# On Efficiency of Markovian Couplings

Mason DiCicco†, joint with Ewan Harlow (U. Wisconsin) and Micheal Dotzel (U. Missouri)
Project Advisor: Iddo Ben-Ari, Mathematics

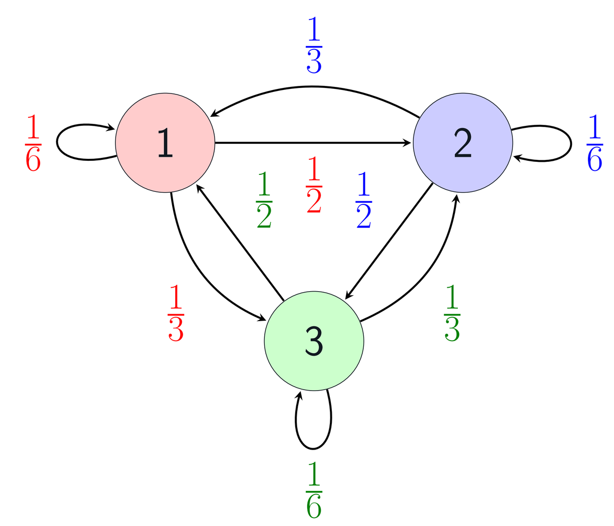
† Support from: UConn Trimble Family SURF award, NSA Markov Chains REU 2018

UCONN
COLLEGE OF LIBERAL ARTS AND SCIENCES

## Markov Chains (MCs)

A Random process $\mathbf{X} = (X_t)_{t\in\mathbb{Z}_+}$ taking values in a finite set $\mathcal{S}$ with the

$$\begin{aligned}\mathbb{P}(X_{t+1} = j | X_0 = i_0, \dots, X_t = i_t) &= \mathbb{P}(X_1 = j | X_0 = i_t)\\ &:= p(i_t, j) \quad \leftarrow \quad \textbf{transition function}\end{aligned} \tag{1}$$




## Total Variation Distance

**Definition 1.** *Let $\mu_1$ and $\mu_2$ be two probability distributions on $\mathcal{S}$. The total variation distance between the distribution of $X_t$ under $X_0 \sim \mu_1$ and the distribution of $X_t$ under $X_0 \sim \mu_2$, is defined as*

$$d_t(\mu_1, \mu_2) = \max_{A\subseteq\mathcal{S}} \left(\mathbb{P}_{\mu_1}(X_t \in A) - \mathbb{P}_{\mu_2}(X_t \in A)\right).$$

If, say, $\mu_1$ is $\delta_i$ for some state $i$, we write $d_t(i, \mu_2)$ for $d_t(\delta_i, \mu_2)$, etc.

## Ergodic Theorem for MCs

A transition function $p$ is

- **Irreducible** if for any two states $i, j$, there exists a $t > 0$ such that $p^t(i, j) > 0$. This means that it is possible to get from any state to any other state using only transitions of positive probability
- **Aperiodic** if for every state $i$, $\gcd\{t : p^t(i, i) > 0\} = 1$.

**Theorem 1** (Ergodic Theorem for MCs)**.**

1. *If $p$ is irreducible, then it possess a unique* **stationary distribution** *$\pi$: $\mathbb{P}_\pi(X_t = j) = \pi(j)$ for all $n$, or, equivalently, $\pi p = \pi$.*
2. *If $p$ is irreducible and aperiodic, there exists a constant $\rho \in [0, 1)$ such that*
$$\max_i d_t(i, \pi) \le \max_{i,j} d_t(i, j) \asymp \rho^n. \tag{2}$$

Note.

- From linear algebra, the smallest $\rho$ satisfying (2) is the norm of the subdominant eigenvalue for $p$.

## The Method of Coupling

**Definition 2.** *Let $p$ be a transition function on a state space $\mathcal{S}$. An $\mathcal{S}\times\mathcal{S}$-valued process $(\mathbf{X}, \mathbf{Y})$ is called a* **coupling for** *$p$ if each marginal process $\mathbf{X} = (X_t)_{t\in\mathbb{Z}_+}$ and $\mathbf{Y} = (Y_t)_{t\in\mathbb{Z}_+}$ is a MC with transition function $p$.*

Suppose $(\mathbf{X}, \mathbf{Y})$ is a coupling. We define the meeting time $\tau$:

$$\tau = \inf\{t \in \mathbb{Z}_+ : X_t = Y_t\}.$$

**Lemma 1** (Aldous' inequality)**.** *Suppose $(\mathbf{X}, \mathbf{Y})$ is a coupling with $(X_0, Y_0) = (i, j)$. Then*

$$d_t(i, j) \le \mathbb{P}(\tau > t). \tag{3}$$

*Thus, coupling is a probabilistic technique for obtaining upper bounds on $d_t(i, j)$ through tail distribution of certain random variables.*

- A "trivial" coupling is independent coupling: the two copies are independent. In this case, the RHS of (3) gives very crude bounds.

A coupling as the Lemma is called **efficient** if

$$d_t(i, j) \asymp \mathbb{P}(\tau > t).$$

**Question.** How close to equality in (3) a coupling can be?

- Griffeath (and later others) proved the existence of a coupling attaining an equality in (3) (AKA a maximal coupling, stronger than efficient). The construction is complex and too hard to use in applications.

### Our work

We restricted the discussion to a special class of analytically tractable couplings in the special case of three-state MCs with symmetric transition functions.

## Greedy Couplings

**Definition 3.** *We call a coupling $(\mathbf{X}, \mathbf{Y})$* **greedy Markovian** *if it satisfies the following two properties:*

- **(Markovian)** *It is a MC: $(X_{t+1}, Y_{t+1})$ is a deterministic function of $(X_t, Y_t)$ and a Uniform-$[0, 1]$ random variable, independent of $(X_0, Y_0), \dots, (X_t, Y_t)$.*
- **(greedy)** *The probability that $Y_{t+1} = X_{t+1}$, conditioned on $(X_t, Y_t) = (x, y)$ is maximized: it is equal to*
$$\sum_j \min\{\mathbb{P}_x(X_1 = j), \mathbb{P}_y(X_1 = j)\} = 1 - d_1(x, y).$$

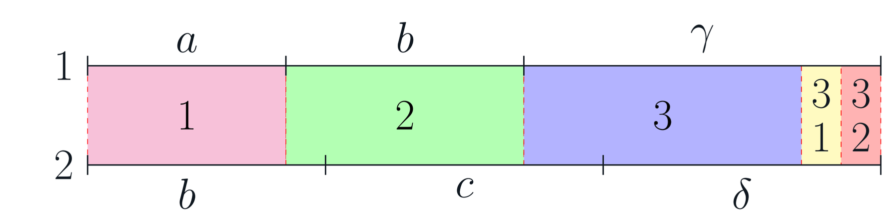
## Our results

Consider a *symmetric* TF of the form

$$p = \begin{bmatrix} a & b & \gamma \\ b & c & \delta \\ \gamma & \delta & \epsilon \end{bmatrix}$$

We analyze all greedy Markovian couplings for $p$.

The diagram below represents the coupling from the state $(1, 2)$ for the coupled process.




1. The diagram corresponds to the particular case $a < b$ and $b < c$.
2. The area of the diagram is 1, representing the probability of all possible configurations for the coupled process in the next step.
3. The three regions on the left have respective areas $a, b$ and $\delta$ representing meeting in the next step at 1,2, and 3 respectively.
4. The area of the remaining two regions represent the probability of not meeting in the next step. This is split into two regions, because
   (a) The leftover probability from the transition $2 \to 1$ is $b - a$.
   (b) The leftover probability from the transition $2 \to 2$ is $c - b$.
   (c) Leftover probability from transitions from 1 is only from the transition $1 \to 3$, and is equal to $\gamma - \delta = (b - a) + (c - b)$.

Summarizing:

$$\begin{tikzcd} \text{(meet)} & (1,2) \arrow[l, "1+a-c"'] \arrow[r, "c-b"] \arrow[dr, "b-a"'] & (3,2) \\ & & (3,1) \end{tikzcd}$$

- Item 4(c) captures a feature which is key to the analysis: if no meeting occurs, one of the chains necessarily transitions to some state determined by structure of the TF.
- Through reduction by symmetry and exhaustive examination of the remaining cases, the following result was obtained:

**Theorem 2.** *Let $p$ be a symmetric $3 \times 3$ irreducible transition function. Let*

$$M_{i,j} = \{k : p_{i,k} > p_{j,k}\}.$$

*Then a greedy Markovian coupling is efficient if and only if*

$$M_{1,2} = M_{2,3} = M_{3,1}.$$