Absorption time and absorption probabilities for a family of multidimensional gambler models

Paweł Lorek and Piotr Markowski

Mathematical Institute, University of Wrocław, pl. Grunwaldzki 2/4, 50-384 Wrocław, Poland.
E-mail address: {Pawel.Lorek,Piotr.Markowski}@math.uni.wroc.pl

Abstract. For a family of multidimensional gambler models we provide formulas for the winning probabilities in terms of parameters of the system and for the distribution of a game duration in terms of eigenvalues of underlying one-dimensional games. These formulas were known for the one-dimensional case – initially proofs were purely analytical, recently probabilistic constructions have been given. Concerning the game duration, in many cases our approach yields sample-path constructions. We heavily exploit intertwining between (not necessarily) stochastic matrices (for game duration results), a notion of Siegmund duality (for winning/ruin probabilities), and a notion of Kronecker products.

1. Introduction

In the one-dimensional gambler's ruin problem two players start a game with the total amount of, say, N dollars and with initial values k and $N - k$. At each step they flip the coin (not necessarily unbiased) to decide who wins a dollar. The game is over when one of them goes bankrupt. There are some fundamental questions related to this process.

Q1 Starting with i dollars, what is the probability of winning?

Q2 Starting with i dollars, what is the distribution (or the structure) of the game duration (*i.e.*, the absorption time)? Or, what is the distribution (or the structure) of the game duration conditioned on winning/losing?

In this paper we will answer above questions for a wide class of multidimensional generalizations of gambler's ruin problem. The proofs will be probabilistic in most cases, utilizing either Siegmund duality or intertwining between chains.

Generalized multidimensional gambler models. In Lorek (2017) we considered the following generalization. There is one player (referred to "we") playing with $d \geq 1$ other players. Our initial assets are $(i_1, \ldots, i_d)$ and assets of consecutive players are $(N_1 - i_1, \ldots, N_d - i_d)$ ($N_j \geq 1$ is the total amount of assets with player j). Then, with probability $p_j(i_j)$ we win one dollar with player j and

with probability $q_j(i_j)$ we lose it. With the remaining probability $1 - \sum_{k=1}^{d}(p_k(i_j) + q_j(i_k))$ we do nothing (*i.e.*, ties are also possible). Once we win completely with player j (*i.e.*, $i_j = N_j$) we do not play with him/her anymore. We lose the whole game if we lose with at least one player, *i.e.*, when $i_j = 0$ for some $j = 1, \ldots, d$. The game can be described more formally as a Markov chain Z with two absorbing states. The state space is $\mathbb{E} = \{(i_1, \ldots, i_d) : 1 \leq i_j \leq N_j, 1 \leq j \leq d\} \cup \{-\infty\}$ (where $-\infty$ means we *lose*). For a convenience denote $p_j(N_j) = q_j(N_j) = 0, j = 1, \ldots, d$. Assume that for all $i_j \in \{1, \ldots, N_j\}, j \in \{1, \ldots, d\}$ we have $p_j(i_j) > 0, q_j(i_j) > 0$ and $\sum_{k=1}^{d}(p_k(i_k) + q_k(i_k)) \leq 1$. With some abuse of notation, we will sometimes write $(i'_1, \ldots, i'_d) = -\infty$. The transitions of the described chain are the following:

$$\mathbf{P}_Z((i_1, \ldots, i_d), (i'_1, \ldots, i'_d)) =$$

$$
\begin{cases}
p_j(i_j) & \text{if} \quad i'_j = i_j + 1, i'_k = i_k, k \neq j, \\
q_j(i_j) & \text{if} \quad i'_j = i_j - 1, i'_k = i_k, k \neq j, \\
\sum_{j:i_j=1} q_j(1) & \text{if} \quad (i'_1, \ldots, i'_d) = -\infty, \\
1 - \sum_{k=1}^{d}(p_k(i_k) + q_k(i_k)) & \text{if} \quad i'_j = i_j, 1 \leq j \leq d, \\
1 & \text{if} \quad (i_1, \ldots, i_d) = (i'_1, \ldots, i'_d) = -\infty.
\end{cases}
\tag{1.1}
$$

The chain has two absorbing states: $(N_1, \ldots, N_d)$ (we *win*) and $-\infty$ (we *lose*). Let

$$\rho((i_1, \ldots, i_d)) = P(\tau_{(N_1, \ldots, N_d)} < \tau_{-\infty} | Z_0 = (i_1, \ldots, i_d)), \tag{1.2}$$

where $\tau_{(i'_1, \ldots, i'_d)} := \inf\{n \geq 0 : Z_n = (i'_1, \ldots, i'_d)\}$. Roughly speaking, $\rho((i_1, \ldots, i_d))$ is the probability of winning starting at $(i_1, \ldots, i_d)$. In Lorek (2017) we derived the formula for this probability, namely

$$\rho((i_1, \ldots, i_d)) = \frac{\prod_{j=1}^{d}\left(\sum_{n_j=1}^{i_j}\prod_{r=1}^{n_j-1}\left(\frac{q_j(r)}{p_j(r)}\right)\right)}{\prod_{j=1}^{d}\left(\sum_{n_j=1}^{N_j}\prod_{r=1}^{n_j-1}\left(\frac{q_j(r)}{p_j(r)}\right)\right)}. \tag{1.3}$$

In this paper we consider a much wider class of d-dimensional games - the chain given in (1.1) is just a special case. For example, within the class we can win/lose in one step with many players. The multidimensional chain is constructed from a variety of one-dimensional chains using Kronecker products. For this class:

- We give expressions for the winning probabilities and prove that it is a product of the winning probabilities corresponding to one-dimensional games. In particular, for a subclass of multidimensional chains, constructed from one-dimensional birth and death chains, the winning probabilities are given in (1.3). The main tool for showing winning probabilities is the Siegmund duality defined for partially ordered state spaces, exploiting the results from Lorek (2018).

- We give formulas for the distributions of the absorption time. In some cases a probability generating function is given, in other cases we show that the absorption time is equal, in distribution, to the absorption time of another chain, which is, in a sense, a multidimensional pure-birth chain. In many cases, the probabilistic proof is given. To show the absorption distribution, we exploit the spectral polynomials given in Fill (2009), and their variations considered in Gong et al. (2012), Mao and Zhang (2017).

To have a feeling on what kind of results related to absorption time we obtain, let us have a look at Figure 1.1 (note that the caption can be fully understood once further sections are read).
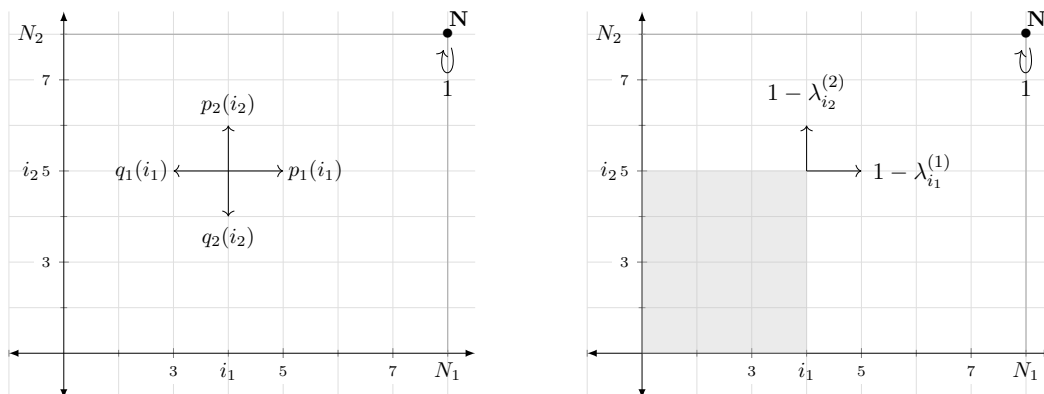
FIGURE 1.1. Sample transitions for the example from Section 6.3 with $d = 2$ and $r = 1$: X^* (left) and $\hat{X}$ (right). State $\mathbf{N} = (N_1, N_2)$ is the only absorbing one in both chains. Probabilities of staying are not depicted. If X^* starts at $(1,1)$, so does the $\hat{X}$ and $T^*_{(1,1),\mathbf{N}} = \hat{T}_{(1,1),\mathbf{N}}$ provided $q_j(1) = 0, j = 1, \ldots, d$. If, say, $\nu^*((i_1, i_2)) = 1$ then the pgf of $T^*_{(i_1,i_2),\mathbf{N}}$ is a mixture of pgfs of $\hat{T}_{(j_1,j_2),\mathbf{N}}$ for $j_1 \leq i_1, j_2 \leq i_2$ (shaded area).

On the left hand side of Fig. 1.1 a chain X^* constructed from two one-dimensional birth and death chains is presented (i-th chain has only one absorbing state $N_i, i = 1, 2$). The chain is constructed in a specific way which results in the bivariate chain with independent moves (either *up, down, left* or *right*). Its transitions are consistent with transitions of a chain given in (1.1) – except there is just one absorbing state $\mathbf{N} = (N_1, N_2)$ (i.e., there is no $-\infty$ state). We will show that the time to absorption of the chain X^* started at (i_1, i_2) is a mixture of times to absorption of a pure-birth chain $\hat{X}$ starting at states (i'_1, i'_2), where $i'_1 \leq i_1, i'_2 \leq i_2$ (shaded area on the right hand side of Fig. 1.1). In particular, if X^* starts in $(1,1)$, so does the chain $\hat{X}$. The chain $\hat{X}$ has also the only absorbing state $\mathbf{N}$, it is pure-birth in the sense, that only *up* and *right* transitions are allowed. The probabilities of its transitions are related to the eigenvalues of one-dimensional birth and death chains from which X^* was constructed.

Remark 1.1. In Lorek (2017) we considered the chain – given in (1.1) – which is constructed from d one-dimensional birth and death chains in a very specific way. The method from this article is much more general, we can construct a variety of multidimensional chains from given d one-dimensional birth and death chains. It is worth mentioning, that even for the case (1.1), the proof is quite different (from the one in Lorek (2017)).

Several variations (including multidimensional ones) of gambler's ruin problem have been considered. Researchers usually study absorption probabilities, absorption time, or both. In Kmet and Petkovšek (2002) authors consider a two-dimensional model (they consider two currencies) and study the expected game duration. In Ross (2009) some multidimensional game is considered: at each step two players are randomly chosen, these players play a regular game, all till one of the players have all the coins. Author derives the probability that a specific player wins, the expected number of turns in total and between two given players. In Rocha and Stern (2004) the following multidimensional game is considered: there are n players, at each step there is one winner who collects $n - 1$ coins from other players, whereas all others lose 1 coin. An asymptotic probability for an individual ruin and dependence of ruin time are studied. In Tzioufas (2019) the multidimensional case is considered, in which with equal probability a unit displacement in any direction is possible. Moments of leaving some ball are considered. In Champagnat et al. (2018) authors present a new probabilistic analysis of distributed algorithm re-considering a variation of a banker

algorithm. Mathematically, it is random walk on a rectangle with specified absorbing states. The results are generalized to the case with many players and resources.

The absorption probability of a given chain may be related to the stationary distribution of some ergodic chain. This relation is given using the *Siegmund duality*, the notion introduced in Siegmund (1976). This is also the tool we use for showing winning probabilities. Already in Lindley (1952) similar duality between some random walks on integers was shown. It was also studied in financial mathematics, where the probability that a dual risk process starting at some level is ruined, is equal to the probability that the stationary queue length exceeds this level (see Asmussen and Albrecher (2010), Asmussen and Sigman (1996)). In all these cases the Siegmund duality was defined for the linear ordering of the state space. The existence of a Siegmund dual for a linearly ordered state space requires stochastic monotonicity of the chain. Recently, in Lorek (2018) we provided *if and only if* conditions for the existence of the Siegmund dual for partially ordered state spaces (roughly speaking, the Möbius monotonicity is required). In this paper, we exploit this duality defined for a coordinate-wise partial ordering.

It is worth mentioning that for one-dimensional gambler models there are several approaches to (each having its advantages and disadvantages) study the winning probability and/or game duration, including conditioning, difference equations (the most common approach to provide the formula for the winning probability in the classical – *i.e.*, the one with constant birth and death rates – gambler's ruin problem), generating functions and martingale-based methods (*e.g.*, Lengyel (2011)), path counting (*e.g.*, Lengyel (2009)).

Absorption time. Consider a one-dimensional game corresponding to the gambler's ruin problem. Let N be the total amount of money. Being at a state $i \in \{2, \ldots, N-1\}$ we can either win one dollar with probability $p(i) > 0$ or lose it with probability $q(i) > 0$, with the remaining probability nothing happens. Assuming $p(1) > 0$ and $p(N) = q(N) = p(0) = q(0) = 0$ the transitions are following:

$$\mathbf{P}_Y(i, i') = \begin{cases} p(i) & \text{if } i' = i+1, \\ q(i) & \text{if } i' = i-1, \\ 1 - (p(i) + q(i)) & \text{if } i' = i. \end{cases} \tag{1.4}$$

States 0 and N are absorbing. Consider two cases:

Case: $q(1) = 0$. Roughly speaking, if started at $i \geq 1$ the chain never reaches 0 and this is actually a birth and death chain on $\{1, \ldots, N\}$ with N being the only absorbing state. Define $T_{a,b} = \inf\{n \geq 0 : Y_n = b \,|\, Y_0 = a\}$. A well known theorem attributed to Keilson (1979) states that the probability generating function pgf of $T_{1,N}$ is the following:

$$\mathsf{pgf}_{T_{1,N}}(u) := \mathbf{E}u^{T_{1,N}} = \prod_{k=1}^{N-1} \left[\frac{(1-\lambda_k)u}{1-\lambda_k u} \right], \tag{1.5}$$

where $-1 \leq \lambda_k < 1, k = 1, \ldots, N-1$ are $N-1$ non-unit eigenvalues of $\mathbf{P}_Y$. The proof was purely analytical. Note that (1.5) corresponds to the sum of N geometric random variables, provided that all eigenvalues are positive (which, in this case, is equivalent to the stochastic monotonicity of the chain). For this case, Fill (2009) gave a probabilistic proof of (1.5) using strong stationary duality and intertwinings between chains. Note that in this case (1.5) can be rephrased as:

Theorem 1.2. *Let X^* be an absorbing chain on $\mathbb{E} = \{1, \ldots, N\}$ starting at 1 with the transition matrix $\mathbf{P}_{X^*}$ given in (1.4) having positive eigenvalues $\lambda_k > 0, k = 1, \ldots, N$. Then $T^*_{1,N}$ has the same*

distribution as $\hat{T}_{1,N}$, *the absorption time of* $\hat{X}$ *on* $\hat{\mathbb{E}} = \mathbb{E}$ *starting at 1 with the transition matrix*

$$\mathbf{P}_{\hat{X}}(i,i') = \begin{cases} 1 - \lambda_i & \text{if } i' = i+1, \\ \lambda_i & \text{if } i' = i, \\ 0 & \text{otherwise.} \end{cases}$$

The chain Y on $\{1, \ldots, N\}$ is called pure-birth if $\mathbf{P}_Y(i,j) = 0$ for $j < i$. Similarly, a multidimensional chain Y on $\mathbb{E} = \{(i_1, \ldots, i_d) : 1 \leq i_j \leq N_j, 1 \leq j \leq d\}$ is said to be **pure-birth** if the probability of decreasing any set of coordinates at one step is 0.

Simply noting that for any $1 < s < N$ we have $T_{1,N} = T_{1,s} + T_{s,N}$ and that $T_{1,s}$ and $T_{s,N}$ are independent (see Cor. 2.1 Gong et al. (2012) for a continuous time version) we have

$$\mathsf{pgf}_{T_{s,N}}(u) := \mathbf{E}u^{T_{s,N}} = \frac{\prod_{k=1}^{N-1}\left[\frac{(1-\lambda_k)u}{1-\lambda_k u}\right]}{\prod_{k=1}^{s-1}\left[\frac{(1-\lambda_k^{\lfloor s \rfloor})u}{1-\lambda_k^{\lfloor s \rfloor}u}\right]}, \tag{1.6}$$

where $\lambda_k^{\lfloor i \rfloor}$ are the eigenvalues of the substochastic $(s-1) \times (s-1)$ matrix

$$\mathbf{P}_Y^{\lfloor s \rfloor}(i,i') = \begin{cases} p(i) & \text{if } i' = i+1, 1 \leq i \leq s-2, \\ q(i) & \text{if } i' = i-1, 2 \leq i \leq s-1, \\ 1 - (p(i) + q(i)) & \text{if } i' = i, 1 \leq i \leq s-1. \end{cases}$$

Case: $q(1) > 0$. In this case, authors in Gong et al. (2012) (different proof is given in Mao and Zhang (2017)) derived formulas for pgf of $T_{s,N}$ and $T_{s,0}$ (more precisely, they derived formulas for continuous time versions), which, in discrete case, are given by

$$\mathsf{pgf}_{T_{s,N}}(u) = \mathbf{E}u^{T_{s,N}} = \rho(s)\frac{\prod_{k=1}^{N-1}\left[\frac{(1-\lambda_k)u}{1-\lambda_k u}\right]}{\prod_{k=1}^{s-1}\left[\frac{(1-\lambda_k^{\lfloor s \rfloor})u}{1-\lambda_k^{\lfloor s \rfloor}u}\right]}, \tag{1.7}$$

$$\mathsf{pgf}_{T_{s,0}}(u) = \mathbf{E}u^{T_{s,0}} = (1 - \rho(s))\frac{\prod_{k=1}^{N-1}\left[\frac{(1-\lambda_k)u}{1-\lambda_k u}\right]}{\prod_{k=1}^{N-s-1}\left[\frac{(1-\lambda_k^{\lceil s \rceil})u}{1-\lambda_k^{\lceil s \rceil}u}\right]},$$

where $\rho(s)$ is the probability of winning (*i.e.*, (1.2) with $d = 1, i_1 = s$) and $\lambda_k^{\lceil s \rceil}$ are the eigenvalues of the substochastic matrix (of the size $N - s - 1$)

$$\mathbf{P}_Y^{\lceil s \rceil}(i,i') = \begin{cases} p(i) & \text{if } i' = i+1, s+1 \leq i \leq N-2, \\ q(i) & \text{if } i' = i-1, s+2 \leq i \leq N-1 \\ 1 - (p(i) + q(i)) & \text{if } i' = i, s+1 \leq i \leq N-1. \end{cases}$$

In this paper we aim at presenting results similar to Theorem 1.2 and to (1.7) for a wide class of multidimensional extensions of gambler's ruin problem.

2. Kronecker product and main results

To state our main results we need to recall a notion of the Kronecker product. Let $\mathbf{A}$ be a matrix of size $n \times m$. Then, for any matrix $\mathbf{B}$ the Kronecker product of the matrices is defined as follows:

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{1m}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots & a_{2m}\mathbf{B} \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}\mathbf{B} & a_{n2}\mathbf{B} & \dots & a_{nm}\mathbf{B} \end{bmatrix}.$$

For square matrices $\mathbf{A}$ and $\mathbf{B}$ it is also convenient to define the Kronecker sum as:

$$\mathbf{A} \oplus \mathbf{B} = \mathbf{A} \otimes \mathbf{I_B} + \mathbf{I_A} \otimes \mathbf{B},$$

where $\mathbf{I_A}$ ($\mathbf{I_B}$) is the identity matrix of the same size as $\mathbf{A}$ ($\mathbf{B}$).

Both, product and sum, are extended as:

$$\bigotimes_{i=1}^{n} \mathbf{A}_i = (\dots((\mathbf{A}_1 \otimes \mathbf{A_2}) \otimes \mathbf{A}_3)\dots) \otimes \mathbf{A}_n = \mathbf{A}_1 \otimes \mathbf{A}_2 \otimes \dots \otimes \mathbf{A}_n$$

and

$$\bigoplus_{i=1}^{n} \mathbf{A}_i = (\dots((\mathbf{A}_1 \oplus \mathbf{A_2}) \oplus \mathbf{A}_3)\dots) \oplus \mathbf{A}_n = \mathbf{A}_1 \oplus \mathbf{A}_2 \oplus \dots \oplus \mathbf{A}_n.$$

Notation. For a convenience, for the given substochastic matrix $\mathbf{P}'_Y$ on $\mathbb{E}' = \{\mathbf{e}_1, \dots, \mathbf{e}_M\}$ by $\mathbf{P}_Y = \mathcal{F}_{\mathbf{e}_0}(\mathbf{P}'_Y)$ we denote a stochastic matrix on $\mathbb{E} = \{\mathbf{e}_0\} \cup \mathbb{E}'$ constructed from $\mathbf{P}'_Y$ in the following way:

$$\mathbf{P}_Y(\mathbf{e}_i, \mathbf{e}_j) = \begin{cases} \mathbf{P}'_Y(\mathbf{e}_i, \mathbf{e}_j) & \text{if} \quad \mathbf{e}_i, \mathbf{e}_j \in \mathbb{E}, \\ 1 - \sum_{\mathbf{e}_k \in \mathbb{E}'} \mathbf{P}'_Y(\mathbf{e}_i, \mathbf{e}_k) & \text{if} \quad \mathbf{e}_i \in \mathbb{E}', \mathbf{e}_j = \mathbf{e}_0, \\ 1 & \text{if} \quad \mathbf{e}_i = \mathbf{e}_j = \mathbf{e}_0. \\ 0 & \text{if} \quad \mathbf{e}_i = \mathbf{e}_0, \mathbf{e}_j \in \mathbb{E}. \end{cases}$$

Similarly, for a stochastic matrix $\mathbf{P}_Y$ on $\mathbb{E} = \{\mathbf{e}_0\} \cup \mathbb{E}'$ let $\mathbf{P}'_Y = \mathcal{F}_{\mathbf{e}_0}^{-1}(\mathbf{P}_Y)$ be a substochastic matrix on $\mathbb{E}'$ resulting from $\mathbf{P}_Y$ by removing the row and the column corresponding to the state $\mathbf{e}_0$.

For a Markov chain Y on $\mathbb{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_M\}$ we say that $A \subseteq \mathbb{E}$ is a communication class if for all $\mathbf{e}, \mathbf{e}' \in A$ we have $\mathbf{P}_Y^n(\mathbf{e}, \mathbf{e}') > 0$ for some $n \geq 0$.

For a given chain Y we define $T_{\nu, \mathbf{e}'} := \inf\{n \geq 0 : Y_n = \mathbf{e}'|Y_0 \sim \nu\}$. Slightly abusing the notation, by $T_{\mathbf{e}, \mathbf{e}'}$ we mean $T_{\nu, \mathbf{e}'}$ with $\nu = \delta_{\mathbf{e}}$. For $\mathbb{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_M\}$ and for $f : \mathbb{E} \to \mathbb{R}$, we define a row vector $\boldsymbol{f} = (f(\mathbf{e}_1), \dots, f(\mathbf{e}_M))$. For $N_j > 0, j = 1, \dots, d$ we define $\mathbf{N} = (N_1, \dots, N_d)$.

2.1. *Absorption probabilities.* Before stating the result, let us provide some intuition behind it. Assume we play with d players (*i.e.*, we have d one-dimensional games), the winning probability playing only with player j is given by $\rho_j(i_j)$, provided we started with i_j dollars. Now, if we play with all the players independently and we define that we win *the whole game* if we win with all the players, lose *the whole game* if we lose with at least one player, then the probability of winning *the whole game* is a product of probabilities *i.e.*, $\prod_{j=1}^{d} \rho_j(i_j)$ (the formula is given in (2.3) below).

However, one can ask the following question: Can we combine the games in *some other* way, so that the resulting winning probability is still of a product form? For example: can we combine $d = 10$ games so that at one step we can play with at most $r = 5$ other players? Can the rules (for combining the games) depend on the current fortune? These type of questions (to which the answers are *yes*) were the motivation for the next theorem, where a wide class of possible combinations is allowed. The examples are provided later in Section 6.

Theorem 2.1. *Fix integers $d \geq 1, m \geq 1$. For $k = 1, \dots, m$ let $\mathbf{A}_k \subseteq \{1, \dots, d\}$. Assume*

- $\forall (1 \leq k \leq m) \ \mathbf{P}_{Z_j^{(k)}} = \mathcal{F}_0(\mathbf{P}'_{Z_j^{(k)}})$ *is a stochastic matrix corresponding to a Markov chain* $Z_j^{(k)}$ *on* $\mathbb{E}_j = \{0, 1, \ldots, N_j\}$ *such that for* $i \in \mathbb{E}_j$ *we have*

$$\rho_j^{(k)}(i) = P(\tau_{N_j} < \tau_0 | Z_j^{(k)}(0) = i) = \rho_j(i). \tag{2.1}$$

  In other words, $Z_j^{(k)}$ *are* m *(* $k = 1, \ldots, m$ *) chains having the same winning probability at every state* i.

- *Let*

$$\mathbb{R}'_{Z_j^{(k)}} = \begin{cases} \mathbf{P}'_{Z_j^{(k)}} & \text{if } j \in \mathbf{A}_k, \\ \mathbf{I}_j & \text{if } j \notin \mathbf{A}_k, \end{cases}$$

  where $\mathbf{I}_j$ *is the identity matrix of size* $N_j \times N_j$.

- *Let* $\mathbf{B}_i, i = 1, \ldots, m$ *be either*
  - *any real numbers (i.e.,* $\mathbf{B}_k \in \mathbb{R}$ *) such that* $\sum_{k=1}^m \mathbf{B}_k = 1$, *or*
  - *square matrices of size* $\prod_{j=1}^d N_j \times \prod_{j=1}^d N_j$ *such that* $\sum_{k=1}^m \mathbf{B}_k = \mathbf{I}$ *(identity matrix of the appropriate size)*

- *The matrix* $\mathbf{P}_Z = \mathcal{F}_{-\infty}(\mathbf{P}_{Z'})$ *with*

$$\mathbf{P}'_Z = \sum_{k=1}^m \mathbf{B}_k \left(\bigotimes_{j \leq d} \mathbb{R}'_{Z_j^{(k)}} \right) \tag{2.2}$$

  is stochastic on $\mathbb{E} = \{-\infty\} \cup \bigotimes_{j \leq d} \mathbb{E}'_j$, *set* $\mathbb{E} \setminus \{\{\mathbf{N}\} \cup \{-\infty\}\}$ *is a communication class.*

Then, the winning probability (i.e., the absorption at $\mathbf{N}$ *) of the Markov chain* Z *on* $\mathbb{E} = \{-\infty\} \cup \{1, \ldots, N_1\} \times \ldots \times \{1, \ldots, N_d\}$ *with the transition matrix* $\mathbf{P}_Z = \mathcal{F}_{-\infty}(\mathbf{P}'_Z)$ *is given by*

$$\rho(i_1, \ldots, i_d) = \prod_{j=1}^d \rho_j(i_j). \tag{2.3}$$

The proof is postponed to Section 4.2.

Note that $\mathbf{P}_{Z_j^{(k)}}$ in Theorem 2.1 are general. If we only know the winning probabilities of $\mathbf{P}_{Z_j^{(k)}}$ (they cannot depend on k), then we know the winning probabilities of Z. Taking $\mathbf{P}_{Z_j^{(k)}}$ corresponding to gambler's ruin game given in (1.4) we have:

Corollary 2.2. *Let* $\mathbf{P}_{Z_j^{(k)}}$ *for* $j = 1, \ldots, d$ *be the birth and death chain given in (1.4). Then, the winning probability of* $\mathbf{P}_Z = \mathcal{F}_{-\infty}(\mathbf{P}'_Z)$ *is given by (1.3).*

Proof: For $\mathbf{P}_{Z_j^{(k)}}$ the winning probability is known (shown in (3.5)), it is

$$\rho_j(i_j) = \frac{\sum\limits_{n_j=1}^{i_j} \prod\limits_{r=1}^{n_j-1} \left(\dfrac{q_j(r)}{p_j(r)} \right)}{\sum\limits_{n_j=1}^{N_j} \prod\limits_{r=1}^{n_j-1} \left(\dfrac{q_j(r)}{p_j(r)} \right)}. \tag{2.4}$$

Assertion of Theorem 2.1 completes the proof. □

The chain Z can be interpreted as a d-dimensional game, with state $(N_1, \ldots, N_d)$ corresponding to winning and state $-\infty$ corresponding to losing.

Remark 2.3. In Lorek (2017, Theorem 2) we showed that the non-negativity of the resulting $\mathbf{P}'_Z$ is not required (for showing a product form formula for the winning probability of the model considered therein) – it is only required that (in our settings) for all $(i'_1, \ldots, i'_d) \in \mathbb{E}$ we have

$$\lim_{n \to \infty} \mathbf{P}'^n_Z((i'_1, \ldots, i'_d), (i_1, \ldots, i_d)) = \pi((i_1, \ldots, i_d)), \qquad \sum_{(i_1, \ldots, i_d) \in \mathbb{E}} \pi((i_1, \ldots, i_d)) = 1.$$

A one-dimensional example was provided in Lorek (2017, Section 4). It is left for a future research to check if the assertion of Theorem 2.1 holds also without the assumption of the non-negativity of $\mathbf{P}'_Z$ (in this paper we focused on stochastic proofs, whenever possible).

2.2. *Absorption time.* Let us start with some motivation. As recalled in the introduction, we have some expressions for the absorption time of a one-dimensional birth and death chain X^* on $1, \ldots, N$ with one absorbing state N. If we start at state 1, this time is expressed in terms of the eigenvalues of the transition matrix (formula (1.5)). Moreover, if the eigenvalues are non-negative (which corresponds to stochastic monotonicity of the chain), we have a stochastic interpretation: its absorption time is equal to the absorption time of a *pure-birth* chain $\hat{X}$, whose transitions involve the aforementioned eigenvalues (formula (1.2)).

In case when we start at $s > 1$, this absorption time of X^* can be expressed in terms of the eigenvalues of the transition matrix and of the truncated substochastic transition matrix (formula (1.6)). Using the duality-based approach given in Fill (2009), it is relatively easy to show that it can be expressed as a mixture of absorption times of a *pure-birth* chain starting at $s' \le s$. To be more precise, the probability generating function of the absorption time of X^* is a mixture of probability generating functions of the absorption time of $\hat{X}$. Moreover, if the mixture coefficients are non-negative, we have stochastic interpretation (a sample-path construction) of this absorption time.

These observations were our motivation for a multidimensional extension. Do similar results hold then? Can we have a similar stochastic interpretation in some cases? How to construct a multidimensional chain out of many one-dimensional birth and death chains, so that the absorption time of the constructed chain can be somehow expressed in terms of pure-birth chains, whose transitions involve eigenvalues of underlying birth and death chains? In the next theorem we provide a wide class of multidimensional chains (ways of constructing such a chain from one-dimensional birth and death chains), for which we are able to express the absorption time in the aforementioned desired way.

Theorem 2.4. *Fix integers $d \ge 1, m \ge 1$. For $k = 1, \ldots, m$ let $A_k \subseteq \{1, \ldots, d\}$. Let $b_i \in \mathbb{R}, i = 1, \ldots, m$ such that $\sum_{k=1}^m b_i = 1$. Let, for $1 \le j \le d$, $\mathbf{P}_{X^*_j}$ be the stochastic matrix corresponding to a birth and death chain X^*_j on $\mathbb{E}_j = \{0, \ldots, N_j\}$ with transitions given in (1.4) with birth rates $p_j(i)$ and death rates $q_j(i)$. Let, for $1 \le j \le d$, $\mathbf{P}'_{X^*_j} = \mathcal{F}_0^{-1}(\mathbf{P}_{X^*_j})$ be the substochastic matrix on $\mathbb{E}'_j = \{1, \ldots, N_j\}$ and*

$$\mathbb{R}'_{X^{*(k)}_j} = \begin{cases} \mathbf{P}'_{X^*_j} & \text{if } j \in A_k, \\[2mm] \mathbf{I}_j & \text{if } j \notin A_k, \end{cases}$$

where $\mathbf{I}_j$ is the identity matrix of size $N_j \times N_j$. I.e., $\mathbb{R}'_{X^{*(k)}_j}$ is either matrix $\mathbf{P}'_{X^*_j}$ or an identity matrix. Let $\lambda_1^{(j)} \le \ldots \le \lambda_{N_j-1}^{(j)} < \lambda_{N_j}^{(j)} = 1$ be the eigenvalues of $\mathbf{P}'_{X^*_j}$.

Assume

A1 *The chains $\mathbf{P}_{X^*_j}, j = 1, \ldots, d$ are stochastically monotone.*

A2 *The matrix* $\mathbf{P}_{X^*} = \mathcal{F}_{-\infty}(\mathbf{P}'_{X^*})$ *with*

$$\mathbf{P}'_{X^*} = \sum_{k=1}^{m} b_k \left(\bigotimes_{j \le d} \mathbb{R}'_{X_j^{*(k)}} \right) \tag{2.5}$$

is a stochastic matrix on $\mathbb{E} = \{-\infty\} \cup \bigotimes_{j \le d} \mathbb{E}'_j$, *set* $\mathbb{E} \setminus \{\{\mathbf{N}\} \cup \{-\infty\}\}$ *is a communication class,* $\mathbf{N} = (N_1, \dots, N_d)$.

A3 *The matrix* $\mathbf{P}_{\hat{X}}$, *given below in* *(2.6)*, *is non-negative.*

Let X^* *be a chain with the above transition matrix* $\mathbf{P}_{X^*}$. *Assume its initial distribution is* ν^*. *The state* $\mathbf{N}$ *is the absorbing state, denote its absorption time by* $T^*_{\nu^*,\mathbf{N}}$.

Then the time to absorption $T^*_{\nu^*,\mathbf{N}}$ *has the following* pgf

$$\mathsf{pgf}_{T^*_{\nu^*,\mathbf{N}}}(s) = \sum_{\hat{\mathbf{e}} \in \mathbb{E}} \hat{\nu}(\hat{\mathbf{e}}) \mathsf{pgf}_{\hat{T}_{\hat{\mathbf{e}},\mathbf{N}}}(s) \left(\prod_{j=1}^{d} \rho_j(1) \right),$$

where $\rho_j(1)$ *is the winning probability of* X_j^* *starting at 1,*

$$\hat{\nu} = \nu^* \bigotimes_{j \le d} \Lambda_j^{-1},$$

Λ_j *are given in* *(3.8)* *calculated for* $\mathbf{P}'_{X_j^*}$ *and* $\hat{T}_{\hat{\mathbf{e}},\mathbf{N}}$ *is the time to absorption for the chain* $\hat{X} \sim (\delta_{\hat{\mathbf{e}}}, \mathbf{P}_{\hat{X}})$ *with:*

$$\mathbf{P}_{\hat{X}}((i_1, \dots, i_d), (i'_1, \dots, i'_d)) =$$

$$\begin{cases}
\displaystyle\prod_{j \in \mathbf{B}} \left(1 - \lambda_{i_j}^{(j)}\right) \sum_{k: \mathbf{B} \subseteq \mathbf{A}_k} \left(b_k \prod_{j \in A_k \setminus B} \lambda_{i_j}^{(j)} \right) & \text{if } i'_j = i_j + 1, \\
& \qquad j \in \mathbf{B} \subseteq \{1, \dots, d\}, \mathbf{B} \ne \emptyset, \\
\displaystyle\sum_{k=1}^{m} b_k \prod_{j \in \mathbf{A}_k} \lambda_{i_j}^{(j)} & \text{if } i'_j = i_j \text{for} j = 1, \dots, d, \\
0 & \text{otherwise.}
\end{cases} \tag{2.6}$$

We also have

$$\forall (\mathbf{e} \in \mathbb{E}) \nu^*(\mathbf{e}) \ne 0 \;\Rightarrow\; \exists (\mathbf{e}' \succeq \mathbf{e}) \hat{\nu}(\mathbf{e}') > 0 \tag{2.7}$$

Moreover, the eigenvalues of $\mathbf{P}_{X^*}$ *and* $\mathbf{P}_{\hat{X}}$ *are the diagonal entries of* $\mathbf{P}_{\hat{X}}$.

Note that $\hat{X}$ is a pure-birth chain. Moreover, at one step it can increase values of coordinates by $+1$ on a set $\mathbf{B}$ such that $\mathbf{B} \subseteq \mathbf{A}_k$, for $k = 1, \dots, m$.

Remark 2.5. In case $b_i \ge 0, i = 1, \dots, m$ (i.e., $(b_1, \dots, b_m)$ is a distribution on $\{1, \dots, m\}$), the matrix $\mathbf{P}_{X^*}$ in assumption **A2** is stochastic (thus **A2** is only about $\mathbb{E} \setminus \{\{\mathbf{N}\} \cup \{-\infty\}\}$ being a communication class) and so is the matrix $\mathbf{P}_{\hat{X}}$ given in (2.6) (i.e., **A3** is fulfilled).

Remark 2.6. Note that the formula for the transitions of the resulting multidimensional chain given in (2.5) is very similar to the formula (2.2), i.e., the one used in Theorem 2.1 (providing results for the winning probabilities). The main difference is that in (2.2) we have some $\mathbf{B}_k$'s which can be either numbers summing up to 1 or some matrices summing up to the identity matrix, whereas in (2.5) they must be numbers summing up to 1. Consequently, the class of the resulting multidimensional chains constructed in Theorem 2.1 is larger than the class of the chains constructed in Theorem 2.4.

Considering initial distribution having whole mass at $(1, \ldots, 1)$ and/or all $q_j(1) = 0, j = 1, \ldots, d$ we have special cases, which we will formulate as a corollary.

Corollary 2.7. *Consider the setup from Theorem 2.4.*

a) *Moreover, assume that $q_j(1) = 0$ for all $j = 1, \ldots, d$. I.e., each X_j^* has actually only one absorbing state (state 0 is not accessible). Then, $\mathbf{N}$ is the only absorbing state of X^*, $\sum_{\mathbf{e} \in \mathbb{E}} \hat{\nu}(\mathbf{e}) = 1$, $\rho_j(1) = 1, j = 1, \ldots, d$ and we have*

$$\mathsf{pgf}_{T_{\nu^*,\mathbf{N}}^*}(s) = \sum_{\hat{\mathbf{e}} \in \mathbb{E}} \hat{\nu}(\hat{\mathbf{e}}) \mathsf{pgf}_{\hat{T}_{\hat{\mathbf{e}},\mathbf{N}}}(s).$$

b) *Moreover, assume that both $q_j(1) = 0$ for all $j = 1, \ldots, d$ and $\nu^*((1, \ldots, 1)) = 1$. Then $T_{(1,\ldots,1),\mathbf{N}}^* \stackrel{d}{=} \hat{T}_{(1,\ldots,1),\mathbf{N}}$, where $\stackrel{d}{=}$ denotes the equality in the distribution.*

c) *Moreover, assume that $\nu^*((1, \ldots, 1)) = 1$. Then assertions of Theorem 2.4 hold with $\hat{\nu}((1, \ldots, 1)) = 1$ and we have*

$$T_{(1,\ldots,1),\mathbf{N}}^* = \begin{cases} \hat{T}_{(1,\ldots,1),\mathbf{N}} & \text{with probability } \prod_{j=1}^d \rho_j(1), \\ -\infty & \text{with probability } 1 - \prod_{j=1}^d \rho_j(1). \end{cases}$$

Sample-path construction. It turns out that when $\hat{\nu}$ resulting from $\hat{\nu} = \nu^* \Lambda^{-1}$ is a distribution (which is always the case in, *e.g.*, Corollary 2.7 *b)* and *c)*), we can have a sample-path construction. I.e., for X^* we can construct, sample path by sample path, a chain $\hat{X}$, so that $T_{\nu^*,\mathbf{N}}^*$ has the distribution expressed in terms of $\hat{T}_{\hat{\nu},\mathbf{N}}$ as stated in Theorem 2.4. The construction is analogous to the construction given in Diaconis and Fill (1990) (paragraph 2.4) - note however that the construction therein was between ergodic chain and its strong stationary dual chain (*i.e.*, the chain with one absorbing state) and the link Λ was a stochastic matrix (it can be substochastic in our case). Having observed $X_0^* = \mathbf{e}_0^*$ (chosen from the distribution ν^*) we set

$$\hat{X}_0 = \hat{\mathbf{e}}_0 \text{ with probability } \frac{\hat{\nu}(\hat{\mathbf{e}}_0)\Lambda(\hat{\mathbf{e}}_0, \mathbf{e}_0^*)}{\nu^*(\mathbf{e}_0^*)}.$$

Then, after choosing $X_1^* = \mathbf{e}_1^*, \ldots, X_{n-1}^* = \mathbf{e}_{n-1}^*$ and $\hat{X}_1 = \hat{\mathbf{e}}_1, \ldots, \hat{X}_n = \hat{\mathbf{e}}_n$ we set

$$\hat{X}_n = \hat{\mathbf{e}}_n \text{ with probability } \frac{\mathbf{P}_{\hat{X}}(\hat{\mathbf{e}}_{n-1}, \hat{\mathbf{e}}_n)\Lambda(\hat{\mathbf{e}}_n, \mathbf{e}_n^*)}{(\mathbf{P}_{X^*}\Lambda)(\hat{\mathbf{e}}_{n-1}, \mathbf{e}_n^*)}.$$

This way we have constructed the chain $\hat{X}$ so that $\Lambda \mathbf{P}_{X^*} = \mathbf{P}_{\hat{X}}\Lambda$ and $\nu^* = \hat{\nu}\Lambda$ with the property that $\hat{X}_n = \hat{\mathbf{e}}_M$ if and only if $X_n^* = \mathbf{e}_M^*$.

Theorem 2.4 is actually neither an extension of (1.6) nor (1.7) to the multidimensional case, since for one-dimensional case the formula for pgf of $T_{s,N}^*$ has a different form, as examples given in Section 6 show.

3. Tools: dualities in Markov chains

Siegmund duality and intertwinings between chains are the key ingredients of our main theorems' proofs.

3.1. *Siegmund duality.* Let X be an ergodic discrete-time Markov chain with the transition matrix $\mathbf{P}_X$ and a finite state space $\mathbb{E} = \{\mathbf{e}_1, \ldots, \mathbf{e}_M\}$ partially ordered by $\preceq$. Denote its stationary distribution by π. We assume that there exists a unique minimal state, say $\mathbf{e}_1$, and a unique maximal state, say $\mathbf{e}_M$. For $A \subseteq \mathbb{E}$, define $\mathbf{P}_X(\mathbf{e}, A) := \sum_{\mathbf{e}' \in A} \mathbf{P}_X(\mathbf{e}, \mathbf{e}')$ and similarly $\pi(A) := \sum_{\mathbf{e} \in A} \pi(\mathbf{e})$.

Define also $\{\mathbf{e}\}^{\uparrow} := \{\mathbf{e}' \in \mathbb{E} : \mathbf{e} \preceq \mathbf{e}'\}$, $\{\mathbf{e}\}^{\downarrow} := \{\mathbf{e}' \in \mathbb{E} : \mathbf{e}' \preceq \mathbf{e}\}$ and $\delta(\mathbf{e}, \mathbf{e}') = \mathbb{1}\{\mathbf{e} = \mathbf{e}'\}$. We say that a Markov chain Z with the transition matrix $\mathbf{P}_Z$ is the **Siegmund dual** of X if

$$\forall(\mathbf{e}_i, \mathbf{e}_j \in \mathbb{E}) \ \forall(n \geq 0) \quad \mathbf{P}_X^n(\mathbf{e}_i, \{\mathbf{e}_j\}^{\downarrow}) = \mathbf{P}_Z^n(\mathbf{e}_j, \{\mathbf{e}_i\}^{\uparrow}). \tag{3.1}$$

In all non-degenerated applications, we can find the *substochastic* matrix $\mathbf{P}'_Z$ fulfilling (3.1). Then we add one extra absorbing state, say $\mathbf{e}_0$, and define $\mathbf{P}_Z = \mathcal{F}_{\mathbf{e}_0}(\mathbf{P}'_Z)$. Note that then $\mathbf{P}_Z$ fulfills (3.1) for all states different from $\mathbb{E}$. This relation also implies that $\mathbf{e}_M$ is an absorbing state in the Siegmund dual, thus Z has two absorbing states. Taking the limits as $n \to \infty$ on both sides of (3.1), we have

$$\pi(\{\mathbf{e}_j\}^{\downarrow}) \quad = \quad \lim_{n \to \infty} \mathbf{P}_Z^n(\mathbf{e}_j, \{\mathbf{e}_i\}^{\uparrow}) \quad = \quad P(\tau_{\mathbf{e}_M} < \tau_{\mathbf{e}_0} | Z_0 = \mathbf{e}_j) = \rho(\mathbf{e}_j). \tag{3.2}$$

The stationary distribution of X is related in this way to the absorption of its Siegmund dual Z.

It is convenient to define the Siegmund duality in a matrix form. Let $\mathbf{C}(\mathbf{e}_i, \mathbf{e}_j) = \mathbb{1}(\mathbf{e}_i \preceq \mathbf{e}_j)$, then the equality (3.1) can be expressed as

$$\mathbf{P}_X^n \mathbf{C} = \mathbf{C}(\mathbf{P}'^{\,n}_Z)^T. \tag{3.3}$$

Relation (3.2) can be rewritten in a matrix form as

$$\boldsymbol{\rho} = \boldsymbol{\pi}\mathbf{C}.$$

The inverse $\mathbf{C}^{-1}$ always exists, usually it is denoted by μ and called the *Möbius function*. To find a Siegmund dual it is enough to find $\mathbf{P}_Z$ fulfilling (3.3) with for $n = 1$.

Let $\preceq := \leq$ be a total ordering on a finite state space $\mathbb{E} = \{1, \dots, M\}$. The chain Y is **stochastically monotone** w.r.t to the total ordering if $\forall i_1 \leq i_2 \ \forall j \ \mathbf{P}_Y(i_2, \{j\}^{\downarrow}) \leq \mathbf{P}_Y(i_1, \{j\}^{\downarrow})$. We have

Lemma 3.1 (Siegmund (1976)). *Let X be an ergodic Markov chain on $\mathbb{E} = \{1, \dots, M\}$ with the transition matrix $\mathbf{P}_X$. The Siegmund dual Z (w.r.t. the total ordering) exists if and only if X is stochastically monotone. In such a case $\mathbf{P}_Z = \mathcal{F}(\mathbf{P}'_Z)$, where*

$$\mathbf{P}'_Z(j, i) = \mathbf{P}_X(i, \{j\}^{\downarrow}) - \mathbf{P}_X(i + 1, \{j\}^{\downarrow})$$

for $i, j \in \mathbb{E}$ (we mean $\mathbf{P}_X(i + 1, \cdot) = 0$).

Since the proof is one line long, we present it.

Proof of Lemma 3.1: The main thing is to show that (3.1) holds for $n = 1$. We have

$$\mathbf{P}'_Z(j, i) \quad = \quad \mathbf{P}'_Z(j, \{i\}^{\uparrow}) - \mathbf{P}'_Z(j, \{i + 1\}^{\uparrow}) = \mathbf{P}_X(i, \{j\}^{\downarrow}) - \mathbf{P}_X(i + 1, \{j\}^{\downarrow}).$$

The latter is non-negative if and only if X is stochastically monotone. $\qquad\qquad\square$

Let X be an ergodic birth and death chain on $\mathbb{E} = \{1, \dots, M\}$ with the transition matrix

$$\mathbf{P}_X(i, i') = \begin{cases} p'(i) & \text{if } i' = i + 1, \\ q'(i) & \text{if } i' = i - 1, \\ 1 - (p'(i) + q'(i)) & \text{if } i' = i, \end{cases} \tag{3.4}$$

where $q'(1) = p'(M) = 0$ and $p'(i) > 0, i = 1, \dots, M - 1, q'(i) > 0, i = 2, \dots, M$. Assume that $p'(i - 1) + q'(i) \leq 1, i = 2, \dots, M$ (what is equivalent to stochastic monotonicity).

It is easily verifiable that when we rename transition probabilities: $p(i) = q'(i), q(i) = p'(i - 1)$, then the transitions $\mathbf{P}_Y$ defined in (1.4) are the transitions of the Siegmund dual of $\mathbf{P}_X$ resulting

from Lemma 3.1. From the known form of the stationary distribution of an ergodic birth and death chain, and from relation (3.2), it follows that for $\mathbf{P}_Y$ given in (1.4) we have

$$\rho(s) = \sum_{k \le s} \pi(s) = \frac{\sum_{n=1}^{s} \prod_{r=1}^{n-1} \left(\frac{q(r)}{p(r)} \right)}{\sum_{n=1}^{M} \prod_{r=1}^{n-1} \left(\frac{q(r)}{p(r)} \right)}. \tag{3.5}$$

3.2. *Intertwinings between absorbing chains.* Let Λ be any nonsingular matrix of size $M \times M$. We say that matrices $\mathbf{P}_{X^*}$ and $\mathbf{P}_{\hat{X}}$ of size $M \times M$ are **intertwined by a link** Λ if

$$\Lambda \mathbf{P}_{X^*} = \mathbf{P}_{\hat{X}} \Lambda.$$

Similarly, we say that vectors $\hat{\nu}$ and ν^* of lengths M are intertwined if

$$\nu^* = \hat{\nu} \Lambda. \tag{3.6}$$

We say that a link Λ is $\mathbf{e}_M^*$-isolated if

$$\Lambda(\hat{\mathbf{e}}, \mathbf{e}_M^*) \quad \begin{cases} \ne 0 & \text{if } \hat{\mathbf{e}} = \hat{\mathbf{e}}_M, \\ = 0 & \text{otherwise .} \end{cases} \tag{3.7}$$

Lemma 3.2. *Let X^* and $\hat{X}$ be Markov chains on $\mathbb{E}^* = \mathbf{e}_0^* \cup \hat{\mathbb{E}}$ and $\hat{\mathbb{E}}$ with transition matrices $\mathbf{P}_{X^*}$ and $\mathbf{P}_{\hat{X}}$ respectively. Moreover, assume X^* has the initial distribution ν^* and two absorbing states: $\mathbf{e}_0^*$ and $\mathbf{e}_M^*$, whereas $\hat{X}$ has one absorbing state $\hat{\mathbf{e}}_M$. Assume that $\mathbf{P}'_{X^*} = \mathcal{F}_{\mathbf{e}_0^*}^{-1}(\mathbf{P}_{X^*})$ and $\mathbf{P}_{\hat{X}}$ are intertwined via an $\mathbf{e}_M^*$-isolated link Λ. Let $\hat{\nu} = \nu^* \Lambda^{-1}$. Then we have*

$$\mathsf{pgf}_{T^*_{\nu^*, \mathbf{e}_M^*}}(s) = \Lambda(\hat{\mathbf{e}}_M, \mathbf{e}_M^*) \sum_{\hat{\mathbf{e}} \in \hat{\mathbb{E}}} \hat{\nu}(\hat{\mathbf{e}}) \mathsf{pgf}_{\hat{T}_{\hat{\mathbf{e}}, \hat{\mathbf{e}}_M}}(s).$$

Proof:

$$\begin{aligned} P(T^*_{\nu^*, \mathbf{e}_M^*} \le t) &= P(X^*(t) = \mathbf{e}_M^*) = \sum_{\mathbf{e}^* \in \mathbb{E}^* \setminus \{\mathbf{e}_0^*\}} \nu^*(\mathbf{e}^*) \mathbf{P}_{X^*}^t(\mathbf{e}^*, \mathbf{e}_M^*) \\ &= \sum_{\hat{\mathbf{e}} \in \mathbb{E}} \sum_{\mathbf{e}^* \in \mathbb{E}^* \setminus \{\mathbf{e}_0^*\}} \hat{\nu}(\hat{\mathbf{e}}) \Lambda(\hat{\mathbf{e}}, \mathbf{e}^*) \mathbf{P}_{X^*}^t(\mathbf{e}^*, \mathbf{e}_M^*) \\ &= \sum_{\hat{\mathbf{e}} \in \mathbb{E}} \sum_{\hat{\mathbf{e}}_2 \in \hat{\mathbb{E}}} \hat{\nu}(\hat{\mathbf{e}}) \mathbf{P}_{\hat{X}}^t(\hat{\mathbf{e}}, \hat{\mathbf{e}}_2) \Lambda(\hat{\mathbf{e}}_2, \mathbf{e}_M^*) \\ &= \Lambda(\hat{\mathbf{e}}_M, \mathbf{e}_M^*) \sum_{\hat{\mathbf{e}} \in \mathbb{E}} \hat{\nu}(\hat{\mathbf{e}}) \mathbf{P}_{\hat{X}}^t(\hat{\mathbf{e}}, \hat{\mathbf{e}}_M). \end{aligned}$$

Now, for pgf we have:

$$\mathsf{pgf}_{T^*_{\nu^*,\mathbf{e}^*_M}}(s) =$$

$$\sum_{k=0}^{\infty} P(T^*_{\nu^*,\mathbf{e}^*_M} = k)s^k = \sum_{k=0}^{\infty}\left(P(T^*_{\nu^*,\mathbf{e}^*_M} \le k) - P(T^*_{\nu^*,\mathbf{e}^*_M} \le k-1)\right)s^k$$

$$= \Lambda(\hat{\mathbf{e}}_M,\mathbf{e}^*_M)\sum_{k=0}^{\infty}\left(\sum_{\hat{\mathbf{e}}\in\hat{\mathbb{E}}}\hat{\nu}(\hat{\mathbf{e}})\mathbf{P}^k_{\hat{X}}(\hat{\mathbf{e}},\hat{\mathbf{e}}_M) - \sum_{\hat{\mathbf{e}}\in\hat{\mathbb{E}}}\hat{\nu}(\hat{\mathbf{e}})\mathbf{P}^{k-1}_{\hat{X}}(\hat{\mathbf{e}},\hat{\mathbf{e}}_M)\right)s^k$$

$$= \Lambda(\hat{\mathbf{e}}_M,\mathbf{e}^*_M)\sum_{\hat{\mathbf{e}}\in\hat{\mathbb{E}}}\hat{\nu}(\hat{\mathbf{e}})\sum_{k=0}^{\infty}\left(\mathbf{P}^k_{\hat{X}}(\hat{\mathbf{e}},\hat{\mathbf{e}}_M) - \mathbf{P}^{k-1}_{\hat{X}}(\hat{\mathbf{e}},\hat{\mathbf{e}}_M)\right)s^k$$

$$= \Lambda(\hat{\mathbf{e}}_M,\mathbf{e}^*_M)\sum_{\hat{\mathbf{e}}\in\hat{\mathbb{E}}}\hat{\nu}(\hat{\mathbf{e}})\sum_{k=0}^{\infty}\left(P(\hat{T}_{\hat{\mathbf{e}},\hat{\mathbf{e}}_M} \le k) - P(\hat{T}_{\hat{\mathbf{e}},\hat{\mathbf{e}}_M} \le k-1)\right)s^k$$

$$= \Lambda(\hat{\mathbf{e}}_M,\mathbf{e}^*_M)\sum_{\hat{\mathbf{e}}\in\hat{\mathbb{E}}}\hat{\nu}(\hat{\mathbf{e}})\sum_{k=0}^{\infty}P(\hat{T}_{\hat{\mathbf{e}},\hat{\mathbf{e}}_M} = k)s^k = \Lambda(\hat{\mathbf{e}}_M,\mathbf{e}^*_M)\sum_{\hat{\mathbf{e}}\in\hat{\mathbb{E}}}\hat{\nu}(\hat{\mathbf{e}})\mathsf{pgf}_{\hat{T}_{\hat{\mathbf{e}},\hat{\mathbf{e}}_M}}(s).$$

□

Corollary 3.3. *Let assumptions of Lemma 3.2 hold and, in addition, let $\hat{\nu}$ be a distribution. Then, we have*

$$T^*_{\nu^*,\mathbf{e}^*_M} = \Lambda(\hat{\mathbf{e}}_M,\mathbf{e}^*_M)\hat{T}_{\hat{\nu},\hat{\mathbf{e}}_M}.$$

From Fill (2009) we can deduce the following lemma.

Lemma 3.4. *Let X^* be a birth and death chain on $\mathbb{E}^* = \{0,\ldots,M\}$ with the transition matrix $\mathbf{P}_{X^*}$ given in (1.4) with two absorbing states: 0 and M. Let $\mathbf{P}'_{X^*} = \mathcal{F}_0^{-1}(\mathbf{P}_{X^*})$. Assume the eigenvalues of $\mathbf{P}'_{X^*}$ are positive, denote them by $0 < \lambda_1 < \ldots < \lambda_M = 1$.*
Define $\mathbf{Q}_1 := \mathbf{I}$ and

$$\mathbf{Q}_k := \frac{(\mathbf{P}'_{X^*} - \lambda_1\mathbf{I})\cdots(\mathbf{P}'_{X^*} - \lambda_{k-1}\mathbf{I})}{(1-\lambda_1)\cdots(1-\lambda_{k-1})}, k = 2,\ldots,M$$

Let Λ be the lower triangular square matrix of size $M \times M$ defined as

$$\Lambda(k,\cdot) = \mathbf{Q}_k(1,\cdot), \quad k = 1,\ldots,M. \tag{3.8}$$

Then, $\mathbf{P}'_{X^*}$ and $\mathbf{P}_{\hat{X}}$ are intertwined by the link Λ, where

$$\mathbf{P}_{\hat{X}}(i,i') = \begin{cases} 1-\lambda_i & \text{if } i' = i+1, \\ \lambda_i & \text{if } i' = i, \\ 0 & \text{otherwise.} \end{cases} \tag{3.9}$$

is a matrix of size $M \times M$.

Note that Lemma 3.4 is similar to Theorem 4.2 in Fill (2009), the difference is that therein Λ is a stochastic matrix, whereas in Lemma 3.4 it can be substochastic (it is strictly substochastic if $q(0) > 0$). An almost identical Λ was considered in Gong et al. (2012), their Proposition 3.3 yields:

Lemma 3.5.

- *The matrices $\mathbf{Q}_k, 1,\ldots,M$ are non-negative and substochastic.*
- *The matrix Λ is non-negative and substochastic, it is lower triangular and*

$$\Lambda(1,1) = 1, \qquad \Lambda(M,M) = \rho(1),$$

  thus Λ is nonsingular.

Remark 3.6. Note that in case X^* has no transition to 0, *i.e.*, $q(1) = 0$, it is actually a chain on $\{1, \ldots, M\}$ and $\mathbf{P}'_{X^*} = \mathcal{F}_0^{-1}(\mathbf{P}_{X^*})$ is a stochastic matrix. Then Λ is a stochastic matrix and $\Lambda(M, M) = 1$.

4. Proofs

4.1. *Properties of the Kronecker product.* In this section we recall some useful properties of the Kronecker product and formulate a lemma relating eigenvectors and eigenvalues of some combination of Kronecker products.

We will exploit the following properties

- bilinearity:
$$\mathbf{A} \otimes (\mathbf{B} + \mathbf{C}) = \mathbf{A} \otimes \mathbf{B} + \mathbf{A} \otimes \mathbf{C}, \tag{P1}$$

- mixed product:
$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD}), \tag{P2}$$

- inverse and transposition:
$$(\mathbf{A} \otimes \mathbf{B})^{-1} = (\mathbf{A})^{-1} \otimes (\mathbf{B})^{-1}, \tag{P3}$$

$$(\mathbf{A} \otimes \mathbf{B})^T = (\mathbf{A})^T \otimes (\mathbf{B})^T. \tag{P4}$$

- eigenvalue and eigenvector:

  having eigenvalues α_j with corresponding left eigenvectors a_j for each matrix $\mathbf{A}_j, j \le n$, we note that $\sum_{j \le n} \alpha_j$ with $\bigotimes_{j \le n} a_j$ and $\prod_{j \le n} \alpha_j$ with $\bigotimes_{j \le n} a_j$ are the eigenvalue and the left eigenvector of $\mathbf{A} = \bigoplus_{j \le n} \mathbf{A}_j$ and $\mathbf{A}' = \bigotimes_{j \le n} \mathbf{A}_j$ respectively. $\tag{P5}$

Last property leads us to the following lemma.

Lemma 4.1. *For all $1 \le j \le n$ and $1 \le i \le m$, let a_j be the left eigenvectors with the corresponding eigenvalues α_j of square matrices $\mathbf{A}_j^{(i)}$ of sizes k_j respectively.*
Let $\mathbf{B}_i, i = 1, \ldots, m$ be square matrices of sizes $\prod_{j=1}^n k_j$ such that $\sum_{i=1}^m \mathbf{B}_i = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix of size $\prod_{j=1}^n k_j$. Then $\prod_{j \le n} \alpha_j$ with $\bigotimes_{j \le n} a_j$ are the eigenvalue and the left eigenvector of $\mathbf{A} = \sum_{i=1}^m (\bigotimes_{j \le n} \mathbf{A}_j^{(i)}) \mathbf{B}_i$.
Similarly, if $b_i, i = 1, \ldots, m$ are real numbers such that $\sum_{i=1}^m b_i$ we have that $\bigotimes_{j \le n} a_j$ is the left eigenvector with the corresponding eigenvalue $\prod_{j \le n} \alpha_j$ of the matrix $\mathbf{A} = \sum_{i=1}^m (\bigotimes_{j \le n} \mathbf{A}_j^{(i)}) b_i$.

Proof: We have

$$
\begin{aligned}
\bigotimes_{j \le n} a_j \sum_{i=1}^m (\bigotimes_{j \le n} \mathbf{A}_j^{(i)}) \mathbf{B}_i &= \sum_{i=1}^m \bigotimes_{j \le n} a_j (\bigotimes_{j \le n} \mathbf{A}_j^{(i)}) \mathbf{B}_i = \sum_{i=1}^m \bigotimes_{j \le n} (a_j \mathbf{A}_j^{(i)}) \mathbf{B}_i \\
&= \sum_{i=1}^m \bigotimes_{j \le n} (a_j \alpha_j) \mathbf{B}_i = \sum_{i=1}^m \prod_{j \le n} \alpha_j \bigotimes_{j \le n} a_j \mathbf{B}_i \\
&= \prod_{j \le n} \alpha_j \bigotimes_{j \le n} a_j \sum_{i=1}^m \mathbf{B}_i = \prod_{j \le n} \alpha_j \bigotimes_{j \le n} a_j.
\end{aligned}
$$

Similarly,

$$\bigotimes_{j \leq n} a_j \sum_{i=1}^{m} (\bigotimes_{j \leq n} \mathbf{A}_j^{(i)}) b_i = \sum_{i=1}^{m} \bigotimes_{j \leq n} a_j (\bigotimes_{j \leq n} \mathbf{A}_j^{(i)}) b_i = \sum_{i=1}^{m} \bigotimes_{j \leq n} (a_j \mathbf{A}_j^{(i)}) b_i$$

$$= \sum_{i=1}^{m} \bigotimes_{j \leq n} (a_j \alpha_j) b_i = \sum_{i=1}^{m} \prod_{j \leq n} \alpha_j \bigotimes_{j \leq n} a_j b_i$$

$$= \prod_{j \leq n} \alpha_j \bigotimes_{j \leq n} a_j \sum_{i=1}^{m} b_i = \prod_{j \leq n} \alpha_j \bigotimes_{j \leq n} a_j.$$

$$\square$$

Substituting stochastic matrices $\mathbf{P}_j^{(i)}$ with stationary distributions π_j (for all $1 \leq j \leq n$, $1 \leq i \leq m$) to matrices $\mathbf{A}_j^{(i)}$ with left eigenvectors α_j (for all $1 \leq j \leq n$, $1 \leq i \leq m$) gives us the following corollary (keeping in mind that 1 is the eigenvalue corresponding to the eigenvector being the stationary distribution):

Corollary 4.2. *Let $\mathbf{P}_j^{(i)}$ be a stochastic matrix of size k_j with π_j being its stationary distribution for all $1 \leq j \leq n, 1 \leq i \leq m$. Let $\mathbf{B}_i$, $1 \leq i \leq m$ be square matrices of sizes $\prod_{j=1}^{n} k_j$ such that $\sum_{i=1}^{m} \mathbf{B}_i = \mathbf{I}$, where $\mathbf{I}$ is the identity of size $\prod_{j=1}^{n} k_j$. Similarly, if b_i, $1 \leq i \leq m$ are real numbers such that $\sum_{i=1}^{m} b_i = 1$, then the stochastic matrices of the form $\sum_{i=1}^{m} (\bigotimes_{j \leq n} \mathbf{P}_j^{(i)}) \mathbf{B}_i$ or $\sum_{i=1}^{m} (\bigotimes_{j \leq n} \mathbf{P}_j^{(i)}) b_i$ have the stationary distribution of the form $\bigotimes_{j \leq n} \pi_j$.*

4.2. *Proof of Theorem 2.1.* We will find an ergodic Markov chain X with the transition matrix $\mathbf{P}_X$ and some partial ordering of the state space (expressed by the ordering matrix $\mathbf{C}$) and show that (3.2) is equivalent to (2.3).

Let $\mathbf{P}_{Z_j}^{(k)}$ (on $\mathbb{E}_j = \{0, \ldots, N_j\}$) be as in the theorem. Let $X_j^{(k)}$ be the ergodic chain on $\mathbb{E}_j' = \{1, \ldots, N_j\}$ with the transition matrix $\mathbf{P}_{X_j}^{(k)}$, such that $Z_j^{(k)}$ is its Siegmund dual w.r.t. the total ordering. *I.e.*, let $\mathbf{C}_j(s, t) = \mathbb{1}(s \leq t)$, and duality means that

$$\mathbf{P}_{X_j}^{(k)} \mathbf{C}_j = \mathbf{C}_j (\mathbf{P}_{Z_j'}^{(k)})^T,$$

where $\mathbf{P}_{Z_j'}^{(k)} = \mathcal{F}_0^{-1}(\mathbf{P}_{Z_j}^{(k)})$. Assumption (2.1) and relation (3.2) imply that for fixed j, the chains $X_j^{(k)}$, $k = 1, \ldots, m$ have the same stationary distribution, denote it by π_j. The relation (3.2) means that $\rho_j = \pi_j \mathbf{C}_j$. On the state space $\mathbb{E} = \bigotimes_{j \leq d} \mathbb{E}_j$ let us introduce the ordering expressed by the matrix $\mathbf{C} = \bigotimes_{j \leq d} \mathbf{C}_j$. From (3.3) we can calculate the matrix $\mathbf{P}_X$:

$$\mathbf{P}_X = \mathbf{C}\mathbf{P}_Z^T \mathbf{C}^{-1} = (\bigotimes_{j \leq d} \mathbf{C}_j) \left(\sum_{k=1}^{m} \mathbf{B}_k (\bigotimes_{j \leq d} \mathbb{R}_{Z_j'}^{(k)}) \right)^T (\bigotimes_{j \leq d} \mathbf{C}_j)^{-1}$$

$$\overset{(P4),(P3)}{=} (\bigotimes_{j \leq d} \mathbf{C}_j) \left(\sum_{k=1}^{m} (\bigotimes_{j \leq d} (\mathbb{R}_{Z_j'}^{(k)})^T) \mathbf{B}_k^T \right) (\bigotimes_{j \leq d} \mathbf{C}_j^{-1})$$

$$\overset{(P1)}{=} \sum_{k=1}^{m} (\bigotimes_{j \leq d} \mathbf{C}_j) \left(\bigotimes_{j \leq d} (\mathbb{R}_{Z_j'}^{(k)})^T \right) \mathbf{B}_k^T (\bigotimes_{j \leq d} \mathbf{C}_j^{-1})$$

$$\overset{(P2)}{=} \sum_{k=1}^{m} (\bigotimes_{j \leq d} \mathbf{C}_j (\mathbb{R}_{Z_j'}^{(k)})^T \mathbf{C}_j^{-1}) (\bigotimes_{j \leq d} \mathbf{C}_j) \mathbf{B}_k^T (\bigotimes_{j \leq d} \mathbf{C}_j^{-1}).$$

Let us define

$$\mathbb{R}_{X_j}^{(k)} = \mathbf{C}_j (\mathbb{R}_{Z_j'}^{(k)})^T \mathbf{C}_j^{-1} = \begin{cases} \mathbf{P}_{X_j}^{(k)} & \text{if } j \in \mathbf{A}_k, \\ \mathbf{I}_j & \text{if } j \notin \mathbf{A}_k. \end{cases}$$

In the case $j \in \mathbf{A}_k$, the distribution π_j is the unique stationary distribution. In the case $j \notin \mathbf{A}_k$, any distribution is an invariant measure, however, we fix it to be π_j. We have

$$\mathbf{P}_X = \sum_{k=1}^m (\bigotimes_{j \leq d} \mathbb{R}_{X_j}^{(k)})(\bigotimes_{j \leq d} \mathbf{C}_j)\mathbf{B}_k^T(\bigotimes_{j \leq d} \mathbf{C}_j^{-1}).$$

From the property (P1) we have that

$$\sum_{k=1}^m (\bigotimes_{j \leq n} \mathbf{C}_j)\mathbf{B}_k^T(\bigotimes_{j \leq n} \mathbf{C}_j^{-1}) = (\bigotimes_{j \leq n} \mathbf{C}_j) \sum_{k=1}^m \mathbf{B}_k^T(\bigotimes_{j \leq n} \mathbf{C}_j^{-1}) = (\bigotimes_{j \leq n} \mathbf{C}_j)(\bigotimes_{j \leq n} \mathbf{C}_j^{-1}) = \mathbf{I},$$

thus Corollary 4.2 implies that $\boldsymbol{\pi} = \bigotimes_{j \leq d} \pi_j$ is the stationary distribution of $\mathbf{P}_X$, thus $\boldsymbol{\rho} = \boldsymbol{\pi}\mathbf{C}$, what is equivalent to (2.3).

$\square$

4.3. *Proof of Theorem 2.4.* To prove the theorem we will construct an $\mathbf{N}$-isolated link Λ, so that $\mathbf{P}_{X^*}'$ and $\mathbf{P}_{\hat{X}}$, given in (2.5) and (2.6) respectively, are intertwined via this link. Consider the matrix $\mathbf{P}_{X_j^*}'$. Define the stochastic matrix $\mathbf{P}_{\hat{X}_j}$ of size $N_j \times N_j$ as:

$$\mathbf{P}_{\hat{X}_j}(i, i') = \begin{cases} 1 - \lambda_i^{(j)} & \text{if } i' = i + 1, \\ \lambda_i^{(j)} & \text{if } i' = i, \\ 0 & \text{otherwise.} \end{cases}$$

Let Λ_j be the link intertwining matrices $\mathbf{P}_{X_j^*}'$ and $\mathbf{P}_{\hat{X}_j}$ given in (3.8). Define

$$\mathbb{R}_{\hat{X}_j^{(k)}} = \begin{cases} \mathbf{P}_{\hat{X}_j} & \text{if } j \in \mathbf{A}_k, \\ \mathbf{I}_j & \text{if } j \notin \mathbf{A}_k. \end{cases}$$

Note that matrices $\mathbf{R}_{X_j^*}'$ and $\mathbb{R}_{\hat{X}_j^{(k)}}$ are also intertwined via Λ_j for any $j = 1, \ldots, d$ and any $k = 1, \ldots, m$. Any link intertwines two identity matrices, which is the case for $j \notin \mathbf{A}_k$. I.e., we have $\Lambda_j \mathbf{R}_{X_j^*}' = \mathbb{R}_{\hat{X}_j^{(k)}} \Lambda_j, j = 1, \ldots, d$. Define

$$\Lambda = \bigotimes_{j \leq d} \Lambda_j.$$

We have

$$\begin{aligned} \Lambda \mathbf{P}_{X^*}' &= \bigotimes_{j \leq d} \Lambda_j \sum_{k=1}^m b_k \left(\bigotimes_{j \leq d} \mathbb{R}_{X_j^{*(k)}}' \right) = \sum_{k=1}^m b_k \left(\bigotimes_{j \leq d} \Lambda_j \mathbb{R}_{X_j^{*(k)}}' \right) \\ &= \sum_{k=1}^m b_k \left(\bigotimes_{j \leq d} \mathbb{R}_{\hat{X}_j^{(k)}} \Lambda_j \right) = \sum_{k=1}^m b_k \left(\bigotimes_{j \leq d} \mathbb{R}_{\hat{X}_j^{(k)}} \right) \bigotimes_{j \leq d} \Lambda_j \end{aligned}$$

Simple calculations yield that $\mathbf{P}_{\hat{X}}$ given in (2.6) can be written as $\sum_{k=1}^m b_k \left(\bigotimes_{j \leq d} \mathbb{R}_{\hat{X}_j^{(k)}} \right)$. Thus, we

have $\Lambda \mathbf{P}_{X^*}' = \mathbf{P}_{\hat{X}} \Lambda$. Now, let us calculate $\hat{\nu} = \nu^* \Lambda^{-1}$ (note that Λ is nonsingular because of the property (P3) and the fact that each $\Lambda_j, j = 1, \ldots, d$ and identity matrices $\mathbf{I}_j$ are nonsingular). In other words, we have $\nu^* = \hat{\nu}\Lambda$. The equation (2.7) holds, since Λ is lower triangular.

Moreover, Λ is $(N_1, \ldots, N_d)$-isolated, since we have

$$\Lambda((i_1, \ldots, i_d), \mathbf{N}) = \prod_{j=1}^{d} \Lambda(i_j, N_j) \overset{(*)}{=} \begin{cases} \prod_{j=1}^{d} \rho_j(1) & \text{if } i_j = N_j \text{ for all j} = 1, \ldots, \text{d}, \\ 0 & \text{otherwise}, \end{cases}$$

where in $\overset{(*)}{=}$ we used Lemma 3.5. Applying Lemma 3.2 completes the proof.

$\square$

5. The outline of an alternative proof of Theorem 2.4: strong stationary duality approach

In Theorem 2.4 we related the absorption time $T^*_{\nu^*, \mathbf{N}}$ of X^* with the absorption time $\hat{T}_{\hat{\nu}, \mathbf{N}}$ of $\hat{X}$. This was done by finding a specific matrix Λ, such that $\Lambda \mathbf{P}_{X^*} = \mathbf{P}_{\hat{X}} \Lambda$, exploiting the existence of such Λ for X^* and $\hat{X}$ being birth and death chains. The exploited Λ is related to spectral polynomials of the stochastic matrix $\mathbf{P}_{X^*}$. Such a link appeared first naturally as a link between an ergodic chain X and an absorbing chain X^*. The proof of Theorem 2.4 in case $q_j(1) = 0$ (i.e., Corollary 2.7 a)) can be different, using intermediate ergodic chain. In this section we will describe its outline.

Strong stationary duality. Let X be an ergodic Markov chain on $\mathbb{E} = \{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$ with the initial distribution ν and the transition matrix $\mathbf{P}_X$. Let $\mathbb{E}^* = \{\mathbf{e}_1^*, \ldots, \mathbf{e}_M^*\}$ be the, possibly different, state space of the absorbing Markov chain X^*, with the transition matrix $\mathbf{P}_{X^*}$ and the initial distribution ν^*, whose unique absorbing state is denoted by $\mathbf{e}_M^*$. Assume that Λ^* is a stochastic $M \times N$ matrix satisfying $\Lambda(\mathbf{e}_M^*, \mathbf{e}) = \pi(\mathbf{e})$. We say that X^* is a *strong stationary dual* (SSD) of X with link Λ^* if

$$\nu = \nu^* \Lambda \quad \text{and} \quad \Lambda^* \mathbf{P}_X = \mathbf{P}_{X^*} \Lambda^*. \tag{5.1}$$

Diaconis and Fill (1990) prove that then the absorption time T^* of X^* is the so called *strong stationary time* for X. This is a random variable T such that X_T has the distribution π and T is independent of X_T. The main application is to study the rate of convergence of an ergodic chain to its stationary distribution, since for such a random variable we always have $d_{TV}(\nu \mathbf{P}_X^k, \pi) \leq sep(\nu \mathbf{P}_X^k, \pi) \leq P(T > k)$, where d_{TV} stands for the *total variation distance*, and *sep* stands for the *separation 'distance'*. For details, see Diaconis and Fill (1990). We say that SSD is **sharp** if T^* corresponds to stochastically the smallest SST, then we have $sep(\nu \mathbf{P}_X^k, \pi) = P(T^* > k)$, the corresponding SST T^* is often called the **fastest strong stationary time (FSST)**.

Strong stationary duality for birth and death chain. Let X be an ergodic birth and death chain on $\mathbb{E} = \{1, \ldots, M\}$, whose time reversal is stochastically monotone with transitions given in (3.4). Diaconis and Fill (1990) show that an absorbing birth and death chain X^* on $\mathbb{E}^* = \mathbb{E}$ with transitions given by

$$\mathbf{P}_{X^*}(i, i-1) = \frac{H(i-1)}{H(i)} p'(i),$$

$$\mathbf{P}_{X^*}(i, i+1) = \frac{H(i+1)}{H(i)} q'(i+1),$$

$$\mathbf{P}_{X^*}(i, i) = 1 - (p'(i) + q'(i+1)),$$

is a *sharp* SSD for X. Here we have

$$H(j) = \sum_{k \leq j} \pi(k), \qquad \Lambda^*(i, j) = \mathbb{1}\{i \leq j\} \frac{\pi(i)}{H(j)}. \tag{5.2}$$

Moreover, starting from an absorbing birth and death chain X^* on $\mathbb{E} = \{1, \ldots, M\}$, whose unique absorbing state is M, Theorem 3.1 in Fill (2009) states that we can find an ergodic chain X (and some stationary distribution π), such that X^* is its *sharp* SSD with the link given in (5.2).

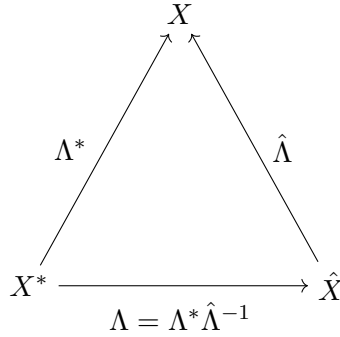FIGURE 5.2. Intertwining between absorbing chains X^* and $\hat{X}$ via ergodic chain X.

Spectral pure-birth chain. Again, let X be an ergodic birth and death chain on $\mathbb{E} = \{1, \ldots, M\}$. Assume its eigenvalues are non-negative, $0 \leq \lambda_1 \leq \ldots \leq \lambda_M = 1$. Then, the chain $\hat{X}$ with transitions given in (3.9) is its *sharp* SSD with the link $\hat{\Lambda}$ given in (3.8).

The outline of an alternative proof. As in Section 4.3, the main idea is to show that two absorbing birth and death chains X_j^* and $\hat{X}_j$ (pure-birth) on $\mathbb{E}_j = \{1, \ldots, N_j\}$ are intertwined by an N_j-isolated link Λ_j. Collecting above findings, we have (skipping conditions on initial distributions):

- Let X_j be an ergodic chain on $\mathbb{E}_j$, whose X_j^* is a sharp SSD, *i.e.*, we have $\Lambda_j^* \mathbf{P}_{X_j} = \mathbf{P}_{X_j^*} \Lambda_j^*$.

- Let $\hat{X}_j$ be a pure-birth sharp SSD for X_j, *i.e.*, we have $\hat{\Lambda}_j \mathbf{P}_X = \mathbf{P}_{\hat{X}_j} \hat{\Lambda}_j$.

It means that absorption times T^* and $\hat{T}$ are equal in distribution (since both X^* and $\hat{X}$ are *sharp* SSDs of X). Mathematically, we have

$$\Lambda_j \mathbf{P}_{X_j^*} = \mathbf{P}_{\hat{X}_j} \Lambda_j, \qquad \text{where} \quad \Lambda_j = \hat{\Lambda}_j \left(\Lambda_j^*\right)^{-1},$$

i.e., X_j^* and $\hat{X}_j$ are intertwined by the link Λ_j, which is N_j-isolated. Intertwining between two absorbing birth and death chains via an ergodic chain is depicted in Fig. 5. Taking $\Lambda = \bigotimes_{j \leq d} \Lambda_j$ and $\hat{\nu} = \nu^* \Lambda^{-1}$ we proceed with the proof of Theorem 2.4 as in Section 4.3.

6. Examples

In first two subsections 6.1 and 6.2 we will present examples on the absorption time of some one-dimensional birth and death chains. Although we mainly focus on multidimensional generalizations, we consider these examples worth presenting. Next two subsections 6.3 and 6.4 contain some non-trivial multidimensional gambler models, for which we either provide results for both, the winning probability and the absorption time (Example 6.3) or only for the winning probability (Example 6.4).

6.1. *A one-dimensional gambler's ruin problem with $N = 3$: calculating $T_{2,3}^*$.* Here we present a simple example for calculating $T_{2,3}^*$ in a one-dimensional gambler's ruin problem using Theorem 2.4. We also check that calculations agree with the formula (1.6).

Example 6.1. Let $d = 1, N_1 = 3$ and $p_1(1) = p_1(2) = p > 0, q_1(1) = q_1(2) = q > 0$ such that $p \neq q$ and $p + q + \sqrt{pq} < 1$. The transition matrix of $\mathbf{P}_{X_1^*}$ is following

$$\mathbf{P}_{X_1^*} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ q & 1-p-q & p & 0 \\ 0 & q & 1-p-q & p \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Then, the pgf of the time to absorption starting at 2 is given by:

$$\mathsf{pgf}_{T_{2,3}^*}(s) = \frac{p(q+p+\sqrt{qp})(-q-p+\sqrt{qp})u(1-u(1-q-p))}{(p^2+qp+q^2)(1-u(1-q-p-\sqrt{qp})(-1+u(1-q-p+\sqrt{qp}))} \tag{6.1}$$

Proof: We have $\mathbf{P}_{X^*} = \mathbf{P}_{X_1^*}$. The eigenvalues of $\mathbf{P}'_{X^*} = \mathcal{F}_0^{-1}(\mathbf{P}_{X^*})$ are $\lambda_1 = 1 - p - q - \sqrt{pq}$, $\lambda_2 = 1 - p - q + \sqrt{pq}$, $\lambda_3 = 1$. The transitions of $\mathbf{P}_{\hat{X}}$ are following

$$\mathbf{P}_{\hat{X}} = \begin{pmatrix} \lambda_1 & 1-\lambda_1 & 0 \\ 0 & \lambda_2 & 1-\lambda_2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Thus,

$$\mathsf{pgf}_{\hat{T}_{1,3}}(s) = \frac{(1-\lambda_1)(1-\lambda_2)s^2}{(1-\lambda_1 s)(1-\lambda_2 s)}, \quad \mathsf{pgf}_{\hat{T}_{2,3}}(s) = \frac{(1-\lambda_2)s}{(1-\lambda_2 s)}.$$

Calculating Λ from (3.8) (for $\mathbf{P}_{X^*}$) we obtain

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 \\ \frac{\sqrt{pq}}{q+p+\sqrt{qp}} & \frac{p}{q+p+\sqrt{qp}} & 0 \\ 0 & 0 & \rho(1) \end{pmatrix}.$$

Calculations yield (we have $\nu^*(2) = 1$)

$$\hat{\nu} = \nu^* \Lambda^{-1} = \left(-\sqrt{\frac{q}{p}}, 1 + \frac{q}{p} + \sqrt{\frac{q}{p}} \right).$$

From (1.6) we have $\rho(i) = \frac{1-\left(\frac{q}{p}\right)^i}{1-\left(\frac{q}{p}\right)^3}, i = 1, 2, 3$. Finally,

$$\mathsf{pgf}_{T_{2,3}^*}(s) = \rho(1) \left(-\sqrt{\frac{q}{p}} \mathsf{pgf}_{\hat{T}_{1,3}}(s) + \left(1 + \frac{q}{p} + \sqrt{\frac{q}{p}} \right) \mathsf{pgf}_{\hat{T}_{2,3}}(s) \right),$$

what can be written as (6.1). On the other hand, (1.7) states that

$$\mathsf{pgf}_{T_{2,3}^*}(s) = \rho(2) \frac{\frac{(1-\lambda_1)(1-\lambda_2)s^2}{(1-\lambda_1 s)(1-\lambda_2 s)}}{\frac{(1-\lambda_1^{\lfloor 2 \rfloor})s}{1-\lambda_1^{\lfloor 2 \rfloor} s}} = \frac{1-\left(\frac{q}{p}\right)^2}{1-\left(\frac{q}{p}\right)^3} \cdot \frac{(1-\lambda_1)(1-\lambda_2)s^2(1-\lambda_1^{\lfloor 2 \rfloor}s)}{(1-\lambda_1 s)(1-\lambda_2 s)(1-\lambda_1^{\lfloor 2 \rfloor})s},$$

where $\lambda_1^{\lfloor 2 \rfloor} = 1 - (p+q)$, which, as can be checked, is equivalent to (6.1). □

6.2. *A one-dimensional gambler's ruin problem related to the Ehrenfest model: calculating $T_{m,N}^*$.* Here we present a concrete example of a birth and death chain on $\mathbb{E} = \{1, \dots, N\}$ with N being the only absorbing state, for which we provide pgf of the absorption time provided the chain started at an arbitrary $m \in \mathbb{E}$. We use Lemma 3.2 to calculate the link Λ. As far as we are aware, this pgf cannot be given using results from Gong et al. (2012) *i.e.*, (1.6). This is because the eigenvalues

of the transition matrix $\mathbf{P}_{X^*}$ are known, but the eigenvalues of the truncated matrix $\mathbf{P}_{X^*}^{[m]}$ are not known for an arbitrary $m \in \mathbb{E}$.

Example 6.2. Let X^* be a Markov chain on the state space $\mathbb{E} = \{1, \ldots, N\}$ with the transition matrix $\mathbf{P}_{X^*}$ of the form:

$$
\mathbf{P}_{X^*}(i,i') = \begin{cases}
\frac{N-i}{2N-2} \frac{\sum_{r=0}^{i-1}\binom{N-1}{r-1}}{\sum_{r=0}^{i}\binom{N-1}{r-1}} & \text{if } i' = i-1, i < N, \\[2ex]
\frac{N-2}{2N-2} & \text{if } i' = i, i < N, \\[2ex]
1 & \text{if } i' = i = N, \\[2ex]
\frac{i}{2N-2} \frac{\sum_{r=0}^{i+1}\binom{N-1}{r-1}}{\sum_{r=0}^{i}\binom{N-1}{r-1}} & \text{if } i' = i+1, \\[2ex]
0 & \text{otherwise.}
\end{cases}
$$

Then the absorption time starting at $m \in \mathbb{E}$ has the following pgf:

$$
\mathsf{pgf}_{T_{m,d}^*}(s) = \sum_{j \leq m} \hat{\nu}(j)\mathsf{pgf}_{\hat{T}_{j,d}}(s), \tag{6.2}
$$

where

$$
\hat{\nu}(j) = \frac{2^{j-1}(-1)^{m+j}(m-j+1)\binom{N-1}{m}\binom{m}{j-1}}{(N-j)\sum_{k=0}^{m-1}\binom{N-1}{k}} \quad \text{and} \quad \mathsf{pgf}_{\hat{T}_{j,N}}(s) = \prod_{k=j}^{N-1}\left[\frac{(1-\frac{k-1}{N-1})s}{1-\frac{k-1}{N-1}s}\right]. \tag{6.3}
$$

In particular, we have

$$
E(T_{m,N}^*) = (N-1)\sum_{j \leq m}\hat{\nu}(j)\sum_{k=j}^{N-1}\frac{1}{N-k}. \tag{6.4}
$$

Proof: Let

$$
\mathbf{P}_{\hat{X}}(i,i') = \begin{cases}
\frac{i-1}{N-1} & \text{if } i' = i, \\[2ex]
\frac{N-i}{N-1} & \text{if } i' = i+1, \\[2ex]
0 & \text{otherwise.}
\end{cases}
$$

To show the result using Lemma 3.2 it is enough to find Λ such that $\Lambda \mathbf{P}_{X^*} = \mathbf{P}_{\hat{X}}\Lambda$ and $\nu^*\Lambda^{-1} = \hat{\nu}$, where $\nu^*(j) = \mathbb{1}\{j = m\}$.

However, since X^* has only one absorbing state, we can – and we will – follow *the outline of an alternative proof* given in Section 5. *I.e.*, we will indicate intermediate ergodic chain X on $\mathbb{E}$ with the transition matrix $\mathbf{P}_X$ and find Λ^* and $\hat{\Lambda}$ such that $\Lambda^*\mathbf{P}_X = \mathbf{P}_{X^*}\Lambda^*$ and $\hat{\Lambda}\mathbf{P}_X = \mathbf{P}_{\hat{X}}\hat{\Lambda}$. Then, we will automatically have $\Lambda = \hat{\Lambda}(\Lambda^*)^{-1}$ and we will show that $\nu^*\Lambda^{-1} = \nu^*\Lambda^*\hat{\Lambda}^{-1} = \hat{\nu}$.

Let X be a chain on $\mathbb{E}$ with the following transition matrix:

$$
\mathbf{P}_X(i,i') = \begin{cases}
\frac{i-1}{2(N-1)} & \text{if } i' = i-1, \\[2ex]
\frac{1}{2} & \text{if } i' = i, \\[2ex]
\frac{N-i}{2(N-1)} & \text{if } i' = i+1, \\[2ex]
0 & \text{otherwise,}
\end{cases}
$$

i.e., X corresponds to the Ehrenfest model of $N-1$ particles with an extra probability (half) of staying (and states are enumerated $1, \ldots, N$, whereas in the classical Ehrenfest model these are

$0, \ldots, N-1$). Its stationary distribution is the binomial distribution $\pi(j) = \frac{1}{2^{N-1}} \binom{N-1}{j-1}$, thus the classical link (cf. (5.2)) is given by

$$\Lambda^*(i,j) = \frac{\binom{N-1}{j-1}}{\sum_{r=0}^{i-1} \binom{N-1}{r}} \mathbb{1}\{j \leq i\},$$

i.e., we have $\Lambda^* \mathbf{P}_X = \mathbf{P}_{X^*} \Lambda$ (X^* is a sharp SSD of X). The eigenvalues of X are known, these are $\frac{i}{N-1}, i = 0, \ldots, N-1$, thus $\hat{X}$ is its pure-birth spectral dual. The link $\hat{\Lambda}$ such that $\hat{\Lambda} \mathbf{P}_X = \mathbf{P}_{\hat{X}} \hat{\Lambda}$ is known (see Eq. (4.6) in Fill (2009)), it is given by

$$\hat{\Lambda}(i,j) = \frac{\binom{i-1}{j-1}}{2^{i-1}}.$$

It can be checked that

$$\hat{\Lambda}^{-1}(i,j) = (-1)^{j-i} 2^{j-1} \binom{i-1}{j-1}.$$

Note that the i-th row of $\hat{\Lambda}^{-1}$ corresponds to the coefficients[1] in expansion of $(2x-1)^{i-1}$. Thus, as outlined in Section 5 we have $\Lambda \mathbf{P}_{X^*} = \mathbf{P}_{\hat{X}} \Lambda$ with $\Lambda = \hat{\Lambda}(\Lambda^*)^{-1}$. We need only to check that $\nu^* \Lambda^{-1} = \nu^* \Lambda^* \hat{\Lambda}^{-1}$ is equal to $\hat{\nu}$ given in (6.3). We have

$$
\begin{aligned}
\hat{\nu}(j) &= (\nu^* \Lambda^* \hat{\Lambda}^{-1})(j) = \sum_k \Lambda^*(m,k) \hat{\Lambda}^{-1}(k,j) \\
&= \sum_{j \leq k \leq m} \frac{\binom{N-1}{k-1}}{\sum_{r=0}^{m-1} \binom{N-1}{r}} (-1)^{j-k} 2^{j-1} \binom{k-1}{j-1} \\
&= \frac{2^{j-1}}{\sum_{r=0}^{m-1} \binom{N-1}{r}} \sum_{j \leq k \leq m} (-1)^{j-k} \binom{N-1}{k-1} \binom{k-1}{j-1} \\
&\overset{(*)}{=} \frac{2^{j-1} \binom{N-1}{j-1}}{\sum_{r=0}^{m-1} \binom{N-1}{r}} \sum_{j \leq k \leq m} (-1)^{j-k} \binom{N-j}{N-k},
\end{aligned}
$$

where in $\overset{(*)}{=}$ we used the identity $\binom{N-1}{k-1} \binom{k-1}{j-1} = \binom{N-1}{j-1} \binom{N-j}{N-k}$. As for the last sum we have

$$
\begin{aligned}
\sum_{j \leq k \leq m} (-1)^{j-k} \binom{N-j}{N-k} &= \sum_{j \leq k \leq m} (-1)^{j-k} \binom{N-j}{k-j} \\
&= \sum_{k=0}^{m-j} (-1)^k \binom{N-j}{k} \overset{(**)}{=} (-1)^{m-j} \binom{N-j-1}{m-j},
\end{aligned}
$$

where in $\overset{(**)}{=}$ we used the identity[2] $\sum_{k=0}^{M} (-1)^k \binom{n}{k} = (-1)^M \binom{n-1}{M}$. Finally,

$$\hat{\nu}(j) = \frac{2^{j-1} (-1)^{m-j} \binom{N-1}{j-1} \binom{N-j-1}{m-j}}{\sum_{r=0}^{m-j} \binom{N-1}{r}},$$

what is equal to (6.3).

Note that the pgf given in (6.3) corresponds to the distribution of $\sum_{k=j}^{N-1} Y_k$, where Y_k is a geometric random variable with parameter $\frac{k-1}{N-1}$ and $Y_1, \ldots, Y_{N-1}$ are independent. We have $EY_k = \frac{N-1}{N-k}$ thus (6.4) follows from (6.2) and (6.3). □

---

[1]The on-line encyclopedia of integer sequences, sequence http://oeis.org/A303872.
[2]See *Partial sums* at https://en.wikipedia.org/wiki/Binomial_coefficient

Note that calculating $\hat{\nu}$ we have actually calculated the link Λ, which is given by

$$\Lambda(i,j) = \begin{cases} \frac{2^{j-1}(-1)^{i+j}(i-j+1)\binom{N-1}{i}\binom{i}{j-1}}{(N-j)\sum_{k=0}^{i-1}\binom{N-1}{k}} & \text{if } j < N, \\ 0 & j = N, i < N, \\ 1 & j = N, i = N. \end{cases}$$

Next two subsections 6.3 and 6.4 contain results for some non-trivial multidimensional gambler models.

6.3. *A multidimensional case, winning probabilities and the absorption time: changing r coordinates at one step in a d-dimensional game.* We will present an example for both Theorems, 2.1 and 2.4. The chains $\mathbf{P}_{Z_j^{(k)}}$ in Theorem 2.1 are quite general, but in this example we consider birth and death chains *i.e.*, we will use $\mathbf{P}_{Z_j^{(k)}} = \mathbf{P}_{X_j^*}$ from Theorem 2.4 (birth and death chains given in (1.4)). Similarly, we have $\mathbb{R}'_{Z_j^{(k)}} = \mathbb{R}'_{X_j^{*(k)}}$ and $\mathbf{P}_Z = \mathbf{P}_{X^*}$.

Example 6.3. The idea of the example is following. We construct a d-dimensional game from one-dimensional games, in such a way, that at one step we play with r other players, where $r \in \{1, \ldots, d\}$. In other words, the multidimensional chain can change at most r coordinates in one step.

Moreover we will take, as $\mathbf{B}_i := b_i$ real numbers. In both theorems let us take $0 < r < d$, $m = \binom{d}{r} + 1$ and $b_k = 1, k = 1, \ldots, m-1, b_m = 1 - \binom{d}{r}$. Let us enumerate combinations of r positive numbers no greater than d in some way (see *e.g.*, Mudrov (1965)). Let $\mathbf{A}_k$ be k-th combination from this numbering, for $k = 1, \ldots, m-1$ and $\mathbf{A}_m = \emptyset$. Then we have

$$\mathbf{P}'_Z = \sum_{k=1}^{m} \mathbf{B}_k \left(\bigotimes_{j \le d} \mathbb{R}'_{Z_j^{(k)}} \right) = \sum_{k=1}^{\binom{d}{r}} \left(\bigotimes_{j \le d} \mathbb{R}'_{Z_j^{(k)}} \right) - \left(\binom{d}{r} - 1 \right) \bigotimes_{j \le d} \mathbf{I}_j. \tag{6.5}$$

We have that $\mathbb{R}'_{Z_j^{(k)}} = \mathbf{P}'_{Z_j}$ if $\{j\} \in \mathbf{A}_k$ and $\mathbb{R}'_{Z_j^{(k)}} = \mathbf{I}_j$ otherwise (for $\{j\} \notin \mathbf{A}_k$), thus $\mathbf{P}'_Z =$

$$\sum_{1 \le i_1 < i_2 < \ldots < i_r \le d} \left(\bigotimes_{j < i_1} \mathbf{I}_j \right) \otimes \mathbf{P}'_{Z_{i_1}} \otimes \left(\bigotimes_{i_1 < j < i_2} \mathbf{I}_j \right) \otimes \ldots \otimes \left(\bigotimes_{i_{r-1} < j < i_r} \mathbf{I}_j \right) \otimes \mathbf{P}'_{Z_{i_r}} \otimes \left(\bigotimes_{i_r < j \le d} \mathbf{I}_j \right) - \left(\binom{d}{r} - 1 \right) \bigotimes_{j \le d} \mathbf{I}_j.$$

In other words, we combine d one-dimensional birth and death chains in such a way, that the resulting d-dimensional chain can change at most r coordinates by ± 1 at one step. We can rewrite this formula for some cases:

- $r = d$, independent games

$$\mathbf{P}'_Z = \bigotimes_{j \le d} \mathbf{P}'_{Z_j}.$$

- $r = d - 1$

$$\mathbf{P}'_Z = \sum_{k=1}^{d} \left(\bigotimes_{j < k} \mathbf{P}'_{Z_j} \right) \otimes \mathbf{I}_k \otimes \left(\bigotimes_{j > k} \mathbf{P}'_{Z_j} \right) - (d-1) \bigotimes_{j \ge d} \mathbf{I}_j.$$

- $r = 2$

$$\mathbf{P}'_Z = \sum_{k=1}^{d} \sum_{i=k+1}^{d} \left(\bigotimes_{j < k} \mathbf{I}_j \right) \otimes \mathbf{P}'_{Z_k} \otimes \left(\bigotimes_{k < j < i} \mathbf{I}_j \right) \otimes \mathbf{P}'_{Z_i} \otimes \left(\bigotimes_{i < j \le d} \mathbf{I}_j \right) - \left(\binom{d}{2} - 1 \right) \bigotimes_{j \ge d} \mathbf{I}_j.$$

- $r = 1$

$$\mathbf{P}'_Z = \sum_{k=1}^{d} \left(\bigotimes_{j < k} \mathbf{I}_j \right) \otimes \mathbf{P}'_{Z_k} \otimes \left(\bigotimes_{j > k} \mathbf{I}_j \right) - (d-1) \bigotimes_{j \ge d} \mathbf{I}_j.$$

This can be rewritten as

$$\mathbf{P}'_Z = \bigoplus_{j \leq d} \mathbf{P}'_{Z_j} - (d-1) \bigotimes_{j \leq d} \mathbf{I}_j.$$

$\mathbf{P}_Z = \mathcal{F}_{-\infty}(\mathbf{P}'_Z)$ are exactly the transition corresponding to the generalized gambler's ruin problem given in (1.1).

In all above cases, the winning probability for the chain $\mathbf{P}_Z$ is given in (1.3). This is since the winning probabilities for $\mathbf{P}_{Z_j}$ are given in (2.4), thus using (2.3) the relation is (1.3) proven.

In all above cases, if we replace $\mathbf{P}'_{Z_j}$ with $\mathbf{P}_{\hat{X}_j}$ and $\mathbf{P}'_Z$ with $\mathbf{P}_{\hat{X}}$, then we have a special cases for formula for $\mathbf{P}_{\hat{X}}$ given in (2.6). If, in addition, we assume that $\nu^*((1,\ldots,1)) = 1$, then from Corollary 2.7 c) we have

$$T^*_{(1,\ldots,1),\mathbf{N}} = \begin{cases} \hat{T}_{(1,\ldots,1),\mathbf{N}} & \text{with probability } \prod_{j=1}^{d} \rho_j(1), \\ +\infty & \text{with probability } 1 - \prod_{j=1}^{d} \rho_j(1). \end{cases}$$

For example, in case $r = 1$ (then we have $m = d+1$ and take $b_k = 1, k = 1,\ldots,d, b_{d+1} = 1-d, \mathbf{A}_k = \{k\}, k = 1,\ldots,d$ and $\mathbf{A}_{d+1} = \emptyset$) we have

$$\mathbf{P}_{\hat{X}}((i_1,\ldots,i_d),(i'_1,\ldots,i'_d)) = \begin{cases} 1 - \lambda_{i_j}^{(j)} & \text{if } i'_j = i_j + 1, \\ 1 - \sum_{j:i_j < N_j} \left(1 - \lambda_{i_j}^{(j)}\right) & \text{if } i'_j = i_j, j = 1,\ldots,d, \\ 0 & \text{otherwise.} \end{cases}$$

Sample transitions for case $d = 2, r = 1$ are depicted in Fig. 1.1.

In Figure 6.3 the transitions of $\hat{X}$ are presented for $d = 3$:

- When $r = 1$ only transitions along blue dotted arrows (••••▸) are possible.
- When $r = 2$ only transitions along blue dotted arrows (••••▸) and green dashed arrows (--▸) are possible.
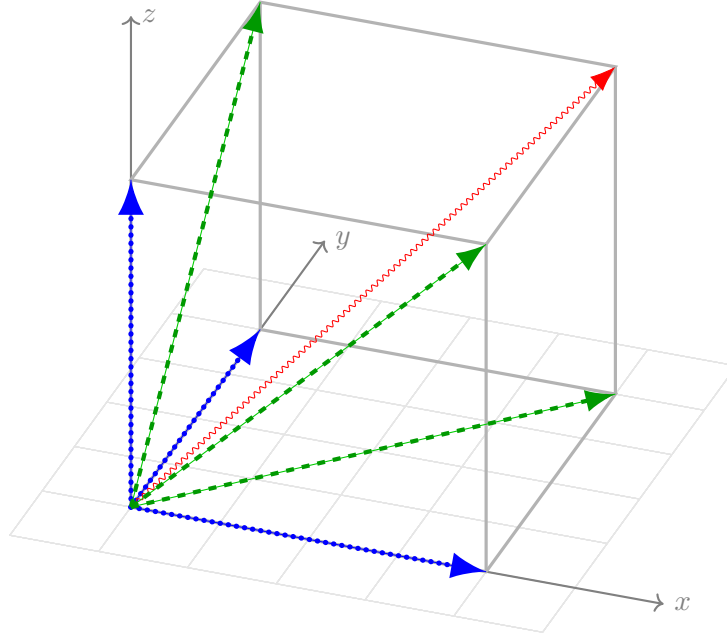- When $r = 3$ all transitions, along blue arrows (••••▸), green dashed arrows (--▸) and red curly arrow (⌇⌇▸) are possible.

FIGURE 6.3. Sample transitions of $\hat{X}$ for the example from Section 6.3 with $d = 3$: transitions for $r = 1$ (blue ▪▪▪▶), $r = 2$ (blue ▪▪▪▶ and green ▪▪▶) and $r = 3$ (blue ▪▪▪▶, green ▪▪▶ and red ⌇⌇▶).

6.4. *A multidimensional case, winning probabilities: changing r (dependent on the current fortune) coordinates at one step in a d-dimensional game.*

   Now we will make use of the possibility that $\mathbf{B}_k$'s appearing in Theorem 2.1 can be matrices. We will provide an extension of the previous Example 6.3 – this time we will only provide the result on the winning probabilities (since the result on the absorption time provided in Theorem 2.4 requires $\mathbf{B}_k$'s to be numbers).

Example 6.4. Before providing intricate details of the example, let us clarify what we aim to achieve. In the previous Example 6.3 we showed how one may construct a multidimensional chain, so that at most $r \le d$ coordinates can be changed in one step (in other words, we can play with at most $r \le d$ players in one step). Now we extend this situation: we want r to be state-dependent. To be more exact: if we are in a state from a set $\mathbf{S}_r$ then we may play with at most r players. For example setting:

$$\mathbf{S}_d = \{(i_1, \ldots, i_d) : \sum_{j=1}^{d} i_j \ge 100\}, \quad \mathbf{S}_2 = \{(i_1, \ldots, i_d) : \sum_{j=1}^{d} i_j < 100\}$$

and $\mathbf{S}_r = \emptyset$ for $r \notin \{2, d\}$ models a game in which:

   • If we have at least 100 dollars we may play with all the players;
   • If we have less than 100 dollars we may play with at most 2 players.

   Consider a partition of the state space into disjunctive sets $\mathbf{S}_r, r = 0, \ldots, d$, i.e., $\mathbb{E} = \bigcup_{r=0}^{d} \mathbf{S}_r$. Each $\mathbf{S}_r$ is a set of states from which we can change at most r coordinates in one step. Let $\mathbf{I}_{\mathbf{S}_r}$ be a matrix with ones only on positions (i, i), where $i \in \mathbf{S}_r$. Let $m_r = \binom{d}{r} + 1$ and set $\mathbf{B}_i^r = \mathbf{I}_{\mathbf{S}_r}$ for $1 \le i < m_r$ and $\mathbf{B}_{m_r}^r = \left(1 - \binom{d}{r}\right) \mathbf{I}_{\mathbf{S}_r}$. Let us enumerate combinations of r positive numbers no greater than d in some way (again, see *e.g.*, Mudrov (1965)). Let $\mathbf{A}_i^r$ be i-th combination from this numbering, for $i = 1, \ldots, m_r - 1$ and $\mathbf{A}_{m_r}^r = \emptyset$.

Now we will renumerate everything to fit to the notation used in Theorem 2.1. For $1 \leq k \leq m$ there exists $1 \leq b \leq d, 1 \leq i \leq m_{b+1}$ such that $k = \left(\sum_{r=1}^{b-1} m_r\right) + i$, we then set $\mathbf{B}_k = \mathbf{B}_{\left(\sum_{r=1}^{b-1} m_r\right)+i} = \mathbf{B}_i^b$ and $\mathbf{A}_k = \mathbf{A}_{\left(\sum_{r=1}^{b-1} m_r\right)+i} = \mathbf{A}_i^b$. In this notation – using (2.2) – the transition matrix of the resulting chain is given by:

$$\mathbf{P}'_Z = \sum_{k=1}^{m} \mathbf{B}_k \left(\bigotimes_{j \leq d} \mathbb{R}'_{Z_j^{(k)}}\right).$$

However, we can rewrite it in a more intuitive way, letting $k_b = \sum_{r=1}^{b} m_r$ we have:

$$\begin{aligned}
\mathbf{P}'_Z &= \sum_{b=1}^{d} \sum_{i=1}^{m_b} \mathbf{B}_{k_{b-1}+i} \left(\bigotimes_{j \leq d} \mathbb{R}'_{Z_j^{(k_{b-1}+i)}}\right) \\
&= \sum_{b=1}^{d} \left\{ \sum_{i=1}^{m_b} \mathbf{I}_{\mathbf{S}_b} \left(\bigotimes_{j \leq d} \mathbb{R}'_{Z_j^{(k_{b-1}+i)}}\right) - \mathbf{I}_{\mathbf{S}_b} \left(\binom{d}{b} - 1\right) \bigotimes_{j \leq d} \mathbf{I}_j \right\} \\
&= \sum_{b=1}^{d} \mathbf{I}_{\mathbf{S}_b} \left\{ \sum_{i=1}^{m_b} \left(\bigotimes_{j \leq d} \mathbb{R}'_{Z_j^{(k_{b-1}+i)}}\right) - \left(\binom{d}{b} - 1\right) \bigotimes_{j \leq d} \mathbf{I}_j \right\}.
\end{aligned}$$

Thus, $\mathbf{P}'_Z$ can be rewritten as $\mathbf{P}'_Z = \sum_{b=1}^{d} \mathbf{I}_{S_b} \mathbf{P}'^b_Z$, where matrices $\mathbf{P}'^b_Z$ correspond to the previous Example 6.3, cf. (6.5). Concluding, Theorem 2.1 implies that the winning probability of the chain $\mathbf{Z}$ with the transition matrix $\mathbf{P}_Z = \mathcal{F}_{-\infty}(\mathbf{P}'_Z)$ is of the product form provided in (1.3).

Acknowledgement

References

Asmussen, S. r. and Albrecher, H. *Ruin probabilities*, volume 14 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, second edition (2010). ISBN 978-981-4282-52-9; 981-4282-52-9. MR2766220.

Asmussen, S. r. and Sigman, K. Monotone stochastic recursions and their duals. *Probab. Engrg. Inform. Sci.*, **10** (1), 1–20 (1996). MR1386851.

Champagnat, N., Schott, R., and Villemonais, D. Probabilistic Non-asymptotic Analysis of Distributed Algorithms. *ArXiv Mathematics e-prints* (2018). arXiv: 1802.02644.

Diaconis, P. and Fill, J. A. Strong stationary times via a new form of duality. *Ann. Probab.*, **18** (4), 1483–1522 (1990). MR1071805.

Fill, J. A. The passage time distribution for a birth-and-death chain: strong stationary duality gives a first stochastic proof. *J. Theoret. Probab.*, **22** (3), 543–557 (2009). MR2530102.

Gong, Y., Mao, Y.-H., and Zhang, C. Hitting time distributions for denumerable birth and death processes. *J. Theoret. Probab.*, **25** (4), 950–980 (2012). MR2993011.

Keilson, J. *Markov chain models—rarity and exponentiality*, volume 28 of *Applied Mathematical Sciences*. Springer-Verlag, New York-Berlin (1979). ISBN 0-387-90405-0. MR528293.

Kmet, A. and Petkovšek, M. Gambler's ruin problem in several dimensions. *Adv. in Appl. Math.*, **28** (2), 107–118 (2002). MR1888839.

Lengyel, T. The conditional gambler's ruin problem with ties allowed. *Appl. Math. Lett.*, **22** (3), 351–355 (2009). MR2483497.

Lengyel, T. Gambler's ruin and winning a series by m games. *Ann. Inst. Statist. Math.*, **63** (1), 181–195 (2011). MR2748940.

Lindley, D. V. The theory of queues with a single server. *Proc. Cambridge Philos. Soc.*, **48**, 277–289 (1952). MR46597.

Lorek, P. Generalized gambler's ruin problem: explicit formulas via Siegmund duality. *Methodol. Comput. Appl. Probab.*, **19** (2), 603–613 (2017). MR3649560.

Lorek, P. Siegmund duality for Markov chains on partially ordered state spaces. *Probab. Engrg. Inform. Sci.*, **32** (4), 495–521 (2018). MR3876149.

Mao, Y.-H. and Zhang, C. Hitting time distributions for birth-death processes with bilateral absorbing boundaries. *Probab. Engrg. Inform. Sci.*, **31** (3), 345–356 (2017). MR3687837.

Mudrov, V. An algorithm for numbering combinations. *USSR Computational Mathematics and Mathematical Physics*, **5** (4), 280–282 (1965). DOI: 10.1016/0041-5553(65)90142-4.

Rocha, A. L. and Stern, F. The asymmetric n-player gambler's ruin problem with equal initial fortunes. *Adv. in Appl. Math.*, **33** (3), 512–530 (2004). MR2081041.

Ross, S. M. A simple solution to a multiple player gambler's ruin problem. *Amer. Math. Monthly*, **116** (1), 77–81 (2009). MR2478757.

Siegmund, D. The equivalence of absorbing and reflecting barrier problems for stochastically monotone Markov processes. *Ann. Probability*, **4** (6), 914–924 (1976). MR431386.

Tzioufas, A. The several dimensional gambler's ruin problem. *Markov Process. Related Fields*, **25** (1), 101–123 (2019). MR3965095.