# On Friedmann's subexponential lower bound for Zadeh's pivot rule*‡

Yann Disser[1,2] and Alexander V. Hopp[1,2]

[1]*Graduate School of Computational Engineering, TU Darmstadt, Germany*
[2]*Department of Mathematics, TU Darmstadt, Germany*

July 12, 2018

**Abstract**

The question whether the Simplex method admits a polynomial time pivot rule remains one of the most important open questions in discrete optimization. Zadeh's pivot rule had long been a promising candidate, before Friedmann (IPCO, 2011) presented a subexponential instance, based on a close relation to policy iteration algorithms for Markov decision processes (MDPs). We investigate Friedmann's lower bound example and exhibit three flaws in his analysis: We show that (a) the initial policy for the policy iteration does not produce the required occurrence records and improving switches, (b) the specification of occurrence records is not entirely accurate, and (c) the sequence of improving switches described by Friedmann does not consistently follow Zadeh's pivot rule. In this paper, we resolve each of these issues. While the first two issues require only minor changes to the specifications of the initial policy and the occurrence records, the third issue requires a significantly more sophisticated ordering and associated tie-breaking rule that are in accordance with the LEAST-ENTERED pivot rule. Most importantly, our changes do not affect the macroscopic structure of Friedmann's MDP, and thus we are able to retain his original result.

**Acknowledgments**

The authors are very grateful to Oliver Friedmann for helpful comments and discussions, as well as support in using his implementation of the original construction to verify our findings.

## 1 Introduction

The Simplex method, originally proposed by Dantzig in 1947 [2], is one of the most important algorithms to solve linear programs in practice. At its core, it operates by maintaining a subset of basis variables while restricting non-basis variables to trivial values, and repeatedly replacing a basis variable according to a fixed *pivot rule* until the objective function value can no longer be improved. Exponential worst-case instances have been devised for many natural pivot rules (see, for example, [1, 4, 6, 7]), and the question whether a polynomial time pivot rule exists remains one of the most important open problems in optimization theory.

Zadeh's LEAST-ENTERED pivot rule [10] was designed to avoid the exponential behavior on known worst-case instances for other pivot rules. The rule is *memorizing* in that it selects a variable to enter the basis that improves the objective function and has previously been selected least often among all improving variables. Indeed, for more than thirty years, Zadeh's rule defied all attempts to construct superpolynomial instances, and it seemed like a promising candidate for a polynomial pivot rule.

*This work is supported by the 'Excellence Initiative' of the German Federal and State Governments and the Graduate School CE at TU Darmstadt.
†Digitally published at the University and State Library Darmstadt
URN: urn:nbn:de:tuda-tuprints-75572
URI: `http://tuprints.ulb.tu-darmstadt.de/id/eprint/7557`
‡Second revised version, first version available at `http://tuprints.ulb.tu-darmstadt.de/id/eprint/6873`

It was a breakthrough when Friedmann eventually presented the first superpolynomial lower bound for Zadeh's pivot rule [3]. His construction uses a connection between the Simplex Algorithm and Howard's Policy Iteration Algorithm [5] for computing optimal policies in Markov decision processes (MDPs). For a given $n \in \mathbb{N}$, Friedmann's construction consists of an MDP of size $\mathcal{O}(n^2)$, an initial policy, and an ordering of the improving switches obeying the LEAST-ENTERED pivot rule. This ordering results in an exponential number of iterations when beginning with the initial policy and repeatedly making improving switches in the specified order.The construction translates into a linear program of the same asymptotic size for which the Simplex Algorithm with Zadeh's pivot rule needs $\Omega(2^n)$ steps. Since the input size is $\mathcal{O}(n^2)$, this in turn results in a superpolynomial lower bound. Recently, an exponential lower bound for Zadeh's pivot rule was found for AUSOs [9], but it is not clear whether this construction can be realized as a linear program.

***Our contribution.*** In this paper, we expose different flaws in Friedmann's lower bound construction and present adaptations to eliminate them. We first show that the chosen initial policy does not produce the claimed occurrence records and improving switches, and propose a modified initial policy that leads to the desired behavior. Second, we observe that the given formula describing the occurrence records (that count the number of times an improving switch was made) is inaccurate, and provide a (small) correction that does not disturb the overall argument. Note that these two modifications are necessary but relatively minor.

Finally, we exhibit a significant problem with the order in which the improving switches are applied in [3]. More precisely, we show that this order does not consistently obey Zadeh's pivot rule, and, in fact, that no consistent ordering exists that updates the MDP level by level in each phase according to a fixed order. This not only rules out Friedmann's ordering, but shows that a fundamentally different approach to ordering improving switches is needed. To amend this issue, we show the existence of an ordering and a tie-breaking rule compatible with the LEAST-ENTERED rule, such that applying improving switches according to the ordering still proceeds along the same macroscopic phases as intended by Friedmann. In this way, we are able to quantitatively retain Friedmann's superpolynomial lower bound on the number of iterations needed by Zadeh's LEAST-ENTERED pivot rule.

***Outline.*** Throughout this paper, we assume some basic familiarity with the construction given in [3] and Markov decision processes in general. We review the most important aspects and notation of [3] in Section 2, and, for convenience, provide a copy of some tables that we rely on in Appendix A. Section 3 treats issues with the initial policy and our adaptation to address them. In Section 4, we correct an inaccuracy concerning some of the occurrence records given in [3]. The main part of this paper is Section 5, where we show that the sequence of improving switches can be reordered such that the order obeys the LEAST-ENTERED rule. Finally, we summarize our findings in Section 6.

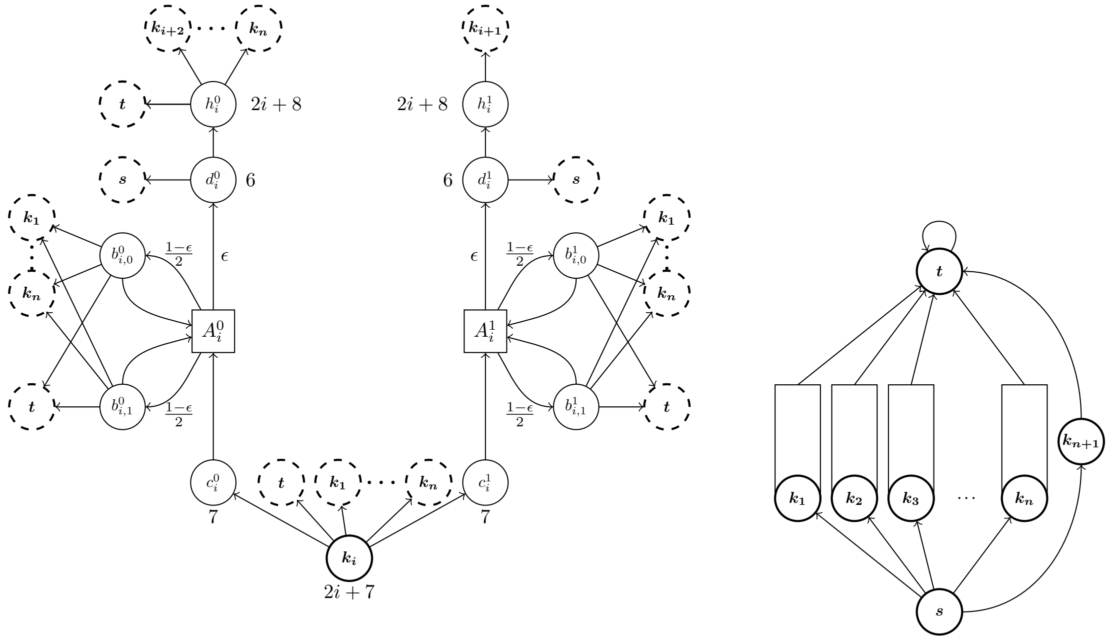
# 2 Preliminaries

## 2.1 Friedmann's lower bound construction

In [3], Friedmann uses the connection between the Simplex Algorithm for linear programming and the Policy Iteration Algorithm for obtaining optimal policies in Markov decision processes. Similarly, we also restrict our discussion to policy iteration for MDPs, with the understanding that results carry over to the Simplex Algorithm. We assume knowledge of MDPs and the connection to the Simplex Algorithm and refer to [8] for more information.

We first establish some notation. Given an MDP, a player-controlled edge $e = (u, v)$ and a policy $\sigma$, we say that $v$ is the *target* of $u$ or $u$ *points to* $v$ if $\sigma(u) = v$. If $\sigma(u) \neq v$, we *switch* $e$ or *switch* $u$ *to* $v$ when we apply the switch $(u, v)$ in $\sigma$. For a policy $\sigma$ and an improving switch $e$ for $\sigma$, we denote the policy obtained by applying the switch $e$ in $\sigma$ by $\sigma[e]$.

Let $n \in \mathbb{N}, n \neq 0$ be fixed. Friedmann's lower bound construction emulates an $n$-bit binary counter by a Markov decision process $G_n$. For every $n$-digit binary number $b = (b_n, \ldots, b_1)$, there is a unique policy $\sigma_b$ for $G_n$ representing $b$. Note we denote the least significant bit by $b_1$, that is, $b = \sum_{i=1}^{n} b_i 2^{i-1}$. The Markov decision process $G_n$ is constructed such that applying the Policy Iteration Algorithm using the LEAST-ENTERED pivot rule enumerates all policies $\sigma_0, \sigma_1, \ldots, \sigma_{2^n-1}$. Due to the pivot rule, the algorithm always chooses an improving switch that was chosen least often until now. More specifically, an *occurrence record* $\phi$ is maintained, and, in every step, a switch minimizing $\phi$ is chosen. The rule does however not




(a) Circular shaped vertices are player-controlled, squares are randomization vertices. Bold vertices are global in that they can be reached from other levels. Dashed vertices do not belong to level $i$. Numbers on edges show the probability of taking this edge, numbers below or next to vertices show the exponent of $(-N)$ of the rewards of edges leaving this vertex. Whenever there is no number, the rewards are 0.

(b) The entry vertices are all connected to the vertices $s$ and $t$. Connections between levels and from the levels to $s$ are not shown here. An additional vertex $k_{n+1}$ is needed for technical reasons.

Figure 1: Level $i$ of $G_n$ (left) and coarse structure of $G_n$ (right).

determine *which* switch minimizing $\phi$ should be chosen if there are multiple candidates. Therefore, an explicit tie-breaking rule may be used in the construction. For an edge $e$ and a policy $\sigma$ for $G_n$ that is calculated during the application of the Policy Iteration Algorithm, we denote the occurrence record of $e$ at the time when $\sigma$ is reached by $\phi^\sigma(e)$. We denote the set of improving switches with respect to the policy $\sigma$ by $I_\sigma$.

For the remainder of this paper, we fix the following notation:

- The set of numbers that can be represented by $n$ bits is denoted by $\mathbb{B}_n$.
- Let $b \in \mathbb{B}_n$. For $i \in \{1, \ldots, n\}$, the $i$-th bit of the binary representation of $b$ is denoted by $b_i$. For $b \neq 0$, we denote the least significant bit of $b$ which is equal to 1 by $\ell(b)$, that is, $\ell(b) := \min\{i \in \{1, \ldots, n\} : b_i \neq 0\}$.
- The unique policy representing $b \in \mathbb{B}_n$ constructed in [3] is denoted by $\sigma_b$.

The process $G_n$ can be interpreted as a "fair alternating binary counter" as follows. Usually, when counting from 0 to $2^n - 1$ in binary, less significant bits are switched more often than more significant bits. As the LEAST-ENTERED pivot rule forces the algorithm to switch all bits equally often, the construction must ensure to operate correctly when all bits are switched equally often. This is achieved by representing every bit by *two* gadgets where only one *actively* represents the bit. The gadgets alternate in actively representing the bit.

The lower bound construction consists of $n$ structurally identical levels, where level $i$ represents the $i$-th bit. A parameter $N \geq 7n+1, N \in \mathbb{N}$ is used for defining the rewards and a parameter $\epsilon \in (0, N^{-(2n+11)})$ is used for defining the probabilities. The $i$-th level is shown in Figure 1(a), the coarse structure of the whole MDP in Figure 1(b).

A number $n_v$ below or next to the name of a vertex $v$ in Figure 1(a) denotes a reward of $(-N)^{n_v}$ associated with every edge leaving $v$. Other edges have a reward of 0. Let $\sigma$ be a policy and $v$ be a vertex.

The *value* $\mathrm{VAL}_\sigma(v)$ of $v$ with respect to $\sigma$ is the expected accumulated reward obtained by an infinite walk starting in $v$. The MDP is constructed such that all values are always finite.

Each level $i$ contains two gadgets attached to the *entry vertex* $k_i$. These gadgets are called *lanes*. We refer to the left lane as *lane 0* and to the right lane as *lane 1*. Lane $j \in \{0,1\}$ of level $i$ contains a randomization vertex $A_i^j$ and two attached cycles with vertices $b_{i,0}^j$ and $b_{i,1}^j$. These gadgets are called *bicycles*, and we identify the bicycle containing vertex $A_i^j$ with that vertex. For a bicycle $A_i^j$, the edges $(b_{i,0}^j, A_i^j), (b_{i,1}^j, A_i^j)$ are called *edges of the bicycle*. For a policy $\sigma$, the bicycle $A_i^j$ is said to be *closed (with respect to $\sigma$)* if and only if $\sigma(b_{i,0}^j) = \sigma(b_{i,1}^j) = A_i^j$. A bicycle that is not closed is called *open*.

The $i$-th level of $G_n$ corresponds to the $i$-th bit of the binary counter. Which of the two bicycles of level $i$ is actively representing the $i$-th bit depends on the setting of the $(i+1)$-th bit. When this bit is equal to 1, bicycle $1A_i^1$ is considered active. Otherwise, bicycle $A_i^0$ is considered active. The $i$-th bit is interpreted as equal to 1 if and only if the active bicycle in level $i$ is closed.

As initial policy, the MDP is provided the policy $\sigma^\star = \sigma_0$ representing 0. Then, a sequence of policies $\sigma_1, \sigma_2, \ldots, \sigma_{2^n-1}$ is enumerated by the Policy Iteration Algorithm using the LEAST-ENTERED pivot rule and an (implicit) tie-breaking rule. For $b \in \mathbb{B}_n, b \neq 0$, the goal is that the policy $\sigma_b$ representing $b$ fulfills the following invariants. These invariants also apply for level $n$ by setting $b_{n+1} \coloneqq 0$ and to $b = 0$ when substituting $k_{\ell(b)}$ with $t$.

1. Exactly the bicycles $A_i^j$ corresponding to bits $b_i = 1$ are closed.
2. For all other bicycles $A_i^j$, it holds that $\sigma_b(b_{i,0}^j) = \sigma_b(b_{i,1}^j) = k_{\ell(b)}$.
3. All entry vertices $k_i$ point to the lane containing the active bicycle if $b_i = 1$ and to $k_{\ell(b)}$ otherwise. Formally, $\sigma_b(k_i) = c_i^j, j = b_{i+1}$ if $b_i = 1$ and $\sigma_b(k_i) = k_{\ell(b)}$ if $b_i = 0$.
4. The vertex $s$ points to the entry vertex corresponding to the least significant set bit.
5. All vertices $h_i^0$ point to the entry vertex of the first level strictly after level $i+1$ corresponding to a bit equal to 1, that is, when $l \coloneqq \min\{j \in \{i+2, \ldots, n\} : b_j = 1\}$, we have $\sigma_b(h_i^0) = k_l$. If no such $l$ exists, $\sigma_b(h_i^0) = t$.
6. The vertex $d_i^j$ points to $h_i^j$ if and only if $b_{i+1} = j$ and to $s$ otherwise.

The Policy Iteration Algorithm is only allowed to switch one edge per iteration. However, the policy $\sigma_{b+1}$ cannot be reached from $\sigma_b$ by performing a single switch. Therefore, intermediate policies need to be introduced for the transition from $\sigma_b$ to $\sigma_{b+1}$. These intermediate policies are divided into six *phases*. In each phase, a different "task" is performed within the construction. We mention here that the following description of the phases partly differs from the informal description given in [3, Pages 8,9]. We explain in detail why our description is different in Section 5. Consider the policy $\sigma_b$ representing some $b \in \mathbb{B}_n$. Let $\ell' \coloneqq \ell(b+1)$.

1. In phase 1, switches inside of some bicycles are performed to keep the occurrence records of the bicycle edges as balanced as possible. For every open bicycle $A_i^j$, at least one of the two edges $(b_{i,0}^j, A_i^j), (b_{i,1}^j, A_i^j)$ is switched. Some inactive bicycles are allowed to switch both of these edges such that their occurrence record can "catch up" with the other edges. In the active bicycle of level $\ell'$, we also switch both edges, as this bicycle needs to be closed with respect to $\sigma_{b+1}$.
2. In phase 2, the new least significant set bit $b_{\ell'}$ is made accessible by the rest of the MDP. Thus, $k_{\ell'}$ is switched to $c_{\ell'}^j$, where $j = (b+1)_{\ell'+1}$ is the lane containing the active bicycle.
3. In phase 3, we perform the "resetting process". The entry vertices of all levels $i$ corresponding to bits with $(b+1)_i = 0$ switch to $k_{\ell'}$. The same is done for all vertices $b_{i,l}^j$ contained in inactive bicycles and all vertices $b_{i,l}^j$ corresponding to levels $i$ with $(b+1)_i = 0$. We discuss this phase in more detail in Section 5.
4. In phase 4, the vertices $h_i^0$ are updated according to the new least significant set bit.
5. In phase 5, we switch $s$ to the entry vertex corresponding to the new least significant set bit, i.e., to $k_{\ell'}$.
6. In phase 6, we update the vertices $d_i^j$ such that $h_i^0$ is the target of $d_i^0$ if and only if $(b+1)_{i+1} = 0$ and $h_i^1$ is the target of $d_i^1$ if and only if $(b+1)_{i+1} = 1$.

## 2.2 Notation related to binary counting

Let $b \in \mathbb{B}_n$. By binary counting, we refer to the process of enumerating the binary representations of all numbers $\tilde{b} \in \{0, 1, \ldots, b\}$ in their natural order. These numbers are used to determine how often and when specific edges of $G_n$ are improving switches and will be applied.

Intuitively, we are interested in *schemes* that we observe when counting from 0 to $b$ in binary, or, more formally, in the set of numbers that *match a scheme* with respect to the following definition.

**Definition 2.1** ([3]). A *scheme* is a set $S \subseteq \mathbb{N} \times \{0, 1\}$. We say that $b \in \mathbb{B}_n$ *matches* $S$ if $b_i = q$ for all $(i, q) \in S$. We define the *match set* $M(b, S) := \{\tilde{b} \in \{0, \ldots, b\} : \tilde{b}_i = q \quad \forall (i, q) \in S\}$ as the set of all numbers between 0 and $b$ that match the scheme $S$.

The next definition introduces the *flip set* with respect to a number $b$, an index $i$ and a scheme $S$. This is a subset of $M(b, S)$ that fixes the $i$ least significant bits in a specific way.

**Definition 2.2** ([3]). Let $b \in \mathbb{B}_n, i \in \{1, \ldots, n\}$ and $S$ be a scheme. We define the *flip set* corresponding to $b, i$ and $S$ as $F(b, i, S) := M(b, S \cup \{(i, 1)\} \cup \{(j, 0) : j \in \{1, \ldots, i-1\}\})$. The *flip number* is defined as $f(b, i, S) := |F(b, i, S)|$. For convenience, we set $F(b, i) := F(b, i, \emptyset)$ and $f(b, i) := f(b, i, \emptyset)$.

Finally, we define the *maximal flip number* with respect to a number $b$, an index $i$ and a scheme $S$. It is the largest number contained in $F(b, i, S)$ smaller than $b$ or 0 if $F(b, i, S) = \emptyset$.

**Definition 2.3** ([3]). Let $b \in \mathbb{B}_n, i \in \{1, \ldots, n\}$ and $S$ a scheme. The *maximal flip number* is defined as $g(b, i, S) := \max(\{0\} \cup \{\tilde{b} : \tilde{b} \in F(b, i, S)\})$.

## 2.3 Imported tables

We briefly describe and summarize the tables introduced in [3] that we use in this work. These tables can also be found in Appendix A.

The first table is [3, Table 2]. For $p \in \{1, \ldots, 6\}$, it defines when a policy $\sigma$ is considered to be a phase $p$ policy. As in [3], we say that a policy $\sigma$ is a phase $p$ policy if every vertex is mapped by $\sigma$ to a choice included in the respective cell of the table. Cells that contain more than one choice indicate that policies of the respective phase are allowed to match any of the choices. As we prove later, there is an issue concerning the side conditions of phase 3.

The next table is [3, Table 3]. For a phase $p$ policy $\sigma$, this table shows subsets $L^p_\sigma$ and supersets $U^p_\sigma$ of the set $I_\sigma$ of improving switches. In general, this table does not show the complete sets of improving switches. We verified that the switches given in the sets $L^1_\sigma$ to $L^5_\sigma$ are in fact improving switches and discuss an issue related to the set $L^6_\sigma$ later.

The last table we use is [3, Table 4]. For $b \in \mathbb{B}_n$, this table contains the occurrence records $\phi^{\sigma_b}$ of the edges with respect to the unique policy representing the number $b$. Again, we found an issue regarding the complicated conditions that we discuss in Section 4.

Other than correcting these issues, we rely on [3, Tables 2,3,4].

# 3 Initial Policy

In this section, we discuss the initial policy $\sigma^\star$ used in [3]. We show that it contradicts several aspects of [3], and discuss how to replace $\sigma^\star$ such that the resulting issues are resolved.

On [3, Page 11], the following is stated regarding the initial policy: "*As designated initial policy $\sigma^\star$, we use $\sigma^\star(d^j_i) = h^j_i$ and $\sigma^\star(\_) = t$ for all other player 0 nodes with non-singular out-degree.*" This initial policy, however, is inconsistent with the sub- and supersets of improving switches given in [3, Table 3] and [3, Lemma 4].[1]

**Issue 3.1.** *The initial policy $\sigma^\star$ for the Markov decision process $G_n$ described in [3, Page 10] contradicts [3, Table 3] since $I_{\sigma^\star} \neq \{(b^j_{i,r}, A^j_i) \colon \sigma^*(b^j_{i,r}) \neq A^j_i\}$.*

Also, the following issue is caused by the initial policy $\sigma^\star$.

[1] This and all other proofs are deferred to Appendix B.

**Issue 3.2.** *When the Policy Iteration Algorithm is started with $\sigma^\star$, either*

1. *[3, Table 4] containing the occurrence records is incorrect for $b = 1$, or*
2. *[3, Table 3] containing the sub- and supersets of $I_{\sigma_b}$ is incorrect for $b = 1$.*

As a consequence, the initial policy $\sigma^\star$ needs to be changed. We propose to use the following policy instead.

**Definition 3.3** (New initial policy $\sigma^*$)**.** We define the following initial policy $\sigma^*$:

- $\sigma^*(d_i^0) := h_i^0$ for all $i \in \{1, \dots, n\}$.
- $\sigma^*(d_i^1) := s$ for all $i \in \{1, \dots, n\}$.
- $\sigma^*(\_) := t$ for all other player-controlled vertices with non-singular out-degree.

This new initial policy is visualized in Figure 2. Note that this policy also represents the number $0$ and [3, Lemma 1] holds for the policy $\sigma^*$. This policy indeed resolves both issues.

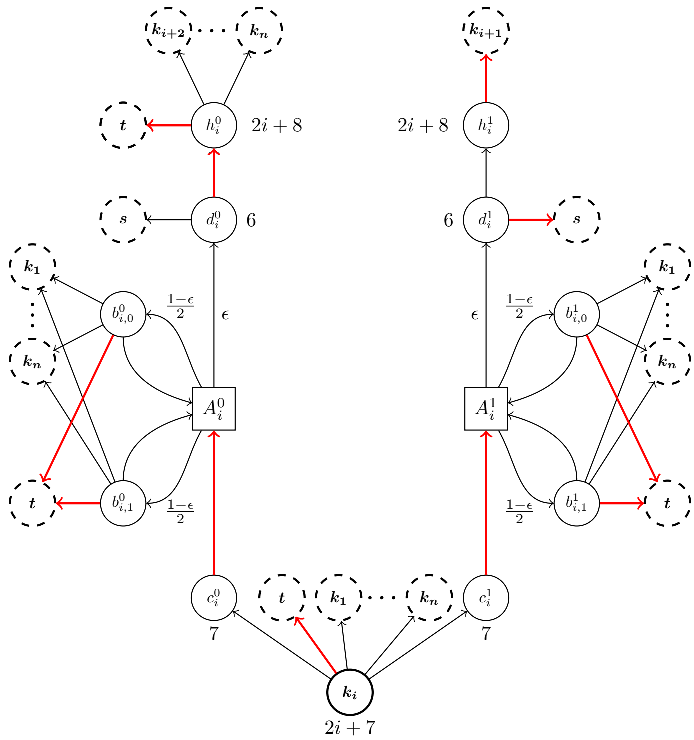



Figure 2: A level of the alternative initial policy $\sigma^*$. Thick red edges correspond to edges of $\sigma^*$.

**Lemma 3.4.** *The set of improving switches for $\sigma^*$ is $I_{\sigma^*} = \{(b_{i,r}^j, A_i^j) : \sigma^*(b_{i,r}^j) \neq A_i^j\}$.*

**Lemma 3.5.** *Starting the Policy Iteration Algorithm with the initial policy $\sigma^*$ avoids Issue 3.2, that is, it does not contradict [3, Tables 3,4] for $b = 1$.*

# 4 Occurrence Records

In this section, we discuss an issue related to the occurrence records of the bicycles as specified in [3, Table 4]. For a fixed $b \in \mathbb{B}_n$ and a fixed bicycle $A_i^j$, we define $g := g(b, i, \{(i+1, j)\})$, $z := b - g - 2^{i-1}$

and $\phi^{\sigma_b}(A_i^j) := \phi^{\sigma_b}(b_{i,0}^j, A_i^j) + \phi^{\sigma_b}(b_{i,1}^j, A_i^j)$. Using these abbreviations, the following is stated regarding the occurrence records:

$$\left|\phi^{\sigma_b}(b_{i,0}^j, A_i^j) - \phi^{\sigma_b}(b_{i,1}^j, A_i^j)\right| \leq 1 \tag{4.1}$$

$$\phi^{\sigma_b}(A_i^j) = \begin{cases} g+1 & \text{if } b_i = 1 \wedge b_{i+1} = j, \\ g+1+2z & \text{if } b_{i+1} \neq j \text{ and } z < \frac{1}{2}(b-1-g), \\ b & \text{otherwise.} \end{cases} \tag{4.2}$$

We discuss an inconsistency regarding Equation (4.2): Assuming that $\phi^{\sigma_b}(A_i^j)$ are given by Equations (4.1) and (4.2) and that the other entries of [3, Table 4] are correct causes the following contradiction.

**Issue 4.1.** *Let $b < 2^{n-k-1} - 1$ for some $k \in \mathbb{N}$. Then, there are edges that have a negative occurrence record according to [3, Table 4].*

We now resolve Issue 4.1. Let $b \in \mathbb{B}_n$ and $A_i^j$ be a bicycle. We show that when applying the switches as described in [3], the occurrence records are given by the following system:

$$|\phi^{\sigma_b}(b_{i,0}^j, A_i^j) - \phi^{\sigma_b}(b_{i,1}^j, A_i^j)| \leq 1 \tag{4.3}$$

$$\phi^{\sigma_b}(A_i^j) = \begin{cases} g+1, & A_i^j \text{ is closed and active} \\ b, & A_i^j \text{ is open and active} \\ b, & A_i^j \text{ is inactive and } b < 2^{i-1} + j \cdot 2^i \\ g+1+2z, & A_i^j \text{ is inactive and } b \geq 2^{i-1} + j \cdot 2^i \end{cases} \tag{4.4}$$

This system of equations properly distinguishes between inactive bicycles that need to catch up with the counter and inactive bicycles that do not need to do so.

Informally, for $b \in \mathbb{B}_n$ the occurrence records once the policy $\sigma_b$ is reached can be described as follows:

- Every closed and active bicycle has an occurrence record corresponding to the last time it was closed.
- Every open and active bicycle has an occurrence record of $b$.
- Inactive bicycles are either "catching up" with other bicycles and thus have an occurrence record less than $b$ or already finished catching up and have an occurrence record of $b$.

Before proving that Equations (4.3) and (4.4) correctly describe the occurrence records, we compare them to [3, Table 4]. Equations (4.1) and (4.3) bounding the difference of the occurrence records within a bicycle are the same and the case of closed and active bicycles is the same. Consider the second condition of Equation (4.2). This case handles inactive bicycles that do not have an occurrence record of $b$. This is handled by the condition $z < \frac{1}{2}(b-1-g)$, which is equivalent to $g+1+2z < b$. However, as shown in Issue 4.1, this condition does not describe inactive bicycles properly. We therefore formulate another condition, regarding the relation between $b$ and $2^{i-1} + j \cdot 2^i$. This condition is used to distinguish inactive bicycles that might need to catch up with the counter because they have already been active once (if $b \geq 2^{i-1} + j \cdot 2^i$), and inactive bicycle that do not need to catch up because they have not been active before. Finally, the case of open and active bicycles concludes our description.

Next, we explain how the improving switches within the bicycles should be applied. This description is a reformulation of the description given in the proof of [3, Lemma 5]. We apply the improving switches according to the following rules for a bicycle $A_i^j$ during phase 1 of the transition from $\sigma_b$ to $\sigma_{b+1}$ (rules are not stated in the order of their application):

1. If $A_i^j$ is open and active, we switch one edge of the bicycle.
2. Let $j := b_{\ell(b+1)+1}$. In addition to the first rule, the second edge of $A_{\ell(b+1)}^j$ is switched.
3. If $A_i^j$ is inactive and $b < 2^{i-1} + j \cdot 2^i$, one edge of the bicycle is switched.

4. If $A_i^j$ is inactive, $b \geq 2^{i-1} + j \cdot 2^i$ and $z < \frac{1}{2}(b-1-g)$, both edges of $A_i^j$ are switched.

5. If $A_i^j$ is inactive, $b \geq 2^{i-1} + j \cdot 2^i$ and $z \geq \frac{1}{2}(b-1-g)$, only one edge of $A_i^j$ is switched.

Applying the improving switches according to these rules indeed yields the occurrence records described by Equations (4.3) and (4.4).

**Theorem 4.2.** *Suppose that the improving switches within the bicycles are applied as described by the rules 1 to 5. Let $b \in \mathbb{B}_n$ and $A_i^j$ be some bicycle. Then, Equations* (4.3) *and* (4.4) *correctly specify the occurrence record $\phi^{\sigma_b}(A_i^j)$.*

# 5 Improving switches of phase 3

In this section, we discuss the application of the improving switches during phase 3. There are two contradictory descriptions in [3] how to apply them. We prove that neither of the given orderings obeys the LEAST-ENTERED rule, even if the issues discussed previously are resolved. We additionally show that a natural adaptation of Friedmann's scheme still does not obey the LEAST-ENTERED rule. We then go on to prove the existence of an ordering and an associated tie-breaking rule that obey the LEAST-ENTERED rule while still producing the intended behavior of Friedmann's construction.

Throughout this section, for a fixed $b \in \mathbb{B}_n$, we use $\ell := \ell(b)$ and $\ell' := \ell(b+1)$ to denote the least significant set bits of $b$ and $b+1$, respectively.

## 5.1 Issues with Friedmann's switching order

In Section 2.1, we stated that during phase 3, improving switches need to be applied for every entry vertex $k_i$ belonging to a level $i$ with $(b+1)_i = 0$. In addition, several bicycles need to be opened. However, according to the informal description given by Friedmann [3, Pages 9–10], both the updates regarding the entry vertices and the updates regarding the bicycles should not be performed for all levels but only those with an index smaller than $\ell'$. To be precise, the following is stated (where $r \in \{0,1\}$ is arbitrary):[2] "*In the third phase, we perform the major part of the resetting process. By resetting, we mean to unset lower bits again, which corresponds to reopening the respective bicycles. Also, we want to update all other inactive or active but not set bicycles again to move to the entry point $k_{\ell'}$. In other words, we need to update the lower entry points $k_z$ with $z < \ell'$ to move to $k_{\ell'}$, and the bicycle nodes $b_{z,l}^j$ to move to $k_{\ell'}$. We apply these switches by first switching the entry node $k_z$ for some $z < \ell'$ and then the respective bicycle nodes $b_{z,r}^j$.*"

Note that we only refer to [3, Table 3] when discussing occurrence records of improving switches since we do not consider the occurrence records of edges $(b_{i,r}^j, A_i^j)$. All discussed results therefore hold independently of the previous findings in Section 4.

We begin by showing an issue regarding the informal description mentioned before.

**Issue 5.1.** *For every $b \in \{1, \ldots, 2^{n-2}-1\}$, the informal description of phase 3 described in [3, Pages 9–10] contradicts [3, Tables 2,4]. It additionally violates the* LEAST-ENTERED *pivot rule during the transition from $\sigma_b$ to $\sigma_{b+1}$ for every $b \in \{3, \ldots, 2^{n-2}-2\}$.*

In other parts of the construction, Friedmann seems to apply the improving switches differently, by not only applying them for levels with a lower index than the least significant set bit but for all levels. Especially, the side conditions specified in [3, Table 2] for defining a phase $p$ policy rely on the fact that these switches are applied for all levels $i$ with $(b+1)_i = 0$. According to the proof of [3, Lemma 5], the switches need to be applied as follows:[2] "*In order to fulfill all side conditions for phase 3, we need to perform all switches from higher indices to smaller indices, and $k_i$ to $k_{\ell'}$ before $b_{i,r}^j$ with $(b+1)_{i+1} \neq j$ or $(b+1)_i = 0$ to $k_{\ell'}$*". However, applying the improving switches in this fashion results in the following issue.

**Issue 5.2.** *Applying the improving switches as described in [3, Lemma 5] does not obey the* LEAST-ENTERED *pivot rule.*

[2]The notation in the quote was adapted from [3] to be in line with our paper.

We can show an even stronger statement. Friedmann applies the improving switches of phase 3 in the following way: During the transition from $\sigma_b$ to $\sigma_{b+1}$, the improving switches are applied "one level after another" where the order of the levels depends on the least significant set bit of $b+1$, that is, $\ell(b+1)$. Our goal is now to show the following: Consider some $l \in \{1, \dots, n-4\}$. When the improving switches of phase 3 are applied level by level according to a fixed ordering $S^l$ during all transitions from $\sigma_b$ to $\sigma_{b+1}$ for which $\ell(b+1) = l$, the LEAST-ENTERED pivot rule is violated at least once.

We therefore prove that an entire class of orderings of the improving switches of phase 3, including Friedmann's, all violate the LEAST-ENTERED pivot rule. This class of orderings consists of all orderings such that the improving switches of phase 3 are applied "level by level", where, during the transition from $\sigma_b$ to $\sigma_{b+1}$, the sequence of levels only depends on the least significant set bit of $b+1$. That is, depending on $\ell(b+1)$, an ordering $S^{\ell(b+1)}$ of the levels 1 to $n$ is considered and when a level $i_1$ appears before a level $i_2$ within $S^{\ell(b+1)}$, all switches in level $i_1$ need to be applied before the improving switches of level $i_2$ are applied. In some sense, this shows that Friedmann's ordering needs to be changed fundamentally, and cannot be fixed by slight adaptation.

**Issue 5.3.** *Suppose that the improving switches of phase 3 are applied one level after another as described above. That is, the ordering of the levels in the transition from $\sigma_b$ to $\sigma_{b+1}$ may only depend on $\ell(b+1)$. Then, the* LEAST-ENTERED *pivot rule is violated.*

## 5.2 Fixing the ordering of the improving switches

In this section we prove the existence of an ordering and an associated tie-breaking rule for the application of the switches of phase 3 that obey the LEAST-ENTERED rule. We then show that these can be used to prove the existence of an ordering and an associated tie-breaking rule that obey the LEAST-ENTERED rule for all phases that produces the intended behavior.

Let $\sigma$ be a phase 3 policy. We need to compare $L^3_\sigma$ and $U^3_\sigma$ since all improving switches that can possibly be applied during phase 3 are contained in $U^3_\sigma$ (by [3, Lemma 4]). This is done via partitioning $U^3_\sigma$. The comparison then enables us to show that there is always a switch contained in $L^3_\sigma$ minimizing the occurrence record. This justifies that "*we will only use switches from $L^p_\sigma$*" [3, Page 12] (at least for phase $p = 3$). We then show the following: All improving switches that should be applied during phase 3 according to [3] can be applied (in a different order) during phase 3, without violating the LEAST-ENTERED pivot rule.

As outlined in Section 2, the transition from $\sigma_b$ to $\sigma_{b+1}$ is partitioned into six phases. During the third phase, the MDP is reset, that is, some bicycles are opened and the targets of some entry vertices are changed. Therefore, a phase 3 policy $\sigma$ is always associated with such a transition and we implicitly consider the underlying transition from $\sigma_b$ to $\sigma_{b+1}$ for the corresponding $b \in \mathbb{B}_n$ when discussing a fixed phase 3 policy.

Now, fix some $b \in \{0, \dots, 2^n - 2\}$. For an edge $e = (v, w)$, we say that the edge *belongs to level $i$* when vertex $v$ is part of level $i$ of the lower bound construction.

We begin by further investigating the occurrence records of switches that should be applied during phase 3, i.e., we analyze the set $L^3_\sigma$ for a phase 3 policy $\sigma$. We first develop an upper bound on the occurrence record of these switches.

**Lemma 5.4.** *Let $\sigma$ be a phase 3 policy. Then $\max_{e \in L^3_\sigma} \phi^\sigma(e) \le f(b, \ell')$.*

Before showing that for all phase 3 policies $\sigma$, there is always an improving switch contained in $L^3_\sigma$ that minimizes the occurrence record, we further discuss the superset $U^3_\sigma$. We observe that $L^6_\sigma$ is contained in this set. We thus need to analyze this set as well. However, there is a small error in the definition of this set that needs to be corrected.

**Issue 5.5.** *For every $b \in \mathbb{B}_n$ with $\ell(b+1) > 1$, there is an improving switch that should be applied in phase 6 of the transition from $\sigma_b$ to $\sigma_{b+1}$ but is not contained in the set $L^6_\sigma$ for any phase 6 policy $\sigma$.*

We propose the following definition for $L^6_\sigma$ and always implicitly consider it henceforth.

**Theorem 5.6.** *For any phase 6 policy $\sigma$, the subset of $I_\sigma$ in [3, Table 3] needs to be*

$$\bar{L}^6_\sigma := \left\{(d_i^0, x) : \sigma(d_i^0) \neq x \wedge \sigma(d_i^0) = \begin{cases} h_i^0, & (b+1)_{i+1} = 1 \\ s, & (b+1)_{i+1} = 0 \end{cases}\right\} \cup \left\{(d_i^1, x) : \sigma(d_i^1) \neq x \wedge \sigma(d_i^1) = \begin{cases} s, & (b+1)_{i+1} = 1 \\ h_i^0, & (b+1)_{i+1} = 0 \end{cases}\right\}.$$

After having fixed Issue 5.5, we return to the discussion of the set $U^3_\sigma$. We partition the superset $U^3_\sigma$ contained in [3, Table 3] for a phase 3 policy $\sigma$ as follows.

$$\begin{aligned}
U^{3,1}_\sigma &:= \{(k_i, k_z) : \sigma(k_i) \notin \{k_z, k_{\ell'}\}, z \leq \ell' \wedge (b+1)_i = 0\} \\
U^{3,2}_\sigma &:= \{(b^j_{i,r}, k_z) : \sigma(b^j_{i,r}) \notin \{k_z, k_{\ell'}\}, z \leq \ell' \wedge (b+1)_i = 0\}\} \\
U^{3,3}_\sigma &:= \{(b^j_{i,r}, k_z) : \sigma(b^j_{i,r}) \notin \{k_z, k_{\ell'}\}, z \leq \ell' \wedge (b+1)_{i+1} \neq j\} \\
U^{3,4}_\sigma &:= \{(h_i^0, k_l) : l \leq \min(\{n+1\} \cup \{j \in \{i+2, \dots, n\} : b_j = 1\})\} \\
U^{3,5}_\sigma &:= \{(s, k_i) : \sigma(s) \neq k_i \wedge i < \ell'\} \\
U^{3,6}_\sigma &:= \{(d_i^j, x) : \sigma(d_i^j) \neq x \wedge i < \ell'\} \\
U^{3,7}_\sigma &:= \left\{(d_i^0, x) : \sigma(d_i^0) \neq x \wedge \sigma(d_i^0) = \begin{cases} h_i^0, & (b+1)_{i+1} = 1 \\ s, & (b+1)_{i+1} = 0 \end{cases}\right\} \\
U^{3,8}_\sigma &:= \left\{(d_i^1, x) : \sigma(d_i^1) \neq x \wedge \sigma(d_i^1) = \begin{cases} s, & (b+1)_{i+1} = 1 \\ h_i^0, & (b+1)_{i+1} = 0 \end{cases}\right\} \\
U^{3,9}_\sigma &:= \{(b^j_{i,l}, A^j_i) : \sigma(b^j_{i,l}) \neq A^j_i\}
\end{aligned}$$

In Lemma 5.4, we showed an upper bound of $f(b, \ell')$ on the occurrence records of the improving switches contained in $L^3_\sigma$ for a phase 3 policy $\sigma$. The next lemma gives a matching lower bound of $f(b, \ell')$ on all improving switches that should be applied *after* phase 3. It will also be used to estimate the occurrence records of possible improving switches contained in $U^3_\sigma$.

**Lemma 5.7.** Let $\sigma$ be a phase 3 policy. Assume that the Policy Iteration Algorithm is started with the policy $\sigma^*$ introduced in Definition 3.3. Then $\min_{e \in L^4_\sigma \cup L^5_\sigma \cup L^6_\sigma} \phi^\sigma(e) \geq f(b, \ell')$.

This lemma can now be used to show that the occurrence records of edges contained in the sets $U^{3,4}_\sigma$ to $U^{3,9}_\sigma$ are too large. To be precise, we show that no improving switch contained in one of these sets will be applied for any phase 3 policy when following the LEAST-ENTERED rule.

**Lemma 5.8.** *Let $\sigma$ be a phase 3 policy. Then, for all edges $e \in L^3_\sigma$ and $\tilde{e} \in I_\sigma \cap \left(U^{3,4}_\sigma \cup \dots \cup U^{3,9}_\sigma\right)$, it holds that $\phi^\sigma(e) \leq \phi^\sigma(\tilde{e})$.*

It remains to analyze the sets $U^{3,1}_\sigma, U^{3,2}_\sigma$ and $U^{3,3}_\sigma$. Our goal is to show that applying certain improving switches contained in $L^3_\sigma$ prevent other switches contained in the three mentioned sets from being applied. To do so, we introduce subsets of $U^{3,1}_\sigma, U^{3,2}_\sigma$ and $U^{3,3}_\sigma$. The idea is to "slice" these sets such that for each such slice, there is an improving switch that prevents the whole slice from being applied.

**Definition 5.9** (Slices)**.** Let $\sigma$ be a phase 3 policy and $i \in \{1, \dots, n\}, j, l \in \{0, 1\}$. Then

- $S^{3,1}_{i,\sigma} := \{(k_i, k_z) : \sigma(k_i) \notin \{k_z, k_{\ell'}\}, z \leq \ell' \wedge (b+1)_i = 0\}$ is called *slice* of $U^{3,1}_\sigma$,
- $S^{3,2}_{i,j,r,\sigma} := \{(b^j_{i,r}, k_z) : \sigma(k_i) \notin \{k_z, k_{\ell'}\}, z \leq \ell' \wedge (b+1)_i = 0\}$ is called *slice* of $U^{3,2}_\sigma$ and
- $S^{3,3}_{i,j,r,\sigma} := \{(b^j_{i,r}, k_z) : \sigma(k_i) \notin \{k_z, k_{\ell'}\}, z \leq \ell' \wedge (b+1)_i \neq j\}$ is called *slice* of $U^{3,3}_\sigma$.

It is easy to verify that the set of all slices of one of the sets $U^{3,1}_\sigma, U^{3,2}_\sigma, U^{3,3}_\sigma$ partitions the corresponding set.

We now show that every switch contained in $L^3_\sigma$ prevents the improving switches contained in certain slices from being applied. This is done by proving that whole sets of improving switches are no longer improving switches when a specific switch is applied.

**Lemma 5.10.** *The following statements hold.*

1. *Let $\sigma$ be the phase 3 policy in which the improving switch $(k_i, k_{\ell'})$ is applied. Let $\sigma'$ be a phase 3 policy of the same transition reached after the policy $\sigma$. Then $I_{\sigma'} \cap S^{3,1}_{i,\sigma'} = \emptyset$.*

2. *Let $\sigma$ be the phase 3 policy in which the improving switch $(b^j_{i,l}, k_{\ell'})$ with $\sigma(b^j_{i,l}) \neq k_{\ell'}$ and $(b+1)_i = 0$ is applied. Let $\sigma'$ be an arbitrary phase 3 policy of the same transition reached after the policy $\sigma$. Then $I_{\sigma'} \cap S^{3,2}_{i,j,l,\sigma'} = \emptyset$.*

3. *Let $\sigma$ be the phase 3 policy in which the improving switch $(b^j_{i,l}, k_{\ell'})$ with $\sigma(b^j_{i,l}) \neq k_{\ell'}$ and $(b+1)_{i+1} \neq j$ is applied. Let $\sigma'$ be an arbitrary phase 3 policy of the same transition reached after the policy $\sigma$. Then $I_{\sigma'} \cap S^{3,3}_{i,j,l,\sigma'} = \emptyset$.*

This enables us to prove the following lemma which allows us to show that it is possible to always choose a switch contained in $L^3_\sigma$ when applying the LEAST-ENTERED pivot rule.

**Lemma 5.11.** *Let $\sigma$ be a phase 3 policy. Then there is an edge $e \in L^3_\sigma \cap \arg\min_{\tilde{e} \in I_\sigma} \phi^\sigma(\tilde{e})$.*

This lemma does not immediately imply that all improving switches of phase 3 can be applied. The reason is that it is not clear why it cannot happen that a phase 4 policy is reached although not all switches of phase 3 were applied yet. However, as shown next, this is possible.

**Theorem 5.12.** *There is an ordering of the improving switches and an associated tie-breaking rule compatible with the* LEAST-ENTERED *pivot rule such that all improving switches contained in $L^3_{\sigma_b}$ are applied and the* LEAST-ENTERED *pivot rule is obeyed during phase 3.*

Although Theorem 5.12 shows that the improving switches of phase 3 can be applied such that the LEAST-ENTERED rule is obeyed, it does not imply that the transition from $\sigma_b$ to $\sigma_{b+1}$ can be executed as intended by Friedmann. That is, it does not imply that the improving switches of the phases 1,2,4,5 and 6 can be applied as intended. This however follows from Theorem 5.12 and a brief analysis of the other phases that can be found in Appendix B.

**Theorem 5.13.** *Fix the transition from $\sigma_b$ to $\sigma_{b+1}$ for some $\sigma \in \mathbb{B}_n$. There is an order in which to apply improving switches during this transition such that the* LEAST-ENTERED *rule is obeyed, and the switches of phase $p$ are applied before any switches of phase $p+1$, for every $p \in \{1, \dots, 5\}$.*

# 6 Conclusion

In this paper we revisited the lower bound example constructed in [3] that yields a subexponential lower bound on the Simplex Algorithm using the LEAST-ENTERED pivot rule. We discussed the example in general and highlighted several issues with Friedmann's analysis in [3]. We proposed alterations regarding the application of the Policy Iteration Algorithm to resolve all of these issues. In particular, we showed that the initial policy for the policy iteration needs to be changed and provided a new initial policy (Section 3). We further showed that the description of occurence records is not entirely accurate and corrected the inaccuracy (Section 4). Most notably, we proved that the order in which Friedmann applies certain improving switches, as well as simple adaptations of this order, are inconsistent with the LEAST-ENTERED rule (Section 5). In addition, we implicitly provided a more involved ordering and associated tie-breaking rule that overcome this issue and produces the intended behavior of Friedmann's construction.

Crucially, our changes retain the macroscopic properties of the construction, and, as a consequence, we are able to recover Friedmann's subexponential lower bound. It remains an open problem to give a lower bound for Zadeh's pivot rule that works with a natural tie-breaking rule.

# References


[1] D. Avis and V. Chvátal. Notes on Bland's pivoting rule. In *Polyhedral Combinatorics: Dedicated to the memory of D.R. Fulkerson*, pages 24–34. Springer, 1978.

[2] G. B. Dantzig. *Linear Programming and Extensions*. Princeton University Press, 1963.

[3] O. Friedmann. A subexponential lower bound for Zadeh's pivoting rule for solving linear programs and games. In *Proceedings of the 15th International Conference on Integer Programming and Combinatoral Optimization (IPCO)*, pages 192–206, 2011.

[4] D. Goldfarb and W. Y. Sit. Worst case behavior of the steepest edge simplex method. *Discrete Applied Mathematics*, 1(4):277–285, 1979.

[5] R. A. Howard. *Dynamic Programming and Markov Processes*. MIT Press, 1960.

[6] R. G. Jeroslow. The simplex algorithm with the pivot rule of maximizing criterion improvement. *Discrete Mathematics*, 4(4):367–377, 1973.

[7] V. Klee and G. J. Minty. How good is the simplex algorithm? In *Inequalities III*, pages 159–175. Academic Press, New York, 1972.

[8] M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc., New York, 2005.

[9] A. Thomas. Exponential Lower Bounds for History-Based Simplex Pivot Rules on Abstract Cubes. In *25th Annual European Symposium on Algorithms (ESA)*, volume 87, page 69(14), 2017.

[10] N. Zadeh. What is the worst case behavior of the simplex algorithm? Technical Report 27, Departments of Operations Research, Stanford, 1980.

# A Imported Tables

This appendix contains tables of [3] used in this paper. They are labeled the same way as in [3] and use the exact notation of [3]. The tables use the alternative notion $\bar{\sigma}$ for referring to the target of a vertex with respect to a policy $\sigma$ that we omitted. It is defined as follows.

| $\sigma(v)$ | $t$ | $k_i$ | $h_*^*$ | $s$ | $A_*^*$ | $c_i^j$ |
|---|---|---|---|---|---|---|
| $\bar{\sigma}(v)$ | $n+1$ | $i$ | $1$ | $0$ | $0$ | $-j$ |

As in [3], we write $\bar{\sigma}(A_i^j) = 1$ if $\sigma(b_{i,0}^j) = A_i^j$ and $\sigma(b_{i,1}^j) = A_i^j$ and $\bar{\sigma}(A_i^j) = 0$ otherwise. In addition, the notation $b' = b + 1$ and $\nu_i^j := \min(\{n+1\} \cup \{j \geq i : b_j = 0\})$ is used.

The first table shows when a policy $\sigma$ is considered a phase $p$ policy.

| Phase | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $\bar{\sigma}(s)$ | $r$ | $r$ | $r$ | $r$ | $r$ | $r'$ |
| $\bar{\sigma}(d_i^0)$ | $1-b_{i+1}$ | $1-b_{i+1}$ | $1-b_{i+1}$ | $1-b_{i+1}$ | $1-b_{i+1}$ | $1-b_{i+1}, 1-b'_{i+1}$ |
| $\bar{\sigma}(d_i^1)$ | $b_{i+1}$ | $b_{i+1}$ | $b_{i+1}$ | $b_{i+1}$ | $b_{i+1}$ | $b_{i+1}, b'_{i+1}$ |
| $\bar{\sigma}(h_i^0)$ | $\nu_{i+2}^n(b)$ | $\nu_{i+2}^n(b)$ | $\nu_{i+2}^n(b)$ | $\nu_{i+2}^n(b), \nu_{i+2}^n(b')$ | $\nu_{i+2}^n(b')$ | $\nu_{i+2}^n(b')$ |
| $\bar{\sigma}(b_{*,*}^*)$ | $0, r$ | $0, r$ | $0, r, r'$ | $0, r'$ | $0, r'$ | $0, r'$ |
| $\bar{\sigma}(A_i^{b_{i+1}})$ | $b_i$ | $*$ | $*$ | $*$ | $*$ | $*$ |
| $\bar{\sigma}(A_i^{b'_{i+1}})$ | $*$ | $b'_i$ | $b'_i$ | $b'_i$ | $b'_i$ | $b'_i$ |

| Phase | 1-2 | 3-4 | 5-6 |
|---|---|---|---|
| $\bar{\sigma}(k_i)$ | $\begin{cases} r & \text{if } b_i = 0 \\ -b_{i+1} & \text{if } b_i = 1 \end{cases}$ | $\begin{cases} r, r' & \text{if } b'_i = 0 \wedge b_i = 0 \\ -b_{i+1}, r' & \text{if } b'_i = 0 \wedge b_i = 1 \\ -b'_{i+1} & \text{if } b'_i = 1 \end{cases}$ | $\begin{cases} r' & \text{if } b'_i = 0 \\ -b'_{i+1} & \text{if } b_i = 1' \end{cases}$ |

| Phase 3 | Side Conditions |
|---|---|
| (a) | $\forall i.([b'_i = 0 \text{ and } (\exists j, l. \bar{\sigma}(b_{i,l}^j) = r')] \text{ implies } \bar{\sigma}(k_i) = r')$ |
| (b) | $\forall i, j.([b'_i = 0, b'_j = 0, \bar{\sigma}(k_i) = r' \text{ and } \bar{\sigma}(k_j) \neq r'] \text{ implies } i > j)$ |

Table 2: Policy Phases where $b' = b + 1, r = \nu_1^n(b), r' = \nu_1^n(b')$ and $*$ is arbitrary

The second table shows subsets $L_\sigma^p$ and supersets $U_\sigma^p$ of the set of improving switches of a phase $p$ policy. Note that this table shows the original Table contained in [3].

| Ph. $p$ | Improving Switches Subset $L_\sigma^p$ | Improving Switches Superset $U_\sigma^p$ |
|---|---|---|
| 1 | $\{(b_{i,l}^j, A_i^j) \mid \sigma(b_{i,l}^j) \neq A_i^j\}$ | $L_\sigma^1$ |
| 2 | $\{(k_{r'}, c_{r'}^{b'_{r'+1}})\}$ | $L_\sigma^1 \cup L_\sigma^2$ |
| 3 | $\{(k_i, k_{r'}) \mid \bar{\sigma}(k_i) \neq r' \wedge b'_i = 0\} \cup$ $\{(b_{i,l}^j, k_{r'}) \mid \bar{\sigma}(b_{i,l}^j) \neq r' \wedge b'_i = 0\} \cup$ $\{(b_{i,l}^j, k_{r'}) \mid \bar{\sigma}(b_{i,l}^j) \neq r' \wedge b'_{i+1} \neq j\}$ | $U_\sigma^4 \cup \{(k_i, k_z) \mid \bar{\sigma}(k_i) \notin \{z, r'\}, z \leq r' \wedge b'_i = 0\} \cup$ $\{(b_{i,l}^j, k_z) \mid \bar{\sigma}(b_{i,l}^j) \notin \{z, r'\}, z \leq r' \wedge b'_i = 0\} \cup$ $\{(b_{i,l}^j, k_z) \mid \bar{\sigma}(b_{i,l}^j) \notin \{z, r'\}, z \leq r' \wedge b'_{i+1} \neq j\}$ |
| 4 | $\{(h_i^0, k_{\nu_{i+2}^n(b')}) \mid \bar{\sigma}(h_i^0) \neq \nu_{i+2}^n(b')\}$ | $U_\sigma^5 \cup \{(h_i^0, k_l) \mid l \leq \nu_{i+2}^n(b')\}$ |
| 5 | $\{(s, k_{r'})\}$ | $U_\sigma^6 \cup \{(s, k_i) \mid \bar{\sigma}(s) \neq i \wedge i < r'\} \cup$ $\{(d_i^j, x) \mid \sigma(d_i^j) \neq x \wedge i < r'\}$ |
| 6 | $\{(d_i^0, x) \mid \sigma(d_i^0) \neq x \wedge \bar{\sigma}(d_i^0) \neq b'_{i+1}\} \cup$ $\{(d_i^1, x) \mid \sigma(d_i^1) \neq x \wedge \bar{\sigma}(d_i^1)' b_{i+1})\}$ | $L_\sigma^1 \cup L_\sigma^6$ |

Table 3: Improving Switches (where $b' = b + 1$ and $r' = \nu_1^n(b')$)

The last table shows the occurrence records of a policy $\sigma_b$ representing the number $b \in \mathbb{B}_n$ according to [3]. In this table, the notation $g^* := g(b, i, \{(i+1, j)\})$ is used.

| Edge $e$ | $(*, t)$ | $(s, k_r)$ | $(h_*^0, k_r)$ |
|---|---|---|---|
| $\phi^b(e)$ | 0 | $f(b, r)$ | $f(b, r)$ |

| Edge $e$ | $(b_{i,*}^j, k_r)$ |
|---|---|
| $\phi^b(e)$ | $f(b, r, \{(i, 0)\}) + f(b, r, \{(i, 1), (i + 1, 1 - j)\})$ |

| Edge $e$ | $(k_i, k_r)$ | $(k_i, c_i^j)$ |
|---|---|---|
| $\phi^b(e)$ | $f(b, r, \{(i, 0)\})$ | $f(b, i, \{(i + 1, j)\})$ |

| Edge $e$ | $(d_i^j, s)$ | $(d_i^j, h_i^j)$ |
|---|---|---|
| $\phi^b(e)$ | $f(b, i + 1) - j \cdot b_{i+1}$ | $f(b, i + 1) - (1 - j) \cdot b_{i+1}$ |

| Complicated Conditions |
|---|
| $\lvert \phi^b(b_{i,0}^j, A_i^j) - \phi^b(b_{i,1}^j, A_i^j) \rvert \leq 1$ |
| $\phi^b(b_{i,0}^j, A_i^j) + \phi^b(b_{i,1}^j, A_i^j) = \begin{cases} g^* + 1 & \text{if } b_i = 1 \text{ and } b_{i+1} = j \\ g^* + 1 + 2 \cdot z & \text{if } b_{i+1} \neq j \text{ and } z := b - g^* - 2^{i-1} < \frac{1}{2}(b - 1 - g^*) \\ b & \text{otherwise} \end{cases}$ |

Table 4: Occurrence Records for the policy $\sigma_b$ representing $b$ (where $r = \ell(b)$)

# B Omitted Proofs

This appendix contains the proofs omitted from the main part of this paper and additional lemmas that are needed for technical reasons.

The following lemma is central for proving most of the statements since it gives important results on the flip set and flip number introduced in Section 2.2.

**Lemma B.1.** *Let $b \in \mathbb{B}_n$ and $i, j \in \{1, \ldots, n\}$. Then the following hold:*

1. *Let $S, S'$ be schemes and $S \subseteq S'$. Then $M(b, S') \subseteq M(b, S)$.*
2. *Let $S, S'$ be schemes and $S \subseteq S'$. Then $f(b, i, S') \leq f(b, i, S)$.*
3. *It holds that $f(b, j) = f(b, j, \{(i, 0)\}) + f(b, j, \{(i, 1)\})$ and $f(b, j) = \lfloor \frac{b + 2^{j-1}}{2^j} \rfloor$.*
4. *Let $i \leq j$ and $S$ be a scheme. Then $f(b, j, S) \leq f(b, i, S)$ and thus $f(b, j) \leq f(b, i)$.*
5. *Let $i < j$. Then $F(b, j) = F(b, j, \{(i, 0)\})$ and thus $f(b, j, \{(i, 0)\}) = f(b, j)$.*

*Proof.* We prove the statements one after another.

1. Let $S, S'$ be schemes such that $S \subseteq S'$. Since every number matching the scheme $S'$ also matches the scheme $S$, it follows that $M(b, S') \subseteq M(b, S)$ for all numbers $b \in \mathbb{B}_n$.
2. This follows directly from (1) and the definition of $f(b, i, S')$.
3. The first statement follows immediately since either $b_i = 0$ or $b_i = 1$ holds for every binary number $b \in \mathbb{B}_n$ and index $i \in \{1, \ldots, n\}$.

   It remains to show that $f(b, j) = \lfloor \frac{b + 2^{j-1}}{2^j} \rfloor$ for $b \in \mathbb{B}_n$ and $j \in \{1, \ldots, n\}$. We observe that $2^{j-1}$ is the smallest number matching the scheme $S_j := \{(j, 1), (j - 1, 0), \ldots, (1, 0)\}$. This implies the statement for $b < 2^{j-1}$. Now, let $m_i$ denote the $i$-th number matching the scheme $S_j$. Then, by the previous argument, $m_1 = 2^{j-1}$. As only numbers ending on the subsequence $(1, 0, \ldots, 0)$ of length $j$ match the scheme $S_j$, we have $m_i = (i - 1) \cdot 2^j + 2^{j-1}$. Since $f(m_i, j) = i$ by definition and

   $$\left\lfloor \frac{m_i + 2^{j-1}}{2^j} \right\rfloor = \left\lfloor \frac{(i - 1) \cdot 2^j + 2^{j-1} + 2^{j-1}}{2^j} \right\rfloor = \left\lfloor \frac{i \cdot 2^j}{2^j} \right\rfloor = i,$$

   we get $f(m_i, j) = \left\lfloor \frac{m_i + 2^{j-1}}{2^j} \right\rfloor$.

Now let $b \in \mathbb{B}_n$ and choose $i \in \mathbb{N}$ such that $b \in [m_i, m_{i+1})$. Then, by the definition of $f(b,j)$, we have $f(b,j) = i$. In addition, by the choice of $i$,

$$\left\lfloor \frac{b + 2^{j-1}}{2^j} \right\rfloor \geq \left\lfloor \frac{m_i + 2^{j-1}}{2^j} \right\rfloor = f(m_i, j) = i \tag{B.1}$$

and

$$\left\lfloor \frac{b + 2^{j-1}}{2^j} \right\rfloor < \left\lfloor \frac{m_{i+1} + 2^{j-1}}{2^j} \right\rfloor = f(m_{i+1}, j) = i + 1. \tag{B.2}$$

By integrality, Equations (B.1) and (B.2) imply that $\left\lfloor \frac{b+2^{j-1}}{2^j} \right\rfloor = i$ and thus,

$$f(b,j) = i = \left\lfloor \frac{b + 2^{j-1}}{2^j} \right\rfloor.$$

4. Let $i \leq j$ and $b \in \mathbb{B}_n$. Let $S_j := \{(j,1), (j-1,0), \ldots, (1,0)\}$ and define $S_i$ analogously. Consider any number $\tilde{b} \leq b$ matching both the schemes $S_j$ and $S$. Then, since $i \leq j$ there needs to be at least one number $\hat{b} \leq \tilde{b}$ matching $S_i$ and $S$. This immediately implies $f(b,j,S) \leq f(b,i,S)$.

   The second inequality follows immediately when setting $S := \emptyset$.

5. Let $i < j$ and define $S_j := \{(j,1), (j-1,0), \ldots, (1,0)\}$. Since $i < j$, we have $(i,0) \in S_j$, immediately implying $F(b,j) = F(b,j,\{(i,0)\})$ . □

## B.1 Proofs of Section 3

This subsection contains the proofs of the statements of Section 3 on the initial policy. The first statement of that section, i.e., Issue 3.1, is proven with the help of the following lemma.

**Lemma B.2.** *None of the edges $(b^1_{i,r}, A^1_i)$ for $i \in \{1, \ldots, n\}$ and $r \in \{0,1\}$ is an improving switch with respect to $\sigma^\star$.*

*Proof.* Fix some $i \in \{1, \ldots, n\}$ and $r \in \{0,1\}$. By definition of $\sigma^\star$, it holds that $\sigma^\star(b^1_{i,r}) = t$. Therefore, $\mathrm{VAL}_{\sigma^\star}(b^1_{i,r}) = \mathrm{VAL}_{\sigma^\star}(\sigma^\star(b^1_{i,r})) = \mathrm{VAL}_{\sigma^\star}(t) = 0$ since all edges starting in $b^1_{i,r}$ have a reward of 0. Analogously, $\mathrm{VAL}_{\sigma^\star}(b^1_{i,1-r}) = 0$. This implies that $(b^1_{i,r}, A^1_i)$ is an improving switch if and only if $\mathrm{VAL}_{\sigma^\star}(A^1_i) > 0$. But, due to $\sigma^\star(k_{i+1}) = t$, it holds that

$$\begin{aligned}
\mathrm{VAL}_{\sigma^\star}(A^1_i) &= \epsilon\,\mathrm{VAL}_{\sigma^\star}(d^1_i) + \frac{1-\epsilon}{2}\mathrm{VAL}_{\sigma^\star}(b^1_{i,l}) + \frac{1-\epsilon}{2}\mathrm{VAL}_{\sigma^\star}(b^1_{i,1-l}) \\
&= \epsilon\,\mathrm{VAL}_{\sigma^\star}(d^1_i) \\
&= \epsilon\left[(-N)^6 + \mathrm{VAL}_{\sigma^\star}(\sigma^\star(d^1_i))\right] \\
&= \epsilon\left[N^6 + \mathrm{VAL}_{\sigma^\star}(h^1_i)\right] \\
&= \epsilon\left[N^6 + (-N)^{2i+8} + \mathrm{VAL}_{\sigma^\star}(k_{i+1})\right] \\
&= \epsilon\left[N^6 + N^{2i+8} + (-N)^{2(i+1)+7} + \mathrm{VAL}_{\sigma^\star}(t)\right] \\
&= \epsilon\left[N^6 + N^{2i+8} - N^{2i+9}\right] < 0,
\end{aligned}$$

as $N \geq 7n + 1 \geq 8$ and $i \geq 1$. Therefore, the edge $(b^1_{i,r}, A^1_i)$ is not an improving switch. □

**Issue 3.1.** *The initial policy $\sigma^\star$ for the Markov decision process $G_n$ described in [3, Page 10] contradicts [3, Table 3] since $I_{\sigma^\star} \neq \{(b^j_{i,r}, A^j_i) \colon \sigma^*(b^j_{i,r}) \neq A^j_i\}$.*

*Proof.* By definition, $\sigma^\star$ is a phase 1 policy. Thus, according to [3, Table 3] and since $L^1_\sigma = U^1_\sigma = I_\sigma$ for all phase 1 policies $\sigma$ by [3, Lemma 4], the set of improving switches is $I_{\sigma^\star} = \{(b^j_{i,r}, A^j_i) | \sigma^\star(b^j_{i,r}) \neq A^j_i\}$. By definition of $\sigma^\star$, it holds that $\sigma^\star(b^1_{i,r}) = t$ which, due to $t \neq A^1_i$, implies $\sigma^\star(b^1_{i,r}) \neq A^1_i$. Therefore, $(b^1_{i,r}, A^1_i) \in I_{\sigma^\star}$ should hold according to [3, Table 3]. But, by Lemma B.2, $(b^1_{i,r}, A^1_i)$ is not an improving switch for any $i \in \{1, \ldots, n\}$ and $r \in \{0,1\}$. Therefore, the initial policy $\sigma^\star$ contradicts [3, Table 3]. □

**Issue 3.2.** *When the Policy Iteration Algorithm is started with $\sigma^\star$, either*

1. *[3, Table 4] containing the occurrence records is incorrect for $b = 1$, or*
2. *[3, Table 3] containing the sub- and supersets of $I_{\sigma_b}$ is incorrect for $b = 1$.*

*Proof.* Consider the first phase 6 policy of the transition from $\sigma^\star$ to $\sigma_1$. Denote this policy by $\sigma$ and fix some index $i \in \{1, \dots, n\}$. Then, by [3, Table 2], we have $\sigma(s) = k_1$ and $\sigma(k_j) = k_1$ for all $j \in \{2, \dots, n\}$. Therefore since the reward of the edge $(s, k_1)$ is equal to zero, this implies that $\mathrm{VAL}_\sigma(s) = \mathrm{VAL}_\sigma(k_1)$, and therefore

$$\begin{aligned}\mathrm{VAL}_\sigma(\sigma(d_i^1)) &= \mathrm{VAL}_\sigma(h_i^1)\\ &= N^{2i+8} + \mathrm{VAL}_\sigma(k_{i+1})\\ &= \underbrace{N^{2i+8} - N^{2i+9}}_{<0} + \underbrace{\mathrm{VAL}_\sigma(k_1)}_{=\mathrm{VAL}_\sigma(s)}\\ &< \mathrm{VAL}_\sigma(s),\end{aligned}$$

hence the edge $(d_i^1, s)$ is an improving switch for every $i \in \{1, \dots, n\}$. Now, we can either

1. apply (some or all) of these improving switches now or
2. we do not apply any of these improving switches now.

Suppose that we apply the switch $(d_i^1, s)$ for every $i \in \{1, \dots, n\}$. By definition, phase 6 ends after these switches are applied and phase 1 of the transition from $\sigma_1$ to $\sigma_2$ begins. But, according to [3, Table 4], it should hold that $\phi^{\sigma_2}(d_i^1, s) = f(b, i+1) - 1 \cdot b_{i+1}$ for all $i \in \{1, \dots, n\}$. Because $b = 1$, we have $b_k = 0$ for all $k \in \{2, \dots, n\}$. In particular, $b_2 = 0$, implying that $\phi^{\sigma_2}(d_i^1, s) = f(b, i+1) = 0$ for all $i \in \{1, \dots, n\}$. This is a contradiction to the fact that we have just switched the edges $(d_i^1, s)$. Note that this argument still holds when we only apply a subset of all of the improving switches.

Now suppose that we do not apply any of the improving switches. Then, all of the edges $(d_i^1, s)$ remain improving switches during phase 1 of the transition from $\sigma_1$ to $\sigma_2$. This however contradicts [3, Table 3]. □

The following two theorems show that the new initial policy $\sigma^\star$ introduced in Definition 3.3 avoids both Issues 3.1 and 3.2.

**Lemma 3.4.** *The set of improving switches for $\sigma^*$ is $I_{\sigma^*} = \{(b_{i,r}^j, A_i^j) \colon \sigma^*(b_{i,r}^j) \neq A_i^j\}$.*

*Proof.* Compared to the original initial policy $\sigma^\star$, the changes can only have an effect on the edges $(d_i^1, h_i^1)$ and $(b_{i,r}^1, A_i^1)$ for $i \in \{1, \dots, n\}, r \in \{0, 1\}$. Thus, it suffices to show that the edge $(d_i^1, h_i^1)$ is not an improving switch for any $i \in \{1, \dots, n\}$ whereas the edges $(b_{i,r}^1, A_i^1)$ are improving switches for all indices $i \in \{1, \dots, n\}$ and $r \in \{0, 1\}$.

Fix some $i \in \{1, \dots, n\}$. By the definition of $\sigma^*$, we have $\sigma^*(d_i^1) = s$ and $\sigma^*(s) = t$. This implies that $\mathrm{VAL}_{\sigma^*}(\sigma^*(d_i^1)) = 0$. In order to show that $(d_i^1, h_i^1)$ is not an improving switch, it thus suffices to show $\mathrm{VAL}_{\sigma^*}(h_i^1) < 0$. This however holds since

$$\begin{aligned}\mathrm{VAL}_{\sigma^*}(h_i^1) &= (-N)^{2i+8} + \mathrm{VAL}_{\sigma^*}(k_{i+1})\\ &= N^{2i+8} + (-N)^{2(i+1)+7} + \mathrm{VAL}_{\sigma^*}(t)\\ &= N^{2i+8} - N^{2i+9} < 0,\end{aligned}$$

due to $N \geq 8$ and $i \geq 1$. Therefore, the edge $(d_i^1, h_i^1)$ is not an improving switch.

Now fix some $r \in \{0, 1\}$. Since it holds that $\mathrm{VAL}_{\sigma^*}(\sigma^*(b_{i,r}^1)) = \mathrm{VAL}_{\sigma^*}(\sigma^*(b_{i,1-r}^1)) = 0$, it suffices to show that $\mathrm{VAL}_{\sigma^*}(A_i^1) > 0$ in order to prove that $(b_{i,r}^1, A_i^1)$ is an improving switch. Due to $\sigma^*(d_i^1) = s$, we have

$$\mathrm{VAL}_{\sigma^*}(A_i^1) = \epsilon\, \mathrm{VAL}_{\sigma^*}(d_i^1) = \epsilon \left[N^6 + \mathrm{VAL}_{\sigma^*}(s)\right] = \epsilon N^6 > 0,$$

so $(b_{i,r}^1, A_i^1)$ is an improving switch. □

**Lemma 3.5.** *Starting the Policy Iteration Algorithm with the initial policy $\sigma^*$ avoids Issue 3.2, that is, it does not contradict [3, Tables 3,4] for $b = 1$.*

*Proof.* Consider the first phase 6 policy of the transition from $\sigma^*$ to $\sigma_1$. Denote this policy by $\sigma$ and fix some index $i \in \{1, \dots, n\}$. Then, $\sigma(d_i^1) = s$ for all $i \in \{1, \dots, n\}$ by the definition of $\sigma^*$ and the application of improving switches. Therefore, none of the edges $(d_i^1, s)$ is an improving switch for any $i \in \{1, \dots, n\}$ and none of these edges can be switched. Thus, once $\sigma_1$ is reached, the occurrence record of these edges is equal to zero as they should be according to [3, Table 4]. This also implies that the edges $(d_i^1, s)$ are no improving switches during the transition from $\sigma_1$ to $\sigma_2$, resolving the contradiction regarding [3, Table 3]. □

## B.2 Proofs of Section 4

This subsection contains the proofs of the statements in Section 4 on the occurrence records of the edges of the bicycles. We first show that the occurrence record of these edges is not correctly described by [3, Table 4].

**Issue 4.1.** *Let $b < 2^{n-k-1} - 1$ for some $k \in \mathbb{N}$. Then, there are edges that have a negative occurrence record according to [3, Table 4].*

*Proof of Issue 4.1.* Let $i \in \{n-k, \dots, n-1\}$ and $j = 1$. By $b < 2^{n-k-1} - 1$ and $i \geq n-k$ it follows that $b < 2^i - 1$. This implies $b_i = 0$ and $b_{i+1} = 0 \neq 1 = j$. Since also $\tilde{b}_{i+1} = 0$ for all $\tilde{b} \leq b$, it follows that $g = g(b, i, \{(i+1, 1)\}) = 0$. In addition, since $b < 2^i - 1$ is equivalent to $2^i > b + 1$, we get

$$2z = 2(b - 2^{i-1}) = 2b - 2^i < 2b - (b+1) = b - 1,$$

or $z < \frac{1}{2}(b-1) = \frac{1}{2}(b-1-g)$. Thus, all conditions for the second case of Equation (4.2) are fulfilled, implying

$$\begin{aligned}
\phi^{\sigma_b}(b_{i,0}^j, A_i^j) + \phi^{\sigma_b}(b_{i,1}^j, A_i^j) &= g + 1 + 2z \\
&= 2z + 1 \\
&= 2(b - 2^{i-1}) + 1 \\
&< 2(2^{n-k-1} - 1 - 2^{i-1}) + 1 \\
&\leq 2(2^{n-k-1} - 1 - 2^{n-k-1}) + 1 \\
&= -1 < 0.
\end{aligned}$$

Hence there must be at least one edge that has a negative occurrence record. □

Our main goal of Section 4 is to show that the occurrence records can be described by the following system of equations:

$$|\phi^{\sigma_b}(b_{i,0}^j, A_i^j) - \phi^{\sigma_b}(b_{i,1}^j, A_i^j)| \leq 1 \tag{4.3}$$

$$\phi^{\sigma_b}(A_i^j) = \begin{cases} g+1, & A_i^j \text{ is closed and active} \\ b, & A_i^j \text{ is open and active} \\ b, & A_i^j \text{ is inactive and } b < 2^{i-1} + j \cdot 2^i \\ g + 1 + 2z, & A_i^j \text{ is inactive and } b \geq 2^{i-1} + j \cdot 2^i \end{cases} \tag{4.4}$$

For the sake of completeness we also restate the rules describing the application of the improving switches within the bicycles.

1. If $A_i^j$ is open and active, we switch one edge of the bicycle.

2. Let $j := b_{\ell(b+1)+1}$. In addition to the first rule, the second edge of $A_{\ell(b+1)}^j$ is switched.

3. If $A_i^j$ is inactive and $b < 2^{i-1} + j \cdot 2^i$, one edge of the bicycle is switched.

4. If $A_i^j$ is inactive, $b \geq 2^{i-1} + j \cdot 2^i$ and $z < \frac{1}{2}(b-1-g)$, both edges of $A_i^j$ are switched.

5. If $A_i^j$ is inactive, $b \geq 2^{i-1} + j \cdot 2^i$ and $z \geq \frac{1}{2}(b-1-g)$, only one edge of $A_i^j$ is switched.

Formally we show the following.

**Theorem 4.2.** *Suppose that the improving switches within the bicycles are applied as described by the rules 1 to 5. Let $b \in \mathbb{B}_n$ and $A_i^j$ be some bicycle. Then, Equations (4.3) and (4.4) correctly specify the occurrence record $\phi^{\sigma_b}(A_i^j)$.*

To simplify the proof, we introduce the following notion. Fix $b \in \mathbb{B}_n$ and a bicycle $A_i^j$. The bicycle $A_i^j$ is called *bicycle of type $k$ with respect to $\sigma_b$* when it fulfills the $k$-th condition mentioned in Equation (4.4) for $\sigma_b$. We additionally establish the following abbreviations and state a lemma that is implicitly contained in the proof of [3, Lemma 5].

- We define $g := g(b, i, \{(i+1, j)\}$, i.e., $g$ is the largest number smaller than or equal to $b$ such that $g_{i+1} = j, g_i = 1$ and $g_l = 0$ for all $l < i$. We define $g' := g(b+1, i, \{(i+1, j)\})$ analogously.
- We define $z := b - g - 2^{i-1}$ and $z' := b + 1 - g' - 2^{i-1}$ analogously.
- We define $\ell := \ell(b)$ and $\ell' := \ell(b+1)$.

**Lemma B.3** ([3])**.** *For every $b \in \mathbb{B}_n, i \in \{1, \ldots, n\}$ with $i \neq \ell(b+1)$ and $j \in \{0,1\}$ we have $g = g'$.*

We also make use of the following lemma.

**Lemma B.4.** *Let $b \in \mathbb{B}_n$ and $A_i^j$ be some bicycle. Then, the bicycle $A_i^j$ was closed at least once during the application of the Policy Iteration Algorithm upto policy $\sigma_b$ if and only if $b \geq 2^{i-1} + j \cdot 2^i$.*

*Proof.* The bicycle $A_i^j$ is closed the first time during the application of the Policy Iteration Algorithm when a number $\tilde{b} \leq b$ is reached such that $\tilde{b}_i = 1, \tilde{b}_{i+1} = j$ and $\tilde{b}_l = 0$ else. This number is exactly $2^{i-1} + j \cdot 2^i$. ☐

With this notation and Lemmas B.3 and B.4 in place, we now prove Theorem 4.2. Whenever we discuss how a bicycle should look like, we implicitly refer to the invariants introduced in Section 2.1 that describe $\sigma_b$.

*Proof of Theorem 4.2.* We prove

$$\phi^{\sigma_b}(A_i^j) \leq b + 1 \tag{B.3}$$

where equality holds if and only if $i = \ell(b)$ and $j = b_{\ell(b)+1}$ alongside since this statement is needed in some cases.

We show both statement via induction on $b$. Let $b = 0$. By the definition of both the original initial policy $\sigma^\star$ and the new initial policy $\sigma^*$, the target of $b_{i,l}^j$ under the corresponding policy is $t$ for all indices $i \in \{1, \ldots, n\}$ and $j, l \in \{0,1\}$. Therefore, all bicycles are open and either active or inactive, regardless which of the two initial policies is considered. As $b = 0$, the inequality $b < 2^{i-1} + j \cdot 2^i$ holds for all $i \in \{1, \ldots, n\}$ and $j \in \{0,1\}$. This implies that every bicycle is either of type 2 or of type 3. Therefore, for Equation (4.4) to hold, the occurrence record of every bicycle needs to be equal to $b = 0$. But, since we consider the initial policies, no improving switch was applied yet. Therefore, $\phi^{\sigma_0}(A_i^j) = 0$ for all bicycles $A_i^j$. Consequently, Equation (4.4) holds. We furthermore observe that there is no least significant set bit $\ell(b)$ since $b = 0$. Hence, since $\phi^{\sigma_0}(A_i^j) = 0 < b + 1$ for all bicycles $A_i^j$, and no bicycle is closed, Equation (B.3) holds as well.

Suppose that the statements holds for all numbers smaller or equal to $b \in \mathbb{B}_n$. We show that the two statements also holds for $b + 1$. We distinguish between the induction hypotheses with respect to Equation (4.4) and Equation (B.3) and always state to which we refer. We discuss Equation (4.3) at the end of the proof.

Fix a bicycle $A_i^j$ for some $i \in \{1, \ldots, n\}$ and $j \in \{0,1\}$. The proof is organized as follows. We distinguish all possible cases of which "state" (*open*, *active*, ...) the bicycle $A_i^j$ could be in with respect to $\sigma_b$. We then investigate of which type the bicycle is with respect to $\sigma_b$ and if this type changes when transitioning to $\sigma_{b+1}$. We then state how many improving switches we need to apply according to our rules and discuss what we need to show such that Equation (4.4) remains valid for the policy $\sigma_{b+1}$.

**Case 1:** $\boldsymbol{A_i^j}$ **is open, active and** $\boldsymbol{i = \ell'}$**.** Then $A_i^j$ is the active bicycle corresponding to the least significant set bit of $b+1$. By construction, it is open with respect to $\sigma_b$ but needs to be closed with respect to $\sigma_{b+1}$. As $b_{\ell'+1} = (b+1)_{\ell'+1}$, the bicycle remains active. Thus, $A_i^j$ is of type 1 with respect to $\sigma_{b+1}$. As we apply rules 1 and 2 and switch both edges of the bicycle, we thus need to show

$$\phi^{\sigma_b}(A_i^j) + 2 = g' + 1.$$

By the induction hypothesis (4.4), we have $\phi^{\sigma_b}(A_i^j) = b$ since $A_i^j$ is a type 2 bicycle with respect to $\sigma_b$. To show Equation (4.4), it therefore suffices to show $b+2 = g'+1$, or, equivalently, $g' = b+1$. This however follows since the binary representations of $g'$ and $b+1$ both end on the subsequence $(b_{\ell'+1}, 1, 0, \dots, 0)$ of length $\ell' + 1$.

Observe that we have $\phi^{\sigma_{b+1}}(A_i^j) = (b+1)+1$ after applying the two switches, hence Equation (B.3) remains valid as well.

**Case 2:** $\boldsymbol{A_i^j}$ **is open and active, but** $\boldsymbol{i \neq \ell'}$**.** We argue that $A_i^j$ remains open and active, i.e., $A_i^j$ is a bicycle of type 2 with respect to $\sigma_{b+1}$. By the definition of open and active, we have $b_i = 0$ and $j = b_{i+1}$. In addition, $b_1 = \dots = b_{\ell'-1} = 1$ since $\ell' = \ell(b+1)$. As all active bicycles corresponding to levels 1 to $\ell' - 1$ are closed in $\sigma$ and $i \neq \ell'$, this implies $i > \ell'$. Due to only the $\ell'$ least significant bits (i.e., the bits $b_1$ to $b_{\ell'}$) being switched, $A_i^j$ remains active with respect to $b+1$. Since the active bicycle of level $\ell'$ is the only bicycle that is open with respect to $\sigma_b$ but closed with respect to $\sigma_{b+1}$, $A_i^j$ remains open. Hence, since $A_i^j$ remains open and active, it is a bicycle of type 2 with respect to $\sigma_{b+1}$. Because we only apply one improving switch in the bicycle $A_i^j$ (rule 1), we therefore need to show that

$$\phi^{\sigma_b}(A_i^j) + 1 = b + 1.$$

By the induction hypothesis (4.4) $\phi^{\sigma_b}(A_i^j) = b$, so $\phi^{\sigma_b}(A_i^j)+1 = b+1$. Therefore, both Equations (4.4) and (B.3) still hold.

**Case 3:** $\boldsymbol{A_i^j}$ **is closed, active and** $\boldsymbol{i > \ell'}$**.** We show that $A_i^j$ is of type 1 with respect to $\sigma_{b+1}$. By the definition of closed and active, $b_i = 1$ and $b_{i+1} = j$. As only bits corresponding to indices smaller than $\ell'$ switch, $A_i^j$ remains active, cf. Case 2. By $i > \ell'$, it also remains closed since only the bits $b_1$ to $b_{\ell'-1}$ switch from 1 to 0 and thus only bicycles corresponding to these levels are opened during phase 3. Therefore, $A_i^j$ is a bicycle of type 1 with respect to $\sigma_{b+1}$ and none of the edges $(b_{i,0}^j, A_i^j), (b_{i,1}^j, A_i^j)$ are switched. We thus need need to show

$$\phi^{\sigma_b}(A_i^j) = g' + 1.$$

By the induction hypothesis (4.4), we have $\phi^{\sigma_b}(A_i^j) = g + 1$. It thus suffices to show $g+1 = g'+1$. Since $i \neq \ell'$, this follows from Lemma B.3. Therefore, Equation (4.4) still holds.

In addition, Equation (B.3) remains valid since $\phi^{\sigma_b}(A_i^j) \leq b$ holds by the induction hypothesis (B.3). Since $\phi^{\sigma_{b+1}}(A_i^j) = \phi^{\sigma_b}(A_i^j)$ holds by the argument above, we obtain $\phi^{\sigma_{b+1}}(A_i^j) < b+1$.

**Case 4:** $\boldsymbol{A_i^j}$ **is closed, active and** $\boldsymbol{i < \ell'}$**.** We show that $A_i^j$ is of type 4 with respect to $\sigma_{b+1}$. Because of $i < \ell'$, the bits $b_i$ and $b_{i+1}$ both switch. Thus, since $i < \ell'$ implies $b_i = 1$, we have $(b+1)_i = 0$. Hence $A_i^j$ is open with respect to $\sigma_{b+1}$. Since $A_i^j$ is active with respect to $\sigma_b$, we have $b_{i+1} = j$ and therefore, because bit $b_{i+1}$ switches, we obtain $(b+1)_{i+1} \neq j$. Thus, $A_i^j$ is inactive with respect to $\sigma_{b+1}$. Since $A_i^j$ is closed, Lemma B.4 implies $b \geq 2^{i-1} + j \cdot 2^i$. Therefore, $A_i^j$ is a bicycle of type 4 with respect to $\sigma_{b+1}$. Because $A_i^j$ is closed, we do not switch the edges of the bicycle and therefore need to show

$$\phi^{\sigma_b}(A_i^j) = g' + 1 + 2z'.$$

By the induction hypothesis (4.4), we have $\phi^{\sigma_b}(A_i^j) = g+1$. Thus, we need to show $g+1 = g'+1+2z'$. Since $i \neq \ell'$, Lemma B.3 implies $g = g'$. It thus suffices to show that $z' = b + 1 - g' - 2^{i-1} = 0$.

From the assumptions $i < \ell'$ and that $A_i^j$ is closed, we get $b_i = 1$. Since $A_i^j$ is also active by assumption, it follows that $j = b_{i+1}$. This implies that $g = (b_n, \dots, b_{i+1}, 1, 0, \dots, 0)$. Therefore, since $i < \ell'$ implies $b = (b_n, \dots, b_{i+1}, 1, 1, \dots, 1)$, we get $b - g = 2^{i-1} - 1$. Because $g = g'$ holds by

Lemma B.3, this can be formulated equivalently, obtaining the equality $z' = b + 1 - g' - 2^{i-1} = 0$. Thus Equation (4.4) remains valid.

As in Case 2, $\phi^{\sigma_{b+1}}(A_i^j) = \phi^{\sigma_b}(A_i^j)$ and since $\phi^{\sigma_b}(A_i^j) \leq b$ holds by the induction hypothesis (B.3), also Equation (B.3) follows.

<u>Case 5:</u> **$A_i^j$ is closed, active and $i = \ell'$.** This cannot happen since both bicycles of level $\ell'$ are open with respect to $\sigma_b$ since $b_{\ell'} = 0$.

<u>Case 6:</u> **$A_i^j$ is closed and inactive.** This cannot happen since closed bicycles are always active (see the invariants described in Section 2.1).

<u>Case 7:</u> **$A_i^j$ is inactive and $b < 2^{i-1} + j \cdot 2^i$.** Then, $A_i^j$ is a bicycle of type 3. We observe that $A_i^j$ being inactive implies that $A_i^j$ is open. We distinguish the type of $A_i^j$ is with respect to $\sigma_{b+1}$, since this is not clear in this case.

It cannot happen that $A_i^j$ is closed with respect to $\sigma_{b+1}$, because the active bicycle of level $\ell'$ is the only bicycle which is open with respect to $\sigma_b$ and closed with respect to $\sigma_{b+1}$, and $A_i^j$ is inactive by assumption.

Suppose that $A_i^j$ is a bicycle of type 3 with respect to $\sigma_{b+1}$. That is, it remains inactive with respect to $\sigma_{b+1}$ and $b + 1 < 2^{i-1} + j \cdot 2^i$ holds. As we apply one improving switch (rule 3), we thus need to show that

$$\phi^{\sigma_b}(A_i^j) + 1 = b + 1.$$

This follows immediately since $\phi^{\sigma_b}(A_i^j) = b$ by the induction hypothesis (4.4).

Suppose that $A_i^j$ is a bicycle of type 2 with respect to $\sigma_{b+1}$. That is, it is active and open with respect to $\sigma_{b+1}$. In this case, we also need to show

$$\phi^{\sigma_b}(A_i^j) + 1 = b + 1,$$

which also follows from the induction hypotheses (4.4).

Suppose that $A_i^j$ is a bicycle of type 4 with respect to $\sigma_{b+1}$. That is, it is inactive with respect to $\sigma_{b+1}$ and $b+1 \geq 2^{i-1} + j \cdot 2^i$. Then, since $b < 2^{i-1} + j \cdot 2^i$, it follows immediately that $b+1 = 2^{i-1} + j \cdot 2^i$. But, by Lemma B.4, this can only happen if the bicycle $A_i^j$ is closed during the transition from $\sigma_b$ to $\sigma_{b+1}$, contradicting the inactivity of $A_i^j$ with respect to $\sigma_b$.

Therefore, $\phi^{\sigma_b}(A_i^j) + 1 = b + 1$ holds in all cases, and both Equation (4.4) and Equation (B.3) stay valid.

<u>Case 8:</u> **$A_i^j$ is inactive, $b \geq 2^{i-1} + j \cdot 2^i$ and $z < \frac{1}{2}(b - 1 - g)$.** Then, $A_i^j$ is a bicycle of type 4 with respect to $\sigma_b$. We show that it also is a bicycle of type 4 with respect to $\sigma_{b+1}$, i.e., that $A_i^j$ is inactive with respect to $\sigma_{b+1}$ and $b+1 \geq 2^{i-1} + j \cdot 2^i$. It then remains to show that $\phi^{\sigma_b}(A_i^j) + 2 = g' + 1 + 2z'$, or, since $\phi^{\sigma_b}(A_i^j) = g + 1 + 2z$ by the induction hypothesis (4.4), that $g + 1 + 2z + 2 = g' + 1 + 2z'$.

Observe that $b+1 \geq 2^{i-1} + j \cdot 2^i$ immediately follows from $b \geq 2^{i-1} + j \cdot 2^i$. Towards a contradiction, assume that $A_i^j$ is active with respect to $\sigma_{b+1}$. Since only bits with an index smaller or equal to $\ell'$ are switched, only inactive bicycles on levels 1 to $\ell' - 1$ can become active. As a consequence, $i < \ell'$.

We next show that $b - g = 2^i + 2^{i-1} - 1$ holds. First assume that $i \neq \ell' - 1$. Then, since $i < \ell' - 1$ and $b = (b_n, \ldots, b_{\ell'+1}, 0, 1, \ldots, 1)$, it follows that $b_{i+1} = 1$. Hence, by the inactivity of $A_i^j$ with respect to $\sigma_b$, we obtain $j = 0$. Therefore,

$$g = (b_n, \ldots, b_{\ell'+1}, 0, 1, \ldots, 1, \underbrace{0}_{g_{i+1}}, \underbrace{1}_{g_i}, 0, \ldots, 0),$$

since we have $g_i = 1$ and $g_{i+1} = j = 0$ by definition of $g$. Consequently, $b - g = 2^i + 2^{i-1} - 1$.

Now assume that $i = \ell' - 1$. We then obtain $b_{i+1} = b_{\ell'} = 0$ and hence, by the inactivity of $A_i^j$, we get $j = 1$. Therefore,

$$g = (\tilde{b}_n, \ldots, \tilde{b}_{\ell'+1}, 1, \underbrace{1}_{g_i = g_{\ell'-1}}, 0, \ldots, 0)$$

where $(\tilde{b}_n, \ldots, \tilde{b}_{\ell'+1}) = (b_n, \ldots, b_{\ell'+1}) - 1$. This implies that $g + 2^i + 2^{i-1} = b + 1$ which is equivalent to $b - g = 2^i + 2^{i-1} - 1$.

Using the identities $b - g = 2^i + 2^{i-1} - 1$ and $\phi^{\sigma_b}(A_i^j) = b + 1 + 2z$ which follows from the induction hypothesis (4.4), we obtain the following estimation for $\phi^{\sigma_b}(A_i^j)$:

$$\begin{aligned}\phi^{\sigma_b}(A_i^j) &= g + 1 + 2(b - g - 2^{i-1}) \\ &= 2b - g - 2^i + 1 \\ &= b + 2^i + 2^{i-1} - 1 - 2^i + 1 \\ &= b + 2^{i-1} > b.\end{aligned}$$

Additionally, by assumption, $z < \frac{1}{2}(b - 1 - g)$, which implies

$$\phi^{\sigma_b}(A_i^j) = g + 1 + 2z < g + 1 + b - 1 - g = b, \tag{B.4}$$

contradicting the previous inequality. Therefore, the assumption of $A_i^j$ being active with respect to $\sigma_{b+1}$ cannot be correct, hence the bicycle must be inactive with respect to $\sigma_{b+1}$ and thus be of type 4.

As discussed before, we now need to show

$$\phi^{\sigma_b}(A_i^j) + 2 = g + 1 + 2z + 2 = g' + 2 + 2z'.$$

We observe that due to the inactivity of $A_i^j$ with respect to $\sigma_{b+1}$, we have $i \neq \ell'$ and therefore, by Lemma B.3, also $g = g'$. Therefore,

$$\begin{aligned}g + 1 + 2z + 2 &= g + 1 + 2b - 2g - 2^i + 2 \\ &= g' + 1 + 2(b + 1) - 2g - 2^i \\ &= g' + 1 + 2z',\end{aligned}$$

hence Equation (4.4) still holds.

It remains to show Equation (B.3). By Equation (B.4), we have $\phi^{\sigma_b}(A_i^j) < b$, and thus, by integrality, $\phi^{\sigma_b}(A_i^j) \leq b - 1$. Thus, $\phi^{\sigma_{b+1}}(A_i^j) = \phi^{\sigma_b}(A_i^j) + 2 \leq b - 1 + 2 = b + 1$ follows since we apply two switches in $A_i^j$.

<u>Case 9:</u> **$A_i^j$ is inactive, $b \geq 2^{i-1} + j \cdot 2^i$ and $z \geq \frac{1}{2}(b - 1 - g)$.** In this case, we do not distinguish the type of $A_i^j$ with respect to $\sigma_{b+1}$. Instead, we show $g + 1 + 2z = b$. This suffices because the bicycle $A_i^j$ cannot become closed and active (i.e., a bicycle of type 1) with respect to $\sigma_{b+1}$ and, by rule 5, the occurrence record of $A_i^j$ increases by 1. Therefore, we do not need to specify the type of $A_i^j$ if we are able to show that its occurrence record before applying the switch is equal to $b$.

To show that $g + 1 + 2z = b$, we need to show $z = \frac{1}{2}(b - 1 - g)$. Towards a contradiction, assume that $z > \frac{1}{2}(b - 1 - g)$. Then, since $A_i^j$ is a bicycle of type 4, by the induction hypothesis (4.4), we have $\phi^{\sigma_b}(A_i^j) = g + 1 + 2z$. Thus

$$\phi^{\sigma_b}(A_i^j) = g + 1 + 2z > g + 1 + b - 1 - g = b,$$

contradicting the induction hypothesis (B.3) requiring $\phi^{\sigma_b}(A_i^j) \leq b$. Therefore, equality holds, implying that $\phi^{\sigma_b}(A_i^j) = g + 1 + (b - 1 - g) = b$. As we apply a single switch, we finally obtain $\phi^{\sigma_b}(A_i^j) + 1 = b + 1$, as claimed.

As we have discussed all possible cases, we successfully showed that the occurrence records given in Equation (4.4) and the estimation given in Equation (B.3) hold. Because the switches can always be applied alternatingly within a bicycle, we can ensure that Equation (4.3) holds at all times during the application of the improving switches. ☐

## B.3 Proofs of Section 5

This subsection contains the proofs of the statement in Section 5 on the application of the improving switches of phase 3. Since we need to analyze the values of the vertices in more detail we need an additional lemma. This lemma is an extraction of some estimations contained in the proof of [3, Lemma 3].

**Lemma B.5.** *Let $\sigma$ be a policy calculated by the Policy Iteration Algorithm during the transition from $\sigma_b$ to $\sigma_{b+1}$. Denote the reward of each edge emanating from vertex $v$ by $\langle v\rangle$. Let*

$$S_i := \sum_{j\in\{i,\dots,n\}:b_j=1} \left(\langle k_j\rangle + \langle c_j^0\rangle + \langle d_j^0\rangle + \langle h_j^0\rangle\right) \quad \text{and} \quad T_i := \sum_{j\in\{i,\dots,n\}:(b+1)_j=1} \left(\langle k_j\rangle + \langle c_j^0\rangle + \langle d_j^0\rangle + \langle h_j^0\rangle\right).$$

*Then,*

$$\begin{aligned}
\mathrm{VAL}_\sigma(s) &\in [S_1, T_1]\\
\mathrm{VAL}_\sigma(k_i) &\in [\langle k_i\rangle + S_1, T_i]\\
\mathrm{VAL}_\sigma(h_i^j) &\in [\langle h_i^j\rangle + \langle k_{i+1}\rangle + S_1, \langle h_i^j\rangle + T_{i+1}]\\
\mathrm{VAL}_\sigma(d_i^j) &\in [\langle d_i^j\rangle + S_1, \langle d_i^j\rangle + \langle h_i^j\rangle + T_{i+1}]\\
\mathrm{VAL}_\sigma(A_i^j) &\in [S_1, \langle d_i^j\rangle + \langle h_i^j\rangle + T_{i+1}]\\
\mathrm{VAL}_\sigma(b_{i,r}^j) &\in [S_1, \langle d_i^j\rangle + \langle h_i^j\rangle + T_{i+1}]\\
\mathrm{VAL}_\sigma(c_i^j) &\in [\langle c_i^j\rangle + S_1, \langle c_i^j\rangle + \langle d_i^j\rangle + \langle h_i^j\rangle + T_{i+1}].
\end{aligned}$$

We first prove that the informal description on how to apply the improving switches is not correct.

**Issue 5.1.** *For every $b\in\{1,\dots,2^{n-2}-1\}$, the informal description of phase 3 described in [3, Pages 9–10] contradicts [3, Tables 2,4]. It additionally violates the* LEAST-ENTERED *pivot rule during the transition from $\sigma_b$ to $\sigma_{b+1}$ for every $b\in\{3,\dots,2^{n-2}-2\}$.*

*Proof of Issue 5.1.* Let $b\in\{1,\dots,2^{n-2}-1\}$. Consider the transition from policy $\sigma_b$ to policy $\sigma_{b+1}$. According to [3, Table 2], for each phase 1 or phase 2 policy $\sigma$, it should hold that $\sigma(k_i)=k_\ell$ if $b_i=0$ and $\sigma(k_i)=c_i^j, j=b_{i+1}$ if $b_i=1$. But, due to $b<2^{n-2}$, we have $\tilde{b}_n=0$ for all $\tilde{b}\in\{0,\dots,b\}$. In particular, $n>\ell(\tilde{b})$ for all of those $\tilde{b}$. Since phase 3 is the only phase in which the target of $k_n$ can be changed, this implies that the target of $k_n$ has never been changed. But for every policy $\sigma$ considered so far, $\sigma(k_n)=t$ held due to $\sigma^\star(k_n)=\sigma^*(k_n)=t$. Since $\sigma_b$ is a phase 1 policy by definition, this contradicts [3, Table 2], even if we change the initial policy as discussed in Section 3. Note that we obtain$\mathrm{VAL}_{\sigma_b}(k_n)=0$ for all $b\in\{1,\dots,2^{n-2}-1\}$ by the same arguments.

As a consequence, the occurrence records of all edges $(k_n,k_i)$ for $i\in\{1,\dots,n-1\}$ are zero. We now discuss how this violates [3, Table 4]. Let $i\in\{1,\dots,\lfloor\log_2(b)\rfloor+1\}$, i.e., consider some $i$ such that $b\geq 2^{i-1}$. Then $\phi^{\sigma_b}(k_n,k_i)=f(b,i,\{(n,0)\})$ should hold according to [3, Table 4]. But, due to $\tilde{b}_n=0$ for all $\tilde{b}\leq b$, we have $f(b,i,\{(n,0)\})=f(b,i)$. Thus, by Lemma B.1 (3) and since $b\geq 2^{i-1}$, we have

$$f(b,i,\{(n,0)\})=f(b,i)=\left\lfloor\frac{b+2^{i-1}}{2^i}\right\rfloor\geq\left\lfloor\frac{2^{i-1}+2^{i-1}}{2^i}\right\rfloor=1.$$

This contradicts the occurrence records of all edges $(k_n,k_i)$ for $i\in\{1,\dots,n-1\}$ being zero.

It remains to show that applying the improving switches as described before contradicts the LEAST-ENTERED rule. We do so by showing that the edge $(k_n,k_1)$ is an improving switch throughout the whole transition from $\sigma_2$ to $\sigma_3$, and discuss the case of $b\in\{3,\dots,2^{n-2}-2\}$ afterwards. By [3, Table 4], $L_\sigma^5=\{(s,k_{\ell'})\}$ for any phase 5 policy $\sigma$. Since only switches contained in the subsets $L_\sigma^p$ are chosen as improving switches, this implies that $(s,k_1)$ is chosen in phase 5 of the transition from $\sigma_2$ to $\sigma_3$. But, since $\ell(1)=\ell(3)=1$, this edge has already been chosen in phase 5 of the transition from $\sigma_0$ to $\sigma_1$. Therefore, the edge has a non-zero occurrence record throughout the transition from $\sigma_2$ to $\sigma_3$. Thus, the result follows once we showed that $(k_n,k_1)$ is an improving switch, since we already observed that it has an occurrence record of zero but is not switched.

Consider $\sigma_b$ for $b=2$. The only set bit in the binary representation of $b$ is $b_2$. As observed before, we have $\sigma_2(k_n)=t$, implying $\mathrm{VAL}_{\sigma_2}(\sigma_2(k_n))=0$. In addition, by Lemma B.5, for every policy $\sigma$ calculated during the transition from $\sigma_2$ to $\sigma_3$, it holds that

$$\begin{aligned}\mathrm{VAL}_{\sigma_2}(k_1) &\geq \langle k_1\rangle + S_1\\ &= (-N)^{2\cdot 1+7} + S_1\\ &\geq \sum_{j\in\{1,\ldots,n\}:b_j=1} \left[(-N)^{2j+7} + (-N)^{2j+8} + (-N)^7 + (-N)^6\right] - N^9\\ &= (-N)^{2\cdot 2+7} + (-N)^{2\cdot 2+8} + (-N)^7 + (-N)^6 - N^9\\ &= N^{12} - N^{11} - N^9 - N^7 - N^6 > 0,\end{aligned}$$

since $N \geq 8$. Thus, $(k_n, k_1)$ is an improving switch during the whole transition from $\sigma_2$ to $\sigma_3$.

Since $\mathrm{VAL}_{\sigma_b}(k_n) = 0$ for all $b \in \{3, \ldots, 2^{n-2} - 2\}$ as discussed before, since $\ell(b) \neq n$ for all of those $b$, and since the values are non-decreasing, $(k_n, k_1)$ remains an improving switch for all $b \in \{3, \ldots, 2^{n-2}-2\}$. We further observe that due to $b \geq 3$, both of the bicycles $A_1^0$ and $A_1^1$ have been closed at least once, see Lemma B.4. This implies that all edges of these bicycles have an occurrence of at least one. Also, at least one of the edges of the inactive bicycle of level 1 is switched when transitioning from $\sigma_b$ to $\sigma_{b+1}$ for any $b \in \mathbb{B}_n$. Because this edge has a non-zero occurrence record whereas the edge $(k_n, k_1)$ has an occurrence record of zero and is an improving switch, this shows that following the informal description contradicts the LEAST-ENTERED pivot rule at least once during the transition from $\sigma_b$ to $\sigma_{b+1}$ for every $b \in \{3, \ldots, 2^{n-1} - 2\}$. □

The next lemmas need a partition of the subset $L_\sigma^3$ of the set of improving switches for a phase 3 policy $\sigma$ into three sets $L_\sigma^{3,1}, L_\sigma^{3,2}$ and $L_\sigma^{3,3}$.

These sets are defined as follows (cf. [3, Table 3]):

- $L_\sigma^{3,1} \coloneqq \{(k_i, k_{\ell'}) : \sigma(k_i) \neq k_{\ell'} \wedge (b+1)_i = 0\}$
- $L_\sigma^{3,2} \coloneqq \{(b_{i,l}^j, k_{\ell'}) : \sigma(b_{i,l}^j) \neq k_{\ell'} \wedge (b+1)_i = 0\}$
- $L_\sigma^{3,3} \coloneqq \{(b_{i,l}^j, k_{\ell'}) : \sigma(b_{i,l}^j) \neq k_{\ell'} \wedge (b+1)_{i+1} \neq j\}$

Note that we use a different notation than Friedmann in order to avoid using the function $\bar{\sigma}$. The following lemmas will be needed within some of the next proofs.

**Lemma B.6.** *Let $b \in \mathbb{B}_n$ and let $\sigma$ be the first phase 3 policy of the transition from $\sigma_b$ to $\sigma_{b+1}$. Then $L_\sigma^3 = L_{\sigma_b}^3$, and $L_{\sigma_b}^3$ is the set of improving switches that should be applied during phase 3 according to [3, Table 2].*

**Lemma B.7.** *Let $\sigma$ be a phase 3 policy and let $e \in L_\sigma^3$. Then $L_{\sigma[e]}^3 = L_\sigma^3 \setminus \{e\}$.*

*Proof.* We only discuss the case $e \in L_\sigma^{3,1}$ – the cases $e \in L_\sigma^{3,2}$ and $e \in L_\sigma^{3,3}$ follow from similar arguments. Let $e \in L_\sigma^{3,1}$. Then, $e = (k_i, k_{\ell'})$ for some $i \in \{1, \ldots, n\}$ with $\sigma(k_i) \neq k_{\ell'}$ and $(b+1)_i = 0$. Hence the improving switch $(k_i, k_{\ell'})$ can be applied in $\sigma$. When the switch $e$ is applied, we have $\sigma[e](k_i) = k_{\ell'}$ for the resulting policy $\sigma[e]$. This immediately implies that $e \notin L_{\sigma[e]}^{3,1}$ and thus $e \notin L_{\sigma[e]}^3$.

Let $\tilde{e} \in L_\sigma^3$ and $\tilde{e} \neq e$. We show that $\tilde{e} \in L_{\sigma[e]}^3$. Since $\tilde{e} \in L_\sigma^3$, we have $\tilde{e} = (x, k_{\ell'})$ where either $x = k_{i'}$ or $x = b_{i',r}^j$ for some $i' \in \{1, \ldots, n\}$ and $r, j \in \{0, 1\}$. In addition, since $\tilde{e} \in L_\sigma^3$, we have $\sigma(x) \neq k_{\ell'}$. The switch $(k_i, k_{\ell'})$ is the only switch that we apply when transitioning from $\sigma$ to $\sigma[e]$. Therefore, $\sigma(x) \neq k_{\ell'}$ implies $\sigma[e](x) \neq k_{\ell'}$ as the target of no vertex other than $k_i$ changes. As furthermore the conditions $(b+1)_i = 0$ and $b_{i+1} \neq j$ remain valid, it follows that $\tilde{e} \in L_{\sigma[e]}^3$. This implies that $L_\sigma^3 \subseteq L_{\sigma[e]}^3 \cup \{e\}$.

Towards a contradiction, assume that there is some edge $\tilde{e} \in L_{\sigma[e]}^3 \cup \{e\}$ but $\tilde{e} \notin L_\sigma^3$. Then, since $e \in L_\sigma^3$, we have that $e \neq \tilde{e}$. Thus, $\tilde{e} = (x, k_{\ell'})$ for some $x$ as in the last case and $\sigma[e](x) \neq k_{\ell'}$. But since $(k_i, k_{\ell'})$ is the only switch that is applied when transitioning from $\sigma$ to $\sigma[e]$, this implies that $\sigma(x) \neq k_{\ell'}$. But then, $e \in L_\sigma^3$ which is a contradiction. We therefore have $L_{\sigma[e]}^3 \cup \{e\} \subseteq L_\sigma^3$ and thus $L_{\sigma[e]}^3 \cup \{e\} = L_\sigma^3$. □

**Corollary B.8.** *Let $\sigma$ be a phase 3 policy and $e \in I_\sigma$ an improving switch for $\sigma$. Let $\sigma'$ be a phase 3 policy reached after $\sigma$ during the same transition. If the switch $e$ was not applied when transitioning from $\sigma$ to $\sigma'$, then $e$ is an improving switch for $\sigma'$.*

The following lemma is crucial to prove Issues 5.2 and 5.3.

**Lemma B.9.** *Let $i \in \{2, \ldots, n-2\}$ and $l < i$. Then, there is a number $b \in \mathbb{B}_n$ with $\ell(b+1) = l$ such that for all $j \in \{i+2, \ldots, n\}$, it holds that $\phi^{\sigma_b}(k_i, k_{\ell'}) < \phi^{\sigma_b}(k_j, k_{\ell'})$ and $(k_i, k_{\ell'}), (k_j, k_{\ell'}) \in L^3_{\sigma_b}$.*

*Proof.* Let $b := 2^i + 2^{l-1} - 1$ and $j \in \{i+2, \ldots, n\}$. Then, $\ell(b+1) = \ell(2^i + 2^{l-1}) = l$ since $i < l$. Furthermore, $j \geq i+2, i > l$ and $i \geq 2$ imply

$$b + 1 = 2^i + 2^{l-1} \leq 2^i + 2^{i-2} \leq 2^{j-2} + 2^{j-4} < 2^{j-1} - 1.$$

Now consider the set $F(b, l)$ containing all $\tilde{b} \leq b$ such that $\ell(\tilde{b}) = l$, see Definition 2.2. We remind here that, by definition, $|F(b, l)| = f(b, l)$. Because $b < 2^{j-1}$, it holds that $\tilde{b}_j = 0$ for all $\tilde{b} \leq b$ and thus, $F(b, l) = F(b, l, \{(j, 0)\})$. Thus, by [3, Table 4] we have

$$\phi^{\sigma_b}(k_j, k_{\ell'}) = \phi^{\sigma_b}(k_j, k_l) = f(b, l, \{(j, 0)\}) = f(b, l).$$

In addition, since $b + 1 < 2^{j-1} - 1$ and thus $(b+1)_j = 0$ and $\sigma_b(k_j) = k_\ell \neq k_{\ell'}$ holds due to the invariants discussed in Section 2, we have $(k_j, k_{\ell'}) \in L^3_{\sigma_b}$. However, because of $b > 2^i, i \geq 2$ and $i > l$, it holds that $\tilde{b} := 2^{i-1} + 2^{l-1} \in F(b, l)$ since $\tilde{b} \leq b$. Additionally, we have that $\tilde{b}_i = 1$. As a consequence, $\tilde{b} \notin F(b, l, \{(i, 0)\}))$. But this implies that $F(b, l, \{(i, 0)\}) \subsetneq F(b, l)$. Since $\phi^{\sigma_b}(k_i, k_l) = f(b, l, \{(i, 0)\})$ and, $|F(b, l\{(i, 0)\})| = f(b, l, \{(i, 0)\})$ hold by [3, Table 4], this implies

$$\phi^{\sigma_b}(k_i, k_{\ell'}) = \phi^{\sigma_b}(k_i, k_l) = f(b, l, \{(i, 0)\}) < f(b, l) = \phi^{\sigma_b}(k_j, k_l) = \phi^{\sigma_b}(k_j, k_{\ell'}).$$

Since $(b+1)_i = b_i = 0$ due to $i > l = \ell(b+1)$ and $\sigma_b(k_i) = k_\ell \neq k_{\ell'}$, we also have $(k_i, k_{\ell'}) \in L^3_{\sigma_b}$. □

With all of these lemmas in place we are now able to prove that the more detailed explanation on how the improving switches should be applied during phase 3 is not correct.

**Issue 5.2.** *Applying the improving switches as described in [3, Lemma 5] does not obey the* Least-Entered *pivot rule.*

*Proof.* According to [3, Lemma 5], the improving switches of phase 3 should be applied as follows[2]: "*[...] we need to perform all switches from higher indices to smaller indices, and $k_i$ to $k_{\ell'}$ before $b^j_{i,l}$ with $(b+1)_{i+1} \neq j$ or $(b+1)_i = 0$ to $k_{\ell'}$*". This description is also further formalized in the side conditions of [3, Table 2].

Let $i \in \{2, \ldots, n-2\}, l < i$ and $j \in \{i+2, \ldots, n-2\}$. By Lemma B.9, there is a number $b \in \mathbb{B}_n$ such that $l = \ell(b+1)$ and $\phi^{\sigma_b}(k_i, k_{\ell'}) < \phi^{\sigma_b}(k_j, k_{\ell'})$. In addition, $(k_i, k_{\ell'}), (k_j, k_{\ell'}) \in L^3_{\sigma_b}$. Therefore, by Lemma B.6, the switch $(k_j, k_{\ell'})$ should be applied before the switch $(k_i, k_{\ell'})$ during the transition from $\sigma_b$ to $\sigma_{b+1}$ when following the description of [3].

Consider the phase 3 policy $\sigma$ of this transition in which the switch $(k_j, k_{\ell'})$ should be applied. Then, since $j > i$ an we "*perform all switches from higher indices to smaller indices*", the switch $(k_i, k_{\ell'})$ was not applied yet. But, by Lemma B.8, it is an improving switch for the current policy $\sigma$. This implies $\phi^{\sigma_b}(k_j, k_{\ell'}) = \phi^{\sigma}(k_j, k_{\ell'})$ and $\phi^{\sigma_b}(k_i, k_{\ell'}) = \phi^{\sigma}(k_i, k_{\ell'})$. Consequently, $\phi^{\sigma}(k_i, k_{\ell'}) < \phi^{\sigma}(k_j, k_{\ell'})$. Thus, since the edge $(k_i, k_{\ell'})$ is an improving switch for $\sigma$ having a lower occurrence record than $(k_j, k_{\ell'})$ and $\sigma$ was chosen as the policy in which $(k_j, k_{\ell'})$ should be applied, the Least-Entered rule is violated. □

The following lemma is used to prove Issue 5.3 by combining it with Lemma B.9.

**Lemma B.10.** *Assume that for any transition, the switches that should be applied during phase 3 were applied in some (possibly changing) order. Let $i \in \{2, \ldots, n-2\}$ and $l < i$. Then there is a number $b \in \mathbb{B}_n$ with $\ell(b+1) = l$ such that $\phi^{\sigma_b}(k_{i+1}, k_{\ell'}) < \phi^{\sigma_b}(b^1_{i,r}, k_{\ell'})$, where $r \in \{0, 1\}$ is arbitrary and $(k_{i+1}, k_{\ell'}), (b^1_{i,r}, k_{\ell'}) \in L^3_{\sigma_b}$.*

*Proof.* Since we assume that the same switches are applied during phase 3, the occurrence records given in [3, Table 4] remain valid. For now, consider some $b \in \mathbb{B}_n$ with $\ell(b+1) = l$. We fix its value later. By [3, Table 4] and since $\ell' = \ell(b+1) = l$,

$$\phi^{\sigma_b}(k_{i+1}, k_{\ell'}) = f(b, \ell', \{(i+1, 0)\})$$

and

$$\phi^{\sigma_b}(b^1_{i,r}, k_{\ell'}) = f(b, \ell', \{(i, 0)\}) + f(b, \ell', \{(i, 1), (i+1, 0)\}).$$

By Lemma B.1 (3),

$$f(b,\ell',\{(i,0)\}) = f(b,\ell',\{(i,0),(i+1,0)\}) + f(b,\ell',\{(i,0),(i+1,1)\}).$$

This implies that $\phi^{\sigma_b}(b_{i,r}^1, k_{\ell'})$ can be formulated equivalently as

$$f(b,\ell',\{(i,0),(i+1,0)\}) + f(b,\ell',\{(i,0),(i+1,1)\}) + f(b,\ell',\{(i,1),(i+1,0)\}).$$

Since $f(b,\ell',\{(i,1),(i+1,0)\}) + f(b,\ell',\{(i,0),(i+1,0)\}) = f(b,\ell',\{(i+1,0)\})$, the whole inequality can thus be formulated as

$$f(b,\ell',\{(i+1,0)\}) < f(b,\ell',\{(i+1,0)\}) + f(b,\ell',\{(i,0),(i+1,1)\}).$$

It thus suffices to find $b \in \mathbb{B}_n$ with $f(b,\ell',\{(i,0),(i+1,1)\}) > 0$, $\ell(b+1) = l$ and $(k_{i+1}, k_{\ell'}), (b_{i,r}^1, k_{\ell'}) \in L_{\sigma_b}^3$.

We show that $b := 2^{i+1} + 2^{l-1} - 1$ fulfills this. We observe that $\ell(b+1) = \ell(2^{i+1} + 2^{l-1}) = l$ since $l < i$. In addition, since $b_{i+1} = 0$, it holds that $\sigma_b(k_{i+1}) = k_\ell \neq k_{\ell'}$. Since also $(b+1)_{i+1} = 0$, we therefore have $(k_{i+1}, k_{\ell'}) \in L_{\sigma_b}^3$. Also, since $(b+1)_{i+1} = 0 \neq 1$ and $\sigma_b(b_{i,r}^1) = k_\ell \neq k_{\ell'}$, we additionally have $(b_{i,\ell'}^1, k_{\ell'}) \in L_{\sigma_b}^3$.

Now consider the number $\tilde{b} := 2^i + 2^{l-1} \in \mathbb{B}_n$. Then, $\tilde{b}_i = 0$ and $\tilde{b}_{i+1} = 1$. Since $\tilde{b} < b$, this implies $f(b,\ell',\{(i,0),(i+1,1)\}) \geq 1$. □

The following issue shows that no consistent ordering exists that updates the MDP level by level in each phase according to a fixed order.

**Issue 5.3.** *Suppose that the improving switches of phase 3 are applied one level after another as described above. That is, the ordering of the levels in the transition from $\sigma_b$ to $\sigma_{b+1}$ may only depend on $\ell(b+1)$. Then, the* LEAST-ENTERED *pivot rule is violated.*

*Proof.* To prove Issue 5.3, we show that applying the improving switches as discussed before violates Zadeh's LEAST-ENTERED rule several times by showing the following statement: Let $S^i$ be an ordering of $\{1,\dots,n\}$ for $i \in \{1,\dots,n\}$. Suppose that the improving switches of phase 3 of the transition from $\sigma_b$ to $\sigma_{b+1}$ are applied in the order defined by $S^{\ell(b+1)}$ for all $b \in \mathbb{B}_n$. Then, for every possible least significant bit $l \in \{1,\dots,n-4\}$, assuming that the ordering $S^l$ obeys the LEAST-ENTERED rule results in a contradiction.

We first observe that Lemma B.9 also holds when the improving switches are applied in some arbitrary order since we always consider the occurrence record with respect to $\sigma_b$.

Fix some $l \in \{1,\dots,n-4\}$. Consider the ordering $S^l = (s_1,\dots,s_n)$. For $k \in \{1,\dots,n\}$, we denote the position of $k$ within $S^l$ by $k^\star$, i.e., $k^\star$ is defined such that $s_{k^\star} = k$. Assume that applying the improving switches level by level according to the ordering $S^l$ obeys the LEAST-ENTERED rule. We show that this assumption yields both $(l+1)^\star < (n-1)^\star$ and $(n-1)^\star < (l+1)^\star$ which clearly is a contradiction.

Let $i \in \{l+1,\dots,n-2\}$. Then, $i > l$ and therefore, by Lemma B.10, there is a number $b \in \mathbb{B}_n$ with $\ell(b+1) = \ell' = l$ and $\phi^{\sigma_b}(k_{i+1}, k_{\ell'}) < \phi^{\sigma_b}(b_{i,r}^1, k_{\ell'})$ such that $(k_{i+1}, k_{\ell'}), (b_{i,r}^0, k_{\ell'}) \in L_{\sigma_b}^3$. Therefore, by Lemma B.6, both switches need to be applied during the transition from $\sigma_b$ to $\sigma_{b+1}$. Because of $\phi^{\sigma_b}(k_{i+1}, k_{\ell'}) < \phi^{\sigma_b}(b_{i,r}^1, k_{\ell'})$, level $i+1$ needs to appear before level $i$ within the ordering $S^l$. Since Lemma B.10 can be applied for all $i \in \{l+1,\dots,n-2\}$, this implies that the sequence $(n-1, n-2, \dots, l+1)$ needs to be a (not necessarily consecutive) subsequence of $S^l$. In particular, $(n-1)^\star < (l+1)^\star$ since we have $l+1 \neq n-1$ by assumption.

Now, let $i = l+1$ and $j \in \{i+2,\dots,n\}$. Then, by Lemma B.9, there is a number $b \in \mathbb{B}_n$ with $\ell(b+1) = l$ such that $\phi^{\sigma_b}(k_i, k_{\ell'}) < \phi^{\sigma_b}(k_{i+2}, k_{\ell'})$ and $(k_i, k_{\ell'}), (k_{i+2}, k_{\ell'}) \in L_{\sigma_b}^3$.

Again, by Lemma B.6, both switches need to be applied during the transition from $\sigma_b$ to $\sigma_{b+1}$. Therefore, for all $i \in \{l+1,\dots,n-2\}$, level $i$ needs to appear before any of the levels level $j \in \{i+2,\dots,n\}$ within $S^l$. But this implies that the sequence $(l+1, l+3, l+4, \dots, n-1, n)$ needs to be a (not necessarily consecutive) subsequence of $S^l$. In particular, $(l+1)^\star < (n-1)^\star$ since $n-1 \geq l+3$ as we have $l \leq n-4$ by assumption. This however contradicts $(n-1)^\star < (l+1)^\star$.

Therefore, applying the improving switches level by level according to the ordering $S^l$ does not obey the LEAST-ENTERED rule. □

The following statements are used to prove that there is a way to apply the improving switches of phase 3 such that the LEAST-ENTERED rule is obeyed. We begin with an estimation of the occurrence records of these switches.

**Lemma 5.4.** *Let $\sigma$ be a phase 3 policy. Then* $\max_{e\in L_\sigma^3} \phi^\sigma(e) \le f(b,\ell')$.

*Proof.* As discussed previously, $L_\sigma^3$ can be partitioned into three subsets $L_\sigma^{3,1}, L_\sigma^{3,2}$ and $L_\sigma^{3,3}$. It thus suffices to distinguish three cases. The last two cases can be discussed together as the occurrence records of edges contained in $L_\sigma^{3,2}$ and $L_\sigma^{3,3}$ are the same, see [3, Table 4].

<u>Case 1:</u> $\boldsymbol{e \in L_\sigma^{3,1}}$**.** Then, $e = (k_i, k_{\ell'})$, where $\sigma(k_i) \neq k'_\ell$ and $(b+1)_i = 0$ holds. The first of these conditions implies that the switch $e$ was not applied yet during the transition from $\sigma_b$ to $\sigma_{b+1}$. We therefore have $\phi^\sigma(k_i, k_{\ell'}) = \phi^{\sigma_b}(k_i, k_{\ell'})$. Since $\phi^{\sigma_b}(e) = f(b, \ell', \{(i,0)\})$ by [3, Table 4], this implies $\phi^\sigma(e) = f(b,\ell',\{(i,0)\})$. By Lemma B.1 (3), we therefore have

$$\phi^\sigma(e) = f(b,\ell',\{(i,0)\}) = f(b,\ell') - f(b,\ell',\{i,1\}) \le f(b,\ell').$$

<u>Case 2:</u> $\boldsymbol{e \in L_\sigma^{3,2}}$ **or** $\boldsymbol{e \in L_\sigma^{3,3}}$**.** Then, $e = (b_{i,r}^j, k_{\ell'})$, where $\sigma(b_{i,r}^j) \neq k_{\ell'}$ and we either have $(b+1)_i = 0$ if $e \in L_\sigma^{3,2}$ or $(b+1)_{i+1} \neq j$ if $e \in L_\sigma^{3,3}$. The first condition implies that the switch $e$ was not applied yet during the transition from $\sigma_b$ to $\sigma_{b+1}$. Hence $\phi^\sigma(b_{i,r}^j, k_{\ell'}) = \phi^{\sigma_b}(b_{i,r}^j, k_{\ell'})$. Since

$$\phi^{\sigma_b}(e) = f(b,\ell',\{(i,0)\}) + f(b,\ell',\{(i,1),(i+1,1-j)\})$$

by [3, Table 4], this implies

$$\phi^\sigma(e) = f(b,\ell',\{(i,0)\}) + f(b,\ell',\{(i,1),(i+1,1-j)\}).$$

By Lemma B.1 (2), it also holds that $f(b,\ell',\{(i,1),(i+1,1-j)\}) \le f(b,\ell',\{(i,1)\})$. Thus, we obtain

$$\phi^\sigma(e) \le f(b,\ell',\{(i,0)\} + f(b,\ell',\{(i,1)\}) = f(b,\ell'). \qquad \square$$

The next statements briefly point out a minor mistake contained in the definition of the set $L_\sigma^6$ and corrects it.

**Issue 5.5.** *For every $b \in \mathbb{B}_n$ with $\ell(b+1) > 1$, there is an improving switch that should be applied in phase 6 of the transition from $\sigma_b$ to $\sigma_{b+1}$ but is not contained in the set $L_\sigma^6$ for any phase 6 policy $\sigma$.*

*Proof.* Fix some $b \in \mathbb{B}_n$ such that $\ell' = \ell(b+1) > 1$. Consider the vertex $d_{\ell'-1}^0$. We show that the switch $(d_{\ell'-1}^0, s)$ needs to be applied during phase 6 of the transition from $\sigma_b$ to $\sigma_{b+1}$ but is not contained in $L_\sigma^6$ for any phase 6 policy $\sigma$. By analyzing [3, Table 2] and the function $\bar\sigma$ that is used in this table, it can be shown that $b_\ell = 0$ implies $\sigma_b(d_{\ell'-1}^0) = h_i^0$. Since the $\ell'$-th bit switches during the transition from $\sigma_b$ to $\sigma_{b+1}$, by [3, Table 2], $\sigma_{b+1}(d_{\ell'-1}^0) = s$ needs to hold. Therefore, $(d_{\ell'-1}^0, s)$ needs to be an improving switch for some policy $\sigma$ calculated during the transition from $\sigma_b$ to $\sigma_{b+1}$.

Towards a contradiction, assume that there was a policy $\sigma$ in which the switch $(d_{\ell'-1}^0, s)$ should be applied. Since the subsets of phase 6 policies are the only subsets that can contain this switch, $\sigma$ needs to be a phase 6 policy. By [3, Lemma 4], $(d_{\ell'-1}^0, s) \in L_\sigma^6$ then holds for this policy $\sigma$. Again analyzing the function $\bar\sigma$, it can be shown that due to $(d_{\ell'-1}^0, s) \in L_\sigma^6$, both $\sigma(d_{\ell'-1}^0) \neq s$ and $\sigma(d_{\ell'-1}^0) = s$ need to hold. This is clearly a contradiction. As a consequence, there is no policy $\sigma$ for which the switch $(d_{\ell'-1}^0, s)$ should be applied. $\square$

**Theorem 5.6.** *For any phase 6 policy $\sigma$, the subset of $I_\sigma$ in [3, Table 3] needs to be*

$$\bar L_\sigma^6 := \left\{(d_i^0, x) : \sigma(d_i^0) \neq x \wedge \sigma(d_i^0) = \begin{cases} h_i^0, & (b+1)_{i+1} = 1 \\ s, & (b+1)_{i+1} = 0 \end{cases}\right\} \cup$$
$$\left\{(d_i^1, x) : \sigma(d_i^1) \neq x \wedge \sigma(d_i^1) = \begin{cases} s, & (b+1)_{i+1} = 1 \\ h_i^0, & (b+1)_{i+1} = 0 \end{cases}\right\}.$$

*Proof.* Let $\sigma$ be a phase 6 policy. We show that when we assume that the switch $(d_{\ell'-1}^0, s)$ was not applied yet, it holds that $(d_{\ell'-1}^0, s) \in \bar L_\sigma^6$. For all other edges contained in $\bar L_\sigma^6$, the statement can be shown in a similar way.

As discussed when proving Issue 5.5, $\sigma_b(d_{\ell'-1}^0) = h_i^0$ holds and $\sigma_{b+1}(d_{\ell'-1}^0) = s$ needs to hold. Since $(d_{\ell'-1}^0, s)$ was not applied yet by assumption, $\sigma(d_{\ell'-1}^0) = \sigma_b(d_{\ell-1'}^0) = h_i^0$. In particular, $\sigma(d_{\ell'-1}^0) \neq s$. But we have$(b+1)_{\ell'-1+1} = (b+1)_{\ell'} = 1$ by the definition of $\ell'$. Therefore, $e \in \bar L_\sigma^6$. $\square$

We now proceed to show bounds on the occurrence records of the improving switches that should be applied in or after phase 3.

**Lemma 5.7.** Let $\sigma$ be a phase 3 policy. Assume that the Policy Iteration Algorithm is started with the policy $\sigma^*$ introduced in Definition 3.3. Then $\min_{e\in L^4_\sigma\cup L^5_\sigma\cup L^6_\sigma}\phi^\sigma(e)\geq f(b,\ell')$.

*Proof.* The policy $\sigma$ is calculated after the policy $\sigma_b$. Thus, $\phi^\sigma(e)\geq\phi^{\sigma_b}(e)$ holds for all edges $e$. It therefore suffices to show $\phi^{\sigma_b}(e)\geq f(b,\ell')$ for all $e\in L^4_\sigma\cup L^5_\sigma\cup L^6_\sigma$. Note that the conditions that we give here are not exactly the same as those given in [3], since we omit the additional notation $\bar{\sigma}$. They are, however, equivalent. We distinguish three cases.

**Case 1:** $\boldsymbol{e\in L^4_\sigma}$**.** Then, by [3, Table 3], $e=(h^0_i,k_{\nu^n_{i+2}(b+1)})$ for some $i\in\{1,\dots,n\}$ and, in addition, $\sigma(h^0_i)\notin\{k_{\nu^n_{i+2}(b+1)},t\}$. Since $\sigma(h^0_i)\neq t$ and by the way the improving switches are applied, there needs to be a next bit equal to 1 with an index of at least $i+2$.

Since $\ell'$ is the least significant bit of $b+1$, we have $b_j=(b+1)_j$ for all $j\in\{\ell'+1,\dots,n\}$. Therefore, the bit equal to 1 with an index of at least $i+2$ does not change if $i\geq\ell'-1$. More formally, $\nu^n_{j+2}(b)=\nu^n_{j+2}(b+1)$ holds for all $j\in\{\ell'-1,\dots,n-2\}$. Thus, $i\leq\ell'-2$ needs to hold since otherwise, $\sigma(h^0_i)=k_{\nu^n_{i+2}(b+1)}$, contradicting that $(h^0_i,k_{\nu^n_{i+2}(b+1)})$ is an improving switch. As also $(b+1)_j=0$ for $j<\ell'$, it follows that $\nu^n_{j+2}(b+1)=\ell'$ for all $j\in\{1,\dots,\ell'-2\}$. Thus, $e=(h^0_i,k_{\ell'})$ for some $i\in\{1,\dots,\ell'-2\}$, and, by [3, Table 4],

$$\phi^{\sigma_b}(e)=\phi^{\sigma_b}(h^0_i,k_{\ell'})=f(b,\ell').$$

**Case 2:** $\boldsymbol{e\in L^5_\sigma}$**.** Then, by [3, Table 3] and since $L^5_\sigma=\{(s,k_{\ell'})\}$, we have $e=(s,k_{\ell'})$. Therefore, because $\phi^{\sigma_b}(s,k_{\ell'})=f(b,\ell')$ by [3, Table 4] it holds that

$$\phi^{\sigma_b}(e)=\phi^{\sigma_b}(s,k_{\ell'})=f(b,\ell').$$

**Case 3:** $\boldsymbol{e\in L^6_\sigma}$**.** By Lemma 5.6, it holds that

$$L^6_{\sigma_b}=\{(d^1_{\ell'-1},h^1_{\ell'-1}),(d^0_{\ell'-1},s)\}\cup\{(d^0_i,h^0_i),(d^1_i,s):i\in\{1,\dots,\ell'-2\}\}.$$

Since $L^6_\sigma\subseteq L^6_{\sigma_b}$ can be obtained by a result similar to Lemma B.7 it suffices to show the inequality for all $e\in L^6_{\sigma_b}$.

First, let $e=(d^0_{\ell'-1},s)$. Then, by [3, Table 4],

$$\phi^{\sigma_b}(d^0_{\ell'-1},s)=f(b,(\ell'-1)+1)+j\cdot b_{i+1}=f(b,\ell'-1+1)-0\cdot b_{i+1}=f(b,\ell').$$

Analogously, for $e=(d^1_{\ell'-1},h^1_{\ell'-1})$,

$$\phi^{\sigma_b}(d^1_{\ell'-1},h^1_{\ell'-1})=f(b,(\ell'-1)+1)+(1-j)\cdot b_{i+1}=f(b,\ell')-0\cdot b_{i+1}=f(b,\ell').$$

Therefore, $\phi^{\sigma_b}(e)\geq f(b,\ell')$ holds for $e\in\{(d^0_{\ell'-1},s),(d^1_{\ell'-1},h^1_{\ell'-1})\}$.

Let $e=(d^1_i,s)$ for some $i\in\{1,\dots,\ell'-2\}$. Then, $e$ is an improving switch if and only if the $(i+1)$-th bit switches from 1 to 0. We observe that the first transition in which $(d^1_i,s)$ is an improving switch is therefore the transition from $\sigma_{2^{i+1}-1}$ to $\sigma_{2^{i+1}}$. As the Policy Iteration Algorithm is initialized with the policy representing the number 0, the number $b\in\mathbb{B}_n$ is represented after $b$ many transitions. Therefore, $e$ is an improving switch every $2^{i+1}$-th transition as the $i+1$ least significant bits are all equal to 0 again once the number $b=2^{i+1}$ is reached.

We now interpret $\phi^{\sigma_b}(e)$ as a "counter", which increases during the application of the Policy Iteration Algorithm. By what we just discussed, this counter increases every $2^{i+1}$ transitions and is initialized with zero. In contrast to this, the "counter" $f(b,\ell')$ increases the first time when the number $2^{\ell'-1}$ is reached. But then, after another $2^{\ell'-1}$ transitions the number $2^{\ell'}$ is reached and $\ell(2^{\ell'})=\ell'+1$. Therefore, it takes another $2^{\ell'-1}$ transitions until the counter $f(b,\ell')$ increases another time. In short, the counter $f(b,\ell')$ increases every $2^{\ell'}$ iterations, excluding the first increase which is reached after $2^{\ell'-1}$ iterations. Since $i+1\leq\ell'-1$ follows immediately from $i\leq\ell'-2$, this shows that whenever the counter $f(b, \ell')$ is increased, the counter $\phi^{\sigma_b}(e)$ must have been increased at least once before or in the same iteration. Therefore, $\phi^{\sigma_b}(e) \geq f(b, \ell')$.

The statement follows for $e = (d_i^0, h_i^0)$ by the same arguments in the following way. The switch $(d_i^1, s)$ is applied whenever the $(i+1)$-th bit is no longer equal to 1. The switch $(d_i^0, h_i^0)$ is applied whenever the $(i+1)$-th bit becomes 0. Both of these happen whenever the $(i+1)$-th bit switches from 1 to 0 and thus, the same arguments used before can be applied. ☐

**Lemma 5.8.** *Let $\sigma$ be a phase 3 policy. Then, for all edges $e \in L_\sigma^3$ and $\tilde{e} \in I_\sigma \cap \left(U_\sigma^{3,4} \cup \dots \cup U_\sigma^{3,9}\right)$, it holds that $\phi^\sigma(e) \leq \phi^\sigma(\tilde{e})$.*

*Proof.* Let $\sigma$ be a phase 3 policy and let $e \in L_\sigma^3$. Then, $\phi^\sigma(e) \leq f(b, \ell')$ by Lemma 5.4. It thus suffices to show $\phi^\sigma(\tilde{e}) \geq f(b, \ell')$ for all $\tilde{e} \in I_\sigma \cap \left(U_\sigma^{3,4} \cup \dots \cup U_\sigma^{3,9}\right)$. We distinguish in which of the sets $U_{3,k}$ the switch $\tilde{e}$ is contained.

**Case 1: $\tilde{e} \in U_\sigma^{3,4}$.** Then $\tilde{e} = (h_i^0, k_l)$ for some $l \leq \nu_{i+2}^n(b+1)$, where $\nu_{i+2}^n(b+1)$ again denotes the first bit equal to 1 with an index of at least $i+2$. If there is no such bit, $\nu_{i+2}^n(b+1)$ is equal to $n+1$. By [3, Table 4], we have $\phi^{\sigma_b}(\tilde{e}) = f(b, l)$ and since $\sigma$ is reached after $\sigma_b$, also $\phi^\sigma(\tilde{e}) \geq \phi^{\sigma_b}(\tilde{e}) = f(b, l)$. First, assume that $l \leq \ell'$. Then, by Lemma B.1 (4), it holds that $f(b, l) \geq f(b, \ell')$, implying $\phi^\sigma(\tilde{e}) \geq f(b, \ell')$.

Now assume that $l > \ell'$. We show that this results in a contradiction. To be precise, we show that $(h_i^0, k_l)$ is not an improving switch in this case, i.e., we show $\mathrm{VAL}_\sigma(\sigma(h_i^0)) \geq \mathrm{VAL}_\sigma(k_l)$. To simplify the notation, let $\nu := \nu_{i+2}^n(b+1)$.

First observe that $\sigma(h_i^0) \in \{t, k_{i+2}, \dots, k_n\}$, see Figure 1a. Therefore $\nu \neq n+1$ needs to hold since the edge $(h_i^0, k_{n+1})$ does not exist. In addition, by the definition of $\nu$ and the invariants discussed in Section 2.1, $\sigma(h_i^0) = k_\nu$. We thus need to show $\mathrm{VAL}_\sigma(k_\nu) \geq \mathrm{VAL}_\sigma(k_l)$.

Since $l > \ell'$ by assumption and $\nu \geq l$ by the choice of $\tilde{e}$, also $\nu > \ell'$. Therefore, since $\ell'$ is the least significant set bit of $b+1$, we have $b_j = (b+1)_j$ for all $j \in \{\nu, \dots, n\}$. This implies that during phase 1, no bicycle of one of these levels was opened and the target of none of the vertices $k_\nu, \dots, k_n$ was changed during phase 2. Therefore, using the notation introduced in Lemma B.5, it holds that $\mathrm{VAL}_\sigma(k_\nu) = S_\nu$. By the same lemma, we also get $\mathrm{VAL}_\sigma(k_l) \leq T_l$. Thus, using $b_j = (b+1)_j$ for all $j > \ell'$ and $l > \ell'$, we obtain,

$$\begin{aligned}\mathrm{VAL}_\sigma(k_l) \leq T_l &= \sum_{j \in \{l, \dots, n\} : (b+1)_j = 1} [N^{2j+8} - N^{2j+7} - N^7 + N^6] \\ &= \sum_{j \in \{l, \dots, n\} : b_j = 1} [N^{2j+8} - N^{2j+7} - N^7 + N^6].\end{aligned}$$

By definition, $\nu$ is the smallest index larger than or equal to $i+2$ such that the corresponding bit of $b+1$ is equal to 1. Since $\sigma(h_i^0) \in \{t, k_{i+2}, \dots, k_n\}$, also $l \geq i+2$ needs to hold, see Figure 1a. Therefore, since $l \leq \nu$, this implies that $b_l = b_{l+1} = \dots = b_{\nu-1} = 0$ and using the previous inequality we obtain

$$\begin{aligned}\mathrm{VAL}_\sigma(k_l) &\leq \sum_{j \in \{l, \dots, n\} : b_j = 1} [N^{2j+8} - N^{2j+7} - N^7 + N^6] \\ &= \sum_{j \in \{\nu, \dots, n\} : b_j = 1} [N^{2j+8} - N^{2j+7} - N^7 + N^6] \\ &= S_\nu \\ &= \mathrm{VAL}_\sigma(k_\nu).\end{aligned}$$

This shows that $\tilde{e}$ is not an improving switch. Thus, $l > \ell'$ implies $\tilde{e} \notin I_\sigma$. Therefore, $\phi^\sigma(\tilde{e}) \geq f(b, \ell')$ holds for all $\tilde{e} \in I_\sigma \cap U^{3,4}$.

**Case 2: $\tilde{e} \in U_\sigma^{3,5}$.** Then $\tilde{e} = (s, k_i)$ for some $i < \ell'$ and $\sigma(s) \neq k_i$. Therefore, by [3, Table 4], we have that $\phi^{\sigma_b}(s, k_i) = f(b, i)$. Since $\sigma$ is reached after $\sigma_b$, we also have $\phi^\sigma(s, k_i) \geq \phi^{\sigma_b}(s, k_i)$. Since, by assumption, $i < \ell'$ and by Lemma B.1 (4), this implies

$$\phi^\sigma(s, k_i) \geq \phi^{\sigma_b}(s, k_i) = f(b, i) \geq f(b, \ell').$$

**Case 3:** $\tilde{e} \in U_\sigma^{3,6}$. Then $\tilde{e} = (d_i^j, x)$ for $x \in \{s, h_i^j\}$ where $i \in \{1, \dots, \ell'\}, j \in \{0,1\}$ and $\sigma(d_i^j) \neq x$. First, assume that $x = s$. Then $\sigma(d_i^j) \neq x = s$, implying $\sigma(d_i^j) = h_i^j$. Since $i < \ell'$, it holds that $b_{i+1} = 1$ for $i \neq \ell' - 1$ and $b_{i+1} = 0$ for $i = \ell' - 1$. In addition, the target vertex of $d_i^j$ can only be changed during phase 6 of the current transition and was thus not changed yet. This implies that we either have $d_i^j = d_i^1$ and $\sigma(d_i^1) = h_i^1$ for some $i < \ell', i \neq \ell' - 1$ or $d_i^j = d_{\ell'-1}^0$ and $\sigma(d_i^j) = \sigma(d_{\ell'-1}^0) = h_{\ell'-1}^0$ if $i = \ell' - 1$. For these switches we however already showed in the proof of Lemma 5.7 that $\phi^{\sigma_b}(\tilde{e}) \geq f(b, \ell')$ holds and thus, also $\phi^\sigma(\tilde{e}) \geq f(b, \ell')$ holds.

Now assume that $x = h_i^j$, that is, $\tilde{e} = (d_i^j, h_i^j)$. Analogously to the case $x = s$ it can then be shown that we either have $h_i^j = h_i^0$ and $\sigma(d_i^0) = h_i^0$ for $i < \ell' - 1$ or $h_i^j = h_i^1$ and $\sigma(d_i^1) = h_i^1$ for $i = \ell' - 1$. Again, these edges have already been investigated in the proof of Lemma 5.7 and the inequality $\phi^\sigma(\tilde{e}) \geq f(b, \ell')$ was shown there.

**Case 4:** $\tilde{e} \in U_\sigma^{3,7}$ **or** $\tilde{e} \in U_\sigma^{3,8}$. Since $U_\sigma^{3,7}, U_\sigma^{3,8} \subseteq L_\sigma^6$ and $\phi^\sigma(\tilde{e}) \geq f(b, \ell')$ holds for all edges $\tilde{e} \in L_\sigma^6$ by Lemma 5.7, the statement follows immediately.

**Case 5:** $\tilde{e} \in U_\sigma^{3,9}$. The set $U_\sigma^{3,9}$ contains edges that are improving switches since phase 1. We thus refer to Section 4 and the description of the application of these edges. We need to investigate the occurrence record of switches that we could have applied during phase 1 but did not apply. By the rules I to V and Theorem 4.2, we only switched one instead of two edges within a bicycle $A_i^j$ when $\phi^{\sigma_b}(A_i^j) = b$ held at the beginning of phase 1. Since we always chose to switch the edge with the lower occurrence record in a bicycle and their occurrence records differ at most by one by Equation (4.3), this implies that for any $\tilde{e} = (b_{i,l}^j, A_i^j) \in U_\sigma^{3,9}$ with $\sigma(b_{i,l}^j) \neq A_i^j$ the equality $\phi^{\sigma_b}(b_{i,l}^j, A_i^j) = \lceil \frac{b}{2} \rceil = \lfloor \frac{b+1}{2} \rfloor$ needs to hold. Since $\lfloor \frac{b+1}{2} \rfloor = f(b,1)$ holds by Lemma B.1 (3), $\ell' \geq 1$, and Lemma B.1 (4), we obtain

$$\phi^\sigma(\tilde{e}) \geq \phi^{\sigma_b}(\tilde{e}) = f(b,1) \geq f(b, \ell'). \qquad \square$$

The following statement shows that applying certain improving switches prevents other possible improving switches from being applied.

**Lemma 5.10.** *The following statements hold.*

1. *Let $\sigma$ be the phase 3 policy in which the improving switch $(k_i, k_{\ell'})$ is applied. Let $\sigma'$ be a phase 3 policy of the same transition reached after the policy $\sigma$. Then $I_{\sigma'} \cap S_{i,\sigma'}^{3,1} = \emptyset$.*

2. *Let $\sigma$ be the phase 3 policy in which the improving switch $(b_{i,l}^j, k_{\ell'})$ with $\sigma(b_{i,l}^j) \neq k_{\ell'}$ and $(b+1)_i = 0$ is applied. Let $\sigma'$ be an arbitrary phase 3 policy of the same transition reached after the policy $\sigma$. Then $I_{\sigma'} \cap S_{i,j,l,\sigma'}^{3,2} = \emptyset$.*

3. *Let $\sigma$ be the phase 3 policy in which the improving switch $(b_{i,l}^j, k_{\ell'})$ with $\sigma(b_{i,l}^j) \neq k_{\ell'}$ and $(b+1)_{i+1} \neq j$ is applied. Let $\sigma'$ be an arbitrary phase 3 policy of the same transition reached after the policy $\sigma$. Then $I_{\sigma'} \cap S_{i,j,l,\sigma'}^{3,3} = \emptyset$.*

*Proof.* We show the first statement in detail and only sketch the proof of the other two statements since all of them use the same arguments.

1. Let $\sigma'$ be an arbitrary phase 3 policy reached after $\sigma$. Let $\tilde{e} \in S_{i,\sigma'}^{3,1}$. We show that $\tilde{e}$ is not an improving switch with respect to $\sigma'$.

   We observe that due to the application of $e$ in $\sigma$ and since $\sigma'$ is reached after $\sigma$, we have $\sigma'(k_i) = k_{\ell'}$. Since $\tilde{e} \in S_{i,\sigma'}^{3,1}$, we have $\tilde{e} = (k_i, k_z)$ for some $z \leq \ell'$ such that $\sigma'(k_i) \neq k_z$. It thus suffices to show that $\mathrm{VAL}_{\sigma'}(k_{\ell'}) \geq \mathrm{VAL}_{\sigma'}(k_z)$. Since $\sigma'$ is a phase 3 policy, by Lemma B.5,

$$\mathrm{VAL}_{\sigma'}(k_z) \leq \sum_{j \in \{z, \dots, n\} : (b+1)_j = 1} \left[ (-N)^{2j+8} + (-N)^{2j+7} + (-N)^7 + (-N)^6 \right]. \tag{B.5}$$

   However, since $\sigma[e]$ is also a phase 3 policy, the active bicycle of level $\ell'$ was already closed (phase 1) and the vertex $k_{\ell'}$ points towards the lane containing the active bicycle (phase 2). In addition, since $\ell' = \ell(b+1)$, no bicycle corresponding to a level $j > \ell'$ was opened as $b_j = (b+1)_j$ for these indices. This implies

$$\mathrm{VAL}_{\sigma[e]}(k_{\ell'}) = \sum_{j \in \{\ell', \dots, n\} : (b+1)_j = 1} \left[(-N)^{2j+8} + (-N)^{2j+7} + (-N)^7 + (-N)^6\right]. \tag{B.6}$$

As the values of the vertices are non-decreasing during the application of the Policy Iteration Algorithm, we have $\mathrm{VAL}_{\sigma'}(k_{\ell'}) \geq \mathrm{VAL}_{\sigma[e]}(k_{\ell'})$. Since $(b+1)_j = 0$ for all $j < \ell'$, combining Equations (B.5) and (B.6) yields

$$\begin{aligned}
\mathrm{VAL}_{\sigma'}(k_{\ell'}) &\geq \mathrm{VAL}_{\sigma[e]}(k_{\ell'}) \\
&= \sum_{j \in \{\ell', \dots, n\} : (b+1)_j = 1} \left[(-N)^{2j+8} + (-N)^{2j+7} + (-N)^7 + (-N)^6\right] \\
&= \sum_{j \in \{1, \dots, n\} : (b+1)_j = 1} \underbrace{\left[(-N)^{2j+8} + (-N)^{2j+7} + (-N)^7 + (-N)^6\right]}_{>0 \quad \forall j \geq 1} \\
&\geq \sum_{j \in \{z, \dots, n\} : (b+1)_j = 1} \left[(-N)^{2j+8} + (-N)^{2j+7} + (-N)^7 + (-N)^6\right] \\
&\geq \mathrm{VAL}_{\sigma'}(k_z)
\end{aligned}$$

Thus, $\mathrm{VAL}_{\sigma'}(k_{\ell'}) \geq \mathrm{VAL}_{\sigma'}(k_z)$ and $\tilde{e} = (k_i, k_z)$ is not an improving switch for the policy $\sigma'$.

2. We need to show that for every phase 3 policy $\sigma'$ reached after applying $e = (b_{i,r}^j, k_{\ell'})$ in $\sigma$, no switch contained in $S_{i,j,r,\sigma'}^{3,2}$ is an improving switch. Let $\sigma'$ be an arbitrary phase 3 policy reached after $\sigma$. Then, $\sigma[e](b_{i,r}^j) = k_{\ell'}$ and thus $\mathrm{VAL}_{\sigma[e]}(b_{i,r}^j) = \mathrm{VAL}_{\sigma[e]}(k_{\ell'})$. Since any edge $\tilde{e} \in S_{i,j,r,\sigma'}^{3,2}$ is of the form $\tilde{e} = (b_{i,r}^j, k_z)$ for some $z \leq \ell'$, it therefore suffices to show $\mathrm{VAL}_{\sigma'}(k_{\ell'}) \geq \mathrm{VAL}_{\sigma'}(k_z)$. This however follows by the same estimations used in the first case.

3. This is proven analogously to 2. □

The next lemma shows that it is possible to always find an improving switch during phase 3 that should be applied and that minimizes the occurrence record. This lemma is then used together with some of the previous lemmas to finally prove that there it is possible to apply the improving switches of phase 3 while obeying Zadeh's LEAST-ENTERED pivot rule.

**Lemma 5.11.** *Let $\sigma$ be a phase 3 policy. Then there is an edge $e \in L_\sigma^3 \cap \arg\min_{\tilde{e} \in I_\sigma} \phi^\sigma(\tilde{e})$.*

*Proof.* We first observe that $I_\sigma^3 \neq \emptyset$ for any phase 3 policy $\sigma$ since the set of improving switches is empty if and only if $\sigma$ is an optimal policy. Let $e \in \arg\min_{\tilde{e} \in I_\sigma} \phi^\sigma(e)$.

Since $L_\sigma^3 \subseteq I_\sigma \subseteq U_\sigma^3$ by [3, Lemma 4], either $e \in L_\sigma^3$ or $e \in U_\sigma^3 \setminus L_\sigma^3$. Assume that the second case holds, since the statement follows directly otherwise. We observe that since $U_\sigma^{3,1}, \dots, U_\sigma^{3,9}$ form a partition of $U_\sigma^3$, there is exactly one $k \in \{1, \dots, 9\}$ with $e \in U_\sigma^{3,k}$.

Assume that $k \in \{4, \dots, 9\}$. Then, by Lemma 5.8, $\phi^\sigma(e) \geq \phi^\sigma(\tilde{e})$ for all $\tilde{e} \in L_\sigma^3$ since $e \in I_\sigma$. Since $e$ minimizes the occurrence record, this implies $\phi^\sigma(e) = \phi^\sigma(\tilde{e})$ for all $\tilde{e} \in L_\sigma^3$. Thus there is an edge $\tilde{e} \in L_\sigma^3$ minimizing the occurrence record, so $\tilde{e} \in \arg\min_{\tilde{e} \in I_\sigma} \phi^\sigma(\tilde{e}) \cap L_\sigma^3$.

Now assume that $k \in \{1, 2, 3\}$. We analyze these cases one after another.

<u>**Case 1:**</u> $\boldsymbol{e \in U_\sigma^{3,1}}$. In this case $e = (k_i, k_z)$ for some indices $i \in \{1, \dots, n\}$ and $z \in \{1, \dots, \ell'\}$ such that $\sigma(k_i) \notin \{k_z, k_{\ell'}\}$ and $(b+1)_i = 0$. Thus $e \in S_{i,\sigma}^{3,1}$. First assume that the switch $(k_i, k_{\ell'})$ was not applied yet. Then, $\phi^\sigma(k_i, k_{\ell'}) = \phi^{\sigma_b}(k_i, k_{\ell'})$ and, by [3, Table 4], also $\phi^{\sigma_b}(k_i, k_{\ell'}) = f(b, \ell', \{(i, 0)\})$. Together with $z \leq \ell'$ and Lemma B.1 (4), this implies

$$\phi^\sigma(k_i, k_{\ell'}) = \phi^{\sigma_b}(k_i, k_{\ell'}) = f(b, \ell', \{(i, 0)\}) \leq f(b, z, \{(i, 0)\}) = \phi^{\sigma_b}(e) \leq \phi^\sigma(e).$$

Since $e$ is chosen such that it minimizes the occurrence records among all improving switches, it holds that $\phi^\sigma(k_i, k_{\ell'}) = \phi^\sigma(e)$. This however implies $(k_i, k_{\ell'}) \in \arg\min_{\tilde{e} \in I_\sigma} \phi^\sigma(\tilde{e})$. Therefore, the statement follows from $(k_i, k_{\ell'}) \in L_\sigma^3$.

It remains to show that $(k_i, k_{\ell'})$ was not applied yet. Towards a contradiction, assume that it was applied before in this transition. Then there was another phase 3 policy $\sigma'$ reached before $\sigma$ such that $(k_i, k_{\ell'})$ was applied in $\sigma'$. But then, by Lemma 5.10, we have $I_\sigma \cap S^{3,1}_{i,\sigma} = \emptyset$ since the policy $\sigma$ is reached after $\sigma'$. This is a contradiction since $e \in I_\sigma$ and $e \in S^{3,1}_{i,\sigma}$.

**Case 2:** $\boldsymbol{e \in U^{3,2}_\sigma}$**.** In this case $e = (b^j_{i,r}, k_z)$ for some $i \in \{1, \dots, n\}$ and some $z \in \{1, \dots, \ell'\}$ such that $\sigma(b^j_{i,l}) \notin \{k_z, k_{\ell'}\}$ and $(b+1)_i = 0$. Hence, $e \in S^{3,2}_{i,j,r,\sigma}$. First assume that the improving switch $(b^j_{i,r}, k_{\ell'})$ was not applied yet. Then, since $z \leq \ell'$, by [3, Table 4] and by Lemma B.1 (4),

$$\begin{aligned}
\phi^\sigma(b^j_{i,r}, k_{\ell'}) &= \phi^{\sigma_b}(b^j_{i,l}, k_{\ell'}) \\
&= f(b, \ell', \{(i,0)\}) + f(b, \ell', \{(i,1), (i+1, 1-j)\}) \\
&\leq f(b, z, \{(i,0)\}) + f(b, z, \{(i,1), (i+1, 1-j)\}) \\
&= \phi^{\sigma_b}(e) \\
&\leq \phi^\sigma(e).
\end{aligned}$$

Since $e$ is chosen such that it minimizes the occurrence records among all improving switches, we have $\phi^\sigma(b^j_{i,r}, k_{\ell'}) = \phi^\sigma(e)$. This implies that $(b^j_{i,r}, k_{\ell'}) \in \arg\min_{\tilde{e} \in I_\sigma} \phi^\sigma(\tilde{e})$. Therefore, the statement follows from $(b^j_{i,r}, k_{\ell'}) \in L^3_\sigma$.

It remains to show that $(b^j_{i,r}, k_{\ell'})$ was not applied yet. However, assuming that this switch was applied before results in the same contradiction as in the last case when applying Lemma 5.10.

**Case 3:** $\boldsymbol{e \in U^{3,3}_\sigma}$**.** This follows analogously to the previous case. □

**Theorem 5.12.** *There is an ordering of the improving switches and an associated tie-breaking rule compatible with the* LEAST-ENTERED *pivot rule such that all improving switches contained in* $L^3_{\sigma_b}$ *are applied and the* LEAST-ENTERED *pivot rule is obeyed during phase 3.*

*Proof.* Let $\sigma$ denote the first phase 3 policy of the transition from $\sigma_b$ to $\sigma_{b+1}$. Then, $L^3_\sigma = L^3_{\sigma_b}$ by Lemma B.6. By Lemma 5.11, there is an edge $e_1 \in L^3_\sigma$ minimizing the occurrence record $I_\sigma$. By Lemma B.7, applying this switch results in a new phase 3 policy $\sigma[e_1]$ such that $L^3_{\sigma[e_1]} = L^3_\sigma \setminus \{e_1\}$. Now, again by Lemma 5.11, there is an edge $e_2 \in L^3_{\sigma[e_1]}$ minimizing the occurrence record $I_{\sigma[e_1]}$.

We can now apply the same argument iteratively until we reach a phase 3 policy $\hat{\sigma}$ such that $\left|L^3_{\hat{\sigma}}\right| = 1$ while only applying switches contained in $L^3_{\sigma_b}$. Then, by construction and by Lemma 5.11, $(e_1, e_2, \dots)$ defines an ordering of the edges of $L^3_{\sigma_b}$ and an associated tie-breaking rule that always follow the LEAST-ENTERED rule. When the policy $\hat{\sigma}$ with $\left|L^3_{\hat{\sigma}}\right| = 1$ is reached, applying the remaining improving switch results in a phase 4 policy. Then, all improving switches contained in $L^3_{\sigma_b}$ were applied and the LEAST-ENTERED pivot rule was obeyed. □

Note that the ordering used in the proof of Theorem 5.12 avoids Issue 5.3: As we proved in Issue 5.3, it is not possible to apply the improving switches in level $\ell(b+1)$ and level $n-1$ consistently such that all switches of level $\ell(b+1)$ are applied before any switch of level $n+1$ is applied and vice versa. By further analyzing the proof of Issue 5.3, it can be shown that the same holds for the improving switches of other levels. Our ordering always chooses an improving switch that minimizes the occurrence record regardless of the level, and in particular does not apply improving switches level by level in an order that only depends on the least significant set bit.

Using Theorem 5.12 we can now show the following final theorem.

**Theorem 5.13.** *Fix the transition from* $\sigma_b$ *to* $\sigma_{b+1}$ *for some* $\sigma \in \mathbb{B}_n$*. There is an order in which to apply improving switches during this transition such that the* LEAST-ENTERED *rule is obeyed, and the switches of phase* $p$ *are applied before any switches of phase* $p+1$*, for every* $p \in \{1, \dots, 5\}$.

However, proving this theorem requires analyzing the phases 1,2,4,5 and 6. Note that we do not analyze this phase as detailed as phase 3 and use a lot of the results and descriptions given in [3].

For the remainder of this appendix, fix some $b \in \mathbb{B}_n$ and consider the transition from $\sigma_b$ to $\sigma_{b+1}$. As always, we define $\ell := \ell(b)$ and $\ell' := \ell(b+1)$.

**Lemma B.11.** *Let $\sigma$ be a phase 1 policy for $G_n$. Then there is an $e \in L^1_\sigma$ such that $\phi^\sigma(e) \le \min_{\tilde{e} \in U^1_\sigma \cap I_\sigma} \phi^\sigma(\tilde{e})$.*

*Proof.* By [3, Table 4], we have $L_\sigma = I_\sigma = U_\sigma$ for any phase 1 policy $\sigma$, immediately implying the statement. □

**Lemma B.12.** *Let $\sigma$ be a phase 2 policy for $G_n$. Then there is an $e \in L^2_\sigma$ such that $\phi^\sigma(e) \le \min_{\tilde{e} \in U^2_\sigma \cap I_\sigma} \phi^\sigma(\tilde{e})$.*

*Proof.* Let $\sigma$ be a phase 2 policy and $e \in L^2_\sigma$. Then, by [3, Table 3], $L^2_\sigma$ only contains one edge. We thus have $e = (k_{\ell'}, c^j_{\ell'})$ where $j = (b+1)_{\ell'+1}$. Since $U^2_\sigma = L^1_\sigma \cup L^2_\sigma$ and $I_\sigma \subset U^2_\sigma$, it therefore suffices to show that $\phi^\sigma(e) \le \min_{\tilde{e} \in L^1_\sigma} \phi^\sigma(\tilde{e})$.

So let $\tilde{e} \in L^1_\sigma$. Then, $\tilde{e} = (b^j_{i,r}, A^j_i)$ for some $i \in \{1, \dots, n\}$ and $j \in \{0, 1\}$ such that $\sigma(b^j_{i,r}) \neq A^j_i$. By this condition, we have that the improving switch $\tilde{e}$ was not applied during phase 1, since this would imply $\sigma(b^j_{i,r}) = A^j_i$. As we have already discussed in the proof of Lemma 5.8, we therefore have $\phi^\sigma(\tilde{e}) = f(b, 1)$. But then, since $\phi^\sigma(e) = f(b, \ell', \{(\ell'+1, j)\})$ by [3, Table 4] and by using Lemma B.1 (2,4), we get

$$\phi^\sigma(e) = f(b, \ell', \{(\ell'+1, j)\}) \le f(b, \ell') \le f(b, 1) = \phi^\sigma(\tilde{e}).$$ □

**Lemma B.13.** *Let $\sigma$ be a phase 4 policy for $G_n$. Then there is an $e \in L^4_\sigma$ with $\phi^\sigma(e) \le \min_{\tilde{e} \in U^4_\sigma \cap I_\sigma} \phi^\sigma(\tilde{e})$.*

*Proof.* Let $e \in L_{\sigma^4}$. As proven in Lemma 5.7, Case 1, we then have $\phi^\sigma(e) = f(b, \ell')$. It therefore suffices to show that $\phi^\sigma(e) \ge f(b, \ell')$ for all $e \in U^4_\sigma \cap I_\sigma$.

We observe that $U^4_\sigma = U^{3,4}_\sigma \cup \cdots \cup U^{3,9}_\sigma$. We also observe that we already proved $\phi^{\sigma'}(\tilde{e}) \ge f(b, \ell')$ for all $\tilde{e} \in I_{\sigma'} \cap U^4_{\sigma'}$ in the proof of Lemma 5.8 when $\sigma'$ is a phase 3 policy. The statement thus follows for phase 4 policies by applying the same arguments. □

**Lemma B.14.** *Let $\sigma$ be a phase 5 policy for $G_n$. Then there is an $e \in L^5_\sigma$ such that $\phi^\sigma(e) \le \min_{\tilde{e} \in U^5_\sigma \cap I_\sigma} \phi^\sigma(\tilde{e})$.*

*Proof.* Let $e \in L^5_\sigma$. Then $e = (s, k_{\ell'})$ since the set $L^5_\sigma$ only contains one edge,. Thus, by [3, Table 4], we have $\phi^\sigma(e) = f(b, \ell')$ . It thus suffices to show that $\phi^\sigma(\tilde{e}) \ge f(b, \ell')$ holds for all $\tilde{e} \in U^5_\sigma$. This however can again be shown by the same arguments used in the proof of Lemma 5.8 since $U^5_\sigma = U^{3,5}_\sigma \cup \cdots \cup U^{3,9}_\sigma$. □

**Lemma B.15.** *Let $\sigma$ be a phase 6 policy for $G_n$. Then there is an $e \in L^6_\sigma$ such that $\phi^\sigma(e) \le \min_{\tilde{e} \in U^6_\sigma \cap I_\sigma} \phi^\sigma(\tilde{e})$.*

*Proof.* Let $e \in L^6_\sigma$. Since $U^6_\sigma = L^1_\sigma \cup L^6_\sigma$ it suffices to show that $\phi^\sigma(e) \le \phi^\sigma(\tilde{e})$ for all $\tilde{e} \in L^1_\sigma$, since this implies that there is always a switch contained in $L^6_\sigma$ minimizing the occurrence record among all switches in $U^6_\sigma = L^1_\sigma \cup L^6_\sigma$.

Therefore let $\tilde{e} \in L_{\sigma^1}$. As shown in the proof of Lemma 5.8, we have that $\phi^\sigma(\tilde{e}) = f(b, 1)$ since $\tilde{e} \in L^6_\sigma$ and $\tilde{e}$ was not applied during phase 1. Since $e \in L^6_\sigma$, we either have $e = (d^j_i, s)$ for some $i \in \{1, \dots, n\}$ and $j \in \{0, 1\}$, implying that $\phi^\sigma(e) = f(b, i+1) - j \cdot b_{i+1}$ or $e = (d^j_i, h^j_i)$ for some $i \in \{1, \dots, n\}, j \in \{0, 1\}$, implying $\phi^\sigma(e) = f(b, i+1) - (1-j) \cdot b_{i+1}$. However, using Lemma B.1 (4) and $i \ge 1$ we obtain

$$\phi^\sigma(e) \le f(b, i+1) \le f(b, 2) \le f(b, 1) = \phi^\sigma(\tilde{e}).$$ □

*Proof of Theorem 5.13.* Consider the initial phase 1 policy $\sigma_b$ representing $b$. By Lemma B.11 there is a switch contained in $L^1_{\sigma_b}$ minimizing the occurrence record among all improving edges. Thus, this switch can be applied without violating Zadeh's LEAST-ENTERED pivot rule. By [3, Lemma 5], the resulting policy is either a phase 2 policy or another phase 1 policy. In the latter case we can apply the same argument again. After applying a finite number of improving switches we thus obtain a phase 2 policy $\sigma^2$. By Lemma B.12 we can apply the single improving switch contained in $L^2_{\sigma^2}$ without violating the LEAST-ENTERED pivot rule. By [3, Lemma 5], the resulting policy is then a phase 3 policy. As proven in Theorem 5.12 we can now apply all improving switches that should be applied during phase 3 and obtain a phase 4 policy $\sigma^4$.

By Lemma B.13 there is a switch contained in $L^4_{\sigma^4}$ minimizing the occurrence record among all improving edges since $I_{\sigma^4} \subseteq U^4_{\sigma^4}$ by [3, Lemma 4]. The resulting policy is then either another phase 4 policy or a phase 5 policy. In the first case we can apply the same argument again. We therefore obtain a phase 5 policy after applying a finite number of improving switches.

By Lemmas B.14 and B.15, the same arguments used for phase 4 can now be used for phase 5 and 6, concluding our proof. □