



An exponential lower bound for Zadeh's pivot rule

Yann Disser¹ · Oliver Friedmann² · Alexander V. Hopp³

Received: 16 April 2021 / Accepted: 10 June 2022 / Published online: 19 July 2022

© The Author(s) 2022

Abstract

The question whether the Simplex Algorithm admits an efficient pivot rule remains one of the most important open questions in discrete optimization. While many natural, deterministic pivot rules are known to yield exponential running times, the random-facet rule was shown to have a subexponential running time. For a long time, Zadeh's rule remained the most prominent candidate for the first deterministic pivot rule with subexponential running time. We present a lower bound construction that shows that Zadeh's rule is in fact exponential in the worst case. Our construction is based on a close relation to the Strategy Improvement Algorithm for Parity Games and the Policy Iteration Algorithm for Markov Decision Processes, and we also obtain exponential lower bounds for Zadeh's rule in these contexts.

Keywords Simplex method · Zadeh's rule · Lower bound · Parity games · Markov decision processes · Strategy improvement · Policy iteration

Mathematics Subject Classification 68Q25 · 90C05 · 90C40

The work of Yann Disser and Alexander V. Hopp was supported by the 'Excellence Initiative' of the German Federal and State Governments and the Graduate School CE at TU Darmstadt. Alexander V. Hopp's work was mainly contributed during his time at the Department of Mathematics at the TU Darmstadt.

✉ Yann Disser
disser@mathematik.tu-darmstadt.de

Oliver Friedmann
oliver.friedmann@gmail.com

Alexander V. Hopp
alexander.hopp@merckgroup.com

¹ Department of Mathematics, TU Darmstadt, Darmstadt, Germany

² Department of Informatics, LMU Munich, München, Germany

³ Merck KGaA, Darmstadt, Germany

1 Introduction

The quest for discovering the best pivot rule for the Simplex Algorithm [5] remains one of the most important challenges in discrete optimization. In particular, while several other *weakly* polynomial algorithms for solving Linear Programs have been proposed in the past [3, 9, 27–29], no fully “combinatorial” algorithm with strongly polynomial running time is known to date – in fact, the question whether such an algorithm exists is contained in Smale’s list of 18 mathematical problems for the century, among other famous unsolved problems like the Riemann hypothesis and the P versus NP problem [36]. The Simplex Algorithm is inherently combinatorial and may yield a strongly polynomial algorithm if a suitable pivot rule exists. The question what theoretical worst-case running time can be achieved with a pivot rule for the Simplex Algorithm is closely related to the question what the largest possible (combinatorial) diameter of a polytope is, and, in particular, to the weak Hirsch conjecture which states that the diameter is polynomially bounded [5, 35, 41].

For a variety of natural pivot rules, exponential worst-case examples were found soon after the Simplex Algorithm was proposed [1, 20, 30]. These examples are highly imbalanced in the sense that they cause some improving directions to be selected by the pivot rule only rarely, while others are selected often. Randomized pivot rules were proposed as a way to average out the behavior of the Simplex Algorithm and to thus avoid imbalanced behavior. The hope that this may lead to a better worst-case performance was met when subexponential upper bounds were eventually established for the random-facet pivot rule [22, 25, 32]. Other promising candidates for efficient pivot rules were deterministic “pseudo-random” rules that balance the behavior of the algorithm explicitly by considering all past decisions in each step, instead of obliviously deciding for improvements independently. The two most prominent examples of such pivot rules are Cunningham’s rule [4] which fixes an order of all possible improvement directions at the start and, in each step, picks the next improving direction in this order in round robin fashion, and Zadeh’s rule [43] which picks an improving direction chosen least often so far in each step. By design, bad examples are much more difficult to construct for these more balanced pivoting rules, and it took more than 30 years until the first lower bounds were established. Eventually, a subexponential lower bound was shown for the random-facet rule [17, 18, 21] and the random-edge rule [17, 33]. Most recently, a subexponential lower bound was shown for Zadeh’s rule [7, 15], and an exponential lower bound for Cunningham’s rule [2]. An exponential lower bound for Zadeh’s rule is known on Acyclic Unique Sink Orientations [39], but it is unclear whether the corresponding construction can be realized as a Linear Program. This means that Zadeh’s rule remained the only promising candidate for a deterministic pivot rule to match the subexponential running time of the random-facet rule.

Local search algorithms similar to the Simplex Algorithm are important in other domains, like Vöge and Jurdziński’s Strategy Improvement Algorithm for Parity Games [42] and Howard’s Policy Iteration Algorithm for Markov Decision Processes [23]. Much like the Simplex Algorithm, these algorithms rely on a pivot rule that determines which local improvement to perform in each step. And much like for the Simplex Algorithm, many natural deterministic pivot rules for these algorithms have been shown to be exponential [2, 10, 12, 13], while a subexponential upper bound

has been shown for the random-facet rule [16, 18, 25, 26, 32]. Again, Zadeh's rule remained as a promising candidate for a deterministic subexponential pivot rule.

Our results and techniques. In this paper, we give the first exponential lower bound for Zadeh's pivot rule for the Strategy Improvement Algorithm for Parity Games, for the Policy Iteration Algorithm for Markov Decision Processes, and for the Simplex Algorithm. This closes a long-standing open problem by eliminating Zadeh's pivot rule as a candidate for a deterministic, subexponential pivot rule in each of these three areas (up to tie-breaking). We note that while the lower bound for the Simplex Algorithm is arguably our most important result, the lower bounds for Parity Games and Markov Decision Processes are important in their own right and complement previous results in these areas [10, 12, 13, 16].

Our lower bound construction is based on the technique used in [2, 15] (among others). In particular, we construct a Parity Game that forces the Strategy Improvement Algorithm to emulate a binary counter by enumerating strategies corresponding to the natural numbers 0 to $2^n - 1$. The construction is then converted into a Markov Decision Process that behaves similarly (but not identically) regarding the Policy Iteration Algorithm. Finally, using a well-known transformation, the Markov Decision Process can be turned into a Linear Program for which the Simplex Algorithm mimics the behavior of the Policy Iteration Algorithm. We remark that we use an artificial, but systematic and polynomial time computable, tie-breaking rule for the pivot step whenever Zadeh's rule does not yield a unique improvement direction. Importantly, while the tie-breaking rule is carefully crafted to simplify the analysis, conceptually, our construction is not based on exploiting the tie-breaking rule. Note that it cannot be avoided to fix a tie-breaking rule when analyzing Zadeh's pivot rule, in the sense that, for every Markov Decision Process of size n , a tie-breaking rule tailored to this instance exists, such that the Policy Iteration Algorithm takes at most n steps [14, Cor. 4.79].

Roughly speaking, much like the subexponential construction in [15], our construction consists of multiple levels, one for each bit of the counter. The subexponential construction of [15] requires each level to connect to the level of the least significant set bit of the currently represented number, which yields a quadratic number m of edges in the construction, which in turn leads to a lower bound of $2^{\Omega(n)} = 2^{\Omega(\sqrt{m})}$, i.e., a subexponential bound in the size $\Theta(m)$ of the construction. In contrast, our construction only needs each level to connect to one of the first two levels, depending on whether the currently represented number is even or odd. Very roughly, this is the key idea of our result, since it allows us to reduce the size of the construction to $\Theta(n)$, which leads to an exponential lower bound. However, to make this change possible, many other technical details have to be addressed, and, in particular, we are no longer able to carry the construction for Parity Games over as-is to Markov Decision Processes.

A challenge when constructing a lower bound for Zadeh's rule is to keep track not only of the exact sets of improving directions in each step, but also of the exact number of times every improving direction was selected in the past. In contrast, the exponential lower bound construction for Cunningham's rule [2] "only" needs to keep track of the next improving direction in the fixed cyclic order. As a consequence, the full proof of our result is very technical, because it requires us to consider *all*

possible improvements in every step, and there are many transitional steps between configurations representing natural numbers. In this paper, we give an exact description of our construction and an outline of our proof. A complete and detailed proof can be found in the full version [6]. Importantly, our construction has been implemented and tested empirically for consistency with our formal treatment, see Appendix A.¹ The resulting animations of the execution for $n = 3$ resp. $n = 4$, which take 160 resp. 466 steps, are available online [24].

2 Parity games and strategy improvement

A *Parity Game (PG)* is a two player game that is played on a directed graph where every vertex has at least one outgoing edge. Formally, it is defined as a tuple $G = (V_0, V_1, E, \Omega)$, where $V_0 \cap V_1 = \emptyset$, $(V_0 \cup V_1, E)$ is a directed graph and $\Omega : V_0 \cup V_1 \rightarrow \mathbb{N}$ is the *priority function*. The set V_p is the set of vertices of player $p \in \{0, 1\}$ and the set $E_p := \{(v, w) \in E : v \in V_p\}$ is the set of edges of player $p \in \{0, 1\}$. We let $V := V_0 \cup V_1$. A *play* in G is an infinite walk in the graph. The *winner* of a play is determined by the highest priority that occurs infinitely often along the walk. If this priority is even, player 0 wins, otherwise, player 1 wins.

Formally, a play in G can be described by a pair of *strategies*. A strategy for player p is a function that chooses one outgoing edge for each vertex of player p . To be precise, a (*deterministic positional*) strategy for player p is a function $\sigma : V_p \rightarrow V$ that selects for each vertex $v \in V_p$ a target vertex $\sigma(v)$ such that $(v, \sigma(v)) \in E_p$ for all $v \in V_p$. Throughout this paper we only consider deterministic positional strategies and henceforth simply refer to them as *strategies*. Two strategies σ, τ for players 0,1 and a starting vertex v then define a unique play starting at v with the corresponding walk being determined by the strategies of the players. A play can thus fully be described by a tuple (σ, τ, v) and is denoted by $\pi_{\sigma, \tau, v}$. A player 0 strategy σ is *winning* for player 0 at vertex v , if player 0 is the winner of every game $\pi_{\sigma, \tau, v}$, regardless of τ . Winning strategies for player 1 are defined analogously. A fundamental result in the theory of Parity Games is that, for every starting vertex, there always is a winning strategy for exactly one of the two players. The computational problem of *solving* a parity game consists in finding the corresponding partition of V .

Theorem 1 (e.g. [11, 31]) *In every Parity Game, V can be partitioned into winning sets (W_0, W_1) , where player p has a positional winning strategy for each starting vertex $v \in W_p$.*

2.1 Vertex valuations, the strategy improvement algorithm, and sink games

We now discuss the Strategy Improvement Algorithm of Vöge and Jurdziński [42] and its theoretical background. We discuss the concept of *vertex valuations* and define a

¹ We used two separate implementations to double-check our construction. Firstly, we used the PGSOLVER library [19] as a black box with our construction as input. Secondly, a new policy iteration tool for sink games was implemented from scratch [37], again using our concrete construction only as an input. Both implementations confirm the intended behavior, including quantitative details, for instances up to size $n = 5$.

special class of games that our construction belongs to, called *sink games*, and define vertex valuations for this class of games. We refer to [2, 14] for a more in-depth and general discussion of these topics.

Fix a pair σ, τ of strategies for players 0, 1, respectively. The idea of vertex valuations is to assign a valuation to every $v \in V$ that encodes how “profitable” vertex v is for player 0. By defining a suitable pre-order on these valuations, this enables us to compare the valuations of vertices and “improve” the strategy σ by changing the target $\sigma(v)$ of a vertex v to a more “profitable” vertex $w \neq \sigma(v)$ with $(v, w) \in E$. Since there are only finitely many strategies and vertices, the strategy of player 0 can only be improved a finite number of times, eventually resulting in a so-called *optimal strategy* for player 0. It is known (e.g., [14, 42]) that an optimal strategy can then be used to determine the winning sets W_0, W_1 of the Parity Game and thus to solve the game.

Formally, vertex valuations are given as a totally ordered set (U, \leq) . For every pair of strategies σ, τ , we are given a function $\mathcal{E}_{\sigma, \tau} : V \rightarrow U$ assigning vertex valuations to vertices. Since U is totally ordered, this induces a preorder of the vertices for fixed strategies σ, τ . To eliminate the dependency on the player 1 strategy, we define the vertex valuation of v with respect to σ by $\mathcal{E}_\sigma(v) := \min_{\prec} \mathcal{E}_{\sigma, \tau}(v)$ where the minimum is taken over all player 1 strategies τ . Formally, if $\mathcal{E}_{\sigma, \tau}(\tau(v)) \leq \mathcal{E}_\sigma(v)$ for all $(v, u) \in E_1$, then the player 1 strategy τ is called *counterstrategy* for σ . It is well-known that counterstrategies exist and can be computed efficiently [42]. For a strategy σ , an arbitrary but fixed counterstrategy is denoted by τ^σ .

We can extend this ordering to a partial ordering of strategies by defining $\sigma \leq \sigma'$ if and only if $\mathcal{E}_\sigma(v) \leq \mathcal{E}_{\sigma'}(v)$ for all $v \in V$. We write $\sigma \triangleleft \sigma'$ if $\sigma \leq \sigma'$ and $\sigma \neq \sigma'$. Given a strategy σ , a strategy σ' with $\sigma \triangleleft \sigma'$ can be obtained by applying an *improving switch*. Intuitively, an improving switch is an edge such that including e in σ improves the strategy with respect to \leq . Formally, let $e = (v, u) \in E_0$ and $\sigma(v) \neq u$. We define the strategy σe via $\sigma e(v') := \sigma(v)$ if $v' \neq v$ and $\sigma e(v) := u$. The edge e is *improving* for σ if $\sigma \triangleleft \sigma e$ and we denote the set of improving switches for σ by I_σ .

The Strategy Improvement Algorithm now operates as follows. Given an initial strategy ι , apply improving switches until a strategy σ^* with $I_{\sigma^*} = \emptyset$ is reached. Such a strategy is called *optimal* and a strategy is optimal if and only if $\sigma \triangleleft \sigma^*$ for all player 0 strategies [42]. The running time of this algorithm highly depends on the order in which improving switches are applied – a point that we discuss in more detail later.

This terminology allows us to introduce a special class of Parity Games, called *sink games*. This class allows for an easy definition of the vertex valuations as discussed after the definition.

Definition 1 A Parity Game $G = (V_0, V_1, E, \Omega)$ together with an initial player 0 strategy ι is a *Sink Game* if the following two statements hold.

1. There is a vertex $t \in V_1$ with $(t, t) \in E$ and $\Omega(t) = 1$ reachable from all vertices. In addition, $\Omega(v) > \Omega(t)$ for all $v \in V \setminus \{t\}$. This unique vertex t is called the *sink* of the sink game.
2. For each player 0 strategy σ with $\iota \leq \sigma$ and each vertex v , every play $\pi_{\sigma, \tau^\sigma, v}$ ends in t .

Let $G = (V_0, V_1, E, \Omega)$ and ι define a Sink Game. To simplify the presentation, assume that Ω is injective. Since G is a Sink Game, every play $\pi_{\sigma, \tau^\sigma, v}$ in G can be

represented as the walk $\pi_{\sigma, \tau^\sigma, v} = v, v_2, \dots, v_k, (t)^\infty$. In particular, a play can be identified with its *path component* v, v_2, \dots, v_k . Now, defining $\Xi_\sigma(v)$ as the path component of $\pi_{\sigma, \tau^\sigma, v}$ is a well-studied choice of vertex valuations. To give a total ordering of the vertex valuations, it thus suffices to give an ordering of all subsets of V .

Let $M, N \subseteq V, M \neq N$. Intuitively, M is better than N for player 0 if it contains a vertex with large even priority not contained in N , or if it there is a vertex with large odd priority contained in N but not in M . Formally, $v \in M \Delta N$ is called *most significant difference of M and N* if $\Omega(v) > \Omega(w)$ for all $w \in M \Delta N, w \neq v$. The most significant difference of M and N is denoted by $\Delta(M, N)$ and allows us to define an ordering $<$ on the subsets of V . For $M, N \subseteq V, M \neq N$ we define

$$N < M \iff [\Delta(M, N) \in M \wedge \Delta(M, N) \text{ is even}] \vee [\Delta(M, N) \in N \wedge \Delta(M, N) \text{ is odd}].$$

Note that $<$ is a total ordering as we assume Ω to be injective. We mention here that injectivity is not necessary – the most significant difference of any two vertex valuations being unique suffices.

The following theorem summarizes the most important aspects related to Parity Games, vertex valuations and improving switches. Note that the construction of vertex valuations given here is a simplified version of the general concept of vertex valuations used for Parity Games. It is, however, in accordance with the general construction and we refer to [14] for a more detailed discussion.

Theorem 2 [42] *Let $G = (V_0, V_1, E, \Omega)$ be a Sink Game and σ be a player 0 strategy.*

1. *The vertex valuations of a player 0 strategy are polynomial-time computable.*
2. *There is an optimal player 0 strategy σ^* with respect to the ordering \triangleleft .*
3. *If $I_\sigma = \emptyset$, then σ is optimal.*
4. *We have $I_\sigma = \{(v, w) \in E_0 : \Xi_\sigma(\sigma(v)) \triangleleft \Xi_\sigma(w)\}$ and $\sigma \triangleleft \sigma e$ for all $e \in I_\sigma$.*
5. *Given an optimal player 0 strategy, the winning sets W_0 and W_1 of player 0 and player 1 can be computed in polynomial time.*

3 Lower bound construction

In this section, we describe a PG $S_n = (V_0, V_1, E, \Omega)$ such that the Strategy Improvement Algorithm performs at least 2^n iterations when using Zadeh’s pivot rule and a specific tie-breaking rule. Before giving a formal definition, we give a high-level intuition of the main idea of the construction. A simplified visualization of the construction is given in Fig. 1.

The key idea is that S_n simulates an n -digit binary counter. We thus introduce notation related to binary counting. It will be convenient to consider counter configurations with more than n bits, where unused bits are zero. In particular, we always interpret bit $n + 1$ as 0. Formally, we denote the set of n -bit configurations by $\mathcal{B}_n := \{b \in \{0, 1\}^\infty : b_i = 0 \ \forall i > n\}$. We start with index one, hence a counter configuration $b \in \mathcal{B}_n$ is a tuple (b_n, \dots, b_1) . Here, b_1 is the least and b_n is the most significant bit. The *integer value* of $b \in \mathcal{B}_n$ is $\sum_{i=1}^n b_i 2^{i-1}$. We identify the integer value of b with its counter configuration and use the natural ordering of \mathbb{N} to order counter

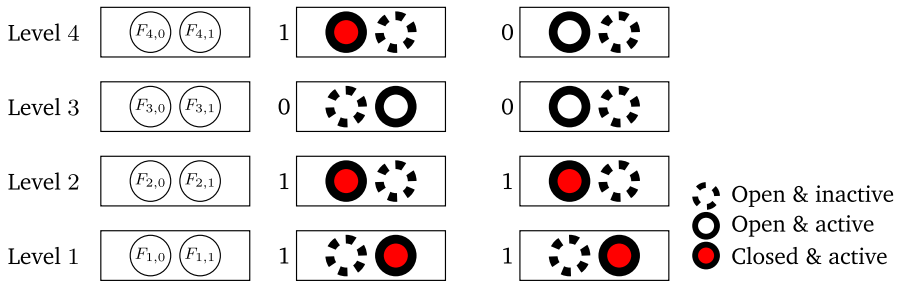


Fig. 1 Visualization of the general structure of the binary counter for $n = 4$. Level 1 encodes the least significant bit and level 4 encodes the most significant bit. The left picture shows the cycle centers and their distribution in the levels. The two figures on the right give examples for settings of the cycles representing the numbers 11 and 3, respectively

configurations. For $\mathbf{b} \in \mathcal{B}_n$, $\mathbf{b} \neq 0$, we define $\nu(\mathbf{b}) := \min\{i \in \{1, \dots, n\} : b_i = 1\}$ to be the *least significant set bit* of \mathbf{b} .

The PG S_n consists of n (nearly) identical levels and each level encodes one bit of the counter. Certain strategies and corresponding counterstrategies in S_n are then interpreted as binary numbers. If the Strategy Improvement Algorithm enumerates at least one player 0 strategy per $\mathbf{b} \in \mathcal{B}_n$ before finding the optimal strategy, it enumerates at least 2^n strategies. Since the game has size linear in n , this then establishes the exponential lower bound.

The main challenge is to obey Zadeh’s pivot rule as it forces the algorithm to only use improving switches used least often during the execution. Intuitively, a counter obeying this rule needs to switch bits in a “balanced” way. However, counting from 0 to $2^n - 1$ in binary does not switch individual bits equally often. For example, the least significant bit is switched every time and the most significant bit is switched only once. The key idea to overcome this obstacle is to have a substructure in each level that contains two gadgets. These gadgets are called *cycle centers*. In every iteration of the algorithm, only one of the cycle centers is interpreted as encoding the bit of the current level. This enables us to perform operations within the other cycle center without losing the interpretation of the bit being equal to 0 or 1. This is achieved by an alternating encoding of the bit by the two cycle centers.

We now provide more details. Consider some level i , some $\mathbf{b} \in \mathcal{B}_n$ and denote the cycle centers of level i by $F_{i,0}$ and $F_{i,1}$. One of them now encodes b_i . Which of them represents b_i depends on b_{i+1} , since we always consider $F_{i,b_{i+1}}$ to encode b_i . This cycle center is called the *active* cycle center of level i , while $F_{i,1-b_{i+1}}$ is called *inactive*. A cycle center can additionally be *closed* or *open*. These terms are used to formalize when a bit is interpreted as 0 or 1. To be precise, b_i is interpreted as 1 if and only if $F_{i,b_{i+1}}$ is closed. In this way, cycle centers encode binary numbers. Since bit $i + 1$ switches every second time bit i switches, counting from 0 to $2^n - 1$ in binary then results in an alternating and balanced usage of both cycle centers of any level as required by Zadeh’s pivot rule.

Table 1 Edges and vertex priorities of S_n

Vertex	Successors	Priority
b_i	g_i, b_{i+1}	3
$e_{i,j,k}$	b_2, g_1	3
g_i	$F_{i,0}, F_{i,1}$	$2i + 9$
$h_{i,0}$	b_{i+2}	$2i + 10$
$s_{i,j}$	$h_{i,j}, b_1$	$10 - 2 \cdot j$
$d_{i,j,k}$	$F_{i,j}, e_{i,j,k}$	3
$F_{i,j}$	$d_{i,j,0}, d_{i,j,1}, s_{i,j}$	$6 - 2 \cdot j$
$h_{i,1}$	g_{i+1}	$2i + 10$

We now describe the construction of a parity game that implements this idea in detail. Fix some $n \in \mathbb{N}$. The vertex sets V_0, V_1 of the underlying graph are composed as follows:

$$V_0 := \{b_i, g_i : i \in \{1, \dots, n\}\} \cup \{d_{i,j,k}, e_{i,j,k}, s_{i,j} : i \in \{1, \dots, n - j\}, j, k \in \{0, 1\}\}$$

$$V_1 := \{F_{i,j} : i \in \{1, \dots, n - j\}, j \in \{0, 1\}\} \cup \{h_{i,j} : i \in \{1, \dots, n - j\}, j \in \{0, 1\}\} \cup \{t\}$$

The priorities of the vertices and their sets of outgoing edges are given by Table 1. Note that every vertex $v \in V_0$ has at most two outgoing edges. For convenience of notation, we henceforth identify the node names b_i and g_i for $i > n$ with t . The graph can be separated into n levels, where the levels $i < n - 1$ are structurally identical and the levels $n - 1$ and n differ slightly from the other levels. The i -th level is shown in Fig. 2, the complete graph of S_3 is shown in Fig. 3.

The general idea of the construction is the following. Certain pairs of player 0 strategies σ and counterstrategies τ^σ are interpreted as representing a number $b \in \mathcal{B}_n$. Such a pair of strategies induces a path starting at b_1 and ending at t , traversing the

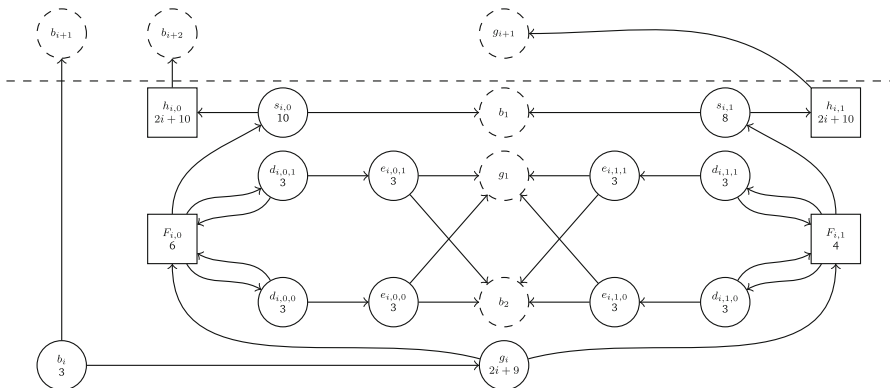


Fig. 2 Level i of S_n for $i \in \{1, \dots, n - 2\}$. Circular vertices are player 0 vertices, rectangular vertices are player 1 vertices. Labels below vertex names denote their priorities. Dashed vertices do not (necessarily) belong to level i

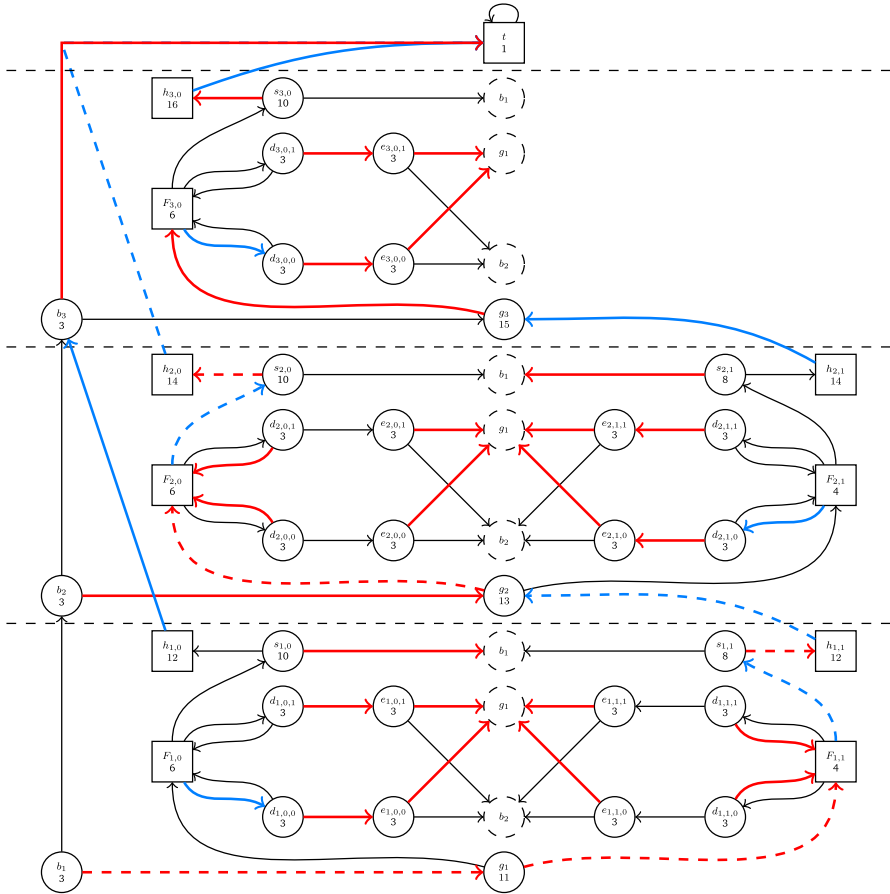


Fig. 3 The graph S_3 together with a canonical strategy representing the number 3 in the graph S_3 . The dashed copies of the vertices g_1 , b_1 and b_2 all refer to the corresponding vertices of levels 1 and 2. Red edges belong to the strategy of player 0, blue edges belong to the counterstrategy of player 1. The dashed edges indicate the spinal path

levels $i \in \{1, \dots, n\}$ with $b_i = 1$ while ignoring levels with $b_i = 0$. This path is called the *spinal path* with respect to $b \in \mathcal{B}_n$. Ignoring and including levels in the spinal path is controlled by the *entry vertex* b_i of each level $i \in \{1, \dots, n\}$. To be precise, when b is represented, the entry vertex of level i is intended to point towards the *selector vertex* g_i of level i if and only if $b_i = 1$. Otherwise, i.e., when $b_i = 0$, level i is ignored and the entry vertex b_i points towards the entry vertex of the next level.

Consider a level $i \in \{1, \dots, n - 2\}$. Attached to the selector vertex g_i are the *cycle centers* $F_{i,0}$ and $F_{i,1}$ of level i . As described at the beginning of this section, these player 1 vertices are the main structures used for interpreting whether the bit i is equal to one. They alternate in encoding bit i . As discussed before, this is achieved by interpreting the *active cycle center* $F_{i,b_{i+1}}$ as encoding b_i while the *inactive cycle center* $F_{i,1-b_{i+1}}$ does not interfere with the encoding. This enables us to manipulate

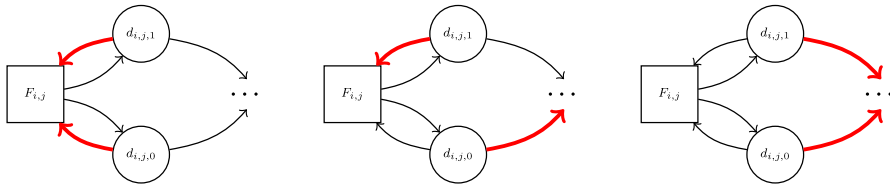


Fig. 4 A closed, a halfopen and an open cycle center. Edges of the strategy of player 0 are marked in red. Choices of player 1 are not depicted. Dots represent that the cycle vertices point to some unspecified vertex

the inactive part of a level without losing the encoded value of b_i . To this end, the selector vertex g_i is used to ensure that the active cycle center is contained in the spinal path.

As discussed previously, a cycle center $F_{i,j}$ can have different configurations. To be precise, it can be *closed*, *halfopen*, or *open*. The configuration of $F_{i,j}$ is defined via the *cycle vertices* $d_{i,j,0}$ and $d_{i,j,1}$ of the cycle center and the two *cycle edges* $(d_{i,j,0}, F_{i,j})$ and $(d_{i,j,1}, F_{i,j})$. More precisely, $F_{i,j}$ is *closed* with respect to a player 0 strategy σ if both cycle vertices point towards the cycle center, i.e., when $\sigma(d_{i,j,0}) = \sigma(d_{i,j,1}) = F_{i,j}$. If this is the case for exactly one of the two edges, the cycle center $F_{i,j}$ is called *halfopen*. A cycle that is neither closed nor halfopen is called *open*. An example of the different configurations is given in Fig. 4.

In addition, the cycle center is connected to its *upper selection vertex* $s_{i,j}$. It connects the cycle center $F_{i,j}$ with the first level via $(s_{i,j}, b_1)$ and with either level $i + 1$ or $i + 2$ via $(s_{i,j}, h_{i,j})$ via the respective edge $(h_{i,0}, b_{i+2})$ or $(h_{i,1}, g_{i+1})$ (depending on j). This vertex is thus central in allowing $F_{i,j}$ to get access to either the beginning of the spinal path or the next level of the spinal path.

We next discuss the cycle vertices. If their cycle centers are not closed, these vertices still need to be able to access the spinal path. The valuation of vertices along this path is usually very high and it is almost always very profitable for player 0 vertices to get access to this path. Since the cycle vertices cannot obtain access via the cycle center (as this would, by definition, close the cycle center) they need to “escape” the level in another way. This is handled by the *escape vertices* $e_{i,j,0}$ and $e_{i,j,1}$. The escape vertices are used to connect the levels with higher indices to the first two levels and thus enable each vertex to access the spinal path. To be precise, they are connected with the entry vertex of level 2 and the selector vertex of level 1. In principle, the escape vertices will point towards g_1 when the least significant set bit of the currently represented number has the index 1 and towards b_2 otherwise.

We now formalize the idea of a strategy encoding a binary number by defining the notion of a *canonical strategy*. Note that the definition also includes some aspects that are purely technical, i.e., solely required for some proofs, and do not have an immediate intuitive interpretation.

Definition 2 Let $\mathfrak{b} \in \mathcal{B}_n$. A player 0 strategy σ for the Parity Game S_n is called a *canonical strategy for \mathfrak{b}* if it has the following properties.

1. All escape vertices point to g_1 if $b_1 = 1$ and to b_2 if $b_1 = 0$.
2. The following hold for all levels $i \in \{1, \dots, n\}$ with $b_i = 1$:

- (a) Level i needs to be accessible, i.e., $\sigma(b_i) = g_i$.
 - (b) The cycle center $F_{i, b_{i+1}}$ needs to be closed while $F_{i, 1-b_{i+1}}$ must not be closed.
 - (c) The selector vertex of level i needs to select the active cycle center, i.e., $\sigma(g_i) = F_{i, b_{i+1}}$.
3. The following hold for all levels $i \in \{1, \dots, n\}$ with $b_i = 0$:
- (a) Level i must not be accessible and needs to be “avoided”, i.e., $\sigma(b_i) = b_{i+1}$.
 - (b) The cycle center $F_{i, b_{i+1}}$ must not be closed.
 - (c) If the cycle center $F_{i, 1-b_{i+1}}$ is closed, then $\sigma(g_i) = F_{i, 1-b_{i+1}}$.
 - (d) If neither of the cycle centers $F_{i, 0}, F_{i, 1}$ is closed, then $\sigma(g_i) = F_{i, 0}$.
4. Let $b_{i+1} = 0$. Then, level $i + 1$ is not accessible from level i , i.e., $\sigma(s_{i,0}) = h_{i,0}$ and $\sigma(s_{i,1}) = b_1$.
5. Let $b_{i+1} = 1$. Then, level $i + 1$ is accessible from level i , i.e., $\sigma(s_{i,0}) = b_1$ and $\sigma(s_{i,1}) = h_{i,1}$.
6. If $b < 2^n - 1$, then both cycle centers of level $v(b + 1)$ are open.

We use σ_b to denote a canonical strategy for $b \in \mathcal{B}_n$. A canonical strategy representing $(0, 1, 1)$ in S_3 is shown in Fig. 3.

As mentioned before, the main structure that is used to determine whether a bit is interpreted as being set are the cycle centers. In fact, any configuration of the cycle centers can be interpreted as an encoded number in the following way.

Definition 3 Let σ be a player 0 strategy for S_n . Then, the *induced bit state* $\beta^\sigma = (\beta_n^\sigma, \dots, \beta_1^\sigma)$ is defined inductively as follows. We define $\beta_n^\sigma = 1$ if and only if $\sigma(d_{n,0,0}) = \sigma(d_{n,0,1}) = F_{n,0}$ and $\beta_i^\sigma = 1$ if and only if $\sigma(d_{i, \beta_{i+1}^\sigma, 0}) = \sigma(d_{i, \beta_{i+1}^\sigma, 1}) = F_{i, \beta_{i+1}^\sigma}$ for $i < n$

This definition is in accordance with our interpretation of encoding a number as $\beta^{\sigma_b} = b$ if σ_b is a canonical strategy for b .

4 Lower bound for policy iteration on MDPs

In this section we discuss the Markov Decision Process (MDP) that is constructed analogously to the PG S_n . We discuss how this MDP allows the construction of a Linear Program (LP) such that the results obtained for the MDP carry over to the LP formulation. The main idea is to replace player 1 by the “random player” and to choose the probabilities in such a way that applying improving switches in the MDP behaves nearly the same way as in the PG. Note that we continue to use the same language for valuations, strategies and so on in MDP context, although other notions (like policy instead of strategy) are more common.

We give a brief introduction to the theory of MDPs (see also [2]). Similarly to a PG, an MDP is formally defined by its underlying graph (V_0, V_R, E, r, p) . Here, V_0 is the set of vertices controlled by player 0 and V_R is the set of randomization vertices. We let $V := V_0 \cup V_R$. For $p \in \{0, R\}$, we define $E_p := \{(v, w) : v \in V_p\}$. The set E_0 then corresponds to possible choices that player 0 can make, and each such choice

is assigned a *reward* by the reward function $r : E_0 \rightarrow \mathbb{R}$. The set E_R corresponds to probabilistic transitions and transition probabilities are specified by the function $p : E_R \rightarrow [0, 1]$ with $\sum_{u: (v,u) \in E_R} p(v, u) = 1$.

As for S_n , a (*player 0*) *strategy* is a function $\sigma : V_0 \rightarrow V$ that selects for each vertex $v \in V_0$ a target corresponding to an edge, i.e., such that $(v, \sigma(v)) \in E_0$. There are several computational tasks that can be investigated for MDPs. They are typically described via an *objective*. We consider the *expected total reward objective* for MDPs which can be formulated using vertex valuations in the following sense. Given an MDP, we define the *vertex valuations* $\mathcal{E}_\sigma^M(*)$ with respect to a strategy σ as the solution (if it exists) of the following set of equations:

$$\mathcal{E}_\sigma^M(u) := \begin{cases} r(u, \sigma(u)) + \mathcal{E}_\sigma^M(\sigma(u)), & u \in V_0, \\ \sum_{v: (u,v) \in E_R} p(u, v) \cdot \mathcal{E}_\sigma^M(v), & u \in V_R. \end{cases}$$

We also impose the condition that the values sum up to 0 on each irreducible recurrent class of the Markov chain defined by σ , yielding uniqueness [2]. We intentionally use very similar notation as for vertex valuations in the context of Parity Games since this allows for a unified treatment.

We now discuss the *Policy Iteration Algorithm* and refer to [23] for further details. Similarly to the Strategy Improvement Algorithm for PGs, this algorithm starts with some initial policy $\iota = \sigma_0$. In each step i , it generates a strategy σ_i by changing the target vertex $\sigma_{i-1}(v)$ of some vertex $v \in V_0$ to some vertex w with $\mathcal{E}_{\sigma_i}^M(w) > \mathcal{E}_{\sigma_{i-1}}^M(\sigma_{i-1}(v))$. For an arbitrary strategy σ , such an edge $(v, w) \in E_0$ with $w \neq \sigma(v)$ but $\mathcal{E}_\sigma^M(w) > \mathcal{E}_\sigma^M(\sigma(v))$ is called *improving switch* and the set of improving switches is denoted by I_σ . The term *optimal strategy* is defined as in PG context. In particular, a strategy σ is optimal if and only if $I_\sigma = \emptyset$. Moreover, applying an improving switch cannot decrease the valuation of any vertex. That is, if $e = (v, w) \in I_\sigma$ and σe denotes the strategy obtained after applying e to σ , then $\mathcal{E}_{\sigma e}^M(v') \geq \mathcal{E}_\sigma^M(v')$ for all $v' \in V$ and $\mathcal{E}_{\sigma e}^M(v) > \mathcal{E}_\sigma^M(v)$. Since there are only finitely many strategies, the algorithm thus generates a finite sequence $\sigma_0, \sigma_1, \dots, \sigma_N$ with $I_{\sigma_N} = \emptyset$.

We now discuss how the counter introduced in Sect. 3 is altered to obtain an MDP M_n . A sketch of level i of M_n can be found in Fig. 5. First, all player 1 vertices are replaced by randomization vertices. This is a common technique used for obtaining MDPs that behave similarly to given PGs and was used before (e.g., [2, 10]). While the ideas used in the transformations are similar, there is no standard reduction from PGs to MDPs preserving all properties.

In our construction, all cycle centers $F_{i,j}$ and all vertices $h_{i,j}$ are now randomization vertices. As vertices of the type $h_{i,j}$ have only one outgoing edge, the probability of this edge is set to 1. For defining the probabilities of the cycle edges, we introduce a small parameter $\varepsilon > 0$ and defer its exact definition to later. The idea is to use ε to make the probabilities of edges $(F_{i,j}, s_{i,j})$ very small by setting $p(F_{i,j}, s_{i,j}) = \varepsilon$ and $p(F_{i,j}, d_{i,j,k}) = \frac{1-\varepsilon}{2}$ for $k \in \{0, 1\}$. Then, the valuation of $s_{i,j}$ can only contribute significantly to the valuation of $F_{i,j}$ if the cycle center is closed. If the cycle center is not closed, then the contribution of this vertex can often be neglected. However, there are situations in which even this very low contribution has a significant impact on the

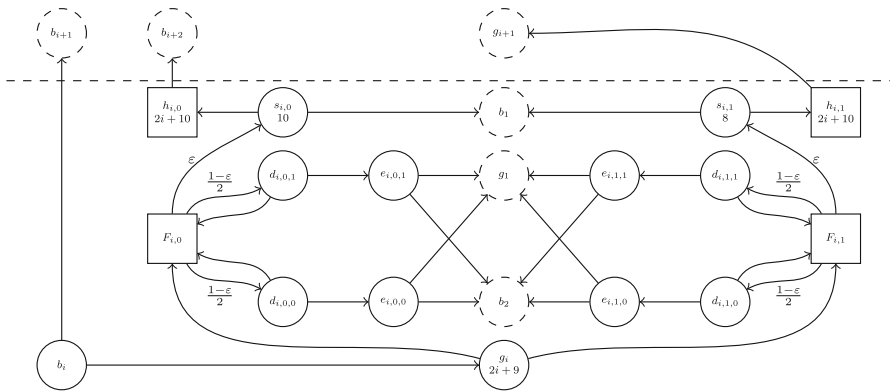


Fig. 5 Level i of the MDP. Circular vertices are vertices of the player, rectangular vertices are randomization vertices. Numbers below vertex names, if present, encode priorities Ω . If a vertex has priority $\Omega(v)$, then a reward of $\langle v \rangle := (-N)^{\Omega(v)}$ is associated with every edge leaving this vertex

valuation of the cycle center. For example, if $F_{i,0}$ and $F_{i,1}$ are both open for σ , then $\mathcal{E}_\sigma^M(F_{i,0}) > \mathcal{E}_\sigma^M(F_{i,1})$ if and only if $\mathcal{E}_\sigma^M(s_{i,0}) > \mathcal{E}_\sigma^M(s_{i,1})$. This sometimes results in a different behavior of the MDP when compared to the PG. We discuss this later in more detail.

Second, all player 0 vertices remain player 0 vertices. Each player 0 vertex is assigned the same priority as in S_n . This priority is now used to define the rewards of the edges leaving a vertex. More precisely, if we denote the priority of $v \in V_0$ by $\Omega(v)$, then we define the reward of any edge leaving v as $\langle v \rangle := (-N)^{\Omega(v)}$, where $N \geq 7n$ is a large and fixed parameter. Note that the reward of an edge thus only depends on its starting vertex. The reward function that is defined in that way then has the effect that vertices with an even priority are profitable while vertices with an odd priority are not profitable. In addition, the profitability of a vertex is better (resp. worse) the higher its priority is. By choosing a sufficiently large parameter N , it is also ensured that rewards are sufficiently separated. For example, the profitability of some vertex v with even priority cannot be dominated by traversing many vertices with lower but odd priorities. In principle, this ensures that the MDP behaves very similarly to the PG.

Having introduced the parameter N , we now fix the parameter ε such that $\varepsilon < (N^{2n+11})^{-1}$. Note that both parameters can be encoded by a polynomial number of bits with respect to the parameter n . By defining the reward of the edge (t, t) as 0, this completely describes the MDP.

We now provide more details on the aspects where the PG and the MDP differ. One of the main differences between the PG and the MDP is the set of canonical strategies. Consider a strategy σ representing some $b \in \mathcal{B}_n$, some level i and the two cycle centers $F_{i,0}, F_{i,1}$. In PG context, both vertices have an even priority and the priority of $F_{i,0}$ is larger than the priority of $F_{i,1}$. Thus, if both cycle centers escape the level, the valuation of $F_{i,0}$ is better than the valuation of $F_{i,1}$. Consequently, if $\sigma(g_i) \neq F_{i,0}$, then $(g_i, F_{i,0})$ is improving for σ . In some sense, this can be interpreted as the PG “preferring” $F_{i,0}$ over $F_{i,1}$. A similar, but not the same, phenomenon occurs

in MDP context. If both cycle centers $F_{i,0}$ and $F_{i,1}$ are in the same “state”, then the valuation of the two upper selection vertices $s_{i,0}, s_{i,1}$ determines which cycle center has the better valuation. It turns out that the valuation of $s_{i,b_{i+1}}$ is typically better than the valuation of $s_{i,1-b_{i+1}}$. It is in particular not true that the valuation of $s_{i,0}$ is typically better than the valuation of $s_{i,1}$. Hence, the MDP “prefers” vertices $F_{i,b_{i+1}}$ over vertices $F_{i,1-b_{i+1}}$. We thus adjust the definition of a canonical strategy in MDP context in the following way.

Definition 4 Let $b \in \mathcal{B}_n$. A player 0 strategy σ for the MDP M_n is called a *canonical strategy for b* if it has the properties defined in Definition 2 where Property 3.(d) is replaced by the following: If neither of the cycle centers $F_{i,0}, F_{i,1}$ is closed, then $\sigma(g_i) = F_{i,b_{i+1}}$.

5 Lower bound for the simplex algorithm and linear programs

Following the arguments of [2, 15], we now discuss how the MDP can be transformed into an LP such that the results obtained for the Policy Iteration Algorithm may be transferred to the Simplex Algorithm. This transformation makes use of the *unichain condition*. This condition (see [34]) states that the Markov Chain obtained from each strategy σ has a single irreducible recurrent class. Unfortunately, the MDP constructed previously does not fulfill the unichain condition. As we prove in Lemma 1, it however fulfills a *weak* version of the unichain condition. This weak version states that the *optimal* policy has a single irreducible recurrent class and does not demand this to be true for every strategy. This implies that the same LP which can be obtained by transforming an MDP fulfilling the unichain condition can be used. We refer to [14, 40] for more details.

We thus return to the discussion for MDPs fulfilling the unichain condition. Optimal policies for MDPs fulfilling this condition can be found by solving the following Linear Program:

$$\begin{aligned}
 & \max \sum_{(u,v) \in E_0} r(u, v) \cdot x(u, v) \\
 & \text{s.t. } \sum_{(u,v) \in E} x(u, v) - \sum_{\substack{(v,w) \in E_0 \\ (w,u) \in E_R}} p(w, u) \cdot x(v, w) = 1 \quad \forall u \in V_0 \quad (\text{P}) \\
 & x(u, v) \geq 0 \quad \forall (u, v) \in E_0
 \end{aligned}$$

The variable $x(u, v)$ for $(u, v) \in E_0$ represents the probability (or frequency) of using the edge (u, v) . The constraints of (P) ensure that the probability of entering a vertex u is equal to the probability of exiting u . It is not difficult to see that the basic feasible solutions of (P) correspond directly to strategies of the MDP, see, e.g., [2]. For each strategy σ we can define a feasible setting of the variables $x(u, v)$ with $(u, v) \in E_0$ such that $x(u, v) > 0$ only if $\sigma(u) = v$. Conversely, for every basic feasible solution of (P), we can define a corresponding policy σ . It is well-known that

the policy corresponding to an optimal basic feasible solution of (P) is an optimal policy for the MDP (see, e.g., [2, 34]).

Our MDP only fulfills the weak unichain condition. If it is provided an initial strategy that has the same single irreducible recurrent class as the optimal policy, then the same Linear Program introduced above can be used [14]. This follows since all considered basic feasible solutions will have the same irreducible recurrent class by monotonicity. We refer to [40] for more details.

6 Lower bound proof

6.1 The approach and basic definitions

In this section we outline the proof for the exponential lower bound on the running time of the Strategy Improvement resp. Policy Iteration Algorithm using Zadeh's pivot rule and a strategy-based tie-breaking rule. We discuss the following key components separately before combining them into our main result.² In the following, we use the notation G_n to simultaneously refer to S_n and M_n . If a statement or definition only holds for either S_n or M_n , we explicitly state this.

1. We first define an initial strategy ι such that the pair (G_n, ι) defines a sink game in PG context resp. has the weak unichain condition in MDP context. We also formalize the idea of counting how often an edge has been applied as improving switch.
2. We then state and discuss the tie-breaking rule. Together with the initial strategy, this completely describes the application of the improving switches performed by the Strategy Improvement resp. Policy Iteration Algorithm. Further statements, proofs and explanations that are provided in Appendix C thus only serve to prove that the algorithms and the tie-breaking rule indeed behave as intended.
3. We then focus on a single transition from a canonical strategy σ_b to the next canonical strategy σ_{b+1} . During such a transition, many improving switches need to be applied and thus many intermediate strategies need to be considered. These strategies are divided into five *phases*, depending on the configuration of G_n induced by the encountered strategies.
4. To prove that the tie-breaking rule indeed proceeds along the described phases, we need to specify how often player 0 edges are applied as improving switches, which is formalized by an *occurrence record*. We explicitly describe the occurrence records for canonical strategies.
5. Finally, we combine the previous aspects to prove that applying the respective algorithms with Zadeh's pivot rule and our tie-breaking rule yields an exponential number of iterations.

We begin by providing the initial strategy ι for G_n . In principle, the initial strategy is a canonical strategy for 0 in the sense of Definition 2 resp. 4.

² Formal proofs can be found in Appendix C and the full version [6].

Definition 5 The initial player 0 strategy $\iota: V_0 \mapsto V$ is defined as follows:

v	$b_i (i < n)$	b_n	g_i	$d_{i,j,k}$	$e_{i,j,k}$	$s_{i,0}$	$s_{i,1} (i < n)$
$\iota(v)$	b_{i+1}	t	$F_{i,0}$	$e_{i,j,k}$	b_2	$h_{i,0}$	b_1

We further introduce the notion of a *reachable* strategy. A strategy σ' is reachable from some strategy σ if it can be produced by the Strategy Improvement Algorithm starting from σ and applying a finite number of improving switches. Note that the notion of reachability does not depend on the pivot rule or the tie-breaking rule and that every strategy calculated by the Strategy Improvement resp. Policy Iteration Algorithm is reachable by definition.

Definition 6 Let σ be a player 0 strategy for G_n . The set of all strategies that can be obtained from σ by applying an arbitrary sequence of improving switches is denoted by Λ_σ . A strategy σ' is *reachable from σ* if $\sigma' \in \Lambda_\sigma$.

Note that reachability is a transitive property and that we include $\sigma \in \Lambda_\sigma$ for convenience. The i -th level of the initial strategy is shown in Fig. 6. The initial strategy is chosen such that G_n and ι define a sink game in S_n resp. have the weak unichain condition in M_n .

Lemma 1 For all $n \in \mathbb{N}$, the game G_n and the initial player 0 strategy ι define a Sink Game with sink t in PG context, resp. have the weak unichain condition in MDP context.

As Zadeh’s pivot rule is a memorizing pivot rule, the algorithms need to maintain information about how often edges have been applied as improving switches. During the execution of the algorithms, we thus maintain an *occurrence record* $\phi^\sigma: E_0 \mapsto \mathbb{R}$ that specifies how often an improving switch was applied since the beginning of the

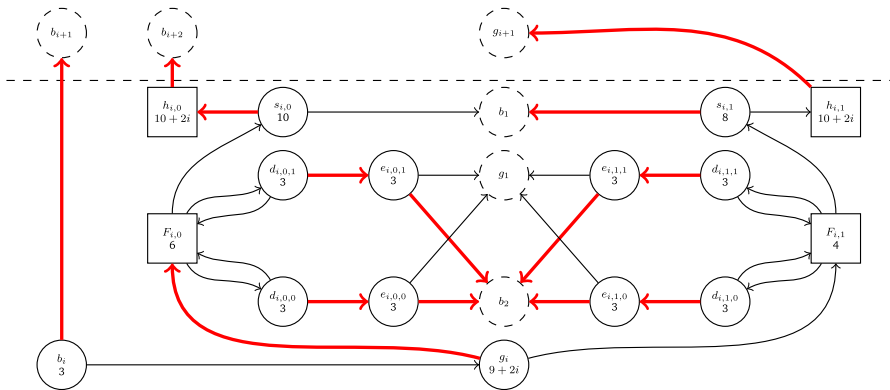


Fig. 6 The initial strategy ι (red edges) in level i for $i \in \{1, \dots, n - 2\}$

algorithms. Formally, we define $\phi^l(e) := 0$ for every edge $e \in E_0$, i.e., the occurrence record with respect to the initial strategy is equal to 0. Then, whenever the algorithms apply an edge e , the occurrence record of e is increased by 1.

6.2 The tie-breaking rule

We now discuss the tie-breaking rule. It specifies which edge to apply if there are multiple improving switches that minimize the occurrence record for the current strategy. Note that, in contrast to many classical pivot rules, fixing a specific tie-breaking rule cannot easily be avoided for Zadeh’s pivot rule, since it is unavoidable that occurrence records occasionally coincide between multiple improving switches. Intuitively, asking for a lower bound construction to work for all tie-breaking rules might be comparably difficult to asking for a construction that works for all pivot rules in the first place.

We rely on a structurally simple tie-breaking rule that is in principle implemented as an ordering of the set E_0 and depends on the current strategy σ as well as the occurrence records. In fact, it turns out to be sufficient to specify a pre-order of E_0 that can be extended to a total order arbitrarily. Whenever the algorithms have to break ties, they then choose the first edge according to this ordering. For convenience, there is one small exception from this behavior. During the transition from the canonical strategy representing 1 towards the canonical strategy representing 2, one improving switch e (which we do not specify yet) has to be applied earlier than during other transitions. The reason is that the occurrence records of several edges, including e , are still zero at this point in time, which leads to unwanted behavior. In later iterations, e is the unique improving switch minimizing the occurrence record whenever it has to be applied, so no special treatment is necessary.

Let σ be a player 0 strategy for G_n . Henceforth, we use the symbol $*$ as a wildcard. More precisely, when using the symbol $*$, this means any suitable index or vertex (depending on the context) can be inserted for $*$ such that the corresponding edge exists. For example, the set $\{(e_{*,*,*}, *)\}$ would then denote the set of all edges starting in escape vertices. Using this notation, we define the following sets of edges.

- $\mathbb{G} := \{(g_i, F_{i,*})\}$ is the set of all edges leaving selector vertices.
- $\mathbb{E}^0 := \{(e_{i,j,k}, *) : \sigma(d_{i,j,k}) \neq F_{i,j}\}$ is the set of edges leaving escape vertices whose cycle vertices do not point towards their cycle center. Similarly, $\mathbb{E}^1 := \{(e_{i,j,k}, *) : \sigma(d_{i,j,k}) = F_{i,j}\}$ is the set of edges leaving escape vertices whose cycle vertices point towards their cycle center.
- $\mathbb{D}^1 := \{(d_{*,*,*}, F_{*,*})\}$ is the set of cycle edges and $\mathbb{D}^0 := \{(d_{*,*,*}, e_{*,*,*})\}$ is the set of the other edges leaving cycle vertices.
- $\mathbb{B}^0 := \bigcup_{i=1}^{n-1} \{(b_i, b_{i+1})\} \cup \{(b_n, t)\}$ is the set of all edges between entry vertices. The set $\mathbb{B}^1 := \{(b_*, g_*)\}$ of all edges leaving entry vertices and entering selection vertices is defined analogously and $\mathbb{B} := \mathbb{B}^0 \cup \mathbb{B}^1$ is the set of all edges leaving entry vertices.
- $\mathbb{S} := \{(s_{*,*}, *)\}$ is the set of all edges leaving upper selection vertices.

We next define two pre-orders based on these sets. However, we need to define finer pre-orders for the sets \mathbb{E}^0 , \mathbb{E}^1 , \mathbb{S} and \mathbb{D}^1 first.

Informally, the pre-order on \mathbb{E}^0 forces the algorithms to favor switches of higher levels and to favor $(e_{i,0,k}, *)$ over $(e_{i,1,k}, *)$ in S_n and $(e_{i,\beta_{i+1}^\sigma,k}, *)$ over $(e_{i,1-\beta_{i+1}^\sigma,k}, *)$ in M_n . For a formal description let $(e_{i,j,x}, *)$, $(e_{k,l,y}, *) \in \mathbb{E}^0$. In S_n , we define $(e_{i,j,x}, *) <_\sigma (e_{k,l,y}, *)$ if either $i > k$, or $i = k$ and $j < l$. In M_n , we define $(e_{i,j,x}, *) <_\sigma (e_{k,l,y}, *)$ if either $i > k$, or $i = k$ and $j = \beta_{i+1}^\sigma$.

Similarly, the pre-order on \mathbb{S} also forces the algorithm to favor switches of higher levels. Thus, for $(s_{i,j}, *)$, $(s_{k,l}, *) \in \mathbb{S}$, we define $(s_{i,j}, *) <_\sigma (s_{k,l}, *)$ if $i > k$.

We now describe the pre-order for \mathbb{E}^1 . Let $(e_{i,j,x}, *)$, $(e_{k,l,y}, *) \in \mathbb{E}^1$.

1. The first criterion encodes that switches contained in higher levels are applied first. Thus, if $i > k$, then $(e_{i,j,x}, *) <_\sigma (e_{k,l,y}, *)$.
2. If $i = k$, then we consider the states of the cycle centers $F_{i,j}$ and $F_{k,l} = F_{i,1-j}$. If exactly one cycle center of level i is closed, then the improving switches within this cycle center are applied first.
3. Consider the case where $i = k$ but no cycle center of level i is closed. Let $t^\rightarrow := b_2$ if $v(b+1) > 1$ and $t^\rightarrow := g_1$ if $v(b+1) = 1$. If there is exactly one halfopen cycle center escaping to t^\rightarrow in level i , then switches within this cycle center have to be applied first.
4. Assume that none of the prior criteria applied. This includes the case where both cycle centers are in the same state, and $i = k$ holds in this case. Then, the order of application depends on whether we consider S_n or M_n . In S_n , improving switches within $F_{i,0}$ are applied first. In M_n , improving switches within F_{i,β_{i+1}^σ} are applied first.

We next give a pre-order for \mathbb{D}_1 . Let $(d_{i,j,x}, F_{i,j})$, $(d_{k,l,y}, F_{k,l}) \in \mathbb{D}_1$.

1. The first criterion states that improving switches that are part of open cycles are applied first. We thus define $(d_{i,j,x}, F_{i,j}) <_\sigma (d_{k,l,y}, F_{k,l})$ if $\sigma(d_{k,l,1-y}) = F_{k,l}$ but $\sigma(d_{i,j,1-x}) \neq F_{i,j}$.
2. The second criterion states the following. Among all halfopen cycle centers, improving switches contained in cycle centers such that the bit of the level the cycle center is part of is equal to zero are applied first. If the first criterion does not apply, we thus define $(d_{i,j,x}, F_{i,j}) <_\sigma (d_{k,l,y}, F_{k,l})$ if $\beta_k^\sigma > \beta_i^\sigma$.
3. The third criterion states that among all partially closed cycle centers, improving switches inside cycle centers contained in lower levels are applied first. If none of the first two criteria apply, we thus define $(d_{i,j,x}, F_{i,j}) <_\sigma (d_{k,l,y}, F_{k,l})$ if $k > i$.
4. The fourth criterion states that improving switches within the active cycle center are applied first within one level. If none of the previous criteria apply, we thus define $(d_{i,j,x}, F_{i,j}) <_\sigma (d_{k,l,y}, F_{k,l})$ if $\beta_{k+1}^\sigma \neq l$ and $\beta_{i+1}^\sigma = j$.
5. The last criterion states that edges with last index equal to zero are preferred within one cycle center. That is, if none of the previous criteria apply, we define $(d_{i,j,x}, F_{i,j}) <_\sigma (d_{k,l,y}, F_{k,l})$ if $x < y$. If this criterion does not apply either, the edges are incomparable.

We now define the pre-order $<_\sigma$ and the tie-breaking rule, implemented by an ordering of E_0 .

Definition 7 Let σ be a player 0 strategy for G_n and $\phi^\sigma : E^0 \rightarrow \mathbb{N}_0$ be an occurrence record. We define the pre-order $<_\sigma$ on E_0 by defining the set-based pre-order

$$\mathbb{G} <_\sigma \mathbb{D}^0 <_\sigma \mathbb{E}^1 <_\sigma \mathbb{B} <_\sigma \mathbb{S} <_\sigma \mathbb{E}^0 <_\sigma \mathbb{D}^1$$

where the sets $\mathbb{E}^0, \mathbb{E}^1, \mathbb{S}$ and \mathbb{D}^1 are additionally pre-ordered as described before. We extend the pre-order to an arbitrary but fixed total ordering and denote the corresponding order also by $<_\sigma$. We define the following tie-breaking rule: *Let I_σ^{\min} denote the set of improving switches with respect to σ that minimize the occurrence record. Apply the first improving switch contained in I_σ^{\min} with respect to the ordering $<_\sigma$ with the following exception: If $\phi^\sigma(b_1, b_2) = \phi^\sigma(s_{1,1}, h_{1,1}) = 0$ and $(b_1, b_2), (s_{1,1}, h_{1,1}) \in I_\sigma$, apply $(s_{1,1}, h_{1,1})$ immediately.*

We usually just use the notation $<$ to denote the ordering if it is clear from the context which strategy is considered and whether $<$ is defined via the default or special pre-order.

Lemma 2 *Given a strategy $\sigma \in \Lambda_t$ and an occurrence record $\phi^\sigma : E^0 \rightarrow \mathbb{N}_0$, the tie-breaking rule can be evaluated in polynomial time.*

6.3 The phases of a transition and the application of improving switches

As explained earlier, the goal is to prove that Zadeh’s pivot rule with our tie-breaking rule enumerates at least one strategy per number $b \in \mathcal{B}_n$. This is proven in an inductive fashion. That is, we prove that given a canonical strategy σ_b for $b \in \mathcal{B}_n$, the algorithms eventually calculate a canonical strategy σ_{b+1} for $b + 1$. This process is called a *transition* and each transition is partitioned into up to five phases. In each phase, a different “task” is performed in order to obtain the strategy σ_{b+1} . These tasks are, for example, the opening and closing of cycle centers, updating the escape vertices or adjusting some of the selection vertices.

Depending on whether we consider S_n or M_n and $v(b + 1)$, there can be 3,4 or 5 different phases. Phases 1,3 and 5 always take place while Phase 2 only occurs if $v(b + 1) > 1$, as it updates the target vertices of some selection vertices $s_{i,j}$ with $i < v(b + 1)$. The same holds for Phase 4, although this phase only exists if we consider S_n . In M_n , we apply the corresponding switches already in Phase 3 and there is no separate Phase 4.

We now give a detailed description of the phases. For the sake of the presentation we only describe the main function of each phase and omit switches that are applied for technical reasons. Furthermore, we abbreviate $v := v(b + 1)$ whenever b is clear from the context. See Fig. 7 for a full example of the phases in a transition between two canonical strategies.

1. During Phase 1, cycle centers are closed such that the induced bit state of the final strategy is $b + 1$. Furthermore, several cycle edges are switched such that the occurrence records of these switches are as balanced as possible. In the end of the phase, the cycle center $F_{v,(b+1)_{v+1}}$ is closed and either Phase 2 or Phase 3 begins.

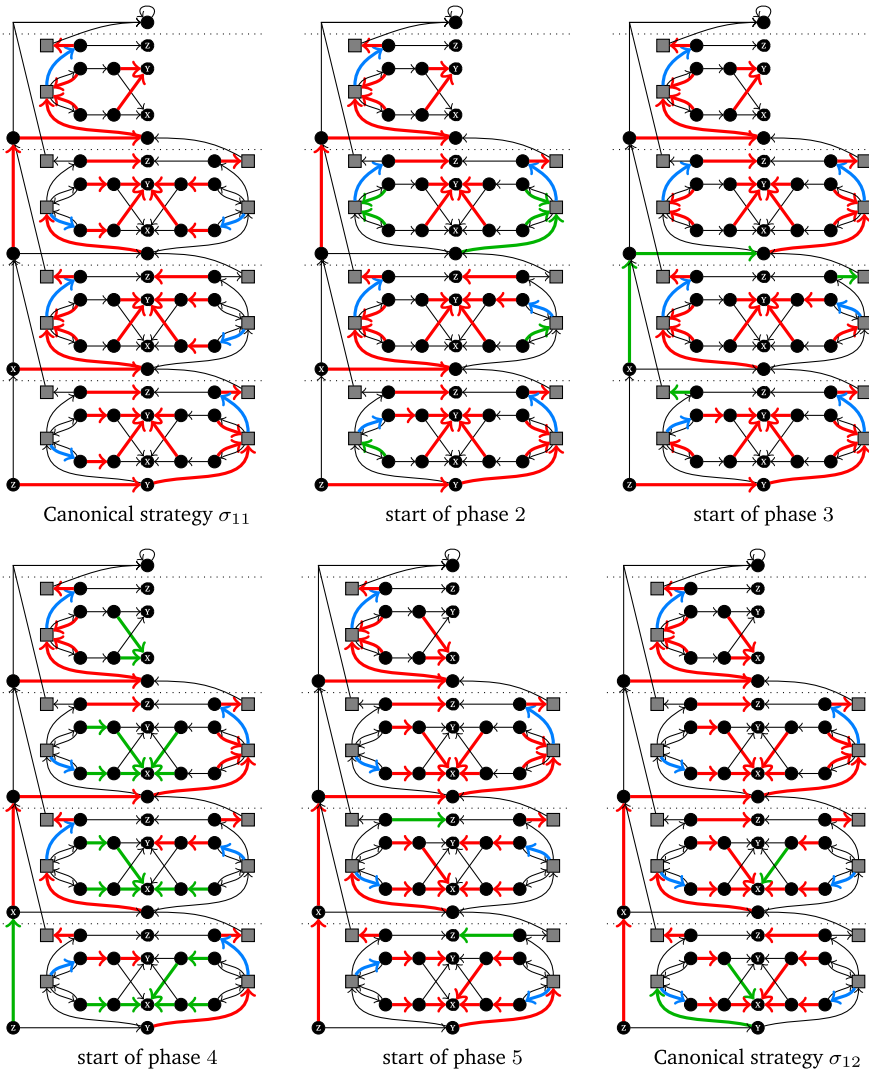


Fig. 7 Visualization of the phases encountered when transitioning from σ_{11} to σ_{12} in S_4 . Each subfigure shows the state of the counter at the beginning of a phase. Red and green edges correspond to the strategy of player 0, blue edges to the strategy of player 1. Green edges show improving switches that were applied in the preceding phase. The edges of vertices with an outdegree of 1 are not marked as these are chosen in every strategy. All occurrences of vertices with the same label refer to the same unique vertex

2. During Phase 2, the upper selection vertices $s_{i,j}$ for $i \in \{1, \dots, v - 1\}$ and $j = (b + 1)_{i+1}$ change their targets to $h_{i,j}$. This is necessary as the induced bit state of the strategy is now equal to $b + 1$. Also, the entry vertices of these levels are switched towards the entry vertex of the next level. Since $b_i = b_{i+1}$ for all $i \neq 1$ if $v(b + 1) = 1$, these operations only need to be performed if $v(b + 1) > 1$.

3. Phase 3 is partly responsible for applying improving switches involving escape vertices. Since $v(b) \neq v(b + 1)$, all escape vertices need to change their target vertices. In Phase 3, some (but not all) escape vertices perform the corresponding switch. Also, for some of these escape vertices $e_{i,j,k}$, the switch $(d_{i,j,k}, e_{i,j,k})$ is applied. This later enables the application of the switch $(d_{i,j,k}, F_{i,j})$ which is necessary to balance the occurrence records of the cycle edges. At the end of this phase, depending on v , either (b_1, g_1) or (b_1, b_2) is applied. In M_n , the switches described for Phase 4 are also applied during Phase 3.
4. During Phase 4, the upper selection vertices $s_{i,j}$ for $i \in \{1, \dots, v - 1\}$ and $j \neq (b + 1)_{i+1}$ change their targets to b_1 . Updating the upper selection vertices is necessary since they need to give their cycle centers access to the spinal path. Similar to Phase 2, these switches are only performed if $v(b + 1) > 1$.
5. During Phase 5, the remaining escape vertices switch their targets and some of the cycle vertices switch to their cycle centers. This phase ends once all improving switches at the escape vertices are performed, yielding a canonical strategy for $b + 1$.

We next give the formal definition of the different phases. For this, we need to introduce a strategy-based parameter $\mu^\sigma \in \{1, \dots, n + 1\}$. This parameter μ^σ is called the *next relevant bit* of the strategy σ . Before defining this parameter formally, we briefly explain its importance and how it can be interpreted.

As described in Sect. 3, both cycle centers of level i alternate in encoding bit i . Therefore, the selection vertex g_i needs to select the correct cycle center and the entry vertex b_i should point towards g_i if and only if bit i is equal to one (see Definition 2). In particular, the selection vertex g_{i-1} of level $i - 1$ needs to be in accordance with the entry vertex b_i of level i if bit $i - 1$ is equal to one. That is, it should not happen that $\sigma(b_i) = g_i$ and $\sigma(b_{i+1}) = g_{i+1}$ but $\sigma(g_i) = F_{i,0}$. However, we cannot guarantee that this does not happen for some intermediate strategies. Therefore, we need to perform some operations within the levels i and $i - 1$ and define μ^σ as the lowest level higher than any level that is set “incorrectly” in that sense. If there are no such levels, the parameter denotes the index of the lowest level with $\sigma(b_i) = b_{i+1}$. The parameter can thus be interpreted as an indicator encoding where “work needs to be done next”. As its formal definition is rather complex, we introduce an additional notation $\bar{\sigma}$ that somehow “encodes” σ using integer numbers. As this notation is however a pure formal tool, we do not fully introduce it here. The definition is provided in Appendix B. Formally, the next relevant bit is now defined as follows.

Definition 8 Let $\sigma \in \Lambda_t$. The set $\mathcal{I}^\sigma := \{i \in \{1, \dots, n\} : \bar{\sigma}(b_i) \wedge \bar{\sigma}(g_i) \neq \bar{\sigma}(b_{i+1})\}$ is the *set of incorrect levels*. The *next relevant bit* μ^σ of the strategy σ is defined by

$$\mu^\sigma := \begin{cases} \min(\{n\} \cup \{i' > \max\{i \in \mathcal{I}^\sigma\} : \bar{\sigma}(b_{i'}) \wedge \bar{\sigma}(g_{i'}) = \bar{\sigma}(b_{i'+1})\}) & \text{if } \mathcal{I}^\sigma \neq \emptyset, \\ \min(\{i \in \{1, \dots, n + 1\} : \sigma(b_i) = b_{i+1}\}) & \text{if } \mathcal{I}^\sigma = \emptyset. \end{cases}$$

Note that we interpret expressions of the form $x \wedge y = z$ as $x \wedge (y = z)$. Using the next relevant bit μ^σ , we now give a formal definition of the phases. Formally, a strategy belongs to one of the five phases if it has a certain set of properties.

These properties can be partitioned into several categories and are described in Table 5 in Appendix B. Each of these properties depends on either the level or the cycle center. Properties **Bac1**, **Bac2** and **Bac3** involve the entry vertices \mathbf{b} , Properties **Usv1** and **Usv2** involve the **Upper Selection Vertices** and Properties **Esc1**, **Esc2**, **Esc3**, **Esc4** and **Esc5** the **Escape vertices**. In addition, Properties **Rel1** and **Rel2** involve the next **relevant bit** μ^σ , Properties **Cc1**, **Cc2** the **Cycle Centers** and Property **Sv1** the **Selection Vertices**.

For most of the phases, there are additional *special* conditions that need to be fulfilled. The corresponding conditions simplify the distinction between different phases and allow for an easier argumentation or description of statements. However, we do not discuss these here as they are solely needed for technical reasons and we refer to Appendix B for details. We also provide the table used for the exact definition of the phases there.

Definition 9 (*Phase- k -strategy*) Let $\mathbf{b} \in \mathcal{B}_n$, $\sigma \in \Lambda_i$ and $k \in \{1, \dots, 5\}$. The strategy σ is a *Phase- k -strategy* for \mathbf{b} if it has the properties of the k -th column of Table 4 for the respective indices as well as the special conditions of the respective phase (if there are any).

6.4 The occurrence records

We next describe the actual occurrence records that occur when applying the Strategy Improvement resp. Policy Iteration Algorithm. To do so, we need to introduce notation related to binary counting.

The number of applications of specific edges in level i as improving switches depends on the last time the corresponding cycle centers were closed or how often they were closed. We thus define $\text{fl}(\mathbf{b}, i, \cdot)$ as the number of numbers smaller than \mathbf{b} with least significant set bit having index i . To quantify how often a specific cycle center was closed, we introduce the *maximal flip number* and the *maximal unflip number*. Let $\mathbf{b} \in \mathcal{B}_n$, $i \in \{1, \dots, n\}$ and $j \in \{0, 1\}$. Then, we define the *maximal flip number* $\text{mfn}(\mathbf{b}, i, \{(i+1, j)\})$ as the largest $\tilde{\mathbf{b}} \leq \mathbf{b}$ with $v(\tilde{\mathbf{b}}) = i$ and $\tilde{b}_{i+1} = j$. Similarly, we define the *maximal unflip number* $\text{mufn}(\mathbf{b}, i, \{(i+1, j)\})$ as the largest $\tilde{\mathbf{b}} \leq \mathbf{b}$ with $\tilde{b}_1 = \dots = \tilde{b}_i = 0$ and $\tilde{b}_{i+1} = j$. If there are no such numbers, then $\text{mfn}(\mathbf{b}, i, \{(i+1, j)\}) := 0$, $\text{mufn}(\mathbf{b}, i, \{(i+1, j)\}) := 0$. If we do not impose the condition that bit $i+1$ needs to be equal to j then we omit the term in the notation, i.e., $\text{mfn}(\mathbf{b}, i) = \max(\{0\} \cup \{\mathbf{b}' \leq \mathbf{b} : v(\mathbf{b}') = i\})$ and $\text{mufn}(\mathbf{b}, i)$ is defined analogously.

These notations enable us to properly describe the occurrence records. We however do not describe the occurrence record for every strategy σ produced by the considered algorithms. Instead, we only give a description of the occurrence records for canonical strategies. When discussing the application of the improving switches, we later prove the following: Assuming that the occurrence records are described correctly for $\sigma_{\mathbf{b}}$, they are also described correctly for $\sigma_{\mathbf{b}+1}$ when improving switches are applied according to Zadeh's pivot rule and our tie-breaking rule.

Theorem 3 Let $\sigma_{\mathbf{b}}$ be a canonical strategy for $\mathbf{b} \in \mathcal{B}_n$ and assume that improving switches are applied as described in Sect. 6.3. Then Table 2 describes the occurrence records with respect to $\sigma_{\mathbf{b}}$.

Table 2 Occurrence records for the canonical strategy $\sigma_{\mathbf{b}}$

Edge e	$\phi^{\sigma_{\mathbf{b}}}(e)$	Edge e	$\phi^{\sigma_{\mathbf{b}}}(e)$
$(e_{i,j,k}, g_1)$	$\left\lceil \frac{\mathbf{b}}{2} \right\rceil$	(b_i, g_i)	$\text{fl}(\mathbf{b}, i)$
$(e_{i,j,k}, b_2)$		(b_i, b_{i+1})	$\text{fl}(\mathbf{b}, i) - b_i$
$(s_{i,j}, h_{i,j})$	$\text{fl}(\mathbf{b}, i + 1) - (1 - j) \cdot b_{i+1}$	$(d_{i,j,k}, e_{i,j,k})$	$\leq \begin{cases} \phi^{\sigma_{\mathbf{b}}}(e_{i,j,k}, g_1), & b_1 = 0 \\ \phi^{\sigma_{\mathbf{b}}}(e_{i,j,k}, b_2), & b_1 = 1 \end{cases}$
$(s_{i,j}, b_1)$	$\text{fl}(\mathbf{b}, i + 1) - j \cdot b_{i+1}$	$(g_i, F_{i,j})$	$\leq \min_{k \in \{0,1\}} \phi^{\sigma_{\mathbf{b}}}(d_{i,j,k}, F_{i,j})$
Condition	$\phi^{\sigma_{\mathbf{b}}}(d_{i,j,k}, F_{i,j})$	Tolerance	
$b_i = 1 \wedge b_{i+1} = j$	$\left\lceil \frac{\text{mfn}(\mathbf{b}, i, \{(i+1, j)\}) + 1 - k}{2} \right\rceil$	0	
$b_i = 0 \vee b_{i+1} \neq j$	$\min \left(\left\lceil \frac{\mathbf{b} + 1 - k}{2} \right\rceil, \ell^{\mathbf{b}}(i, j, k) + t_{\mathbf{b}} \right)$	$t_{\mathbf{b}} \in \begin{cases} \{0\}, & i = 1 \vee b_i = 1 \\ \{0, 1\}, & i \neq 1 \wedge b_1 = 0 \\ \{-1, 0, 1\}, & i \neq 1 \wedge b_1 = 1 \end{cases}$	
$\ell^{\mathbf{b}}(i, j, k) := \left\lceil \frac{\text{mfn}(\mathbf{b}, i, \{(i+1, j)\}) + 1 - k}{2} \right\rceil + \mathbf{b} - \mathbf{1}_{j=0} \text{mfn}(\mathbf{b}, i + 1) - \mathbf{1}_{j=1} \text{mfn}(\mathbf{b}, i + 1)$			

Depending on the edge, we either give the exact occurrence record, give an upper bound, or we give the occurrence record up to a certain tolerance. A parameter $t_{\mathbf{b}}$ fulfilling the assumptions for the case $b_i = 0 \vee b_{i+1} \neq j$ is called *feasible* for \mathbf{b}

We now give some intuition for the occurrence records of parts of Table 2. As the occurrence records of most of the edges are much more complicated to explain, we omit an intuitive description of their occurrence records here. Let $\sigma_{\mathbf{b}} \in \Lambda_i$ be a canonical strategy for $\mathbf{b} \in \mathcal{B}_n$.

Consider some edge (b_i, g_i) . This edge is applied as an improving switch whenever bit i switches from 0 to 1. That is, it is applied if and only if we transition towards some $\mathbf{b}' \in \mathcal{B}_n$ with $v(\mathbf{b}') = i$ and $\mathbf{b}' \leq \mathbf{b}$. Therefore, $\phi^{\sigma_{\mathbf{b}}}(b_i, g_i) = \text{fl}(\mathbf{b}, i)$. Now consider (b_i, b_{i+1}) . This edge is only applied as an improving switch when bit i switches from 1 to 0. This can however only happen if bit i switched from 0 to 1 earlier. That is, applying (b_i, b_{i+1}) can only happen when (b_i, g_i) was applied before. Also, we can only apply the switch (b_i, g_i) again after bit i has been switched back to 0 again, i.e., after (b_i, b_{i+1}) was applied. Consequently, $\phi^{\sigma_{\mathbf{b}}}(b_i, b_{i+1}) = \phi^{\sigma_{\mathbf{b}}}(b_i, g_i) - b_i = \text{fl}(\mathbf{b}, i) - b_i$.

Next, consider some edge $(s_{i,j}, h_{i,j})$ and fix $j = 1$ for now. This edge is applied as an improving switch if and only if bit $i + 1$ switches from 0 to 1. Hence, as discussed before, $\phi^{\sigma_{\mathbf{b}}}(s_{i,j}, h_{i,j}) = \text{fl}(\mathbf{b}, i + 1)$. Now let $j = 0$. The switch $(s_{i,0}, h_{i,0})$ is applied whenever bit $i + 1$ switches from 1 to 0. This requires the bit to have switched from 0 to 1 before. Therefore, $\phi^{\sigma_{\mathbf{b}}}(s_{i,0}, h_{i,0}) = \phi^{\sigma_{\mathbf{b}}}(s_{i,1}, h_{i,1}) - b_{i+1} = \text{fl}(\mathbf{b}, i + 1) - b_{i+1}$. Further note that the switch $(s_{i,j}, b_1)$ is applied in the same transitions in which the switch $(s_{i,1-j}, h_{i,1-j})$ is applied. Hence, $\phi^{\sigma_{\mathbf{b}}}(s_{i,j}, h_{i,j}) = \text{fl}(\mathbf{b}, i + 1) - (1 - j) \cdot b_{i+1}$ and $\phi^{\sigma_{\mathbf{b}}}(s_{i,j}, b_1) = \text{fl}(\mathbf{b}, i + 1) - j \cdot b_{i+1}$.

Finally consider some edge $(e_{i,j,k}, g_1)$. This edge is applied as an improving switch whenever the first bit switches from 0 to 1. Since 0 is even, this happens once for every odd number smaller than or equal to \mathbf{b} , i.e., $\left\lceil \frac{\mathbf{b}}{2} \right\rceil$ times. Since the switch $(e_{i,j,k}, b_2)$ is

applied during each transition in which the switch $(e_{i,j,k}, g_1)$ is not applied, we have $\phi(e_{i,j,k}, g_1) = \mathfrak{b} - \left\lceil \frac{\mathfrak{b}}{2} \right\rceil = \left\lfloor \frac{\mathfrak{b}}{2} \right\rfloor$ as $\mathfrak{b} \in \mathbb{N}$.

6.5 Proving the lower bound

We now sketch our proof that applying the Strategy Improvement resp. Policy Iteration Algorithm with Zadeh's pivot rule and our tie-breaking rule introduced in Definition 7 takes an exponential number of iterations. This is shown in an inductive fashion as follows. Assume that we are given a canonical strategy $\sigma_{\mathfrak{b}}$ for some $\mathfrak{b} \in \mathcal{B}_n$ that fulfills a certain set of conditions. Then, applying improving switches according to Zadeh's pivot rule and our tie-breaking rule produces a canonical strategy $\sigma_{\mathfrak{b}+1}$ for $\mathfrak{b} + 1$ that fulfills the same conditions. If we can then show that the initial strategy ι is a canonical strategy for 0 having the desired properties, our bound on the number of iterations follows. Most of these properties are however rather complicated and are only needed for technical reasons. We thus do not discuss them here and refer to Appendix B for a detailed overview. These conditions are called *canonical conditions* and are the following:

1. The occurrence records $\phi^{\sigma_{\mathfrak{b}}}$ are described correctly by Table 2.
2. $\sigma_{\mathfrak{b}}$ has Properties (Or1)_{*,*,*} to (Or4)_{*,*,*} related to the occurrence records.
3. Any improving switch was applied at most once per previous transition.

As a basis for the following proofs and statements, we give a characterization of the set of improving switches for canonical strategies. We furthermore prove that canonical strategies are Phase-1-strategies for the corresponding number. We use the notation $\sigma \rightarrow \sigma'$ to denote the sequence of strategies calculated by the algorithms when starting with σ until they eventually reach σ' . Throughout this section let $\nu := \nu(\mathfrak{b} + 1)$.

Lemma 3 *Let $\sigma_{\mathfrak{b}} \in \Lambda_i$ be a canonical strategy for $\mathfrak{b} \in \mathcal{B}_n$. Then $\sigma_{\mathfrak{b}}$ is a Phase-1-strategy for \mathfrak{b} and $I_{\sigma_{\mathfrak{b}}} = \{(d_{i,j,k}, F_{i,j}) : \sigma_{\mathfrak{b}}(d_{i,j,k}) \neq F_{i,j}\}$.*

As claimed, ι is in fact a canonical strategy and fulfills the canonical conditions.

Lemma 4 *The initial strategy ι is a canonical strategy for 0 fulfilling the canonical conditions.*

We now discuss how the main statement is proven. Consider some canonical strategy $\sigma_{\mathfrak{b}}$ for some $\mathfrak{b} \in \mathcal{B}_n$. We prove that applying improving switches to $\sigma_{\mathfrak{b}}$ using Zadeh's pivot rule and our tie-breaking rule produces a specific Phase- k -strategy for \mathfrak{b} for every $k \in \{1, \dots, 5\}$. These strategies are typically the first strategy of the corresponding phase and have several nice properties that help in proving the statement. The properties of these strategies and the set of improving switches of these strategies are summarized in Tables 6 and 7 in Appendix B. We finally prove that applying improving switches to the Phase-5-strategy that is calculated by the Strategy Improvement resp. Policy Iteration Algorithm then yields a canonical strategy $\sigma_{\mathfrak{b}+1}$ for $\mathfrak{b} + 1$ fulfilling the canonical conditions. Since ι is a canonical strategy for 0 fulfilling the canonical conditions, we can then apply the statement iteratively to prove the main theorem of this paper.

Before formally stating this idea as a theorem, we discuss the individual phases. Consider Phase 1. In this phase, (mainly) cycle edges $(d_{i,j,k}, F_{i,j})$ are applied. The last switch of this phase closes the cycle center $F_{\nu,(\mathfrak{b}+1)_{\nu+1}}$. We will show that after the application of this switch, the induced bit state of the strategy is $\mathfrak{b} + 1$. As a consequence, the resulting strategy is not a Phase-1-strategy. It then depends on the parity of \mathfrak{b} whether this strategy belongs to Phase 2 or Phase 3. If $\nu > 1$, then the next strategy is a Phase 2 strategy. If $\nu = 1$, then the next strategy is a Phase-3-strategy.

Now consider the beginning of Phase 2. The properties of the corresponding strategy we refer to here are given by Tables 6 and 7. Since the induced bit state of the strategy changed from \mathfrak{b} to $\mathfrak{b} + 1$, the entry vertices of all levels $i \leq \nu$ need to be adjusted. Accordingly, the upper selection vertices of all levels $i < \nu$ need to be adjusted. This is reflected by the set of improving switches, containing the edges (b_ν, g_ν) and $(s_{\nu-1,1}, h_{\nu-1,1})$. Moreover, edges $(d_{*,*,*}, F_{*,*})$ that were improving for $\sigma_\mathfrak{b}$ and have not been applied yet remain improving and these edges have relatively high occurrence records. Also, Table 2 describes the occurrence records of the edges that were just applied when interpreted for $\mathfrak{b} + 1$. Note that we explicitly exclude the switches $(g_i, F_{i,j})$ here. The reason is that it is hard to show that the bound on the occurrence records of these edges is valid by only considering a single transition. We will thus show that these switches cannot be applied too often during $\iota \rightarrow \sigma_\mathfrak{b}$ for any $\mathfrak{b} \in \mathcal{B}_n$ after discussing the single phases in detail.

As discussed previously, the targets of the upper selection vertices are not set correctly for $\mathfrak{b} + 1$ if $\nu > 1$. This is partly handled during Phase 2. More precisely, the improving switches $(s_{i,(\mathfrak{b}+1)_{i+1}}, h_{i,(\mathfrak{b}+1)_{i+1}})$ are applied during Phase 2 for all $i < \nu$. Furthermore, the target vertices of all entry vertices b_2 to b_ν are updated during Phase 2. After applying all of these switches, we then have a Phase-3-strategy. Since none of these switches needs to be applied if $\nu = 1$, we reach such a Phase-3-strategy directly after Phase 1 in that case.

The Phase 3 strategy obtained at this point shares several properties with the Phase 2 strategy described earlier. For example, if $(d_{*,*,*}, F_{*,*})$ was improving for $\sigma_\mathfrak{b}$, it is still improving at the beginning of Phase 3 and has a “high” occurrence record. Furthermore, all edges $(e_{*,*,*}, b_2)$ resp. $(e_{*,*,*}, g_1)$ are improving now. The explanation for this is that the spinal path already contains the correct levels with respect to $\mathfrak{b} + 1$ (with the exception of the corresponding edge starting at b_1) and has thus a very good valuation. Note that this requires that the vertices $s_{i,(\mathfrak{b}+1)_{i+1}}$ were updated for $i < \nu$. The vertices b_2 resp. g_1 are thus very profitable, implying that all edges leading directly towards these vertices become improving.

It turns out that switches $(e_{*,*,*}, *)$ then minimize the occurrence record and are applied next. Due to the tie-breaking rule, the algorithms then only apply switches $(e_{i,j,k}, *)$ with $\sigma(d_{i,j,k}) = F_{i,j}$. If $(\mathfrak{b} + 1)_i \neq 1$ or $(\mathfrak{b} + 1)_{i+1} \neq j$, applying a switch $(e_{i,j,k}, *)$ can then make the edge $(d_{i,j,k}, e_{i,j,k})$ improving. However, the occurrence record of these edges is bounded from above by the occurrence record of the just applied switch $(e_{i,j,k}, *)$. Consequently, such a switch is then applied next.

At this point, there is a major difference between S_n and M_n . In S_n , the application of such a switch has no influence on the valuation of the corresponding cycle center $F_{i,j}$. The reason is that the player 1 controlled cycle center can simply choose the other cycle vertex (or, in case that the cycle center was closed, $s_{i,j}$). This is not true

in M_n as the valuation of $F_{i,j}$ is then directly affected if the valuation of the cycle vertices changes. Thus, the valuation of cycle centers increases in M_n . This might unlock improving switches involving upper selection in lower levels which will then be applied directly as they have low occurrence records. The set of switches that will be applied in that fashion exactly corresponds to the set of switches applied during Phase 4 in S_n and can also only occur if $\nu > 1$.

To summarize Phase 3, all switches $(e_{i,j,k}, *)$ with $\sigma(d_{i,j,k}) = F_{i,j}$ are applied. If such an application makes an edge $(d_{i,j,k}, e_{i,j,k})$ improving, then this switch is also applied. After all of these switches are applied, the tie-breaking rule then chooses to apply (b_1, b_2) resp. (b_1, g_1) , concluding Phase 3. The application of this final switch makes many edges $(d_{*,*,*}, F_{*,*})$ improving. The exact set of switches that is applied is rather complicated and depends on $b + 1$ and whether or not $b + 1$ is a power of two. The application of the switch (b_1, b_2) resp. (b_1, g_1) then results in either a Phase-4-strategy or a Phase-5-strategy for b .

If $\nu > 1$, then the algorithms produce a Phase-4-strategy in S_n . During Phase 4, it then applies the improving switches $(s_{i,*}, b_1)$ for $i \leq \nu - 1$. The final switch that will be applied is the switch $(s_{1,*}, b_1)$, resulting in a Phase-5-strategy. In particular, the algorithms always produce a Phase-5-strategy at some point.

We now discuss Phase 5. To satisfy the definition of a canonical strategy, the switches $(e_{i,j,k}, *)$ not applied during Phase 3 need to be applied. However, some switches of the form $(d_{i,j,k}, F_{i,j})$ will be applied first as they have very low occurrence records. Once these switches have been applied, switches of the form $(e_{i,j,k}, *)$ minimize the occurrence record and will be applied.

The next statement describes that we in fact encounter these different phases. Note that we use the term of the “next feasible row” as certain phases may not be present in certain cases. Thus, the term “the next row” may not always be accurate.

Lemma 5 *Let $\sigma_b \in \Lambda_b$ be a canonical strategy for b fulfilling the canonical conditions. Let $\sigma \in \Lambda_{\sigma_b}$ be a strategy obtained by applying a sequence of improving switches to σ_b . Let σ have the properties of row k of Table 6 and let I_σ be described by row k of Table 7 for some $k \in \{1, \dots, 5\}$. Then, applying improving switches according to Zadeh’s pivot rule and our tie-breaking rule produces a strategy σ' that is described by the next feasible rows of Tables 6 and 7.*

The following statement is implied by Lemma 5. It states that applying improving switches to a canonical strategy fulfilling the canonical conditions yields the next canonical strategy which then also fulfills the canonical conditions. It also allows us to prove our main theorem.

Lemma 6 *Let σ_b be a canonical strategy for $b \in \mathcal{B}_n$ fulfilling the canonical conditions. After applying a finite number of improving switches according to Zadeh’s pivot rule and our tie-breaking rule, the Strategy Improvement resp. Policy Iteration Algorithm calculates a strategy σ_{b+1} with the following properties.*

1. $I_{\sigma_{b+1}} = \{(d_{i,j,k}, F_{i,j}) : \sigma_{b+1}(d_{i,j,k}) \neq F_{i,j}\}$.
2. The occurrence records of all edges are described by Table 2 when interpreted for $b + 1$.

3. σ_{b+1} is a canonical strategy for $b + 1$ and has Properties (Or1)_{*,*,*} to (Or4)_{*,*,*}.
4. When transitioning from σ_b to σ_{b+1} , every improving switch was applied at most once.

Theorem 4 *Applying the Strategy Improvement Algorithm with Zadeh's pivot rule and the strategy-based tie-breaking rule described in Definition 7 on the game G_n of size $O(n)$ needs at least 2^n iterations when using ι as the initial player 0 strategy.*

7 Conclusion

In this paper, we have shown that Zadeh's pivot rule has an exponential worst-case running time in PG, MDP, and Simplex context. Together with previous results, we now have a complete picture regarding the worst-case performance of all traditional pivot rules (up to tie-breaking). This means that new pivot rules will have to be introduced and analyzed in order to make further progress towards the question whether an efficient pivot rule that is independent of tie-breaking exists. In particular, addressing the following questions might be the next step:

1. Can the known lower bound constructions for history-based pivot rules be generalized to eliminate a larger class of pivot rules? In particular, can our construction be generalized to arbitrary tie-breaking in Zadeh's rule?
2. Can we devise new natural pivot rules that systematically rule out the general ideas behind the recent lower bound constructions for history-based pivot rules?

Another approach to eliminating pivot rules is the analysis of their inherent complexity. For example, it was shown recently that predicting the behavior of the Simplex Algorithm with Dantzig's original pivot rule is computationally intractable [8, 11]. It would be interesting to investigate whether these results carry over to Zadeh's pivot rule, and, in particular, whether a unified approach emerges that encompasses a broader class of pivot rules with a single construction.

Acknowledgements We would like to thank Felix Solms for his implementation of a sink game solver [37], which helped us to iron out quantitative details of our construction.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

A Implementations and verification of the results

We briefly discuss two independent implementations that were used to verify the technical details of our construction quantitatively. Both implementations are publicly available [19, 37].

The first implementation is based on the PGSolver software developed in OCaml by Oliver Friedmann [19]. This implementation was developed alongside the theory and was used to verify our construction for parity games. It was also used to create the animations of the counter for $n = 3$ and $n = 4$ levels [24]. The PGSolver provides a general framework for solving parity games which was previously used to verify several lower bound constructions, e.g., [2, 12, 13, 15–17]. The only algorithmic part of the implementation that was modified for our construction is the definition of our specific tie-breaking rule. Otherwise, we only implemented some automated checks, and provided an input file that specifies our parity game as a graph with vertex priorities and a partition of the vertices into player 0 and player 1 vertices, together with the initial strategy. We applied the general solver as a black box to run the Strategy Improvement Algorithm with Zadeh's pivot rule and our tie-breaking for this input to check (by hand) every single step of the algorithm for $n \in \{3, 4\}$ levels against our theoretical predictions. In particular, we verified every single application of an improving switch and every occurrence record throughout the execution. We performed similar manual checks of many steps for $n = 5$ and some selection of steps for $n \in \{6, \dots, 9\}$, focusing on potentially critical special cases and canonical strategies. As the construction already needs 466 iterations for $n = 4$ levels [24] and 1228 iterations for $n = 5$, it was not feasible to verify larger examples step-by-step.

In addition, we implemented automated checks of occurrence records, as well as other technical details. Specifically, we checked “well-behavedness” of strategies (for a definition see full version [6]). As the number of iterations becomes very large already for relatively small values of n , we performed such automated checks for examples with up to $n = 10$ levels. For $n = 10$, the strategy improvement algorithm already requires more than 80.000 iterations and it was not feasible to perform additional verifications using this implementation.

The second implementation was developed by Felix Solms in Java as a part of his Master's thesis [37]. This implementation executes the Strategy Improvement Algorithm with Zadeh's pivot rule and our specific tie-breaking for a given sink game. More precisely, it computes the counter-strategy for player 1 for every strategy of player 0, uses this to compute the set of improving switches, and computes the improving switch that Zadeh's pivot rule would select. Based on this, the implementation allows to navigate step-by-step through the execution of the algorithm and provides a visualization of the set of improving switches in each step. Importantly, both the algorithms and the input files (and format) were developed from scratch and altogether independently from the first implementation. Specifics of our construction (beyond tie-breaking) are limited to an input file and thus cleanly separated from the implementation of the Strategy Improvement Algorithm. The implementation was used to manually confirm our predicted behavior for $n \in \{3, 4, 5\}$.

Table 3 Strategy notation and expressions. For convenience, $\bar{\sigma}(b_{n+1}) := 0$

Symbol	Encoded boolean expression
$\bar{\sigma}(b_i)$	$\sigma(b_i) = g_i$
$\bar{\sigma}(s_{i,j})$	$\sigma(s_{i,j}) = h_{i,j}$
$\bar{\sigma}(g_i)$	$\sigma(g_i) = F_{i,1}$
$\bar{\sigma}(d_{i,j,k})$	$\sigma(d_{i,j,k}) = F_{i,j}$
$\bar{\sigma}(e_{i,j,k})$	$\sigma(e_{i,j,k}) = b_2$
$\bar{\sigma}(s_i)$	$\bar{\sigma}(s_i, \bar{\sigma}(g_i))$
$\bar{\sigma}(d_{i,j})$	$\bar{\sigma}(d_{i,j,0}) \wedge \bar{\sigma}(d_{i,j,1})$
$\bar{\sigma}(d_i)$	$\bar{\sigma}(d_i, \bar{\sigma}(g_i))$
$\bar{\sigma}(eg_{i,j})$	$\bigvee_{k=0,1} [\neg \bar{\sigma}(d_{i,j,k}) \wedge \neg \bar{\sigma}(e_{i,j,k})]$
$\bar{\sigma}(eb_{i,j})$	$\bigvee_{k=0,1} [\neg \bar{\sigma}(d_{i,j,k}) \wedge \bar{\sigma}(e_{i,j,k})]$
$\bar{\sigma}(eg_i)$	$\bar{\sigma}(eg_i, \bar{\sigma}(g_i))$
$\bar{\sigma}(eb_i)$	$\bar{\sigma}(eb_i, \bar{\sigma}(g_i))$

B Abbreviations and tables

In this appendix we explain all abbreviations used in the proofs and properties of strategies introduced in the main part. For all of these explanations let $\sigma \in \Lambda_i$ be some strategy. Table 3 contains an overview of several boolean expressions. These expressions are either true (i.e., equal to 1) or false (i.e., equal to 0). They are used to have a compact representation of the state of the counter and to compare and link the configurations of different vertices.

We now define the properties used for defining the term 'Phase- k -strategy'. As discussed in the main part, each property is a boolean expression that might depend on one or two parameters. Further note that the properties also might depend on whether we consider PG or MDP context.

$$\sigma(s_{i,\beta_{i+1}^\sigma}) = h_{i,\beta_{i+1}^\sigma} \wedge \sigma(s_{i,1-\beta_{i+1}^\sigma}) = b_1 \tag{Usv1}$$

$$\sigma(s_{i,j}) = h_{i,j} \tag{Usv2}$$

$$\sigma(s_{i,\beta_{i+1}^\sigma}) \neq h_{i,\beta_{i+1}^\sigma} \wedge \sigma(s_{i,1-\beta_{i+1}^\sigma}) \neq b_1 \tag{Usv3}$$

$$\nexists i: \sigma(b_{i-1}) = g_{i-1} \wedge \bar{\sigma}(b_i) \neq \bar{\sigma}(g_{i-1}) \tag{Rel1}$$

$$\mu^\sigma = v(\beta^\sigma) \tag{Rel2}$$

$$i < \mu^\sigma \Rightarrow \neg \bar{\sigma}(d_{i,0}) \vee \neg \bar{\sigma}(d_{i,1}) \tag{Cc1}$$

$$\bar{\sigma}(d_v) \wedge \bar{\sigma}(g_v) = (b + 1)_{v+1} \tag{Cc2}$$

$$\bar{\sigma}(d_{i,j}) \Rightarrow j \neq \beta_{i+1}^\sigma \tag{Cc3}$$

$$\text{MDP-context: } \bar{\sigma}(g_i) = 1 - \beta_{i+1}^\sigma \Rightarrow \bar{\sigma}(d_{i,1-\beta_{i+1}^\sigma}) \tag{Sv1}$$

$$\text{PG-context: } \bar{\sigma}(g_i) = 1 \Rightarrow \bar{\sigma}(d_{i,1})$$

$$\beta_1^\sigma = 0 \Rightarrow \sigma(e_{*,*,*}) = b_2 \tag{Esc1}$$

$$\beta_1^\sigma = 1 \Rightarrow \sigma(e_{*,*,*}) = g_1$$

$$\sigma(e_{*,*,*}) = g_1 \tag{Esc2}$$

$$\bar{\sigma}(eg_{i,j}) \wedge \neg \bar{\sigma}(eb_{i,j}) \tag{Esc3}$$

$$\bar{\sigma}(eb_{i,j}) \wedge \neg \bar{\sigma}(eg_{i,j}) \tag{Esc4}$$

$$\bar{\sigma}(eg_{i,j}) \wedge \bar{\sigma}(eb_{i,j}) \tag{Esc5}$$

$$\bar{\sigma}(b_i) = \bar{\sigma}(d_{i,\beta_{i+1}^\sigma}^\sigma) \tag{Bac1}$$

$$\bar{\sigma}(b_i) \Rightarrow \bar{\sigma}(g_i) = \beta_{i+1}^\sigma \tag{Bac2}$$

$$\bar{\sigma}(b_i) \Rightarrow \neg \bar{\sigma}(d_{i,1-\beta_{i+1}^\sigma}^\sigma) \tag{Bac3}$$

Before giving the definition of the phases, we introduce additional properties. These properties are not used for the definition of the phases and are related to the occurrence records of cycle edges. These properties in particular yield insights regarding the parameter t_b used in Table 2. The occurrence records of cycle edges are rather complicated, hence these additional properties that help us in proving that Table 2 describes their occurrence records correctly. They are furthermore part of the canonical conditions, a set of properties every canonical strategy has. If the considered binary number b is clear from the context, we let $m := \lfloor (b + 1)/2 \rfloor$ as this number turns out to provide a bound on the maximum occurrence record of the improving switches for σ_b .

$$\sigma(d_{i,j,k}) = F_{i,j} \wedge (b_i = 0 \vee b_{i+1} \neq j) \implies \phi^\sigma(d_{i,j,k}, F_{i,j}) < m \tag{Or1}$$

$$\beta_i^\sigma = 0 \vee \beta_{i+1}^\sigma \neq j \implies [\phi^\sigma(d_{i,j,k}, F_{i,j}) = \ell^b(i, j, k) + 1 \Leftrightarrow \sigma(d_{i,j,k}) = F_{i,j}] \tag{Or2}$$

$$\phi^\sigma(d_{i,j,k}, F_{i,j}) = \ell^b(i, j, k) - 1 \wedge \phi^\sigma(d_{i,j,k}, F_{i,j}) \neq \left\lfloor \frac{b + 1 - k}{2} \right\rfloor \tag{Or3}$$

Table 4 Definition of the phases. The entries show for which set of indices the strategy has the corresponding property resp. whether the strategy has the property at all

Property	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5
(Bac1) _i	$i \in [n]$	$i > \mu^\sigma$	$i > 1$	$i \in [n]$	$i \in [n]$
(Bac2) _i	$i \in [n]$	$i \geq \mu^\sigma$	$i > 1$	$i \in [n]$	$i \in [n]$
(Bac3) _i	$i \in [n], i \neq \nu$	$i > \mu^\sigma$	$i > 1, i \neq \mu^\sigma$	$i \in [n]$	$i \in [n]$
(Usv1) _i	$i \in [n]$	$i \geq \mu^\sigma$	$i \geq \mu^\sigma$	$i \geq \nu$	$i \in [n]$
(Usv2) _{i,j}	-	$(i, 1 - \beta_{i+1}^\sigma) : i < \mu^\sigma$	$(i, *) : i < \mu^\sigma$	$(i, \beta_{i+1}^\sigma) : i < \nu$	-
(Esc1)	True	-	-	-	False*
(Esc2)	-	True	-	-	-
(Esc4) _{i,j}	-	-	-	S ₁	-
(Esc5) _{i,j}	-	-	-	S ₂	-
(Rel1)	True	-	-	True	True
(Rel2)	-	True	True	False	False
(Cc1) _i	$i \in [n]$	$i \in [n]$	$i \in [n]$	$i \in [n]$	$i \in [n]$
(Cc2)	-	-	-	True	True
(Sv1) _i	$i \in [n]$	$i \in [n]$	-	-	-*
$\beta^\sigma =$	\mathfrak{b}	$\mathfrak{b} + 1$	$\mathfrak{b} + 1$	$\mathfrak{b} + 1$	$\mathfrak{b} + 1$
Special	Phase 2: Phase 3: Phase 4: Phase 5:	$\exists i < \mu^\sigma : (\text{Usv3})_i \wedge \neg(\text{Bac2})_i \wedge \neg(\text{Bac3})_i$ A Phase 3 strategy that has Property Cc2 is called <i>proper</i> Phase-3-strategy $\exists i < \nu(\mathfrak{b} + 1) : (\text{Usv2})_{i, 1 - \beta_{i+1}^\sigma}$ *If a strategy has Property Esc1 and there is an index i such that it does not have Property Sv1 _i , it is defined as a Phase-5-strategy			

The last row contains assumptions and further properties used for the definition of the phases. We define $[n] = \{1, \dots, n\}$ and $\nu := \nu(\mathfrak{b} + 1)$

$$\iff \mathfrak{b} \text{ is odd, } \nexists l \in \mathbb{N} : \mathfrak{b} + 1 = 2^l, i = \nu(\mathfrak{b} + 1), j \neq \mathfrak{b}_{i+1}, k = 0.$$

$$\sigma(d_{i,j,k}) \neq F_{i,j} \implies \phi^\sigma(d_{i,j,k}, F_{i,j}) \in \{\mathfrak{m} - 1, \mathfrak{m}\} \tag{Or4}$$

Table 4 is used to define the five phases, see Definition 9. It contains one column listing all properties and one column per phase. A strategy is called a Phase- k -strategy if it has the properties listed in the corresponding column (resp. if it has the property for the respective indices if the property depends on one or two indices). A ‘-’ signifies that it is not specified whether the strategy has the corresponding property. It thus may or may not have it.

Table 5 contains sets of pairs of indices (i, j) that are used for defining the phases or within several later proofs and statements.

The next table gives an overview over the strategies at the beginning of each phase. More precisely, it contains properties that the strategies at the beginning of the different phases have. It also gives an overview over the different combinations of phases, contexts and parameters ν that can and cannot occur. We also introduce the following notation. For strategies σ, σ' with $\sigma \in \Lambda_{\sigma'}$, the sequence of improving switches that the Strategy Improvement Algorithm applies when starting with σ until it reaches σ' is denoted by $\mathfrak{A}_\sigma^{\sigma'}$.

The next table gives an overview over the set of improving switches at the beginning of the different phases. It thus in particular describes the same strategies whose properties are summarized in 6. It also shows which phases can actually occur depending on whether we consider PG or MDP context and depending on the least significant set bit of the next number.

Table 5 Sets used for the definition of the phases

$S_1 =$	$\begin{aligned} & \{(i, 1 - \beta_{i+1}^\sigma) : i \leq \nu\} - 1 \cup \\ & \{(i, 1 - \beta_{i+1}^\sigma) : i \in \{\nu, \dots, m - 1\} \wedge \beta_i^\sigma = 0\} \cup \\ & \begin{cases} \emptyset & \exists k \in \mathbb{N} : \mathfrak{b} + 1 = 2^k \\ \{(\nu, 1 - \beta_{\nu+1}^\sigma)\}, \nexists k \in \mathbb{N} : \mathfrak{b} + 1 = 2^k \end{cases} \end{aligned}$
$S_2 =$	$\begin{aligned} & \{(i, \beta_{i+1}^\sigma) : i \leq \nu(\mathfrak{b} + 1) - 1\} \cup \\ & \{(i, 1 - \beta_{i+1}^\sigma) : i \in \{\nu(\mathfrak{b} + 1) + 1, \dots, m\} \wedge \beta_i^\sigma = 1\} \cup \\ & \{(i, \beta_{i+1}^\sigma) : i \in \{\nu(\mathfrak{b} + 1), \dots, m - 1\} \wedge \beta_i^\sigma = 0\} \cup \\ & \{(i, k) : i > m, k \in \{0, 1\}\} \cup \\ & \begin{cases} \{(\nu(\mathfrak{b} + 1), 1)\} & \exists k \in \mathbb{N} : \mathfrak{b} + 1 = 2^k \\ \emptyset, & \nexists k \in \mathbb{N} : \mathfrak{b} + 1 = 2^k \end{cases} \end{aligned}$
$S_3 =$	$\begin{aligned} & \{(i, 1 - \beta_{i+1}^\sigma) : i \in \{1, \dots, u\}\} \cup \\ & \{(i, 1 - \beta_{i+1}^\sigma) : i \in \{u + 1, \dots, m\} \wedge \beta_i^\sigma = 1\} \cup \\ & \{(i, \beta_{i+1}^\sigma) : i \in \{u + 1, \dots, m - 1\} \wedge \beta_i^\sigma = 0\} \cup \\ & \{(i, k) : i > m, k \in \{0, 1\}\} \cup \{(u, \beta_{u+1}^\sigma)\} \end{aligned}$
$S_4 =$	$\{(i, 1 - \beta_{i+1}^\sigma) : i \in \{u + 1, \dots, m - 1\} \wedge \beta_i^\sigma = 0\}$

We use the abbreviations $m = \max\{i : \beta_i^\sigma = 1\}$ and $u = \min\{i : \beta_i^\sigma = 0\}$ as well as $\nu = \nu(\mathfrak{b} + 1)$

C Proofs for the main statements

This appendix contains a selection of proofs and proof sketches for the statements of the main part of the paper. Throughout this appendix, $n \in \mathbb{N}$ is a fixed natural number, $\mathfrak{b} \in \mathcal{B}_n$ and $\nu := \nu(\mathfrak{b} + 1)$. Also, if not stated otherwise, we assume $i \in \{1, \dots, n\}$, $j, k \in \{0, 1\}$ to be arbitrary but fixed indices. To further simplify notation, we define $\sum(\mathfrak{b}, i) := \sum_{\ell < i} \mathfrak{b}_\ell 2^{\ell-1}$. Also, we add an upper index S respectively M when discussing vertex valuations in order to distinguish whether the argument holds for the case $G_n = S_n$ respectively $G_n = M_n$. If an argument is applicable in both cases, this is marked by an upper index “*”.

We prove that the Strategy Improvement resp. Policy Iteration Algorithm behave as intended as follows. We begin by stating lemmas that are needed for several proofs. Then, there is one part for every single phase. In each part, we provide lemmas and, eventually, tables formalizing the application of single improving switch during the corresponding phases. We then use these lemmas and tables to prove that applying improving switches to a canonical strategy $\sigma_{\mathfrak{b}}$ for \mathfrak{b} yields a Phase- k -strategy for the respective $k \in \{1, \dots, 5\}$ as described by Tables 6 and 7. We finally prove that we obtain a canonical strategy $\sigma_{\mathfrak{b}+1}$ for $\mathfrak{b} + 1$ that fulfills the canonical conditions.

The statements provided here are proven in the full version [6] and marked with (★). To simplify notation, we also refine the notation for describing the states of cycle centers.

Definition 10 Let σ be a strategy and $t^\rightarrow \in \{g_1, b_2\}$. Then, a cycle center $F_{i,j}$ is t^\rightarrow -open if $\sigma(d_{i,j,*}) = e_{i,j,*}$ and $\sigma(e_{i,j,*}) = t^\rightarrow$. It is t^\rightarrow -halfopen if $\sigma(d_{i,j,k}) = e_{i,j,k}$, $\sigma(e_{i,j,k}) = t^\rightarrow$ and $\sigma(d_{i,j,1-k}) = F_{i,j}$ for some $k \in \{0, 1\}$. Finally, it is mixed if $\bar{\sigma}(eb_{i,j}) \wedge \bar{\sigma}(eg_{i,j})$. The corresponding term is called *state* of $F_{i,j}$.

The proofs require several additional lemmas. They usually either describe specific strategies and the application of specific improving switches or terms used for the description of the occurrence records. We do not prove these lemmas here as their

Table 6 Properties that specific Phase- k -strategies have

Phase	$\nu = 1$	$\nu > 1$	
		PG	MDP
1	Canonical strategy for b fulfilling the canonical conditions		
2	-	$\sigma(d_{i,j,k}) \neq F_{i,j} \Rightarrow \phi^\sigma(d_{i,j,k}, F_{i,j}) = m$ $(g_i, F_{i,j}) \in \mathfrak{A}_{\sigma_b}^\sigma \Rightarrow [b_i = 0 \wedge b_{i+1} \neq j] \vee i = \nu$ and $F_{i,j}$ is closed $\mathfrak{A}_{\sigma_b}^\sigma \subseteq \mathbb{D}^1 \cup \mathbb{G}$ $\sigma(g_\nu) = F_{\nu, \beta_{\nu+1}^\sigma}$ and $\sigma(g_i) = F_{i, 1-\beta_{i+1}^\sigma}$ for all $i < \nu$ $i < \nu \Rightarrow \bar{\sigma}(d_i)$ and Property $Usv3_i$	
3	$\sigma(d_{i,j,k}) \neq F_{i,j} \Rightarrow \phi^\sigma(d_{i,j,k}, F_{i,j}) = \phi^{\sigma^v}(d_{i,j,k}, F_{i,j}) = m$ $\sigma(s_{i,*}) = h_{i,*}, \sigma(g_i) = F_{i, 1-\beta_{i+1}^\sigma}$ and $\bar{\sigma}(d_i)$ for all $i < \nu$ as well as $\sigma(g_\nu) = F_{\nu, \beta_{\nu+1}^\sigma}$ $\mathfrak{A}_{\sigma_b}^{\sigma^3} \subseteq \mathbb{D}^1 \cup \mathbb{G} \cup \mathbb{S} \cup \mathbb{B}$ and $(g_i, F_{i,j}) \in \mathfrak{A}_{\sigma_b}^\sigma \Rightarrow [b_i = 0 \wedge b_{i+1} \neq j] \vee i = \nu$ and $F_{i,j}$ is closed		
4	-	$\mu^\sigma = \min\{i: \beta_i^\sigma = 0\}$ $\sigma(d_{i,j,k}) = F_{i,j} \Leftrightarrow \beta_i^\sigma = 1 \wedge (b+1)_{i+1} = j$ $(i, j) \in S_1 \Rightarrow \bar{\sigma}(eb_{i,j}) \wedge \neg \bar{\sigma}(eg_{i,j})$ $(i, j) \in S_2 \Rightarrow \bar{\sigma}(eb_{i,j}) \wedge \bar{\sigma}(eg_{i,j})$ $(g_i, F_{i,j}) \in \mathfrak{A}_{\sigma_b}^\sigma \Rightarrow [b_i = 0 \wedge b_{i+1} \neq j] \vee i = \nu$ $(g_i, F_{i,j}) \in \mathfrak{A}_{\sigma_b}^\sigma \Rightarrow F_{i,j}$ is closed $\sigma(e_{i,j,k}) = b_2 \Rightarrow \phi^\sigma(d_{i,j,k}, F_{i,j}) = \phi^{\sigma^v}(d_{i,j,k}, F_{i,j})$ $\sigma(e_{i,j,k}) = b_2 \Rightarrow \phi^\sigma(d_{i,j,k}, F_{i,j}) = m$	
1	$\mu^\sigma = \min\{i: \beta_{i+1}^\sigma = 0\}$ $\sigma(d_{i,j,k}) = F_{i,j} \Leftrightarrow \beta_i^\sigma = 1 \wedge (b+1)_{i+1} = j$ $(g_i, F_{i,j}) \in \mathfrak{A}_{\sigma_b}^\sigma \Rightarrow [b_i = 0 \wedge b_{i+1} \neq j] \vee i = \nu$ $\sigma(e_{i,j,k}) = t^\rightarrow \Rightarrow \phi^\sigma(d_{i,j,k}, F_{i,j}) = \phi^{\sigma^v}(d_{i,j,k}, F_{i,j}) = m$ $(i, j) \in S_4 \Rightarrow \bar{\sigma}(eg_{i,j}) \wedge \neg \bar{\sigma}(eb_{i,j})$ $(i, j) \in S_1 \Rightarrow \bar{\sigma}(eb_{i,j}) \wedge \neg \bar{\sigma}(eg_{i,j})$ $(i, j) \in S_3 \Rightarrow \bar{\sigma}(eg_{i,j}) \wedge \bar{\sigma}(eb_{i,j})$ $(i, j) \in S_2 \Rightarrow \bar{\sigma}(eb_{i,j}) \wedge \bar{\sigma}(eg_{i,j})$ $i < \nu \Rightarrow \bar{\sigma}(g_i) = 1 - \beta_{i+1}^\sigma$		
1	Canonical strategy for $b+1$ fulfilling the canonical conditions		

To simplify notation, we define $t^\rightarrow := g_1$ if $\nu = 1$ and $t^\rightarrow := b_2$ if $\nu > 1$. A '-' signifies that the corresponding combination of phase, context and ν does not occur during the execution of the algorithm. We do not interpret the header as the first row of this table

Table 7 Improving switches at the beginning of different phases when starting with a canonical strategy σ_b for b

Phase	$\nu = 1$	$\nu > 1$	
		PG	MDP
1	$\mathfrak{D}^\sigma := \{(d_{i,j,k}, F_{i,j}): \sigma(d_{i,j,k}) \neq F_{i,j}\}$		
2	-	$\mathfrak{D}^\sigma \cup \{(b_\nu, g_\nu), (s_{\nu-1,1}, h_{\nu-1,1})\}$	
3	$\mathfrak{D}^\sigma \cup \{(b_1, g_1)\} \cup \{(e_{*,*,*}, g_1)\}$	$\mathfrak{D}^\sigma \cup \{(b_1, b_2)\} \cup \{(e_{*,*,*}, b_2)\}$	
4	-	$\mathfrak{E}^\sigma \cup \{(s_{\nu-1,0}, b_1)\} \cup \{(s_{i,1}, b_1): i \leq \nu - 2\} \cup X_0 \cup X_1$ -	
5	$\mathfrak{E}^\sigma \cup \bigcup_{\substack{i=\mu^{\sigma^v}+1 \\ \beta_i^\sigma=0}}^{m-1} \{(d_{i,1-\beta_{i+1}^\sigma}, s_{i,*}, F_{i,1-\beta_{i+1}^\sigma})\}$	$\mathfrak{E}^\sigma \cup \bigcup_{i=1}^{\nu-1} \{d_{i,1-\beta_{i+1}^\sigma}, s_{i,*}, F_{i,1-\beta_{i+1}^\sigma}\} \cup X_0 \cup X_1$	
1	\mathfrak{D}^σ		
$X_k := \begin{cases} \emptyset, & b+1 \text{ is a power of two} \\ \{(d_{\nu,1-\beta_\nu^\sigma}, F_{\nu,1-\beta_\nu^\sigma})\} \cup \bigcup_{\substack{i=\nu+1 \\ \beta_i=0}}^{m-1} \{(d_{i,1-\beta_{i+1}^\sigma}, F_{i,1-\beta_{i+1}^\sigma})\}, & \text{otherwise} \end{cases}$			

We use $\nu := \nu(b+1)$ and $m := \max\{i: \sigma(b_i) = g_i\}$ to simplify the notation. We further define $\mathfrak{E}^\sigma := \{(d_{i,j,k}, F_{i,j}), (e_{i,j,k}, b_2): \sigma(e_{i,j,k}) = g_1\}$ if $\nu > 1$ and, analogously, $\mathfrak{E}^\sigma := \{(d_{i,j,k}, F_{i,j}), (e_{i,j,k}, g_1): \sigma(e_{i,j,k}) = b_2\}$ if $\nu = 1$. Note that we do not interpret 1 as a power of two

proofs are rather technical and involved and state them prior to the statements using them.³

Vertex valuations and technical details

Proving our main results requires a lot of technical work. For every intermediate step, it is necessary to investigate the strategies, improving switches and occurrence records. As we are fully aware of the length and technical complexity of our proof, we have decided to omit the investigation of the vertex valuations which can be found in [6, Appendix A]. Furthermore, we decided to omit straight-forward but tedious technical details in the proofs presented here. We refer to the full version [6] for all omitted details.

Basic statements and statements independent of different phases

We begin by showing that the pair (G_n, ι) defines a Sink Game if $G_n = S_n$ resp. that it has the weak unichain condition if $G_n = M_n$.

Lemma 1 *For all $n \in \mathbb{N}$, the game G_n and the initial player 0 strategy ι define a Sink Game with sink t in PG context, resp. have the weak unichain condition in MDP context.*

Proof If $G_n = S_n$, it suffices to prove that the game is completely won by player 1 and that $\pi_{\iota, \tau^t, v}$ ends in t for all $v \in V$ (see e.g. [14]). Similarly, if $G_n = M_n$, it suffices to prove that the single irreducible recurrent class of the initial and the optimal policy is t . It is easy to verify the conditions related to the initial policy and to verify that the following strategy is optimal and that it has the desired properties: □

v	$d_{i,j,k}$	$e_{i,j,k}$	$g_i, i < n$	g_n	$s_{i,0}, i < n$	$s_{n,0}$	$s_{i,1}$	b_i
$\sigma^*(v)$	$F_{i,j}$	g_1	$F_{i,1}$	$F_{n,0}$	b_1	$h_{n,0}$	$h_{i,1}$	g_i

We next show that the tie-breaking rule is computationally tractable.

Lemma 2 *Given a strategy $\sigma \in \Lambda_l$ and an occurrence record $\phi^\sigma : E^0 \rightarrow \mathbb{N}_0$, the tie-breaking rule can be evaluated in polynomial time.*

Proof Let $\sigma \in \Lambda_l$. Identifying the subsets of E_0 can be done by iterating over E_0 and checking $\sigma(v)$ for all $v \in E_0$. Therefore, the pre-order of the sets can be calculated in polynomial time. Since expanding the chosen pre-order to a total order is possible in polynomial time [38], we can compute the tie-breaking rule in polynomial time. Whenever the tie-breaking rule needs to be considered, the algorithm needs to iterate over the chosen ordering. Since this can also be done in time polynomial in the input, the tie-breaking rule can be applied in polynomial time. Also, handling the exception described in Definition 7 can be done in polynomial time. □

³ The (complete) proof can be found in the full version [6].

The first main lemma states that canonical strategies are Phase-1-strategies and also give a characterization of their set of improving switches.

Lemma 3 *Let $\sigma_b \in \Lambda_l$ be a canonical strategy for $b \in \mathcal{B}_n$. Then σ_b is a Phase-1-strategy for b and $I_{\sigma_b} = \{(d_{i,j,k}, F_{i,j}) : \sigma_b(d_{i,j,k}) \neq F_{i,j}\}$.*

Proof By the definition of a canonical strategy, σ_b has all of the properties defining a Phase-1-strategy. It thus suffices to show $I_{\sigma_b} = \mathcal{D}^\sigma$. We thus have to prove that $\sigma(d_{i,j,k}) \neq F_{i,j}$ implies $\mathcal{E}_\sigma^*(F_{i,j}) > \mathcal{E}_\sigma^*(e_{i,j,k})$ and that there are no other improving switches. This can be checked by simple, but tedious calculations (see full version [6, p.54 ff.] for details). □

We next prove that the initial strategy ι fulfills the canonical conditions that we impose on canonical strategies at the beginning of Sect. 6.5.

Lemma 4 *The initial strategy ι is a canonical strategy for 0 fulfilling the canonical conditions.*

Proof As no improving switch was applied yet and it is obvious that ι is a canonical strategy for 0, it suffices to prove that ι has Properties (Or1)_{*,*,*} to (Or4)_{*,*,*}. Let $i \in \{1, \dots, n\}, j, k \in \{0, 1\}$. First, ι has Property Or1 _{i,j,k} as $\iota(d_{i,j,k}) = e_{i,j,k}$. In addition, $0 = \phi^l(d_{i,j,k}, F_{i,j}) < 1 \leq \ell^b(i, j, k) + 1$, so ι has Property Or2 _{i,j,k} . Moreover, $\phi^l(d_{i,j,k}, F_{i,j}) = 0 = \lfloor \frac{1-k}{2} \rfloor = \lfloor \frac{b+1-k}{2} \rfloor$, hence the premise of Property Or3 _{i,j,k} is always incorrect. Thus, ι has Property Or3 _{i,j,k} . Since it is immediate that ι has Property Or4 _{i,j,k} , the statement follows. □

Before discussing the application of the improving switches during the individual phases, we provide several technical lemmas that are used to determine the exact occurrence records of the cycle vertices. We remind here that $m = \lfloor (b + 1)/2 \rfloor$.

Lemma 5 *Let σ_b be a canonical strategy for b such that its occurrence records are described by Table 2. Assume that σ_b has Properties (Or1)_{*,*,*} to (Or4)_{*,*,*}. Then, the following hold.*

1. Let $i \in \{1, \dots, n\}, j \in \{0, 1\}$ such that $b_i = 0$ or $b_{i+1} \neq j$. Then $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) \leq m$.
2. Let $j := b_{v+1}$. It holds that $\phi^{\sigma_b}(d_{v,j,0}, F_{v,j}) = m$. In addition, $\phi^{\sigma_b}(d_{v,j,1}, F_{v,j}) = m$ if $v = 1$ and $\phi^\sigma(d_{v,j,1}, F_{v,j}) = m - 1$ if $v > 1$.
3. If $i = 1$, then $\sigma_b(d_{i,1-b_{i+1},*}) \neq F_{i,1-b_{i+1}}$ and $\phi^{\sigma_b}(d_{i,1-b_{i+1},0}, F_{i,1-b_{i+1}}) = m$.

In the following, we use $\mathbf{1}_{x=y}$ as an abbreviation for the function that is equal to one if $x = y$ and equal to zero otherwise.

Lemma 6 *Let $b \in \mathcal{B}_n$ and $i \in \{1, \dots, n\}, j \in \{0, 1\}$ such that $b_i = 0$ or $b_{i+1} \neq j$. Then,*

$$\begin{aligned} & \mathbf{1}_{j=0} \text{mfn}(b, i + 1) - \mathbf{1}_{j=1} \text{mfn}(b, i + 1) \\ &= \mathbf{1}_{j=0} \text{mfn}(b + 1, i + 1) - \mathbf{1}_{j=1} \text{mfn}(b + 1, i + 1). \end{aligned}$$

Moreover, if $i \neq v$, then $\ell^b(i, j, k) + 1 = \ell^{b+1}(i, j, k)$.

Lemma 7 *If $\mathbf{1}_{j=0}\text{mfn}(b, i + 1) + \mathbf{1}_{j=1}\text{mufn}(b, i + 1) = 0$, then $\ell^b(i, j, k) \geq b$. Otherwise, the following hold:*

Setting of bits	$b_i = 1 \wedge b_{i+1} = 1 - j$	$b_i = 0 \wedge b_{i+1} = j$	$b_i = 0 \wedge b_{i+1} = 1 - j$
$\ell^b(i, j, k)$	$= \left\lceil \frac{b + \sum(b, i) + 1 - k}{2} \right\rceil$	$= \left\lceil \frac{b + 2^{i-1} + \sum(b, i) + 1 - k}{2} \right\rceil$	$= \left\lceil \frac{b - 2^{i-1} + \sum(b, i) + 1 - k}{2} \right\rceil$

Reaching a phase-2-strategy

We begin by providing lemmas and proofs describing the application of improving switches during Phase 1 for $\nu > 1$ as these are the circumstances under which there is a Phase 2. In this phase, cycle edges $(d_{*,*,*}, F_{*,*})$ and edges $(g_{*,*}, F_{*,*})$ are applied.

We first provide an overview describing the application of individual improving switches during Phase 1.³ We interpret each row of this table stating that if a strategy σ fulfills the given conditions, applying the given switch e results in a strategy σe that has the claimed properties. For convenience, conditions specifying the improving switch, resp. the level or cycle center corresponding to the switch, are contained in the second column. Note that we also include one improving switch that technically belongs to Phase 2. This is included as Table 8 then contains all statements necessary to prove that applying improving switches to σ_b yields the Phase-2-strategy that is described in Tables 6 and 7. Henceforth, we let $\mathfrak{D}^\sigma := \{(d_{i,j,k}, F_{i,j}) : \sigma(d_{i,j,k}) \neq F_{i,j}\}$ to simplify notation.

The following lemma now describes the application of cycle edges in Phase 1 in more detail. It in principle summarizes the first three rows of Table 8.

Table 8 Improving switches applied during Phase 1

Conditions for σ	Switch e	Properties of σe
$F_{i,j}$ is open and $I_\sigma = \mathfrak{D}^\sigma$ $G_n = S_n$ and $I_\sigma = \widehat{\mathfrak{D}}^\sigma$	$(d_{i,j,k}, F_{i,j})$ $(d_{i,1-b_{i+1},k}, F_{i,1-b_{i+1}})$ $i \neq 1$	Phase-1-strategy for b and $I_{\sigma e} = \widehat{\mathfrak{D}}^{\sigma e}$ Phase-1-strategy for b $I_{\sigma e} = \mathfrak{D}^{\sigma e}$
$I_\sigma = \widehat{\mathfrak{D}}^\sigma$ and $\sigma(g_i) = F_{i,b_{i+1}}$ $\sigma(d_{i,1-b_{i+1},1-k}) = F_{i,1-b_{i+1}}$	$(d_{i,1-b_{i+1},k}, F_{i,1-b_{i+1}})$ $b_i = 0$	Phase-1-strategy for b $I_{\sigma e} = \mathfrak{D}^{\sigma e} \cup \{(g_i, F_{i,1-b_{i+1}})\}$
$I_\sigma = \widehat{\mathfrak{D}}^\sigma \cup \{(g_i, F_{i,1-b_{i+1}})\}$ $F_{i,j}$ is closed	$(g_i, F_{i,1-b_{i+1}})$ $i \neq 1 \wedge b_i = 0$	Phase-1-strategy for b $I_{\sigma e} = I_\sigma \setminus \{e\} = \mathfrak{D}^{\sigma e}$
$I_\sigma = \mathfrak{D}^\sigma$ $\sigma(d_{\nu,b_{\nu+1},1-k}) = F_{\nu,b_{\nu+1}}$	$(d_{\nu,b_{\nu+1},k}, F_{\nu,b_{\nu+1}})$	$\nu = 1 \Rightarrow$ Phase-3-strategy for b $\nu = 1 \wedge \sigma(g_\nu) = F_{\nu,b_{\nu+1}}$ imply $I_{\sigma e} = \mathfrak{D}^{\sigma e} \cup \{(b_1, g_1)\} \cup \{(e_{*,*,*}, g_1)\}$ $\nu > 1 \Rightarrow$ Phase-2-strategy for b $\nu > 1 \wedge \sigma(g_\nu) = F_{\nu,b_{\nu+1}}$ imply $I_{\sigma e} = \mathfrak{D}^{\sigma e} \cup \{(b_\nu, g_\nu), (s_{\nu-1,1}, h_{\nu-1,1})\}$ $\sigma(g_\nu) \neq F_{\nu,b_{\nu+1}}$ implies $I_{\sigma e} = \mathfrak{D}^{\sigma e} \cup \{(g_\nu, F_{\nu,b_{\nu+1}})\}$
Phase-2-strategy for b and $\nu > 1$ $I_\sigma = \mathfrak{D}^\sigma \cup \{(g_\nu, F_{\nu,b_{\nu+1}})\}$	$(g_\nu, F_{\nu,b_{\nu+1}})$	Phase-2-strategy for b $I_{\sigma e} = \mathfrak{D}^{\sigma e} \cup \{(b_\nu, g_\nu), (s_{\nu-1,1}, h_{\nu-1,1})\}$

For convenience, we always assume $\sigma \in A_l$ and that σ is a Phase-1-strategy for b if not stated otherwise. We thus also always have $\sigma e \in A_l$

Lemma 8 *Let $\sigma \in \Lambda_{\sigma_b}$ be a Phase-1-strategy for b with $I_\sigma = \mathfrak{D}^\sigma$. Let $e := (d_{i,j,k}, F_{i,j}) \in I_\sigma, I_{\sigma_b}$ with $\phi^\sigma(e) = \phi^{\sigma_b}(e) = m - 1$. Assume that σ_b fulfills the canonical conditions. Then $\sigma_e \in \Lambda_t$ is a Phase-1-strategy for b . Furthermore, if $\sigma(d_{i,j,1-k}) = F_{i,j}, j = 1 - b_{i+1}\sigma(g_i) = F_{i,1-j}$ and $b_i = 0$, then $I_{\sigma_e} = (I_\sigma \setminus \{e\}) \cup \{(g_i, F_{i,j})\}$. Otherwise, $I_{\sigma_e} = I_\sigma \setminus \{e\}$. In addition, the occurrence record of e with respect to σ_e is described by Table 2 when interpreted for $b + 1$.*

We begin by considering the case $\nu > 1$. We prove that the Strategy Improvement resp. Policy Iteration Algorithm produces a Phase-2-strategy $\sigma^{(2)}$ as described by the corresponding rows of Tables 6 and 7. We furthermore prove that Table 2 (interpreted for $b + 1$) characterizes the occurrence record of every improving switch that is applied is when transitioning from σ_b to $\sigma^{(2)}$. As analogous statements will be shown for every phase, this later enables us to simplify the proof of Theorem 3 significantly.

Lemma 7 *Let σ_b be a canonical strategy for $b \in \mathcal{B}_n$ with $\nu > 1$ fulfilling the canonical conditions. After applying a finite number of improving switches, the Strategy Improvement resp. Policy Iteration Algorithm produces a Phase-2-strategy $\sigma^{(2)} \in \Lambda_t$ as described by the corresponding rows of Tables 7 and 6.*

Proof Let $j := b_{\nu+1}$ and observe that $b_{\nu+1} = (b + 1)_{\nu+1}$. Since σ_b is a canonical strategy, we have $\sigma_b(d_{\nu,j,*}) \neq F_{\nu,j}$. Moreover, $I_{\sigma_b} = \mathfrak{D}^{\sigma_b}$ and σ_b is a Phase-1-strategy for b by Lemma 3, implying $(d_{\nu,j,*}, F_{\nu,j}) \in I_{\sigma_b}$. By Lemma 5, $(d_{\nu,j,0}, F_{\nu,j})$ maximizes the occurrence record among all improving switches and $\phi^{\sigma_b}(d_{\nu,j,1}, F_{\nu,j}) = m - 1$. By Property Or4_{*,*,*}, I_{σ_b} can be partitioned into $I_{\sigma_b} = I_{\sigma_b}^{<m} \cup I_{\sigma_b}^m$ where $e \in I_{\sigma_b}^{<m}$ if $\phi^{\sigma_b}(e) = m - 1$ and $e \in I_{\sigma_b}^m$ if $\phi^{\sigma_b}(e) = m$. If $I_{\sigma_b}^{<m} \neq \emptyset$, then a switch contained in this set is applied first due to Zadeh’s pivot rule. By applying Lemma 8 iteratively, we can apply improving switches $e \in I_{\sigma_b}^{<m}$ until we either reach a strategy σ with $I_\sigma^{<m} = \emptyset$ or until an edge $(g_i, F_{i,j})$ becomes improving. By Lemma 8, Table 2 (interpreted for $b + 1$) describes the occurrence record of all switches applied in the process. Using Table 2 and by investigating under which condition an edge $(g_i, F_{i,j})$ can become improving, it is easy to verify that such a switch is applied directly after becoming improving and that Lemma 8 can be applied afterwards.³ Note that we do *not* prove yet that Table 2 describes its occurrence record, this is done at the end of the appendix as it requires insights about the other phases.

We observe that $(g_i, F_{i,j})$ can only be applied if $F_{i,j}$ was closed by applying $(d_{i,j,k}, F_{i,j})$. This implies that either $(d_{i,j,0}, F_{i,j}), (d_{i,j,1}, F_{i,j}) \in I_{\sigma_b}^{<m}$ or $\sigma(d_{i,j,1-k}) = \sigma_b(d_{i,j,1-k}) = F_{i,j}$. The first case can only happen for $i = \nu$ and $j = 1 - b_{\nu+1}$. Thus, if a switch $(g_i, F_{i,j})$ is applied, then either $i = \nu$ or $\sigma_b(d_{i,j,1-k}) = F_{i,j}$ and $(d_{i,j,k}, F_{i,j}) \in I_{\sigma_b}^{<m}$.

Thus let σ be a Phase-1-strategy σ with $I_\sigma^{<m} = \emptyset$ and assume that $\mathbb{G} \cap I_\sigma = \emptyset$. Further note that $\mathfrak{A}_{\sigma_b}^\sigma \subseteq \mathbb{D}^1 \cup \mathbb{G}$ and that $(g_i, F_{i,j}) \in \mathfrak{A}_{\sigma_b}^\sigma$ implies $b_i = 0 \wedge b_{i+1} \neq j$ and that the previous arguments hold independent of ν . In particular, we implicitly proved the following corollary.

Corollary 1 *Let σ_b be a canonical strategy and $\sigma_b(d_{i,j,k}) \neq F_{i,j}$. If $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) < m$, then $(d_{i,j,k}, F_{i,j})$ is applied during Phase 1.*

We now discuss the application of switches contained in $I_{\sigma_b}^m$, so consider a Phase-1-strategy σ with $I_\sigma = \{e = (d_{i,j,k}, F_{i,j}) : \phi^\sigma(e) = m\} = \mathfrak{D}^\sigma$. It is easy to verify

that there are no open cycle centers with respect to σ by proving $\sigma(d_{*,*,1}) = F_{*,*}$.³ We prove that $e := (d_{v,b_{v+1},0}, F_{v,b_{v+1}})$ is applied next. By Lemma 5, $\phi^{\sigma_b}(e) = m$. In addition, $\sigma_b(d_{v,b_{v+1},0}) \neq F_{v,b_{v+1}}$ as σ_b is a canonical strategy for b . As only edges with an occurrence record less than m were applied so far, this implies $e \in I_\sigma$. Since all improving switches have the same occurrence records, it is sufficient to show that no other improving switch is ranked lower by the tie-breaking rule. As proven before, there are no open cycle centers. Hence, the ordering of the edges is based on the bit of the levels, the index of the levels and whether the cycle center is active. To be precise, the first switch according to the tie-breaking rule is the improving switch contained in the active cycle center of the lowest level with a bit equal to 0. This edge is precisely $e = (d_{v,b_{v+1},0}, F_{v,b_{v+1}})$. It is easy to verify that the occurrence record of e is described by Table 2 when interpreted for $b + 1$ after the application.³

By row 5 of Table 8, σe is a Phase-2-strategy for b . If $\sigma(g_v) \neq F_{v,b_{v+1}}$, then $(g_v, F_{v,b_{v+1}})$ minimizes the occurrence record among all improving switches. Due to the tie-breaking rule, this switch is then applied next, and this application is formalized in row 6 of Table 8.

Let $\sigma^{(2)}$ denote the strategy obtained after applying $(g_v, F_{v,b_{v+1}})$ if $\sigma(g_v) \neq F_{v,b_{v+1}}$ resp. after applying $(d_{v,b_{v+1},0}, F_{v,b_{v+1}})$ if $\sigma(g_v) = F_{v,b_{v+1}}$. Then, $I_{\sigma^{(2)}} = \mathfrak{D}^{\sigma^{(2)}} \cup \{(b_v, g_v)\} \cup \{(s_{v-1,1}, h_{v-1,1})\}$ by row 5 resp. 6 of Table 8. Furthermore, $\sigma^{(2)}$ has Property **Usv3_i** for all $i < v$ since σ_b has Property **Usv1_i** for all $i < v$ and $b_i = 1 - (b + 1)_i$ for $i \leq v$. In addition, $\sigma^{(2)}(d_{i,j,k}) \neq F_{i,j}$ implies $\phi^{\sigma^{(2)}}(d_{i,j,k}, F_{i,j}) = m$ by Corollary 1. Moreover, since we did not apply any improving switch $(d_{*,*,*}, e_{*,*,*})$ and $b_i = 1 - \beta_{i+1}^\sigma$ for all $i < v$, we have $\sigma(g_i) = 1 - \beta_{i+1}^\sigma$ as well as $\bar{\sigma}(d_{i,1-\beta_{i+1}^\sigma})$ for all $i < v$. \square

We furthermore implicitly proved the following two corollaries. We later show that the condition $v > 1$ can be dropped in the first corollary.

Corollary 2 *Let σ_b be a canonical strategy for b and let $v > 1$. Then, during Phase 1, the switch $(g_i, F_{i,j})$ is applied if and only if $F_{i,j}$ is closed during Phase 1, $\sigma_b(g_i) \neq F_{i,j}$ and $i \neq v$. A cycle center can only be closed during Phase 1 if either $i = v$ or if $\sigma_b(d_{i,j,k}) = F_{i,j}$, $\phi^{\sigma_b}(d_{i,j,1-k}, F_{i,j}) < m$ for some $k \in \{0, 1\}$, $j \neq b_{i+1}$ and $\sigma_b(b_i) = b_{i+1}$.*

Corollary 3 *Let $\sigma^{(2)}$ be the Phase-2-strategy calculated by the Strategy Improvement resp. Policy Iteration Algorithm when starting with a canonical strategy σ_b fulfilling the canonical conditions as per Lemma 7. Then, Table 2 specifies the occurrence record of every improving switch applied so far when interpreted for $b + 1$, excluding switches $(g_*, F_{*,*})$, and each switch was applied once.*

Reaching a phase-3-strategy

Depending on v , Phase 3 is either reached after applying the improving switches of Phase 2 or directly after Phase 1. During Phase 2, the entry vertices b_i and some of the selection vertices $s_{i,*}$ of levels $i \leq v$ are updated. We again provide an overview describing the application of individual improving switches during Phase 2.³ We also

include one row that technically does not belong to Phase 2, handling a special case that might occur at the beginning of Phase 3.

We furthermore need the following lemma that is used to determine the values of several expressions related to binary counting.

Lemma 9 *It holds that $\text{fl}(\mathbf{b}, i) = \lfloor \frac{\mathbf{b}+2^{i-1}}{2^i} \rfloor$ and $\text{fl}(\mathbf{b} + 1, i) = \text{fl}(\mathbf{b}, i) + \mathbf{1}_{i=v}$. In addition, for indices $i_1, i_2 \in \{1, \dots, n\}$ with $i_1 < i_2$ and $\mathbf{b} \geq 2^{i_1-1}$ imply $\text{fl}(\mathbf{b}, i_1) > \text{fl}(\mathbf{b}, i_2)$. Furthermore, if $k := \frac{\mathbf{b}+1}{2^{v-1}}$ and $x \in \{1, \dots, v-1\}$, then $\text{fl}(\mathbf{b}, v-x) = k \cdot 2^{x-1}$.*

The following lemma now states that we reach a Phase-3-strategy for \mathbf{b} under all circumstances. If $v = 1$, then this follows by analyzing Phase 1 in a similar fashion as done when proving Lemma 7. It in fact turns out that nearly the identical arguments can be applied. If $v > 1$, then we use that lemma to argue that we obtain a Phase-2-strategy. We then investigate Phase 2 in detail and prove that we also obtain a Phase-3-strategy.

Lemma 8 *Let $\sigma_{\mathbf{b}} \in \Lambda_{\iota}$ be a canonical strategy for $\mathbf{b} \in \mathcal{B}_n$ fulfilling the canonical conditions. After applying a finite number of improving switches, the Strategy Improvement resp. Policy Iteration Algorithm produces a proper Phase-3-strategy $\sigma^{(3)} \in \Lambda_{\iota}$ as described by the corresponding rows of Tables 6 and 7.*

Proof Consider the case $v = 1$ first. As shown in the proof of Lemma 7, we can partition $I_{\sigma_{\mathbf{b}}}$ into $I_{\sigma_{\mathbf{b}}}^{<m}$ and $I_{\sigma_{\mathbf{b}}}^m$. Since Lemma 8 also applies for $v = 1$, the same arguments imply that the algorithm calculates a Phase-1-strategy $\sigma \in \Lambda_{\iota}$ with $I_{\sigma} = \{e = (d_{i,j,k}, F_{i,j}) : \phi^{\sigma}(e) = \mathbf{m}\} = \mathcal{D}^{\sigma}$. We can again deduce $\mathfrak{A}_{\sigma_{\mathbf{b}}}^{\sigma} \subseteq \mathbb{D}^1 \cup \mathbb{G}$ and that $(g_i, F_{i,j}) \in \mathfrak{A}_{\sigma_{\mathbf{b}}}^{\sigma}$ implies $\mathbf{b}_i = 0 \wedge \mathbf{b}_{i+1} \neq j$ or $i = v$. We can further assume that there are no indices i, j with $(g_i, F_{i,j}) \in I_{\sigma}$. Also, by Lemma 8, the occurrence records of edges $(d_{i,j,k}, F_{i,j}) \in \mathfrak{A}_{\sigma_{\mathbf{b}}}^{\sigma}$ is described by Table 2 when interpreted for $\mathbf{b} + 1$.

Since all improving switches have the same occurrence records, their order of application depends on the tie-breaking rule. Due to the first criterion, improving switches contained in open cycle centers are applied first. Hence, a sequence of strategies is produced until a strategy is reached such that there are no more open cycle centers. Note that all produced strategies are Phase-1-strategies for \mathbf{b} , reachable from ι by row 1 of Table 8. Also, by the tie-breaking rule, $(d_{i,j,0}, F_{i,j})$ is applied as improving switch in an open cycle center. By the same arguments used when proving Lemma 7, the second switch of $F_{v,\mathbf{b}_{v+1}}$ is applied next and, possibly, $(g_v, F_{v,\mathbf{b}_{v+1}})$ is applied afterwards. Let σ denote the obtained strategy. We prove that the occurrence records of edges $(d_{i,j,k}, F_{i,j})$ applied so far is specified by Table 2 when interpreted for $\mathbf{b} + 1$.

As argued previously, each such switch is contained in a cycle center open for $\sigma_{\mathbf{b}}$. Consider such a cycle center. If the occurrence record of one of its cycle edges is $m - 1$, then the application of the corresponding switch is described by Lemma 8 and we do not need to consider it here. Also, due to the tie-breaking rule, we do not apply an improving switch contained in halfopen cycle centers (with the exception of $F_{v,\mathbf{b}_{v+1}}$) as we only consider switches contained in $I_{\sigma_{\mathbf{b}}}^{\max}$. We may thus assume that $F_{i,j}$ is open with respect to $\sigma_{\mathbf{b}}$ and that both cycle edges have an occurrence record of m . We only consider the case $i \neq v$ or $i = v \wedge j \neq \mathbf{b}_{i+1}$ here.³ By the tie-breaking rule, $e = (d_{i,j,0}, F_{i,j})$ is then applied. Let σ denote the strategy in which e is applied.

Since b is even, $\phi^{\sigma e}(e) = m + 1 = \lfloor \frac{b+2}{2} \rfloor$. It thus suffices to show that there is a parameter t_{b+1} feasible for $b + 1$ such that $\lfloor \frac{b}{2} \rfloor + 1 \leq \ell^{b+1}(i, j, k) + t_{b+1}$. By the choice of i and j , Lemma 6 implies $\ell^b(i, j, k) + 1 = \ell^{b+1}(i, j, k)$. Therefore, $\phi^\sigma(e) + 1 \leq \ell^b(i, j, k) + t_b + 1 \leq \ell^{b+1}(i, j, k) + t_b$ for some t_b feasible for b . Since b is even, Property Or4 $_{i,j,0}$ implies $\phi^{\sigma b}(e) \neq \ell^b(i, j, k) - 1$. In addition, by Property Or2 $_{i,j,0}$, $t_b \neq 1$ as this would imply $\sigma_b(d_{i,j,0}) = F_{i,j}$, contradicting our assumption. Consequently, $t_b = 0$ and we can thus choose $t_{b+1} = 0$ as parameter.

Let $\sigma^{(3)}$ denote the strategy obtained after closing $F_{\nu, b_{\nu+1}}$ resp. after applying $(g_\nu, F_{\nu, b_{\nu+1}})$ if it becomes improving. Note that the last row of Table 9 applies to $(g_\nu, F_{\nu, b_{\nu+1}})$. Then, by row 5 of Table 8 resp. the last row of Table 9 and our previous arguments, $\sigma^{(3)}$ has all properties listed in the respective rows of Tables 6 and 7. Furthermore, as we used the same arguments, Corollary 2 is also valid for $\nu = 1$ and we can drop the assumption $\nu > 1$.

Consider the case $\nu > 1$, implying $b \geq 1$. By Lemma 7, applying improving switches to σ_b yields a Phase-2-strategy σ for b with $\sigma \in \Lambda_t$ and $I_\sigma = \mathfrak{D}^\sigma \cup \{(b_\nu, g_\nu), (s_{\nu-1,1}, h_{\nu-1,1})\}$. By Table 2, it holds that $\phi^\sigma(b_\nu, g_\nu) = \text{fl}(b, \nu) = \phi^\sigma(s_{\nu-1,1}, h_{\nu-1,1})$. By Lemma 9, $\text{fl}(b, \nu) = \lfloor \frac{b+2\nu-1}{2} \rfloor$. Since $\nu > 1$ and $b \geq 1$, this implies $\text{fl}(b, \nu) \leq \lfloor \frac{b+2}{4} \rfloor = \lfloor \frac{b+1}{4} \rfloor \leq \lfloor \frac{b+1}{2} \rfloor$. By Lemma 7, any $(d_{i,j,k}, F_{i,j}) \in I_\sigma$ has an occurrence record of m . Thus, by the tie-breaking rule, (b_ν, g_ν) is applied next. Let $\sigma' := \sigma[(b_\nu, g_\nu)]$. It is easy to verify that σ has the properties of row 1 of Table 9. Consequently, σ' is a Phase-2-strategy for b with $\sigma' \in \Lambda_t$. By Lemma 9, $\phi^{\sigma'}(b_\nu, g_\nu) = \text{fl}(b, \nu) + 1 = \text{fl}(b + 1, \nu)$, so Table 2 describes the occurrence record of (b_ν, g_ν) with respect to $b + 1$. The set of improving switches for σ' now depends on ν , see row 1 of Table 9.

- Let $\nu = 2$. Then $I_{\sigma'} = \mathfrak{D}^{\sigma'} \cup \{(b_1, b_2), (s_{1,1}, h_{1,1})\} \cup \{(e_{*,*,*}, b_2)\}$. By Table 2, we have $\phi^{\sigma'}(e_{*,*,*}, b_2) = \lfloor \frac{b}{2} \rfloor$, $\phi^{\sigma'}(b_1, b_2) = \text{fl}(b, 1) - 1$ and $\phi^{\sigma'}(s_{1,1}, h_{1,1}) = \text{fl}(b, 2)$. Since b is odd, Lemma 9 implies $\text{fl}(b, 1) - 1 = \lfloor \frac{b}{2} \rfloor$. Consequently, $\phi^{\sigma'}(b_1, b_2) = \phi^{\sigma'}(e_{i,j,k}, b_2)$. There are two cases. If $b = 1$, then $\text{fl}(b, 2) =$

Table 9 Improving switches applied during Phase 2

Properties of σ	Switch e	Properties of σe
$I_\sigma = \mathfrak{D}^\sigma \cup \{(b_\nu, g_\nu), (s_{\nu-1,1}, h_{\nu-1,1})\}$ Properties Rel1, Cc2 Property Usv3, for all $i < \nu$	(b_ν, g_ν)	$\nu \neq 2$ implies $I_{\sigma e} = (I_\sigma \setminus \{e\}) \cup \{(b_{\nu-1}, b_\nu), (s_{\nu-2,0}, h_{\nu-2,0})\}$ $\nu = 2$ implies $I_{\sigma e} = \mathfrak{D}^{\sigma e} \cup \{(b_1, b_2), (s_{1,1}, h_{1,1})\} \cup \{(e_{*,*,*}, b_2)\}$
$i' < \mu^\sigma \Rightarrow F_{i', \bar{\sigma}(g_{i'})}$ is closed Property Usv3 $_{i'}$ for all $i' \leq i$ Properties (Bac1) $_{\mu^\sigma}$, Bac1 $_{i+1}$, Cc2	$(s_{i,j}, h_{i,j})$ $i < \mu^\sigma$ $j = \beta^\sigma_{i+1}$	$i \neq 1 \Rightarrow$ Phase-2-strategy for b $i = 1 \Rightarrow$ proper Phase-3-strategy for b $I_{\sigma e} = I_\sigma \setminus \{e\}$
$i' < \mu^\sigma \Rightarrow F_{i', \bar{\sigma}(g_{i'})}$ is closed Property Usv3 $_{i'}$ for all $i' \leq i$ $i' > i \Rightarrow$ Property Bac1 $_{i'}$ \wedge (Bac2) $_{i'}$ $i' > i, i' \neq \mu^\sigma \Rightarrow$ Property Bac3 $_{i'}$	(b_i, b_{i+1}) $i > 1$ $i < \mu^\sigma$	Phase-2-strategy for b $i \neq 2 \Rightarrow I_{\sigma e} = (I_\sigma \setminus \{e\}) \cup \{(b_{i-1}, b_i), (s_{i-2,0}, h_{i-2,0})\}$ $i = 2 \Rightarrow I_{\sigma e} = (I_\sigma \setminus \{e\}) \cup \{(b_1, b_2)\} \cup \{(e_{*,*,*}, b_2)\}$
Phase-3-strategy for b and $\nu = 1$ $I_\sigma = \mathfrak{D}^\sigma \cup \{(g_\nu, F_{\nu, b_{\nu+1}})\}$	$(g_\nu, F_{\nu, b_{\nu+1}})$	Proper Phase-3-strategy for b $I_{\sigma e} = \mathfrak{D}^{\sigma e} \cup \{(b_1, g_1)\} \cup \{(e_{*,*,*}, g_1)\}$

For convenience, we always assume $\sigma \in \Lambda_t$, that σ is a Phase-2-strategy for b and that $\nu > 1$ if not stated otherwise. We thus also always have $\sigma e \in \Lambda_t$.

$\lfloor \frac{b+2}{4} \rfloor = 0 = \lfloor \frac{b}{2} \rfloor$ and $\phi^{\sigma'}(s_{1,1}, h_{1,1}) = \phi^{\sigma'}(b_1, b_2)$. But then, $(s_{1,1}, h_{1,1})$ is applied next as this is the situation in which the exception is applied. If $b > 1$, then $v = 2$ implies $b \geq 5$. But this implies $\phi^{\sigma'}(s_{1,1}, h_{1,1}) < \phi^{\sigma'}(b_1, b_2)$, so $(s_{1,1}, h_{1,1})$ is applied next. Consequently, $e := (s_{1,1}, h_{1,1})$ is applied next in any case. Using Lemma 7 and our previous arguments, it is easy to verify that the requirements of row 2 of Table 9 are fulfilled.³ Hence, $\sigma := \sigma'[e]$ is a proper Phase-3-strategy for b with $\sigma \in \Lambda_i$ and $I_\sigma = I_{\sigma'} \setminus \{e\} = \mathcal{D}^\sigma \cup \{(b_1, b_2)\} \cup \{(e_{*,*,*}, b_2)\}$. Since $\phi^{\sigma'[e]}(e) = \text{fl}(b, 2) + 1 = \lfloor \frac{b+2}{4} \rfloor + 1 = \lfloor \frac{(b+1)+2}{4} \rfloor = \text{fl}(b+1, 2)$ since b is odd, Table 2 describes the occurrence record of $(s_{1,1}, h_{1,1})$ with respect to $b+1$. Since we did not apply any improving switch $(g_*, F_{*,*})$ or $(d_{*,*,*}, e_{*,*,*})$, the conditions on cycle centers in levels below v hold for $\sigma^{(3)}$ as they held for $\sigma^{(2)}$. Therefore, σ is a strategy as described by the respective rows of Table 6 and 7.

- Let $v > 2$, implying $b \neq 1$. Then, $I_{\sigma'} = \mathcal{D}^{\sigma'} \cup \{(b_{v-1}, b_v), (s_{v-1,1}, h_{v-1,1}), (s_{v-2,0}, h_{v-2,0})\}$ by the first row of Table 9. By Table 2, we then have $\phi^{\sigma'}(b_{v-1}, b_v) = \text{fl}(b, v-1) - 1$ and $\phi^{\sigma'}(s_{v-1,1}, h_{v-1,1}) = \text{fl}(b, v)$. In addition, $\phi^{\sigma'}(s_{v-2,0}, h_{v-2,0}) = \text{fl}(b, v-1) - 1$. Hence, both edges (b_{v-1}, b_v) and $(s_{v-2,0}, h_{v-2,0})$ minimize the occurrence record. By the tie-breaking rule, the switch $e := (b_{v-1}, b_v)$ is now applied. It is now easy to verify that the application of e can be described by row 3 of Table 9.³ Hence, by our previous arguments and row 3 of Table 9, $\sigma := \sigma'[e]$ is a Phase-2-strategy for b that has Property Cc2 as well as Property Bac1_{*i*} and (Bac2)_{*i*} for all $i \geq v-1$ and Property Bac3_{*i*} for all $i > v-1, i \neq v$. In addition, $\bar{\sigma}(d_i)$ for all $i < v$ and σ has Property Usv3_{*i*} for all $i < v-1$. Furthermore, Lemma 9 implies $\phi^\sigma(e) = \text{fl}(b, v-1) - 1 + 1 = \text{fl}(b, v-1) = \text{fl}(b+1, v-1) - (b+1)_{v-1}$, so Table 2 describes the occurrence record of e with respect to $b+1$. By row 3 of Table 9, $v-1 > 2$ implies $I_\sigma = \mathcal{D}^\sigma \cup \{(s_{v-1,1}, h_{v-1,1}), (s_{v-2,0}, h_{v-2,0}), (b_{v-2}, b_{v-1}), (s_{v-3,0}, h_{v-3,0})\}$. Similarly, $v-1 = 2$ implies $I_\sigma = \mathcal{D}^\sigma \cup \{(e_{i,j,k}, b_2)\} \cup \{(b_1, b_2), (s_{2,1}, h_{2,1}), (s_{1,0}, h_{1,0})\}$. We prove that $e := (s_{v-1,1}, h_{v-1,1}) \in I_\sigma$ is applied next. By the definition of $v, b = k \cdot 2^{v-1} - 1$ for some $k \in \mathbb{N}$. By Table 2, Lemma 9 and using $v > 2$, we obtain the following:

$$\begin{aligned} \phi^\sigma(s_{v-1,1}, h_{v-1,1}) &= \text{fl}(b, v) = \left\lfloor \frac{k \cdot 2^{v-1} - 1 + 2^{v-1}}{2^v} \right\rfloor = \left\lfloor \frac{k}{2} \right\rfloor \\ \phi^\sigma(s_{v-2,0}, h_{v-2,0}) &= \text{fl}(b, v-1) - 1 = k \cdot 2^0 - 1 = k - 1 \\ \phi^\sigma(s_{v-3,0}, h_{v-3,0}) &= \phi^\sigma(b_{v-2}, b_{v-1}) \\ &= \text{fl}(b, v-2) - 1 = k \cdot 2^{2-1} - 1 = 2k - 1 \\ \phi^\sigma(e_{i,j,k}, b_2) &= \left\lfloor \frac{b}{2} \right\rfloor = \left\lfloor \frac{k \cdot 2^{v-1} - 1}{2} \right\rfloor = \left\lfloor k \cdot 2^{v-2} - \frac{1}{2} \right\rfloor = 2^{v-2}k - 1. \end{aligned}$$

If $k > 2$, then $(s_{v-1,1}, h_{v-1,1})$ is the uniquely minimized the occurrence records. If $k \leq 2$, then the occurrence records of $(s_{v-1,1}, h_{v-1,1})$ and $(s_{v-2,0}, h_{v-2,0})$ are identical and lower than the occurrence record of any other improving switch. Since the tie-breaking rule applies improving switches at selection vertices contained in higher levels first, $(s_{v-1,1}, h_{v-1,1})$ is also applied first then. Consequently,

$e := (s_{v-1,1}, h_{v-1,1})$ is applied next in any case.

We prove that σ fulfills the conditions of row 2 of Table 9. By our previous arguments, it suffices to prove that σ has Property **Usv** 3_{v-1} . As $\beta_v^\sigma = 1$, this however follows since $(s_{v-1,1}, h_{v-1,1}) \in I_\sigma$ and since σ has Property **Usv** $2_{i,0}$ by the definition of a Phase-2-strategy. By our previous arguments and row 2 of Table 9, σe then has Properties **(Usv2)** $_{v-1,1}$, **(Cc2)**, **(Bac1)** $_v$ and Property **Usv** $3_{i,1-\beta_{i+1}^\sigma}$ for all $i < v - 1$. Furthermore, $I_{\sigma e} = I_\sigma \setminus \{e\}$. More precisely, $I_{\sigma[e]} = \mathfrak{D}^{\sigma e} \cup \{(s_{v-2,0}, h_{v-2,0}), (b_{v-2}, b_{v-1}), (s_{v-3,0}, h_{v-3,0})\}$ if $v - 1 > 2$ and $v > 2$ implies $I_{\sigma e} = \mathfrak{D}^{\sigma e} \cup \{(e_{*,*,*}, b_2)\} \cup \{(b_1, b_2), (s_{1,0}, h_{1,0})\}$. Also note that $v > 2$ implies $\phi^{\sigma e}(s_{v-1,1}, h_{v-1,1}) = \text{fl}(\mathfrak{b}, v) + 1 = \text{fl}(\mathfrak{b} + 1, v)$ by Lemma 9, so Table 2 specifies its occurrence record with respect to $\mathfrak{b} + 1$.

Consider the case $v - 1 > 2$. We argue that applying improving switches according to Zadeh’s pivot rule and our tie-breaking rule then results in a sequence of strategies such that we finally obtain a strategy σ' with $I_{\sigma'} = \mathfrak{D}^{\sigma'} \cup \{(e_{*,*,*}, b_2)\} \cup \{(b_1, b_2), (s_{1,0}, h_{1,0})\}$. For any $x \in \{2, \dots, v - 2\}$, Lemma 9 implies

$$\begin{aligned} \phi^{\sigma e}(s_{v-x,0}, h_{v-x,0}) &< \phi^{\sigma e}(b_{v-x}, b_{v-(x+1)}) \\ &= \phi^{\sigma e}(s_{v-(x-1),0}, h_{v-(x-1),0}) < \phi^{\sigma e}(e_{i,j,k}, b_2). \end{aligned} \tag{1}$$

Thus, $(s_{v-2,0}, h_{v-2,0})$ is applied next. It is easy to verify that σe meets the requirements of row 2 of Table 9, so it can be used to describe the application of $(s_{v-2,0}, h_{v-2,0})$.

Let σ' denote the strategy obtained. Then $I_{\sigma'} = \mathfrak{D}^{\sigma'} \cup \{(b_{v-2}, b_{v-1}), (s_{v-3,0}, h_{v-3,0})\}$. Also, the occurrence record of the edge $(s_{v-2,0}, h_{v-2,0})$ is described by Table 2 with respect to $\mathfrak{b} + 1$. By Eq. (1) and the tie-breaking rule, (b_{v-2}, b_{v-1}) is applied next. Similar to the previous cases, it is easy to check that row 3 of Table 9 applies to this switch. We thus obtain a strategy σ such $v - 2 \neq 2$ implies $I_\sigma = \mathfrak{D}^\sigma \cup \{(s_{v-3,0}, h_{v-3,0}), (b_{v-3}, b_{v-2}), (s_{v-4,0}, h_{v-4,0})\}$ and $v - 2 = 2$ implies $I_\sigma = \mathfrak{D}^\sigma \cup \{(e_{*,*,*}, b_2)\} \cup \{(b_1, b_2), (s_{1,0}, h_{1,0})\}$. In either case, the occurrence record of (b_{v-2}, b_{v-1}) is described by Table 2 with respect to $\mathfrak{b} + 1$.

In the first case, we can now apply the same arguments again iteratively as Eq. (1) remains valid for σ' and $x \in \{2, \dots, v - 3\}$. After applying a finite number of improving switches we thus obtain a Phase-2-strategy $\sigma \in \Lambda_l$ with $I_\sigma = \mathfrak{D}^\sigma \cup \{(e_{*,*,*}, b_2)\} \cup \{(b_1, b_2), (s_{1,0}, h_{1,0})\}$. Furthermore, σ has Property **Bac1** $_i$, Property **Bac2** $_i$ and Property **Usv** $2_{i,\beta_{i+1}^\sigma}$ for all $i > 1$ as well as Property **Bac3** $_i$ for all $i > 1, i \neq \mu^\sigma$ and Property **Cc2**. In addition, it holds that $\bar{\sigma}(g_i) = 1 - \beta_{i+1}^\sigma$ and $\bar{\sigma}(d_{i,1-\beta_{i+1}^\sigma})$ for all $i < v$ and the occurrence records of all edges applied so far (with the exception of switches $(g_*, F_{*,*})$) is described by Table 2 with respect to $\mathfrak{b} + 1$. Note that this also holds if $v - 1 = 2$.

Consequently, σ meets the requirements of row 2 of Table 9. As $v > 2$, we have $\beta_2^\sigma = 0$. By Table 2, it holds that $\phi^\sigma(s_{1,0}, h_{1,0}) = \text{fl}(\mathfrak{b}, 2) - 1 < \text{fl}(\mathfrak{b}, 1) - 1 = \phi^\sigma(b_1, b_2)$ as well as $\text{fl}(\mathfrak{b}, 2) - 1 = \lfloor \frac{\mathfrak{b}+2}{4} \rfloor - 1 < \lfloor \frac{\mathfrak{b}}{2} \rfloor = \phi^\sigma(e_{i,j,k}, b_2)$. Hence, the switch $e = (s_{1,0}, h_{1,0})$ is applied next and by row 2 of Table 9, $\sigma^{(3)} := \sigma e$ is a proper Phase-3-strategy for \mathfrak{b} with $I_{\sigma^{(3)}} = \mathfrak{D}^{\sigma^{(3)}} \cup \{(e_{i,j,k}, b_2)\} \cup \{(b_1, b_2)\}$.

We thus always obtain a strategy as described by the corresponding rows of Tables 6 and 7. \square

We henceforth use $\sigma^{(3)}$ to refer to the Phase-3-strategy described by Lemma 8. Note that we implicitly proved the following corollaries where the second follows by Corollary 3. Further note that we can drop the condition $\nu > 1$ in Corollary 2.

Corollary 4 *No cycle center is open with respect to $\sigma^{(3)}$.*

Corollary 5 *Let $\sigma^{(3)}$ be the Phase-3-strategy calculated by the Strategy Improvement resp. Policy Iteration Algorithm when starting with a canonical strategy σ_b fulfilling the canonical conditions as per Lemma 8. Then, Table 2 specifies the occurrence record of every improving switch applied so far when interpreted for $b + 1$, excluding switches $(g_*, F_{*,*})$, and each such switch was applied once.*

Reaching either phase 4 or phase 5

We now discuss the application of improving switches during Phase 3, which highly depends on whether we have $G_n = S_n$ or $G_n = M_n$ and on the least significant set bit of $b + 1$. In principle, the escape vertices are reset with respect to the new least significant bit and the targets of several cycle vertices are updated accordingly. As usual, we provide an overview describing the application of individual improving switches during Phase 3.³ To simplify and unify the arguments, we define $t^\rightarrow := b_2$ if $\nu > 1$ and $t^\rightarrow := g_1$ if $\nu = 1$. Similarly, let $t^\leftarrow := g_1$ if $\nu > 1$ and $t^\leftarrow := b_2$ if $\nu = 1$. We furthermore define $\mathcal{E}^\sigma := \{(d_{i,j,k}, F_{i,j}), (e_{i,j,k}, t^\rightarrow) : \sigma(e_{i,j,k}) = t^\leftarrow\}$.

In addition, we make use of the following lemma.

Lemma 10 *Let $G_n = M_n$. Let σ be a proper Phase-3-strategy for b with $\sigma \in \Lambda_i$. Let $j := 1 - \beta_{i+1}^\sigma$. Assume $e := (d_{i,j,k}, e_{i,j,k}) \in I_\sigma$ and $\sigma(e_{i,j,k}) = t^\rightarrow$ for some $k \in \{0, 1\}$. Assume that there are no other indices x, y, z with $(d_{x,y,z}, e_{x,y,z}) \in I_\sigma$, that $F_{i,j}$ is closed and that σ fulfills the following assumptions:*

1. *If $\beta_i^\sigma = 0$, then $\sigma(g_i) = F_{i,j}$ and $F_{i,1-j}$ is t^\leftarrow -halfopen.*
2. *$i < \mu^\sigma$ implies $[\sigma(s_{i,j}) = h_{i,j}$ and $\sigma(s_{i',j'}) = h_{i',j'} \wedge \bar{\sigma}(d_{i'})$ for all $i' < i, j' \in \{0, 1\}]$ and that the cycle center $F_{i',1-\bar{\sigma}(g_{i'})}$ is t^\leftarrow -halfopen for all $i' < i$. In addition, $i < \mu^\sigma - 1$ implies $\bar{\sigma}(eb_{i+1})$.*
3. *$i' > i$ implies $\sigma(s_{i,1-\beta_{i'+1}^\sigma}) = b_1$.*
4. *$i' > i$ and $\beta_{i'}^\sigma = 0$ imply that either $[\bar{\sigma}(g_{i'}) = \beta_{i'+1}^\sigma$ and $F_{i,0}, F_{i,1}$ are mixed] or $[\bar{\sigma}(g_{i'}) = 1 - \beta_{i'+1}^\sigma, F_{i',1-\beta_{i'+1}^\sigma}$ is t^\rightarrow -open and $F_{i',\beta_{i'+1}^\sigma}$ is mixed] and*
5. *$i' > i$ and $\beta_{i'}^\sigma = 1$ imply that $F_{i',1-\beta_{i'+1}^\sigma}$ is either mixed or t^\rightarrow -open.*

Then σe is a proper Phase-3-strategy for b with $\sigma e \in \Lambda_i$ and $I_{\sigma e} = [I_\sigma \cup \{(s_{i,j}, b_1)\}] \setminus \{e\}$ if $i < \mu^\sigma$ and $I_{\sigma e} = I_\sigma \setminus \{e\}$ otherwise.

The next lemma describes the application of switches of the type $(d_{i,j,k}, e_{i,j,k})$ for several cases.

Lemma 11 *Let $\sigma \in \Lambda_{\sigma^{(3)}}$ be a proper Phase-3-strategy for b obtained through the application of a sequence $\mathcal{A}_{\sigma^{(3)}} \subseteq \mathbb{E}^1 \cup \mathbb{D}^0$ of improving switches. Assume that the*

Table 10 Improving switches applied during Phase 3

Properties of σ	Switch e	Properties of σe
If $G_n = S_n$: Property $\text{Usv}2_{i',*} \forall i' < \mu^\sigma$ If $G_n = M_n$: $\sigma(s_{i',*}) = b_1 \Rightarrow \bar{\sigma}(eb_{i',*}) \wedge \neg \bar{\sigma}(eg_{i',*}) \forall i' < \mu^\sigma$	$(e_{i,j,k}, t^\rightarrow)$	proper Phase-3-strategy for \mathfrak{b} $\sigma(d_{i,j,1-k}) = e_{i,j,1-k}$ $\forall [\sigma(d_{i,j,1-k}) = F_{i,j} \wedge j \neq \beta_{i+1}^\sigma]$ imply $I_{\sigma e} = (I_\sigma \setminus \{e\}) \cup \{(d_{i,j,k}, e_{i,j,k})\}$ $\sigma(d_{i,j,1-k}) = F_{i,j} \wedge j = \beta_{i+1}^\sigma$ imply $I_{\sigma e} = I_\sigma \setminus \{e\}$
$G_n = S_n, \nu > 1$ and $I_\sigma = \bar{e}^\sigma \cup \{(b_1, b_2)\}$ $\sigma(d_{i,j,k}) = F_{i,j} \Leftrightarrow \beta_i^\sigma = 1 \wedge \beta_{i+1}^\sigma = j$ (Esc4) $_{i,j}$ resp. (Esc5) $_{i,j}$ for all $(i, j) \in S_2$ resp. $(i, j) \in S_2$ $i < \mu^\sigma \Rightarrow \sigma(s_{i,*}) = h_{i,*}$ $G_n = M_n, \nu > 1$ and $I_\sigma = \bar{e}^\sigma$	(b_1, b_2)	Phase-4-strategy for \mathfrak{b} with $\mu^{\sigma e} = 1$ $I_{\sigma e} = (I_\sigma \setminus \{e\}) \cup \{(s_{\nu-1,0}, b_1)\}$ $\cup \{(s_{i,1}, b_1) : i \leq \nu - 2\} \cup X_0 \cup X_1$
$\sigma(d_{i,j,k}) = F_{i,j} \Leftrightarrow \beta_i^\sigma = 1 \wedge \beta_{i+1}^\sigma = j$ (Esc4) $_{i,j}$ resp. (Esc5) $_{i,j}$ for all $(i, j) \in S_1$ resp. $(i, j) \in S_2$ Property $\text{Usv}1_i$ for all $i \in \{1, \dots, n\}$	(b_1, b_2)	Phase-5-strategy for \mathfrak{b} with $\mu^{\sigma e} = 1$ $I_{\sigma e} = (I_\sigma \setminus \{e\}) \cup X_0 \cup X_1$ $\cup \{(d_{i,-1, \beta_{i+1}^\sigma}, F_{i,1-\beta_{i+1}^\sigma}) : i < \nu\}$
$\nu = 1$ and $I_\sigma = \bar{e}^\sigma \cup \{(b_1, g_1)\}$ $\sigma(d_{i,j,k}) = F_{i,j} \Leftrightarrow \beta_i^\sigma = 1 \wedge \beta_{i+1}^\sigma = j$ (Esc3) $_{i,j}$ resp. (Esc5) $_{i,j}$ for all $(i, j) \in S_4$ resp. $(i, j) \in S_3$	(b_1, g_1)	Phase-5-strategy with $\mu^{\sigma e} = u$ $I_{\sigma e} = (I_\sigma \setminus \{e\})$ $\cup \bigcup_{i'=u+1}^{m-1} \{(d_{i,1-\beta_{i+1}^\sigma}, F_{i,1-\beta_{i+1}^\sigma})\}$ $\beta_i^\sigma = 0$
$G_n = M_n$ $F_{i,j}$ is t^\leftarrow -halfopen, $F_{i,1-j}$ is t^\rightarrow -open, $\sigma(g_i) = F_{i,1-j}$ $\sigma(e_{i,j,k}) = t^\rightarrow$ $G_n = M_n$ and $\sigma(e_{i,j,k}) = t^\rightarrow$	$(d_{i,j,k}, e_{i,j,k})$ $\beta_i^\sigma = 0$ $j = \beta_{i+1}^\sigma$	Proper Phase-3-strategy for \mathfrak{b} $I_{\sigma e} = I_\sigma \setminus \{e\}$
$F_{i,j}$ is t^\rightarrow -halfopen $\beta_i^\sigma = 0 \Rightarrow \sigma(g_i) = F_{i,j} \wedge F_{i,1-j}$ is t^\leftarrow -halfopen $G_n = M_n$	$(d_{i,j,k}, e_{i,j,k})$ $j = 1 - \beta_{i+1}^\sigma$	Proper Phase-3-strategy for \mathfrak{b} $I_{\sigma e} = (I_\sigma \setminus \{e\})$
$\nu > 1$ $\bar{\sigma}(eb_{i,j}) \wedge \neg \bar{\sigma}(eg_{i,j})$	$(s_{i,j}, b_1)$ $i < \nu$ $j = 1 - \beta_{i+1}^\sigma$	Proper Phase-3-strategy for \mathfrak{b} $I_{\sigma e} = (I_\sigma \setminus \{e\})$

For convenience, we always assume $\sigma \in \Lambda_l$ and that σ is a proper Phase-3-strategy for \mathfrak{b} if not stated otherwise, implying that always $\sigma e \in \Lambda_l$. The definition of the sets X_k, S_i can be found in Tables 7, 6

conditions of row 1 of Table 10 are fulfilled for each intermediate strategy σ' . Let $e = (d_{i,j,k}, e_{i,j,k}) \in I_\sigma$ be the switch applied next and assume $\sigma(e_{i,j,k}) = t^\rightarrow, \beta_i^\sigma = 0 \vee \beta_{i+1}^\sigma \neq j$ and $I_\sigma \cap \mathbb{D}^0 = \{e\}$. Further assume that either $i \geq \nu$ or that $G_n = S_n$. Then σe is a proper Phase-3-strategy for \mathfrak{b} with $I_{\sigma e} = (I_\sigma \setminus \{e\})$.

The next lemma now summarizes the application of improving switches during Phase 3. Depending on whether we consider PG or M_n and depending on ν , we then either obtain a Phase-4-strategy or a Phase-5-strategy for \mathfrak{b} .

Lemma 9 Let $\sigma_b \in \Lambda_l$ be a canonical strategy for $\mathfrak{b} \in \mathcal{B}_n$ fulfilling the canonical conditions. After applying a finite number of improving switches, the Strategy Improvement resp. Policy Iteration Algorithm produces a strategy σ with the following properties: If $\nu > 1$, then σ is a Phase- k -strategy for \mathfrak{b} , where $k = 4$ if $G_n = S_n$ and $k = 5$ if $G_n = M_n$. If $\nu = 1$, then σ is a Phase-5-strategy for \mathfrak{b} . In any case, $\sigma \in \Lambda_l$ and σ is described by the corresponding rows of Tables 6 and 7.

Proof By Lemma 8, applying improving switches according to Zadeh’s pivot rule and our tie-breaking rule yields a proper Phase-3-strategy $\sigma^{(3)} \in \Lambda_l$ described by the corresponding row of Table 6 and 7. We begin by describing Phase 3 informally.

For every cycle vertex, either $\sigma^{(3)}(d_{i,j,k}) = F_{i,j}$ or $\sigma^{(3)}(d_{i,j,k}) = e_{i,j,k}$ and $(d_{i,j,k}, F_{i,j}) \in I_{\sigma^{(3)}}$. It will turn out that only switches corresponding to cycle vertices of the first type are applied during Phase 3. For each such cycle vertex $d_{i,j,k}$, the switch $(e_{i,j,k}, t^\rightarrow)$ will be applied. If $(\mathfrak{b} + 1)_i = 0$ or $(\mathfrak{b} + 1)_{i+1} \neq j$, then $(d_{i,j,k}, e_{i,j,k})$ becomes improving and is applied next. This then goes on until all such improving

switches have been applied. During this procedure, it might happen that an edge $(s_{i',*}, b_1)$ with $i' < v$ becomes improving after applying some switch $(d_{i,j,k}, e_{i,j,k})$ if we consider $G_n = M_n$ and if we have $v > 1$. In this case, the corresponding switch is applied immediately. Finally, (b_1, b_2) resp. (b_1, g_1) is applied, resulting in a Phase-4-strategy in S_n if $v > 1$ and in a Phase-5-strategy otherwise.

We now formalize this behavior. Using Lemma 8 and Table 2, it is easy to verify that switches $(e_{*,*,*}, t^{\rightarrow})$ minimize the occurrence record among all improving switches.³ By the tie-breaking rule, a switch $(e_{i,j,k}, t^{\rightarrow})$ with $\sigma^{(3)}(d_{i,j,k}) = F_{i,j}$ is thus applied next. Since $\sigma^{(3)}(s_{i',*}) = h_{i',*}$ for all $i' < \mu^{\sigma^{(3)}}$ by Lemma 8, the statement of row 1 of Table 10 can be applied.

In addition, it is easy to verify that the characterization given in this row implies that we have $I_{\sigma e} = (I_{\sigma} \setminus \{e\}) \cup \{(d_{i,j,k}, e_{i,j,k})\}$ if $\beta_i^{\sigma} = 0 \vee \beta_{i+1}^{\sigma} \neq j$ and $I_{\sigma e} = I_{\sigma} \setminus \{e\}$ else.³ Consequently, by the tie-breaking rule and row 1 of Table 10, improving switches $(e_{*,*,*}, t^{\rightarrow}) \in \mathbb{E}^1$ are applied until a switch with $\beta_i^{\sigma} = 0 \vee \beta_{i+1}^{\sigma} \neq j$ is applied. Note that the occurrence record of each applied switch is described by Table 2 when interpreted for $b + 1$ since $\lfloor \frac{b}{2} \rfloor + 1 = m$ if b is odd and $\lfloor \frac{b}{2} \rfloor + 1 = \lfloor \frac{b+1}{2} \rfloor$ if b is even. By row 1 of Table 10, $(d_{i,j,k}, e_{i,j,k})$ then becomes improving. As $(d_{i,j,k}, e_{i,j,k}) \notin \mathfrak{A}_{\sigma_b}^{\sigma}$ and since switches of the type $(e_{*,*,*}, t^{\rightarrow})$ minimize the occurrence record, Table 2 and the tie-breaking rule imply that $(d_{i,j,k}, e_{i,j,k})$ is applied next. In particular, an edge $(d_{i,j,k}, e_{i,j,k})$ is applied immediately if it becomes improving and this requires that $(e_{i,j,k}, t^{\rightarrow})$ was applied earlier. Thus, their occurrence record is described by Table 2 when interpreted for $b + 1$. Therefore, the application of improving switches $(e_{*,*,*}, t^{\rightarrow})$ is described by row 1 of Table 10 and whenever an edge $(d_{*,*,*}, e_{*,*,*})$ becomes improving, its application is described by Lemma 11.

Let $G_n = S_n$. Then, row 1 of Table 10 and Lemma 11 can be applied until a strategy σ is reached such that all improving switches $(e_{i,j,k}, t^{\rightarrow})$ with $\sigma^{(3)}(d_{i,j,k}) = F_{i,j}$ were applied. As $(d_{i,j,k}, e_{i,j,k})$ was applied if and only if $\beta_i^{\sigma} = 0 \vee \beta_{i+1}^{\sigma} = j$, this implies that $\sigma(d_{i,j,k}) = F_{i,j}$ is equivalent to $\beta_i^{\sigma} = 1 \wedge \beta_{i+1}^{\sigma} = j$). Consequently, every cycle center is closed or escapes towards t^{\rightarrow} . In addition, an edge $(d_{i,j,k}, F_{i,j})$ is an improving switch exactly if the switch $(e_{i,j,k}, t^{\rightarrow})$ was not applied. Consequently, $I_{\sigma} = \{(d_{i,j,k}, F_{i,j}), (e_{i,j,k}, t^{\rightarrow}) : \sigma(e_{i,j,k}) = t^{\leftarrow}\} \cup \{(b_1, t^{\rightarrow})\}$. Now, as $\phi^{\sigma}(b_1, t^{\rightarrow}) = \phi^{\sigma}(e_{*,*,*}, t^{\rightarrow})$ and $\mathbb{E}^1 = \emptyset$, the switch $e := (b_1, t^{\rightarrow})$ is applied next due to the tie-breaking rule. We prove that we can apply row 2 resp. 4 of Table 10, implying the statement follows for $G_n = S_n$ with arbitrary v and for $G_n = M_n$ with $v = 1$.

We thus prove the following. If $v > 1$, then $\bar{\sigma}(eb_{i,j}) \wedge \neg \bar{\sigma}(eg_{i,j})$ for all $(i, j) \in S_1$ and, in addition, $\bar{\sigma}(eb_{i,j}) \wedge \bar{\sigma}(eg_{i,j})$ for all $(i, j) \in S_2$. If $v = 1$, then $\bar{\sigma}(eg_{i,j}) \wedge \neg \bar{\sigma}(eb_{i,j})$ for all $(i, j) \in S_4$ and $\bar{\sigma}(eb_{i,j}) \wedge \bar{\sigma}(eg_{i,j})$ for all $(i, j) \in S_3$. We begin by observing that the definition of the sets S_1 to S_4 implies that $\beta_i^{\sigma} = 0 \vee \beta_{i+1}^{\sigma} \neq j$ for all of the relevant indices. Also, $(e_{i,j,k}, t^{\rightarrow}) \in \mathfrak{A}_{\sigma^{(3)}}^{\sigma}$ if and only if $\sigma^{(3)}(d_{i,j,k}) = F_{i,j}$. Thus, $(e_{i,j,k}, t^{\rightarrow}) \notin \mathfrak{A}_{\sigma^{(3)}}^{\sigma}$ if and only if $\sigma^{(3)}(d_{i,j,k}) = e_{i,j,k}$. Since $e_{i,j,k}$ has an outdegree of 2, this implies $(\sigma(e_{i,j,k}) = t^{\leftarrow} \Leftrightarrow \sigma^{(3)}(d_{i,j,k}) = e_{i,j,k})$. In particular, due to $\beta_i^{\sigma} = 0 \vee \beta_{i+1}^{\sigma} \neq j$, the switch $(d_{i,j,k}, e_{i,j,k})$ was then also applied. Hence, if

there is a $k \in \{0, 1\}$ with $\sigma^{(3)}(d_{i,j,k}) = e_{i,j,k}$, then $\bar{\sigma}(eg_{i,j})$ if $\nu > 1$ resp. $\bar{\sigma}(eb_{i,j})$ if $\nu = 1$.

Since every cycle center is closed or escapes to t^\rightarrow with respect to σ , either $\bar{\sigma}(eb_{i,j}) \wedge \bar{\sigma}(eg_{i,j})$ or $\bar{\sigma}(eb_{i,j}) \wedge \neg \bar{\sigma}(eg_{i,j})$ or $\bar{\sigma}(d_{i,j})$ for all cycle centers $F_{i,j}$ if $\nu > 1$. Similarly, for $\nu = 1$, either $\bar{\sigma}(eb_{i,j}) \wedge \bar{\sigma}(eg_{i,j})$ or $\bar{\sigma}(eg_{i,j}) \wedge \neg \bar{\sigma}(eb_{i,j})$ or $\bar{\sigma}(d_{i,j})$ for all cycle centers $F_{i,j}$. Consequently, the cycle centers $F_{i,j}$ with $\bar{\sigma}(eb_{i,j}) \wedge \bar{\sigma}(eg_{i,j})$ are exactly the cycle centers that contain a vertex $d_{i,j,k}$ with $\sigma_b(d_{i,j,k}) \neq F_{i,j}$ and such that $(d_{i,j,k}, F_{i,j})$ was *not* applied during Phase 1. By Lemma 8, all improving switches $(d_{i,j,k}, F_{i,j})$ not applied in Phase 1 had $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) = m$. By Corollary 4, it thus suffices to prove that there is a $k \in \{0, 1\}$ with $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) = m$ to prove $\bar{\sigma}(eb_{i,j}) \wedge \bar{\sigma}(eg_{i,j})$. Analogously, to prove $\bar{\sigma}(eb_{i,j}) \wedge \neg \bar{\sigma}(eg_{i,j})$ resp. $\bar{\sigma}(eg_{i,j}) \wedge \neg \bar{\sigma}(eb_{i,j})$, it suffices to show that $F_{i,j}$ was closed at the end of Phase 1.

We now prove the corresponding statements for the case $\nu > 1$ and omit the case $\nu = 1$.³ Let $m = \max\{i : \sigma(b_i) = g_i\}$ and $u = \min\{i : \sigma(b_i) = b_{i+1}\}$.

1. We prove that $\phi^{\sigma_b}(d_{i,j,0}, F_{i,j}) = m$ for all $(i, j) \in S_2$.

- Let $i \leq \nu - 1$ and $j = \beta_{i+1}^\sigma$. Then, $b_{i+1} \neq (b + 1)_{i+1} = \beta_{i+1}^\sigma$ as $i \leq \nu - 1$, so $j \neq b_{i+1}$. Thus, there is a feasible t_b for b with $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) = \min\left(\left\lfloor \frac{b+1-k}{2} \right\rfloor, \ell^b(i, j, k) + t_b\right)$. However, since $i \leq \nu - 1$, $b_i = 1$ and thus $t_b = 0$ is the only feasible parameter. It thus suffices to show $\ell^b(i, j, 0) \geq m$. Since $b_i = 1 \wedge j \neq b_{i+1}$, this follows from Lemma 7.
- Let $i \in \{\nu + 1, \dots, m\}$, $\beta_i^\sigma = 1$ and $j = 1 - \beta_{i+1}^\sigma$. Since $i > \nu$ implies $\beta_i^\sigma = b_i$ and $\beta_{i+1}^\sigma = b_{i+1}$, $\ell^b(i, j, 0) \geq m$ follows as in the previous case.
- Let $i \in \{\nu, \dots, m - 1\} \wedge \beta_i^\sigma = 0$ and $j = \beta_{i+1}^\sigma$. Since $i + 1 > \nu$ implies that we have $\beta_{i+1}^\sigma = b_{i+1}$, $b_{\nu-1} = 1$ and $\nu \geq 2$, we obtain $\ell^b(i, j, 0) > m + 1$ by an easy calculation. Thus, $\ell^b(i, j, k) + t_b > m$ for every t_b feasible for b , implying $\phi^{\sigma_b}(d_{i,j,0}, F_{i,j}) = m$.
- Let $i > m$ and $j \in \{0, 1\}$. Then, $\text{mfn}(b, i + 1) = \text{mufn}(b, i + 1) = 0$ since $b'_i = 0$ for all $b' \leq b$. Hence, by Lemma 7, $\ell^b(i, j, k) \geq b$. Consequently, $\phi^{\sigma_b}(d_{i,j,0}, F_{i,j}) = m$.
- Assume $b + 1 = 2^l$ for some $l \in \mathbb{N}$. Then $\nu = l + 1$ and $b_\nu = 0$. This implies $\text{mfn}(b, \nu, \{(\nu, 1)\}) = \text{mfn}(b, \nu + 1) = \text{mufn}(b, \nu + 1) = 0$, hence $\phi^{\sigma_b}(d_{i,j,0}, F_{i,j}) = m$.

2. We prove that either $\bar{\sigma}_b(d_{i,j})$ or $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) < m$ for both $k \in \{0, 1\}$ for all $(i, j) \in S_1$.

- Let $i \leq \nu - 1$ and $j = 1 - \beta_{i+1}^\sigma$. Then $b_i = 1$ and $j = 1 - \beta_{i+1}^\sigma = b_{i+1}$. Hence $F_{i,j}$ was closed with respect to σ_b .
- Let $i \in \{\nu, \dots, m - 1\}$, $\beta_i^\sigma = 0$ and $j = 1 - \beta_{i+1}^\sigma$. Then $b_i = \beta_i^\sigma = 0$, $\beta_{i+1}^\sigma = b_{i+1}$ and $\beta_i^\sigma = 0$ implies $i \neq \nu$. In particular, $\nu \leq i - 1$ and $b_\nu = 0$. Using Lemma 7, this implies $\ell^b(i, j, k) \leq \left\lfloor \frac{b+1-k}{2} \right\rfloor - 1$. This implies $\ell^b(i, j, k) + 1 \leq \left\lfloor \frac{b+1-k}{2} \right\rfloor$, hence $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) \leq \ell^b(i, j, 1) + 1$. If this inequality is strict, the statement follows. If the inequality is tight, then $\sigma_b(d_{i,j,k}) = F_{i,j}$ by Property Or2 _{i,j,k} and Property Or1 _{i,j,k} thus implies $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) < m$.

- Assume that there is no $l \in \mathbb{N}$ with $b + 1 = 2^l$ and let $i = v$ and $j = 1 - b_{v+1}$. Since b is odd, Property **Or**_{3*i,j,0*} implies $\phi^{\sigma^b}(d_{i,j,0}, F_{i,j}) < m$. For $k = 1, b$ being odd implies $\phi^{\sigma^b}(d_{i,j,1}, F_{i,j}) \leq \lfloor \frac{b+1-1}{2} \rfloor = \lfloor \frac{b}{2} \rfloor < m$.

By row 2 of Table 10, the statement follows for $G_n = S_n$ if $v > 1$. In addition, by row 4 of Table 10, the statement follows for $v = 1$.

It remains to consider the case $v > 1$ if $G_n = M_n$, implying $t^{\rightarrow} = b_2$ and $t^{\leftarrow} = g_1$. Using the same argumentation as before, row 1 of Table 10 and Lemma 11 imply that improving switches within levels $i \geq v$ are applied until we obtain a proper Phase-3-strategy σ for b with

$$I_{\sigma} = \{(d_{i,j,k}, F_{i,j}) : i < v \wedge \sigma(d_{i,j,k}) \neq F_{i,j}\} \cup \{(d_{i,j,k}, F_{i,j}), (e_{i,j,k}, b_2) : i \geq v \wedge \sigma(e_{i,j,k}) = g_1\} \cup \{(b_1, b_2)\}.$$

As no cycle center in any level $i' < v$ was opened yet, the switch $e = (e_{i,j,k}, b_2)$ with $i = v - 1, j = 1 - \beta_{i+1}^{\sigma}$ and $k \in \{0, 1\}$ is applied next. Since $\sigma(d_{i,j,k}) = F_{i,j}$, row 1 of Table 10 implies $I_{\sigma e} = (I_{\sigma} \setminus \{e\}) \cup \{(d_{i,j,k}, e_{i,j,k})\}$. Due to the tie-breaking rule, $(d_{i,j,k}, e_{i,j,k})$ is applied next. We prove that σ meets the requirements of Lemma 10.

There are no other indices i', j', k' with $(d_{i',j',k'}, e_{i',j',k'}) \in I_{\sigma e}$. Also, as no such switch was applied yet in any level below level $i, F_{i,j}$ is closed for σe as it was closed for $\sigma^{(3)}$ by Lemma 8. As $i < v$ and $b_i = 1 \wedge b_{i+1} \neq \beta_{i+1}^{\sigma e}$, Definition 4 implies $\sigma_b(g_i) = F_{i,j}$. By the same arguments used when discussing the case $G_n = S_n$, it can be proven that $F_{i,1-j}$ was not closed during Phase 1 as $(i, 1 - j) \in S_2$. Consequently, $\sigma e(g_i) = F_{i,j}$ follows from Corollary 2. By the tie-breaking rule, no improving switch involving $F_{i,1-j}$ was applied yet. Therefore, $\sigma e(e_{i,1-j,*}) = \sigma^{(3)}(e_{i,1-j,*}) = g_1$ as well as $\sigma e(d_{i,1-j,*}) = \sigma^{(3)}(d_{i,1-j,*})$. By Corollary 4, $F_{i,1-j}$ cannot be open for $\sigma^{(3)}$, so it is not open for σe . Therefore, as $\beta_i^{\sigma e} = 0$ and $1 - j = \beta_{i+1}^{\sigma e}$, it is g_1 -halfopen. Thus, the first requirement of Lemma 10 is met.

By Lemma 8 and since $(s_{i',*}, h_{i',*}) \notin \mathfrak{A}_{\sigma^{(3)}}^{\sigma e}$ for any $i' < v, \sigma e(s_{i',*}) = \sigma^{(3)}(s_{i',*}) = h_{i',*}$ for all $i' < v$. Furthermore, $i' < v$ implies $b_{i'} = 1$ and no improving switch $(d_{*,*,*}, e_{*,*,*})$ below level v was applied yet. Consequently, $\bar{\sigma}e(d_{i'})$ for all $i' < v$. Now consider some cycle center $F_{i',j'}$ where $i' < i$ and $j' = 1 - \bar{\sigma}e(g_{i'})$. We prove that $F_{i',j'}$ is g_1 -halfopen. The cycle center $F_{i',\beta_{i'+1}^{\sigma e}}$ is not closed while $F_{i',1-\beta_{i'+1}^{\sigma e}}$ is closed due to $1 - \beta_{i'}^{\sigma e} = b_{i'}$. Thus, by Corollary 2 and the same arguments used before, $\bar{\sigma}e(g_{i'}) = \bar{\sigma}_b(g_{i'}) = 1 - \beta_{i'+1}^{\sigma e}$ and, in particular, $j' = \beta_{i'+1}^{\sigma e}$. However, by Corollary 4 and the tie-breaking rule, this implies that $F_{i',j'}$ is g_1 -halfopen as before. Thus, the second requirement of Lemma 10 is met.

The third requirement is met as $i' > i = v - 1$ and since σe has Property **Usv**_{1*i'*}.

Consider the fourth requirement. Let $i' > i$ and $\beta_{i'}^{\sigma} = 0$. Then, due to the tie-breaking rule, all improving switches $(e_{i',j',k'}, b_2)$ with $\sigma^{(3)}(d_{i',j',k'}) = F_{i',j'}$ have already been applied. Since $\beta_{i'}^{\sigma} = 0, F_{i',\beta_{i'+1}^{\sigma}}$ cannot have been closed for $\sigma^{(3)}$. If both cycle centers of level i' were g_1 -halfopen for $\sigma^{(3)}$, then they are mixed for σ , and $\sigma(g_{i'}) = \sigma^{(3)}(g_{i'}) = \sigma_b(g_{i'}) = F_{i',\beta_{i'+1}^{\sigma}}$. If $F_{i',1-\beta_{i'+1}^{\sigma}}$ is closed for $\sigma^{(3)}$, then $F_{i',\beta_{i'+1}^{\sigma}}$ can only be g_1 -halfopen for $\sigma^{(3)}$. Consequently, by Corollary 2 resp. Definition 4,

$\bar{\sigma}(g_{i'}) = 1 - \beta_{i'+1}^\sigma$. Furthermore, $F_{i', 1 - \beta_{i'+1}^\sigma}$ is then b_2 -open and $F_{i', \beta_{i'+1}^\sigma}$ is b_2 -halfopen (for σ). Thus, the fourth requirement is met.

By the same argument, if $i' > i$ and $\beta_{i'+1}^\sigma = 1$, then $F_{i', 1 - \beta_{i'+1}^\sigma}$ is b_2 -open if it was closed for $\sigma^{(3)}$ and mixed if it was g_1 -halfopen. Thus, the fifth and final requirement is met.

Therefore, the application of $(d_{i,j,k}, e_{i,j,k})$, yields a proper Phase-3-strategy $\sigma \in \Lambda_i$ for \mathfrak{b} with $I_\sigma = (I_{\sigma_e} \setminus \{(d_{i,j,k}, e_{i,j,k})\}) \cup \{(s_{i,j}, b_1)\}$. We prove $\phi^\sigma(s_{i,j}, b_1) < \phi^\sigma(e_{i,j,k}, b_2) = \lfloor \frac{\mathfrak{b}}{2} \rfloor$, implying that $(s_{i,j}, b_1)$ is applied next. It is easy to verify that $(s_{i,j}, b_1) \notin \mathfrak{A}_{\sigma_{\mathfrak{b}}}^\sigma$. Consequently, by Table 2 and as $i = \nu - 1$ and $j = 1 - \beta_{i+1}^\sigma = 0$, $\phi^\sigma(s_{i,j}, b_1) = \text{fl}(\mathfrak{b}, i + 1) - j \cdot \mathfrak{b}_{i+1} = \text{fl}(\mathfrak{b}, \nu) \leq \lfloor \frac{\mathfrak{b}+2}{4} \rfloor < m$ if $\mathfrak{b} \geq 3$ since $\nu \geq 2$. If $\mathfrak{b}_1 = 1$, then $(s_{i,j}, b_1)$ is also the next switch applied as the tie-breaking rule then ranks $(s_{i,j}, b_1)$ higher than any switch $(e_{*,*,*}, b_2)$. Since $(e_{i,j,k}, b_2), (d_{i,j,k}, e_{i,j,k}) \in \mathfrak{A}_{\sigma_{\mathfrak{b}}}^\sigma$ and since $F_{i,j}$ was closed $(e_{i,j,k}, b_2)$ was applied, we have $\bar{\sigma}(e_{i,j,k}) \wedge \neg \bar{\sigma}(e_{g_{i,j}})$. Therefore, the last row of Table 10 describes the application of $e = (s_{i,j}, b_1)$. Consequently, σe is a proper Phase-3-strategy with $I_{\sigma e} = I_\sigma \setminus \{e\}$ and $\phi^{\sigma e}(s_{i,j}, b_1) = \text{fl}(\mathfrak{b}, \nu) + 1 = \text{fl}(\mathfrak{b} + 1, \nu)$ by Lemma 9. Thus, Table 2 describes the occurrence record of $(s_{i,j}, b_1)$ when interpreted for $\mathfrak{b} + 1$. Since $F_{i,j}$ is b_2 -halfopen for σe whereas $F_{i, 1-j}$ is g_1 -halfopen, $(e_{i,j, 1-k}, b_2)$ is applied next. By the first row of Table 10, this application unlocks $(d_{i,j, 1-k}, e_{i,j, 1-k})$. Using our previous arguments and observations, it is easy to verify that $(d_{i,j, 1-k}, e_{i,j, 1-k})$ is applied next and that its application is described by the second-to-last row of Table 10. The tie-breaking rule then chooses to apply $(e_{i, 1-j, k}, b_2) \in \mathbb{E}^1$ next. By row 1 of Table 10, $(d_{i, 1-j, k}, e_{i, 1-j, k})$ then becomes improving and is applied next. Its application is described by row 5 of Table 10. After applying this switch, we then obtain a strategy σ with

$$I_\sigma = \{(d_{i,j,k}, F_{i,j}) : i < \nu - 1 \wedge \sigma(d_{i,j,k}) \neq F_{i,j}\} \cup \{(d_{i,j,k}, F_{i,j}), (e_{i,j,k}, b_2) : i \geq \nu - 1 \wedge \sigma(e_{i,j,k}) = g_1\} \cup \{(b_1, b_2)\}.$$

It is easy to verify that the same arguments can be applied iteratively as applying a switch $(s_{i',j'}, b_1)$ with $i' < \nu$ always requires to open the corresponding cycle center $F_{i',j'}$ first. Thus, after finitely many iterations, we obtain a strategy σ with $I_\sigma = \{(d_{i,j,k}, F_{i,j}), (e_{i,j,k}, b_2) : \sigma(e_{i,j,k}) = g_1\} \cup \{(b_1, b_2)\}$. By the same arguments used for the case $G_n = S_n$, the conditions of the third row of Table 10 are met, so we obtain a strategy as described by the corresponding rows of Tables 6 and 7. \square

We henceforth use $\sigma^{(3)}$ to refer to the Phase-3-strategy described by Lemma 8. As before, we implicitly proved the following corollary which follows from Corollary 5.

Corollary 6 *Let $\sigma^{(4)}$ be the Phase-4-strategy calculated by the Strategy Improvement resp. Policy Iteration Algorithm when starting with a canonical strategy $\sigma_{\mathfrak{b}}$ fulfilling the canonical conditions as per Lemma 9. Then, Table 2 specifies the occurrence record of every improving switch applied so far when interpreted for $\mathfrak{b} + 1$, excluding switches $(g_*, F_{*,*})$, and each such switch was applied once.*

Reaching phase 5 if there is a phase 4

As shown by Lemma 9, we do not always obtain a Phase-5-strategy immediately after Phase 3 as we have to apply improving switches involving selection vertices $s_{i,*}$ in levels $i < v$ if $G_n = S_n$. We thus prove that we also reach a Phase 5 strategy after applying these switches. Consequently, we always reach a Phase 5 strategy. The application of all of these switches is captured by the following lemma.

Lemma 12 *Let $G_n = S_n$. Let $\sigma \in \Lambda_i$ be a Phase-4-strategy for $\mathfrak{b} \in \mathcal{B}_n$ with $v > 1$. Assume $e := (s_{i,1-\beta_{i+1}^\sigma}, b_1) \in I_\sigma$ for some $i < v$. Further assume that σ has Property $\text{Usv}1_{i'}$ for all $i' > i$ and that $\sigma(d_{i',j',k'}) = F_{i',j'} \Leftrightarrow \beta_{i'}^\sigma = 1 \wedge \beta_{i'+1}^\sigma = j'$. Moreover, assume that $i' < v$ implies $\bar{\sigma}(g_{i'}) = 1 - \beta_{i'+1}^\sigma$ and that $i' < i$ implies $\sigma(s_{i',*}) = h_{i',*}$. If $(s_{i',1-\beta_{i'+1}^\sigma}, b_1) \in I_\sigma$ for some $i' < i$, then σe is a Phase-4-strategy for \mathfrak{b} . Otherwise, it is a Phase-5-strategy for \mathfrak{b} . In either case, $I_{\sigma e} = (I_\sigma \setminus \{e\}) \cup \{(d_{i,j,0}, F_{i,j}), (d_{i,j,1}, F_{i,j})\}$.*

We now state the lemma describing the application of the improving switches during Phase 4.

Lemma 10 *Let $\sigma_{\mathfrak{b}} \in \Lambda_i$ be a canonical strategy for $\mathfrak{b} \in \mathcal{B}_n$ fulfilling the canonical conditions. After applying a finite number of improving switches, the Strategy Improvement resp. Policy Iteration Algorithm produces a Phase-5-strategy $\sigma^{(5)} \in \Lambda_i$ as described by the corresponding rows of Tables 6 and 7.*

Proof By Lemma 9, it suffices to consider the case $G_n = S_n$ and $v > 1$. By Lemma 9, the Strategy Improvement Algorithm calculates a Phase-4-strategy σ for \mathfrak{b} with $\sigma \in \Lambda_i$ and

$$I_\sigma = \{(d_{i,j,k}, F_{i,j}), (e_{i,j,k}, b_2) : \sigma(e_{i,j,k}) = g_1\} \cup \{(s_{v-1,0}, b_1)\} \cup \{(s_{i,1}, b_1) : i \leq v - 2\} \cup X_0 \cup X_1.$$

We show that $(s_{v-1,0}, b_1)$ is applied next. Assume that $\mathfrak{b} + 1$ is a power of two, implying $\mathfrak{b} = 2^{v-1} - 1$. It is easy to verify that $(s_{v-1,0}, b_1)$ and $(s_{v-2,1}, b_1)$ both minimize the occurrence record if $\mathfrak{b} > 1$.³ If $\mathfrak{b} = 1$, then the switches $(e_{i,j,k}, b_2)$ with $\sigma(e_{i,j,k}) = g_1$ also minimize the occurrence record. Due to the tie-breaking rule, $(s_{v-1,0}, b_1)$ is however applied next in any case.

If $\mathfrak{b} + 1$ is not a power of two, then $\mathfrak{b} \geq 2^v + 2^{v-1} - 1$ and $\mathfrak{b} \geq 6$, implying $\lfloor \frac{\mathfrak{b}+2}{4} \rfloor < m$ as well as $\lfloor \frac{\mathfrak{b}+2}{4} \rfloor < \lfloor \frac{\mathfrak{b}}{2} \rfloor$. It is again easy to verify that this implies that $(s_{v-1,0}, b_1)$ minimizes the occurrence record.³ Thus, the switch $e = (s_{v-1,0}, b_1)$ is applied next in any case. We now prove that Lemma 12 describes the application of e . Since σ is a Phase-4-strategy and since $i' > i = v - 1$ implies $i' \geq v$, σ has Property $\text{Usv}1_{i'}$ for all $i' > i$. By Lemma 9, it follows that σ also meets the other requirements of Lemma 12.

Consider the case $v = 2$ first. Then, applying $e = (s_{1,0}, b_1)$ yields a Phase-5-strategy and $\phi^{\sigma e}(e) = \text{fl}(\mathfrak{b}, v) + 1 = \text{fl}(\mathfrak{b} + 1, v)$ by Lemma 9. Hence, Table 2 describes the occurrence record of e with respect to $\mathfrak{b} + 1$. In addition, we then have

$$I_{\sigma e} = (I_\sigma \setminus \{e\}) \cup \{(d_{1,0,0}, F_{1,0}), (d_{1,0,1}, F_{1,0})\}$$

$$= \{(d_{i,j,k}, F_{i,j}), (e_{i,j,k}, b_2) : \sigma e(e_{i,j,k}) = g_1\} \\ \cup \{(d_{i,1-\beta_{i+1}^\sigma, k}, F_{i,1-\beta_{i+1}^\sigma}) : i \leq v - 1\} \cup X_0 \cup X_1.$$

Since σe is a Phase-5-strategy, it has Property **Rel1**, implying $\mu^{\sigma e} = u = \min\{i : \beta_i^{\sigma e} = 0\}$. Thus, σe has all properties listed in the corresponding rows of Tables 6 and 7.

Before discussing the case $v > 2$, we discuss edges $(d_{i,j,k}, F_{i,j})$ that become improving when a switch $(s_{i,j}, b_1)$ with $i < v$ and $j = 1 - \beta_{i+1}^\sigma$ is applied, see Lemma 12. Since $i < v$ implies $1 - \beta_{i+1}^\sigma = b_{i+1}$, their cycle centers $F_{i,j}$ were closed for σ_b . Therefore, their occurrence record might be very low. However, their occurrence records are not “too low” in the sense that they interfere with the improving switches applied during Phase 4. More precisely, it can be shown that $i < v$ and $j = b_{i+1}$ imply $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) > \lfloor \frac{b+2}{4} \rfloor - 1$.³

We now consider case $v > 2$. We obtain $\phi^{\sigma e}(e) = \text{fl}(b + 1, v)$ as before. Furthermore,

$$I_{\sigma e} = \{(d_{i,j,k}, F_{i,j}), (e_{i,j,k}, b_2) : \sigma e(e_{i,j,k}) = g_1\} \\ \cup \{(s_{i,1}, b_1) : i \leq v - 2\} \cup \{(d_{v-1,0,0}, F_{v-1,0}), (d_{v-1,0,1}, F_{v-1,0})\}$$

by Lemma 12. We show that the switches $(s_{v-2,1}, b_1), \dots, (s_{1,1}, b_1)$ are applied next and in this order. To simplify notation, we denote the current strategy by σ . By Table 2, it holds that $\phi^\sigma(s_{i,1}, b_1) = \text{fl}(b, i + 1) - 1$ for all $i \leq v - 2$. Hence $\phi^\sigma(s_{v-2,1}, b_1) < \dots < \phi^\sigma(s_{1,1}, b_1)$ by Lemma 9. It thus suffices to show that the occurrence record of $\phi^\sigma(s_{1,1}, b_1)$ is smaller than the occurrence record of any switch improving for σ and any improving switch that might be unlocked by applying some switch $(s_{i,1}, b_1)$ for $i \leq v - 2$.

The second statement follows since $\phi^{\sigma_b}(s_{1,1}, b_1) = \text{fl}(b, 2) - 1 = \lfloor \frac{b+2}{4} \rfloor - 1$ and since the occurrence record of any edge that becomes improving is bounded by $\lfloor \frac{b+2}{4} \rfloor$ as discussed earlier. It thus suffices to show the first statement. Let $e := (d_{i,j,k}, F_{i,j}) \in I_\sigma$ such that $\sigma(e_{i,j,k}) = g_1$. By Lemma 8 and Lemma 9, this implies $\phi^\sigma(e) = m = \text{fl}(b, 1)$. Then, $v > 2$ implies $\text{fl}(b, 1) > \text{fl}(b, v - 1)$, hence $\phi^\sigma(e) < \phi^\sigma(s_{1,1}, b_1)$. Next let $e := (e_{i,j,k}, b_2) \in I_\sigma$ such that $\sigma(e_{i,j,k}) = g_1$. Then, since b is odd, Table 2 implies $\phi^\sigma(e) = \lfloor \frac{b}{2} \rfloor = m - 1 = \text{fl}(b, 1) - 1$, hence $\phi^\sigma(e) > \phi^\sigma(s_{1,1}, b_1)$. If $b + 1$ is not a power of two, we need to show this estimation for some more improving switches. But this can be shown by the easy calculations similar to the calculations necessary when discussing the application of $(s_{v-1,0}, b_1)$.

Consequently, the switches $(s_{v-1}, b_1), \dots, (s_{1,1}, b_1)$ are applied next by the Strategy Improvement Algorithm and in this order. It is easy to verify that the requirements of Lemma 12 are always met, so this lemma describes the application of these switches. It is also easy to check that the occurrence record of these edges is described by Table 2 after applying them. Let σ denote the strategy obtained after applying $(s_{1,1}, b_1)$. Then σ is a Phase-5-strategy for b with $\sigma \in A_i$ and $\mu^\sigma = \min\{i : \beta_i^\sigma = 0\}$. This further

implies

$$I_\sigma = \{(d_{i,j,k}, F_{i,j}), (e_{i,j,k}, b_2) : \sigma(e_{i,j,k}) = g_1\} \\ \cup \{(d_{i,1-(b+1)_{i+1,*}}, F_{i,1-(b+1)_{i+1}}) : i \leq \nu - 1\} \cup X_0 \cup X_1$$

Further note that $\sigma(e_{i,j,k}) = g_1$ still implies $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) = \phi^\sigma(d_{i,j,k}, F_{i,j}) = m$ since the corresponding switches are improving since the end of Phase 1. Also, every improving switch was applied at most once and we proved that the occurrence record of any improving switch that was applied is described correctly by Table 2 when interpreted for $b + 1$. Since no improving switch involving cycle vertices were applied, we also still have $\sigma^\omega(d_{i,j,k}) = F_{i,j}$ if and only of $(b + 1)_i = 1$ and $(b + 1)_{i+1} = j$. Hence, all conditions listed in the corresponding rows of Tables 6 and 7 are fulfilled, proving the statement. \square

We henceforth use $\sigma^{(3)}$ to refer to the Phase-3-strategy described by Lemma 8. As before, we implicitly proved the following corollary which follows from Corollaries 5 and 6.

Corollary 7 *Let $\sigma^{(5)}$ be the Phase-5-strategy calculated by the Strategy Improvement resp. Policy Iteration Algorithm when starting with a canonical strategy σ_b fulfilling the canonical conditions as per Lemma 10. Then, Table 2 specifies the occurrence record of every improving switch applied so far when interpreted for $b + 1$, excluding switches $(g_*, F_{*,*})$, and each such switch was applied once.*

Reaching a canonical strategy for $b + 1$

We now prove Lemma 5. It states that applying improving switches according to Zadeh’s pivot rule and our tie-breaking rule produces the strategies described by Tables 7 and 6. In particular, it states that we obtain a canonical strategy for $b + 1$ fulfilling the canonical conditions. We split the proof as follows. First, we observe that it only remains to prove Lemma 5 for $k = 5$. We then prove that we obtain a canonical strategy for $b + 1$ and show that this strategy fulfills the canonical conditions afterwards. These two statements thus imply Lemma 5 for $k = 5$.

We first provide an overview describing the application of improving switches during Phase 5.³

Table 11 Improving switches applied during Phase 5

Properties of σ	Switch e	Properties of σe
$\sigma(b_i) = b_{i+1}$ $\bar{\sigma}(g_i) = 1 - \beta_{i\pm 1}^\sigma$ $\beta_i^\sigma = 0$	$(d_{i,j,k}, F_{i,j})$ $i \neq 1 \wedge j = 1 - \beta_{i\pm 1}^\sigma$	Phase-5-strategy for b $I_{\sigma e} = I_\sigma \setminus \{e\}$ $\sigma(e_{*,*,*}) = t^{-\tau} \wedge (\text{SvI})_{i'} \bar{\forall} i' \in [n]$
$\nu = 1 \Rightarrow \bar{\sigma}(eg_{i,j}) \wedge \neg \bar{\sigma}(eg_{i,j})$ $\nu > 1 \Rightarrow \bar{\sigma}(eb_{i,j}) \wedge \neg \bar{\sigma}(eg_{i,j})$ $\mu^\sigma = 1 \Rightarrow [i' \geq i \Rightarrow \bar{\sigma}(d_{i',*}) \vee (\bar{\sigma}(eb_{i',*}) \wedge \neg \bar{\sigma}(eg_{i',*}))]$	$(g_i, F_{i,j})$	\Rightarrow Phase-1-strategy for $b + 1$ Else Phase-5-strategy for b $I_{\sigma e} = I_\sigma \setminus \{e\}$

For convenience, we always assume $\sigma \in A_l$ and that σ is a Phase-5-strategy for b . Note that we thus also always have $\sigma e \in A_l$

The following lemma describes the application of switches $(e_{i,j,k}, t^{\rightarrow})$ during Phase 5. As it is rather involved, we state it separately and do not include it into Table 11.

Lemma 13 *Let $\sigma \in \Lambda_l$ be a Phase-5-strategy for $\mathfrak{b} \in \mathcal{B}_n$. Let $e := (e_{i,j,k}, t^{\rightarrow}) \in I_\sigma$ and let $F_{i,j}$ be mixed. Assume that $\nu > 1 \Rightarrow \neg \bar{\sigma}(eg_{i,1-j})$ and $\nu = 1 \Rightarrow \neg \bar{\sigma}(eb_{i,1-j})$ if $G_n = S_n$ and $j = 1$. Similarly, assume that $\nu > 1 \Rightarrow \neg \bar{\sigma}(eg_{i,1-j})$ and $\nu = 1 \Rightarrow \neg \bar{\sigma}(eb_{i,1-j})$ if $G_n = M_n$ $j = 1 - \beta_{i+1}^\sigma$. Moreover, assume that $\nu = 2$ implies $\sigma(g_1) = F_{1,0}$ if $G_n = S_n$. Then the following hold.*

1. *If there are $(i', j', k') \neq (i, j, k)$ with $(e_{i',j',k'}, t^{\rightarrow}) \in I_\sigma$ or if there is an i' such that σ does not have Property **Sv1** $_{i'}$, then σe is a Phase-5-strategy for \mathfrak{b} .*
2. *If there are no $(i', j', k') \neq (i, j, k)$ with $(e_{i',j',k'}, t^{\rightarrow}) \in I_\sigma$ and if σ has Property **Sv1** $_{*}$, then σe is a Phase-1-strategy for $\mathfrak{b} + 1$.*
3. *If $G_n = S_n$,*

$$(g_i, F_{i,j}) \in I_{\sigma e} \iff \beta_i^{\sigma e} = 0 \wedge \bar{\sigma e}(g_i) = 1 \wedge j = 0 \wedge \begin{cases} \bar{\sigma}(eb_{i,1-j}), & \nu > 1 \\ \bar{\sigma}(eg_{i,1-j}), & \nu = 1 \end{cases}.$$

If $G_n = M_n$

$$\begin{aligned} (g_i, F_{i,j}) \in I_{\sigma e} &\iff \beta_i^{\sigma e} = 0 \wedge \bar{\sigma e}(g_i) = 1 - \beta_{i+1}^\sigma \wedge j \\ &= \beta_{i+1}^\sigma \wedge \begin{cases} \bar{\sigma}(eb_{i,1-j}), & \nu > 1 \\ \bar{\sigma}(eg_{i,1-j}), & \nu = 1 \end{cases}. \end{aligned}$$

If the corresponding conditions are fulfilled, then $I_{\sigma e} = (I_\sigma \setminus \{e\}) \cup \{(d_{i,j,1-k}, F_{i,j}), (g_i, F_{i,j})\}$. Otherwise, $I_{\sigma e} = (I_\sigma \setminus \{e\}) \cup \{(d_{i,j,1-k}, F_{i,j})\}$.

We also use the following statement for comparing valuations of cycle centers.

Lemma 14 *Let $G_n = M_n$. Let $\sigma \in \Lambda_l$ be a Phase- k -strategy, $k \in \{1, \dots, 5\}$, for some $\mathfrak{b} \in \mathcal{B}_n$ having Property **Usv1** $_i$ and Property **Bac1** $_{i+1}$ for some $i \in \{1, \dots, n\}$. If $F_{i,0}$ and $F_{i,1}$ are in the same state and if either $i \geq \nu$ or σ has Property **Rel1**, then $\Xi_\sigma^M(F_{i,\beta_{i+1}^\sigma}) > \Xi_\sigma^M(F_{i,1-\beta_{i+1}^\sigma})$.*

We now prove that we obtain a canonical strategy $\sigma_{\mathfrak{b}+1}$ for $\mathfrak{b} + 1$.

Lemma 11 *Let $\sigma_{\mathfrak{b}} \in \Lambda_l$ be a canonical strategy for \mathfrak{b} fulfilling the canonical conditions. Then, applying improving switches according to Zadeh’s pivot rule and our tie-breaking rule produces a canonical strategy $\sigma_{\mathfrak{b}+1} \in \Lambda_l$ for $\mathfrak{b} + 1$ with $I_{\sigma_{\mathfrak{b}+1}} = \{(d_{i,j,k}, F_{i,j}) : \sigma_{\mathfrak{b}+1}(d_{i,j,k}) \neq F_{i,j}\}$.*

Proof By Lemma 10, applying improving switches according to Zadeh’s pivot rule and our tie-breaking rule yields a Phase-5-strategy $\sigma^{(5)}$ for \mathfrak{b} with $\sigma^{(5)} \in \Lambda_l$ and $\mu^{\sigma^{(5)}} = u = \min\{i : \beta_i^{\sigma^{(5)}} = 0\}$. Let $m := \max\{i : \beta_i^\sigma = 1\}$ and $\sigma := \sigma^{(5)}$.

Consider the case $\nu = 1$. We first investigate the occurrence record of the improving switches.

1. Let $e = (d_{i,j,k}, F_{i,j})$ resp. $e = (e_{i,j,k}, g_1)$ with $\sigma(e_{i,j,k}) = b_2$. Then $\phi^\sigma(e) = \lfloor \frac{b+1}{2} \rfloor$ by Lemma 10 resp. $\phi^{\sigma^{(5)}}(e) = \phi^{\sigma_b}(e) = \lceil \frac{b}{2} \rceil = \lfloor \frac{b+1}{2} \rfloor$.
2. Let $e = (d_{i,j,k}, F_{i,j})$ with $\beta_i^\sigma = 0, i \in \{u + 1, \dots, m - 1\}$ and $j = 1 - \beta_{i+1}^\sigma$. Then, $b_i = 0$ and $j = 1 - b_{i+1}$ since $i \geq u + 1 > 1$ and $v = 1$. In addition, $b_1 = 0$ and, due to $i > u$, there is at least one $l \in \{2, \dots, i - 1\}$ with $(b + 1)_l = b_l = 0$. Consequently, Lemma 7 yields $\ell^b(i, j, k) = \lceil \frac{b-2^{i-1} + \sum_{2 \leq l \leq i-1} (b_l + 1) - k}{2} \rceil \leq \lceil \frac{b-3-k}{2} \rceil = \lfloor \frac{b-k}{2} \rfloor - 1$. Since there is a t_b feasible for b , it holds that $\phi^{\sigma_b}(e) = \min(\lfloor \frac{b+1-k}{2} \rfloor, \ell^b(i, j, k) + t_b)$. We thus distinguish the following cases.
 - (a) Assume $\phi^{\sigma_b}(e) = \ell^b(i, j, k) + 1$. Then, by Property Or2 $_{i,j,k}$, $\sigma_b(d_{i,j,k}) = F_{i,j}$ and e was not applied yet. Consequently, Property Ori $_{i,j,k}$ implies $\phi^\sigma(e) = \phi^{\sigma_b}(e) < m$.
 - (b) Assume $\phi^{\sigma_b}(e) = \ell^b(i, j, k)$. Then, $\phi^{\sigma_b}(e) \leq \phi^\sigma(e) \leq \lfloor \frac{b-k}{2} \rfloor - 1 < m$ as well as $\sigma_b(d_{i,j,k}) \neq F_{i,j}$ by Property Or2 $_{i,j,k}$. This implies $\phi^{\sigma_b}(e) = m - 1$ by Property Or4 $_{i,j,k}$. Hence, by Corollary 1, e was applied during Phase 1. Consequently, $\phi^\sigma(e) = \phi^{\sigma_b}(e) + 1 = m$.
 - (c) The case $\phi^{\sigma_b}(e) = \ell^b(i, j, k) - 1$ cannot occur since $t_b = -1$ is not feasible as b is even.
 - (d) Assume $\phi^{\sigma_b}(e) = \lfloor \frac{b+1-k}{2} \rfloor$ and $\lfloor \frac{b+1-k}{2} \rfloor \neq \ell^b(i, j, k)$ and $\lfloor \frac{b+1-k}{2} \rfloor \neq \ell^b(i, j, k) + 1$. This implies $\lfloor \frac{b+1-k}{2} \rfloor < \ell^b(i, j, k)$ since $\phi^{\sigma_b}(e) = \min(\lfloor \frac{b+1-k}{2} \rfloor, \ell^b(i, j, k) + t_b)$. But this is a contradiction since $\ell^b(i, j, k) \leq \lfloor \frac{b-k}{2} \rfloor - 1$.

Thus, the occurrence records of all improving switches are bounded by $\lfloor \frac{b+1}{2} \rfloor$ and some switches have an occurrence record of exactly $\lfloor \frac{b+1}{2} \rfloor$. Consequently, improving switches $(d_{i,j,k}, F_{i,j})$ with $i \in \{u + 1, \dots, m - 1\}, \beta_i^\sigma = 0$ and $j = 1 - \beta_{i+1}^\sigma$ are applied first. Further note that an edge $e = (d_{i,j,k}, F_{i,j})$ can only fulfill these assumptions if $\sigma_b(d_{i,j,k}) = F_{i,j}$, implying that e was not applied during Phase 1. It furthermore implies $\phi^\sigma(e) = \phi^{\sigma_b}(e) = \ell^b(i, j, k) + 1$. Consider such an edge e and assume $\phi^\sigma(e) < m$. We show that row 1 of Table 11 applies to e . We thus prove $\sigma(b_i) = b_{i+1}, j = 1 - \beta_{i+1}^\sigma, \bar{\sigma}(g_i) = 1 - \beta_{i+1}^\sigma$ and $i \neq 1$. The first two statements follow directly since σ is a Phase-5-strategy and $\beta_i^\sigma = 0$ as well as by the choice of j . Also, $i \neq 1$ follows from $i \geq u + 1 > 1$. It thus suffices to show $\bar{\sigma}(g_i) = 1 - \beta_{i+1}^\sigma$.

This can be shown by using Corollaries 2 and 1.³ Hence, let $e = (d_{i,j,k}, F_{i,j}) \in I_\sigma$ be an improving switch minimizing the occurrence record. Then, $\sigma[e]$ is a Phase-5-strategy for b with $\sigma e \in \Lambda_t$ and $I_{\sigma e} = I_\sigma \setminus \{e\}$ by row 1 of Table 11. By Lemma 6 and the choice of i and $j, \ell^b(i, j, k) + 1 = \ell^{b+1}(i, j, k)$. In particular,

$$\phi^{\sigma'}(e) = \ell^b(i, j, k) + 1 + 1 = \ell^{b+1}(i, j, k) + 1 \leq m \leq \left\lfloor \frac{(b + 1) + 1 - k}{2} \right\rfloor.$$

Thus, by choosing the parameter $t_{b+1} = 1$, which is feasible since $i \neq 1$, the occurrence record of e is described by Table 2 when interpreted for $b + 1$.

Now, the same arguments can be used for all improving switches $e' \in \mathbb{D}^1 \cap I_{\sigma'}$ having an occurrence record smaller than m . All of these switches are thus applied and their occurrence records are specified by Table 2 when interpreted for $b + 1$. After the application of these switches, we obtain a Phase-5-strategy σ for b with $\sigma \in \Lambda_t$ and

$$I_{\sigma} = \{(d_{i,j,k}, F_{i,j}), (e_{i,j,k}, g_1) : \sigma(e_{i,j,k}) = b_2\} \cup \bigcup_{\substack{i=u+1 \\ \beta_i^{\sigma}=0}}^{m-1} \left\{ e = (d_{i,1-\beta_{i+1}^{\sigma}}, F_{i,1-\beta_{i+1}^{\sigma}}) : \phi^{\sigma}(e) = m \right\}. \tag{2}$$

In particular, all improving switches have an occurrence record of m . Thus, the tie-breaking rule now applies a switch of the type $(e_{*,*,*}, g_1)$. Let $e = (e_{i,j,k}, g_1)$ denote the switch applied next. We prove that Lemma 13 applies to this switch. First, we show that $F_{i,j}$ is mixed. Since $e = (e_{i,j,k}, g_1) \in I_{\sigma}$ implies $(d_{i,j,k}, F_{i,j}) \in I_{\sigma}$, we have $\bar{\sigma}(eb_{i,j})$. In particular, $F_{i,j}$ is not closed, so $\beta_i^{\sigma} = 0 \vee \beta_{i+1}^{\sigma} \neq j$. Consequently, $(i, j) \in S_3$ or $(i, j) \in S_4$. By Lemma 10, $\bar{\sigma}(eb_{i,j})$, and as no improving switch $(e_{*,*,*}, b_2)$ was applied during $\sigma^{(5)} \rightarrow \sigma$, we need to have $(i, j) \in S_3$, implying the statement. We now prove that $j = 1$ implies $\neg\bar{\sigma}(eb_{i,1-j})$ if $G_n = S_n$. Since $j = 1$, we need to prove $\neg\bar{\sigma}(eb_{i,0})$. If $F_{i,0}$ is closed, then the statement follows. If $F_{i,0}$ is not closed, then $\beta_i^{\sigma} = 0 \vee \beta_{i+1}^{\sigma} \neq j$ as $F_{i,1}$ cannot be closed by the choice of e . Consequently, $(i, 0) \in S_3$ or $(i, 0) \in S_4$. In the second case, $\neg\bar{\sigma}^5(eb_{i,0})$ by Lemma 10 and the statement follows as no improving switch $(e_{*,*,*}, b_2)$ was applied during $\sigma^{(5)} \rightarrow \sigma$. Consider the case $(i, 0) \in S_3$. Then, by Lemma 10, $F_{i,0}$ and $F_{i,1}$ are mixed with respect to $\sigma^{(5)}$. Thus, as we consider the case $G_n = S_n$, the tie-breaking rule must have applied the improving switches $(e_{i,0,*}, g_1)$ prior to $(e_{i,1,k}, g_1)$, implying the statement. Note that “ $j = 1 - \beta_{i+1}^{\sigma} \implies \neg\bar{\sigma}(eb_{i,1-j})$ if $G_n = M_n$ ” follows by the same arguments and since the tie-breaking rule applies improving switches $(e_{i,\beta_{i+1}^{\sigma}}, g_1)$ first.

These arguments can be applied for any improving switch $(e_{*,*,*}, g_1)$. Thus, Lemma 13 applies to the switch e . Observe that $\phi^{\sigma e}(e)$ is specified by Table 2 when interpreted for $b + 1$ as $v = 1$ implies $\lceil \frac{b}{2} \rceil + 1 = \lceil \frac{b+1}{2} \rceil$. If the conditions listed in the fourth case of Lemma 13 are fulfilled, then $I_{\sigma e} = (I_{\sigma e} \setminus \{e\}) \cup \{(d_{i,j,1-k}, F_{i,j}), (g_i, F_{i,j})\}$. If these conditions are not fulfilled, then $I_{\sigma e} = (I_{\sigma e} \setminus \{e\}) \cup \{(d_{i,j,1-k}, F_{i,j})\}$. In particular, $\tilde{e} := (d_{i,j,1-k}, F_{i,j})$ becomes improving in either case. We prove that \tilde{e} has an occurrence record of m or $m + 1$.

By the characterization of I_{σ} , we have $(d_{i,j,k}, F_{i,j}) \in I_{\sigma}$, implying $\sigma(d_{i,j,k}), \sigma e(d_{i,j,k}) \neq F_{i,j}$. Since σ is a Phase-5-strategy for b , this implies $\beta_i^{\sigma} = 0 \vee \beta_{i+1}^{\sigma} \neq j$. Assume that \tilde{e} was applied previously in this transition. It is not possible that \tilde{e} was applied during Phase 5 since this would imply $\sigma e(d_{i,j,1-k}) = F_{i,j}$, contradicting that $F_{i,j}$ is mixed. Consequently, \tilde{e} was applied during Phase 1. Thus, $\sigma_b(d_{i,j,1-k}) \neq F_{i,j}$ and $\phi^{\sigma_b}(\tilde{e}) \in \{m - 1, m\}$ by Property Or $4_{i,j,1-k}$. This implies $\phi^{\sigma e}(\tilde{e}) \in \{m, m + 1\}$ as claimed. Now assume that \tilde{e} was not applied previously, implying $\phi^{\sigma e}(\tilde{e})$

$= \phi^{\sigma_b}(\tilde{e})$. Consider the case $\sigma_b(d_{i,j,1-k}) \neq F_{i,j}$. Then, by Property Or4 $_{i,j,1-k}$, $\phi^{\sigma_b}(\tilde{e}) \in \{m - 1, m\}$. But this implies $\phi^{\sigma_b}(\tilde{e}) = m$ by Corollary 1 as claimed. Thus assume $\sigma_b(d_{i,j,1-k}) = F_{i,j}$. Then, by Property Or1 $_{i,j,1-k}$ and Property Or2 $_{i,j,1-k}$, we have $\phi^{\sigma_b}(\tilde{e}) = \ell^b(i, j, 1 - k) + 1 < m$. Using Lemma 7 and by distinguishing between the cases $\beta_i^{\sigma_e} = 1 \wedge \beta_{i+1}^{\sigma_e} \neq j$, $\beta_i^{\sigma_e} = 0 \wedge \beta_{i+1}^{\sigma_e} = j$ and $\beta_i^{\sigma_e} = 0 \wedge \beta_{i+1}^{\sigma_e} \neq j$, it is easy to verify that this always yields a contradiction.³ We thus have just proven the following.

Corollary 8 *If the edge $(d_{i,j,1-k}, F_{i,j})$ becomes improving during Phase 5 due to the application of $(e_{i,j,k}, g_1)$, then the corresponding strategy has Property Or4 $_{i,j,1-k}$.*

Now, assume that $(g_i, F_{i,j})$ becomes improving when applying $(e_{i,j,k}, g_1)$. Using the conditions stated in Lemma 13, Corollary 2 and Lemma 10, it is easy to verify that $(g_i, F_{i,j}) \notin \mathcal{A}_{\sigma_b}^{\sigma_e}$.³

Since $\phi^{\sigma_e}(g_i, F_{i,j}) = \phi^{\sigma_b}(g_i, F_{i,j})$, we have $\phi^{\sigma_e}(g_i, F_{i,j}) \leq \phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) = m$ by Table 2. As argued previously, $\phi^{\sigma_e}(d_{i,j,1-k}, F_{i,j}) \geq m$. Therefore, the occurrence record of any improving switch except $(g_i, F_{i,j})$ is at least m . Thus, $(g_i, F_{i,j})$ either uniquely minimizes the occurrence record or has the same occurrence record as all other improving switches. Consequently, by the tie-breaking rule, $(g_i, F_{i,j})$ is applied next in either case.

We prove that row 2 of Table 11 applies to this switch. Since $v = 1, \mu^{\sigma_e} = u > 1$ and $\beta_i^{\sigma_e} = 0$, it suffices to prove $\bar{\sigma}e(eg_{i,j}) \wedge \neg\bar{\sigma}e(eb_{i,j})$. But this follows as we applied $(e_{i,j,k}, g_1)$ earlier and since $F_{i,j}$ was mixed when this switch was applied. Observe that the following corollary holds due to the conditions which specify when a switch $(g_i, F_{i,j})$ is unlocked, independent on v .

Corollary 9 *If an improving switch $(g_i, F_{i,j})$ is applied during Phase 5, then the created strategy has Property Sv1 $_i$.*

Let σ denote the strategy obtained after applying $(e_{i,j,k}, g_1)$ and (eventually) $(g_i, F_{i,j})$. For now, assume that there are indices i', j', k' with $(e_{i',j',k'}, g_1) \in I_\sigma$. Then, by Lemma 13 resp. row 2 of Table 11, σ is a Phase-5-Strategy for b . By our previous discussion, all improving switches have an occurrence record of at least m . Among all improving switches with an occurrence record of exactly m , the tie-breaking rule then decides which switch to apply. There are two types of improving switches. Each switch is either of the form $(d_{i,j,k}, F_{i,j})$ or of the form $(e_{i,j,k}, g_1)$ with $\sigma(d_{i,j,k}) = e_{i,j,k}$. Since every edges $(e_{i,j,k}, g_1)$ minimizes the occurrence record among all improving switches, one of these edges is chosen. But then, the same arguments used previously can be used again. More precisely, Lemma 13 applies to this switch, making the edge $(d_{i,j,1-k}, F_{i,j})$ and eventually also $(g_i, F_{i,j})$ improving. Also, Corollaries 8 and 9 apply to these switches and another switch of the form $(e_{*,*,*}, g_1)$ is applied afterwards. Thus, inductively, all remaining switches of the form $(e_{i,j,k}, g_1)$ are applied.

Each of these applications creates the improving switch $(d_{i,j,1-k}, F_{i,j})$ and might make $(g_i, F_{i,j})$ improving. In the latter case, the corresponding switch is then applied immediately. Let σ denote the strategy that is reached before the last improving switch of the form $(e_{*,*,*}, g_1)$ is applied. We argue that this switch is $e := (e_{1,1-\beta_2^\sigma,k}, g_1)$ for

some $k \in \{0, 1\}$ and that σ has Property **Sv1** $_i$ for all $i \in \{1, \dots, n\}$. As the tie-breaking rule applies improving switches in higher levels first, it suffices to prove that there there is a $k \in \{0, 1\}$ such that $e \in I_{\sigma^{(5)}}$. This however follows from Lemma 10 as $v = 1$ implies $(1, \beta_2^\sigma) \in S_3$. It remains to prove that σ has Property **Sv1** $_i$ for all $i \in \{1, \dots, n\}$. If $\beta_i^\sigma = 1$, then this follows from the definition of a Phase-5-strategy. If $\beta_i^\sigma = 0$ and $(g_i, F_{i,j}) \in \mathcal{X}_{\sigma^{(5)}}^\sigma$, then this follows from Corollary 9. Thus, let $\beta_i^\sigma = 0$ and $(g_i, F_{i,j}) \notin \mathcal{X}_{\sigma^{(5)}}^\sigma$, implying $i \neq 1$ since $v = 1$. We now prove the following statement. If $\bar{\sigma}(g_i) = 1$ (if $G_n = S_n$) resp. $\bar{\sigma}(g_i) = 1 - \beta_{i+1}^\sigma$ (if $G_n = M_n$) and $-\bar{\sigma}(d_{i,1})$ (if $G_n = S_n$), resp. $-\bar{\sigma}(d_{i,1-\beta_{i+1}^\sigma})$ (if $G_n = M_n$), then $(g_i, F_{i,0}) \in I_\sigma$ resp. $(g_i, F_{i,\beta_{i+1}^\sigma}) \in I_\sigma$. Note that this proves the statement as $I_\sigma \cap \mathbb{G} = \emptyset$.

Thus, let $j := 0$ (if $G_n = S_n$) resp. $j := \beta_{i+1}^\sigma$ (if $G_n = M_n$) and assume $-\bar{\sigma}(d_{i,1-j})$. It suffices to prove $\mathcal{E}_\sigma^*(F_{i,j}) \succ \mathcal{E}_\sigma^*(F_{i,1-j})$. As $\sigma(e_{i',j',k'}) = g_1$ for all $(i', j', k') \neq (1, \beta_2^\sigma, k)$, $i \neq 1$ and $\mu^\sigma = u \neq 1$, the cycle centers are either closed or escape only to g_1 . Note that a closed cycle center has $\mathcal{E}_\sigma^*(F_{i,j}) = \mathcal{E}_\sigma^*(s_{i,j})$ due to Lemma 1. Let $G_n = S_n$. If both cycle centers escape towards g_1 , then the statement follows since

$$\begin{aligned} \mathcal{E}_\sigma^S(F_{i,0}) &= \{F_{i,0}, d_{i,0,*}, e_{i,0,*}, b_1\} \cup \mathcal{E}_\sigma^S(g_1) \triangleright \{F_{i,1}, d_{i,1,*}, e_{i,1,*}, b_1\} \cup \mathcal{E}_\sigma^S(g_1) \\ &= \mathcal{E}_\sigma^S(F_{i,1}) \end{aligned}$$

due to the priorities of $F_{i,0}$ and $F_{i,1}$. As $\beta_i^\sigma = 0$, only $F_{i,1-\beta_{i+1}^\sigma}$ can be closed. Assume that we have $j = 0 = 1 - \beta_{i+1}^\sigma$. Then, by Property **Usv1** $_i$ and since $\sigma(b_1) = g_1$, the statement follows since $\mathcal{E}_\sigma^S(F_{i,0}) = \{s_{i,0}, b_1\} \cup \mathcal{E}_\sigma^S(g_1)$ and $\mathcal{E}_\sigma^S(F_{i,1}) = \{F_{i,1}, d_{i,1,k}, e_{i,1,k}, b_1\} \cup \hat{\mathcal{E}}_\sigma^S(g_1)$ for some $k \in \{0, 1\}$. Now assume $j = 0 = \beta_{i+1}^\sigma$. Then, $F_{i,1-j} = F_{i,1}$ is closed, contradicting the assumption of the statement. Let $G_n = M_n$. If both cycle centers are g_1 -open or g_1 -halfopen, then the statement follows by Lemma 14 since $i > v$. If one of the cycle centers is g_1 -open and one cycle center is g_1 -halfopen, then the statement follows by an easy calculation.³ Since only $F_{i,1-\beta_{i+1}^\sigma} = F_{i,1-j}$ can be closed in level i , the statement then follows by the same argument used for the case $G_n = S_n$.

Thus, Lemma 13 applies to $(e_{1,\beta_2^\sigma,k}, g_1)$. Let $\sigma_{b+1} := \sigma e$. Then, by Lemma 13, σ_{b+1} is a Phase-1-strategy for $b + 1$ with $\sigma_{b+1} \in \Lambda_i$. Note that, since every edge was applied at most once during $\sigma_b \rightarrow \sigma^{(5)}$ by Lemma 10 and since no edge applied during $\sigma^{(5)} \rightarrow \sigma_{b+1}$ was applied earlier, every edge was applied at most once as improving switch during $\sigma_b \rightarrow \sigma_{b+1}$. Also note that we implicitly proved the following corollary.

Corollary 10 *Let $v = 1$ and let σ_{b+1} denote the strategy obtained after the application of the final improving switch $(e_{*,**,} g_1)$. Then $(d_{i,j,k}, F_{i,j}) \in \mathcal{X}_{\sigma^{(5)}}^{\sigma_{b+1}}$ if and only if $\sigma_b(d_{i,j,k}) = F_{i,j}, \phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) < m, i \in \{u + 1, \dots, m - 1\}, \beta_i^{\sigma_{b+1}} = 0$ and $j = 1 - \beta_{i+1}^{\sigma_{b+1}}$. In addition, σ_{b+1} has Property **Or2** $_{i,j,k}$*

We prove that σ_{b+1} is a canonical strategy for $b + 1$ with $I_{\sigma_{b+1}} = \mathfrak{D}^{\sigma_{b+1}}$. To simplify notation, let $\sigma := \sigma_{b+1}$. We first prove $I_\sigma = \{(d_{i,j,k}, F_{i,j}) : \sigma(d_{i,j,k}) \neq F_{i,j}\}$. Consider the strategy $\sigma^{(5)}$. Using the characterization of the strategy that was obtained after having applied all switches $(d_{i,j,k}, F_{i,j})$ with an occurrence record smaller than m (see Eq. (2)), we obtain

$$I_\sigma = \{(d_{i,j,*}, F_{i,j}) : \sigma^{(5)}(e_{i,j,*}) = b_2\} \cup \bigcup_{\substack{i=u+1 \\ \beta_i^\sigma=0}}^{m-1} \left\{ e = (d_{i,1-\beta_{i+1}^\sigma}, F_{i,1-\beta_{i+1}^\sigma}) : \phi^{\sigma^{(5)}}(e) = m \right\}.$$

In particular, $I_\sigma \subseteq \{(d_{i,j,k}, F_{i,j}) : \sigma(d_{i,j,k}) \neq F_{i,j}\}$ and every improving switch has an occurrence record of at least m . To prove $\{(d_{i,j,k}, F_{i,j}) : \sigma(d_{i,j,k}) \neq F_{i,j}\} \subseteq I_\sigma$, let $e := (d_{i,j,k}, F_{i,j})$ with $\sigma(d_{i,j,k}) \neq F_{i,j}$. It thus suffices to show $\mathcal{E}_\sigma^*(F_{i,j}) \succ \mathcal{E}_\sigma^*(e_{i,j,k})$. Property **Esc1** and $\nu = 1$ imply $\bar{\sigma}(eg_{i,j}) \wedge \neg\bar{\sigma}(eb_{i,j})$. Furthermore, $\mu^\sigma = \min\{i' : \sigma(b_{i'}) = b_{i'+1}\} \neq 1$ as σ has Property **Rel1**. It is easy to verify that this implies $\mathcal{E}_\sigma^S(F_{i,j}) = \{F_{i,j}\} \cup \mathcal{E}_\sigma^S(e_{i,j,k})$, hence $\mathcal{E}_\sigma^S(F_{i,j}) \triangleright \mathcal{E}_\sigma^S(e_{i,j,k})$. For $G_n = M_n$, it suffices to prove $\mathcal{E}_\sigma^M(s_{i,j}) > \mathcal{E}_\sigma^M(g_1)$ as this implies $\mathcal{E}_\sigma^M(F_{i,j}) > \mathcal{E}_\sigma^M(g_1)$. Since $\sigma(d_{i,j,k}) \neq F_{i,j}$, either $\beta_i^\sigma = 0$ or $\beta_{i+1}^\sigma \neq j$. In the second case, Property **Usv1_i** implies that we have $\sigma(s_{i,j}) = b_1$ and the statement follows since $\mathcal{E}_\sigma^M(s_{i,j}, b_1) = \langle s_{i,j} \rangle + \mathcal{E}_\sigma^M(g_1)$ due to $\sigma(b_1) = g_1$. Thus assume $\beta_i^\sigma = 0 \wedge \beta_{i+1}^\sigma = j$. Then, the statement follows since $\mathcal{E}_\sigma^M(s_{i,j}) = \langle s_{i,j}, h_{i,j} \rangle + \mathcal{E}_\sigma^M(b_{i+1})$ by Property **Bac1_{i+1}** and since $\langle s_{i,j}, h_{i,j} \rangle > \sum_{\ell < i} \langle g_\ell, s_{\ell, \bar{\sigma}(g_\ell)}, h_{\ell, \bar{\sigma}(g_\ell)} \rangle$.

We now prove that σ is a canonical strategy for $b + 1$. Since σ is a Phase-1-strategy for $b + 1$, we have $b + 1 = \beta^\sigma$. Consider the conditions listed in Definition 2 resp. 4. Condition 1 is fulfilled since $\sigma(e_{*,*,*}) = g_1$ and $\nu = 1$. Condition 2(a) is fulfilled since $\beta_i^\sigma = (b + 1)_i = 1$ implies $\sigma(b_i) = g_i$ by Property **Bac1_i**. Consider condition 2(b). If $(b + 1)_i = 1$, then $F_{i,(b+1)_{i+1}}$ is closed by Property **Bac1_i**. It is easy to verify that $(b + 1)_i = 1$ implies that $F_{i,1-(b+1)_{i+1}}$ cannot be closed, hence Condition 2(b) is fulfilled.³ Condition 2(c) is fulfilled by $\beta^\sigma = b + 1$ and Property **Bac2_{*}**. Conditions 3(a) and 3(b) are fulfilled since σ has Property **Bac1_{*}**. Consider condition 3(c). We prove that $(b + 1)_i = 0$ and $\bar{\sigma}(d_{i,j})$ imply $\sigma(g_i) = F_{i,j}$ where $j = 1 - (b + 1)_{i+1}$. By Lemma 1, $F_{i,j}$ being closed implies $\hat{\mathcal{E}}_\sigma^*(F_{i,j}) = \hat{\mathcal{E}}_\sigma^*(s_{i,j})$. Thus, $\mathcal{E}_\sigma^*(F_{i,j}) = \llbracket s_{i,j} \rrbracket \oplus \mathcal{E}_\sigma^*(g_1)$ by the choice of j and since $\nu = 1$. It can be shown that $\mu^\sigma \neq 1$, $\bar{\sigma}(eg_{i,1-j})$, $\neg\bar{\sigma}(eb_{i,1-j})$ and $1 - j = \beta_{i+1}^\sigma$ implies $\mathcal{E}_\sigma^*(F_{i,1-j}) = \{F_{i,1-j}, d_{i,1-j,k}, e_{i,1-j,k}, b_1\} \cup \mathcal{E}_\sigma^*(g_1)$ for some $k \in \{0, 1\}$.³ But this implies $\sigma(g_i) = F_{i,j}$ since $(g_i, F_{i,1-j}) \in I_\sigma$ otherwise, contradicting $I_\sigma = \{(d_{i,j,k}, F_{i,j}) : \sigma(d_{i,j,k}) = F_{i,j}\}$.

Consider condition 3(d) and let $j := 0$ if $G_n = S_n$ and $j := \beta_{i+1}^\sigma$ if $G_n = M_n$. It suffices to prove $\mathcal{E}_\sigma^*(F_{i,j}) \succ \mathcal{E}_\sigma^*(F_{i,1-j})$ if none of the cycle centers is closed. For $G_n = M_n$, this follows from Lemma 14 or an easy calculation if $G_n = M_n$ since $i \geq 1 = \nu$. For $G_n = S_n$, this follows from $\Omega(F_{i,0}) > \Omega(F_{i,1})$ and since both priorities are even. Conditions 4 and 5 follow as σ has Property **Usv1_{*}**. For condition 6, let $i := \nu(b + 2)$, $j := (b + 1)_{i+1}$ and $k \in \{0, 1\}$. Since $\nu(b + 1) = 1$, we have $i \geq 2$ and $b_i = (b + 1)_i = 0$ as well as $b_{i+1} = (b + 1)_{i+1} = j$. We prove $\sigma(d_{i,j,k}) \neq F_{i,j}$. For the sake of contradiction, let $\sigma(d_{i,j,k}) = F_{i,j}$. Then, by the choice of i and j and Lemma 10, $(d_{i,j,k}, F_{i,j}) \in \mathcal{A}_{\sigma^{(5)}}^*$. Thus, by Corollary 10 and Property **Or2_{i,j,k}**, it holds that $\phi^{\sigma^{(5)}}(d_{i,j,k}, F_{i,j}) < m$ and $\phi^{\sigma^b}(d_{i,j,k}, F_{i,j}) = \ell^b(i, j, k) + 1$. But, by Lemma 7, we have $\ell^b(i, j, k) = \lceil \frac{b+2i-1+\sum(b,i)+1-k}{2} \rceil \geq \lceil \frac{b+3-k}{2} \rceil = \lfloor \frac{b+2-k}{2} \rfloor$, which is a contradiction. Hence, $\sigma(d_{i,j,k}) \neq F_{i,j}$.

Now consider the case $\nu > 1$. Then, b is odd and $m = \lfloor \frac{b}{2} \rfloor + 1$. By Lemma 10, applying improving switches according to Zadeh’s pivot rule and our tie-breaking rule yields a Phase-5-strategy σ for b with $\sigma \in \Lambda_l$ and $\mu^\sigma = u$. In addition

$$I_\sigma = \{(d_{i,j,k}, F_{i,j}), (e_{i,j,k}, b_2) : \sigma(e_{i,j,k}) = g_1\} \cup \bigcup_{i=1}^{\nu-1} \{(d_{i,1-\beta_{i+1}^\sigma}, F_{i,1-\beta_{i+1}^\sigma})\} \cup X_0 \cup X_1,$$

where, X_k is defined as in Table 7. We now investigate the occurrence records of the improving switches. Note that $\mathcal{E}_\sigma^*(g_1) < \mathcal{E}_\sigma^*(b_2)$ since $\sigma(e_{i,j,k}) = g_1$ implies $(e_{i,j,k}, b_2) \in I_\sigma$.

1. Let $e = (d_{i,j,k}, F_{i,j})$ with $\sigma(e_{i,j,k}) = g_1$. Then, by Lemma 10, $\phi^\sigma(e) = \phi^{\sigma_b}(e) = m$.
2. Let $e = (e_{i,j,k}, b_2)$ with $\sigma(e_{i,j,k}) = g_1$. Then, by Table 2, $\phi^\sigma(e) = \phi^{\sigma_b}(e) = \lfloor \frac{b}{2} \rfloor = m - 1$.
3. Let $e = (d_{\nu,j,k}, F_{\nu,j})$ with $j := 1 - \beta_{\nu+1}^\sigma$ for some $k \in \{0, 1\}$. This edge is only improving if $b + 1$ is not a power of two. Note that this implies $\mathbf{1}_{j=0} \text{mf}_n(b, \nu + 1) + \mathbf{1}_{j=1} \text{muf}_n(b, \nu + 1) \neq 0$. Since $b_\nu = 0 \wedge b_{\nu+1} \neq j$, Lemma 7 thus implies

$$\ell^b(\nu, j, k) = \begin{cases} m, & k = 0 \\ m - 1, & k = 1 \end{cases}.$$

Since $b + 1$ is not a power of two, the parameter $t_b = -1$ is not feasible by Property Or3 $_{i,j,k}$. Hence $\phi^{\sigma_b}(d_{\nu,j,k}, F_{\nu,j}) = m - k$. This implies $(d_{\nu,j,1}, F_{\nu,j}) \in \mathcal{A}_{\sigma_b}^\sigma$ by Corollary 1. Consequently, $\phi^\sigma(d_{\nu,j,k}, F_{\nu,j}) = m$ for both $k \in \{0, 1\}$.

4. Let $e = (d_{i,j,k}, F_{i,j})$ with $i \in \{\nu + 1, \dots, m - 1\}$, $\beta_i^\sigma = 0$, $j := 1 - \beta_{i+1}^\sigma$ and $k \in \{0, 1\}$. This edge is only improving if $b + 1$ is not a power of two. Since $i > \nu$, $\beta_i^\sigma = 0$ implies $b_i = 0 \wedge b_{i+1} \neq j$. Also, $i < m$ implies $\mathbf{1}_{j=0} \text{mf}_n(b, i + 1) + \mathbf{1}_{j=1} \text{muf}_n(b, i + 1) \neq 0$ since $j = 1 - \beta_{i+1}^\sigma$ and $b \geq 1$ by the choice of i . It is easy to verify that $b_\nu = 0$ then implies $\ell^b(i, j, k) \leq m - 1$.³ There are two cases. If $\sigma_b(d_{i,j,k}) = F_{i,j}$, then $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) = \ell^b(i, j, k) + 1 \leq m - 1$ by Property Or1 $_{i,j,k}$. If $\sigma_b(d_{i,j,k}) \neq F_{i,j}$, then $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) = \ell^b(i, j, k) \leq m - 1$. In the first case, e was not applied during Phase 1 and $\phi^{\sigma_b}(e) = \phi^\sigma(e) \leq m - 1$. In the second case, $\phi^{\sigma_b}(e) = m - 1$ by Property Or4 $_{i,j,k}$. Then, e was applied during Phase 1, implying $\phi^\sigma(d_{i,j,k}, F_{i,j}) = m$.
5. Let $e = (d_{i,j,k}, F_{i,j})$ with $i \leq \nu(b + 1) - 1$ and $j := 1 - \beta_{i+1}^\sigma$. Then, bit i and bit $i + 1$ switched during $\sigma_b \rightarrow \sigma^{(5)}$. In particular, $F_{i,j}$ was closed with respect to σ_b and consequently $(d_{i,j,k}, F_{i,j}) \notin \mathcal{A}_{\sigma_b}^\sigma$. Hence, by Table 2, $\phi^\sigma(e) = \phi^{\sigma_b}(e) = \lfloor \frac{b-2^{i-1}+3-k}{2} \rfloor$. We now distinguish several cases.

- For $i = 1$, $\phi^\sigma(e) = \lfloor \frac{b+2-k}{2} \rfloor = m$ independent of k .

- For $i = 2$, $\phi^\sigma(e) = \lfloor \frac{b+1-k}{2} \rfloor$, so $\phi^\sigma(e) = m$ if $k = 0$ and $\phi^\sigma(e) = m - 1$ if $k = 1$.
- For $i = 3$, $\phi^\sigma(e) = \lfloor \frac{b-1-k}{2} \rfloor$, so $\phi^\sigma(e) = m - 1$ if $k = 0$ and $\phi^\sigma(e) = m - 2$ if $k = 1$.
- For $i > 3$, it is easy to see that the occurrence record is always strictly smaller than $m - 1$.

We partition I_σ into three subsets. A switch $e \in I_\sigma$ is called *type 1 switch* if $\phi^\sigma(e) = m$, *type 2 switch* if $\phi^\sigma(e) = m - 1$ and *type 3 switch* if $\phi^\sigma(e) < m - 1$. By Zadeh’s pivot rule, type 3 switches are applied first. Thus, let $e \in I_\sigma$ be a type 3 switch and note that this implies $e = (d_{i,j,k}, F_{i,j})$ where either $i < v - 1, j = 1 - \beta_{i+1}^\sigma$ or $i \in \{v+1, \dots, m-1\}, \beta_i^\sigma = 0, j := 1 - \beta_{i+1}^\sigma$ as well as $\sigma_b(d_{i,j,k}) = F_{i,j}$. In particular, by Property Or2 _{i,j,k} , these switches fulfill $\phi^{\sigma_b}(e) = \ell^b(i, j, k) + 1$. We prove that applying e can be described by row 1 of Table 11. Since it is easy to verify $i \neq 1$ and $\sigma(b_i) = b_{i+1}$ using the previous explanations and that any improving switch of type 3 has $\sigma(b_i) = b_{i+1}$, we only show $\bar{\sigma}(g_i) = 1 - \beta_{i+1}^\sigma$. By Lemma 10, this holds for all $i \leq v - 1$. It thus suffices to prove this for $i \in \{v + 1, \dots, m - 1\} \wedge \beta_i^\sigma = 0$. This can be shown by proving that $\bar{\sigma}(g_i) = \beta_{i+1}^\sigma$ implies $(g_i, F_{i,1-\beta_{i+1}^\sigma}) \in I_\sigma$, contradicting the given characterization of I_σ .³

Thus, all requirements for applying row 1 of Table 11 are met. We next show that Table 2 specifies the occurrence record of $e = (d_{i,j,k}, F_{i,j})$ its application when interpreted for $b + 1$. First let $i \in \{v + 1, \dots, m - 1\}, \beta_i^\sigma = 0$ and $j = 1 - \beta_{i+1}^\sigma$. Since e is a type 3 switch, this implies $\phi^{\sigma_b}(e) = \ell^b(i, j, k) + 1$ as we pointed out in the previous paragraph. Since e is applied, the statement follows since $\ell^{b+1}(i, j, k) = \ell^b(i, j, k) + 1$ by Lemma 9. Now let $i \leq v - 1$. Then, $F_{i,j}$ was closed with respect to σ_b and $j = b_{i+1} = 1 - \beta_{i+1}^\sigma$. It is easy to verify that this implies that we have $\phi^{\sigma_b}(e) = \lceil \frac{b-\sum(b,i)+1-k}{2} \rceil$. Since $(b + 1)_i = 0 \wedge (b + 1)_{i+1} \neq j$ and the switch e is applied, it suffices to prove $\ell^{b+1}(i, j, k) = \lceil \frac{b-\sum(b,i)+1-k}{2} \rceil$ as we can then choose $t^{b+1} = 1$ as feasible. This however follows by an easy calculation.³ Note that we do not discuss yet that choosing this parameter is in accordance with Properties (Or1) _{$*,*,*$} to (Or4) _{$*,*,*$} . Note that e being a type 3 switch thus implies $\phi^{\sigma_e}(e) \leq m - 1 = \lfloor \frac{b+1+1}{2} \rfloor - 1$. Hence, σ_e has Property Or1 _{i,j,k} .

Corollary 11 *Let $v > 1$. Every switch $e = (d_{i,j,k}, F_{i,j})$ with $\phi^{\sigma_b}(e) < m - 1$ (i.e., every switch of type 3) is applied during Phase 5, and the obtained strategy has Property Or1 _{i,j,k} .*

Now, the first row of Table 11 and the corresponding arguments can be applied for every improving switch of type 3. Thus, we obtain a Phase-5-strategy $\sigma \in \Lambda_t$ such that any improving switch is of type 1 or 2. The next improving switch that is applied has an occurrence record of $m - 1$, i.e., it is of type 2. Since any switch is either of the form $(d_{i,j,k}, F_{i,j})$ or $(e_{i,j,k}, b_2)$ and since the latter switches are of type 2, some switch $e = (e_{i,j,k}, b_2)$ is applied next due to the tie-breaking rule. We prove that Lemma 13 describes the application of this switch. We begin by proving that $F_{i,j}$ is mixed. Since only improving switches of type 3 were applied, $\sigma(e_{i,j,k}) = g_1$

implies $\sigma(d_{i,j,k}) = e_{i,j,k}$. Consequently, $\bar{\sigma}(eg_{i,j})$. In particular, $F_{i,j}$ is not closed, so $\beta_i^\sigma = 0 \vee \beta_{i+1}^\sigma = j$. Thus, either $(i, j) \in S_1$ or $(i, j) \in S_2$. By Lemma 10, $\bar{\sigma}(eg_{i,j})$ and as no switch $(e_{*,*,*}, g_1)$ was applied during $\sigma^{(5)} \rightarrow \sigma$, we need to have $(i, j) \in S_2$, implying that $F_{i,j}$ is mixed. We go on and prove that $j = 1$ resp. $j = 1 - \beta_{i+1}^\sigma$ (depending on whether $G_n = S_n$ or $G_n = M_n$) implies $\neg\bar{\sigma}(eg_{i,1-j})$. Consider the case $G_n = S_n$ and thus $j = 1$ first. We prove $\neg\bar{\sigma}(eg_{i,0})$. If $F_{i,0}$ is closed, then the statement follows. If it is not closed, then $\beta_i^\sigma = 0 \vee \beta_{i+1}^\sigma \neq 0$. Consequently, either $(i, 0) \in S_1$ or $(i, 0) \in S_2$. In the first case, $\neg\bar{\sigma}(eg_{i,0})$ follows from Lemma 10 as no improving switch $(e_{*,*,*}, b_2)$ was applied during $\sigma^{(5)} \rightarrow \sigma$, so assume $(i, 0) \in S_2$. Then, by Lemma 10, both cycle centers $F_{i,0}, F_{i,1}$ were mixed for $\sigma^{(5)}$. Thus, as we consider $G_n = S_n$, the tie-breaking rule must have applied the improving switches $(e_{i,0,*}, b_2)$ prior to $(e_{i,j,k}, b_2)$, implying $\neg\bar{\sigma}(eg_{i,0})$. If $G_n = M_n$, $\neg\bar{\sigma}(eg_{i,1-j})$ follows by the same arguments as the tie-breaking rule applied the improving switches $(e_{i,\beta_{i+1}^\sigma,*}, b_2)$ first. Finally, as no improving switch $(g_*, F_{*,*})$ was applied during $\sigma^{(5)} \rightarrow \sigma$, $v = 2$ implies $\sigma(g_1) = F_{1,0}$ if $G_n = S_n$ by Lemma 10.

Thus, all requirements of Lemma 13 are met. In addition, Table 2 describes the occurrence record of e when interpreted for $b + 1$ since $\phi^{\sigma^b}(e) = \phi^{\sigma^b}(e) + 1 = \lfloor \frac{b}{2} \rfloor + 1 = m$. Now, by Lemma 13, $(d_{i,j,1-k}, F_{i,j}) \in I_{\sigma^b}$ and $(g_i, F_{i,j})$ might become improving for σ^b . The strategy σ^b is now either a Phase-5-strategy for b or a Phase-1-strategy for $b + 1$. By similar arguments used for proving Corollary 8, it is easy to verify the following corollary.³

Corollary 12 *If $(d_{i,j,1-k}, F_{i,j})$ becomes improving during Phase 5 after the application of $(e_{i,j,k}, b_2)$, then the corresponding strategy has Property Or4 _{$i,j,1-k$} and $\min_{k' \in \{0,1\}} \phi^{\sigma^b}(d_{i,j,k'}, F_{i,j}) \leq m - 1$.*

We now use Corollary 12 to prove that $(g_i, F_{i,j})$ is applied next if it becomes improving. Note that σ^b is a Phase-5-strategy for b if such a switch is unlocked since it does not have Property Sv1 _{i} then. Let σ denote the current strategy and note that we consider a strategy that was obtained by applying an improving switch $(e_{i,j,*}, b_2)$ according to Lemma 13.

Due to the tie-breaking rule and Corollary 12, it suffices to show $\phi^\sigma(g_i, F_{i,j}) \leq m - 1$. Since Table 2 and Corollary 12 yield $\phi^{\sigma^b}(g_i, F_{i,j}) \leq \min_{k' \in \{0,1\}} \phi^{\sigma^b}(d_{i,j,k'}, F_{i,j}) \leq m - 1$, it suffices to prove $(g_i, F_{i,j}) \notin \mathfrak{A}_{\sigma^b}^\sigma$. By Lemma 13, $e \in I_\sigma$ if and only if $\beta_i^\sigma = 0, \bar{\sigma}(eb_{i,1-j})$ and $[j = 0 \wedge \bar{\sigma}(g_i) = 1]$ if $G_n = S_n$ resp. $[j = \beta_{i+1}^\sigma \wedge \bar{\sigma}(g_i) = 1 - \beta_{i+1}^\sigma]$ if $G_n = M_n$. Let, for the sake of contradiction, $(g_i, F_{i,j}) \in \mathfrak{A}_{\sigma^b}^\sigma$. The conditions on j and $\bar{\sigma}(g_i)$ imply $(g_i, F_{i,j}) \notin \mathfrak{A}_{\sigma^{(5)}}^\sigma$. Since $\beta_i^\sigma = 0$ implies $i \neq v$, also $(g_i, F_{i,j}) \neq (g_v, F_{v,*})$. Thus, by Lemma 10, $b_i = 0 \wedge b_{i+1} \neq j$. Consequently, $0 = b_i = \beta_{i+1}^\sigma = (b + 1)_{i+1}$ and $j = 1 - b_{i+1}$. Since all bits below level v have $b_i = 1 \wedge (b + 1)_i = 0$, this implies $i > v$. Therefore, $b_{i+1} = (b + 1)_{i+1} = 1 - j$ and in particular $j = 1 - \beta_{i+1}^\sigma$. This is a contradiction if $G_n = M_n$ as $j = \beta_{i+1}^\sigma$, hence consider the case $G_n = S_n$. Then, $j = 1 - \beta_{i+1}^\sigma = 0$, implying $\beta_{i+1}^\sigma = 1$. Thus, $i \in \{v + 1, \dots, m - 1\}, \beta_i^\sigma = 0$ and $j = 1 - \beta_{i+1}^\sigma$, implying $(i, j) \in S_1$. Therefore, $\bar{\sigma}^5(eb_{i,j}) \wedge \neg\bar{\sigma}^5(eg_{i,j})$, contradicting $(e_{i,j,k}, b_2), (d_{i,j,k}, e_{i,j,k}) \in I_{\sigma^{(5)}}$. Thus, $(g_i, F_{i,j}) \notin \mathfrak{A}_{\sigma^b}^{\sigma^{(5)}}$, implying $\phi^\sigma(g_i, F_{i,j}) \leq m - 1$.

Due to the tie-breaking rule, $(g_i, F_{i,j})$ is thus applied next. We prove that the row 2 of Table 10 applies to our situation.

First, $\beta_i^\sigma = 0$ follows from the conditions of Lemma 13. Second, $\bar{\sigma}(eb_{i,j}) \wedge \neg\bar{\sigma}(eg_{i,j})$ follows as the cycle center $F_{i,j}$ was mixed earlier and since we just applied $(e_{i,j,k}, b_2)$. To prove that we have $\bar{\sigma}(d_{i',j'}) \vee [\bar{\sigma}(eb_{i',j'}) \wedge \neg\bar{\sigma}(eg_{i',j'})]$ holds for all $i' \geq i$ and $j \in \{0, 1\}$, fix some $i' \geq i$ and $j' \in \{0, 1\}$. If $\beta_{i'}^\sigma = 1 \wedge \beta_{i'+1}^\sigma = j'$, the statement follows from Property **Bac1** $_{i'}$. We may hence assume $\beta_{i'}^\sigma = 0 \vee j' \neq \beta_{i'+1}^\sigma$ and that $F_{i',j'}$ is not closed. Then, by Lemma 10, either $\bar{\sigma}(eb_{i',j'}) \wedge \bar{\sigma}(eg_{i',j'})$ or $\bar{\sigma}(eb_{i',j'}) \wedge \neg\bar{\sigma}(eg_{i',j'})$. Assume, for the sake of contradiction, that the first case was true and note that this implies $i' \neq i$. Then, for some $k \in \{0, 1\}$, $\sigma(e_{i',j',k}) = g_1$ and $\sigma(d_{i',j',k}) = e_{i',j',k}$. This in particular implies $(e_{i',j',k}, b_2) \in I_\sigma$. This is however a contradiction to the fact that we apply improving switches according to the tie-breaking rule since $i' > i$ implies that the switch $(e_{i',j',k}, b_2)$ is applied before the switch $(e_{i,j',k}, b_2)$.

Hence, all requirements of the second row of Table 10 are met and the statement can be applied. Further note that the strategy obtained after applying the switch has Property **Sv1** $_i$ due to the conditions described in Lemma 13. In particular, Corollary 9 also holds for $v > 1$.

After the application of $(e_{i,j,k}, b_2)$ (or $(g_i, F_{i,j})$ if it becomes improving), the tie-breaking rule determines which switch is applied next. Since $(d_{i,j,1-k}, F_{i,j})$ has an occurrence record of at least $m - 1$, another switch of the type $(e_{*,*,*}, b_2)$ is applied. But then, the same arguments used above can be used again. That is, we can apply some switch $(e_{i',j',k}, b_2)$, making $(d_{i',j',1-k'}, F_{i',j'})$ improving, and eventually making $(g_{i'}, F_{i',j'})$ improving as well. The switch $(g_{i'}, F_{i',j'})$ is applied immediately (if it becomes improving) whereas the other switch is not applied. Then, inductively, all remaining switches of the form $(e_{*,*,*}, b_2)$ are applied.

Let σ denote the strategy that is reached after applying the final improving switch $(e_{i,j,k}, b_2)$. We can then show that Property **Sv1** $_i$ holds for all $i \geq 2$. Furthermore, if $(g_1, F_{1,j})$ does not become improving, it is easy to prove that Property **Sv1** $_1$ holds at well.³ The same is true after the application of this switch if it becomes improving. In any case, we obtain a strategy that has Property **Sv1** $_i$ for all $i \in \{1, \dots, n\}$. As the proof is similar to the case $v = 1$, we omit it here.³

Thus, by Lemma 13 resp. the row 2 of Table 11, σ is a Phase-1-strategy for $b + 1$ with $\sigma \in \Lambda_i$.

We now prove that σ is a canonical strategy for $b + 1$ with $I_\sigma = \{(d_{i,j,k}, F_{i,j}) : \sigma(d_{i,j,k}) \neq F_{i,j}\}$. We begin by proving $I_\sigma = \{(d_{i,j,k}, F_{i,j}) : \sigma(d_{i,j,k}) \neq F_{i,j}\}$. Let $\sigma^{(5)}$ denote the Phase-5-strategy of Lemma 10 with $\sigma \in \Lambda_{\sigma^{(5)}}$. It is easy to verify that I_σ can be partitioned as

$$\begin{aligned}
 I_\sigma &= \{(d_{i,j,*}, F_{i,j}) : \sigma^{(5)}(e_{i,j,*}) = g_1\} \cup \{(d_{v,1-\beta_{v+1}^\sigma}, F_{v,1-\beta_{v+1}^\sigma})\} \\
 &\cup \left\{ e = (d_{i,1-\beta_{i+1}^\sigma}, F_{i,1-\beta_{i+1}^\sigma}) : i \in \{v + 1, \dots, m - 1\}, \beta_i^\sigma = 0, \phi^{\sigma_5}(e) = m - 1 \right\} \\
 &\cup \left\{ e = (d_{i,1-\beta_{i+1}^\sigma}, F_{i,1-\beta_{i+1}^\sigma}) : i < v, \phi^{\sigma^{(5)}}(e) \geq m - 1 \right\},
 \end{aligned}$$

if $b+1$ is not a power of two. A similar partition can be derived if $b+1$ is a power of two. In particular, $I_\sigma \subseteq \{(d_{i,j,k}, F_{i,j}) : \sigma(d_{i,j,k}) \neq F_{i,j}\}$. We prove that $e = (d_{i,j,k}, F_{i,j})$ implies $e \in I_\sigma$ if $\sigma(d_{i,j,k}) \neq F_{i,j}$.

If $\sigma^{(5)}(e_{i,j,k'}) = g_1$ for some $k' \in \{0, 1\}$, then $e \in I_\sigma$ as one of the cycle edges of $F_{i,j}$ is improving for $\sigma^{(5)}$ while the other becomes improving after applying $(e_{i,j,k'}, b_2)$. Thus let $\sigma^{(5)}(e_{i,j,*}) = b_2$, implying $\neg\bar{\sigma}^5(eg_{i,j})$. Then, by Lemma 10, $\bar{\sigma}^5(d_{i,j})$ or $\bar{\sigma}^5(eb_{i,j}) \wedge \neg\bar{\sigma}^5(eg_{i,j})$. In the first case, $\beta_i^{\sigma^{(5)}} = 1 \wedge \beta_{i+1}^{\sigma^{(5)}} = j$ by Lemma 10. But this implies $\bar{\sigma}(d_{i,j})$ since σ is a Phase-5-strategy for b and thus has Property Bac1_{*i*}. This however contradicts $\sigma(d_{i,j,k}) \neq F_{i,j}$. Hence, assume that $\bar{\sigma}^5(eb_{i,j}) \wedge \neg\bar{\sigma}^5(eg_{i,j})$. Then, by Lemma 10, $(i, j) \in S_1$. We distinguish three cases.

1. Let $(i, j) \in \{(i, 1 - \beta_{i+1}^\sigma) : i \leq v - 1\}$. If $\phi^{\sigma^{(5)}}(e) < m - 1$, then e was an improving switch of type 3 for $\sigma^{(5)}$ and thus applied during Phase 5. But this contradicts $\sigma(d_{i,j,k}) \neq F_{i,j}$ since no switch $(d_{*,*,*}, e_{*,*,*})$ is applied during Phase 5. This implies that we need to have $(i, j) \in \{(i, 1 - \beta_{i+1}^\sigma) : i \leq v - 1, \phi^{\sigma^{(5)}}(e) \geq m - 1\}$, hence $e \in I_\sigma$.
2. Let $(i, j) \in \{(i, 1 - \beta_{i+1}^\sigma) : i \in \{v + 1, \dots, m - 1\}, \beta_i^\sigma = 0\}$ which can only occur if $b + 1$ is not a power of 2. As proved when discussing $I_{\sigma^{(5)}}$, we then either have $\sigma_b(d_{i,j,k}) = F_{i,j}$, implying $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) \leq m - 1$ or $\sigma_b(d_{i,j,k}) \neq F_{i,j}$ and $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) = m - 1$. Consider the first case. If the inequality is strict, the switch was applied previously during Phase 5, yielding a contradiction. Otherwise, $(d_{i,j,k}, F_{i,j}) \in I_\sigma$. In the second case, the switch was applied during Phase 1, hence it was a switch of type 1 during Phase 5, also implying $(d_{i,j,k}, F_{i,j}) \in I_\sigma$.
3. Finally, let $i = v \wedge j = 1 - \beta_{v+1}^\sigma$ which only needs to be considered if $b + 1$ is not a power of 2. In this case we however have $e \in I_{\sigma^{(5)}}$, implying $e \in I_\sigma$.

Thus, $e \in I_\sigma$ in all case, proving the statement.

We now prove that σ is a canonical strategy for $b + 1$. Note that we heavily use that σ is a Phase-5-strategy for b . We thus refer to 4 for an overview over all properties that σ has. First, we have $\beta^\sigma = b + 1$. Thus, condition 1 follows since $\sigma(e_{*,*,*}) = b_2$ and $v > 1$. This also implies that conditions 2(a), 2(c), 3(a) and 3(b) are fulfilled as σ has Property Bac1* and Property Bac2*. Consider condition 2(b). Since $(b + 1)_i = 1$ implies that $F_{i, (b+1)_{i+1}}$ is closed, we prove that $F_{i, 1 - (b+1)_{i+1}}$ is not closed. Let $j := 1 - (b + 1)_{i+1}$. Then, by Lemma 10, $\sigma^{(5)}(d_{i,j,*}) = e_{i,j,*}$ and it suffices to prove $(d_{i,j,0}, F_{i,j}) \notin \mathfrak{A}_{\sigma^{(5)}}$. As such a switch is applied during $\sigma^{(5)} \rightarrow \sigma$ if and only if it is of type 3 by Corollary 11, we prove $\phi^{\sigma^{(5)}}(d_{i,j,0}, F_{i,j}) \geq m - 1$. This follows directly if $\mathbf{1}_{j=0} \text{mfn}(b, i + 1) + \mathbf{1}_{j=1} \text{mufn}(b, i + 1) = 0$ since this implies $\ell^b(i, j, k) \geq b$. Thus suppose that this term is not 0. Then, $\ell^b(i, j, k) = \left\lceil \frac{b + \sum(b,i) + 1 - k}{2} \right\rceil \geq \left\lceil \frac{b + 2 - k}{2} \right\rceil = \left\lfloor \frac{b + 1 - k}{2} \right\rfloor + 1$ since $b_1 = 1$ and by the choice of i and j . But this implies $\ell^b(i, j, 0) \geq m + 1$, hence $\phi^{\sigma^{(5)}}(d_{i,j,0}, F_{i,j}) \geq m$.

Consider condition 3(c) and let $j := 1 - (b + 1)_{i+1}$. It is easy to prove that σ has condition 3(c) by proving that $(b + 1)_i = 0$ and $F_{i,j}$ being closed imply $\Xi_\sigma^*(F_{i,j}) > \Xi_\sigma^*(F_{i, 1-j})$. The reason is that this implies that we need to have $\sigma(g_i) = F_{i,j}$ due to $I_\sigma = \{(d_{i,j,k}, F_{i,j}) : \sigma(d_{i,j,k}) \neq F_{i,j}\}$. Next, consider condition 3 (d) and consider a level i with $(b + 1)_i = 0$. Let $j := 0$ resp. $j := \beta_{i+1}^\sigma$ depending on whether $G_n = S_n$

resp. $G_n = M_n$. We prove that $\mathcal{E}_\sigma^*(F_{i,j}) \succ \mathcal{E}_\sigma^*(F_{i,1-j})$ if none of the two cycle centers is closed. In M_n , this either follows from Lemma 14 since σ has Property **Rel1** or by an easy calculation. In S_n , this follows since $\Omega(F_{i,0}) > \Omega(F_{i,1})$ and as these priorities are even.

Property **Usv1** implies that σ fulfills conditions 4 and 5 for all indices. Finally, consider condition 6 and let $i = \nu(b + 2)$, $j = \beta_{\nu(b+2)+1}^\sigma$. By the same argument used for condition 3(c), it suffices to prove $\phi^\sigma(d_{i,j,k}, F_{i,j}) \geq m - 1$ for both $k \in \{0, 1\}$. This however follows from $\nu(b + 2) = 1$. Hence, σ is a canonical strategy for $b + 1$. \square

Note that we implicitly proved the following corollary.

Corollary 13 *Let σ_{b+1} be the canonical strategy for $b + 1$ calculated by the Strategy Improvement resp. Policy Iteration Algorithm when starting with a canonical strategy σ_b fulfilling the canonical conditions as per Lemma 7. Then, Table 2 specifies the occurrence record of every improving switch applied until reaching σ_{b+1} , excluding switches $(g_*, F_{*,*})$, when interpreted for $b + 1$, and each such switch was applied once.*

It remains to prove that the canonical strategy σ_{b+1} fulfills the canonical conditions. By Corollary 13, it suffices to prove that it has Properties **(Or1)**_{*,*,*} to **(Or4)**_{*,*,*} and that Table 2 specifies the occurrence records of all edges that were not applied.

The following statement is required when discussing Properties **(Or1)**_{*,*,*} to **(Or4)**_{*,*,*}. It states that the occurrence record of the cycle edges of $F_{\nu(b+2), 1-(b+2)}$ are large if b is even and will be used repeatedly.

Lemma 15 *Let $b \in \mathcal{B}_n$ be even, $i := \nu(b+2)$ and $j := 1 - (b+2)_{i+1}$. If $b+2$ is a power of 2, then $\phi^{\sigma_b}(d_{i,j,*}, F_{i,j}) = m$. Otherwise, $\phi^{\sigma_b}(d_{i,j,0}, F_{i,j}) = m \wedge \phi^{\sigma_b}(d_{i,j,1}, F_{i,j}) = m - 1$. In any case, $\sigma_b(d_{i,j,k}) \neq F_{i,j}$ for both $k \in \{0, 1\}$.*

We now prove that the canonical strategy σ_{b+1} for $b + 1$ has Properties **(Or1)**_{*,*,*} to **(Or4)**_{*,*,*}.

Lemma 12 *Let σ_{b+1} denote the canonical strategy calculated by the Strategy Improvement resp. Policy Iteration Algorithm as described by Lemma 11. Then σ_{b+1} has Properties **(Or1)**_{*,*,*} to **(Or4)**_{*,*,*}.*

Proof To simplify notation, let $\sigma := \sigma_{b+1}$. We first prove that σ has Properties **(Or1)**_{*,*,*} to **(Or4)**_{*,*,*} and discuss Property **Or3**_{*,*,*} at the end. We only prove the statement for the case $\nu > 1$. The case $\nu = 1$ follows by using similar arguments.³

Consider Property **Or4** _{i,j,k} . We prove that any improving switch has an occurrence record of either m or $m - 1$ as $m = \lfloor \frac{b+1+1}{2} \rfloor$ due to $\nu > 1$. Any $e \in I_\sigma$ was either improving for $\sigma^{(5)}$ or became improving during Phase 5, i.e., when transitioning from $\sigma^{(5)}$ to σ . As in the proof of Lemma 11, all improving switches not applied during Phase 5 had an occurrence record of at least $m - 1$. More precisely, this was shown implicitly when giving the characterization of the improving switches. Also, the occurrence records of these edges are at most m , proving the statement for these edges. For improving switches that were unlocked during Phase 5, the statement follows by Corollary 12. Hence, σ has Property **Or4** _{i,j,k} for all indices.

We prove that σ has Property **Or2** $_{i,j,k}$ and Property **Or1** $_{i,j,k}$. Consider some indices i, j with $\beta_i^\sigma = 0 \vee \beta_{i+1}^\sigma \neq j$. We prove that $\sigma(d_{i,j,k}) = F_{i,j}$ is equivalent to $\phi^\sigma(d_{i,j,k}, F_{i,j}) = \ell^{b+1}(i, j, k) + 1$.

Let $\sigma(d_{i,j,k}) = F_{i,j}$. Then, since $\sigma^{(5)}(d_{i,j,k}) \neq F_{i,j}$ by the choice of i and j , the switch was applied during $\sigma_5 \rightarrow \sigma$. Consequently, it was not applied before Phase 5 as switches are applied at most once by Corollary 13. Thus, $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) = \phi^{\sigma^{(5)}}(d_{i,j,k}, F_{i,j}) < m - 1$. But this implies $\sigma_b(d_{i,j,k}) = F_{i,j}$ since the switch would have been applied in Phase 1 otherwise. Consequently, by Lemma 9

$$\begin{aligned} \phi^\sigma(d_{i,j,k}, F_{i,j}) &= \phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) + 1 = \ell^b(i, j, k) + 1 + 1 \\ &= \ell^{b+1}(i, j, k) + 1 \leq m - 1. \end{aligned}$$

Hence Property **Or1** $_{i,j,k}$ also holds. Now, let $\phi^\sigma(d_{i,j,k}, F_{i,j}) = \ell^{b+1}(i, j, k) + 1$. We prove that this implies $\sigma(d_{i,j,k}) = F_{i,j}$. We first observe that $\phi^\sigma(d_{i,j,k}, F_{i,j}) = \ell^{b+1}(i, j, k) + 1 \leq \left\lfloor \frac{b+1+1-k}{2} \right\rfloor$ implies $\ell^{b+1}(i, j, k) \leq \left\lfloor \frac{b-k}{2} \right\rfloor$. By Lemma 7, we thus need to have $\beta_{i+1}^\sigma = 1 - j$. Consider the case $b_i = 0 \wedge b_{i+1} \neq j$. Then, $t\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) = \min\left(\left\lfloor \frac{b+1-k}{2} \right\rfloor, \ell^b(i, j, k) + t_b\right)$ for some t_b feasible for b . Assume $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) \neq \ell^b(i, j, k) + t_b$ for all feasible parameters and note that this implies $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) = \left\lfloor \frac{b+1-k}{2} \right\rfloor$. Then $\phi^\sigma(d_{i,j,k}, F_{i,j}) < \ell^b(i, j, k) + 1$, implying

$$\begin{aligned} \ell^{b+1}(i, j, k) &= \ell^b(i, j, k) + 1 > \left\lfloor \frac{b+1-k}{2} \right\rfloor + 1 \\ &= \left\lfloor \frac{b+3-k}{2} \right\rfloor \geq \left\lfloor \frac{b+1+1-k}{2} \right\rfloor \end{aligned}$$

which is a contradiction. Consequently, $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) = \ell^b(i, j, k) + t_b$ for some feasible t_b . Assume $\phi^\sigma(d_{i,j,k}, F_{i,j}) = \ell^b(i, j, k)$. Then $\phi^\sigma(d_{i,j,k}, F_{i,j}) = \ell^{b+1}(i, j, k) + 1 = \ell^b(i, j, k) + 2 = \phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) + 2$, implying that the switch would have been applied twice during $\sigma_b \rightarrow \sigma$. This is a contradiction. The same contradiction follows if we assume $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) = \ell^b(i, j, k) - 1$. Hence, it holds that $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) = \ell^b(i, j, k) + 1$, implying $\sigma_b(d_{i,j,k}) = F_{i,j}$. Since $\ell^b(i, j, k) = \ell^{b+1}(i, j, k) - 1$, this also implies that the switch was indeed applied during the transition. However, $\sigma_b(d_{i,j,k}) = F_{i,j}$ implies that the switch was not applied during Phase 1 of that transition. But then it must have been applied in Phase 5, implying $\sigma(d_{i,j,k}) = F_{i,j}$.

We now show that the same holds if $b_i = 1$ and $b_{i+1} = j$, implying $i < v$. It is then easy to calculate that this yields $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) = \ell^{b+1}(i, j, k)$.³ Since $\phi^\sigma(d_{i,j,k}, F_{i,j}) = \ell^{b+1}(i, j, k) + 1$, this implies that the switch was applied during Phase 5 of $\sigma_b \rightarrow \sigma$. Consequently, $\sigma(d_{i,j,k}) = F_{i,j}$.

It remains to prove that σ has Property **Or3** $_{*,*,*}$. As a reminder, Property **Or3** $_{i,j,k}$ states that $\phi^\sigma(d_{i,j,k}, F_{i,j}) = \ell^{b+1}(i, j, k) - 1 \wedge \phi^\sigma(d_{i,j,k}, F_{i,j}) \neq \left\lfloor \frac{b+1+1-k}{2} \right\rfloor$ if and only if $b + 1$ is odd, $b + 2$ is not a power of 2, $i = v(b + 2)$, $j \neq (b + 2)_{i+1}$

and $k = 0$. We first prove the “if” part. Since $b + 1$ is odd, b is even. As $b + 2$ is not a power of b by assumption, $\phi^{\sigma_b}(d_{i,j,0}, F_{i,j}) = m$ and $\phi^{\sigma_b}(d_{i,j,1}, F_{i,j}) = m - 1$ as well $\sigma_b(d_{i,j,k}) \neq F_{i,j}$ for both $k \in \{0, 1\}$ by Claim 15. Now consider Phase 1 of $\sigma_b \rightarrow \sigma$. Then, $(d_{i,j,1}, F_{i,j})$ is applied in this phase by Corollary 1. Thus, by the tie breaking rule, $(d_{i,j,0}, F_{i,j})$ is not applied during Phase 1. Since no switch with an occurrence record of m is applied during Phase 5, the switch is also not applied during Phase 5. Consequently, $\phi^\sigma(d_{i,j,0}, F_{i,j}) = \phi^{\sigma_b}(d_{i,j,0}, F_{i,j}) = m = \lfloor \frac{b+1+1}{2} \rfloor - 1$ since $b + 1$ is odd. It thus remains to show $\ell^{b+1}(i, j, 0) = \lfloor \frac{b+1+1}{2} \rfloor$. Since $b + 1$ is odd, $v(b + 2) \neq v$ and $b_i = 0$. Hence, by Lemma 9, $\ell^{b+1}(i, j, 0) = \ell^b(i, j, 0) + 1 = \lfloor \frac{b}{2} \rfloor + 1 = \lfloor \frac{b+1+1}{2} \rfloor$. Thus, the “if” part is fulfilled. Now, the “only if” part can be shown by proving that each of the conditions is necessary by proving the following five statements:³

1. If $j := (b + 2)_{i+1}$, then either $\phi^\sigma(e) \neq \ell^{b+1}(i, j, k) - 1$ or $\phi^\sigma(e) = \lfloor \frac{b+1+1-k}{2} \rfloor$.
2. If $i \neq v(b + 2)$ and $j \neq (b + 2)_{i+1}$, then either $\phi^\sigma(e) \neq \ell^{b+1}(i, j, k) - 1$ or $\phi^\sigma(e) = \lfloor \frac{b+1+1-k}{2} \rfloor$.
3. If $b + 1 \in \mathcal{B}_n$ is even, $i = v(b + 2)$ and $j \neq (b + 2)_{i+1}$, then $\phi^\sigma(e) = \lfloor \frac{b+1+1-k}{2} \rfloor$.
4. If $b + 1 \in \mathcal{B}_n$ is odd, $i := v(b + 2)$, $j := 1 - (b + 2)_{i+1}$, $k \in \{0, 1\}$ and $b + 2$ is a power of two, then $\phi^\sigma(d_{i,j,k}, F_{i,j}) = \lfloor \frac{b+1+1-k}{2} \rfloor$.
5. If $b \in \mathcal{B}_n$ is even, $i = v(b + 2)$, $j \neq (b + 2)_{i+1}$, $k = 1$ and $b + 2$ is not a power of two, then $\phi^\sigma(e) = \lfloor \frac{b+1+1-k}{2} \rfloor$.

□

We now prove that Table 2 specifies the occurrence records with respect to the canonical strategy σ_{b+1} for $b + 1$ when it is interpreted for $b + 1$. Note that this then in particular implies Theorem 3.

Lemma 13 *Let σ_{b+1} be the canonical strategy for $b + 1$ calculated by the Strategy Improvement resp. Policy Iteration Algorithm when starting with a canonical strategy σ_b fulfilling the canonical conditions as per Lemma 7. Then, Table 2 specifies the occurrence record of all edges.*

Proof There are two types of edges. Each edge was either applied as improving switch when transitioning from σ_b to σ_{b+1} or was not applied as an improving switch. We already proved that Table 2 specifies the occurrence records of all improving switch that were applied, with the exception of switches $(g_*, F_{*,*})$. It thus suffices to consider these switches as well as switches that were not applied when transitioning from σ_b to σ_{b+1} .

As usual, we do not always explicitly state that σ has certain properties due to being a Phase-5-strategy and refer to Table 4 for an overview. We begin by identifying the edges that were not applied as improving switches and prove that their occurrence record is described by Table 2. We only prove the statement for $v > 1$ here.³ We first prove the statement for all edges that are not of the type $(d_{i,j,k}, F_{i,j})$.

1. Consider edges of the type $(b_i, *)$. Since $v > 1$, (b_i, b_{i+1}) for $i \in \{1, \dots, v-1\}$ and (b_v, g_v) were applied. Let $e = (b_i, b_{i+1})$ and $i \geq v$. Then $\phi^\sigma(e) = \text{fl}(b, i) - b_i = \text{fl}(b+1, i) - (b+1)_i$ since either $\text{fl}(b, i) = \text{fl}(b+1, i)$ and $b_i = (b+1)_{i+1}$ (if $i > v$) or $\text{fl}(b+1, i) = \text{fl}(b, i) + 1$ and $b_i = 0$, $(b+1)_i = 1$ (if $i = v$). Let $e = (b_i, g_i)$ for $i \neq v$. Then, by Lemma 9, $\phi^\sigma(e) = \text{fl}(b, i) = \text{fl}(b+1, i)$.
2. Consider some edge $(g_i, F_{i,j})$ that was not applied during $\sigma_b \rightarrow \sigma$. Then, the upper bound remains valid as it can only increase.
3. Consider some vertex $s_{i,j}$. Since $v > 1$, the edges $(s_{v-1,1}, h_{v-1,1}), (s_{v-1,0}, b_1)$ as well as the edges $(s_{i,0}, h_{i,0}), (s_{i,1}, b_1)$ for $i \in \{1, \dots, v-2\}$ were switched. It thus suffices to consider indices $i \geq v$. This implies $\phi^\sigma(s_{i,j}, b_1) = \text{fl}(b, i+1) - j \cdot b_{i+1} = \text{fl}(b+1, i+1) - j \cdot (b+1)_{i+1}$ by the choice of i . Similarly, $\phi^\sigma(s_{i,j}, h_{i,j}) = \text{fl}(b, i+1) - (1-j)b_{i+1} = \text{fl}(b+1, i+1) - (1-j)(b+1)_{i+1}$.
4. For $e = (e_{i,j,k}, g_1)$, Table 2 implies $\phi^\sigma(e_{i,j,k}, g_1) = \left\lfloor \frac{b}{2} \right\rfloor = \left\lfloor \frac{b+1}{2} \right\rfloor$ since $v > 1$.
5. Consider some $e = (d_{i,j,k}, e_{i,j,k})$. We need to prove $\phi^\sigma(e) \leq \phi^\sigma(e_{i,j,k}, g_1) = \left\lfloor \frac{b}{2} \right\rfloor = \left\lfloor \frac{b+1}{2} \right\rfloor$ since b is odd. But this follows since $\phi^\sigma(e) \leq \phi^{\sigma_b}(e) + 1 \leq \left\lfloor \frac{b}{2} \right\rfloor + 1 = \left\lfloor \frac{b+2}{2} \right\rfloor = \left\lfloor \frac{b+1}{2} \right\rfloor$.

Consider some $e = (d_{i,j,k}, F_{i,j})$ that was not switched. We distinguish the following cases.

1. Let $(b_i = 1 \wedge b_{i+1} = j)$ and $((b+1)_i = 1 \wedge (b+1)_{i+1} = j)$. Then, since any intermediate strategy had Property **Bac1** _{i} , $F_{i,j}$ was always closed during $\sigma_b \rightarrow \sigma$. Thus $i \neq v$, implying $\text{mfn}(b, i, \{(i+1, j)\}) = \text{mfn}(b+1, i, \{(i+1, j)\})$. Therefore, $\phi^\sigma(e)$ is described by Table 2.
2. Let $(b_i = 1 \wedge b_{i+1} = j)$ and $(b+1)_i = 0$, implying $i < v$. Then bit $i+1$ also switched, so $(b+1)_i = 0 \wedge (b+1)_{i+1} \neq j$. Consequently, e was not switched during Phase 1 since $F_{i,j}$ was closed with respect to any intermediate strategy due to Property **Bac1** _{i} . It is however possible that such a switch is applied during Phase 5. Since $i \leq v-1$, this switch is applied if and only if $\phi^{\sigma_b}(e) < m-1$. We may thus assume $\phi^{\sigma_b}(e) \geq m-1 = \left\lfloor \frac{b-1}{2} \right\rfloor$ and only need to consider e if $\left\lfloor \frac{\text{mfn}(b, i, \{(i+1, j)\}) + 2 - k}{2} \right\rfloor \geq \left\lfloor \frac{b-1}{2} \right\rfloor$. This inequality holds if and only if one of the following three cases applies:

- $\text{mfn}(b, i, \{(i+1, j)\}) + 2 - k \geq b - 1$
- $\text{mfn}(b, i, \{(i+1, j)\}) + 2 - k$ is even and $\text{mfn}(b, i, \{(i+1, j)\}) + 2 - k = b - 2$.
- $\text{mfn}(b, i, \{(i+1, j)\}) + 2 - k$ is odd and $\text{mfn}(b, i, \{(i+1, j)\}) + 2 - k = b$.

These assumptions can only hold if $i \in \{1, 2\} \vee (i = 3 \wedge k = 0)$. It thus suffices to consider three more cases. For $i = 1$, we obtain $\ell^{b+1}(i, j, k) = \left\lfloor \frac{b+1-1+1-k}{2} \right\rfloor = \left\lfloor \frac{b+1-k}{2} \right\rfloor = \phi^\sigma(e)$. Similarly, for $i = 2$, we obtain $\ell^{b+1}(i, j, k) = \left\lfloor \frac{b+1-2+1-k}{2} \right\rfloor = \left\lfloor \frac{b-k}{2} \right\rfloor = \left\lfloor \frac{b+1-k}{2} \right\rfloor = \phi^\sigma(e)$. Finally, for $i = 3$ and $k = 0$, we obtain $\ell^{b+1}(i, j, k) = \left\lfloor \frac{b+1-4+1-k}{2} \right\rfloor = \left\lfloor \frac{b-2-k}{2} \right\rfloor = \phi^\sigma(e)$. Hence, for all three cases, choosing the parameter $t_{b+1} = 0$ yields the desired characterization of $\phi^\sigma(e)$.

3. Let $(b_i = 0 \wedge b_{i+1} \neq j)$ and $((b + 1)_i = 0 \wedge (b + 1)_{i+1} \neq j)$, implying $i > v$. First assume $\mathbf{1}_{j=0}\text{mfn}(b, i + 1) + \mathbf{1}_{j=1}\text{mufn}(b, i + 1) = 0$. Then $\ell^b(i, j, k) \geq b$ by Lemma 7, implying $\phi^{\sigma_b}(e) = \lfloor \frac{b+1-k}{2} \rfloor$. Since b is odd, $\lfloor \frac{b+1-1}{2} \rfloor < m$. Hence, $(d_{i,j,1}, F_{i,j})$ was applied during Phase 1 of $\sigma_b \rightarrow \sigma$ and $e = (d_{i,j,0}, F_{i,j}) \notin \mathfrak{A}_{\sigma_b}^\sigma$. Thus, since $\ell^{b+1}(i, j, k) \geq b + 1$ by the choice of i , choosing $t_{b+1} = 0$ yields the desired characterization.
 Now assume $\mathbf{1}_{j=0}\text{mfn}(b, i + 1) + \mathbf{1}_{j=1}\text{mufn}(b, i + 1) \neq 0$, implying $i < m = \max\{i : \beta_i^\sigma = 1\}$. Using $i > v \geq 2$, it is easy to verify that this yields $\ell^b(i, j, k) \leq \lfloor \frac{b+1-k}{2} \rfloor - 1$.³ If $\sigma_b(d_{i,j,k}) \neq F_{i,j}$, this implies $\phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) \leq \ell^b(i, j, k) \leq \lfloor \frac{b+1-k}{2} \rfloor - 1$. Then, by Corollary 1 the switch was applied during Phase 1. We may hence assume $\sigma_b(d_{i,j,k}) = F_{i,j}$, implying that we have $\phi^{\sigma_b}(e) = \ell^b(i, j, k) + 1 \leq \lfloor \frac{b+1-k}{2} \rfloor$ as well as $\phi^{\sigma_b}(e) \leq m - 1$ by Property Or1 _{i,j,k} . As we assume $e \notin \mathfrak{A}_{\sigma_b}^\sigma$, it suffices to consider the case $\phi^\sigma(e) = \phi^{\sigma_b}(e) = m - 1$ since e is applied during Phase 5 otherwise (see Corollary 12). Since $\ell^{b+1}(i, j, k) = \ell^b(i, j, k) + 1$ by Lemma 9, choosing $t_{b+1} = 0$ yields the desired characterization.
4. Let $(b_i = 0 \wedge b_{i+1} \neq j)$ and $((b + 1)_i = 1 \wedge (b + 1)_{i+1} \neq j)$, i.e., $i = v$. The statement follows by the same argument used earlier if $\mathbf{1}_{j=0}\text{mfn}(b, i + 1) + \mathbf{1}_{j=1}\text{mufn}(b, i + 1) = 0$. Hence let $\mathbf{1}_{j=0}\text{mfn}(b, i + 1) + \mathbf{1}_{j=1}\text{mufn}(b, i + 1) \neq 0$, implying $\ell^b(v, j, k) = \lfloor \frac{b+1-k}{2} \rfloor$. Since σ_b is a canonical strategy for b , we have $\sigma_b(d_{i,j,k}) \neq F_{i,j}$. Assume $\phi^{\sigma_b}(e) = \ell^b(i, j, k)$. Then $\phi^\sigma(e) = \lfloor \frac{b+1-k}{2} \rfloor$ and the same arguments used in the third case can be used to show the statement. Assume $\phi^\sigma(e) = \ell^b(i, j, k) - 1$. Then $\phi^{\sigma_b}(e) = m - 1$ since we need to have $k = 0$ by Property Or3 _{i,j,k} . But this implies that e was switched during Phase 1 and that we do not need to consider it here.
5. Finally, let $(b_i = 0 \wedge b_{i+1} = j)$. We only need to consider the case $(b + 1)_i = 0 \wedge (b + 1)_{i+1} = j$, implying $i > v$. If $\mathbf{1}_{j=0}\text{mfn}(b, i + 1) + \mathbf{1}_{j=1}\text{mufn}(b, i + 1)$, the statement follows by the same arguments made earlier. Otherwise, we can also use the previous same arguments since $\ell^b(i, j, k) > \lfloor \frac{b+1-k}{2} \rfloor$ implies $\phi^\sigma(e) = \lfloor \frac{b+1-k}{2} \rfloor$.

It remains to investigate edges $(g_*, F_{*,*})$. We prove that Table 2 specifies their occurrence records by inductively proving the following statement: *Let $\sigma_b \in \mathcal{A}_i$ be a canonical strategy for $b \in \mathcal{B}_n$ calculated by the Strategy Improvement Algorithm. Then $\phi^{\sigma_b}(g_i, F_{i,j}) \leq \min_{k \in \{0,1\}} \phi^{\sigma_b}(d_{i,j,k}, F_{i,j})$.*

For simplicity, we interpret the improving switch $(g_v, F_{v,b_{v+1}})$ that might technically be applied at the beginning of Phase 2 or 3 as an improving switch that is applied during Phase 1. We prove the statement via induction on b . We briefly discuss how the statement can be proven for $i \neq 1$ and do not discuss the formal detail or the case $i = 1$.³ As we consider several earlier strategies during the inductive step, we show that the statement holds for all $b \leq 2^i$.

Thus, let $b \leq 2^i =: \tilde{b}$ and consider some edge $e := (g_i, F_{i,j})$. Then $\tilde{b}_{i+1} = 1$ and $b'_{i+1} = 0$ for all $b' \leq \tilde{b}$. We prove that e was applied at most once when transitioning

from ι to $\sigma_{\tilde{b}}$ and that this application can only happen during $\sigma_{\tilde{b}-1} \rightarrow \sigma_{\tilde{b}}$. The statement then follows since it is easy to verify that the occurrence records of the cycle edges of $F_{i,j}$ are both at least one.

Since $\iota(g_i) = 0$, e cannot have been applied during Phase 1 of any transition encountered during the sequence $\iota \rightarrow \sigma_{\tilde{b}}$ as the choice of \tilde{b} implies that there is no $b' \leq \tilde{b}$ with $b'_i = 1 \wedge b'_{i+1} = 1$. It is also easy to show that this implies that it cannot happen that the cycle center $F_{i,j}$ was closed during Phase 1 if $j = 1 - b'_{i+1}$. The switch $(g_i, F_{i,j})$ can thus only have been applied during some Phase 5. However, since $\iota(g_i) = 0$ and due to the choice of \tilde{b} , this can only happen when transitioning from $\sigma_{\tilde{b}-1}$ to $\sigma_{\tilde{b}}$.

Thus, the statement holds for σ_b with $b \leq 2^i$. Now, assume that it holds for all $b' < b$ where $b > 2^i$. We prove that the statement also holds for σ_b . Fix some edge $e := (g_i, F_{i,j})$ and consider the strategy σ_{b-1} . We begin by arguing that several cases do not need to be considered.

First of all, every improving switch is applied at most once per transition. The statement thus follows by the induction hypothesis if $\min_{k \in \{0,1\}} \phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) > \min_{k \in \{0,1\}} \phi^{\sigma_{b-1}}(d_{i,j,k}, F_{i,j})$. We thus assume

$$\min_{k \in \{0,1\}} \phi^{\sigma_b}(d_{i,j,k}, F_{i,j}) = \min_{k \in \{0,1\}} \phi^{\sigma_{b-1}}(d_{i,j,k}, F_{i,j}). \tag{3}$$

Similarly, if e is not applied during $\sigma_{b-1} \rightarrow \sigma_b$, then the statement also follows by the induction hypothesis. We thus assume $e \in \mathcal{A}_{\sigma_{b-1}}^{\sigma_b}$.

These observations give first structural insights on $b - 1$ and b . First, if $b_i = 1 \wedge (b - 1)_i = 1$, then it is not possible to apply e during $\sigma_{b-1} \rightarrow \sigma_b$. Second, if $b_i = 1 \wedge (b - 1)_i = 0$, then $i = \nu(b)$. By Definition 2 resp. 4, both cycle centers of level $\nu(b)$ are open for σ_{b-1} . Hence, Corollary 2 implies that $F_{i,j}$ is closed during $\sigma_{b-1} \rightarrow \sigma_b$ by applying both switches $(d_{i,j,0}, F_{i,j})$ and $(d_{i,j,1}, F_{i,j})$. But then, Eq. (3) is not fulfilled and the statement follows. This implies that it suffices to consider the case $b_i = 0$.

It can then be show that these assumptions imply that the occurrence record of the edges $(d_{i,j,*}, F_{i,j})$ is “large”.³ To be precise, that Eq. (3), $e \in \mathcal{A}_{\sigma_{b-1}}^{\sigma_b}$ and $b_i = 0$ imply

$$\min_{k \in \{0,1\}} \phi^{\sigma_{b-1}}(d_{i,j,k}, F_{i,j}) \geq \left\lfloor \frac{b}{2} \right\rfloor - 1. \tag{4}$$

It then suffices to prove $\phi^{\sigma_{b-1}}(g_i, F_{i,j}) < \left\lfloor \frac{b}{2} \right\rfloor - 1$ to complete the proof.

We begin by proving that we cannot have $(b - 1)_i = 1$. Let, for the sake of contradiction, $(b - 1)_i = 1$. Then, as $b_i = 0$, we have $i < \nu(b)$ and consequently $(b - 1)_{i+1} \neq b_{i+1}$. It further implies that b is even. Then, since $e \in \mathcal{A}_{\sigma_{b-1}}^{\sigma_b}$ by assumption, it was applied during Phase 5 of $\sigma_{b-1} \rightarrow \sigma_b$. This implies $j = 0$ in S_n resp. $j = b_{i+1} = 1 - (b - 1)_{i+1}$ in M_n . Consider M_n first. Then, since $(b - 1)_i = 1 \wedge j = 1 - (b - 1)_{i+1}$ imply $\ell^{b-1}(i, j, k) \geq \left\lfloor \frac{b-k}{2} \right\rfloor + 1$ by Lemma 7, we obtain $\phi^{\sigma_{b-1}}(d_{i,j,k}, F_{i,j}) = \left\lfloor \frac{b-k}{2} \right\rfloor$.

In addition, since $\phi^{\sigma_{b-1}}(d_{i,j,k}, F_{i,j}) \neq \ell^{b-1}(i, j, k) + 1$ for both $k \in \{0, 1\}$, $F_{i,j}$ is then open with respect to σ_{b-1} by Property Or2 $_{i,j,*}$. This implies that $(d_{i,j,1}, F_{i,1})$ is applied during Phase 1 of $\sigma_{b-1} \rightarrow \sigma_b$. But then $\min_{k \in \{0,1\}} \phi^{\sigma_{b-1}}(d_{i,j,k}, F_{i,j}) < \min_{k \in \{0,1\}} \phi^{\sigma_b}(d_{i,j,k}, F_{i,j})$, contradicting our assumption. Now consider the case $G_n = S_n$. If $j = 0 = 1 - (b - 1)_{i+1}$, then the statement follows by the same arguments. This is the case if and only if $i < \nu(b) - 1$, so let $i = \nu(b) - 1$. By Definition 2, this implies $\sigma_{b-1}(g_i) = F_{i,0}$. Since $F_{i,0}$ is then closed during Phase 1 of the transition $\sigma_{b-1} \rightarrow \sigma_b$ and since $(g_i, F_{i,1})$ cannot be applied during Phase 5 if $G_n = S_n$, this is a contradiction.

It thus suffices to consider the case $(b - 1)_i = 0$. To simplify notation, we denote the binary number obtained by subtracting 1 from a binary number (b'_n, \dots, b'_1) by $[b'_n, \dots, b'_1] - 1$. Then, b and $b - 1$ can be represented as

$$\begin{aligned}
 b &= (b_n, \dots, b_{i+1}, \mathbf{0}, b_{i-1}, \dots, b_1), \\
 b - 1 &= (b_n, \dots, b_{i+1}, \mathbf{0}, [b_{i-1}, \dots, b_1] - 1)
 \end{aligned}$$

where bit i is marked in **bold**. The idea of the proof is now the following. We define two smaller numbers that are relevant for the application of $(g_i, F_{i,j})$. We use these numbers and the induction hypothesis to prove that even if $(g_i, F_{i,j})$ was applied during (nearly) all of these transitions, the bound that we claim still holds.

We thus define $\tilde{b} = ([b_n, \dots, b_{i+1}] - 1, \mathbf{1}, 1, \dots, 1)$ and $\tilde{b} = ([b_n, \dots, b_{i+1}] - 1, \mathbf{1}, 0, \dots, 0)$ where bit i is again marked in **bold**. Note that these numbers are well-defined since $b \geq 2^i$.

Consider \tilde{b} . Let $\mathfrak{N}(\tilde{b}, b - 1)$ denote the number of applications of $(g_i, F_{i,j})$ when transitioning from $\sigma_{\tilde{b}}$ to σ_{b-1} . Then, since $b'_i = 1$ for all $b' \in \{\tilde{b}, \dots, \tilde{b}\}$, we have $\mathfrak{N}(\tilde{b}, b - 1) = \mathfrak{N}(\tilde{b}, b - 1)$. We thus can describe the occurrence record of $(g_i, F_{i,j})$ as

$$\begin{aligned}
 \phi^{\sigma_{b-1}}(g_i, F_{i,j}) &= \mathfrak{N}(0, b - 1) = \mathfrak{N}(0, \tilde{b}) + \mathfrak{N}(\tilde{b}, b - 1) \\
 &= \phi^{\sigma_{\tilde{b}}}(g_i, F_{i,j}) + \mathfrak{N}(\tilde{b}, b - 1).
 \end{aligned}$$

Our goal is to bound the two terms on the right-hand side. Due to the induction hypothesis, the first term can be bounded by $\lfloor \frac{\tilde{b}}{2} \rfloor$. Since every improving switch is applied at most once per transition by Corollary 13, we have $\mathfrak{N}(\tilde{b}, b - 1) \leq (b - 1) - \tilde{b}$. However, this upper bound is not strong enough. It can however be proven that $(g_i, F_{i,j})$ was not applied during all transition from \tilde{b} to $b - 1$. More precisely, $\mathfrak{N}(\tilde{b}, b - 1) \leq (b - 1) - \tilde{b} - 1$. Combining these results and using $\tilde{b} = b - \sum (b, i) - 1$ and $\tilde{b} = b - \sum (b, i) - 2^{i-1}$ yields the statement.³

We can now combine the previous arguments to prove Lemma 5. Note that this implies Lemma 6.

Lemma 5 *Let $\sigma_b \in \Lambda_i$ be a canonical strategy for b fulfilling the canonical conditions. Let $\sigma \in \Lambda_{\sigma_b}$ be a strategy obtained by applying a sequence of improving switches to σ_b . Let σ have the properties of row k of Table 6 and let I_σ be described by row k*

of Table 7 for some $k \in \{1, \dots, 5\}$. Then, applying improving switches according to Zadeh's pivot rule and our tie-breaking rule produces a strategy σ' that is described by the next feasible rows of Tables 6 and 7.

Proof By Lemmas 7, 8, 9 and 10, it suffices to consider the case $k = 5$. By Lemma 11, applying improving switches to a canonical strategy σ_b for b produces a canonical strategy σ_{b+1} for $b + 1$. It suffices to prove that σ_{b+1} fulfills the canonical conditions.

By Corollary 13, Table 2 correctly specifies the occurrence record of all edges when interpreted for $b + 1$. By Lemma 12, σ_{b+1} has Properties (Or1)_{*,*,*} to (Or4)_{*,*,*}. By Corollary 13, each improving switch was applied at most once when transitioning from $\sigma_b \rightarrow \sigma_{b+1}$. Consequently, σ_{b+1} fulfills the canonical conditions. \square

This now enables us to prove our final theorem.

Theorem 4 *Applying the Strategy Improvement Algorithm with Zadeh's pivot rule and the strategy-based tie-breaking rule described in Definition 7 on the game G_n of size $O(n)$ needs at least 2^n iterations when using ι as the initial player 0 strategy.*

Proof This follows by Lemma 4 and by applying Lemma 6 iteratively. \square

References

1. Avis, D., Chvátal, V.: Notes on Bland's pivoting rule. In: Polyhedral Combinatorics: Dedicated to the memory of D.R. Fulkerson, pp. 24–34. Springer, Heidelberg (1978)
2. Avis, D., Friedmann, O.: An exponential lower bound for Cunningham's rule. *Math. Program.* **161**(1), 271–305 (2017)
3. Bertsimas, D., Vempala, S.: Solving convex programs by random walks. *J. of the ACM* **51**(4), 540–556 (2004)
4. Cunningham, W.H.: Theoretical properties of the network simplex method. *Math. of Oper. Res.* **4**(2), 196–208 (1979)
5. Dantzig, G.B.: *Linear Programming and Extensions*. Princeton University Press (1963)
6. Disser, Y., Friedmann, O., Hopp, A.V.: An Exponential Lower Bound for Zadeh's pivot rule (2020). Full version. Available at <https://arxiv.org/pdf/1911.01074.pdf>
7. Disser, Y., Hopp, A.V.: On Friedmann's subexponential lower bound for Zadeh's pivot rule. In: Proceedings of the 20th Conference on Integer Programming and Combinatorial Optimization (IPCO) pp. 168–180 (2019)
8. Disser, Y., Skutella, M.: The simplex algorithm is NP-mighty. *ACM Trans. on Algorithms* **15**(1), 5–19 (2018)
9. Dunagan, J., Vempala, S.: A simple polynomial-time rescaling algorithm for solving linear programs. *Math. Program.* **114**(1), 101–114 (2007)
10. Fearnley, J.: Exponential lower bounds for policy iteration. In: Proceedings of the 37th International Colloquium on Automata, Languages, and Programming (ICALP) pp. 551–562 (2010)
11. Fearnley, J., Savani, R.: The complexity of the simplex method. In: Proceedings of the 47th Annual ACM on Symposium on Theory of Computing (STOC) pp. 201–208 (2015)
12. Friedmann, O.: An exponential lower bound for the parity game strategy improvement algorithm as we know it. In: Proceedings of the 24th Annual IEEE Symposium on Logic In Computer Science (LICS) pp. 145–156 (2009)
13. Friedmann, O.: An exponential lower bound for the latest deterministic strategy iteration algorithms. *Logical Methods in Computer Sci.* **7**(3:19), 1–42 (2011)
14. Friedmann, O.: Exponential Lower Bounds for Solving Infinitary Payoff Games and Linear Programs. PhD thesis, University of Munich (2011)

15. Friedmann, O.: 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 Combinatorial Optimization (IPCO) pp. 192–206 (2011)
16. Friedmann, O., Hansen, T.D., Zwick, U.: A subexponential lower bound for the random facet algorithm for parity games. In: Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA) pp. 202–216 (2011)
17. Friedmann, O., Hansen, T.D., Zwick, U.: Subexponential lower bounds for randomized pivoting rules for the simplex algorithm. In: Proceedings of the 43rd Annual ACM Symposium on Theory of Computing (STOC) pp. 283–292 (2011)
18. Friedmann, O., Hansen, T.D., Zwick, U.: Errata for: A subexponential lower bound for the random facet algorithm for parity games (2014). CoRR, abs/1410.7871
19. Friedmann, O., Lange, M.: PGSolver, version 4.1 (June 2017). Available at <https://github.com/tcsprojects/pgsolver>
20. Goldfarb, D., Sit, W.Y.: Worst case behavior of the steepest edge simplex method. *Discrete Applied Math.* **1**(4), 277–285 (1979)
21. Hansen, T.D.: Worst-case Analysis of Strategy Iteration and the Simplex Method. PhD thesis, Department Office Computer Science, Aarhus University (2012)
22. Hansen, T.D., Zwick, U.: An improved version of the random-facet pivoting rule for the simplex algorithm. In: Proceedings of the 47th Annual ACM on Symposium on Theory of Computing (STOC) pp. 209–218 (2015)
23. Howard, R.A.: *Dynamic Programming and Markov Processes*. MIT Press (1960)
24. Animation of the counter for 3 and 4 levels (2019). <https://oliverfriedmann.com/downloads/add/zadeh-exponential-animation-3.pdf> <https://oliverfriedmann.com/downloads/add/zadeh-exponential-animation-4.pdf>
25. Kalai, G.: A subexponential randomized simplex algorithm. In: Proceedings of the 24th Annual ACM Symposium on Theory of Computing (STOC) pp. 475–482 (1992)
26. Kalai, G.: Linear programming, the simplex algorithm and simple polytopes. *Math. Program.* **79**, 217–233 (1997)
27. Karmarkar, N.: A new polynomial-time algorithm for linear programming. In: Proceedings of the 16th Annual ACM Symposium on Theory of Computing (STOC), pp. 302–311 (1984)
28. Kelner, J.A., Spielman, D.A.: A randomized polynomial-time simplex algorithm for linear programming. In: Proceedings of the 38th Annual ACM Symposium on Theory of Computing (STOC) pp. 51–60 (2006)
29. Khachiyan, L.G.: Polynomial algorithms in linear programming. *USSR Comput. Math. and Math. Phys.* **20**(1), 53–72 (1980)
30. Klee, V., Minty, G.J.: How good is the simplex algorithm? In: *Inequalities III*, pp. 159–175. Academic Press, New York (1972)
31. Küsters, R.: Memoryless determinacy of parity games. In: Grädel, Erich, Thomas, Wolfgang, Wilke, Thomas (eds.) *Automata Logics, and Infinite Games*. Springer, Berlin Heidelberg (2002)
32. Matoušek, J., Sharir, M., Welzl, E.: A subexponential bound for linear programming. *Algorithmica* **16**(4–5), 498–516 (1996)
33. Matoušek, J., Szabó, T.: Random edge can be exponential on abstract cubes. *Adv. in Math.* **204**(1), 262–277 (2006)
34. Puterman, M.L.: *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons Inc, New York (2005)
35. Santos, F.: A counterexample to the Hirsch conjecture. *Annals of Math.* **176**(1), 383–412 (2012)
36. Smale, S.: Mathematical problems for the next century. *The Math. Intell.* **20**(2), 7–15 (1998)
37. Solms, Felix: Implementing a sink game solver. Master's thesis, TU Darmstadt (2021). implementation available at <https://www2.mathematik.tu-darmstadt.de/~disser/data/VisualSinkGameSolver.zip>
38. Szpilrajn, E.: Sur l'extension de l'ordre partiel. *Fundamenta Math.* **16**, 386–389 (1930)
39. Thomas, A.: Exponential lower bounds for history-based simplex pivot rules on abstract cubes. In: Proceedings of the 25th Annual European Symposium on Algorithms (ESA) pp. 69(14) (2017)
40. Tijms, H.C.: *A First Course in Stochastic Models*. John Wiley & Sons, Ltd (2004)
41. Todd, M.J.: An improved Kalai-Kleitman bound for the diameter of a polyhedron. *SIAM J. on Discrete Math.* **28**(4), 1944–1947 (2014)

42. Vöge, J., Jurdziński, M.: A discrete strategy improvement algorithm for solving parity games. In: *Computer Aided Verification*, pp. 202–215 (2000)
43. Zadeh, N.: What is the worst case behavior of the simplex algorithm? Technical Report 27. Departments of Operations Research, Stanford (1980)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.