Simulations of the growth of cities

Nga Nguyen

UNIVERSITÉ D'ORLÉANS

ÉCOLE DOCTORALE SCIENCES ET TECHNOLOGIES

LABORATOIRE : MAPMO

THÈSE présentée par :

Thi Thuy Nga NGUYEN

soutenue le : **08 January 2014**

pour obtenir le grade de : **Docteur de l'université d'Orléans**

Discipline/ Spécialité : **Mathématiques**

MODÉLISATION DE LA CROISSANCE DES VILLES

THÈSE DIRIGÉE PAR :

Michel ZINSMEISTER	Professeur, Université d'Orléans - Directeur de thèse
Athanasios BATAKIS	MCF, Université d'Orléans - Co-directeur de thèse

RAPPORTEURS :

Pierre NOLIN	Professeur, ETH Zürich,
Dominique BADARIOTTI	Directeur de Recherche, Université de Strasbourg

---

JURY :

Dominique BADARIOTTI	Directeur de Recherche, Université de Strasbourg
Athanasios BATAKIS	MCF, Université d'Orléans
Aline BONAMI	Professeur, Université d'Orléans
Laurent BOUDIN	MCF, Université Pierre et Marie Curie
Bertrand DUPLANTIER	Directeur de Recherche, IPhT, Saclay
Denis GREBENKOV	Chargé de Recherche, Ecole Polytechnique
Jean-Marc ZANINETTI	Professeur, Université d'Orléans
Michel ZINSMEISTER	Professeur, Université d'Orléans

Remerciements

Ce manuscrit conclut trois ans de travail, je tiens en ces quelques lignes à exprimer ma reconnaissance envers tous ceux qui de près ou de loin y ont contribué.

En premier lieu, je tiens à remercier Michel Zinsmeister et Athanassios Batakis, directeur et co-directeur de thèse, pour leur encadrement, leurs conseils, ainsi que pour leurs encouragements, leur enthousiasme et leur confiance. Je leur suis reconnaissante pour leur aide précieuse dans la relecture et la correction de ma thèse. Je souhaite néanmoins remercier plus particulièrement Nassos pour son attention de tout instant sur mes travaux, pour sa patience, sa disponibilité à comprendre et à communiquer et sans oublier ses précieuses intuitions.

Je remercie cordialement Messieurs Pierre Nolin et Dominique Badariotti d'avoir accepté d'être rapporteurs de cette thèse. Je souhaite adresser également mes remerciements à Madame Aline Bonami, Messieurs Laurent Boudin, Bertrand Duplantier, Denis Grebenkov et Jean-Marc Zaninetti d'avoir fait l'honneur d'examiner mon travail. Leurs nombreuses remarques et suggestions ont amélioré la qualité de ce mémoire.

J'exprime mes sincères remerciements à l'équipe du projet TRUC (Transport, Réseaux, Urbanisme, Croissance). Les discussions que j'ai pu avoir durant les réunions d'équipe m'ont beaucoup apporté. Cette thèse n'aurait pas été possible sans leur aide. Je tiens à remercier particulièrement Chloë Legrand pour toutes nos discussions et ses données géographiques indispensables à mon travail.

Une grande partie de cette thèse était effectuée en commun avec Michel Zinsmeister, Bertrand Duplandtier et Chi Nguyen. Je leur suis très reconnaissante pour cette collaboration. Je remercie tout particulièrement Chi pour avoir relu la première partie et pour m'avoir donné de précieux conseils.

Je ne peux également oublier de remercier mon laboratoire MAPMO pour son soutien logistique et matériel qui ne fait jamais défaut aux doctorants. Je souhaite notamment remercier l'équipe d'Analyse et Systèmes Dynamiques . Leur activité scientifique dynamique m'a permis de découvrir la richesse du monde de la recherche.

Je remercie mes amis, Chi, Tam, Ly, Ky, Nhat, Thuy, Phuong et Lan pour leurs conseils, leur soutien et leurs qualités de cœur au cours de ces trois années éloignées de ma famille. Je souhaite remercier spécialement Bao pour m'avoir accompagnée et écoutée tout au long de la thèse. Un sincère remerciement à Nicole pour m'avoir aidée en français.

Mes remerciements s'adressent enfin à mes parents et mon frère pour leur amour et leur soutien sans faille.

Contents

Introduction

Modeling of urban growth is an interesting and important problem for economy and for society in general. It gives foundation to predict development trend of urban (population, transport system, commercial areas ...) in order to give reasonable policies for urbanism. Scientists from many fields (geography, physics, mathematics ....) have studied this problem.
For a physicist or a mathematician, looking only at the map, this problem is a particular case of 2D growth problem. A very famous 2D growth problem is DLA (diffusion-limited aggregation) for which cluster at time $(n+1)$ is obtained from the one at time n by adding a pixel chosen on the n^{th} cluster with probability law being harmonic measure at infinity.
If this model is very realistic for electrodeposition or some other physical phenomena, it is not suitable for cities growth: there are quantified reasons for this (see in chapter 4) but also obvious one, i.e., that the DLA clusters, do not look like cities (see figure 1).

There is a mathematical general model to describe 2D growth phenomena. This is Loewner differential equation. Very roughly, Loewner has shown how to describe the growth of a cluster as a process "driven" by a function of the time with values in the unit circle. Loewner developed this theory for solving Bieberbach conjecture for the coefficient $n = 3$ (see in chapter 2).
Much later in 1999, Oded Schramm revived Loewner equation with the intuition of taking in Loewner problem a driving function of the form $\lambda(t) = e^{i\sqrt{\kappa}B_t}$ where B_t is a standard 1-D Brownian motion. This may be generalized to $\lambda(t) = e^{iL_t}$ where L_t is a Lévy process: in this way we keep the Markov property of the process but allow discontinuous driving functions, a fact which translates into dendritic growth, as in DLA for instance.

In chapter one, we recall Loewner differential equation and define Stochastic Loewner evolution of O. Schramm under three variants: whole-plane, radial, and chordal. We also define Lévy Loewner evolution in the same vein.

In chapter 2, we come back to Bieberbach coefficients for whole-plane LLE and study what can be said for the coefficients which are now random variables. We observe in particular that $\mathbb{E}(|a_n|^2) = 1$ for $\kappa = 6$ and $\mathbb{E}(|a_n|^2) = n$ for $\kappa = 2$ which is a new phenomena that we now completely understand via Beliaev-Smirnov equation [DNNZ12].

Now a very famous theorem of Smirnov identifies SLE_6 as boundary of critical

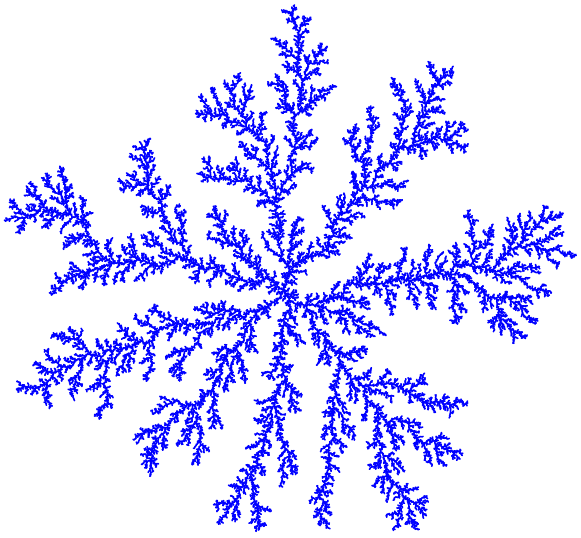


Figure 1: An example of DLA.

percolation clusters [Smi01]. It happens that a variant of percolation is the main ingredient of a growth model for cities that was introduced by Makse et al [MAJB$^+$98] and used for the simulation of the growth of Berlin.
In chapter 3, we study in detail the model of gradient percolation which was first introduced by Sapoval et al [SRG85] and studied mathematically by Nolin [Nol08a]. In the model where the percolation probability depends exponentially on the distance of the site to the center, we show that boundary of the largest cluster stay in an area near the curve at critical percolation with high probability and its expected length will be estimated according to length of this curve.

In chapter 4, we give a detailed account of how gradient percolation may be used to model city growth. We discuss the needed parameters and also the coefficient of correlation which is needed to obtain realistic simulation fractal dimension of frontier. In the end of this chapter, we illustrate all these notions with simulation of the city of Baltimore, made possible by data found on the website.

Chapter 5 is a description of a work connected with APR TRUC, a grant we obtained for a pluridisciplinary study in particular of the SCoT (in French, it means le schéma de cohérence territoriale) of Montargois-en-Gâtinais. The chapter begins with a geographical description of the area of Montargis; we have studied population, constructions, land lots, and transport system together with natural elements like rivers, forests, risk zone, etc.... Later we use correlated gradient percolation to model the growth of the city and this modeling had been made possible by the

accurate knowledge of the data of Montargis. We proposed two modified models. With the first model depending on local density we can simulate Montargis in the future and in the past. The second one is built on land lots which added natural factors which affect urban growth such as rivers, roads system, forests and risk zone. We end the chapter with some evaluations showing the future works of this approach.

Chapter 1

Planar Growth Processes

1.1 Loewner equation

In 1923 Charles Loewner, in order to prove Bieberbach conjecture for $n = 3$ (see below for detail on Bieberbach conjecture), invented a very important tool to study planar growth processes named after him Loewner differential equation. This equation involves Riemann mapping theorem for simply-connected domains and we begin with a brief introduction to this subject.

1.1.1 Simply connected domains

An arc in a metric space X is a continuous mapping $\gamma : [a, b] \subset \mathbb{R} \to X$. Such an arc is said to be closed if $\gamma(a) = \gamma(b)$. Two arcs γ_1, γ_2 defined on the same interval $[a, b]$ are said to be *homotopic* if there exists $\Gamma : [a, b] \times [0, 1] \to X$ continuous such that

$$\forall s \in [a, b],\ \Gamma(s, 0) = \gamma_1(s),\ \ \Gamma(s, 1) = \gamma_2(s).$$

Definition 1.1.1. The space X is called simply-connected if it is connected and if every closed arc $\gamma : [a, b] \to X$ is homotopic to a constant arc $\gamma_0 : [a, b] \to \gamma(a)$.

When X is a plane domain we have the following equivalent characterizations of simply connected domains:

Theorem 1.1.2. *For a connected open subset Ω of $\mathbb{C}$ the followings are equivalent:*

i. Ω is simply connected,

ii. $\overline{\mathbb{C}} \setminus \Omega$ is connected,

iii. For any closed arc γ whose image lies in Ω and any $z \notin \Omega$, $\mathrm{Ind}(z, \gamma) = 0$.

We recall that $\mathrm{Ind}(z, \gamma)$ stands for the variation of the argument (measured in number of turns) of $\gamma(t) - z$ along $[a, b]$. When γ is piecewise C^1 this quantity is also equal to

$$\frac{1}{i2\pi} \int_a^b \frac{\gamma'(s)}{\gamma(s) - z} \mathrm{d}s = \frac{1}{i2\pi} \int_\gamma \frac{1}{\zeta - z} \mathrm{d}\zeta.$$

Theorem 1.1.3 (Riemann)**.** *Let Ω be a simply-connected proper subdomain of $\mathbb{C}$ and $w \in \mathbb{C}$. Then there exists a unique biholomorphic map $g : \Omega \to \mathbb{D}$ such that $g(w) = 0$, $g'(w) > 0$.*

An equivalent statement is that there exists a unique holomorphic bijection $f : \mathbb{D} \to \Omega$ sending 0 to $z_0 \in \Omega$ and $f'(0) > 0$. This specific map f will be called the *Riemann map* for z_0.
We will also use a slightly different version of Riemann mapping theorem.
A set K is called a *CCF-set* if it is compact, connected, with connected complement (we say then full), containing 0 but not reduced to this point. And its complement $\Omega = \overline{\mathbb{C}} \setminus K$, containing ∞, is called a *CCF-domain*. In order to state a Riemann mapping theorem for this domains we consider the reference CCF-domain $\Delta = \overline{\mathbb{C}} \setminus \overline{\mathbb{D}}$. We will recall holomorphicity at ∞ for a mapping fixing ∞, using the complex structure at ∞.

Definition 1.1.4. If $\Omega = \overline{\mathbb{C}} \setminus K$, where K is a CCF-compact, and $f : \Omega \to \overline{\mathbb{C}} \setminus \{0\}$ is a mapping fixing ∞, we say that f is holomorphic at ∞ if the mapping

$$\tilde{f(z)} = \frac{1}{f(1/z)}$$

is holomorphic at 0.

We now introduce another version of Riemann mapping theorem:

Theorem 1.1.5. *If K is a CCF-compact there exists a unique holomorphic bijection $f : \Delta \to \Omega$ such that $f(\infty) = \infty$ and $f'(\infty) > 0$, we call it the Riemann map.*

In this statement we have defined $f'(\infty) = \lim_{z \to \infty} z f(z)$.

1.1.2 Caratheodory convergence theorem

Definition 1.1.6. Let U_n be a sequence of open sets in $\mathbb{C}$ containing 0. Let V_n be the connected component of the interior of $\bigcap_{k \geq n} U_k$ containing 0. The *kernel* of the sequence is defined to be the union of the V_n's, provided it is non-empty; otherwise it is defined to be $\{0\}$. Thus the kernel is either a connected open set containing 0 or the one point set $\{0\}$.

For the case of domains containing ∞, the definition of the kernel is similar. The kernel of a sequence Ω_n of CCF-domains is the union of all domains $U \subset \overline{\mathbb{C}}$ such that $\infty \in U$ and $U \subset \Omega_n$ for n large enough. If no such domain exists we say that the kernel is $\{\infty\}$.
The sequence is said to converge to a kernel if each subsequence has the same kernel.
We now recall the Caratheodory convergence theorem.

Theorem 1.1.7 (Caratheodory convergence theorem)**.** *Let (f_n) be a sequence of holomorphic univalent functions on the unit disk $\mathbb{D}$, normalized so that $f_n(0) = 0$ and $f_n'(0) > 0$. Then f_n converges uniformly on compacta in $\mathbb{D}$ to a function f if and only if $U_n = f_n(\mathbb{D})$ converges to its kernel and this kernel is not $\mathbb{C}$. If the kernel is $\{0\}$, then $f = 0$. Otherwise the kernel is a connected open set U, f is univalent on $\mathbb{D}$ and $f(\mathbb{D}) = U$.*

There is another version of this theorem for the case of domains containing ∞.

Theorem 1.1.8 (Caratheodory convergence theorem for domains containing ∞)**.** *Let Ω_n be a sequence of CCF-domains and f_n the corresponding sequence of Riemann maps. Then the sequence f_n is uniformly convergent on compact subsets of Δ if and only if Ω_n converges in the sense of Caratheodory to a kernel distinct from $\overline{\mathbb{C}} \setminus \{0\}$. If Ω_n converges and Ω denotes its kernel then*

i. If $\Omega = \infty$ then $f_n \to \infty$ uniformly on compact subsets of Δ.

ii. If $\Omega = \overline{\mathbb{C}} \setminus \{0\}$ then $f_n \to 0$ uniformly on compact subsets of $\Delta \setminus \{\infty\}$.

iii. Otherwise, f_n converges to f, the Riemann mapping of Ω.

1.1.3 Whole-plane Loewner equation

Definition 1.1.9. Let f and g be holomorphic univalent functions on the unit disk $\mathbb{D}$ with $f(0) = 0 = g(0)$. f is said to be *subordinate* to g (denoted by $f \prec g$) if and only if there is a univalent mapping $\varphi : \mathbb{D} \to \mathbb{D}$ fixing 0 such that

$$f(z) = g(\varphi(z))$$

for all $|z| < 1$.

Definition 1.1.10. The family $(f_t)_{t \geq 0}$ of holomorphic univalent mappings from $\mathbb{D}$ to $\mathbb{C}$ is called a ***Loewner chain*** if:

i. $f_t(z) = e^t z + a_2(t)z^2 + a_3(t)z^3 + \cdots$,

ii. $f_s \prec f_t$ if $0 \leq s \leq t$.

We now recall that a Jordan curve or a simple closed curve in the plane $\mathbb{R}^2$ is the image C of an injective continuous map of a circle into the plane, $\varphi : S^1 \to \mathbb{R}^2$. A *Jordan arc* in the plane is the image of an injective continuous map of a closed interval into the plane.
We develop a variant of the Loewner process called whole-plane Loewner process (figure 1.1). We will describe this process in the simple case of a slit domain.
Let a function $\gamma : [0, +\infty) \to \mathbb{C}$ be a Jordan arc joining $\gamma(0)$ to ∞ and not passing through the origin 0. For each $t > 0$, we define a slit domain $\Omega_t = \mathbb{C} \setminus \gamma([t, \infty))$, a simply connected domain containing 0.
We can thus consider a Riemann mapping $f_t : \mathbb{D} \to \Omega_t$ satisfying $f_t(0) = 0$ and $f_t'(0) > 0$. By the Caratheodory convergence theorem it can be seen that $t \mapsto f_t$ is

continuous in the Caratheodory topology. There is no loss of generality in assuming that $f_0'(0) = 1$ and, by changing the time if necessary, that $f_t'(0) = e^t$.
One of the main contributions of Loewner is a considerable strengthening of this latter result. He proved that the map $t \mapsto f_t$ is actually absolutely continuous and in particular differentiable almost everywhere, a fact which makes possible the following statement:

Theorem 1.1.11 (Loewner 1923)**.** *Let $f_t(z)$ defined as above, then there exists a continuous function $\lambda : [0, +\infty) \to \partial\mathbb{D}$ such that almost everywhere in t we have for all $z \in \mathbb{D}$,*

$$\frac{\partial f_t(z)}{\partial t} = z\frac{\partial f_t(z)}{\partial z}\frac{\lambda(t) + z}{\lambda(t) - z},\ z \in \mathbb{D}. \tag{1.1.1}$$

The equation (1.1.1) is called the *Loewner equation* (or the Loewner PDE).
Notice that the slit domain Ω_t implies that for every t, this probability measure must be a Dirac mass at point $\lambda(t) = f_t^{-1}(\gamma(t))$. In this case, it is said that the Loewner chain $(f_t)_{t\geq 0}$ associated with Ω_t is driven by the function $\lambda(t)$.

The importance of Loewner equation lies in the fact that it can be reversed: given a function $\lambda(t)$ which is right continuous with left limits at every point of $\mathbb{R}_+$

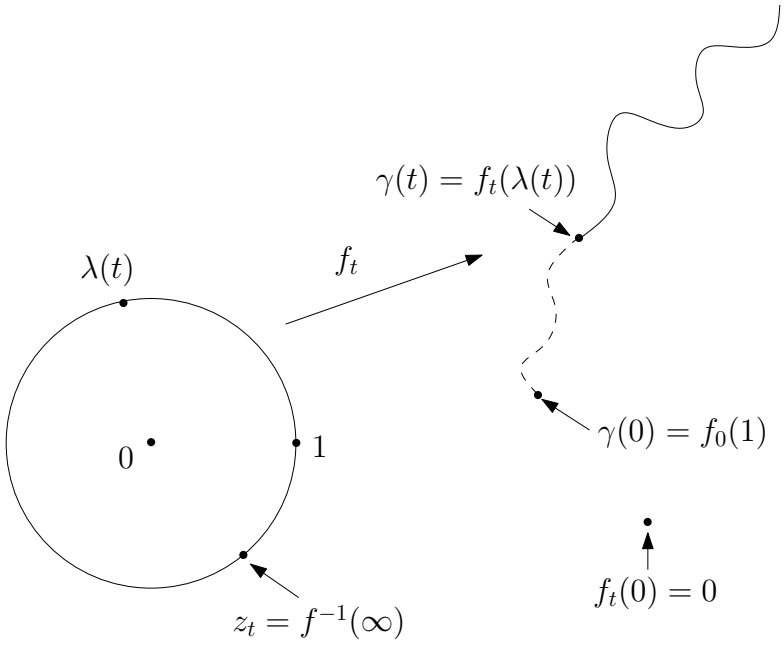


Figure 1.1: Loewner map $f_t(z)$ from the unit disc $\mathbb{D}$ to the slit domain $\Omega_t = \mathbb{C} \setminus \gamma([t, \infty))$ (here slit by a single curve $\gamma([t, \infty))$ for $\kappa \leq 4$) with $f_t(0) = 0$ for all nonnegative t. At $t = 0$, the driving function $\lambda(0) = 1$ so that the image of $z = 1$ is at the tip $\gamma(0) = f_0(1)$ of the curve.

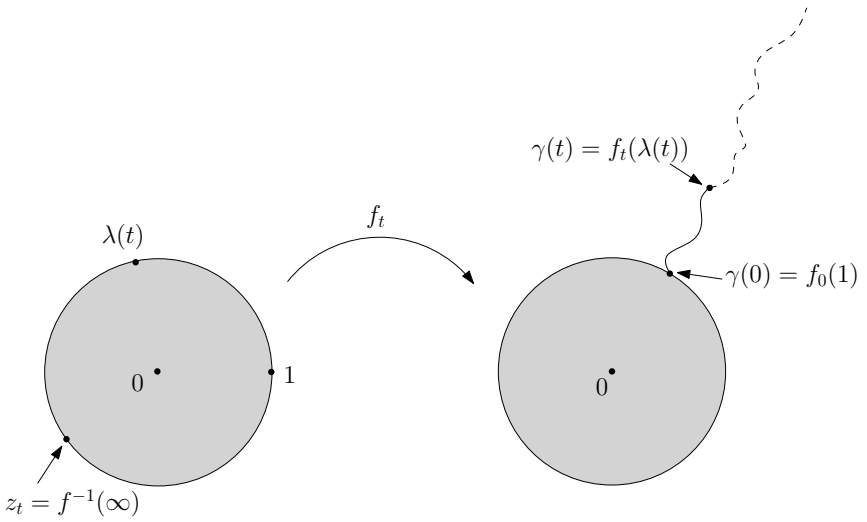


Figure 1.2: Loewner map $f_t(z)$ from CCF-domain $\Delta = \overline{\mathbb{C}} \setminus \overline{\mathbb{D}}$ to the slit domain $\Omega_t = \Delta \setminus \gamma([0,t])$ with $f_t(\infty) = \infty$ for all nonnegative t.

and values in the unit disc, then the equation (1.1.1) has a solution $f_t(z)$ being the Riemann mapping of a domain $\Omega_t = f_t(\mathbb{D})$ and the family $(\Omega_t)_{t>0}$ is increasing in t.

1.1.4 Radial Loewner equation

As in the whole-plane case, we introduce some concepts before defining the Loewner equation.

Definition 1.1.12. Let $f, g : \Delta \to \mathbb{C}$ be two holomorphic univalent functions. We say that f is *subordinate* to g (denoted by $f \prec g$) if and only if there exists $\varphi : \Delta \to \Delta$ holomorphic and fixing ∞ such that

$$f(z) = g(\varphi(z))$$

for all $z \in \Delta$.

Definition 1.1.13. The family $(f_t)_{t\geq 0}$ of holomorphic univalent mappings from Δ to $\mathbb{C}$ is called a *Loewner chain* if:

i. $f_t(z) = e^t z + a_2(t)z^2 + a_3(t)z^3 + \cdots$,

ii. $f_t \prec f_s$ if $0 \leq s \leq t$.

In a similar manner as in the whole-plane case we define the Loewner equation in the radial case as follows (see [Zin07] for more detail):

Let $\gamma : [0, +\infty) \to \Delta$ be a Jordan arc joining $\gamma(0)$ to ∞ and not containing the origin 0. For each $t > 0$, we define a slit domain $\Omega_t = \Delta \setminus \gamma([0,t])$, a CCF-domain containing ∞.

Let f_t be the Riemann map of Ω_t such that $f_t(\infty) = \infty$. We may assume that $f_t'(\infty) = \lim_{z\to\infty} \frac{f_t(z)}{z} = e^t$.
We can see that f_t^{-1} extends by continuity to $\gamma(t)$ and $\lambda(t) = f_t^{-1}(\gamma(t)) \in \partial\Delta$. Moreover the function $t \mapsto \lambda(t)$ may be shown to be continuous and f_t satisfies the Loewner equation

$$\frac{\partial f_t(z)}{\partial t} = -z\frac{\partial f_t(z)}{\partial z}\frac{\lambda(t)+z}{\lambda(t)-z},\ z \in \Delta. \tag{1.1.2}$$

Conversely, given $\lambda : \mathbb{R}+ \to \partial\mathbb{D}$ a continuous function then equation (1.1.2) has a solution $(f_t)_{t>0}$ and f_t maps Δ onto Ω_t, where $(\Omega_t)_{t>0}$ is decreasing.

1.1.5 Chordal Loewner equation

The Loewner processes driven by a regulated function that we have encountered so far were the radial ones, starting at a point on the boundary of Δ and going to ∞ which is an interior point of Δ. The variant of Loewner process we want to introduce here starts at a boundary point and head on to another boundary point. The convenient geometry for this new family of processes is the upper half-plane $\{y > 0\}$, the starting and target points being respectively 0 and ∞.

Let $\gamma : \mathbb{R}_+ \to \{y \geq 0\}$ be continuous and injective such that $\gamma(0) = 0$ and $\gamma(t) \in \{y > 0\}$ for $t > 0$. Let $\Omega_t = \{y > 0\} \setminus \gamma([0,t])$. There exists unique conformal map $g_t : \Omega_t \to \{y > 0\}$ such that $g_t(z) = z + \frac{k(t)}{z} + ...$ and by changing time we may assume that $k(t) = 2t$. The map g_t satisfies the Loewner differential equation

$$\frac{\partial g_t(z)}{\partial t} = \frac{2}{g_t(z) - \lambda(t)}$$

where $\lambda(t) = g_t(\gamma(t))$, and $f_t = g_t^{-1}$ satisfies the following equation

$$\frac{\partial f_t(z)}{\partial t} = \frac{\partial f_t(z)}{\partial z}\frac{2}{\lambda(t)-z}. \tag{1.1.3}$$

The function λ is continuous from $\mathbb{R}_+$ into $\mathbb{R}$. As in the radial case we will say that the Loewner process is driven by $\lambda(t)$.
Conversely, given a continuous function $\lambda : \mathbb{R}_+ \to \mathbb{R}$, the equation (1.1.3) can be solved and f_t represents a normalized, in the sense of $g_t(z) = z + \frac{k(t)}{z} + ...$, Riemann map from $\{y > 0\}$ onto Ω_t and $(\Omega_t)_{t>0}$ decreases with t.

1.2 Stochastic Loewner Evolution

The stochastic Loewner evolution (or Schamm-Loewner evolution) with parameter κ (SLE$_\kappa$), discovered by Oded Schramm (2000), is a family of random planar

curves that have been proven to be the scaling limit of a variety of two-dimensional lattice models in statistical mechanics. Given a parameter κ and a domain in the complex plane U, it gives a family of random curves in U, with κ controlling how much the curve turns. SLE_κ is the Loewner process driven by the function

$$\lambda(t) = e^{i\sqrt{\kappa}B_t} \tag{1.2.1}$$

in the whole-plane and radial case, and

$$\lambda(t) = \sqrt{\kappa}B_t \tag{1.2.2}$$

in the chordal case. $\kappa \in [0, \infty)$ and B_t is a standard, one-dimensional Brownian motion characterized by the three fundamental properties:

i. Stationarity: if $0 \leq s \leq t$, then $B_t - B_s$ has the same law as B_{t-s};

ii. Markov property: if $0 \leq s \leq t$, then $B_t - B_s$ is independent of B_s;

iii. Gaussianity: B_t has a normal distribution with mean 0 and variance t.

SLE_κ is conjectured or proved to describe the scaling limit of various stochastic processes in the plane, the most famous one is critical percolation which is introduced in the next section.
Figure 1.3 gives samples of chordal SLE_κ for 3 different values of κ. The SLE_2-curve (figure 1.3a) corresponds to the loop-erased random walk and it is a simple curve. The SLE_4-curve in (figure 1.3b) is still a simple one. In fact, for all $0 \leq \kappa \leq 4$, the curve is almost surely simple (does not intersect itself) and only intersects the real axis at $t = 0$. Meanwhile, in SLE_6 (figure 1.3c), the curve intersects the real axis and intersects itself. In this case, $\Omega_t = \{y > 0\} \setminus K_t$, where K_t is the bounded component of $\{y > 0\} \setminus \gamma([0, t])$. The process SLE_6 has been proven to be the scaling limit of critical percolation on the triangular lattice [Smi01].

A Lévy process is a stochastic process that is only assumed to satisfy the first two of above properties of Brownian motion. The essential difference with Brownian motion is that jumps are then allowed. The characteristic function of a Lévy process L_t has the form

$$\mathbb{E}(e^{i\xi L_t}) = e^{-t\eta(\xi)} \tag{1.2.3}$$

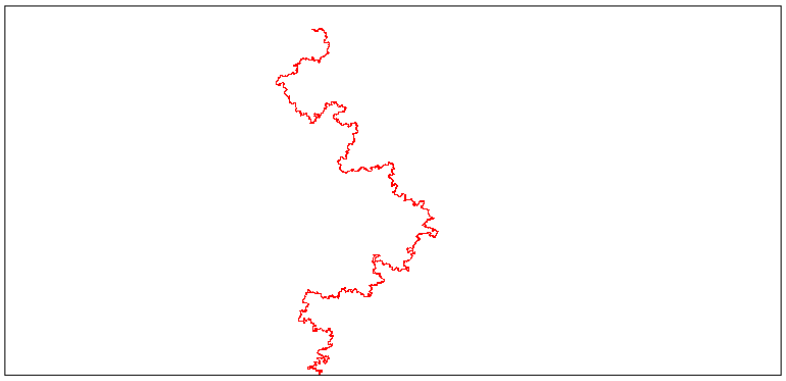
where η (called the Lévy symbol) is a continuous complex function of $\xi \in \mathbb{R}$, satisfying $\eta(0) = 0$ and $\eta(-\xi) = \overline{\eta(\xi)}$. SLE_κ corresponds to a Gaussian characteristic function and is a Lévy process with symbol

$$\eta(\xi) = \kappa\xi^2/2. \tag{1.2.4}$$

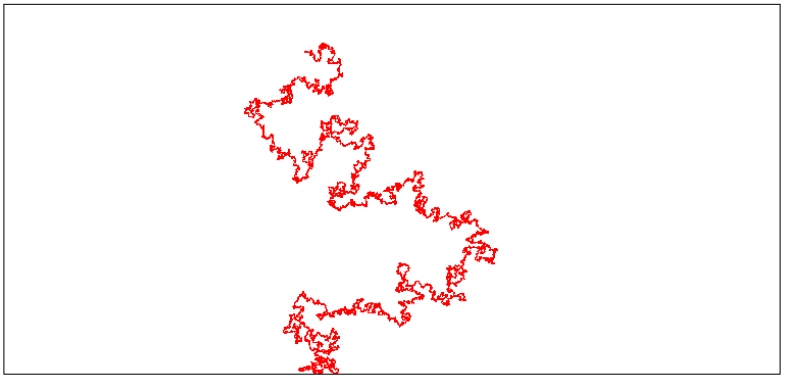
More generally, the function

$$\eta(\xi) = \kappa|\xi|^\alpha/2, \ \alpha \in (0, 2] \tag{1.2.5}$$

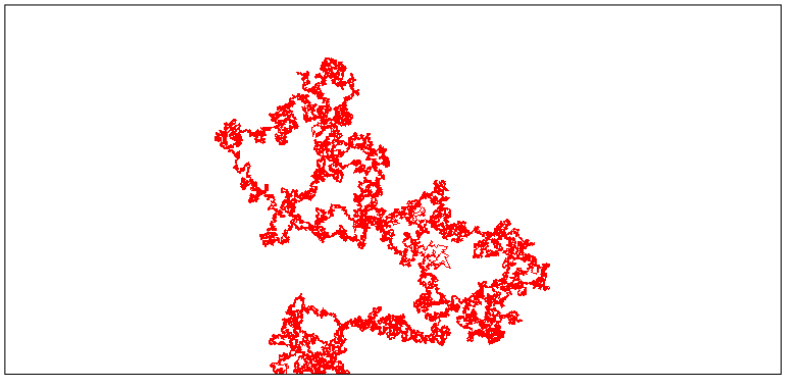
is the Lévy symbol of the so-called $\alpha-$*stable process*. As we shall see, all the quantities that we will consider depend only on the values of the Lévy symbol at *integer*



(a) Chordal SLE_2



(b) Chordal SLE_4



(c) Chordal SLE_6

Figure 1.3: Chordal SLE_κ. Source: Tom Kennedy's website [Ken].

arguments; for this reason we shall use the "sequence" notation: $\eta_k := \eta(k)$, $k \in \mathbb{Z}$. In the whole-plane SLE_κ, the associated conformal maps, obeying (1.1.1), are denoted by f_t, and in this work, we study their coefficients $a_n(t)$, which are random variables, defined by the normalized series expansion:

$$f_t(z) = e^t \left(z + \sum_{n \geq 2} a_n(t) z^n \right). \tag{1.2.6}$$

In the next part we start with the computation, in terms of the Lévy symbols $\eta_k, k \in \mathbb{Z}$, of $\mathbb{E}(a_n)$ for all n, and of $\mathbb{E}(|a_n|^2)$ for small n, for a general Lévy-Loewner evolution process f_t. We firstly introduce a discrete process (percolation) that is related to SLE_κ with $\kappa = 6$.

1.3 Percolation

The percolation exploration path was introduced by Schramm in 1999 [Sch00]. He proposed that this path converges in distribution to the trace of chordal SLE_6. This conjecture was proved by Smirnov [Smi01] for site percolation on triangular lattice in 2001. Combined with earlier work of Harry Kesten [Kes87], this led to the determination of many of the critical exponents for percolation [SW01]. This breakthrough, in turn, allowed to further analyze many aspects of this model. We will give the definitions below to study more this problem.
Consider a two-dimensional triangular lattice $\mathcal{T}$, we can think of its sites as the elementary cells of a regular hexagonal lattice $\mathcal{H}$. Each site of $\mathcal{T}$ is colored blue or yellow. We say that two hexagons are *neighbors* if they have a common edge. A sequence of hexagons with two consecutive sites are neighbors will be called a *$\mathcal{T}$-path*. If its first and last hexagon are neighbors, the path will be called a *$\mathcal{T}$-loop*. A set D of hexagons is *connected* if any two of them can be joined by a $\mathcal{T}$-path contained in D. A finite set D of hexagons is called *simply connected* if both D and $\mathcal{T} \setminus D$ are connected. For a simply connected set D, the set of the hexagons that do not belong to D but are adjacent to hexagons in D is called *external site boundary* or *s-boundary* denoted by ΔD; and when D is considered as a domain of $\mathbb{C}$, its topological boundary is denoted ∂D. If a simply connected set D has ΔD being a $\mathcal{T}$-loop it will be called a *Jordan set*.
For a Jordan set $D \subset \mathcal{T}$, a vertex $x \in \partial D$ is called an *e-vertex* if there exists an edge incident on $x \notin \partial D$ does not belong to a hexagon in D. Given two e-vertices x, y in ∂D, the counterclockwise path from x to y of ∂D, denoted by $\partial_{x,y} D$, is called the *right boundary* and the remaining part, denoted by $\partial_{y,x} D$, is the *left boundary*. Analogously for ΔD, $\Delta_{x,y} D$ is the *right s-boundary* and $\Delta_{y,x} D$ is the *left s-boundary*. A *cluster* is a maximal, connected, monochromatic subset of $\mathcal{T}$; we will distinguish between blue and yellow clusters (figure 1.4). A *boundary* of a cluster is the set of edge of $\mathcal{H}$ that surround the cluster (the thick black path in figure 1.4). We can see that the edges in this boundary are always between two different color hexagons.

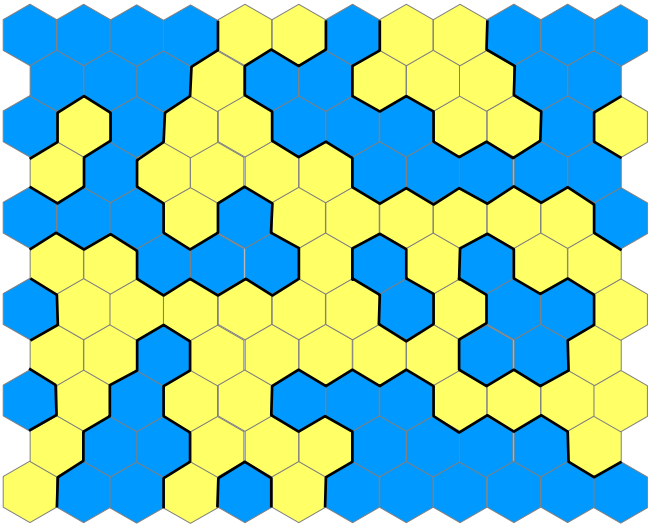


Figure 1.4: The clusters on the triangular lattice $\mathcal{T}$ with each site is represented by a hexagon of the hexagonal lattice $\mathcal{H}$. In the critical percolation, choosing the color (blue or yellow) of each site is randomly with equal probability. The thick black paths are the cluster boundaries; there are some small loops, the other boundaries extend beyond the finite window.

A boundary path γ is a sequence of distinct edges belonging to the boundary of a cluster such that two consecutive edges have a common vertex. For each edge in boundary of cluster we can thus associate a direction in such a way that the hexagon to the right of edge with the direction is blue.

For a Jordan set D with ∂D defined as the boundary of a cluster and two e-vertices x, y in ∂D, there is a unique boundary path γ from x to y (see figure 1.5) which separates the blue and the yellow clusters. We call it a percolation exploration path in D from x to y and denote by $\gamma_{D,x,y}$. Notice that $\gamma_{D,x,y}$ does not depend on the color of hexagons in ΔD, hence we can color blue all hexagons in $\Delta_{x,y}D$ and yellow all those in $\Delta_{y,x}D$. This path is constructed dynamically by the process defined below.

We will start with the edge $e_x \notin \partial D$ incident in x, oriented in the direction of x. At each step there are two edges to choose (left or right with respect to the current direction) and both of them belong to the same hexagon ξ. If ξ is not colored we will decide its color by tossing a fair coin and then if ξ is blue the edge to the left is chosen and if ξ is yellow the edge to the right is chosen. If the hexagon we arrive on is already colored, the continuous edge will be chosen according to the rule above. The exploration stops when it reaches y.

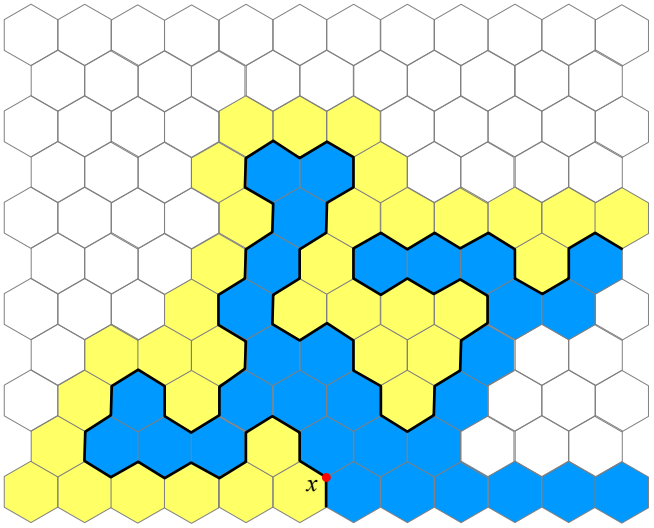


Figure 1.5: A percolation exploration process starts at x in a domain of the hexagonal lattice with the bottom row is in the s-boundary (with the blue right and yellow left). The hexagons (not in the bottom row) are colored during the exploration process. The thick path between yellow and blue hexagons is the exploration path.

Chapter 2

The Coefficient Problem of Stochastic Loewner Evolution

2.1 Loewner's method

As we mentioned in the last section, Loewner introduced his famous equation in order to solve Bieberbach conjecture for $n = 3$. Let us recall what Bieberbach conjecture is.

Consider a holomorphic injective (univalent) function in the open unit disk $\mathbb{D} := \{z : |z| < 1\}$ with Taylor series of the form

$$f(z) = \sum_{n\geq 0} a_n z^n.$$

What can we say about the coefficients a_n? In 1916, Bieberbach conjectured that

$$|a_n| \leq n|a_1|, \tag{2.1.1}$$

and proved the inequality for $n = 2$.
This conjecture (2.1.1) was proven by Louis de Branges in 1985. In 1923, Charles Loewner proved $|a_3| \leq 3|a_1|$ using his (Loewner) equation and the following method.

By expanding both sides of Loewner's equation (1.1.1) as power series, and identifying coefficients, for all $n \geq 2$, one leads to the set of equations

$$\begin{aligned} \dot{a}_n(t) - (n-1)a_n(t) &= 2\sum_{p=1}^{n-1}(n-p)a_{n-p}(t)\bar{\lambda}^p(t) \\ &= 2\sum_{k=1}^{n-1} k a_k(t)\bar{\lambda}^{n-k}(t); \end{aligned} \tag{2.1.2}$$

the dot means a t-derivative, and $\bar{\lambda}(t)$ means the complex conjugate of $\lambda(t)$. Specifying for $n = 2, 3$ gives

$$\dot{a}_2 - a_2 = 2\bar{\lambda}, \tag{2.1.3}$$
$$\dot{a}_3 - 2a_3 = 4a_2\bar{\lambda} + 2\bar{\lambda}^2. \tag{2.1.4}$$

The first differential equation (2.1.3) (together with the uniform bound, $\forall t \geq 0, |a_2(t)| \leq C_2 < +\infty$), yields

$$a_2(t) = -2e^t \int_t^{+\infty} e^{-s}\bar{\lambda}(s)ds. \tag{2.1.5}$$

This gives a proof for the $n = 2$ case of Bieberbach conjecture. In a similar way, the second one (2.1.4) leads to

$$a_3(t) = -4e^{2t} \int_t^{+\infty} e^{-2s}a_2(s)\bar{\lambda}(s)ds - 2e^{2t} \int_t^{+\infty} e^{-2s}\bar{\lambda}^2(s)ds,$$

We can see that the first integral is of the form $\int_t^{\infty} u_2(s)\dot{u}_2(s)ds = -u_2^2(t)/2$, where $u_2(s) := e^{-s}a_2(s)$. The formula for a_3 then reduces to

$$a_3(t) = 4e^{2t} \left(\int_t^{+\infty} e^{-s}\bar{\lambda}(s)ds \right)^2 - 2e^{2t} \int_t^{+\infty} e^{-2s}\bar{\lambda}^2(s)ds. \tag{2.1.6}$$

From this expression, we can show that $|a_3| \leq 3$ (see Appendix 5.1.2 in [DNNZ12] for detail) and get estimates of coefficients for small order in the next section.

2.2 The coefficients for small order

The aim of this chapter is to revisit coefficient problem for SLE or LLE processes. The coefficients $a_n := a_n(0)$ of these random processes then become random variables and we wish to compute their expectation and their second moment (from which we may derive the variance), $\mathbb{E}(a_n)$, $\mathbb{E}\left(|a_n|^2\right)$.

2.2.1 Expectation of $f_0(z)$

In this first section, we give an explicit expression for the expectations of the coefficients $a_n(t)$ of the expansion (1.2.6) in the Lévy-Loewner setting, thereby obtaining the expectation of the map, $[f_0]$, and its derivative.
The differential recursion (2.1.2) in previous section then becomes, for $\lambda(t) := e^{iL_t}$, and in terms of the auxiliary function $u_n(t)$,

$$u_n(t) := a_n(t)e^{-(n-1)t} \tag{2.2.1}$$
$$\dot{u}_n(t) = 2\sum_{k=1}^{n-1} kX_t^{n-k}u_k(t), \tag{2.2.2}$$

where X_t is defined as

$$X_t := e^{-t-iL_t}. \tag{2.2.3}$$

The recursion (2.2.2) can be rewritten under the simpler form:

$$\dot{u}_n = X_t[\dot{u}_{n-1} + 2(n-1)u_{n-1}]. \tag{2.2.4}$$

Recall that $u_1 = a_1 = 1$, while the next term of this recursion, as already seen in Eqs. (2.1.5), is

$$u_2(t) = -2\int_t^{+\infty} ds X_s. \tag{2.2.5}$$

Similarly, we can write the general solution u_n, for $n \geq 2$, under the form

$$u_n(t) = -2\int_t^{+\infty} ds X_s v_n(s), \tag{2.2.6}$$

with $v_2(s) = 1$, and rewrite the differential equation (2.2.4) as an *integral* equation

$$v_n(t) = X_t v_{n-1}(t) - 2(n-1)\int_t^{+\infty} ds X_s v_{n-1}(s). \tag{2.2.7}$$

Define then the multiplicative and integral operators $\mathcal{X}$ and $\mathcal{J}$ such that

$$\mathcal{X}v(t) := X_t v(t), \tag{2.2.8}$$

$$\mathcal{J}v(t) := -2\int_t^{+\infty} ds X_s v(s). \tag{2.2.9}$$

The solution to (2.2.5), (2.2.6) and (2.2.7) can then be written as the operator product

$$\begin{aligned} u_n &= \mathcal{J} \circ [\mathcal{X} + (n-1)\mathcal{J}] \circ \cdots \circ (\mathcal{X} + 2\mathcal{J})\mathbb{1} \\ &= \mathcal{J}\prod_{k=1}^{n-2} \circ (\mathcal{X} + (k+1)\mathcal{J})\mathbb{1}, \end{aligned} \tag{2.2.10}$$

where $\mathbb{1}(= v_2)$ is the constant function equal to 1 on $\mathbb{R}^+$.

Next, recall the strong Markov property of the Lévy process, which implies the *identity in law:* $\forall s \geq t, L_s \overset{(\text{law})}{=} L_t + \tilde{L}_{s-t}$, where $\tilde{L}_{s'}$ is an independent copy of the Lévy process, also started at $\tilde{L}_0 = 0$. Therefore, the process X_t (2.2.3) is, in law,

$$X_s \overset{(\text{law})}{=} X_t \tilde{X}_{s-t}, \forall s \geq t \tag{2.2.11}$$

where $\tilde{X}_{s'} := e^{-s'-\tilde{L}_{s'}}, s' \geq 0$, is an independent copy of that process, with $\tilde{X}_0 = 1$. The operator $\mathcal{J}$ (2.2.9) can then be written as

$$\mathcal{J}v(t) \overset{(\text{law})}{=} -2X_t\int_0^{+\infty} ds \tilde{X}_s v(s+t) \tag{2.2.12}$$

$$= \mathcal{X} \circ \tilde{\mathcal{J}}v(t), \tag{2.2.13}$$

with $\tilde{\mathcal{J}}v(t) := -2\int_0^{+\infty} ds\tilde{X}_s v(s+t)$. By iteration of the use of the Markov property, Eq. (2.2.10) can be rewritten as

$$\begin{aligned} u_n &\overset{(\text{law})}{=} \mathcal{J}\circ\big[\mathcal{X}\big(1+(n-2)\tilde{\mathcal{J}}^{[n-1]}\big)\big]\circ\cdots\circ\big[\mathcal{X}\big(1+2\tilde{\mathcal{J}}^{[1]}\big)\big]\mathbb{1} \\ &\overset{(\text{law})}{=} \mathcal{J}\prod_{k=1}^{n-2}\circ[\mathcal{X}\big(1+(k+1)\tilde{\mathcal{J}}^{[k]}\big)]\mathbb{1}, \end{aligned} \tag{2.2.14}$$

where the integral operators $\tilde{\mathcal{J}}^{[k]}, k=1,\cdots,n-2$, involve successive *independent* copies, $\tilde{X}^{[k]}_{s_k}, k=1,\cdots,n-2$, of the original exponential Lévy process X_s. We therefore arrive at the following explicit representation of the solution (2.2.10)

$$u_n(t) \overset{(\text{law})}{=} -2\int_t^{+\infty} dsX_s^{n-1}\prod_{k=1}^{n-2}\left(1-2(k+1)\int_0^{+\infty} ds_k\big(\tilde{X}^{[k]}_{s_k}\big)^k\right). \tag{2.2.15}$$

As mentioned in the introduction, the *conjugate* whole-plane Lévy-Loewner evolution $e^{-iL_t}f_t\big(e^{iL_t}z\big)$ should have the *same law* as $f_0(z)$. At order n, we are thus interested in the stochastically rotated coefficients:

$$e^{i(n-1)L_t}a_n(t) = (X_t)^{-(n-1)}u_n(t).$$

Using again the identity in law (2.2.11) in (2.2.15), we arrive at

$$\begin{aligned} e^{i(n-1)L_t}a_n(t) &\overset{(\text{law})}{=} \\ &-2\int_0^{+\infty} ds\tilde{X}_s^{n-1}\prod_{k=1}^{n-2}\left(1-2(k+1)\int_0^{+\infty} ds_k\big(\tilde{X}^{[k]}_{s_k}\big)^k\right) \\ &\overset{(\text{law})}{=} a_n(0), \end{aligned} \tag{2.2.16}$$

which, as it must, no longer depends of t.
All factors in (2.2.16) involve successive independent copies of the Lévy process, and their expectations can now be taken *independently*. Recalling the form (1.2.3) of the Lévy characteristic function, we have $\mathbb{E}[(\tilde{X}_s)^k] = e^{-(\eta_k+k)s}$. Thus

$$\begin{aligned} \mathbb{E}[a_n(0)] &= -2\int_0^{+\infty} ds\,\mathbb{E}[\tilde{X}_s^{n-1}]\prod_{k=1}^{n-2}\left(1-2(k+1)\int_0^{+\infty} ds_k\big(\mathbb{E}\big[\big(\tilde{X}^{[k]}_{s_k}\big)^k\big]\right) \\ &= -2\,\frac{1}{\eta_{n-1}+n-1}\prod_{k=1}^{n-2}\left(1-\frac{2(k+1)}{\eta_k+k}\right). \end{aligned} \tag{2.2.17}$$

We finally obtain:

Theorem 2.2.1. *For $n\geq 2$, setting $a_n := a_n(0)$,*

$$a_n(0) \overset{(\text{law})}{=} e^{i(n-1)L_t}a_n(t), \tag{2.2.18}$$

$$\mathbb{E}(a_n) = -2\frac{\prod_{k=1}^{n-2}(\eta_k-k-2)}{\prod_{k=1}^{n-1}(\eta_k+k)}. \tag{2.2.19}$$

Corollary 2.2.1. *The expected conformal map $\mathbb{E}[f_0(z)]$ of the whole-plane Lévy-Loewner evolution, in the setting of Theorem 2.2.1, is polynomial if there exists a positive k such that $\eta_k = k+2$, has radius of convergence 1 for an α-stable Lévy process of symbol $\eta_n = \kappa n^\alpha/2$, $\alpha \in (0,2]$, except for the Cauchy process $\alpha = 1, \kappa = 2$, where $\mathbb{E}[f_0(z)] = ze^{-z}$.*

Proof. From Theorem 2.2.1, $\mathbb{E}[f_0(z)]$ is polynomial if there exists $k \in \mathbb{N}$ such that $\eta_k = k+2$, as all $\mathbb{E}(a_n)$ then vanish for $n \geq k+2$. Otherwise, use D'Alembert's criterion, applied here to

$$\lim_{n\to\infty} \frac{|\mathbb{E}(a_{n+1})|}{|\mathbb{E}(a_n)|} = \lim_{n\to\infty} \frac{|\eta_{n-1} - n - 1|}{|\eta_n + n|} = 1,$$

for an α-stable symbol, $\eta_n = \kappa|n|^\alpha/2, \forall \alpha \in (0,2]$, except if $\alpha = 1$ and $\kappa = 2$, for which the limit vanishes. In that case, Eq. (2.2.19) gives $\mathbb{E}(a_n) = (-1)^{n-1}/(n-1)!$ for $n \geq 2$, thus $\mathbb{E}[f_0(z)] = ze^{-z}$ and $\mathbb{E}[f_0'(z)] = (1-z)e^{-z}$. □

In the last paragraph we have computed $\dot{u_2} = 2X_t$ and we deduce that

$$\mathbb{E}(u_2) = -\frac{2}{1+\eta_1},$$

which is the induction hypothesis for $j = 2$. Assume that the formula is valid up to $n-1$: by the results of the preceding section and with the notations there, $\mathbb{E}(u_{n-1})$ may be written as a linear combination of integrals that we have denoted by

$$(\alpha_1, \alpha_1) \dots (\alpha_j, \alpha_j),$$

i.e.

$$\mathbb{E}(u_{n-1}) = \sum_I c_I (\alpha_1, \alpha_1) \dots (\alpha_j, \alpha_{j_I}),$$

where $\alpha_1 + \ldots + \alpha_j = n-2$. Using (2.2.2) we may write

$$u_n = -\int_t^\infty X_s \dot{u}_{n-1} ds - 2(n-1) \int_t^\infty X_s u_{n-1} ds$$

and

$$\begin{aligned} -\mathbb{E}(\int_t^\infty X_s \dot{u}_{n-1} ds) &= \sum c_I (\alpha_1 + 1, \alpha_1 + 1)(\alpha_2, \alpha_2) \dots (\alpha_{j_I}, \alpha_{j_I}) \\ &= \sum c_I \frac{(\alpha_2, \alpha_2) \dots (\alpha_{j_I}, \alpha_{j_I})}{\eta_{n-1} + n - 1} \\ &= \frac{\eta_{n-2} + n - 2}{\eta_{n-1} + n - 1} \mathbb{E}(u_{n-1}). \end{aligned}$$

We compute with the same method

$$-2(n-1)\mathbb{E}(\int_t^\infty X_s u_{n-1} ds) = -2(n-1)\frac{\mathbb{E}(u_{n-1})}{\eta_{n-1} + n - 1}$$

and the proof follows.

2.2.2 Moments of order 2

Proposition 2.2.2. *For Lévy-Loewner processes, we have, setting here* $a_2 := a_2(0)$,

$$\mathbb{E}\left(|a_2|^2\right) = \frac{4}{1+\eta_1}.$$

In SLE case

$$\mathbb{E}\left(|a_2|^2\right) = \Re\left(\frac{8}{2+\kappa}\right).$$

Proof. It is sufficient to prove the LLE case.
From (2.1.5), we can write

$$\begin{aligned}|a_2|^2 &= 2\int_0^\infty e^{-s+iL_s}\int_s^\infty e^{-s'-iL_{s'}}ds'ds + 2\int_0^\infty e^{-s+iL_s}\int_0^s e^{-s'-iL_{s'}}ds'ds\\ &= 2\int_0^\infty e^{-s}\int_s^\infty e^{-s'-i(L_{s'}-L_s)}ds'ds + 2\int_0^\infty e^{-s}\int_0^s e^{-s'+i(L_s-L'_s)}ds'ds.\end{aligned}$$

Using now the characteristic function (1.2.3) for $L_s - L_{s'}$, and taking care of the relative order of s and s', we get

$$\mathbb{E}\left(|a_2|^2\right) = 2\int_0^\infty e^{-s}\int_s^\infty e^{-s'-(s'-s)\eta_1}ds'ds + 2\int_0^\infty e^{-s}\int_0^s e^{-s'-(s-s')\overline{\eta_1}}ds'ds,$$

and the result follows. □

For calculations involving the third order term a_3 as given by (2.1.6), and in order to avoid repetitions, we have computed at once $\mathbb{E}(|a_3 - \mu a_2^2|^2)$, where μ is a real constant. Let us state here the result.

Proposition 2.2.3. *If* μ *is a real coefficient then*

$$\mathbb{E}\left(\left|a_3 - \mu a_2^2\right|^2\right) =$$
$$\Re\left(\frac{16(1-\mu)^2(4+\eta_2)}{(1+\eta_1)(2+\eta_2)(3+\eta_1)} - \frac{16(1-\mu)(2+\eta_1)}{(1+\eta_1)(2+\eta_2)(3+\eta_1)} + \frac{2}{2+\eta_2} + \frac{8(1-\mu)(1-2\mu)}{(\overline{\eta_1}+1)(\eta_1+3)}\right).$$

In the real η *case:*

$$\mathbb{E}\left(\left|a_3 - \mu a_2^2\right|^2\right) = \frac{32(1-\mu)^2(3+\eta_2) - 8(1-\mu)(6+2\eta_1+\eta_2) + 2(1+\eta_1)(3+\eta_1)}{(1+\eta_1)(2+\eta_2)(3+\eta_1)}.$$

In the SLE case (i.e., for $\eta_\ell = \frac{\kappa}{2}\ell^2$*):*

$$\mathbb{E}\left(\left|a_3 - \mu a_2^2\right|^2\right) = \frac{(108 - 288\mu + 192\mu^2) + (88 - 208\mu + 128\mu^2)\kappa + \kappa^2}{(1+\kappa)(2+\kappa)(6+\kappa)}.$$

Proof. We write

$$e^{-4t}\left|a_3-\mu a_2^2\right|^2 = 16(1-\mu)^2 I_1 - 16(1-\mu)\Re I_2 + 4I_3,$$

where

$$\begin{aligned}
I_1 &= \int_t^\infty\int_t^\infty\int_t^\infty\int_t^\infty e^{-(s_1+s_2+s_3+s_4)}\overline{\lambda}(s_1)\lambda(s_2)\overline{\lambda}(s_3)\lambda(s_4)ds_1ds_2ds_3ds_4,\\
I_2 &= \int_t^\infty\int_t^\infty\int_t^\infty e^{-(s_1+s_2+2s_3)}\overline{\lambda}(s_1)\overline{\lambda}(s_2)\lambda(s_3)^2ds_1ds_2ds_3,\\
I_3 &= \int_t^\infty\int_t^\infty e^{-2(s_1+s_2)}\overline{\lambda}(s_1)^2\lambda(s_2)^2ds_1ds_2.
\end{aligned}$$

From now on we set the parameter $t=0$ in the above formula. The computation of I_3 follows the same lines as the one in Proposition 2.2.2 and we find

$$\mathbb{E}\left(I_3\right) = \Re\left(\frac{1}{2\left(2+\eta_2\right)}\right).$$

In order to compute $\mathbb{E}(I_2)$ we have to use the strong Markov property. First, by symmetry, we may write

$$I_2 = 2\int_{s_1=0}^\infty\int_{s_2=s_1}^\infty\int_{s_3=0}^\infty e^{-(s_1+s_2+2s_3)}e^{i(L_{s_3}-L_{s_1})}e^{i(L_{s_3}-L_{s_2})}ds_1ds_2ds_3,$$

and we cut this integral into $I_2 = 2(I_{2,1}+I_{2,2}+I_{2,3})$, where in $I_{2,1}$ (resp. in $I_{2,2}, I_{2,3}$), s_3 lies in $[0,s_1]$ (resp. in $[s_1,s_2],[s_2,\infty)$).
For $I_{2,1}$, write

$$e^{i(L_{s_3}-L_{s_1})}e^{i(L_{s_3}-L_{s_2})} = e^{-2i(L_{s_1}-L_{s_3})}e^{-i(L_{s_2}-L_{s_1})},$$

so that the Markov property can be used to get its expectation: $e^{-\overline{\eta_2}(s_1-s_3)}e^{-\overline{\eta_1}(s_2-s_1)}$. From this, the value of $\mathbb{E}(I_{2,1})$ easily follows as

$$\mathbb{E}\left(I_{2,1}\right) = \frac{1}{4(1+\overline{\eta_1})(2+\overline{\eta_2})}.$$

Similar considerations lead to

$$\begin{aligned}
\mathbb{E}\left(I_{2,2}\right) &= \frac{1}{4(1+\overline{\eta_1})(3+\eta_1)},\\
\mathbb{E}\left(I_{2,3}\right) &= \frac{1}{4(2+\eta_2)(3+\eta_1)}.
\end{aligned}$$

Combining these computations, we get

$$\Re\mathbb{E}\left(I_2\right) = \Re\left(\frac{1}{2(1+\eta_1)(2+\eta_2)} + \frac{1}{2(1+\overline{\eta_1})(3+\eta_1)} + \frac{1}{2(2+\eta_2)(3+\eta_1)}\right).$$

The computation of I_1 follows the same lines. First, by symmetry,

$$I_1 = 4\int_0^\infty\int_{s_1}^\infty\int_0^\infty\int_{s_3}^\infty e^{-(s_1+s_2+s_3+s_4)}e^{i(L_{s_3}-L_{s_1})}e^{i(L_{s_4}-L_{s_2})}ds_1ds_2ds_3ds_4.$$

We then split this integral into the sum of six pieces associated with the domains

I. $s_3 < s_4 < s_1 < s_2$;

II. $s_3 < s_1 < s_4 < s_2$;

III. $s_3 < s_1 < s_2 < s_4$;

IV. $s_1 < s_3 < s_4 < s_2$;

V. $s_1 < s_3 < s_2 < s_4$;

VI. $s_1 < s_2 < s_3 < s_4$.

Clearly, the respective contributions of (I) and (VI), (II) and (V), (III) and (IV), are complex conjugate of each other. The same arguments as above give, in a short-hand notation,

$$\begin{aligned}\mathbb{E}(\mathrm{I}) &= \frac{1}{4(1+\overline{\eta_1})(2+\overline{\eta_2})(3+\overline{\eta_1})},\\ \mathbb{E}(\mathrm{II}) &= \frac{1}{8(1+\overline{\eta_1})(3+\overline{\eta_1})},\\ \mathbb{E}(\mathrm{III}) &= \frac{1}{8(1+\eta_1)(3+\overline{\eta_1})}.\end{aligned}$$

Altogether, we get

$$\mathbb{E}\left(I_1\right) = \Re\left(\frac{2}{(1+\eta_1)(2+\eta_2)(3+\eta_1)} + \frac{1}{(1+\eta_1)(3+\eta_1)} + \frac{1}{(1+\overline{\eta_1})(3+\eta_1)}\right).$$

□

With $\mu = 0$ we immediately get the following corollary,

Corollary 2.2.2. *For Lévy-Loewner processes with η real, we have*

$$\mathbb{E}(|a_3|^2) = \frac{1}{(1+\eta_1)(3+\eta_1)}\left[24 + 2\frac{(\eta_1-1)(\eta_1-3)}{2+\eta_2}\right]. \tag{2.2.20}$$

In the SLE case:

$$\mathbb{E}(|a_3|^2) = \frac{108+88\kappa+\kappa^2}{(1+\kappa)(2+\kappa)(6+\kappa)}.$$

From these results we noticed that: $\mathbb{E}\left(a_2\right) = \mathbb{E}\left(a_3\right)$=1 for $\eta_1 = 3$, corresponding to $\kappa = 6$.

Remark 2.2.4. In the second term of the expression of $\mathbb{E}\left(|a_n|^2\right)$ (with $n = 2, 3$), we note the presence of the factors $\left(\eta_1 - 1\right)\left(\eta_1 - 3\right)$. Therefore, for $n = 2, 3$, with $\eta_1 = 1$ or 3 (corresponding to $\kappa = 2$ or 6), the result no longer depends on η_2 and

i.

$$\mathbb{E}\left(|a_n|^2\right) = 1,$$

for $\eta_1 = 3$ ($\kappa = 6$) and

ii.

$$\mathbb{E}\left(|a_n|^2\right) = n,$$

for $\eta_1 = 1$ ($\kappa = 2$).

These patterns were further checked by calculations done by hand for $n = 4$ and later were confirmed at higher orders that may be reached by computations using computer, as we shall see in the following section in the following section.

2.3 Computational experiments

As one may see, these computations become more and more involved. Moreover, it seems difficult to find a closed formula for all terms. This section is devoted to the description of an algorithm that we have implemented on MATLAB to compute $\mathbb{E}(|a_n|^2)$. This algorithm is divided into two parts: the first encodes the computation of a_n, while the second uses it to compute $\mathbb{E}(|a_n|^2)$. Since the important cases of SLE and α-stable processes both have real Lévy symbols η, we restrict the study to the latter case.
For the encoding of a_n, we observe that they are linear combinations of successive integrals of the form

$$\int_t^\infty ds_1\, e^{-i\alpha_1 L_{s_1} - \beta_1 s_1} \int_{s_1}^\infty ds_2\, e^{-i\alpha_1 L_{s_2} - \beta_2 s_2} \dots \int_{s_{k-1}}^\infty ds_k\, e^{-i\alpha_k L_{s_k} - \beta_k s_k}. \tag{2.3.1}$$

Their expectations are encoded as

$$(\alpha_1, \beta_1) \dots (\alpha_k, \beta_k) \quad (1 \le k \le n), \tag{2.3.2}$$

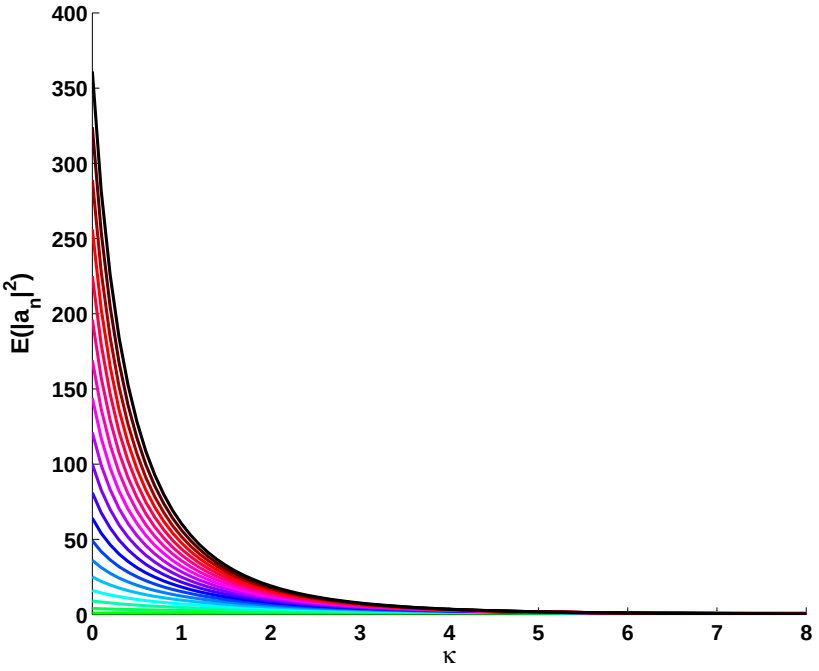
and are explicitly computed by using as above the strong Markov property and the Lévy characteristic function (1.2.3):

$$(\alpha_1, \beta_1) \dots (\alpha_k, \beta_k) = \prod_{j=0}^{k-1} \left[\beta_k + \beta_{k-1} + \dots + \beta_{k-j} + \eta(\alpha_k + \alpha_{k-1} + \dots + \alpha_{k-j})\right]^{-1}.$$

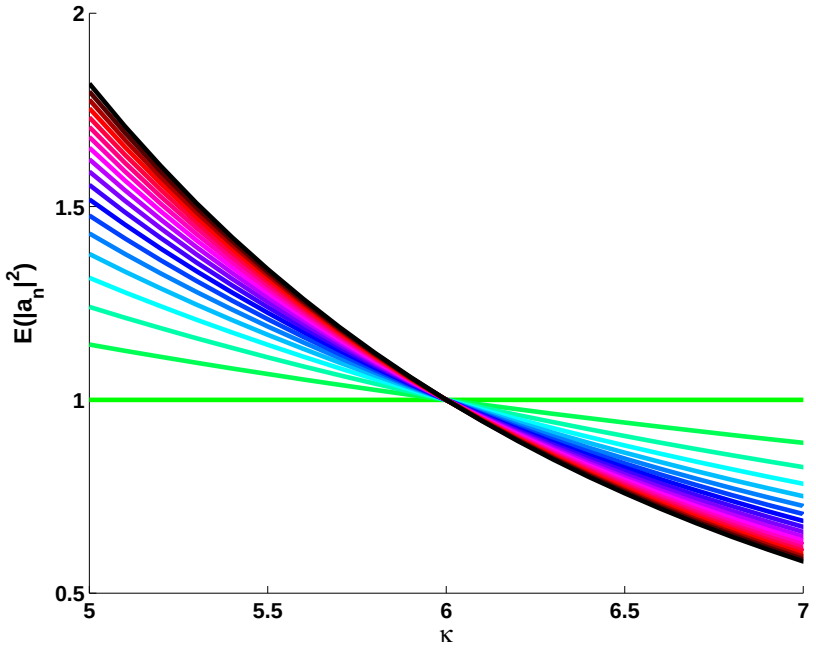
Next, in order to compute $|a_n|^2$, we need to evaluate the expectation of products of integrals such as (2.3.1) with complex conjugate of others, that we symbolically denote by

$$[(\alpha_1, \beta_1) \dots (\alpha_k, \beta_k); (-\alpha'_1, \beta'_1) \dots (-\alpha'_\ell, \beta'_\ell)] \quad (1 \le k, \ell \le n). \tag{2.3.3}$$

The product integrals may be written as a sum of $\binom{k+\ell}{k}$ ordered integrals with $k+\ell$ variables: the k first ones and the ℓ last ones are ordered and the number of ordered integrals corresponds to the number of ways of shuffling k cards in the left hand with ℓ cards in the right hand. This sum is quite large and, in order to systematically compute it, we write its expectation as the sum of expectations of integrals of the

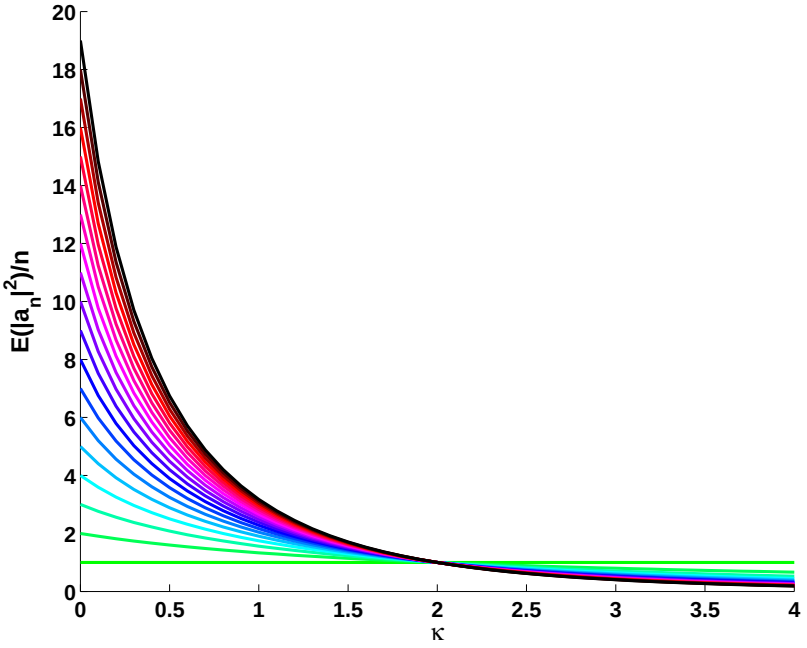


(a)

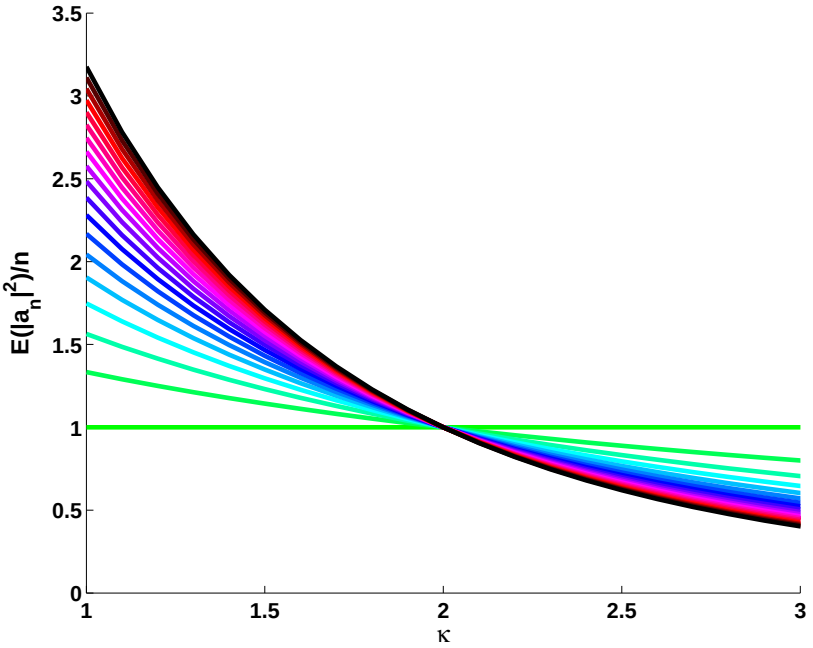


(b)

Figure 2.1: Graphs of the SLE_κ map $\kappa \mapsto \mathbb{E}\left(|a_n|^2\right)$ (a), and its zooming in near $\kappa = 6$ (b) for $n = 1, \cdots, 19$.



(a)



(b)

Figure 2.2: Graphs of the SLE_κ map $\kappa \mapsto \frac{\mathbb{E}\left(|a_n|^2\right)}{n}$ (a), and its zooming in near $\kappa = 2$ (b) for $n = 1, \cdots, 19$.

form (2.3.2) that begin with a term of type (α_1, β_1) or with a term of type $(-\alpha_1', \beta_1')$, thus reducing the work to a computation at lower order.
Using dynamic programing, we performed computations (formal up to $n = 8$ and numerical up to $n = 19$) on a usual computer in the next section. The graphs given in figure 2.1a for the SLE_κ map $\kappa \mapsto \mathbb{E}\left(|a_n|^2\right)$, for $n = 1, \cdots, 19$ illustrate the phenomena described above, in particular the striking constant value $\mathbb{E}\left(|a_n|^2\right) = 1$ for $\kappa = 6$ (figure 2.1b). And the figure 2.2 show the case $\mathbb{E}\left(|a_n|^2\right) = n$ for $\kappa = 2$.

2.4 The coefficients for high order

Here are the results for a_3, a_4 and a_5 in the LLE-case:

$$\begin{aligned}
\mathbb{E}\left(|a_3|^2\right) &= \frac{3!2^2}{(\eta_1+1)(\eta_1+3)} + \frac{2(\eta_1-1)(\eta_1-3)}{(\eta_1+1)(\eta_1+3)(\eta_2+2)};\\
\mathbb{E}\left(|a_4|^2\right) &= \frac{4!2^3}{(\eta_1+1)(\eta_1+3)(\eta_1+5)} + \frac{4(\eta_1-1)(\eta_1-3)\eta_2(\eta_2-4)(\eta_1+3)}{3(\eta_1+1)(\eta_1+3)(\eta_1+5)(\eta_2+2)(\eta_2+4)(\eta_3+3)}\\
&+ \frac{32(\eta_1-1)(\eta_1-3)}{(\eta_1+1)(\eta_1+3)(\eta_1+5)(\eta_2+2)(\eta_2+4)};\\
\mathbb{E}\left(|a_5|^2\right) &= \frac{5!2^4}{(\eta_1+1)(\eta_1+3)(\eta_1+5)(\eta_1+7)}\\
&+ \frac{4(\eta_1-1)(\eta_1-3)\eta_2(\eta_2-4)(\eta_1+3)(\eta_3+1)(\eta_3-5)(\eta_1+3)(\eta_1+5)(\eta_2+4)}{4(\eta_1+1)(\eta_1+3)(\eta_1+5)(\eta_1+7)(\eta_2+2)(\eta_2+4)(\eta_2+6)(\eta_3+3)(\eta_3+5)(\eta_4+4)}\\
&+ \frac{(\eta_1-1)(\eta_1-3)Q}{(\eta_1+1)(\eta_1+3)(\eta_1+5)(\eta_1+7)(\eta_2+2)(\eta_2+4)(\eta_2+6)(\eta_3+3)(\eta_3+5)},\\
Q &= \frac{4}{3}(24\eta_1^2\eta_2^2 + 9\eta_1^2\eta_2\eta_3^2 + 72\eta_1^2\eta_2\eta_3 + 39\eta_1^2\eta_2 + 36\eta_1^2\eta_3^2 + 288\eta_1^2\eta_3 + 520\eta_1^2 + 19\eta_1\eta_2^3\eta_3\\
&+ 77\eta_1\eta_2^3 + 56\eta_1\eta_2^2\eta_3 + 472\eta_1\eta_2^2 - 36\eta_1\eta_2\eta_3^2 - 816\eta_1\eta_2\eta_3 - 3660\eta_1\eta_2 - 144\eta_1\eta_3^2\\
&- 1152\eta_1\eta_3 - 2160\eta_1 + 75\eta_2^3\eta_3 + 285\eta_2^3 + 348\eta_2^2\eta_3^2 + 2952\eta_2^2\eta_3 + 6420\eta_2^2 + 3507\eta_2\eta_3^2\\
&+ 26184\eta_2\eta_3 + 43245\eta_2 + 8460\eta_3^2 + 67680\eta_3 + 126900).
\end{aligned}$$

In each expression for $\mathbb{E}\left(|a_n|^2\right)$, and after the first term there notice that the presence of the common factors $(\eta_1 - 1)(\eta_1 - 3)$ in denominators. The first term equals 1 for $\eta_1 = 3$ (or $\kappa = 6$), or equals n for $\eta_1 = 1$ (or $\kappa = 2$).
We have the general formula of $\mathbb{E}\left(|a_n|^2\right)$ for all $n \geq 3$:

$$\mathbb{E}(|a_n|^2) = \frac{n!2^{n-1}}{\prod_{j=0}^{n-2}(\eta_1+1+2j)} + (\eta_1-1)(\eta_1-3)\frac{Q_n}{\prod_{i=1}^{n-1}\prod_{j=0}^{n-1-i}(\eta_i+i+2j)} \quad (2.4.1)$$

with

$$Q_n = Q_n^1 + Q_n^2(\eta_{n-1} + n - 1)$$

and

$$Q_n^1 = \frac{4}{n-1}\prod_{i=2}^{n-2}((\eta_i-2)^2-i^2)\prod_{i=1}^{n-3}\prod_{j=1}^{n-i-2}(\eta_i+i+2j).$$

$Q_2^2 = 0$, Q_n^2 $(n \geq 3)$ we still don't know the formula now.
Motivated by the observations above, we were able to conjecture and prove for $n < 20$ the following theorem [DNNZ12].

Theorem 2.4.1. *Let $(f_t)_{t\geq 0}$ be the whole-plane Loewner process driven by the Lévy process L_t with Lévy symbol η. We write*

$$f_t(z) = e^t\left(z + \sum_{n\geq 2} a_n(t)z_n\right),$$

and $a_n = a_n(0)$. Then:

i. If $\eta_1 = 3$ we have

$$\mathbb{E}\left(|a_n|^2\right) = 1,\ \forall n \geq 1;$$

this case covers SLE_6.

ii. If $\eta_1 = 1$, $\eta_2 = 4$, we have

$$\mathbb{E}\left(|a_n|^2\right) = n,\ \forall n \geq 1;$$

this case covers SLE_2.

Later we have succeeded in giving a rigorous proof of this theorem.
The results for a_4 to a_8 in the SLE-case, and their graphs in terms of κ:

$$\mathbb{E}\left(|a_4|^2\right) = \frac{8}{9}\frac{\kappa^5+104\kappa^4+4576\kappa^3+18288\kappa^2+22896\kappa+8640}{(\kappa+10)(3\kappa+2)(\kappa+6)(\kappa+1)(\kappa+2)^2};$$

$$\begin{aligned}\mathbb{E}\left(|a_5|^2\right) =& (27\kappa^8+3242\kappa^7+194336\kappa^6+6142312\kappa^5+42644896\kappa^4\\ &+119492832\kappa^3+153156096\kappa^2+87882624\kappa+18144000)\\ &/[36(\kappa+14)(3\kappa+2)(\kappa+10)(2\kappa+1)(\kappa+6)(\kappa+3)(\kappa+1)(\kappa+2)^2];\end{aligned}$$

$$\begin{aligned}\mathbb{E}\left(|a_6|^2\right) =& \frac{2}{225}(216\kappa^{10}+29563\kappa^9+2062556\kappa^8+90749820\kappa^7+2277912280\kappa^6\\ &+16419864848\kappa^5+50825787744\kappa^4+76716664128\kappa^3\\ &+58263304320\kappa^2+21233664000\kappa+2939328000)\\ &/[(\kappa+18)(3\kappa+2)(\kappa+14)(2\kappa+1)(\kappa+10)(\kappa+6)(5\kappa+2)\\ &(\kappa+3)(\kappa+1)(\kappa+2)^2];\end{aligned}$$

$$\begin{aligned}\mathbb{E}\left(|a_7|^2\right) = &\frac{1}{8100}(27000\kappa^{15} + 4479353\kappa^{14} + 373838334\kappa^{13} + 20594712527\kappa^{12}\\ &+ 787796136854\kappa^{11} + 19121503739240\kappa^{10} + 221861771218136\kappa^9\\ &+ 1386550697705712\kappa^8 + 5130607642056896\kappa^7 + 11854768997862912\kappa^6\\ &+ 17547915006086400\kappa^5 + 16725481436226816\kappa^4 + 10110569026936320\kappa^3\\ &+ 3711483045734400\kappa^2 + 749049576192000\kappa + 63371911680000)\\ &/[(\kappa+22)(3\kappa+1)(5\kappa+2)(\kappa+18)(2\kappa+1)(\kappa+14)(3\kappa+2)\\ &(\kappa+10)(\kappa+6)(\kappa+5)(\kappa+3)(\kappa+1)^2(\kappa+2)^3]\,;\end{aligned}$$

$$\begin{aligned}\mathbb{E}\left(|a_8|^2\right) = &\frac{2}{99225}(729000\kappa^{18} + 143757261\kappa^{17} + 14031668642\kappa^{16} + 906444920407\kappa^{15}\\ &+ 42715714646750\kappa^{14} + 1476227672190480\kappa^{13} + 34674813906653712\kappa^{12}\\ &+ 471116720002819536\kappa^{11} + 3802657434377773600\kappa^{10}\\ &+ 19218418658636100992\kappa^9 + 63191729416067875840\kappa^8\\ &+ 138392538501661946112\kappa^7 + 204258207932541043200\kappa^6\\ &+ 203508494170475323392\kappa^5 + 135640094878259859456\kappa^4\\ &+ 59063686024095313920\kappa^3 + 16005106174366310400\kappa^2\\ &+ 2435069931098112000\kappa + 158176291553280000)\\ &/[(7\kappa+2)(5\kappa+2)(\kappa+26)(3\kappa+1)(\kappa+22)(2\kappa+1)(\kappa+18)(\kappa+14)\\ &(3\kappa+2)(\kappa+10)(\kappa+5)(\kappa+3)(\kappa+6)^2(\kappa+1)^2(\kappa+2)^3]\,.\end{aligned}$$

These results call for two observations:

i. Somehow surprisingly, all the coefficients of the polynomial expansions in κ are positive;

ii. For $\kappa \to \infty$ (or $\eta \to \infty$), these expectations vanish as κ^{-1}.

Chapter 3

Gradient percolation

3.1 Introduction

We study some aspects of *gradient percolation* for site percolation on triangular lattice. In this model, the coloring probability depends on the location of the site. We adapt proofs of P. Nolin to establish results predicted by physicists for this model concerning critical exponents.

Let us consider, to fix ideas, an infinite triangular lattice, or equivalently, a regular hexagonal paving. The vertices of this lattices are colored independently black or white with probability p and $1-p$ respectively. There exist a critical value p_c of p, such that when $p > p_c$, there is an infinite connected component of black sites a.s. and when $p < p_c$, there is, a.s. an infinite connected component of white sites. The value p_c is called *threshold percolation* (or *critical percolation*) of this standard percolation model [Gri99].

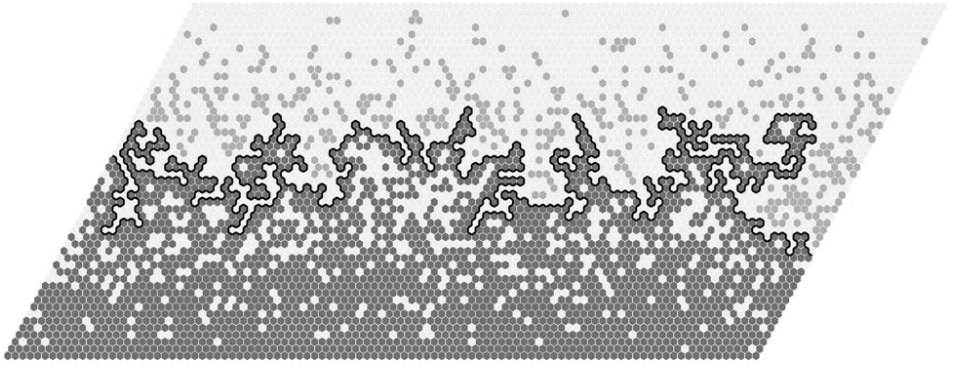


Figure 3.1: A simulation of P. Nolin for gradient percolation with $p(z) = y(z)/N$ on the triangular lattice of height $N = 50$ and of length $\ell_N = 100$.

In [Nol08a], Nolin considered an inhomogeneous percolation model: given a large integer N, take the horizontal strip $[0, \ell_N] \times [0, N]$ where $\ell_N > N$ and consider

triangular lattice inside the strip. Each site $z = (x, y)$ on the triangular lattice is colored black or white independently with a probability depending on the position of z, $p(z) = \frac{y}{N}$ with y-coordinate equal to $y(z) \in [0, N]$, while $y < 0$, $p(z) = 0$ with $y > N$, $p(z) = 1$. This function is linear, continuous, and decreasing by y. In this configuration, there is almost surely a (unique) infinite black connected component and a unique infinite white connected component. Furthermore, the external boundary of the white cluster and the external boundary of the black cluster coincide. This separating path is called the *percolation front*.

Nolin establishes results concerning localization and length of external boundary (the black curve in figure 3.1) between connected components of black sites and white sites (i.e. apparating cells). For large N, on a triangular strip lattice with length ℓ_N, the percolation front will tend to be localized near the line $y/N = 1/2$, where probability $p(z)$ is close to threshold percolation. The width of the trip covering the front is $N^{4/7}$ and the expectation length of the front approximates $N^{3/7}\ell_N$. Furthermore, the length of the front is close to its expected value.
These results are proven for the case that the occupation probability function is linear. Here we show that these results hold with a nonlinear probability function $p(z) = e^{-\lambda r_z}$, with λ is a parameter and r_z is distance to a given point.

3.2 Homogeneous percolation

We recall some facts of homogeneous percolation (or standard percolation) that we will use later.

3.2.1 Setting

In this chapter, we will consider site percolation in two dimensions on the triangular lattice on oblique coordinates, with origin at 0 and the basis given by 1 and $e^{i\pi/3}$ (see figure 3.2). The parallelogram R with vertices $a_j + b_k e^{i\pi/3} (j, k = 1, 2)$ will be denoted by $[a_1, a_2] \times [b_1, b_2]$. Its interior and boundary will be denoted by $\mathring{R}$ and ∂R, respectively. We will consider $\|z\|$, the (discrete) infinity norm of z and $d(z, z') := \|z - z'\|$, the associated distance. For the site $z = (z_1, z_2)$, we will use the rhombus $S_n(z) := [z_1 - n, z_1 + n] \times [z_2 - n, z_2 + n]$ and refer to $S_n(0)$ simply as S_n.

Recall that percolation is a random (black and white) coloring of the faces of the hexagonal lattice. In the standard homogeneous case with a constant $p \in [0, 1]$, it is known for a large class of lattices that there is a phase transition at a certain critical parameter $p_c \in (0, 1)$ (its value depends on type of lattice): when $p \leq p_c$ corresponding to sub-critical percolation there is (a.s.) no infinite cluster, while for $p > p_c$ (super-critical percolation) there is (a.s.) a unique infinite cluster. This value, p_c, is called the critical probability. A celebrated result of Kesten [Kes80] asserts that $p_c = 1/2$ on the triangular lattice.
We declare each site z to be occupied/black with probability p, vacant/white with

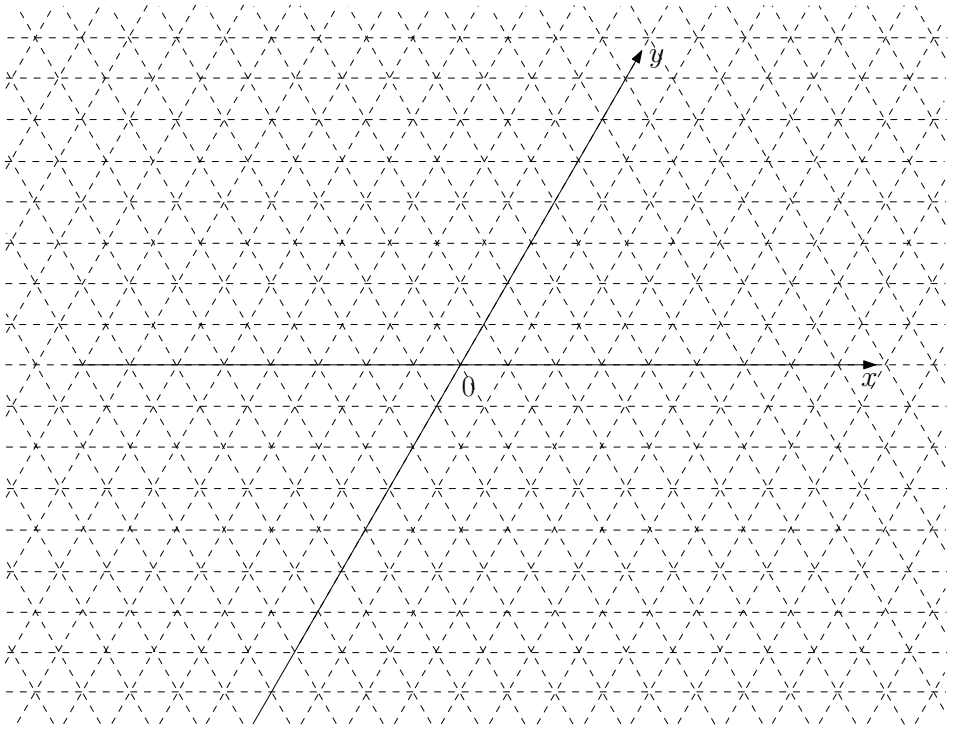


Figure 3.2: The triangular lattice.

probability $1-p$, independently of the other sites. The corresponding probability measure on the set of configuration will be referred to as $\mathbb{P}_p$ and $\mathbb{E}_p$ will denote the expectation.

A concept which is often used in this chapter is crossing. A left-right (or horizontal) occupied *crossing* of the parallelogram $[a_1, a_2] \times [b_1, b_2]$ is simply a black path inside the parallelogram connecting its left side to its right side. However, with this definition, the existence of a crossing in two parallelograms sharing a side is not independent. In order to ensure this independence, we add the condition on its extremities: a crossing is only composed by the black connected sites in the interior of a parallelogram except its extremities which may be either black or white. A vacant crossing is composed by white sites.

In the following, we are interested in the connectivity properties of the set of occupied sites. The connected components are called *clusters*. The term "interface" is used for a curve on the dual hexagonal lattice bordered by occupied sites on one side, and by vacant sites on the other side (the boundary of a finite cluster for instance).

As usual in the statistical physics literature, for two positive function f and g, the notation $f \asymp g$ means that there exist two constants $C_1,\ C_2 > 0$ such that $C_1 g \leq f \leq C_2 g$; while $f \approx g$ means that $\log f / \log g \to 1$ when $p \to 1/2$ or when $n \to \infty$.

3.2.2 Critical exponents

Let us consider a fixed integer $j \geq 2$. For $m < n$ we define the event $A^j(m,n)$ that there exist j disjoint monochromatic paths (*arms*) from ∂S_m to ∂S_n that are not all of the same color (each path is either completely black or completely white, and there is at least one white path and one black path). In [SW01], the following result is proven.

Proposition 3.2.1 ([SW01]). *For any fixed $j \geq 2$, and for all large enough n (i.e. $m \geq j$),*

$$\mathbb{P}_{1/2}\left(A^j(m,n)\right) \approx n^{-\left(j^2-1\right)/12}, \tag{3.2.1}$$

when $n \to \infty$.

Recall also the partial order on the set of events (colorings of vertices) $\Omega : \omega \leq \omega'$ iff whenever a vertex is black in ω it is also black in ω'. The event A is called *increasing* if $I_A(\omega) \leq I_A(\omega')$ whenever $\omega \leq \omega'$, where I_A is the indicator function of A. We call A *decreasing* if its complement $\bar{A}$ is increasing. For example, if x and y are vertices, then the event that there is an black path from x to y is an increasing event.

FKG inequality (Harris / Fortuin, Kasteleyn, Ginibre - 1960). If A and B are increasing events

$$\mathbb{P}(A \cap B) \geq \mathbb{P}(A)\mathbb{P}(B).$$

The behavior of percolation at the critical point, exhibited in the Kesten's paper [Kes82], links some of the previous arm exponents to other critical exponents describing the behavior of connectivity probabilities near $p = p_c$.

Let us introduce some notations that will be used later frequently. Denote $\mathcal{C}_H\left([a_1,a_2]\times[b_1,b_2]\right)$ (resp. $\mathcal{C}_V\left([a_1,a_2]\times[b_1,b_2]\right)$) the event that there exists a horizontal (resp. vertical) occupied crossing of the parallelogram $[a_1,a_2]\times[b_1,b_2]$, and $\mathcal{C}_H^*$, $\mathcal{C}_V^*$ the same events with vacant crossings.
Continuously, we define a quantity called *characteristic length* $L(p)$ [Kes82]. For each fixed $\epsilon_0 > 0$

$$L(p) = L\left(p,\epsilon_0\right) = \begin{cases} \min\{n \text{ s.t. } \mathbb{P}_p(\mathcal{C}_H([0,n]\times[0,n])) \leq \epsilon_0\}, & \text{when } p < 1/2 \\ \min\{n \text{ s.t. } \mathbb{P}_p(\mathcal{C}_H^*([0,n]\times[0,n])) \leq \epsilon_0\}, & \text{when } p > 1/2. \end{cases} \tag{3.2.2}$$

The Russo-Seymour-Welsh (RSW) theory (cf. theorem 6.1 in [Kes82]) implies that: for each $k \geq 1$, there exists $\delta_k > 0$ (depending only on ϵ_0) such that

$$\forall n \leq L(p) \qquad \mathbb{P}_p(\mathcal{C}_H([0,kn]\times[0,n])) \geq \delta_k. \tag{3.2.3}$$

This bound also holds with horizontal vacant crossings for symmetry reasons. This estimate is established in [Nol08b].
We will have the following important result.

Lemma 3.2.2. *If ϵ_0 has been chosen sufficiently small, then there exists a constant $C > 0$ such that for all n and all $p < 1/2$,*

$$\mathbb{P}_p\left(\mathcal{C}_H\left([0,n]\times[0,n]\right)\right) \leq Ce^{-n/L(p)}. \tag{3.2.4}$$

Once more, the proof of this lemma is given in Nolin's paper [Nol08a] and some variants of it have previously been used and mentioned in Kesten's paper [Kes82].

Proposition 3.2.3. *When $p \to 1/2$,*

$$L(p) \approx |p - 1/2|^{-4/3}. \tag{3.2.5}$$

Remark 3.2.4. Kesten showed in [Kes82] that for any fixed ϵ_1 and ϵ_2 with $0 < \epsilon_1, \epsilon_2 \leq \epsilon_0$,

$$L\left(p, \epsilon_1\right) \approx L\left(p, \epsilon_2\right). \tag{3.2.6}$$

3.3 Description of the model

Let us now describe the model itself. We consider a site-percolation model on the infinite triangular lattice $\mathcal{S}$ with nonlinear occupation probability of a site z

$$f\left(z\right) = e^{-\lambda\|z\|}, \tag{3.3.1}$$

where $\lambda > 0$ is the density gradient. This function is used in simulations of the growth of cities in the next chapter.
At each site z on the lattice, we assign a random number u_z with uniform distribution on the interval $[0, 1]$. Color or status of this site depends on the occupation probability $f(z)$. It is colored black (with probability $f(z)$) if $u_z \leq f(z)$ and white (vacant) if $u_z > f(z)$. Clearly, the sites close to center 0 will be mostly occupied (black) ($p \simeq 1$) and the sites being far away center will be vacant (white) ($p \simeq 0$). The function (3.3.1) is continuous and decreasing $\|z\|$. Hence, when $p_c = 1/2$ there exists a $r_0 > 0$ such that $\|z_c\| = r_0$ and $f\left(z_c\right) = 1/2$, i.e., using (3.3.1),

$$r_0 = \lambda^{-1}\log\left(2\right). \tag{3.3.2}$$

As mentioned above, there is a transition at p_c. In the sub-critical region, $\{z, \|z\| > r_0\}$, there only exist isolated clusters. The cluster containing the center of lattice is called the largest cluster (figure 3.3). Notice that this cluster is not entirely contained in super-critical region.

The following results concern the front of this cluster.
From observation of simulations, we can see that size of this cluster is finite. For a formal proof of this, we recall the Borel-Cantelli lemma:

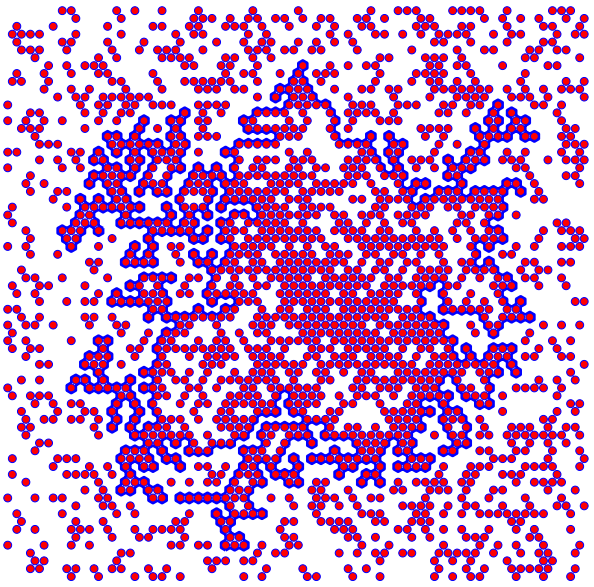


Figure 3.3: The biggest cluster on triangular lattice using our model with $\lambda = 0.02$. The blue line is the external perimeter of the cluster.

Lemma 3.3.1 (Borel-Cantelli)**.** *Let (E_n) is a sequence of events in some probability space and*

$$\sum_{n=1}^{\infty} \mathbb{P}(E_n) < \infty,$$

then

$$\mathbb{P}(E_n \ i.o.) = 0,$$

where i.o. stands for "infinitely often".

Consider an infinite lattice with origin 0 and a parallelogram S_n with its boundary $\partial S_n = \{z : \|z\| = n\}$. Let A_n be the event that there is at least a black point on ∂S_n. We can estimate the probability of this event

$$\begin{aligned}\mathbb{P}(A_n) &\leq \sum_{z\in\partial S_n} \mathbb{P}(z \text{ is black}) = \sum_{z\in\partial S_n} \mathbb{P}(u_z \leq f(z)) \\ &= \sum_{z\in\partial S_n} e^{-\lambda\|z\|} \leq 8ne^{-\lambda n}.\end{aligned}$$

From that we estimate the sum

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) \leq \sum_{n=1}^{\infty} 8ne^{-\lambda n} < +\infty. \tag{3.3.3}$$

By Borel-Cantelli lemma it follows that

$$\mathbb{P}\left(A_n \, i.o.\right) = 0. \tag{3.3.4}$$

From this equation we obtain that a.s. there exists $N > 0$ such that for all $n > N$ there are no black point on ∂S_n. And hence the number of black points on an infinite lattice is finite.

We can also deduce lower estimate of the probability that the largest cluster C is contained in S_N :

$$\begin{aligned} \mathbb{P}(C \subset S_N) &\geq 1 - \sum_{n=N+1}^{\infty} \mathbb{P}\left(A_n\right) \geq 1 - \sum_{n=N+1}^{\infty} n e^{-\lambda n} \\ &\geq 1 - \frac{8}{\lambda}\left(N + 1 + \frac{1}{\lambda}\right) e^{-\lambda(N+1)}. \end{aligned} \tag{3.3.5}$$

We consider a sufficiently large N, and, in order to normalize distances, we choose $\lambda = \frac{\tau}{N}$, with $\tau > 0$ is a given parameter. The function (3.3.1) becomes

$$f(z) = e^{-\|z\|\tau/N}, \tag{3.3.6}$$

and the distance at critical point p_c

$$r_0 = N \log\left(2\right)/\tau. \tag{3.3.7}$$

Assume that the largest cluster is totally contained in lattice of size $[-N, N] \times [-N, N]$ (which happens with positive and bounded below probability if $\tau \geq 1$).

According to physical result of Sapoval et al in [SRG85], when $\tau = 1$, the front of the largest cluster is localized around the boundary of parallelogram $[-r_0, r_0] \times [-r_0, r_0]$ and its width, denoted by σ_f, is a function of concentration gradient λ

$$\sigma_f \sim \lambda^{-\frac{\nu}{1+\nu}}, \tag{3.3.8}$$

with $\nu = 4/3$.

With this model, we will prove the above result of Sapoval et al of localization of front of the largest cluster and estimate its length in the next sections.

3.4 Some results

Recall that our setting is realized on a triangular lattice $\mathcal{S}$. Denote $S_m = [-m, m] \times [-m, m]$ the parallelogram of size $2m \times 2m$ and $S_{m,n} = S_n \setminus S_m$, $n > m$, the annulus between two parallelograms. Since we will use often $S_{r_0 - N^\alpha, r_0 + N^\alpha}$ in the next sections, for the sake of simplicity, we will refer to it simply as $S_{r_0 \pm N^\alpha}$.
All of results will relate to front of the largest cluster. We adopt the following definition.

Definition 3.4.1. A *front* will be any closed interface ρ_λ containing the origin 0 on dual hexagonal lattice of $\mathcal{S}$ that is bordered by an closed occupied curve on one side and a closed vacant curve on the other side.

3.4.1 Localization

To start with, we consider $\mathcal{R}_\lambda$, the cluster containing the origin 0 and, instead of the front, the furthest closed occupied curve from the center of lattice $\mathcal{R}_\lambda$. Note that the sites outside $\mathcal{R}_\lambda$ form a curve so that the outside boundary ρ_λ of $\mathcal{R}_\lambda$ (the path on dual hexagonal lattice bordering it outside) is a front.

Theorem 3.4.2 (Localization of the front.)**.** *For all $\delta > 0$, there exists a $\delta' > 0$ such that for sufficiently large N, $\lambda_N = \frac{\tau}{N}$ $(\tau \geq 1)$,*

i.

$$\mathbb{P}\left(\mathcal{R}_{\lambda_N} \subseteq S_{r_0 \pm \lambda_N^{-(4/7-\delta)}}\right) \leq e^{-N^{\delta'}} \tag{3.4.1}$$

ii.

$$\mathbb{P}\left(\mathcal{R}_{\lambda_N} \nsubseteq S_{r_0 \pm \lambda_N^{-(4/7+\delta)}}\right) \leq e^{-N^{\delta'}} \tag{3.4.2}$$

Proof. i. Firstly, we prove that $\mathcal{R}_{\lambda_N}$ is contained in $S_{r_0 \pm \lambda_N^{-(4/7-\delta)}}$ or $S_{r_0 \pm (N/\tau)^{4/7-\delta}}$ with small probability.
Divide $S_{r_0 \pm (N/\tau)^{4/7-\delta}}$ into eight parts as in figure 3.4 with four rhombi of size $2(N/\tau)^{4/7-\delta}$ and four parallelograms of size $2\left(r_0 - (N/\tau)^{4/7-\delta}\right) \times 2(N/\tau)^{4/7-\delta}$. Notice that, occupation probability in these parallelograms is the same. Therefore, without loss of generality, we can consider the top parallelogram

$$\left[-r_0 + (N/\tau)^{4/7-\delta}, r_0 - (N/\tau)^{4/7-\delta}\right] \times \left[r_0 - (N/\tau)^{4/7-\delta}, r_0 + (N/\tau)^{4/7-\delta}\right].$$

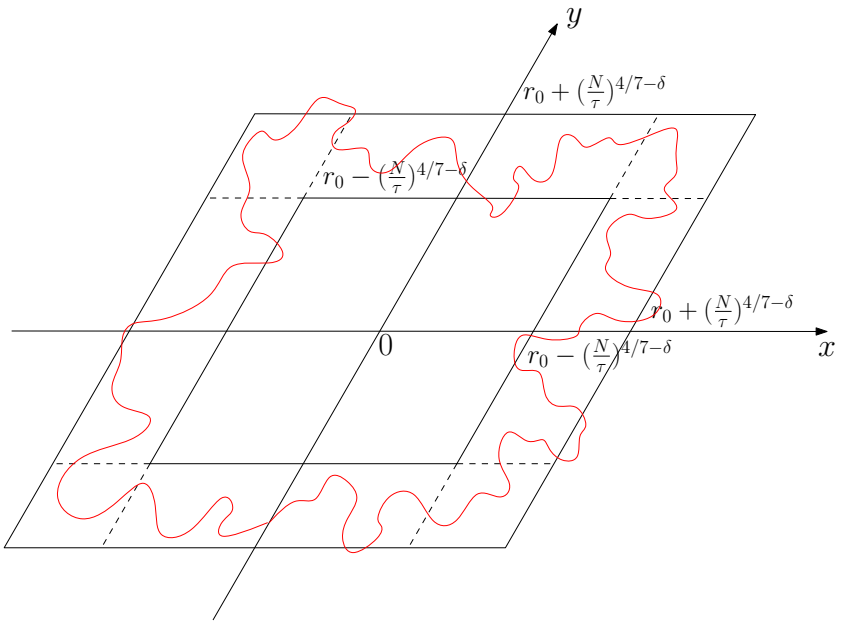


Figure 3.4: Annulus $S_{r_0 \pm \left(\frac{N}{\tau}\right)^{4/7-\delta}}$ with its divisions and a closed curve.

By the disjoint construction, we have

$$\mathbb{P}\left(\mathcal{R}_{\lambda_N} \subseteq S_{r_0\pm(N/\tau)^{4/7-\delta}}\right) \leq \left(\mathbb{P}\left(\mathcal{C}_H\left(\left[-r_0+(N/\tau)^{4/7-\delta}, r_0-(N/\tau)^{4/7-\delta}\right]\right.\right.\right. \\ \left.\left.\left. \times \left[r_0-(N/\tau)^{4/7-\delta}, r_0+(N/\tau)^{4/7-\delta}\right]\right)\right)\right)^4. \quad (3.4.3)$$

Consider the parallelogram

$$\left[-r_0+(N/\tau)^{4/7-\delta}, r_0-(N/\tau)^{4/7-\delta}\right] \times \left[r_0-(N/\tau)^{4/7-\delta}, r_0+(N/\tau)^{4/7-\delta}\right];$$

we divide it into disjoint rhombi of the form

$$\left[-r_0+(2i+1)(N/\tau)^{4/7-\delta}+i, -r_0+(2i+3)(N/\tau)^{4/7-\delta}+i\right] \\ \times \left[r_0-(N/\tau)^{4/7-\delta}, r_0+(N/\tau)^{4/7-\delta}\right],$$

for $i=0,1,2,...$ We can take at least $\frac{r_0}{3(N/\tau)^{4/7-\delta}}$ such rhombi. All of them are contained in the region with occupancy probability smaller than

$$p = f\left(r_0-(N/\tau)^{4/7-\delta}\right) = e^{-\frac{\tau}{N}\left(r_0-(N/\tau)^{4/7-\delta}\right)} = p_c e^{(N/\tau)^{-3/7-\delta}}.$$

Each of them thus has (independently with the others) a vertical vacant (white) crossing with probability larger than

$$\mathbb{P}_p\left(\mathcal{C}_V^*\left(\left[0, 2(N/\tau)^{4/7-\delta}\right]^2\right)\right).$$

By using (3.2.5) and Taylor expansion of $e^{(N/\tau)^{-3/7-\delta}}$ we can get

$$L(p) = \left|p_c - p_c e^{(N/\tau)^{-3/7-\delta}}\right|^{-4/3} \approx (N/\tau)^{4/7+4\delta/3},$$

and hence

$$L(p) \gg 2(N/\tau)^{4/7-\delta}.$$

By symmetry we have

$$\mathbb{P}_p\left(\mathcal{C}_V^*\left(\left[0, 2(N/\tau)^{4/7-\delta}\right]^2\right)\right) = \mathbb{P}_p\left(\mathcal{C}_H^*\left(\left[0, 2(N/\tau)^{4/7-\delta}\right]^2\right)\right).$$

According to the definition of $L(p)$ (3.2.2), we have

$$\mathbb{P}_p\left(\mathcal{C}_V^*\left(\left[0, 2(N/\tau)^{4/7-\delta}\right]^2\right)\right) > \epsilon_0.$$

By independence, we obtain

$$\mathbb{P}\left(\mathcal{C}_H\left(\left[-r_0+(N/\tau)^{4/7-\delta}, r_0-(N/\tau)^{4/7-\delta}\right] \times \left[r_0-(N/\tau)^{4/7-\delta}, r_0+(N/\tau)^{4/7-\delta}\right]\right)\right) \\ \leq \left(\mathbb{P}\left(\mathcal{C}_H\left(\left[0, 2(N/\tau)^{4/7-\delta}\right]^2\right)\right)\right)^{r_0/(3(N/\tau)^{4/7-\delta})} \\ \leq \left(1-\mathbb{P}_p\left(\mathcal{C}_V^*\left(\left[0, 2(N/\tau)^{4/7-\delta}\right]^2\right)\right)\right)^{r_0/(3(N/\tau)^{4/7-\delta})} \\ \leq (1-\epsilon_0)^{r_0/(3(N/\tau)^{4/7-\delta})}.$$

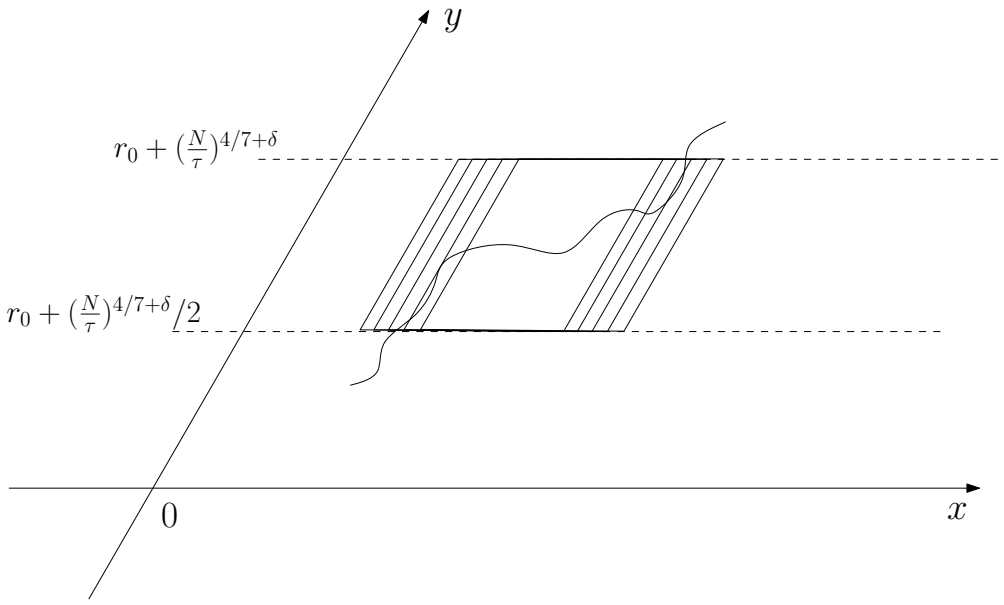


Figure 3.5: $\mathcal{R}_\lambda$ crosses one of rhombi vertically or horizontally.

Using the formula of r_0 (3.3.7), the right hand side of the inequality (3.4.3) gives

$$\mathbb{P}\left(\mathcal{R}_{\lambda_N} \subseteq S_{r_0 \pm (N/\tau)^{4/7-\delta}}\right) \leq \left((1-\epsilon_0)^{r_0/\left(3(N/\tau)^{4/7-\delta}\right)}\right)^4.$$

Since $1 - \epsilon_0 < e^{-\epsilon_0}$ for $0 < \epsilon_0 < 1$,

$$\mathbb{P}\left(\mathcal{R}_{\lambda_N} \subseteq S_{r_0 \pm (N/\tau)^{4/7-\delta}}\right) \leq e^{-4\epsilon_0 r_0/\left(3(N/\tau)^{4/7-\delta}\right)}.$$

Therefore, there exists a $\delta' > 0$ such that

$$\mathbb{P}\left(\mathcal{R}_{\lambda_N} \subseteq S_{r_0 \pm (N/\tau)^{4/7-\delta}}\right) \leq e^{-N^{\delta'}}$$

ii. In the second part, we prove that $\mathcal{R}_{\lambda_N}$ is localized in $S_{r_0 \pm \lambda_N^{-(4/7+\delta)}}$ or $S_{r_0 \pm (N/\tau)^{4/7+\delta}}$ with high probability.

Consider $S_{r_0 \pm (N/\tau)^{4/7+\delta}}$ and its divisions as above. Assume that $\mathcal{R}_{\lambda_N}$ is not entirely contained in $S_{r_0 \pm (N/\tau)^{4/7+\delta}}$. There are thus some points outside $S_{r_0+(N/\tau)^{4/7+\delta}}$ or some points inside $S_{r_0-(N/\tau)^{4/7+\delta}}$.
Without loss of generality, we can further assume that there are some points outside $S_{r_0+(N/\tau)^{4/7+\delta}}$ and some of them are localized above the top side in the region

$$\left[-r_0 - (N/\tau)^{4/7+\delta}, r_0 + (N/\tau)^{4/7+\delta}\right] \times \left[r_0 + (N/\tau)^{4/7+\delta}, \infty\right).$$

We face the following alternative:

(a) $\mathcal{R}_{\lambda_N}$ visits

$$\left[-r_0-(N/\tau)^{4/7+\delta}, r_0+(N/\tau)^{4/7+\delta}\right]\times\left[r_0-(N/\tau)^{4/7+\delta}/2, r_0+(N/\tau)^{4/7+\delta}/2\right]$$

at some points.
We cover this parallelogram into the following (non-disjoint) rhombi:

$$\begin{aligned}\left[-r_0-(N/\tau)^{4/7+\delta}/2+i, -r_0+i+1-(N/\tau)^{4/7+\delta}/2\right]\\ \times\left[r_0+(N/\tau)^{4/7+\delta}/2, r_0+(N/\tau)^{4/7+\delta}\right],\end{aligned}$$

for $i=0,1,...$
It is easy to see that $\mathcal{R}_{\lambda_N}$ will have to cross one of them vertically or horizontally (figure 3.5). There are at most $2\left(r_0+(N/\tau)^{4/7+\delta}\right)$ such rhombi and they are in a zone with occupation probability less than

$$\begin{aligned}p' = f\left(r_0+(N/\tau)^{4/7+\delta}/2\right) &= e^{-\frac{\tau}{N}\left(r_0+\frac{N}{\tau}\right)^{4/7+\delta}/2)}\\ &= p_c e^{-(N/\tau)^{-3/7+\delta}/2}.\end{aligned}$$

A crossing of this parallelogram thus occurs with probability less than

$$\begin{aligned}2\left(r_0+(N/\tau)^{4/7+\delta}\right)&\mathbb{P}_{p'}\left(\mathcal{C}_H\left(\left[0,(N/\tau)^{4/7+\delta}/2\right]^2\right)\cup\mathcal{C}_V\left(\left[0,(N/\tau)^{4/7+\delta}/2\right]^2\right)\right)\\ &\leq 4\left(r_0+(N/\tau)^{4/7+\delta}\right)\mathbb{P}_{p'}\left(\mathcal{C}_H\left(\left[0,(N/\tau)^{4/7+\delta}/2\right]^2\right)\right)\\ &\leq C\left(r_0+(N/\tau)^{4/7+\delta}\right)e^{-(N/\tau)^{4/7+\delta}/(2L(p'))}\end{aligned}$$

by lemma 3.2.2.
Using (3.2.5) and Taylor expansion of $e^{-(N/\tau)^{-3/7+\delta}}$,

$$L(p')\approx\left|p_c-p_c\left(1-(N/\tau)^{-3/7+\delta}/2\right)\right|^{-4/3}\approx(N/\tau)^{4/7-4\delta/3}.$$

Therefore, *under the assumption "there are some points of $\mathcal{R}_{\lambda_N}$ above the top side of $S_{r_0+(N/\tau)^{4/7+\delta}}$"* the front $\mathcal{R}_{\lambda_N}$ visits

$$\left[-r_0-(N/\tau)^{4/7+\delta}, r_0+(N/\tau)^{4/7+\delta}\right]\times\left[r_0-(N/\tau)^{4/7+\delta}/2, r_0+(N/\tau)^{4/7+\delta}/2\right]$$

with probability smaller than

$$C\left(r_0+(N/\tau)^{4/7+\delta}\right)e^{-(N/\tau)^{4/7+\delta}/(2(N/\tau)^{4/7-4\delta/3})}=C\left(r_0+(N/\tau)^{4/7+\delta}\right)e^{-(N/\tau)^{7\delta/3}/2}.$$

Clearly, since $r_0+(N/\tau)^{4/7+\delta}$ (with $r_0=(N/\tau)\log(2)$) does not grow too fast, the probability of the considered event tends to 0 subexponentially fast.

(b) $\mathcal{R}_{\lambda_N}$ stays constantly above the trip

$$\left[-r_0-(N/\tau)^{4/7+\delta}, r_0+(N/\tau)^{4/7+\delta}\right]\times\left[r_0-(N/\tau)^{4/7+\delta}/2, r_0+(N/\tau)^{4/7+\delta}/2\right]$$

where the occupied probability is also smaller than p'. $\mathcal{R}_{\lambda_N}$ will cross one of the rhombi

$$\begin{aligned}&\left[-r_0-(N/\tau)^{4/7+\delta}, -r_0-(N/\tau)^{4/7+\delta}/2\right]\\ &\qquad\times\left[j+r_0+(N/\tau)^{4/7+\delta}/2, j+1+r_0+(N/\tau)^{4/7+\delta}/2\right],\end{aligned}$$

$j=1,2,...$
forming a column, vertically or horizontally with probability less than

$$\begin{aligned}&\mathbb{P}_{p'}\left(\mathcal{C}_H\left(\left[0,(N/\tau)^{4/7+\delta}/2\right]^2\right)\cup\mathcal{C}_V\left(\left[0,(N/\tau)^{4/7+\delta}/2\right]^2\right)\right)\\ &\le 2\mathbb{P}_{p'}\left(\mathcal{C}_H\left(\left[0,(N/\tau)^{4/7+\delta}/2\right]^2\right)\right)\\ &\le Ce^{-(N/\tau)^{7\delta/3}/2}.\end{aligned}$$

From (3.3.5) we get that

$$\mathbb{P}(\mathcal{R}_{\lambda_N}\not\subset S_{N^k})\le\frac{8N}{\tau}\left(N^k+1+\frac{N}{\tau}\right)e^{-\frac{\tau}{N}(N^k+1)},$$

with $k>1$.
The probability in this case is less than

$$\begin{aligned}&2\mathbb{P}_{p'}\left(\mathcal{C}_H\left(\left[0,(N/\tau)^{4/7+\delta}/2\right]^2\right)\right)+\mathbb{P}(\mathcal{R}_{\lambda_N}\not\subset S_{N^k})\\ &\le CN^ke^{-(N/\tau)^{7\delta/3}/2}+\frac{8N}{\tau}\left(N^k+1+\frac{N}{\tau}\right)e^{-\frac{\tau}{N}(N^k+1)}\end{aligned}$$

This probability also tends to 0 subexponentially fast.

The same conclusion can be drawn for other cases, when some points of $\mathcal{R}_{\lambda_N}$ are localized outside $S_{r_0+(N/\tau)^{4/7+\delta}}$ below the bottom edge,to the left of the left edge or to the right of the right edge of this parallelogram.
In the case $\mathcal{R}_{\lambda_N}$ is inside $S_{r_0-(N/\tau)^{4/7+\delta}}$ at some points, the argument is identical: consider instead the vacant (white) crossing bordering $\mathcal{R}_{\lambda_N}$.

□

3.4.2 The uniqueness

In the previous subsection, we focused on the furthest closed occupied curve covering 0, $\mathcal{R}_{\lambda_N}$. Clearly, the results also hold with the nearest vacant curve covering 0, $\mathcal{R}^*_{\lambda_N}$. Recall that $\mathcal{R}_{\lambda_N}$ is bordered outside by a vacant curve so that its external boundary ρ_{λ_N} is a front. Similarly, $\mathcal{R}^*_{\lambda_N}$ is bordered by an occupied curve and its

internal boundary $\rho^*_{\lambda_N}$ is also a front. Note that $\rho^*_{\lambda_N}$ is always inside ρ_{λ_N}.

Consider lattice $S_{r_0+(N/\tau)^{4/7+\delta}}$. From the localization theorem, we see that $\mathcal{R}^*_{\lambda_N}$ stays inside this lattice with high probability. To show the uniqueness of the front we can show that these fronts coincide with probability close to 1. This is equivalent to verifying that $\mathcal{R}^*_{\lambda_N}$ is connected to the boundary of $S_{r_0+(N/\tau)^{4/7+\delta}}$ $(\partial S_{r_0+(N/\tau)^{4/7+\delta}})$ by a vacant path (or that $\mathcal{R}_{\lambda_N}$ is connected to the origin 0 by an occupied path). We will prove that this occurs with a very large probability.

Proposition 3.4.3. *There exists a $\delta'' > 0$ such that for all sufficiently large N, $\lambda_N = \frac{\tau}{N}$,*

$$\mathbb{P}\left(\rho_{\lambda_N} \equiv \rho^*_{\lambda_N}\right) \geq 1 - e^{-N^{\delta''}}. \tag{3.4.4}$$

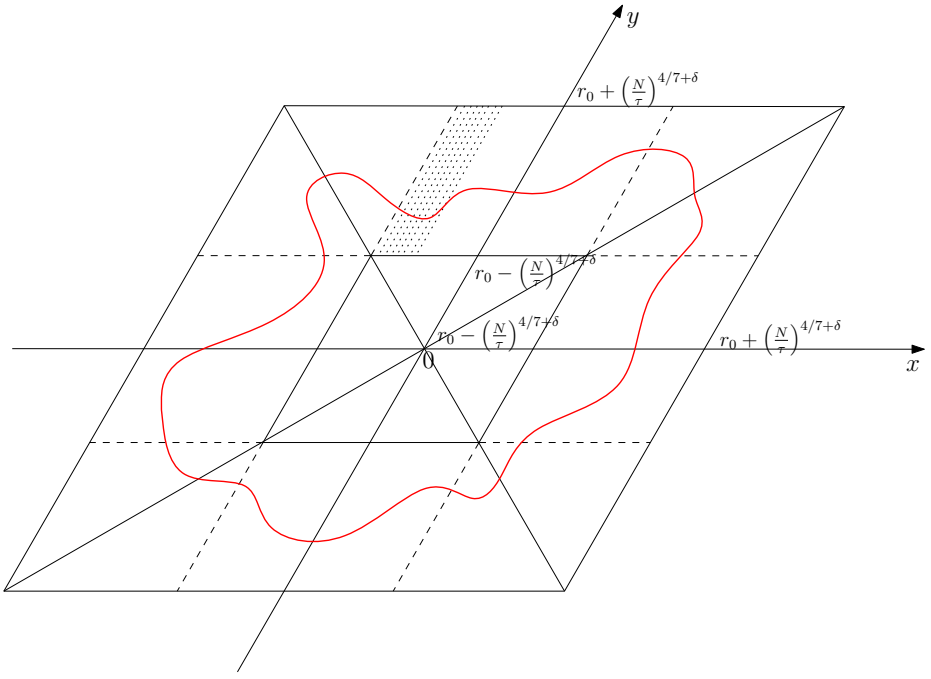


Figure 3.6: Four divisions (triangles) of $S_{r_0+\left(\frac{N}{\tau}\right)^{4/7+\delta}}$, each of them having vertex at the origin and determined by a side of $S_{r_0+\left(\frac{N}{\tau}\right)^{4/7+\delta}}$. The front is visualized by a red curve outside $S_{r_0-\left(\frac{N}{\tau}\right)^{4/7+\delta}}$.

Proof. As mentioned above, to prove this proposition we will work with $\mathcal{R}^*_{\lambda_N}$ and show that $\mathcal{R}^*_{\lambda_N}$ is not connected to boundary of lattice $\partial S_{r_0+(N/\tau)^{4/7+\delta}}$ by a vacant crossing with a very small probability on the condition $\mathcal{R}^*_{\lambda_N}$ is contained in parallelogram $S_{r_0+(N/\tau)^{4/7+\delta}}$. All crossings which are mentioned below will be vacant.

For that purpose, we divide $S_{r_0+(N/\tau)^{4/7+\delta}}$ into eight parts (figure 3.6) involving four parallelograms:

$$\begin{aligned}
S_N^0 &= [-r_0+(N/\tau)^{4/7+\delta}, r_0-(N/\tau)^{4/7+\delta}]\times[0, r_0+(N/\tau)^{4/7+\delta}],\\
S_N^1 &= [-r_0+(N/\tau)^{4/7+\delta}, r_0-(N/\tau)^{4/7+\delta}]\times[-r_0-(N/\tau)^{4/7+\delta}, 0],\\
S_N^2 &= [-r_0-(N/\tau)^{4/7+\delta}, 0]\times[-r_0+(N/\tau)^{4/7+\delta}, r_0-(N/\tau)^{4/7+\delta}],\\
S_N^3 &= [0, r_0+(N/\tau)^{4/7+\delta}]\times[-r_0+(N/\tau)^{4/7+\delta}, r_0-(N/\tau)^{4/7+\delta}];
\end{aligned}$$

and four rhombi

$$\begin{aligned}
S_N^4 &= [-r_0-(N/\tau)^{4/7+\delta}, -r_0+(N/\tau)^{4/7+\delta}]\times[r_0-(N/\tau)^{4/7+\delta}, r_0+(N/\tau)^{4/7+\delta}],\\
S_N^5 &= [r_0-(N/\tau)^{4/7+\delta}, r_0+(N/\tau)^{4/7+\delta}]\times[r_0-(N/\tau)^{4/7+\delta}, r_0+(N/\tau)^{4/7+\delta}],\\
S_N^6 &= [r_0-(N/\tau)^{4/7+\delta}, r_0+(N/\tau)^{4/7+\delta}]\times[-r_0-(N/\tau)^{4/7+\delta}, -r_0+(N/\tau)^{4/7+\delta}]\\
S_N^7 &= [-r_0-(N/\tau)^{4/7+\delta}, -r_0+(N/\tau)^{4/7+\delta}]\times[-r_0-(N/\tau)^{4/7+\delta}, -r_0+(N/\tau)^{4/7+\delta}].
\end{aligned}$$

Denote $\mathcal{B}_N^l = \partial S_{r_0+(N/\tau)^{4/7+\delta}} \cap \partial S_N^l$, $l = 0, ..., 7$. The part of $\mathcal{R}_{\lambda_N}^*$ crossing S_N^l is called $R_{\lambda_N}^{*l}$ (they are not necessarily disjoint).
In this proof we only consider the vacant crossings so we will use the probability parameter $p^* = 1-p$. Clearly, the probability that $\mathcal{R}_{\lambda_N}^*$ is not connected to $\partial S_{r_0+(N/\tau)^{4/7+\delta}}$ is smaller than

$$\begin{aligned}
&\prod_{l=0}^{7}\left(\mathbb{P}\left(R_{\lambda_N}^{*l} \text{ is not connected to } \mathcal{B}_N^l \text{ by a vacant crossing in } S_N^l\right)\right)\\
&\leq \prod_{l=0}^{3}\left(\mathbb{P}\left(R_{\lambda_N}^{*l} \text{ is not connected to } \mathcal{B}_N^l \text{ by a vacant crossing in } S_N^l\right)\right).
\end{aligned}$$

We restrain ourselves to the part of $\mathcal{R}_{\lambda_N}^*$ contained in the four parallelograms. For symmetry reasons, probability in these strips is the same. For the shake of simplicity, we fix $l = 0$ and try to find a upper bound for the probability that $R_{\lambda_N}^{*0}$ is not connected to $\mathcal{B}_N^0$ by a vacant path in S_N^0 (considered $\mathcal{B}_N^0$ is the top side of $S_{r_0+(N/\tau)^{4/7+\delta}}$).
As the previous part, we can see that $\mathcal{R}_{\lambda_N}^*$ is localized outside $S_{r_0-(N/\tau)^{4/7+\delta}}$ with high probability. For $\epsilon := \delta/4$, we divide the strip S_N^0 into $(N/\tau)^{3\epsilon}/6$ disjoint substrips of length $3(N/\tau)^{4/7+\epsilon}$ (it is not necessarily entirely covered),

$$\begin{aligned}
P_N^i = \left[-r_0+(N/\tau)^{4/7+\delta}+n_N^i, -r_0+(N/\tau)^{4/7+\delta}+3(N/\tau)^{4/7+\epsilon}+n_N^i\right]&\\
\times\left[0, r_0+(N/\tau)^{4/7+\delta}\right],&
\end{aligned}$$

with $(i = 0, ..., (N/\tau)^{3\epsilon}/6-1)$.
Let $r_{\lambda_N}^i$ be the "lowest" part of $R_{\lambda_N}^{*0}$ inside P_i^N. We are interested in the probability that $r_{\lambda_N}^i$ is connected to $\mathcal{B}_N^0$ by a vacant crossing inside P_N^i.

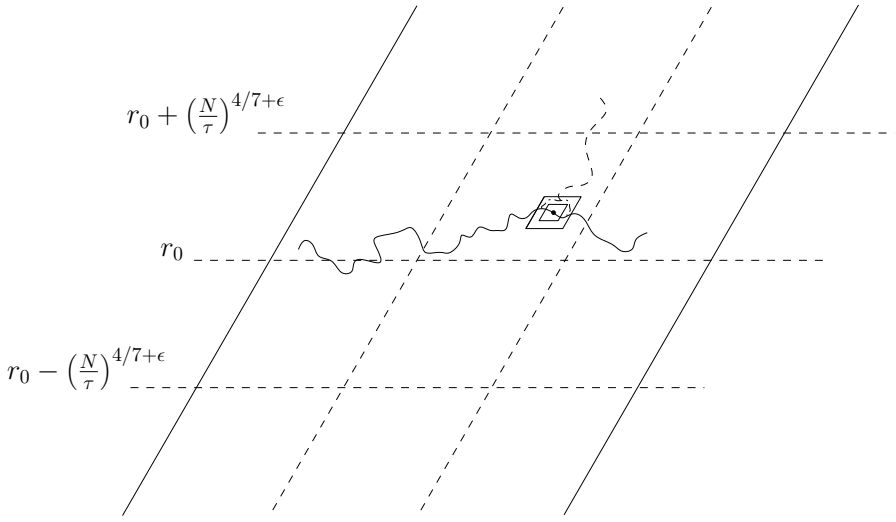


Figure 3.7: The circuit in annulus $S_{2\left(\frac{N}{\tau}\right)^{4/7-\epsilon}}(z) \backslash \mathring{S}_{\left(\frac{N}{\tau}\right)^{4/7-\epsilon}}(z)$ around z.

Fix an i (for notational convenience, we take $i = 0$ and $n_N^i = 0$). Firstly, note that, with probability at least $1/2$, there exists an occupied top-to-bottom crossing of the rhombus $\left[(N/\tau)^{4/7+\epsilon} - a_0, 2(N/\tau)^{4/7+\epsilon} - a_0\right] \times \left[r_0 - (N/\tau)^{4/7+\epsilon}, r_0\right]$ (supercritical region), with $a_0 = r_0 - (N/\tau)^{4/7+\delta}$ so that the "highest" point z on $r^0_{\lambda_N}$ in the middle part $\left[(N/\tau)^{4/7+\epsilon} - a_0, 2(N/\tau)^{4/7+\epsilon} - a_0\right] \times \left[0, r_0 + (N/\tau)^{4/7+\delta}\right]$ of the trip P_N^0 lies above the line $y = r_0$ with probability at least $1/2$.
Following [Nol08a], we will show that, in the case when $r^0_{\lambda_N}$ passes above the line $y = r_0$, the conditional probability that it is connected to the part of $\mathcal{B}_N^0$ in subtrip P_N^0 by a vacant crossing is bounded away from 0 by a quantity of order $(N/\tau)^{-2\epsilon}$.
To prove that, we choose z (above $y = r_0$) and define the annulus

$$S_{2(N/\tau)^{4/7-\epsilon}}(z) \backslash \mathring{S}_{(N/\tau)^{4/7-\epsilon}}(z)$$

around z described in figure 3.7. Since the annulus is contained in the region where vacant probability p^* is larger than

$$\begin{aligned} p_1^* = 1 - f\left(r_0 - 2(N/\tau)^{4/7-\epsilon}\right) &= 1 - e^{-\frac{\tau}{N}\left(r_0 - 2\frac{N}{\tau})^{4/7-\epsilon}\right)} \\ &= 1 - p_c e^{2(N/\tau)^{-3/7-\epsilon}}. \end{aligned}$$

Characteristic length corresponding to this value of the parameter is of order

$$L(p_1^*) = \left|p_c - 1 + p_c\left(1 + 2(N/\tau)^{-3/7-\epsilon}\right)\right|^{-4/3} \approx (N/\tau)^{4/7+4\epsilon/3} \gg (N/\tau)^{4/7-\epsilon}.$$

From RSW theory [Gri99], there is a probability of at least δ_4^4 to observe a vacant circuit in this annulus. Now, we want to know how this circuit is connected to $\mathcal{B}_N^0$ in P_N^0. Note that the part of the circuit that is above $r^0_{\lambda_N}$, together with $r^0_{\lambda_N}$, contains a vacant circuit around the segment $I_N = z + \left[-(N/\tau)^{4/7-\epsilon}, (N/\tau)^{4/7-\epsilon}\right] \times \left\{(N/\tau)^{4/7-\epsilon}\right\}$. We need the following lemma for critical percolation.

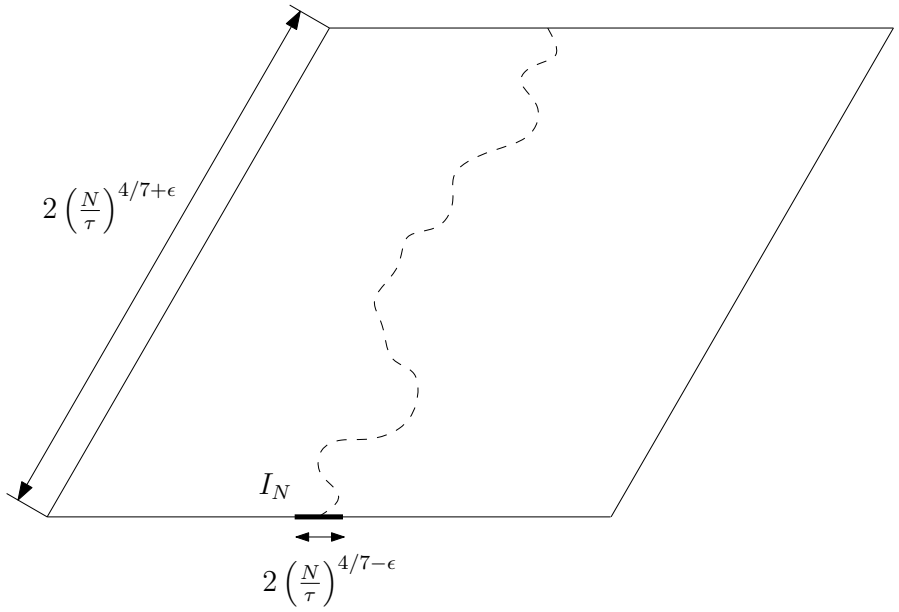


Figure 3.8: The crossing connect I_N to th top side of a rhombus.

Lemma 3.4.4. *Consider the rhombus*

$$\left[-(N/\tau)^{4/7+\epsilon}, (N/\tau)^{4/7+\epsilon}\right] \times \left[r_0, r_0 + 2(N/\tau)^{4/7+\epsilon}\right]$$

and the subinterval $I_N = \left[-(N/\tau)^{4/7-\epsilon}, (N/\tau)^{4/7-\epsilon}\right] \times \{r_0\}$ *on the bottom edge of this rhombus. Then the event* $\mathcal{C}^*_{V,I_N}\left(\left[-(N/\tau)^{4/7+\epsilon}, (N/\tau)^{4/7+\epsilon}\right] \times \left[r_0, r_0 + 2(N/\tau)^{4/7+\epsilon}\right]\right)$ *that there exists a vertical vacant crossing connecting* I_N *(figure 3.8) to the top edge has a probability of at least*

$$\mathbb{P}_{1/2}\left[\mathcal{C}^*_{V,I_N}\left(\left[-(N/\tau)^{4/7+\epsilon}, (N/\tau)^{4/7+\epsilon}\right] \times \left[r_0, r_0 + 2(N/\tau)^{4/7+\epsilon}\right]\right)\right] \geq C(N/\tau)^{-2\epsilon} \tag{3.4.5}$$

for some constant $C > 0$ *(depending neither on* N, τ *nor on* ϵ*).*

Proof. Consider the parallelogram $\left[0, (N/\tau)^{4/7+\epsilon}\right] \times \left[r_0, r_0 + 2(N/\tau)^{4/7+\epsilon}\right]$ and cover its bottom edge by less than $(N/\tau)^{2\epsilon}$ disjoint intervals

$$I^j_N = \left[n'^j_N - (N/\tau)^{4/7+\epsilon}, n'^j_N + (N/\tau)^{4/7+\epsilon}\right] \times \{r_0\}$$

of length $2(N/\tau)^{4/7-\epsilon}$. From RSW theorem, we see that there is a $\delta_2 > 0$ such that

$$\begin{aligned}
&\sum_j \mathbb{P}_{1/2}\left[\mathcal{C}^*_{V,I^j_N}\left(\left[0, (N/\tau)^{4/7+\epsilon}\right] \times \left[r_0, r_0 + 2(N/\tau)^{4/7+\epsilon}\right]\right)\right] \\
&\quad \geq \mathbb{P}_{1/2}\left[\mathcal{C}^*_V\left(\left[0, (N/\tau)^{4/7+\epsilon}\right] \times \left[r_0, r_0 + 2(N/\tau)^{4/7+\epsilon}\right]\right)\right] \\
&\quad \geq \delta_2.
\end{aligned}$$

But, for each j we have

$$\begin{aligned}
&\mathbb{P}_{1/2}\left[\mathcal{C}^*_{V,I_N^j}\left(\left[0,(N/\tau)^{4/7+\epsilon}\right]\times\left[r_0,r_0+2(N/\tau)^{4/7+\epsilon}\right]\right)\right]\\
&\quad\le \mathbb{P}_{1/2}\left[\mathcal{C}^*_{V,I_N^j}\left(\left[n'^j_N-(N/\tau)^{4/7+\epsilon},n'^j_N+(N/\tau)^{4/7+\epsilon}\right]\times\left[r_0,r_0+2(N/\tau)^{4/7+\epsilon}\right]\right)\right]\\
&\quad= \mathbb{P}_{1/2}\left[\mathcal{C}^*_{V,I_N}\left(\left[-(N/\tau)^{4/7+\epsilon},(N/\tau)^{4/7+\epsilon}\right]\times\left[r_0,r_0+2(N/\tau)^{4/7+\epsilon}\right]\right)\right].
\end{aligned}$$

From these inequalities we can obtain

$$(N/\tau)^{2\epsilon}\times\mathbb{P}_{1/2}\left[\mathcal{C}^*_{V,I_N}\left(\left[-(N/\tau)^{4/7+\epsilon},(N/\tau)^{4/7+\epsilon}\right]\times\left[r_0,r_0+2(N/\tau)^{4/7+\epsilon}\right]\right)\right]\ge\delta_2.$$

□

By using the FKG inequality, we get that, for each $i\le(N/\tau)^{3\epsilon}/6-1$, the probability that $R^{*0}_{\lambda_N}$ is connected to the part of $\mathcal{B}^0_N$ in P^i_N by a vacant crossing is bounded from below by $C'(N/\tau)^{-2\epsilon}$, independently for each i as following:

$$\begin{aligned}
&\mathbb{P}\left(r^0_{\lambda_N}\text{ is connected }\mathcal{B}^0_N\text{ by a vacant path in }P^i_N\right)\\
&\ge\mathbb{P}\left(\left\{r^0_{\lambda_N}\text{ is connected }\mathcal{B}^0_N\text{ by a vacant path in }P^i_N\right\}\right.\\
&\qquad\left.\cap\left\{r^0_{\lambda_N}\text{ passes above line }y=r_0\text{ at some }z\right\}\right)\\
&\ge\frac{1}{2}\mathbb{P}\left(\left\{\text{for some }z\text{ on }\{y=r_0\}\text{ there exists a circuit in }S_{2(N/\tau)^{4/7-\epsilon}}(z)\backslash\mathring{S}_{(N/\tau)^{4/7-\epsilon}}(z)\right\}\right.\\
&\qquad\left.\cap\left\{\text{ the circuit connect to }\mathcal{B}^0_N\text{ in }P^i_N\right\}\right)\\
&\ge\frac{1}{2}\delta_4^4C(N/\tau)^{-2\epsilon}\ge C'(N/\tau)^{-2\epsilon}.
\end{aligned}$$

We denote by A_{λ_N} the event that $\mathcal{R}^*_{\lambda_N}$ is connected to $\partial S_{r_0+(N/\tau)^{4/7+\delta}}$ by a vacant crossing and by $A^c_{\lambda_N}$ the complement of A_{λ_N}. On the conditions of $\mathcal{R}^*_{\lambda_N}$ stays inside $S_{r_0+(N/\tau)^{4/7+\delta}}$, it follows

$$\begin{aligned}
\mathbb{P}\left(A_{\lambda_N}\right)&=1-\mathbb{P}\left(A^c_{\lambda_N}\right)\\
&\ge1-\left(\mathbb{P}\left(R^{*0}_{\lambda_N}\text{ is not connected to }\mathcal{B}^0_N\text{ in }S^0_N\right)\right)^4\\
&\ge1-\left(1-C'(N/\tau)^{-2\epsilon}\right)^{4(N/\tau)^{3\epsilon}/6}\\
&\ge1-\left(e^{-C'(N/\tau)^{-2\epsilon}}\right)^{\frac{4}{6}(N/\tau)^{3\epsilon}}\\
&\ge1-e^{-N^{\epsilon'}}
\end{aligned}$$

for some positive ϵ'.

We can conclude that

$$\begin{aligned}
\mathbb{P}\left(\rho_{\lambda_N}\equiv\rho^*_{\lambda_N}\right)&\ge\mathbb{P}\left(A_{\lambda_N}\mid\mathcal{R}^*_{\lambda_N}\subset S_{r_0+(N/\tau)^{4/7+\delta}}\right)\mathbb{P}\left(\mathcal{R}^*_{\lambda_N}\subset S_{r_0+(N/\tau)^{4/7+\delta}}\right)\\
&\ge\left(1-e^{-N^{\epsilon'}}\right)\left(1-e^{-N^{\delta'}}\right)\\
&\ge1-e^{-N^{\delta''}},
\end{aligned}$$

for some positive δ''. □

The lemma 3.4.4 still holds with the event $\mathcal{C}_{H,I_N}$ with I_N on the top side of the rhombus $\left[(N/\tau)^{4/7+\epsilon}, (N/\tau)^{4/7+\epsilon}\right] \times \left[r_0, r_0 + 2(N/\tau)^{4/7+\epsilon}\right]$.
In the next sections, we will consider the unique front and denote it by $\mathcal{F}_{\lambda_N}$. It is exactly the set of the edges from which two arms can be drawn - one occupied to the center 0 and one vacant to infinity.

3.4.3 Length of the front

In this part we will study the length of the front T_{λ_N}, that is, its number of edges. As the previous part, we need a two arm probability estimate for this purpose. The expression *two arms* here refers to one occupied (black) arm and one vacant (white) arm. To simplify notation, we denote an edge by e and x_e is one of two neighboring sites which is assigned to e arbitrarily and permanently.
For a parallelogram R and a site v contained in its interior, we define the event of existence of four arms from v to ∂R:

$$\begin{aligned}\Gamma_4(v,R) =&\{\text{there exist four paths } r_1,...,r_4 \text{ from } \partial v \text{ to } \partial R, \text{ ordered clockwisely,}\\ &\text{such that } r_1, r_3 \text{ are occupied and } r_2, r_4 \text{ are vacant}\},\end{aligned}$$

where ∂v is the set of vertices neighboring v and ∂R is the boundary of R.
The equivalent two arms event is

$$\Gamma_2(v,R) = \{\text{there exist an occupied path } r_1 \text{ and a vacant path } r_2 \text{ from } \partial v \text{ to } \partial R\}.$$

For a probability distribution $\hat{P}$ on the vertices of the lattice we will say that $\hat{P}$ between $\mathbb{P}_p$ and $\mathbb{P}_q$ whenever the coloring probability of any site z is beween p and q. Nolin proved the following lemma in his paper [Nol08a],

Lemma 3.4.5. *Uniformly in p, $\hat{P}$ between $\mathbb{P}_p$ and $\mathbb{P}_{1-p}$, $n \leq L(p)$, we have*

$$\hat{P}\left[\Gamma_2\left(0, S_n\right)\right] \asymp \mathbb{P}_{1/2}\left[\Gamma_2\left(0, S_n\right)\right]. \tag{3.4.6}$$

We will need to extend the paths r_1, r_2 outside R. For this purpose, we will define the event Δ_2 (figure 3.9). We first consider the strips

$$\begin{aligned}\mathcal{A}(1,k) &:= \left[-2^{k-1}, 2^{k-1}\right] \times \left[-2^k, -2^{k-1}\right]\\ \mathcal{A}(2,k) &:= \left[-2^{k-1}, 2^{k-1}\right] \times \left[2^{k-1}, 2^k\right].\end{aligned}$$

Define, for a site v in $S_{2^{k-1}}$

$$\begin{aligned}\Delta_2\left(v, S_{2^k}\right) =&\{\Gamma_2\left(v, S_{2^k}\right) \text{ occurs, with } r_i \cap \left(S_{2^k} \backslash \mathring{S}_{2^{k-1}}\right) \subseteq \mathcal{A}(i,k),\\ &\text{and there exist an occupied horizontal crossing of } \mathcal{A}(1,k)\\ &\text{and a vacant horizontal crossing of } \mathcal{A}(2,k)\}.\end{aligned}$$

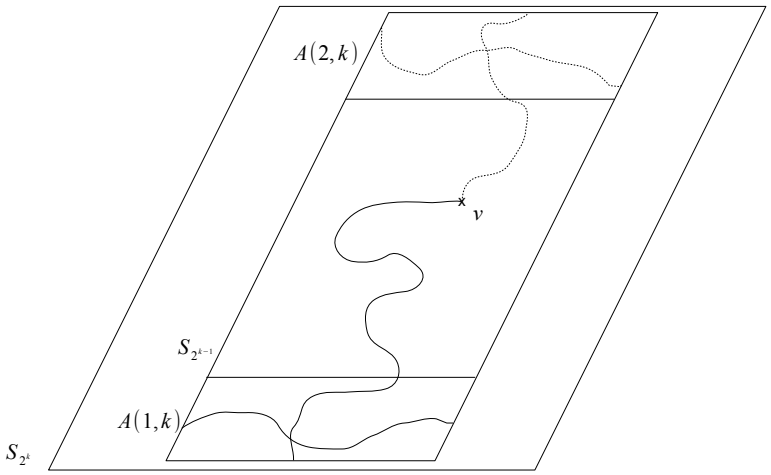


Figure 3.9: The illustration of the event $\Delta_2(v, S_{2^k})$.

We similarly define, for a parallelogram R' contained in the interior of S_{2^k},

$$\begin{aligned}\tilde{\Gamma}_2(S_{2^k}, R') = \{&\text{There exists an occupied path } r_1 \text{ and a vacant path } r_2\\ &\text{from } \partial R' \text{ to the bottom and top edges, respectively, of } S_{2^k}\\ &\text{which are (with the exception of their extremities on } \partial R')\\ &\text{contained in } S_{2^k} \backslash R'\}.\end{aligned}$$

These events will be used in the proof in the next subsection. Moreover, the probablities of these events are comparable:

Remark 3.4.6. (analog of Lemma 6 in [Kes87]). There exist constants C_1 and C_2 such that

$$C_1 \hat{P}(\Gamma_2(0, S_{2^k})) \leq \hat{P}\left(\tilde{\Gamma}_2(S_{2^k}, S_{2^j})\right) \hat{P}(\Gamma_2(v, S_{2^j})) \leq C_2 \hat{P}(\Delta_2(0, S_{2^k})) \tag{3.4.7}$$

uniformly in $\hat{P}$ between $\mathbb{P}_p$ and $\mathbb{P}_{1-p}$ for all p and $j \leq k-2$, $2^k \leq L(p)$.

We will continuously study properties of the front in the "critical region". Firstly, we will estimate the expectation of the length of the front T_{λ_N}.

Proposition 3.4.7. *For all $\epsilon > 0$, we have, for N sufficiently large, $\lambda_N = \frac{\tau}{N}$ ($\tau \geq 1$),*

$$\lambda_N^{-(3/7-\epsilon)} r_0 \leq \mathbb{E}[T_{\lambda_N}] \leq \lambda_N^{-(3/7+\epsilon)} r_0 \tag{3.4.8}$$

Proof. We have

$$\mathbb{E}[T_{\lambda_N}] = \sum_{e \in \mathcal{S}} \mathbb{P}(e \in \mathcal{F}_{\lambda_N}). \tag{3.4.9}$$

- First of all, we consider the upper bound. Choose $\epsilon' = \epsilon/4$; we have

$$\mathbb{E}\left[T_{\lambda_N}\right] \leq 6\left|B_{\lambda_N}\right| \mathbb{P}\left(\mathcal{F}_{\lambda_N} \nsubseteq S_{r_0 \pm (N/\tau)^{4/7+\epsilon'}}\right) + \sum_{e \in S_{r_0 \pm (N/\tau)^{4/7+\epsilon'}}} \mathbb{P}\left(e \in \mathcal{F}_{\lambda_N}\right) \quad (3.4.10)$$

with the number of occupied points $|B_{\lambda_N}| < +\infty$ (see equation (3.3.3)). Following the theorem 3.4.2 we obtain that the first term tends to 0 subexponentially fast. Therefore, it is only necessary to verify summation over the edges inside $S_{r_0 \pm (N/\tau)^{4/7+\epsilon'}}$ (the second term of the inequality).
For $e \in S_{r_0 \pm (N/\tau)^{4/7+\epsilon'}}$,

$$\mathbb{P}\left(e \in \mathcal{F}_{\lambda_N}\right) \leq \mathbb{P}\left[\Gamma_2\left(x_e, S_{(N/\tau)^{4/7-2\epsilon'}}(x_e)\right)\right]$$

and $S_{(N/\tau)^{4/7-2\epsilon'}}(x_e) \subseteq S_{r_0 \pm 2(N/\tau)^{4/7+\epsilon'}}$, the percolation probability p in this box is contained in the zone

$$\begin{aligned} f\left(r_0 + 2(N/\tau)^{4/7+\epsilon'}\right) \leq p \leq f\left(r_0 - 2(N/\tau)^{4/7+\epsilon'}\right) \\ p_c e^{-2(N/\tau)^{-3/7+\epsilon'}} \leq p \leq p_c e^{2(N/\tau)^{-3/7+\epsilon'}} \end{aligned}$$

The corresponding characteristic length

$$L(p) \approx \left|2(N/\tau)^{-3/7+\epsilon'}\right|^{-4/3} \approx (N/\tau)^{4/7-4\epsilon'/3} \gg (N/\tau)^{4/7-2\epsilon'},$$

and by using lemma 3.4.5 and proposition 3.2.1, we can get that

$$\begin{aligned} \mathbb{P}\left[\Gamma_2\left(x_e, S_{(N/\tau)^{4/7-2\epsilon'}}(x_e)\right)\right] &\asymp \mathbb{P}_{1/2}\left[\Gamma_2\left(x_e, S_{(N/\tau)^{4/7-2\epsilon'}}(x_e)\right)\right] \\ &\approx \left((N/\tau)^{4/7-2\epsilon'}\right)^{-1/4} \\ &\ll (N/\tau)^{-1/7+\epsilon'}. \end{aligned}$$

Put all estimates back in the inequality (3.4.10): with N large enough we get

$$\begin{aligned} \sum_{e \in S_{r_0 \pm (N/\tau)^{4/7+\epsilon'}}} \mathbb{P}\left(e \in \mathcal{F}_{\lambda_N}\right) &\leq 6\left|S_{r_0 \pm (N/\tau)^{4/7+\epsilon'}}\right| (N/\tau)^{-1/7+\epsilon'} \\ &\leq 6\left[(2r_0+2)\,2(N/\tau)^{4/7+\epsilon'}\right](N/\tau)^{-1/7+\epsilon'} \\ &\leq (N/\tau)^{3/7+\epsilon} r_0. \end{aligned}$$

- Next, we turn to the lower bound.
Choose $\epsilon' := \epsilon/6$ and take an edge $e \in S_{r_0 \pm (N/\tau)^{4/7-\epsilon'}}$. For such an edge, we would like to estimate the probability of having two arms, one occupied to

center and one vacant to boundary of S_N, so we will use events Δ_2 and Γ_2. Indeed, take j such that $(N/\tau)^{4/7-\epsilon'} < 2^j < 2(N/\tau)^{4/7-\epsilon'}$, the probability of having two arms is at least

$$\mathbb{P}\left(\Delta_2\left(x_e, S_{2^j}(x_e)\right)\right)$$

These paths can be extended so that they go out of $S_{r_0\pm(N/\tau)^{4/7+\epsilon'}}$. Lemma 3.4.4 implies that this can be done with probability at least

$$\left(C(N/\tau)^{-2\epsilon'}\right)^2 = C'(N/\tau)^{-4\epsilon'}.$$

On the other hand, these sites stay in $S_{r_0\pm 3(N/\tau)^{4/7-\epsilon'}}$ so the percolation probability

$$\begin{aligned} f\left(r_0 + 3(N/\tau)^{4/7-\epsilon'}\right) \leq p \leq f\left(r_0 - 3(N/\tau)^{4/7-\epsilon'}\right) \\ p_c e^{-3(N/\tau)^{-3/7-\epsilon'}} \leq p \leq p_c e^{3(N/\tau)^{-3/7-\epsilon'}} \end{aligned}$$

and the corresponding characteristic length

$$L(p) \approx (N/\tau)^{4/7+4\epsilon'/3}.$$

From that we can get that

$$\begin{aligned} \mathbb{P}\left(\Delta_2\left(x_e, S_{2^j}(x_e)\right)\right) &\asymp \mathbb{P}\left(\Gamma_2\left(x_e, S_{2^j}(x_e)\right)\right) \\ &\asymp \mathbb{P}_{1/2}\left(\Gamma_2\left(x_e, S_{2^j}(x_e)\right)\right) \\ &\geq \mathbb{P}_{1/2}\left(\Gamma_2\left(0, S_{2(N/\tau)^{4/7-\epsilon'}}\right)\right) \end{aligned}$$

and the proposition 3.2.1 implies that

$$\mathbb{P}_{1/2}\left(\Gamma_2\left(0, S_{2(N/\tau)^{4/7-\epsilon'}}\right)\right) \approx \left(2(N/\tau)^{4/7-\epsilon'}\right)^{-1/4} \gg (N/\tau)^{-1/7}. \tag{3.4.11}$$

With the edges in $S_{r_0\pm(N/\tau)^{4/7-\epsilon'}}$ like that and for N large enough we can construct two arms going out of $S_{r_0\pm(N/\tau)^{4/7+\epsilon'}}$ with probability at least $(N/\tau)^{-1/7-4\epsilon'}$. As the theorem 3.4.2, the front will be stay in $S_{r_0\pm(N/\tau)^{4/7+\epsilon'}}$ with probability $1-\epsilon_N$, for some ϵ_N (independent of e) tending to 0 subexponentially fast. We obtain the following conclusion for the lower bound, with N large enough,

$$\begin{aligned} \sum_{e\in\mathcal{S}} \mathbb{P}\left(e \in \mathcal{F}_{\lambda_N}\right) &\geq \sum_{e\in S_{r_0\pm(N/\tau)^{4/7-\epsilon'}}} \mathbb{P}\left(e \in \mathcal{F}_{\lambda_N}\right) \\ &\geq \left|S_{r_0\pm(N/\tau)^{4/7-\epsilon'}}\right| \left((N/\tau)^{-1/7-4\epsilon'} - \epsilon_N\right) \\ &\geq \left[4\left(r_0+1\right)(N/\tau)^{4/7-\epsilon'}\right](N/\tau)^{-1/7-4\epsilon'} \\ &\geq r_0(N/\tau)^{3/7-\epsilon}. \end{aligned}$$

□

The next property of the front considered is variance of the length, $\mathrm{Var}\left[T_{\lambda_N}\right]$.

Proposition 3.4.8. *For all $\epsilon > 0$, we have, for N sufficiently large, $\lambda_N = \frac{\tau}{N}$,*

$$\mathrm{Var}\left[T_{\lambda_N}\right] \leq \lambda_N^{-(10/7+\epsilon)} r_0. \tag{3.4.12}$$

Proof. For an edge, we denote

$$F_e = \{\text{the event } e \in \mathcal{F}_{\lambda_N}\}.$$

We can write as following

$$T_{\lambda_N} = \sum_e \mathbb{I}_{F_e}$$

and

$$\mathbb{E}\left[T_{\lambda_N}\right] = \sum_e \mathbb{P}\left(F_e\right).$$

From that we can get that

$$\begin{aligned}
\mathrm{Var}\left[T_{\lambda_N}\right] = \mathbb{E}\left[T_{\lambda_N}^2\right] - \left(\mathbb{E}\left[T_{\lambda_N}\right]\right)^2 &= \mathbb{E}\left[\left(\sum_e \mathbb{I}_{F_e}\right)^2\right] - \left(\sum_e \mathbb{P}\left(F_e\right)\right)^2 \\
&= \mathbb{E}\left[\sum_e \sum_f \mathbb{I}_{F_f} \mathbb{I}_{F_e}\right] - \sum_e \mathbb{P}\left(F_e\right) \sum_f \mathbb{P}\left(F_f\right) \\
&= \sum_{e,f} \mathbb{P}\left(F_e \cap F_e\right) - \sum_e \mathbb{P}\left(F_e\right) \sum_f \mathbb{P}\left(F_f\right) \\
&= \sum_{e,f} \left[\mathbb{P}\left(F_e \cap F_e\right) - \mathbb{P}\left(F_e\right) \mathbb{P}\left(F_f\right)\right].
\end{aligned}$$

Take $\epsilon' = \epsilon/8$, we can restrict the summation to the edges $e, f \in S_{r_0 \pm (N/\tau)^{4/7+\epsilon'}}$

$$\begin{aligned}
\mathrm{Var}\left[T_{\lambda_N}\right] \leq 6 \left|B_{\lambda_N}\right|^2 \mathbb{P}\left(\mathcal{F}_{\lambda_N} \nsubseteq S_{r_0 \pm (N/\tau)^{4/7+\epsilon'}}\right) \\
+ \sum_{e,f \in S_{r_0 \pm (N/\tau)^{4/7+\epsilon'}}} \left[\mathbb{P}\left(F_e \cap F_e\right) - \mathbb{P}\left(F_e\right) \mathbb{P}\left(F_f\right)\right].
\end{aligned} \tag{3.4.13}$$

Similarly to the above, we see that the first term tends to 0 subexponentially fast and it is replaced by ϵ_N.
We will replace F_e by the event $\tilde{F}_e$ which depends only on sites in a box around e of size $(N/\tau)^{4/7+\epsilon'}$ and is defined as

$$\tilde{F}_e := \left\{2 \text{ arms } e \rightsquigarrow \text{ top and bottom sides of } \partial S_{(N/\tau)^{4/7+\epsilon'}}\left(x_e\right)\right\}.$$

We thus have

$$\mathrm{Var}\left[T_{\lambda_N}\right] = \sum_{e,f}\left[\mathbb{P}\left(\tilde{F}_e \cap \tilde{F}_f\right) - \mathbb{P}\left(\tilde{F}_e\right)\mathbb{P}\left(\tilde{F}_f\right)\right] + \epsilon_N,$$

with ϵ_N replacing the first term in (3.4.13) and tending to 0 subexponentially fast. We fix an edge e and the box $S_{2(N/\tau)^{4/7+\epsilon'}}(x_e)$. We can easily see that if $f \notin S_{2(N/\tau)^{4/7+\epsilon'}}(x_e)$ then

$$\mathbb{P}\left(\tilde{F}_e \cap \tilde{F}_f\right) - \mathbb{P}\left(\tilde{F}_e\right)\mathbb{P}\left(\tilde{F}_f\right) = 0.$$

So, we will only consider $f \in S_{2(N/\tau)^{4/7+\epsilon'}}(x_e)$.
Now we will estimate the summation

$$\sum_{f \in S_{2(N/\tau)^{4/7+\epsilon'}}(x_e)} \mathbb{P}\left(\tilde{F}_e \cap \tilde{F}_f\right).$$

For notational convenience, let $d = d(x_e, x_f)$.

- If $d > (N/\tau)^{4/7-2\epsilon'}/4$, by using the proposition 3.2.1 for the boxes $S_{(N/\tau)^{4/7-2\epsilon'}/8}(x_e)$ and $S_{(N/\tau)^{4/7-2\epsilon'}/8}(x_f)$ we get

$$\begin{aligned}
\mathbb{P}\left(\tilde{F}_e \cap \tilde{F}_f\right) &\leq \mathbb{P}\left(\Gamma_2\left(x_e, S_{(N/\tau)^{4/7-2\epsilon'}/8}(x_e)\right) \cap \Gamma_2\left(x_f, S_{(N/\tau)^{4/7-2\epsilon'}/8}(x_f)\right)\right) \\
&= \mathbb{P}\left(\Gamma_2\left(x_e, S_{(N/\tau)^{4/7-2\epsilon'}/8}(x_e)\right)\right)\mathbb{P}\left(\Gamma_2\left(x_f, S_{(N/\tau)^{4/7-2\epsilon'}/8}(x_f)\right)\right) \\
&\leq \left[\left((N/\tau)^{4/7-2\epsilon'}\right)^{-1/4}\right]^2 \\
&\leq (N/\tau)^{-2/7+2\epsilon'},
\end{aligned}$$

  for N large enough. We thus have

$$\begin{aligned}
\sum_{f \in \left(S_{2(N/\tau)^{4/7+\epsilon'}}(x_e) \setminus S_{(N/\tau)^{4/7+\epsilon'}/4}(x_e)\right)} & \mathbb{P}\left(\tilde{F}_e \cap \tilde{F}_f\right) \\
&\leq 6\left(4(N/\tau)^{4/7+\epsilon'} + 1\right)^2 (N/\tau)^{-2/7+2\epsilon'} \\
&\leq (N/\tau)^{6/7+5\epsilon'}.
\end{aligned}$$

- With $4 \leq d \leq (N/\tau)^{4/7-2\epsilon'}/4$, we see that

$$\begin{aligned}
\mathbb{P}\left(\tilde{F}_e \cap \tilde{F}_f\right) \leq \mathbb{P}\left(\Gamma_2\left(x_e, S_{d/2}(x_e)\right)\right)\mathbb{P}\left(\Gamma_2\left(x_f, S_{d/2}(x_f)\right)\right) \\
\times \mathbb{P}\left(\tilde{\Gamma}_2\left(S_{2(N/\tau)^{4/7-2\epsilon'}}(x_e), S_{2d}(x_e)\right)\right).
\end{aligned}$$

Using the remark 3.4.6, we can estimate the first and the third terms

$$\begin{aligned}\mathbb{P}\left(\Gamma_2\left(x_e, S_{d/2}\left(x_e\right)\right)\right)\mathbb{P}\left(\tilde{\Gamma}_2\left(S_{2(N/\tau)^{4/7-2\epsilon'}}\left(x_e\right), S_{2d}\left(x_e\right)\right)\right) &\leq C_1\mathbb{P}\left(\Delta_2\left(x_e, S_{2(N/\tau)^{4/7-2\epsilon'}}\left(x_e\right)\right)\right)\\ &\leq C_1\mathbb{P}\left(\Gamma_2\left(x_e, S_{2(N/\tau)^{4/7-2\epsilon'}}\left(x_e\right)\right)\right)\\ &\leq (N/\tau)^{-1/7+\epsilon'}.\end{aligned}$$

Hence, by using the fact that there are at most $C_2 j$ edges at a distance j from e and the proposition 3.2.1 we can get

$$\begin{aligned}&\sum_{f\in\left(S_{(N/\tau)^{4/7-2\epsilon'}/4}(x_e)\setminus S_4(x_e)\right)} \mathbb{P}\left(\tilde{F}_e\cap\tilde{F}_f\right)\\ &\leq \sum_{j=4}^{(N/\tau)^{4/7-2\epsilon'}/4} C_2 j(N/\tau)^{-1/7+\epsilon'}\mathbb{P}\left(\Gamma_2\left(x_f, S_{j/2}\left(x_f\right)\right)\right)\\ &\leq C_3(N/\tau)^{-1/7+\epsilon'}\sum_{j=4}^{(N/\tau)^{4/7-2\epsilon'}/4} j(j)^{-1/4+\epsilon'}\\ &\leq C_4(N/\tau)^{-1/7+\epsilon'}\left((N/\tau)^{4/7-2\epsilon'}/4\right)\left((N/\tau)^{4/7-2\epsilon'}/4\right)^{3/4+\epsilon'}\\ &\leq (N/\tau)^{6/7+2\epsilon'}.\end{aligned}$$

- Finally, we consider the edges at a distance $d < 4$ then

$$\begin{aligned}\sum_{f\in S_4(x_e)}\mathbb{P}\left(\tilde{F}_e\cap\tilde{F}_f\right) &\leq C_5\mathbb{P}\left(\tilde{F}_e\right)\\ &\leq C_5\mathbb{P}\left(\Gamma_2\left(x_e, S_4\left(x_e\right)\right)\right)\\ &\leq C_5(N/\tau)^{-1/7+\epsilon'}\\ &\ll (N/\tau)^{6/7+2\epsilon'}.\end{aligned}$$

Summing three above contributions, we get that

$$\sum_{f\in S_{2(N/\tau)^{4/7+\epsilon'}}(x_e)}\mathbb{P}\left(\tilde{F}_e\cap\tilde{F}_f\right) \leq (N/\tau)^{6/7+6\epsilon'}.$$

Hence,

$$\begin{aligned}\mathrm{Var}\left[T_{\lambda_N}\right] &\leq 6\left[3\left(2(N/\tau)^{4/7+\epsilon'}+1\right)^2+6\left(2(N/\tau)^{4/7+\epsilon'}+1\right)\left(r_0-(N/\tau)^{4/7+\epsilon'}\right)\right]\\ &\qquad\times(N/\tau)^{6/7+6\epsilon'}+\epsilon_N\\ &\leq (N/\tau)^{10/7+\epsilon}r_0.\end{aligned}$$

□

From the propositions 3.4.7 and 3.4.3 we can immediately get that

$$\mathrm{Var}\left[T_{\lambda_N}\right] = o\left(\mathbb{E}\left[T_{\lambda_N}\right]^2\right) \tag{3.4.14}$$

and the following theorem

Theorem 3.4.9. *We have*

$$\frac{T_{\lambda_N}}{\mathbb{E}\left[T_{\lambda_N}\right]} \longrightarrow 1 \quad \textit{in } L^2 \textit{ as } N \to \infty. \tag{3.4.15}$$

Chapter 4

Simulation of the growth of cities

4.1 Modeling of Urban Growth

4.1.1 Introduction

A model, in simple sense, is a representation of a physical system to simulate reality. It allows scientists to have ability to predict the future evolution of the systems, to study system dynamics and to give hints for data collection and design of experiments. Models are basically built by consideration of the pertinent physical principles, operated by logic, modified by experimental judgment and plain intuition. According to M. Batty , "models act as a vehicle to enable experimentation with theory in a predictive sense, and to enhance understanding which may be prior to predictions of situations".

Urban modeling is the process of defining, building, and applying models for specific purposes in physical planning. It has a great influence on the social and human geographies built around location theory and spatial analysis.
The modeling of urbanism has begun with von Thunen's study of the spatial distribution of production in 1826 [vT26]. The prominent regional scientist Walter Isard has called him "the father of location theorists". In the model, he compares the relationships between markets, productions and distance; he also shows how market's processing could determine how land in different locations would be used. It essentially argued that industries are located according to the balance between their spatial patterns of demand and supply. While its generalization to populations sought to show how cities were structured hierarchically from the largest to the smallest according to demand in their hinterlands for the services they provided. After that, Weber (1909) [Web09] formulated a least cost theory of industrial location which tries to explain and predict the located pattern of the industry at a macro-scale. He assumes that an industry will choose its location based on min-

imizing transportation cost of raw materials and markets, minimizing labor costs, and maximizing agglomeration economies. The savings which would be made if the three things were located together are calculated for each plant.

The central place theory, which was created by the German geographer Walter Christaller (1933) [Chr33], attempted to explain the reasons behind the distribution patterns, size, and number of cities and towns around the world. He tested the model on Southern Germany and came to the conclusion that people gather together in cities to share goods and ideas and that they exist for purely economic reason. This has been linked to industrial location theory in a coherent economic framework by Losch (1943, 1954) [Los43]. He established inter-urban theory based on the ideas that systems of cities were also organized spatially as overlapping hierarchical fields while it was picked up by those concerned with the shape or morphology of cities. And armed with ideas about how gravitation and potential might condition human location, transportation modeling began in the early 1950s closely followed by its extension to embrace land use.

Urban modeling began to grow strongly in the 1950s after the appearance of the electronic digital computer and computers entered commerce in the form of mainframes, engineers. However, the foundation of computation was invented before by a Russian economist, Wassily Leontieff, who emigrated to the United Sates after that, with the input-output model of the economy in the 1920s. Based on this and the appearance of digital conceptions, policy markets began to think about how to use all of them in their problem-solving and decision-making.

There have been many classifications of urban models from different perspectives. Batty (2009) [Bat09] has suggested to group them into three main classes: (1) land-use-transportation model; (2) urban dynamics model; and (3) cellular automata, agent-based model and micro-simulation. However, this classification has several problems; for example, cellular automata and agent-based models are different, and there is no neural networks and fractal based modeling. Meanwhile, Pooyandeh et al. (2007) [PMAS07] have classified spatio-temporal urban models into two distinct classes: (1) complexity model, and (2) temporal GIS model. Complexity model contains CA based modeling, agent based modeling, neural network based modeling and fractal based modeling. However, they have also overlooked several other models, for example, theoretical models. From that, in the book [Bha10], B. Bhatta has suggested a classification:

i. Theoretical models.
ii. Aggregate-level urban dynamics models.
iii. Complexity science-based models (ANN)-based models.
    (a) Cell-based dynamics models.
    (b) Agent-based models.

(c) Artificial neural network.

(d) Fractal geometry-based models.

iv. Rule-based land-use and transport models.

Most of these models was used in modeling a static or an aggregate-level dynamic urban. However, urban is a dynamic system. This opens an interesting research in simulating the urban growth and predicting future urban dynamics. Along with the development of computer technology and information, input data in urban models are also widened and more copious. This allows models of urban dynamic to be improved and strengthened.
One of approachs being developed is *top-down* usually used in land-use and transport models. This approach consists in a multi-scale analysis of the city area from the smallest scale to the biggest and allows to take under account local geographic specificities [Weg11]. Multi-level, multi-scale model is a challenging necessity in modeling growth of city. However, this issue seems to be highly complicated. Our model, which is described in the next parts, is a disaggregated simple scale model based on percolation.

In recent years, modeling of urban growth has become important to city planners, economists, ecologists and resource managers. There are two main reasons of urban growth modeling: (1) the need to improve understanding of the cause-effect relationships in urban dynamic, (2) contribute to decision of urban growth management.
From the above statements, modeling of urbanism appears to be an attracting problem and there are many different approaches. If considered on geometry aspect, urban structure looks irregular, complex and difficult to describe by Euclidean geometries. Since the appearance of a new geometry, fractal that has been christened by Benoit Mandelbrot [Man82], we have had one more method in modeling the morphology of city.

4.1.2 Urban as Fractal

One of interested aspects of urban system is urban spatial structure which concerns the arrangement of public and private space in cities and the degree of connectivity and accessibility. There are three famous models developed by scientists of University of Chicago to help to explain where different types of people tend to live in an urban area - the concentric zone [Bur25], sector [Hoy39], and multiple nuclei models [HU45]. The three models describing the internal social structure of cities were developed in Chicago.

In multiple nuclei models, geographers C. D. Harris and E. L. Ullman (1945) consider a city as a multi-centers system in which the city contains more than one center around which activities revolve. In fact, the centers and their hinterlands

which form this hierarchy have many elements in common in functional terms which are repeated across several spatial scales, and in this sense, districts of different sizes at different levels in the hierarchy have a similar structure. For example, a farm-house is the point of attraction for different fields and lands of an agricultural domain, but several farms together make a homogeneous pattern in a village's territory. At a higher level, a market town attracts population and activities from surrounding villages, and a regional capital is a major center of attraction for several of those elementary farming districts. Moreover, the growth of cities not only occurs through the addition of units of development at the most basic scale, but through increasing specialization of key centers. These mechanisms of urban growth also ensure that the city is stable.
Cities are created from demand for exchange of ideas and material goods, and city size depends upon the level at which the city exists in the entire hierarchy of size from the smallest hamlet to the most global city. But large cities grow from the tiniest seeds, and depend on interaction of population, economic productions, consumption, and the others in the markets. Such structures which repeat themselves at different levels of the hierarchy and which in turn are associated with different scales and sizes are said to be self-similar. This is also a property of a new geometry which is introduced in the next part.

When mentioning to geometry, we usually talk of Euclidean geometry which based on the straingt line and dimension. But there are many natural objects which are so irregular and fragmented that it is difficult to describe them by Euclidean geometry such as a cloud, a mountain, a coastline, or a tree,.... From that, fractal geometry was developed and has become popular. The term "*fractal*" was first used by mathematician Benoît Mandelbrot in 1975, based on the Latin fractus meaning "broken" or "fractured". In essence, a fractal is an object whose irregularity, as a non-smooth form, is repeated across many scales, and in this sense, systems such as cities which manifest discrete self-similarity are ideal candidates for such study.
An important feature of fractal beyond self-similarity is *fractal dimension* which is greater than its topological dimension. A fractal dimension is a ratio providing a statistical index of complexity comparing how detail in a pattern (strictly speaking, a fractal pattern) changes with the scale at which it is measured. In a rough sense, it measures "how many points" lie in a given set. A fractal dimension does not have to be an integer. One non-trivial example is the fractal dimension of a Koch curve (figure 4.1). It has a topological dimension of 1, but its fractal dimension is a number between one and two ($\log(4)/\log(3)$). This problem will be studied more in the last section of this chapter.

In the middle of the 1980s, Michael Batty and Paul Longley firstly applied fractal geometry into studying and modeling the growth of city [BL86]. A few years later, Pierre Frankhauser gave a more general overview of this geometry in urban geography [Fra91]. Fractals are used to define urbanisation rules and to apply these rules at

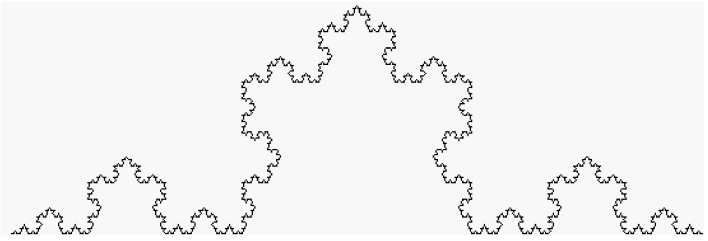


Figure 4.1: Koch curve.

different scales of the urban areas (see in [Fra08, FP10, CVC+11, Bad05b, Bad05a]).

4.1.3 Modeling of Urban Growth

As mentioned above, urban has the properties of fractals structure. Like all natural objects, cities also grow and evolve. So, how fractals can describe dynamic structures which grow and change through time?
In fact, cities evolve through the cumulative addition and deletion of basic units, cells or particles on a determined space (figure 4.2). Such units may be buildings, population, transportation networks .... All of them exist, interact in a urban system, and create its growth.

In the beginning, diffusion limited aggregation (DLA) (figure 4.4) has been con-

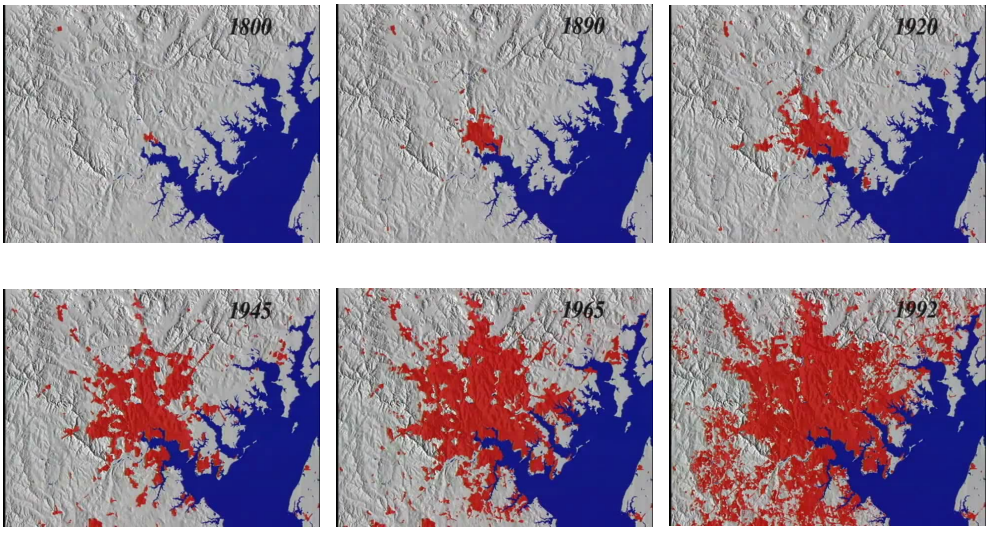


Figure 4.2: The growth of Baltimore from 1800 to 1992 (source: NASA's website [NAS96].

sidered by Mike Batty [BL94] to describe urban growth. The model, proposed by Witten and Sander in 1981, is the process whereby particles undergoing a random walk due to Brownian motion cluster together to form aggregates of such particles. It is a simple model that was able to generate fractal structure whose self-similarity was dendritic or tree-like. From calculating fractal dimension of some urban structures in [BL94], one notes that all the dimensions lie between 1 and 2 and most of these values are greater than 1.5, most lying between 1.6 and 1.8 around dimension value of DLA, 1.7. However, this model does not satisfy some observed properties of urban system. Most of urban structures are actually more compact at center than cluster of DLA. Besides, DLA model does not account for any specific constraints on its development, for example, people can not live on the sea, roads, rivers, ... Moreover, in this model there is only one connected cluster, while a city is actually a complex system of related components.

To overcome some of the limitations of DLA, Batty is proposed using dielectric breakdown model (DBM), a modified DLA, in modeling urban growth of Cardiff (figure 4.3). DBM, developed by Niemeyer, Pietronero, and Weismann in 1984, is combination of DLA and electric field. This model enables us to generate cities of many different shapes and degrees of compactness. In addition, the physical constraints of rivers and sea were considered in the simulations.
Despite the fact that simulations of DBM are more realistic, they remain some limitations. Cluster of the model does not satisfy the complex property of urban system and there is difference in their density distribution. The density of cluster in simulations decreases from center as a power law

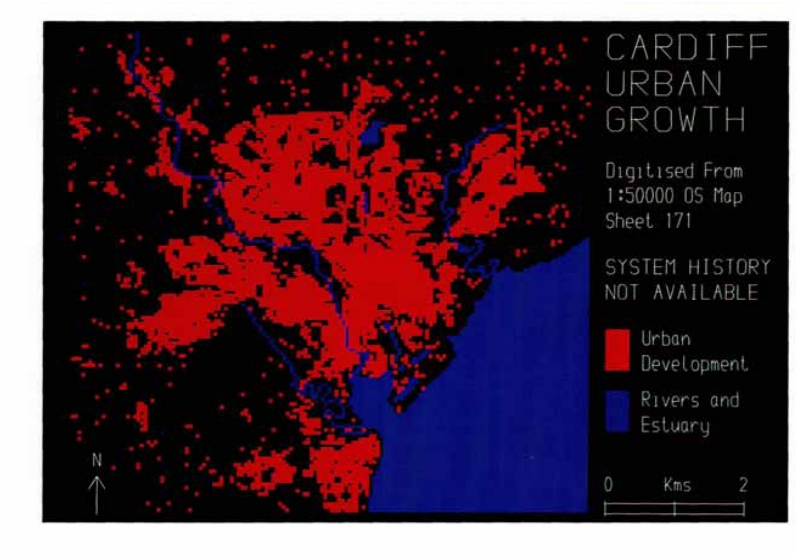
$$\rho(r) \sim r^{D-2}$$

where r is the radial distance from center of cluster, and $1 < D < 2$ is fractal dimension of cluster of simulation. Meanwhile, actual urban data have been more commonly found to fit an exponential decay [Cla51]

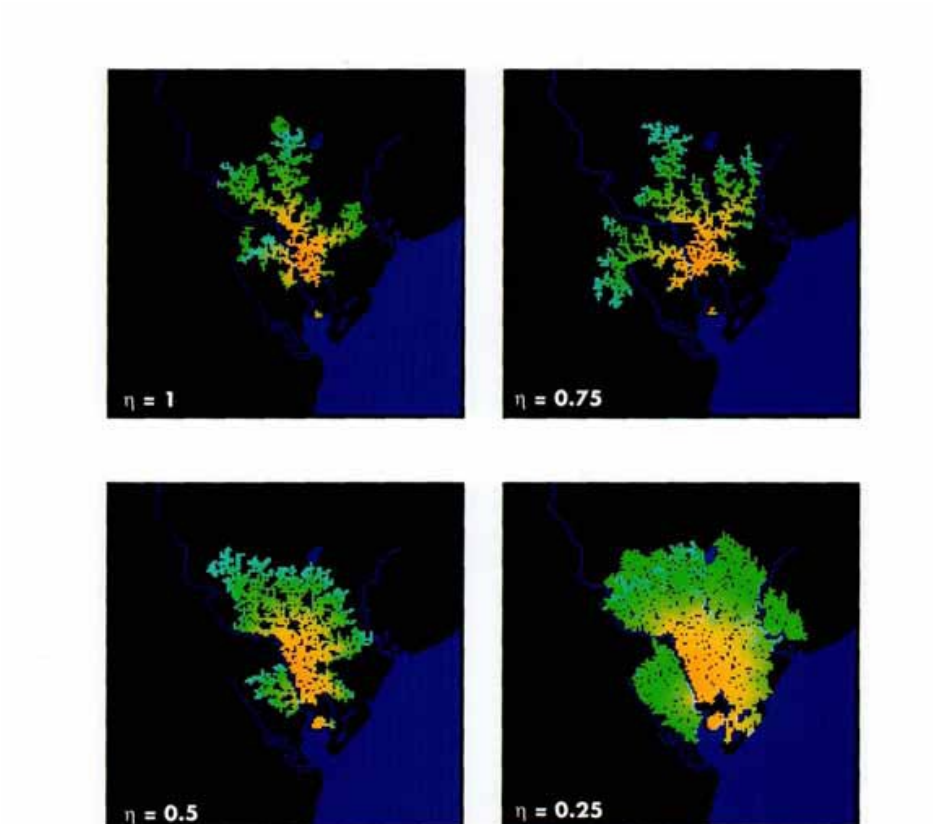
$$\rho(r) \sim e^{-\lambda r}$$

with λ is the *density gradient*.

Makse et al [MAJB+98] proposed using the correlated gradient percolation model in the presence of a gradient in simulation of Berlin city with amazing results (see figure 4.5). The model generates an urban as a complex system and more realistic cities. This will be studied in the next sections.
However, this model is random and does not describe a given city. Additionally, it has not considered physical constraints of urban system yet. In our research, we consider a given city at determined time to study and take it as a starting point for the next simulations. Moreover, we study interaction of population and effect of other factors such as rivers, roads,... on this city.



(a) Cardiff



(b) Simulations

Figure 4.3: The urban area and the simulations of urban growth of Cardiff in [BL94].

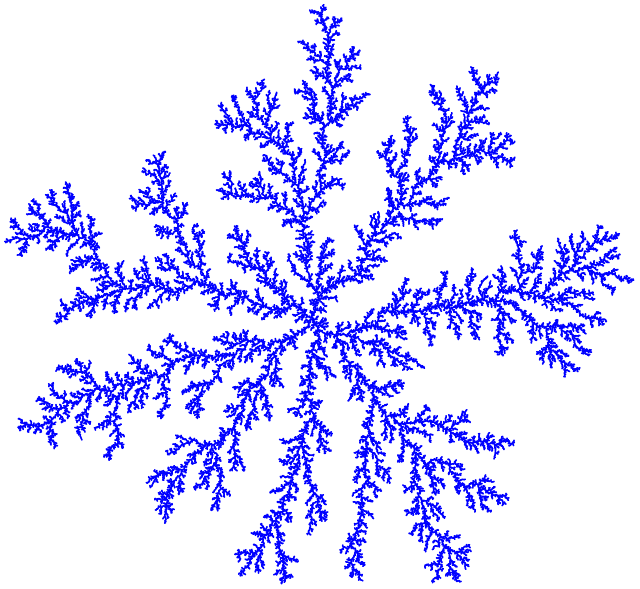


Figure 4.4: An example of DLA with 50000 particles.

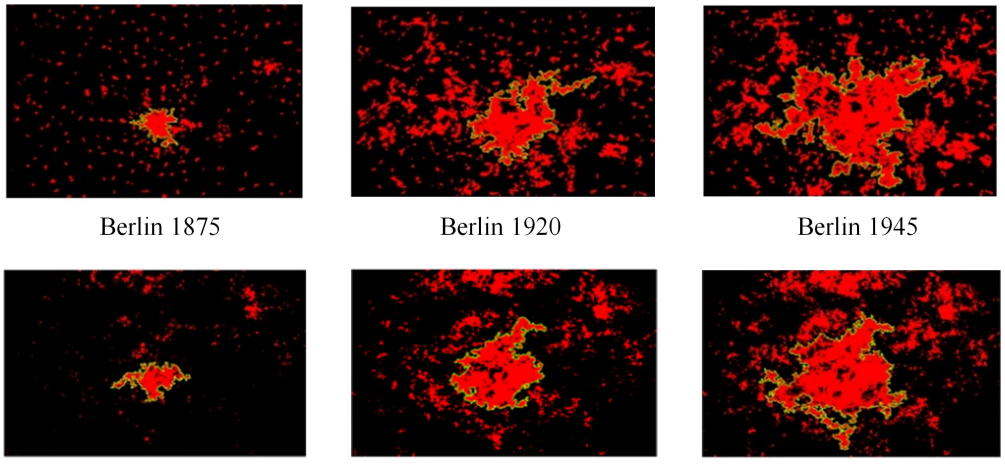


Figure 4.5: The results of simulation of Makse et al for Berlin city.

In the next parts of this chapter, some mathematical basics from percolation to correlation gradient percolation models will be represented. After that, the adapted correlation gradient percolation models are studied and applied in simulating Baltimore (a city of United States of America). In the next chapter we will study more detail urban system of Montargis (a commune in France) and give models to simulate its growth.

4.2 Percolation Models

Percolation is a random model introduced by John Hammerslay in 1957 to describe the flow of a fluid or gas through a porous medium of small channels which may or may not let gas or fluid to pass. During the last five decades, percolation theory has brought new understanding and techniques to a broad range of topics in physics, materials science, complex networks, geography, ...

In mathematics, percolation theory describes the behavior of connected clusters in a random graph. Percolation model can be defined on any lattice in any dimension. In two-dimensions, their model amounts to the following. Let $\mathbb{Z}^2$ be the plane square lattice of size $n \times n$ and p a number satisfying $0 \leq p \leq 1$. We examine each edge of $\mathbb{Z}^2$ in turn, and declare this edge to be *open* with probability p and *closed* otherwise, independently of all other edges. Therefore, for a given p, what is the probability that an open path exists from the top to the bottom? The behavior of this probability for large n is of primary interest. This problem is called now *bond percolation*. In a slightly different mathematical model for obtaining a random graph, a vertex, usually called *site*, is *occupied* with probability p or *empty* (in which case its edges are removed) with probability $1 - p$; the corresponding problem is called *site percolation*. Percolation problem is challenging both in mathematical physics and probability theory.

4.2.1 Site Percolation

As mentioned above, site percolation is a process that uses a probability to decide the state of a site in a lattice. For each site $z = (i, j)$ in a square lattice size $N \times M$, we create a random number $u(z)$, called *occupancy variable*, having uniform probability distribution between $[0, 1]$. This site is occupied if $u(z)$ less than a fixed *occupation probability*, p. If p is small, only isolated clusters exist and on the opposite, if p is near 1, most of the lattice sites are occupied, and the occupied sites will form an infinite cluster. Such a cluster is said to be a *spanning cluster*. Because there is no spanning cluster for small p and there is one for p near 1, there must be an intermediate value of p at which it first appears. If we consider the model in an infinite lattice, there exists a well defined *critical point* p_c such that:

- if $p < p_c$ there is no infinite cluster a.s., this is called *sub-critical* regime;

- $p = p_c$, *critical* regime, there is no infinite cluster a.s. with $d = 2$ or $d \geq 19$. In others cases, we still don't know;
- and $p > p_c$ there is a unique infinite cluster *super-critical* regime a.s. However, this finite clusters are generally small.

Harry Kesten proved that in 2-dimensions, on triangular lattice, $p_c = 1/2$ [Kes82].

Site percolation is a model having many applications. However, because the occupation probability p is a constant the density on the lattice is stable, this cannot be adapted to model the growth of cities whose population density $\rho(z)$ decrease from center as an exponential law [Cla51].

4.2.2 Gradient Percolation

Gradient Percolation is a model of inhomogeneous percolation introduced by B. Sapoval, M. Rosso, and J.-F. Gouyet [SRG85] in 1985. In that, the occupation probability p of each site z is not a constant value, but a function decreasing from 1 to 0 by the distance from z to the original, or to a side of lattice. In this model, there is always a cluster of the connected occupied sites in the region that $p(z)$ is greater than critical point p_c and there is also a cluster of the empty sites with $p(z)$ is less than p_c.

Consider Euclidean distance between $z = (i, j)$ and $z_0 = (i_0, j_0)$:

$$d(z, z_0) = |z - z_0| = \sqrt{(i - i_0)^2 + (j - j_0)^2}. \tag{4.2.1}$$

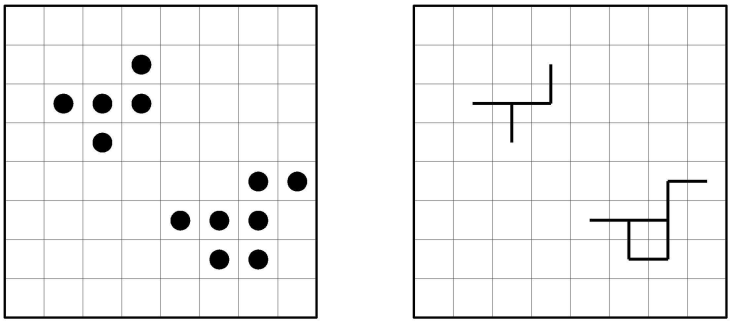
We provide some examples of this model in figure 4.8. It is clear that the density of cluster depends on the function of occupation probability. In figure 4.8a the compact core is too large and in figure 4.8b it is too small and indiscernible, they are not suitable for cities. In the meantime, the compact core in figure 4.8c looks like a city. This suits Clark's urban density which is an exponential law with respect to distance to the center.

According to the studies of some cities, C. Clark [Cla51] proposed that the population density $\rho(z)$ of the real urban systems satisfies the relation

$$\rho(z) = \rho_0 e^{-\lambda r_z}, \tag{4.2.2}$$

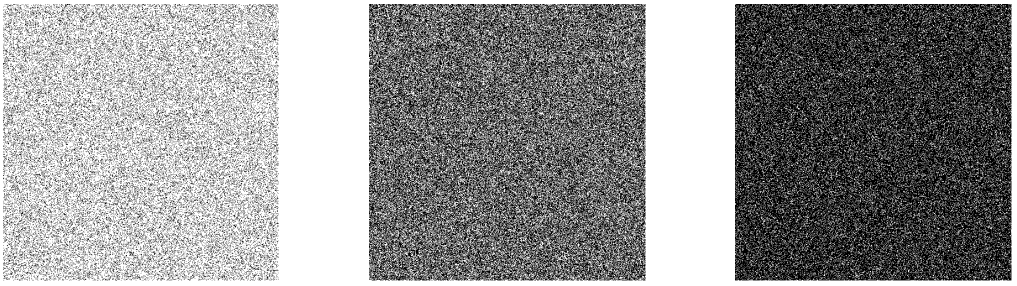
with $r_z = d(z, z_0)$ is the distance from z to center of lattice, λ is the *density gradient* in interval $[0, 1]$. The density gradient quantifies the extent of the urban spread around the central core, or the size of the largest cluster. If we consider this model to describe a urban system then the sites on lattice will be fundamental units of the city such as land use, road, river, ... The probability that a unit occupies a given position decreases as distance from center of city.
Therefore, we can assume the exponential function of probability as

$$p(z) = \frac{\rho(z)}{\rho_0} = e^{-\lambda r_z}. \tag{4.2.3}$$



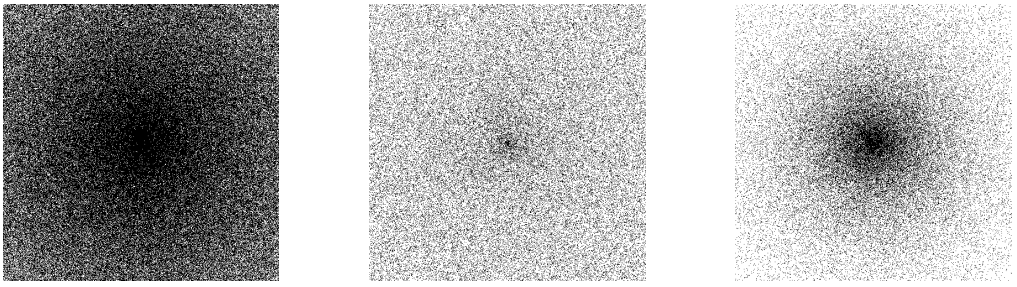
(a) Site-percolation (b) Bond-percolation

Figure 4.6: Two main types of percolation



(a) $p = 0.2$ (b) $p = 0.59$ (c) $p = 0.8$

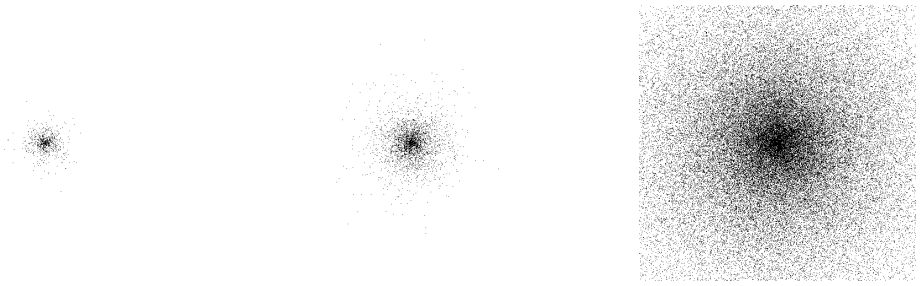
Figure 4.7: Site-Percolation with different probability p on square lattice of size 501×501.



(a) $p(r_z) = 1 - \frac{r_z}{\sqrt{2}N}$ (b) $p(r_z) = r_z^{-0.3}$ (c) $p(r_z) = 1.01^{-r_z}$

Figure 4.8: Gradient percolation with different probability functions of the distance r_z from site z to center of square lattice of size 501×501.

The evolution of the largest cluster with different λ (figure 4.9) looks like the growth of a city. With small λ ($\lambda \approx 0$), the central cluster is large (correspond to city developed), and vice versa, the larger the value of λ ($\lambda \approx 1$) is, the smaller central cluster is.



(a) $\lambda = 0.1$ (b) $\lambda = 0.05$ (c) $\lambda = 0.01$

Figure 4.9: Gradient percolation with different exponential functions of probability $p(r) = e^{-\lambda r}$ and λ on square lattice of size 501×501.

4.2.3 Correlated Gradient Percolation

Observations of reality suggest that the development units are not positioned randomly. The units seem to be occupied with probability depending on the presence of their neighbors. It means that there exists a correlation between the units. Near an occupied unit, the probability of additional development is higher and decreases as we move away from it. We argue that these rules of placement affect forming and growth of urban areas. Starting from this assumption, we introduce *correlation gradient percolation* model.

This model is a gradient percolation modified to introduce correlations among occupancy variables by using a method proposed in [MHSS96], a modification of the Fourier filtering method (Ffm) suitable for large system. After that, we will obtain correlated random numbers depending on the distance between two sites. In this part we will present the mathematical basics of Ffm and how to apply it into our gradient percolation model.

4.2.3.1 Fourier Filtering Method

The purpose of Ffm in our model is generating a sequence of correlated random numbers from given uncorrelated occupancy variables.
We start by defining the Ffm in dimension 1. Firstly, consider a stationary sequence of N uncorrelated random numbers $\{u_i\}_{1,\dots,N}$ with Gaussian distribution.

The brackets $\langle . \rangle$ will denote an average with respect to a Gaussian distribution. The *correlation function* $K(\ell)$ of the sequence $\{u_i\}$ is defined as following:

$$K(\ell) = \langle u(i)u(i+\ell)\rangle\,, \tag{4.2.4}$$

where $u(i) := u_i$, $i = 1, ..., N$. If this function is independent of i, we can put $i' = i - \ell$ and get that

$$\begin{aligned} K(\ell) &= \langle u(i)u(i+\ell)\rangle = \langle u(i')u(i'+\ell)\rangle \\ &= \langle u(i-\ell)u(i)\rangle = \langle u(i)u(i-\ell)\rangle\,, \end{aligned}$$

or

$$K(\ell) = K(-\ell). \tag{4.2.5}$$

We define the asymptotic equivalence ($\sim$) as following

$$f(n) \sim g(n) : \lim_{n\to\infty} \frac{f(n)}{g(n)} = 1. \tag{4.2.6}$$

If the sequence $\{u_i\}$ is independent, we have

$$K(\ell) = \left\langle u^2(0)\right\rangle \delta_{\ell,0},$$

with $\delta_{\ell,0}$ is *Kronecker delta*,

$$\delta_{\ell,0} = \begin{cases} 1 & \text{if } \ell = 0 \\ 0 & \text{if } \ell \neq 0. \end{cases}$$

We can take the uncorrelated random numbers $\{u_i\}$ with $\mathbb{E}(u) = 0$ and $\mathbb{E}(u^2) = 1$. From these numbers, we want to generate a new correlated sequence $\eta(i)$ with a long-range power-law correlation function $C(\ell)$

$$C(\ell) = \langle \eta(i)\eta(i+\ell)\rangle \sim \ell^{-\alpha} \quad (\ell \to \infty)\,. \tag{4.2.7}$$

Here, α is the correlation exponent, and the long-range correlations are relevant for $0 < \alpha < d$ (in this case $d = 1$).
The spectral density $S(q)$ defined as the Fourier transform of $C(\ell)$ in the equation (4.2.7) [Rei65]

$$S(q) = \hat{C}(\ell) = \langle \hat{\eta}(q)\hat{\eta}(-q)\rangle \tag{4.2.8}$$

where $\{\hat{\eta}(q)\}$ corresponds to the Fourier transform coefficients of $\{\eta(i)\}$.
From this equation we can get

$$\hat{\eta}(q) = \sqrt{S(q)}\hat{u}(q), \tag{4.2.9}$$

where $\hat{u}(q)$ is Fourier transform coefficients of $u(i)$.
Finally, we get the correlated sequence $\eta(i)$ in the equation (4.2.7) by getting inverse

Fourier transform.

Of course, the function $C(\ell) = \ell^{-\alpha}$ has a singularity at $\ell = 0$. To deal with this, we replace (4.2.7) with a slightly modified correlation function [MHSS96] that has the desired power-law behavior for large ℓ , and is well-defined at the origin,

$$C(\ell) \equiv (1 + \ell^2)^{-\alpha/2}. \tag{4.2.10}$$

This function is defined in the interval $[-N/2, ..., N/2]$ with periodic boundary conditions, $C(\ell) = C(\ell + N)$. It will be used in our model later.

The actual numerical algorithm for Ffm consists of the following steps:

i. Generate a one-dimensional sequence $\{u(i)\}$ of uncorrelated random numbers with Gaussian distribution.

ii. Obtain $\{\hat{\eta}(q)\}$ by using (4.2.10), (4.2.8), and (4.2.9).

iii. Calculate the inverse Fourier transform of $\hat{\eta}(q)$ to obtain $\eta(i)$, the sequence in real space with the desired power-law correlation function which asymptotically behaves as (4.2.7).

This algorithm can be applied in two-dimensional case by replacing $\{u(i)\}$ with a 2-D sequence $\{u(z)\}$, $z = (i, j)$, $i, j = 1, ..., N$, and correlation function becomes

$$C(\ell) = \langle \eta(z)\eta(z') \rangle \equiv (1 + \ell^2)^{-\alpha/2}, \tag{4.2.11}$$

where, $\ell = d(z, z')$ as (4.2.1) and the correlation exponent $0 < \alpha < 2$.

4.2.3.2 Correlation Gradient Percolation (CGP)

Let us embed these modifications in our model, namely, the use of gradient percolation instead of uniform and the use of correlated site probabilities. We see that if the uncorrelated variables $u(z)$ have Gaussian distribution, then so do $\eta(z)$ after Ffm. How do we use this correlated sequence in percolation model?

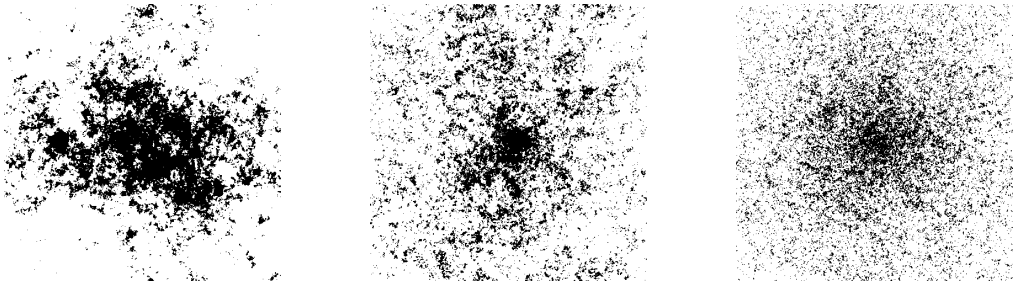
With $z \in [1, M] \times [1, N]$, we consider the correlated sequence η and its *distribution function*

$$F(\eta(z)) = \mathbb{P}\left(\eta \leq \eta(z)\right) = \int_{-\infty}^{\eta(z)} f(t)dt, \tag{4.2.12}$$

where $f(t)$ is probability density of η.
Since η has Gaussian distribution, $F(\eta)$ is strictly increasing and continuous. Put $Y = F(\eta)$, then Y is a random variable. Take $y \in [0, 1]$, we see that

$$\mathbb{P}\left(Y \leq y\right) = \mathbb{P}\left(F(\eta) \leq y\right) = \mathbb{P}\left(\eta \leq F^{-1}(y)\right) = F\left(F^{-1}(y)\right) = y.$$

This implies that Y has uniform distribution on $[0, 1]$.



(a) $\lambda = 0.01, \alpha = 0.1$ (b) $\lambda = 0.01, \alpha = 0.7$ (c) $\lambda = 0.01, \alpha = 1.7$

Figure 4.10: Correlation Gradient Percolation with the same $\lambda = 0.01$ and different α on square lattice of size 1001×1001.

Therefore, we can compare the distribution function of random variables and occupation probability p in the percolation model. Consider the occupation probability function of site z on a square lattice size $M \times N$

$$p(z) = e^{-\lambda r_z} \tag{4.2.13}$$

where λ is is the density gradient and $r_z = d(z, z_0)$ is the distance from z to center z_0 of lattice (defined in (4.2.1)).

On the square lattice L of size $M \times N$, a site z is black and marked 1 (occupied) or white with marked 0 (vacant) with probability in equation(4.2.13) as following:

$$L(z) = \begin{cases} 1 & \text{if } F\left(\eta(z)\right) < p(z) \\ 0 & \text{otherwise.} \end{cases} \tag{4.2.14}$$

Finally, using CGP model we get the simulations (figure 4.10) with the same density gradient λ and different correlation exponents α. While λ controls the size of cluster, α (in the equation (4.2.11)) deals with the concentration, the smaller the value of α is, the higher concentration (or the correlation) of cluster is.

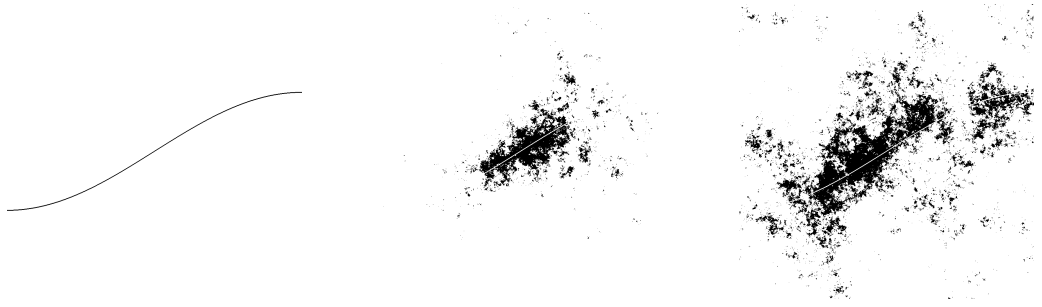
The algorithm of CGP model consists of the following steps:

i. Generate a correlated sequence $\{\eta(z)\}$ from a uncorrelated random sequence $\{u(z)\}$ with Gaussian distribution by using modified Fourier filtering method.

ii. Find $F(\eta(z))$.

iii. Calculate the occupation probability function in (4.2.13) and simulate $L(z)$ by (4.2.14).

Besides the center, we consider CGP model with attending of the others factors (as river). For this modification, the distance r_z in the equation (4.2.13) will be added the distance to river. It becomes the function of two distances $r_z = f(r_z^c, r_z^r)$



(a) The river (b) CGP with $r_z = r_z^c + r_z^r$ (c) CGP with $r_z = r_z^c/5 + r_z^r$

Figure 4.11: Simulation of Correlation Gradient Percolation with the different functions of distance, $p(z) = e^{-0.009 r_z}$ and $\alpha = 0.05$ on the square lattice of size 1000×1000.

(r_z^c is the distance to center and r_z^r is that to river). It can be seen in the figure 4.11 that the distance function affects correlation and density of the cluster in the simulations. In the figure 4.11c, the cluster is concentrated along the river and bigger than cluster in 4.11b.

4.2.4 Properties of Correlation Gradient Percolation

We will study some properties of CGP model. Consider site-percolation, we will review some definitions.

4.2.4.1 The largest cluster

First of all, *clusters* are typically defined as irreducible topological spaces of occupied sites, pathways-connected by occupied bonds. A set of vacant sites connected outside of a cluster is called *dual cluster*. And the path between cluster and its dual is its frontier (or boundary) defined as following:

Definition 4.2.1. The *frontier* (or the external perimeter) F of a cluster C is a set of occupied sites in C such that these sites have at least one neighbors in dual cluster C^{-1}

$$F(C) = \left\{ z \in C : \exists z' \in C^{-1}, z' \text{ is neighbor of } z \right\} \tag{4.2.15}$$

There are different dual clusters which depend on how vacant sites connect and kind of lattice. For example, on square lattice, if the vacant sites only connect by four edge of square we have a dual cluster, meanwhile if these sites connect by four edge and four vertices we will obtain dual cluster with *fjords* as figure 4.12. In this figure, the black points are connected in a cluster, the blue squares are fjords and the red squares are lakes.

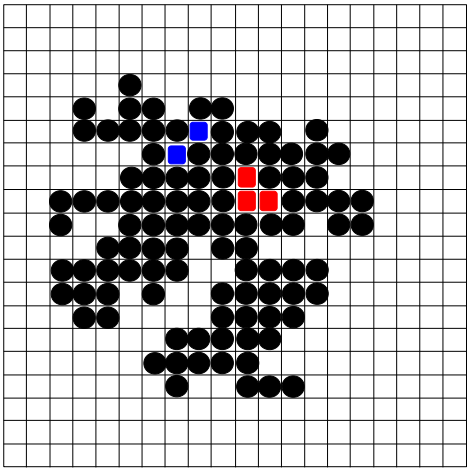


Figure 4.12: The cluster is the set of black circles, the dual cluster is the set of white squares and dual cluster with fjords is white and blue squares.

We have two kinds of frontier, without and with fjords, corresponding to two dual cluster. From figure 4.13, an example of these frontiers, it is easy to see that the frontier with fjords is more complex than the one without. We will study their dimension in the next part.

Definition 4.2.2. The *length* L of a frontier of a cluster C is the number of sites on this one.

$$L(C) = \#F(C) \tag{4.2.16}$$

Consider the largest cluster generated by CGP model, we can see that it stays around the center of lattice with compactness close to center (at position that $p > p_c$) and the isolated clusters in the further positions (where $p < p_c$). It is thus very interesting to study the frontier of the largest cluster.

4.2.4.2 Width of the frontier

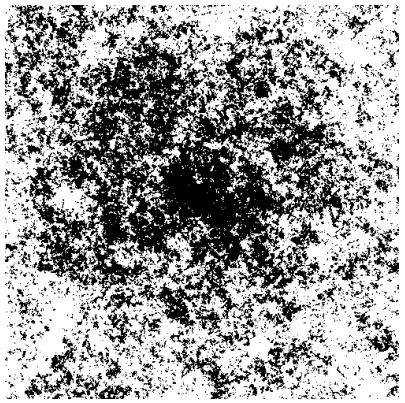
Firstly, we consider the width of the frontier denoted by σ_f and defined as following:

$$\sigma_f = \left\langle (r - r_f)^2 \right\rangle^{1/2} \tag{4.2.17}$$

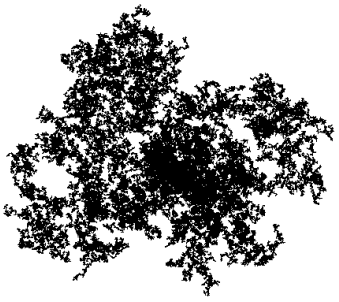
where $r_f = \langle r \rangle$, and r is the distance from a site on the external perimeter of the largest cluster to the center point.
Following [SRG85] and the results in chapter 3, this width is a function of concentration gradient λ

$$\sigma_f \sim \lambda^{-\nu/(1+\nu)} \tag{4.2.18}$$

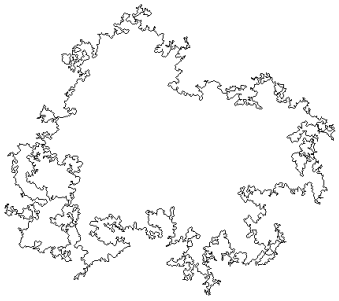
with ν is the connectedness length exponent and its value $\nu = \frac{4}{3}$ corresponds to the uncorrelated percolation problem.



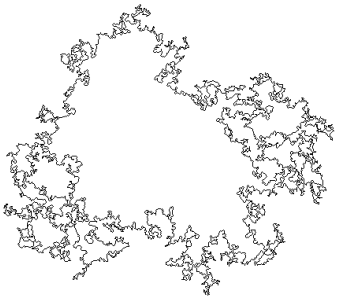
(a) Simulation



(b) The biggest cluster



(c) The frontier without fjords



(d) The frontier with fjords

Figure 4.13: An example for frontier of cluster in the cases with and without fjords on square lattice of size 501×501 simulated by CGP model with correlation exponent $\alpha = 0.7$ and density gradient $\lambda = 0.0052$.

Using definition (4.2.17), the figure 4.14 shows that the width σ_f is a function of λ. It behaves as $\lambda^{-4/7}$ in the uncorrelated case (corresponding to $\nu = \frac{4}{3}$ case), but in strong correlated case the function decreases more rapidly. This indicates that the connectedness length exponent ν depends on the correlation exponent α.
As mentioned above the correlation exponent α influences the concentration and the shape of the largest cluster, it thus causes the change of the frontier as well as its width. By using the equation (4.2.18) on the simulations of CGP we find the changes of ν (see figure 4.15). In particular, $\nu(\alpha)$ is $\frac{4}{3}$ with $\alpha = 2$ (the uncorrelated case) and increases when the system has strong correlations ($\alpha \to 0$).

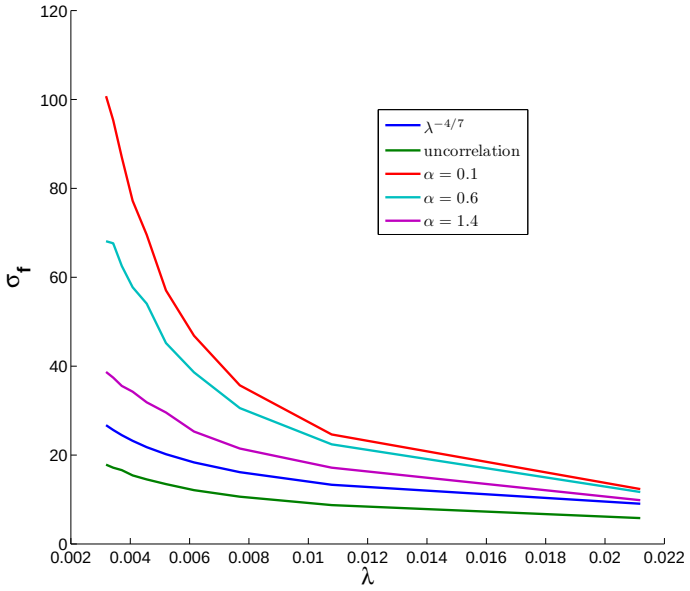


Figure 4.14: The width of the frontier of the largest cluster for a function of density gradient λ, decays as $\lambda^{-4/7}$ in uncorrelated case but exponent varies in the correlated case.

4.2.4.3 Fractal dimension

Percolation clusters become self-similar precisely at the critical exponent p_c for sufficiently large length scales. With this structure, the fractal dimension is interesting to study.

As mentioned above, fractal dimension is a quantitative measure of the structure of fractals. It quantifies the complexity of fractal patterns or sets as a ratio of the change in detail to the change in scale. It may be non-integer valued. Felix Hausdorff (1868-1942) and Abram Besicovitch(1891-1970) demonstrated many curves having dimension between 1 and 2 related to the varying amounts of information they contain. Before defining Hausdorff-dimension we will introduce some related definitions.

Consider a metric space X with distance function d. The *diameter* of a set $E \subset X$ is defined

$$|E| := \sup \{d(x, y) : x, y \in E\} .$$

Let $E \subset X$, a subset of a metric space, for every $\delta > 0$, a countable collection subsets $\{F_i\}$ of E is said to be a δ-cover of E if

$$E \subseteq \bigcup_i F_i \quad \text{and} \quad 0 < |F_i| < \delta, \quad \forall i.$$

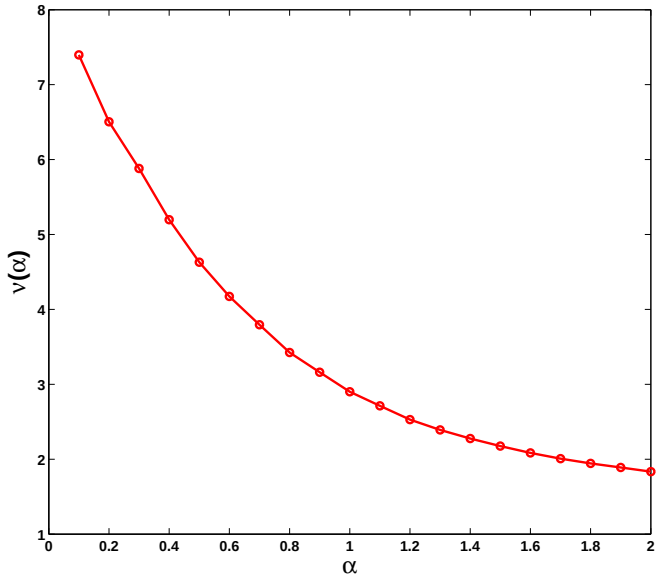


Figure 4.15: Connected length exponent $\nu(\alpha)$ as a function of the correlation exponent α calculated from (4.2.17) and (4.2.18).

And for all $d > 0$, we define the function d-dimensional Hausdorff measure of E

$$H_\delta^d(E) = \inf\left\{\sum_i |F_i|^d : \{F_i\} \text{ is a } \delta\text{-cover of } E\right\}.$$

The Hausdorff dimension may then be defined as

Definition 4.2.3. Let E be a subset of a metric space X, the Hausdorff dimension (or Hausdorff-Besicovitch dimension) is

$$\dim_H(E) = \inf\left\{d \geq 0 : \lim_{\delta\to 0} H_\delta^d(E) = 0\right\}. \tag{4.2.19}$$

A different notion of dimension can be provided from a generalization of the concept of topological dimension applied to Euclidean objects. In Euclidean geometry, let Θ be a bounded subset of $\mathbb{R}^n$ and $N_\Theta(r)$ be the minimum number of balls of radius r required to cover Θ, then the simplified version of Hausdorff dimension can be defined as following

$$\dim_{box}(\Theta) = \limsup_{r\to 0} \frac{\log N_\Theta(r)}{\log(1/r)}. \tag{4.2.20}$$

This dimension is called upper Minkowski-Bouligand dimension (Schroeder 1991). In fractal geometry, the Minkowski-Bouligand dimension, also known as Minkowski dimension or box-counting dimension, is a way of determining the fractal dimension

of a set S in a Euclidean space $\mathbb{R}^n$, or more generally in a metric space (X, d).

The upper Minkowski dimension is always greater or equal to the Hausdorff dimension, but they need not be the same.

We note that the largest cluster in Correlation Gradient Percolation model has fractal structure only near its frontier (external perimeter). At distance close to the center, where $p(z) > p_c$, the cluster becomes compact, it is thus nonfractal. And for further distance $p(z) < p_c$, only small isolated clusters exist so that they are not fractal either. For more details we refer to [MAJB$^+$98].
We will study two kinds of the fractal dimension: dimension of cluster near its perimeter on length scales smaller than the width of the frontier σ_f and dimension of frontier. The first, fractal dimension of cluster D is estimated by the "mass-radius" relation

$$M(R) \sim R^D,$$

with $M(R)$ is the number of black (occupied) sites of cluster inside the region of radius R. We found that $D = 1.89$ corresponding to the fractal dimension of uncorrelated percolation cluster and this value is independent of the correlation.

The second, the *dimension of frontier* d_f is calculated by the box-counting method as in Minkowski dimension. Let $L(\epsilon)$ be the number of small squares of size ϵ that cover all of the frontier of cluster, this dimension is calculated by the following relation

$$L(\epsilon) \sim \epsilon^{-d_f}. \tag{4.2.21}$$

Applying this method we calculate the dimension of frontier with and without fjords of cluster with occupation probability function $p(z) = e^{-\lambda r_z}$ on a square lattice. From the figure 4.16 we can see that there is a small variation of the fractal dimension of the external perimeter of the largest cluster with the value between 1.3 and 1.4 in the case without fjords, it is around the value $d_f = \frac{4}{3}$ of cluster at critical point in standard percolation. If we take into account fjords, these values are between 1.38 and 1.54 (see the figure 4.17) and increase in the value of α but their difference is not too much. Hence, this does not rule out the fact that it may be independent of the parameters of correlation α and gradient λ.

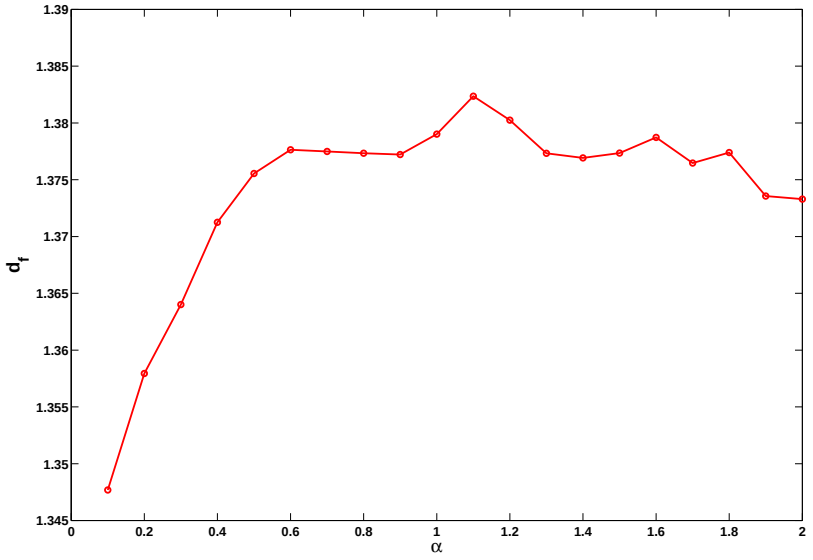


Figure 4.16: Fractal dimension of frontier without fjords of the largest cluster by α with $\lambda = 0.004$ on square lattice of size 1000×1000.

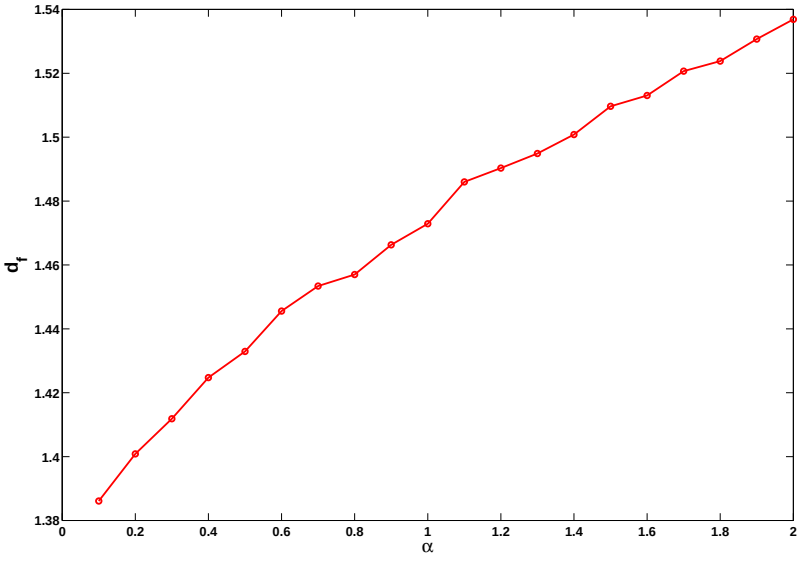


Figure 4.17: Fractal dimension of frontier with fjords of the largest cluster by α with $\lambda = 0.004$ on square lattice of size 1000×1000..

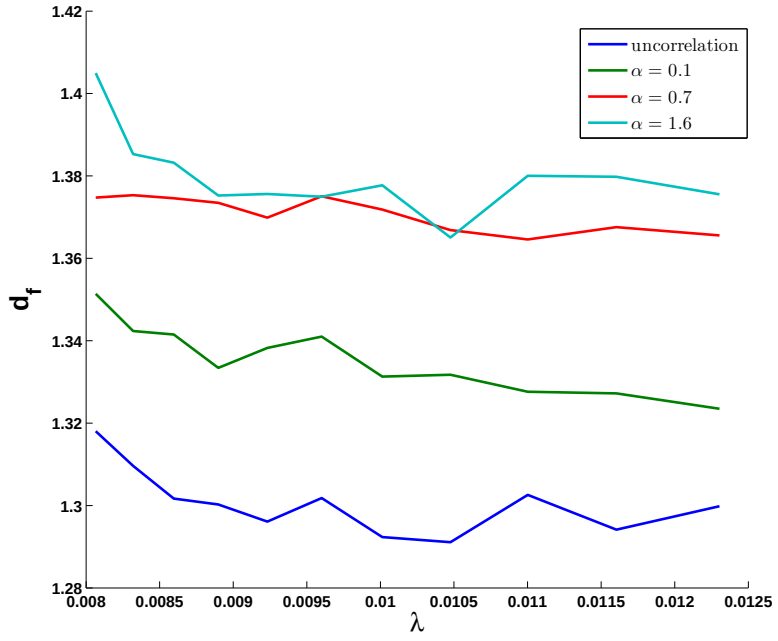


Figure 4.18: Fractal dimension of frontier without fjords of the largest cluster by λ with different α on square lattice of size 1000×1000.

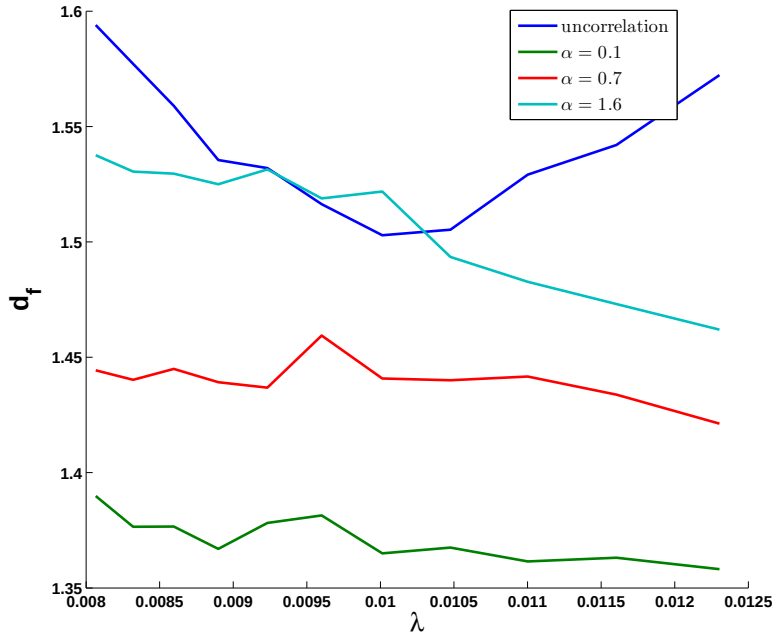


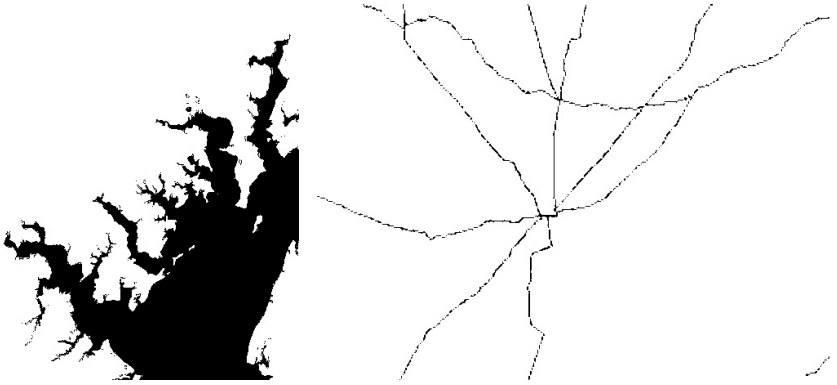
Figure 4.19: Fractal dimension of frontier with fjords of the largest cluster by λ with different α on square lattice of size 1000×1000..

4.3 Simulation of the Growth of Baltimore

In this section, we adapt CGP model in modeling the growth of Baltimore city. The input data which are got from a film on NASA's website [NAS96] contain main geographical location, population growth and the main road system of Baltimore in 200 years from 1792 to 1992.

4.3.1 Urban system of Baltimore

Baltimore is the largest city in the U.S. state of Maryland and the 26^{th} largest city in the country. It located on the Mid-Atlantic coast and was built at the mouth of the Patapsco River (figure 4.20a), which empties directly into the Chesapeake Bay. Data of Baltimore gotten from NASA's website contain urban growth in the region around Baltimore from 1792 to 1992 and major roads. Figure 4.20b is major roads in 1792, these roads have some changes during growth process.



(a) The mouth of the Patapsco River near Baltimore

(b) Road system 1792

Figure 4.20: The input data of simulations of Baltimore got from NASA's website [NAS96].

In reality, in the first years, urban system of Baltimore grows very slow. However, from 1900, the growth is stronger and after the 1930s the growth explodes (see in figure 4.21).

The density gradient λ, in figure 4.23, is decreasing by time. As mentioned above in CGP model, the change of this value affects evolution of black points in simulations. From this observation, we choose λ in the model is a function depending on step of simulation, $\lambda^i = \frac{c}{\sqrt{i}}$, with c is a given constant.

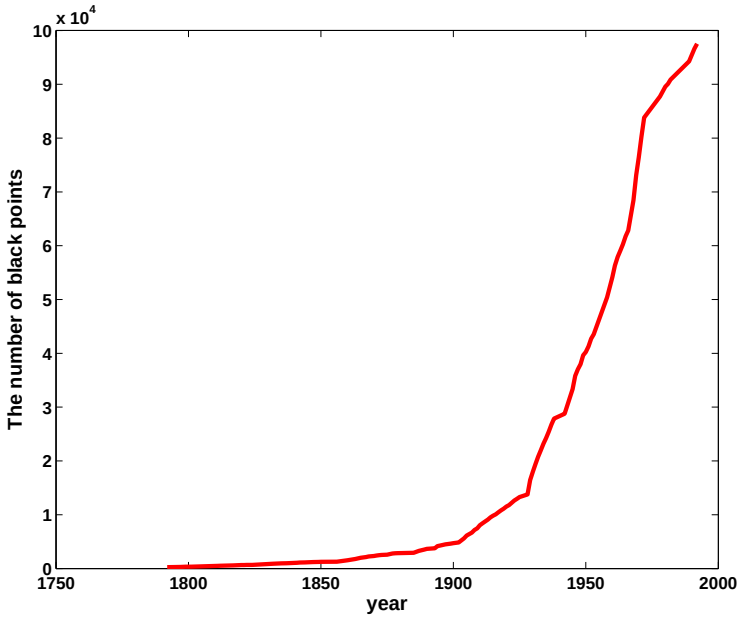


Figure 4.21: Evolution of Baltimore from 1792 to 1992.

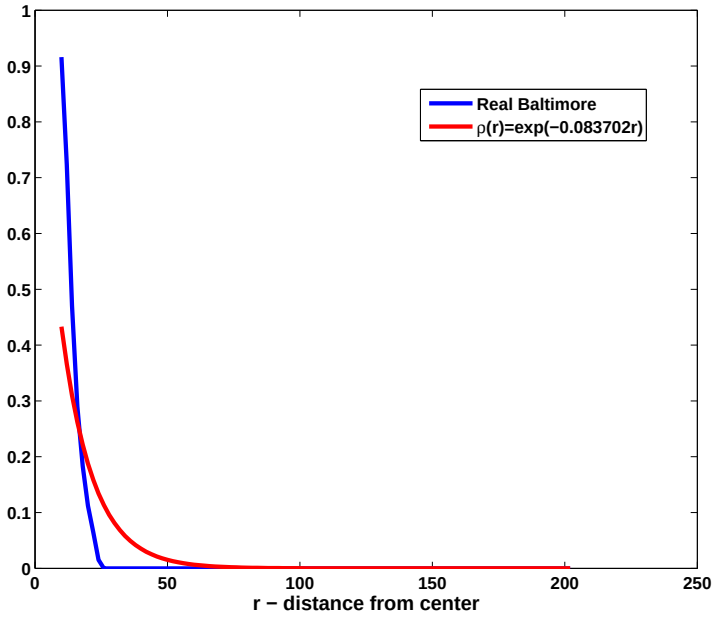


Figure 4.22: Density of Baltimore in 1822 at a distance to center r corresponds to function $e^{-0.08r}$.

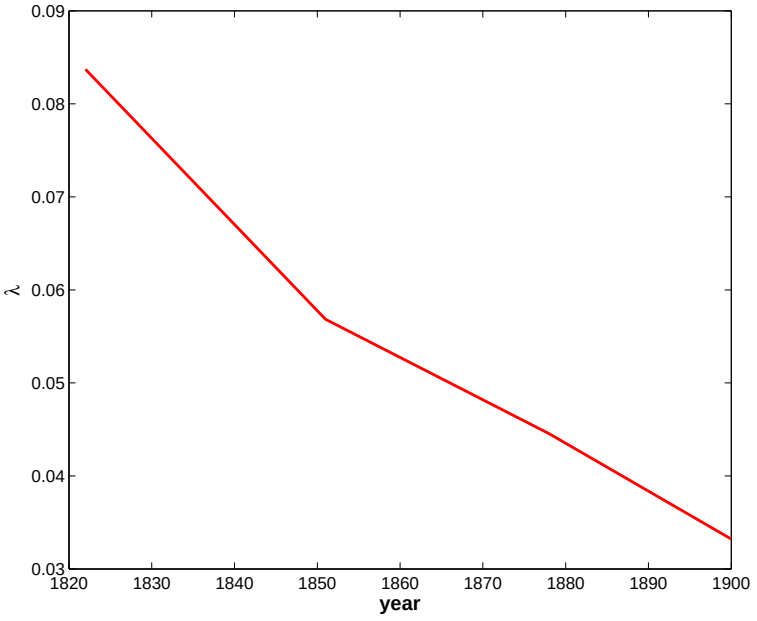


Figure 4.23: Density gradient λ computed from real Baltimore (in 1822, 1851, 1878, 1900).

4.3.2 Model

Consider CGP with a strong correlation exponent $\alpha = 0.001$ and occupancy probability function as follows:

$$p_i(z) = \frac{1}{C} e^{-\lambda_1^i (d(z,z_0)+d_{\min}(z,road))} \sum_{z_1 \in B} e^{-\lambda_2^i d(z,z_1)} \tag{4.3.1}$$

where:

- C is a constant,
- $d(x, y)$ is distance between x and y,
- $d_{\min}(z, road)$ is minimum distance from z to major roads (in figure 4.20b).
- λ_1^i, λ_2^i are the parameters,
- B is the set of black points and the main road system,
- z_0 is the center that has the biggest density.

From observation of Baltimore density in reality (see figure 4.22) we choose density gradient around 0.08 as follows:

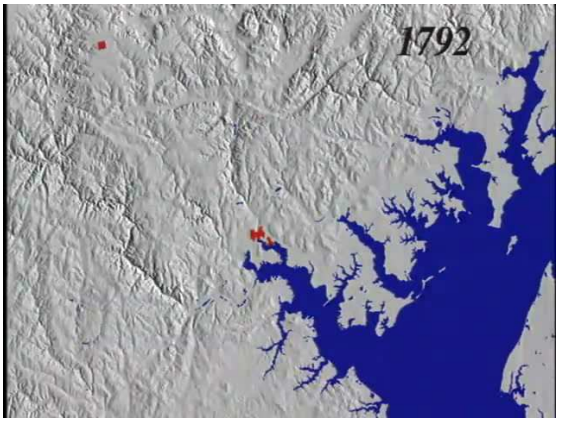
$$p_i(z) = \frac{1}{C} e^{-\frac{0.08}{\sqrt{i}} (d(z,z_0)+d_{\min}(z,road))} \sum_{z_1 \in B} e^{-\frac{0.06}{\sqrt{i}} d(z,z_1)}. \tag{4.3.2}$$

In order to adapt the model in simulating the growth of Baltimore, we assume that an occupied point won't change its status (becomes vacant) in future. Therefore, the CGP model is only used in vacant points. This is the way that simulations done.

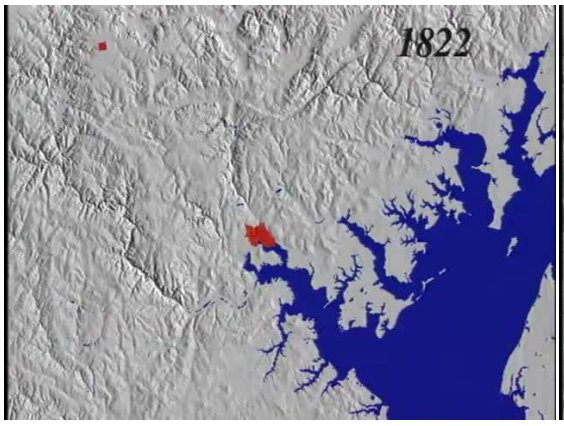
4.3.3 Results

Baltimore 1792 (figure 4.24a) is taken as a starting point with the presentation of the road system (figure 4.20b) in the simulations.

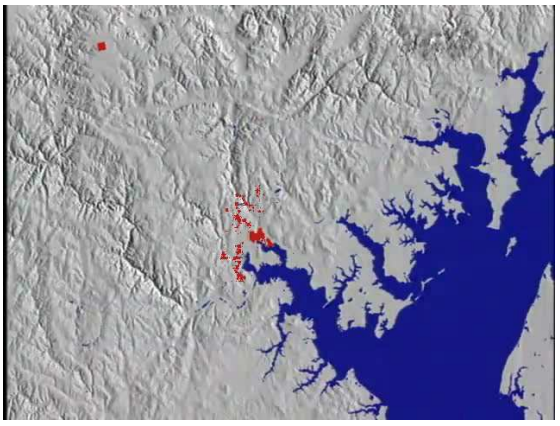
Figure 4.28 shows that evolution of simulations is not exactly similar with evolution of Baltimore in reality (figure 4.21). The slow growth of the first period (after 1900) and the exploded growth after the 1930s are not represented in simulations. In fact, there are other factors which affect the growth of urban system such as: the development of science and technology, epidemic diseases, war ... which we do not control.



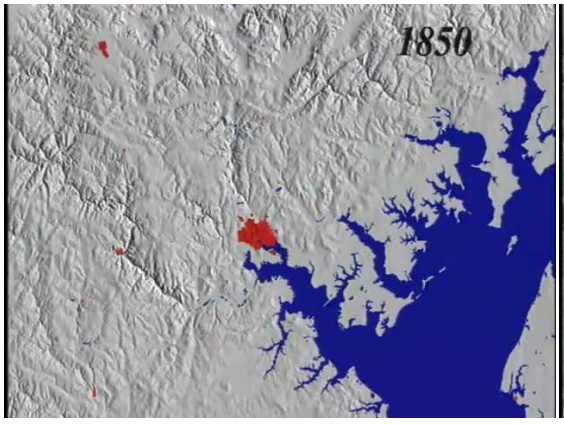
(a) Baltimore 1792



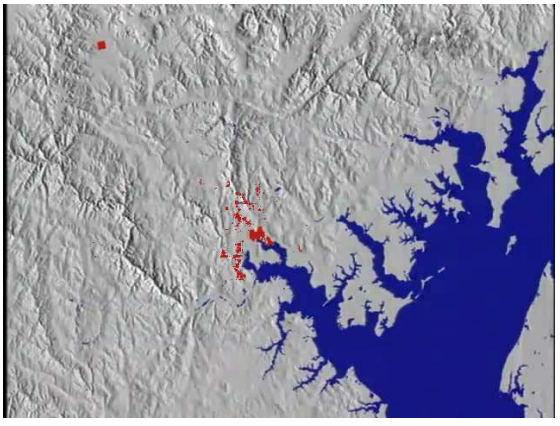
(b) Baltimore 1822



(c) Simulation 12

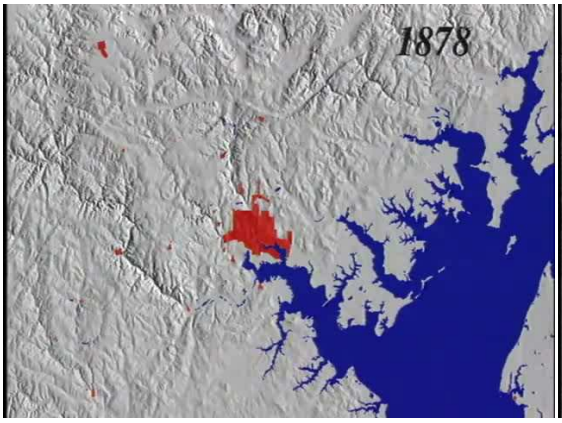


(d) Baltimore 1850

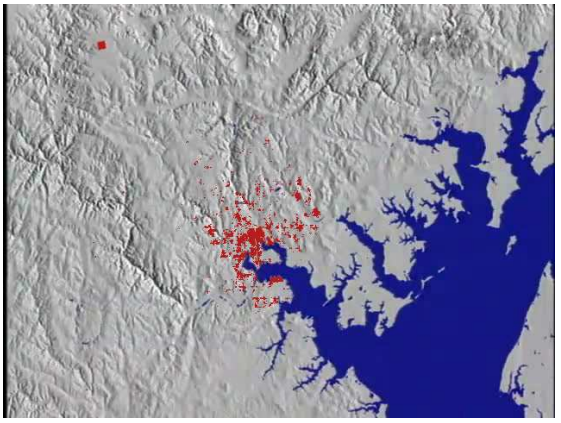


(e) Simulation 15

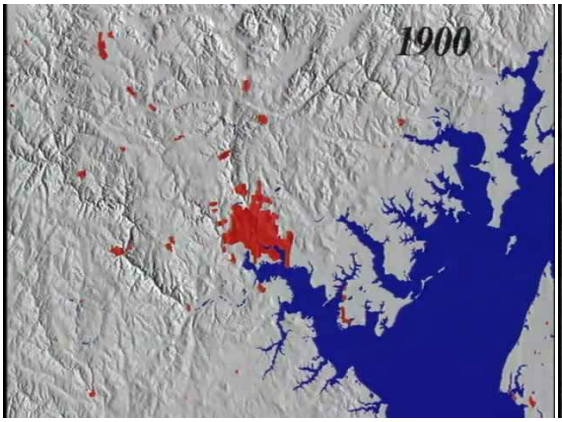
Figure 4.24: Real Baltimore and its simulation.



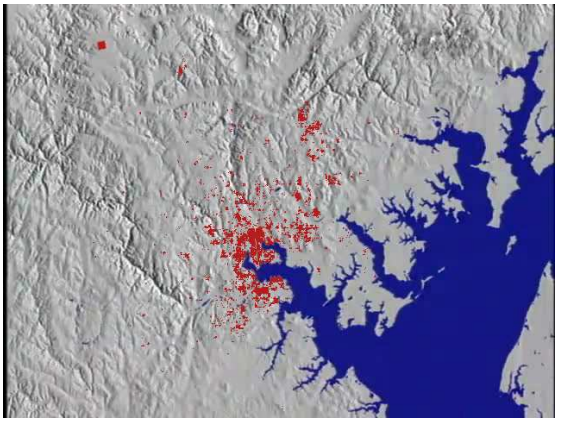
(a) Baltimore 1878



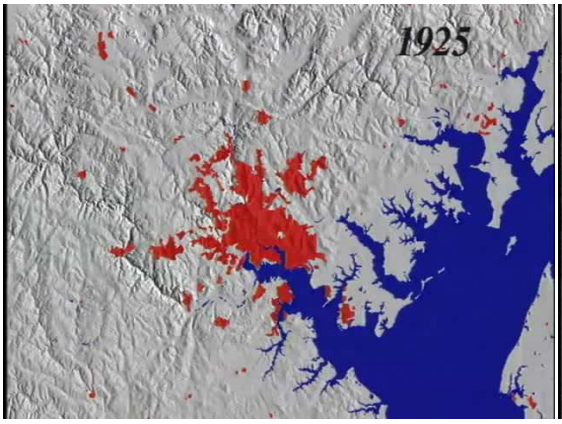
(b) Simulation 18



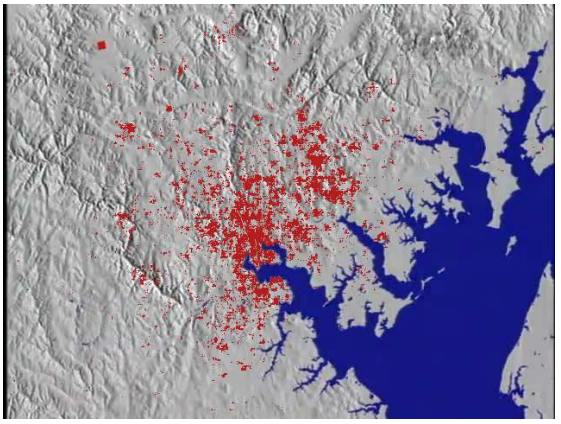
(c) Baltimore 1900



(d) Simulation 24

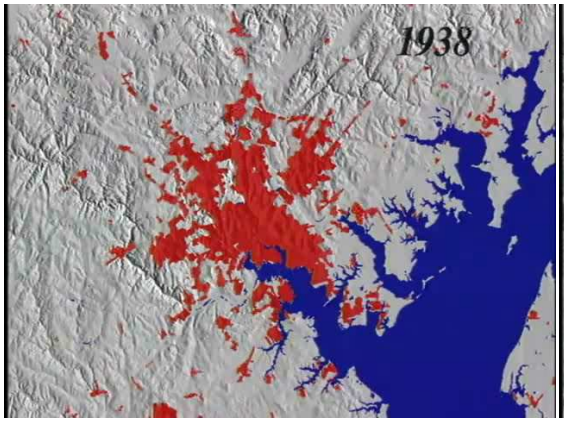


(e) Baltimore 1925

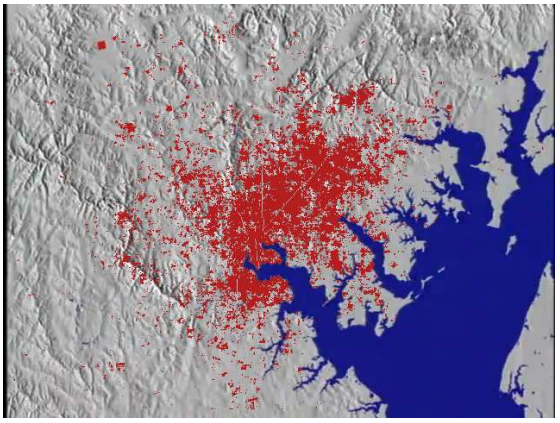


(f) Simulation 37

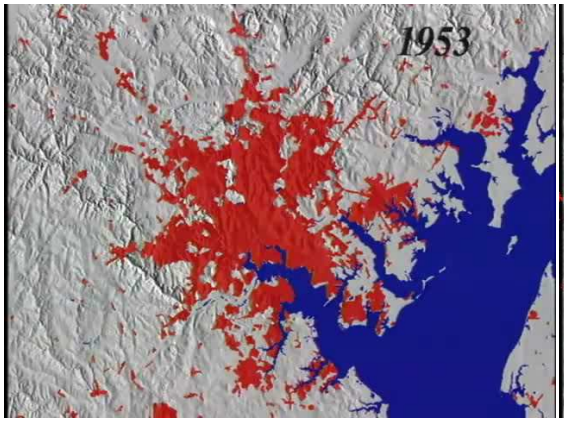
Figure 4.25: Real Baltimore and its simulation (cont)



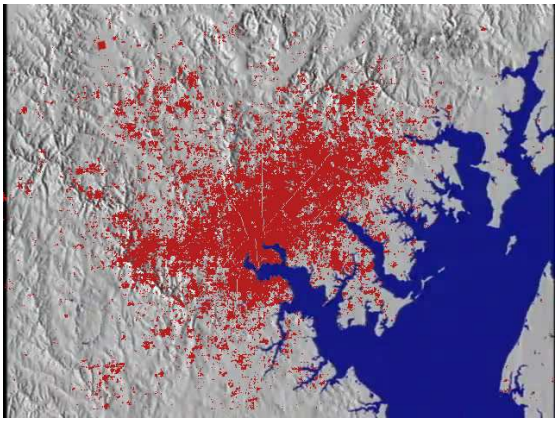
(a) Baltimore 1938



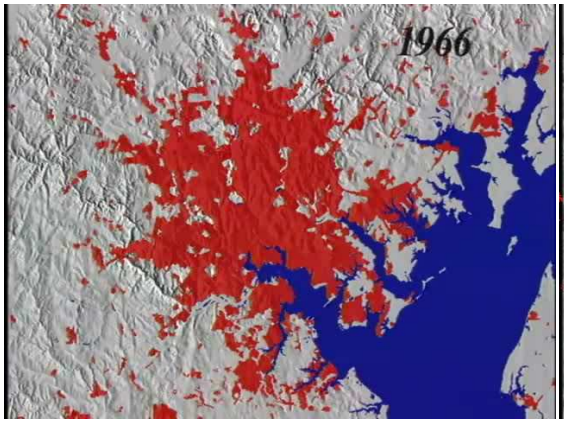
(b) Simulation 43



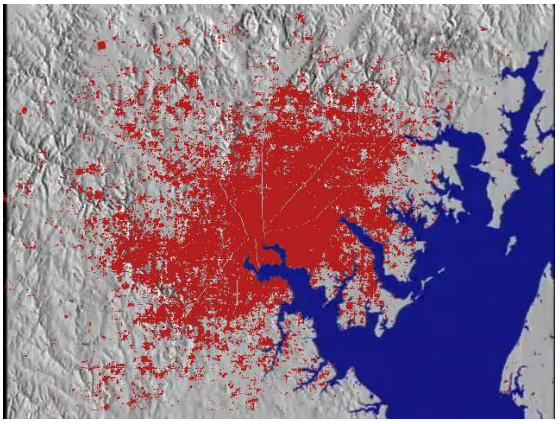
(c) Baltimore 1952



(d) Simulation 53

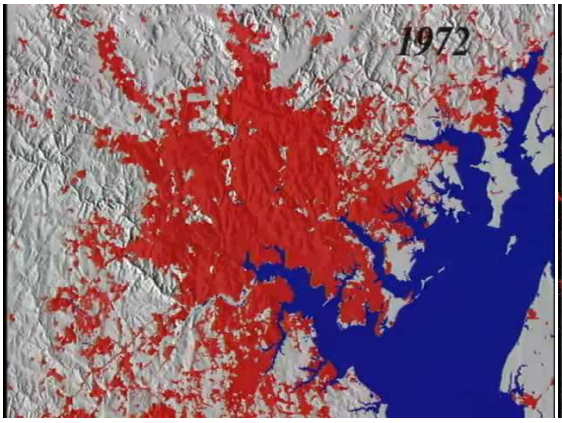


(e) Baltimore 1966

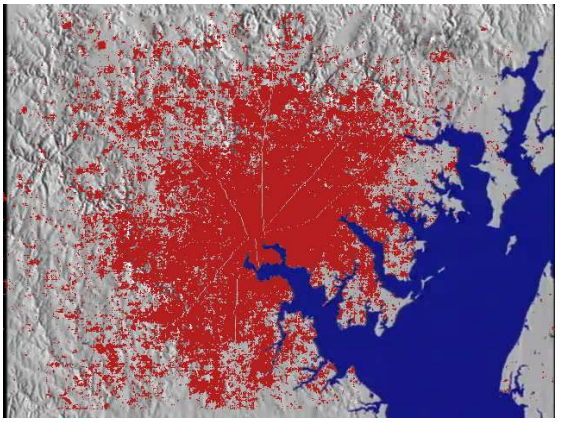


(f) Simulation 56

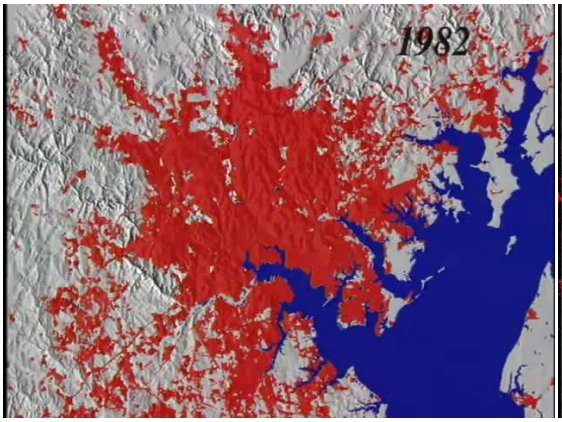
Figure 4.26: Real Baltimore and its simulation (cont)



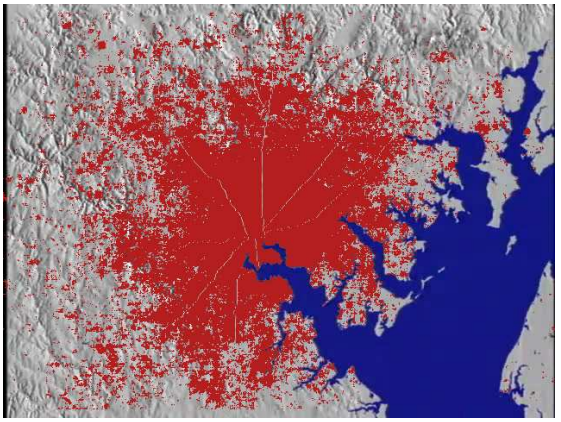
(a) Baltimore 1972



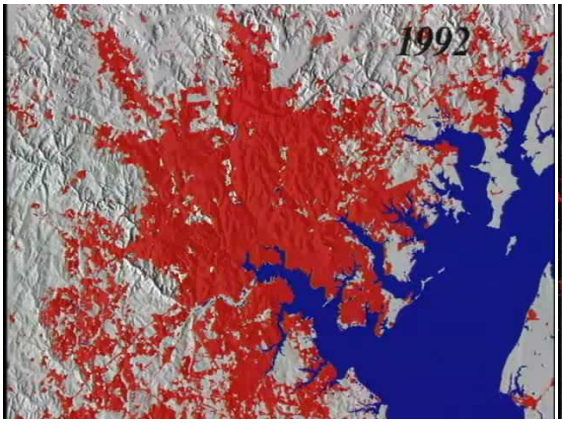
(b) Simulation 62



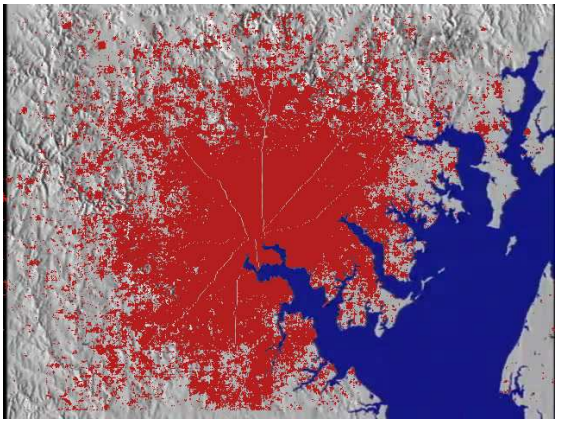
(c) Baltimore 1982



(d) Simulation 64



(e) Baltimore 1992



(f) Simulation 66

Figure 4.27: Real Baltimore and its simulation (cont)

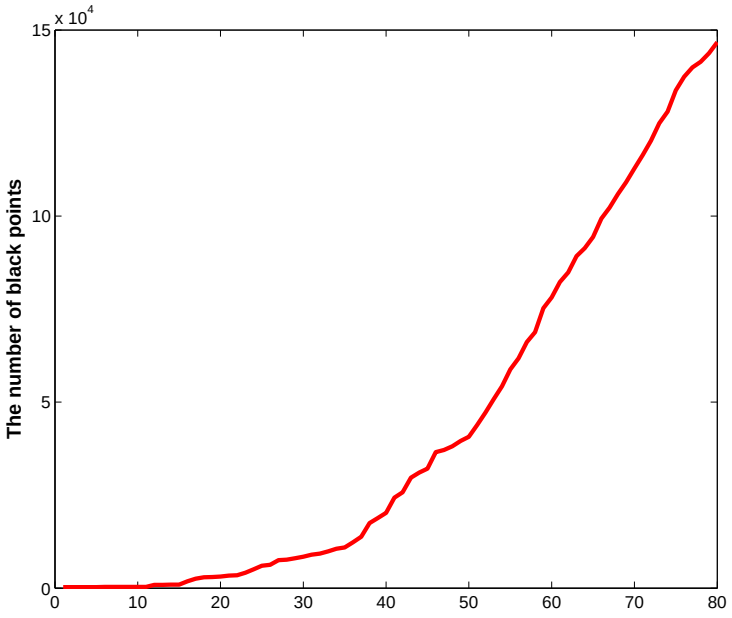


Figure 4.28: Evolution of 80 simulations for Baltimore.

Chapter 5

Simulation of the growth of Montargis

5.1 Introduction

In this chapter, two models will be used to simulate city growth with some differences. In both of these two, correlated gradient percolation model will be applied to simulate the growth of Montargis, a commune in France. The simulations are direct applications of the models and use real data of Montargis. In the first model, we use the density of population at given time as the occupation probability function. The model is built for both of the future and the past. Meanwhile, some factors of urban system are put in the second model, that gives more realistic simulations.
All of works are part of the project TRUC (Transport, Réseaux, Urbanisme, Croissance) supported by Centre Region. In this project, four laboratories (MAPMO, CRMD, CEDETE of Orléans University and CITERES of Tours University) worked on developing a model of territorial urban sprawl. A partnership agreement with the DDT Loiret (Direction Départementale des Territoires du Loiret) leads to choose SCoT of Montargois-en-Gâtinais (figure 5.1) as field experimental model.
We will briefly introduce geography of Montargis as well as data used in the simulations. Administrative data and statistics are got from Corine Land Cover (CLC), IGN (l'information grandeur nature), INSEE (National Institute of Statistics and Economic Studies - France), and the General Directorate of Public Finance (Direction Générale des Finances Publiques, DGFiP) are analyzed by Jean-Marc Zaninetti, geography professor, and Chloé Legrand, study engineer, working in CEDETE laboratory.

5.2 Geography of Montargis

Montargis is a commune in Loiret department in north-central France on the Loing river. The town is located about 110 km south of Paris and 70 km east of

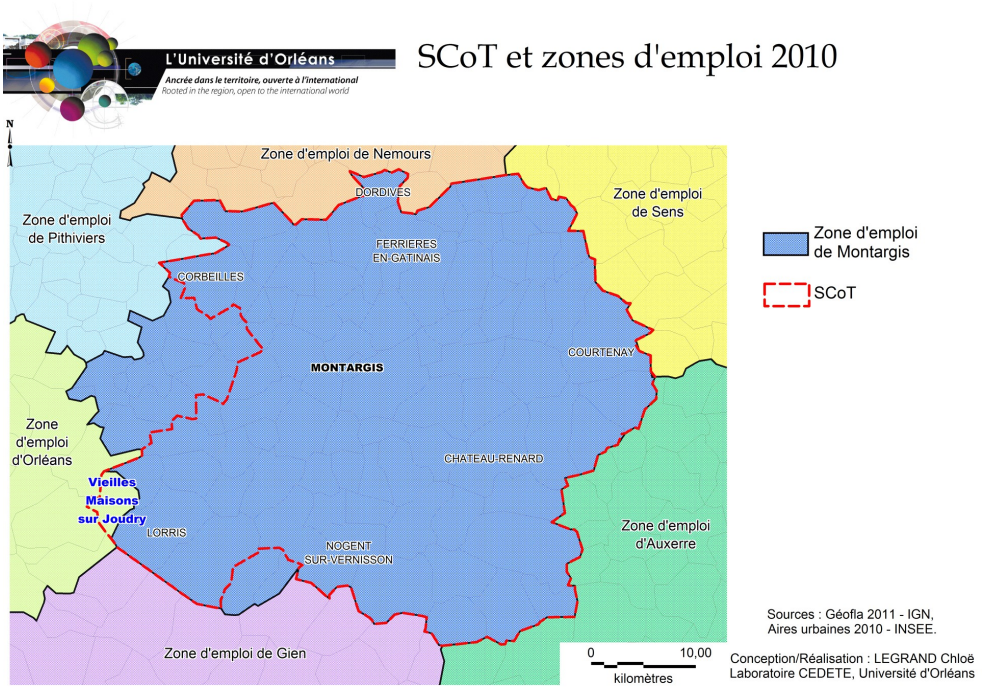


Figure 5.1: The boundary of SCoT of Montargois-en-Gâtinais is rather close to the employment area of Montargis. This area is redefined in 2010 by INSEE and DARES. The employment area is fully included in the Loiret, however it has some differences with the area of SCoT.

Orléans in the Gâtinais. Montargis is the second largest city in Loiret, after Orléans. It is near a large forest, and contains light industry and farming, including saffron. The area SCoT of Montargois-en-Gâtinais is rather close to the employment area of Montargis but they have some differences. There are 15 communes in the employment area of Montargis not included in the SCoT of Montargois-en-Gâtinais. Meanwhile, the commune Vieilles Maisons sur Joudry located in the extreme southwest of the area of SCoT is included in the employment area of Orléans and stays out of Montargis. Montargis and its urban area are the main center in the SCoT of Montargois-en-Gâtinais.
We will introduce some natural elements and populations of the area of SCoT of Montargois-en-Gâtinais (referred to as SCoT's Montargis for simplicity).

5.2.1 Population, Buildings and Land lots

There is a natural relationship between population and buildings. Of course, people live in buildings and from distribution of the buildings we can extract (very) approximative data for distribution of population. Nevertheless, clearly the type and

size of buildings need to be taken into account if we wish to correlate population and buildings. In this sense, although urban growth and 2-D morphology are important issues studied for centuries, they still form an incomplete approach to understand city dynamics.

Land is occupied by the buildings which were constructed within the limits of land lots obeying to local governments' laws. This implies that we have to use real data in order to simulate cities, and global paving by squares, hexagons or any other pattern is insufficient.

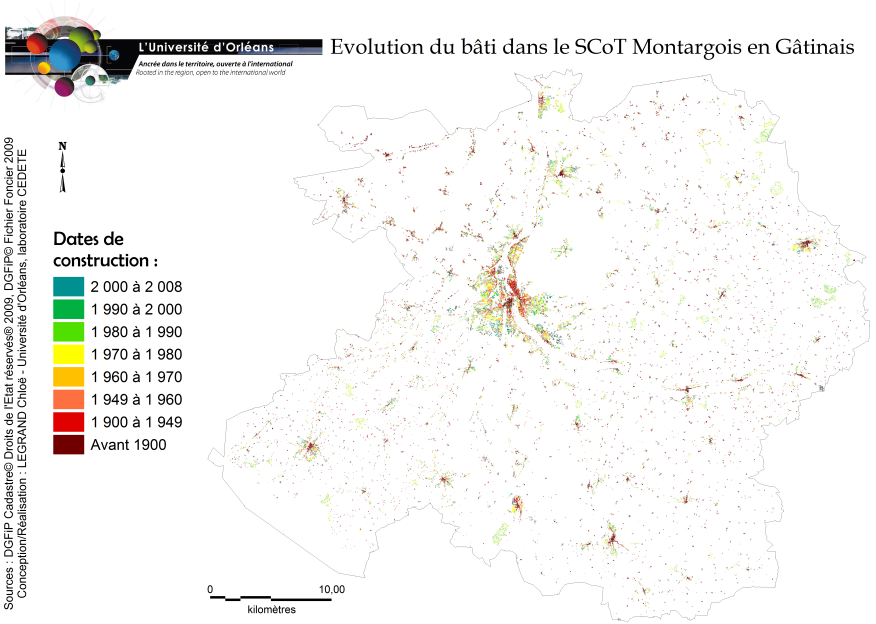
From data of DGFiP we re-constructed the evolution of buildings within the perimeter of SCoT's Montargis in figure 5.2. The main mass of the buildings within SCoT's Montargis is constructed between 1970 and 1980. In fact, from the early 1900's until 1960 the number of buildings did not change much; then it explosed until 1980 when the growth of buildings was slowed down again. The peak of the growth in the 1970s mostly concerned Montargis city.

SCoT's Montargis has 192 816 land lots (got from MAJIC files of DGFiP). Occupation time for a land lot is the time of construction of the oldest buildings on this lot. An "occupied" land lot can have one or more than one buildings. In Montargis city, most of the land lots are small and constructed with buildings for living purposes (see figure 5.3). Meanwhile, in further areas sizes of lots are larger and some of them are farms (the south of SCoT).

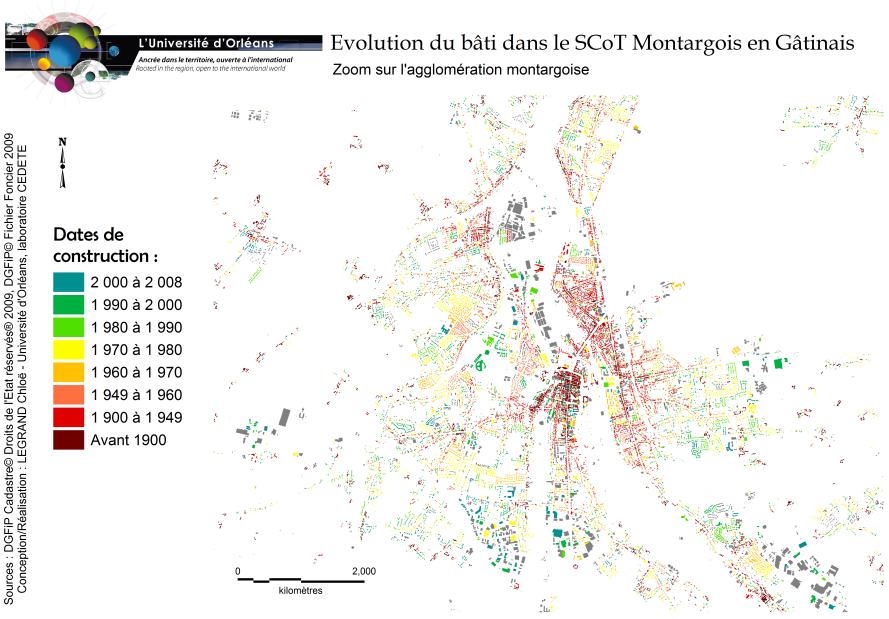
Before 1949, buildings were concentrated on both sides of Loing river and these constructions have continued grafting up until now. This place has become the center of the area of SCoT's Montargis.

From data of occupied land lots, time of buildings were considered. Note that there is a big difference between evolution of land lots/buildings and population. In figure 5.4, the evolution of population is relatively linear except of period from 2007 to 2008. On the other hand growth of premises and population, constructions were more increasing in the 1970s: between 1968 and 1990 there are 85 576 inhabitants in 36 512 premises in 1968 and 102 976 inhabitants in 61 155 premises in 1990.

In figure 5.5, the graph confirms that there is a more rapid evolution of premises, occupied land lots and their areas between 1968 and 1982. After the inflection in the 1980s, the curves which are almost parallel since 1990, showing that the population and the occupied land lots evolve together.

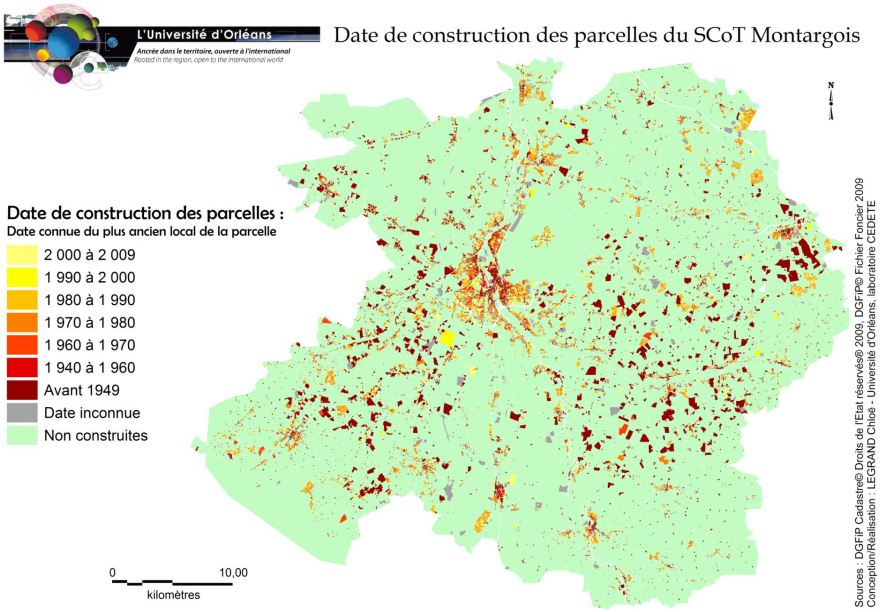


(a) The buildings in the area of Montargis

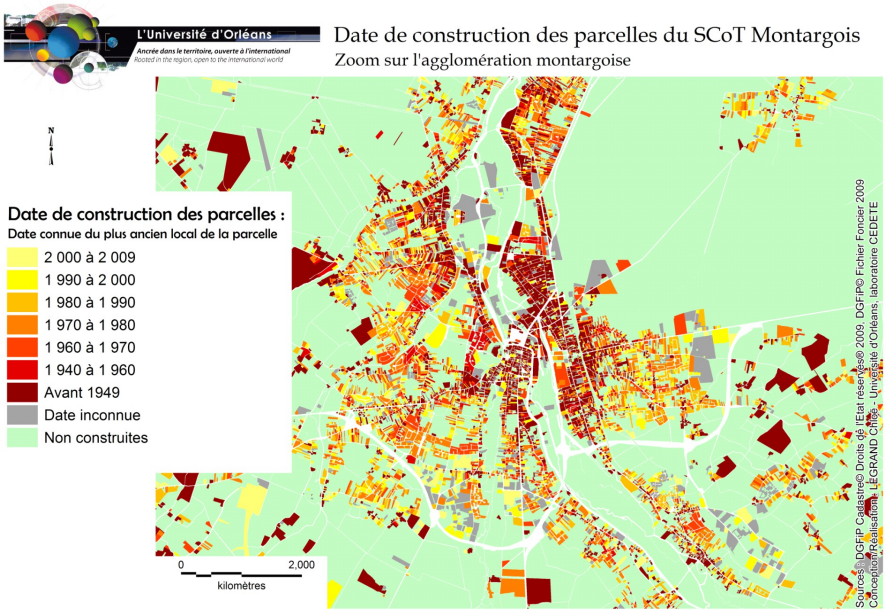


(b) The zoom of the area of Montargis at Montargis commune

Figure 5.2: The buildings and time when they were constructed in the area of Montargis and its zoom (around the center).



(a) The occupied land lots in the area of Montargis



(b) The zoom of the area of Montargis at Montargis commune

Figure 5.3: The occupied land lots and time when they were constructed in the area of Montargis and its zoom (around the center).

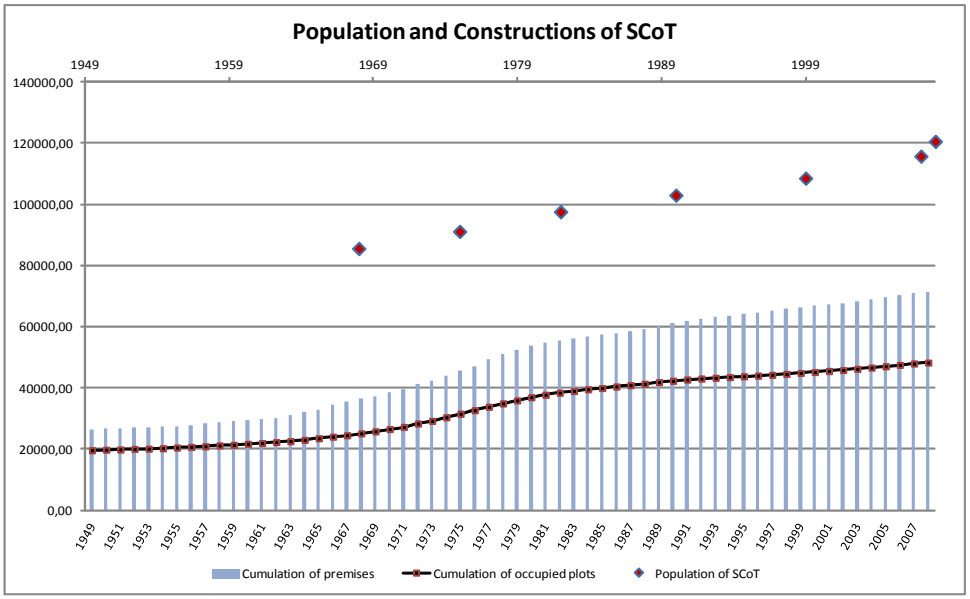


Figure 5.4: The evolution of population and occupied land lots of Montargis from 1949 to 2008.

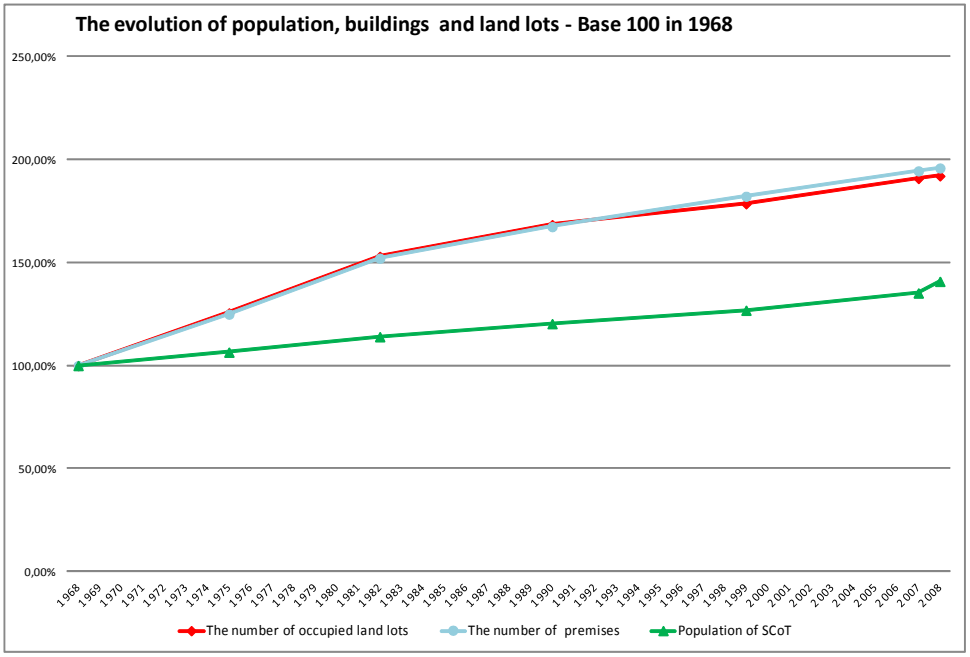


Figure 5.5: The evolution of population, occupied land lots and premises of Montargis from 1968 to 2008.

5.2.2 The natural effects

As shown above, the buildings concentrate around the center of Montargis on the Loing river and the occupated area grow rapidly. Besides that, this growth also occurs at the other centers and along the axis (roads and rivers).
From figure 5.6 we can see that 50% of buildings of Montargis are within 500m of the rivers, while 50% of the area of SCoT's Montargis is located within one km from the rivers and over 70% of buildings are in this area.

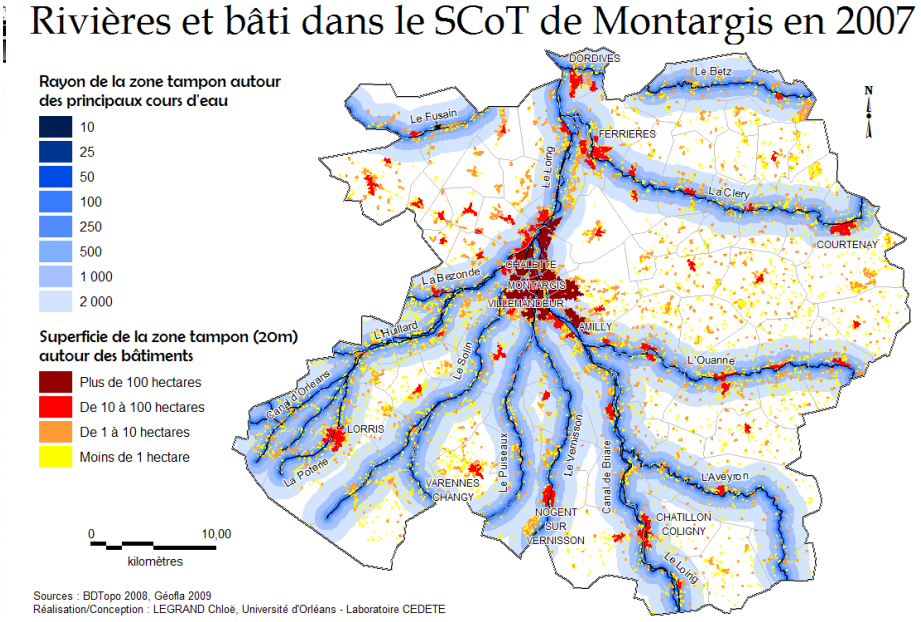
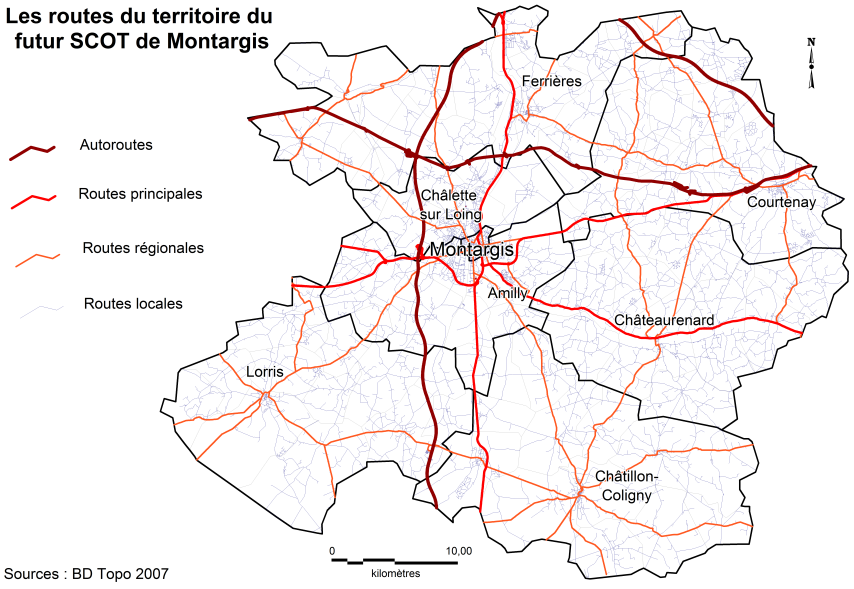


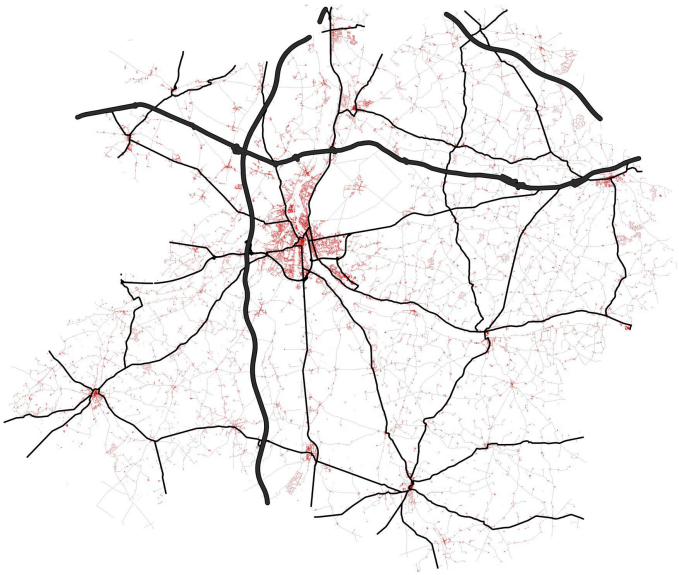
Figure 5.6: River system with its around zone and the buildings of SCoT's Montargis in 2007.

Road system is an important factor in urban dynamics. When concentration of population increases in an area, the road system will develop in this same area but also connections to neighboring areas. On the other hand, if a place has a convenient transports network it will attract habitants as well as commercial activities.
The urban area of Montargis is well served by road networks which includes N7 national road to Paris, RN60 national road to Orléans, A6 and A77 highway, and a complex road system (figure 5.7).

Accessibility of a place represents the degree of facility to reach this place from one or more other places by a person who can use all existing means of transport. It is measured by distance or by time to get from a set of nodes to another. A *node* is selected on a road according to its importance in the urban system or its position with respect to other roads (interchanges, crossroads, ...).



(a) The road system of SCoT's Montargis

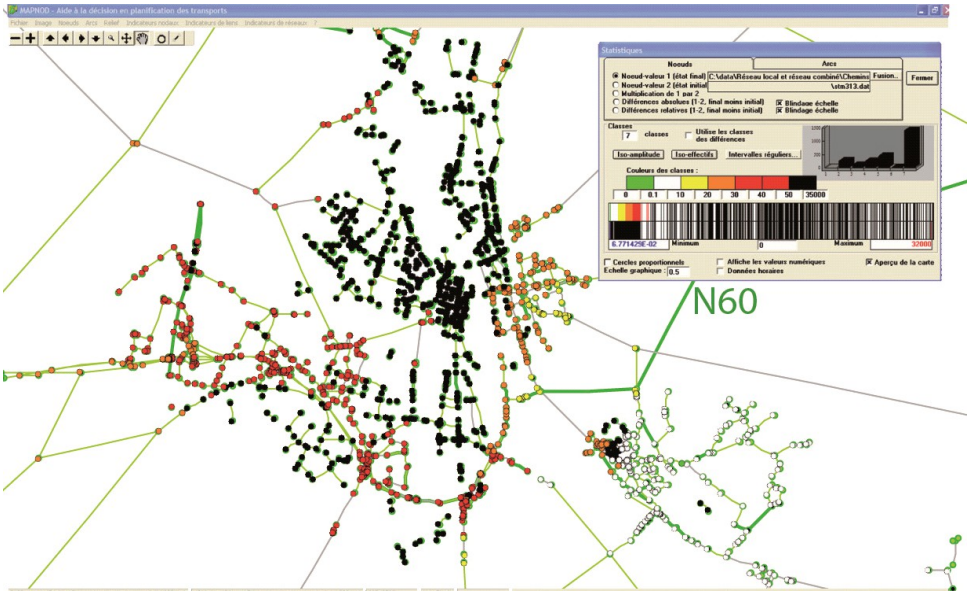


(b) The road system and buildings

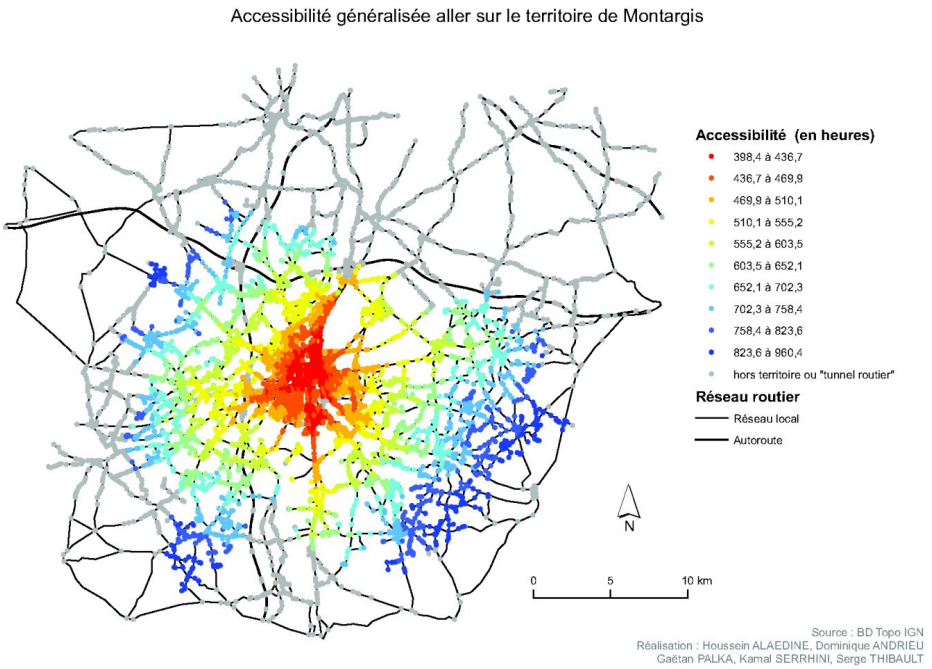
Figure 5.7: The occupied land lots and time when they were constructed in the area of Montargis and its zoom (around the center).

This issu has been studied by Houssein Alaeddine, Dominique Andrieu, Gaëtan Palka, Kamal Serrhini and Serge Thibault (laboratory CITERES of Tours University). Figure 5.8 shows time to get in the center of Montargis city from the outer area.

Besides the attractive zones, in the urban system there are the zones with high risk/noises/traffic, unattractive to construct. Other zones such as flood zone and forest impose their particular signatures and are taken under account in our simulations.
In figure 5.9, $40,38$ km^2 or 25% of the area of the SCoT is tree-covered and there are not many buildings into or near forests. Notice that grownd impact of these forests does not change in time significantly.
Habitants concentrate along both sides of Loing river especially in Montargis city. However, the adjacent region of this river is of high risk to constructions because of flood. We call it "flood zone". In figure 5.10, the red zone corresponds to very high hazard area and the construction possibilities are extremely limited there; orange zone includes natural areas and dispersed areas, in this region the possibilities of new construction is very limited, although the hazard is from low to medium; the yellow corresponds to low-urbanized areas where the hazard is also from low to medium, in this region the urbanization is controlled as much as possible. There are 1001 buildings in flood zone which are 109 in the red zone, 290 in the orange, and 602 in the yellow one.



(a)



(b)

Figure 5.8: The accessibility to Montargis city of its near regions.

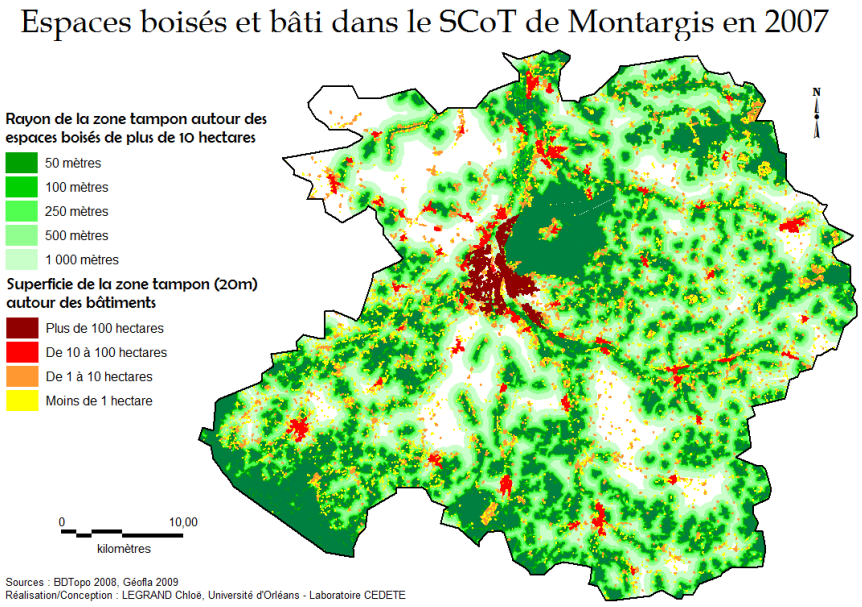


Figure 5.9: The forest area and the buildings of Montargis in 2007.

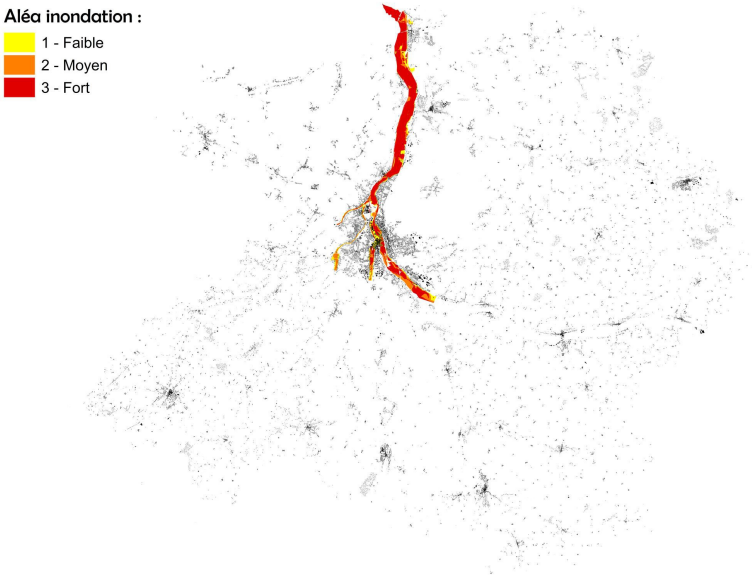


Figure 5.10: The area of the flooding along Loing river.

5.3 Simulations of the growth of Montargis

5.3.1 Centrality and density

The city center is the (often historical) area of a city where commerce, entertainment, shopping and political power are concentrated. A city center is often the first settled part of a city, which can make it the most historical part of a city.
In our model, we assume that the center is the point which is the highest local density. The local density is density of the black points (buildings) in the ball radius 50 units (points). In figure 5.11, the peak, the point $[3119, 4035]$ on an image of size 8905×7441, is considered as the center of SCoT's Montargis later.

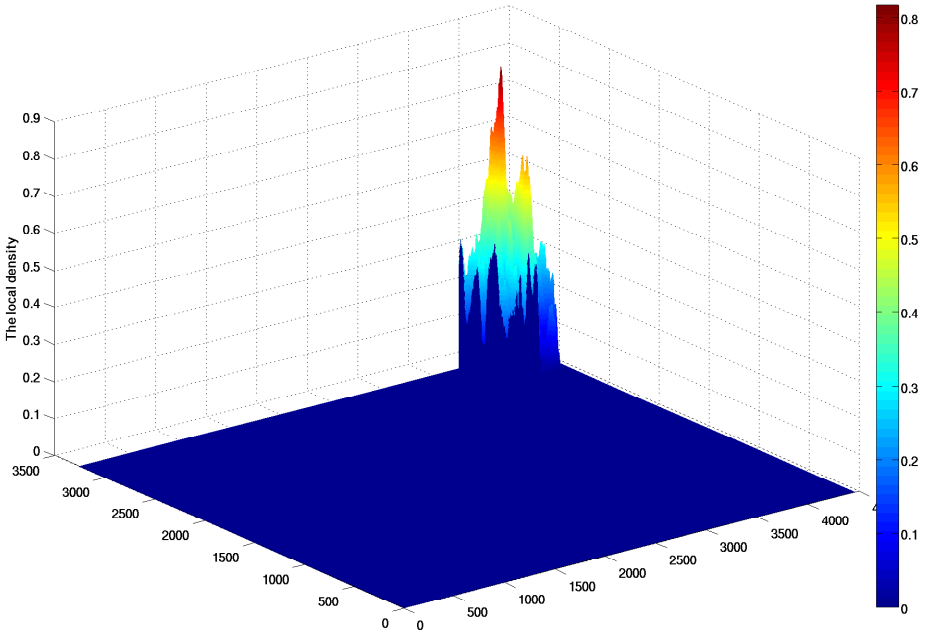


Figure 5.11: The local density of SCoT's Montargis in the balls of radius 50 points, the red peak is point $[3119, 4035]$.

Let us put in detail two random models aiming to simulate Montargis' growth.

5.3.2 Model 1: Differentiel Equation Density model

We first check density of SCoT's Montargis. The graph in figure 5.12 shows that the distribution of density of buildings of Montargis decays as an exponential function of square root of distance to center r. If we only consider the relation between one point and its neighbors, this graph behaves like the exponential function of $\sqrt{r}$.

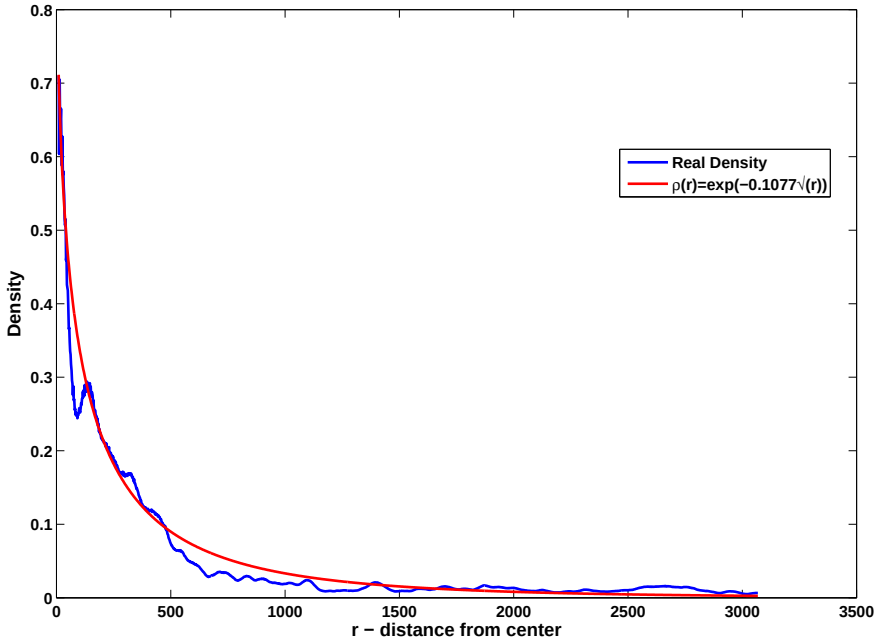


Figure 5.12: The average local density of SCoT's Montargis by distance to center r. It is near the exponential function of $\sqrt{r}$, $\rho(r) = e^{-\lambda\sqrt{r}}$.

This is the reason we choose the equation (5.3.1) in this model. The unusual decrease in the first part of real density is caused by the presence of the big forest near Montargis city. Besides, in areas at distance from 500 to 1500 units of the center, the density is rather low but in further away it seems growing up again. This is explained by the presence of other small urban centers in SCoT's Montargis.
From the above observation, it is clear that density can not be uniformly radially determined. We propose a model (implying density) where the occupancy probability function of a point depends on its neighbors and not only on distance to the center.

We next introduce some notations and concepts in this model. As a first step, consider the square lattice. For each point $z = (i, j)$, $i = 1, ..., M$; $j = 1, ..., N$ (a pixel on the image), we define a binary matrix $\mathcal{M}$ of size $M \times N$:

$$\mathcal{M}(z) = \begin{cases} 1 & \text{if } z \text{ is black or occupied,} \\ 0 & \text{if } z \text{ is white or vacant.} \end{cases}$$

This matrix will be used to visualize a city or a simulation.
In urban system, it is easy to see that if an empty position (no building) has many neighbors surrounding it, it will be occupied with high probability. Due to this attractive parameter, we will define the occupancy probability function used in this

modeling as the weighted density

$$p(z) = \frac{1}{C} \sum_{z \in \mathcal{M}_0} e^{-\lambda \sqrt{r_z}} \tag{5.3.1}$$

where C is a constant, λ is the density gradient (mentioned in the previous section), $\mathcal{M}_0$ is a binary matrix with value 1 at constructed points of Montargis at given time and r_z is distance of site z to the center of lattice. The CGP model, with the function (5.3.1) applied in the simulations, is a random model. This means that in the realisation matrix $\mathcal{M}$, the vacant sites becomes occupied ($\mathcal{M}(z) = 0$ changes $\mathcal{M}(z) = 1$) randomly. Note that constructions are not destroyed in this model.
Let us denote by $A_{\mathcal{M}}$ the set of the black points and by $L_{\mathcal{M}}$ the set of all points

$$A_{\mathcal{M}} = \{z : \mathcal{M}(z) = 1\},$$
$$L_{\mathcal{M}} = \{z : z \text{ is a point on } \mathcal{M}\}.$$

We define the concentration function $\rho(z)$ as following:

$$\rho(z) = \mathcal{M}(z) * \delta(z) = \sum_{z' \in A_{\mathcal{M}}} \delta(z - z') \tag{5.3.2}$$

where $(*)$ is convolution operator and

$$\delta(z) = \frac{e^{-\sqrt{|z|}}}{C_z}, \quad C_z = \int_{L_{\mathcal{M}}} e^{-\sqrt{|z|}} dz.$$

Let $\mathcal{R}(z)$ denote the expectation of the concentration function

$$\mathcal{R}(z) = \mathbb{E}(\rho(z)) = \mathbb{E}(\mathcal{M}(z)) * \delta(z). \tag{5.3.3}$$

It is clear that $\mathcal{R}(z)$ belongs to the interval $[0, 1]$.
We next turn to modelisation of a city in the future and also in the past.

5.3.2.1 Forward model

Recall that, by assumption, a building that is constructed will not destroyed, hence the city grows in an irreversible manner (cluster growth). This means that we will simulate the growth of city by leaving inchanged the occupied sites, replacing the vacant sites (buildings) on the matrix of Montargis at given time by the occupied sites. So, all of the current black points will be black in future, just only white points will be changed. Here $\mathcal{M}_1$ stands for the city in future. The growth model can be formalized as following:

$$\mathcal{M}_1(z) = \mathcal{M}_0(z) + \mathcal{H}_1(z) \tag{5.3.4}$$

where $\mathcal{H}_1(z)$ is a binary matrix with white at the black points of $\mathcal{M}_0$. Here the correlation gradient percolation will be applied to find $\mathcal{H}_1$ on the white points of

matrix $\mathcal{M}_0$.
From (5.3.1) and (5.3.2), it is easy to check that

$$p(z) = \rho_0(z) = \mathcal{M}_0(z) * \delta(z),$$

and this value is not random, but depends on $\mathcal{M}_0$. The function ρ will be used in the simulations in this model.
In CGP model, the value of $\mathcal{H}_1(z)$ is 1 with a probability $\rho_1(z)$ and on the contrary $\mathcal{H}_1(z) = 0$ with probability $1 - \rho_1(z)$, with $\rho_1(z)$ unknown and random,

$$\rho_1(z) = \mathcal{M}_1(z) * \delta(z),$$

and its expectation is denoted by $\mathcal{R}_1(z)$,

$$\mathcal{R}_1(z) = \mathbb{E}(\rho_1(z)). \tag{5.3.5}$$

Thus $\mathcal{H}_1(z)$ is a random matrix and its expectation is

$$\mathbb{E}\left(\mathcal{H}_1\left(z\right)\right) = \left(1 - \mathcal{M}_0\right)\left(z\right)\mathcal{R}_1\left(z\right). \tag{5.3.6}$$

By taking the expectation then putting convolution into the equation (5.3.4) we get:

$$\mathbb{E}\left(\mathcal{M}_1\left(z\right)\right) * \delta\left(z\right) = \mathbb{E}\left(\mathcal{M}_0\left(z\right)\right) * \delta\left(z\right) + \mathbb{E}\left(\mathcal{H}_1\left(z\right)\right) * \delta\left(z\right),$$

or

$$\mathbb{E}\left(\rho_1(z)\right) = \mathbb{E}\left(\rho_0(z)\right) + \mathbb{E}\left(\mathcal{H}_1\left(z\right)\right) * \delta\left(z\right).$$

Using the equations (5.3.5) and (5.3.6), we obtain

$$\mathcal{R}_1\left(z\right) = \mathcal{R}_0\left(z\right) + \left[\left(1 - \mathcal{M}_0\right)\left(z\right)\mathcal{R}_1\left(z\right)\right] * \delta\left(z\right), \tag{5.3.7}$$

with $\mathcal{R}_0\left(z\right) = \rho_0(z)$.
Next, we will solve this equation by defining an operator $\mathcal{T}$

$$\mathcal{T}\mathcal{R}\left(z\right) = \left[\left(1 - \mathcal{M}_0\right)\left(z\right)\mathcal{R}\left(z\right)\right] * \delta\left(z\right). \tag{5.3.8}$$

It is easy to check that $\mathcal{T}$ is a linear positive operator. It follows that the equation (5.3.7) can be rewritten as follows:

$$\mathcal{R}_1\left(z\right) = \mathcal{R}_0\left(z\right) + \mathcal{T}\mathcal{R}_1\left(z\right).$$

Therefore,

$$\mathcal{R}_1\left(z\right) = \left(\mathbb{I} - \mathcal{T}\right)^{-1}\mathcal{R}_0\left(z\right). \tag{5.3.9}$$

The constant function 1 is a solution of the equation (5.3.9). In the simulations, $\mathcal{R}\left(z\right)$ is used as probability function in CGP model. After a long time, when $\mathcal{R}\left(z\right) = 1$, the building will be full on the map (or all of sites on lattice are occupied). But, how the city is at time t? To find this we define the function $\mathcal{R}_t\left(z\right)$ at time t:

$$\mathcal{R}_t\left(z\right) = \left(\mathbb{I} - t\mathcal{T}\right)^{-1}\mathcal{R}_0\left(z\right), \tag{5.3.10}$$

with $t \in [0,1]$. In this function, when $t = 0$ the concentration does not change $\mathcal{R}_t(z) = \mathcal{R}_0(z)$, while in $t = 1$, $\mathcal{R}_t(z)$ concentration equals 1.
We put t into the equation (5.3.7) to get that:

$$\mathcal{R}_t(z) = \mathcal{R}_0(z) + t\left[(1 - \mathcal{M}_0)(z)\,\mathcal{R}_t(z)\right] * \delta(z).$$

We then replace $\mathcal{R}_t$ in the right-hand side of the equation by the left-hand side we have

$$\mathcal{R}_t(z) = \mathcal{R}_0(z) + \sum_{n\geq 1} t^n \left[(1 - \mathcal{M}_0)^n(z)\,\mathcal{R}_0(z)\right] * \delta^{(n)}(z). \tag{5.3.11}$$

This function has been used as the occupation probability function in the simulations of the growth of Montargis in the results section. A version of this method can (and will) also be used as the model to simulate the city in the past.

5.3.2.2 Backward model

Similarly to the previous model, we assume that a black point will stay black in the future. However, we will change the above model to be able to simulate a city in the past by using dual matrix instead of $\mathcal{M}$.
Consider matrix $\mathcal{M}'_0 = 1 - \mathcal{M}_0$, at each site z, if $\mathcal{M}_0 = 1$ then $\mathcal{M}'_0 = 0$ and vice-versa. The growth of $\mathcal{M}'_0$ can be written:

$$\mathcal{M}'_{-1}(z) = \mathcal{M}'_0(z) + \mathcal{H}'_{-1}(z).$$

The probability that a point z becomes black in this model is $\mathcal{R}'(z)$. We can see that,

$$\mathcal{R}'(z) = 1 - \mathcal{R}(z). \tag{5.3.12}$$

Thus we realize the same way to simulate $\mathcal{H}'_{-1}$ with the probability function $\mathcal{R}'_{-1}(z)$ to get

$$\mathcal{R}'_{-1}(z) = \mathcal{R}'_0(z) + \left[(1 - \mathcal{M}'_0)(z)\,\mathcal{R}'_{-1}(z)\right] * \delta(z).$$

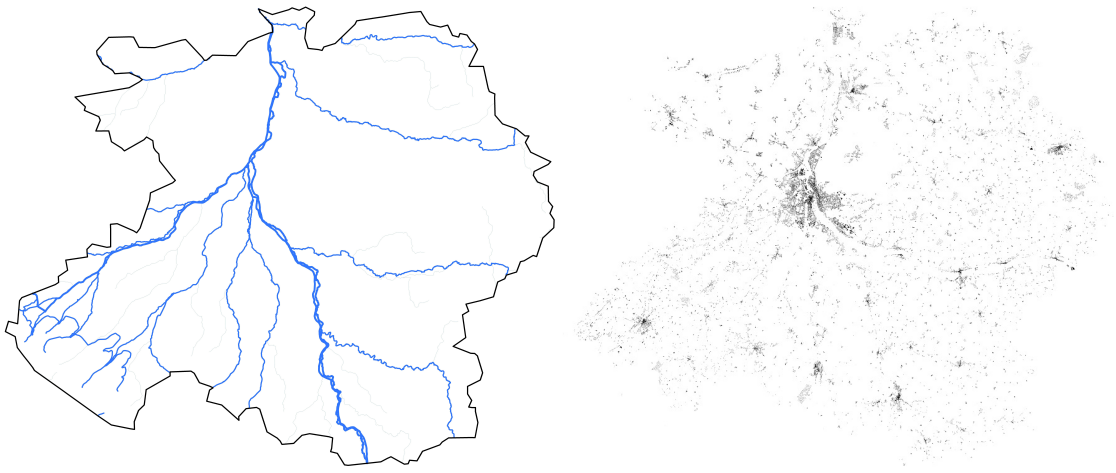
From the equation (5.3.12), this equation can be rewritten:

$$\mathcal{R}_{-1}(z) = \mathcal{R}_0(z) - \left[\mathcal{M}_0(z)\,(1 - \mathcal{R}_{-1}(z))\right] * \delta(z).$$

Similarly to equation (5.3.11), we can obtain the occupation probability function by time t

$$\mathcal{R}_t(z) = \mathcal{R}_0(z) - \sum_{n\geq 1} t^n \left[\mathcal{M}_0^n(z)\,(1 - \mathcal{R}_0(z))\right] * \delta^{(n)}(z)$$

with $t \in [-1,0]$.
By using CGP model with the occupancy probability functions as in (5.3.11) and (5.3.2.2) with strong correlation exponent $\alpha = 0.005$, we have simulated the growth of Montargis in future and also shown how it is in past. All of them will be illustrated in the following section.



(a) The rivers of Montargis (b) Montargis in 2007

Figure 5.13: The river and the constructed building of Montargis at 2007, the data are got from the project TRUC.

5.3.2.3 Results

The simulations base on the data (buildings) of Montargis from the project Transport-Réseau-Urbanisation-Croissance (TRUC) and are realized with MATLAB program.
Besides the buildings, we consider the influence of the rivers (figure 5.13a) on the growth of the city. We assume that the effect of the rivers on the occupied probability is the same as the buildings. The rivers are thus considered as black points like the buildings.
Firstly, the forward model is applied on Montargis 2007 (Figure 5.13b) as the starting point ($t = 0$). The black points in the starting matrix contain the constructed buildings at 2007 and the rivers system of Montargis. The simulations are done with a time sequence of scale $dt = 0.01$, we get some at $t = 0.01, 0.02, 0.03$ in the figure 5.14.

Secondly, consider Montargis 2007 as the starting point and use the backward model to simulate for the past at $t = -0.01, -0.02$ (figure 5.15). We see that the number of buildings (black points) decrease very fast at the first simulation.

Finally, we use the simulation of backward model at $t = -0.02$ (figure 5.15b) as the starting point to find how it is in future by using the forward model with $t' = 0.01, 0.02$ (figure 5.16).

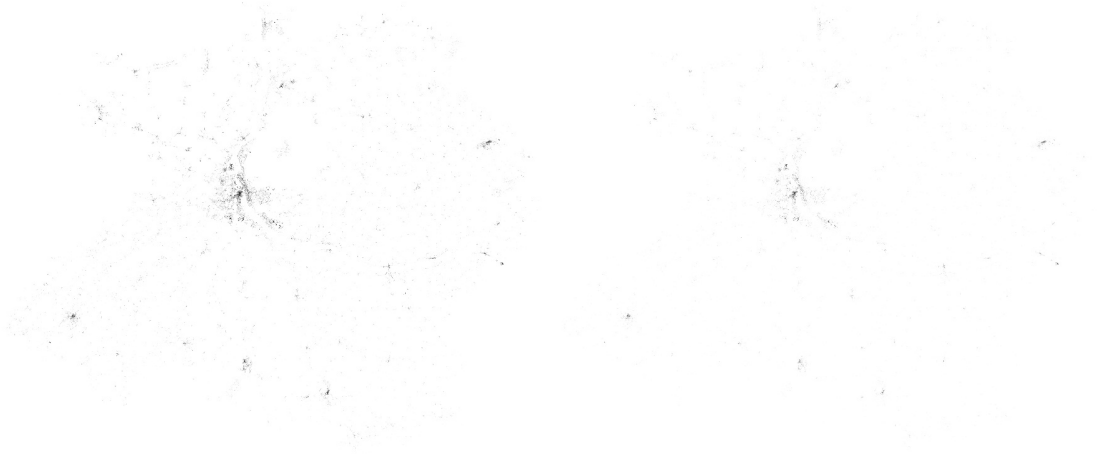
(a) Montargis in 2007

(b) $t = 0.01$

(c) $t = 0.02$

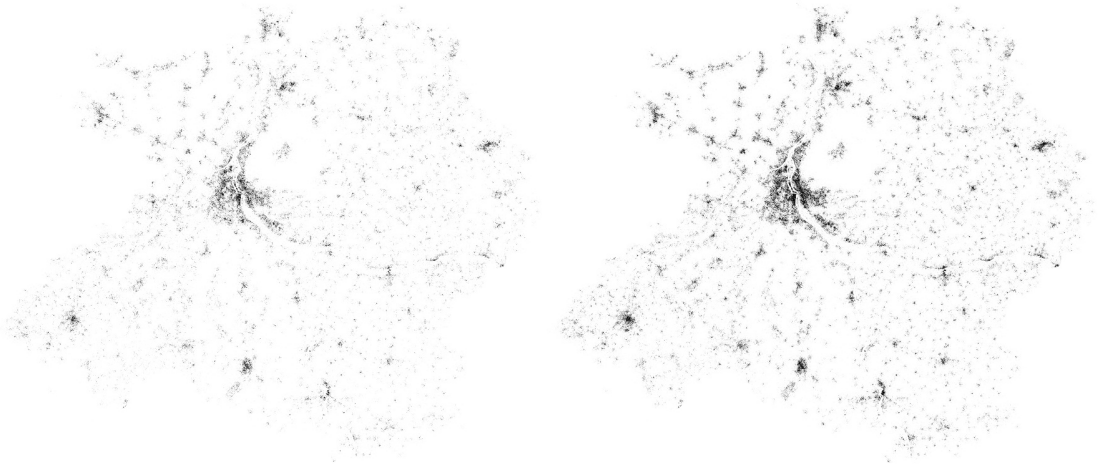
(d) $t = 0.03$

Figure 5.14: The simulations in future of Montargis using CGP model with probability function (5.3.11) in different time t.



(a) $t = -0.01$ (b) $t = -0.02$

Figure 5.15: The simulations for the past of Montargis from 2007 using the backward model in different time t.



(a) $t = 0.01$ (b) $t = 0.02$

Figure 5.16: The simulations for the future of the simulation 5.15b as the starting point using the forward model with different t.

5.3.3 Model 2

In model 1, we considered the distribution of population (or buildings) to model the growth of population in future or the reduction in past on Montargis. Besides, in reality, there are many factors interacting in urban system such as the minimum around area of a building; the accessibility, the commercial areas, the schools, ...; the geographical location that is difficult to live or construct; .... We will study these factors and add them to the second model to simulate the growth of Montargis.

5.3.3.1 Model

Lot (or plot) is a land with a determined area and boundary, owned by an individual or organization (see the figure 5.17). On that, owners can construct the buildings to live or to rent. A lot that has buildings is called the occupied plot and empty (or vacant) if it has nothing inside.
The plots are of different size (we are no longer in the square lattice paradigm). The big plots which are usually far from the center of city can be divided into the smaller ones. However, due to the cost of division, a big plot is often divided all at once to sell or construct. In fact, we can admit that a big plot will be divided when it has high potential to be sold and hence high occupation probability. We can thus assume that an occupied plot means that it is not possible to build any more or that all its divisions were constructed. This also meets previous hypothesis that an occupied land lot will maintain its status in future (it does not become vacant or empty). These land lots are the objects of our study. They replace sites (points) in the previous model.

In addition to this point, this model does not consider size of buildings in land lots. In fact, there are different types of buildings with different size and height.
Moreover, public buildings (such as commercial ones, schools, hospitals ...) are also not differentiated in the model. The attraction of these buildings is assumed to be

(a) A photograph (b) The map with the plots

Figure 5.17: An example of the plots on Montagis with the buildings (the pink blocks) inside (source from `geoportail.gouv.fr`).

the same with other buildings for living.

To summarize, we adapt CGP model on land lots with assumptions:

- The change status of a land lot (vacant to occupied) occurs with a probability related not only to distance to the center of the city but also to local density of the urban system and accessibility,
- An occupied land lot will maintain occupied, there is not changing from occupied to vacant,
- Not distinguish the different types of buildings in land lots in both of size and purpose of use.

In the next part we will establish this model. First of all, we define the binary matrix $\mathcal{M}$ of size $M \times N$ containing center points of land lots of Montargis, $z \in [1, M] \times [1, N]$ (size of Montargis) as following:

$$\mathcal{M}(z) = \begin{cases} 1 & \text{if } z \text{ is a occupied land lot,} \\ 0 & \text{if } z \text{ is empty.} \end{cases}$$

This matrix illustartes the simulations of growth of Montargis on the next section.

Of course, in real life, the occupied probability of an empty plot depends on not only its neighbors but also many factors. One of them is the accessibility of a set of nodes (figure 5.18b). These nodes (mainly cross-roads) are categorized into five levels representing the total time to reach a node from the others. The locations from which it is easy to go to other positions will be more attractive to the habitants. Conversely, in areas such as the forests or floodplain (see in figure 5.18c) will be less prefered for living.

Therefore we construct the occupied probability function at step i as following:

$$p_i(z) = \frac{1}{C} \sum_{z_1 \in B} e^{-\lambda_1^i d(z,z_0) - \lambda_2 d(z,z_1) - \lambda_3^i d(z,z_2) t_{node}(z_2) - \lambda_4^i d(z,z_3)} \tag{5.3.13}$$

where

- C is a constant,
- $d(x, y)$ is the distance between x and y, defined in the equation (4.2.1),
- λ_1^i, λ_2, λ_3^i, λ_4^i are parameters depended on step i,
- B is the set of the black points (occupied sites) and road system,
- z_0 is center point of city,
- z_2 is the nearest node of z,
- z_3 is the nearest distance to river of z,
- $t_{node}(z_2)$ is time of accessibility to center from z_2.

As mentioned above, Montargis has an important forest area and these area affect the growth of buildings. Under the influence of the forest (but also inside risk

zones), the occupation probability decreases by some amount C_0,

$$p_i^*(z) = \begin{cases} \min\{p_i(z) - C_0; 0\} & \text{if } z \in S_F, \\ p_i(z) & \text{if } z \notin S_F, \end{cases} \tag{5.3.14}$$

where:

- S_F is the set of plots in forest and floodplain zone,
- C_0 is a constant (we take $C_0 = 0.03$ here), and
- $p_i(z)$ is from the equation (5.3.13).

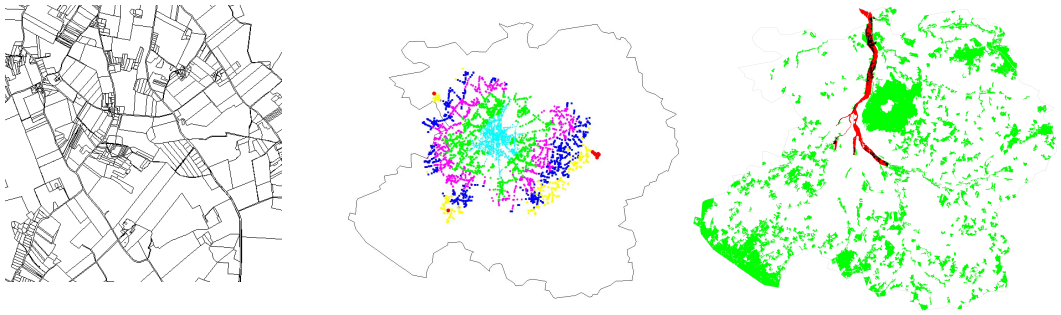
With z is the center of a plot, the detail of the occupied probability function (5.3.14) is

$$p_i(z) = \frac{1}{C} \sum_{z_1 \in B} e^{-0.0001 i^{-1/4} d(z,z_0) - 0.03 d(z,z_1) - 0.0001 i^{-1/4} d(z,z_2) t_{Noeud}(z_2) - 0.001 i^{-1/4} d(z,z_3)}. \tag{5.3.15}$$

By the above assumptions, CGP model is used on vacant land lots which become to be occupied with probability (5.3.13). The correlation exponent in the model is strong: $\alpha = 0.001$. With starting point is Montargis 1900 considered as $\mathcal{M}_0$, a simulation is performed on its previous step.

5.3.3.2 Results

By taking Montargis 1900 (figure 5.19) as a starting point we obtain the simulations for the growth of Montargis for the years after 1900 in the figures 5.20 and after 2007 in figure 5.21.



(a) Land lots (b) Accessibility (c) Risk zone

Figure 5.18: The factors that affect the growth of the city (Montargis) are added into the occupancy probability function in the simulations. The plots is used in the model through their center points; the accessibilty is considered by time at node with five levels; the risk zone contain forest (green) and floodplain (red).

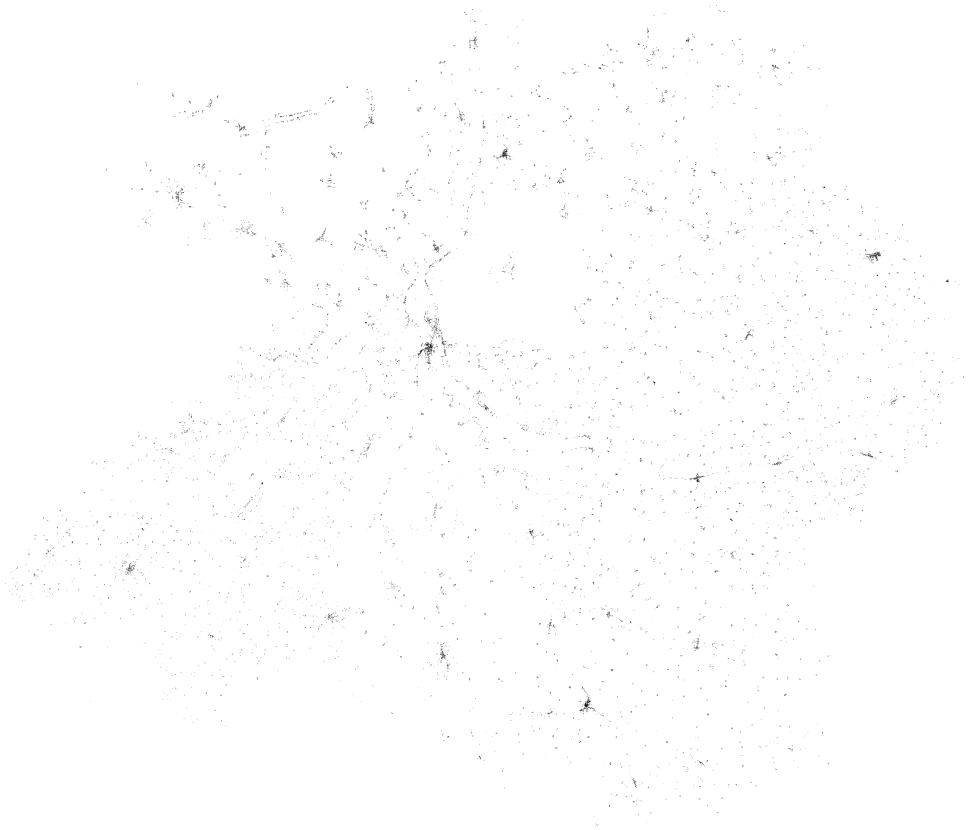
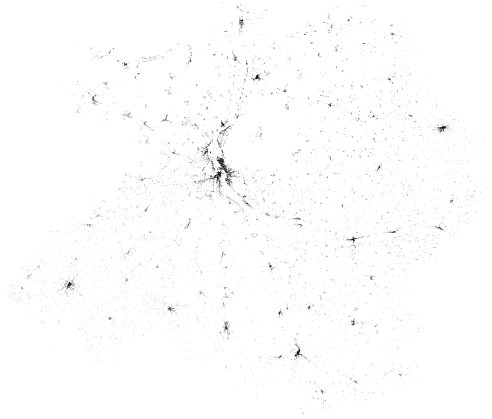
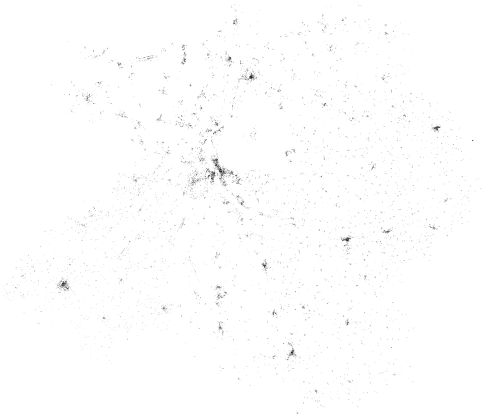


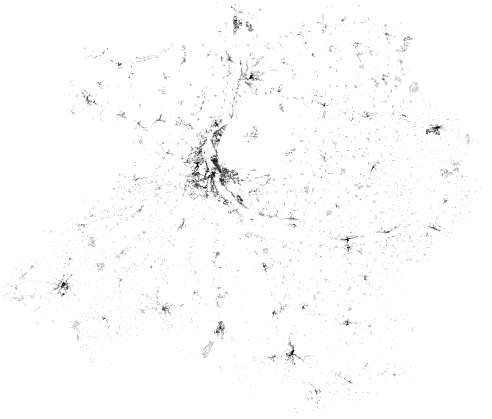
Figure 5.19: Montargis 1900



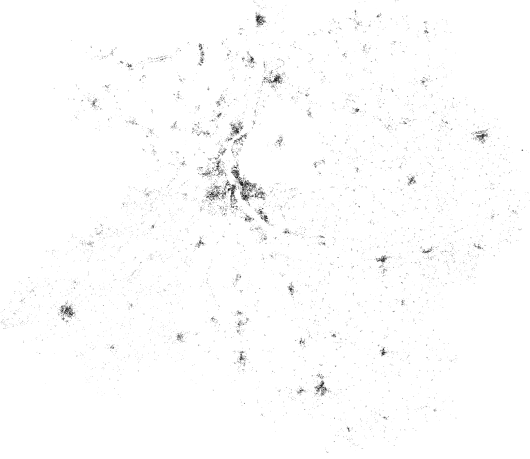
(a) Montargis 1949



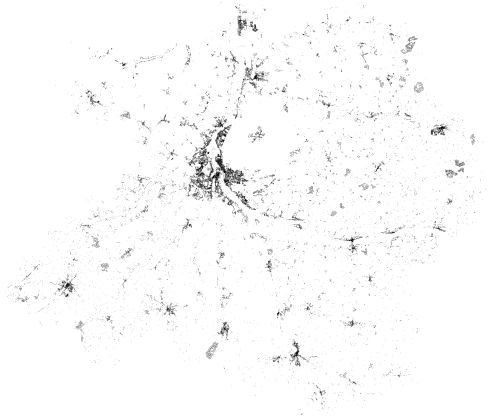
(b) Simulation



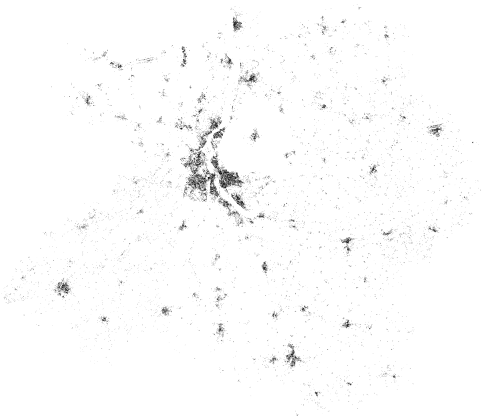
(c) Montargis 1979



(d) Simulation

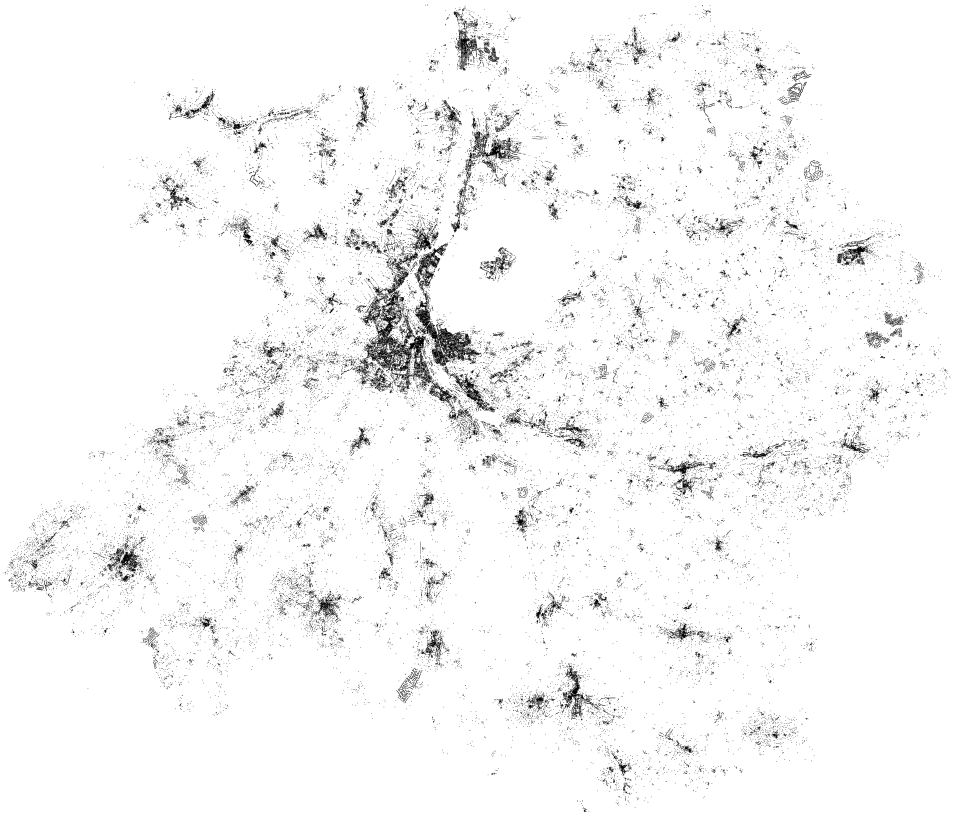


(e) Montargis 2007



(f) Simulation

Figure 5.20: The simulations et reality of Montargis.



(a)

Figure 5.21: The simulations of Montargis in future.

5.3.3.3 Comparisons

A comparison is performed by overlapping real Montargis 2008 on its simulation. There are some regions near the boundary where the simulation does not match speed of growth (slower) while in north region of center and near river the growth of simulation is more rapid.

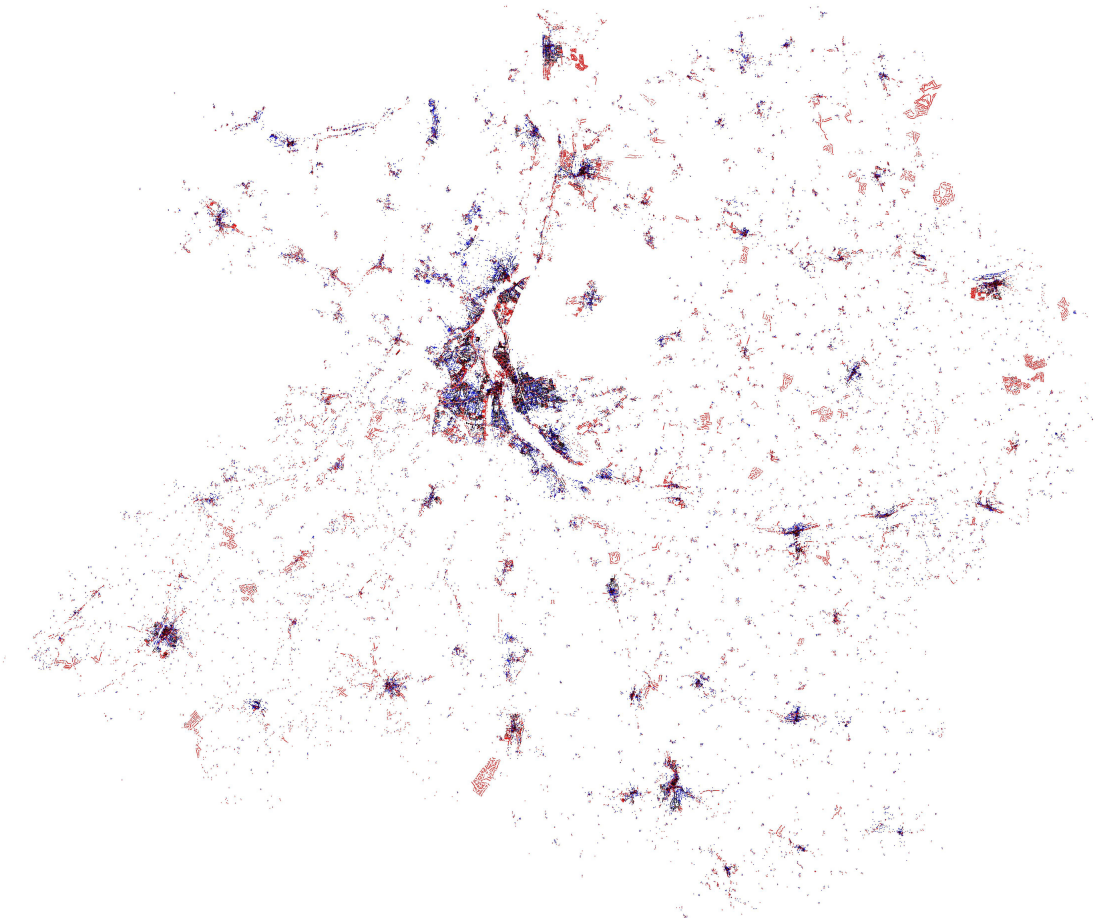


Figure 5.22: The overlap between of real (red) and simulation (blue) of Montargis 2008.

With the number of constructed plots in reality of each years from 1900 to 2008, we find the year having the nearest number of these with each simulation that we do. Then we use the evolution of the number of constructed plots to compare them by all of years from 1900 to 2008 (in the figure 5.23).

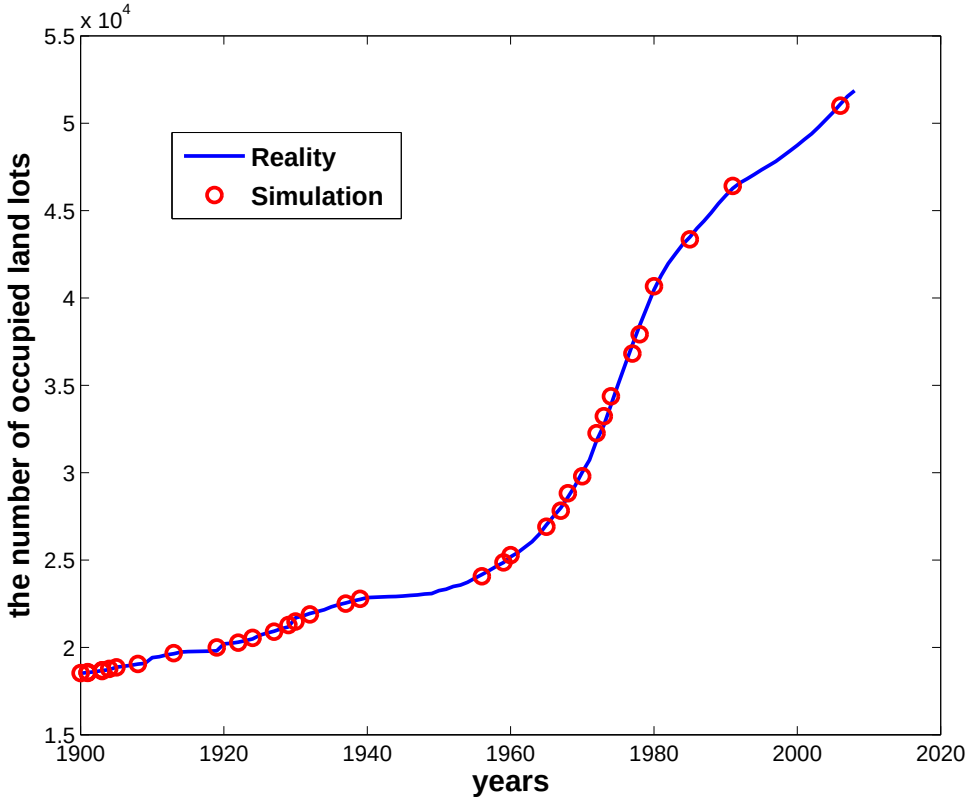


Figure 5.23: The evolution of the number of constructed plots of real Montargis and its simulations from 1900 to 2008.

5.3.3.4 The error of simulations

In additional to the above overlap, we give some numerical camparisons between Montargis 2007 and its simulation.
Let $f_r(z)$, $f_s(z)$ be the density functions of a plot z in real Montargis and its simulations respectively; C be a constant,

$$f_j(z) = \frac{1}{C} \sum_{z' \in B_j} e^{-\lambda_2 d(z,z')}, \quad j = \{r, s\}$$

with $C = \max_z \{f_r(z); f_s(z)\}$ and $\lambda_2 = 0.03$.
The mathematical errors of the simulations:

i. The maximal error of density:

$$\epsilon_{max} = \max_z |f_s(z) - f_r(z)| = 0.25.$$

ii. The distribution of errors

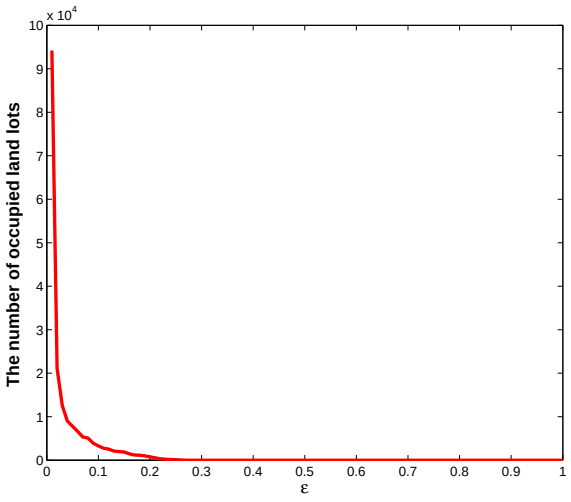


Figure 5.24: The distribution of errors.

iii. The mean error

$$\epsilon_{mean} = \frac{1}{n} \sum_z |f_s(z) - f_r(z)| = 0.031.$$

iv. The mean squared error

$$\epsilon_{MSE} = \frac{1}{n} \sum_z |f_s(z) - f_r(z)|^2 = 0.003.$$

v. The other statistics
- The number of superimposed points of reality and simulation: 30229,
- The number of points in reality but not in simulation: 20778,
- The number of points in simulation but not in reality: 21637.

Chapter 6

Conclusions and Perspectives

6.1 Conclusions

In this thesis we proposed new random models of simulation of city growth. These models were developed from known models using correlated gradient percolation with different modifications taking care of geographical constraints.
In the first model, with assumption of perpetuation of occupied points (or black points), density of buildings at current time is used as the probability function in occupancy decision of a vacant point in next simulation. From this model, we propose a backward one to model city in the past. However, this leads to mismatch between forward and backward model, as shown in chapter 5. According to our predictions, this is caused by unexpected effects of natural factors (such as the big forest near Montargis city and many others, the complex river system and its risk problem...) which were not considered.
In a refined model the simulations are carried out on land lots with presence of other elements in urban system such as attraction of city center, effect of rivers, transport system, forest and flood zone beyond density, with the same assumption in the previous model, but with building blocks land lots instead of simple pixels. We have compared simulations with real city in the end of chapter 5 and they give rather good results. A remaining problem of this model is that it cannot control time between two steps in simulations with reality, we start simulating at 2007 and we do not know the time that simulation of the next step corresponds to in reality. Moreover, in this model we fixed size of each land lot while in reality, large ones may be divided. We have not controlled this fact yet. However, because of division's expenditure, the owner of large land lot will divide it into many ones simultaneously when it is attracting. And then, we can reasonably assume that this land lot is occupied. One more problem is that road system is taken from data of 2008 for all simulations. But road system changes with time in a beyond our control way.
In chapter 4, we also carry out some simulations for Baltimore city with more simple model (consider density as a probability function with the presence of main road system). Due to the lack of geographical data, these simulations do not match with

real city but look similar, as if cities were random systems with given distribution.

In addition to these models, we study some mathematical aspects of gradient percolation concerning the boundary of clusters on triangular lattice in chapter 3. It is an extension of P. Nolin's results for exponential case. With probability function $p(r) = e^{-\lambda r}$ the boundary of the largest cluster tends to be localized near the curve with distance to center r_0 at critical percolation $e^{-\lambda r_0} = 1/2$. Its width is a function of density gradient, $\lambda^{-4/7}$. This result is numerically illustrated in chapter 4 where we study properties of correlated gradient percolation. In addition to this point, we also prove the uniqueness of the front and show that boundary's length tends to $\lambda^{-3/7} r_0$.

Another kind of growth process is studied in this thesis: we present a program to compute numerical results of coefficients of SLE (Stochastic Loewner evolution) and LLE (Lévy Loewner evolution) processes. These processes were invented by Oded Shramm who has had the idea of introducing randomness in Loewner theory of planar growth. In this thesis we go back to the original motivation of Loewner which was Bieberbach conjecture about coefficients of univalent functions. More precisely we study coefficient problems for SLE and LLE. We build a program allowing to compute L^2 norms of these coefficients which are random variables. These results provided the basis for conjecturing a general case (theorem 2.4.1) of expected values of their absolute values squared. We have discovered the remarkable fact (which has since found a rigorous proof) that for respectively $\kappa = 6$ and 2, these norms are respectively 1 and $\sqrt{n}$ for the n'th coefficient of whole-plane SLE.

6.2 Perspectives

Regarding future work there are still many interesting research questions to explore for further improving the growth of cities models.

- First of all, differentiel equation density model could be systematically improved by adding other factors of urban system into it; this could lead to a program that could be used by urbanists and decision-makers.
- With the second model, it would be nice to find the way to take into account the time factor.
- From the results of the growth of buildings, we have an ambitious development of modeling of road system's growth. For this purpose, it is necessary to understand growth laws of this system.
- A future work is to model the interactions between cities which may be considered as a network.
- Another work is solving problem of land lot's division. It requires researches and statistics in reality to know division's law.

In addition to these points, more theoretical problems arise for gradient percolation and coefficient problem.

- From estimation of length of front in the end of chapter 3 we can mathematically find fractal dimension of this front in gradient percolation case. One more question is that these results of the front remain on hold with other kinds of lattices beyond triangular lattice.
- For the coefficients of SLE, LLE process, finding the general form of their L^2 is one challenge. The computing program is stopping at order $n = 19$ because of long execute time. If the algorithm is improved, we may expect results of higher orders.

Appendix A

C Code for the coefficient problem

This section will present C code to calculate expectation of coefficients' absolute values squared of SLE process.

In the first part, we construct structure of a "term" with B contains values of (α, β) in integrals (2.3.1) and A contains coefficients front of B. These values is computed by Loewner method.

```
typedef struct {
    long nC;
    int C[16384][14];
    int B[16384][14];
} term;

double matrix[50][101];
```

Function mySLE uses Loewner method to compute term of a_n.

```
term mySLE(int n)
{
    int k, j;
    unsigned long i;
    term A, temp;
    temp.nC = 1;
    temp.C[0][0] = 1; temp.C[0][1] = -2;
    temp.B[0][0] = 1; temp.B[0][1] = 1;

    if(n==2) return(temp);

    for(k=3;k<=n;k++)
    {
        A.nC = (unsigned long)(pow(2,k-2));
        i = 0;
        while(i<A.nC)
        {
            A.C[i][0] = temp.C[(unsigned long)(floor(i/2))][0];
            A.C[i+1][0] = temp.C[(unsigned long)(floor(i/2))][0]+1;
            A.C[i+1][1] = -2*(k-1);
            A.B[i][0] = temp.B[(unsigned long)(floor(i/2))][0];
            A.B[i][1] = temp.B[(unsigned long)(floor(i/2))][1]+1;
            A.B[i+1][0] = temp.B[(unsigned long)(floor(i/2))][0]+1;
            A.B[i+1][1] = 1; A.B[i+1][2] = temp.B[(unsigned long)(floor(i/2))↩
                ][1];
```

```
            for(j=2;j<=temp.B[(unsigned long)(floor(i/2))][0];j++)
            {
                A.B[i][j] = temp.B[(unsigned long)(floor(i/2))][j];
                A.B[i+1][j+1] = temp.B[(unsigned long)(floor(i/2))][j];
            }
            for(j=1;j<=temp.C[(unsigned long)(floor(i/2))][0];j++)
            {
                A.C[i][j] = temp.C[(unsigned long)(floor(i/2))][j];
                A.C[i+1][j+1] = temp.C[(unsigned long)(floor(i/2))][j];
            }
            i = i+2;
        }
        temp = A;
    }
    return(A);
}
```

Function exp2 uses dynamic programing to calculate $\mathbb{E}(|a_n|^2)$.

```
double exp2(int n,float K)
{
    if(n==0){
        printf("\n Error input.\n");
        return(0);
    }
    if(n==1) return(1);

    term A;
    A = mySLE(n);

    unsigned long i,j;
    int ii,jj,m,l, ij, ji,Sm,Sl;
    double mA[n][n], S, t, tempS;

    S = 0;

    for(i=0;i<A.nC;i++){
        m = A.B[i][0];
        for(j=0;j<=i;j++){
            l = A.B[j][0];
            mA[m][l] = 1;

            Sm = 0;
            for(ii=m-1;ii>=0;ii--){
                Sm = Sm+A.B[i][ii+1];
                t = double(Sm+Sm*Sm*K/2);
                mA[ii][l] = mA[ii+1][l]/t;
                Sl=0;
                for(jj=l-1;jj>=0;jj--){
                    Sl = Sl+A.B[j][jj+1];
                    mA[m][jj] = mA[m][jj+1]/(double(Sl+Sl*Sl*K/2));
                    mA[ii][jj] = (mA[ii+1][jj]+mA[ii][jj+1])/(double(Sm+Sl+(Sm-Sl↩
                        )*(Sm-Sl)*K/2));
                }
            }
            tempS = mA[0][0];
            if(i==j) {
                for(ij=1;ij<=A.C[i][0];ij++)
                    tempS = (long double)(A.C[i][ij]*tempS*A.C[i][ij]);
                S = S + tempS;
            }
            else{
                for(ij=1;ij<=A.C[i][0];ij++)
                    tempS = tempS*(long double)(A.C[i][ij]);
```

```
                for(ji=1;ji<=A.C[j][0];ji++)
                    tempS = tempS*(long double)(A.C[j][ji]);
                S = S + 2*tempS;
            }
        }
    }
    return(S);
}
```

```
int main(void)
{
    int n, i, j, nK;

    nK = 101;
    n = 8;

    int K = 6;
    for(i=0;i<n;i++){
        matrix[i][0] = exp2(i+1,K);
    }
    return(0);
}
```

Appendix B

Matlab Code for simulation of the growth of cities

B.1 Model of density

```
function concentration(k,time,init)
%% function calculate concentration by time t
%% k: R_k is computed
%% time: mark simulate for future or past
%% init: if init = 1 create R0, init = 0 don't

MR = imread('Datas/bwRouteRivere.bmp'); % roads and midle of river are marked 0;
if init==1
    M0 = imread('Datas/bwBati.bmp');
    M0 = and(M0,MR);% take road, river in
    M = im2bw(1-M0);
    [nR nC] = size(M0);
    [X,Y] = meshgrid(-nR+1:nR-1,-nC+1:nC-1);
    D = sqrt(X.^2+Y.^2);
    pBP = exp(-sqrt(D));
    R0 = MyConv(M,pBP);
    dlmwrite('Simulations/R0.txt',R0);
end

if k==1
    if init==0
        M0 = imread('Datas/bwBati.bmp');
        R0 = dlmread('Simulations/R0.txt');
    end
else
    if time==1
        R0 = dlmread(['Simulations/RF',num2str(k-1),'.txt']);
        M0 = imread(['Simulations/IF',num2str(k-1),'.bmp']);
    else
        M0 = imread(['Simulations/IB',num2str(k-1),'.bmp']);
        R0 = dlmread(['Simulations/RB',num2str(k-1),'.txt']);
    end
end

L = M0; % matrice of buildings
R = R0;
M0 = and(M0,MR);% take road, river in
```

```
dt = 0.1;  % priod between 2 point of time
N = 5; % number of priod compute R
[nR nC] = size(M0);

[X,Y] = meshgrid(-nC+1:nC-1,-nR+1:nR-1);
D_2 = sqrt(X.^2+Y.^2);
pBP = exp(-sqrt(D_2));

if time==1
    MB = imread('Datas/bwLimite.bmp'); % matrix outside boundary marked 1
    MH = imread('Datas/bwAutoroutes.bmp'); % highway marked 0
    M0 = and(M0,MH); % get highway in marked 0
    M0 = M0-MB; % get out boundary
    M0(M0<0) = 0;
    M0 = im2bw(M0);
else
    M0 = 1-M0;
    R0 = 1-R0;
end
M = R0;

for i=1:N
    M = M0.*M;
    M = MyConv(M,pBP);
    R = R + time*dt^i*M;
end
R(R>1) = 1;
R(R<0) = 0;

if time==1
    dlmwrite(['Simulations/RF',num2str(k),'.txt'],R);
else
    dlmwrite(['Simulations/RB',num2str(k),'.txt'],R);
end

x0 = 3109; y0 = 4018;
alpha = 0.005;
[x,y] = meshgrid(1-y0:nC-y0,-x0+1:nR-x0);
D_1 = sqrt(x.^2+y.^2);
MCo = (D_1.^2+1).^(-alpha/2);
v = FFM2D(nR,nC,MCo,[x0,y0]); % take random and correlation
t = reshape(v,1,[]);
[mu,sigma] = normfit(t);

if time==1
    for i=1:nR
        for j=1:nC
            if M0(i,j)==1
                theta = norminv((R(i,j)),mu,sigma);
                if v(i,j)<theta
                    L(i,j) = 0;
                end
            end
        end
    end
    name = ['Simulations/IF',num2str(k),'.bmp'];
else
    for i=1:nR
        for j=1:nC
            if M0(i,j)==1
                theta = norminv(1-(R(i,j)),mu,sigma);
                if v(i,j)<theta
                    L(i,j) = 1;
                end
```

```
            end
        end
    end
    name = ['Simulations/IB',num2str(k),'.bmp'];
end
```

```
function growth(index)
%% simulating for Montargis using my model

%% 1.Input
lambda_1 = 0.02/sqrt(index); % Bati and Rivers
lambda_2 = 0.05; % black point
lambda_3 = 0.01/sqrt(index); % Noeuds

if index==1
    M0 = imread('Datas\bwBati.bmp');
else
    M0 = imread(['Simulations\S',num2str(index-1),'.bmp']);
end
River = 1-imread('Datas\bwRouteRivere.bmp');
M1 = 1-and(M0,River);  % matrix with Rivers as population

[nR,nC] = size(M0);
x0 = 3137; y0 = 4060; % center that computed from findCenter.m

%% 2.Finding delta function (z) = e^-(|z-z0|^x) (x<1) and Fourier transform
%% of it

[x,y] = meshgrid(-y0+1:nC-y0,-x0+1:nR-x0);
[X,Y] = meshgrid(-nC+1:nC-1,-nR+1:nR-1);

D_1 = sqrt(x.^2+y.^2);
D_2 = sqrt(X.^2+Y.^2);
p = exp(-D_1*lambda_1); % distance to center, roads and rivers
pBP = exp(-D_2*lambda_2); % black points

convBP = MyConv(M1,pBP);

tN = dlmread('Datas\timeDistNoeuds.txt'); % time of Noeuds the nearest
pNoeud = exp(-lambda_3*tN);

%% 3.Finding probability of population R0 = e^(-lambda_1*d_0)*sum(e^(-lambda_2*↩
    d_1))*sum(e^(-lambda_3*d_2))

R0 = convBP.*pNoeud.*p;
Max = max(max(R0));
R0 = R0./Max;

%% 4.Create matrix correlation with center of lattice

alpha = 0.005;
S1 = (D_2+1).^(-alpha/2);
S = abs(fft2(S1));

%% 5.Fourier transform to find mu & sigma
u = randn(2*nR-1,2*nC-1);
u1 = fft2(u);
v1 = sqrt(S).*u1;
v = real(ifft2(v1));
v = v(nR-x0+1:2*nR-x0,nC-y0+1:2*nC-y0);
t = reshape(v,1,[]);
[mu,sigma] = normfit(t);
```

```
%% Compute R by time t
R = R0;

%% 6.Main simulation for Baltimore city
% L = ones(nR,nC);
MH = imread('Datas\bwAutoroutes.bmp'); % Matrix contains Boundary, Highway, and ↩
    River marked 0.
ML = imread('Datas\bwLimite.bmp');  % Matrix contains Roads marked 0.
MF = imread('Datas\bwForest.bmp');
MI = imread('Datas\bwInondable.bmp');
L = M0;
C = 0.03;
for i=1:nR
    for j=1:nC
        if M0(i,j)==1 && MH(i,j)==1 && ML(i,j)==0
            if MF(i,j)==0 || MI(i,j)==0
                R(i,j) = max(R(i,j)-C,0);
            end
            theta = norminv(R(i,j),mu,sigma);
            if v(i,j)<theta
                L(i,j) = 0;
            end
        end
    end
end
imwrite(L,['Simulations\S',num2str(index),'.bmp']);
```

```
function R = MyConv(A,B)
%% function compute convolution of 2 matrix use FFT2
%% dimension the same with A

[rA,cA] = size(A);
[rB,cB] = size(B);

nr = rA+rB-1; nc = cA+cB-1;

r0 = ceil((nr-rA)/2);
c0 = ceil((nc-cA)/2);

N = pow2(nextpow2(nr));
M = pow2(nextpow2(nc));

fftw('planner', 'estimate');
fA = fft2(A,N,M);
fB = fft2(B,N,M);
fC = fA.*fB;
C = ifft2(fC,N,M);
R = C(r0+1:rA+r0,c0+1:cA+c0);
```

```
function v = FFM2D(N,M,MCo,mid)

u = randn(N,M);

u1 = fft2(u);

S = fft2(MCo);

v1 = sqrt(S).*u1;
v = ifft2(v1);
v = circshift(v,-mid);
```

B.2 Model on land lots

```
function urbanParcelle(index)
%% simulating for Montargis using my model on land lots
%% 1.Input
if index>0
    lambda_1 = 0.0001*(index^(-1/4)); % Center
    lambda_2 = 0.025; % black point
    lambda_3 = 0.0001*(index^(-1/4)); % Noeuds
    lambda_4 = 0.0007*(index^(-1/4)); % roads and river
    Ci = 0.3; % parametre for Inondable & Foret
else
    lambda_1 = 0.0001*sqrt(-index); % Center
    lambda_2 = 0.03; % black point
    lambda_3 = 0.0001*sqrt(-index); % Noeuds
    lambda_4 = 0.001*sqrt(-index); % roads and river
end

if index>0
    Max = 150000*exp(-sqrt(index));
end

if index<0
    if index==-1
        M0 = xlsread('Data\mParCen2008.xlsx');
    else
        M0 = xlsread(['Simulations\simNR',num2str(index+1),'.xlsx']);
    end
else
    if index==1
        M0 = xlsread('Data\mParCen1900.xlsx'); % matrix building on Parcelles (↩
            nx3)
    else
        M0 = xlsread(['Simulations\simNR',num2str(index-1),'.xlsx']);
    end
end

n = size(M0,1);
x0 = 3115; y0 = 4048; % center that computed from findCenter.m

%% 2.Finding delta function (z) = e^-(|z-z0|^x) (x<1) and Fourier transform
%% of it
D_1 = sqrt((M0(:,1)-x0).^2+(M0(:,2)-y0).^2); % distance to centre nx1
p = exp(-D_1*lambda_1);

convBP = zeros(n,1);  % correlation with black parcelles nx1
M_B = M0(M0(:,3)==1,1:2);
for i=1:n
    D_2 = sqrt((M_B(:,1)-M0(i,1)).^2 + (M_B(:,2)-M0(i,2)).^2);
    convBP(i) = sum(exp(-lambda_2*D_2));
end
tN = dlmread('Data\timeDistNoeudsPar.txt'); % time of Noeuds the nearest
pNoeud = exp(-lambda_3*tN); % nx1

dR = dlmread('Data\DistRoadRiver.txt'); % nx1
dR = xlsread('Data\DistRiver.xlsx');
pR = exp(-lambda_4*dR);

%% 3.Finding probability R0 = e^(-lambda_1*d_0)*sum(e^(-lambda_2*d_1))*sum(e^(-↩
    lambda_3*d_2))*sum(e^(-lambda_4*d_3))
R0 = convBP.*pNoeud.*pR.*p; % an array length n
if index<0
    Max = max(R0);
```

```
end
R0 = R0/Max;

%% 4.Create matrix correlation with center of lattice
% Follow the condition of correlation, matric MCo must be symetric and
% periodic (padding 0 for 2*N-1) but it's impossible beacause of error "Out of ↩
    memory"
alpha = 0.001;
MCo = (D_1.^2+1).^(-alpha/2);

%% 5.Fourier transform to find mu & sigma
u = randn(n,1);
u1 = fft(u);
S = abs(fft(MCo));

v1 = sqrt(S).*u1;
v = real(ifft(v1));
[mu,sigma] = normfit(v);

R = R0;

%% 6.Main simulation for Montargis
if index<0
    if index==-1
        MB1 = imread('Data\bwParBati.bmp'); % matrix bati on parcelles
    else
        MB1 = imread(['Simulations\simNR',num2str(index+1),'.bmp']);
    end
else
    if index==1
        MB1 = imread('Data\bwBati1900.bmp'); % matrix bati on parcelles
    else
        MB1 = imread(['Simulations\simNR',num2str(index-1),'.bmp']);
    end
end

L = M0;
    for i=1:n
        if M0(i,3)~=1 && M0(i,3)~=2 % black point & no building
            if M0(i,3)==3 || M0(i,3)==4 % Inondable & foret
                R(i) = max(R(i)-Ci,0);
            end
            theta = norminv(R(i),mu,sigma);
            if v(i)<theta
                L(i,3) = 1;
                MB1(M0(i,1)-2:M0(i,1)+2,M0(i,2)-2:M0(i,2)+2) = 0;
            end
        end
    end
imwrite(MB1,['Simulations\simNR',num2str(index),'.bmp']);
xlswrite(['Simulations\simNR',num2str(index),'.xlsx'],L);
```

Bibliography

[AFTY08] Jean-Philippe Antoni, Pierre Frankhauser, Cécile Tannier, and Samy Youssoufi. Modelling urbanization to simulate prospective scenarios: a comparative approach. In *International Conference of Territorial Intelligence, Huelva 2007. Papers on territorial intelligence and governance, participative action-research and territorial development*, pages 277–295, Huelva, Espagne, 2008.

[Bad05a] Dominique Badariotti. Autosimilarité et autoréférence en urbanisme. Quelques exemples ... In 13èmes journées de Rochebrune. Comité scientifique., editor, *Traces, Enigmes, problèmes : émergence et construction du sens.*, Rencontres interdisciplinaires sur les systèmes complexes naturels et artificiels, pages 1–9. Ecole nationale des télécommunications, 2005.

[Bad05b] Dominique Badariotti. Des fractales pour l'urbanisme? *Cahiers de géographie du Québec*, 49(137), 2005.

[Bat09] M. Batty. Urban modeling. In N. Thrift and R. Kitchin, editors, *International Encyclopedia of Human Geography*, pages 51–58. Elsevier, Oxford, 2009.

[BBM07] Dominique Badariotti, Arnaud Banos, and Diego Moreno. Conception d'un automate cellulaire non stationnaire à base de graphe pour modéliser la structure spatiale urbaine: le modèle Remus. *Cybergeo: European Journal of Geography*, (2), 2007.

[Bha10] Basudeb Bhatta. *Analysis of Urban Growth and Sprawl from Remote Sensing Data*. Springer, 2010.

[BL86] M. Batty and P. Longley. The fractal simulation of urban structure. *Environment and Planning A*, 18(9), 1986.

[BL94] M. Batty and P. Longley. *Fractal Cities*. Academic Press, San Diego, 1994.

[Bur25] Ernest W. Burgess. The growth of the city: an introduction to a research project. In *The Trend of Population*, pages 85–97. American Sociological Society, 1925.

[BW02] Dominique Badariotti and Christiane Weber. La mobilité résidentielle en ville. Modélisation par automates cellulaires et système multi-agents à Bogota. *L'Espace Géographique*, (403), 2002.

[Chr33] Walter Christaller. *Central Place in Southern Germany.* Englewood Cliffs, N.J. : Prentice-Hall, english edition 1966 edition, 1933.

[Cla51] C. Clark. Urban population densities. *Journal of the Royal Statistical Society*, 114(4):490–496, 1951.

[CVC+11] Geoffrey Caruso, Gilles Vuidel, Jean Cavailhès, Pierre Frankhauser, Dominique Peeters, and Isabelle Thomas. Morphological similarities between DBM and a microeconomic model of sprawl. *Journal of Geographical Systems*, 13(1):31–48, 2011.

[DNNZ12] Bertrand Duplantier, Thi Phuong Chi Nguyen, Thi Thuy Nga Nguyen, and Michel Zinsmeister. The coefficient problem and multifractality of whole-plane SLE and LLE. *available at arXiv:1211.2451 [math-ph]*, 2012.

[FP10] Pierre Frankhauser and Denise Pumain. *Fractals and Geography*, pages 281–311. ISTE, 2010.

[Fra91] P. Frankhauser. Aspects fractals des structures urbaines. *Espace géographique*, 19(1), 1991.

[Fra08] Pierre Frankhauser. Fractal geometry for measuring and modelling urban patterns. In Sergio Albeverio, Denise Andrey, Paolo Giordano, and Alberto Vancheri, editors, *The Dynamics of Complex Urban Systems*, pages 213–243. Physica-Verlag HD, 2008.

[Gri99] G. Grimmett. *Percolation.* Springer, New York, 2nd edition, 1999.

[Hoy39] Homer Hoyt. *The structure and growth of residential neighborhoods in American cities.* Federal Housing Administration Wash, 1939.

[HU45] C. D. Harris and E. L. Ullman. The nature of cities. *Annals of the American Academy of Political and Social Science*, 242, 1945.

[Ken] Tom Kennedy. Pictures of Chordal SLE. `http://math.arizona.edu/~tgk/`.

[Kes80] Harry Kesten. The critical probability of bond percolation on the square lattice equals 1/2. *Comm. Math. Phys.*, 74:41–59, 1980.

[Kes82] Harry Kesten. *Percolation Theory for Mathematicians*. Birkhaüser, 1982.

[Kes87] Harry Kesten. Scaling relation for 2d-percolation. *Comm. Math. Phys.*, 109:109–156, 1987.

[Los43] August Losch. *The Economics of Location*. New Haven,: Yale University Press, english edition 1954 edition, 1943.

[MAJB+98] H. A. Makse, J. S. Andrade Jr., M. Batty, S. Havlin, and H. E. Stanley. Modeling urban growth patterns with correlated percolation. *Phys. Rev. E*, 58(6), 1998.

[Man82] B. Mandelbrot. *The Fractal Geometry of Nature*. W. H. Freeman and Company, New York, 1982.

[MHSS96] H. A. Makse, S. Havlin, M. Schwartz, and H. E. Stanley. Method for generating long-range correlations for large systems. *Phys. Rev. E*, 53(5), 1996.

[NAS96] Goddard Space Flight Center NASA. Scientific Visualization Studio. `http://svs.gsfc.nasa.gov/vis/a000000/a000000/a000099/index.html`, Avril 1996. Video.

[Nol08a] Pierre Nolin. Critical exponents of planar gradient percolation. *Ann. Probab.*, 36:031131, 2008.

[Nol08b] Pierre Nolin. *Near-critical percolation in two dimensions, and related models*. PhD thesis, Université Paris-Sud, 2008.

[PMAS07] M. Pooyandeh, S. Mesgari, A. Alimohammadi, and R. Shad. A comparison between complexity and temporal GIS models for spatio-temporal urban applications. In O. Gervasi and M. Gavrilova, editors, *ICCSA 2007, LNCS 4706, Part II*, pages 308–321. Springer, 2007.

[Rei65] R. Reif. *Fundamentals of Statistical and Thermal Physics*. McGraw-Hill, New York, June, 1965.

[Sch00] Oded Schramm. Scaling limits of loop-erased random walks and uniform spanning trees. *Israel Journal of Mathematics*, 118(1):221–288, 2000.

[Smi01] Stanislav Smirnov. Critical percolation in the plane: conformal invariance, Cardy's formula, scaling limits. *C. R. Acad. Sci. Paris, Ser. I Math.*, 333:239–244, 2001.

[SRG85] B. Sapoval, M. Rosso, and J.-F. Gouyet. The fractal nature of a diffusion front and the relation to percolation. *Phys. Rev. Lett.*, 46, 1985.

[SW01] Stanislav Smirnov and Wendelin Werner. Critical exponents for two-dimensional percolation. *Math. Res. Lett.*, 8:729–744, 2001.

[vT26] Johann Heinrich von Thunen. *The Isolated State*. Oxford: Pergammon Press, english edition 1966 edition, 1826.

[Web09] Alfred Weber. *Theory of the Location of Industries*. IL: University of Chicago Press, Chicago, 1909.

[Weg11] M. Wegener. From macro to micro - How much micro is too much? *Transport Reviews*, 31(2), 2011.

[Zin07] Michel Zinsmeister. Modeling of phenomena in natural sciences. `http://www.univ-orleans.fr/MAPMO/membres/zins/Vietnam/PUF/giangvien2007-2008-french.htm`, December 2007. Lecture.

List of Figures

Thi Thuy Nga NGUYEN

MODÉLISATION DE LA CROISSANCE DES VILLES

Résumé :

Dans cette thèse nous proposons and nous mettons en application plusieurs modèles décrivant le croissance et la morphologie du tissu urban. Le premier de ces modèles est issu de la percolation en gradient (correlée) déjà proposé de la littérature. Le second, inédit, fait appel à un équation différentielle stochastique. Nos modèles sont paramétrisables: les paramètres que nous avons choisi d'appliquer sont naturels et tiennent compte de l'accessibilité des sites. Le résultat des simulations est conforme à la réalité du terrain. Par ailleurs, nous étudions la percolation en gradient: nous démontrons , suivant Nolin, que la frontière de cluster principal se situe dans un voisinage de la courbe critique et nous estimons ses longueurs et largeurs. Enfin, nous menons une étude du processus de croissance SLE. Nous calculons (preuve assistée par ordinateur) l'espérance des carrés des modules pour SLE_2 and SLE_6. Ces résultats sont liés à la conjecture de Bieberbach.

Keywords : modélisation, croissance des villes, percolation en gradient, percolation critique, SLE .

SIMULATION OF THE GROWTH OF CITIES

Abstract :

In this thesis we propose and test models that describe the growth and morphology of cities. The first of these models is used from previously developed correlated gradient percolation model. The second model is related to a stochastic differential equation and has never been proposed before. Both models are parameterizable. The parameters we chose in applications are well justified by physical observations: proximily to axes and accessibility of sites. The result is consistent with actual data. We also study the gradient percolation as a mathematical object. We prove, following Nolin's ideas, that the front of gradient percolation cluster is localised in a neighborhood of the critical curve with width and length depending on density gradient. Finally, we also study SLE growth processes. We calculate (computer assisted demonstration) the expected value of square of moduli for SLE_2 and SLE_6 related to the Bieberbach conjecture.

Keywords : modeling, urban growth, gradient percolation, critical percolation, SLE .

MAPMO UMR 6628 - CNRS
Rue de Chartres - BP6759
45067 ORLÉANS CEDEX