Markov Chains for Linear Extensions, the Two-Dimensional Case

Stefan Felsner   Lorenz Wernisch*
felsner@inf.fu-berlin.de   wernisch@inf.fu-berlin.de
Freie Universität Berlin,
Fachbereich Mathematik, Institut für Informatik,
Takustraße 9, 14195 Berlin,
Germany.

Abstract. We study the generation of uniformly distributed linear extensions using Markov chains. In particular we show that monotone coupling from the past can be applied in the case of linear extensions of two-dimensional orders. For width two orders a mixing rate of $O(n^3 \log n)$ is proved. We conjecture that this is the mixing rate in the general case and support the conjecture by empirical data. On the course we obtain several nice structural results concerning Bruhat order and weak Bruhat order of permutations.

1 Introduction

Markov chains have been proven useful for random generation and approximate counting of combinatorial objects, see Kannan [Kan94]. Among the best studied objects in this area are linear extensions of partial orders. The interest in random linear extensions has several sources. For example, it has applications in the general sorting problem (see Brightwell et al. [BFT95]). The problems of randomly generating and counting linear extensions correspond to the generation of random points and volume computations for the order polytope. This polytope is a good first instance for the application of algorithms that work on polyhedra or more general convex bodies (e.g., see Lovász [Lov92]). In their paper about volume approximation for polyhedra Dyer et al. [DFK89] already mention as consequence the approximation of the number of linear extensions. Karzanov and Khachiyan [KK91] study a natural Markov chain on the set of linear extensions of an order. Their algorithm returns an approximately uniformly distributed linear extension of an n element order after $O(n^6 \log n)$ steps.

Here is the process analyzed in [KK91] for generating a random linear extension of order P: start from an arbitrary linear extension $L = (x_1, .., x_n)$ of P. Pick a position $i \in \{1, .., n-1\}$ uniformly at random. If elements x_i and x_{i+1} are incomparable in P interchange them with probability $1/2$ otherwise leave L unchanged. Repeat this step a large number of times. From the theory of Markov chains it follows easily that the resulting linear extension is almost uniformly distributed. In [BW91b] Brightwell and

*Lorenz Wernisch has been supported by the Graduiertenkolleg "Algorithmische Diskrete Mathematik"

Winkler show how to extend this to a polynomial approximation scheme for the problem of counting linear extensions. In the same paper this counting problem is shown to be $\#P$-complete.

Recently, Propp and Wilson [PW95] proposed a stopping rule for Markov chains called *coupling from the past*. This rule can be used to sample according to the exact stationary distribution of a Markov chain. For coupling from the past to be practical the Markov chain has to obey a certain monotonicity property. In this manuscript we show that the above described Markov chain has the required monotonicity in the special case of two-dimensional orders. We report on experiments that support the conjecture that coupling yields a uniformly distributed linear extension in expected $O(n^3 \log n)$ steps. The conjecture is also supported by a proof of this coupling time for orders of width two.

The paper is organized as follows. In the next section we review some basic facts on Markov chains and describe how coupling from the past works. Section 3 exemplifies the framework with the easy example of linear extensions of width two orders. In Section 4 we show that the *path order* on the linear extensions allows monotone coupling for two-dimensional orders to work. In Section 5 we discuss mixing and coupling times and report on empirical studies.

2 Coupling from the past

Let $\{X_t\}_{t\geq 0}$ be an irreducible and aperiodic (i.e., ergodic) Markov chain on state space S with transition matrix $P = (p_{ij})$. From the theory of Markov chains it follows that there is a unique probability distribution π satisfying $\sum_i \pi_i p_{ij} = \pi_j$ for all states j. Hence, if π is the distribution of X_s then π is the distribution of X_t for all $t \geq s$, therefore π is called the *stationary distribution* of the Markov chain. It is known that if we run an ergodic Markov chain for a large number T of steps the probability that the final state is i converges to π_i. Much work has been done to analyze the convergence rate of combinatorial Markov chains, i.e., to estimate the time T the chain has to run so that the distribution of X_T and π are close, e.g., in total variation distance. For an overview see Kannan [Kan94].

The *coupling from the past* framework gives a stopping rule that allows to sample with unbiased stationary distribution. First we extend our time axis to the past and adopt the convention that the state returned by the sampling procedure will always be the state at time 0. This is achieved by starting the chain with X_{-T}. Now assume a routine RandomMap that returns at each call a function $f : S \to S$ so that, for all $i, j \in S$, $\Pr(f(i) = j) = p_{ij}$. If $\{f_t\}_{-T\leq t}$ is a sequence of functions produced by calls to RandomMap then for every $X_{-T} \in S$ the sequence $\{X_t\}_t$ with $X_{t+1} = f_t(X_t)$ is a Markov chain with transition matrix P. If T is large the distribution of X_0 will be close to π. The crucial observation, however, is the following. Suppose t' is such that the composed map $F_{-t'} = f_{-1} \circ f_{-2} \circ \ldots f_{-t'}$ is a constant map and i_* is the unique value in the image of the composed map then for all $T \geq t'$ and all initial states X_{-T} the state returned by the sampling procedure is i_*. Imagine the sequence $\{f_t\}$ to be extended to $-\infty$ and a π distributed random state $X_{-\infty}$. The sequence $\{X_t\}_t$ with $X_{t+1} = f_t(X_t)$ is a Markov chain with transition matrix P. Consequently, X_t is π distributed for every t. Such informal reasoning suggests that state i_* returned by the procedure is π distributed too.

The algorithm is as follows: At time t let $f_{-t} \leftarrow$RandomMap and consider $F_{-t} =$

$F_{-(t-1)} \circ f_{-t}$. As soon as F_{-t} becomes constant, return the constant state i_*. With the tacit assumption that the probability is 1 that the algorithm returns a value at all, we have the following theorem (see [PW95]).

Theorem 1 *If P is an ergodic transition matrix the state returned by coupling from the past is distributed according to the stationary distribution of the Markov chain.*

As stated the algorithm is far from practical. We now review, from [PW95] the *monotone Monte Carlo* algorithm for sampling according to π. Suppose there exists an ordering $\leq$ on the state space S with elements $\hat{0}, \hat{1} \in S$ such that $\hat{0} \leq x \leq \hat{1}$ for all $x \in S$. Moreover, every f ever produced by RANDOMMAP has the property that $x \leq y$ implies $f(x) \leq f(y)$. It is clear that in this setting F_{-t} is a constant map exactly if $F_{-t}(\hat{0}) = F_{-t}(\hat{1})$. Hence, instead of observing a huge state space we may restrict the observation to just two states.

3 Random linear extensions of orders of width two

The *width* of an order P is the minimum number of chains of a partition of the elements of P into chains. It is classical that a minimum chain partition of P can be found efficiently. Let P be of width two partitioned into chains $A = \{a_1 < a_2 < \ldots < a_k\}$ and $B = \{b_1 < b_2 < \ldots < b_{n-k}\}$. A linear extension $L = x_1, x_2, \ldots, x_n$ of P is a permutation of the elements such that for all $x < y$ in P element x comes before y in L. Let L be a linear extension of a width two order. With L associate a 0-1 sequence by

$$\text{sequence}(L,i) = \begin{cases} 0 & \text{if } x_i \in A, \\ 1 & \text{if } x_i \in B, \end{cases} \quad \text{for } 1 \leq i \leq n.$$

Let us also define

$$\text{ones}(L,k) = \sum_{1 \leq i \leq k} \text{sequence}(L,i), \quad \text{for } 1 \leq k \leq n.$$

Lemma 2 *The mapping $L \to \text{sequence}(L)$ is one-to-one and the image of the mapping is a subset of $\binom{[n]}{k}$.* □

There is a nice geometric characterization of the image of the mapping: visualize sequence(L) as a lattice path. Starting at the origin $(0,0)$ scan sequence(L) from left to right and draw a south-east diagonal step for each 0 and a north-east diagonal step for each 1 found. The paths end in point $(n, n-2k)$ (see Fig. 1).

Comparability relations between elements of A and B lead to forbidden areas for the paths corresponding to linear extensions. If $a_i < b_j$ the ith occurence of 0 in sequence(L) must be left of the jth occurence of 1 for all L, i.e., $\text{ones}(L, i+j-1) < j$. Symmetrically, relation $b_i < a_j$ translates to $\text{ones}(L, i+j-1) \geq i$. Thus the forbidden areas are $\vee$ and $\wedge$ shaped regions as shown in Fig. 2.

Fact 1. A lattice path from $(0,0)$ to $(n, n-2k)$ corresponds to a linear extension of P iff it respects the forbidden areas of all comparable pairs $a_i \in A$ and $b_j \in B$.

Lattice paths carry a natural 'below' ordering $\leq_p$ as shown in Fig. 4. Formally, sequence(S) $\leq_p$ sequence(T) iff $\text{ones}(S,k) \leq \text{ones}(T,k)$ for $1 \leq k \leq n$. Note that the

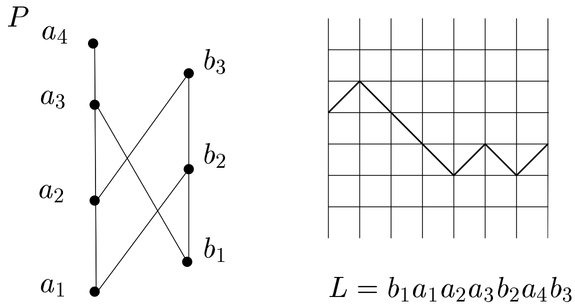

Figure 1: A linear extension and the corresponding path.

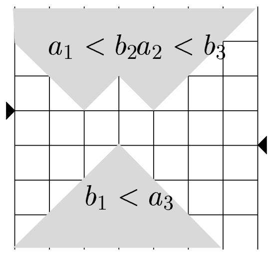

Figure 2: Forbidden wedges for paths of linear extensions of P.

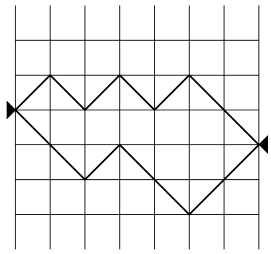

Figure 3: $\leq_p$-maximal and $\leq_p$-minimal lattice path.

forbidden areas for lattice path of linear extensions of P can be characterized in terms of a $\leq_p$-maximal lattice path τ_1 and a $\leq_p$-minimal lattice path τ_0 (see Fig. 3). Fact 1 can be restated as: a lattice path ρ from $(0, 0)$ to $(n, n-2k)$ corresponds to a linear extension of P iff $\tau_0 \leq_p \rho \leq_p \tau_1$.

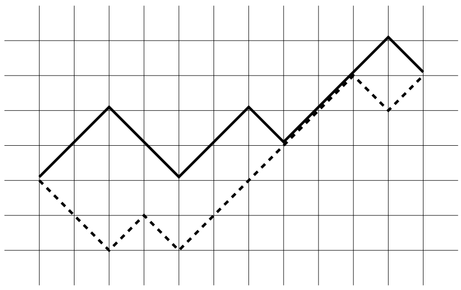

Figure 4: Path 11001101110 is above path 00101111101.

We come back to monotone coupling fom the past. Define a family of $2(n-1)$ random maps, two for each $i = 1..n-1$. The *upward flip at position i* changes '01' at positions $i, i+1$ in sequence(L) into '10' if allowed. The *downward flip at position i* changes '10' at positions $i, i+1$ in sequence(L) into '01' if allowed. (See Fig. 5). The next fact states the monotonicity of these random maps:

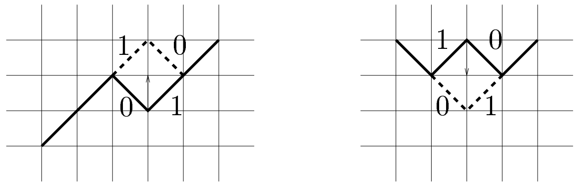

Figure 5: An upward and a downward flip.

Fact 2. Let τ' and ρ' be obtained by an upward (downward) flip at some position. If $\tau \leq_p \rho$ then $\tau' \leq_p \rho'$.

Theorem 1 and the monotonicity of the Markov-chain provide us with a simple method of generating a uniformly distributed linear extension of P: run the Markov-chain from the past until $F_t(\tau_0) = F_t(\tau_1)$, i.e., until τ_0 and τ_1 have coupled.

Theorem 3 *The expected number of steps it takes for τ_0 and τ_1 to couple is $O(n^3 \log n)$.*

This bound is almost sharp since a lower bound of $\Omega(n^3)$ for the expected number of steps can be given. See a forthcoming paper of Wilson [Wil95] for the details.

4 Linear extensions of two-dimensional orders

A classical theorem of Dushnik and Miller [DM41] states that every order is the intersection of its linear extensions. A set R of linear extensions of order P is called a *realizer* iff P is the intersection of the linear extensions in R. The *dimension* of P is the minimum number of linear extensions in a realizer. Let σ_1, σ_2 be a realizer of a two-dimensional order P on elements $1,..,n$. Rename x as $\sigma_1^{-1}(x)$ and let $\pi = \sigma_1^{-1}(\sigma_2(x))$. It follows that $x < y$ in P iff $x < y$ as integers and $\pi^{-1}(x) < \pi^{-1}(y)$ (denoted by $x \operatorname{bef}_\pi y$). This allows the identification of two-dimensional orders with n elements and permutations in S_n. Henceforth, we casually write P_π instead of P to indicate that P corresponds to π, i.e., that $\{\mathrm{id}, \pi\}$ is a realizer of P.

Pause for a moment to note that orders of width two are two-dimensional: the linear extensions corresponding to the maximum τ_1 and the minimum path τ_0 are the linear extensions of a realizer.

An *inversion* of a permutation σ is a pair (i, j) with $i < j$ and $j \operatorname{bef}_\sigma i$. Let $I(\sigma)$ be the set of inversions of σ. A *sorting transposition of $\sigma \in S_n$ at positions i and j* is a permutation $\sigma' = \mathrm{sort}(\sigma, i, j) \in S_n$ which is the same as σ save $\sigma'_i = \sigma_j$ and $\sigma'_j = \sigma_i$ if $\sigma_i > \sigma_j$ (otherwise, $\sigma' = \sigma$). Similarly, an *unsorting transposition* at positions i and j is a permutation $\sigma' = \mathrm{unsort}(\sigma, i, j)$ with $\sigma'_i = \sigma_j$ and $\sigma'_j = \sigma_i$ if $\sigma_i > \sigma_j$ (otherwise, $\sigma' = \sigma$). A *sorting (unsorting) adjacent transposition at position* i is one with $j = i + 1$.

The *weak Bruhat order* $(S_n, \leq_{wB})$ on the permutations S_n of length n relates two permutations $\tau \leq_{wB} \sigma$ if τ can be obtained from σ by a sequence of sorting adjacent transpositions (see Fig. 6).

There is a simple interpretation of the weak Bruhat order.

Lemma 4 *All inversions of τ are inversions of σ if and only if $\sigma \geq_{wB} \tau$.* □

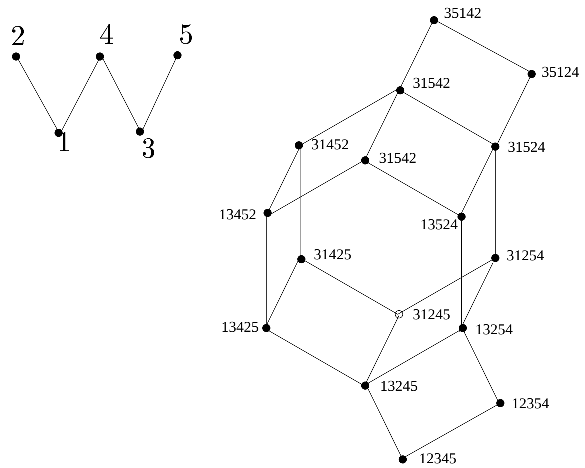

Figure 6: The order P_{35142} and its linear extensions in weak Bruhat order.

Proof. If $\sigma \geq_{wB} \tau$ then τ is obtained from σ by sorting transpositions; but sorting transpositions can only reduce the number of inversions. On the other hand, starting with σ we bubble the element τ_1 to the first position by adjacent transpositions. All these adjacent transpositions must be sorting. Otherwise, σ contains an inversion that is not in τ. Next, continue with τ_2, etc. □

It is also easy to identify linear extensions in the weak Bruhat order.

Lemma 5 *The set of linear extensions of P_π is the ideal of the weak Bruhat order generated by π, i.e., σ is a linear extension of P_π iff $\sigma \leq_{wB} \pi$.*

Proof. A moments thought shows that σ is a linear extension of P_π iff $I(\sigma) \subseteq I(\pi)$. □

This is a special case of a more general relation between linear extensions of orders and convex subsets of the weak Bruhat order studied in more detail by Björner and Wachs [BW91a] and Reuter [Reu96].

Note that in the width two case the relation $<_p$ and the weak Bruhat relation on the corresponding permutations coincide. It is thus tempting to suspect that a suitable generalization of Fact 4 will hold with $\leq_p$ replaced by $\leq_{wB}$. This is not quite the case as the following example shows: $312 \leq_{wB} 321$ and $\mathrm{sort}(312, 1, 2) = 132$ and $\mathrm{sort}(321, 1, 2) = 231$ but $132 \not\leq_{wB} 231$.

Similarly to the definitions of Section 3 we define for a permutation $\sigma \in S_n$ the *path at level l* by

$$\mathrm{sequence}_l(\sigma, i) = \begin{cases} 0 & \text{if } \sigma_i < l, \\ 1 & \text{if } \sigma_i \geq l, \end{cases} \quad \text{for } 1 \leq i \leq n.$$

Hence, $\mathrm{sequence}_l(\sigma, i)$ indicates whether σ reaches level l at position i (see Fig. 7). Represent sequences of ones and zeros by graphical paths as shown in Fig. 7. A one (zero) in the sequence at level l corresponds to an upward (downward) step in the path of level l.

Again we define

$$\mathrm{ones}_l(\sigma, k) = \sum_{1 \leq i \leq k} \mathrm{sequence}_l(\sigma, i), \quad \text{for } 1 \leq k \leq n.$$

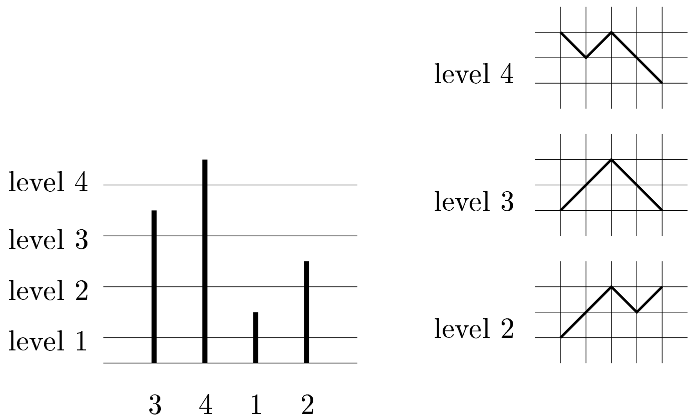

Figure 7: Path diagrams of permutation 3412

Note that $\mathrm{ones}_l(\sigma, n) = n - l + 1$ for every $\sigma \in S_n$. The *path order* $(S_n, \leq_p)$ on S_n relates two permutations $\tau \leq_p \sigma$ if $\mathrm{ones}_l(\tau, i) \leq \mathrm{ones}_l(\sigma, i)$ for all $1 \leq l, i \leq n$. This means that the paths of σ are above that of τ *on all levels* (see Figs. 7 and 4).

A sorting adjacent transposition at position k corresponds to an upward flip in the paths of all levels, i.e., a '10' configuration is converted to a '01' configuration (all other configurations are unchanged). Similarly, an unsorting adjacent transposition changes a '01' configuration to a '10' configuration (see Fig. 5).

From Figs. 4 and 5 it is clear that if one path is above the other this remains so after both paths are flipped upwards or downwards at the same position. Since an adjacent transposition corresponds to flips at the same position on all levels we have

Lemma 6 *Let σ' and τ' be obtained from σ and τ by an adjacent transposition at the same position. If $\sigma \geq_p \tau$ then $\sigma' \geq_p \tau'$.* □

The relation $\sigma \geq_{wB} \tau$ means that the paths of τ can be obtained from that of σ by downward flips (i.e., sorting adjacent transpositions).

Lemma 7 *If $\sigma \geq_{wB} \tau$ then $\sigma \geq_p \tau$.* □

Note that the converse is not true. For example, $231 \geq_p 132$ but $231 \not\geq_{wB} 132$ since the interchange of 2 and 3 cannot be obtained by adjacent transpositions that are sorting. But it is possible to characterize the relation $\geq_p$ in terms of transpositions.

The *Bruhat order* $(S_n, \leq_B)$ on the permutations of S_n relates two permutations $\tau \leq_{wB} \sigma$ if τ can be obtained from σ by a sequence of sorting (not necessarily adjacent) transpositions.

Lemma 8 *For two permutations σ and τ, $\sigma \geq_B \tau$ if and only if $\sigma \geq_p \tau$.*

Proof. Note that if σ' is obtained from σ by a sorting transposition then $\mathrm{ones}_l(\sigma, k) \geq \mathrm{ones}_l(\sigma', k)$ for all $1 \leq l, k \leq n$. Hence, $\sigma \geq_B \tau$ implies $\sigma \geq_p \tau$.

To prove the converse we find a sorting transposition that changes σ to σ' such that $\sigma' \geq_p \tau$. It is easily seen that such a transposition reduces the number of inversions. To be more precise, if $\sigma_i > \sigma_j$ with $i < j$ are interchanged then the number of inversions reduces exactly by the cardinality of the set $\{k \mid i < k \leq j \text{ and } \sigma_j \leq \sigma_k < \sigma_i\}$. Hence, we are done by induction on the number of inversions.

To find the searched sorting transposition let i be minimal such that $\sigma_i \neq \tau_i$ (then $\sigma_i > \tau_i$). Such i exists if $\sigma \neq \tau$. Next let $j > i$ be minimal with $\tau_i \leq \sigma_j < \sigma_i$. Such j exists: since $\sigma_k = \tau_k$ for all $k < i$, there must be a $k > i$ with $\sigma_k = \tau_i < \sigma_i$ (of course, σ_j need not be this σ_k). Then σ' is obtained from σ by a sorting transposition of σ at positions i and j. It remains to show that $\sigma' \geq_p \tau$.

We define the function

$$\phi_l(k) = \text{ones}_l(\sigma, k) - \text{ones}_l(\tau, k), \quad 1 \leq l, k \leq n,$$

which counts the overhead of ones in σ over that in τ in the first k positions and which is nonnegative for all l and k if and only if $\sigma \geq_p \tau$.

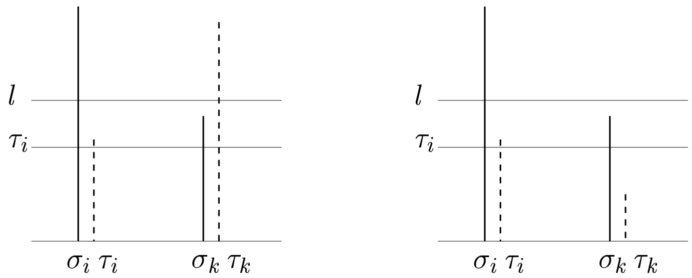

Figure 8: Selectively decrease (left) or increase (right) $\phi_l(k-1)$.

Assume $\tau_i < l \leq \sigma_i$. Then we show $\phi_l(k) > \phi_{\tau_i}(k) \geq 0$, for $i \leq k < j$. This is true for $k = i$ since $\phi_l(i) = 1$ and $\phi_{\tau_i}(i) = 0$ (recall that σ and τ coincide in the first $i-1$ positions). Now assume $\phi_l(k-1) > \phi_{\tau_i}(k-1)$ and observe that $\phi_l(k)$ is obtained from $\phi_l(k-1)$ by an increase of -1, 0, or 1 depending on σ_k and τ_k. But to selectively decrease $\phi_l(k-1)$ without decreasing $\phi_{\tau_i}(k-1)$ (see left part of Fig. 8) or to selectively increase $\phi_{\tau_i}(k-1)$ without increasing $\phi_l(k-1)$ (see right part of Fig. 8) we need $\tau_i \leq \sigma_k < l \leq \sigma_i$, which for $k < j$ is in contradiction to the definition of j.

When we replace σ by σ' then $\phi_l(k)$ is reduced by 1 only for $\tau_i \leq \sigma_j < l \leq \sigma_i$ and $i \leq k < j$. By the statememt of the preceding paragraph, $\phi(k)$ is strictly positive for such k before the replacement and hence $\phi_l(k)$ remains nonnegative after it. □

We will consider transpositions that are under the control of a permutation π. We say that an *unsorting adjacent transposition of a permutation σ at position i under π* is a permutation $\sigma' = \text{unsort}_\pi(\sigma, i) \in S_n$ which is the same as σ save $\sigma'_i = \sigma_{i+1}$ and $\sigma'_{i+1} = \sigma_i$ if $\sigma_{i+1} > \sigma_i$ and $\text{pos}(\sigma_{i+1}, \pi) < \text{pos}(\sigma_i, \pi)$ (i.e., $\sigma_{i+1} \text{ bef}_\pi\, \sigma_i$), (otherwise, $\sigma' = \sigma$).

The central observation that is useful in showing the monotonicity of the linear extensions of some permutation π under adjacent transpositions allowed under π is stated in the following lemma.

Lemma 9 *Suppose $\pi \geq_{wB} \sigma, \tau$ and $\sigma \geq_p \tau$ for permutations $\pi, \sigma, \tau \in S_n$ and suppose that* $\text{unsort}_\pi(\sigma, k) = \sigma$ *whereas* $\text{unsort}_\pi(\tau, k) = \tau' \neq \tau$. *Then $\sigma \geq_p \tau'$.*

Proof. Assume that the consequence of the lemma is not true, i.e., that $\sigma \not\geq_p \tau'$. Since $\sigma \geq_p \tau$ there must be a level l with $\text{ones}_l(\sigma, k-1) = \text{ones}_l(\tau, k-1) = \text{ones}_l(\tau', k-1)$ and $\text{sequence}_l(\sigma, k) = \text{sequence}_l(\tau, k) = 0$ but $\text{sequence}_l(\tau', k) = 1$ (a path of τ on some level is no longer under the path of σ as τ makes an upward flip to τ', see Fig. 9).

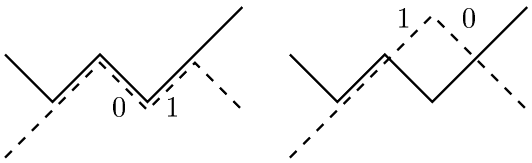

Figure 9: τ is below σ, τ' is above σ.

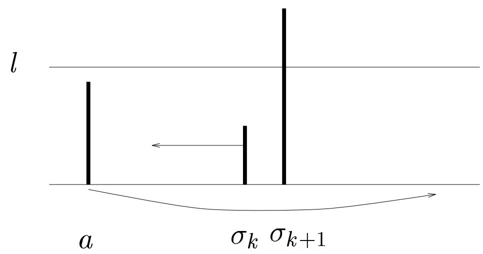

Figure 10: Position of a in σ.

Hence, we have $\sigma_k, \tau_k < l$ and $\sigma_{k+1}, \tau_{k+1} \geq l$ as well as $\sigma_k \operatorname{bef}_\pi \sigma_{k+1}$ and $\tau_{k+1} \operatorname{bef}_\pi \tau_k$. These facts as well as the fact that $\text{ones}_i(\sigma, j) \geq \text{ones}_j(\tau, j)$, $1 \leq i, j \leq n$, will be used throughout the proof.

We also know that either σ_k differs from τ_k or σ_{k+1} differs from τ_{k+1}. Furthermore, since $\sigma_k \operatorname{bef}_\pi \sigma_{k+1}$, either $\text{pos}(\sigma_k, \tau) < k$ or $\text{pos}(\sigma_{k+1}, \tau) > k+1$. We only have to consider the first case. The second case is reduced to the first when the transformation $\gamma \to \widehat{\gamma}$ with $\widehat{\gamma}_i = n + 1 - \gamma_{n-i}$ is applied to π, σ, and τ, since the relations $\leq_{wB}$ and $\leq_p$ are preserved under this transformation and σ_k, σ_{k+1} exchange their roles. Thus, in the sequel we assume $\text{pos}(\sigma_k, \tau) < k$.

We first show the following statement (see Fig. 10).

> *There exists an a with* $\text{pos}(a, \sigma) < k$, $\text{pos}(a, \tau) \geq k$, *and* $\sigma_k < a < l$. *Moreover,* $\tau_{k+1} \operatorname{bef}_\pi \sigma_k$.

Suppose that for all $a > \sigma_k$ with $\text{pos}(a, \sigma) < k$ we have $\text{pos}(a, \tau) < k$, then, since additionally $\text{pos}(\sigma_k, \tau) < k$, on level σ_k we have $\text{ones}_{\sigma_k}(\tau, k-1) > \text{ones}_{\sigma_k}(\sigma, k-1)$, a contradiction. Since $\text{ones}_{\sigma_k}(\sigma, k-1) \geq \text{ones}_{\sigma_k}(\tau, k-1) \geq 1$ there is also at least one a as described above. Such an a also fulfills $a \operatorname{bef}_\sigma \sigma_k$ (so, by Lemma 4, $a \operatorname{bef}_\pi \sigma_k$) and $a = \tau_k$ or $\tau_{k+1} \operatorname{bef}_\tau a$ (so $\tau_{k+1} \operatorname{bef}_\pi a$) implying that $\tau_{k+1} \operatorname{bef}_\pi \sigma_k$.

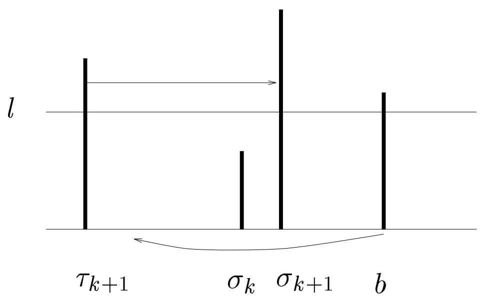

Figure 11: Position of b in σ.

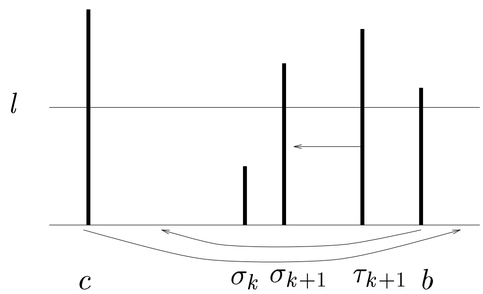

Figure 12: Position of b and c in σ.

Next we show a complementary statement (see Fig. 11).

There exists a b with $\mathrm{pos}(b,\sigma) \geq k+1$, $\mathrm{pos}(b,\tau) < k$, *and* $\sigma_{k+1} \geq b \geq l$.

We consider two cases. First, assume $\tau_{k+1} \leq \sigma_{k+1}$ (see Fig. 11). If $\mathrm{pos}(\tau_{k+1},\sigma) \geq k+1$, we have $\sigma_{k+1} \,\mathrm{bef}_\pi\, \tau_{k+1}$ (Lemma 4) which together with the first statement above gives $\sigma_{k+1} \,\mathrm{bef}_\pi\, \sigma_k$, a contradiction. Hence, $\mathrm{pos}(\tau_{k+1},\sigma) < k$. But $\mathrm{pos}(\tau_{k+1},\tau) = k+1$ and $\mathrm{ones}_l(\sigma,k-1) = \mathrm{ones}_l(\tau,k-1)$ and so there must be at least one $b \geq l$ more with $\mathrm{pos}(b,\sigma) \geq k+1$ and $\mathrm{pos}(b,\tau) < k$ than there are $x \geq l$ with $\mathrm{pos}(x,\sigma) < k$ and $\mathrm{pos}(x,\tau) \geq k+1$. We know that if $\tau_{k+1} < n$, $\mathrm{ones}_{\tau_{k+1}+1}(\sigma,k-1) \geq \mathrm{ones}_{\tau_{k+1}+1}(\tau,k-1)$. So it is impossible that for all such b's $b > \tau_{k+1}$. Hence, there is one with $\sigma_{k+1} \geq \tau_{k+1} \geq b$.

Now assume $\tau_{k+1} > \sigma_{k+1}$. In a first step we find a $c \geq \tau_{k+1}$ with $\mathrm{pos}(c,\sigma) < k$ and $\mathrm{pos}(c,\tau) \geq k+1$ (see Fig. 12). If $\mathrm{pos}(\tau_{k+1},\sigma) < k$ we simply set $c = \tau_{k+1}$. Otherwise, $\mathrm{pos}(\tau_{k+1},\sigma) > k+1$ (note here that $\tau_{k+1} \neq \sigma_k$, $\tau_{k+1} \neq \sigma_{k+1}$). Since $\mathrm{pos}(\tau_{k+1},\tau) = k+1$ but $\mathrm{ones}_i(\sigma,k+1) \geq \mathrm{ones}_i(\tau,k+1)$, for all $i \leq \tau_{k+1}$, there must be a $c > \tau_{k+1}$ as searched for. Similarly as in the preceding paragraph, since $\mathrm{ones}_l(\sigma,k-1) = \mathrm{ones}_l(\tau,k-1)$, there must be a $b \geq l$ with $\mathrm{pos}(b,\sigma) \geq k+1$, $\mathrm{pos}(b,\tau) < k$, to compensate for such c. Of course, $\mathrm{ones}_i(\sigma,k+1) \geq \mathrm{ones}_i(\tau,k+1)$ on levels i with $\sigma_{k+1} < i \leq \tau_{k+1}$, and we may assume that $b \leq \sigma_{k+1}$.

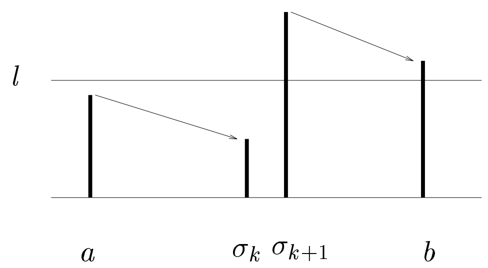

Figure 13: Position of a and b in σ.

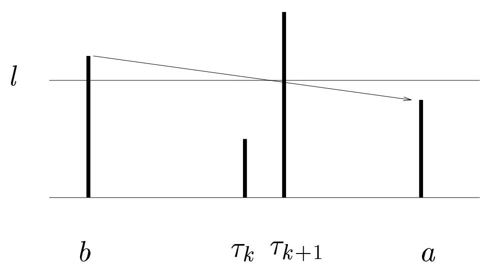

Figure 14: Position of a and b in τ.

With the two statements we can now prove the lemma. The first statement implies $a \,\mathrm{bef}_\sigma\, \sigma_k$. The second one implies $\sigma_{k+1} = b$ or $\sigma_{k+1} \,\mathrm{bef}_\sigma\, b$ (see Fig. 13). Both imply $b \,\mathrm{bef}_\tau\, a$ (see Fig. 14). Consequently, by Lemma 4, $\sigma_{k+1} \,\mathrm{bef}_\pi\, \sigma_k$, a contradiction. □

It is now simple to proof the montonicity of adjacent transpositions.

Theorem 10 *Suppose $\pi \geq_{wB} \sigma, \tau$ and $\sigma \geq_p \tau$. If σ', τ' are obtained from σ, τ by a (sorting or unsorting) adjacent transposition at the same position under π then $\sigma' \geq \tau'$.*

Proof. Lemma 6 can be applied in the case of a sorting transposition. In the case of an unsorting tranposition it can also be applied if $\sigma \neq \sigma'$ and $\tau \neq \tau'$. If $\tau \neq \tau'$ and $\sigma = \sigma'$ use Lemma 9. If $\tau = \tau'$ and $\sigma \neq \sigma'$, by Lemma 7, $\sigma' \geq_{wB} \sigma$ implies $\sigma' \geq_p \sigma$ and the theorem follows. □

5 Remarks on the coupling time

The monotone Monte Carlo algorithm describeded after Theorem 1 together with Theorem 10 allows to sample uniformly from the linear extensions of P_π by observing traces of (id, π) from the past until they couple. The function RandomMap consists of a randomly chosen position where to make an adjacent transposition and a choice whether to sort or unsort.

The problem remaining is to estimate the number of Markov steps needed for the coupling. But note that this is only of theoretical interest, since the algorithm stops as soon as coupling occurs. In contrast, estimations for the mixing time of runs of Markov chains are an essential criterion when to stop a corresponding (forward) algorithm if one wants to obtain with high probability a good approximation of the stationary distribution.

If we set π to the permutation $n, n-1, \ldots, 1$ (i.e., all permutations are allowed as linear extensions) an analysis similar to that leading to Theorem 3 gives a coupling time of $O(n^3 \log n)$ (the only difference is a slightly larger constant since the area between the paths of *all* levels must be considered).

We are not able to give a similar sharp bound for the coupling time in the case where π is set to an arbitrary permutation. In Fig. 15 we tried to fit polynomials of degree three to the outcome of experiments comparing coupling in the whole space of permutations (un-

Length	unrestricted	weak Bruhat
20	7071	3687
30	30129	12757
40	73710	31224
50	153020	61374
60	290506	120568
80	753161	263692
100	1575570	514710

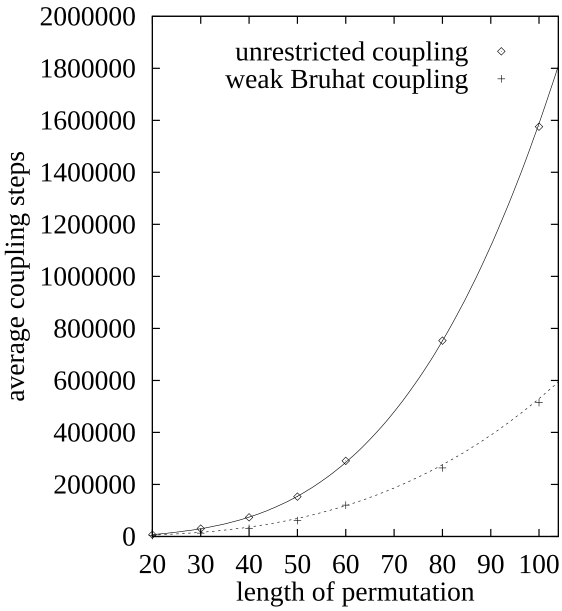

Figure 15: Average number of unrestricted and weak Bruhat coupling steps.

restricted) to coupling restricted to linear extensions of some permutation (weak Bruhat). The method was to generate 10 random permutations (of course, by backward coupling) of a given length and to couple them with the identity permutation. We think that these experimental results give good reasons to suggest the following conjecture.

Conjecture 1 *The expected coupling time of the above chain in the weak Bruhat ideal of some permutation π of length n is $O(n^3 \log n)$.*

In [BW91b] it is shown that the counting problem for linear extensions is $\#P$-complete. On the other hand, details are given of an algorithm for approximate counting of linear extensions. This algorithm uses the approximate uniform generation of linear extensions as a subroutine. In the case of two-dimensional orders monotone coupling certainly yields a faster approximative count. However, we don't know the complexity of the exact counting problem.

Problem 1 *Prove #P-completeness of the counting problem for linear extensions of two-dimensional orders or show that they can be counted efficiently.*

Nevertheless, a polynomial bound on the coupling time of the process can be given. A general result of Wilson and Propp (Section 5 from [PW95]) relates expected coupling time to expected mixing time. Combined with the estimate of Karzanov and Khachiyan [KK91] for the mixing time of the Markov chain we obtain a bound of $O(n^6 \log^2 n)$ for the coupling time.

The proof for the mixing rate in the work of Karzanov and Khachiyan is based on a bound for the conductance of the transition graph of the Markov chain. This transition graph is the cover graph of the weak Bruhat ordering of the linear extensions as shown in Fig. 6. Define the *conductance* of a graph as

$$c = \min_{X \subset V} \frac{C(X, \overline{X})}{\min(|X|, |\overline{X}|)}$$

where $C(X, \overline{X})$ is the number of edges from X to its complement $\overline{X}$. Note that this definition differs from the usual one by a factor of $1/(2n-2)$. For a discussion of the relation of conductance to mixing rate see, e.g., Kannan [Kan94]. Using geometric arguments about convex bodies Karzanov and Khachiyan prove $c \geq \frac{1}{n\sqrt{2n}}$. It is believed that the conductance of linear extension graphs is at least $2/n$ (which would improve the bound for the coupling time to $O(n^5 \log^2 n)$). Known examples where $2/n$ is attained are disjoint unions of a singleton element x, an odd length chain C, and an arbitrary order Q, where X is the set of linear extensions with x preceding the middle element of C. Beside the consequence for the mixing rate we consider this an important and nice conjecture:

Conjecture 2 *$2/n$ is a lower bound for the conductance of the graph of linear extensions of every n element order.*

Let x, y be incomparable elements of P the (x, y)*-cut* of the linear extension graph of P is the partition $X, \overline{X}$ with $X = \{L : x \operatorname{bef}_L y\}$. To prove a slightly weaker bound of $1/(n-1)$ it would suffice to show that the worst cuts with respect to conductance are (x, y)-cuts as is the case in the above examples with $c = \frac{2}{n}$.

Lemma 11 *The conductance of (x, y)-cuts is at least $1/(n-1)$.*

Proof. Consider the sample space of linear extensions of P with the uniform distribution. For $1 \leq i \leq n-2$ let a_i be the probability that $x \operatorname{bef} y$ in L with $(i-1)$ elements in between and let b_i be the probability that $y \operatorname{bef} x$ with $(i-1)$ elements in between. Kahn and Saks in their seminal paper about balancing pairs [KS84] establish relations implying:

$$(1) \quad \sum_{i=1}^{n-1} a_i + \sum_{i=1}^{n-1} b_i = 1.$$

(2) $a_1 = b_1$.

(3) $a_1 + b_1 \geq a_2 + b_2$.

(4) If $a_1 \geq a_2$ then $a_i \geq a_{i+1}$ for all i and if $b_1 \geq b_2$ then $b_i \geq b_{i+1}$ for all i.

Property (4) follows from the log-concavity of the sequences which is proved using the Alexandrov/Fenchel inequalities for mixed volumes.

Assume w.l.o.g. that $\sum a_i \geq \sum b_i$ and note that the conductance of the (x, y)-cut is $b_1/\sum b_i$. If $b_1 \geq b_2$ then by (4) $\sum b_i \leq (n-1)b_1$, hence, $c \geq 1/(n-1)$. If $b_1 \leq b_2$ then by (3) $a_1 \geq a_2$ wich implies $\sum a_i \leq (n-1)a_1$, hence $c = b_1/\sum b_i \geq a_1/\sum a_i \geq 1/(n-1)$. $\square$

References

[BFT95] G. Brightwell, S. Felsner, and T. Trotter. Balancing pairs and the cross product conjecture. *Order*, 12:321–335, 1995.

[BW91a] Björner and M. Wachs. Permutation statistics and linear extensions of posets. *J. of Comb. Th. (A)*, 58:85–114, 1991.

[BW91b] G. Brightwell and Winkler. Counting linear extensions. *Order*, 8:225–242, 1991.

[DFK89] M. Dyer, A. Frieze, and R. Kannan. A random polynomial time algorithm for approximating the volume of convex bodies. In *21th Annual ACM STOC*, pages 375–381, 1989.

[DM41] B. Dushnik and E. Miller. Partially ordered sets. *Am. J. Math.*, 63:600–610, 1941.

[Kan94] R. Kannan. Markov chains and polynomial time algorithms. In *Proc. 35th Ann. IEEE Symp. on Foundations of Computer Science*, pages 656–671, 1994.

[KK91] A. Karzanov and L. Khachiyan. On the conductance of order markov chains. *Order*, 8:7–15, 1991.

[KS84] Kahn and Saks. Balancing poset extensions. *ORDER*, 1984.

[Lov92] L. Lovász. How to compute the volume? *Jber. d. Dt. Math.-Verein. (Jubiläumstagung 1990)*, pages 138–151, 1992.

[PW95] J.G. Propp and D.W. Wilson. Exact sampling with coupled markov chains and applications to statistical mechanics. *Random Structures and Algorithms*, 1995.

[Reu96] K. Reuter. Linear extension of posets as abstract convex sets. *Preprint*, 1996.

[Wil95] D. W. Wilson. personal communication, 1995.