Weak CSR expansions and transience bounds in max-plus algebra

Glenn Merlet [a], Thomas Nowak [b,1], Sergeĭ Sergeev [c,*,2]

[a] *Université d'Aix-Marseille, CNRS, IML, 13288 Marseille, France*
[b] *École Normale Supérieure, Département d'Informatique, 75230 Paris, France*
[c] *University of Birmingham, School of Mathematics, Edgbaston B15 2TT, UK*

A R T I C L E   I N F O

A B S T R A C T

This paper aims to unify and extend existing techniques for deriving upper bounds on the transient of max-plus matrix powers. To this aim, we introduce the concept of weak CSR expansions: $A^t = CS^t R \oplus B^t$. We observe that most of the known bounds (implicitly) take the maximum of (i) a bound for the weak CSR expansion to hold, which does not depend on the values of the entries of the matrix but only on its pattern, and (ii) a bound for the $CS^t R$ term to dominate.

To improve and analyze (i), we consider various cycle replacement techniques and show that some of the known bounds for indices and exponents of digraphs apply here. We also show how to make use of various parameters of digraphs. To improve and analyze (ii), we introduce three different kinds of weak CSR expansions (named after Nachtigall, Hartman–Arguelles, and Cycle Threshold). As a result, we obtain a collection of bounds, in general incomparable to one another, but better than the bounds found in the literature.

---

* Corresponding author.
  E-mail addresses: glenn.merlet@univ-amu.fr (G. Merlet), thomas.nowak@ens.fr (T. Nowak), sergiej@gmail.com (S. Sergeev).

1. Introduction

Max-plus algebra is a version of linear algebra developed over the max-plus semiring, which is the set $\mathbb{R}_{\max} = \mathbb{R} \cup \{-\infty\}$ equipped with the multiplication $a \otimes b = a + b$ and the addition $a \oplus b = \max(a, b)$. This semiring has zero $\mathbf{0} := -\infty$ (neutral with respect to $\oplus$) and unity $\mathbf{1} := 0$ (neutral with respect to $\otimes$), and each element μ except for $\mathbf{0}$ has an inverse $\mu^- := -\mu$ satisfying $\mu \otimes \mu^- = \mu^- \otimes \mu = \mathbf{1}$. Taking powers of scalars in $\mathbb{R}_{\max}$ means ordinary multiplication: $\lambda^{\otimes t} := t \cdot \lambda$.

The max-plus arithmetic is extended to matrices in the usual way, so that $(AB)_{ij} = \bigoplus_k a_{ik} \otimes b_{kj} = \max_k (a_{ik} + b_{kj})$ for matrices $A = (a_{ij})$ and $B = (b_{ij})$ of compatible sizes. In this paper, all matrix multiplications are to be understood in the max-plus sense. For multiplication by a scalar and for taking powers of scalars we will write the sign $\otimes$ explicitly, while for the matrix multiplication it will be always omitted.

Historically, max-plus algebra first appeared to analyze production systems driven by the dynamics

$$x_i(k+1) = \max_j \big(x_j(k) + a_{ij} \big). \tag{1}$$

Thus, repeated application of matrix $A = (a_{ij})$ in max-plus algebra to an initial vector $x(0)$ computes the vectors $x(k)$. Here $x(k)$ is typically a vector consisting of n real components, expressing the times of certain events happening during the kth production cycle. According to dynamics (1), event i has to wait until all the preceding events j happen and the necessary time delays a_{ij} have passed, so that event i can then occur as early as possible. Such situation is usual in train scheduling, working plan analysis, and synchronization of multiprocessor systems [3,5,15]. Recently, Charron-Bost et al. [9] have shown that also the behavior of link reversal algorithms used for routing, scheduling, resource allocation, leader election, and distributed queuing can be described by a recursion of the form (1).

In this paper, we investigate the sequence of max-plus matrix powers $A^t = \overbrace{AAA \cdots A}^{t \text{ times}}$. Cohen et al. [10] proved that this sequence eventually exhibits a periodic regime whenever A is *irreducible*, i.e., whenever the digraph $\mathcal{D}(A)$ described by A is strongly connected: there exists a positive integer γ and a nonnegative integer T such that

$$\forall t \geq T : \quad A^{t+\gamma} = \lambda^{\otimes \gamma} \otimes A^t, \tag{2}$$

where $\lambda = \lambda(A)$ is the unique max-plus eigenvalue of A. This result is known as the Cyclicity Theorem. The smallest T that can be chosen in (2) is called the *transient* of A; we denote it by $T(A)$.

Since it satisfies $x(t) = A^t v$, every max-plus linear dynamical system, i.e., every sequence $x(t)$ satisfying (1) is periodic in the same sense whenever A is irreducible. Its transient $T(A, v)$ in general depends on v and is always upper-bounded by $T(A)$.

Bounds on the transients were obtained by Hartmann and Arguelles [14], Bouillard and Gaujal [4], Soto y Koelemeijer [27], Akian et al. [2], and Charron-Bost et al. [8]. Those bounds are incomparable because they depend on different parameters of A or assume different hypotheses. However they all appear, at least in the proofs, as the maximum of a first bound independent of the values of the entries of A and a second bound taking those values into account. The first motivation for this paper was to find a common ground for these bounds in order to understand, unify, combine, and improve them.

Schneider [20] observed that the Cyclicity Theorem can be written in the form of a *CSR expansion*, which was subsequently formulated by Sergeev [23]: there exists a nonnegative integer T such that

$$\forall t \geq T: \quad A^t = \lambda^{\otimes t} \otimes CS^t R, \tag{3}$$

where the matrices C, S, and R are defined in terms of A and fulfill $CS^{t+\gamma}R = CS^t R$ for all $t \geq 0$. In an earlier work, considering infinite-dimensional matrices, Akian, Gaubert and Walsh [2, Section 7] gave a similar formulation originating in the preprints of Cohen et al. [10].

Because of the periodicity of the sequence $CS^t R$, the smallest T satisfying (3) is $T(A)$.

Later, Sergeev and Schneider [24] proved that for t large enough, A^t is the sum (in the max-plus sense) of terms of the form $\lambda_i^{\otimes t} \otimes C_i S_i^t R_i$. This sum, which we call *CSR decomposition* and is closely related to the decomposition introduced in Nachtigall [19], has two remarkable properties: it holds for reducible matrices as well as irreducible ones, and the CSR decomposition holds for $t \geq 3n^2$, a bound that does not depend on the values of the entries of A.

As a common ground of transience bounds and CSR decomposition, we propose the new concept of *weak CSR expansions*. We suggest that all existing techniques for deriving transience bounds implicitly use the idea that eventually we have

$$\forall t \geq T: \quad A^t = \left(\lambda^{\otimes t} \otimes CS^t R\right) \oplus B^t, \tag{4}$$

where C, S, and R are defined as in the CSR expansion, and B is obtained from A by setting several entries (typically, all entries in several rows and columns) to $\mathbf{0}$. In this case, we say that B is *subordinate* to A. Call the smallest T for which (4) holds the *weak CSR threshold* of A with respect to B and denote it by $T_1(A, B)$.

This quantity heavily depends on the choice of B, i.e., on which entries are set to $\mathbf{0}$. If we choose $B = (\mathbf{0})$, then we recover the ordinary CSR expansion and we have $T_1(A, B) = T(A)$. If $\mathcal{D}(B)$ is acyclic, then $B^n = (\mathbf{0})$ and $T(A) \leq \max(T_1(A, B), n)$. More generally $T(A) \leq \max(T_1(A, B), T_2(A, B))$, where $T_2(A, B)$ is the least integer satisfying

$$\forall t \geq T: \quad \lambda^{\otimes t} \otimes \left(CS^t R\right) \geq B^t.$$

Analogously, we call $T_2(A, B, v)$ the least integer satisfying

$$\forall t \geq T : \quad \lambda^{\otimes t} \otimes \left(CS^t Rv \right) \geq B^t v.$$

We claim that the bounds in [4,8,14,27] implicitly are of this type, for various choices of B and various ways to bound T_1 and T_2.

We next summarize the contents of the remaining part of this paper. In Section 2, we recall notions and results of max algebra, focusing on its relation to weighted digraphs. In Section 3, we introduce three schemes of defining B, and thereby weak CSR expansions: the Nachtigall scheme, the Hartmann–Arguelles scheme, and the cycle threshold scheme. The first scheme is implicitly used in [2,4,8,27], the second one is derived from [14], and the third one is completely new.

In Section 4 we state some bounds on $T_1(A, B)$ and $T_2(A, B)$, thus on T that we obtain in this paper. Those bounds strictly improve the ones in [4,8,14,27]. Moreover they can be combined in several ways. Notably, for the three schemes defined in Section 3, we bound the weak CSR threshold $T_1(A, B)$ by the *Wielandt number*

$$\text{Wi}(n) = \begin{cases} 0 & \text{if } n = 1 \\ (n-1)^2 + 1 & \text{if } n > 1 \end{cases} \tag{5}$$

(named in honor of [28]). The bound $\text{Wi}(n)$ is optimal because it is the worst case transient of powers of Boolean matrices, i.e., matrices with entries **0** and **1** (see Remarks 3.1 and 4.2). We also recover another optimal bound for Boolean matrices due to Dulmage and Mendelsohn [11] that does not only depend on n but also on some graph parameter. The section also includes examples to compare the different bounds.

In Section 5, we compare our results to some bounds found in the literature.

In Section 6, we explain the strategy of the proof, which leads us to introduce a graph theoretic quantity, which we name *cycle removal threshold* of a graph and state bounds on $T_1(A, B)$ that depend on this quantity for some graphs.

In Sections 7 and 8, we prove the results stated in Section 6 to bound $T_1(A, B)$ in terms of the cycle removal threshold.

In Section 9 we bound the cycle removal threshold. First we recall the bounds of [7] that depend on several parameters of $\mathcal{D}(A)$ and use the ideas of Hartman and Arguelles [14] to give a new bound depending on less parameters. Then, we introduce a new technique leading to other two bounds on $T_1(A, B)$.

In Section 10 we prove the bounds on $T_2(A, B)$.

In Section 11, we recall some bounds on the index of Boolean matrices to be used in some bounds on T_1.

The technique of local reduction, originating from Akian, Gaubert and Walsh [2, Section 7], is recalled in Section 12. We show that this technique can be combined with any of the CSR schemes described in Section 3.

2. Preliminaries

2.1. Walks in weighted digraphs

Let us recall the optimal walk interpretation of matrix powers in max algebra. This is the fact that the entries of a matrix power A^t are equal to maximum weights of walks of length t in the digraph associated to matrix A.

To a square matrix $A = (a_{ij}) \in \mathbb{R}_{\max}^{n \times n}$ we associate an edge-weighted digraph $\mathcal{D}(A)$ with set of nodes $N = \{1, 2, \ldots, n\}$ and set of edges $E \subseteq N \times N$ containing a pair (i, j) if and only if $a_{ij} \neq \mathbf{0}$; the weight of an edge $(i, j) \in E$ is defined to be $p(i, j) = a_{ij}$. A *walk* W in $\mathcal{D}(A)$ is a finite sequence $(i_0, i_1, \ldots, i_L)$ of adjacent nodes of $\mathcal{D}(A)$. We define its length $l(W) = L$ and weight $p(W) = a_{i_0, i_1} \otimes a_{i_1, i_2} \cdots \otimes a_{i_{t-1}, i_t}$. A closed walk is a walk whose start node i_0 coincides with its end node i_L. Closed walks are often called circuits in the literature. There exists an empty closed walk at every node of length 0 and weight $\mathbf{1} = 0$.

The multiplicity of an edge e in W is the number of k's such that $(i_k, i_{k+1}) = e$. A *subwalk* of walk W is a walk V such that the edges of V appear in W with larger multiplicity. A subwalk of W is a *proper* subwalk if it is not equal to W.

A closed walk is a *cycle* if it does not contain any nonempty closed walk as a proper subwalk. A walk is a *path* if it does not contain a nonempty cycle as a subwalk.

An elementary result of graph theory states that a walk can always be split into a path and some cycles. Reciprocally, union of edges of one path and some cycles can always be reordered into a walk provided the graph with all the edges and nodes of those walks is connected. The best way to see this is in terms of multigraph $M(W)$ defined by a walk W.

For a set $\mathcal{W}$ of walks, we write $p(\mathcal{W})$ for the supremum of walk weights in $\mathcal{W}$. Denote by $\mathcal{W}^t(i \to j)$ the set of all walks from i to j of length t and write $A^t = (a_{ij}^{(t)})$. It is immediate from the definitions that

$$a_{ij}^{(t)} = p\big(\mathcal{W}^t(i \to j)\big). \tag{6}$$

When we do not want to restrict the lengths of walks, we define the set $\mathcal{W}(i \to j)$ of all walks connecting i to j. An analog of $(I - A)^{-1}$ in max-plus algebra is the *Kleene star*

$$A^* = I \oplus A \oplus A^2 \oplus A^3 \oplus \ldots, \tag{7}$$

where I is the max-plus identity matrix. It follows from the optimal walk interpretation (6) that series (7) converges if and only if $p(Z) \leq \mathbf{1}$ for all closed walks Z in $\mathcal{D}(A)$, in which case it can be truncated as $A^* = I \oplus A \oplus \ldots \oplus A^{n-1}$. If we denote $A^* = (a_{ij}^*)$, it is again immediate that

$$a_{ij}^* = p\big(\mathcal{W}(i \to j)\big). \tag{8}$$

The *maximum cycle mean* of $A \in \mathbb{R}_{\max}^{n \times n}$ is defined by

$$\lambda(A) = \max\big\{p(Z)^{\otimes 1/l(Z)} \mid Z \text{ is a nonempty cycle in } \mathcal{D}(A)\big\}. \tag{9}$$

Because every closed walk is composed of cycles, we could replace "cycle" by "closed walk" in definition (9). The maximum cycle mean $\lambda(A)$ is equal to the greatest max-algebraic eigenvalue of A, i.e., a $\mu \in \mathbb{R}_{\max}$ such that there exists a nonzero vector x satisfying $A \otimes x = \mu \otimes x$. Nonempty closed walks of weight $\lambda(A)$ are called *critical*, and so are the nodes and edges on these walks. The subgraph of $\mathcal{D}(A)$ consisting of the set of critical nodes N_c and the set of critical edges E_c is called the *critical graph* of A and is denoted by $\mathcal{G}^c(A) = (N_c, E_c)$. A useful fact (used throughout the paper) is that every nonempty closed walk in $\mathcal{G}^c(A)$ is critical.

As we will see, the behavior of max-algebraic matrix powers is eventually dominated by the walks that visit the critical graph. The set of such walks in $\mathcal{W}^t(i \to j)$ will be denoted by $\mathcal{W}^t(i \xrightarrow{\mathcal{G}^c} j)$. More generally, for a node k and a subgraph $\mathcal{D}$ of $\mathcal{D}(A)$ we write

$$\mathcal{W}^t(i \xrightarrow{k} j) = \bigcup_{t_1+t_2=t} \big\{W_1 W_2 \mid W_1 \in \mathcal{W}^{t_1}(i \to k), W_2 \in \mathcal{W}^{t_2}(k \to j)\big\},$$

$$\mathcal{W}^t(i \xrightarrow{\mathcal{D}} j) = \bigcup_{k \in \mathcal{D}} \mathcal{W}^t(i \xrightarrow{k} j) \quad \text{and} \quad \mathcal{W}(i \xrightarrow{\mathcal{D}} j) = \bigcup_{t \geq 0} \mathcal{W}^t(i \xrightarrow{\mathcal{D}} j).$$

2.2. Cyclicity of digraphs

A digraph $\mathcal{G} = (N, E)$ is *strongly connected* if there exists a walk from i to j for all nodes $i, j \in N$. A *strongly connected component (s.c.c.)* of $\mathcal{G}$ is a maximal strongly connected subgraph of $\mathcal{G}$. Digraph $\mathcal{G}$ is called *completely reducible* if there are no edges between distinct s.c.c.'s of $\mathcal{G}$. The critical graph $\mathcal{G}^c(A)$ will be the most important example of this.

Matrix A is called *irreducible* if its associated digraph is strongly connected, and *reducible* otherwise. Further, it is called completely reducible if so is the associated digraph.

The *cyclicity* $\gamma(\mathcal{G})$ of a strongly connected digraph $\mathcal{G}$ is the greatest common divisor of the lengths of its closed walks. If $\mathcal{G}$ is not strongly connected, its cyclicity $\gamma(\mathcal{G})$ is the least common multiple of the cyclicities of its s.c.c.'s. It is well-known that any two lengths of walks on $\mathcal{G}$ both starting at some node i and both ending at some node j are congruent modulo $\gamma(\mathcal{G})$. Moreover, if $\mathcal{G}$ is strongly connected, there exist walks from i to j of all lengths that are large enough and that are congruent to some t_{ij} modulo $\gamma(\mathcal{G})$.

We call a subgraph $\mathcal{G}$ of $\mathcal{G}^c(A)$ a *representing subgraph* if $\mathcal{G}$ is completely reducible and every s.c.c. of $\mathcal{G}^c(A)$ contains exactly one s.c.c. of $\mathcal{G}$. The cyclicity $\gamma(\mathcal{G})$ of a representing subgraph of $\mathcal{G}^c(A)$ is always a multiple of $\gamma(\mathcal{G}^c(A))$. Hence Eq. (2) also holds with $\gamma = \gamma(\mathcal{G})$ instead of $\gamma(\mathcal{G}^c(A))$.

2.3. Visualization and max-balancing

The maximum cycle mean $\lambda(A)$ also appears as the least $\mu \in \mathbb{R}_{\max}$ such that there exists a finite vector x satisfying $Ax \leq \mu \otimes x$. When $\mu = \lambda(A)$, we can take $x_i = \bigoplus_{j=1}^{n} (\lambda^-(A) \otimes A)_{ij}^*$, that is, the component-wise maximum of all columns of $(\lambda^-(A) \otimes A)^*$. Setting D, resp. D^-, to be the diagonal matrix with entries $d_{ii} = x_i$ and $d_{ij} = \mathbf{0}$ for $i \neq j$, resp. $d_{ii}^- := x_i^-$ and $d_{ij} = \mathbf{0}$ for $i \neq j$, we obtain for $B = D^-(\lambda^- \otimes A)D$ that $\mathcal{G}^c(B) = \mathcal{G}^c(A)$ (as unweighted digraphs) and

$$b_{ij} \leq \mathbf{1} \quad \text{for all } i, j,$$
$$b_{ij} = \mathbf{1} \quad \text{for all } i, j \text{ in } \mathcal{G}^c(B). \tag{10}$$

When (10) holds we say that B is *visualized*: it exhibits the edges of the critical graph. A diagonal matrix D such that $B = D^-(\lambda^-(A) \otimes A)D$ is visualized and $\mathcal{G}^c(B) = \mathcal{G}^c(A)$ is also called a Fiedler–Pták scaling [12] of A. In this case, we call B a *visualization* of A.

Fiedler–Pták scalings were described in more detail by Sergeev, Schneider and Butkovič [25] using Kleene stars and max algebra. Butkovič and Schneider [6] described applications to various kinds of nonnegative similarity scalings. A Fiedler–Pták scaling, particularly interesting to us, is called the *max-balancing*. It was described by Schneider and Schneider [21]:

Theorem 2.1. *(See Schneider and Schneider [21].) For all $A \in \mathbb{R}_{\max}^{n \times n}$ there exists a visualization B of A satisfying the following equivalent properties:*

1. *(Cycle cover) For all edges (i, j) in $\mathcal{D}(B)$ there exists a cycle Z in $\mathcal{D}(B)$ containing (i, j) such that all edges of Z have weight at least b_{ij}.*
2. *(Max-balancing) For all sets $M \subseteq \{1, \ldots, n\}$, we have: $\max_{i \in M, j \notin M} b_{ij} = \max_{i \notin M, j \in M} b_{ij}$.*

2.4. CSR expansions, weak CSR expansions

For any $A \in \mathbb{R}_{\max}^{n \times n}$ and any subgraph $\mathcal{G}$ of $\mathcal{G}^c(A)$ with no trivial s.c.c., we set $M = ((\lambda(A)^- \otimes A)^{\gamma(\mathcal{G})})^*$ and define the matrices $C, S, R \in \mathbb{R}_{\max}^{n \times n}$ by

$$c_{ij} = \begin{cases} m_{ij} & \text{if } j \text{ is in } \mathcal{G} \\ \mathbf{0} & \text{otherwise,} \end{cases} \qquad r_{ij} = \begin{cases} m_{ij} & \text{if } i \text{ is in } \mathcal{G} \\ \mathbf{0} & \text{otherwise,} \end{cases}$$
$$s_{ij} = \begin{cases} \lambda(A)^- \otimes a_{ij} & \text{if } (i, j) \text{ is in } \mathcal{G} \\ \mathbf{0} & \text{otherwise.} \end{cases} \tag{11}$$

When the dependency on $\mathcal{G}$ needs to be emphasized, we write $C_{\mathcal{G}}$, $S_{\mathcal{G}}$ and $R_{\mathcal{G}}$ instead of C, S and R.

Essentially C and R can be regarded as sub-matrices of M extracted from the columns, resp. the rows of M with indices in $\mathcal{G}$. If A is visualized, then matrix S is exactly the

associated Boolean matrix of $\mathcal{G}$. The matrices C, S, and R are called the *CSR terms* of A with respect to $\mathcal{G}$.

The following is a CSR version of the Cyclicity Theorem.

Theorem 2.2. *(See [20,23].) Let $A \in \mathbb{R}_{\max}^{n \times n}$ be irreducible and let C, S, R be the CSR terms of A with respect to $\mathcal{G}^c(A)$. Then for all $t \geq T(A)$:*

$$A^t = \lambda(A)^{\otimes t} \otimes CS^t R.$$

As it is shown below, Theorem 2.2 also holds with $\mathcal{G}^c(A)$ replaced by some representing subgraph $\mathcal{G}$ of $\mathcal{G}^c(A)$.

Note that this theorem implies periodicity of A^t after $T(A)$, because the sequence of matrices $CS^t R$ is *purely periodic*, i.e., periodic from the very beginning. In fact, this statement is more generally true for all completely reducible (and hence also for all representing) subgraphs of $\mathcal{G}^c(A)$:

Proposition 2.3. *(See [24].) Let $A \in \mathbb{R}_{\max}^{n \times n}$ be irreducible and C, S, R be the CSR terms of A with respect to some completely reducible subgraph $\mathcal{G}$ of the critical graph $\mathcal{G}^c(A)$. Then the sequence of matrices $CS^t R$ is purely periodic.*

This fact was shown by Sergeev and Schneider [24], where CSR terms with respect to completely reducible subgraphs of $\mathcal{G}^c(A)$ were studied in detail. It can also be deduced from Theorem 6.1 proved below.

A *weak CSR expansion* of A is an expansion of the form (4) where C, S, R are CSR terms with respect to some representing subgraph of $\mathcal{G}^c(A)$ and $\mathcal{D}(B)$ is a subgraph of $\mathcal{D}(A)$ disjoint to $\mathcal{G}^c(A)$. In particular, the result of Theorem 2.2 is also a weak CSR expansion (take B equal to the max-plus zero matrix).

By iteration of weak CSR expansions, we recover the CSR decomposition of A introduced in [24]. Bounds on T_1 give bounds on the time from which A^t admits such a decomposition. (See Corollary 4.3.)

3. Weak CSR schemes

In this section, we introduce our three schemes for weak CSR expansions and discuss their relation. We define them in terms of the subgraph $\mathcal{D}$ of $\mathcal{D}(A)$ whose edges denote the indices that are set to $\mathbf{0}$ in the subordinate matrix B. More explicitly:

$$b_{ij} = \begin{cases} \mathbf{0} & \text{if } i \text{ or } j \text{ is a node of } \mathcal{D} \\ a_{ij} & \text{else.} \end{cases} \tag{12}$$

The three schemes are:

1. *Nachtigall scheme.* Here, the subgraph $\mathcal{D} = \mathcal{G}^c(A)$. We denote the resulting matrix B by B_{N}.

This scheme is consistent with the expansion introduced by Nachtigall [19], which was studied by Molnárová [18] and Sergeev and Schneider [24]. It was used by almost all authors who studied matrix transients [2,4,8,27], excluding Hartmann and Arguelles [14].

2. *Hartmann–Arguelles scheme.* This scheme is defined in terms of the max-balancing $V = (v_{ij})$ of A. Given $\mu \in \mathbb{R}_{\max}$, define the *Hartmann–Arguelles threshold graph* $\mathcal{T}^{ha}(\mu)$ induced by all edges (i, j) in $\mathcal{D}(A) = \mathcal{D}(V)$ with $v_{ij} \geq \mu$. For $\mu = \lambda(A) = \lambda(V)$ we have $\mathcal{T}^{ha}(\mu) = \mathcal{G}^c(A) = \mathcal{G}^c(V)$. Let μ^{ha} be the maximum of $\mu \leq \lambda(A)$ such that $\mathcal{T}^{ha}(\mu)$ has an s.c.c. that does not contain any s.c.c. of $\mathcal{G}^c(A)$. If no such μ exists, then $\mu^{ha} = \mathbf{0}$ and $\mathcal{T}^{ha}(\mu^{ha}) = \mathcal{D}(A)$.

   The subgraph $\mathcal{D} = \mathcal{G}^{ha}$ defining B in the Hartmann–Arguelles scheme is the union of the s.c.c. of $\mathcal{T}^{ha}(\mu^{ha})$ intersecting $\mathcal{G}^c(A)$. We denote this matrix B by B_{HA}. Observe that $\lambda(B_{\mathrm{HA}}) = \mu^{ha}$ and the graphs $\mathcal{T}^{ha}(\mu)$, for all μ, are completely reducible due to max-balancing (more precisely, the cycle cover property).

3. *Cycle threshold scheme.* For $\mu \in \mathbb{R}_{\max}$, define the *cycle threshold graph* $\mathcal{T}^{ct}(\mu)$ induced by all nodes and edges belonging to the cycles in $\mathcal{D}(A)$ with mean weight greater or equal to μ. Again, for $\mu = \lambda(A)$ we have $\mathcal{T}^{ct}(\mu) = \mathcal{G}^c(A)$. Let μ^{ct} be the maximum of $\mu \leq \lambda(A)$ such that $\mathcal{T}^{ct}(\mu)$ has an s.c.c. that does not contain any s.c.c. of $\mathcal{G}^c(A)$. If no such μ exists, then $\mu^{ct} = \mathbf{0}$ and $\mathcal{T}^{ct}(\mu^{ct})$ is equal to $\mathcal{D}(A)$.

   The subgraph $\mathcal{D} = \mathcal{G}^{ct}$ defining B in the cycle threshold scheme is the union of the s.c.c. of $\mathcal{T}^{ct}(\mu^{ct})$ intersecting $\mathcal{G}^c(A)$. This matrix B will be denoted by B_{CT}. We again observe that $\lambda(B_{\mathrm{CT}}) = \mu^{ct}$.

Remark 3.1. Since $\mathcal{G}^c(A) \subset \mathcal{D}$, we see that $\lambda(B) < \lambda(A)$.

In particular, if A is an irreducible Boolean matrix, then $\mathcal{G}^c(A) = \mathcal{D}(A)$ and $B = (\mathbf{0})$ for all schemes, thus $T_1(A, B_{\mathrm{N}}) = T_1(A, B_{\mathrm{HA}}) = T_1(A, B_{\mathrm{CT}}) = T(A)$.

The weak CSR thresholds hence are generalizations of the transient of irreducible Boolean matrices, which has been investigated in the literature under the name *index (of convergence)* of A, or also *exponent* in case of primitive matrices. See Section 11 for a brief account.

Proposition 3.2. *The matrices B_{N} and B_{HA} can be computed in polynomial time. The computation of the threshold graphs $\mathcal{T}^{ct}(0)$ is NP-hard.*

Proof. The computation of B_{N} relies on the computation of $\mathcal{G}^c(A) = (N_c, E_c)$, for which we can exploit the well-known criterion $a_{ij} a_{ji}^* = \mathbf{1} \Leftrightarrow (i, j) \in E_c(A)$ (when $\lambda(A) = \mathbf{1}$). This yields complexity at most $O(n^3)$.

Concerning B_{HA}, Schneider and Schneider [21] proved that a max-balancing of A can be computed in polynomial time (at most $O(n^4)$). The same order of complexity is added if we "brutally" examine at most n^2 threshold graphs (for each of them, the strongly connected components found in $O(n^2)$ time). A better complexity result can be derived from the work of Hartmann–Arguelles [14].

To show NP-hardness of the computation of $\mathcal{T}^{ct}(\mu)$, we reduce the Longest Path Problem [13, p. 213, ND29] to it. Consider the Longest Path Problem as a decision problem that takes as input an edge-weighted digraph with integer weights, a pair of nodes (i, j) with $i \neq j$ in the digraph, and an integer K. The output is YES if there exists a path of weight at least K from i to j. The output is NO if there is none. Observe that if $i \neq j$, then by inserting the edge (j, i) with weight $-K$, the Longest Path Problem can be polynomially reduced to the problem of calculating $\mathcal{T}^{ct}(0)$ by checking whether the new edge (j, i) belongs to $\mathcal{T}^{ct}(0)$. □

The relation between these schemes is as follows. The cycle threshold scheme is most precise, while the Nachtigall scheme is the coarsest. We measure this in terms of the size of B and the value $\lambda(B)$.

Proposition 3.3. B_{CT} *is subordinate to* B_{HA}*, which is subordinate to* B_{N}*. In particular,*

$$\lambda(B_{\mathrm{CT}}) \leq \lambda(B_{\mathrm{HA}}) \leq \lambda(B_{\mathrm{N}}).$$

Proof. Evidently both $\mathcal{D}(B_{\mathrm{CT}})$ and $\mathcal{D}(B_{\mathrm{HA}})$ are subgraphs of $\mathcal{D}(B_{\mathrm{N}})$, which is extracted from all non-critical nodes. This implies $\lambda(B_{\mathrm{CT}}) \leq \lambda(B_{\mathrm{N}})$ and $\lambda(B_{\mathrm{HA}}) \leq \lambda(B_{\mathrm{N}})$.

We show that $\mathcal{D}(B_{\mathrm{CT}})$ is a subgraph of $\mathcal{D}(B_{\mathrm{HA}})$. For this we can assume that the whole digraph is max-balanced, and notice first that $\mathcal{T}^{ha}(\mu) \subseteq \mathcal{T}^{ct}(\mu)$ for any value of μ. We also have that $\mathcal{T}^{ha}(\mu_1) \supseteq \mathcal{T}^{ha}(\mu_2)$ and $\mathcal{T}^{ct}(\mu_1) \supseteq \mathcal{T}^{ct}(\mu_2)$ for any $\mu_1 \leq \mu_2$. Now consider the value μ^{ct}. The components of $\mathcal{T}^{ct}(\mu^{ct})$ which do not contain the components of $\mathcal{G}^c(A)$, have the property that any other cycle intersecting with them has a strictly smaller cycle mean. It follows that all edges of these components have weight μ^{ct}. Indeed, suppose that there is a component containing an edge with a different weight. In this component, any cycle that contains this edge also has an edge with weight strictly greater than μ^{ct}. The cycle cover property implies that there is a cycle containing this edge, where this edge has the smallest weight. The mean of that cycle is strictly greater than μ^{ct}, a contradiction. But then $\mathcal{T}^{ha}(\mu^{ct})$ contains these components as its s.c.c.'s. In particular they do not contain the components of $\mathcal{G}^c(A)$, hence $\mu^{ct} \leq \mu^{ha}$.

If $\mu := \mu^{ct} = \mu^{ha}$ then $\mathcal{T}^{ha}(\mu) \subseteq \mathcal{T}^{ct}(\mu)$, while we have shown that the components of $\mathcal{T}^{ct}(\mu)$ not containing the components of $\mathcal{G}^c(A)$ are also components of $\mathcal{T}^{ha}(\mu)$. It follows that $\mathcal{G}^{ha} \subseteq \mathcal{G}^{ct}$.

If $\mu^{ct} < \mu^{ha}$ then we obtain that

$$\mathcal{G}^{ct} \supseteq \mathcal{T}^{ct}\big(\mu^{ha}\big) \supseteq \mathcal{T}^{ha}\big(\mu^{ha}\big) \supseteq \mathcal{G}^{ha},$$

thus $\mathcal{G}^{ha} \subseteq \mathcal{G}^{ct}$ in any case, hence $\mathcal{D}(B_{\mathrm{CT}}) \subset \mathcal{D}(B_{\mathrm{HA}})$. □

The following example shows that all three schemes can differ and, moreover, that the thresholds $T_1(A, B_{\mathrm{N}})$, $T_1(A, B_{\mathrm{HA}})$ and $T_1(A, B_{\mathrm{CT}})$ can all differ.
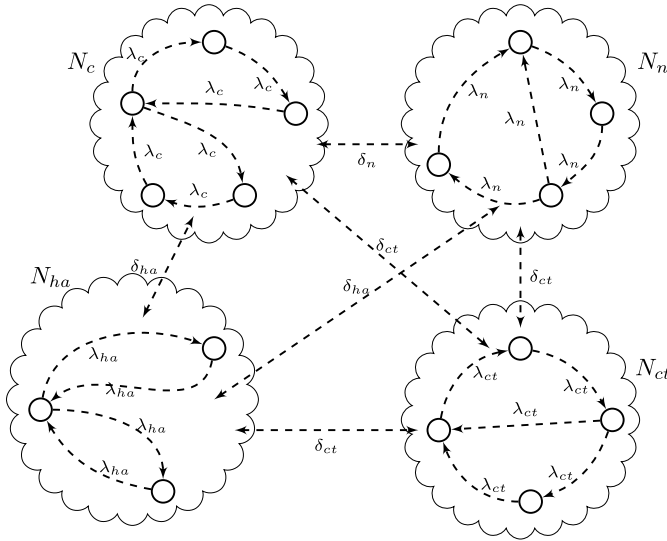
Fig. 1. A schematic sketch of $\mathcal{D}(A)$ of Example 3.4 (the general case).

Example 3.4. Consider a matrix

$$A = \begin{pmatrix} 0 & 0 & -1 & -\infty & -7 \\ 0 & 0 & -1 & -\infty & -7 \\ -1 & -1 & -1 & -3 & -7 \\ -3 & -\infty & -\infty & -2 & -7 \\ -7 & -7 & -7 & -7 & -3 \end{pmatrix}. \tag{13}$$

In this example we have $\lambda(A) = 0$, it is visualized and, moreover, max-balanced. The matrices B_{N}, resp. B_{HA} and B_{CT} are formed by setting the first 2 rows and columns, resp. the first 3 and 4 rows and columns to $\mathbf{0} = -\infty$, and the corresponding values are $\lambda(B_{\mathrm{N}}) = -1$, $\lambda(B_{\mathrm{HA}}) = -2$ and $\lambda(B_{\mathrm{CT}}) = -3$. The corresponding thresholds are $T_1(A, B_{\mathrm{N}}) = 2$, $T_1(A, B_{\mathrm{HA}}) = 3$ and $T_1(A, B_{\mathrm{CT}}) = 4$: all different. The periodicity threshold of $(A^{\otimes t})_{t \geq 1}$ is equal to $T(A) = 5$, which is the same as $T_2(A, B_{\mathrm{N}}) = T_2(A, B_{\mathrm{HA}}) = T_2(A, B_{\mathrm{CT}})$.

Let us provide a class of examples that generalizes the example above to arbitrary dimension. For any matrix A in this class of examples, all three schemes are different but the corresponding thresholds $T_1(A, B)$ may coincide.

Consider a matrix A such that the node set N of $\mathcal{D}(A)$ is partitioned into $N = N_c \cup N_n \cup N_{ha} \cup N_{ct}$, see Fig. 1. For each $x \in \{c, n, ha, ct\}$, the nodes in N_x form a strongly connected graph where all edges have weight λ_x. We set $\lambda_c > \lambda_n > \lambda_{ha} > \lambda_{ct}$. For each set N_x with $x \in \{n, ha, ct\}$, we assume that there is at least one edge from N_x to some set N_y with $\lambda_y > \lambda_x$, and one edge from one of such N_y to N_x. With this assumption, it can be shown that $\mathcal{D}(A)$ is strongly connected. Let us also assume that all such edges (from N_x and to N_x) have the same weight δ_x. Observe that for the matrix

of (13), we have $N_c = \{1, 2\}$, $N_n = \{3\}$, $N_{ha} = \{4\}$ and $N_{ct} = \{5\}$; $\lambda_c = 0$, $\lambda_n = -1$, $\lambda_{ha} = -2$ and $\lambda_{ct} = -3$; $\delta_n = -1$, $\delta_{ha} = -3$ and $\delta_{ct} = -7$.

Assume that δ_x satisfies

$$\delta_x \leq \min\big(\lambda_x, \min\{\delta_y \mid \lambda_y > \lambda_x\}\big). \tag{14}$$

Then $\mathcal{D}(A)$ is also max-balanced (since it can be shown that each edge (i, j) with $i \neq j$ is on a cycle where it has the smallest weight).

We also see that $\lambda_c = \lambda(A)$, $\lambda_n = \lambda(B_{\mathrm{N}})$, while λ_{ha} and λ_{ct} are "candidates" for $\lambda(B_{\mathrm{HA}})$ and $\lambda(B_{\mathrm{CT}})$, respectively. To enforce the correct behavior of threshold graphs and ensure that $\lambda(B_{\mathrm{HA}}) = \lambda_{ha}$ and $\lambda(B_{\mathrm{CT}}) = \lambda_{ct}$, we set:

1) $\delta_n = \lambda_n$;
2) $\delta_{ha} = \lambda_{ha} - s$, where s is chosen in such a way that the inequality

$$\big(\ell(Z) - 2\big) \cdot \lambda_n + 2(\lambda_{ha} - s) \geq \ell(Z) \cdot \lambda_{ha}$$

holds at least for one cycle Z containing one node in N_{ha} and the rest in $N_c \cup N_n$;
3) δ_{ct} not greater than δ_{ha} (for the sake of max-balancing) and such that the mean weight of each cycle containing a node of N_{ct} and a node of $N \backslash N_{ct}$ is strictly less than λ_{ct}.

Observe, in particular, that condition 2) ensures that $\mathcal{T}^{ct}(\mu)$ does not gain any new component as μ decreases from λ_c to λ_{ha} so that $\lambda_{ct} < \lambda(B_{\mathrm{HA}})$, and that condition 3) ensures $\lambda_{ct} = \lambda(B_{\mathrm{CT}})$. Note that (13) satisfies conditions 1)–3).

4. Main results

In this section, we present the main results of this paper. The bounds of this section use the following graph parameters of a digraph $\mathcal{D}$:

- *size* $|\mathcal{D}|$: the number of nodes of $\mathcal{D}$,
- *circumference* $\mathrm{cr}(\mathcal{D})$: the greatest length of a cycle in graph $\mathcal{D}$,
- *cab driver's diameter* $\mathrm{cd}(\mathcal{D})$: the greatest length of a path in $\mathcal{D}$,
- *max-girth* $\hat{\mathrm{g}}(\mathcal{G})$: the greatest *girth*, i.e., shortest cycle length, of strongly connected components of $\mathcal{D}$
- *max-cyclicity* $\hat{\gamma}(\mathcal{G})$: the greatest cyclicity of strongly connected components of $\mathcal{D}$.

The computation of the circumference $\mathrm{cr}(\mathcal{D})$ and the cab driver's diameter $\mathrm{cd}(\mathcal{D})$ are both NP-hard in the number of nodes of $\mathcal{D}$. However, they can be upper bounded by $|\mathcal{D}|$ and $|\mathcal{D}| - 1$, respectively.

Denote by $\|A\|$ the difference between the largest and the smallest finite (i.e., $\neq \mathbf{0}$) entry of A, and by n_B the size of the smallest submatrix of B containing all its finite entries.

We explained in the introduction that $T(A) \leq \max(T_1(A, B), T_2(A, B))$. Our main results are bounds on T_1 and T_2. All of them are mutually incomparable.

Theorem 4.1. *For any matrix $A \in \mathbb{R}_{\max}^{n \times n}$, if $B = B_\mathrm{N}$ or $B = B_\mathrm{HA}$, we have the following bounds*

$$T_1(A, B) \leq \mathrm{Wi}(n) \tag{15}$$

$$T_1(A, B) \leq \hat{\mathrm{g}}(n - 2) + n \tag{16}$$

$$T_1(A, B) \leq (\hat{\mathrm{g}} - 1)(\mathrm{cr} - 1) + (\hat{\mathrm{g}} + 1)\,\mathrm{cd} \tag{17}$$

where $\hat{\mathrm{g}} = \hat{\mathrm{g}}(\mathcal{G}^c(A))$, $\mathrm{cr} = \mathrm{cr}(\mathcal{D}(A))$, $\mathrm{cd} = \mathrm{cd}(\mathcal{D}(A))$.

$$T_1(A, B) \leq \hat{\gamma}(n - 2) + n - n_c + \mathrm{ep}\big(\mathcal{G}^c(A)\big) \tag{18}$$

$$T_1(A, B) \leq (\hat{\gamma} - 1)(\mathrm{cr} - 1) + (\hat{\gamma} + 1)\,\mathrm{cd} + \mathrm{ep}\big(\mathcal{G}^c(A)\big) \tag{19}$$

where $\hat{\gamma} = \hat{\gamma}(\mathcal{G}^c(A))$, $n_c = |\mathcal{G}^c(A)|$ (i.e., the number of critical nodes) and $\mathrm{ep}(\mathcal{G}^c(A))$ is the exploration penalty of $\mathcal{G}^c(A)$ (see Definition 6.4).

The exploration penalty $\mathrm{ep}(\mathcal{G}^c(A))$ is a quantity that depends only on the critical graph and can be bounded by its index, see Section 11 for further details.

Remark 4.2. As we noted in Remark 3.1, those bounds apply to the transient of Boolean matrices. We thus recover the bound of Wielandt [28] in (15) and the bound of Dulmage and Mendelsohn [11] in (16). Notice that (16) implies (15) if $\hat{\mathrm{g}} \leq n - 1$. The remaining case is trivial for Boolean matrices because there is only one such matrix, but in the non-Boolean case we need a different strategy (Proposition 9.4 below).

Bound (15) is optimal in the sense that the bound is reached for any n, as was already noted in [28], while bound (16) is reached if and only if $\hat{\mathrm{g}}$ and n are coprime (see [26]).

Iterating the process of weak CSR expansion, we get the following improvement of [24, Theorem 4.2]:

Corollary 4.3 *(CSR decomposition). For any matrix $A \in \mathbb{R}_{\max}^{n \times n}$, there are matrices C_i, S_i, R_i defined by induction with S_i diagonally similar to Boolean periodic matrices and some scalars $\lambda_i \in \mathbb{R}$, where i varies between 1 and $K \leq n$ such that we have:*

$$\forall t \geq \min\big(\mathrm{Wi}(n), (n - 2)\,\mathrm{cr}\big(\mathcal{D}(A)\big) + n\big), \quad A^t = \bigoplus_{k=1}^{K} \lambda_i^{\otimes t} C_i S_i^t R_i.$$

Theorem 4.4. *For any matrix $A \in \mathbb{R}_{\max}^{n \times n}$, we have the following bounds*

$$T_1(A, B_{\mathrm{CT}}) \leq \mathrm{Wi}(n) \tag{20}$$

$$T_1(A, B_{\mathrm{CT}}) \leq (n-1)\,\mathrm{cr} + \min(n, \mathrm{cd} + \mathrm{cr} + 1) \tag{21}$$

$$T_1(A, B_{\mathrm{CT}}) \leq (\mathrm{cd} + \mathrm{cr} - 1)\,\mathrm{cr} + \mathrm{cd} + 1 \tag{22}$$

where $\mathrm{cr} = \mathrm{cr}(\mathcal{D}(A))$ and $\mathrm{cd} = \mathrm{cd}(\mathcal{D}(A))$.

The proof of those theorems is explained in Section 6 and performed in Sections 7 and 8. Now we also state bounds on $T_2(A, B)$ and on $T_2(A, B, v)$.

Theorem 4.5. *Let $A \in \mathbb{R}_{\max}^{n \times n}$ be irreducible and let B be subordinate to A. Denote $\mathrm{cd}_B := \mathrm{cd}(\mathcal{D}(B))$ and $\hat{\gamma} = \hat{\gamma}(\mathcal{G}^c(A))$.*

If $\lambda(B) = \mathbf{0}$, then $T_2(A, B) \leq \mathrm{cd}_B + 1 \leq n_B$. Otherwise, we have the following bounds

$$
\begin{aligned}
T_2(A, B) &\leq \frac{(n^2 - n + 1)(\lambda(A) - \min_{ij} a_{ij}) + \mathrm{cd}_B(\max_{ij} b_{ij} - \lambda(B))}{\lambda(A) - \lambda(B)} \\
&\leq \frac{n^2 - n + 1}{\lambda(A) - \lambda(B)}\|A\| + \mathrm{cd}_B
\end{aligned} \tag{23}
$$

$$
\begin{aligned}
T_2(A, B) &\leq \frac{(\hat{\gamma}(n-1) + n)(\lambda(A) - \min_{ij} a_{ij}) + \mathrm{cd}_B(\max_{ij} b_{ij} - \lambda(B))}{\lambda(A) - \lambda(B)} \\
&\leq \frac{\hat{\gamma}(n-1) + n}{\lambda(A) - \lambda(B)}\|A\| + \mathrm{cd}_B
\end{aligned} \tag{24}
$$

$$
\begin{aligned}
T_2(A, B) &\leq \frac{((\hat{\gamma} - 1)\,\mathrm{cr} + (\hat{\gamma} + 1)\,\mathrm{cd})(\lambda(A) - \min_{ij} a_{ij}) + \mathrm{cd}_B(\max_{ij} b_{ij} - \lambda(B))}{\lambda(A) - \lambda(B)} \\
&\leq \frac{(\hat{\gamma} - 1)\,\mathrm{cr} + (\hat{\gamma} + 1)\,\mathrm{cd}}{\lambda(A) - \lambda(B)}\|A\| + \mathrm{cd}_B \,.
\end{aligned} \tag{25}
$$

If A has only finite entries, then we have:

$$T_2(A, B) \leq \frac{2(\lambda(A) - \min_{ij} a_{ij}) + (\lambda(B) - \min_{ij} b_{ij})}{\lambda(A) - \lambda(B)} \leq \frac{3\|A\|}{\lambda(A) - \lambda(B)} \tag{26}$$

$$T_2(A, B) \leq \frac{2(\lambda(A) - \min_{ij} a_{ij})}{\lambda(A) - \lambda(B)} + \mathrm{cd}_B \leq \frac{2\|A\|}{\lambda(A) - \lambda(B)} + \mathrm{cd}_B \,. \tag{27}$$

The following theorem generalizes Proposition 5 of [8] and Theorem 3.5.12 of [27].

Theorem 4.6. *Let $A \in \mathbb{R}_{\max}^{n \times n}$ be irreducible, B be subordinate to A and v be a vector with only finite entries, i.e., $v \in \mathbb{R}^n$.*

If $\lambda(B) = \mathbf{0}$, then $T_2(A, B, v) \leq T_2(A, B) \leq \mathrm{cd}_B + 1 \leq n_B$. Otherwise, we have the following bound:

$$T_2(A, B, v) \leq \frac{\|v\| + (n-1)\|A\|}{\lambda(A) - \lambda(B)} \tag{28}$$

If A has only finite entries, then we have:

$$T_2(A, B, v) \leq \frac{\|v\| + (\lambda(A) - \min_{ij} a_{ij}) + (\lambda(B) - \min_{ij} b_{ij})}{\lambda(A) - \lambda(B)} \leq \frac{2\|A\| + \|v\|}{\lambda(A) - \lambda(B)}. \tag{29}$$

The proofs of Theorems 4.5 and 4.6 are deferred to Section 10.

Remark 4.7. The bounds on T_1 are quadratic in n, but even if one fixes the size of the entries (for instance entries are $-\infty, 0$ or 1), the general bounds on T_2 have degree 4, because $\frac{1}{\lambda(A) - \lambda(B)}$ can be as big as $\|A\|(n^2 - 1)/4$. (Take two cycles with length $(n+1)/2$ and $(n-1)/2$ which both have weight 1.)

For the same reason, both bounds are quadratic if all entries are finite.

5. Comparison to previous transience bounds

Hartmann and Arguelles [14] proved one transience bound for irreducible max-plus matrices and one for irreducible max-plus systems with finite initial vector. These two bounds are, respectively,

$$\max\left(2n^2, \frac{2n^2}{\lambda(A) - \lambda(B_{\mathrm{HA}})}\|A\|\right) \quad \text{and} \quad \max\left(2n^2, \frac{\|v\| + n\|A\|}{\lambda(A) - \lambda(B_{\mathrm{HA}})}\right).$$

Combination of our bounds in (20) and (23), respectively (28), yields bounds that are strictly lower than that of Hartmann and Arguelles. Note that our results, being more detailed, allow a considerably more fine-grained analysis of the transient phase. For instance, there exist matrices for which $\lambda(B_{\mathrm{CT}}) = \mathbf{0}$ but $\lambda(B_{\mathrm{HA}}) \neq \mathbf{0}$ (cf. Example 3.4). Our bounds show in particular that the transients of these matrices and systems are at most $\mathrm{Wi}(n)$, which cannot be deduced from previous bounds, including that of Hartmann and Arguelles.

Bouillard and Gaujal [4] and Akian et al. [2] gave transience bounds for irreducible matrices in the case that the cyclicity of the critical graph is equal to 1. They explained how to extend their bounds to arbitrary cyclicities, but that reduction involves multiplying the bound by the cyclicity of the critical graph or its subgraph. Akian et al. [2] derive bounds for the periodicity transient of $\{a_{ij}^{(t)}\}_{t \geq 1}$ for fixed i, j instead of the whole matrix powers, and show that their bounding techniques extend to the case of matrices of infinite dimensions. We discuss the relation of this approach to weak CSR expansions in more detail in Section 12.

Soto y Koelemeijer [27] (Theorem 3.5.12) established a transience bound for matrices whose entries are all finite. In our notation, it reads

$$\max\left(2n^2, \frac{2\|A\|}{\lambda(A) - \lambda(B_{\mathrm{N}})} + n + 1\right).$$

Combination of our bounds in (20) and (27) yields a bound of $\max(\mathrm{Wi}(n), 2\|A\|/(\lambda(A) - \lambda(B_{\mathrm{CT}})) + \mathrm{cd}_B)$, which is strictly lower. In many cases, it is even better to use (26).

Charron-Bost et al. [8] gave two transience bounds for systems. They also explained how to transform transience bounds for systems into transience bounds for matrices. Combination of our bounds (17), (19), and (28) yields bounds that are strictly lower than those of [8].

6. Proof strategy

In this section, we outline the proof of the bounds on T_1 stated in Theorems 4.1 and 4.4. Moreover, we provide some general statements that can be used to get a better bound if more information on the matrix is available.

In all proofs, we assume $\lambda(A) = \mathbf{1}$ (replacing A by $\lambda(A)^- \otimes A$ if necessary).

The first stage of the proof is the following representation theorem for CS^tR expansions.

Theorem 6.1 *(CSR and walks). Let $A \in \mathbb{R}_{\max}^{n \times n}$ be a matrix with $\lambda(A) = \mathbf{1}$ and C, S, R be the CSR terms of A with respect to some completely reducible subgraph $\mathcal{G}$ of the critical graph $\mathcal{G}^c(A)$.*

Let γ be a multiple of $\gamma(\mathcal{G})$ and $\mathcal{N}$ a set of critical nodes that contains one node of every s.c.c. of $\mathcal{G}$.

Then we have, for any i, j and $t \in \mathbb{N}$:

$$\left(CS^tR\right)_{ij} = p\big(\mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{N}} j)\big) \tag{30}$$

where $\mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{N}} j) := \{W \in \mathcal{W}(i \xrightarrow{\mathcal{N}} j) \mid l(W) \equiv t \pmod{\gamma}\}$.

The proof of this theorem is deferred to Section 7.

Observe that it implies Proposition 2.3 as well as the following corollary.

Corollary 6.2. *CS^tR depends only on the set of s.c.c.'s of $\mathcal{G}^c(A)$ intersecting with $\mathcal{G}$.*

Let $\mathcal{G}_1, \ldots, \mathcal{G}_l$ be the s.c.c. of $\mathcal{G}^c(A)$ with node sets $N_1, \ldots, N_l$, and let $C_{\mathcal{G}_1}, S_{\mathcal{G}_1}, R_{\mathcal{G}_1}$ be the CSR terms defined with respect to $\mathcal{G}_1$. For $\nu = 2, \ldots, l$, we define a subordinate matrix $A^{(\nu)}$ by setting the entries of A with rows and columns in $N_1 \cup \ldots \cup N_{\nu-1}$ to $\mathbf{0}$, and let $C_{\mathcal{G}_\nu}, S_{\mathcal{G}_\nu}, R_{\mathcal{G}_\nu}$ be the CSR terms defined with respect to $\mathcal{G}_\nu$ in $A^{(\nu)}$.

Corollary 6.3. *If $\mathcal{G}_1, \cdots, \mathcal{G}_l$ are the s.c.c.'s of $\mathcal{G}^c(A)$, then we have:*

$$CS^t R = \bigoplus_{\nu=1}^{l} C_{\mathcal{G}_\nu} S^t_{\mathcal{G}_\nu} R_{\mathcal{G}_\nu}. \qquad (31)$$

Proof. Using Theorem 6.1, observe that the set of walks $\mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{G}^c(A)} j)$, where γ is the cyclicity of $\mathcal{G}^c(A)$, can be decomposed into the sets $\mathcal{W}_\nu$ consisting of walks in $\mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{G}_\nu} j)$ that do not visit any node of $\mathcal{G}_1, \ldots, \mathcal{G}_{\nu-1}$, for $\nu = 1, \ldots, l$ (in particular, $\mathcal{W}_1 = \mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{G}_1} j)$). $\quad\square$

Corollary 6.3, which will be useful in the final section of the paper, and Corollary 4.3 are different examples of the CSR decomposition schemes considered by Sergeev and Schneider [24].

If $\mathcal{D}$ is the graph defining B in (12), it contains $\mathcal{G}^c(A)$ and by the optimal walk interpretation (6), we have:

$$
\begin{aligned}
a_{ij}^{(t)} &= b_{ij}^{(t)} \oplus p\big(\mathcal{W}^t(i \xrightarrow{\mathcal{D}} j)\big) \\
&= b_{ij}^{(t)} \oplus p\big(\mathcal{W}^t(i \xrightarrow{\mathcal{G}^c(A)} j)\big) \oplus p\big(\mathcal{W}^t(i \xrightarrow{\mathcal{D}} j) \setminus \mathcal{W}^t(i \xrightarrow{\mathcal{G}^c(A)} j)\big).
\end{aligned}
$$

The proof that $T_1(A, B) \leq T$ has two parts:

1. Scheme-dependent part: shows that for $t \geq T$ we have

$$p\big(\mathcal{W}^t(i \xrightarrow{\mathcal{D}} j) \setminus \mathcal{W}^t(i \xrightarrow{\mathcal{G}^c(A)} j)\big) \leq p\big(\mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{G}^c(A)} j)\big). \qquad (32)$$

2. Scheme-independent part: shows that for $t \geq T$ we have

$$p\big(\mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{G}^c(A)} j)\big) \leq p\big(\mathcal{W}^t(i \xrightarrow{\mathcal{G}^c(A)} j)\big). \qquad (33)$$

By Theorem 6.1, we have $p(\mathcal{W}^t(i \xrightarrow{\mathcal{G}^c(A)} j)) \leq p(\mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{G}^c(A)} j)) = (CS^t R)_{ij}$. Thus, (32) implies $A^t \leq B^t \oplus CS^t R$, while (33) implies $A^t \geq B^t \oplus CS^t R$.

Let us go deeper into the strategy for each part.

1. The scheme-dependent part goes as follows:
   (a) For $B = B_{\mathrm{N}}$, $\mathcal{D} = \mathcal{G}^c(A)$ and there is nothing to prove.
   (b) For $B = B_{\mathrm{HA}}$, we take a walk W with maximal weight in $\mathcal{W}^t(i \xrightarrow{\mathcal{D}} j) \setminus \mathcal{W}^t(i \xrightarrow{\mathcal{G}^c(A)} j)$ and a closed walk V from a node of W to $\mathcal{G}^c(A)$ and back whose edges have weight greater than or equal to the greatest weight of the edges in W. Then, we insert $V^{\gamma(\mathcal{G}^c(A))}$ (i.e., $\gamma(\mathcal{G}^c(A))$ copies of V) in W, and remove as many cycles of the new walk as possible, preserving the length modulo $\gamma(\mathcal{G}^c(A))$ until we get a walk $\tilde{W}$ with length at most t.

As a result, we thus replaced some edges of W by edges with greater weight and removed other edges, so $p(\tilde{W}) \geq p(W)$.

(c) For $B = B_{\mathrm{CT}}$, we also take a walk W with maximal weight in $\mathcal{W}^t(i \xrightarrow{\mathcal{D}} j) \setminus \mathcal{W}^t(i \xrightarrow{\mathcal{G}^c(A)} j)$ but now we replace some cycles of W by some copies of a cycle with greater mean weight, to get a new walk with length t. We therefore introduce the concept of a "staircase" of cycles, and Lemma 8.4 will ensure us that we can iterate this process and eventually reach a critical node.

Note that we need to remove cycles before we replace them and to have some steps with non-critical cycles, which explains why the bound for $T_1(A, B_{\mathrm{CT}})$ are larger than the one for $T_1(A, B_{\mathrm{HA}})$ and $T_1(A, B_{\mathrm{N}})$. However, the worst case remains $\mathrm{Wi}(n)$.

2. By Theorem 6.1, to have (33), it is enough to prove that for each s.c.c. H of $\mathcal{G}^c(A)$ there is a $\gamma \in \mathbb{N}$ and a set of nodes $\mathcal{N} \subset H$ such that

$$p\big(\mathcal{W}^t(i \xrightarrow{H} j)\big) \geq p\big(\mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{N}} j)\big). \tag{34}$$

To ensure that Eq. (34) is satisfied, we use the following steps:

(a) For each s.c.c. H of $\mathcal{G}^c(A)$, choose $\mathcal{N} \subset H$ and γ a multiple of $\gamma(H)$ and take a walk W such that $p(W) = p(\mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{N}} j))$.

(b) Remove as many cycles as possible from W, keeping it in $\mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{N}} j)$.

(c) Insert critical cycles so that the new walk has length t.

Since $\lambda(A) \leq \mathbf{1}$, steps 2(b) and 2(c) cannot strictly increase the weight of the walk, so (34) is satisfied.

It is clear from the strategy that the main point is to remove cycles from a given walk, while preserving the length modulo some given integer. This will be the subject of Section 9. We will use three different tactics, one of them is completely new. Different bounds depending on different parameters arise from different choices of $\mathcal{N}$ and γ in step 2(a) and different tactics in step 2(b). To reach the (optimal) Wielandt number $\mathrm{Wi}(n)$, we have to combine two of them.

To state general results, we introduce two graph-theoretic quantities.

Definition 6.4. Let $\mathcal{D}$ be a subgraph of $\mathcal{D}(A)$ and $\gamma \in \mathbb{N}$.

1. The *cycle removal threshold* $T^\gamma_{cr}(\mathcal{G})$ (resp. the *strict cycle removal threshold* $\tilde{T}^\gamma_{cr}(\mathcal{G})$) of $\mathcal{G}$ is the smallest nonnegative integer T for which the following holds: for all walks $W \in \mathcal{W}(i \xrightarrow{\mathcal{G}} j)$ with length $\geq T$, there is a walk $V \in \mathcal{W}(i \xrightarrow{\mathcal{G}} j)$ obtained from W by removing cycles (resp. at least one cycle) and possible inserting cycles of $\mathcal{G}$ such that $l(V) \equiv l(W) \pmod{\gamma}$ and $l(V) \leq T$.

2. The *exploration penalty* $\mathrm{ep}^\gamma(i)$ of a node $i \in \mathcal{G}^c(A)$ is the least $T \in \mathbb{N}$ such that for any multiple t of γ greater than or equal to T, there is a closed walk on $\mathcal{G}^c(A)$ with length t starting at i.

The *exploration penalty* $\mathrm{ep}^\gamma(\mathcal{G})$ of $\mathcal{G} \subseteq \mathcal{G}^c(A)$ is the maximum of the $\mathrm{ep}^\gamma(i)$ for $i \in \mathcal{G}$. We further set $\mathrm{ep}(\mathcal{G}^c(A)) = \max_l \mathrm{ep}^{\gamma(\mathcal{G}_l^c)}(\mathcal{G}_l^c)$, which is the quantity used in Theorem 4.1.

Obviously, $T_{cr}^\gamma(\mathcal{G}) \leq \tilde{T}_{cr}^\gamma(\mathcal{G}) \leq T_{cr}^\gamma(\mathcal{G}) + 1$ but it will be useful to have both definitions. Bounds on ep^γ are given in Section 11 while T_{cr}^γ is investigated in Section 9. We can already notice the following.

First, $\mathrm{ep}^\gamma(i)$ is finite if and only if γ is a multiple of the cyclicity of its s.c.c. in $\mathcal{G}^c(A)$. Second, if γ is multiplied by an integer, then $\mathrm{ep}^\gamma(i)$ decreases but $T_{cr}^\gamma(\mathcal{G})$ increases. Third, for fixed γ, $\mathrm{ep}^\gamma(\mathcal{G})$ decreases when $\mathcal{G}$ increases. Finally, $\mathrm{ep}^\gamma(i) = 0$ if and only if there is a critical closed walk with length γ at i. In particular, for any cycle Z, we have

$$\mathrm{ep}^{l(Z)}(Z) = 0.$$

This gives two extremal choices for $\mathcal{G}$ and γ: either $\mathcal{G}$ is an s.c.c. of $\mathcal{G}^c(A)$ and γ is its cyclicity, or $\mathcal{G}$ is a critical cycle and γ is its length.

The first choice is used in [4], the second one in [14] and both choices in [8]. Here we systematically test those two choices. The first one is used to prove the bounds in Theorem 4.1 that depend on $\mathrm{ep}(\mathcal{G}^c(A))$. The second one is used for the other bounds on T_1.

If other choices prove to be useful under additional assumptions on $\mathcal{D}(A)$, one can apply Proposition 6.5 with other parameters.

The strategy explained in this section leads to the following proposition, which implies Theorems 4.1 and 4.4, except for (20).

Proposition 6.5 *(From cycle removal to weak CSR). Let A be a square matrix and $\mathcal{G}$ be a representing subgraph of $\mathcal{G}^c(A)$ with s.c.c.'s $\mathcal{G}_1, \cdots, \mathcal{G}_m$, and let γ_l be multiples of $\gamma(\mathcal{G}_l)$.*

(i) *If $B \in \{B_N, B_{HA}\}$, then $T_1(A, B) \leq \max_l(T_{cr}^{\gamma_l}(\mathcal{G}_l) - \gamma_l + 1 + \mathrm{ep}^{\gamma_l}(\mathcal{G}_l))$.*
(ii) *If $B = B_{CT}$, then $T_1(A, B) \leq \max\{\tilde{T}_{cr}^{l(Z)}(Z) \mid Z \text{ cycle in } \mathcal{G}^{ct}\}$.*

This proposition is proved in Section 8. The bounds of Theorem 4.1 with $\hat{\gamma}(\mathcal{G}^c(A))$ and $\mathrm{ep}(\mathcal{G}^c(A))$ can be improved if one knows more of the structure of $\mathcal{G}^c(A)$.

7. Proof of Theorem 6.1

Let $A, \mathcal{G}, \mathcal{N}, \gamma, t, i, j$ be as in the statement of Theorem 6.1.
We first prove:

$$\left(CS^t R\right)_{ij} \leq p\left(\mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{N}} j)\right). \tag{35}$$

By definition of C, S and R, there are walks W_1, W_2 and W_3 such that $(CS^t R)_{ij} = p(W_1 W_2 W_3)$ and

$$l(W_1) \equiv l(W_3) \equiv 0 \pmod{\gamma(\mathcal{G})}, \qquad W_2 \subset \mathcal{G} \quad \text{and} \quad l(W_2) = t. \tag{36}$$

Let k be the start node of W_2. By hypotheses, k is critical and there is a node l of $\mathcal{N}$ in the same s.c.c. H of $\mathcal{G}^c(A)$ as k. Thus there are walks W_4 and W_5 with only critical edges, going from k to l and from l to k respectively. Thus, $W_4 W_5$ is a circuit of $\mathcal{G}^c(A)$ and $p(W_4) + p(W_5) = 0$.

Let G be the s.c.c. containing k in $\mathcal{G}$. As $G \subseteq H$, $\gamma(H)$ divides $\gamma(G)$, thus also $\gamma(\mathcal{G})$ and γ. Hence $\gamma(H)$ divides $l(W_1)$ and $l(W_3)$. It also divides $l(W_4 W_5)$ and we have

$$L = l(W_1) + l(W_3) + l(W_4) + l(W_5) \equiv 0 \quad (\text{mod } \gamma(H)).$$

Therefore, for $m \in \mathbb{N}$ large enough, there is a closed walk W_6 on H starting at k with length $m\gamma - L$.

Set $W = W_1 W_4 W_6 W_5 W_2 W_3$. By construction $W \in \mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{N}} j)$ and $p(W) = p(W_1 W_2 W_3) = (C S^t R)_{ij}$, so (35) is proved.

It remains to show:

$$\left(C S^t R\right)_{ij} \geq p\big(\mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{N}} j)\big). \tag{37}$$

By definition of $\mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{N}} j)$ there are a node $l \in \mathcal{N}$ and two walks V_1 and V_2 going from i to l and from l to j respectively such that $l(V_1 V_2) \equiv t \pmod{\gamma}$ and $p(V_1) + p(V_2) = p(\mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{N}} j))$.

Let k be a node of $\mathcal{G}$ in the same s.c.c. H of $\mathcal{G}^c(A)$ as l. As above, there are critical walks W_4 and W_5, going from k to l and from l to k respectively and $\gamma(H)$ divides γ.

Let V_3 be a closed walk in $\mathcal{G}$ with start node k, whose length is $\geq t + \gamma$. Let V_4 be its shortest prefix such that $l(V_1) + l(W_5) + l(V_4) \equiv 0 \pmod{\gamma}$ and V_5 be the complementary (i.e., $V_3 = V_4 V_5$). Let W_2 be the prefix of length t of V_5 and V_6 be its complementary ($V_5 = W_2 V_6$, $V_3 = V_4 W_2 V_6$).

Set $W_1 = V_1 W_5 V_4$ and $W_3 = V_6 V_3^{(\gamma-1)} (W_4 W_5)^{(\gamma-1)} W_4 V_2$. By construction W_1, W_2 satisfy (36). Moreover, we have

$$W_1 W_2 W_3 = V_1 W_5 V_4 W_2 V_6 V_3^{(\gamma-1)} (W_4 W_5)^{(\gamma-1)} W_4 V_2$$
$$= V_1 W_5 V_3^{\gamma} (W_4 W_5)^{(\gamma-1)} W_4 V_2$$

so $l(W_1 W_2 W_3) \equiv l(V_1) + l(V_2) \equiv 0 \pmod{\gamma}$ and W_3 also satisfies (36).

On the other hand $W_5 V_3^{\gamma} (W_4 W_5)^{(\gamma-1)} W_4$ is a critical closed walk, so it has weight 0 and $p(W_1 W_2 W_3) = p(V_1) + p(V_2) = p(\mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{N}} j))$, so (37) is proved.

8. Proof of Proposition 6.5

In this section, we prove Proposition 6.5, following the strategy described in Section 6.

8.1. Scheme independent part

In this section, we prove the following lemma.

Lemma 8.1 *(Scheme independent part). Let A be a square matrix with $\lambda(A) = 0$ and $\mathcal{G}$ be a representing subgraph of $\mathcal{G}^c(A)$ with s.c.c. $\mathcal{G}_1, \ldots, \mathcal{G}_m$ and γ_l be multiples of $\gamma(\mathcal{G}_l)$, for $l = 1, \ldots, m$.*

For any $t \geq \max_l(T_{cr}^{\gamma_l}(\mathcal{G}_l) - \gamma_l + 1 + \mathrm{ep}^{\gamma_l}(\mathcal{G}_l))$ and any i, j, inequality (33)*, with $\mathcal{G}$ instead of $\mathcal{G}^c(A)$, holds for $\gamma = \mathrm{lcm}_l \, \gamma_l$.*

Proof. Indeed, any walk $W \in \mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{G}} j)$ is in $\mathcal{W}^{t,\gamma_l}(i \xrightarrow{\mathcal{G}_l} j)$ for some l. By definition of $T_{cr}^{\gamma_l}(\mathcal{G}_l)$, there is a walk $V \in \mathcal{W}^{t,\gamma_l}(i \xrightarrow{\mathcal{G}_l} j)$ with length at most $T_{cr}^{\gamma_l}(\mathcal{G}_l)$ and $p(V) \geq p(W)$.

If $t \geq T_{cr}^{\gamma_l}(\mathcal{G}_l) - \gamma_l + 1 + \mathrm{ep}^{\gamma_l}(\mathcal{G}_l)$, then $t - l(V) \geq \mathrm{ep}^{\gamma_l}(\mathcal{G}_l) - \gamma_l + 1$. Since $t - l(V)$ and $\mathrm{ep}^{\gamma_l}(\mathcal{G}_l)$ are both multiples of γ_l, it implies $t - l(V) \geq \mathrm{ep}^{\gamma_l}(\mathcal{G}_l)$, so there is a closed walk on $\mathcal{G}^c(A)$ with length $t - l(V)$ at each node of $\mathcal{G}_l$. Inserting such a walk in V where it reaches $\mathcal{G}_l$, we get a new walk $\tilde{W} \in \mathcal{W}^t(i \xrightarrow{\mathcal{G}_l} j) \subseteq \mathcal{W}^t(i \xrightarrow{\mathcal{G}} j)$ with $p(\tilde{W}) = p(V) \geq p(W)$. $\square$

8.2. Hartmann and Arguelles scheme

In this section, we perform step 1 of the strategy in the case $B = B_{\mathrm{HA}}$. We prove the following lemma.

Lemma 8.2. *Let A be a square matrix with $\lambda(A) = 0$ and $\mathcal{G}$ be a representing subgraph of $\mathcal{G}^c(A)$ with s.c.c. $\mathcal{G}_1, \cdots, \mathcal{G}_m$ and γ_l be multiples of $\gamma(\mathcal{G}_l)$ for $l = 1, \ldots, m$.*

For any $t \geq \max_l(T_{cr}^{\gamma_l}(\mathcal{G}_l) - \gamma_l + 1)$ and any i, j, inequality (32) *holds with $\gamma = \mathrm{lcm}_l \, \gamma_l$ and $\mathcal{D} = \mathcal{G}^{ha}$ (the graph defining B_{HA} in Section 3).*

Proof. We assume without loss of generality that A is max-balanced.

Let W be a walk with maximal weight in $\mathcal{W}^t(i \xrightarrow{\mathcal{D}} j) \setminus \mathcal{W}^t(i \xrightarrow{\mathcal{G}^c(A)} j)$. We show that there exists a walk $\tilde{W} \in \mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{G}^c(A)} j)$ with $p(\tilde{W}) \geq p(W)$.

Denote the maximum weight of edges in W by $\mu(W)$. Define the graph

$$\tilde{\mathcal{D}} := \begin{cases} \mathcal{G}^{ha} & \text{if } \mu(W) \leq \mu^{ha}, \\ \mathcal{T}^{ha}(\mu(W)) & \text{otherwise.} \end{cases} \tag{38}$$

By the definition of Hartmann–Arguelles threshold graphs, $\mathcal{G}^c(A) \subseteq \tilde{\mathcal{D}} \subseteq \mathcal{G}^{ha}$. In both cases of (38), walk W contains a node k of digraph $\tilde{\mathcal{D}}$, which is completely reducible (due to the max-balancing).

Let $W = W_1 \cdot W_2$ with W_1 ending at node k. By definition of $\tilde{\mathcal{D}}$, there exists a critical node ℓ in the same s.c.c. H of $\tilde{\mathcal{D}}$ as k. Moreover, H contains a whole s.c.c. $\mathcal{G}_l^c(A)$ of $\mathcal{G}^c(A)$, and hence also a component $\mathcal{G}_l$ of the representing subgraph $\mathcal{G}$. Hence we can choose ℓ in $\mathcal{G}_l$.

Let V_1 be a walk in $\tilde{\mathcal{D}}$ from k to ℓ and V_2 be a walk in $\tilde{\mathcal{D}}$ from ℓ to k. Set $V = V_1 V_2$ and $W_3 = W_1 \cdot V^{\gamma_l} \cdot W_2$. By the definition of the cycle replacement threshold, there

exists a walk $\tilde{W} \in \mathcal{W}^{t,\gamma_l}(i \xrightarrow{\mathcal{G}_l} j)$ obtained from W_3 by removing cycles and possibly inserting cycles from $\mathcal{G}_l$ such that $l(\tilde{W}) \leq T_{cr}^{\gamma_l}(\mathcal{G}_l) \leq t + \gamma_l - 1$. Since $l(W_3) \equiv t \pmod{\gamma_l}$, it implies $l(\tilde{W}) \leq t$.

Recall that since A is max-balanced and $\lambda(A) = 0$, all edges have nonpositive weights, and the weight of each edge of $\tilde{\mathcal{D}}$ is not smaller than that of any edge of W. Each edge of W is either removed, kept or replaced by an edge of $\tilde{\mathcal{D}}$ in $\tilde{W}$, thus we conclude that $p(\tilde{W}) \geq p(W)$. This shows

$$p\big(\mathcal{W}^t(i \xrightarrow{\mathcal{D}} j) \setminus \mathcal{W}^t(i \xrightarrow{\mathcal{G}^c(A)} j)\big) \leq \max_l p\big(\mathcal{W}^{t,\gamma_l}(i \xrightarrow{\mathcal{G}_l} j)\big).$$

However, Theorem 6.1 implies that

$$p\big(\mathcal{W}^{t,\gamma_l}(i \xrightarrow{\mathcal{G}_l} j)\big) = p\big(\mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{G}_l^c(A)} j)\big)$$

for each l and hence

$$\max_l p\big(\mathcal{W}^{t,\gamma_l}(i \xrightarrow{\mathcal{G}_l} j)\big) = \max_l p\big(\mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{G}_l^c(A)} j)\big) = p\big(\mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{G}^c(A)} j)\big),$$

and this concludes the proof. $\square$

Proposition 6.5(i) now follows from Lemmas 8.1 and 8.2.

8.3. Cycle threshold scheme

In this section, we perform step 1 of the strategy in the case $B = B_{\mathrm{CT}}$. We prove the following lemma.

Lemma 8.3. *Let A be a square matrix with $\lambda(A) = 0$.*

For any $t \geq \max\{\tilde{T}_{cr}^{l(Z)}(Z) | Z \text{ cycle of } \mathcal{G}^{ct}\}$ and any i, j, inequality (32) holds with $\gamma = \gamma(\mathcal{G}^c(A))$ and $\mathcal{D} = \mathcal{G}^{ct}$ the graph defining B_{CT} in Section 3.

A finite sequence of cycles $Z_1, \ldots, Z_m$ in $\mathcal{G}$ is called a *staircase* in $\mathcal{G}$ if, for all $1 \leq s \leq m-1$, Z_s and Z_{s+1} share a node, $p(Z_s)/l(Z_s) \leq p(Z_{s+1})/l(Z_{s+1})$ and, moreover, the cycle mean of Z_{s+1} is the greatest among all the cycles sharing a node with Z_s.

Lemma 8.4. *Let $\mu > \mu^{ct}$ and Z be a cycle in $\mathcal{T}^{ct}(\mu)$ or $\mu = \mu^{ct}$ and Z be a cycle in $\mathcal{G}^{ct}(\mu)$ with $p(Z)/l(Z) = \mu$. Then there exists a staircase $Z_1, \ldots, Z_m$ in $\mathcal{T}^{ct}(\mu)$ such that $Z_1 = Z$ and Z_m is critical.*

Proof. Suppose by contradiction that no such staircase exists. Let $Z_1, \ldots, Z_m$ be a staircase in $\mathcal{T}^{ct}(\mu)$ such that $Z_1 = Z$ and $p(Z_m)/l(Z_m)$ is maximal.

Denote $\mu' = p(Z_m)/l(Z_m)$, so $\mu' < \lambda(A)$. If the s.c.c. of $\mathcal{T}^{ct}(\mu')$, in which Z_m lies, contains a cycle of mean weight strictly greater than μ', then we can build a staircase

with a greater cycle mean of the final cycle, a contradiction. If that component of $\mathcal{T}^{ct}(\mu')$ does not contain a cycle of mean weight strictly greater than μ', this is a contradiction to the definition of μ^{ct} and the fact that $\mu' \geq \mu^{ct}$. Thus we must have $\mu' = \lambda(A)$. $\quad\square$

Proof of Lemma 8.3 and Proposition 6.5(ii). Let $t \geq \max_Z \tilde{T}_{cr}(Z)$ and let $W \in \mathcal{W}^t(i \rightarrow j)$ visiting a node of $\mathcal{G}^{ct}$ but no critical node.

Denote by $\nu(W)$ the largest cycle mean of subcycles of W. We assume in the following that $\nu(W)$ is maximal among all $W \in \mathcal{W}^t(i \rightarrow j)$ with $p(W) = a_{ij}^{(t)}$. We prove Lemma 8.3 by showing $\nu(W) = \lambda(A)$. Assume that $\nu(W) < \lambda(A)$, and define

$$\tilde{\mathcal{D}} := \begin{cases} \mathcal{G}^{ct} & \text{if } \nu(W) \leq \mu^{ct}, \\ \mathcal{T}^{ct}(\nu(W)) & \text{otherwise.} \end{cases} \tag{39}$$

By the definition of cycle threshold graphs, $\mathcal{G}^c(A) \subseteq \tilde{\mathcal{D}} \subseteq \mathcal{G}^{ct}$.

By Lemma 8.4, there exists a staircase $Z_1, \ldots, Z_m$ in $\tilde{\mathcal{D}}$ such that Z_1 has $p(Z_1) = \nu(W)$ and shares a node with W, and Z_m is critical. We inductively define walks $W_0, \ldots, W_m$ as follows: Set $W_0 = W$. For $1 \leq \ell \leq m$, let $\mathcal{G}$ be the subgraph of $\mathcal{D}(A)$ induced by Z_ℓ. By definition of $\tilde{T}_{cr}$, there is a walk $V \in \mathcal{W}^{t,l(Z_\ell)}(i \xrightarrow{Z_\ell} j)$ obtained from $W_{\ell-1}$ by removing at least one cycle and inserting at least one cycle in $\mathcal{G}$ (i.e., one copy of Z_ℓ) such that $l(V) \leq \tilde{T}_{cr}^{l(Z_\ell)}(Z_\ell) \leq t$. Now define W_ℓ as walk V after inserting enough copies of Z_ℓ, to have $l(W_\ell) = t$. Thus Z_ℓ is a subwalk of W_ℓ for all ℓ, and walk W_m contains a critical node.

We now show that $p(W_\ell) \geq p(W_{\ell-1})$ on each step. For this we will prove by induction that, on each step, the mean weight of $Z_{\ell+1}$ is not less than that of any cycle (and hence closed walk) in W_ℓ. The base of induction ($\ell = 0$) is due to the definition of $\tilde{\mathcal{D}}$. In general, observe that the cycles in W_ℓ are 1) Z_ℓ and cycles using the edges of Z_ℓ, 2) cycles that were already in $W_{\ell-1}$. For the latter cycles we use the inductive assumption, while the cycles using edges of Z_ℓ share a common node with it and hence their mean weight does not exceed that of $Z_{\ell+1}$ by the definition of staircase.

Setting $\tilde{W} = W_m$ we obtain $\tilde{W} \in \mathcal{W}^t(i \xrightarrow{\mathcal{G}^c(A)} j)$ and $p(\tilde{W}) \geq p(W)$, thus Lemma 8.3 and Proposition 6.5(ii) are proved. $\quad\square$

9. Cycle removal

9.1. Cycle removal threshold

In this section, we state some bounds on $T_{cr}^\gamma(\mathcal{G})$ for some subgraphs $\mathcal{G}$ of $\mathcal{D}(A)$. Those bounds are achieved by three different methods, one of them is new. Recall $\mathrm{cr}(\mathcal{D}(A))$, $\mathrm{cd}(\mathcal{D}(A))$ and other parameters (Section 4).

Let us now recall an elementary application of the pigeonhole principle. The origins of this lemma were briefly discussed by Aigner and Ziegler [1, p. 133]. In the context of max-algebraic matrix powers, it was considered for the first time by Hartmann and Arguelles [14]. It is in the heart of almost all of our cycle reductions.
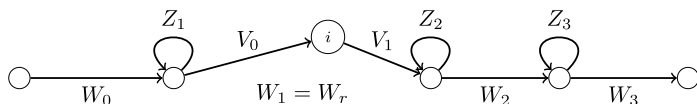
Fig. 2. Walk W in the proof of Proposition 9.2 ($m = 3$ and $r = 1$).

Lemma 9.1. *Let $a_1, \ldots, a_m$ be integers. Then there exists a nonempty subset $I \subseteq \{1, \ldots, m\}$ of indices such that the sum $\sum_{i \in I} a_i$ is a multiple of m.*

One of the bounds that we use is in fact proved in [7] (see also [8, Theorem 2]). The proof is recalled for the reader's convenience.

Proposition 9.2. *(See Lemma 20 of [7].) For any $A \in \mathbb{R}_{\max}^{n \times n}$, any node i and any integer γ, we have:*

$$T_{cr}^{\gamma}\big(\{i\}\big) \leq (\gamma - 1)\,\mathrm{cr} + (\gamma + 1)\,\mathrm{cd}, \tag{40}$$

where $\mathrm{cd} = \mathrm{cd}(\mathcal{D}(A))$ *and* $\mathrm{cr} = \mathrm{cr}(\mathcal{D}(A))$.

Proof. Let W be a walk going through i. Write this walk as $W = W_0 \cdot Z_1 \cdots Z_m \cdot W_m$ where (i) all Z_s are nonempty cycles, (ii) node i is a node of the walk W_r, and (iii) m is maximal. Write also $W_r = V_0 V_1$ so that i is the end of V_0 and the start of V_1. The whole configuration is shown in Fig. 2.

If a subset $S \subseteq \{1, \ldots, m\}$ of indices such that γ divides $\sum_{s \in S} l(Z_s)$ cannot be chosen, then by Lemma 9.1 $m < \gamma - 1$, and the walks

$$W_1, \ldots, W_{r-1}, V_0, V_1, W_{r+1}, \ldots, W_m$$

are paths (otherwise m is not maximal), which implies that $l(W) \leq (\gamma - 1)\,\mathrm{cr} + (\gamma + 1)\,\mathrm{cd}$.

If $l(W) > (\gamma - 1)\,\mathrm{cr} + (\gamma + 1)\,\mathrm{cd}$, then such a subset of cycles can be chosen, and a strictly shorter subwalk of the same length modulo γ is obtained by cycle deletion, hence the claim. $\quad\square$

Proposition 9.2 implies that $T_{cr}^{l(Z)}(Z) \leq (l(Z) - 1)\,\mathrm{cr} + (l(Z) + 1)\,\mathrm{cd}$ and $T_{cr}^{\gamma(\mathcal{G}^c)}(\mathcal{G}^c) \leq 2\gamma(\mathcal{G}^c)(n-1) + \gamma(\mathcal{G}^c) - 1$ for any s.c.c. $\mathcal{G}^c \subset \mathcal{G}^c(A)$ but both bounds can be improved using various methods.

The first bound is improved in Section 9.2, following a method used in [14], which leads to:

Proposition 9.3. *For $A \in \mathbb{R}_{\max}^{n \times n}$, Z a cycle of $\mathcal{D}(A)$ and γ a divisor of $l(Z)$, we have:*

$$T_{cr}^{\gamma}(Z) \leq \big(n - 1 - l(Z) + \gamma\big)\,\mathrm{cr} + \mathrm{cd} + l(Z), \tag{41}$$

where $\mathrm{cd} = \mathrm{cd}(\mathcal{D}(A))$ *and* $\mathrm{cr} = \mathrm{cr}(\mathcal{D}(A))$.

Table 1
Expressions of Proposition 6.5 (with $l(Z)$).

Prop.	$T_{cr}^{l(Z)}(Z) - l(Z) + 1$	$\tilde{T}_{cr}^{l(Z)}(Z)$
9.2	$(l(Z) - 1)(\mathrm{cr} - 1) + (l(Z) + 1)\,\mathrm{cd}$	$(l(Z) - 1)\,\mathrm{cr} + (l(Z) + 1)\,\mathrm{cd} + 1$
9.3	$(n - 1)\,\mathrm{cr} + \mathrm{cd} + 1$	$n\,\mathrm{cr} + \mathrm{cd} + 1$
9.4	$\mathrm{Wi}(n)$	$n^2 - n + 1$
9.5	$l(Z)(n - 2) + n$	$l(Z)(n - 1) + n$

Table 2
Expressions of Proposition 6.5 (with γ).

Prop.	$T_{cr}^{\gamma(\mathcal{G}^c)}(\mathcal{G}^c) - \gamma(\mathcal{G}^c) + 1$
9.2	$(\gamma - 1)(\mathrm{cr} - 1) + (\gamma + 1)\,\mathrm{cd}$
9.5	$\gamma(n - 1) + n - |\mathcal{G}^c|$

Table 3
How to deduce the bounds on T_1.

Bound on $T_1(A, B)$	$\mathcal{N}$	γ	Prop.
(15)	Z s.t. $l(Z) = \mathrm{g}(\mathcal{G}^c)$	$l(Z) = \mathrm{g}(\mathcal{G}^c)$	9.4, 9.5
(16)	Z s.t. $l(Z) = \mathrm{g}(\mathcal{G}^c)$	$l(Z) = \mathrm{g}(\mathcal{G}^c)$	9.5
(17)	$i \in Z$ s.t. $l(Z) = \mathrm{g}(\mathcal{G}^c)$	$l(Z) = \mathrm{g}(\mathcal{G}^c)$	9.2
(18)	$\mathcal{G}^c$	$\gamma(\mathcal{G}^c)$	9.5
(19)	$i \in \mathcal{G}^c$	$\gamma(\mathcal{G}^c)$	9.2
(21)	Z in staircase or Z critical	$l(Z)$	9.3, 9.5
(22)	any i in any Z	$l(Z)$	9.2

This method also leads to:

Proposition 9.4. *For $A \in \mathbb{R}_{\max}^{n \times n}$ and Z a cycle with length n of $\mathcal{D}(A)$, we have $\tilde{T}_{cr}^n(Z) \leq n^2 - n + 1$.*

The second bound is improved in Section 9.3 thanks to a new method, which leads to:

Proposition 9.5. *For $A \in \mathbb{R}_{\max}^{n \times n}$ and $\mathcal{G}$ a subgraph of $\mathcal{D}(A)$ with n_1 nodes, we have:*

$$\forall \gamma \in \mathbb{N}, \quad T_{cr}^\gamma(\mathcal{G}) \leq \gamma n + n - n_1 - 1.$$

Tables 1 and 2 show the bounds obtained for a critical cycle Z or an s.c.c. $\mathcal{G}^c$ of $\mathcal{G}^c(A)$. Here the first column contains proposition number, $\gamma = \gamma(\mathcal{G}^c)$, and other parameters refer to $\mathcal{D}(A)$. Note that $l(Z) = n$ in the case of Proposition 9.4.

Proof of Theorems 4.1 and 4.4. Theorems 4.1 and 4.4 are combinations of the bounds in Tables 1 and 2 with Proposition 6.5. For each s.c.c. $\mathcal{G}^c$ of $\mathcal{G}^c(A)$, Table 3 explains which choices of $\mathcal{N}$, γ and proposition to bound $T_{cr}^\gamma(\mathcal{N})$ should be made.

To obtain bounds (15)–(17) we take, for the representing subgraph $\mathcal{G}$ in Proposition 6.5, any collection of critical cycles such that each s.c.c. of $\mathcal{G}^c(A)$ contains exactly

one cycle of the collection and each cycle has the minimal length in the corresponding s.c.c. In the case of (18) and (19), we set $\mathcal{G} = \mathcal{G}^c(A)$. Bounds (21) and (22) can be obtained from the last column of Table 3. Note that (21) is obtained as the minimum of two bounds.

The only difficult case is bound (20). Indeed, in the worst case, cycle Z with length n, we only get $\tilde{T}^n_{cr}(Z) \leq n^2 - n + 1$ by Proposition 9.4 instead of $\mathrm{Wi}(n)$. Thus, Proposition 6.5 would give $T_1 \leq n^2 - n + 1$ instead of $T_1 \leq \mathrm{Wi}(n)$ and we have to go into more details. The proof of (20) is thus postponed to the end of the next subsection. $\quad\square$

9.2. Cycle removal by cycle decomposition

In this section, we present and improve the method of [14] to prove Propositions 9.3 and 9.4. It will also be used to prove that $T_1(A, B_{\mathrm{CT}}) \leq \mathrm{Wi}(n)$ (Eq. (20)) at the end of the next subsection. For any set of walks W_α with $\alpha \in S$ for S a subset of natural numbers, let us denote by $\mathcal{G}(\bigcup_{\alpha \in S} W_\alpha)$ the subgraph of $\mathcal{D}(A)$ consisting of all nodes and edges that belong to some walk W_α, $\alpha \in S$.

Proof of Propositions 9.3 and 9.4. To any walk $W \in \mathcal{W}(i \xrightarrow{Z} j)$, we apply the following procedure, adapted from [14].

1. We choose a decomposition of the walk $W \in \mathcal{W}(i \to j)$ into a path P and a set of cycles Z_α for $\alpha \in S$ (with S a subset of natural numbers). Note that P may be empty. If it is, walk W is closed. Then, it has the same start and end node.
   We denote by n_W the number of nodes that appear at least once in W and by cd_W the maximum length of an acyclic walk whose edges belong to W.
2. We take a subset R_1 of S with $|R_1| \leq n - l(Z)$ such that $\mathcal{G}(P \cup Z \bigcup_{\alpha \in R_1} Z_\alpha)$ is connected and contains all nodes appearing in W. This is possible because the connection of $\mathcal{G}(P \cup Z)$ with all the nodes of W can be ensured by adding at most $n - l(Z)$ edges of W to $P \cup Z$, and hence by adding to it at most $n - l(Z)$ cycles Z_α, for $\alpha \in S$.
3. Let R_2 be a result of recursively removing from $S \setminus R_1$ sets of indices whose corresponding cycles have a combined length that is a multiple of γ.
   By Lemma 9.1, $|R_2| \leq \gamma - 1$. Let R be $R_1 \cup R_2$.
   Set $\mathrm{cr}_W = \max_{\alpha \in R} l(Z_\alpha)$ (circumference of the walk W).
4. If $\mathcal{G}_0 = \mathcal{G}(P \cup \bigcup_{\alpha \in R} Z_\alpha)$ is connected, then we build a walk $V \in \mathcal{W}(i \xrightarrow{Z} j)$ by starting from P and successively inserting (in some order) all cycles Z_α with $\alpha \in R$.
5. Otherwise, we build $V \in \mathcal{W}(i \xrightarrow{Z} j)$ by starting from P and successively inserting (in some order) all cycles Z_α with $\alpha \in R$, and Z.

By construction, $l(V) \equiv l(W) \pmod{\gamma}$ in both cases. Let us bound the length of W. If $\mathcal{G}_0$ is connected,

$$l(V) = l(P) + \sum_{\alpha \in R} l(Z_\alpha) \le \mathrm{cd}_W + \mathrm{cr}_W \big(n - l(Z) + \gamma - 1\big)$$

$$= \mathrm{cr}_W \big(n - l(Z) + \gamma - 2\big) + (\mathrm{cd}_W + \mathrm{cr}_W) \qquad (42)$$

If $\mathcal{G}_0$ is not connected, we have $l(V) \le \mathrm{cr}_W(n - l(Z) + \gamma - 2) + (\mathrm{cd}_W + \mathrm{cr}_W) + l(Z)$.

But there is some $\hat{\alpha} \in R$ such that $l(P) + l(Z_{\hat{\alpha}}) \le n_W - 1$, because otherwise every Z_α with $\alpha \in R$ would share a node with P. Because $|R \setminus \{\hat{\alpha}\}| \le n - l(Z) + \gamma - 2$, we have

$$l(V) = l(Z) + l(P) + l(Z_{\hat{\alpha}}) + \sum_{\substack{\alpha \in R \\ \alpha \ne \hat{\alpha}}} l(Z_\alpha)$$

$$\le l(Z) + (n_W - 1) + \big(n - l(Z) + \gamma - 2\big)\, \mathrm{cr}_W \qquad (43)$$

Finally, we have

$$l(V) \le l(Z) + \big(n - l(Z) + \gamma - 2\big)\, \mathrm{cr}_W + \min(n_W - 1, \mathrm{cr}_W + \mathrm{cd}_W)$$

if M_0 is not connected, and

$$l(V) \le l(Z) + \big(n - l(Z) + \gamma - 2\big)\, \mathrm{cr}_W + \big(\mathrm{cr}_W + \mathrm{cd}_W - l(Z)\big).$$

This gives the following

Lemma 9.6. *For any cycle Z, any divisor γ of $l(Z)$ and any walk $W \in \mathcal{W}(i \xrightarrow{Z} j)$, there is a walk $V \in \mathcal{W}(i \xrightarrow{Z} j)$ with length at most $l(Z) + (n - l(Z) + \gamma - 2)\, \mathrm{cr}_W + \max(\min(n_W - 1, \mathrm{cr}_W + \mathrm{cd}_W), \mathrm{cr}_W + \mathrm{cd}_W - l(Z))$ obtained by removing cycles from W and possibly inserting Z such that $l(V) \equiv l(W) \pmod{\gamma}$.*

Moreover, if no copy of Z is inserted then $l(V) \le \mathrm{cr}_W(n - l(Z) + \gamma - 1) + \mathrm{cd}_W$.

Using that $\mathrm{cr}_W \le \mathrm{cr}(\mathcal{D}(A))$ and $\mathrm{cd}_W \le \mathrm{cd}(\mathcal{D}(A))$, we get Proposition 9.3.

When $l(Z) = \gamma = n$, R_1 is empty and the cycles in R_2 have length at most $n - 1$ (otherwise they would be removed). So we use (42) with $\mathrm{cr}_W \le n - 1$, and we obtain $l(V) \le (n-1)(n-1) + n - 1 = n^2 - n$. Hence $T_{cr}^n(Z) \le n^2 - n$ and $\tilde{T}_{cr}(Z) \le n^2 - n + 1$. Proposition 9.4 is proved. $\square$

9.3. Cycle removal by arithmetic method

In this section, we present a new method to bound T_{cr} leading to Proposition 9.5. We begin with:

Lemma 9.7. *Let $\gamma \in \mathbb{N}$ and let $W \in \mathcal{W}(i \to j)$. Then there exists a walk $W' \in \mathcal{W}(i \to j)$ obtained from W by removing cycles such that $l(W') \equiv l(W) \pmod{\gamma}$ and each node appears at most γ times in W'.*

Proof. Consider W as a sequence of adjacent nodes $(i_0, \cdots, i_L)$, where L is the length of the walk.

If a given node appears twice, first as i_a and then as i_b and if $a \equiv b \pmod{\gamma}$, then the subwalk $(i_0, \cdots, i_a, i_{b+1}, \cdots, i_L)$ is strictly shorter than W and has the same length modulo γ.

Iterating this process, we get a sequence of subwalks of W. Since the sequence of length is strictly decreasing, the sequence is finite and we denote the last walk by W'.

Obviously, $l(W') \equiv l(W) \pmod{\gamma}$ and a node does appear twice as i_a and i_b only if $a \not\equiv b \pmod{\gamma}$, so the pigeonhole principle implies that it appears at most γ times (otherwise there would exist i_a and i_b with $a \equiv b \pmod{\gamma}$). $\quad\square$

Proof of Proposition 9.5. We take $W \in \mathcal{W}(i \xrightarrow{\mathcal{G}} j)$ and construct a subwalk V with length at most $\gamma n + n - n_1 - 1$ by the following steps.

1. Find the first occurrence of a node of $\mathcal{G}$ in W, and denote this node by k. Let W_1 be the subwalk of W connecting i to k, and let W_2 be the remaining subwalk. So we have

$$W_1 \in \mathcal{W}(i \to k), \qquad W_2 \in \mathcal{W}(k \to j), \qquad l(W_1) + l(W_2) = l(W) \qquad (44)$$

2. As long as there is a node ℓ that appears twice in W_1 and at least once in W_2, we can write $W_1 = U_1 \cdot U_2 \cdot U_3$ and $W_2 = V_1 \cdot V_2$, where U_1, U_2, V_1 end with ℓ and U_2, U_3, V_2 start with ℓ. Thus, we can replace W_1 by $U_1 \cdot U_3$ and W_2 by $V_1 \cdot U_2 \cdot V_2$. Eq. (44) still holds, but now i appears only once in W_1. Step 2 is over when all nodes that appear more than once in W_1 do not appear in W_2. Let us denote the resulting walks by W_3 and W_4 respectively.

3. Apply Lemma 9.7 to W_3 and W_4, obtaining W_1' and W_2' respectively.

4. Set $V = W_1' \cdot W_2'$.

Obviously, $l(V) \equiv l(W_1) + l(W_2) \equiv l(W) \pmod{\gamma}$.

Now we take a node of V and bound the number of its appearances.

(1) If it is a node of $\mathcal{G} \setminus \{k\}$, then it appears only in W_2', thus at most γ times. k appears once in W_1', as ending node, and at most γ times in W_2'. In the concatenation of the walks, one occurrence disappears, so all nodes of $\mathcal{G}$ appear at most γ times.

(2) If it is a node of W_2', then it is also a node of W_4, it appears at most once in W_3, thus also in W_1'. Therefore it appears at most $\gamma + 1$ times in V.

(3) If it is not a node of W_2', then it appears only in W_1', thus at most γ times.

The total number of appearances of all nodes in V is at most $(\gamma + 1)(n - n_1) + \gamma n_1 = \gamma n + (n - n_1)$, so $l(V)$ is bounded by $\gamma n + (n - n_1 - 1)$, as claimed. $\quad\square$

Proof of bound (20). To prove bound (20), we apply Lemma 8.1 as before and the difficulty only comes from Lemma 8.3 that is not good enough.

To prove that inequality (32) holds with $t \geq \mathrm{Wi}(n)$ and $\mathcal{D} = \mathcal{G}^{ct}$, we do as in the proof of Proposition 6.5: we remove cycles from a walk W to replace them by cycles with greater weight, following the staircase given by Lemma 8.4. In this process, the walks to reduce have no critical node.

We apply Lemma 9.6 with $\gamma = l(Z)$ and we obtain

$$l(V) \leq (n-2)\,\mathrm{cr}_W + l(Z) + \max\big(n_W - 1, \mathrm{cr}_W + \mathrm{cd}_W - l(Z)\big)$$
$$\leq n_W - 1 + n = (n-1)n_W + n - 1.$$

Since W has no critical node, $n_W \leq n - 1$, and this bound is less than $\mathrm{Wi}(n)$ except when $n_W = n - 1$.

But in this last case one has a critical loop on the only critical node and the rest of the nodes are in W. Let Z be the penultimate cycle of the staircase, it shares nodes with W and contains the unique critical node. The weight of this cycle is greater than or equal to that of all cycles in W. Applying Proposition 9.5 with $\mathcal{G} = Z$ and $\gamma = 1$, it is possible to reduce the walk to a length at most $2n - l(Z) - 1$, insert Z and then as many critical loops as necessary to get back to a walk with length t.

This is possible if $t \geq 2n - 1$. Thus, Eq. (20) holds true for any n. $\square$

10. Proof of Theorems 4.5 and 4.6

Theorem 4.5 follows from the bounds on T_{cr} together with the following proposition.

Proposition 10.1. *Let A be an irreducible matrix, $\mathcal{G}$ be a representing subgraph of $\mathcal{G}^c(A)$ with cyclicity γ and B be subordinate to A such that $\lambda(B) \neq \mathbf{0}$. Then*

$$T_2(A,B) \leq \frac{T_{cr}^\gamma(\mathcal{G})(\lambda(A) - \min_{kl} a_{kl}) + (\max_{kl} b_{kl} - \lambda(B))\,\mathrm{cd}(\mathcal{D}(B))}{\lambda(A) - \lambda(B)}.$$

If moreover A has only finite entries, then Eqs. (26) and (27) hold.

We begin with the following lemmas.

Lemma 10.2. *Let $A \in \mathbb{R}_{\max}^{n \times n}$ be an irreducible matrix, and C, S, R be defined relative to any completely reducible $\mathcal{G} \subseteq \mathcal{G}^c(A)$. For any B subordinate to A and any t, if $b_{ij}^{(t)}$ is finite, then $(CS^tR)_{ij}$ is finite too.*

Proof. If $b_{ij}^{(t)}$ is finite, so is $a_{ij}^{(t)}$. By the optimal walk interpretation (6), there is a walk W connecting i to j, of length t, such that $p(W) = a_{ij}^{(t)}$. As A is irreducible, there is a closed walk V containing i and a node k of $\mathcal{G}$. If γ is the cyclicity of $\mathcal{G}$ then $V^\gamma W \in \mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{G}} j)$, and $(CS^tR)_{ij} \neq \mathbf{0}$ by (30). $\square$

Lemma 10.3. *For any $B \in \mathbb{R}_{\max}^{n \times n}$ and any $t \in \mathbb{N}$, let $\tilde{B}$ be $B - \lambda(B)$, we have:*

$$B^t \leq t\lambda(B) \otimes \tilde{B}^* \text{ and } \tilde{b}_{ij}^* \leq \text{cd}\big(\mathcal{D}(B)\big) \quad \Big(\max_{kl} b_{kl} - \lambda(B)\Big) \leq (n_B - 1)\|B\|$$

If B has only finite entries, then $\tilde{b}_{ij}^* \leq (\lambda(B) - \min_{kl} b_{kl})$.

Proof. The first part of the claim immediately follows from the optimal walk interpretation (6) and (8).

For the second part, observe that $\tilde{b}_{ij}^*$ is equal to $p(W) - \lambda(B)l(W)$ for some walk W connecting i to j in B. As $b_{ji} \neq \mathbf{0}$ we have $p(W) \leq \lambda(B)(l(W) + 1) - b_{ji}$, hence $\tilde{b}_{ij}^* \leq \lambda(B) - b_{ji}$ and the second part of the claim. $\square$

Lemma 10.4. *Let $A \in \mathbb{R}_{\max}^{n \times n}$ be a matrix with $\lambda(A) = \mathbf{1}$, C, S, R be defined relatively to $\mathcal{G}^c(A)$, let $\mathcal{G}$ be a representing subgraph of $\mathcal{G}^c(A)$ and γ be a multiple of the cyclicity of $\mathcal{G}$. For any $t \in \mathbb{N}$, the finite entries of $CS^t R$ satisfy*

$$\big(CS^t R\big)_{ij} \geq T_{cr}^\gamma(\mathcal{G}) \min_{kl} a_{kl}. \tag{45}$$

If A has only finite entries, then for all i, j we have:

$$\big(CS^t R\big)_{ij} \geq 2\min_{ij} a_{ij} \tag{46}$$

$$\big(CS^t R\big)_{ij} \geq 2\min_{ij} a_{ij} + \tilde{b}_{ij}^* + \text{cd}_B \lambda(B) \tag{47}$$

$$\big(CS^t Rv\big)_i \geq \min_{ij} a_{ij} + \min_j v_j. \tag{48}$$

Before proving this lemma, let us state another one to use for the matrices with finite entries.

Lemma 10.5. *Let A be a matrix with $\lambda(A) = \mathbf{1}$, then for any integer m there is a walk W_0 with length m and nonnegative weight on $\mathcal{D}(A)$.*

Proof. Let Z be a critical cycle of A. Since $l(Z^m) = l(Z) \cdot m$, there exist walks $W_1, \cdots, W_{l(Z)}$ of length m such that $W_1 \cdots W_{l(Z)} = Z^m$. Since $\sum_l p(W_l) = p(Z)t = 0$, there is a W_k with nonnegative $p(W_k)$. $\square$

Proof of Lemma 10.4. We first show inequality (45). By the optimal walk interpretation (30) we have $(CS^t R)_{ij} = \max\{p(W) : W \in \mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{G}} j)\}$ for any walk W. If $(CS^t R)_{ij}$ is finite then the walk set $\mathcal{W}^{t,\gamma}(i \xrightarrow{\mathcal{G}} j)$ is nonempty and contains a walk with the length bounded by $T_{cr}^\gamma(\mathcal{G})$, hence (45).

To prove inequality (46), let us assume that A has only finite entries, and that $t \geq 2 + n$ (using that the sequence $\{CS^t R\}_{t \geq 1}$ is periodic).

Apply Lemma 10.5 with $m = t - 2$ and set $W = (i, r) \cdot W_0 \cdot (s, j)$, where r, resp. s, is the beginning node, resp. the end node of W_0. By the optimal walk interpretation (30), we get $(CS^t R)_{ij} \geq p(W) \geq a_{ir} + a_{sj} \geq 2 \min_{kl} a_{kl}$.

The inequalities (47) and (48) are proved similarly. For (47), select a walk V with minimal length among those with weight $\tilde{b}_{ij}^*$ on $\mathcal{D}(\tilde{B})$ and a walk W_0 with nonnegative $p(W_0)$ and length $t - l(V) - 2$. Set $W = (i, r) \cdot W_0 \cdot (s, i) \cdot V$ and get

$$\left(CS^t R\right)_{ij} \geq p(W) \geq a_{ir} + a_{si} + p(V) \geq 2 \min_{k\ell} a_{k\ell} + \tilde{b}_{k\ell}^* + \lambda(B) \operatorname{cd}_B .$$

For (48), select a walk W_0 with nonnegative $p(W_0)$ and length $t - 1$ and set $W = (i, r) \cdot W_0$ (where r is the beginning node of W_0). $\square$

Proof of Proposition 10.1 and Theorem 4.5. Assume that $\lambda(A) = 0$ and t is greater than one of the bounds. We want to prove that equation

$$t\lambda(A) \otimes \left(CS^t R\right)_{ij} \geq t\lambda(B) \otimes \tilde{b}_{ij}^{(t)} \tag{49}$$

holds for all i, j.

By Lemma 10.2, if $(CS^t R)_{ij} = \mathbf{0}$ then $\tilde{b}_{ij}^t = \mathbf{0}$ and there is nothing to prove. So we can assume that $(CS^t R)_{ij}$ is finite, in which case we can use the inequalities of Lemmas 10.4 and 10.3, which show that (49) follows when we have

$$t\lambda(A) + T_{cr}^\gamma(\mathcal{G})\left(\min_{kl} a_{kl} - \lambda(A)\right) \geq t\lambda(B) + \operatorname{cd}\left(\mathcal{D}(B)\right)\left(\max_{kl} b_{kl} - \lambda(B)\right),$$

$$t\lambda(A) + 2\left(\min_{kl} a_{kl} - \lambda(A)\right) \geq t\lambda(B) + \left(\lambda(B) - \min_{kl} b_{kl}\right), \tag{50}$$

in the general case (the first inequality) and in the case of finite entries (the second inequality). If t is greater than one of the required bounds, then one of the inequalities (50) holds, and (49) follows.

To obtain Theorem 4.5 it remains to deduce the shorter parts of (23)–(25) from the longer ones. Observe that all the longer parts of the bounds are of the form

$$\frac{n_1(\lambda(A) - a_{ij}) + \operatorname{cd}_B(a_{kl} - \lambda(B))}{\lambda(A) - \lambda(B)} \tag{51}$$

for some i, j, k, l, where n_1 is greater than cd_B. Using $n_1 > \operatorname{cd}_B$, expression (51) can be bounded by

$$\frac{(n_1 - \operatorname{cd}_B)(\lambda(A) - a_{ij}) + \operatorname{cd}_B(a_{kl} - a_{ij} + \lambda(A) - \lambda(B))}{\lambda(A) - \lambda(B)}$$

$$\leq \frac{(n_1 - \operatorname{cd}_B)\|A\| + \operatorname{cd}_B(\|A\| + \lambda(A) - \lambda(B))}{\lambda(A) - \lambda(B)} = n_1 \frac{\|A\|}{\lambda(A) - \lambda(B)} + \operatorname{cd}\left(\mathcal{D}(B)\right).$$

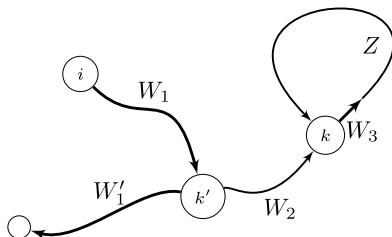This completes the proof of all the bounds of Theorem 4.5. $\square$

Fig. 3. Walks $\tilde{V} = W_1 W_1'$ and $W = W_1 W_2 Z^r W_3$ in proof of Theorem 4.6.

It remains to prove Theorem 4.6. We do it by generalizing the proof of [8, Proposition 5].

Proof of Theorem 4.6. The case $\lambda(A) = \mathbf{0}$ is trivial. In the rest of the prove, we assume $\lambda(A) = \mathbf{1}$ by replacing A with $\lambda(A)^- \otimes A$.

We denote by Δ and δ the greatest and smallest edge weight in $\mathcal{D}(A)$, respectively. We have $\|A\| = \Delta - \delta$. If $\Delta = \delta$, then $\mathcal{G}^c(A) = \mathcal{D}(A)$ and hence $B^t \leq A^t \leq CS^t R$ by the optimal walks interpretations (6) and (30).

We hence assume $\Delta \neq \delta$ in the rest of the proof. The assumption $\lambda(A) = 0$ implies $\delta \leq \lambda(B) \leq 0 \leq \Delta$.

Denote by $v_{\max}$ and $v_{\min}$ the greatest and smallest entry of v, respectively. It is $\|v\| = v_{\max} - v_{\min}$.

Let $t \geq (\|v\| + (n-1)\|A\|)/(-\lambda(B))$. We show $CS^t Rv \geq B^t v$.

Let i be a node of $\mathcal{D}(A)$. Let V be a walk in $\mathcal{D}(B)$ of length t starting at i, and let $\tilde{V}$ be the remaining walk after repeated cycle deletion. Let W_2 be a shortest path connecting some node k' of $\tilde{V}$ to a critical node k and let W_1 be the prefix of $\tilde{V}$ ending at k' and let $\tilde{V} = W_1 \cdot W_1'$. See Fig. 3 for an illustration of these walks. We obtain

$$p(V) \leq p(\tilde{V}) + \lambda(B) \cdot \left(t - l(\tilde{V})\right) \leq p\left(W_1 \cdot W_1'\right) - \delta \cdot l\left(W_1 \cdot W_1'\right) + \lambda(B) \cdot t$$
$$\leq p(W_1) + \|A\| \cdot l\left(W_1'\right) - \delta \cdot l(W_1) - \|v\| - \|A\| \cdot (n-1), \tag{52}$$

using that $\lambda(B)t \leq -(\|v\| + \|A\|(n-1))$.

Let Z be a critical cycle starting at k and set $r = \lfloor (t - l(W_1 \cdot W_2))/l(Z) \rfloor$. Then let W_3 be the prefix of Z of length $t - l(W_1 \cdot W_2 \cdot Z^r)$, which is between 0 and $l(Z) - 1$. Setting $W = W_1 \cdot W_2 \cdot Z^r \cdot W_3$, we have

$$p(W) \geq p(W_1 \cdot W_2 \cdot W_3) \geq p(W_1) + \delta \cdot l(W_2 \cdot W_3) \tag{53}$$

and hence, because $l(W_1) + l(W_2) + l(W_3) + l(W_1') \leq n - 1$,

$$p(V) \leq p(W_1) - \|A\| \cdot \left(l(W_1) + l(W_2) + l(W_3)\right) - \delta \cdot l(W_1) - \|v\| \tag{54}$$
$$\leq p(W_1) + \delta \cdot l(W_2 \cdot W_3) - \|v\| \leq p(W) - \|v\|. \tag{55}$$

Since $p(W) + v_{\min} \le (CS^tRv)_i$ by the walk interpretation (30) of CSR terms, we have $(CS^tRv)_i \ge (B^tv)_i$, which concludes the proof.

The claim for the case that all entries of A are finite follows from Lemmas 10.3 and 10.4. $\square$

11. Cycle insertion

In this section, we state some bounds on ep^γ.

The exploration penalty has been introduced in [8], where the following is proven.

Proposition 11.1. *(See Theorem 3 of [8].) Let $\mathcal{G}$ be a strongly connected graph with cyclicity γ and girth* g. *Its exploration penalty* ep^γ *satisfies:*

$$\mathrm{ep}^\gamma \le 2\frac{\mathrm{g}}{\gamma}|\mathcal{G}| - \frac{\mathrm{g}}{\gamma} - 2\,\mathrm{g} + \gamma.$$

Since $\mathrm{ep}^\gamma(\mathcal{G})$ is bounded by $\mathrm{ind}(\mathcal{G})$, it is also possible to use the following bounds.

Proposition 11.2. *Let $\mathcal{G}$ be a strongly connected graph. Its index $\mathrm{ind}(\mathcal{G})$ is related to its girth* g *and its cyclicity γ by the following inequalities:*

$$\mathrm{ind}(\mathcal{G}) \le \mathrm{Wi}\big(|\mathcal{G}|\big),$$
$$\text{where } \mathrm{Wi}(1) = 0 \text{ and } \mathrm{Wi}(r) = (r-1)^2 + 1 \text{ otherwise.} \tag{56}$$
$$\mathrm{ind}(\mathcal{G}) \le \gamma\,\mathrm{Wi}(r) + s,$$
$$\text{where } r \text{ is the quotient of the division of } |\mathcal{G}| \text{ by } \gamma \text{ and } s \text{ its remainder.} \tag{57}$$
$$\mathrm{ind}(\mathcal{G}) \le |\mathcal{G}| + \big(|\mathcal{G}| - 2\big)\mathrm{g}. \tag{58}$$

Bound (57) can be traced back to a work of Wielandt [28]. Bound (57) is due to Schwarz [22], but a more comprehensive explanation was given by Shao and Li [16]. Bound (58) was originally proved by Dulmage and Mendelsohn [11] for primitive matrices but the case of a non-primitive matrix also follows (for instance) from Theorem 4.1 by Remark 4.2. Other bounds on $\mathrm{ind}(\mathcal{D})$ can be also found in the literature.

As noticed by Kim [17], the same method as in the proof of (57) by Shao and Li [16] applied to (58) instead of (56) gives:

$$\mathrm{ind}(\mathcal{G}) \le \gamma r + (r-2)\,\mathrm{g} + s \le |\mathcal{G}|\left(1 + \frac{\mathrm{g}}{\gamma}\right) - 2\mathrm{g}. \tag{59}$$

Let us derive from (58) the following:

Proposition 11.3. *Let $\mathcal{G}^c(A)$ have n_c nodes, maximal girth $\hat{\mathrm{g}}$, h s.c.c.'s, of which h_2 have at least 2 nodes. Then any s.c.c. $\mathcal{G}^c$ of $\mathcal{G}^c(A)$ satisfies*

$$\mathrm{ep}^{\gamma(\mathcal{G}^c)}\big(\mathcal{G}^c\big) \le n_c - h - h_2 + 2 + (n_c - h - h_2)\hat{\mathrm{g}}.$$

Together with this bound, Eq. (18) generalizes and improves the bound of [4]. Actually, in [4] $(h + h_2)$ is replaced by $2h$, but this is due to a mistake: they use that $|\mathcal{G}_l^c| - 2$ is non-negative for each s.c.c. $\mathcal{G}_l^c$, which fails to be true if $|\mathcal{G}_l^c| = 1$.

Proof. Observe that any s.c.c. with at least two nodes has at most $n_c - h - h_2 + 2$ nodes (because there are $h - h_2$ nodes in s.c.c.'s of size 1 and $2(h_2 - 1)$ nodes in other s.c.c.'s of size at least 2). The Dulmage–Mendelsohn bound then implies, for every index i in a component $\mathcal{G}^c$ of $\mathcal{G}^c(A)$ with at least two nodes, that

$$\mathrm{ep}^{\gamma(\mathcal{G}^c)}(i) \leq \mathrm{ind}(\mathcal{G}^c) \leq n_c - h - h_2 + 2 + \hat{\mathrm{g}}(n_c - h - h_2).$$

Observe that $n_c \geq h + h_2$. Also, since the right-hand side of that inequality is at least 2, the bound also trivially holds for i in a single node component of $\mathcal{G}^c(A)$. $\quad\square$

12. Local reductions

Every weak CSR expansion gives rise to *local weak CSR expansions* that can take the following forms:

$$
\begin{aligned}
a_{ij}^{(t)} &= \left(CS^t R\right)_{ij} \oplus b_{ij}^{(t)}, \quad \text{for } t \geq \tilde{\tau}(i,j), \\
a_{ij}^{(t)} &= \left(CS^l R\right)_{ij} \oplus b_{ij}^{(t)}, \quad \text{for } t \equiv l \pmod{\gamma} \text{ and } t \geq \tilde{\tau}(i,j,l), \\
\left(A^t v\right)_i &= \left(CS^l R v\right)_i \oplus \left(B^t v\right)_i, \quad \text{for } t \geq \tilde{\tau}(i,v), \\
\left(A^t v\right)_i &= \left(CS^l R v\right)_i \oplus \left(B^t v\right)_i, \quad \text{for } t \equiv l \pmod{\gamma} \text{ and } t \geq \tilde{\tau}(i,l,v).
\end{aligned}
\tag{60}
$$

In connection with these schemes, define the following subsets:

$$
\begin{aligned}
J(i,j) &:= \left\{ s \colon a_{is}^* a_{sj}^* < \min_l \left(CS^l R\right)_{ij} \right\}, \\
J(i,j,l) &:= \left\{ s \colon a_{is}^* a_{sj}^* < \left(CS^l R\right)_{ij} \right\}, \\
J(i,v) &:= \left\{ s \colon \bigoplus_j a_{is}^* a_{sj}^* v_j < \min_l \left(CS^l R v\right)_i \right\}, \\
J(i,l,v) &:= \left\{ s \colon \bigoplus_j a_{is}^* a_{sj}^* v_j < \left(CS^l R v\right)_i \right\}.
\end{aligned}
\tag{61}
$$

Remark 12.1. Unless $i = s = j$, $a_{is}^* a_{sj}^*$ is the biggest weight of a walk connecting i to j via s. It follows from Theorem 6.1 and this optimal walk interpretation that $i, j \notin J(i,j), J(i,j,l)$ and $i \notin J(i,v), J(i,l,v)$. Moreover, if some critical s belongs to one of the sets defined here, then so do all nodes in its s.c.c. of $\mathcal{G}^c(A)$, since for each pair of nodes in the same s.c.c. of $\mathcal{G}^c(A)$ we can find a closed walk in $\mathcal{G}^c(A)$ containing both of them.

Now let $\tilde{\mathcal{G}}^c(A)$ be the remainder of the critical graph, without the s.c.c. with indices in J, for $J = J(i,j)$, $J(i,j,l)$, $J(i,v)$ or $J(i,l,v)$.

Redefine $\tilde{C}, \tilde{S}$ and $\tilde{R}$ using $\tilde{\mathcal{G}}^c(A)$ instead of $\mathcal{G}^c(A)$, and the subordinate matrix' $\tilde{A}$ of A where all rows and columns with indices in J are canceled, instead of A. Redefine $\tilde{B}$ as a subordinate of $\tilde{A}$ whose indices are in $\mathcal{D}(B)$ (but not in J). This procedure will be referred to as *local reduction* of a weak CSR expansion. When $J = J(i,j)$, or resp. $J = J(i,j,l)$, $J = J(i,v)$ or $J = J(i,v,l)$, this will be called (i,j)-reduction, or resp. (i,j,l)-reduction, (i,v)-reduction or (i,l,v)-reduction.

Theorem 12.2. *Let $A \in \mathbb{R}_{\max}^{n \times n}$, B subordinate to A and the integer numbers $\tilde{\tau}(i,j)$, $\tilde{\tau}(i,j,l)$, $\tilde{\tau}(i,v)$ and $\tilde{\tau}(i,l,v)$, for $i,j \in \{1, \ldots, n\}$ and $v \in \mathbb{R}_{\max}^n$, satisfy* (60)*. Corresponding to the definitions of J given in* (61)*, we have*

$$a_{ij}^{(t)} = (\tilde{C}\tilde{S}^t\tilde{R})_{ij} \oplus \tilde{b}_{ij}^{(t)}, \quad \textit{for } t \geq \tilde{\tau}(i,j),$$

$$a_{ij}^{(t)} = (\tilde{C}\tilde{S}^l\tilde{R})_{ij} \oplus \tilde{b}_{ij}^{(t)}, \quad \textit{for } t \equiv l \ (\mathrm{mod}\ \gamma) \textit{ and } t \geq \tilde{\tau}(i,j,l),$$

$$(A^t v)_i = (\tilde{C}\tilde{S}^l\tilde{R}v)_i \oplus (\tilde{B}^t v)_i, \quad \textit{for } t \geq \tilde{\tau}(i,v),$$

$$(A^t v)_i = (\tilde{C}\tilde{S}^l\tilde{R}v)_i \oplus (\tilde{B}^t v)_i, \quad \textit{for } t \equiv l \ (\mathrm{mod}\ \gamma) \textit{ and } t \geq \tilde{\tau}(i,l,v), \tag{62}$$

with $\tilde{C}, \tilde{S}, \tilde{R}$ and $\tilde{B}$ defined in the local reduction procedure.

Proof. We prove the theorem only in the case of (i,j)-reduction, i.e., in the first case of (62) corresponding to the first case of (60) and (61). All other cases are similar. Let N_B, resp. N_c be the set of nodes of $\mathcal{D}(B)$, resp. $\mathcal{G}^c(A)$.

Define the subordinate matrix A' of A formed by setting to $\mathbf{0}$ all rows and columns with indices in $J(i,j) \cap N_B$. We first show that the first equation of (60) for $a_{ij}^{(t)}$ holds also with CSR terms and B defined from A' instead of A. First, recall that the weights of walks going through $s \in J(i,j) \cap N_B$ are less than $\min_l (CS^l R)_{ij}$, and $(CS^l R)_{ij}$ is the greatest weight of any walk with certain length constraint, connecting i to j via a critical node. Defining $(CS^l R)_{ij}$ from A' amounts to canceling all walks going through $s \in J(i,j) \cap N_B$ and contributing to $(CS^l R)_{ij}$. Since such walks have low weight, $(CS^l R)_{ij}$ does not decrease, for any l, when defined from the subordinate matrix A', so it is exactly the same. Next, observe that (since the weights of walks going through $s \in J(i,j) \cap N_B$ are less than $\min_l (CS^l R)_{ij}$) we can replace $b_{ij}^{(t)}$ by $\tilde{b}_{ij}^{(t)}$ in the first equation of (60).

We next show that the CSR term defined from A' can be reduced. Use expansion (31) of the CSR terms defined from A', where the first terms in (31) are defined from the components of $\mathcal{G}^c(A)$ with indices in $J(i,j)$ (these components can be taken in any order). The sum of these terms expresses $p(\mathcal{W}^t(i \xrightarrow{J(i,j) \cap N_c} j))$ (with walk sets defined in $\mathcal{D}(A')$) for all large enough t. Since these walk weights are strictly less than the entries of CSR, all those terms in expansion (31) with indices in $J(i,j)$ can be canceled. The remaining part of expansion (for the entry i,j) sums up to the reduced CSR term defined from the subordinate matrix $\tilde{A}$ (as defined in the reduction procedure). $\square$

Notice that the proofs of Theorems 4.5 and 4.6 in Section 10 work with walks in $\mathcal{W}^t(i \to j)$. Hence the corresponding bounds can be combined with all reductions of Theorem 12.2. In particular, local reductions may lead to smaller B and $\lambda(B)$ when i and j or i and v are fixed. Moreover, they can also result in decrease of the initial bounds on $\tilde{\tau}$ based on the cycle removal threshold, since some of the critical components get removed.

We now also recall a bound of a type that can be found in Akian, Gaubert and Walsh [2], and Bouillard and Gaujal [4], formulated here for the case of (i, j, l)-reduction.

Proposition 12.3. *Suppose that $A \in \mathbb{R}_{\max}^{n \times n}$ is irreducible, with $\lambda(A) = 1$, and take $i, j \in \{1, \ldots, n\}$ and $l > 0$. Let γ be the cyclicity of $\mathcal{G}^c(A)$, and let $\tilde{\tau}(i, j, l)$ be an integer such that*

$$a_{ij}^{(t)} = \left(\tilde{C}\tilde{S}^l\tilde{R}\right)_{ij} \oplus \tilde{b}_{ij}^{(t)}, \quad t \equiv l \pmod{\gamma}, \ t \geq \tilde{\tau}(i, j, l), \tag{63}$$

where the terms $\tilde{C}, \tilde{S}, \tilde{R}$ and matrix $\tilde{B}$ are obtained by the (i, j, l)-reduction of some weak CSR expansion. Let

$$T(i, j, l) = \min\left\{t : \lambda^{\otimes t}(\tilde{B}) \otimes \left(\lambda^{-}(\tilde{B}) \otimes \tilde{B}\right)_{ij}^* \leq \left(\tilde{C}\tilde{S}^l\tilde{R}\right)_{ij}\right\}. \tag{64}$$

Then the transient $\tau(i, j, l)$ for which

$$a_{ij}^{(t)} = \left(\tilde{C}\tilde{S}^l\tilde{R}\right)_{ij}, \quad t \equiv l \pmod{\gamma}, \ t \geq \tau(i, j, l), \tag{65}$$

satisfies $\tau(i, j, l) \leq \max(\tilde{\tau}(i, j, l), T(i, j, l))$.

Proof. We only need to show that $\tilde{b}_{ij}^{(t)} \leq \lambda^{\otimes t}(\tilde{B}) \otimes (\lambda^{-}(\tilde{B}) \otimes \tilde{B})_{ij}^*$. Indeed, this follows after dividing both parts of this inequality (in max-plus sense) by $\lambda^{\otimes t}(\tilde{B})$ and using the optimal walk interpretation of $(\lambda^{-}(\tilde{B}) \otimes \tilde{B})^t$ and $(\lambda^{-}(\tilde{B}) \otimes \tilde{B})^*$. $\quad\square$

References

[1] Martin Aigner, Günther M. Ziegler, Proofs from The Book, second edition, Springer, Heidelberg, 2001.
[2] Marianne Akian, Stéphane Gaubert, Cormac Walsh, Discrete max-plus spectral theory, in: G.L. Litvinov, V.P. Maslov (Eds.), Idempotent Mathematics and Mathematical Physics, in: Contemporary Mathematics, vol. 377, AMS, 2005, pp. 53–77.
[3] François-Louis Baccelli, Guy Cohen, Geert Jan Olsder, Jean-Pierre Quadrat, Synchronization and Linearity: An Algebra for Discrete Event Systems, John Wiley and Sons, 1992.

[4] Anne Bouillard, Bruno Gaujal, Coupling time of a (max, plus) matrix, in: Proceedings of the Workshop on Max-Plus Algebra at the 1st IFAC Symposium on System Structure and Control, Elsevier, Amsterdam, 2001, pp. 335–400.

[5] Peter Butkovič, Max-linear Systems: Theory and Algorithms, Springer, 2010.

[6] Peter Butkovič, Hans Schneider, Applications of max-algebra to diagonal scaling of matrices, Electron. J. Linear Algebra 13 (2005) 262–273.

[7] Bernadette Charron-Bost, Matthias Függer, Thomas Nowak, On the transience of linear max-plus dynamical systems, E-print arXiv:1111.4600, 2011.

[8] Bernadette Charron-Bost, Matthias Függer, Thomas Nowak, New transience bounds for long walks, E-print arXiv:1209.3342, 2012.

[9] Bernadette Charron-Bost, Matthias Függer, Jennifer L. Welch, Josef Widder, Full reversal routing as a linear dynamical system, in: Adrian Kosowski, Masafumi Yamashita (Eds.), SIROCCO 2011, in: LNCS, vol. 6796, Springer, Heidelberg, 2011, pp. 101–112.

[10] Guy Cohen, Didier Dubois, Jean-Pierre Quadrat, Michel Viot, Analyse du comportement périodique de systèmes de production par la théorie des dioïdes, INRIA, Rapport de Recherche no. 191, Février 1983.

[11] Anthony L. Dulmage, Nathan S. Mendelsohn, Gaps in the exponent set of primitive matrices, Illinois J. Math. 8 (4) (1964) 642–656.

[12] Miroslav Fiedler, Vlastimil Pták, Diagonally dominant matrices, Czechoslovak Math. J. 17 (3) (1967) 420–433.

[13] Michael R. Garey, David S. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness, Freeman, San Francisco, 1979.

[14] Mark Hartmann, Cristina Arguelles, Transience bounds for long walks, Math. Oper. Res. 24 (2) (1999) 414–439.

[15] Bernd Heidergott, Geert-Jan Olsder, Jacob Van Der Woude, Max-Plus at Work, Princeton Univ. Press, 2005.

[16] Jia-yu Shao, Qiao Li, On the index of convergence of an irreducible Boolean matrix, Linear Algebra Appl. 97 (1987) 185–210.

[17] Ki-Hang Kim, An extension of the Dulmage–Mendelsohn theorem, Linear Algebra Appl. 27 (1979) 187–197.

[18] Monika Molnárová, Computational complexity of Nachtigall's representation, Optimization 52 (2003) 93–104.

[19] Karl Nachtigall, Powers of matrices over an extremal algebra with applications to periodic graphs, Math. Methods Oper. Res. 46 (1997) 87–102.

[20] Hans Schneider, Private communication, 2009.

[21] Hans Schneider, Michael H. Schneider, Max-balancing weighted directed graphs and matrix scaling, Math. Oper. Res. 16 (1991) 208–222.

[22] Štefan Schwarz, On a sharp estimation in the theory of binary relations on a finite set, Czechoslovak Math. J. 20 (1970) 703–714.

[23] Sergeǐ Sergeev, Max algebraic powers of irreducible matrices in the periodic regime: an application of cyclic classes, Linear Algebra Appl. 431 (6) (2009) 1325–1339.

[24] Sergeǐ Sergeev, Hans Schneider, CSR expansions of matrix powers in max algebra, Trans. Amer. Math. Soc. 364 (2012) 5969–5994, E-print arXiv:0912.2534.

[25] Sergeǐ Sergeev, Hans Schneider, Peter Butkovič, On visualization scaling, subeigenvectors and Kleene stars in max algebra, Linear Algebra Appl. 431 (12) (2009) 2395–2406, E-print arXiv:0808.1992.

[26] Jia-yu Shao, On the exponent of a primitive digraph, Linear Algebra Appl. 64 (1985) 21–31.

[27] Gerardo Soto y Koelemeijer, On the behavior of classes of min-max-plus systems, Ph.D. thesis, Delft Univ. of Technology, The Netherlands, 2003.

[28] Helmut Wielandt, Unzerlegbare, nicht negative Matrizen, Math. Z. 52 (1) (1950) 642–645.