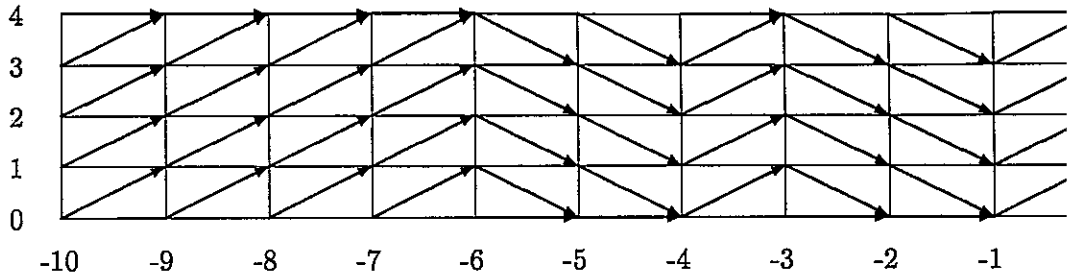

Figure 3.2.9: CFTP for the symmetric random walk in Example 3.2.9

The usual MCMC approach is to run the chain $X_n = F_{0,n}(i)$ for n large. Define

$$Y_n(i) := F_{-n,0}(i) \stackrel{\mathcal{D}}{=} X_n \text{ given } X_0 = i.$$

Notice that $\{Y_n, n \geq 0\}$ is not a Markov chain. Define

$$T_c := \inf\{n; Y_n(i) \text{ do not depend on } i\}.$$

Then, it is easy to see that

$$Y_\infty = \lim_{n\to\infty} Y_n(i) = Y_{T_c}(i) \sim \pi.$$

We have that T_c is *coalescence time* and it is a stopping time in the reverse filtration:

$$T_c = \inf\{n; F_{-n,0} \text{is constant}\}$$

and $F_{-n,0}(i) = Y_\infty$ has the desired distribution π.

In the example, showed in Figure 3.2.9 $T_c = 10$ and $F_{-10,0}(i) = 1$ is an unbiased sample from π.

Notice that if we run the chain forward in time until coupling of all trajectories we have that the only possible coupling states are 0 and 4. In this case,

$$\mathbb{P}[X_{\tau^*} = 0] = \mathbb{P}[X_{\tau^*} = 4] = 1/2.$$

Let $\ldots,, U_{-3}, U_{-2}, U_{-1}, U_0$ be independent identically distributed random variables and $\phi(\cdot,\cdot)$ a deterministic function such that

$$\mathbb{P}[\phi(i, U_0) = j] = P_{i,j}$$

for all $i, j \in \mathcal{S}$. Define

$$F_{m,n} = \phi(\phi(\phi(\ldots, U_{-m}), U_{-m+1}), \ldots, U_{-n}).$$

If, with probability one, there exists a T such that $F_{T,0}$ is constant, which value we denote by $\phi(\ldots, U_{-2}, U_{-1}, U_0)$. Then,

$$\phi(\ldots, U_{-2}, U_{-1}, U_0) \sim \pi.$$

Monotonicity

The CFTP algorithm has been designed for use with finite state space $\mathbf{X}$. The problem is to verify coalescence when the state space is very large or even infinite. Assume that we can update the chain using a monotone rule ϕ:

$$x \leq y \Rightarrow \phi(x, U_0) \leq \phi(y, U_0)$$

almost surely with respect to U_0.

Moreover, assume that $\mathcal{S}$ has maximal $\hat{1}$ and minimal $\hat{0}$ such that $\hat{0} \leq x \leq \hat{1}$ for all $x \in \mathcal{S}$ in some partial ordering. In this case only the extremal states have to be followed. In actual simulations τ is not really computed. In general $F_{[-n,0]}(x, \underline{V})$ is computed for all x and for different —but not all— values of n (for instance powers of 2) up to the first time $F_{[-n,0]}$ is constant in x. Notice that the number of simulation involved is equal to $2(1 + 2 + 4 + \ldots) < 2^{k+2}$, (2 comes from following 2 trajectories). However, this is close to the optimum since T_c has to exceed 2^{k-1}, therefore to verify the coalescence we have to run $2.2^{k-1} = 2^k$ simulations. Consequently, the procedure of doubling until "overshooting" is within a factor of 4 from the best algorithm that prevent "overshooting".

Pseudo-code:

$T \leftarrow 1$
  $U_{-1} \sim U(0,1)$
repeat
  $upper \leftarrow \hat{1}$
  $lower \leftarrow \hat{0}$
  for $n = -T$ to -1
    $upper \leftarrow \phi(upper, U_n)$
    $lower \leftarrow \phi(lower, U_n)$
  $T \leftarrow 2T$
  for $n = -T$ to $-T/2$
    $U_n \sim U(0,1)$
until $upper = lower$
return $upper$

Remark: It is never to much stressing the fact that the same uniform random variables are used in each loop, that is, U_{-1} is going to be generated only once to go from step -1 to 0.

3.2.2 Jump processes in continuous-time

The previous algorithm can be trivially adapted for invariant measures of Markov jump processes with an embedded ergodic Markov chain. Let (Y_t) be a process of this type, with finite state space $\mathbf{X}$, rates $Q(x,y)$, $x, y \in \mathbf{X}$, and (unique) invariant measure μ. It is convenient to consider another process, with rescaled transition times, having the same invariant measure. The new process has transition times given by a Poisson process $N(t)$ of rate $\lambda = \max_x \sum_y Q(x,y)$ and transitions determined by an skeleton Markov chain $\underline{X}$ with transition probabilities

$$P(x,y) = \begin{cases} \lambda^{-1} Q(x,y) & \text{if } y \neq x \\ 1 - \lambda^{-1} \sum_{z \neq x} Q(x,z) & \text{if } y = x \end{cases} . \tag{3.2.10}$$

We have that

$$Y_t \stackrel{\mathcal{D}}{=} X_{N(t)} . \tag{3.2.11}$$

Therefore, the invariant measure for $\underline{X}$ coincides with that of the original process (Y_t). It is therefore enough to proceed as in the discrete-time case.

3.2.3 Dominated CFTP

In the case of unbounded infinite (or very large) state space, it is not possible to use the above described method. Kendall (1998) introduced a modification in Propp and Wilson's algorithm in order to apply it to point processes. The idea, however, is not limited to this case and has been used to generate from continuous unbounded state space (Green and Murdoch, 1999). It is also called horizontal CFTP (Kendall and Møller, 1999) and coupling into and from the past (Wilson, 2000).

The idea is to find another Markov chain $\{C_t, t \in \mathbb{Z}\}$ — chosen in such way that we know how to generate exactly from its invariant measure — that dominates the chain under study. Assume, without loss of generality, that the state space has a minimal state $\hat{0}$, but not a maximal state. In this case, the ingredients of the algorithm are:

- A coupling that guarantees that if for some t we have $C_t \geq X_t$ that the same is true for all subsequent times. Usually this is reached using the same variables $\underline{V}$ to update both chains at the same time. That is, there exist ϕ_1 and ϕ_2 such that $(X_{t+1}, C_{t+1}) = (\phi_1(X_t, U_t), \phi_2(C_t, U_t))$ and if $x \leq c$ we have $\phi_1(x, u) \leq \phi_2(x, u)$ for all $u \in [0, 1]$.

- For any value $x \in \mathcal{S}$ and time $t < 0$, there exists a.s. an $s < t$ such that $X_t \leq C_t$ if $X_u = x$ for $u < s$.

- We can simulate directly from the invariant distribution of C_t.

- Given C_t the conditional distribution of (C_{t-1}, U_{t-1}) is known and we can sample from this distribution. That is, we can simulate C_t into the past. This can be obtained easily if C_t is reversible.

In this case, we can use the CFTP algorithm based on generating an upper process in the same way that vertical CFTP.

3.2.4 Point processes

Kendall (1997, 1998) introduced dominated CFTP algorithm to obtain simulations of processes that can be obtained as weighted Boolean models using quermass integrals. These include the *area-interaction* processes considered by Baddeley and van Lieshout (1995) described in Section 1.4.1, where the point process is produced by the germs of a Boolean model Θ under the weighting

$$\gamma^{-\text{area}(\Theta)}.$$

Its Radon-Nikodym derivative (restricted to a bounded window Λ) is given by (*cf.* (1.4.4))

$$\mu_\Lambda(dN) \;=\; \frac{\kappa^{N(\Lambda)}\,\phi^{-m_d(N\oplus G)}}{Z_\Lambda(\kappa,\phi)}\,\mu^0_\Lambda(dN)\,, \tag{3.2.12}$$

where κ and ϕ are positive parameters, $Z_\Lambda(\kappa,\phi)$ is a normalizing constant and $N\oplus G$ is the *coverage process* given by

$$N\oplus G \;:=\; \bigcup_{x\in N}\{x+G\}\,. \tag{3.2.13}$$

The attractive processes can be simulated using CFTP methods in the presence of monotonicity, when models can be sandwiched between a "maximal" and a "minimal" weighted Boolean models. In fact, through a minor modification the algorithm is also applicable to repulsive point processes Kendall (1997). We describe his scheme.

Consider the space-time Boolean model of cylinders constructed in Section 2.

Finite-volume construction

Now, fix $-T<0$; for $t\in[-T,0]$ we are going to follow the evolution of three processes: $\eta^{\max}_{-T}(t)$, $\eta^{\min}_{-T}(t)$ and $\eta_{-T}(t)$ on Λ. Each process will have initial configuration $\eta^{\min}_{-T}(-T)\subset\eta_{-T}(-T)\subset\eta^{\max}_{-T}(-T)$ and they will use the finite set of marked cylinders

$$\mathcal{C}^\Lambda_{-T}=\{C\in\mathcal{C};\text{Basis}\,(C)\in\Lambda,\text{Life}\,(C)\cap[-T,0]\neq\emptyset\}. \tag{3.2.14}$$

The initial "maximal" and "minimal" configurations are defined by

$$\eta_{-T}^{\max}(-T) = \{\text{Basis}\,(C); C \in \mathcal{C}_{-T}^{\Lambda}, \text{Life}\,(C) \ni -T\} \tag{3.2.15}$$

$$\eta_{-T}^{\min}(-T) = \{\text{Basis}\,(C); C \in \mathcal{C}_{-T}^{\Lambda}, \text{Life}\,(C) \ni -T, \text{Flag}\,(C) \leq \phi^{-m_d(G)}\} \tag{3.2.16}$$

while $\eta_{-T}(-T)$ can be any arbitrary subset of $\eta_{-T}^{\max}(-T)$ and superset of $\eta_{-T}^{\min}(-T)$.

Using the graphical construction defined in Section 2.1.1, it is possible to couple monotonically the trajectories $\eta_{-T}^{\max}(t)$ and $\eta_{-T}^{\min}(t)$ for all $-T \leq t \leq 0$ until the coalescence time

$$T_C := \min\{T; \eta_{-T}^{\max}(0) = \eta_{-T}^{\min}(0)\}. \tag{3.2.17}$$

In general, the algorithm is run for fixed times $T_1 < T_2 < T_3 < \ldots$, with $T_N = e^N$.

The modification needed for the repulsive case, is that in **F.V.3** (Section 2.1.1),

- a point ξ is added to $\eta_{-T}^{\max}(t)$ if the mark

$$z > \phi^{m_d(G) - m_d((\xi+G)\setminus(\eta_{-T}^{\min}(t-)\oplus G))}.$$

- a point ξ is added $\eta_{-T}^{\min}(t)$ if the mark

$$z > \phi^{m_d(G) - m_d((\xi+G)\setminus(\eta_{-T}^{\max}(t-)\oplus G))}.$$

The above references analyze processes in a finite window with *fixed* boundary conditions. On the other hand, it seems interesting to consider finite windows of an infinite-volume distribution. The only mention to this is by Kendall (1997), who points out a scheme that requires that the underlying Boolean model do not exhibit percolation in space-time. In this case, the CFTP method can be extended by looking at $[-T, 0] \times [-K, K]^d$ for ever increasing T and K. The lack of percolation ensures that eventually

the area-interaction process will not be affected by whatever boundary conditions are imposed at time $-T$ and outside $[-K, K]^2$. This lack of percolation argument is the same used in Section 1.6.1 for continuous unbounded loss networks, as before an oriented-percolation argument can lead to a scheme that can be applied to a broad regimen.

3.2.5 User Impatience Bias.

The coupling from the past algorithm possesses the impatient-user bias. That is, it has a running time which is not independent of the state sampled, thus if the user aborts a long run of the algorithm a bias is introduced. The following simple example is presented in Thönnes (1999). Consider the Markov chain X with state space $\{0, 1, 2\}$ and transition matrix

$$P = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \\ 1/2 & 0 & 1/2 \end{pmatrix}.$$

The stationary distribution is given by $\pi = (2/5, 1/5, 2/5)$. We can simulate X using the following rule:

$$\phi(x, U) = \begin{cases} 0, & \text{for } x = 0, 2 \text{ and } U \leq 1/2 \\ \min\{x+1, 2\}, & \text{for } x = 0, 2 \text{ and } U > 1/2 \\ 2, & \text{for } x = 1 \end{cases}$$

where U is a $U(0, 1)$ random variable. Notice that this rule is not monotone and to run CFTP we need to follow the simulation for all 3 states.

Denote by

$$\phi(x, U_0, U_1, \ldots, U_T) = \phi(\phi(\ldots\phi(\phi(x, U_0), U_1), \ldots), U_T) \tag{3.2.18}$$

the state of the chain at time 0 starting the chain at time $-T$ in state x.

Suppose now that the user always terminates a run of CFTP after I iterations without obtaining coalescence. In this case, if coalescence is attained at time T we are sampling from the distribution defined by

$$\mathbb{P}(\phi(x, U_1, \ldots, U_T) \in A \,|\, T < I) \tag{3.2.19}$$

which is different from π since

$$\mathbb{P}(\phi(x, U_1, \ldots, U_T) = r \,|\, T < I) = \frac{\mathbb{P}(T < I \text{ and } \phi(x, U_1, \ldots, U_T) = r)}{\mathbb{P}(T < I)}. \tag{3.2.20}$$

Let $N = 2^{I-1}$, then by combinatorial arguments we can show that

$$\mathbb{P}(\phi(x, U_1, \ldots, U_T) = r \,|\, T < I) = \frac{\sum_{k=2}^{N} (1/2)^{k-1} [1/2 P_{0r}^{(N-k)} + 1/2 P_{2r}^{(N-k)}]}{1 - (1/2)^{N-1}} \tag{3.2.21}$$

where $P^{(k)}$ is the k-step transition matrix. Therefore, for $I \geq 3$ we have

$$\begin{aligned}
\mathbb{P}(\phi(x, U_1, \ldots, U_T) = 0 \,|\, T < I) &= \frac{2}{5}\Big[\frac{1 - 2^{-N}}{1 - 2^{-N+1}}\Big] \\
\mathbb{P}(\phi(x, U_1, \ldots, U_T) = 1 \,|\, T < I) &= \frac{1}{5}\Big[\frac{1 - 2^{-N}}{1 - 2^{-N+1}}\Big] \\
\mathbb{P}(\phi(x, U_1, \ldots, U_T) = 2 \,|\, T < I) &= 1 - \frac{3}{5}\Big[\frac{1 - 2^{-N}}{1 - 2^{-N+1}}\Big].
\end{aligned}$$

Although the bias decreases with I as expected, the sample will always be biased.

3.3 Fill's interruptible algorithm

Fill (1998) describes a perfect sampling scheme based on a rejection sampling method, which protects against the user impatience bias. Consider a Markov chain on a partially ordered finite state space $(\mathcal{S}, \preceq)$ with stationary distribution μ. Suppose that the state space has a maximum element $\hat{1}$ and a minimal element $\hat{0}$. The transition matrix P is such that its *time reversal* $\tilde{P}$ defined by

$$\tilde{P}(x, y) = \frac{\mu(y) P(y, x)}{\mu(x)} \tag{3.3.1}$$

for x such that $\mu(x) > 0$ is monotone with respect to $\leq$. As $\tilde{P}$ is monotone, it is well-known, *e.g.* Theorem 1 in Kamae, Krengel and O'Brien (1977), that there exists an upward kernel $K_{(x,y)}(\cdot,\cdot)$ such that

$$\tilde{P}(y,y') = \sum_{x' \in \mathcal{S}} \tilde{P}(x,x') K_{(x,y)}(x',y') \tag{3.3.2}$$

for all $y \in \mathcal{S}$.

It is assumed that it is possible to sample from the measure $K_{(x,y)}(x',\cdot)$ whenever $x \leq y$ and $\tilde{P}(x,x') > 0$.

Definition 3.3.3 *A monotone transition rule for a transition matrix $\tilde{P}$ on a partially ordered space $(\mathcal{S},\leq)$ is a measurable function $\tilde{f} : \mathcal{S} \times \mathcal{U} \to \mathcal{S}$ together with a random variable U taking values in an arbitrary probability space $\mathcal{U}$ such that:*
(i) $\tilde{f}(x,u) \leq \tilde{f}(y,u)$ for all $u \in \mathcal{U}$ whenever $x \leq y$;
(ii) $\mathbb{P}(\tilde{f}(x,U) \in \cdot) = \tilde{P}(x,\cdot)$ for all $x \in \mathcal{S}$.

If $\tilde{P}$ has a monotone transition rule $\tilde{f}$ then we can take as upward kernel

$$K_{(x,y)}(x',y') := \mathbb{P}(\tilde{f}(y,U) = y' \mid \tilde{f}(x,U) = x'), \tag{3.3.4}$$

for all $y' \in \mathcal{S}$ when $x \leq y$ and $\tilde{P}(x,x') > 0$.

Algorithm 3.3.5 *Fill's algorithm consists in three steps:*

1. *Start X in $\hat{0}$ and runs it for t steps. Record the obtained trajectory $(X_0 = \hat{0}, X_1, \ldots, X_t = z)$.*

2. *Reverse the obtained trajectory in time leading to the time-reversed trajectory (which is regarded as a $\tilde{P}$ trajectory conditioned to start at z and to end at $\hat{0}$)*

$$(\tilde{X}_0 = z, \tilde{X}_1, \ldots, \tilde{X}_t = \hat{0}) = (X_t = z, X_{t-1}, \ldots, X_0 = \hat{0}).$$

3. *A second Markov chain $\tilde{Y}$ is simulated for t steps using the upward kernels $K_{(x,y)}(\cdot,\cdot)$ together with the time-reversed trajectory. The initial state of $\tilde{Y}$ is set to be $\hat{1}$. Then $\tilde{Y}_k$ for $k = 1,\ldots,t$ is simulated according to the kernel*

$$K_{(\tilde{X}_{k-1},\tilde{Y}_{k-1})}(\tilde{X}_k,\cdot).$$

If $\tilde{Y}_t = \hat{0}$ the proposed sample z is accepted.

4. *If the sample is not accepted, reinitiate the process with $t+1$ independently.*

Notice that this algorithm is based on the well-known acceptance-rejection algorithm for sampling, see Ripley (1987).

(a) Generate an observation from $P^t(\hat{0},\cdot)$;

(b) Find a constant c such that

$$c > \frac{\pi(z)}{P^t(\hat{0},z)}, \qquad \text{for all } z \in \mathcal{S} \text{ such that } P^t(\hat{0},z) > 0.$$

(c) Accept z as an observation from π with probability $c^{-1}\pi(z)/P^t(\hat{0},z)$.

Notice that for this algorithm to work there is a couple of questions to be answered.

- How to choose c? By definition of $\tilde{P}$ we know that

$$\frac{\pi(z)}{P^t(\hat{0},z)} = \frac{\pi(\hat{0})}{\tilde{P}^t(z,\hat{0})} \tag{3.3.6}$$

and by the monotonicity of $\tilde{P}$, we can choose

$$c = \frac{\pi(\hat{0})}{\tilde{P}^t(\hat{1},\hat{0})}. \tag{3.3.7}$$

Thus, step (c) says to accept z as an observation from π with probability

$$\frac{\tilde{P}^t(\hat{1},\hat{0})}{\pi(\hat{0})} \times \frac{\pi(z)}{P^t(\hat{0},z)} = \frac{\tilde{P}^t(\hat{1},\hat{0})}{\pi(\hat{0})} \times \frac{\pi(\hat{0})}{\tilde{P}^t(z,\hat{0})} = \frac{\tilde{P}^t(\hat{1},\hat{0})}{\tilde{P}^t(z,\hat{0})} \tag{3.3.8}$$

- How to design a coin-flip with probability of heads equals to $\tilde{P}^t(\hat{1},\hat{0})/\tilde{P}^t(z,\hat{0})$? Running the coupled-reversed chain starting at $\hat{1}$ for t steps, if the $\tilde{Y}_t = \hat{0}$, the coin flips head. In fact,

$$\mathbb{P}(\tilde{Y}_t = \hat{0} \mid \tilde{X}_0 = z, \tilde{X}_t = \hat{0}, \tilde{Y}_0 = \hat{1}) = \frac{\tilde{P}^t(\hat{1},\hat{0})}{\tilde{P}^t(z,\hat{0})}. \tag{3.3.9}$$

In this case we can construct the process $\underline{V} = (\mathbb{V}_t, t \geq 1)$ where $V_t = ((X_0,\ldots,X_t),(\tilde{Y}_0,\ldots,\tilde{Y}_t))$ are independently generated by the rejection algorithm described above.

$$\tau = \min\{t; \tilde{Y}_t = \hat{0}\} \tag{3.3.10}$$

and

$$\Phi_{\underline{V}}(t) = X_t. \tag{3.3.11}$$

Fill suggests the the algorithm should be run using t as powers of 2 similarly to CFTP.

3.3.1 Application to attractive spin systems

Consider an attractive spin system with attractive equilibrium measure π. Consider the Gibbs sampler (heat-bath algorithm) with uniform random update for attractive spins systems as described in Section 1.5. In this case, the chain is reversible with $\tilde{P} = P$ which is monotone since the system is attractive, then (1.5.2) gives us a monotone transition rule $\phi = \tilde{f}$. IA algorithm becomes:

1. Start the chain σ at $\sigma^{\min} \equiv -1$ and run it for t steps using ϕ as an updating rule. We obtain

$$(\sigma_0 \equiv -1, \sigma_1, \ldots, \sigma_{t-1}, \sigma_t = \zeta).$$

2. Construct the time reversed trajectory

$$(\tilde{\sigma}_0 = \zeta, \tilde{\sigma}_1 = \sigma_{t-1}, \ldots, \tilde{\sigma}_t \equiv -1).$$

3. Simulate a second Markov chain $\tilde{\eta}$ for t steps starting at $\tilde{\eta}_0 \equiv +1$ using the following rule:

   (a) When $\tilde{\sigma}_{n-1} \neq \tilde{\sigma}_n$ (they disagree at a unique site i),

   - If $\tilde{\sigma}_{n-1}(i) = -1$ and $\tilde{\sigma}_n(i) = +1$ then set $\tilde{\eta}_n(i) = +1$;
   - If $\tilde{\sigma}_{n-1}(i) = +1$ and $\tilde{\sigma}_n(i) = -1$ then set

$$\begin{cases} \tilde{\eta}_n(i) = -1, & \text{with probability } \frac{\pi(\tilde{\eta}^-_{n-1})}{\pi(\tilde{\eta}^-_{n-1})+\pi(\tilde{\eta}^+_{n-1})} \frac{\pi(\tilde{\sigma}_{n-1})+\pi(\tilde{\sigma}_n)}{\pi(\tilde{\sigma}_n)} \\ \tilde{\eta}_n(i) = +1, & \text{with probability } 1 - \frac{\pi(\tilde{\eta}^-_{n-1})}{\pi(\tilde{\eta}^-_{n-1})+\pi(\tilde{\eta}^+_{n-1})} \frac{\pi(\tilde{\sigma}_{n-1})+\pi(\tilde{\sigma}_n)}{\pi(\tilde{\sigma}_n)} \end{cases}$$

   (b) $\tilde{\sigma}_{n-1} = \tilde{\sigma}_n = \tilde{\sigma}^*$, then the computation of the conditional probability is a little messy, but we can overcome this problem by noticing that $\tilde{\sigma}_n = \sigma_{t-n}$ and $\tilde{\sigma}_{n-1} = \sigma_{t-n+1}$ and this transition was produced in step 1 by generating $U \sim U(0,1)$ and $V \sim U(\Lambda)$ independent and updating

$$\tilde{\sigma}^* = \phi(\tilde{\sigma}^*, U, V).$$

   So we can store (U, V) and use it to set

$$\tilde{\eta}_n = \phi(\tilde{\eta}_{n-1}, U, V).$$

4. If $\tilde{\eta}_t \equiv -1$ then the proposed sample ζ is accepted, if not reinitiate the process at Step 1.

3.3.2 Point processes

Thönnes (1999) uses IA algorithm to simulate from penetrable sphere model without impatient user bias. Her argument follows Fill's interruptible algorithm and uses a forward construction and a backward checking.

The objective is to simulate from the bi-dimensional point process $(N, M) \in \mathcal{N} \times \mathcal{N}$ described in Section 1.4.3. The crucial observation is that under the model (1.4.5), the the conditional distribution of N given M is a homogeneous Poisson process with intensity β_1 on $\Lambda \setminus (M \oplus G)$. Where G is a

sphere of radius R. Similarly, the conditional distribution of M given N is a homogeneous Poisson process with intensity β_2 on $\Lambda \setminus (N \oplus G)$.

In this case, a convenient Markov chain to be used is the Gibbs sampler. Given the configuration of the process at time t to be (n_t, m_t), then in the next step

$$m_{t+1} \sim \beta_2 \text{ Poisson on } \Lambda \setminus (n_t \oplus G) \tag{3.3.12}$$

$$n_{t+1} \sim \beta_1 \text{ Poisson on } \Lambda \setminus (m_{t+1} \oplus G). \tag{3.3.13}$$

Notice that, this Markov chain has an uncountable state space we can partially order this space by considering

$$(n, m) \leq (n', m') \qquad \text{if} \qquad n \subset n' \quad \text{and } m \supset m'.$$

It is easy to check that the Gibbs sampler defined by (3.3.12) and (3.3.13) defines a monotone transition rule f given by

$$f(n, m, V_1, V_2) = (n', m') \tag{3.3.14}$$

where V_1 and V_2 are independent Poisson point processes on Λ with rates β_1 and β_2 respectively and $n' = V_1 \setminus (m \oplus G)$ and $m' = V_2 \setminus (n' \oplus G)$. Notice that

$$f(n, m, V_1, V_2) \leq f(n', m', V_1, V_2) \quad \text{whenever} \quad (n, m) \leq (n', m'). \tag{3.3.15}$$

In fact, if $(n_1, m_1) = f(n, m, V_1, V_2)$, and $(n_1', m_1') = f(n', m', V_1, V_2)$, then

$$m_1 = V_1 \setminus (n \oplus G) \supset m_1' = V_1 \setminus (n' \oplus G), \text{ since } n \oplus G \subset n' \oplus G$$

and

$$n_1 = V_2 \setminus (m_1 \oplus G) \subset n_1' = V_2 \setminus (m_1' \oplus G), \text{ since } m_1 \oplus G \supset m_1' \oplus G.$$

However, the state space $\mathcal{N} \times \mathcal{N}$ does not have a maximal or minimal element. Instead, Häggström et al. (1999) call an element (n, m) *quasi-maximal* if

$$\Lambda \subset n \oplus G \qquad \text{and} \qquad m = \emptyset. \tag{3.3.16}$$

Similarly, the element (n, m) is quasi-minimal if

$$n = \emptyset \qquad \text{and} \qquad \Lambda \subset m \oplus G. \tag{3.3.17}$$

It is easy to check that if (n^0, m^0) is a minimal state and (n^1, m^1) is a maximal state, then for an arbitrary configuration (n, m), if we call (N^0, M^0), (N, M) and (N^1, M^1) the Markov chains obtained with initial states $(N_0^0, M_0^0) = (n^0, m^0)$, $(N_0, M_0) = (n, m)$ and $(N_0^1, M_0^1) = (n^1, m^1)$ respectively, we have

$$(N_n^0, M_n^0) \leq (N_n, M_n) \leq (N_n^1, M_n^1), \text{ for all } n \geq 1. \tag{3.3.18}$$

In fact, if $(N_1^0, M_1^0) = f(n^0, m^0, V_1, V_2)$, $(N_1, M_1) = f(n, m, V_1, V_2)$ and $(N_1^1, M_1^1) = f(n^1, m^1, V_1, V_2)$, then

$$M_1^0 = V_1 \text{ and } N_1^0 = V_2 \setminus (V_1 \oplus G)$$

$$M_1 = V_1 \setminus (n \oplus G) \text{ and } N_1 = V_2 \setminus (M_1 \oplus G)$$

and

$$M_1^1 = V_1 \setminus \Lambda = \emptyset \text{ and } N_1^1 = V_2$$

Then,

$$M_1^0 \supset M_1 \supset M_1^1 \quad \text{and} \quad N_1^0 \subset N_1 \subset N_1^1.$$

This proves (3.3.18) for $n = 1$, for $n > 1$ it is a consequence of the monotonicity of f.

Notice that, the Markov chain defined by the monotone rule f defined by (3.3.14), is reversible. However, the two-step rule (3.3.12) and (3.3.13) is not reversible. The two step rule for the reversed chain is:

$$\tilde{n}_{t+1} \sim \beta_1 \text{ Poisson on } \Lambda \setminus (\tilde{m}_t \oplus G) \tag{3.3.19}$$

$$\tilde{m}_{t+1} \sim \beta_2 \text{ Poisson on } \Lambda \setminus (\tilde{n}_{t+1} \oplus G). \tag{3.3.20}$$

Fill's algorithm

1. Start the chain (n, m) at a quasi-minimal point (n^0, m^0) and run it for t steps using (3.3.12) and 3.3.13 as a two step update rule. We obtain

$$((n_0, m_0) = (n^0, m^0), (n_1, m_1), \ldots, (n_{t-1}, m_{t-1}), (n_t, m_t) = (n^*, m^*)).$$

2. Construct the time reversed trajectory

$$((\tilde{n}_0, \tilde{m}_0) = (n^*, m^*), (\tilde{n}_1, \tilde{m}_1) = (n_{t-1}, m_{t-1}), \ldots, (\tilde{n}_t, \tilde{m}_t) = (n^0, m^0)).$$

3. Simulate a second Markov chain $(\tilde{v}, \tilde{w})$ for t steps starting at a quasi-maximal state using the following rule:

   - If at time n we have the following transitions on the time-reversed chain $(\tilde{n}_{n-1}, \tilde{m}_{n-1}) \to (\tilde{n}_{n-1}, \tilde{m}_n) \to (\tilde{n}_n, \tilde{m}_n)$, then let

$$v_n \sim \beta_1 \text{ Poisson on } \tilde{m}_{n-1} \oplus G \tag{3.3.21}$$

$$\tilde{v}_n \leftarrow [\tilde{n}_n \cup v_n] \setminus [\tilde{w}_{n-1} \oplus G] \tag{3.3.22}$$

$$\tilde{w}_n \leftarrow \tilde{m}_n \setminus [\tilde{v}_n \oplus G]. \tag{3.3.23}$$

4. If $(\tilde{v}_t, \tilde{w}_t)$ is quasi-minimal then the proposed sample ζ is accepted, if not reinitiate the process at Step 1.

3.4 Backward-Forward Algorithm

3.4.1 Spatial point processes

Consider first the case of the simulation of the invariant measure of a spatial birth and death process on $\mathbb{R}^d$ as described in Section 2 with finite birth rate λ such that

$$\bar{\lambda} = \sup_{x,\eta} \lambda(x, \eta) < \infty \tag{3.4.1}$$

and unit death rate.

Moreover, suppose that λ is G-local, that is, there exists a compact convex set G such that $\lambda(x, n_1) = \lambda(x, n_2)$ if n_1 and n_2 coincide inside G. Then, by (2.1.30) we know that condition $\bar{\lambda} < (m^d(G))^{-1}$ is sufficient for the birth and death process to be ergodic and its invariant measure μ is absolutely continuous with respect to the law of a $\bar{\lambda}$-homogeneous Poisson process on $\mathbb{R}^d$. In the case of the area-interaction point process

described in Section 1.4.1, G is a compact convex set and $\bar{\lambda} = \kappa$. For the attractive case, $\lambda(x\eta) = \kappa\phi^{-m_d((x+G)\setminus(\eta\oplus G))}$ and for the repulsive case $\lambda(x\eta) = \kappa\phi^{m_d(G)-m_d((x+G)\setminus(\eta\oplus G))}$. For the Strauss process, G is the ball centered at the origin with radius r, $\bar{\lambda} = e^{\beta_1}$ and $\lambda(x,\eta) = e^{\beta_1}e^{\beta_2\eta(x+G)}$. In this case, our approach requires a truncation of the form $S(N,\Lambda) \to \min\{S(N,\Lambda), K\}$, for some $K > 0$, to guarantee uniformly bounded birth-rates.

The objective is to simulate from μ restricted to a finite-box Λ. The outline of the scheme is:

1. Generate the *free* process α as a $\bar{\lambda}$-homogeneous Poisson process on λ according to Algorithm 1.2.3.

2. Construct the clan of ancestors of all points of α.

3. Apply the *deterministic* finite-volume "cleaning procedure" described in Section 2.1.1 to decide which points of α are going to be kept.

Algorithm 3.4.2 Construction of the clan of ancestors

(i) Generate the free *process* $\alpha_0 = \{x_1, x_2, \ldots, x_R\}$ *as a* $\bar{\lambda}$*-homogeneous Poisson process on* Λ *according to Algorithm 1.2.3.*

(ii) Generate $S_1^0, \ldots, S_R^0$ *independent mean one exponential random variables and construct the following cylinders:*

$$C_0 = \{(x_i + G) \times [-S_i^0, 0]; i = 1, 2, \ldots, R\}. \tag{3.4.3}$$

(iii) Consider the following fattening of C_0*, a subset of* $\mathbb{R}^d \times (-\infty, 0]$

$$\Lambda_0 = \cup_{C\in C_0}(\text{Basis}\,(C) \oplus G) \times \text{Life}\,(C) \tag{3.4.4}$$

where $\text{Basis}\,(C) \oplus G = \{x + y; x \in \text{Basis}\,(C), y \in G\}$.

(iv) Set $\ell = 1$. *Generate a* $\bar{\lambda}$*-homogeneous Poisson process* $\{\xi_1^1, \xi_2^1, \ldots, \xi_{R_1}^1\}$ *on* Λ_0 *according to Algorithm 1.2.3.*

(v) *Generate* $S_1^1, \ldots, S_{R_1}^1$ *independent mean one exponential random variables and construct the following cylinders:*

$$C_1 = \{(\xi_i^\ell + G) \times [-S_i^\ell, 0]; i = 1, 2, \ldots, R_\ell\}. \tag{3.4.5}$$

(vi) *Consider the following fattening of* C_ℓ*, a subset of* $\mathbb{R}^d \times (-\infty, 0]$

$$\Lambda_\ell = \cup_{C \in C_\ell} (\text{Basis}\,(C) \oplus G) \times life(C) \tag{3.4.6}$$

(vii) *Generate a* $\bar{\lambda}$*-homogeneous Poisson process* $\{(\xi_1^{\ell+1}, T_1^{\ell+1}), (\xi_2^{\ell+1}, T_2^{\ell+1}), \ldots, (\xi_{R_\ell}^{\ell+}$ *on* $\Lambda_\ell \setminus \Lambda_{\ell-1}$ *according to Algorithm 1.2.3.*

(viii)

- *If* $R_{\ell+1} = 0$*, set the clan of ancestors of* α

$$A^\alpha := \cup_{i=1}^{\ell+1} C_i \tag{3.4.7}$$

and stop.

- *If not, set* $\ell = \ell + 1$*. Generate* $S_1^\ell, \ldots, S_{R_\ell}^\ell$ *independent mean one exponential random variables and construct the following cylinders:*

$$C_\ell = \{(\xi_i^\ell + G) \times [-S_i^\ell, T_i^\ell]; i = 1, 2, \ldots, R_\ell\}. \tag{3.4.8}$$

and go back to (vi).

A slight modification of the algorithm allows to simulate the penetrable sphere model described in Section 1.4.3.

1. Generate the *free* process α as a marked point process on λ according to Section 1.3 where the location process is a $\beta_1 + \beta_2$-homogeneous Poisson process and the marks are independent with

$$P(1) = \frac{\beta_1}{\beta_1 + \beta_2} \quad \text{and} \quad P(2) = \frac{\beta_2}{\beta_1 + \beta_2}.$$

2. Construct the clan of ancestors of all points of α.

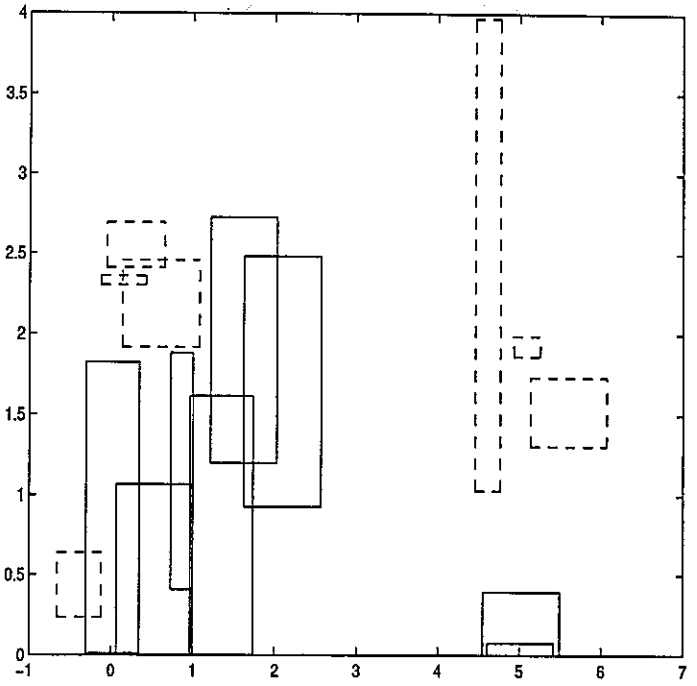

Figure 3.4.8: The *free process*

3. Apply a modification of the *deterministic* finite-volume "cleaning procedure" described in Section 2.1.1 to decide which points of α are going to be kept.

   FV.3a. If t_i is a birth time, that is, $t_i = s$ for some $(\xi, s, l, z) \in \mathbf{C}^\Lambda[0,t]$ then we do not add the point (ξ, z) if $z = 1$ and $d(\xi, \eta_{t-}(\cdot, 1)) \leq R$ or if $z = 2$ and $d(\xi, \eta_{t-}(\cdot, 2)) \leq R$: we set

   $$\eta^\Lambda_{t_i} = \eta^\Lambda_{t_i-}. \tag{3.4.9}$$

   Otherwise the point (ξ, z) is added: we set

   $$\eta^\Lambda_{t_i} = \eta^\Lambda_{t_i-} \cup \{(\xi, z)\}. \tag{3.4.10}$$

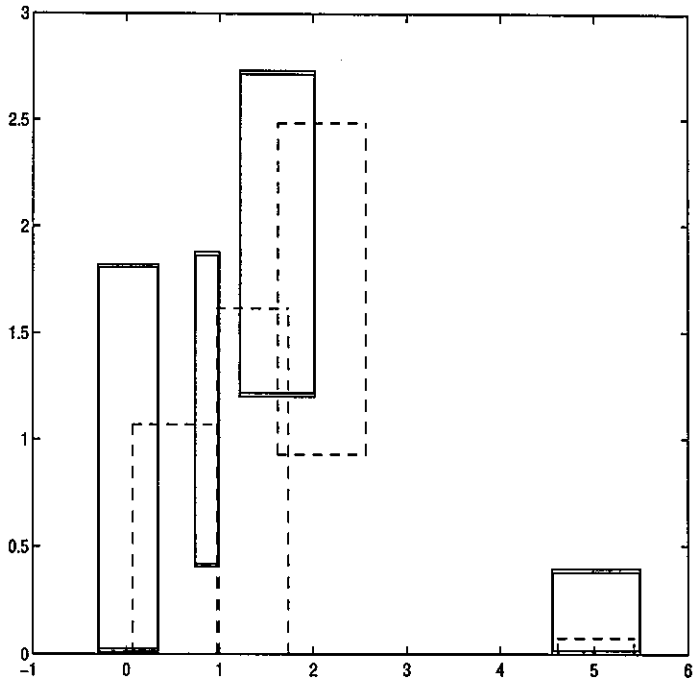

Figure 3.4.10: Cleaned process

3.4.2 Peierls contours of the ferromagnetic Ising model

The above algorithm do not translate immediately to the Ising model because the fattening of the cylinders cannot be done before the generation of the contours. In fact, practical limitations prevent even the inclusion of all possible sizes in the simulation, in fact the mere enumeration of the possible contours is beyond reach when more than a few dozens of links are involved. Also, this scheme suffers from the user impatient bias as described for the CFTP. In this case, we do have a simulation scheme from the distribution μ conditioned on two events, letting $K =$ "maximum perimeter of bases of cylinders in the clan" we sample from (using the notation of Definition 3.1.1)

$$\mathbb{P}\Big(\Phi_{\underline{V}}(\tau) \in A \;\Big|\; \{K < k\} \cap \{\tau < S\}\Big) \tag{3.4.11}$$

where $k = 30$ and $S =$"the maximum time left in order to have the results ready for the next congress", for instance.

In fact, our approach also admits a joint realization (η, ξ) with the right marginal distributions such that $\eta = \xi$ if $K < k$ and $\tau < S$, and such that $\mathbb{P}(\{K \geq k\} \cap \{\tau > S\})$ goes to zero exponentially fast in S and in the cutoff of the length of the contours (30 in our example). Slightly more precisely,

$$\mathbb{P}\Big(\{K \geq k\} \cap \{\tau > S\}\Big) \leq \mathrm{O}\Big(\alpha^{\tau} \times \sup_{x} \pi_x(K > k)\Big). \tag{3.4.12}$$

This follows from the sub-criticality of the majorizing branching process. For the Ising model, for instance, $\pi_x(K > k) = \mathrm{O}(e^{-\beta k})$.

Fernández, Ferrari and Garcia (2000) propose a non-homogeneous time-backwards construction of these clans based on a result proven in Section 4.5.1 of Fernández et al. (1999). It is shown there that the clan of ancestors of a family of cylinders can be obtained combing *back* in time and generating births of ancestors with an appropriate rate. In fact, this rate is equal to the rate density of the free process multiplied by an exponential time factor ensuring that the ancestor has a lifespan large enough to actually be an ancestor. This time factor involves the time-distance to the birth of existing cylinders, which can be expressed through the following function. For a finite region Λ and a finite set of cylinders $\mathbf{H}$, let the set of basis of the potential ancestors of $\mathbf{H}$ and $\Lambda \times \{0\}$ be defined by

$$\mathbf{G}(\mathbf{H}, \Lambda) := \Big\{\theta \in \mathbf{G} : \mathrm{Basis}\,(C') \not\sim \theta, \text{ for some } C' \in \mathbf{H}\Big\} \bigcup \Big\{\theta \in \mathbf{G} : \theta \cap \Lambda \neq \emptyset\Big\} \tag{3.4.13}$$

and for a given individual $\theta \in \mathbf{G}(\mathbf{H}, \Lambda)$,

$$\mathrm{TI}(\mathbf{H}, \Lambda, \theta) = \min\Big\{\mathrm{Birth}\,(C') : C' \in \mathbf{H}, \mathrm{Basis}\,(C') \not\sim \theta\Big\} \tag{3.4.14}$$

with the convention $\min \emptyset = 0$. By definition, $\mathrm{TI}(\mathbf{H}, \Lambda, \theta) \leq 0$.

The outline of the scheme is:

1. Generate in Λ the "free process" of contours ξ with distribution μ^0_Λ, product of Poisson random variables with mean $e^{-\beta|\gamma|}$ for $\gamma \in \mathbf{G}$ such that $|\gamma| \leq K$ and $\gamma \cap \Lambda \neq \emptyset$ (*cf.* (1.5.11)). Let $\xi = \{\gamma_1, \gamma_2, \ldots, \gamma_R\}$.

2. Construct the clan of ancestors of all contours of ξ.

3. Beginning with the first ancestor (the one first born), erase all incompatible contours.

Algorithm 3.4.15 An algorithm to construct the backwards clan of a finite region

The combination of (3.4.13)/(3.4.14) can be translated into the following explicit algorithm. We do it first for the case of countable number of individuals and indicate at the end of this section how to proceed in the continuous case. To generate $\mathbf{A}^{\Lambda,0}$*:*

BFA.1. *Set* $\ell = 0$ *and* $\tau_0 = 0$. *Generate* $S_1^0, S_2^0, \ldots, S_R^0$ *independent mean one exponential random variables.*

Set

$$\mathcal{C}_0 = \{(\gamma_i, 0, S_i^0); i = 1, 2, \ldots, R\}. \tag{3.4.16}$$

BFA.2. *For each* $\gamma \in G(\mathcal{C}_\ell, \Lambda)$ *generate an independent random variable* $\tau(\gamma)$ *such that*

$$\mathbb{P}(\tau(\gamma) > t) = 1 - \exp(-\nu_\gamma(s)) \tag{3.4.17}$$

where

$$\nu_\gamma(s) = e^{-\beta|\gamma|} e^{-s + \mathrm{TI}(\mathcal{C}_\ell, \Gamma, \gamma)} \mathbf{1}\{s > \tau_\ell\}. \tag{3.4.18}$$

Notice that $\tau(\gamma)$ *may be infinity.*

BFA.3. *Let Let* $\ell = \ell + 1$ *and* $\tau_\ell = \inf\{\tau(\gamma); \gamma \in G(\mathcal{C}_\ell, \Lambda)\}$.

- *If* $\tau_\ell < \infty$, *call* $\widetilde{\gamma}$ *be such that* $\tau_\ell = \tau(\widetilde{\gamma})$. *Let*

$$\mathcal{C}_\ell = \mathcal{C}_{\ell-1} \cup \{(\widetilde{\gamma}, -\tau_\ell, \tau_\ell + \mathrm{TI}(\mathcal{C}_{\ell-1}, \Lambda, \widetilde{\gamma}) + S^\ell)\} \tag{3.4.19}$$

where S^ℓ *is an exponentially distributed mean one random variable generated independently of everything else. Go back to* **BFA.2**.

- *If $\tau_\ell = \infty$, let the clan of ancestors of ξ be defined as*

$$\mathbf{A}^{\xi} := \mathcal{C}_\ell \tag{3.4.20}$$

and stop.

The above algorithm can be improved by removing the generation of the free process and beginning with the empty configuration. The algorithm will generate the cylinders and we can find which ones survive at time zero. See Section 4.2 (Step 1) of Fernández, Ferrari and Garcia (2000).

In the continuous loss network described in Section 1.6.1, time and space can not be in general separated. Instead of steps **BFA.2** and **BFA.3** above we must consider a random sample w of π and the events (x, s) of a Poisson process with rate

$$\nu(x,s) = \kappa\, e^{-s+\mathrm{TI}(\mathcal{C}_\ell,\Lambda,\gamma)}\, \mathbf{1}\{s > \tau_\ell\}\, \mathbf{1}\{\gamma \in \mathbf{G}(\mathcal{C}_\ell,\Lambda)\} \tag{3.4.21}$$

where $\gamma = (x, x+w)$.

For a finite window Λ the total rate is finite, hence these events can be well ordered by looking to the time coordinate. If the set of these events is not empty, we take τ_ℓ to be the minimal time coordinate (it is strictly positive with probability one) and denote $\widetilde{\gamma}$ the associated interval $(x, x+w)$. If the Poisson process with rate density (3.4.21) yields no event we take $\tau_\ell = \infty$. We then continue as in **BFA.3**.

3.5 Concluding remarks

There are several different approaches to perfect simulation of spatial point processes, every week a new procedure is proposed either improving old ones or suggesting new ideas. As it can be seem from the examples given above, none of them is better than the others in absolute terms. For example, as pointed by Møller (personal communication), CFTP is much more efficient

than BFA for the Strauss process in a finite region. First, it can be applied to a much broader regimen. Second, exploring the structure of a model such as the repulsive behavior in the Strauss process and using upper and lower processes as described in Kendall and Møller (1999) give a lower coalescence time than the stopping time needed for BFA. However, for infinite-volume regions, it is less efficient since it involves a coupling for maximal and minimal configurations and also a limit procedure in time and space. Fill's IA has the advantage of having no "user impatient bias", however it requires monotonicity of the reversed process. In fact, we can think as all of the methods to be, in some sense, complementary to each other. Simulations based on CFTP and IA can be applied to much broader class of processes. Nevertheless, they need specific conditions such as finite volume or monotonicity. On the other hand, although BFA has a much smaller range of validity it has the advantage of its generality. Moreover, it is a powerful theoretical tool. Probabilistic arguments (successive dominations by oriented percolation, life-and-death and branching processes) yield all the properties obtained via usual cluster expansions —except analyticity— in a larger region and in a more intuitive and concrete way: convergence of the series is replaced by sub-criticality of a branching process, mixing and central-limit properties are a consequence of lack of percolation.

Bibliography

Abbé, E. (1879). Über blutkörper-zahlung, *Jena Z. Med. Naturwiss* **13**: 98–105. New Series 6.

Aizenman, M., Bricmont, J. and Lebowitz, J. (1987). Percolation of minority spins in high-dimensional Ising models, *J. Statist. Phys.* **49**: 859–865.

Aldous, D. and Fill, J. (1999). *Reversible Markov Chains and Random Walks on Graphs*, Draft book, http://www.stat.berkeley.edu/~aldous/book.html.

Baddeley, A. J. and Møller, J. (1989). Neasrest-neighbour markov point processes and random sets, *Int. Statist. Rev.* **57**: 90–121.

Baddeley, A. J. and van Lieshout, M. N. M. (1995). Area-interaction point processes, *Ann. Inst. Statist. Math.* **47**(4): 601–619.

Bateman, H. (1910). Note on the probability distribution of α-particles, *Philos. Mag.* **20**(6): 704–707. Note to E. Rutherford and H. Geiger, The probability variations in the distribution of α-particles. *Philos. Mag.*, **20**(6), pp.698–704.

Clifford, P. and Nicholls, G. (1994). Comparison of birth-and-death and metropolis-hastings markov chain monte carlo for the strauss process, Unpublished manuscript. `http://www.math.auckland.ac.nz/~nicholls/linkfiles/papers.html`.

Daley, D. J. and Vere-Jones, D. (1988). *An introduction to the theory of point processes*, Springer-Verlag, New York.

Dobrushin, R. L. (1965). Existence of a phase transition in the two-dimensional and three-dimensional Ising models, *Soviet Physics Dokl.* **10**: 111–113.

Dobrushin, R. L. (1996). Perturbation methods of the theory of Gibbsian fields, *Lectures on probability theory and statistics (Saint-Flour, 1994)*, Springer, Berlin, pp. 1–66.

Erlang, A. K. (1909). The theory of probabilities and telephone conversations, *Nyt. Tidsskr. Mat. B* **20**: 33–41. [Reprinted in E. Brockmeyer, H.L. Halstrom and Arne Jensen (1948), *The life and Works of A.K. Erlang*, Copenhagen Telephone Company, Copenhagen, pp.131–137.].

Feller, W. (1968). *An Introduction to Probability Theory and its Applications, Vol. I*, John Wiley & Sons Inc., New York.

Fernández, R., Ferrari, P. A. and Garcia, N. L. (1998). Measures on contour, polymer or animal models. A probabilistic approach, *Markov Process. Related Fields* 4(4): 479–497. I Brazilian School in Probability (Rio de Janeiro, 1997).

Fernández, R., Ferrari, P. A. and Garcia, N. L. (1999). Loss network representation of Peierls contours, *Preprint* p. http://www.ime.usp.br/~pablo.

Fernández, R., Ferrari, P. A. and Garcia, N. L. (2000). Perfect simulation for interacting point processes, loss networks and ising models, *Preprint. Los Alamos archive: Mathematics, abstract math.PR/9911162t.*

Ferrari, P. A. and Garcia, N. L. (1998). One-dimensional loss networks and conditioned $M/G/\infty$ queues, *J. Appl. Probab.* **35**(4): 963–975.

Fill, J. A. (1998). An interruptible algorithm for perfect sampling via Markov chains, *Ann. Appl. Probab.* **8**(1): 131–162.

Foss, S. G. and Tweedie, R. L. (1998). Perfect simulation and backward coupling, *Comm. Statist. Stochastic Models* **14**(1-2): 187–203. Special issue in honor of Marcel F. Neuts.

Garcia, N. L. (1995). Birth and death processes as projections of higher-dimensional Poisson processes, *Adv. in Appl. Probab.* **27**(4): 911–930.

Green, P. J. and Murdoch, D. J. (1999). Exact sampling for Bayesian inference: towards general purpose algorithms, *Bayesian statistics, 6 (Alcoceber, 1998)*, Oxford Univ. Press, New York, pp. 301–321.

Häggström, O. and Nelander, K. (1998). Exact sampling from anti-monotone systems, *Statist. Neerlandica* **52**(3): 360–380.

Häggström, O., van Lieshout, M.-C. N. M. and Møller, J. (1999). Characterization results and Markov chain Monte Carlo algorithms including exact simulation for some spatial point processes, *Bernoulli* **5**(4): 641–658.

Hall, P. (1988). *Introduction to the theory of coverage processes*, John Wiley & Sons Inc., New York.

Kamae, T., Krengel, U. and O'Brien, G. L. (1977). Stochastic inequalities on partially ordered spaces, *Ann. Probability* **5**(6): 899–912.

Karr, A. F. (1986). *Point Processes and their Statistical Inference*, Marcel Dekker, New York.

Kelly, F. P. (1991). Loss networks, *Ann. Appl. Probab.* **1**(3): 319–378.

Kelly, F. P. and Ripley, B. D. (1976). A note on Strauss's model for clustering, *Biometrika* **63**(2): 357–360.

Kendall, W. and Møller, J. (1999). Perfect Metropolis-Hastings simulation of locally stable point processes, *Technical Report R-99-2001*, Department of Mathematical Sciences, Aalborg University. To appear in *Advances in Applied Probability.*

Kendall, W. S. (1997). On some weighted Boolean models, *Proceedings of the International Symposium on Advances in Theory and Applications of Random Sets (Fontainebleau, 1996)*, World Sci. Publishing, River Edge, NJ, pp. 105–120.

Kendall, W. S. (1998). Perfect simulation for the area-interaction point process, *Probability towards 2000 (New York, 1995)*, Springer, New York, pp. 218–234.

Khinchin, A. Y. (1955). *Mathematical Methods in the Theory of Queueing (in Russian)*, Vol. 49, Trudy Mat. Inst. Steklov. Translated (1960) Griffin, London.

Kurtz, T. (1989). Stochastic processes as projections of poisson random measures. special invited paper at ims meeting, washington, d.c., *Unpublished.*

Lebowitz, J. L. and Gallavotti, G. (1971). Phase transitions in binary lattice gases, *J. Mathematical Phys.* **12**: 1229–1233.

Lebowitz, J. L. and Mazel, A. E. (1998). Improved Peierls argument for high-dimensional Ising models, *J. Statist. Phys.* **90**(3-4): 1051–1059.

Liggett, T. M. (1985). *Interacting particle systems*, Springer-Verlag, New York.

McKendrick, A. G. (1914). Studies on the theory of continuous probabilities with specieal reference to its bearing on natural phenomena of a progressive nature, *Proc. London Math. Soc.* **13**(2): 401–416.

McKendrick, A. G. (1926). The application of mathematics to medical problems, *Proc. Edinburgh Math. Soc.* **44**: 98–130.

Mira, A., Møller, J. and Roberts, G. O. (1999). Perfect slice sampler, *Technical Report R-99-2020*, Department of Mathematical Sciences, Aalborg University.

Møller, J. (1989). On the rate of convergence of spatial birth-and-death processes, *Ann. Inst. Statist. Math.* **41**(3): 565–581.

Møller, J. (1999). Markov chain Monte Carlo and spatial point processes., *in* O. Barndorff-Nielsen, W. Kendall and M. van Lieshout (eds), *Stochastic Geometry: Likelihood and Computations*, number 80 in *Monographs on Statistics and Applied Probability*, Chapman and Hall/CRC, Boca Raton, pp. 141–172.

Møller, J. and Nicholls, G. (1999). Perfect simulation for sample-based inference, *Technical Report R-99-2011*, Department of Mathematical Sciences, Aalborg University.

Neyman, J. (1939). On a new class of "contagious" distributions applicable in entomology and bacteriology, *Ann.Math. Statist.* **10**: 35–57.

Nyeman, J. and Scott, E. (n.d.). Statistical approaches to prblems of cosmology, *J. Royal Statist. Soc. B* **20**: 1–29.

Palm, C. (1943). Intensitätsschwankungen im fernsprechverkehr, *Ericsson Techniks* 44: 1–189.

Poisson, S. (n.d.). *Recherches sur la Probabilité des Jugements en Matière Criminelle er en Matière Civile, Précédées des Règles Générales du Calcul des Probabilités*, Bachalier, Paris.

Preston, C. J. (1977). Spatial birth-and-death processes, *Bull. Int. Statist. Inst.* **46**: 371–391.

Prohorov, Y. V. (1956). Convergence of random processes and limit theorems in probability theory (in russian), *Teor. Veroyat. Primen.* **1**: 177–238. Translation in *Theory Probab. Appl.*, **1**, pp. 157–214.

Propp, J. G. and Wilson, D. B. (1996). Exact sampling with coupled Markov chains and applications to statistical mechanics, *Proceedings of the Seventh International Conference on Random Structures and Algorithms (Atlanta, GA, 1995)*, Vol. 9, pp. 223–252.

Ripley, B. D. (1977). Modelling spatial patterns, *J. Roy. Statist. Soc. Ser. B* **39**(2): 172–212. With discussion.

Ripley, B. D. (1987). *Stochastic simulation*, John Wiley & Sons Inc., New York.

Ruelle, D. (1971). Existence of a phase transition in a continuous classical system, *Phys. Rev. Lett.* **27**: 1040–1041.

Seidel, H. (1876). Über die probabilitäten solcher ereignisse welche nur seiten vorkommen, obgleich sie unbeschränkt oft möglich sind., *Sitzungsber. Math. Phys. Cl. Akad. Wiss. München* **6**: 44–50.

Strauss, D. J. (1975). A model for clustering, *Biometrika* **62**(2): 467–475.

Thönnes, E. (1999). Perfect simulation of some point processes for the impatient user, *Adv. in Appl. Probab.* **31**(1): 69–87.

Widow, B. and Rowlinson, J. S. (1970). New model for the study of liquid-vapor phase transitions, *J. Chem. Phys.* **52**: 1670–1684.

Wilson, D. B. (2000). How to couple from the past using a read-once source of randomness, *Random Structures and Algorithms* **16**(1): 85–113. arXiv:math.PR/9910050.

Yule, G. (1924). A mathematical theory of evolution, based onthe conclusions of dr. j.c. willis, *Philos. Trans. B* **213**: 21–87.

Impresso na Gráfica do

pelo Sistema Xerox / 5390