Tight Bounds for Consensus Systems Convergence*

Pierre-Yves Chevalier, Julien M. Hendrickx, Raphaël M. Jungers

January 20, 2016

Abstract

We analyze the asymptotic convergence of all infinite products of matrices taken in a given finite set, by looking only at finite or periodic products.

It is known that when the matrices of the set have a common nonincreasing polyhedral norm, all infinite products converge to zero if and only if all infinite periodic products with period smaller than a certain value converge to zero, and bounds exist on that value [15].

We provide a stronger bound holding for both polyhedral norms and polyhedral seminorms. In the latter case, the matrix products do not necessarily converge to 0, but all trajectories of the associated system converge to a common invariant space. We prove our bound to be tight, in the sense that for any polyhedral seminorm, there is a set of matrices such that not all infinite products converge, but every periodic product with period smaller than our bound does converge.

Our work is motivated by problems in consensus systems, where the matrices are stochastic (non-negative with rows summing to one), and hence always share a same common nonincreasing polyhedral seminorm. In this particular case, we show that for the dimension of the space $n \geq 8$, our new bound is smaller than the previously known bound by a multiplicative factor of $\frac{3}{2\sqrt{\pi n}}$.

Our technique is based on an analysis of the combinatorial structure of the face lattice of the unit ball of the nonincreasing seminorm. The bound we obtain is equal to half the size of the largest antichain in this lattice. Explicitly evaluating this quantity may be challenging in some cases. We therefore link our problem with the Sperner property: the property that, for some graded posets, – in this case the face lattice of the unit ball of the norm – the size of the largest antichain is equal to the size of the largest rank level.

We show that this property holds for the polytope obtained when treating sets of stochastic matrices, and that our bound can then be easily evaluated in that case which motivated our study. However, we show that some other sets of matrices with invariant polyhedral seminorms lead to posets that do not have that Sperner property.

*All authors are members of ICTEAM, Université catholique de Louvain, Belgium. R. M. Jungers is an F.R.S.-FNRS research associate. Their work is supported by the Belgian Network DYSCO, funded by the Belgian government and the Concerted Research Action (ARC) of the French Community of Belgium. \{pierre-yves.chevalier, julien.hendrickx, raphael.jungers\}@uclouvain.be. The authors would also like to thank professors Jürgen Eckhoff, Samuel Fiorini and Jean-Paul Doignon for their helpful ideas and comments.
Introduction

We consider the problem of determining the stability of matrix sets, that is, determining whether or not all infinite products of matrices from a given set converge to zero, or more generally to a common invariant subspace. This problem appears in several different situations in control engineering, computer science, and applied mathematics. For instance, the stability of matrix sets characterizes the stability of switching dynamical systems [14], which have numerous application in control [12][14][19]. Stability of matrix sets is instrumental in proving the continuity of certain wavelet functions [8][13]. Somewhat surprisingly, it also helped establishing the best known asymptotic bounds on the number of \( \alpha \)-power-free binary words of length \( n \), a central problem in combinatorics on words [3][16].

Deciding the stability of a matrix set is notoriously difficult and the decidability of this problem is not known. The related problem of the existence of an infinite product whose norm diverges is undecidable [5]. However, it is possible to decide stability when the set has the finiteness property, that is, when there is a bound \( k \) such that the existence of an infinite non-converging product implies the existence of an infinite non-converging periodic product with period smaller than or equal to \( k \). Indeed, when such a bound \( k \) exists, checking the stability of the set can be done by checking the stability of all products whose length is smaller than or equal to \( k \). In this work, we look for the smallest valid bound \( k \).

A similar question is particularly relevant in the context of consensus problems. These systems are models for groups of agents trying to agree on some common value by an iterative process. Each agent has a value \( x_i \) which it updates by computing the weighted average of values of agents with which it can communicate. Consensus systems have attracted considerable attention due to their numerous applications in control of vehicle formations [2], flocking [6][13] or distributed sensing [17][22]. These systems typically have varying communication networks due to e.g. communication failures, or to the movements of the agents. This leads to systems whose (linear) dynamics may switch at each time-step. When a set of possible linear dynamics is known, one fundamental question is whether convergence can be guaranteed for any switching sequence [4].

Consensus systems can be modeled by discrete-time linear switching systems of the form \( x(t+1) = A_t x(t) \), where the transition matrices \( A_t \) are stochastic (nonnegative matrices whose rows sum to 1) because the agents always compute weighted averages. In that case, the products certainly do not converge to zero, since products of stochastic matrices remain stochastic. However, one can wonder whether the agents asymptotically converge to the same value. Deciding whether a consensus system converges for any sequence of transition matrices and any initial condition corresponds thus to determining whether all products of stochastic matrices converge to a rank one matrix of the form \( 1 y^\top \), where \( y \in \mathbb{R}^n \), \( y^\top 1 = 1 \). This particularization to stochastic matrices has many other applications, including nonhomogeneous Markov chains, and probabilistic automata [18].

It is easy to verify that the seminorm \( \|x\|_P = \frac{1}{2} (\max_i x_i - \min_i x_i) \) is nonincreasing under the action of any stochastic matrix, that is,

\[
\forall x \in \mathbb{R}^n, \ A \text{ stochastic, } \|Ax\|_P \leq \|x\|_P.
\]

The function \( \| \cdot \|_P \) is a polyhedral seminorm in the sense that its unit ball (e.g. the set \( \{ x : \|x\|_P \leq 1 \} \) is defined by a finite number of linear inequalities. It turns out that
the existence of an nonincreasing polyhedral seminorm provides strong information on the asymptotic convergence of products of these matrices.

For sets of matrices with a common invariant polyhedral norm, there exists a bound $k$ such that the existence of an infinite non-converging product implies the existence of an infinite non-converging periodic product with period at most $k$. This was first established by Lagarias and Wang [15]. The authors also give an explicit value for the bound (namely half the number of faces of the unit ball of the norm).

The particular case of stochastic matrices has been analyzed earlier in the context of non-homogeneous Markov chains [18], and later in the context of consensus systems [4]. A finiteness result has been known since Paz [18], who proved that all infinite products converge to a rank one matrix if and only if a certain condition on all products of length $B = \frac{1}{2}(3^n - 2^{n+1} + 1)$ is satisfied. This result can be interpreted as a bound $k$ such as described above for the particular case of stochastic matrices. In our recent paper [7], we showed that this bound can be derived from the result of Lagarias and Wang applied to the seminorm $\|\cdot\|_P$ described above.

Our Contribution

In this work, we consider a general problem that includes these particular cases: we study matrix sets for which there exists a polyhedral seminorm which is nonincreasing for all matrices in the given set, and we wonder whether long products of these matrices are asymptotically contractive. We improve all the bounds previously known in the particular cases, and prove that our bound is tight. More precisely, we answer the following question:

**Question 1.** Let $\|\cdot\|$ be a polyhedral seminorm; what is the smallest $k$ such that for any set $\Sigma$ for which $\|\cdot\|$ is nonincreasing, the existence of an infinite non-contracting product implies the existence of an infinite periodic non-contracting product with period smaller than or equal to $k$?

Our analysis relies on the fact that the convergence of the dynamical system is encapsulated in a discrete representation by a dynamical system on the face lattice of the polyhedral (semi)norm. Then, our bound relies on a careful study of the combinatorial structure of the trajectories in this discrete structure.

The improvement over the previously known bound depends on the seminorm. In the case of stochastic matrices, the improvement is a multiplicative factor of $\frac{3}{2\sqrt{n}}$ when $n$, the dimension of the space, is larger than or equal to 8.

## 1 Problem Setting

Let $\Sigma = \{A_1, \ldots, A_m\}$ be a matrix set. We investigate the convergence of the switching system

$$x(t + 1) = A_{\sigma(t)}x(t)$$

where $\sigma : \mathbb{N} \mapsto \{1, \ldots, m\} : t \mapsto \sigma(t)$ is an infinite sequence of indices. More precisely, our goal is to characterize conditions under which, for all initial conditions $x(0)$ and sequence $\sigma$ of matrices, the trajectory generated by (1) converges to some given subspace of $\mathbb{R}^n$ invariant under all matrices in $\Sigma$. 
An important particular case is the convergence of all trajectories to 0, which corresponds to the stability of the switched systems, and was studied in [14, 20]. Another one concerns consensus systems: In these systems, \( n \) agents each update their value \( x_i \) by computing a weighted average of values \( x_j \) of agents with which they can communicate, and these possibilities of communication may change with time. The update of the vector \( x \) of values can be written under the form (1) with stochastic matrices \( A_{\sigma(t)} \) since updates are weighted averages. The question is then to determine if \( x(t) \) always converges to a vector in which all agents have the same value, that is a vector of the form \( a1 \).

The stability of System (1) is equivalent [14] to the convergence to zero of all infinite products

\[ \ldots A_{\sigma(t)} \ldots A_{\sigma(2)} A_{\sigma(1)}, \]

while for consensus, the convergence to consensus for any sequence of matrix and any initial condition is equivalent to the convergence of all infinite products to rank one matrices of the form \( 1y^T \), where \( y \in \mathbb{R}^n \), \( y^T 1 = 1 \).

We say that a product \( \ldots A_{\sigma(2)} A_{\sigma(1)} \) is periodic if the sequence \( \sigma \) is periodic. We recall that a seminorm on \( \mathbb{R}^n \) is an application \( \| \| \) that has the following properties:

- \( \forall x \in \mathbb{R}^n, a \in \mathbb{R}, \|ax\| = |a|\|x\| \)
- \( \forall x, y \in \mathbb{R}^n, \|x + y\| \leq \|x\| + \|y\| \).

We call a polyhedral seminorm a seminorm whose unit ball is a polyhedron, that is, a set that can be defined by a finite set of linear inequalities

\[ \{ x : \|x\| \leq 1 \} = \{ x : \forall i, b_i^T x \leq c_i \}. \]

We say that a seminorm \( \| \| \) is nonincreasing with respect to a matrix \( A \) if

\[ \forall x \in \mathbb{R}^n, \|Ax\| \leq \|x\|. \]

Geometrically, this correspond to its unit ball being invariant

\[ A\{ x : \|x\| \leq 1 \} \subset \{ x : \|x\| \leq 1 \}. \]

We say that a seminorm \( \| \| \) is nonincreasing with respect to a set \( \Sigma \) of matrices if it is nonincreasing with respect to each of the matrices in \( \Sigma \). We say that a matrix \( A \) contracts a seminorm \( \| \| \) if \( \forall x \in \mathbb{R}^n, \|Ax\| < \|x\| \). We say that an infinite product \( \ldots A_{\sigma(2)} A_{\sigma(1)} \) contracts a seminorm \( \| \| \) if there is \( t \) such that

\[ A_{\sigma(t)} \ldots A_{\sigma(2)} A_{\sigma(1)} \{ x : \|x\| \leq 1 \} \subset \text{int}(\{ x : \|x\| \leq 1 \}). \]

One can easily verify that if all products of length \( k \) of matrices in \( \Sigma \) contract a seminorm \( \| \| \), then all trajectories \( x(t) \) of the corresponding switched system (1) asymptotically approach the set \( \{ x : \|x\| = 0 \} \), and that their distance to that set decay exponentially. In particular, if \( \| \| \) is a norm, \( x \) converges exponentially to 0. And if \( \| \| \) is the seminorm \( \|x\|_P = \frac{1}{2}(\max_i x_i - \min_i x_i) \) – a seminorm that is nonincreasing for stochastic matrices – then \( x \) approaches the consensus space set \( \{ a1 \} \), and we have proved in previous work [7] that each trajectory actually converges in that case to one specific point in that set, as opposed to just approaching the set. For these reasons, we will investigate contraction of seminorms, keeping in mind that this question is intimately related to that of convergence.
**Question 1.** Let $\| \cdot \|$ be a polyhedral seminorm; what is the smallest $k$ such that for any set $\Sigma$ for which $\| \cdot \|$ is nonincreasing, the existence of an infinite non-contracting product implies the existence of an infinite periodic non-contracting product with period smaller than or equal to $k$?

**2 General Case**

In this section, we answer Question 1. We find that the answer is equal to half the size of the largest antichain in the face lattice of the unit ball of the nonincreasing seminorm. We also link the size of the largest antichain with the Sperner property. We provide an example of a norm whose face lattice of its unit ball does not have the Sperner property.

We start by recalling some definitions. As regards poset and polyhedron terminology, we use the definitions of [23]. A partially ordered set is a set $P$ with a binary relation $\preceq$ that is transitive, antisymmetric and reflexive. We also note $x \prec y$ the relation $x \preceq y$ and $x \neq y$. An antichain is a set $X \subset P$ of elements that are not comparable:

$$\forall x, y \in X, x \not\prec y \text{ and } y \not\prec x.$$ 

A poset $(P, \preceq)$ is called graded if it can be equipped with a rank function $r : P \mapsto \mathbb{N}$ such that

$$x \preceq y \Rightarrow r(x) \leq r(y)$$

and

$$(y \prec x \text{ and } \not\exists z, y \prec z \prec x) \Rightarrow r(x) = r(y) + 1.$$ 

A poset is called a lattice if any pair of elements has a unique infimum and a unique supremum.

**Definition 1** (Faces of a Polyhedron). A non-empty subset $F$ of a polyhedron $Q$ is called a face or closed face if $F = Q$, if $F = \emptyset$ or if it can be represented as $F = Q \cap \{x : b^\top x = c\}$ where $b \in \mathbb{R}^n$, $c \in \mathbb{R}$ are such that

$$\forall x \in Q, b^\top x \leq c.$$ 

If the face contains exactly $d + 1$ affinely independent points, that is, points $u_0, u_1, \ldots, u_d$ such that $u_1 - u_0, u_2 - u_0, \ldots, u_d - u_0$ are linearly independent, we call $d$ the dimension of the face. A proper face is a face that is neither the polyhedron itself nor the empty face. An open face is the relative interior of a face.

It is well known that faces of any dimension are intersections of facets and their number is therefore finite. It is also known that any polyhedron decomposes into a disjoint union of open faces.

**Definition 2** (Face Lattice). Given a centrally symmetric polyhedron $Q$, we call face lattice the poset $(P, \subseteq)$ where $P$ is the set of (closed) faces of $Q$ and $F_1 \subseteq F_2$ is the inclusion relation. In the case of polytopes, it is well-known that this poset is a graded lattice. This is also the case for unbounded centrally symmetric polyhedra. A rank function is given by the dimension of the faces and $d_{\min} - 1$ for the empty face, where $d_{\min}$ is the lowest dimension of faces of $Q$. 

4
Definition 3 (Width of a Poset). We call the width \( W(P) \) of a poset \( P \) the number of elements of the largest antichain of \( P \). We also write \( W(Q) \) for the width of the face lattice of the polyhedron \( Q \).

The following Lemma by Lagarias and Wang allows abstracting Question 1 as a combinatorial problem, as it shows that matrices in \( \Sigma \) can be completely represented (for our purpose) as functions sending each face of the invariant polyhedron onto another one.

**Lemma 1** (Lagarias and Wang [15]). Let \( \Sigma \) be a finite set of matrices having a common invariant polyhedron \( Q \). Then, for any \( A \in \Sigma \) and any open face \( O_1 \) of \( Q \), there exists exactly one open face \( O_2 \) (possibly int(\( Q \))) such that \( AO_1 \subseteq O_2 \).

**Theorem 1.** For any seminorm \( ||.|| \), the answer to Question 1 is equal to the half of the width of the face lattice of the unit ball of the seminorm:

\[
  k^* = \frac{W(B)}{2}, \quad \text{with } B = \{ x : ||x|| \leq 1 \}.
\]

**Proof.** **Claim 1:** \( k^* \leq \frac{W(B)}{2} \)

We first prove that \( k^* \) is finite. Let us suppose that there exist an infinite non-contracting product \( \ldots A_{\sigma(2)} A_{\sigma(1)} \ldots \) and a point \( x_0 \) such that

\[
  \forall i, A_{\sigma(i)} \ldots A_{\sigma(1)} x_0 \notin \text{int}(B).
\]

Since the number of faces is finite, we have that there is a proper open face \( F \) and indices \( i < j \) such that

\[
  A_{\sigma(i)} \ldots A_{\sigma(1)} x_0 \in F \quad \text{and} \quad A_{\sigma(j)} \ldots A_{\sigma(1)} x_0 \in F.
\]

By Lemma 1, we have

\[
  A_{\sigma(j)} \ldots A_{\sigma(i+1)} F \subseteq F.
\]

Therefore, the infinite power of \( A_{\sigma(j)} \ldots A_{\sigma(i+1)} \) is an infinite periodic non-contracting product, proving that \( k^* \) is finite.

We now prove the \( k^* \leq \frac{W(B)}{2} \). Let \( \ldots P^* P^* P^* \) be the infinite non-contracting product with the smallest period \( k^* \) and

\[
  P^* = A_{\sigma^*(k^*)} \ldots A_{\sigma^*(1)}.
\]

Let \( O_1 \) be an open face of \( B \) such that

\[
  \forall t \geq 0, \quad (P^*)^t O_1 \subseteq \text{int}(B)
\]

(such a face exists thanks to Lemma 1 and the fact that \( \ldots P^* P^* \) is non-contracting), let \( O_2 \) be the open face containing \( A_{\sigma^*(1)} O_1 \) (by Lemma 1 there is exactly one such face), \( O_3 \) containing \( A_{\sigma^*(2)} A_{\sigma^*(1)} O_1 \) up to \( O_{k^*} \) containing \( A_{\sigma^*(k^*-1)} \ldots A_{\sigma^*(1)} O_1 \). Let also \( F_1 = \text{cl}(O_1) \), \( \ldots, F_{k^*} = \text{cl}(O_{k^*}) \).
We prove that \( \{F_1, \ldots, F_k\} \) is an antichain in the face lattice of \( \mathcal{B} \). Suppose, to obtain a contradiction, that for some \( i, j \) with \( i > j \), \( F_i \subseteq F_j \). Then
\[
A_{\sigma(i-1)} \ldots A_{\sigma(j)} F_j \subseteq F_i \subseteq F_j
\]
hence
\[
\forall t \geq 0, \ (A_{\sigma(i-1)} \ldots A_{\sigma(j)})^t F_j \subseteq F_j.
\]
This contradicts the assumption that \( \ldots P^* P^* \ldots \) is the infinite periodic non-contracting product with the smallest period. Similarly, if for some \( i < j \), \( F_i \subseteq F_j \), then
\[
\forall t \geq 0, \ (A_{\sigma(i-1)} A_{\sigma(1)} A_{\sigma(k^*)} \ldots A_{\sigma(j)})^t F_j \subseteq F_j,
\]
and again we would have a contradiction. Therefore, there is no \( i \neq j \) such that \( F_i \subseteq F_j \), and the set of faces \( \{F_1, \ldots, F_k^*\} \) is an antichain.

We now prove that \( \{O_1, \ldots, O_{k^*}\} \) does not contain opposite faces. Indeed, if it contains opposite faces \( O_i = -O_j \), then the product \( A_{\sigma(i-1)} \ldots A_{\sigma(j)} \) is shorter than \( P^* \) and its powers are non-contracting, contradicting our assumption that \( P^* P^* \ldots \) is the non-contracting infinite product with the smallest period.

The set of faces \( \{F_1, \ldots, F_k^*\} \) is an antichain and does not contain opposite faces. Since each proper face has an opposite face, every antichain that does not contain opposite faces has at most \( \frac{W(B)}{2} \) elements. Therefore, \( k^* \leq \frac{W(B)}{2} \).

Claim 2: \( k^* \geq \frac{W(B)}{2} \). We construct a set of matrices such that the infinite non-contracting product that has the smallest period has a period equal to \( \frac{W(B)}{2} \). Let \( X = \{F_1, \ldots, F_l\} \) be the largest antichain in the face lattice, let \( \{F_1, \ldots, F_m\} \) be the largest subset of \( X \) that does not contain opposite faces and let \( O_1, \ldots, O_m \) be the corresponding open faces.

By definition, each proper face is the intersection of \( \mathcal{B} \) with a hyperplane
\[
F_i = \mathcal{B} \cap \{x : b_i^\top x = c_i\}
\]
such that \( \mathcal{B} \) is in one halfspace defined by the hyperplane:
\[
\mathcal{B} \subseteq \{x : b_i^\top x \leq c_i\}.
\]
Furthermore, \( c_i \neq 0 \). If \( c_i = 0 \), then \( \mathcal{B} \subseteq \{x : b_i^\top x \leq 0\} \) and because \( \mathcal{B} \) is the unit ball of a seminorm, it is centrally symmetric around the origin, \( \mathcal{B} \subseteq \{x : -b_i^\top x \leq 0\} \) and therefore \( F_i = \mathcal{B} \cap \{x : b_i^\top x = 0\} = \mathcal{B} \) and \( F_i \) is not a proper face. Because \( c_i \neq 0 \), we can scale \( b_i \) and \( c_i \) to have \( \forall i, \ c_i = 1 \).

By taking any \( v_i \) in the open face \( O_{(i \mod m)+1} \) and defining
\[
A_i = v_i b_i^\top \text{ and } \Sigma = \{A_1, \ldots, A_m\},
\]
we have
\[
\forall i, \ A_i F_i = A_i (\mathcal{B} \cap \{x : b_i^\top x = 1\})
\]
\[
\subseteq A_i \{x : b_i^\top x = 1\}
\]
\[
= \{A_i x : b_i^\top x = 1\}
\]
\[
= \{v_i b_i^\top x : b_i^\top x = 1\}
\]
\[
= \{v_i\}
\]
\[
\subseteq O_{(i \mod m)+1}.
\]
Because $\mathcal{B}$ is centrally symmetric, we have that $-F_i$ is a face and that it is equal to \{ $x : -b_i^T x = 1$ \}. We then have

$$
\forall i, A_i(\mathcal{B} \setminus (F_i \cup -F_i)) = A_i(\mathcal{B} \cap \{ x : -1 < b_i^T x < 1 \}) \\
\subseteq \{ v_i b_i^T x : -1 < b_i^T x < 1 \} \\
= \{ \lambda v_i : -1 < \lambda < 1 \} \\
\subseteq \{ \lambda y : -1 < \lambda < 1, y \in \mathcal{B} \}
$$

(3)

By (2) and (3), for any $j \neq (i \mod m) + 1$ and any subset $S$ of $\mathcal{B}$,

$$
A_j A_i S \subseteq A_j (\text{int}(\mathcal{B})) \cup O_{(i \mod m)+1} = A_j \text{int}(\mathcal{B}) \cup A_j O_{(i \mod m)+1} \subseteq \text{int}(\mathcal{B}).
$$

Therefore,

$$
\ldots A_{(h+2 \mod m)+1} A_{(h+1 \mod m)+1} A_{(h \mod m)+1} A_h
$$

is the only infinite non-contracting product starting with $A_h$. Because the matrices are $A_1, \ldots, A_m$ are all different, for any $h$, it has a period of $m$. That is, the non-contracting infinite product with the smallest period has a period of $m$. With $X$ being the largest antichain, we obtain $m = \frac{W(\mathcal{B})}{2}$ and we conclude that $k^* \geq \frac{W(\mathcal{B})}{2}$.

Given an explicit value to the size of the largest antichain may prove difficult in some cases However, since a set of faces of same dimension always constitute an antichain, we have the following lower bound

$$
k^* = \frac{W(\mathcal{B})}{2} \geq \max_i f_i,
$$

where $f_i$ is the number of faces of dimension $i$. Equality holds if the face lattice of $Q$ has the so-called Sperner property.

**Definition 4 (Sperner Property [10]).** A graded poset is said to have the Sperner property if no antichain in it is larger than the width of the poset.

Not all face lattices have the Sperner property. In appendix A, Example 1 presents a seminorm that has a unit ball whose face lattice does not have the Sperner property.

### 3 Stochastic Matrices

We now investigate sets of stochastic matrices, with respect to which the following seminorm is always nonincreasing

$$
\|x\|_P = \frac{1}{2} (\max_i x_i - \min_i x_i).
$$

The (polyhedral) unit ball of that seminorm

$$
\mathcal{P} = \left\{ x : \frac{1}{2} (\max_i x_i - \min_i x_i) \leq 1 \right\},
$$

is thus invariant under multiplication by any stochastic matrix. We prove that the face lattice of this polyhedron has the Sperner property, allowing us to compute an explicit value for our bound $k^*$. 

7
Definition 5 (Upper and Lower Shadow [10]). Let \((P, \leq)\) be a graded poset and let \(S \subset P\) be such that \(\exists k\) such that \(\forall x \in S, \text{rank}(x) = k\). We call the upper shadow
\[
\nabla(S) = \{x \in P : \exists y \in S, y \leq x, \text{rank}(x) = k + 1\}.
\]
Similarly, we define the lower shadow
\[
\Delta(S) = \{x \in P : \exists y \in S, x \leq y, \text{rank}(x) = k - 1\}.
\]

\(P\) has no face of dimension 0 because \(\forall x \in P, a \in \mathbb{R}, x + a1 \in P\). The face of dimension \(n\) is equal to \(P\) itself. Each face of dimension \(2 \leq d \leq n - 1\) of \(P\) can be written as
\[
F = \{x \in \partial P : \forall i \in S_m, x_i = \min_j x_j, \forall i \in S_M, x_i = \max_j x_j\}
\]
for certain disjoint non-empty sets \(S_m, S_M \in \{1, \ldots, n\}\) with \(|S_m \cup S_M| = n - (d - 1)|\).

Therefore, the lower shadow of each single face \(F\) of \(P\) of dimension \(2 \leq d \leq n - 1\) contains
\[
|\Delta(\{F\})| = 2(d - 1)
\]
elements (the faces obtained by adding an element to either \(S_m\) or \(S_M\)) and the upper shadow has
\[
|\nabla(\{F\})| = n - d + 1 \text{ or } |\nabla(\{F\})| = n - d
\]
elements (the faces obtained by removing an element from either \(S_m\) or \(S_M\) while keeping them both non-empty).

Theorem 2. The face lattice of \(P\) has the Sperner property.

Proof. Let \(S = \{F_1, \ldots, F_{|S|}\}\) be any set of faces of \(P\) of the same dimension \(d\), that is, a subset of a level in the face lattice. Let \(E_+\) be the set of pairs of faces of \(S\) and \(\nabla(S)\) being neighbors to each other:
\[
E_+ = \{(F_1, F_2) : F_1 \in S, F_2 \in \nabla(\{F_1\})\}.
\]
Since the upper shadow of each element of \(S\) has at least \(n - d\) elements, we have
\[
|E_+| \geq |S|(n - d).
\]
Since the lower shadow of each element of \(\nabla S\) contains exactly \(2d\) elements – not all of which belonging to \(S\) –, we have
\[
|E_+| \leq |\nabla(S)|2d.
\]
Combining the two inequalities, we obtain,
\[
\forall d \leq \frac{n}{3}, \quad |\nabla(S)| \geq |S|, \quad |\nabla(S)| \geq |S| \frac{n - d}{2d}.
\]
(4)

By a similar reasoning, we obtain
\[
\forall d \geq \frac{n + 4}{3}, \quad |\Delta(S)| \geq |S|, \quad |\Delta(S)| \geq |S| \frac{2(d - 1)}{n - d + 2}.
\]
(5)
Let now \( X \) be the largest antichain, let \( d^- \) be the smallest dimension in which \( X \) contains elements and let \( S^- \) be the intersection of the antichain with the level \( d^- \). If \( d^- \leq \frac{n}{3} \), Equation (4) tells us that the antichain

\[
(X \setminus S^-) \cup \nabla(S^-)
\]

has at least as many elements as \( X \). We can repeat this process until the antichain contains only faces of dimension strictly larger than \( \frac{n}{3} \). Similarly we use (5) to obtain an antichain with at least as many elements of rank strictly smaller than \( \frac{n+4}{3} \). Since

\[
\frac{n}{3} < d^* < \frac{n+4}{3}
\]

has a unique integer solution \( d^* = \lfloor n/3 \rfloor + 1 \), the final antichain contains only faces of dimension \( d^* \).

### 3.1 New Finiteness Bound for Consensus

We have computed in [7] that the number of faces of dimension \( d \) of \( P \) is

\[
f_d = \frac{n}{d-1}(2^{n-d+1} - 2).
\]

Therefore, the answer to Question \( 1 \) for the case of stochastic matrices is

\[
k^* = \frac{f_{d^*}}{2} = \frac{n}{\lfloor n/3 \rfloor}(2^{n-\lfloor n/3 \rfloor} - 1).
\]

As we discussed in Section 1, contraction of the seminorm \( \| . \|_P \) is intimately linked with convergence. More formally, we have the following proposition.

**Proposition 1.** Let \( \Sigma \) be a set of stochastic matrices. System (1) converges to consensus for any initial condition and any sequence of transition matrices from \( \Sigma \) if and only if it converges to consensus for any initial condition and any periodic sequence of transition matrices from \( \Sigma \), whose period is smaller than or equal to \( k^* = \left( \frac{n}{\lfloor n/3 \rfloor} \right)(2^{n-\lfloor n/3 \rfloor} - 1) \).

**Proof.** This is a consequence of the definition of \( k^* \), of [7, Proposition 1.a], and of the computation of \( k^* \).

An infinite periodic product is simply the infinite power of a finite product. The convergence of powers of finite products can be verified by checking that the modulus of the second eigenvalue of the product is smaller than 1 [7]. The convergence for all products with period bounded by \( k^* \) can thus be verified by checking the second eigenvalue of every products of length smaller than or equal to \( k^* \).

A finiteness result such as Proposition 1 was known [4, 18] with \( B = \frac{3n - 2n+1}{2} \) instead of \( k^* \). We show in Appendix B that for \( n \geq 8 \),

\[
k^* \leq \frac{3}{2\sqrt{\pi n}}B;
\]

so that our bound \( k^* \) improves the previously known bound by a multiplicative factor of \( \frac{3}{2\sqrt{\pi n}} \).

\(^1\)We recall that we investigate convergence to a rank one matrix and that the largest eigenvalue of every stochastic matrix is equal to one and corresponds to the eigenvector 1.
Figure 1: In blue: ratio between our bound $k^*$ and the previous bound $B$. This ratio is below one (red) and close to $\frac{3}{2\sqrt{\pi n}}$ (green).

**Conclusion**

Deciding the asymptotic convergence of long matrix products has many applications in engineering and computer science \[3, 8\]. In this paper, we have studied this problem for the case where the given set of matrices admits a nonincreasing polyhedral seminorm, and one wonders whether all long products of these matrices send the state space onto points whose seminorm is equal to zero (the so-called consensus problem is a particular case of this setting). We have significantly improved the available bound by leveraging the combinatorial structure of (an abstraction of) the dynamical system described by these matrices.

We see several further directions for our work: a major tool in our analysis is Lemma 1 from Lagarias and Wang. In their paper, these authors also provide a similar result when the invariant set is not a polyhedron, but has a more involved algebraic structure (namely, piecewise analytic). We believe that our analysis could be applied further to this case, but it is not clear whether there would be particular relevant applications in that setting.
References

[1] J. M. Anthonisse and H. Tijms. Exponential convergence of products of stochastic matrices. *Journal of Mathematical Analysis and Its Applications*, 598:360–364, 1977.

[2] B. Bamieh, M. R. Jovanovic, P. Mitra, and S. Patterson. Coherence in large-scale networks: dimension dependent limitations of local feedback. *IEEE Transactions on Automatic Control*, 57(9):2235–2249, 2012.

[3] V. D. Blondel, J. Cassaigne, and R. M. Jungers. On the number of $\alpha$-power-free words for $2 < \alpha \leq 7/3$. *Theoretical Computer Science*, 410:2823–2833, 2009.

[4] V. D. Blondel and A. Olshevsky. How to decide consensus? A combinatorial necessary and sufficient condition and a proof that consensus is decidable but NP-hard. *SIAM Journal on Control and Optimization*, 52(5):2707–2726, 2014.

[5] V. D. Blondel and J. N. Tsitsiklis. The boundedness of all products of a pair of matrices is undecidable. *Systems and Control Letters*, 41:135–140, 2000.

[6] V.D. Blondel, J.M. Hendrickx, A. Olshevsky, and J.N. Tsitsiklis. Convergence in multiagent coordination, consensus, and flocking. In *Proceedings of the 44th IEEE Conference on Decision and Control*, 2005.

[7] P.-Y. Chevalier, J. M. Hendrickx, and R. M. Jungers. Efficient algorithms for the consensus decision problem. To appear in *SIAM Journal on Control and Optimization*, 2015.

[8] I. Daubechies and J. C. Lagarias. Sets of matrices all infinite products of which converge. *Linear Algebra and its Applications*, 161:227–263, 1992.

[9] J. Eckhoff. Combinatorial properties of f-vectors of convex polytopes. *Nordisk Matematisk Tidsskrift*, 54:146–159, 2006.

[10] Konrad Engel. *Sperner Theory*. Cambridge University Press, 1997.

[11] B. Grünbaum. Convex polytopes. In *Graduate Text in Mathematics*. Springer, New York, 2003.

[12] E. A. Hernandez-Vargas, R. H. Middleton, and P. Colaneri. Optimal and mpc switching strategies for mitigating viral mutation and escape.

[13] A Jadbabel and J Lin. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48(6):988–1001, 2003.

[14] R. M. Jungers. *The Joint Spectral Radius: Theory and Applications*. Springer, 2009.

[15] J. C. Lagarias and Y. Wang. The finiteness conjecture for the generalized spectral radius of a set of matrices. *Linear Algebra and its Applications*, 214:17–42, 1995.

[16] B. E. Moision, A. Orlitsky, and P. H. Siegel. On codes that avoid specified differences. *IEEE Transactions on Information Theory*, 47:433–442, 2001.
[17] R. Olfati-Saber and J. S. Shamma. Consensus filters for sensor networks and distributed sensor fusion. In Proceedings of the 44th IEEE Conference on Decision and Control.

[18] A. Paz. Introduction to Probabilistic Automata. Academic Press, New York, 1971.

[19] R. Shorten, F. Wirth, and D. Leith. A positive systems model of tcp-like congestion control: asymptotic results. IEEE/ACM Transactions on Networking, 14(6), 2006.

[20] R. Shorten, F. Wirth, O. Mason, K. Wulff, and C. King. Stability criteria for switched and hybrid systems. SIAM Review, 49(4):545–592, 2007.

[21] J. Wolfowitz. Products of indecomposable, aperiodic, stochastic matrices. Proceedings of the American Mathematical Society, 15:733–736, 1963.

[22] L. Xiao, S. Boyd, and S.-J. Kim. Distributed average consensus with least-mean-square deviation. Journal of Parallel and Distributed Computation, 67(1):33–46, 2007.

[23] B. Ziegler. Lectures on polytopes. In Graduate Text in Mathematics. Springer, New York, 1995.
Appendices

A Example of a norm that does not lead to Sperner property

We construct a norm for which the face lattice of the unit ball does not have the Sperner property. To construct our norm, we construct a polytope \( P \) that is centrally symmetric around the origin. The norm of a vector \( x \) is then simply defined as the scalar \( \|x\| \) such that \( \frac{x}{\|x\|} \) is in the boundary of the polytope. It can be easily verified that this application is a norm. In particular, the triangular inequality follows from the convexity of the polytope.

Before presenting the example, we need some definitions.

**Definition 6** (Simplicial and simple polytopes). An \( n \)-dimensional simplicial polytope is a polytope whose facets contain exactly \( n \) vertices. An \( n \)-dimensional simple polytope is a polytope whose vertices are adjacent to exactly \( n \) edges.

**Definition 7** (F-vector). The f-vector of a polytope of dimension \( n \) is the \( n \)-dimensional vector whose \( i \)th element is equal to the number of faces of dimension \( i - 1 \).

For example, the f-vector of a cube is equal to \((8 \quad 12 \quad 6)\). We give here an example of a centrally symmetric polytope of dimension 4 whose face lattice does not have the Sperner property.

**Example 1.** We use the connected sum \( \# \) presented in [9, 23]. The idea of the construction is to cut off a vertex from a simple polytope \( P' \), to apply a suitable projective transformation \( T \) to the rest of \( P' \) and to glue it on a facet of a simplicial polytope \( P \). The transformation \( T \) can be chosen such that the result \( P \# P' \) is a convex polytope whose faces are those of \( P \) except the facet on which the transformed sliced \( P' \) was glued, and those of \( P' \) except the cut off vertex. In particular, all open faces of \( P' \), except the vertex that has been cut off, remain present (albeit modified) and do not intersect with \( P \). Similarly, all faces of \( P \), except the face on which the modified \( P' \) has been glued, remain present and do not intersection with the modified \( P' \). In turn, the f-vector of \( P \# P' \) is equal to

\[
\begin{align*}
f + f' - \begin{pmatrix} 1 & 0 & \ldots & 0 & 1 \\
0 & 1 & \ldots & f_{n-1} & f'_{n-1} - 1
\end{pmatrix},
\end{align*}
\]

where \( f \) is the f-vector of \( P \) and \( f' \) that of \( P' \). We use this construction twice to glue two hypercubes

\[
P' = \{x : \|x\|_\infty \leq 1\}
\]
on opposite faces of an hyperoctahedron

\[
P = \{x : \|x\|_1 \leq 1\}.
\]

The result is a convex polytope with f-vector

\[
\begin{align*}
2f + f' - \begin{pmatrix} 2 & 0 & 0 & 2 \\
2 & 8 & 32 & 24 & 8
\end{pmatrix} &= 2 \begin{pmatrix} 16 & 32 & 24 & 8 \\
8 & 24 & 32 & 16
\end{pmatrix} + \begin{pmatrix} 2 & 0 & 0 & 2 \\
8 & 24 & 32 & 16
\end{pmatrix} - \begin{pmatrix} 2 & 0 & 0 & 2 \\
8 & 24 & 32 & 16
\end{pmatrix} \\
&= \begin{pmatrix} 38 & 88 & 80 & 30 \end{pmatrix}.
\end{align*}
\]

\(^2\)that is, a bounded polyhedron [23]
To construct a large antichain in the face lattice, we can take the 64 faces of dimension 1 from the hypercubes and the 32 faces of dimension 2 of the hyperoctahedron. This set of faces is an antichain in the face lattice because:

- Each face of dimension 1 of the hypercubes contains two vertices. At least one of them is not the cut off vertex. After transformation, this vertex does not belong to $P$. Therefore no face of $P$ can contain a transformed face of dimension 1 of the hypercubes.

- The faces of dimension 2 of $P$ cannot be subsets of the faces of dimension 1 of the transformed hypercube, because of the dimension.

- The faces of dimension 2 of $P$ are different and therefore cannot be included in one another.

- Similarly, the faces of the modified hypercube cannot be included in one another.

We conclude that the set that we constructed is an antichain in the face lattice. We obtain an antichain of $96 > 88$ elements, which proves that the polytope does not have the Sperner property. The polytope can be chosen to be centrally symmetric (for instance, $P\#P' \cup -(P\#P')$). To illustrate, Figures 2, 3, 4 and 5 present the same construction in dimension 3. The difference is that the antichain that is obtained is not large enough and the polytope is Sperner.

Figure 2: sliced cube. In green a set of faces making up an antichain in the face lattice and with none of these faces being entirely in the new face $C \cap \{x_1 + x_2 + x_3 \leq 2\}$.

Figure 3: A transformation of the sliced cube, preserving its combinatorial structure.

Technical details

We now present the transformation that makes Example 1 possible. We prove that we indeed obtain a convex polytope with the structure that we claimed. We start by recalling some basics on projective transformations, a class of transformations that preserve colinearity.
Definition 8 (Projective transformation, nonsingularity, permissibility \[11\]). A projective transformation is a transformation of the form

\[ T(x) = \frac{Ax + b}{c^\top x + \delta}, \]

with \( A \in \mathbb{R}^{n \times n}, b, c \in \mathbb{R}^n, \delta \in \mathbb{R} \) and at least one of \( c \) and \( \delta \) being different from zero. The transformation is called nonsingular if

\[ \begin{pmatrix} A & b \\ c^\top & \delta \end{pmatrix} \]

is regular. The transformation is called permissible for a polytope \( Q \) if

Definition 9 (Combinatorial equivalence \[11\]). Two polytopes are said combinatorially equivalent if their face lattices are the same.

Theorem 3 (Second part of Assertion 3 of page 38 of \[11\]). If \( T \) is a nonsingular projective transformation, permissible for \( Q \), then \( Q \) and \( T(Q) \) are combinatorially equivalent.

We now present formally the 4-dimensional construction. The hyperoctahedron \( O \) has equation

\[ \forall b \in \{0, 1\}^4, \ b^\top x \leq 1. \]

The sliced hypercube \( C \) has equation

\[ \forall i \in \{1, 2, 3, 4\}, \ |x_i| \leq 1 \]

\[ \sum_{i=1}^{4} x_i \leq 3. \]

We apply to the sliced cube the projective transformation

\[ T(x) = \frac{(-I - \frac{2}{3} \mathbf{1} \mathbf{1}^\top) x + 3 \mathbf{1}}{3.1^\top x - 10}. \]

We want to prove that

\[ T(C) \cup O \cup -T(C) \]

is a polytope and that its f-vector is \( 2f + f' - (2 \ 0 \ 0 \ 2) \) as claimed.
Lemma 2. The set $T(C)$ is a full-dimensional polytope with the same combinatorial structure as $C$.

Proof. We use Theorem 3 and we prove the non-singularity and the permissibility for polytope $C$. Non-singularity. We have to prove that
\[
\begin{pmatrix}
-I - \frac{2}{3} 11^\top & 3.1 \\
3.1^\top & -10
\end{pmatrix}
\]
is full rank. Indeed the first block
\[-I - \frac{2}{3} 11^\top\]
is negative definite. On the other hand, the Schur complement
\[-I - \frac{2}{3} 11^\top + \frac{9}{10} 11^\top = -I + \frac{7}{30} 11^\top\]
is diagonally dominant and therefore also negative definite. In turn, the matrix is negative definite and hence, nonsingular. Permissibility.
\[
\forall x \in C, \ 3.1^\top x - 10 \neq 0,
\]
as a consequence of
\[
\sum_{i=1}^{4} x_i \leq 3.
\]
\[\square\]

Proposition 2. The set $T(C) \cup \mathcal{O} \cup -T(C)$ is a convex polytope.

Proof. Vertices and facets of the transformed sliced cube. The vertices of the transformed sliced cube $T(C)$ are
\[
T\left(\begin{pmatrix} 0 & 1 & 1 & 1 \end{pmatrix}\right) = \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}
\]
\[
T\left(\begin{pmatrix} -1 & 1 & 1 & 1 \end{pmatrix}\right) = \frac{1}{6} \begin{pmatrix} 4 & 1 & 1 & 1 \end{pmatrix}
\]
\[
T\left(\begin{pmatrix} -1 & -1 & 1 & 1 \end{pmatrix}\right) = \frac{1}{5} \begin{pmatrix} 2 & 2 & 1 & 1 \end{pmatrix}
\]
\[
T\left(\begin{pmatrix} -1 & -1 & -1 & 1 \end{pmatrix}\right) = \frac{1}{24} \begin{pmatrix} 8 & 8 & 8 & 5 \end{pmatrix}
\]
\[
T\left(\begin{pmatrix} -1 & -1 & -1 & -1 \end{pmatrix}\right) = \frac{10}{33} \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}
\]
and the points obtained by permuting elements of these vectors. We define $e^i$ to be the vector with its $i^{th}$ element equal to 1 and the others equal to zero. One can verify that the facet inequalities of $T(C)$ are the following inequalities (every $1 \leq i \leq 4$ generates different facets).
\[
(1 - e^i)^\top x \leq 1
\]
\[
\begin{pmatrix} 8.1 + e^i \\ 10 \end{pmatrix}^\top x \leq 1
\]
\[
\sum_{j=1}^{4} x_j \geq 1.
\]
All facets of $O$ except two are facets of $T(C) \cup O \cup -T(C)$. We notice that

$$\forall x \in T(C), \ b \in \{0,1\}^4 \setminus \{(1 \ 1 \ 1 \ 1)\}, \ b^\top x \leq 1$$

$$\forall x \in T(C), \ \sum_{i=1}^{4} x_i \geq 1,$$

which means that each inequality of the form $O$

$$b^\top x \leq 1$$

with

$$b \in \{0,1\}^4 \setminus \{(1 \ 1 \ 1 \ 1)\}$$

is satisfied by any element of $T(C)$. Hence, any facet except that with

$$b = (1 \ 1 \ 1 \ 1)$$

of the octahedron is satisfied by any point of $T(C)$. By a similar argument on points of $-T(C)$, we have that all facets of $O$ except those with $b = (1 \ 1 \ 1 \ 1)$ and $b = (-1 \ -1 \ -1 \ -1)$ are still facets of $T(C) \cup O \cup -T(C)$. All facets of $T(C)$ except one are facets of $T(C) \cup O \cup -T(C)$. It is easy to verify that, for every $1 \leq i \leq 4$,

$$\forall x \in O \cup -T(C), \ (1 - e^i)^\top x \leq 1$$

$$\forall x \in O \cup -T(C), \ \left(\frac{8.1 + e^i}{10}\right)^\top x \leq 1$$

which again means that any facet of $T(C)$ except the facet by which it is glued to $O$ is still a face of $T(C) \cup O \cup -T(C)$. **Convexity.** Convexity is now easy to prove. Let us prove, for example, that every convex combination of a point in $T(C)$ and one in $O$ except the facet by which they are glued (i.e., the facet $(1 \ 1 \ 1 \ 1) x \leq 1$ of $O$, and the facet $(1 \ 1 \ 1 \ 1) x \geq 1$ of $T(C)$). Therefore every convex combination of $x$ and $y$ satisfies these same facet constraints. If this convex combination satisfies $(1 \ 1 \ 1 \ 1) x \leq 1$ then it belongs to $O$. Otherwise, it belongs to $T(C)$. In both cases, it belongs to $T(C) \cup O \cup -T(C)$.

The value of the f-vector follows from the convexity and from the fact that faces originating from $C$ and those originating from $O$ are different.

**B Comparison of bounds**

**Proposition 3.** For stochastic matrices, for $n \geq 8$, our bound $k^*$ improves the previously known bound $B$ by a multiplicative factor equal to $\frac{3}{2\sqrt{\pi n}}$:

$$k^* \leq \frac{3}{2\sqrt{\pi n}} B.$$
Proof. We will make use of the Gamma function $\Gamma(k)$. When $k$ is integer, we have that $\Gamma(k) = (n-1)!$. We will also use the Stirling formula to bound the Gamma function:

$$\sqrt{2\pi k} \left(\frac{k}{e}\right)^k e^{\frac{k}{12k+1}} < \Gamma(k+1) < \sqrt{2\pi k} \left(\frac{k}{e}\right)^k e^{1/k}.$$  

Let us start the proof by upper bounding $\left(\frac{n}{\lfloor \frac{n}{3} \rfloor}\right)^n$, which appears in the expression of $k^*$.

$$\left(\frac{n}{\lfloor \frac{n}{3} \rfloor}\right)^n = \frac{n!}{\lfloor \frac{n}{3} \rfloor! \left\lceil \frac{2n}{3} \right\rceil!} \leq \frac{n!}{\Gamma(\frac{4}{3}+1)\Gamma(\frac{\frac{2n}{3}}{3}+1)}$$

$$\leq \frac{\sqrt{2\pi n} \left(\frac{n}{\frac{2n}{3}}\right)^{\frac{n}{3}} \left(\frac{\frac{2n}{3}}{e}\right)^{\frac{2n}{3}} e^{\frac{1}{12\left(\frac{n}{2n}\right)^{3}+1} e^{\frac{1}{12\left(\frac{n}{2n}\right)^{3}+1}}}}{\sqrt{(2\pi)^2 \frac{2n}{3} \left(\frac{n}{3} \frac{1}{n}\right)^{\frac{n}{3}} \left(\frac{2n}{3} \frac{1}{n}\right)^{\frac{2n}{3}} e^{\frac{1}{12n} e^{-\frac{1}{12n} - \frac{1}{12n+1}}}}$$

$$= \frac{3\sqrt{\pi n} \left(\frac{1}{n}\right)^{\frac{n}{3}} \left(\frac{2n}{3} \frac{1}{n}\right)^{\frac{2n}{3}} e^{-\frac{112n^2-12n+1}{384n^3 + 144n^2 + 12n}}}{2\sqrt{\pi n} 2^{\frac{2n}{3}} + 1}$$

$$= \frac{3\sqrt{\pi n} \left(\frac{1}{n}\right)^{\frac{n}{3}} \left(\frac{2n}{3} \frac{1}{n}\right)^{\frac{2n}{3}} e^{-\frac{112n^2-12n+1}{384n^3 + 144n^2 + 12n}}}{2\sqrt{\pi n} 2^{\frac{2n}{3}} + 1}$$

We can now bound $k^*$:

$$k^* = \left(\frac{n}{\frac{2n}{3}}\right)^{\frac{n}{3}} \left(\frac{2n}{3} - 1\right) \leq \frac{3^{n+1}}{2\sqrt{\pi n} 2^{\frac{2n}{3}} + 1} e^{-\frac{112n^2-12n+1}{384n^3 + 144n^2 + 12n}} \left(\frac{2n}{3} - 1\right)$$

$$= \frac{3^{n+1}}{2\sqrt{\pi n} 2^{\frac{2n}{3}} + 1} e^{-\frac{112n^2-12n+1}{384n^3 + 144n^2 + 12n}} \left(\frac{2n}{3} - 1\right)$$

The first factor is the improvement over the old bound. The second factor is a quantity that, for most values of $n$ is slightly larger than the previously known bound $B = \frac{3^n - 2^{n+1} + 2}{2}$ because $\frac{3}{2} \approx 1.89 < 2$. The third factor is a quantity slightly smaller than 1, which, as we will see, allows to compensate for the excess of the second factor. For $n \geq 8$, we have

$$3^{8n^3 + 144n^2 + 12n} \leq 4n(112n^2 + 12n - 1)$$

hence

$$e^{-\frac{112n^2-12n+1}{384n^3 + 144n^2 + 12n}} \leq e^{-\frac{1}{11n}}$$
and

\[ k^* \leq \frac{3}{2\sqrt{\pi n}} \left( \frac{3}{2} \right)^n 2^{n-\frac{1}{4}} e^{-\frac{1}{4n}}. \]

Since we can bound the exponential from below

\[ e^{\frac{1}{4n}} \geq 1 + \frac{1}{4n} \]

we have

\[ k^* \leq \frac{3}{2\sqrt{\pi n}} \frac{3^n - 2 \left( \frac{3}{2^n} \right)^n 2^{n-\frac{1}{4}} e^{-\frac{1}{4n}}}{4n + 1}. \]

We have now

\[ \frac{k^*}{\frac{3}{2\sqrt{\pi n}}} - B \leq -\frac{3^n}{2} \frac{1}{4n + 1} + 2^n - \left( \frac{3}{2^n} \right)^n 4n + 1 - \frac{1}{2} \quad (6) \]

We can prove numerically that the right-hand side is negative for \( 8 \leq n \leq 20 \), which proves the proposition for these values of \( n \).

| \( n \) | right-hand side of (6) |
|-------|-----------------------|
| 8     | -1.7112               |
| 9     | -53.7213              |
| 10    | -263.6607             |
| 11    | -994.8405             |
| 12    | -3.3609e+03           |
| 13    | -1.0699e+04           |
| 14    | -3.2857e+04           |
| 15    | -9.8629e+04           |
| 16    | -2.9167e+05           |
| 17    | -8.5405e+05           |
| 18    | -2.4847e+06           |
| 19    | -7.1993e+06           |
| 20    | -2.0808e+07           |

For \( n > 20 \), we prove that the right-hand side of

\[ \frac{k^*}{\frac{3}{2\sqrt{\pi n}}} - B \leq -\frac{3^n}{2} \frac{1}{4n + 1} + 2^n \]

is negative. Hence, we want to prove that

\[ \frac{3^n}{2} \frac{1}{4n + 1} \geq 2^n \]

\[ \left( \frac{3}{2} \right)^n \geq 2(4n + 1) \]

which is obvious for \( n > 20 \) and proves the proposition for these values of \( n \). As we proved it earlier for \( 8 \leq n \leq 20 \), the proposition is now entirely proven. \( \square \)