Ann Inst Stat Math (2013) 65:269–293
DOI 10.1007/s10463-012-0368-7

The algebra of reversible Markov chains

Giovanni Pistone · Maria Piera Rogantin

Received: 12 April 2011 / Revised: 12 January 2012 / Published online: 15 July 2012

Abstract For a Markov chain, both the detailed balance condition and the cycle Kolmogorov condition are algebraic binomials. This remark suggests to study reversible Markov chains with the tool of Algebraic Statistics, such as toric statistical models. One of the results of this study is an algebraic parameterization of reversible Markov transitions and their invariant probability.

Keywords Reversible Markov chain · Algebraic statistics · Toric ideal

1 Introduction

On a finite state space V, we consider q-reversible (quasi-reversible) Markov matrices, i.e. Markov matrices P with elements denoted $P_{v\to w},\ v, w \in V$, such that $P_{v\to w} = 0$ if, and only if, $P_{w\to v} = 0,\ v \neq w$.

The support of P is the simple graph $\mathscr{G} = (V, \mathscr{E})$, where $\overline{vw},\ v \neq w$, is an edge if, and only if, $P_{v\to w}$ and $P_{w\to v}$ are both positive. We associate to each edge $\overline{vw}$ the two directed arcs $v \to w$ and $w \to v$ to get a directed graph without loops (i.e. arcs from v to v) that we denote by $\mathscr{D} = (V, \mathscr{A})$, see Fig. 1. The neighborhood of v is $\mathrm{N}(v)$, the degree of v is $d(v)$, the set of arcs leaving v is $\mathrm{out}(v)$, the set of arcs entering v is $\mathrm{in}(v)$.

G. Pistone was supported by de Castro Statistics Initiative, Collegio Carlo Alberto, Moncalieri, Italy.

G. Pistone (✉)
Collegio Carlo Alberto, Via Real Collegio 30, 10024 Moncalieri, Italy
e-mail: giovanni.pistone@carloalberto.org

M. P. Rogantin
DIMA Università di Genova, Via Dodecaneso 35, 16146 Genova, Italy
e-mail: rogantin@dima.unige.it

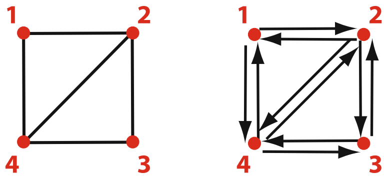

Fig. 1 Running example: undirected graph $\mathscr{G}$ and directed graph $\mathscr{D}$

Viceversa, given a connected graph $\mathscr{G} = (V, \mathscr{E})$, a Markov matrix P is q-reversible with structure $\mathscr{G}$ if the following two conditions hold: $P_{v\to w} = 0$ if $\overline{vw} \notin \mathscr{E}$; $P_{v\to w} = 0$ if, and only if, $P_{w\to v} = 0$ for all edges $\overline{vw} \in \mathscr{E}$. Such a Markov matrix is characterized by its restriction to non diagonal elements, that is by a mapping $P\colon \mathscr{A} \to \mathbb{R}_+$ such that $\sum_{w\in \mathrm{N}(v)} P_{v\to w} \leq 1,\ v \in V$. As we assume $P_{v\to w} = 0$ if $\overline{vw} \notin \mathscr{E}$, the diagonal elements are computed as $P_{v\to v} = 1 - \sum_{w\in \mathrm{N}(v)} P_{v\to w}$. The support of P will be a sub-graph of $\mathscr{G}$. We are going to need this detailed classification in the following.

For any set I, we denote by $\mathbb{R}^I$, respectively, $\mathbb{R}^I_+$, the vector space of real, respectively, non-negative real, functions defined on I. Each element is represented as a column array whose row's names are the elements of the set I. An element of $\mathbb{R}^n$, $n = |I|$, would be a vector without row's names. With this notation, $\Delta(I) = \left\{x \in \mathbb{R}^I_+ : \sum_{i\in I} x_i = 1\right\}$ denotes the (flat) simplex on the index set I, while $S(I) = \left\{x \in \mathbb{R}^I_+ : \sum_{i\in I} x_i \leq 1\right\}$ denotes the (solid) simplex over the index set I.

The set of Markov transitions $P = (P_e)_{a\in\mathscr{A}}$ is parameterized by the product of simplexes $\times_{v\in V} S(\mathrm{N}(v))$.

A Markov matrix P on V satisfies the *detailed balance* condition if there exists $\kappa(v) > 0,\ v \in V$, such that

$$\kappa(v) P_{v\to w} = \kappa(w) P_{w\to v}, \quad v, w \in V.$$

It follows that P is q-reversible and that $\pi(v) = \kappa(v)/\sum_{v\in V}\kappa(v)$ is an invariant probability with full support. Equivalently, the Markov chain $(X_n)_{n\geq 0}$ with invariant probability π and transition matrix $[P_{v\to w}]$, $v, w \in V$, has *reversible* bivariate joint distribution

$$\mathrm{P}\left(X_n = v, X_{n+1} = w\right) = \mathrm{P}\left(X_n = w, X_{n+1} = v\right), \quad v, w \in V, \quad n \geq 0. \tag{1}$$

Such a Markov chain (MC), and its transition matrix, are called reversible. Reversible Markov Chains are relevant in Statistical Physics, e.g. in the theory of entropy production, and in Applied Probability, e.g. the simulation method Monte Carlo Markov Chain (MCMC). The main aim of this paper is to find useful parameterizations of the reversible Markov matrices of a given structural graph.

In Sect. 2, we review some basics from Dobrushin et al. (1988), Kelly (1979), Strook (2005, Ch 5), Diaconis and Rolles (2006), Hastings (1970), Peskun (1973), and Liu (2008).

In Sect. 3, we discuss the algebraic theory prompted by the detailed balance condition. The results pertain to the area of Algebraic Statistics (see, e.g. Pistone et al. 2001; Drton et al. 2009; Gibilisco et al. 2010). Previous results in the same algebraic spirit were presented in Suomela (1979) and Mitrophanov (2004). We believe the results here are new; some proofs depend on classical notions of graph theory that are reviewed in some detail because of our particular context.

The discussion and the conclusions are briefly presented in Sect. 4.

2 Background

2.1 Reversible Markov process

The reversibility of the bivariate joint distribution in Eq. (1) gives a second parameterization of reversible Markov chains. In fact, a stochastic process $(X_n)_{n\geq 0}$ with state space V is 2-reversible and 2-stationary if, and only if, Eq. (1) holds. When the process is a Markov chain, the distribution depends on the bivariate distributions only. In particular, the process is 1-stationary: by summing over $w \in V$, we have

$$\mathrm{P}(X_n = v) = \mathrm{P}(X_{n+1} = v) = \pi(v), \quad v \in V, n \geq 0.$$

Let V_2 be the set of all subset of V of cardinality 2. The elements of V_2 are the edges of the full graph on V. The following parameterization of the two-dimensional distributions has been used in Diaconis and Rolles (2006):

$$\begin{aligned} \theta_v &= \mathrm{P}(X_n = v, X_{n+1} = v), \quad v \in V, \\ \theta_{\overline{vw}} &= \mathrm{P}(X_n = v, X_{n+1} = w) + \mathrm{P}(X_n = w, X_{n+1} = v) \\ &= 2\,\mathrm{P}(X_n = v, X_{n+1} = w), \quad \overline{vw} \in V_2. \end{aligned} \tag{2}$$

The number of parameters is $N + \binom{N}{2} = \binom{N+1}{2}$; moreover it holds

$$1 = \sum_{v,w\in V} \mathrm{P}(X_n = v, X_{n+1} = w) = \sum_{v\in V} \theta_v + \sum_{\overline{vw}\in V_2} \theta_{\overline{vw}},$$

hence $\theta = (\theta_V, \theta_{V_2})$ belongs to the simplex $\Delta(V \cup V_2)$.

Given an undirected graph $\mathcal{G} = (V, \mathcal{E})$ such that $\mathrm{P}(X_n = v, X_{n+1} = w) = 0$ if $\overline{vw} \notin \mathcal{E}$, then the vector of parameters $\theta = (\theta_V, \theta_{\mathcal{E}})$ belongs to the convex set $\Delta(V \cup \mathcal{E})$. We note that the vertices V are identified with loops of the transitions because $\theta_v = \mathrm{P}(X_n = v, X_{n+1} = v)$.

The marginal probability π can be written using the θ parameters:

$$\pi(v) = \sum_{w\in V} P(X_n = v, X_{n+1} = w) = \theta_v + \frac{1}{2} \sum_{w\in \mathrm{N}(v)} \theta_{\overline{vw}},$$

or, in matrix form,

$$\pi = \theta_V + \frac{1}{2}\Gamma\theta_{\mathscr{E}},$$

where Γ is the incidence matrix of the graph $\mathscr{G}$.

Example 1 (Running example) Consider the graph $\mathscr{G} = (V, \mathscr{E})$ with $V = \{1, 2, 3, 4\}$ and $\mathscr{E} = \left\{\overline{12}, \overline{23}, \overline{34}, \overline{14}, \overline{24}\right\}$, see left side of Fig. 1. Here,

$$\Gamma = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} \overline{12} & \overline{23} & \overline{34} & \overline{14} & \overline{24} \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} \end{array}.$$

All π's are obtained this way if we admit positive probability on loops.

Proposition 1 1. *The map*

$$\gamma : \Delta(V \cup \mathscr{E}) \ni \theta = \begin{bmatrix} \theta_V \\ \theta_{\mathscr{E}} \end{bmatrix} \longmapsto \pi = \left[I_V \; \tfrac{1}{2}\Gamma\right] \begin{bmatrix} \theta_V \\ \theta_{\mathscr{E}} \end{bmatrix} \in \Delta(V)$$

is a surjective Markov map.

2. *The image of* $(0, \theta_{\mathscr{E}})$, $\theta_{\mathscr{E}} \in \Delta(\mathscr{E})$, *is the convex hull of the half points of each edge of the simplex* $\Delta(V)$ *whose vertices are connected in* $\mathscr{G}$.

Proof 1. Each probability π is the image of $\theta = (\pi, 0_{\mathscr{E}})$.

2. If $\theta_V = 0$, then $\pi = \sum_{w \in \mathrm{N}(v)} \frac{1}{2}\Gamma_{\overline{vw}}\theta_{\overline{vw}}$, where $\sum_{w \in \mathrm{N}(v)} \theta_{\overline{vw}} = 1$ and $\Gamma_{\overline{vw}}$ is the $\overline{vw}$-column of Γ. Hence, $\frac{1}{2}\Gamma_{\overline{vw}}$ is the the middle point of the v- and the w-vertex of the simplex $\Delta(V)$. □

Item 1 of the proposition leaves open the question of the existence of an element θ such that $\theta_{\mathscr{E}} > 0$ for each $\pi > 0$. This is discussed in the next subsection. Item 2 shows that, while all π's can be obtained if loops are allowed ($\theta_V \geq 0$), only a convex subset of $\Delta(V)$ is obtained if we do not allow for loops ($\theta_V = 0$).

2.2 From a positive π to positive transitions

Given π, the fiber $\gamma^{-1}(\pi)$ is contained in an affine space parallel to the subspace

$$\ker\left[I_V \; \tfrac{1}{2}\Gamma\right] = \left\{\theta \in \mathbb{R}^{V \cup \mathscr{E}} : \theta_V + \frac{1}{2}\Gamma\theta_{\mathscr{E}} = 0\right\}.$$

Each fiber $\gamma^{-1}(\pi)$, $\pi > 0$, contains special solutions. The solution $(\pi, 0_{\mathscr{E}})$ is not of interest because we want $\theta_{\mathscr{E}} > 0$. If the graph has full connections, $\mathscr{G} = (V, V_2)$, there is the independence solution $\theta_v = \pi(v)^2$, $\theta_{\overline{vw}} = 2\pi(v)\pi(w)$.

If $\pi(v) > 0$, $v \in V$, a strictly positive solution is obtained as follows. Let $d(v)$ be the degree of the vertex v and define a transition probability by $A(v, w) = 1/(2d(v))$

if $\overline{vw} \in \mathcal{E}$, $A(v,v) = 1/2$, and $A(v,w) = 0$ otherwise. A is the transition matrix of a random walk on the graph $\mathcal{G}$, stopped with probability 1/2. Define a probability on $V \times V$ with $Q(v,w) = \pi(v)A(v,w)$. If $Q(v,w) = Q(w,v)$, $v, w \in V$, we are done: we have found a 2-reversible probability with marginal π and such that $Q(v,w) > 0$ if, and only if $\overline{vw} \in \mathcal{E}$. Otherwise, if $Q(v,w) \neq Q(w,v)$ for some $v, w \in V$, we turn to the following Hastings-Metropolis construction.

Proposition 2 *Let Q be a probability on $V \times V$ such that $Q(v,w) > 0$ if, and only if, $\overline{vw} \in \mathcal{E}$ or $v = w$. Write $\pi(v) = \sum_w Q(v,w)$. Given $f :]0,1[\times]0,1[\to]0,1[$ a symmetric function such that $f(x,y) \leq x \wedge y$, then*

$$P(v,w) = \begin{cases} f(Q(v,w), Q(w,v)) & \text{if } v \neq w, \\ \pi(v) - \sum_{w:\, w\neq v} P(v,w) & \text{if } v = w. \end{cases}$$

is a 2-reversible probability on $V \times V$ such that $\pi(v) = \sum_w P(v,w)$ and $P(v,w) > 0$ if, and only if, $\overline{vw} \in \mathcal{E}$.

Proof For $\overline{vw} \in \mathcal{E}$ we have $P(v,w) = P(w,v) > 0$, otherwise zero. As $P(v,w) \leq Q(v,w)$, $v \neq w$, it follows

$$\begin{aligned} P(v,v) &= \pi(v) - \sum_{w:\, w\neq v} P(v,w) \\ &\geq \sum_w Q(v,w) - \sum_{w:\, w\neq v} Q(v,w) \\ &= Q(v,v) > 0. \end{aligned}$$

We have $\sum_w P(v,w) = \pi(v)$ by construction and, in particular, $P(v,w)$ is a probability on $V \times V$. □

Remark 1 1. The proposition applies to
   (a) $f(x,y) = x \wedge y$. This is the standard Hastings choice.
   (b) $f(x,y) = xy/(x+y)$. This was suggested by Barker.
   (c) $f(x,y) = xy$. In fact, as $y < 1$, we have $xy < x$.
2. Given a joint probability P, the corresponding parameters

$$\theta_{\overline{vw}} = 2P(v,w) \quad \text{and} \quad \theta_v = P(v,v)$$

are strictly positive for $\overline{vw} \in \mathcal{E}$ and $v \in V$, otherwise zero. We have shown the existence of a mapping from π in the interior of $\Delta(V)$ to a vector of parameters θ in the interior of $\Delta(V \cup \mathcal{E})$.

2.3 Parameterization of reversible Markov matrices

A q-reversible Markov matrix P supported on a graph $\mathcal{G}$ is parameterised by its non-zero extradiagonal values $P_{v\to w}$, i.e. by the elements of $\times_{v\in V} S^\circ(\mathrm{N}(v))$, where S°

denotes the open solid simplex. As $\mathscr{G}$ is connected, the invariant probability of the Markov matrix P is unique, therefore the joint 2-distribution is uniquely defined. If moreover the Markov matrix is reversible, the joint 2-distribution is symmetric and the θ parameters are computed from Eq. (2). Viceversa, given the θ's, the transition matrix is given by

$$P_{v\to w} = \frac{\mathrm{P}\left(X_n = v, X_{n+1} = w\right)}{\mathrm{P}\left(X_n = v\right)} = \frac{\theta_{\overline{vw}}}{2\theta_v + \sum_{z\in \mathrm{N}(v)} \theta_{\overline{vz}}}. \tag{3}$$

The mapping $\theta \mapsto (P_{v\to w} \colon (v \to w) \in \mathscr{A})$ is a rational mapping and the number of degrees of freedom is $\#V + \#\mathscr{E} - 1$.

Denoting $2\theta_v + \sum_{z\in D(v)} \theta_{\overline{vz}}$ by $\kappa(v)$, from (3), the detailed balance condition follow:

$$\kappa(v)P_{v\to w} = \kappa(w)P_{w\to v} \quad \text{and} \quad \sum_v \kappa(v) = 1.$$

2.4 A reversible Markov matrix is an auto-adjoint operator

The detailed balance condition $\pi(v)P_{v\to w} = \pi(w)P_{w\to v}$ is equivalent to P being adjoint as a linear operator on $L^2(\pi)$

$$\langle Pf, g\rangle_\pi = \langle f, Pg\rangle_\pi, \quad f, g \in L^2(\pi),$$

where $\langle Pf, g\rangle_\pi = \sum_v \left(\sum_w P_{v\to w} f(w)\right) g(v)\pi(v)$.

As $L^2(\pi)$ is isomorphic to the canonical Euclidian space $\mathbb{R}^V$ via the linear mapping

$$I \colon L^2(\pi) \ni f \mapsto \operatorname{diag}(\pi)^{1/2} f \in \mathbb{R}^V,$$

where $\operatorname{diag}(\pi)^{1/2}$ is the diagonal matrix whose (v, v) entry is $\pi(v)^{1/2}$, the Markov matrix P is mapped to the symmetric matrix

$$S = I \circ P \circ I^{-1} = \operatorname{diag}(\pi)^{1/2} P \operatorname{diag}(\pi)^{-1/2}.$$

This implies that a reversible Markov matrix is diagonalizable; in particular the left eigenvector π and the right eigenvector 1 of P are both mapped to the eigenvector $(\pi(v)^{1/2} \colon v \in V)$ of S.

Each element of the symmetric matrix S is positive if, and only if, the corresponding element of P is positive. Each reversible Markov matrix P with invariant probability π is parameterized by a unique symmetric matrix S.

Viceversa, let $s(v, w) = s(w, v) > 0$ be defined for $\overline{vw} \in \mathscr{E}$. Extend s to all pairs $v \neq w$, $\overline{vw} \notin \mathscr{E}$ by $s(v, w) = 0$. If

$$\sum_{w \colon w\neq v} s(v, w)\sqrt{\pi(w)} \leq \sqrt{\pi(v)}, \quad v \in V, \tag{4}$$

we can define

$$S(v, v) = \frac{1}{\sqrt{\pi(v)}} \left(\sqrt{\pi(v)} - \sum_{w \neq v} s(v, w) \sqrt{\pi(w)} \right) \geq 0$$

to get a symmetric non-negative matrix $S = [s(v, w)],\ v, w \in V$. The matrix

$$P = \operatorname{diag}(\pi)^{-1/2} S \operatorname{diag}(\pi)^{1/2}$$

has non-negative entries and it is a Markov matrix because

$$\sum_w P_{v \to w} = \sum_w \pi(v)^{-1/2} s(v, w) \pi(w)^{1/2} = 1, \quad v \in V.$$

The matrix P satisfies the detailed balance equations

$$\pi(v) P_{v \to w} = \pi(v)^{1/2} \pi(w)^{1/2} s(v, w) = \pi(w) P_{w \to v}.$$

We can rephrase the computations above as follows.

Proposition 3 *The set of Markov matrices which have structure $\mathcal{G}$ and are reversible with invariant probability π is parameterized by*

$$\begin{cases} P_{v \to w} = \pi(v)^{-1/2} \pi(w)^{1/2} s(v, w), & (v \to w) \in \mathcal{A}, \\ P_{v \to v} = 1 - \sum_{w \in \mathrm{N}(v)} P_{v \to w} \end{cases}$$

the polytope of all positive weight functions $s : \mathcal{E} \to \mathbb{R}_{>0}$ and (unnormalized) positive probability π satisfying the inequalities (4).

2.5 Kolmogorov's theorem

Let $\mathcal{G} = (V, \mathcal{E})$ be a connected graph. For each closed path $\omega = v_0 v_1 \ldots v_n v_0$, we denote by $r(\omega)$ the reversed path $r(\omega) = v_0 v_n \ldots v_1 v_0$. The Kolmogorov's characterization of reversibility based on closed paths is well known. However, we give here a variation of the proof by Suomela (1979) as it is an introduction to the algebraic arguments in the next section. The proof has been modified to allow null transitions. If $\gamma = v_0 v_1 \cdots v_{n-1} v$ is any path connecting v_0 to v we write P^γ to denote the product of transitions along γ, i.e. $P^\gamma = \prod_{i=1,n} P_{v_{i-1} \to v_i}$.

Theorem 1 (Kolmogorov's theorem) *The Markov irreducible matrix P is reversible if, and only if, for all closed paths ω*

$$P^\omega = P^{r(\omega)}. \tag{5}$$

Proof Assume that the process is reversible. By multiplying together all detailed balance equations

$$\kappa(v_i)P_{v_i\to v_{i+1}} = \kappa(v_{i+1})P_{v_{i+1}\to v_i}, \quad i=0,1,\dots n, v_{n+1}=v_0,$$

and clearing the κ's we obtain (5).

Viceversa, assume that all closed path have property (5). Fix a vertex v_0 and consider a generic path γ from v_0 to v. First we prove that there exists a positive constant $\kappa(v)$, depending only on v, such that $P^\gamma = \kappa(v)P^{r(\gamma)}$. In fact, for any other path $\gamma' = v_0v'_1\dots v'_{n'}v$ with the same endpoints v_0 and v, $\gamma r(\gamma')$ is a closed path. Denoting by $k'(v)$ the corresponding constant, Kolmogorov's condition (5) implies $k(v)=k'(v)$. Moreover, for any vertex w connected with v, consider the path γw and the corresponding constant $k(w)$. We have:

$$\begin{aligned} P^\gamma P_{v\to w}P_{w\to v} &= k(v)P^{r\gamma}P_{v\to w}P_{w\to v}\\ P^{\gamma w}P_{w\to v} &= k(v)P^{r(\gamma w)}P_{v\to w}\\ k(w)P^{r(\gamma w)}P_{w\to v} &= k(v)P^{r(\gamma w)}P_{v\to w}\end{aligned}$$

i.e. the detailed balance condition on w and v. □

In the next section, we will discuss the algebraic interpretation of Kolmogorov condition.

3 Algebraic theory

The present section is devoted to the algebraic structure implied by the Kolmogorov's theorem for reversible Markov chains. We refer mainly to the textbook by Bollobás (1998) for graph theory, and to the textbooks Cox et al. (1997) and Kreuzer and Robbiano (2000) for computational commutative algebra. The theory of toric ideals is treated in detail in Sturmfels (1996) and Bigatti and Robbiano (2001). General references for algebraic methods in Stochastics are, e.g. Drton et al. (2009), Gibilisco et al. (2010). Graver bases are presented in Sturmfels (1996) and Onn (2011).

3.1 Kolmogorov's ideal

We denote by $\mathscr{G}=(V,\mathscr{E})$ an undirected graph. We split each edge into two opposite arcs to get a connected directed graph (without loops) denoted by $\mathscr{D}=(V,\mathscr{A})$. The arc going from vertex v to vertex w is denoted by $v\to w$ or $(v\to w)$. The graph $\mathscr{D}$ is such that $(v\to v)\notin\mathscr{A}$ and $(v\to w)\in\mathscr{A}$ if, and only if, $(w\to v)\in\mathscr{A}$. Because of our application to Markov chains, we want two arcs on each edge, see Example 1.

The *reversed* arc is the image of the 1-to-1 function $r:\mathscr{A}\to\mathscr{A}$ defined by $r(v\to w)=(w\to v)$. A *path* is a sequence of vertices $\omega = v_0v_1\cdots v_n$ such that $(v_{k-1}\to v_k)\in\mathscr{A}$, $k=1,\dots,n$. The reversed path is denoted by $r(\omega)=v_nv_{n-1}\cdots v_0$. Equivalently, a path is a sequence of inter-connected arcs $\omega = a_1\dots a_n$, $a_k=(v_{k-1}\to v_k)$, and $r(\omega)=r(a_n)\dots r(a_1)$.

A *closed path* $\omega = v_0 v_1 \cdots v_{n-1} v_0$ is any path going from a vertex v_0 to itself; $r(\omega) = v_0 v_{n-1} \cdots v_1 v_0$ is the reversed closed path. In a closed path, any vertex can be the initial and final vertex. If we do not distinguish any initial vertex, the equivalence class of paths is called a *circuit*. A closed path is *elementary* if it has no proper sub-closed-path, i.e. if does not meet twice the same vertex except the initial one v_0. The circuit of an elementary closed path is a *cycle*. We denote by $\mathscr{C}$ the set of cycles of $\mathscr{D}$.

Consider the commutative indeterminates $P = [P_{v \to w}]$, $(v \to w) \in \mathscr{A}$, and the polynomial ring $k[P_{v \to w} : (v \to w) \in \mathscr{A}]$, i.e. the set of all polynomials in the indeterminates P and coefficients in the number field k.

For each path $\omega = a_1 \cdots a_n$, where $a_k \in \mathscr{A}$, $k = 1, \ldots, n$, we define the monomial term P^ω,

$$\omega = a_1 \cdots a_n \mapsto P^\omega = \prod_{k=1}^{n} P_{a_k}.$$

For each $a \in \mathscr{A}$, let $N_a(\omega)$ be the number of traversals of the arc a by the path ω. Hence,

$$P^\omega = \prod_{a \in \mathscr{A}} P_a^{N_a(\omega)}.$$

Note that $\omega \mapsto P^\omega$ is a representation of the non-commutative concatenation of arcs on the commutative product of indeterminates. Two closed paths associated to the same circuit are mapped to the same monomial term because they have the same traversal counts. The monomial term of a cycle is square-free because no arc is traversed twice.

Figure 2 presents six cycles of the running example. This list of cycles is larger than a basis of cycles in the undirected graph $\mathscr{G}$, for instance $\{\omega_A, \omega_B\}$. We will see below that all directed cycles are needed for the algebraic argument.

Definition 1 (K-ideal) The *Kolmogorov's ideal* or *K-ideal* of the graph $\mathscr{G}$ is the ideal of they ring $k[P_{v \to w} : (v \to w) \in \mathscr{A}]$ generated by the binomials $P^\omega - P^{r(\omega)}$, where ω is *any circuit*. The K-variety is the k-affine variety of the K-ideal.

Remark 2 The equivalence class of the closed path $\omega = v_0 v_1 v_0$ is, according to our definition, a special circuit, in particular a cycle. The corresponding binomial

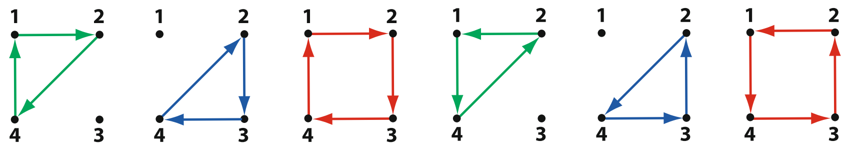

Fig. 2 The six cycles (not back and forth) of the graph on Fig. 1: $\omega_A = (1 \to 2)(2 \to 4)(4 \to 1)$, $\omega_B = (2 \to 3)(3 \to 4)(4 \to 2)$, $\omega_C = (1 \to 2)(2 \to 3)(3 \to 2)(4 \to 1)$, $\omega_D = (\omega_A)$, $\omega_E = r(\omega_B)$, $\omega_F = r(\omega_C)$

$P^{\omega} - P^{r(\omega)}$ is zero, because $P_{v_0 \to v_1} P_{v_1 \to v_0} - P_{v_1 \to v_0} P_{v_0 \to v_1} = 0$. If a closed path ω contains $v_0 v_1 v_0$, then $P^{\omega} - P^{r(\omega)} = P_{v_0 \to v_1} P_{v_1 \to v_0} (P^{\omega_1} - P^{r(\omega_1)})$, where ω_1 is the path obtained with the shortening $v_0 v_1 v_0 \to v_0$. It follows that the Kolmogorov ideal is generated by paths that do not show this back-and-forth behaviour. *In the rest of the paper all circuits and cycles are assumed to have this property.*

Our main application concerns the real case $k = \mathbb{R}$, but the combinatorial structure of the K-ideal does not depend on the choice of a specific field. An interesting choice for computations could be the Galois field $k = \mathbb{Z}_2$.

Proposition 4 (Examples of K-ideals) *Let the Markov matrix P with structure $\mathscr{G}$ be reversible.*

1. *The real vector $P_{v \to w}$, $(v \to w) \in \mathscr{A}$, is a point of the intersection of the variety of the K-ideal with $S(\mathscr{D}) = \times_{v \in V} S(v)$, where*

$$S(v) = \left\{ P_a \in \mathbb{R}_+^{\text{out}(v)} : \sum_{a \in \text{out}(v)} P_a(w) \leq 1 \right\}.$$

2. *Let $(X_n)_{n \geq 0}$ be the stationary Markov chain with transition P. Then the real vector of joint probabilities $p(v, w) = \mathrm{P}\,(X_n = u, X_{n+1} = v)$, $(v \to w) \in \mathscr{A}$, is a point in the intersection of the K-variety and the simplex*

$$S(\mathscr{A}) = \left\{ p \in \mathbb{R}_+^{\mathscr{A}} : \sum_{a \in \mathscr{A}} P(a) \leq 1 \right\}.$$

Proof 1. It is the first part of the Kolmogorov's theorem.

2. Let $\omega = v_0 \dots v_n v_0$ be a closed path. If π is the stationary probability, by multiplying the Kolmogorov's equations by the product of the initial probabilities at each transition, we obtain

$$\begin{aligned} &\pi(v_0)\pi(v_1)\cdots\pi(v_n) P_{v_0 \to v_1} \cdots P_{v_n \to v_0} \\ &\quad = \pi(v_0)\pi(v_n)\cdots\pi(v_n) P_{v_0 \to v_n} \cdots P_{v_1 \to v_0}, \end{aligned}$$

hence

$$p(v_0, v_1) p(v_1, v_2) \cdots p(v_n, v_0) = p(v_0, v_n) p(v_n, v_{n-1}) \cdots p(v_1, v_0).$$

However, in this case the Kolmogorov's equations are trivially satisfied as $p(v, w) = p(w, v)$. □

The K-ideal has a finite basis because of the Hilbert's basis theorem. Precisely, a finite basis is obtained by restricting to cycles, which are finite in number. We underline that here we consider all the cycles, not just a generating set of cycles. The related result by Mitrophanov (2004, Th. 1) is discussed in the next subsection.

Proposition 5 (Cycle basis of the K-ideal) *The K-ideal is generated by the set of binomials $P^{\omega} - P^{r(\omega)}$, where ω is cycle.*

Proof Remember we consider only circuits without back-and-forth of the type vwv. Let $\omega = v_0 v_1 \cdots v_0$ be a closed path which is not elementary and consider the least $k \geq 1$ such that $v_k = v_{k'}$ for some $k' < k$. Then the sub-path ω_1 between the k'th vertex and the kth vertex is an elementary closed path and the residual path $\omega_2 = v_0 \cdots v_{k'} v_{k+1} \cdots v_0$ is closed and shorter than the original one. The arcs of ω are in 1-to-1 correspondence with the arcs of ω_1 and ω_2, hence $N_a(\omega) = N_a(\omega_1) + N_a(\omega_2)$, $a \in \mathscr{A}$. The procedure can be iterated and stops in a finite number of steps. Hence, given any closed path ω, there exists a finite sequence of cycles $\omega_1, \ldots, \omega_l$, such that the list of arcs in ω is partitioned into the lists of arcs of the ω_i's. From $P^{\omega_i} - P^{r(\omega_i)} = 0$, $i = 1, \ldots, l$, it follows

$$P^{\omega} = \prod_{i=1}^{l} P^{\omega_i} = \prod_{i=1}^{l} P^{r(\omega_i)} = P^{r(\omega)}.$$

□

The K-ideal is generated by a finite set of binomials and this set has the same number of elements as the set of undirected cycles of $\mathscr{G}$. This is an involved way to prove all stationary Markov processes on a tree are reversible (Kelly 1979, Lemma 1.5).

The cycle basis of Proposition 5 belongs to the special class of bases, namely Gröbner bases. We refer to the textbooks Cox et al. (1997) and Kreuzer and Robbiano (2000) for a detailed discussion. We review the basic definitions of this theory, which is based on the existence of a monomial order $\succ$, i.e. a total order on monomial terms which is compatible with the product. Given such an order, the leading term $\mathrm{LT}(f)$ of the polynomial f is defined. A generating set is a *Gröbner basis* if the set of leading terms of the ideal is generated by the leading terms of monomials of the generating set. A Gröbner basis is *reduced* if the coefficient of the leading term of each element of the basis is 1 and no monomial in any element of the basis is in the ideal generated by the leading terms of the other element of the basis. The Gröbner basis property depends on the monomial order. However, a generating set is said to be a *universal* Gröbner basis if it is a Gröbner basis for all monomial orders.

The finite algorithm for testing the Gröbner basis property depends on the definition of *syzygy*. The syzygy of two polynomials f and g is the polynomial

$$\mathrm{Syz}(f, g) = \frac{\mathrm{LT}(g)}{\gcd(\mathrm{LT}(f), \mathrm{LT}(g))} f - \frac{\mathrm{LT}(f)}{\gcd(\mathrm{LT}(f), \mathrm{LT}(g))} g.$$

A generating set of an ideal is a Gröbner basis if, and only if, it contains the syzygy $\mathrm{Syz}(f, g)$ whenever it contains the polynomials f and g (see Cox et al. 1997, Ch 6 or Kreuzer and Robbiano 2000, Th. 2.4.1 p. 111).

Proposition 6 (Universal G-basis) *The cycle basis of the K-ideal is a reduced universal Gröbner basis.*

Proof Choose any monomial order $\succ$ and let ω_1 and ω_2 be two cycles with $\omega_i \succ r(\omega_i)$, $i = 1, 2$. Assume first they do not have any arc in common. In such a case $\gcd(P^{\omega_1}, P^{\omega_2}) = 1$ and the syzygy is

$$\begin{aligned}&\mathrm{Syz}(P^{\omega_1} - P^{r(\omega_1)}, P^{\omega_2} - P^{r(\omega_2)})\\ &\quad = P^{\omega_2}(P^{\omega_1} - P^{r(\omega_1)}) - P^{\omega_1}(P^{\omega_2} - P^{r(\omega_2)}) = P^{\omega_1}P^{r(\omega_2)} - P^{r(\omega_1)}P^{\omega_2},\end{aligned}$$

which belongs to the K-ideal.

Let now α be the common part, that is $\gcd(P^{\omega_1}, P^{\omega_2}) = P^{\alpha}$. The syzygy of $P^{\omega_1} - P^{r(\omega_1)}$ and $P^{\omega_2} - P^{r(\omega_2)}$ is

$$P^{\omega_1-\alpha}P^{r(\omega_2)} - P^{\omega_2-\alpha}P^{r(\omega_1)} = P^{r\alpha}(P^{\omega_1-\alpha}P^{r(\omega_2)-r\alpha} - P^{\omega_2-\alpha}P^{r(\omega_1)-r\alpha}),$$

which again belongs to the K-ideal because $\omega_1 - \alpha + r(\omega_2 - \alpha)$ is a cycle. In fact $\omega_1 - \alpha$ and $\omega_2 - \alpha$ have in common the extreme vertices, corresponding to the extreme vertices of α. Notice that α is the common part of ω_1 and ω_2 only if it is traversed in the same direction by the both cycle. The previous proof does not depend on the choice of the leading term of the binomials, therefore the Gröbner basis is universal. The Gröbner basis is reduced because no monomial of a cycle can divide a monomial of a different cycle. □

Example 2 (Running example continue) Figure 3 is an illustration of the proof. In fact, from

$$\begin{aligned}P_{1\to2}P_{2\to4}P_{4\to1} &= P_{1\to4}P_{4\to2}P_{2\to1}\\ P_{1\to2}P_{2\to3}P_{3\to4}P_{4\to1} &= P_{1\to4}P_{4\to3}P_{3\to2}P_{2\to1}\end{aligned}$$

it follows

$$P_{4\to1}P_{1\to2}P_{2\to1}P_{1\to4}\,(P_{2\to3}P_{3\to4}p_{4\to2} - P_{2\to4}P_{4\to3}P_{3\to2}) = 0$$

which, in turn, gives the binomial of ω_B if $P_{4\to1}P_{1\to2}P_{2\to1}P_{1\to4} \neq 0$ and, therefore, the factor $P^{\alpha}P^{r(\alpha)}$ can be cleared. This is confirmed by the use of a symbolic algebraic software such as CoCoA, see CoCoATeam (online). This computation shows that the ω_B equation does not belong to the ideal generated by the ω_A, ω_C equations unless we add the condition $P_{4\to1}P_{1\to2}P_{2\to1}P_{1\to4} \neq 0$. Notice that ω_A and ω_C are the cycles obtained from the spanning tree $3 \to 4 \to 1 \to 2$.

Example 3 (Running example: Monomial basis of the quotient ring) Take any order on vertexes, e.g. $1 \prec 2 \prec 3 \prec 4$, and derive a lexicographic order on arcs: $1 \to 2 \prec 1 \to 4 \prec 2 \to 1 \prec 2 \to 3 \prec 2 \to 4 \prec 3 \to 2 \prec 3 \to 4 \prec 4 \to 1 \prec 4 \to 2 \prec 4 \to 3$. Take the same order on indeterminates P_a, $a \in \mathscr{A}$, and the lexicographic order on monomials. We check that the leading terms of the binomials in the G-basis are $P^{r(\omega_A)}$, $P^{r(\omega_B)}$, $P^{r(\omega_C)}$, see Figure 2. The exponents of the leading terms of the

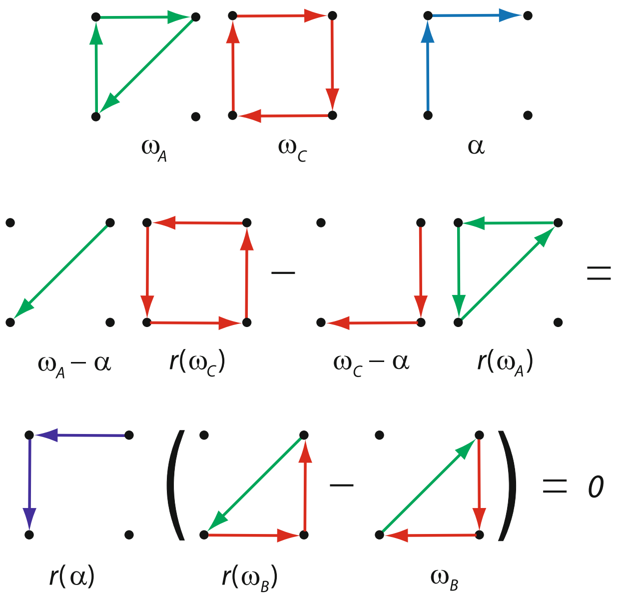

Fig. 3 Running example: illustration of the proof of Proposition 6

G-basis are

$\mathscr{A}$	$1\to 2$	$1\to 4$	$2\to 1$	$2\to 3$	$2\to 4$	$3\to 2$	$3\to 4$	$4\to 1$	$4\to 2$	$4\to 3$
$N(r(\omega_A))$	0	1	1	0	0	0	0	0	1	0
$N(r(\omega_B))$	0	0	0	0	1	1	0	0	0	1
$N(r(\omega_C))$	0	1	1	0	0	1	0	0	0	1

Each monomial $P^N = \prod_{a\in\mathscr{A}} P_a^{N_a}$, $N \in \mathbb{Z}_{\geq}^{\mathscr{A}}$, is reduced by the K-ideal to a monomial whose exponent does not contain any of the counts in the table. E.g. $P_{1\to 4}P_{2\to 1}P_{3\to 2}$ is an element of the monomial basis of the quotient of the polynomial ring by the K-ideal.

3.2 Cycle and cocycle spaces

We adapt to our context some standard tools of algebraic graph theory, namely the cycle an cocycle spaces, see e.g. (Bollobás 1998, II.3).

Let $\mathscr{C}$ be the set of cycles without back-and-forth. For each such cycle $\omega \in \mathscr{C}$ we define the *cycle vector* of ω to be $z(\omega) = (z_a(\omega) : a \in \mathscr{A})$, where

$$z_a(\omega) = \begin{cases} +1 & \text{if } a \text{ is an arc of } \omega, \\ -1 & \text{if } r(a) \text{ is an arc of } \omega, \\ 0 & \text{otherwise.} \end{cases}$$

We observe that the definition above makes no sense for a back-and-forth, this case being of no use for our purpose.

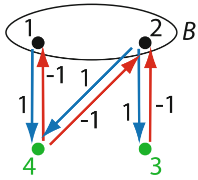

Fig. 4 Example of cocycle vector. All arcs not shown take value 0

Note that $z_{r(a)}(\omega) = -z_a(\omega)$. If z^+ and z^- are the positive and the negative part of z, respectively, then $z_a^+(\omega) = N_a(\omega)$ and $z_a^-(\omega) = N_a(r(\omega))$. It follows that $P^\omega = P^{N(\omega)} = P^{z^+(\omega)} = \prod_{a\in\mathscr{A}} P_a^{z_a^+(\omega)}$ and

$$P^\omega - P^{r(\omega)} = P^{z^+(\omega)} - P^{z^-(\omega)}. \tag{6}$$

More generally, the definition can be is extended to any circuit ω by defining

$$z_a(\omega) = N_a(\omega) - N_{r(a)}(\omega).$$

The equality $z^+(\omega) = N(\omega)$ holds if, and only if, $a \in \omega$ implies $r(a) \notin \omega$, $a \in \mathscr{A}$.

Let $Z(\mathscr{D})$ be the *cycle space*, i.e. the vector space generated in $\mathbb{R}^{\mathscr{A}}$ by the cycle vectors.

For each proper subset B of the set of vertices, $\emptyset \neq B \subsetneq V$ we define the *cocycle vector* of B to be $u(B) = (u_a(B) : a \in \mathscr{A})$, with

$$u_a(B) = \begin{cases} +1 & \text{if } a \text{ exits from } B, \\ -1 & \text{if } a \text{ enters into } B, \\ 0 & \text{otherwise.} \end{cases} \qquad a \in \mathscr{A}$$

See an example in Fig. 4. Note that $u_{r(a)}(B) = -u_a(B)$.

Let $U(\mathscr{D})$ be the *cocycle space*, i.e. the vector space generated in $\mathbb{R}^{\mathscr{A}}$ by the cocycle vectors. Let U be the matrix whose rows are the cocycle vectors $u(B)$, $\emptyset \neq B \subsetneq V$. The matrix $U = [u_a(B)]_{\emptyset\neq B\subsetneq V, a\in\mathscr{A}}$ is the *cocycle matrix*.

The cycle space and the cocycle space are orthogonal in $\mathbb{R}^{\mathscr{A}}$. In fact, for each cycle vector $z(\omega)$ and cocycle vector $u(B)$, we have

$$z_{r(a)}(\omega)u_{r(a)}(B) = (-z_a(\omega))(-u_a(B)) = z_a(\omega)u_a(B), \quad a \in \mathscr{A},$$

so that

$$\begin{aligned} z(\omega)\cdot u(B) &= \sum_{a\in\mathscr{A}} z_a(\omega)u_a(B) = \sum_{a\in\omega} z_a(\omega)u_a(B) + \sum_{r(a)\in\omega} z_a(\omega)u_a(B) \\ &= 2\sum_{a\in\omega} z_a(\omega)u_a(B) = 2\left[\sum_{a\in\omega, u_a(B)=+1} 1 - \sum_{a\in\omega, u_a(B)=-1} 1\right] = 0. \end{aligned}$$

Table 1 Running example

	$1\to 2$	$1\to 4$	$2\to 3$	$2\to 4$	$3\to 4$	$2\to 1$	$4\to 1$	$3\to 2$	$4\to 2$	$4\to 3$
$\overline{12}$	1	0	0	0	0	1	0	0	0	0
$\overline{14}$	0	1	0	0	0	0	1	0	0	0
$\overline{23}$	0	0	1	0	0	0	0	1	0	0
$\overline{24}$	0	0	0	1	0	0	0	0	1	0
$\overline{34}$	0	0	0	0	1	0	0	0	0	1
{1}	**1**	**1**	**0**	**0**	**0**	**−1**	**−1**	**0**	**0**	**0**
{2}	−1	0	1	1	0	1	0	−1	−1	0
{3}	**0**	**0**	**−1**	**0**	**1**	**0**	**0**	**1**	**0**	**−1**
{4}	0	−1	0	−1	−1	0	1	0	1	1
{12}	**0**	**1**	**1**	**1**	**0**	**0**	**−1**	**−1**	**−1**	**0**
{13}	1	1	−1	0	1	−1	−1	1	0	−1
{14}	1	0	0	−1	−1	−1	0	0	1	1
{23}	−1	0	0	1	1	1	0	0	−1	−1
{24}	−1	−1	1	0	−1	1	1	−1	0	1
{34}	0	−1	−1	−1	0	0	1	1	1	0
{123}	0	1	0	1	1	0	−1	0	−1	−1
{124}	0	0	1	0	−1	0	0	−1	0	1
{134}	1	0	−1	−1	0	−1	0	1	1	0
{234}	−1	−1	0	0	0	1	1	0	0	0
$z(\omega_A)$	1	−1	0	1	0	−1	1	0	−1	0
$z(\omega_B)$	0	0	1	−1	1	0	0	−1	1	−1

Model matrix A and a basis of $\mathbb{Z}(\mathscr{D})$

It is shown, for example, in the previous references that the cycle space is the orthogonal complement of the cocycle space for undirected graphs. In our setting, it is the orthogonal complement relative to the subspace of vectors x such that $x_{r(a)} = -x_a$. As we are interested in elements of the cycle space with integer entries, i.e. those elements $z = (z_a \colon a \in \mathscr{A})$ of the cycle space that can be exponents of a monomial $P^z = \prod_{a\in\mathscr{A}} P_a^{z_a}$, we are going to use the following matrix encoding of our problem.

Definition 2 (Model matrix) Consider the matrix $E = [E_{e,a}]_{e\in\mathscr{E},a\in\mathscr{A}}$ whose element $E_{e,a}$ in position (e, a) is 1 if the arc a is one of the directions of the edge e, zero otherwise. Let U be the cocycle matrix. The *model matrix* is the block matrix

$$A = \begin{bmatrix} E \\ U \end{bmatrix}.$$

It follows $Z(\mathscr{D}) = \ker A$. A *lattice basis* of the lattice $\mathbb{Z}(\mathscr{D}) = \ker A \cap \mathbb{Z}^{\mathscr{A}}$ is a linear basis of $\ker A$ with integer entries.

The matrix A has dimension $\#\mathscr{E} + \#V - 1$. In fact E can be re-arranged as $[I_{\#\mathscr{E}} | I_{\#\mathscr{E}}]$, with $I_{\#\mathscr{E}}$ the identity matrix, and U has $\#V - 1$ linearly independent rows, the dimension of the cocycle space.

Example 4 (Running example continue) Table 1 shows the matrix A and a lattice basis of $\mathbb{Z}(\mathscr{D})$ computed with CoCoA. The top matrix is the E matrix; the bottom matrix is the U matrix, where three linearly independent rows are highlighted. The two bottom row vectors are the lattice basis.

3.3 Toric ideal

We want to show that the K-ideal is the toric ideal of the model matrix A, see Definition 2. Basic definitions and theory are in (Sturmfels, 1996, Ch 4), see also Bigatti and Robbiano (2001).

Consider the polynomial ring $\mathbb{Q}[P_a \colon a \in \mathscr{A}]$ and the Laurent polynomial ring $\mathbb{Q}[t_e^{\pm 1}, t_B^{\pm 1} \colon e \in \mathscr{E}, \emptyset \neq B \subsetneq V]$, together with their homomorphism h defined by

$$h\colon P_a \longmapsto \prod_e t_e^{E_{e,a}} \prod_B t_B^{u_a(B)} = t^{A(a)}. \tag{7}$$

As $E_{e,a} = E_{e,r(a)}$ for all edge e and arc a, the first factor in (7) is a symmetric function $s(v,w) = \prod_e t_e^{E_{e,v\to w}} = s(w,v)$. We could write

$$h\colon P_{v\to w} \longmapsto s(v,w) \prod_B t_B^{u_{v\to w}(B)}. \tag{8}$$

The kernel $I(A)$ of h is called the *toric ideal* of A,

$$I(A) = \left\{ f \in \mathbb{Q}[P_a : a \in \mathscr{A}] : f(t^{A(a)} \colon a \in \mathscr{A}) = 0 \right\}.$$

The toric ideal $I(A)$ is a prime ideal and the binomials

$$P^{z^+} - P^{z^-}, \quad z \in \mathbb{Z}^{\mathscr{A}}, \quad Az = 0,$$

are a generating set of $I(A)$ as a $\mathbb{Q}$-vector space. A finite generating set of the ideal is formed by selecting a finite subset of such binomials. The basis we find is a Graver basis.

We recall the definition of Graver basis as it is given in De Loera et al. (2008) and we apply it to the *cycle lattice* $\mathbb{Z}(\mathscr{D})$. We introduce a partial order and its set of minimal elements as follows.

Definition 3 (*Graver basis*) Let z_1 and z_2 be two element of the cycle lattice $\mathbb{Z}(\mathscr{D})$.

1. z_1 is *conformal* to z_2, relation denoted by $z_1 \sqsubseteq z_2$, if the component-wise product is non-negative (i.e. z_1 and z_2 are in the same quadrant) and $|z_1| \leq |z_2|$ component-wise, i.e. $z_{1,a} z_{2,a} \geq 0$ and $z_{1,a} \leq z_{2,a}$ for all $a \in \mathscr{A}$.
2. A *Graver basis* of $\mathbb{Z}(\mathscr{D})$ is the set of the minimal elements with respect to the conformity partial order $\sqsubseteq$.

Proposition 7

1. *For each cycle vector* $z \in \mathbb{Z}(\mathscr{D})$, $z = \sum_{\omega \in \mathscr{C}} \lambda(\omega) z(\omega)$, $\lambda(\omega) \in \mathbb{Q}$, *there exist cycles* $\omega_1, \ldots, \omega_n \in \mathscr{C}$ *and positive integers* $\alpha(\omega_1), \ldots, \alpha(\omega_n)$, *such that* $z^+ \geq z^+(\omega_i)$, $z^- \geq z^-(\omega_i)$, $i = 1, \ldots, n$, *and*

$$z = \sum_{i=1}^{n} \alpha(\omega_i) z(\omega_i).$$

2. *The set* $\{z(\omega)\colon \omega \in \mathscr{C}\}$ *is a* Graver basis *of* $\mathbb{Z}(\mathscr{D})$.

Proof 1. For all $\omega \in \mathscr{C}$ we have $-z(\omega) = z(r(\omega))$, so that we can assume all the $\lambda(\omega)$'s to be non-negative. Notice also that we can arrange things in such a way that at most one of the two direction of each cycle has a non-zero coefficient. We define

$$\mathscr{A}_+(z) = \{a \in \mathscr{A} : z_a > 0\}, \quad \mathscr{A}_-(z) = \{a \in \mathscr{A} : z_a < 0\},$$

and consider two subgraph of $\mathscr{D}$ with a the set of arcs restricted to $\mathscr{A}_+$, $\mathscr{A}_-$, respectively. We note that $r\mathscr{A}_+(z) = \mathscr{A}_-(z)$ and $r\mathscr{A}_-(z) = \mathscr{A}_+(z)$; in particular, both $\mathscr{A}_+(z)$, $\mathscr{A}_-(z)$ are not empty. We drop from now on the dependence on z for ease of notation.

We show first there is a cycle whose arcs are in $\mathscr{A}_+$. If not, if a cycle of full graph $\mathscr{D}$ has one arc in $\mathscr{A}_+$, it would exists vertex v such that $\operatorname{out}(v) \cap \mathscr{A}_+ = \emptyset$ while $\operatorname{in}(v) \cap \mathscr{A}_+ \neq \emptyset$. Let $u(v)$ be the cocycle vector of $\{v\}$; we derive a contradiction to the assumption $z \cdot u(v) = 0$. In fact,

$$\begin{aligned} z \cdot u(v) &= \sum_{a \in \mathscr{A}_+} z_a u_a(v) + \sum_{a \in \mathscr{A}_-} z_a u_a(v) \\ &= \sum_{a \in \mathscr{A}_+} z_a u_a(v) + \sum_{a \in \mathscr{A}_+} z_{r(a)} u_{r(a)}(v) = 2 \sum_{a \in \mathscr{A}_+} z_a u_a(v) \neq 0 \end{aligned}$$

because each of the terms $z_a u_a(v)$, $a \in \mathscr{A}_+$, is either 0 or equal to $-z_a < 0$ if $a \in \operatorname{in}(v)$.

For each cycle ω_1 in $\mathscr{A}_+$ let $\alpha(\omega_1) \geq 1$ be an integer such that $z^+ - \alpha(\omega_1) z^+(\omega_1) \geq 0$ and it is zero for at least one a. The vector $z^1 = z - \alpha(\omega_1) z(\omega_1)$ belongs to the cycle space $\mathbb{Z}(\mathscr{D})$, and moreover $\mathscr{A}_+(z^1) \subset \mathscr{A}_+(z)$.

By repeating the same step a finite number of times we obtain a new representation of z in the form $z = \sum_{i=1}^{n} \alpha(\omega_i) z(\omega_i)$ where the support of each $\alpha(\omega_i) z^+(\omega_i)$ is contained in $\mathscr{A}_+$. It follows

$$z^+ = \sum_{i=1}^{n} \alpha(\omega_i) z^+(\omega_i) \quad \text{and} \quad z^- = \sum_{i=1}^{n} \alpha(\omega_i) z^-(\omega_i). \tag{9}$$

2. In the previous decomposition each $z(\omega_i)$, $i = 1, \ldots, n$, is conformal to z. In fact, from $z^+ \geq z^+(\omega_i)$ and $z^- \geq z^-(\omega_i)$, it follows $z_a z_a(\omega_i) = z_a^+ z_a^+(\omega_i) - z_a^- z_a^-(\omega_i) \geq 0$ and $|z_a(\omega_i)| = z_a^+(\omega_i) - z_a^-(\omega_i) \leq z_a^+ + z_z^- = |z_a|$. Therefore $z(\omega_i) \sqsubseteq z$. □

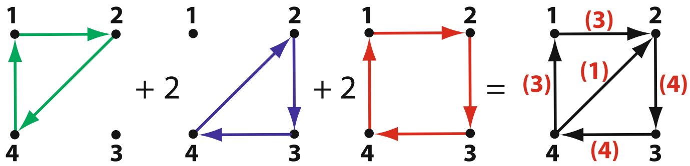

Fig. 5 Running example. Cycle space

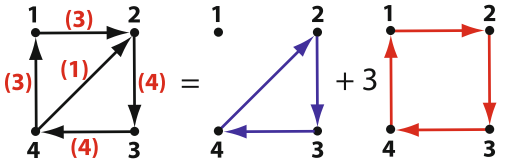

Fig. 6 Running example. Computation of the conformal representation of the z of Fig. 5

Example 5 (Running example continue) We give an illustration of the previous proof. Consider the cycle vectors

	1 → 2	2 → 1	2 → 3	3 → 2	3 → 4	4 → 3	4 → 1	1 → 4	2 → 4	4 → 2
$z(\boldsymbol{\omega_A}) =$ (	1	−1	0	0	0	0	1	−1	1	−1)
$z(\boldsymbol{\omega_B}) =$ (	0	0	1	−1	1	−1	0	0	−1	1)
$z(\boldsymbol{\omega_C}) =$ (	1	−1	1	−1	1	−1	1	−1	0	0)

and the element of the cycle space $z = z(\omega_A) + 2z(\omega_B) + 2z(\omega_C)$, see Figure 5. We have

$$
\begin{aligned}
z &= z(\boldsymbol{\omega_A}) + 2z(\boldsymbol{\omega_B}) + 2z(\boldsymbol{\omega_C}) = (3, -3, 4, -4, 4, -4, 0, 0, -1, 1)\\
z^+ &= z^+(\boldsymbol{\omega_B}) + 3z^+(\boldsymbol{\omega_C}) = (3, 0, 4, 0, 4, 0, 0, 0, 0, 1)
\end{aligned}
$$

as it is illustrated in Fig. 6.

Theorem 2 (The K-ideal is toric)

1. *The K-ideal is the toric ideal of the matrix A.*
2. *The binomials of the cycles form a Graver basis of the K-ideal.*

Proof 1. For each cycle ω the cycle vector $z(\omega)$ belongs to $\mathbb{Z}(\mathscr{D})$. From Eq. (6), $P^{z^+(\omega)} - P^{z^-(\omega)} = P^{\omega} - P^{r(\omega)}$, therefore the K-ideal is contained in the toric ideal $I(A)$ because of Proposition 5.

To prove the equality we must show that each binomial in $I(A)$ belongs to the K-ideal. From Proposition 7.1, it follows that

$$
P^{z^+} - P^{z^-} = \prod_{i=1}^{n} (P^{z^+(\omega_i)})^{\alpha(\omega_i)} - \prod_{i=1}^{n} (P^{z^-(\omega_i)})^{\alpha(\omega_i)}
$$

belongs to the K-ideal.

2. The Graver basis of a toric ideal is the set of binomials whose exponents are the positive and negative parts of a Graver basis of $\mathbb{Z}(\mathscr{D})$. From Propositions 7 and the previous item the proof follows. □

3.4 Non-zero K-ideal

The knowledge that the K-ideal is toric is relevant, because the homomorphism definition in Eq. (8) provides a parametric representation of the variety. In particular, the strictly positive P_a, $a \in \mathscr{A}$, are given by:

$$\begin{aligned} P_{v\to w} &= s(v,w)\prod_B t_B^{u_{v\to w}(B)} \\ &= s(v,w)\prod_{B:\, v\in B, w\notin B} t_B \prod_{B:\, w\in B, v\notin B} t_B^{-1}, \quad s(v,w)>0, \quad t_B>0. \end{aligned} \tag{10}$$

We observe that the first set of parameters, $s(v,w)$, is a function of the edge, while the second set of parameters, t_B, represents the deviation from symmetry. In fact, as

$$P_{w\to v} = s(w,v)\prod_B t_B^{u_{w\to v}(B)} = s(v,w)\left(\prod_B t_B^{u_{v\to w}(B)}\right)^{-1},$$

we have $P_{v\to w} = \left(\prod_B t_B^{u_{v\to w}(B)}\right)^2 P_{w\to v}$, so that $P_{v\to w} = P_{w\to v}$ if, and only if,

$$\left(\prod_B t_B^{u_{v\to w}(B)}\right)^2 = 1, \quad v \in V.$$

As the rows of E are linearly independent, the $s(v,w)$'s parameters carry $\#\mathscr{E}$ degrees of freedom to represent a generic symmetric matrix. The second set of parameters is not identifiable because the rows of the U matrix are not linearly independent. The parameterization (10) can be used to derive an explicit form of the invariant probability. All properties of the parameterization are collected in the following Proposition.

Theorem 3 *Consider the strictly non-zero points on the K-variety.*

1. *The symmetric parameters $s(e)$, $e \in \mathscr{E}$, are uniquely determined in Eq. (10). The parameters t_B, $\emptyset \neq B \subsetneq V$ are confounded by* $\ker U = \{U^t t = 0\}$.
2. *An identifiable parameterization is obtained by taking a subset of parameters corresponding to linearly independent rows, denoted by t_B, $B \subset \mathscr{S}$:*

$$P_{v\to w} = s(v,w)\prod_{B\subset\mathscr{S}:\, v\in B, w\notin B} t_B \prod_{B\subset\mathscr{S}:\, w\in B, v\notin B} t_B^{-1}. \tag{11}$$

3. *The detailed balance equations, $\kappa(v)P_{v\to w} = \kappa(w)P_{w\to v}$, are verified by*

$$\kappa(v) \propto \prod_{B\colon v\in B} t_B^{-2}. \tag{12}$$

Proof 1. We have $\log P = E^t s + U^t t$ for $P = (P_{v\to w}\colon (v \to w) \in \mathscr{A})$, $s = (s(e)\colon e \in \mathscr{E})$, $t = (t_B\colon \emptyset \neq B \subsetneq V)$. If $E^t s_1 + U^t t_1 = E^t s_2 + U^t t_2$, then $E^t(s_1 - s_2) = 0$ because the rows of E are orthogonal to the rows of U. Hence, $s_1 = s_2$ because E has full rank. Finally, $U^t t_1 = U^t t_2$.
2. The sub-matrix of A formed by E and by the rows of U in $\mathscr{S}$ has full rank.
3. Using Eqs. (10), we have:

$$\kappa(v)\, s(v,w) \prod_{B\colon v\in B, w\notin B} t_B \prod_{B\colon w\in B, v\notin B} t_B^{-1} = \kappa(w)\, s(v,w) \prod_{B\colon w\in B, v\notin B} t_B \times \prod_{B\colon v\in B, w\notin B} t_B^{-1}$$

which is equivalent to

$$\kappa(v) \prod_{B\colon v\in B, w\notin B} t_B^2 = \kappa(w) \prod_{B\colon w\in B, v\notin B} t_B^2.$$

By multiplying both terms in the equality by $\prod_{B\colon v\in B, w\in B} t_B^2$, we obtain

$$\kappa(v) \prod_{B\colon v\in B} t_B^2 = \kappa(w) \prod_{B\colon w\in B} t_B^2,$$

so that $\kappa(v) = \prod_{B\colon v\in B} t_B^{-2}$ depends only on v and satisfy the detailed balance condition. □

We are now in the position of stating an algebraic version of Kolmogorov's theorem.

Definition 4 The *detailed balance ideal* is the ideal of the ring

$$\mathbb{Q}[\kappa(v) : v \in V, P_{v\to w}, (v \to w) \in \mathscr{A}]$$

generated by the polynomials

$$\prod_{v\in V} \kappa(v) - 1,$$
$$\kappa(v)P_{v\to w} - \kappa(w)P_{v\to w}, \quad (v \to w) \in \mathscr{A}.$$

The first polynomial in the list of generators, $\prod_{v\in V} \kappa(v) - 1$, assures that the κ's are not zero.

Theorem 4

1. *A point $P = [P_{v \to w}]_{v \to w \in \mathscr{A}}$ with non-zero components belongs to the variety of the K-ideal if, and only if, there exists $\kappa = (\kappa(v) : v \in V)$ such that (κ, P) belongs to the variety of the detailed balance ideal.*
2. *The detailed balance ideal is a toric ideal.*
3. *The K-ideal is the κ-elimination ideal of the detailed balance ideal.*

Proof 1. One direction is the Kolmogorov's theorem. The other direction is a rephrasing of Item 3 of Theorem 3.
2. This ideal is the kernel of the homomorphism defined by (8), i.e. $P_{v \to w} \longmapsto s(v, w) \prod_B t_B^{u_{v \to w}(B)}$ together with $\kappa(v) \mapsto \prod_{B : v \in B} t_B^{-2}$.
3. The elimination ideal is generated by dropping the parametric equations of the indeterminates to be eliminated. □

3.5 Parameterization of reversible transitions

The parameterization in Theorem 3, Item 2, is to be compared to that in Proposition 3. Both split the parameter space into a representation of the invariant probability and a generic symmetric function. The latter is a special case of the former. In fact, it is obtained by the use of the basis of the cocycle space where each B is the set containing one vertex.

The parameterization of reversible Markov matrices is obtained by adding to the representation in Eq. (11) the relevant inequalities. Let us check first the degrees of freedom. As a reversible Markov matrix supported by a connected graph $\mathscr{G}$ is parameterized by the joint 2-distributions of the stationary Markov chain, the number of degrees of freedom is $\#V + \#\mathscr{E} - 1$, i.e. the cyclotomic number of the graph $\mathscr{G}$. In the parameterization of Proposition 3 the probability π carries $\#V - 1$ degrees of freedom, therefore s must carry $\#\mathscr{E}$ degrees of freedom.

Theorem 5 *Let P be a matrix with supporting graph $\mathscr{G} = (V, \mathscr{E})$. Let $\mathscr{S}$ be a family of subsets of V such that the cocycle vectors u_B, $B \in \mathscr{S}$, span the cocycle space.*

1. *P is a reversible Markov matrix if, and only if, there exists a non-negative symmetric function $s : V \times V \to \mathbb{R}_{>0}$ which is supported on the edges $\mathscr{E}$ and there exist positive parameters $t_B > 0$, $B \in \mathscr{S}$, such that*

$$P_{v \to w} = s(v, w) \prod_{B \in \mathscr{S} : v \in B, w \notin B} t_B \prod_{B \in \mathscr{S} : w \in B, v \notin B} t_B^{-1}, \quad (v \to w) \in \mathscr{A} \tag{13}$$

and the invariant probability is proportional to $\kappa = \prod_{B \ni v} t_B^{-2}$.

2. *For $v \in V$,*

$$\prod_{B \ni v} t_B^{-1} \geq \sum_{w \in N(v)} s(v, w) \prod_{B \ni w} t_B^{-1}.$$

3. *The parameters $s(e)$, $e \in \mathscr{E}$ and $\kappa(v)$, $v \in V$, are identifiable, while the parameters t_B, $B \in \mathscr{S}$, are identifiable if u_B, $B \in \mathscr{S}$, is a basis of the cocycle space.*

Proof 1. It follows from (11) and (12) and

$$\kappa(w)^{1/2}\kappa(v)^{-1/2} = \frac{\prod_{B:\, v\in B} t_B}{\prod_{B:\, w\in B} t_B} = \prod_{B\subset\mathscr{S}:\, v\in B, w\notin B} t_B \prod_{B\subset\mathscr{S}:\, w\in B, v\notin B} t_B^{-1}.$$

2. It follows from Eq. (13) as P is a transition probability.
3. Assume there exists two set of parameters $t^{(i)}$, $s^{(i)}$, $i = 1, 2$ giving the same P, and define $t = t^{(1)}/t^{(2)}$, $s = s^{(1)}/s^{(2)}$, $\kappa(v) = \kappa^{(1)}(v)/\kappa^{(2)}(v)$. It follows

$$1 = s(v, w) \prod_{B\in\mathscr{S}:\, v\in B, w\notin B} t_B \prod_{B\in\mathscr{S}:\, w\in B, v\notin B} t_B^{-1}, \quad (v \to w) \in \mathscr{A},$$
$$1 = s(w, v) \prod_{B\in\mathscr{S}:\, w\in B, v\notin B} t_B \prod_{B\in\mathscr{S}:\, v\in B, w\notin B} t_B^{-1}, \quad (w \to v) \in \mathscr{A},$$

hence $s(v, w)s(w, v) = s(v, w)^2 = 1$. In turn we get

$$1 = \left(\prod_{B\in\mathscr{S}:\, v\in B, w\notin B} t_B \prod_{B\in\mathscr{S}:\, w\in B, v\notin B} t_B^{-1}\right)^2 = \frac{\kappa(v)}{\kappa(w)}.$$

The identifiability of the t_B's follows from

$$\log P_{v\to w} = \log s(v, w) + \sum_{B\in\mathscr{S}} (\log t_B) u_B.$$

□

Example 6 (Running example continue) An over-parameterization of two transition probabilities of the K-variety is:

$$P_{3\to 4} = s(3, 4)\; t_{\{3\}}\; t_{\{1,3\}}\; t_{\{2,3\}}\; t_{\{1,2,3\}}\; t_{\{4\}}^{-1}\; t_{\{1,4\}}^{-1}\; t_{\{2,4\}}^{-1}\; t_{\{1,2,4\}}^{-1},$$
$$P_{4\to 3} = s(3, 4)\; t_{\{4\}}\; t_{\{1,4\}}\; t_{\{2,4\}}\; t_{\{1,2,4\}}\; t_{\{3\}}^{-1}\; t_{\{1,3\}}^{-1}\; t_{\{2,3\}}^{-1}\; t_{\{1,2,3\}}^{-1}.$$

By choosing the cocycle basis $\mathscr{S} = \{\{1\}, \{3\}, \{1, 2\}\}$, we have:

$$\begin{cases} \kappa(1) = t_{\{1\}}^{-2} t_{\{1,2\}}^{-2} \\ \kappa(2) = t_{\{1,2\}}^{-2} \\ \kappa(3) = t_{\{3\}}^{-2} \\ \kappa(4) = 1 \end{cases} \iff \begin{cases} t_{\{1\}} = \kappa(1)^{-1/2}\,\kappa(2)^{1/2} \\ t_{\{3\}} = \kappa(3)^{-1/2} \\ t_{\{1,2\}} = \kappa(2)^{-1/2} \end{cases}.$$

The transition matrix parameterized by $s(e)$, $e \in \mathscr{E}$ and t_B, $S \in \mathscr{S}$ is

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array}\begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} \star & s(1,2)\, t_{\{1\}}^{-1} & 0 & s(1,4)\, t_{\{1\}}^{-1}\, t_{\{1,2\}}^{-1} \\ s(1,2)\, t_{\{1\}} & \star & s(2,3)\, t_{\{1,2\}}^{-1}\, t_{\{3\}} & s(2,4)\, t_{\{1,2\}}^{-1} \\ 0 & s(2,3)\, t_{\{3\}}^{-1}\, t_{\{1,2\}} & \star & s(3,4)\, t_{\{3\}}^{-1} \\ s(1,4)\, t_{\{1\}}\, t_{\{1,2\}} & s(2,4)\, t_{\{1,2\}} & s(3,4)\, t_{\{3\}} & \star \end{bmatrix} \end{array},$$

where the diagonal terms are uniquely defined if the following set of inequalities is true

$$\begin{cases} 1 \geq s(1,2)t_{\{1,2\}} + s(1,4) \\ 1 \geq s(1,2)t_{\{1\}}t_{\{1,2\}} + s(2,3)t_{\{3\}} + s(2,4) \\ 1 \geq s(2,3)t_{\{1,2\}} + s(3,4) \\ 1 \geq s(1,4)t_{\{1\}}t_{\{1,2\}} + s(2,4)t_{\{1,2\}} + s(3,4)t_{\{3\}} \end{cases}.$$

4 Discussion

The algebraic analysis of statistical models of the type $p \propto t^A$ where A is an integer matrix has been first introduced in Geiger et al. (2006), where an implicit binomial form of the model is derived from the monomial form.

Our analysis differs from that in two respects. First, we move backwards from the binomial form represented by the Kolmogorov's condition to the monomial form. Second, we parameterize the transition probabilities, so that the normalization requires more than one constant.

Our parameterization of a reversible Markov matrix is based on a generic weight function $s(e)$ on the edges $e \in \mathscr{E}$ of the structure graph $\mathscr{G}$ and a monomial form of the invariant probability. When the basis of the cocycle space is given by the verteces of the graph, the parameterization is identical to that the classical form $P_{v\to w} = \pi(v)^{-1/2}\pi(w)^{1/2}s(v,w)$. The monomial form of the unnormalized invariant probability $\kappa(v) = \prod_{B\in\mathscr{B}:\, v\in B} t_B^{-2}$ suggests the use of a family of sets $\mathscr{B}$ smaller than a cocycle basis $\mathscr{S}$ in order to get a parsimonious statistical model. For example, if the graph $\mathscr{G}$ is a square $N \times N$ grid, a coarse-grained model could use $n \times n$ sub-grids, $1 < n < N$.

One distinct advantage of the implicit binomial form is its ability to fully describe the closure of its strictly positive part, i.e. the extended exponential family. The computation of a Hilbert basis of the non-negative integer kernel of the cocycle matrix U leads to a parameterization of the extended exponential family as in Malagò and Pistone (2010), see also Rauh et al. (2011). However, in this case the border of the model appears to consist simply on the deletion of edges in the support graph.

It follows from general properties of toric ideals that a Graver basis is a universal Gröbner basis and that a universal Gröbner basis is a Markov basis, Sturmfels (1996). The Markov basis property is related with the connectedness of random walks on the fibers of A, see Diaconis and Sturmfels (1998) and subsequent literature on MCMC simulation. In this case it would be a simulation of a random reversible Markov matrix.

Finally, the knowledge of a Graver basis for the K-ideal provides efficient algorithms for discrete optimization, see De Loera et al. (2008), Onn (2011).

Acknowledgments During this research we have discussed various items with S. Sullivan, B. Sturmfels, G. Casnati and we want to thank them for providing us with supporting advice. Preliminary versions of this paper have been presented at the workshop WOGAS2, Warwick University 2010 and the conference CREST-SBM 2, Osaka 2010. We wish to thank G. Letac for bringing reference Suomela (1979) to our attention and F. Rigat for the reference Peskun (1973). Also they thank S. Onn for pointing out the relevance of his own work on Graver bases. After the publication of the first draft on the ArXiv, Y. A. Mitrophanov provided us with the reference to his own work on the subject, we where not aware of and that with gratefully acknowledge.

References

Bigatti, A., Robbiano, L. (2001). Toric ideals. *Matemática Contemporânea, 21*, pp. 1–25. 16th School of Algebra, Part II (Portuguese) (Brasília, 2000).

Bollobás, B. (1998). Modern graph theory. In: *Graduate Texts in Mathematics*, vol. 184. New York: Springer.

CoCoATeam (online). CoCoA: a system for doing Computations in Commutative Algebra. Available at http://cocoa.dima.unige.it.

Cox, D., Little, J., O'Shea, D. (1997). Ideals, varieties, and algorithms: An introduction to computational algebraic geometry and commutative algebra, 2nd edn. In: *Undergraduate Texts in Mathematics*. New York: Springer.

De Loera, J. A., Hemmecke, R., Onn, S., Weismantel, R. (2008). n-fold integer programming. *Discrete Optimization. The Journal of Combinatorial Operations Research, 5*, 231–241.

Diaconis, P., Rolles, S. W. W. (2006). Bayesian analysis for reversible Markov chains. *The Annals of Statistics, 34*, 1270–1292.

Diaconis, P., Sturmfels, B. (1998). Algebraic algorithms for sampling from conditional distributions. *The Annals of Statistics, 26*, 363–397.

Dobrushin, R. L., Sukhov, Y. M., Fritts, Ĭ. (1988). A. N. Kolmogorov—founder of the theory of reversible Markov processes. *Akademiya Nauk SSSR i Moskovskoe Matematicheskoe Obshchestvo. Uspekhi Matematicheskikh Nauk, 43*, 167–188.

Drton, M., Sturmfels, B., Sullivant, S. (2009). Lectures on algebraic statistics. In: *Oberwolfach Seminars*, vol. 39. Basel: Birkhäuser Verlag.

Geiger, D., Meek, C., Sturmfels, B. (2006). On the toric algebra of graphical models. *The Annals of Statistics, 34*, 1463–1492.

Gibilisco, P., Riccomagno, E., Rogantin, M. P., Wynn, H. P. (eds) (2010). *Algebraic and geometric methods in statistics*. Cambridge: Cambridge University Press.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika, 57*, 97–109.

Kelly, F. P. (1979). *Reversibility and stochastic networks*. Wiley Series in Probability and Mathematical Statistics. Chichester: Wiley.

Kreuzer, M., Robbiano, L. (2000). *Computational commutative algebra. 1*. Berlin: Springer.

Liu, J. S. (2008). *Monte Carlo strategies in scientific computing*. Springer Series in Statistics. New York: Springer.

Malagò, L., Pistone, G. (2010). A note on the border of an exponential family. ArXiv:1012.0637v1.

Mitrophanov, A. Y. (2004). Reversible Markov chains and spanning trees. *The Mathematical Scientist, 29*, 107–114.

Onn, S. (2011). Theory and Applications of n-fold Integer Programming. In: *Mixed integer non-linear programming*, no. 154 in Frontier Science, (pp. 559–593). IMA.

Peskun, P. H. (1973). Optimum Monte-Carlo sampling using Markov chains. *Biometrika, 60*, 607–612.

Pistone, G., Riccomagno, E., Wynn, H. P. (2001). Algebraic statistics: Computational commutative algebra in statistics. In: *Monographs on Statistics and Applied Probability*, vol. 89. Boca Raton, FL: Chapman & Hall/CRC.

Rauh, J., Kahle, T., Ay, N. (2011). Support sets in exponential families and oriented matroid theory. *International Journal of Approximate Reasonoing, 52*, 613–626. Special Issue Workshop on Uncertainty Processing WUPES09.

Strook, D. W. (2005). An Introduction to Markov Processes. In: *Graduate Texts in Mathematics*, No. 230. Berlin: Springer.

Sturmfels, B. (1996). *Gröbner bases and convex polytopes*. Providence, RI: American Mathematical Society.

Suomela, P. (1979). Invariant measures of time-reversible Markov chains. *Journal of Applied Probability, 16*, 226–229.