L. Fribourg, S. Messika and C. Picaronny

Traces of Randomized Distributed Algorithms As Markov Fields. Application to Rapid Mixing

Research Report LSV–03–10, July 2003

Laboratoire Spécification et Vérification

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Ecole Normale Supérieure de Cachan
61, avenue du Président Wilson
94235 Cachan Cedex France

http://www.lsv.ens-cachan.fr/Publis/
Research Report LSV-03-10, Lab. Spécification et Vérification, CNRS & ENS de Cachan, France, July 2003

Traces of Randomized Distributed Algorithms As Markov Fields – Application to Rapid Mixing

Laurent Fribourg, Stéphane Messika and Claudine Picaronny

LSV, CNRS & ENS de Cachan,
61 av. du Prés. Wilson
94235 Cachan cedex, France
e-mail: {fribourg,messika,picaro}@lsv.ens-cachan.fr,
fax: 33.1.47.40.75.21

Abstract. Given a randomized distributed algorithm $\mathcal{A}$ that operates on a ring (or linear array) of N processes, we use the fact that the traces of computation via $\mathcal{A}$ form a Markov field (i.e., a random field governed by local interactions), and that $\mathcal{A}$ has a unique stationary distribution if the Markov field has no phase transition. We then use van den Berg's sufficient condition on "paths of disagreement" for guaranteeing the absence of phase transition of Markov fields, as an original tool for proving the uniqueness of the stationary distribution π and the rapid mixing of $\mathcal{A}$ towards π. This new method of proving rapid mixing is illustrated on Herman's self-stabilizing algorithm of mutual exclusion and the classical random walk on a ring.

1 Introduction

In this paper we consider randomized distributed algorithms that operate on a ring (circular array) of N finite state processes $M_1, \cdots, M_N$, where N is a parameter. Each process M_i is a finite-state machine with set of states Q. The configuration of the system is given by the string $q_1 \cdots q_N$ of the states of the processes. We consider here synchronous executions: at each step, all the processes execute synchronously a randomized action, and change their state accordingly. A transition then leads the system from a configuration of the form $q_1 \cdots q_N$ to $q'_1 \cdots q'_N$. An infinite computation trace via a given distributed algorithm $\mathcal{A}$ is represented as an infinite grid of the form:

$$\begin{pmatrix} q_1 & q_2 & \cdots & q_N \\ q'_1 & q'_2 & \cdots & q'_N \\ q''_1 & q''_2 & \cdots & q''_N \\ & & \cdots & \end{pmatrix}$$

where the value of the (i,j)-th coordinate (or "site") of the grid is the state of the j-th machine after step i (with the convention that the indexing of i starts at 0). If we focus on the set of computation traces via $\mathcal{A}$ which start from the same

initial configuration, we get an infinite random field, with a Q-valued random variable for each site.

Such a random field is a *Markov field*, i.e., a field governed by local interactions: for each site s, the conditional distribution of the value of s, given the values of all the other sites, depends only on the values of the sites adjacent to s.

We then use one criterion of [11] (called here "van den Berg's criterion") for ensuring the absence of "phase transition" of the Markov field, which corresponds here to the uniqueness of the stationary distribution π of $\mathcal{A}$. Besides, van den Berg's criterion ensures in our context the rapid mixing of $\mathcal{A}$ to π.

Comparison with related recent work on application of van den Berg's method to rapid mixing.

van den Berg's criterion involves the notion of "disagreement path" [11]. Given a Markov field, and two images X and Y of this random field, a disagreement occurs at site s between X and Y if $X(s) \neq Y(s)$. A path of disagreement $s_1 \sim s_2 \sim \cdots \sim s_k$ between X and Y is a sequence of distinct sites which are adjacent for the Markov field and all disagree ($X(s_i) \neq Y(s_i)$ for all $1 \leq i \leq k$). van den Berg's criterion says that, if the probability of a disagreement path of length m decreases exponentially with m, then there is no phase transition for the Markov field. van den Berg used the idea for showing the absence of phase transition in the monomer-dimer model [12]. He also used this spatial information (exponential decay of correlations between sites) to infer a temporal information: the rapid mixing of the Glauber dynamics for random sampling the monomer-dimer model [13]. Dyer and Frieze on the one hand, Molloy on the other hand also used the spatial notion path disagreement to infer the rapid mixing of the Glauber dynamics, but for random sampling in the graph colouring problem [3,10]. Recently, Dyer et al. used the notion of disagreement path for proving in a combinatorial way the relationship between spatial information (exponential decay of correlations between sites of a give graph) and temporal information (rapid mixing of the Glauber dynamics for random sampling an appropriate distribution of configurations of this graph) [4]. In our context, the relationship between spatial and temporal information is straightforward because of the type of the graph we study: it is a grid representing the trace of execution of the randomized algorithm. Also, the dynamics under study is not Glauber dynamics but the fully parallel dynamics. Our problem is also simplified by the fact that the randomized algorithm concerns only linear topology, and by the fact that a site is updated according to its current state and the state of the *left* neighbour. This allows us to focus on *S-SE* disagreement paths, that are path that move downward in a vertical ("South") or oblique ("South-East") manner.

Plan of the paper. After some preliminaries on distributed randomized algorithms (Sec. 2), we regard the traces of computation via $\mathcal{A}$ as a Markov field (Sec. 3). We then use van den Berg's condition for showing the absence of phase transition in Markov fields, as a new tool for proving the rapid mixing of

randomized algorithms (Sec. 4). This is illustrated on Herman's algorithm and a classical random walk on a ring (Sec. 5). We conclude in Sec. 6.

2 Distributed Algorithms As Markov Chains

We focus in this paper on the simple topology of linear networks, i.e. rings of finite-state machines. We assume that the communication between machines is done through the reading of neighbors' state. A *randomized distributed algorithm* is a triple $\mathcal{A} = (N, \xrightarrow{\mathcal{A}}, Q)$ where N is the number of machines of the system, $\xrightarrow{\mathcal{A}}$ is a state transition algorithm, and Q is a finite set of states taken by the machines (see, e.g., [9]). For the sake of simplicity, we will suppose henceforth that Q contains only two elements 0 and 1. The expression $\neg q$ stands for 1 (resp. 0) when q is 0 (resp. 1). Calculations on indices i of machines are done modulo N (more precisely: $i - 1 = N$ if $i = 1$). Each machine M_i can observe the state q_{i-1} of its **left** neighbor M_{i-1}. (In particular machine M_1 can read the state of M_N.) The algorithm is *uniform* in the sense that $\xrightarrow{\mathcal{A}}$ and Q are common to every machine. A *configuration* of $\mathcal{A}$ is an N-tuple of machine states; if the current state of machine M_i is $q_i \in Q$, then the configuration of the system is an element $\overline{q} = (q_1, \cdots, q_N)$ of Q^N. The state transition algorithm $\xrightarrow{\mathcal{A}}$ is given as a set of $|Q|^2$ *guarded commands* of the form:

`IF` $x_{i-1} = \alpha_k \wedge x_i = \beta_k$ `THEN` $x'_i = \gamma_k$ (resp. $\neg\gamma_k$)
with probability p (resp. $1 - p$).

Hereabove (α_k, β_k) ranges over the values of $Q \times Q$. For all configuration $(q_1, \cdots, q_N)$ and all position j $(1 \leq j \leq N)$, there is a unique "enabled" command, i.e. a command of index, say k, such that $q_{j-1} = \alpha_k$ and $q_j = \beta_k$ hold. The execution of an enabled command modifies the state q_j of machine M_j, which becomes γ_k (resp. $\neg\gamma_k$) with probability p (resp. $1 - p$).

We consider hereafter the case where the algorithm is executed *synchronously*,[1] where all the machines of the ring are processed simultaneously at each step: a transition thus leads from $\overline{q} = (q_1, \cdots, q_N)$ to $\overline{q}' = (q'_1, \cdots, q'_N)$, where q'_j is the result of executing the command enabled at position j $(1 \leq j \leq N)$. Such a transition is written $\overline{q} \xrightarrow{\mathcal{A}} \overline{q}'$. The probability associated to this transition is written $p(\overline{q} \xrightarrow{\mathcal{A}} \overline{q}')$. We have: $p(\overline{q} \xrightarrow{\mathcal{A}} \overline{q}') = \prod_{j=1}^{N} p_j$, where p_j is the probability associated with the command executed at position j. We write $\overline{q} \xrightarrow{\mathcal{A}}{}^{*} \overline{q}'$ if there is a finite sequence of transitions from $\overline{q}$ to $\overline{q}'$.

Example 1. Herman's mutual exclusion algorithm [7], is defined by:

`IF` $x_{i-1} = 0 \wedge x_i = 1$ `THEN` $x'_i = 0$ (resp. 1)	with $p = 1$ (resp. 0)
`IF` $x_{i-1} = 1 \wedge x_i = 0$ `THEN` $x'_i = 1$ (resp. 0)	with $p = 1$ (resp. 0)
`IF` $x_{i-1} = 0 \wedge x_i = 0$ `THEN` $x'_i = 0$ (resp. 1)	with $p = \frac{1}{2}$ (resp. $\frac{1}{2}$)

[1] We could also have chosen a "roundabout" execution where one machine by step is processed in increasing order: M_1, then $M_2, \cdots$.

IF $x_{i-1} = 1 \wedge x_i = 1$ THEN $x'_i = 1$ (resp. 0) with $p = \frac{1}{2}$ (resp. $\frac{1}{2}$).
(or, in a more compact way,:
IF $x_{i-1} \neq x_i$ THEN $x'_i = \neg x_i$
IF $x_{i-1} = x_i$ THEN $x'_i = x_i$ (resp. $\neg x_i$) with $p = \frac{1}{2}$.)
The command is deterministic when $x_{i-1} \neq x_i$, and probabilistic otherwise. Consider for example the configuration $\overline{q} = 10\mathbf{000}$. Letters in bold indicate that they are subjet to probabilistic transitions. We have $\overline{q} \xrightarrow{\mathcal{A}} 01abc$ with probability $(1/2)^3$, for any $(a, b, c) \in Q^3$.

Algorithm $\mathcal{A}$ can be seen as a Markov chain $\{x^t\}_{t\geq 0}$ whose state space is Q^N. The value of x^t is the configuration of the ring after t steps. The i-th component of x^t, denoted by $x^t(i)$, is the state of the i-th machine after t steps. The transition probability, is such that:

$Pr(x^{t+1} = \overline{q}' \mid x^t = \overline{q}) \;=\; p(\overline{q} \xrightarrow{\mathcal{A}} \overline{q}')$.

Let L be a finite set and let X_1 and X_2 be two L-valued random variables with distribution ρ_1 and ρ_2 respectively. The *variational distance* of X_1 and X_2 (or, equivalently, of ρ_1 and ρ_2) is defined by

$d(\rho_1, \rho_2) = 1/2 \sum_{a\in L} |\rho_1(a) - \rho_2(a)|$.

Let $Pr(x^u = . \mid x^0 = T_0)$ be the distribution of the Markov chain $\{x^t\}_{t\geq 0}$ at time u, given that it starts in configuration $x^0 = T_0$. Recall some basic notions (see, e.g., [8]):

Definition 2. Given a Markov chain $\{x^t\}_{t\geq 0}$ (associated with $\mathcal{A}$) of unique stationary distribution π and an initial configuration T_0, the *mixing time* of $\{x^t\}_{t\geq 0}$ from T_0 is the quantity

$\tau_{T_0}(\varepsilon) = min\{t : \; d(Pr(x^u = . \mid x^0 = T_0),\; \pi) \leq \varepsilon \;\text{ for all } u \geq t\}$.

We say that $\mathcal{A}$ is *rapidly mixing* if, for any $T_0 \in Q^N$, $\tau_{T_0}(\varepsilon)$ is bounded by some polynomial in N and $ln(\varepsilon^{-1})$.

In the following, we give a sufficient condition that ensures that the stationary distribution π is unique, and that $\mathcal{A}$ is rapidly mixing, using the notion of "locally stable" coupling.

Definition 3. A *coupling* of $\mathcal{A}$, is a pair process $\{(x^t, y^t)\}_{t\geq 0}$ such that:

(i) Both x^t and y^t evolve according to the probabilistic transitions of $\mathcal{A}$, given initial configurations $x^0 = T_0$ and $y^0 = U_0$ respectively; and

(ii) If $x^t = y^t$, then $x^{t+1} = y^{t+1}$.

A coupling $\{(x^t, y^t)\}_{t\geq 0}$ is *locally stable* iff, for all $1 \leq i \leq N$, for all $t \geq 0$:

$(x^t(i-1) = y^t(i-1) \wedge x^t(i) = y^t(i)) \Rightarrow x^{t+1}(i) = y^{t+1}(i)$.

We will use $(x, y)^t(i)$ as an abbreviation for $(x^t(i), y^t(i))$. Under this notation, local stability means that, for all $q, r, s \in Q$, all $1 \leq i \leq N$ and all $t \geq 0$:

$((x, y)^t(i-1) = (q, q) \wedge (x, y)^t(i) = (r, r)) \Rightarrow (x, y)^{t+1}(i) = (s, s)$.

Example 4. Let $\{x^t\}_{t\geq 0}$ and $\{y^t\}_{t\geq 0}$ be two copies of Herman's algorithm (see example 1). We define a coupling as follows :

Case 1: $(x,y)^t(i) = (x,y)^t(i+1) = (0,0)$ or $(x,y)^t(i) = (x,y)^t(i+1) = (1,1)$.

Case 2: $(x,y)^t(i) = (x,y)^t(i+1) = (0,1)$ or $(x,y)^t(i) = (x,y)^t(i+1) = (1,0)$.

In both Cases 1 and 2, the coupling forces machines $x^t(i+1)$ and $y^t(i+1)$ to do the same random choice, so that, at step $t+1$:

$(x,y)^{t+1}(i+1) = (1,1)$ with probability $\frac{1}{2}$, or

$(x,y)^{t+1}(i+1) = (0,0)$ with probability $\frac{1}{2}$.

It is easy to see that this coupling ensures local stability, i.e., for all $q, r, s \in \{0,1\}$:

$((x,y)^t(i) = (q,q) \wedge (x,y)^t(i+1) = (r,r)) \Rightarrow (x,y)^{t+1}(i+1) = (s,s)$.

This is because, in the deterministic situation at time t, i.e., if $(x,y)^t(i) = (0,0) \wedge (x,y)^t(i+1) = (1,1)$ (resp.: $(x,y)^t(i) = (1,1) \wedge (x,y)^t(i+1) = (0,0)$), we have at time $t+1$: $(x,y)^{t+1}(i+1) = (0,0)$ (resp.: $(x,y)^{t+1}(i+1) = (1,1)$).

3 Computation Traces As A Markov Field

3.1 Random fields

An $\mathcal{A}$-*trace* $\mathcal{T}$ of length m starting with a given configuration T_0, is a sequence of configurations $T_0, T_1, \cdots, T_m$ in Q^N such that $T_{l-1} \xrightarrow{\mathcal{A}} T_l$ for all $1 \leq l \leq m$. Such a trace can be represented as a grid of size $(m+1) \times N$, where line 0 corresponds to the initial configuration T_0, and (l,i)-th coordinate ($1 \leq l \leq m$, $1 \leq i \leq N$) corresponds to the state of the i-th machine after l-th step. The probability associated with $\mathcal{T}$ is $\prod_{l=1}^{m} p(T_{l-1} \xrightarrow{\mathcal{A}} T_l)$. It represents the probability, given an initial configuration T_0, that $\mathcal{A}$ generates the sequence $T_1, \cdots, T_m$ of configurations after m steps.

More precisely, let $S = \{0, 1, \cdots, m\} \times \{1, \cdots, N\}$ denote the set of coordinates (or *sites*) of $\mathcal{T}$. Let S_l be the subset $\{(l,1), \cdots, (l,N)\}$ of S, for every $0 \leq l \leq m$. Let $S'' = S_1 \cup \cdots \cup S_{m-1}$ and $S' = S'' \cup S_m$ (hence: $S = S_0 \cup S' = S_0 \cup S'' \cup S_m$). We have:

Definition 5. Given a Markov chain $\{x^t\}_{t\geq 0}$ (associated with $\mathcal{A}$) starting by a given configuration $x^0 = T_0$, and an integer $m \geq 0$, the *random field X of $\mathcal{A}$-traces of length m* starting with T_0 is a collection $X = \{X(l,i)\}_{(l,i)\in S}$ of Q-valued random variables such that, for all $(l,i) \in S$: $X(l,i) = x^l(i)$.

Let Ω_S denote the set of all maps from S to Q. A trace can be regarded as an element of Ω_S, and denoted by $X = \{X(s)\}_{s\in S}$ if it concerns the random field, and by $\mathcal{T}, \mathcal{U}, \cdots$ if we consider a fixed trace. We will also use this notation for elements of Ω_B, where B is a subset of S. For $C \subset B \subset S$, $\eta \in \Omega_B$ and $\eta' \in \Omega_C$, we write "$\eta \equiv \eta'$ on C" if the restriction of η to the set of sites of C coincides with η'. For $B \subset S$, $s \in S$ and $\eta \in \Omega_B$, expression $\eta(s)$ denotes the value of η at site s, and $\eta(B)$ is a shorthand notation for $\{\eta(s)\}_{s\in B}$.

Define, for each $T_0 \in \Omega_{S_0}$, the conditional probability distribution on $\Omega_{S'}$:

$\mathcal{P}^{\mathcal{A}}_{S'}(.,T_0) = Pr(X \equiv . \text{ on } S' \mid X \equiv T_0 \text{ on } S_0)$.
We have, for all $\mathcal{T}' \in \Omega_{S'}$:
$\mathcal{P}^{\mathcal{A}}_{S'}(\mathcal{T}',T_0) = p(T_0 \xrightarrow{\mathcal{A}} T_1)p(T_1 \xrightarrow{\mathcal{A}} T_2)\cdots p(T_{m-1} \xrightarrow{\mathcal{A}} T_m)$,
where T_l is the restriction of $\mathcal{T}'$ on S_l $(1 \leq l \leq m)$.

Example 6. Consider Herman's algorithm. Suppose that $N = 5$ and $m = 3$. The following grid

$$\begin{pmatrix} 1\,0\,0\,0\,0 \\ 0\,1\,1\,1\,0 \\ 1\,0\,1\,0\,1 \\ 0\,1\,0\,1\,0 \end{pmatrix}$$

represents trace $\mathcal{T} = T_0T_1T_2T_3$ in which T_l corresponds to line l of the grid $(0 \leq l \leq 3)$. The probability of going from line 0 to 1 (resp. from 1 to 2) is $(1/2)^3$. The probability of going from line 2 to 3 is $1/2$. Therefore $\mathcal{P}^{\mathcal{A}}_{S'}(\mathcal{T}',T_0) = (1/2)^3(1/2)^3 1/2 = 2^{-7}$ where $\mathcal{T}'$ is $T_1T_2T_3$.

For $T_m \in \Omega_{S_m}$, let:

$$\mathcal{P}^{\mathcal{A}}_{S',S_m}(T_m,T_0) = \sum_{\mathcal{T}' \in \Omega_{S'}:\ \mathcal{T}' \equiv T_m \ on \ S_m} \mathcal{P}^{\mathcal{A}}_{S'}(\mathcal{T}',T_0).$$

$\mathcal{P}^{\mathcal{A}}_{S',S_m}(.,T_0)$ is called the *marginal distribution* on S_m of $\mathcal{P}^{\mathcal{A}}_{S'}(.,T_0)$.

Proposition 7. *We have, for all $T_0, T_m \in Q^N$:*
$\mathcal{P}^{\mathcal{A}}_{S',S_m}(T_m,T_0) = Pr(x^m = T_m \mid x^0 = T_0)$.

In words: the marginal distribution on S_m of the probability distribution of X is equal to the distribution of x^m.

Proof:

$$\mathcal{P}^{\mathcal{A}}_{S',S_m}(T_m,T_0) = \sum_{\mathcal{T}' \in \Omega_{S'}:\ \mathcal{T}' \equiv T_m \ on \ S_m} \mathcal{P}^{\mathcal{A}}_{S'}(\mathcal{T}',T_0)$$

$$= \sum_{U_1 \in Q^N, \cdots, U_{m-1} \in Q^N} Pr(X \equiv (U_1U_2\cdots U_{m-1}T_m) \texttt{ on } S' \mid X \equiv T_0 \texttt{ on } S_0)$$

$$= \sum_{U_1 \in Q^N, \cdots, U_{m-1} \in Q^N} p(T_0 \xrightarrow{\mathcal{A}} U_1)p(U_1 \xrightarrow{\mathcal{A}} U_2)\cdots p(U_{m-1} \xrightarrow{\mathcal{A}} T_m)$$

$$= Pr(x^m = T_m \mid x^0 = T_0). \qquad \square$$

Definition 8. Let T_0, U_0 be configurations of Q^N, and $\{(x^t,y^t)\}_{t\geq 0}$ a coupling of $\mathcal{A}$ starting with (T_0,U_0). The *induced coupling of $\mathcal{A}$-traces of length m* is the $\Omega_S \times \Omega_S$-valued random variable (X,Y) such that, for all $(l,i) \in S$, we have:
$X(l,i) = x^l(i)$ and $Y(l,i) = y^l(i)$.

Note that, if the coupling $\{(x^t,y^t)\}_{t\geq 0}$ is locally stable, then the induced coupling (X,Y) is such that, for all $(l,i) \in S'$:
$(X(l-1,i-1) = Y(l-1,i-1) \wedge X(l-1,i) = Y(l-1,i)) \Rightarrow X(l,i) = Y(l,i)$.

3.2 Neighborhood

For a site $s \in S'$, the *neighborhood* is defined as follows:

- For $s = (l, i) \in S''$, the neighborhood of s, denoted $\mathcal{N}_s$, is the 6-tuple of sites $\langle (l-1, i-1), (l-1, i), (l, i-1), (l, i+1), (l+1, i), (l+1, i+1) \rangle$.

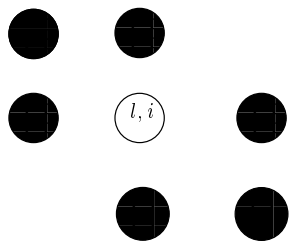


Fig. 1. Neighborhood for $s \in S''$.

- For $s = (m, i) \in S_m$, the neighborhood of s, denoted $\mathcal{O}_s$, is the couple of sites $\langle (m-1, i-1), (m-1, i) \rangle$.

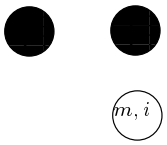


Fig. 2. Neighborhood for $s \in S^m$.

The neighborhood of $s \in S''$ (resp. $s \in S_m$) is depicted in black on Fig. 1 (resp. Fig. 2). We say that two sites r and s of S (not both in S_0) are *neighbors* or *adjacent*, and write $r \sim s$, when:

- $s \in \mathcal{N}_r$ if $r \in S''$, or
- $s \in \mathcal{O}_r$ if $r \in S_m$,

or symmetrically.

For $m = 3$ and $N = 5$, the graph of adjacents sites (l, i), with $0 \leq l \leq m = 3$ and $1 \leq i \leq N = 5$, is represented on Fig. 3. (For simplicity, the edges between site $(l, 1)$ and sites $(l, 5), (l-1, 5)$, with $l = 1, 2, 3$, have not been drawn.)

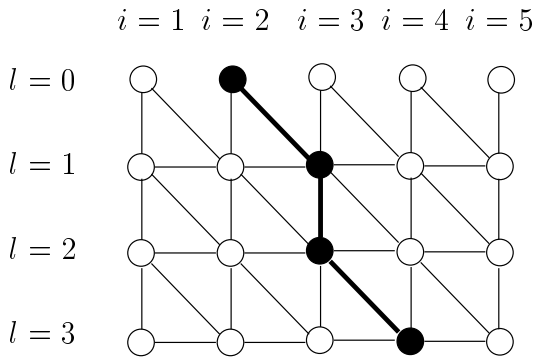


Fig. 3. Graph of adjacent sites.

By a *path* from S_i to S_j, we mean a sequence of distinct sites $s_1, s_2, \cdots, s_k$ in which the consecutive sites are adjacent (i.e.: $s_1 \sim s_2 \sim \cdots \sim s_k$), and with $s_1 \in S_i$ and $s_k \in S_j$.

An *S-SE path* from S_i to S_j is a path $s_1 \sim s_2 \sim \cdots \sim s_k$ from S_i to S_j that goes only in "South" or "South-East" directions, i.e., such that, for all h $(1 \leq h \leq k-1)$, if site s_h is (l, i), then s_{h+1} is $(l+1, i)$ (South site) or $(l+1, i+1)$ (South-East site). An example of S-SE path is drawn in bold on Figure 3.

It is easy to see that the field of $\mathcal{A}$-traces is a random field where the distribution of the value of site $s \in S''$ (resp. $s \in S_m$) is influenced only by the values of the adjacent sites of $\mathcal{N}_s$ (resp. $\mathcal{O}_s$). (See [5] for details.) This means that the field of $\mathcal{A}$-traces is a *Markov field*. Formally:

Proposition 9. *For all $m > 0$ and all $T_0 \in \Omega_{S_0}$, the random field X of $\mathcal{A}$-traces of length m starting from T_0, is a Markov field for the neighborhood system $\{\mathcal{N}_s\}_{s \in S''} \cup \{\mathcal{O}_s\}_{s \in S^m}$, i.e., for all $s \in S''$ (resp. $s \in S_m$), all $q \in Q$ and all $\eta \in \Omega_{S \setminus \{s\}}$:* $Pr(X(s) = q \mid X(S \setminus \{s\}) = \eta) = Pr(X(s) = q \mid X(\mathcal{N}_s) = \eta(\mathcal{N}_s))$ *(resp.* $Pr(X(s) = q \mid X(S \setminus \{s\}) = \eta) = Pr(X(s) = q \mid X(\mathcal{O}_s) = \eta(\mathcal{O}_s))$.

4 A Sufficient Condition For Rapid Mixing

4.1 van den Berg's condition

In order to prove the absence of "phase transition" of Markov fields, van den Berg introduced a criterion based on the notion of "path of disagreement" (see [11]). The absence of phase transition corresponds in our context to the fact that $\mathcal{A}$ has a unique stationary distribution π (see [5]). Furthermore, in our context, van den Berg's criterion ensures not only the uniqueness of π, but the rapid mixing of $\mathcal{A}$ to π, as shown below.

If $\mathcal{T}, \mathcal{U} \in \Omega_S$, we say that the pair $(\mathcal{T}, \mathcal{U})$ has a *path of disagreement* from S_0 to S_m if there is a sequence $r_1 \sim \cdots \sim r_k$ of sites in S with $r_1 \in S_0$, $r_k \in S_m$ and $\mathcal{T}(r_h) \neq \mathcal{U}(r_h)$ for $h = 0, \cdots, k$. The following result holds in the general case of Markov fields (see [12]; cf [11]).

Proposition 10. *For all $m > 0$, $r \in S_0$, and all pair $T_0, U_0 \in \Omega_{S_0}$ with $T_0(s) = U_0(s)$ if $s \neq r$, we have:*

$$d(\mathcal{P}^{\mathcal{A}}_{S',S_m}(.,T_0),\ \mathcal{P}^{\mathcal{A}}_{S',S_m}(.,U_0)) \leq (\mathcal{P}^{\mathcal{A}}_{S'}(.,T_0) \times \mathcal{P}^{\mathcal{A}}_{S'}(.,U_0))$$
$$(\exists \text{ path of disagreement from } S_0 \text{ to } S_m)$$

In words, Proposition 10 says that the variational distance between the marginal distributions on S_m of the two distributions $\mathcal{P}^{\mathcal{A}}_{S'}(.,T_0)$ and $\mathcal{P}^{\mathcal{A}}_{S'}(.,U_0)$ is smaller than or equal to the probability that, if we randomly and independently choose two elements in $\Omega_{S'}$, one with the first distribution above, the other with the second distribution, this pair of elements has a pair of disagreement from S_0 to S_m. van den Berg and Maes have extended this result in the case where, instead of independently choosing two independent elements in $\Omega_{S'}$, one considers a coupling of the two distributions that satisfies certain "monotony" conditions (see [14]). However, the result does not extend to *any* coupling, as shown by Häggström's counterexample [6].

Let us show that, in our context of $\mathcal{A}$-traces, (a variant of) van den Berg's result holds when the coupling (X, Y) is induced by a "locally stable" coupling $\{(x^t, y^t)\}_{t \geq 0}$. Furthermore: on the one hand, van den Berg's result will be enhanced by focusing to S-SE paths (of disagreement); on the other hand, it will be weakened by considering arbitrary initial configurations T_0 and U_0, that may differ in more than one site r. Using the fact that $\mathcal{P}^{\mathcal{A}}_{S',S_m}(T_m, T_0) = Pr(x^m = T_m \mid x^0 = T_0)$ (see Proposition 7), we adapt Proposition 10 as follows:

Proposition 11. *For each $m > 0$, each pair $T_0, U_0 \in \Omega_{S_0}$ and each locally stable coupling $\{(x^t, y^t)\}_{t \geq 0}$ of $\mathcal{A}$ with $(x^0, y^0) = (T_0, U_0)$, the induced coupling (X, Y) of $\mathcal{A}$-traces of length m (whose distribution we denote by μ) is such that:*

$$d(Pr(x^m = . \mid x^0 = T_0),\ Pr(y^m = . \mid y^0 = U_0))$$
$$\leq\ \mu(\exists\ S\text{-}SE \text{ path of disagreement from } S_0 \text{ to } S_m).$$

The proof of Proposition 11 relies on the following:

Lemma 12. *For each $m > 0$, each pair $T_0, U_0 \in \Omega_{S_0}$ and each locally stable coupling $\{(x^t, y^t)\}_{t \geq 0}$ of $\mathcal{A}$ with $(x^0, y^0) = (T_0, U_0)$, the induced coupling (X, Y) of $\mathcal{A}$-traces of length m (whose distribution we denote by μ) is such that:*

$x^m \neq y^m$ iff there is an S-SE path of disagreement from S_0 to S_m (μ-a.s.).

Proof of Proposition 11

$$\begin{aligned} &d(Pr(x^m = . \mid x^0 = T_0),\ Pr(y^m = . \mid y^0 = U_0)) \\ &\quad \leq Pr(x^m \neq y^m \mid (x^0, y^0) = (T_0, U_0)) && \text{(by "coupling lemma")} \\ &\quad \leq \mu(\exists\ S\text{-}SE \text{ path of disagreement from } S_0 \text{ to } S_m) && \text{(by Lemma 12).} \end{aligned}$$

□

Definition 13. The field of $\mathcal{A}$-traces is *rapidly agreeing* if for each $m > 0$ and each pair $T_0, U_0 \in \Omega_{S_0}$, there exists a locally stable coupling $\{(x^t, y^t)\}_{t\geq 0}$ of $\mathcal{A}$ with $(x^0, y^0) = (T_0, U_0)$ such that the induced coupling (X, Y) of $\mathcal{A}$-traces of length m (whose distribution we denote by μ) satisfies:

$\mu(\exists\ S\text{-}SE$ path of disagreement from S_0 to $S_m) \leq Kexp(-c\ m)$,
where K and $1/c$ are polynomial in N.

We have:

Theorem 14. *If the field of $\mathcal{A}$-traces is rapidly agreeing, then $\mathcal{A}$ has a unique stationary distribution π, and $\mathcal{A}$ is rapidly mixing.*

Proof: Let us first show ad absurdum that there is only one stationary distribution. If there were two distinct stationary distributions π_1 and π_2, the distance after m steps between the distribution issued from an initial distribution equal to π_1 and the distribution issued from an initial distribution equal to π_2, would be greater than a positive constant, and would not decrease exponentially with m. By Proposition 11, this would contradict the assumption of rapid agreement.

Let π be the unique stationary distribution. By proposition 11 and rapid agreement, we have:

$d(Pr(x^m = . \mid x^0 = T_0), \pi)$
$\leq\ \mu(\exists\ S\text{-}SE$ path of disagreement from S_0 to $S_m) \leq Kexp(-cm)$.
The mixing time $\tau(\varepsilon)$ is thus bounded by the relation $Kexp(-c\tau(\varepsilon)) \leq \varepsilon$, i.e.:

$\tau \geq \frac{ln(\varepsilon^{-1})}{c} + \frac{ln(K)}{c}$, which shows that τ is polynomial in N and $ln(\varepsilon^{-1})$. □

5 Applications

5.1 Herman's self-stabilizing algorithm of mutual exclusion.

Let us come back to Herman's algorithm (see example 1). In a configuration, a "token" at position i ($1 \leq i \leq N$) corresponds to the presence of two contiguous states of the same value (00 or 11) at position $i-1$ and i. We suppose now that N is odd. In this case, any configuration contains always at least one token. Given a configuration, each step of execution makes every token move one place to the right with probability $\frac{1}{2}$, or stay at the same place. It is easy to show that the number of tokens never increases. The number of tokens may decrease when a token "collides" the token to its right: an overlapping of two tokens (of the form 000 or 111) then occurs, and, at next step, these two collided tokens disappear simultaneously with probability 1/4. The evolution of the distance between two consecutive tokens may be seen as a random walk on an interval, so the event of collision is of probability 1, and all the tokens but one eventually disappear with probability 1. The set $\mathcal{L}$ of configurations with one token corresponds to the set of "legal configurations" (see [2]). This set is "strongly connected" or "ergodic": one can go from any element of $\mathcal{L}$ to any one via a finite set of transitions. It is easy to show that the associated stationary distribution π is the uniform distribution over the elements of $\mathcal{L}$. (More precisely, $\pi(\overline{q}) = 1/(2N)$ if $\overline{q} \in \mathcal{L}$, and

$\pi(\overline{q}) = 0$ otherwise.) Such an ergodic set $\mathcal{L}$ (and accordingly, such a stationary distribution π), is unique. This uniqueness corresponds to the property of "self-stabilization" according to Dijkstra's terminology (see [2,5]). We now show:

Theorem 15. *For N odd, the field of traces of Herman's algorithm is rapidly agreeing.*

By Theorem 14, this gives a new proof of the fact that Herman's algorithm is self-stabilizing (i.e, that the stationary distribution π is unique). In other terms, whatever the initial configurations T_0 is (in particular, whatever the number of tokens T_0 contains), the executions converges with probability 1 to a legal configuration (i.e., a configuration with a single token). Such a configuration corresponds to the case of "mutual exclusion" where no more than one machine enjoys the resource. Theorem 14 guarantees furthermore that the distribution of configurations is "ε-close" to the uniform distribution π over legal configurations, after N^2 steps. To our knowledge, the hitting time of Herman's algorithm was known to be quadratic but it is the first time that a quadratic upper bound is obtained for the mixing time.

The idea of the proof of Theorem 15 is as follows. Let us consider two $\mathcal{A}$-traces X and Y induced by the coupling (x, y) of example 4, that do not agree in some site in S_0 ($x^0 \neq y^0$). Let us compute the probability of the existence of a disagreement path on (X, Y).

Consider two "extreme" positions of disagreement i_0 and j_0 of S_0, i.e. sites of disagreement ($x^0(i_0) \neq y^0(i_0)$, $x^0(j_0) \neq y^0(j_0)$) such that all the sites of S_0 "outside" $[i_0, j_0]$ or "inside" $]i_0, j_0[$ agree. Without loss of generality, suppose: $1 \leq i_0 \leq j_0 \leq N$ and $x^0(k) = y^0(k)$, for all $k, 1 \leq k \leq i_0 - 1$ or $j_0 + 1 \leq k \leq N$.

By local stability, all sites between $j_0 + 2$ and $i_0 - 1$ still agree at step 1: for all k with $0 \leq k \leq i_0 - 1$ or $j_0 + 2 \leq k \leq N$, $(x, y)^1(k) = (q_k, q_k)$ for some $q_k \in Q$. Consider now the limit positions i_0 and $j_0 + 1$.

At i_0, the value of $(x, y)^1$ depends on the values of $(x, y)^0$ at $i_0 - 1$ and i_0. Now: $(x, y)^0(i_0 - 1) = (q, q)$ and $(x, y)^0(i_0) = (q', \neg q')$ for some $q, q' \in Q$. To fix the ideas, suppose: $(x, y)^0(i_0 - 1) = (0, 0)$ and $(x, y)^0(i_0) = (0, 1)$. (The other cases are similar.) At position i_0, there is a token in x_0 (of value 0) but no token in y_0. The value of $x^1(i_0)$ is 0 (resp. 1) with probability $\frac{1}{2}$ while the value of $y^1(i_0)$ is 0 with probability 1. Hence x^1 and y^1 agree (resp. disagree) at i_0 with probability $\frac{1}{2}$. In case of agreement, the position of extreme left disagreement moves (at least) one place right, and becomes $i_0 + 1$ in S_1 (SE-move). In case of disagreement, it remains equal to i_0 in S_1 (S-move).

Likewise, at the leftmost agreement site $j_0 + 1$, there is a token in x^0 and no token in y^0, or vice-versa. So, at step 1, sites x^1 and y^1 disagree (resp. agree) at site $j_0 + 1$ with probability $\frac{1}{2}$. In case of disagreement, the extreme right position of disagreement moves one place right, and becomes $j_0 + 1$ in S_1 (SE-move); in case of agreement, it is (at most) equal to j_0 (S-move).

Therefore the site of extreme left disagreement follows an S-move with probability $\frac{1}{2}$, or (in the worst case) an SE-move with probability $\frac{1}{2}$ from S_0 to S_1.

Likewise, the site of extreme right disagreement follows (in the worst case) an S-step with probability $\frac{1}{2}$, or an S-SE-step with probability $\frac{1}{2}$ from S_0 to S_1.

Repeating the reasoning, we define two extreme S-SE-paths of disagreement $L : l_0 \sim l_1, \cdots$ and $R : r_0 \sim r_1, \cdots$ until one of the two possible limit cases occurs: case of "total disagreement", or case of "total agreement". The case of "total disagreement" appears when, at step t, no position agrees ($l_t = (t, i_t), r_t = (t, j_t)$ with $j_t = i_t + (N-1)$). Since N is odd, there is always at least one token in x^t. In case of total disagreement, the position of of a token, say 00, in x^t corresponds to a position of token, say 11, in y^t. So, at step $t+1$, such a position agrees with probability 1, according to our coupling. A case of total disagreement at step t is thus "transient", since one always comes back at step $t+1$ to the general situation where at least one site agrees. On the other hand, a case of a total agreement at step t (all sites agree in $(x, y)^t$) is definitive, and corresponds to the definitive stopping of L and R (as well as any possible disagreement path between them): all the sites always continue to agree between t and m.

The left (resp. right) S-SE path L (resp. R) of disagreement is represented in bold (dotted) line in Figure 4. All the sites agree "outside" the zone they delimite. Inside the zone, the S-SE-paths are generally (but not necessarily) paths of disagreement. This zone is the *zone of disagreement*.

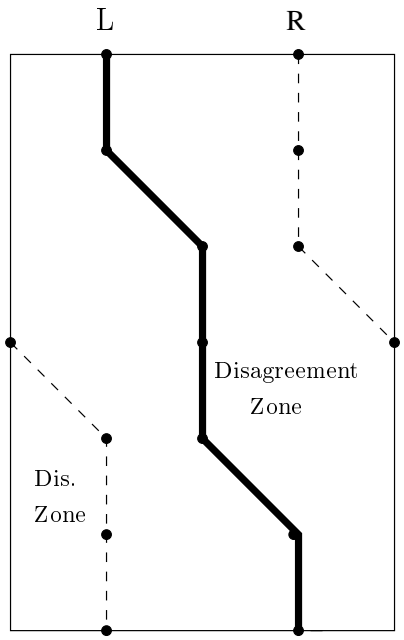


Fig. 4. Zone of disagreement

Each path of disagreement in (X, Y) is an *S-SE*-path of length m and probability $(1/2)^m$. In order to evaluate (an upper bound of) the probability Pr of of a disagreement path, it suffices, to evaluate the probability of the existence of such a couple (L, R) of *S-SE*-paths of length m. This mainly reduces to evaluate the number of couples (L, R). In Appendix 1, we explain how to calculate this number, and show that the mixing time of Herman's algorithm is quadratic in N.

5.2 Rapid mixing of a random walk on a ring.

The former example (Herman's algorithm) corresponds to an algorithm originating from the community of distributed computing. One interesting feature of this algorithm is the fact that the corresponding Markov chain is *non*-reversible. We show now that our combinatorial method on disagrement paths also applies with a good precision to the evaluation of mixing time of more classical (reversible) Markov chains. This is illustrated on the classical random walk on a ring. The mixing time is known to be quadratic (see, e.g., Diaconis and Saloff-Coste [1], p. 732). Any configuration is of the form $(0, \cdots, 0, 1, 0, \cdots, 0)$. The configuration is determined by the position of the (unique) 1. The randomized command is of the form:

`IF` $(x_{i-1}, x_i, x_{i+1}) = (0, 1, 0)$ `THEN` $(x'_{i-1}, x'_i, x'_{i+1}) = (1, 0, 0)$ (resp. $(0, 0, 1)$) with probability $\frac{1}{2}$.

Although this form of command is not exactly of the same form as described in section 2 (since the updated value of a state depends also on the current value of the *right* neighbour), it is easy to extend our method on disagreement paths in a straightforward manner. (In particular, Proposition 11 still holds, provided that we replace the notion of *S-SE*-path by that of *S-SE-SW*-path.)

Let us now explain how the method of disagreement path allows us to retrieve the quadratic upper bound. We define a locally stable coupling as follows. For all $1 \leq i \leq N$, for all $t \geq 0$:

If $(x, y)^t(i) = (1, 1)$ then $(x, y)^{t+1}(i + 1) = (1, 1)$ with probability $\frac{1}{2}$, or $(x, y)^{t+1}(i - 1) = (1, 1)$ with probability $\frac{1}{2}$.

We have:

Theorem 16. *The field of traces of the random walk on a ring is rapidly agreeing.*

The idea of the proof is similar as before: we estimate the probability of disagreement paths (of length m) on the induced coupling (X, Y) of $\mathcal{A}$-traces, by introducing the notion of zone of disagreement. Disagreement paths are here *S-SE-SW*-paths (instead of *S-SE*-paths, as in Herman's algorithm). The problem is now to enumerate the number of *S-SE-SW*-paths. The same kind of calculation as the one done for Herman's algorithm leads us to a quadratic upper bound on the mixing time of the random walk. For details, see Appendix 2.

6 Final Remarks

In this paper, we have used (a variant of) van den Berg's notion of path of disagreement to show the rapid mixing of Markov chains for linear graphs. Our method presents several original features:

- It enhances the notion of disagreement path by focusing on top-down paths (*S-SE*-paths or *S-SE-SW*-paths).
- It combines the notion of disagreement path with the notion of coupling, without the classical requirement of existence of an ordering on the configurations. (The only requirement on the coupling is here a natural property of "local stability", ensuring the equivalence between the existence of a disagreement site s and the existence of a disagreement path going to s).
- It applies to non-reversible Markov chains, and proves as a side-effect the property of "self-stabilization" (i.e., uniqueness of stationary distribution).
- It is an elementary method the application of which relies essentially on a combinatorial calculation.
- It works with a fully parallel (or "synchronous") dynamics; note that it could be adapted to a systematic scan (or "roundabout") dynamics as well.

First examples of application, as those presented above, tend to show that the method offers a good precision for upper bounding the mixing time of randomized algorithms on linear graphs (rings or linear arrays).

References

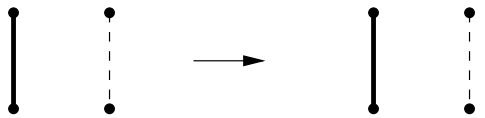
1. P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *The Annals of Applied Probability*, 6(3):695–750, 1996.
2. E.W. Dijkstra. Self-stabilizing systems in spite of distributed control. *Communications of the ACM*, 17(11):643–644, Nov. 1974.
3. M. Dyer and A. Frieze. Randomly colouring graphs with lower bounds on girth and maximal degree. In *Proc. of the 42nd Annual Symp. on Foundations of Computer Science (FOCS)*, pages 579–587, 2001.
4. M. Dyer, A. Sinclair, E. Vigoda, and D. Weitz. Mixing in Time and Space for Lattice Spin Systems: A Combinatorial View. Technical report, TR-2003-05, Computer Science, The University of Chicago, USA, March 2003. (An earlier version appeared in the conference RANDOM 2002, pp. 149-163.).
5. L. Fribourg, S. Messika, and Picaronny C. On the absence of phase transition in randomized distributed algorithms. Technical report, LSV-03-07, ENS de Cachan, France, April 2003. Available on `http://www.lsv.ens-cachan.fr/~fribourg/publis.php`.
6. O. Haggstrom. A note on disagreement percolation. *Rand. Structures Algorithms*, 18:267–278, 2001.
7. T. Herman. Probabilistic self-stabilization. *IPL*, 35(2):63–67, June 1990.
8. M. Jerrum and A. Sinclair. *The Markov chain Monte Carlo method: an approach to approximate counting and integration*. In "Approximation Algorithms for NP-hard Problems", D.S.Hochbaum ed., PWS Publishing, 1996.

9. H. Kakugawa and M. Yamashita. Uniform and self-stabilizing token rings allowing unfair daemon. *IEEE Trans. Parallel and Distributed Systems*, 8(2), 1997.
10. M. Molloy. The Glauber dynamics on colourings of a graph with high girth and maximim degree. In *Proc. of the 34th ACM Symposium on Theory of Computing (STOC)*, pages 91–98, 2002.
11. J. van den Berg. A uniqueness condition for Gibbs measure with application to the 2-dimensional Ising antiferromagnet. *Commun. Math. Phys.*, 152:161–166, 1993.
12. J. van den Berg. On the absence of phase transition in the monomer-dimer model. *Perplexing Problems in Probability (Festschrift in honor of Harry Kesten), M. Branson and R. Durrett eds.*, pages 185–195, 1999.
13. J. van den Berg and R. Brouwer. Random Sampling for the Monomer-Dimer Model on a Lattice. Technical report, PNA-R9917, CWI, The Netherlands, 1999.
14. J. van den Berg and C. Maes. Disagreement percolation in the study of markov field. *Ann. Probab.*, 22:749–763, 1994.

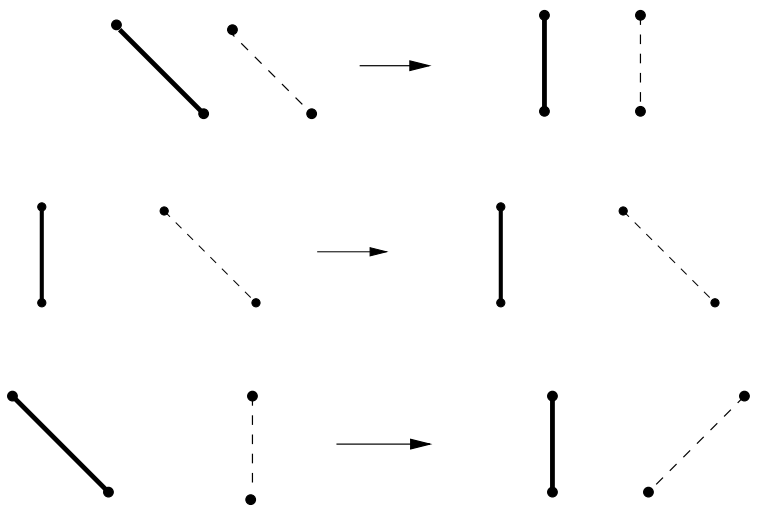
Appendix 1. Proof of Theorem 15.

We use the coupling described in example 4. We want to estimate the probability of S-SE-paths of disagreement (of length m) on the induced coupling (X, Y). More precisely, we defined two extreme S-SE-paths of disagreement L and R of length m inside the rectangular zone of length m and width N (outside these two paths, $x^t(i) = y^t(i)$). We want to calculate the probability of the existence of such couples L and R. Without loss of generality, we suppose that L starts at position $i_0 = 1$ of S_0. We are going to transform L into a path L' (starting from $i_0 = 1$) made of S-steps only; we transform accordingly R into a path R' located at the same distance to L' than R to L. Path R' will be made not only of S and SE moves, but also of SW-moves (i.e., moves from a site (t, j) to $(t+1, j-1)$). Geometrically, we proceed using the following elementary transformations:

a. If the t-th steps of L and R are both S-moves, then the t-th steps of L' and R' are S-moves:



b. If the t-th steps of L and R are both S-E-moves, then the t-th steps of L' and R' are *again* S-moves:
c. If the t-th step of L (resp. R) is an S-move (resp. S-E-move), then the t-th step of L' (resp. R') is an S-move (resp. S-E-move):
d. If the t-th step of L (resp. R) is an S-E-move (resp. S-move), then the t-th step of L' (resp. R') is an S-move (resp. S-O-move):



After this transformation, any *S-SE* path R in the initial disagreement zone becomes an *S-SE-SW* path R' in the rectangular zone of length m and width N. In the original figure, the probability associated with any step of couple paths (L, R) is always $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$. In the transformed figure, the probability associated with a step of (L', R') is $\frac{1}{4}$ when it corresponds to an *SE* or *SW* step of R', and $\frac{1}{4} + \frac{1}{4} = \frac{1}{2}$ when it corresponds to an *S*-step of R' (due to the redundancy of representation in cases a and b). For example, paths L and R of Figure 4 are transformed into L' and R' in Figure 5. In case of total disagreement at step t, the exact transformation of (L, R) into (L', R') should yield an *SW*-move of probability 1 for R' (see section 5.1). For the sake of homogeneity, we actually represent such a move by an *SW*-move of probability 1/4 and an *S*-move of probability 1/2, which contributes to (slightly) overapproximate the disagreement event.

The evaluation of the disagreement event now reduces to the enumeration of all the *S-SE-SW*-paths in the rectangular zone of length m and width N.

We first make abstraction of the *S*-moves of R', and calculate the number of *SE-SW*-paths of length n (with $n \leq m$) (see Figure 6).

Lemma 17. *Let $U_{i,n}$ be the number of* SE-SW*-paths of length n starting from site $(0, i)$ in a rectangular zone of width N. There exists a constant A_N such that, for all $i \in [1, \cdots, N]$ and all $n > 0$:*

$$U_{i,n} \leq A_N (2 \cos \frac{\pi}{N+1})^n.$$

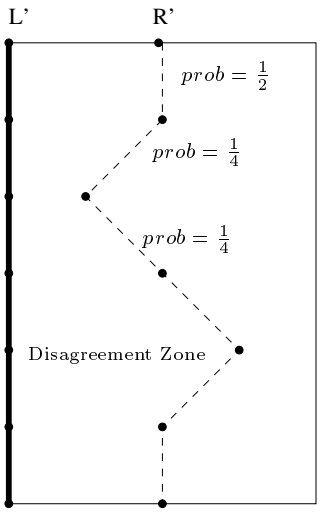


Fig. 5. Zone of disagreement after transformation

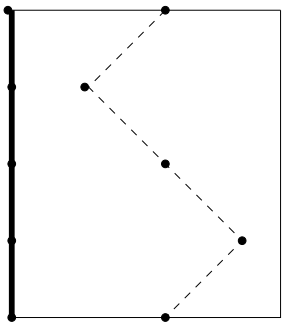


Fig. 6. Rectangular zone after removal of S-moves

Proof. Let $U_n = (U_{1,n}, \cdots, U_{i,n}, \cdots, U_{N,n})$. The sequence $(U_n)_{n\geq 0}$ satisfies the linear recurrence: $U_{n+1} = PU_n$ where the $N \times N$ matrix P is defined by:

$$P = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & 1 & \dots & 0 & 0 \\ 0 & 1 & 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 1 \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

The eigenvalues of this matrix are $(2\cos\frac{k\pi}{N+1})_{1\leq k\leq N}$.Therefore, for all i in $[1,N]$, there exist constants $A_{i,1},\cdots,A_{i,N}$ such that, for all $n>0$:
$U_{i,n}=\sum_{k=1}^{N}A_{i,k}(2\cos\frac{k\pi}{N+1})^n$. We let: $A_N=max\{\sum_{k=1}^{N}|A_{i,k}|\ ;\ i\in[1,N]\}$. □

Proposition 18. *Denote* $\mu_N=\frac{1+\cos\frac{\pi}{N+1}}{2}=(\cos\frac{\pi}{2(N+1)})^2$. *There exists a constant* K_N *such that, for all* $m>0$*:*

$$Pr(\exists \text{ disagreement path of length } \ m)\leq K_N(\mu_N)^m.$$

Proof. Let $\lambda_N=2\cos\frac{\pi}{N+1}$. Let $n\leq m$. In the rectangular zone $[1,\cdots,m]\times[1,\cdots,N]$, the number of *S-SE-SW*-paths of length m, leaving from $(0,i)$, with exactly $m-n$ *S*-steps is $C_m^{m-n}U_{i,n}$. (We choose the $m-n$ *S*-steps and remove them in order to get an *SE-SW*-path of length n.) The probability of such a path is $\frac{1}{4}^n\frac{1}{2}^{m-n}$. Therefore, applying lemma 17, we have:

$$\begin{aligned}Pr(\exists \text{ disagreement path of length } \ m)&\leq A_N\sum_{n=0}^{m}C_m^{m-n}\frac{1}{4}^n\frac{1}{2}^{m-n}\lambda_N^n\\&=A_N(\frac{\lambda_N}{4}+\frac{1}{2})^m\\&=A_N(\frac{\lambda_N+2}{4})^m\\&=A_N(\mu_N)^m.\end{aligned}$$

□

We are now ready to compute the mixing time.

Lemma 19. *As* N *goes to* ∞*, we have:* $-ln\mu_N\sim\frac{\pi^2}{4N^2}$.

Proposition 20. *Herman's algorithm is rapidly agreeing. Furthermore, its mixing time is quadratic in* N.

Proof. $Pr(\exists$ disagreement path of length $t)\leq K_N exp(-ct)$,
where $c=-ln(\mu_N)\sim\frac{\pi^2}{4N^2}$ (by Lemma 19). So, by Definition 13, Herman is rapidly agreeing. By Theorem 14, the mixing time is quadratic (as $1/c$). □

Appendix 2. Proof of Theorem 16.

Let us enumerate the number of couples of extreme *S-SE-SW*-paths. As before, we define two extreme *S-SE-SW*-paths of disagreement of length m, $L : l_1 \sim l_2 \sim \cdots \sim l_m$ and $R : r_1 \sim r_2 \sim \cdots \sim r_m$. Similarly, we transform each couple (L, R) into a couple (L', R') such that L' is made only of S-moves, and R' is made of moves of the form: S, SEE, SWW. (An SEE-move goes from a site (t, i) to $(t+1, i+2)$, and similarly for SWW.)

We now consider all the *S-SEE-SWW*-paths in the rectangular zone of length m and width N. Note that, for such a path, the probability of an SEE or SWW-move is $\frac{1}{4}$ but the probability of an S-move is $\frac{1}{4} + \frac{1}{4} = \frac{1}{2}$.

Let $V_{i,n}$ be the number of *SEE-SWW*-paths of length n starting from site $(0, i)$ in a rectangular zone of width N. We have a recurrence relation:

$V_{i,n} = V_{i-2,n-1} + V_{i+2,n-1} \quad \forall\ 2 \leq i \leq N-2$,
therefore the two sequences $V_n^{\text{odd}} = (V_{1,n}, \cdots, V_{2i+1,n}, \cdots) = (V_{2i+1,n})_{1 \leq 2i+1 \leq N}$ and $V_n^{\text{even}} = (V_{2,n}, \cdots, V_{2i,n}, \cdots) = (V_{2i,n})_{1 \leq 2i \leq N}$ satisfy independently the same recurrence relation as sequence U_n of lemma 17. Therefore, using M the integer part of $(N+1)/2$, we obtain the following results: and

Lemma 21. *There exists a constant B_N such that, for all $i \in [1, \cdots, N]$ and all $n > 0$:*

$$V_{i,n} \leq B_N (2 \cos \frac{\pi}{M+1})^n.$$

Proposition 22. *Denote $\nu_N = \frac{1+\cos\frac{\pi}{M+1}}{2} = (\cos\frac{\pi}{2(M+1)})^2$. There exists a constant I_N such that, for all $m > 0$:*

$$Pr(\exists \text{ disagreement path of length } \ m) \leq I_N (\nu_N)^m.$$

We are now ready to compute the mixing time.

Lemma 23. *As N goes to ∞, we have: $-ln\nu_N \sim \frac{\pi^2}{N^2}$.*

Proposition 24. *The random walk on a ring is rapidly agreeing. Furthermore, its mixing time is quadratic in N.*

Proof. $Pr(\exists$ disagreement path of length $t) \leq I_N exp(-ct)$,
where $c = -ln(\nu_N) \sim \frac{\pi^2}{N^2}$. So, by Definition 13, the random walk on a ring is rapidly agreeing. By Theorem 14, the mixing time is quadratic (as $1/c$). □