

Generating Functions of Stochastic L-Systems and Application to Models of Plant Development

Cedric Loi, Paul Henry Cournède

▶ **To cite this version:**

Cedric Loi, Paul Henry Cournède. Generating Functions of Stochastic L-Systems and Application to Models of Plant Development. Fifth Colloquium on Mathematics and Computer Science, 2008, Kiel, Germany. pp.325-338, 10.46298/dmtcs.3574 . hal-01194661

HAL Id: hal-01194661

https://inria.hal.science/hal-01194661

Submitted on 7 Sep 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Fifth Colloquium on Mathematics and Computer Science DMTCS proc. **AI**, 2008, 325–338

Generating Functions of Stochastic L-Systems and Application to Models of Plant Development

Cédric Loi[1,2], Paul Henry Cournède[1,2]

[1]*Ecole Centrale Paris, Laboratory of Applied Mathematics and Systems, 92290 Châtenay Malabry, France*
[2]*INRIA Saclay - Île-de-France, DIGIPLANTE, 91400 Orsay, France*

If the interest of stochastic L-systems for plant growth simulation and visualization is broadly acknowledged, their full mathematical potential has not been taken advantage of. In this article, we show how to link stochastic L-systems to multitype branching processes, in order to characterize the probability distributions and moments of the numbers of organs in plant structure. Plant architectural development can be seen as the combination of two subprocesses driving the bud population dynamics, branching and differentiation. By writing the stochastic L-system associated to each subprocess, we get the generating function associated to the whole system by compounding the associated generating functions. The modelling of stochastic branching is classical, but to model differentiation, we introduce a new framework based on multivariate phase-type random vectors.

Keywords: Stochastic L-system, multitype branching process, phase-type, generating functions, plant development, plant growth models, GreenLab

1 Introduction

Models of plant development (or organogenesis) describe the dynamic creation of organs (leaves, internodes, flowers/fruits) and how they arrange to form plant structures. When the smallest scale of interest is that of organs (and not cells), discrete models are generally used to simulate plant structural development. The parallel rewriting grammar introduced by Lindenmayer (1968) (called L-system) is particularly adapted to model the evolution of branching patterns and its algorithmic power has been broadly taken advantage of since Smith (1984). The stochastic version of this type of grammar gives interesting results from simulation and graphical points of view by increasing the realistic aspect of geometric plants, cf. Prusinkiewicz and Lindenmayer (1990). However, when the aim of development models is not purely geometric but a basis for the study of plant functioning (biomass production and repartition), there is a need to bypass the full simulation of plant structure and to write properly the dynamic equations of development, in order to compute the growth biophysical processes, cf. De Reffye et al. (2008). A solution based on grammar factorization was proposed by de Reffye et al. (2003) and Cournède et al. (2006). It

1365–8050 © 2008 Discrete Mathematics and Theoretical Computer Science (DMTCS), Nancy, France

was also derived in the stochastic case in Kang et al. (2007) and the formalism of stochastic grammars was taken advantage of to compute the generating functions of simple branching plant structures.

The objective of this article is to explore the relationship between stochastic L-Systems and multitype branching processes in order to compute the corresponding probability generating functions and the distribution moments of the number of organs. The formalism is applied to a stochastic model of plant development (GreenLab). The article starts with the presentation of the basic botanical concepts and explains how a stochastic grammar can be derived. The theoretical framework is then introduced and we show how to derive generating functions from stochastic L-systems. It is applied to the two main botanical processes underlying plant development, branching and differentiation sequence of the meristem, which are first studied separately and then composed to form the complete stochastic model. The modelling of stochastic branching is classical but, to model differentiation, we introduce a new framework based on multivariate phase-type random vectors. Finally, the recursive relationship to compute the generating functions of plant structures can be introduced and applied to obtain the moments of the numbers of organs.

2 Botanical concepts to model plant development

As explained in Barthélémy and Caraglio (2007), organogenesis results from the functioning of undifferentiated cells constituting the apical meristem and located at the tip of axes. When in active phase, this meristem forms buds that will develop into new growth units composed of one or several metamers (also called phytomers). A metamer is a botanical entity chosen as the elementary scale to model plant architectural development in this study. It is composed of an internode bearing organs: axillary buds, leaves, flowers. Depending on species, metamers are set in place rhythmically - plants grow by successive shoots of several metamers - or continuously - when meristems keep on functioning and generate metamers one by one. In both cases, the time between the appearances of new shoots defines the architectural *Growth Cycle*. A *Growth Unit* is the set of metamers built by a bud during a growth cycle (only one in the continuous case). These metamers can be of different kinds and ordered according to botanical rules, like acrotony. For example, most temperate trees grow rhythmically, new shoots appearing at spring. If we do not consider polycyclism and neoformation, the architectural growth cycle corresponds to one year.

In this work, we do not consider time scales that are smaller than the architectural growth cycle and we study the development of new growth units as a discrete process. The *Chronological Age* (CA) of a plant (or of an organ) is defined as the number of growth cycles it has existed for.

Since metamers may bear axillary buds, plant architecture develops into a hierarchical branching system. Barthélémy and Caraglio (2007) underlined that architectural units can be grouped into categories characterized by a particular combination of morphological parameters. Thus, the concept of *Physiological Age* (PA) was introduced to represent the different types of growth units and axes. For instance, on coffee trees, there are two types: orthotropic trunk and plagiotropic branches. The main trunk's PA is equal to 1 and the oldest PA denoted by P corresponds to the ultimate state of differentiation for an axis, it is usually short, without branches. We need less than 5 PAs to describe the axis typology of most trees. The apical meristem or bud of an axis is thus characterized by the PA of the growth unit that it may produce and a metamer is characterized by its PA i (which is the PA of the growth unit that it belongs to). Moreover, along an axis, the morphological features of the growth unit may evolve with the age of the apical meristem. This process is described as the meristem sequence of differentiation by Barthélémy and Caraglio (2007), and corresponds to a transition to a superior PA of the meristem.

Since the number of potential axillary buds per metamer can be considered as a fixed botanical data, it is straightforward to deduce the whole plant structure from the population of buds. Moreover, we will often identify the meristem with the bud that it forms, for the sake of simplicity. Therefore, in the following, modelling plant development is equivalent to studying the dynamic evolution of the population of buds. It is mainly driven by the two botanical processes described above: branching resulting from the appearance of lateral buds in growth units and the differentiation sequence of meristems resulting in the change of PAs for terminal buds. As a consequence, at growth cycle n, a bud is characterized by 3 indices: its PA ϕ, the CA k of the axis (which is also the CA of the meristem) and the initial PA of the meristem β. It will be denoted by $b_{n,\phi}^{k,\beta}$. In the sequel, the two botanical processes will be referred to as *branching* and *differentiation*.

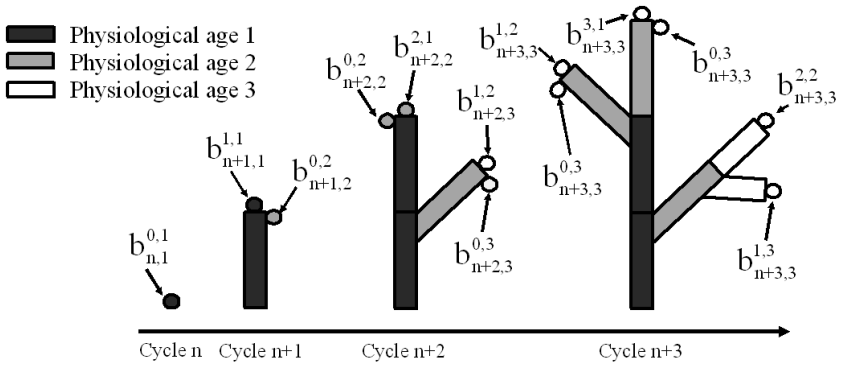


Fig. 1: Example of deterministic development for a bud during 3 cycles. A bud of PA 1 gives a metamer of PA 1 bearing a lateral bud of PA 2 and a terminal bud of PA 1. A bud of PA 2 gives a metamer of PA 2 bearing a lateral bud of PA 3 and a terminal bud of PA 2. A bud of PA 3 gives a metamer of PA 3 and a terminal bud of PA 3. After two cycles, the terminal bud of an axis of PA 1 differentiates and its new PA is 2. After one cycle, the terminal bud of an axis of PA 2 differentiates and its new PA is 3.

The productions of buds deeply vary according to their positions, environmental conditions, plant age...The deterministic modelling of such phenomena is difficult. For this reason, the development processes are considered as stochastic. For the sake of clarity, we only consider in this study growth units with one metamer. However, the same principles would apply for complex growth units with a random number of metamers as described in Kang et al. (2008). The computations of the general case are given in Loi and Cournède (2008). Experimental observations on plant populations and statistical analyses allow the estimation of the parameters of stochastic organogenesis models, cf. de Reffye (1979).

- P_l: bud survival probability. At each growth cycle, a bud may stay alive with probability P_l or die with probability $(1 - P_l)$. It may depend on bud's PA and will thus be denoted by $P_l(i)$ for $1 \leq i \leq P$.
- P_a: bud activity probability. At each growth cycle, if a bud is alive, it may stay dormant with probability $(1 - P_a)$ or produce a new growth unit with probability P_a. P_a may also depend on bud's PA: $P_a(i)$ for $1 \leq i \leq P$.
- $P_{i,j}^b(k)$: production probabilities. If at a given growth cycle, a bud of PA i is active, $P_{i,j}^b(k)$ is the probability that the growth unit it develops into bears k axillary buds of PA j. Botanical constraints usually impose that $0 \leq k \leq B_{i,j}^{max}$.

• λ_i: inverses of average occupation times. During its sequence of differentiation, a meristem stays of PA i for an average period of $1/\lambda_i$.

• $q_{i,j}$: transition probabilities. When a meristem of PA i changes its PA, $q_{i,j}$ is the probability that its new PA equals j. Note that the botanical differentiation sequence imposes that $q_{i,j} = 0$ if $j \leq i$.

In computational models, plants are generally represented as words in a language based on a generative parallel rewriting grammar also called L-system (Lindenmayer (1968), Smith (1984), Prusinkiewicz and Lindenmayer (1990), Françon (1990)). It is particularly adapted to describe the production of buds at discrete time steps. The following section will introduce a probabilistic framework to analyze stochastic models of plant development, based on stochastic L-systems.

3 A probabilistic framework based on stochastic L-systems

In this section, $V = \{v_1, v_2, \cdots, v_m\}$ denotes an alphabet, W the set of all words over V and W^+ the set of nonempty words over V. Let $\mathbf{1}$ be the empty word and "." be the concatenation operator. Then, $(W, ., \mathbf{1})$ is a noncommutative monoid.

3.1 Markov kernel and stochastic L-system

We first recall some basic definitions and properties of the Markov chain theory (Stroock (2005)). Throughout the following all sets are finite or countable.

Definition 3.1 (Transition matrix) *A transition matrix (or kernel) from A to B is a map $k : A \times B \to \mathbb{R}$ such that $k(a, b) \geq 0$ and $\sum_{b \in B} k(a, b) = 1$ for all $a \in A$. A transition matrix or Markov kernel on A is a transition matrix from A to A.*

Definition 3.2 (Markov chain) *A Markov chain with state space S is a stochastic process $(X_n)_{n \in \mathbb{N}}$ on some probability space $(\Omega, \mathcal{F}, P)$, such that $P(X_{n+1} = y \mid X_n = x) = p_{x,y}$ for all $x, y \in S$. $P = (p_{x,y})$ is a Markov kernel on S.*

Theorem 3.1 *For every tupel (P, μ) of Markov kernel P on S and probability measure μ on S there exists a Markov Chain $(X_n)_{n \in \mathbb{N}}$ of kernel P and starting measure $\mu = P(X_0 \in \cdot)$.*

We propose beneath a definition for stochastic 0L-systems on V based on the notion of Markov kernel.

Definition 3.3 (Stochastic 0L-system) *A stochastic 0L-system is a construct $G = \langle \omega_a, \pi \rangle$ where:*

- *$\omega_a \in W^+$ is called an axiom. It represents the structure initiating the growth.*
- *π is a transition matrix from V to W.*

Remark: Definition 3.3 is equivalent to the one given by Prusinkiewicz and Lindenmayer (1990): the set $P_r = \{(x, y) \in V \times W \mid \pi_{x,y} > 0\}$ defines the production rules. It represents the possible evolutions for all letters throughout the L-system.

We can now define a more general class of L-systems called stochastic F0L-system, extending the classical definition of F0L-system (Rozenberg and Salomaa (1980), p. 89) to the stochastic case:

Definition 3.4 (Stochastic F0L-system) *A stochastic F0L-system is a construct $G = \langle A, \pi \rangle$ where:*

- *A is a finite nonempty subset of V (called the set of axioms of G).*
- *for every $\omega_a \in A$, $G[\omega_a] = \langle \omega_a, \pi \rangle$ is a stochastic 0L-system (called component system of G).*

Proposition 3.1 *Let $P = (p_{x,y})$ be the square matrix on W such that, for all $x = x_1x_2\cdots x_n \in W$ with $(x_1, x_2, \cdots, x_n) \in V^n$ and for all $y \in W$:*

$$p_{x,y} = \sum_{\substack{(y_1,y_2,\cdots,y_n)\in W^n,\\ y_1y_2\cdots y_n=y}} \prod_{i=1}^{n} \pi_{x_i,y_i}$$

Then, P is a Markov kernel on W. For a stochastic F0L-system $\langle A, \pi\rangle$, P is called the associated Markov kernel.

Remark: To every component system $G[\omega_a] = \langle \omega_a, \pi\rangle$ we associate a Markov chain $(F_n[\omega_a])_{n\in\mathbb{N}}$ via the associated kernel P of G and the starting measure $\mu = \delta_{\omega_a}$ where δ_{ω_a} is the Dirac measure concentrated on ω_a.

Beside plant topological structure directly deduced from L-Systems, in order to compute plant functioning, the numbers of organs are crucial variables (see for example De Reffye et al. (2008)). To determine them, the order of symbols in words does not play any role, and we can consider the L-systems as commutative. Let R be an equivalence relation on W defined as follows: $w_1Rw_2 \Leftrightarrow$ there exists Π, a permutation on the symbol ranks, such that $\Pi(w_1) = w_2$. Let us denote the quotient set W/R by W^*. From now on, each word $w \in W$ will be assimilated to the ordered representative w^* of its equivalence class (*i.e.* wRw^* and $w^* = \prod_{i=1}^{m} {v_i}^{\alpha_i}$ for $(\alpha_1, \cdots, \alpha_m) \in \mathbb{N}^m$). W^* is isomorphic to $\mathbb{N}^m$. Let Υ be the canonical isomorphism from W^* into $\mathbb{N}^m$. For a word $w = \prod_{i=1}^{m} {v_i}^{\alpha_i}$ in W^*, we have $\Upsilon(w) = (\alpha_1, \cdots, \alpha_m)$. Let us denote by $\Upsilon_i(w)$ the i-th component of $\Upsilon(w)$ (*i.e.* $\Upsilon_i(w) = \alpha_i$).

In the following, the transition matrix π is thus considered as a map from $V \times W^*$ into $\mathbb{R}$.

3.2 Multitype branching processes and stochastic L-systems

Stochastic L-systems are closely related to multitype branching processes. Let us recall first the definition of a multitype branching process (Harris (1963), Athreya and Ney (2004)). Let us consider a population with m types of individuals. Assume a type i individual produces children of all types according to a probability distribution $\{P_i(j) : j = (j_1, \cdots, j_m), j_i \in \mathbb{N}, 1 \le i \le m\}$. Assume all individuals produce offspring independently of each other and of the past history of the process. Let $Z_{n,i}$ be the number of type i individuals in the n-th generation. Let $\{\xi_{n,i}^{(k)} : n \in \mathbb{N}^*, k \in \mathbb{N}, 1 \le i \le m\}$ be independent random vectors in $\mathbb{N}^m$ with $\xi_{n,i}^{(k)}$ having distribution $P_i(.)$.

Definition 3.5 (multitype Galton-Watson branching process) *If the vector $Z_n = (Z_{n,1}, \cdots, Z_{n,m})$ of population sizes in the n-th generation evolves by the recursive relation*

$$Z_{n+1} = \sum_{i=1}^{m} \sum_{k=1}^{Z_{n,i}} \xi_{n,i}^{(k)}, \tag{1}$$

then $(Z_n)_{n\in\mathbb{N}}$ is a multitype Galton-Watson branching process.

The j-th component of $\xi_{n,i}^{(k)}$ represents the number of type j individuals produced by the k-th type i individual in the n-th generation. The set $\{P_i(.)\}_{i\in\{1,\cdots,m\}}$ is called the offspring distribution. Considering this definition, we have the following theorem:

Theorem 3.2 *Let $G = \langle A, \pi\rangle$ be a stochastic F0L-system on $V = \{v_1, \cdots, v_m\}$ with letters acting independently. Let $G[\omega_a] = \langle \omega_a, \pi\rangle$ be a component system of G and $(F_n[\omega_a])_{n\in\mathbb{N}}$ the corresponding Markov chain. Then, $(\Upsilon(F_n[\omega_a]))_{n\in\mathbb{N}}$ is a multitype Galton-Watson branching process whose offspring distributions $\{P_i(.)\}_{i\in\{1,\cdots,m\}}$ are given by:*

$$\forall i \in \{1,\cdots,m\}, \forall (j_1,\cdots,j_m) \in \mathbb{N}^m, \quad P_i(j_1,\cdots,j_m) = \pi_{v_i,\prod_{k=1}^m v_k^{j_k}}.$$

Proof: Let $\{P_i(.)\}_{i\in\{1,\cdots,m\}}$ be a set of probability distributions defined as follows:

$$\forall i \in \{1,\cdots,m\}, \forall (j_1,\cdots,j_m) \in \mathbb{N}^m, \quad P_i(j_1,\cdots,j_m) = \pi_{v_i,\prod_{k=1}^m v_k^{j_k}}.$$

Let $\{\xi_{n,i}^{(k)} : n \in \mathbb{N}^*, k \in \mathbb{N}, 1 \leq i \leq m\}$ be independent random vectors with $\xi_{n,i}^{(k)}$ having distribution $P_i(.)$. Thus, the q-th component of $\xi_{n,i}^{(k)}$ is the random variable that gives the number of type "v_q" symbols in a word generated by a type "v_i" symbol throughout G. Since there are $\Upsilon_i(F_n[\omega_a])$ type "v_i" symbols in $F_n[\omega_a]$, we have:

$$\Upsilon_q(F_{n+1}[\omega_a]) = \sum_{i=1}^{m} \sum_{k=1}^{\Upsilon_i(F_n[\omega_a])} \left(\xi_{n,i}^{(k)}\right)_q$$

and then,

$$\Upsilon(F_{n+1}[\omega_a]) = \sum_{i=1}^{m} \sum_{k=1}^{\Upsilon_i(F_n[\omega_a])} \xi_{n,i}^{(k)}$$

□

N.B.: the converse is true. Giving a multitype Galton-Watson branching process, there is a stochastic L-system whose stochastic production function is entirely determined by the offspring distribution.

Let us now define the generating function associated to a stochastic 0L-system. Let $S = (s_1, \cdots, s_m) \in [0,1]^m$.

Definition 3.6 (generating function associated to a stochastic 0L-system) *Let $G = \langle A, \pi\rangle$ be a stochastic F0L-system on $V = \{v_1, \cdots, v_m\}$ with letters acting independently. Let $G[\omega_a] = \langle \omega_a, \pi\rangle$ be a component system of G and $(F_n[\omega_a])_{n\in\mathbb{N}}$ the corresponding Markov chain. For $n \in \mathbb{N}$, the generating function $\psi_n[\omega_a]$ of $F_n[\omega_a]$ is defined as follows:*

$$\psi_n[\omega_a](S) = \sum_{\chi\in W^*} P(F_n[\omega_a] = \chi) \prod_{i=1}^{m} s_i^{\Upsilon_i(\chi)}.$$

$\psi_1[\omega_a]$ is said to be the generating function associated to $G[\omega_a]$.

By using the classical composition of generating functions (Harris (1963)) for a multitype Galton-Watson branching process, we deduce directly the following theorem:

Theorem 3.3 *Let $G = \langle V, \pi \rangle$ be a stochastic F0L-system on $V = \{v_1, \cdots, v_m\}$ with letters acting independently. For all $v \in V$, let $G[v] = \langle v, \pi \rangle$ be a component system of G and and $(F_n[v])_{n\in\mathbb{N}}$ the corresponding Markov chain. For all $n \in \mathbb{N}$, let $\psi_n[v]$ be the generating function associated to $F_n[v]$. Then,*

$$\forall n \in \mathbb{N}, \forall \omega_a \in V, \quad \psi_{n+1}[\omega_a](S) = \psi_1[\omega_a](\psi_n[v_1](S), \cdots, \psi_n[v_m](S))$$

4 Application to stochastic plant development

The GreenLab model of plant development is based on the botanical modelling concepts recalled in section 2. It can be described by a stochastic F0L-system $G^{tot} = \langle V, \pi^{tot} \rangle$ (Kang et al. (2007)). Let $N \in \mathbb{N}$ be the time during which we observe the growth of the plant. A bud $b_{n,\phi}^{k,\beta}$ is symbolized by $s_\phi^{k,\beta}$ and a metamer of PA j by m_j. V is the union of $\mathcal{S}$=$\{s_\phi^{k,\beta} : (\phi, \beta) \in \{0, \cdots, P\}^2, k \in \{1, \cdots, N\}\}$ and $\mathcal{M}$=$\{m_j : j \in \{1, \cdots, P\}\}$ where $\mathcal{S}$ is the set of buds (nonterminal elements) and $\mathcal{M}$ is the set of metamers (terminal elements). A dead bud will be represented by the empty word $\mathbf{1}$.

The seed can be considered as the initial bud and is thus represented by $s_1^{0,1}$. We consider the component system $G^{tot}[s_1^{0,1}]$ and $(F_n^{tot}[s_1^{0,1}])_{n\in\mathbb{N}}$ the corresponding Markov chain. For $n \in \mathbb{N}$, $F_n^{tot}[s_1^{0,1}]$ is the random variable representing the possible realisations of a plant after n growth cycles. Our objective is to determine the generating function associated to $F_n^{tot}[s_1^{0,1}]$ with Theorem 3.3. Thus, we only need to determine the generating function associated to $F_1^{tot}[s_1^{0,1}]$. However, the development processes underlying GreenLab organogenesis are complex. The dynamics of the population of buds results from the combination of both branching and differentiation (see Section 2) and we do not have an easy access to the distribution of $F_1^{tot}[s_1^{0,1}]$. Thus, the idea is to break down the whole system (branching + differentiation) into two subsystems and study them separately. This is the aim of Sections 4.1 (for branching) and 4.2 (for differentiation). In both cases, we write the corresponding generating function. Section 4.3 combines the two processes by using the results thus obtained and gives the recursive equations for the generating functions derivating from G^{tot}. The recursive equations for the moments of the numbers of metamers are finally deduced.

4.1 Branching

In this section we focus on the GreenLab development model with only branching. The differentiation process is not taken into account. This section gives not only the generating function for the branching process but also a method to build it. The model can be represented by a stochastic F0L-system $G^{br} = \langle V, \pi^{br} \rangle$. V is defined as in Section 4. For every component system $G^{br}[v]$ of G^{br}, the corresponding Markov chain is denoted $(F_n^{br}[v])_{n\in\mathbb{N}}$. In order to illustrate the branching process, let us take the example of coffee trees, plants with stochastic branching and no differentiation. They have two PAs (*i.e.* P=2). A metamer of PA 1 can bear a maximum of two lateral buds of PA 2 (*i.e.* $B_{1,2}^{max} = 2$). A metamer of PA 2 does not bear any lateral bud. At each growth cycle, there are three possible evolutions for a bud:

- case 1: $F_1^{br}[s_\phi^{k,\phi}] = \mathbf{1}$ with $\phi \in \{1, 2\}$, the bud dies.
- case 2: $F_1^{br}[s_\phi^{k,\phi}] = s_\phi^{k+1,\phi}$ with $\phi \in \{1, 2\}$, the bud is still alive but rests, .
- case 3: the bud is active. In that case, it produces a new metamer with lateral buds ($F_1^{br}[s_1^{k,1}] = m_1 s_1^{k+1,1}(s_2^{0,2})^j$ with $j \in \{1, 2\}$) or not ($F_1^{br}[s_\phi^{k,\phi}] = m_\phi s_\phi^{k+1,\phi}$ with $\phi \in \{1, 2\}$).

A probability of occurrence can be given for each of these possible evolutions. Let us now consider the general case. The aim is to write the generating function associated to G^{br}. Let $\psi_1^{br}[s_\phi^{k,\beta}](S, M)$

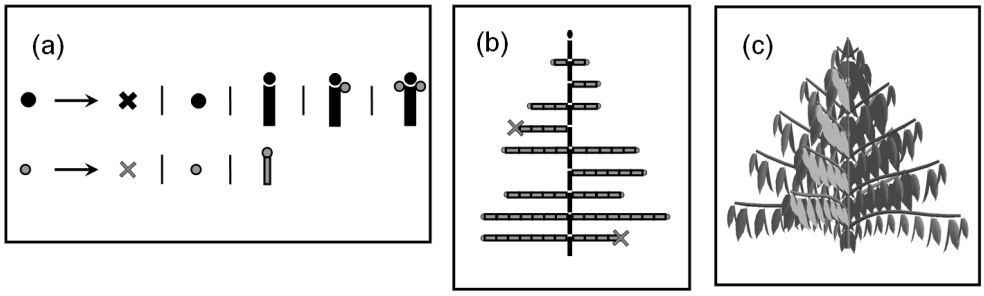


Fig. 2: (a) Production rules of coffee trees: phytomers and buds of PA 1 are in black and that of PA 2 are in gray. The cross symbol represents bud death. (b) An example of topology obtained after 10 growth cycles. (c) 3D representation.

be the generating function associated to the component system $G^{br}[s_\phi^{k,\beta}]$. $S = \left(s_{\phi'}^{k',\beta'}\right)_{k',\beta',\phi'}$ and $M = (m_{\phi'})_{\phi'}$ are vectors respectively on $[0,1]^{(N+1)P^2}$ and $[0,1]^P$. Then,

Theorem 4.1 *The generating function associated to $G^{br}[s_\phi^{k,\beta}]$ is given by*

$$\psi_1^{br}[s_\phi^{k,\beta}](S,M) = 1-P_l(\phi)+P_l(\phi)(1-P_a(\phi))s_\phi^{k+1,\beta}+P_l(\phi)P_a(\phi)m_\phi s_\phi^{k+1,\beta}\sum_{(j_\phi,\ldots,j_P)}\prod_{i=\phi}^{P}P_{\phi,i}^b(j_i)\left(s_i^{0,i}\right)^{j_i}.$$

Proof: Using Definition 3.6, we get the generating function by enumerating all the possible evolutions of $s_\phi^{k,\beta}$.

- case 1: $F_1^{br}[s_\phi^{k,\beta}] = \mathbf{1}$ and the probability of occurrence is $1 - P_l(\phi)$.
- case 2: $F_1^{br}[s_\phi^{k,\beta}] = s_\phi^{k+1,\beta}$ and the probability of occurrence is $P_l(\phi)(1 - P_a(\phi))$.
- case 3: the rules are of type $F_1^{br}[s_\phi^{k,\beta}] = m_\phi s_\phi^{k+1,\beta}\prod_{i=\phi}^{P}\left(s_i^{0,i}\right)^{j_i}$ with probability of occurrence $P_l(\phi)P_a(\phi)\prod_{i=\phi}^{P}P_{\phi,i}^b(j_i)$

Therefore, by summing the three cases, we get the equation of Theorem 4.1. □

4.2 Differentiation

In this section we focus on the differentiation process, corresponding to the change of meristem's PA. The branching process is not taken into account. We proceed as in Section 4.1 and write the stochastic F0L-System $G^{dif} = \left\langle V, \pi^{dif} \right\rangle$ associated to the differentiation process in order to write the corresponding generating function. V is defined as in Section 4. For every component system $G^{dif}[v]$ of G^{dif}, the corresponding Markov chain is denoted $(F_n^{dif}[v])_{n\in\mathbb{N}}$.

To illustrate differentiation, let us consider the example of maize. In standard cultivation conditions, maize is a mono-stem plant, that is to say without ramification. However, we can distinguish two types of

metamers along the stem. The first ones are short and can potentially bear tillers. They are followed by longer phytomers after meristem differentiation. Finally, the meristem ends up by flowering, which terminates the differentiation sequence and the stem development, see Figure 3. The two types of metamers are characterized by two different PAs ($P = 2$). We assume that the terminal bud of an axis of CA $k = 0$

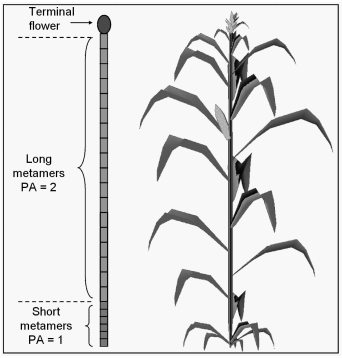


Fig. 3: Differentiation sequence for maize and 3D representation.

can not change its PA immediately. At each growth cycle, there are three possible evolutions for a bud through the differentiation process:

- case 1: $F_1^{dif}[s_\phi^{k,1}] = s_\phi^{k,1}$ with $\phi \in \{1, 2\}$, the bud does not differentiate. The PA of the terminal bud remains the same.
- case 2: $F_1^{dif}[s_1^{k,1}] = s_2^{k,1}$, the bud differentiates and is still alive. The PA of the terminal bud changes and is higher (cf Section 2)
- case 3: $F_1^{dif}[s_\phi^{k,1}] = \mathbf{1}$ with $\phi \in \{1, 2\}$, the bud differentiates and dies. The growth of the axis stops.

A probability of occurrence can be given for each of these evolutions. Let us now consider the general case. In order to write the generating function, we first have to identify the stochastic process underlying the differentiation of the terminal bud.

The aim is to study the evolution of the PA ϕ of a bud with respect to the CA k of the axis that bears the bud. Let $b_{n,\phi}^{k,\beta}$ be a bud of PA ϕ at growth cycle n. The axis that bears the bud is of CA k and was initiated by a bud of PA β. Then, we assume $\phi = \phi(k)$. Moreover, we assume that the occupation time of a bud in the PA j follows an exponential law of parameter λ_j where λ_j is defined in Section 2. It is in keeping with the botanical theory of meristematic differentiation sequence as a sequence of events inducing plant morphological ageing. Therefore, we can represent the discrete process $(\phi(k))_{k\in\mathbb{N}}$ by a finite state space Markov process $(\phi(t))_{t\in\mathbb{R}_+}$.

We choose to model the stochastic behaviour of $(\phi(t))_{t\in\mathbb{R}_+}$ with multivariate phase-type random vectors, which are particularly used in reliability theory (Neuts (1975)). An interesting aspect of phase-type distributions is that they can be written in a closed form, and consequently, various quantities of interest can be evaluated with relative ease (Assaf et al. (1984)). Let $(X(t))_{t\in\mathbb{R}_+}$ be a right-continuous Markov process on a finite state space $E = \{1, \cdots, P, \Delta\}$. Assume $1, \cdots, P$ are transient and Δ is absorbing.

Then, the infinitesimal generator $\mathbf{Q}$ of size $P+1$ by $P+1$ has the form

$$\mathbf{Q} = \left(\begin{array}{c|c} A & -A\mathbf{e} \\ \hline 0 & 0 \end{array}\right)$$

where A is a $P \times P$ matrix and $\mathbf{e}$ is a vector of size P with all its components set to 1.

Let $\Gamma_1, \cdots, \Gamma_n$ be nonempty stochastically closed subsets of E for X. A subset Γ of E is said to be stochastically closed for X if once X enters Γ, it never leaves it. Let a be the starting measure on E with $a(\Delta) = 0$. Let α be the vector of size P defined by $\alpha = (a(1), \cdots, a(P))$. We define $T_k = \inf\{j, X(j) \in \Gamma_k\}$.

Definition 4.1 (multivariate phase-type random vector) *If the following hypotheses are true:*
1 - $\bigcap_{k=1}^{n} \Gamma_k = \Delta$ *,*
2 - the absorption into Δ *is certain,*
3 - $P(T_1 > 0, \cdots, T_n > 0) = 1$*,*
then $(T_1, \cdots, T_n)$ *is called a multivariate phase-type random vector with representation* (α, A)*.*

Considering now the Markov process $(\phi(t))_{t\in\mathbb{R}_+}$, we have $E = \{1, \cdots, P, d\}$ where $1, \cdots, P$ represent the possible PAs (*i.e.* transient states) and d represents a dead bud, or the terminal flowering state as in Maize, (*i.e.* an absorbing state). We assume $\phi(0) \neq d$. In the sequel, it will be important to specify the initial state of the Markov process. Let $(\phi^\beta(t))_{t\in\mathbb{R}}$ be the PA process such that $\phi^\beta(0) = \beta$ with $\beta \in \{1, \cdots, P\}$. Let a^β be the starting measure on E for ϕ^β. We assume that $a^\beta(d) = 0$. Let α^β be the vector of size P defined by $\alpha^\beta = (\alpha^\beta(1), \cdots, \alpha^\beta(P))$. Given that $\phi^\beta(0) = \beta$, we have $\alpha_\beta(j) = \delta_\beta(j)$ where δ_β is the Dirac measure concentrated on β. The infinitesimal generator Q is given by $\{\lambda_j\}_{j\in\{1,\cdots,P\}}$ and the transition probabilities of the underlying Markov chain $\{q_{i,j}\}_{(i,j)\in\{1,\cdots,P\}^2}$. Let us recall that $q_{i,j} = 0$ for $j \leq i$ (Section 2). Then,

$$\mathbf{Q} = \left(\begin{array}{cccc|c} -\lambda_1 & \lambda_1 q_{1,2} & \dots & \lambda_1 q_{1,P} & \lambda_1 q_{1,d} \\ 0 & -\lambda_2 & \dots & \lambda_2 q_{2,P} & \lambda_2 q_{2,m} \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & -\lambda_P & \lambda_P \\ \hline 0 & 0 & \dots & 0 & 0 \end{array}\right) = \left(\begin{array}{c|c} A & -A\mathbf{e} \\ \hline 0 & 0 \end{array}\right)$$

Let $\{\Gamma_k\}_{k\in\{1,\dots,P+1\}}$ be a family of subsets of E defined as follows:

$$\Gamma_k = \begin{cases} \bigcup_{j=k}^{P} \{j\} \cup \{d\} & \text{if } k \in \{1, ..., P\} \\ \{d\} & \text{if } k = P+1 \end{cases}$$

Given $q_{i,j} = 0$ for $j \leq i$, $\{\Gamma_k\}_{k\in\{1,\dots,P+1\}}$ are stochastically closed subsets of E for ϕ^β. T_k^β denotes the time when ϕ^β enters Γ_k:

$$T_k^\beta = \inf\{j, \Phi^\beta(j) \in \Gamma_k\}$$

Then, we have the following proposition:

Proposition 4.1 *For all $\beta \in \{1, \cdots, P\}$, $(T^{\beta}_{\beta+1}, ..., T^{\beta}_{P+1})$ is a multivariate phase-type random vector with representation (α^{β}, A). Moreover, we have $0 < T^{\beta}_{\beta+1} < T^{\beta}_{\beta+2} < ... < T^{\beta}_{P+1}$.*

Proof: For all $\beta \in \{1, \cdots, P\}$, we have to prove the three hypotheses:

1 - it is obvious that $\bigcap_{k=\beta+1}^{P+1} \Gamma_k = d$.

2 - according to Neuts (1975), the absorption into d is certain if and only if A is non singular. This is true because $\det A = (-1)^P \prod_{k=1}^{P} \lambda_k \neq 0$.

3 - given the only nonzero component of α^{β} is the $\beta - th$ one, $T^{\beta}_{k} > 0$ is always true for $k \in \{\beta + 1, \cdots, P + 1\}$. Then, the third hypothesis is true.

According to Definition 4.1, $(T^{\beta}_{\beta+1}, ..., T^{\beta}_{P+1})$ is a multivariate phase-type random vector with representation (α^{β}, A). The second part of the proposition is obvious because $\Gamma_{P+1} \subset \Gamma_P \subset ... \subset \Gamma_1$. □

We will now determine the generating function derivating from G^{dif}. The probability distribution of $F_1^{dif}[s^{k,\beta}_{\phi}]$ is given naturally by the multivariate phase-type distributions.

Let $\psi_1^{dif}[s^{k,\beta}_{\phi}](S)$ be the generating function associated to $G^{dif}[s^{k,\beta}_{\phi}]$. $S = \left(s^{k',\beta'}_{\phi'}\right)_{k',\beta',\phi'}$ is a vector on $[0,1]^{(N+1)P^2}$. Let D be the event $\{T^{\beta}_{\phi} < k-1, T^{\beta}_{\phi+1} > k-1\}$. Let g_k be a diagonal matrix of size P by P whose i-th diagonal element is equal to 1 if $i \in \Gamma_k^c$ and 0 otherwise. Γ_k^c denotes the complement of Γ_k in E. Then,

Theorem 4.2 *The generating function associated to $G^{dif}[s^{k,\beta}_{\phi}]$ is given by*

$$\psi_1^{dif}[s^{k,\beta}_{\phi}](S) = P(T^{\beta}_{P+1} < k|D) + P(T^{\beta}_{\phi+1} > k|D)s^{k,\beta}_{\phi} + \sum_{i=1}^{P-\phi} P(T^{\beta}_{\phi+i} < k, T^{\beta}_{\phi+i+1} > k|D)s^{k,\beta}_{\phi+i}$$

if $1 \leq k \leq N-1$ and by

$$\psi_1^{dif}[s^{0,\beta}_{\beta}](S) = s^{0,\beta}_{\beta}$$

if $k = 0$ and all the probabilities are explicit functions of the matrix A:

$$P(T^{\beta}_{\phi+1} > k|D) = \frac{{}^t(\alpha^{\beta})[g_{\beta+1}e^{A(k-1)} - (g_{\beta+1})^{\mathbf{1}_{\{\beta+1\}}(\phi)}e^{A(k-1)}g_{\phi}]e^{A}g_{\phi+1}\boldsymbol{e}}{{}^t(\alpha^{\beta})[g_{\beta+1}e^{A(k-1)} - (g_{\beta+1})^{\mathbf{1}_{\{\beta+1\}}(\phi)}e^{A(k-1)}g_{\phi}]g_{\phi+1}\boldsymbol{e}}$$

$$P(T^{\beta}_{\phi+i} < k, T^{\beta}_{\phi+i+1} > k|D) = \frac{{}^t(\alpha^{\beta})[g_{\beta+1}e^{A(k-1)}g_{\phi+1} - (g_{\beta+1})^{\mathbf{1}_{\{\beta+1\}}(\phi)}e^{A(k-1)}g_{\phi}]e^{A}[g_{\phi+i+1} - g_{\phi+i}]\boldsymbol{e}}{{}^t(\alpha^{\beta})[g_{\beta+1}e^{A(k-1)} - (g_{\beta+1})^{\mathbf{1}_{\{\beta+1\}}(\phi)}e^{A(k-1)}g_{\phi}]g_{\phi+1}\boldsymbol{e}}$$

Proof: We will proceed as in Section 4.1. For the bud $s^{k,\beta}_{\phi}$, we know that ϕ enters Γ_{ϕ} before $k-1$ but not $\Gamma_{\phi+1}$. It explains why all the probabilities in Theorem 4.2 have the form $P(.|D)$. There are 3 cases for bud differentiation:

- $F_1^{dif}[s^{k,\beta}_{\phi}] = s^{k,\beta}_{\phi}$: the bud does not change its PA. This is the case if ϕ^{β} does not enter Γ_{k+1} before k. Therefore, the probability of occurrence is $P(T^{\beta}_{\phi+1} > k+1|D)$.

- $F_1^{dif}[s_\phi^{k,\beta}] = s_{\phi'}^{k,\beta}$ with $\phi' > \phi$: the bud changes its physiological for ϕ'. It means ϕ^β enters $\Gamma_{\phi'}$ before k but not $\Gamma_{\phi'+1}$. Therefore, the probability of occurrence is $P(T_{\phi'}^\beta < k+1, T_{\phi'+1}^\beta > k+1|D)$.
- $F_1^{dif}[s_\phi^{k,\beta}] = \mathbf{1}$: the bud dies. It means that ϕ^β enters Γ_{P+1} before k. Therefore, the probability of occurrence is $P(T_{P+1}^\beta < k|D)$.

By summing the three cases, we have the expression of $\psi_1^{dif}[s_\phi^{k,\beta}](S)$. The case of $s_\beta^{0,\beta}$ is particular because we assume that a bud that has just initiated a new axis can not change its PA.

Modelling differentiation with multivariate phase-type random vectors is interesting since all probabilities are given explicitly. It was proved in Assaf et al. (1984) that for all $t_{P+1} \geq t_P \geq \cdots \geq t_{\beta+1} \geq 0$,

$$P(T_{\beta+1}^\beta > t_{\beta+1}, \cdots, T_{P+1}^\beta > t_{P+1}) = {}^t(\alpha^\beta) e^{At_{\beta+1}} g_{\beta+1} e^{A(t_{\beta+2}-t_{\beta+1})} \cdots g_P e^{A(t_{P+1}-t_P)} g_{P+1}\mathbf{e}.$$

Thus, given $T_{\beta+1}^\beta < T_{\beta+2}^\beta < ... < T_{P+1}^\beta$, we get the expression of the probabilities of Theorem 4.2. □

4.3 The complete model

In this section, we establish the recursion formulas for the generating functions derivating from the complete L-system $G^{tot} = \langle V, \pi^{tot} \rangle$ defined at the beginning of Section 4. The aim is to write the recursion formulas considering G^{tot} as a composition of the L-systems G^{br} and G^{dif}. Let us define first a compound L-system:

Definition 4.2 (compound L-system) *Let $G = \langle V, \pi \rangle$, $G^1 = \langle V, \pi^1 \rangle$ and $G^2 = \langle V, \pi^2 \rangle$ be stochastic L-systems on the same alphabet V. Let P, P^1 and P^2 be the Markov kernel associated respectively to G, G^1 and G^2. G is said to be the composition of G^1 by G^2 if $P = P^1.P^2$. We write $G = G^1 \circ G^2$.*

The complete GreenLab development model mixes branching and differentiation. We assume the priority is given to the branching process at each growth cycle. Then, each step of G^{tot} begins with a step of G^{br} and goes on with a step of G^{dif}. It is thus obvious that $G^{tot} = G^{br} \circ G^{dif}$. This result leads to the following theorem:

Theorem 4.3 *For all $(\beta, \phi) \in \{1, \cdots, P\}^2$, $S = \left(s_{\phi'}^{k',\beta'}\right)_{k',\beta',\phi'} \in [0,1]^{(N+1)P^2}$ and $M = (m_{\phi'})_{\phi'} \in [0,1]^P$,*

$$\forall n + k \leq N+1, \quad \psi_{n+1}^{tot}[s_\phi^{k,\beta}](S, M) = \psi_1^{br}[s_\phi^{k,\beta}](\boldsymbol{\psi^{dif}}(\boldsymbol{\psi_n^{tot}}(S, M)), M)$$

with $\boldsymbol{\psi_n^{tot}} = \left(\psi_n^{tot}[s_\phi^{k,\beta}]\right)_{k,\beta,\phi}$ and $\boldsymbol{\psi^{dif}} = \left(\psi_1^{dif}[s_\phi^{k,\beta}]\right)_{k,\beta,\phi}$.

Proof: Since $G^{tot} = G^{br} \circ G^{dif}$, the generating function derivating from G^{tot} is the compound function of the generating function derivating from G^{dif} by that of G^{br} *i.e.*:

$$\psi_1^{tot}[s_\phi^{k,\beta}](S, M) = \psi_1^{br}[s_\phi^{k,\beta}](\boldsymbol{\psi^{dif}}(S), M)$$

Then, using Theorem 3.3, we get the result. □

Setting $\psi_1^{tot}[s_\phi^{k,\beta}](S, M) = 0$ for $k + n > N + 1$ and $\boldsymbol{\psi^{br}} = \left(\psi_1^{br}[s_\phi^{k,\beta}]\right)_{k,\beta,\phi}$, we have:

Theorem 4.4 *For all $n \leq N-1$, $\boldsymbol{\psi_{n+1}^{tot}}(S, M) = \boldsymbol{\psi^{br}}(\boldsymbol{\psi^{dif}}(\boldsymbol{\psi_n^{tot}}(S, M)), M)$.*

Let $m_j[s_\phi^{k,\beta}, n]$ be the number of type j metamers in a structure initiated by $s_\phi^{k,\beta}$ after n growth cycles and let M_n be the matrix of size $(N+1)P^2$ by P whose j-th column vector is $\left(E\left[m_j[s_\phi^{k,\beta}, n]\right]\right)_{k,\beta,\phi}$. Let e_k be the vector of size k with all its components set to 1. We deduce the fundamental recursion equation for the expectations of the numbers of organs on each type of structures. It is important to note that the first line of the matrix corresponds to those of the whole plant.

Theorem 4.5 *for* $N > 0$,

$$M_{N+1} = \frac{\partial \boldsymbol{\psi}^{\boldsymbol{br}}}{\partial S}(e_{(N+1)P^2}, e_P)\frac{\partial \boldsymbol{\psi}^{\boldsymbol{dif}}}{\partial S}(e_{(N+1)P^2}, e_P)M_N + \frac{\partial \boldsymbol{\psi}^{\boldsymbol{br}}}{\partial M}(e_{(N+1)P^2}, e_P)$$

Proof: Using the properties of generating functions, we have $M_n = \frac{\partial \boldsymbol{\psi}_n^{\boldsymbol{tot}}}{\partial M}(e_{(N+1)P^2}, e_P)$. Derivating the equation from Theorem 4.4 and taking $(S, M) = (e_{(N+1)P^2}, e_P)$ we get Theorem 4.5. □

In the same way, we can give recursion formulas for the variance. The detailed results, as well as examples of simulations and biological discussions, are given in Loi and Cournède (2008).

5 Conclusion

While the interest of stochastic L-systems for plant growth simulation and visualization is broadly acknowledged, its full mathematical potential to characterize the probability distributions and moments of the numbers of organs in plant structure has not been taken advantage of. The need for an analytic computation of these distributions is crucial in stochastic functional-structural models in order to derive the distribution or the moments of biomass production (cf. Kang et al. (2008) for preliminary results). It has led us to clearly formalize the link between stochastic L-systems and multi-type branching processes, and thus to derive an inductive relationship to compute the associated generating functions. This framework was applied successfully to the GreenLab organogenesis model, by decomposing the development process into two botanical sub-processes, branching and meristem differentiation. For the latter, multivariate phase type random vectors were introduced to describe the stochastic sequence of meristem differentiation. From the inductive relationship on the generating functions of the numbers of organs, we can determine the inductive relationship giving the moments of the distributions.

The next step is the link of stochastic organogenesis and functioning models. It is possible to derive the approximate moments of biomass production thanks to differential statistics. An interesting issue concerns the modelling of the probability distributions governing organogenesis as functions of biomass production or groth rate (cf. Mathieu et al. (2006)). The stochastic processes resulting from these interactions between development and functioning should lead us to improve the proposed formalism.

References

D. Assaf, N. Langberg, T. Savits, and M. Shaked. Multivariate phase-type distributions. *Operations Research*, 32(3):688–702, 1984.

K. Athreya and P. Ney. *Branching Processes*. Dover Publications, 2004.

D. Barthélémy and Y. Caraglio. Plant architecture: a dynamic, multilevel and comprehensive approach to plant form, structure and ontogeny. *Annals of Botany*, 99(3):375–407, 2007.

P.-H. Cournède, M.-Z. Kang, A. Mathieu, J.-F. Barczi, H.-P. Yan, B.-G. Hu, and P. de Reffye. Structural Factorization of Plants to Compute their Functional and Architectural Growth. *Simulation*, 82(7):427–438, 2006.

P. de Reffye. *Modélisation de l'architecture des arbres par des processus stochastiques. Simulation spatiale des modèles tropicaux sous l'effet de la pesanteur. Application au Coffea robusta*. PhD thesis, Université Paris-Sud, Centre d'Orsay, 1979.

P. de Reffye, M. Goursat, J. Quadrat, and B. Hu. The Dynamic Equations of the Tree Morphogenesis Greenlab Model. Technical Report 4877, INRIA, 2003.

P. De Reffye, E. Heuvelink, D. Barthélémy, and P.-H. Cournède. Biological and mathematical concepts for modelling plant growth and architecture. In *Encyclopedia of Ecology*. Jorgensen, D.E. and Fath, B., elsevier edition, 2008.

J. Françon. Sur la modélisation informatique de l'architecture et du développement des végétaux. In *2ème Colloque International: L'Arbre. Institut de Botanique, Montpellier, France*, 1990.

T. Harris. *The theory of branching processes*. Springer, Berlin, 1963.

M.-Z. Kang, P.-H. Cournède, J.-P. Quadrat, and P. de Reffye. A stochastic language for plant topology. In T. Fourcaud and X. Zhang, editors, *Plant growth Modeling, simulation, visualization and their Applications*. IEEE Computer Society (Los Alamitos, California), 2007.

M.-Z. Kang, P.-H. Cournède, P. de Reffye, D. Auclair, and B.-G. Hu. Analytical study of a stochastic plant growth model: application to the greenlab model. *Mathematics and Computers in Simulation*, 78 (1):57–75, 2008.

A. Lindenmayer. Mathematical models for cellular interactions in development. i. filaments with one-sided inputs. *Journal of Theoretical Biology*, 18:280–289, 1968.

C. Loi and P.-H. Cournède. Description of the GreenLab development model with stochastic L-systems and Monte-Carlo simulations. Technical report, INRIA, 2008.

A. Mathieu, P.-H. Cournède, and P. de Reffye. A dynamical model of plant growth with full retroaction between organogenesis and photosynthesis. *ARIMA Journal*, 4:101–107, 2006. URL `http://www-direction.inria.fr/international/arima/004/00407.htm`.

M. F. Neuts. Probability distributions of phase type. *Liber Amicorum Prof. Emeritus H. Florin*, pages 176–206. University of Louvain, Belgium, 1975.

P. Prusinkiewicz and A. Lindenmayer. *The Algorithmic Beauty of Plants*. Springer-Verlag, New-York, 1990.

G. Rozenberg and A. Salomaa. *The Mathematical Theory of L-systems*. Academic Press, New York, 1980.

A. Smith. Plants, fractals and formal languages. *Computer Graphics (SIGGRAPH 84 Conference Proceedings)*, 18(3):1–10, 1984.

D. Stroock. *An introduction to Markov processes*. Springer, 2005.