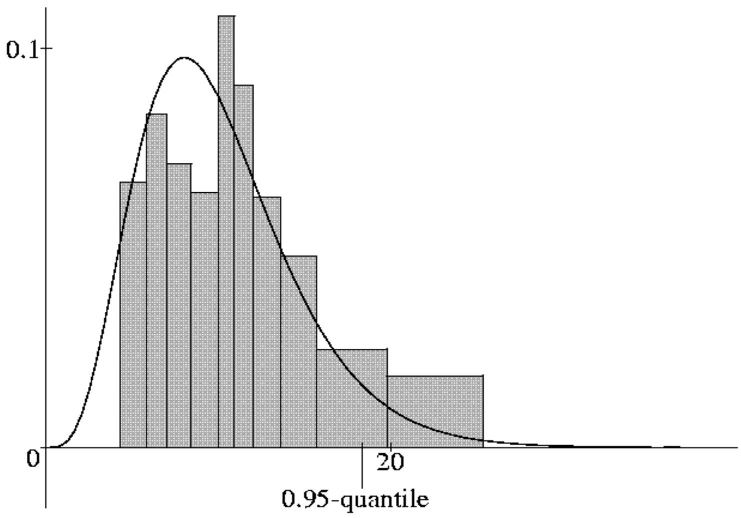
UNIVERSITÄT POTSDAM
Institut für Mathematik

On Time Duality for Markov Chains with Absorbing Boundaries

Peter Keller
Sylvie Roelly
Angelo Valleriani



Mathematische Statistik und Wahrscheinlichkeitstheorie

Universität Potsdam – Institut für Mathematik

Mathematische Statistik und Wahrscheinlichkeitstheorie

On Time Duality for Markov Chains with Absorbing Boundaries

Peter Keller
Max Planck Institute for Colloids and Interfaces, Golm
and Institute of Mathematics, University of Potsdam

Sylvie Roelly
Institute of Mathematics, University of Potsdam

Angelo Valleriani
Max Planck Institute for Colloids and Interfaces, Golm

Preprint 2011/03

Juli 2011

Impressum

Herausgeber: Mathematische Statistik und Wahrscheinlichkeitstheorie
am Institut für Mathematik

Adresse: Universität Potsdam
Am Neuen Palais 10
14469 Potsdam
Telefon: +49-331-977 1500
Fax: +49-331-977 1578
E-mail: neisse@math.uni-potsdam.de

ISSN 1613-3307

Applied Probability Trust (30 June 2011)

ON TIME DUALITY FOR MARKOV CHAINS WITH ABSORBING BOUNDARIES

PETER KELLER,* *Max Planck Institute for Colloids and Interfaces, Golm & Institute of Mathematics, University of Potsdam*

SYLVIE ROELLY,** *Institute of Mathematics, University of Potsdam*

ANGELO VALLERIANI,*** *Max Planck Institute for Colloids and Interfaces, Golm*

Abstract

We consider finite state Markov Chains in continuous time with two absorbing states a and b which can be reached only via a single transition. We show that, under a condition called *nearly reversibility*, the first hitting time of a and the first hitting time of b are equal in distribution for certain choices of initial states: this property will be called *Time Duality.* We note that Birth and Death processes are nearly reversible. Moreover, since nearly reversibility is not necessary for Time Duality to hold, we propose a new necessary and sufficient algebraic condition under which general Markov Chains fulfill Time Duality. In addition we show that a certain invariance of their transition graph under similarity transforms implies Time Duality.

Keywords: continuous time Markov Chain; hitting times; time duality; absorbing boundary

2010 Mathematics Subject Classification: Primary 60J27

Secondary 60J28;60K40

1. Introduction

Hitting times play an important role in many applications of Markov Chains. One example is provided by theories that describe molecular motor proteins e.g. kinesin.

* Postal address: Am Neuen Palais 10,D-14469 Potsdam, Am Mühlenberg 1, D-14476 Potsdam-Golm

** Postal address: Am Neuen Palais 10, D-14469 Potsdam

*** Postal address: Am Mühlenberg 1, D-14476 Potsdam-Golm

In living eukaryotic cells, kinesin "walks" over long, relatively rigid strings called microtubules thereby transporting a cargo from one side of the cell to another. Due to the periodic molecular structure of the microtubules, the steps of the kinesin have all the same length equal to 8 nanometers. Under normal conditions present in living cells, this motor performs a random walk in one dimension with a drift on the microtubule. Experimental studies have led to rather detailed and successful models of the chemical processes in kinesin which control the movement of the motor. These chemical processes are characterized by a network of internal chemical states, see e.g. [8]. The analysis of kinesin and similar motor proteins is made by means of Markov Chains in continuous time (MCc), as the transitions between chemical states are usually interpreted as transition graph of a Markov Chain. The Markov Chain is linear periodic in the sense that copies of the same finite state MCcare connected via bottleneck transitions in a linear fashion, see fig. 1.

Each copy has two distinct special states α and β. These states act as bottlenecks that connect the copies, such that the dynamic has to pass through α to make a transition to the left copy and through β to the copy at the right side, see also fig. 1.

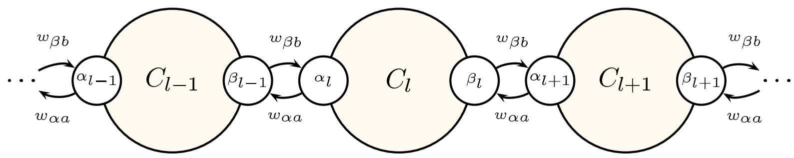


FIGURE 1: Each C_l is connected with two neighboring copies, transitions to the right resp. to the left have to pass through α_l resp. β_l.

For the analysis of such a model first hitting times of the set $\{\alpha_{l+1}, \beta_{l-1}\}$ starting from either α_l resp. β_l are important, as they are interpreted as the time the motor needs to perform a single step forwards resp. backwards.

Lindén and Wallin derive in [7] through statistical analysis of experiments that for a certain class of motor types the first hitting time of α_{l+1} starting in α_l and the first hitting time of β_{l-1} starting in β_l have the same distribution.

To characterize this "Time Duality" (TD) we introduce the following formalism: Let $(X_t)_{t\geq 0}$ be a MCc on a finite state space E with two absorbing states a and b (note that this is equal to the periodic model above stopped at the very instant a neighboring copy is reached). Furthermore let α resp. β be the only state in E communicating with a resp. b, see fig. 2.

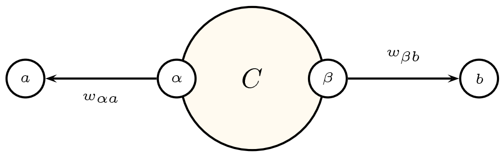


FIGURE 2: The disk denoted with C is a finite collection of states such that every $c_i \in C$ communicates with α and β; the states a and b are absorbing.

A nearly reversibility condition, which will be made precise in section 3, is sufficient for TD, this is already known, see e.g. [6] or [12], section III. We give a new and simple proof based on direct comparison of distributions.

But new models of motor proteins such as in [11] show that the TD appears also in models that do not have the nearly reversibility property.

Time Duality in general can be characterized by an “intertwining relation“. Let $Q_{|a}$ be the infinitesimal generator of the process conditioned to absorption in a and define $Q_{|b}$ analogous (see section 2 for the construction of $Q_{|a}$ and $Q_{|b}$). Then the time duality can be characterized via the existence of a matrix Λ such that the following intertwining relation holds:

$$\Lambda Q_{|a} = Q_{|b}\Lambda. \tag{1}$$

The matrix Λ is also stochastic and can be interpreted as a ”stochastic link“ between $Q_{|a}$ and $Q_{|b}$, see e.g. in [1] and [10] and references therein. A direct way to construct Λ is given in [3], but turns out to be non treatable for theoretical purposes. Therefore our aim is to give a simple criterion to check if a model like in fig. 2 satisfies TD and show that it does not depend on the rates $w_{\alpha a}$ and $w_{\beta b}$.

The organization of the article is as follows. In section 2 we derive expressions for the densities and moment generating functions of τ_a and τ_b for arbitrary initial starting point. Section 3 provides a new direct and simple proof of the time duality in the case of nearly reversibility. As a corollary we get the well known Time Duality result for Birth and Death processes with absorbing boundaries. Section 4 gives a necessary and sufficient condition for Time Duality. We also prove that nearly reversibility is not necessary for TD. In section 5 we use the derived conditions to show that invariance under a certain form of similarity transform of the infinitesimal generator is sufficient for TD and give some non trivial examples to illustrate the connection to the geometry of the associated transition graph.

2. The Model and the Conditioned Dynamics

2.1. The Model

Assume that the MCc $(X_t)_{t\geq 0}$ introduced in section 1 fig. 2, is defined on the finite state space $E = \{a, b\} \cup C \cup \{\alpha, \beta\}$, where C is a finite collection of states such that $C \cup \{\alpha, \beta\}$ is a single communication class and a and b are absorbing states.

We partition the infinitesimal generator Q associated to $(X_t)_{t\geq 0}$ into blocks in the following way:

$$Q := \begin{pmatrix} 0 & 0 & 0 \\ R_{Sa}^\top & S & R_{Sb}^\top \\ 0 & 0 & 0 \end{pmatrix} \tag{2}$$

with row vectors

$$R_{Sa} := w_{\alpha a} e_\alpha := w_{\alpha a}(1, 0, \ldots, 0)$$
$$R_{Sb} := w_{\beta b} e_\beta := w_{\beta b}(0, \ldots, 0, 1)$$

and a matrix S which contains all transition rates between the transient states in $C \cup \{\alpha, \beta\}$. We therefore call S *transientrix.* Without loss of generality we assume that the first row of S encodes the transitions related to α and the last row those of β.

The next lemma can be found in [9], but we rephrase it here giving a new proof which uses spectral and probabilistic arguments.

Lemma 1. *Let Q be the infinitesimal generator of a MCc on a finite state space and T be the submatrix of Q which contains all transitions between the transient states. Then T is invertible.*

Proof. Suppose that T is not invertible; Thus there exists a vector $y \neq 0$ such that $yT = 0$ (i.e. 0 is an eigenvalue of T). Then

$$y \exp(T\xi)\mathbf{1}^\top = y\mathbf{1}^\top$$

for $\mathbf{1} := (1, 1, \dots, 1)$ and each $\xi \in \mathbb{R}$ and thus

$$\lim_{\xi \to \infty} y \exp(T\xi)\mathbf{1}^\top = y\mathbf{1}^\top \neq 0,$$

which is a contradiction to the assumption, as for each $\xi \in \mathbb{R}^+$ the quantity $\exp(T\xi)\mathbf{1}^\top$ is the probability that the process is still in the set of transients for each $\xi \geq 0$.

□

2.2. Conditioned Dynamics

In this section we investigate the behavior of $(X_t)_{t\geq 0}$ conditioned to absorption in either a or b and derive expressions for the laws of the first hitting times τ_a and τ_b according to the initial state.

The conditioning is a "Doob h-transform", see e.g. [5] chapter 8 for an introduction to this technique. Here we compute it directly without explicit use of potential theory. As the calculations are analogue if the process is conditioned to absorption in b instead of a we restrict our attention to the case of absorption in a. The distributions and moment generating functions are derived for both cases in Lemma 3.

As usual define the first hitting times by

$$\begin{aligned}
\tau_{\{a,b\}} &:= \inf\{t \geq 0 : X_t \in \{a, b\}\} \\
\tau_a &:= \inf\{t \geq 0 : X_t = a\} \\
\tau_b &:= \inf\{t \geq 0 : X_t = b\}.
\end{aligned}$$

Obviously $\tau_{\{a,b\}} = \tau_a \wedge \tau_b$.

The random time $\tau_{\{a,b\}}$ is the time until hitting the set of absorbing states $\{a, b\}$. It is a classic result that this time is almost surely finite, i.e.

$$\mathbb{P}(\tau_{\{a,b\}} < \infty | X_0 = i) = 1$$

for any $i \in E$.

Note that although $\tau_{\{a,b\}}$ is almost surely finite, this is not true for τ_a and τ_b. If e.g. the process reaches the state a before b than τ_b is infinite. Therefore the distributions of τ_a and τ_b are defined on $\mathbb{R}^+ \cup \{+\infty\}$.

To shorten notation let $\mathbb{P}_i(X_t = .) := \mathbb{P}(X_t = .|X_0 = i)$.

Lemma 2. *The transition function of* $(X_t)_{t\geq 0}$ *conditioned to absorption in* a *is*

$$\mathbb{P}_i(X_t = j|\tau_a < \tau_b) = \frac{h_a(j)}{h_a(i)}\mathbb{P}_i(X_t = j), \;\; i, j \in E \setminus \{b\} \tag{3}$$

where

$$h_a(i) := \mathbb{P}_i(\tau_a < \tau_b) = w_{\alpha a} e_i (-S)^{-1} e_\alpha^\top. \tag{4}$$

Proof. We use the definition of conditional probability and the Markov Property to derive

$$\begin{aligned}
&\mathbb{P}_i(X_t = j|\tau_a < \tau_b) = \mathbb{P}_i(X_t = j|\exists s \geq 0 : X_{t+s} = a) \\
&\quad = \frac{\mathbb{P}_i(\exists s \geq 0 : X_{t+s} = a|X_t = j)\mathbb{P}_i(X_t = j)}{\mathbb{P}_i(\exists s \geq 0 : X_{t+s} = a)} \\
&\quad = \frac{\mathbb{P}_j(\exists s \geq 0 : X_s = a)}{\mathbb{P}_i(\exists s' \geq 0 : X_{s'} = a)}\mathbb{P}_i(X_t = j) \qquad \text{(MP)} \\
&\quad = \frac{h_a(j)}{h_a(i)}\mathbb{P}_i(X_t = j)
\end{aligned}$$

We derive further

$$h_a(i) = \mathbb{P}_i(\tau_a < \tau_b) = \int_0^\infty e_i \exp(St) R_{Sa}^\top dt = w_{\alpha a} e_i (-S)^{-1} e_\alpha^\top$$

for any $i \in C \cup \{\alpha, \beta\}$ and recognize that $h_a(a) = 1$ and $h_a(b) = 0$. The inverse of S exists by Lemma 1.

□

Remark 1. Lemma 2 shows that the ratio

$$\frac{h_a(j)}{h_a(i)} = \frac{e_i S^{-1} e_\alpha^\top}{e_j S^{-1} e_\alpha^\top}$$

is independent of $w_{\alpha a}$. This implies that the conditional law given in (3) is independent of $w_{\alpha a}$.

We define the $(|E|-2)\times(|E|-2)$-dimensional diagonal matrix

$$\begin{aligned} H_a &:= \operatorname{diag}(h_a(\alpha), h_a(c_1), \ldots, h_a(c_n), h_a(\beta)) \\ &:= \begin{pmatrix} h_a(\alpha) & 0 & \ldots & 0 & 0 \\ 0 & h_a(c_1) & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & h_a(c_n) & 0 \\ 0 & 0 & \ldots & 0 & h_a(\beta) \end{pmatrix}, \end{aligned}$$

where $c_i \in C$ and $n = |C|$. With this definition equality (3) shows that the transientrix $S_{|a}$ of the infinitesimal generator $Q_{|a}$ of the process conditioned to absorption in a is a simple transformation of the original transientrix S, i.e.

$$S_{|a} = H_a^{-1} S H_a. \tag{5}$$

Note that the transformed model inherits the block structure of Q; we fix the notation as

$$Q_{|a} = \begin{pmatrix} 0 & 0 & 0 \\ R_{|a}^\top & S_{|a} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The vector $R_{|a}$ is uniquely determined by the generator property (row sums are zero) therefore

$$R_{|a}^\top = -S_{|a}\mathbf{1}^\top = -H_a^{-1} S \underbrace{H_a \mathbf{1}^\top}_{=-S^{-1}R_a^\top} = H_a^{-1} R_a^\top$$

Lemma 3. *Suppose $\mathbb{P}(X_0 = i) = 1$ with $i \in C \cup \{\alpha, \beta\}$. Then the distributions of $\tau_a \mathbb{1}_{\{\tau_a < \infty\}}$ and $\tau_b \mathbb{1}_{\{\tau_b < \infty\}}$ are absolutely continuous on $\mathbb{R}^+$ and have the densities*

$$\begin{aligned} f_{\tau_a}^i(t) &= \tfrac{1}{e_i(-S)^{-1}e_\alpha^\top} e_i \exp(St) e_\alpha^\top, \quad t > 0, \\ f_{\tau_b}^i(t) &= \tfrac{1}{e_i(-S)^{-1}e_\beta^\top} e_i \exp(St) e_\beta^\top, \quad t > 0. \end{aligned} \tag{6}$$

The moment generating functions are

$$\begin{aligned} M_{\tau_a}^i(u) &= \tfrac{-1}{e_i(-S)^{-1}e_\alpha^\top} e_i (S + uId)^{-1} e_\alpha^\top, \quad u \le 0, \\ M_{\tau_b}^i(u) &= \tfrac{-1}{e_i(-S)^{-1}e_\beta^\top} e_i (S + uId)^{-1} e_\beta^\top, \quad u \le 0. \end{aligned} \tag{7}$$

Proof. By construction of the model $\mathbb{P}_i(\tau_a < \tau_b) + \mathbb{P}_i(\tau_a > \tau_b) = 1$, therefore

$$\begin{aligned}\mathbb{P}_i(\tau_a \in .) &= \mathbb{P}_i(\tau_a > \tau_b)\delta_{\{+\infty\}} + \mathbb{P}_i(\tau_a < \tau_b) f^i_{\tau_a}(t)dt \\ &= h_b(i)\delta_{\{+\infty\}} + h_a(i) f^i_{\tau_a}(t)dt\end{aligned}$$

where $h_b(i) = 1 - h_a(i)$.

Using (3) we further compute

$$\begin{aligned}\mathbb{P}_i(\tau_a \mathbb{1}_{\{\tau_a<\infty\}} \leq t) &= \int_0^t e_i \exp(S_{|a}\xi) R_{|a}^\top d\xi \\ &= \frac{1}{h_a(i)} \int_0^t e_i \exp(S\xi) \underbrace{H_a H_a^{-1}}_{=Id} \underbrace{R_a^\top}_{=w_{\alpha a} e_\alpha^\top} d\xi \\ &= \frac{1}{e_i(-S)^{-1} e_\alpha^\top} \int_0^t e_i \exp(S\xi) e_\alpha^\top d\xi.\end{aligned}$$

and identify the density of τ_a on $\mathbb{R}^+$ as in (6).

We use

$$\exp(tu) e_i \exp(St) = e_i \exp(utId) \exp(St) = e_i \exp((S + uId)t)$$

to compute the moment generating function

$$\begin{aligned}M^i_{\tau_a}(u) = \mathbb{E}_i(e^{\tau_a \mathbb{1}_{\{\tau_a<\infty\}} u}) &= \int_0^\infty \exp(tu) f^i_{\tau_a}(t)dt \\ &= \frac{1}{e_i(-S)^{-1} e_\alpha^\top} e_i \int_{\mathbb{R}^+} \exp(utId) \exp(St)dt \\ &= \frac{1}{e_i(-S)^{-1} e_\alpha^\top} e_i \left[\int_{\mathbb{R}^+} \exp((S + uId)t)dt\right] e_\alpha^\top \\ &= \frac{1}{e_i(-S)^{-1} e_\alpha^\top} e_i (S + uId)^{-1} (\mathbf{0} - Id) e_\alpha^\top.\end{aligned}$$

By Lemma 1 the matrix $S + uId$ is invertible as it is a transientrix for every $u \leq 0$.

The proof is complete by showing the results for τ_b with the same argumentation as for τ_a.

□

3. Nearly Reversibility and Time Duality

The main goal of this article is the comparison of the hitting time τ_a when the MCc starts in β and the hitting time τ_b when the MCc starts in α, see fig. 2. If the process were reversible TD would clearly hold, since reversibility means that for every path through a collection of nodes the reversed path through the same nodes in reversed order has the same probability of appearance and is taken with the same velocity. But, due to the existence of two absorbing states a and b the process can not be reversible. We therefore introduce the less restrictive notion of "nearly reversibility":

Definition 1. (*nearly reversible.*) The MCc with infinitesimal generator $Q = (q_{ij})_{ij}$ given by (2) is *nearly reversible* if there exists $\pi = (\pi(i))_{i \in E \setminus \{a,b\}}, \pi(i) > 0$, such that for every $i, j \in E \setminus \{a, b\}$ the balance condition

$$\pi(i)q_{ij} = q_{ji}\pi(j) \tag{8}$$

holds.

Remark 2. The measure π in the above Definition 1 is not a priori normalized. Nonetheless π can be constructed as a substochastic vector in the following way. If we add "re-entering" rates $w^+ > 0$ and $w^- > 0$ to the system like in figure 3, each state is recurrent. Therefore there exists a unique (positive) stationary probability measure $\tilde{\pi}$ on E associated to the modified infinitesimal generator

$$\tilde{Q} = \begin{pmatrix} -w^+ & R_a & 0 \\ R_{Sa}^\top & S & R_{Sb}^\top \\ 0 & R_b & -w^- \end{pmatrix},$$

where $R_a = w^+ e_\alpha$ and $R_b = w^- e_\beta$. Since Q and $\tilde{Q}$ are identical up to the first and last row, the original process is nearly reversible as soon as $\tilde{\pi}$ is reversible under $\tilde{Q}$ and then $\pi = \tilde{\pi}_{|E \setminus \{a,b\}}$.

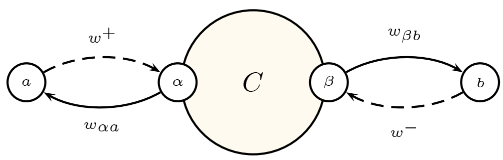


FIGURE 3: The modified model with positive re-entering rates, which ensures that each state is recurrent.

The exact value of $\tilde{\pi}$ does depend on the chosen re-entering rates w^+ and w^-. But, due to the following "generalized Kolmogorov Criterion" the *existence* of a reversible measure under $\tilde{Q}$ does not depend on the re-entering rates. This criterion states that reversibility only depends on products of rates along the cycles of the transition graph of a MCc. A cycle is here to be understood as a finite sequence of distinct states with the only exception that the first and last entry of the sequence are equal. The weight of such a cycle is then the product of the transition rates between successive states of the cycle. The weight can also be zero if any of the transition rates between successive states is zero.

Lemma 4. (generalized Kolmogorov criterion,[4] (1.22).) *An ergodic MCc with infinitesimal generator $\tilde{Q} = (\tilde{q}_{ij})_{ij}$ admits a reversible measure if and only if for all cycles $(i_1, i_2 \ldots, i_{n-1}, i_n, i_1)$, $n \geq 3$,*

$$\tilde{q}_{i_1 i_2} \tilde{q}_{i_2 i_3} \ldots \tilde{q}_{i_{n-1} i_n} \tilde{q}_{i_n i_1} = \tilde{q}_{i_1 i_n} \tilde{q}_{i_n i_{n-1}} \ldots \tilde{q}_{i_3 i_2} \tilde{q}_{i_2 i_1}. \tag{9}$$

If the Kolmogorov-Criterion is fulfilled the reversible measure $\tilde{\pi}$ of $\tilde{Q}$ can easily be constructed as in e.g. [4] p. 21, Theorem 1.7.

Note that cycles of length 2 trivially fulfill this criterion.

The next theorem shows that nearly reversibility is sufficient for TD.

Theorem 1. *If Q is nearly reversible, then TD holds.*

Proof. By nearly reversibility and Remark 2 there exists a strictly positive vector π such that

$$\forall i, j \in C \cup \{\alpha, \beta\} : \pi(i) q_{ij} = \pi(j) q_{ji}.$$

This balance condition can be rewritten in matrix form

$$S^\top = \mathrm{diag}(\pi) S \mathrm{diag}(\pi)^{-1}. \tag{10}$$

A direct comparison of the conditional densities (6) and the use of the similarity transform (10) leads to

$$\begin{aligned} f^{\beta}_{\tau_a}(t) &= \frac{1}{e_\alpha (-S)^{-1} e_\beta^\top} e_\alpha \exp(St) e_\beta^\top = \frac{1}{e_\beta (-S^\top)^{-1} e_\alpha^\top} e_\beta \exp(S^\top t) e_\alpha^\top \\ &= \frac{\pi(\alpha)}{\pi(\beta)} \frac{1}{e_\beta (-S)^{-1} e_\alpha^\top} \frac{\pi(\beta)}{\pi(\alpha)} e_\beta \exp(St) e_\alpha^\top = f^{\alpha}_{\tau_b}(t). \end{aligned}$$

□

Theorem 1 applies to an important class of processes, Birth and Death processes on finite state space with absorbing boundaries.

Corollary 1. *All finite Birth and Death processes on finite state space with non zero birth and death rates and absorbing boundaries satisfy TD.*

Proof. As the only cycles with positive weight are of length 2 every Birth and Death process with absorbing boundaries is nearly reversible and therefore by Theorem 1 TD holds.

□

4. Necessary and Sufficient Condition for Time Duality

In general it is computationally very expensive to calculate the densities of the conditional first hitting times given in (6) as they contain a matrix exponential. Accordingly it is very difficult to check whether TD holds or not for a given model. Thus the idea to rephrase TD via the identity

$$M^{\beta}_{\tau_a}(u) \equiv M^{\alpha}_{\tau_b}(u),\ u \leq 0,$$

of the moment generating functions given in (7), is natural and

$$\frac{e_\alpha S^{-1} e_\beta^\top}{e_\beta S^{-1} e_\alpha^\top} = \frac{e_\alpha (S + uId)^{-1} e_\beta^\top}{e_\beta (S + uId)^{-1} e_\alpha^\top},\ u \leq 0, \tag{11}$$

must hold.

In Theorem 2 we refine this condition to show that TD does not only not depend on the rates $w_{\alpha a}$ and $w_{\beta b}$ (see Remark 1) but also does not depend on the holding times in α and β. The refined condition also enables further exploration of possible conditions that allow TD, see next section.

We now need a further partition of the transientrix S introduced in (2) into blocks which describe the transition behavior between the "boundary states" α, β and the states in C. We therefore write

$$S := \begin{pmatrix} \mathfrak{a} & R_{\alpha\hat{S}} & w_{\alpha\beta} \\ R^\top_{\hat{S}\alpha} & \hat{S} & R^\top_{\hat{S}\beta} \\ w_{\beta\alpha} & R_{\beta\hat{S}} & \mathfrak{b} \end{pmatrix}.$$

with $-\mathfrak{a}, -\mathfrak{b}, w_{\alpha\beta}, w_{\beta\alpha} \in \mathbb{R}^+$ and

$$\mathfrak{a} := -w_{\alpha a} - R_{\alpha\hat{S}}\mathbf{1}^\top - w_{\alpha\beta}, \qquad \mathfrak{b} := -w_{\beta b} - R_{\beta\hat{S}}\mathbf{1}^\top - w_{\beta\alpha}$$

where $\mathbf{1} := (1, 1, \ldots, 1)$. We denote with w_{ij} the transition rate from i to j for $i, j \in E$ and the vectors $R_{\alpha\hat{S}}, R_{\beta\hat{S}}, R_{\hat{S}\beta}, R_{\hat{S}\alpha}$ describe the transitions between $\{\alpha, \beta\}$ and C, while $\hat{S}$ describes the transitions between the states in C.

The infinitesimal generator Q now writes

$$Q := \begin{array}{cc} & \begin{array}{ccccc} a & \alpha & C & \beta & b \end{array} \\ \begin{array}{c} a \\ \alpha \\ C \\ \beta \\ b \end{array} & \left(\begin{array}{c|ccc|c} 0 & 0 & 0 & 0 & 0 \\ \hline w_{\alpha a} & \mathfrak{a} & R_{\alpha\hat{S}} & w_{\alpha\beta} & 0 \\ 0 & R^\top_{\hat{S}\alpha} & \hat{S} & R^\top_{\hat{S}\beta} & 0 \\ 0 & w_{\beta\alpha} & R_{\beta\hat{S}} & \mathfrak{b} & w_{\beta b} \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right) \end{array} \tag{12}$$

Theorem 2. *TD is equivalent to the identity*

$$\forall u \leq 0 : \frac{R_{\alpha\hat{S}}(\hat{S} + uId)^{-1}R^\top_{\hat{S}\beta} - w_{\alpha\beta}}{R_{\beta\hat{S}}(\hat{S} + uId)^{-1}R^\top_{\hat{S}\alpha} - w_{\beta\alpha}} = c. \tag{13}$$

where c is a constant independent of u.

Proof. We already argued that TD is equivalent to identity (11).

We proceed by unveiling the structure of $(S + uId)^{-1}$ in terms of block matrices using the block partition given in (12). Define the following submatrix of $S + uId$:

$$\bar{S}_u := \begin{pmatrix} \hat{S} + uId & R^\top_{\hat{S}\beta} \\ R_{\beta\hat{S}} & \mathfrak{b} + u \end{pmatrix}.$$

The inverse of $\bar{S}_u$ in terms of block matrices can be derived according to [2], p. 73, (86)-(89) via the identities

$$\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}^{-1} = \begin{pmatrix} M_{11}^{-1} + M_{11}^{-1}M_{12}H^{-1}M_{21}M_{11}^{-1} & -M_{11}^{-1}M_{12}H^{-1} \\ -H^{-1}M_{21}M_{11}^{-1} & H^{-1} \end{pmatrix} \tag{14}$$

if M_{11} is invertible resp.

$$\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}^{-1} = \begin{pmatrix} K^{-1} & -K^{-1}M_{12}M_{22}^{-1} \\ -M_{22}^{-1}M_{21}K^{-1} & M_{22}^{-1} + M_{22}^{-1}M_{21}K^{-1}M_{12}M_{22}^{-1} \end{pmatrix} \tag{15}$$

if M_{22} is invertible where

$$H := M_{22} - M_{21}M_{11}^{-1}M_{12}$$

and

$$K := M_{11} - M_{12}M_{22}^{-1}M_{21}.$$

Therefore due to (14)

$$\bar{S}_u^{-1} =$$

$$H^{-1} \begin{pmatrix} H(\hat{S}+uId)^{-1} + (\hat{S}+uId)^{-1}R_{\hat{S}\beta}^\top R_{\beta\hat{S}}(\hat{S}+uId)^{-1} & -(\hat{S}+uId)^{-1}R_{\hat{S}\beta}^\top \\ -R_{\beta\hat{S}}(\hat{S}+uId)^{-1} & 1 \end{pmatrix}$$

with

$$H = \mathfrak{b} + u - R_{\beta\hat{S}}(\hat{S}+uId)^{-1}R_{\hat{S}\beta}^\top$$

Now with (15) we get

$$(S+uId)^{-1} = K^{-1} \begin{pmatrix} 1 & -(R_{\alpha\hat{S}}, w_{\alpha\beta})\bar{S}_u^{-1} \\ -\bar{S}_u^{-1}\begin{pmatrix} R_{\hat{S}\alpha}^\top \\ w_{\beta\alpha} \end{pmatrix} & * \end{pmatrix},$$

where

$$K = \mathfrak{a} + u - (R_{\alpha\hat{S}}, w_{\alpha\beta})\bar{S}_u^{-1}\begin{pmatrix} R_{\hat{S}\alpha}^\top \\ w_{\beta\alpha} \end{pmatrix}.$$

We do not compute the expression for the submatrix $*$ as it is not needed for the completion of the proof. We conclude

$$e_\alpha(S+uId)^{-1}e_\beta^\top = K^{-1}H^{-1}(R_{\alpha\hat{S}}(\hat{S}+uId)^{-1}R_{\hat{S}\beta}^\top - w_{\alpha\beta}),$$
$$e_\beta(S+uId)^{-1}e_\alpha^\top = K^{-1}H^{-1}(R_{\beta\hat{S}}(\hat{S}+uId)^{-1}R_{\hat{S}\alpha}^\top - w_{\beta\alpha}).$$

K and H are real numbers and cancel out if the expressions are inserted into (11), they are the only quantities which contain $w_{\alpha a}$ and $w_{\beta b}$. Moreover if

$$\frac{R_{\alpha\hat{S}}(\hat{S}+uId)^{-1}R_{\hat{S}\beta}^\top - w_{\alpha\beta}}{R_{\beta\hat{S}}(\hat{S}+uId)^{-1}R_{\hat{S}\alpha}^\top - w_{\beta\alpha}} = c$$

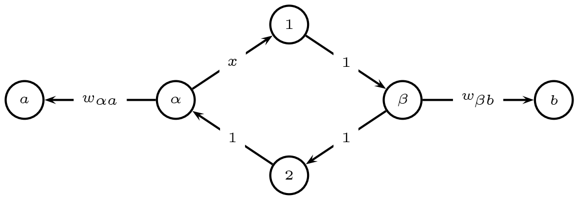


FIGURE 4: Example for a model with TD for *any* choice of the parameter $x > 0$.

for all $u \leq 0$ then obviously

$$\frac{R_{\alpha\hat{S}}\hat{S}^{-1}R^{\top}_{\hat{S}\beta} - w_{\alpha\beta}}{R_{\beta\hat{S}}\hat{S}^{-1}R^{\top}_{\hat{S}\alpha} - w_{\beta\alpha}} = c.$$

□

With the criterion (13) at hand we now can show

Proposition 4.1. *Nearly reversibility is not necessary for TD.*

Proof. We construct a simple MCc without nearly reversibility but with TD. Suppose the infinitesimal generator Q is given by

$$Q := \begin{array}{c} \\ a \\ \alpha \\ 1 \\ 2 \\ \beta \\ b \end{array} \begin{array}{c} \begin{array}{cccccc} a & \alpha & 1 & 2 & \beta & b \end{array} \\ \left(\begin{array}{c|ccccc|c} 0 & 0 & 0 & 0 & 0 & 0 \\ \hline w_{\alpha a} & -w_{\alpha a} - x & x & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 - w_{\beta b} & w_{\beta b} \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \end{array},$$

with real $x > 0$ arbitrary, see also fig. 4 for the transition graph.

According to (9) this model is clearly not nearly reversible, as

$$0 = q_{1\alpha}q_{\alpha 2}q_{2\beta}q_{\beta 1} \neq q_{1\beta}q_{\beta 2}q_{2\alpha}q_{\alpha 1} = x.$$

But TDis fullfilled according to (13):

$$R_{\alpha\hat{S}} = (x, 0), \qquad R^{\top}_{\hat{S}\alpha} = (0, 1), \qquad -\hat{S} = Id,$$
$$R_{\beta\hat{S}} = (0, 1), \qquad R^{\top}_{\hat{S}\beta} = (1, 0), \qquad w_{\alpha\beta} = w_{\beta\alpha} = 0$$

and obviously

$$\forall u \leq 0, \; \frac{R_{\alpha\hat{S}}(-Id+uId)^{-1}R^{\top}_{\hat{S}\beta}}{R_{\beta\hat{S}}(-Id+uId)^{-1}R^{\top}_{\hat{S}\alpha}} = x$$

□

5. Invariance under similarity transform

Nearly reversibility is, like reversibility, shared by only few models: Comparing a path only with its reversed is very restrictive. Note, for example, that if for every trajectory ω_1 from α to b there exists another trajectory ω_2 from β to a, possibly going through other states, with the same transition rates, then TD holds. In the following we give sufficient conditions on the transientrix under which TD holds.

Definition 2. (*invariance under similarity transform.*) We say Q given by (12) is *invariant under similarity transform* if there exists an invertible matrix D such that, for some real numbers $d_1, d_2 > 0$, either

(Type I)
$$\hat{S} = D\hat{S}D^{-1}, \qquad R_{\beta\hat{S}} = d_1 R_{\alpha\hat{S}} D,$$
$$w_{\beta\alpha} = d_1 d_2 w_{\alpha\beta}, \qquad R^{\top}_{\hat{S}\alpha} = d_2 D^{-1} R^{\top}_{\hat{S}\beta},$$

or

(Type II)
$$\hat{S}^{\top} = D\hat{S}D^{-1}, \qquad R_{\beta\hat{S}} = d_1 R_{\hat{S}\beta} D^{\top},$$
$$w_{\beta\alpha} = d_1 d_2 w_{\alpha\beta}, \qquad R^{\top}_{\hat{S}\alpha} = d_2 (D^{-1})^{\top} R^{\top}_{\alpha\hat{S}},$$

holds.

Theorem 3. *Invariance under similarity transform implies TD.*

Proof. Using Theorem 2 and the definition of invariance under similarity transform we obtain:

Type I

$$\frac{R_{\alpha\hat{S}}(\hat{S}+uId)^{-1}R^{\top}_{\hat{S}\beta} - w_{\alpha\beta}}{R_{\beta\hat{S}}(\hat{S}+uId)^{-1}R^{\top}_{\hat{S}\alpha} - w_{\beta\alpha}} = \frac{1}{d_1 d_2} \frac{d_1 d_2 R_{\alpha\hat{S}} D(\hat{S}+uId)^{-1}D^{-1}R^{\top}_{\hat{S}\beta} - d_1 d_2 w_{\alpha\beta}}{R_{\beta\hat{S}}(\hat{S}+uId)^{-1}R^{\top}_{\hat{S}\alpha} - w_{\beta\alpha}}$$
$$= \frac{1}{d_1 d_2}$$

Type II

$$
\begin{aligned}
\frac{R_{\alpha\hat{S}}(\hat{S}+uId)^{-1}R^{\top}_{\hat{S}\beta}-w_{\alpha\beta}}{R_{\beta\hat{S}}(\hat{S}+uId)^{-1}R^{\top}_{\hat{S}\alpha}-w_{\beta\alpha}} &= \frac{1}{d_1d_2}\frac{d_1d_2R_{\alpha\hat{S}}D^{-1}(\hat{S}^{\top}+uId)^{-1}DR^{\top}_{\hat{S}\beta}-d_1d_2w_{\alpha\beta}}{R_{\beta\hat{S}}(\hat{S}+uId)^{-1}R^{\top}_{\hat{S}\alpha}-w_{\beta\alpha}} \\
&= \frac{1}{d_1d_2}\frac{d_1d_2R_{\hat{S}\beta}D^{\top}(\hat{S}+uId)^{-1}(D^{-1})^{\top}R^{\top}_{\alpha\hat{S}}-w_{\beta\alpha}}{R_{\beta\hat{S}}(\hat{S}+uId)^{-1}R^{\top}_{\hat{S}\alpha}-w_{\beta\alpha}} \\
&= \frac{1}{d_1d_2}
\end{aligned}
$$

□

Example 1. Nearly reversible models are invariant under similarity transform of type II: By (10) and (12) there exists a positive vector $\pi := (\pi(\alpha), \pi(C), \pi(\beta))$ such that

$$
\begin{pmatrix} * & R_{\hat{S}\alpha} & w_{\beta\alpha} \\ R^{\top}_{\alpha\hat{S}} & \hat{S}^{\top} & R^{\top}_{\beta\hat{S}} \\ w_{\alpha\beta} & R_{\hat{S}\beta} & * \end{pmatrix} = \operatorname{diag}(\pi(\alpha), \pi(C), \pi(\beta)) \begin{pmatrix} * & R_{\alpha\hat{S}} & w_{\alpha\beta} \\ R^{\top}_{\hat{S}\alpha} & \hat{S} & R^{\top}_{\hat{S}\beta} \\ w_{\beta\alpha} & R_{\beta\hat{S}} & * \end{pmatrix} \operatorname{diag}(\pi(\alpha), \pi(C), \pi(\beta))^{-1}.
$$

Therefore

$$
\begin{aligned}
\hat{S}^{\top} &= \operatorname{diag}(\pi(C))\hat{S}\operatorname{diag}(\pi(C))^{-1}, & w_{\beta\alpha} &= \frac{\pi(\alpha)}{\pi(\beta)}w_{\alpha\beta}, \\
R_{\beta\hat{S}} &= \frac{1}{\pi(\beta)}R_{\hat{S}\beta}\operatorname{diag}(\pi(C))^{\top}, & R^{\top}_{\hat{S}\alpha} &= \pi(\alpha)(\operatorname{diag}(\pi(C))^{-1})^{\top}R^{\top}_{\alpha\hat{S}}.
\end{aligned}
$$

Example 2. Models with nearly symmetric transientrix S are invariant under similarity transform of type II:

$$
\begin{aligned}
\hat{S} &= \hat{S}^{\top}, & R_{\beta\hat{S}} &= d_1R_{\hat{S}\beta}, \\
w_{\beta\alpha} &= d_1d_2w_{\alpha\beta}, & R^{\top}_{\hat{S}\alpha} &= d_2R^{\top}_{\alpha\hat{S}}.
\end{aligned}
$$

This is a relaxed form of symmetric transientrix since for $d_1 = d_2 = 1$, $S = S^{\top}$.

Example 3. Invariance of Type I is satisfied by a class of interest, namely models for which $\hat{S}$ is invariant under a pairwise exchange of rows and columns (in this case, D is a very special orthogonal matrix, a permutation matrix).

The example in fig. 5 is invariant under the transform $D = Id$ — due to the mirror symmetry — and also under

$$
D' = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},
$$

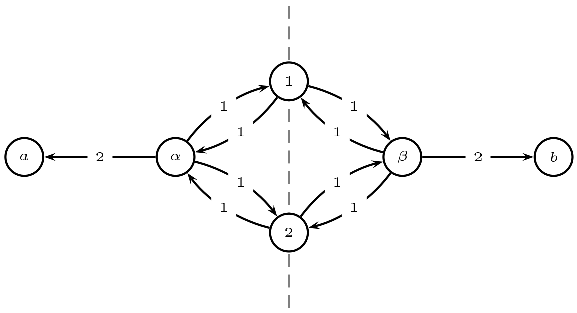


FIGURE 5: A model which is invariant under orthogonal transform. The model is mirror symmetric with respect to the gray dashed line and also rotation symmetric.

for $\hat{S} = -2Id$ and $R_{\alpha\hat{S}} = R_{\hat{S}\alpha} = R_{\hat{S}\beta} = R_{\beta\hat{S}} = (1,1)$, $w_{\alpha\beta} = w_{\beta\alpha} = 0$ — due to the rotation symmetry. Note that this model is nearly reversible too. The model given in Proposition 4.1 is another example with rotation symmetry and also invariant under D'; But in this case nearly reversibility does not hold.

6. Acknowledgment

The authors would like to thank R. Lipowsky for fruitful discussions.

References

[1] DIACONIS, P. AND MICLO, L. (2009). On times to quasi-stationarity for birth and death processes. *Journal of Theoretical Probability* **22,** 558–586.

[2] GANTMACHER, F. R. (1986). *Matrizentheorie.* Springer-Verlag.

[3] ITO, H., AMARI, S.-I. AND KOBAYASHI, K. (1992). Identifiability of hidden markov information sources and their minimum degrees of freedom. *IEEE Transaction on Information Theory* **38,** 324–333.

[4] KELLY, F. P. (1979). *Reversibility and Stochastic Networks.* Wiley.

[5] KEMENY, J. G., SNELL, J. L. AND KNAPP, A. W. (1966). *Denumerable Markov Chains.* The University Series in Higher Mathematics. D. van Nostrand Company, Inc.

[6] Kijima, M. (1988). On passage and conditional passage times for markov chains in continuous time. *Journal of Applied Probability* **25,** 279–290.

[7] Lindén, M. and Wallin, M. (2007). Dwell time symmetry in random walks and molecular motors. *Biophysical Journal* **92,** 3804–3816.

[8] Lipowsky, R., Liepelt, S. and Valleriani, A. (2009). Energy conversion by molecular motors coupled to nucleotide hydrolysis. *Journal of Statistical Physics* **135,** 951–975.

[9] Neuts, M. F. (1994). *Matrix Geometric Solutions in Stochastic Models.* Dover Publications, Inc.

[10] Swart, J. M. (2010). Intertwining of birth and death processes. *ArXiV.org*. arXiv:1004.5515v1[math.PR].

[11] Valleriani, A., Liepelt, S. and Lipowsky, R. (2008). Dwell time distributions for kinesin's mechanical steps. *EPL (Europhysics Letters)* **82,** 28011–p1–28011–p6.

[12] Wang, H. and Qian, H. (2007). On detailed balance and reversibility of semi-markov processes and single molecule enzyme kinetics. *Journal of Mathematical Physics* **48,** 013303–1 – 013303–15.