A constructive fixed point theorem for min-max functions*

Jean Cochet-Terrasson
INRIA, Domaine de Voluceau, 78153 Le Chesnay Cedex, France

Stéphane Gaubert
INRIA, Domaine de Voluceau, 78153 Le Chesnay Cedex, France
Stephane.Gaubert@inria.fr

Jeremy Gunawardena (corresponding author)
BRIMS, Hewlett-Packard Labs, Filton Road, Stoke Gifford, Bristol BS34 8QZ, UK
Tel : +44 117 312 8216
Fax : +44 117 312 9190
Email : jhcg@hplb.hpl.hp.com

July 15, 1999

Abstract

Min-max functions, $F : \mathbb{R}^n \to \mathbb{R}^n$, arise in modelling the dynamic behaviour of discrete event systems. They form a dense subset of those functions which are homogeneous, $F_i(x_1 + h, \cdots, x_n + h) = F_i(x_1, \cdots, x_n) + h$, monotonic, $\vec{x} \leq \vec{y} \Rightarrow F(\vec{x}) \leq F(\vec{y})$, and nonexpansive in the ℓ_∞ norm—so-called topical functions—which have appeared recently in the work of several authors. Our main result characterises those min-max functions which have a (generalised) fixed point, where $F_i(\vec{x}) = x_i + h$ for some $h \in \mathbb{R}$. We deduce several earlier fixed point results. The proof is inspired by Howard's policy improvement scheme in optimal control and yields an algorithm for finding a fixed point, which appears efficient in an important special case. An extended introduction sets the context for this paper in recent work on the dynamics of topical functions.

Keywords: cycle time, discrete event system, fixed point, max-plus semiring, nonexpansive map, nonlinear eigenvalue, Perron-Frobenius, policy improvement, topical function.

*Appears in *Dynamics and Stability of Systems*, Vol. 14 N°4, 1999, pp. 407-433.

1 Introduction

A min-max function $F : \mathbb{R}^n \to \mathbb{R}^n$ is built from terms of the form $x_i + a$, where $1 \leq i \leq n$ and $a \in \mathbb{R}$, by application of finitely many max and min operations in each component. For example,

$$\begin{aligned} F_1(x_1, x_2) &= \max(\min(\max(x_1 + 1, x_2 - 1.2), \max(x_1, x_2 + 2)), \min(x_1 + 0.5, x_2 + 1)) \\ F_2(x_1, x_2) &= \min(\max(x_1 + 7, x_2 + 4.3), \min(x_1 - 5, x_2 - 3)) \ . \end{aligned}$$

(A different notation is used in the body of the paper; see §1.1.) Such functions are homogeneous, $F_i(x_1 + h, \cdots, x_n + h) = F_i(x_1, \cdots, x_n) + h$, monotonic with respect to the usual product ordering on $\mathbb{R}^n$, $\vec{x} \leq \vec{y} \ \Rightarrow \ F(\vec{x}) \leq F(\vec{y})$, and nonexpansive in the ℓ_∞ norm, $\|F(\vec{x}) - F(\vec{y})\| \leq \|\vec{x} - \vec{y}\|$. Functions with these properties have emerged recently in the work of several authors, [2, 23, 28, 31, 40]. We shall follow Gunawardena and Keane and call them topical functions. They include (possibly after suitable transformation) nonnegative matrices, Leontieff substitution systems, dynamic programming operators of games and of Markov decision processes, nonlinear operators arising in matrix scaling problems and demographic modelling and renormalisation operators for certain fractal diffusions, [12, 24]. They also include examples, such as the min-max functions of the present paper, which arise from modelling discrete event systems—digital circuits, computer networks, manufacturing plants—an application discussed in more detail in §1.2.

Any topical function T can be approximated by min-max functions in such a way that some of the dynamical behaviour of T is inherited by its approximations (see Lemma 1.1). In this paper we study the dynamics of min-max functions, motivated partly by the applications to discrete event systems and partly by the need to develop a nonlinear Perron-Frobenius theory for topical functions. For any topical function, a fixed point—or nonlinear eigenvector—is a vector, $\vec{x} = (x_1, \cdots, x_n) \in \mathbb{R}^n$, for which there exists $h \in \mathbb{R}$, such that $F(\vec{x}) = (x_1 + h, \cdots, x_n + h)$. The cycle time vector, $\chi(F) = \lim_{k \to \infty} F^k(\vec{x})/k \in \mathbb{R}^n$, provides the appropriate nonlinear generalisation of the Perron root, or spectral radius (see §1.3). The limit in question does not always exist, [23, Theorem 3.1], and it remains an important open problem to characterise those topical functions for which it does. The nonexpansive property of F guarantees that χ, when it does exist, is independent of the initial condition, $\vec{x}$. The cycle time vector provides a performance measure in discrete event systems, where it measures the asymptotic average system latency (see §1.2).

Unlike the conventional spectral radius, the cycle time is a vector and immediately gives a necessary condition for the existence of a fixed point. If F has a fixed point, then the homogeneity property implies that $\chi(F)$ exists and that $\chi(F) = (h, \cdots, h)$: the cycle time has the same value in each component. (In the context of discrete event systems this represents the fact that the system can only possess an equilibrium state when the average latency of each event is asymptotically the same.) It is interesting to ask whether the converse is true. In other words, whether,

$$\exists \vec{x} \in \mathbb{R}^n, \text{ such that } F(\vec{x}) = (x_1 + h, \cdots, x_n + h) \ \text{ if, and only if, } \ \chi(F) = (h, \cdots, h) \ . \quad (1)$$

In 1994 the third author put forward a conjecture—the Duality Conjecture—on the existence of the cycle time for min-max functions, [19]. The conjecture asserts that any min-max function has a cycle time and that χ, when considered as a functional from min-max functions to $\mathbb{R}^n$, is "almost" a homomorphism of lattices (see [17, §3] for a more precise statement). The conjecture was shown to imply the fixed point result (1) but the method of proof was nonconstructive and gave no algorithm for finding a fixed point, [19]. The question of constructibility is an important one in the applications; for instance, in the study of digital circuits, the solution of the clock schedule verification problem requires calculating the fixed point of a min-max function, [18].

The main result of the present paper is a constructive fixed point theorem for min-max functions which is independent of the Duality Conjecture and of the existence of cycle times for min-max functions. We give a necessary and sufficient condition, in terms of the cycle times of the component "max-only" functions (see below), for any min-max function to have a fixed point. The algorithm based on the proof is not tractable in general but can be made efficient in an important special case. We recover as corollaries of our main theorem the two previous fixed point results for min-max functions, one due to the third author with more restrictive hypotheses, [19, Theorem 3.1], the other due to Olsder, applicable to separated min-max functions (see below) satisfying certain conditions, [33, Theorem 2.1]. Both earlier results were nonconstructive.

The methods of the present paper rely on a special class of min-max functions which can be studied by linear methods, albeit of an unusual nature. This is the class of matrices over the max-plus semiring, $\mathbb{R}_{\max} = \mathbb{R} \cup \{-\infty\}$, where addition and multiplication are defined as max and $+$, respectively, the latter being distributive over the former (see §1.4). Matrices over $\mathbb{R}_{\max}$ correspond to min-max functions in which the min operation is never used, so-called "max-only" functions. Matrix algebra over $\mathbb{R}_{\max}$ has been extensively studied, [1, 9, 10, 29, 32, 42]. Min-max functions can be represented by finite collections of max-plus matrices and the dynamical properties of the latter, known from the linear theory over $\mathbb{R}_{\max}$, can be used to infer those of the former.

After the final draft of this paper was finished, the second and third authors, inspired by the methods used here, proved the Duality Theorem for min-max functions, [13, Theorem 1], and then went on to show the existence of the cycle time for a larger class of topical functions including some arising in Markov decision theory, [12, Theorem 15]. The proofs require some of the technical results of §1.4 but also introduce a number of new ideas. Fixed point results are not elaborated upon in this later work and the results and methods of the present paper remain relevant.

Interest in the class of min-max functions, and in the broader class of topical functions, has come from a number of directions and brings together a number of distinct themes: discrete event systems, Perron-Frobenius theory, max-plus matrix algebra, fixed point theorems for nonexpansive functions, nonlinear dynamics, etc. Because of the recent emergence of this area, we have devoted the remainder of this Introduction to amplifying the outline above. We hope this will give a better sense of the scope of the present work.

Special cases of min-max functions were studied by Olsder in [33]. Min-max functions themselves were introduced in [21]. The present paper incorporates some of the results of [6, 19, 20], as well as new material. The authors gratefully acknowledge discussions with Michael Keane, Roger Nussbaum, Geert-Jan Olsder, Jean-Pierre Quadrat, Colin Sparrow and Sjoerd Verduyn Lunel. They are also grateful to the reviewer and editors for their helpful comments. This work was partially supported by the European Community Frame-

work IV programme through the research network ALAPEDES ("The Algebraic Approach to Performance Evaluation of Discrete Event Systems").

1.1 Min-max functions

We begin with some notation. Vectors in $\mathbb{R}^n$ will be denoted $\vec{x}, \vec{a}, etc.$ For vector valued quantities in general, such as functions $F : X \to \mathbb{R}^n$, the notation F_i will denote component i: $F(x) = (F_1(x), \cdots, F_n(x))$. To avoid clutter, we use x_i for the components of $\vec{x}$. The partial order on $\mathbb{R}$ will be denoted in the usual way by $a \leq b$ but it will be convenient to use infix forms for the lattice operations of least upper bound and greatest lower bound:

$$\begin{aligned} a \vee b &= \mathrm{lub}(a, b) \\ a \wedge b &= \mathrm{glb}(a, b)\,. \end{aligned}$$

(The word "lattice" is used in this paper to refer to a partial order in which any two elements have a least upper bound and a greatest lower bound, [26, §1.1]. We do not require, however, that a lattice has a greatest and a least element.) The same notation will be used for lattices derived from $\mathbb{R}$, such as the function space $X \to \mathbb{R}$. The partial order here is the pointwise ordering on functions: $f \leq g$ if, and only if, $f(x) \leq g(x)$ for all $x \in X$. If $\mathbb{R}^n$ is identified with the set of functions $\{1, \cdots, n\} \to \mathbb{R}$, then this specialises to the product ordering on vectors.

To reduce notational overhead we shall use the following vector-scalar convention: if, in a binary operation or relation, a vector and a scalar are mixed, the relevant operation is performed, or the relevant relation is required to hold, on each component of the vector. For instance, if $h \in \mathbb{R}$ and $\vec{x} \in \mathbb{R}^n$, then $\vec{x} + h$ will denote the vector $(x_1 + h, \cdots, x_n + h)$, and $\vec{x} \leq h$ will imply $x_i \leq h$ for each $1 \leq i \leq n$. Throughout this paper, we shall use h to denote a real number without specifying so explicitly. Formulae such as $\vec{x} = h$ should therefore always be interpreted using the vector-scalar convention: $x_i = h$ for each $1 \leq i \leq n$.

The notation $\|\vec{x}\|$ will denote the ℓ_∞ norm on $\mathbb{R}^n$: $\|\vec{x}\| = |x_1| \vee \cdots \vee |x_n|$. If $F, G : X \to X$, then function composition will be denoted, as usual, by FG: $FG(x) = F(G(x))$.

Definition 1.1 *A min-max function of type $(n, 1)$ is any function $f : \mathbb{R}^n \to \mathbb{R}^1$, which can be written as a term in the following grammar:*

$$f := x_1, x_2, \cdots, x_n \mid f + a \mid f \wedge f \mid f \vee f \qquad (a \in \mathbb{R}). \tag{2}$$

The notation used here is the Backus-Naur form familiar in computer science. The vertical bars separate the different ways in which terms can be recursively constructed. The simplest term is one of the n variables, x_i, thought of as the i-th component function. Given any term, a new one may be constructed by adding $a \in \mathbb{R}$; given two terms, a new one may be constructed by taking a greatest lower bound or a least upper bound. Only these rules may be used to build terms. Of the three terms

$$\begin{gathered} ((((x_1 + 2) \vee (x_2 - 0.2)) \wedge x_3) \vee (x_2 + 3.5)) - 1 \\ x_1 \vee 2 \\ (x_1 + x_2) \wedge (x_3 + 1) \end{gathered}$$

the first is a min-max function but neither the second nor the third can be generated by (2).

We shall assume that $+$ has higher precedence than $\vee$ or $\wedge$, allowing us to write the first example more simply:

$$(((x_1 + 2 \vee x_2 - 0.2) \wedge x_3) \vee x_2 + 3.5) - 1\,.$$

Although the grammar provides a convenient syntax for writing terms, we are interested in them only as functions, $\mathbb{R}^n \to \mathbb{R}$. Terms can therefore be rearranged using the associativity and distributivity of the lattice operations, as well as the fact that addition distributes over both $\wedge$ and $\vee$. The example above can hence be simplified further to

$$(x_1 + 1 \vee x_2 + 2.5) \wedge (x_3 - 1 \vee x_2 + 2.5)\,.$$

It is clear that any term can be reduced in a similar way to a minima of maxima, or, dually, a maxima of minima. We shall discuss the corresponding canonical forms in §2.1.

Definition 1.2 ([21, Definition 2.3]) *A min-max function of type (n,m) is any function $F : \mathbb{R}^n \to \mathbb{R}^m$, such that each component F_i is a min-max function of type $(n,1)$.*

The set of min-max functions of type (n,m) will be denoted $\mathsf{MM}(n,m)$. We shall mostly be concerned with functions of type (n,n), which we refer to as functions of dimension n. It is convenient to single out some special cases. Let $f \in \mathsf{MM}(n,1)$. If f can be represented by a term that does not use $\wedge$, it is said to be max-only. If f can be represented by a term that does not use $\vee$, it is min-only. If f is both max-only and min-only, it is simple. The same terminology extends to functions $F \in \mathsf{MM}(n,m)$ by requiring that each component F_i has the property in question. If $F \in \mathsf{MM}(n,m)$ and each F_i is either max-only or min-only, F is said to be separated. Of the following functions in $\mathsf{MM}(2,2)$,

$$S = \begin{pmatrix} x_1 + 1 \\ x_2 - 1 \end{pmatrix} \qquad T = \begin{pmatrix} x_2 + 1 \\ x_2 - 1 \end{pmatrix} \qquad U = \begin{pmatrix} x_1 + 1 \vee x_2 + 1 \\ x_1 + 2 \end{pmatrix},$$

S and T are both simple and U is max-only. Moreover, $S \wedge T$ is min-only and $(S \vee T) \wedge U$ is separated.

Proposition 1.1 *Let $F, G \in \mathsf{MM}(n,n)$ and $\vec{a}, \vec{x}, \vec{y} \in \mathbb{R}^n$ and $h \in \mathbb{R}$. The following hold.*

1. *$F + \vec{a}$, FG, $F \vee G$ and $F \wedge G$ all lie in $\mathsf{MM}(n,n)$.*
2. Homogeneity: *$F(\vec{x} + h) = F(\vec{x}) + h$.* H
3. Monotonicity: *if $\vec{x} \leq \vec{y}$ then $F(\vec{x}) \leq F(\vec{y})$.* M
4. Nonexpansiveness: *$\|F(\vec{x}) - F(\vec{y})\| \leq \|\vec{x} - \vec{y}\|$.* N

The first three parts follow easily from Definition 1.2, while the fourth is a consequence of the following observation of Crandall and Tartar, [8]. (See also [23, Proposition 1.1] for a proof adapted to the present context.)

Proposition 1.2 *If $F : \mathbb{R}^n \to \mathbb{R}^n$ satisfies* H *then* M *is equivalent to* N.

The homogeneity property suggests a generalisation of the conventional notion of fixed, or periodic, point.

Definition 1.3 *Suppose that $F : \mathbb{R}^n \to \mathbb{R}^n$ satisfies property* H. *We say that $\vec{x} \in \mathbb{R}^n$ is a fixed point of F, if $F(\vec{x}) = \vec{x} + h$ for some $h \in \mathbb{R}$, and that $\vec{x}$ is a periodic point of F with period p, if $\vec{x}$ is a fixed point of F^p, but not of F^k for any $0 < k < p$.*

A fixed point of F in this sense is a fixed point of $F - h$ in the conventional sense. Unless otherwise stated, the phrases "fixed point" and "periodic point" will have the meaning given by Definition 1.3 throughout this paper.

Min-max functions first arose in applications and these applications continue to provide important insights. In the next two sub-sections we review this material.

1.2 Discrete event systems

A discrete event system is, roughly speaking, a system comprising a finite set of events which occur repeatedly: a digital circuit, in which an event might be a voltage change on a wire, from binary 1 to 0 or vice versa; a distributed computer system, in which an event might be the arrival of a message packet at a computer; an automated manufacturing plant, in which an event might be the completion of a job on a machine. Discrete event systems are ubiquitous in modern life and are the focus of much interest in engineering circles, [1, 7, 15, 25]. They are dynamical systems, in the sense that they evolve in time, but their analysis leads to quite different mathematics to that used to model dynamic behaviour in continuous and differentiable systems.

If n is the number of events in the system, let $\vec{x} \in \mathbb{R}^n$ be such that x_i is the time of first occurrence of event i, relative to some arbitrary origin of time when the system is started. (It is worth noting that the existence of such a global clock is not always a reasonable assumption: distributed systems sometimes operate on the basis of only local time references.) Suppose further that the system can be modelled in such a way that, for some function $F : \mathbb{R}^n \to \mathbb{R}^n$, $F_i(\vec{x})$ gives the time of next occurrence of event i. In this case the dynamic behaviour of the system can be modelled by the discrete dynamic system F.

It might be thought that a model of this kind is too simplified to occur in practice. This is not the case. For instance, the problem of clock schedule verification in digital circuits leads directly to such a model. To each "latched synchronous" circuit may be associated an element $F \in \mathsf{MM}(n, n)$, where n is one more than the number of storage latches. (Practical circuits may have as many as 10^4 latches.) The clock schedule verification problem for the circuit may be solved by finding a fixed point of F. We shall not discuss this application further here; the reader should consult [18, 37, 39] for more details. It does underline, however, the importance of understanding when fixed points of min-max functions exist and how to calculate them when they do. These are the main concerns of the present paper.

The axioms of homogeneity and monotonicity have an appealing interpretation in the context introduced above. Homogeneity is tantamount to saying that the origin of the global time reference is irrelevant. Monotonicity asserts that delaying some events cannot speed-up any part of the system. This latter condition is intuitively reasonable and is often observed in practice but the complexity of the real world does throw up examples in which it fails. It is natural, in the light of these observations, to focus attention on functions satisfying these two properties. By virtue of Proposition 1.2, these are necessarily nonexpansive in the ℓ_∞ norm.

Definition 1.4 ([23, Definition 1.1]) *A topical function is any function $F : \mathbb{R}^n \to \mathbb{R}^n$ satisfying properties* H *and* M.

Topical functions have appeared in the work of a number of authors, [2, 23, 28, 31, 40]. Current work suggests that they provide a mathematically appealing model for discrete event systems, [17, 2]. For serious applications in this area, the dynamics of a single function must be extended in several ways. First, by considering the semigroup generated by a set of functions, $\{F(\alpha) \mid \alpha \in A\}$, [11, 38] as in the theory of automata, [34]. This allows for the possibility of nondeterminism: if the system is in state $\vec{x}$, it may evolve to any of the states $F(\alpha)(\vec{x})$. For instance, demanding £20 from an automatic cash machine may sometimes result in two ten pound notes and sometimes in one ten and two fives. A second extension comes by taking $F(\alpha)$ to be a random variable from some suitable measure space into the space of topical functions. This permits stochastic behaviour to be modelled: in a digital circuit it is conventional to consider only the maximum or minimum delays through a component (the manufacturer provides a data book which lists these values) but in a distributed computer system the time taken by a message packet will vary widely and a probabilistic approach is more appropriate, [1, Chapter 7], [36]. Notwithstanding these extensions, the dynamical behaviour of a single topical function remains largely unknown and leads to a number of open problems, [24].

What role do min-max functions play within the larger class of topical functions? In turns out to be an unexpectedly central one, as shown by the following observation of Gunawardena, Keane and Sparrow.

Lemma 1.1 ([24]) *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a topical function and let $S \subseteq \mathbb{R}^n$ be any finite set of points. There exists $H \in \mathsf{MM}(n, n)$ such that $T \leq H$ and $T(\vec{u}) = H(\vec{u})$ for all $\vec{u} \in S$.*

It follows that min-max functions approximate topical functions, not only in the topological sense that $\mathsf{MM}(n, n)$ is dense in the set of topical functions (in the compact-open topology), but also in a lattice theoretic sense: any topical function is the lower envelope of a family of min-max functions. More importantly, this approximation preserves some aspects of the dynamics. Using the notation of Lemma 1.1, it follows from property M that $T^k \leq H^k$. In particular, the cycle time vector of T will be bounded above by that of H (provided both exist). It also follows from Lemma 1.1 that every periodic orbit of a topical function is the orbit of some min-max function. Lemma 1.1 and its consequences provide one of the principal motivations for the present paper: to study min-max functions as a foundation for analysing topical functions.

We have presented topical functions as arising naturally from attempts to find a mathematical model for discrete event systems. However, they also have intrinsic mathematical interest because they include a number of classical examples which have been extensively studied in quite different contexts.

1.3 Topical functions and cycle times

Let $\mathbb{R}^+$ denote the positive reals: $\mathbb{R}^+ = \{x \in \mathbb{R} \mid x > 0\}$. The whole space, $\mathbb{R}^n$, can be put into bijective correspondence with the positive cone, $(\mathbb{R}^+)^n$, via the mutually inverse functions $\exp : \mathbb{R}^n \to (\mathbb{R}^+)^n$ and $\log : (\mathbb{R}^+)^n \to \mathbb{R}^n$, which do exp and log on each component: $\exp(\vec{x})_i = \exp(x_i)$, for $\vec{x} \in \mathbb{R}^n$, and $\log(\vec{x})_i = \log(x_i)$, for $\vec{x} \in (\mathbb{R}^+)^n$. Let A be any $n \times n$ matrix all of whose entries are nonnegative. Elements of $\mathbb{R}^n$ can be thought of as column vectors and A acts on them on the left as $A\vec{x}$. We further suppose the nondegeneracy condition that no row of A is zero:

$$\forall 1 \leq i \leq n, \ \ \exists 1 \leq j \leq n, \ \text{ such that } \ A_{ij} \neq 0\,. \tag{3}$$

In this case, A maps the positive cone onto itself, $A : (\mathbb{R}^+)^n \to (\mathbb{R}^+)^n$. Let $\mathcal{E}(A) : \mathbb{R}^n \to \mathbb{R}^n$ denote the conjugate $\mathcal{E}(A)(\vec{x}) = \log(A(\exp(\vec{x})))$. Clearly, $\mathcal{E}(AB) = \mathcal{E}(A)\mathcal{E}(B)$, so that the dynamics of A and $\mathcal{E}(A)$ are entirely equivalent.

The point of this is that $\mathcal{E}(A)$ is always a topical function: property H is the additive equivalent of the fact that A commutes with scalar multiplication, while property M follows from the nonnegativity of A. We see that the dynamics of topical functions includes as a special case that of nonnegative matrices; in other words, Perron-Frobenius theory. It can be shown that a number of classical examples in optimal control, game theory and mathematical economics also give rise to topical functions. The geography of the space of topical functions is discussed in more detail in [12, 24].

If $\vec{x} \in \mathbb{R}^n$ is a fixed point of $\mathcal{E}(A)$, so that $\mathcal{E}(A)(\vec{x}) = \vec{x} + h$, then $\exp(\vec{x})$ is an eigenvector of A with eigenvalue $\exp(h)$. Fixed points of $\mathcal{E}(A)$ therefore correspond bijectively to positive eigenvectors of A. That is, to eigenvectors lying in the positive cone. What about the eigenvalue? Can this also be generalised to the nonlinear context? A clue to doing this came from the applications.

A frequent demand from system designers is to estimate performance, [3, 18]. If the system can be modelled by a single function, $F : \mathbb{R}^n \to \mathbb{R}^n$, as described above, an estimate can be made on the basis of the time elapsed between successive occurrences: $F(\vec{x}) - \vec{x}$. Better still is an average over several occurrences:

$$(F^k(\vec{x}) - F^{k-1}(\vec{x}) + \cdots + F(\vec{x}) - \vec{x})/k\,.$$

Letting $k \to \infty$, we get $\lim_{k\to\infty} F^k(\vec{x})/k$. This is a vector quantity, which measures the asymptotic average latency in each component. Does this limit exist?

Lemma 1.2 *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ satisfy property N. If $\lim_{k\to\infty} F^k(\vec{x})/k$ exists somewhere, then it exists everywhere and has the same value.*

Proof Suppose that, for some $\vec{x} \in \mathbb{R}^n$, $\lim_{k\to\infty} F^k(\vec{x})/k = \vec{a}$ and let $\vec{y}$ be another point of $\mathbb{R}^n$. Choose $\epsilon > 0$. By property N, for all sufficiently large k,

$$\|\vec{a} - F^k(\vec{y})/k\| \leq \|\vec{a} - F^k(\vec{x})/k\| + \|F^k(\vec{x})/k - F^k(\vec{y})/k\| \leq \epsilon + \|\vec{x} - \vec{y}\|/k\,.$$

From which the result follows immediately.

$\square$

Definition 1.5 *Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ satisfy property* N. *The cycle time vector of F, denoted $\chi(F) \in \mathbb{R}^n$, is defined to be*

$$\lim_{k\rightarrow\infty} F^k(\vec{x})/k \tag{4}$$

when this limit exists for some $\vec{x} \in \mathbb{R}^n$, and to be undefined otherwise.

Suppose that $F = \mathcal{E}(A)$ and that A has a positive eigenvector with eigenvalue λ, which must necessarily be real. (Perron-Frobenius theory tells us that λ equals the spectral radius of A.) It then follows, as above, that $\chi(\mathcal{E}(A)) = \log(\lambda)$. We see from this that χ is a nonlinear vector generalisation of the spectral radius. It can be shown that if A is any nonnegative matrix satisfying (3) then χ exists and can be determined in terms of the spectral radii of the irreducible components of A, [24]. Indeed, by using $\mathcal{E}^{-1}$ on (4), we see that the cycle time vector corresponds to the usual spectral radius formula, albeit disintegrated into individual components.

The cycle time vector immediately yields a necessary condition for a fixed point. Suppose that F is a topical function with a fixed point, so that $F(\vec{x}) = \vec{x}+h$. By repeated application of H, we see that $F^k(\vec{x}) = \vec{x} + k.h$. Hence, χ does exist and $\chi(F) = h$. We recall from the vector-scalar convention that this means each component of $\chi(F)$ has the same value h, which, by Lemma 1.2, is independent of the choice of fixed point. As discussed in the Introduction, the converse result, (1), also holds when F is a min-max function, [12, 13].

Not all topical functions have cycle times, [23, Theorem 3.1]. It remains an important open problem to characterise those that do, [24]. Furthermore, one cannot expect such a strong fixed point theorem as (1) even for those topical functions that do have cycle times. If

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

then $\chi(\mathcal{E}(A)) = (0, 0)$ but A does not have a positive eigenvector. This example raises a number of issues which take us beyond the scope of the present paper and we defer further discussion to later work.

1.4 Max-only functions and max-plus matrices

It should be clear from the remarks before Definition 1.2 that if $f \in \mathsf{MM}(n, 1)$ is max-only, it can be reduced to the form

$$f = x_1 + a_1 \vee \cdots \vee x_n + a_n$$

where the absence of a term $x_j + a_j$ is indicated by setting $a_j = -\infty$. This can be thought of as an element adjoined to $\mathbb{R}$ which is less than any real number and acts as an absorbing element for addition: $a + (-\infty) = -\infty$. Hence, if $F \in \mathsf{MM}(n, n)$ is max-only, it can be represented by a matrix A with entries in $\mathbb{R} \cup \{-\infty\}$:

$$\begin{array}{rcl} F_1 & = & x_1 + A_{11} \vee \cdots \vee x_n + A_{1n} \\ \vdots & \vdots & \vdots \\ F_n & = & x_1 + A_{n1} \vee \cdots \vee x_n + A_{nn}\,. \end{array} \tag{5}$$

Since each component of F must have some value, A satisfies a nondegeneracy property formally similar to (3):

$$\forall 1 \leq i \leq n, \ \ \exists 1 \leq j \leq n, \ \text{ such that } \ A_{ij} \neq -\infty \,. \tag{6}$$

Suppose that the algebraic operations on $\mathbb{R} \cup \{-\infty\}$ are now redefined so that the sum operation becomes maximum and the multiplication operation becomes addition. The element $-\infty$ then becomes a zero for sum, while 0 becomes a unit for multiplication. Since addition distributes over maximum, this structure forms a semiring, called the max-plus semiring, and denoted $\mathbb{R}_{\max}$. If vectors in $\mathbb{R}^n$ are thought of as column vectors, (5) can be rewritten as a matrix equation

$$F(\vec{x}) = A\vec{x} \tag{7}$$

in which the matrix operations are interpreted in $\mathbb{R}_{\max}$. It follows that $F^k(\vec{x}) = A^k\vec{x}$ and the dynamics of F reduce to matrix algebra, albeit of an unusual sort. We have, in effect, linearised an apparently nonlinear problem.

Cuninghame-Green was perhaps the first to realise the implications of matrix algebra over max-plus, [9]. Since that time the idea has been rediscovered and redeveloped several times and there are now several standard texts on the subject, [1, 4, 10, 29, 42]. For recent overviews, see [14, 17].

In this paper we shall not adopt max-plus notation. That is, $+$ and $\times$ will always have their customary meanings. We shall use $\vee$ and $+$ for the corresponding max-plus operations. Similarly, 0 will always have its customary meaning and we shall use $-\infty$ for the zero in $\mathbb{R}_{\max}$. If A and B are, respectively, $n \times p$ and $p \times m$ matrices over $\mathbb{R}_{\max}$, then AB will always mean the matrix product over $\mathbb{R}_{\max}$:

$$(AB)_{ij} = \bigvee_{1 \leq k \leq p} A_{ik} + A_{kj} \ \ .$$

(Recall that $+$ has higher precedence than $\vee$.) The customary ordering on $\mathbb{R}$ extends to $\mathbb{R}_{\max}$ in the obvious way, so that $-\infty \leq x$ for all $x \in \mathbb{R}_{\max}$. The same symbol is used for the product ordering on vectors: if $\vec{x}, \vec{y} \in (\mathbb{R}_{\max})^n$ then $\vec{x} \leq \vec{y}$ if, and only if, $x_i \leq y_i$ for all i. An $n \times n$ matrix over $\mathbb{R}_{\max}$, A, acts on the whole space $(\mathbb{R}_{\max})^n$ and it is easy to see that it is monotonic with respect to the product ordering: if $\vec{x} \leq \vec{y}$ then $A\vec{x} \leq A\vec{y}$. We recall that $\vec{x} \in (\mathbb{R}_{\max})^n$ is an eigenvector of A for the eigenvalue $h \in \mathbb{R}_{\max}$, if $A\vec{x} = \vec{x} + h$. If A satisfies the nondegeneracy condition (6), so that A can also be considered as a min-max function, then fixed points of A correspond bijectively to eigenvectors of A lying in $\mathbb{R}^n$. (This restriction is formally similar to that needed for nonnegative matrices and their eigenvectors in §1.3.) In this paper, the word "eigenvector" will indicate an element of $(\mathbb{R}_{\max})^n$ while the phrase "fixed point" will imply that the element in question lies in $\mathbb{R}^n$.

We need to recall various standard results in max-plus theory. The reader seeking more background should consult [1, Chapter 3].

Let A be an $n \times n$ matrix over $\mathbb{R}_{\max}$. The precedence graph of A, denoted $\mathcal{G}(A)$, is the directed graph with labelled edges which has nodes $\{1, \cdots, n\}$ and an edge from j to i if, and only if, $A_{ij} \neq -\infty$. The label on this edge is then the real number A_{ij}. (Some authors use the opposite convention for the direction of edges.) We shall denote an edge from j to i by $i \leftarrow j$. A path from i_m to i_1 is a sequence of nodes $i_1, \cdots, i_m$ such that $1 < m$ and

$i_j \leftarrow i_{j+1}$ for $1 \leq j < m$. A circuit is a path which starts and ends at the same node: $i_1 = i_m$. A circuit is elementary if the nodes $i_1, \cdots, i_{m-1}$ are all distinct. A node j is upstream from i, denoted $i \Leftarrow j$, if either $i = j$ or there is a path in $\mathcal{G}(A)$ from j to i. (A node is always upstream from itself.) A circuit g is upstream from node i, denoted $i \Leftarrow g$, if some node on the circuit is upstream from i. The weight of a path p, $|p|_{\mathsf{w}}$, is the sum of the labels on the edges in the path:

$$|p|_{\mathsf{w}} = \sum_{j=1}^{m-1} A_{i_j i_{j+1}} .$$

It follows from this that matrix multiplication has a nice interpretation in terms of path weights: A^s_{ij} is the maximum weight among all paths of length s from j to i. The length of a path, $|p|_\ell$, is the number of edges in the path: $|p|_\ell = m - 1$. If g is a circuit, its cycle mean, denoted $\mathsf{m}(g)$ is defined by $\mathsf{m}(g) = |g|_{\mathsf{w}}/|g|_\ell$. If A is an $n \times n$ matrix over $\mathbb{R}_{\max}$, let $\mu(A) \in (\mathbb{R}_{\max})^n$ be defined by

$$\mu_i(A) = \max\{\mathsf{m}(g) \mid i \Leftarrow g\} . \tag{8}$$

This is well defined: although there may be infinitely many circuits in $\mathcal{G}(A)$, only the elementary ones are needed to determine $\mu(A)$. By convention, the maximum of an empty set is taken to be $-\infty$. Hence, if there are no circuits upstream from node i, $\mu_i(A) = -\infty$. If A satisfies the nondegeneracy condition (6) then every node has an upstream circuit and so $\mu(A) \in \mathbb{R}^n$.

It is convenient at this point to single out the functions $\mathsf{t}, \mathsf{b} : \mathbb{R}^n \to \mathbb{R}$ given by

$$\begin{array}{rcl} \mathsf{t}(\vec{x}) & = & x_1 \vee \cdots \vee x_n \\ \mathsf{b}(\vec{x}) & = & x_1 \wedge \cdots \wedge x_n . \end{array}$$

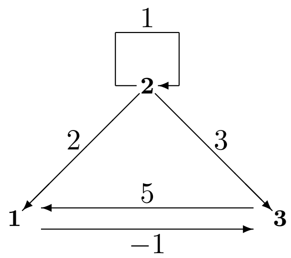
If c is any vector valued quantity, we shall often simplify this notation by writing $\mathsf{t}c$ and $\mathsf{b}c$ in place of $\mathsf{t}(c)$ and $\mathsf{b}(c)$, respectively. It follows from (8) that $\mathsf{t}\mu(A)$ is the maximum cycle mean over all circuits. A critical circuit is an elementary circuit with cycle mean $\mathsf{t}\mu(A)$.

Before proceeding further, it may be helpful to see an example. The max-only function

$$\begin{array}{rcl} F_1(x_1, x_2, x_3) & = & x_2 + 2 \vee x_3 + 5 \\ F_2(x_1, x_2, x_3) & = & x_2 + 1 \\ F_3(x_1, x_2, x_3) & = & x_1 - 1 \vee x_2 + 3 \end{array} \tag{9}$$

has associated max-plus matrix, precedence graph and μ vector shown below

$$\begin{pmatrix} -\infty & 2 & 5 \\ -\infty & 1 & -\infty \\ -1 & 3 & -\infty \end{pmatrix}$$



$$\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} . \tag{10}$$

The maximum cycle mean is 2 and $1 \leftarrow 3 \leftarrow 1$ is the unique critical circuit.

Proposition 1.3 *If $F \in \mathsf{MM}(n,n)$ is max-only and A is the associated matrix over $\mathbb{R}_{\max}$, then $\chi(F)$ exists and $\chi(F) = \mu(A)$.*

Proof Let $\mu_1(A) = h$. Suppose initially that $h = 0$ and consider the sequence of numbers $\alpha(s) = (A^s\vec{0})_1$. It follows from one of the remarks above that we can interpret $\alpha(s)$ as the maximum weight among paths in $\mathcal{G}(A)$ of length s which terminate at node 1. If we consider any path terminating at node 1 then the only positive contribution to the weight of the path can come from those edges which are not repeated on the path: a repeated edge would be contained in a circuit, whose contribution to the path weight is at most 0. Since there are only finitely many edges, the weight of any path must be bounded above by $\sum_{A_{ij}>0} A_{ij}$. Hence $\alpha(s)$ is bounded above. Since $h = 0$, we know that there is some circuit upstream from node 1 whose weight is 0. Call this circuit g. For s sufficiently large, we can construct a path, $p(s)$, terminating at node 1 whose starting point cycles round the circuit g. The weight of this path can only assume a finite set of values because $|g|_{\mathsf{w}} = 0$. Since $\alpha(s)$ is the path of maximum weight of length s, it follows that $\alpha(s) \geq |p(s)|_{\mathsf{w}}$ and so $\alpha(s)$ is also bounded below. We have shown that there exist $m, M \in \mathbb{R}$ such that, for all $s \geq 0$, $m \leq \alpha(s) \leq M$. It follows immediately that $\lim_{s\to\infty} \alpha(s)/s = 0$. Hence,

$$\lim_{s\to\infty} (F^s(\vec{0}))_1/s = \mu_1(A).$$

If $h \neq 0$ then replace F by $G = F - h$. G is also a max-only function and if B is its associated matrix, then $B_{ij} = A_{ij} - h$. Hence $\mu_1(B) = 0$ and we can apply the argument above to show that $\lim_{s\to\infty}(G^s(\vec{0}))_1/s = 0$. But since $F = G + h$, it follows from property H that $\lim_{s\to\infty}(F^s(\vec{0}))_1/s = h = \mu_1(A)$. The same argument can be applied to any component of F and the result follows.

□

If F has a fixed point, so that $F(\vec{x}) = \vec{x} + h$, then $h = \mu(A)$. In particular, $h = \mathsf{t}\mu(A)$, the maximum cycle mean over all circuits in $\mathcal{G}(A)$. This is the eigenvalue associated to any eigenvector of A lying in $\mathbb{R}^n$. It is the analogue for max-plus matrices of the Perron root, or spectral radius, for nonnegative matrices, [1, Theorem 3.23].

Suppose that A is an $n \times n$ matrix over $\mathbb{R}_{\max}$. Suppose further that $\mathsf{t}\mu(A) = 0$, so that all circuits of $\mathcal{G}(A)$ have nonpositive weight. Since any path p in $\mathcal{G}(A)$, with $|p|_\ell \geq n$, must contain a circuit, it is not difficult to see that

$$(A^s)_{ij} \leq A_{ij} \vee \cdots \vee (A^n)_{ij} \tag{11}$$

for all $s \geq n$. Let $(A^+)_{ij} = \sup\{(A^s)_{ij} \mid 1 \leq s\}$, which is well defined as an element of $\mathbb{R}_{\max}$ by the previous observation. (It is well-known in max-plus theory that, $A^+ = A \vee \cdots \vee A^n$, [1, Theorem 3.20], but we shall not need this here.) Note that it is still possible for $(A^+)_{ij} = -\infty$, since it may be the case that there are no paths from j to i. Let $\mathcal{C}(A) \subseteq \{1, \cdots, n\}$ be the set of those nodes of $\mathcal{G}(A)$ which lie on some critical circuit. Let P_A be the $\mathbb{R}_{\max}$ matrix defined as follows:

$$(\mathsf{P}_A)_{ij} = \bigvee_{u\in\mathcal{C}(A)} (A^+)_{iu} + (A^+)_{uj} \ . \tag{12}$$

P_A is sometimes called a spectral projector, [1, §3.7.3]. Part 3 of Lemma 1.4 will show that it encapsulates information about the eigenvectors of A.

The next lemma is a standard result in max-plus matrix theory.

Lemma 1.3 ([1, Theorem 3.105]) *Suppose that A is an $n \times n$ matrix over $\mathbb{R}_{\max}$ such that $\mathsf{t}\mu(A) = 0$. Then $(\mathsf{P}_A)A = A(\mathsf{P}_A) = \mathsf{P}_A$ and $(\mathsf{P}_A)^2 = \mathsf{P}_A$.*

The next lemma collects together a number of useful observations. Some of them are well-known in max-plus theory, [1, Chapter 3], but do not appear in a convenient form in the literature.

Lemma 1.4 *Suppose that A is an $n \times n$ matrix over $\mathbb{R}_{\max}$ such that $\mathsf{t}\mu(A) = 0$. Suppose further that $\vec{x}, \vec{y} \in (\mathbb{R}_{\max})^n$. The following statements hold.*

1. *If $A\vec{x} \le \vec{x}$ then $\mathsf{P}_A\vec{x} \le \vec{x}$.*
2. *$A\vec{x} = \vec{x}$ if, and only if, $\mathsf{P}_A\vec{x} = \vec{x}$.*
3. *The image of $\mathsf{P}_A : (\mathbb{R}_{\max})^n \to (\mathbb{R}_{\max})^n$ is the eigenspace of A for the eigenvalue 0.*
4. *If $A\vec{x} = \vec{x}$, $A\vec{y} = \vec{y}$ and $x_i = y_i$ for all $i \in \mathcal{C}(A)$, then $\vec{x} = \vec{y}$.*
5. *If $i \in \mathcal{C}(A)$ then $(\mathsf{P}_A)_{ii} = 0$.*
6. *If $A\vec{x} \le \vec{x}$ then $(A\vec{x})_i = x_i$ for all $i \in \mathcal{C}(A)$.*
7. *If $\mu(A) = 0$ then $\mathsf{P}_A : \mathbb{R}^n \to \mathbb{R}^n$.*

Proof 1. Since $A\vec{x} \le \vec{x}$, it follows that $A^s\vec{x} \le \vec{x}$ and so $A^+\vec{x} \le \vec{x}$. Hence $(A^+)_{uj} + x_j \le (A^+\vec{x})_u \le x_u$. Choose $1 \le i \le n$. Then, by (12), $(\mathsf{P}_A\vec{x})_i \le \bigvee_{u \in \mathcal{C}(A)} (A^+)_{iu} + x_u \le (A^+\vec{x})_i \le x_i$. Hence $\mathsf{P}_A\vec{x} \le \vec{x}$ as required.

2. If $\mathsf{P}_A\vec{x} = \vec{x}$ then it follows immediately from Lemma 1.3 that $A\vec{x} = \vec{x}$. So suppose that $A\vec{x} = \vec{x}$. By part 1, $\mathsf{P}_A\vec{x} \le \vec{x}$. Choose $1 \le i \le n$. If $x_i = -\infty$, then certainly $(\mathsf{P}_A\vec{x})_i = x_i$, so we may assume that $x_i > -\infty$. For each s, $\vec{x} = A^s\vec{x}$. Hence there exists $1 \le j \le n$ such that $x_i = (A^s)_{ij} + x_j$. $(A^s)_{ij}$ is the weight of some path of length s from node j to node i. If we choose $s = n$, then this path must contain a circuit g. It is not difficult to see that, because $x_i > -\infty$, we must have $\mathsf{m}(g) = 0$. It follows that there exists $v \in \mathcal{C}(A)$ such that, for some s, $x_i = (A^s)_{iv} + x_v$. But then,

$$x_i \le (A^+)_{iv} + x_v \le \bigvee_{u \in \mathcal{C}(A)} (A^+)_{iu} + x_u \le (A^+\vec{x})_i = x_i \ ,$$

the last equality holding because $\vec{x}$ is evidently an eigenvector of A^+. It follows that $x_i = \bigvee_{u \in \mathcal{C}(A)} (A^+)_{iu} + x_u$. Hence, by (12), $(\mathsf{P}_A\vec{x})_i = \bigvee_{u \in \mathcal{C}(A)} (A^+)_{iu} + (A^+\vec{x})_u = x_i$. Hence $\mathsf{P}_A\vec{x} = \vec{x}$.

3. According to Lemma 1.3, $(\mathsf{P}_A)^2 = \mathsf{P}_A$. Hence the image of P_A coincides with the set of eigenvectors of P_A. By part 2, the eigenvectors of P_A are exactly the eigenvectors of A.

4. Choose $1 \le i \le n$. Since $\vec{x}$ and $\vec{y}$ are eigenvectors of A, they are also eigenvectors of A^+. Hence, since $x_u = y_u$ for all $u \in \mathcal{C}(A)$,

$$(\mathsf{P}_A\vec{x})_i = \bigvee_{u \in \mathcal{C}(A)} (A^+)_{iu} + x_u = \bigvee_{u \in \mathcal{C}(A)} (A^+)_{iu} + y_u = (\mathsf{P}_A\vec{y})_i \ .$$

By part 2, $\vec{x} = \vec{y}$.

5. If $i \in \mathcal{C}(A)$ then there exists some k such that $(A^k)_{ii} = 0$. Hence $(A^+)_{ii} = 0$. It then follows from (12) that $(\mathsf{P}_A)_{ii} = 0$.

6. Since $A\vec{x} \leq \vec{x}$ it follows that $A^k\vec{x} \leq A^{k-1}\vec{x}$. Hence, by (11), $A^+\vec{x} \leq A\vec{x} \leq \vec{x}$. If $i \in \mathcal{C}(A)$ then as in the previous part, $(A^+)_{ii} = 0$. Hence, $x_i = (A^+)_{ii} + x_i \leq (A^+\vec{x})_i \leq x_i$. It follows that $(A\vec{x})_i = x_i$ as required.

7. If $\mu(A) = 0$ then there must be a critical circuit upstream from every node of $\mathcal{G}(A)$. Hence, for any $1 \leq i \leq n$, there exists some $k \in \mathcal{C}(A)$, such that $A_{ik} > -\infty$. It follows from Lemma 1.3 and part 5 that $(\mathsf{P}_A)_{ik} \geq A_{ik} + (\mathsf{P}_A)_{kk} > -\infty$. Hence $\mathsf{P}_A : \mathbb{R}^n \to \mathbb{R}^n$, as required.

$\square$

Proposition 1.4 *Let $F \in \mathsf{MM}(n,n)$ be max-only. The fixed point result* (1) *holds for F.*

Proof If F has a fixed point then clearly $\chi(F) = h$. So suppose $\chi(F) = h$. Assume first that $h = 0$. Let A be the max-plus matrix associated to F. By Proposition 1.3 we see that $\mu(A) = 0$ and, in particular, $\mathsf{t}\mu(A) = 0$. Let $\vec{c} = \mathsf{P}_A(0,\cdots,0)$. By Lemma 1.3, $\vec{c}$ is an eigenvector of A with eigenvalue 0. Furthermore, since $\mu(A) = 0$, part 7 of Lemma 1.4 shows that $\vec{c} \in \mathbb{R}^n$. Hence F has a fixed point. If $h \neq 0$ then the same reasoning shows that $(F - h)$ has a fixed point: $(F-h)(\vec{c}) = \vec{c}$. Hence, $F(\vec{c}) = \vec{c} + h$, as required.

$\square$

Dual results to those above hold for min-only functions. To each such function is associated a matrix over the min-plus semiring: $\mathbb{R}_{\min} = \mathbb{R} \cup \{+\infty\}$ with minimum as sum and addition as multiplication. $\mathbb{R}_{\min}$ is isomorphic, as a semiring, to $\mathbb{R}_{\max}$ and any result holding over one of them has a dual over the other in which the roles of max and min are interchanged. We leave it to the reader to formulate these and any associated definitions; we shall not state them separately. It will be helpful, however, to use a different notation for the dual of the μ-vector. If B is an $n \times n$ matrix over $\mathbb{R}_{\min}$, which satisfies the nondegeneracy condition dual to (6), $\eta(B) \in \mathbb{R}^n$ will denote the vector of minimun upstream cycle means in $\mathcal{G}(B)$:

$$\eta_i(B) = \min\{\mathsf{m}(g) \mid i \Leftarrow g\}\,.$$

If F is the corresponding min-only function, then by Proposition 1.3, $\chi(F) = \eta(B)$.

With this preparation, we are now in a position to study the main concerns of the present paper.

2 Fixed points of min-max functions

The main goal of this section is to derive a constructive fixed point theorem for min-max functions. The proof of this occupies the second sub-section. We then develop an algorithm for an important special case and discuss its complexity. While we cannot determine this completely, experience shows the algorithm to be efficient. In the final sub-section we show that our general fixed point theorem leads to a straightforward proof of an earlier fixed point result due to Olsder, [33]. We begin by setting up the conditions which enter into the fixed point result.

2.1 Rectangularity

Definition 2.1 *Let $F \in \mathsf{MM}(n,m)$. A subset $S \subseteq \mathsf{MM}(n,m)$ is said to be a max-representation of F if S is a finite set of max-only functions such that $F = \bigwedge_{H \in S} H$.*

It should be clear from the remarks before Definition 1.2 that every min-max function has a max-representation and a (dual) min-representation. Since we know the cycle time vectors of max-only functions, we can estimate that of F, when it exists. Suppose that $\chi(F)$ does exist. For any $H \in S$, $F \leq H$. By property M, $\chi(F) \leq \chi(H)$. Hence,

$$\chi(F) \leq \bigwedge_{H \in S} \chi(H)\,. \tag{13}$$

A max-only representation therefore gives an upper estimate for the cycle time. This estimate can be used to develop an alternative condition for fixed points. The first difficulty is that there are many different max-representations of a given min-max function and the corresponding estimates may differ. The min-max function

$$\begin{array}{rcl} F_1(x_1,x_2,x_3) & = & (x_2+2 \vee x_3+5) \wedge x_1 \\ F_2(x_1,x_2,x_3) & = & x_2+1 \wedge x_3+2 \\ F_3(x_1,x_2,x_3) & = & x_1-1 \vee x_2+3 \end{array} \tag{14}$$

has both the max-representation

$$\left\{ \begin{pmatrix} x_2+2 \vee x_3+5 \\ x_2+1 \\ x_1-1 \vee x_2+3 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_3+2 \\ x_1-1 \vee x_2+3 \end{pmatrix} \right\}$$

and the max-representation

$$\left\{ \begin{pmatrix} x_2+2 \vee x_3+5 \\ x_3+2 \\ x_1-1 \vee x_2+3 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2+1 \\ x_1-1 \vee x_2+3 \end{pmatrix} \right\}.$$

The cycle time vectors of the constituent max-only functions can be calculated by the methods of the previous section. We leave it to the reader to show that they are, in the order in which they appear above,

$$\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 2.5 \\ 2.5 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 2.5 \\ 2.5 \\ 2.5 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

It follows that the estimate (13) gives, for the first max-representation, $(0,1,2)$, while for the second, $(0,1,1)$.

To get the best estimate, the information in all the max-representations of F must be used. Observe that the set of min-max functions $\mathsf{MM}(n,m)$ has a natural representation as an m-fold Cartesian product: $\mathsf{MM}(n,m) = \mathsf{MM}(n,1) \times \cdots \times \mathsf{MM}(n,1)$. If $S \subseteq A_1 \times \cdots \times A_m$ is a subset of such a Cartesian product, let $\pi_i(S) \subseteq A_i$ denote its projection on the i-th factor.

Definition 2.2 *The rectangularisation of S, denoted* $\mathsf{Rec}(S)$*, is defined by*

$$\mathsf{Rec}(S) = \pi_1(S) \times \cdots \times \pi_m(S)\,.$$

S is said to be rectangular if $S = \mathsf{Rec}(S)$.

It is, of course, always the case that $S \subseteq \mathsf{Rec}(S)$. It is also clear that $\pi_i(S) = \pi_i(\mathsf{Rec}(S))$. It follows that, if $S \subseteq \mathsf{MM}(n, m)$ is finite, then

$$\bigwedge_{H \in S} H = \bigwedge_{H \in \mathsf{Rec}(S)} H \quad \text{and} \quad \bigvee_{H \in S} H = \bigvee_{H \in \mathsf{Rec}(S)} H\,, \tag{15}$$

since the partial order on $\mathsf{MM}(n, m)$ is defined componentwise. Furthermore, if S contains only max-only functions, then so does $\mathsf{Rec}(S)$. It is worth observing that neither of the max-representations used above was rectangular.

Suppose that $S \subseteq P$, where $(P, \leq)$ is a partially ordered set. Denote by $\mathsf{Min}(S)$ the subset of least elements of S,

$$\mathsf{Min}(S) = \{x \in S \mid y \in S,\ y \leq x \implies y = x\}\,,$$

and by $\mathsf{Max}(S)$ the corresponding set of greatest elements. If S is finite and $x \in S$, then there exist $u \in \mathsf{Min}(S)$ and $v \in \mathsf{Max}(S)$ such that $u \leq x \leq v$.

Now suppose that P is a product partial order: $P = A_1 \times \cdots \times A_m$, with the partial order on P defined componentwise from those on the A_i.

Lemma 2.1 *Let $S_i \subseteq A_i$ be finite subsets for $1 \leq i \leq m$. Then*

$$\mathsf{Min}(S_1 \times \cdots \times S_m) = \mathsf{Min}(S_1) \times \cdots \times \mathsf{Min}(S_m)\,.$$

Proof It is clear that both $L = \mathsf{Min}(S_1 \times \cdots \times S_m)$ and $R = \mathsf{Min}(S_1) \times \cdots \times \mathsf{Min}(S_m)$ are irredundant: no two elements are related by the partial order. If $x \in S_1 \times \cdots \times S_m$ then, by definition of the least element subset, we can find $u \in L$ such that $u \leq x$. By a similar argument on each component, we can find $v \in R$ such that $v \leq x$. It follows easily that $L = R$.

□

Theorem 2.1 *Let $F \in \mathsf{MM}(n, m)$ and suppose that $S, T \subseteq \mathsf{MM}(n, m)$ are rectangular max-representations of F. Then $\mathsf{Min}(S) = \mathsf{Min}(T)$.*

Proof For $m = 1$ this is a restatement of one of the main results of an earlier paper, which asserts the existence of a canonical form for min-max functions, [21, Theorem 2.1]. Now suppose that $m > 1$. Since $\pi_i(S)$ and $\pi_i(T)$ are evidently max-representations of F_i, it follows from the first case that $\mathsf{Min}(\pi_i(S)) = \mathsf{Min}(\pi_i(T))$. But then, since S and T are rectangular, it follows from Lemma 2.1 that

$$\mathsf{Min}(S) = \mathsf{Min}(\pi_1(S)) \times \cdots \times \mathsf{Min}(\pi_m(S)) = \mathsf{Min}(\pi_1(T)) \times \cdots \times \mathsf{Min}(\pi_m(T)) = \mathsf{Min}(T)\,,$$

as required.

□

Corollary 2.1 *Let $F \in \mathsf{MM}(n,n)$ and suppose that $S, T \subseteq \mathsf{MM}(n,n)$ are rectangular max-representations of F. Then*

$$\bigwedge_{H \in S} \chi(H) = \bigwedge_{G \in T} \chi(G)\,.$$

Proof Since χ is monotonic, it must be the case that $\bigwedge_{H \in \mathsf{Min}(S)} \chi(H) = \bigwedge_{H \in S} \chi(H)$. The result follows immediately from Theorem 2.1.

□

The max-representations used for example (14) have identical rectangularisations, obtained by taking the union of the two representations. The best estimate for the cycle time of F, on the basis of Corollary 2.1, is therefore $\chi(F) \leq (0,1,1)$.

Suppose that $F \in \mathsf{MM}(n,n)$. Let $S, T \subseteq \mathsf{MM}(n,n)$ be rectangular max and min respresentations, respectively, of F. If $G \in T$ and $H \in S$ then clearly $G \leq H$ and so $\chi(G) \leq \chi(H)$. It follows that

$$\bigvee_{G \in T} \mathsf{b}\chi(G) \leq \bigvee_{G \in T} \chi(G) \leq \bigwedge_{H \in S} \chi(H) \leq \bigwedge_{H \in S} \mathsf{t}\chi(H)\,. \tag{16}$$

Furthermore, by (13) and its dual, $\chi(F)$, if it exists, must have a value intermediate between the two innermost terms.

The Duality Conjecture asserted that, if F is any min-max function, and S and T are rectangular max and min representations, respectively, of F, then

$$\bigvee_{G \in T} \chi(G) = \bigwedge_{H \in S} \chi(H)\,.$$

It is easy to see that, in this case, $\chi(F)$ must exist and have the same value. For the min-max function (14), this gives $\chi(F) = (0,1,1)$. The inequalities (16) also suggest conditions for the existence of a fixed point.

Proposition 2.1 *Suppose that S, T are rectangular max and min representations, respectively, of F and suppose in addition that F has a fixed point, where $F(\vec{x}) = \vec{x} + h$. Then (16) collapses to an equality.*

Proof Suppose that $F(\vec{x}) = \vec{x} + h$. We know in this case that χ does exist and that $\chi(F) = h$. Since S is rectangular, there must be some $H \in S$ such that $H(\vec{x}) = \vec{x} + h$. But then $h = \chi(H) = \mathsf{t}\chi(H)$. It follows that $h = \bigwedge_{H \in S} \mathsf{t}\chi(H)$. The dual argument, using the rectangularity of T, shows that $h = \bigvee_{G \in T} \mathsf{b}\chi(G)$. The result follows.

□

We see from this that equality of the outermost terms in (16) is a necessary condition for F to have a fixed point. It was shown in [19] that this is also a sufficient condition. There are two difficulties with this result. Firstly, it is nonconstructive, which is a major handicap in the applications of min-max functions. Secondly, it requires information from both a max representation and a min representation of F, which is difficult to obtain for any given function. In particular, we do not know how to deduce Olsder's Theorem from the result in [19].

Proposition 2.1 also shows that the condition $\bigwedge_{H\in S} \chi(H) = h$, and its dual, are both necessary for F to have fixed point. The main result of this paper is to show that either of these weaker conditions is also sufficient. This fixed point result, Theorem 2.2 below, suffers from neither of the defects just mentioned.

2.2 The fixed point theorem

It will be convenient, from this point onwards, to drop any notational distinction between max-only or min-only functions and their associated matrices. If A is a max-only or min-only function, we shall use the same symbol to stand for its associated $\mathbb{R}_{\max}$ or $\mathbb{R}_{\min}$ matrix, respectively. Furthermore, taking advantage of Proposition 1.3, we shall use $\chi(A)$ interchangeably with $\mu(A)$ and $\eta(A)$.

If U and V are sets, let $U\backslash V$ denote the complement of V in U: $U\backslash V = \{i \in U \mid i \notin V\}$. The next result is the key technical lemma of this section.

Lemma 2.2 *Suppose that $F \in \mathsf{MM}(n,n)$ and that $S \subseteq \mathsf{MM}(n,n)$ is a rectangular max-representation of F. Choose any family of n functions in S: $A_1, \cdots, A_n \in S$. There exists a function $K \in S$ such that $\mathsf{t}\chi(K) = \bigvee_{1\leq i\leq n} \chi_i(A_i)$.*

Proof Let $h_i = \mu_i(A_i)$ and $h = \bigvee_{1\leq i\leq n} h_i$. We have to find $K \in S$ such that $\mathsf{t}\chi(K) = h$. We can assume, without loss of generality, that $h_1 = h$. Let $U_i \subseteq \{1, \cdots, n\}$ be the subset of nodes upstream from i in $\mathcal{G}(A_i)$:

$$U_i = \{k \in \{1, \cdots, n\} \mid i \Leftarrow k \text{ in } \mathcal{G}(A_i)\}\,.$$

By convention, a node is always upstream from itself, so that $i \in U_i$. It follows that the sets $\{U_i\}$ provide a cover of $\{1, \cdots, n\}$: $U_1 \cup \cdots \cup U_n = \{1, \cdots, n\}$. Let $V_r = U_1 \cup \cdots \cup U_r$ for $1 \leq r \leq n$. The sets $\{V_r\}$ provide a filtration of $\{1, \cdots, n\}$: $U_1 = V_1 \subseteq \cdots \subseteq V_n = \{1, \cdots, n\}$. Define a function $\ell : \{1, \cdots, n\} \to \{1, \cdots, n\}$ by the filtration level at which a number first appears:

$$\ell(i) = \begin{cases} 1 & \text{if } i \in V_1 \\ r & \text{if } i \in V_r\backslash V_{r-1} \text{ for } r > 1 \end{cases}$$

Now define a new matrix K according to the following rule: $K_{ij} = (A_{\ell(i)})_{ij}$. Since S is rectangular, $K \in S$. It remains to show that K has the required property.

Let i, j be nodes of $\mathcal{G}(K)$ such that $i \leftarrow j$. Let $\ell(i) = r$, so that $i \in U_r$. By construction of U_r, $r \Leftarrow i$ in $\mathcal{G}(A_r)$. Since $i \leftarrow j$ in $\mathcal{G}(K)$, it must be the case that $K_{ij} \neq -\infty$ and so $(A_r)_{ij} \neq -\infty$. Hence, $i \leftarrow j$ in $\mathcal{G}(A_r)$ and therefore also $r \Leftarrow j$. It follows that $j \in U_r$. But then $\ell(j) \leq r$. We have shown that if $i \leftarrow j$ in $\mathcal{G}(K)$, then $\ell(i) \geq \ell(j)$.

Suppose that $g = i_1 \leftarrow \cdots \leftarrow i_m$ is a circuit in $\mathcal{G}(K)$, where $m > 1$ and $i_1 = i_m$. Let $\ell(i_1) = r$. By the previous paragraph, it must be the case that $\ell(i_j) = r$ for $1 \leq j \leq m$. Hence g is also a circuit in $\mathcal{G}(A_r)$ and furthermore $r \Leftarrow g$. But then, by virtue of (8), $\mathsf{m}(g) \leq \mu_r(A_r) = h_r$. In particular, $\mathsf{m}(g) \leq h$. Since g was chosen arbitrarily, it follows that $\mathsf{t}\mu(K) \leq h$.

Finally, let g be a critical circuit of A_1 upstream from node 1. Since we assumed that $h = h_1$, it follows that $\mathsf{m}(g) = h$. Every node on g and every node on the path from g

to 1 are in U_1. By construction of K, g is also upstream from node 1 in $\mathcal{G}(K)$. Hence, $\mathsf{t}\mu(K) \geq h$. It follows that $\mathsf{t}\mu(K) = h$, as claimed. This completes the proof.

$\square$

Lemma 2.2 has a number of useful consequences which we collect in the following Lemmas.

Lemma 2.3 *Under the same conditions as Lemma 2.2, the function* $\{1, \cdots, n\} \times S \to \mathbb{R} : (i, H) \to \chi_i(H)$ *has a saddle point:*

$$\mathsf{t}\left(\bigwedge_{H \in S} \chi(H)\right) = \bigwedge_{H \in S} \mathsf{t}\chi(H)\,.$$

Proof It is well known that half the conclusion always holds; we briefly recall the argument. Choose $1 \leq j \leq n$. For any $H \in S$, $\chi_j(H) \leq \mathsf{t}\chi(H)$. Hence, $\bigwedge_{H\in S} \chi_j(H) \leq \bigwedge_{H\in S} \mathsf{t}\chi(H)$. Since j was chosen arbitrarily, it follows that $\mathsf{t}\left(\bigwedge_{H\in S} \chi(H)\right) \leq \bigwedge_{H\in S} \mathsf{t}\chi(H)$.

Let $H_i \in S$ be a max-only function for which $\chi_i(H_i) = \bigwedge_{H\in S} \chi_i(H)$. Let $h = \bigvee_{1\leq i\leq n} \chi_i(H_i)$. It follows from Lemma 2.2 that there exists $K \in S$ such that $\mathsf{t}\chi(K) = h$. Hence,

$$\bigwedge_{H \in S} \mathsf{t}\chi(H) \;\leq\; h \;=\; \mathsf{t}\left(\bigwedge_{H \in S} \chi(H)\right).$$

The result follows.

$\square$

Lemma 2.4 *Under the same conditions as Lemma 2.2, if* $\bigwedge_{H\in S} \chi(H) = h$*, there exists* $K \in S$ *such that* $\chi(K) = h$.

Proof It follows from Lemma 2.3 that $h = \mathsf{t}\left(\bigwedge_{H\in S} \chi(H)\right) = \bigwedge_{H\in S} \mathsf{t}\chi(H)$. Let $K \in S$ be such that $\mathsf{t}\chi(K) = h$. Then,

$$h = \bigwedge_{H \in S} \chi(H) \leq \chi(K) \leq h\,. \tag{17}$$

It follows that $\chi(K) = h$, as required.

$\square$

The next result is the main theorem of this section. It follows in detail an argument given by Cochet-Terrasson and Gaubert in [6]. The additional ingredient which appears here is Lemma 2.2, in the guise of Lemma 2.4, which allows a stronger result to be derived than that in [6].

The proof is based on a min-max analogue of Howard's policy improvement algorithm for stochastic control problems with average or ergodic cost (see, for example, [41, Ch. 31–33],[35]). Typically, Howard's algorithm finds a fixed point of $F(\vec{x}) = \bigwedge_{u\in U} \vec{c}_u + P_u\vec{x}$ where U is a finite set and, for all $u \in U$, $\vec{c}_u \in \mathbb{R}^n$ is a cost vector and P_u is a row-stochastic matrix. (In this paragraph matrix operations are to be interpreted in the usual algebra.) It is easy to see using Proposition 1.1 that functions of this form are in fact topical. At each step, Howard's algorithm selects a function $A(\vec{x}) = \vec{c} + P\vec{x}$ in $S = \mathsf{Rec}\{\vec{c}_u + P_u\vec{x} \mid u \in U\}$ and finds a fixed point of it. It is necessary to assume that such a fixed point can be found, which is the case, for instance, if each matrix P_u is positive. If this point is not also a fixed point of F, then the function A is replaced by $A' \in S$ which satisfies $F(\vec{x}) = A'(\vec{x})$ and the process is repeated. Under appropriate conditions it can be shown that this leads, after finitely many steps, to a fixed point of F.

The convergence proof for the traditional Howard algorithm relies on a form of maximum principle: algebraically, the fact that the inverse of $I - P$ is monotone, for a nonnegative matrix P whose spectral radius is strictly less than one. The analogue of this in the proof below is the monotonicity property of the spectral projector which appears as part 1 of Lemma 1.4.

Theorem 2.2 *Let $F \in \mathsf{MM}(n,n)$ and suppose that $S, T \in \mathsf{MM}(n,n)$ are rectangular and, respectively, a max-representation and a min-representation of F. The following conditions are equivalent.*

1. *F has a fixed point with $F(\vec{x}) = \vec{x} + h$.*
2. *$\bigwedge_{H \in S} \chi(H) = h$.*
3. *$\bigvee_{G \in T} \chi(G) = h$.*

Proof It follows from Proposition 2.1 that *1* implies both *2* and *3*. Assume that *2* holds. We shall deduce *1*. The fact that *3* also implies *1* follows by a dual argument.

We may assume, as usual, that $h = 0$. It follows from Lemma 2.4 that there is $A_1 \in S$ such that $\chi(A_1) = 0$. By Proposition 1.4, A_1 has a fixed point: $A_1(\vec{a}_1) = \vec{a}_1$. Hence, $F(\vec{a}_1) \leq \vec{a}_1$. Since S is rectangular, we can find $A_2 \in S$ such that $A_2(\vec{a}_1) = F(\vec{a}_1)$. We can ensure, furthermore, that if $F_i(\vec{a}_1) = (\vec{a}_1)_i$, then $(A_2)_i = (A_1)_i$. Since $A_2(\vec{a}_1) \leq \vec{a}_1$, it follows by property M that $\mu(A_2) \leq 0$ and so, by a similar argument to (17) that $\mu(A_2) = 0$. As a consequence, it follows from part 6 of Lemma 1.4, that $(A_2(\vec{a}_1))_i = (\vec{a}_1)_i$ for all $i \in \mathcal{C}(A_2)$. Hence, $F_i(\vec{a}_1) = (A_2(\vec{a}_1))_i = (\vec{a}_1)_i$, and so, by construction, $(A_2)_i = (A_1)_i$ for all $i \in \mathcal{C}(A_2)$. It is then not difficult to see that $\mathcal{C}(A_2) \subseteq \mathcal{C}(A_1)$.

Since $\mu(A_2) = 0$, it also follows that A_2 has a fixed point. By part 7 of Lemma 1.4, $\mathsf{P}_{A_2} : \mathbb{R}^n \to \mathbb{R}^n$. Hence we may choose the fixed point of A_2 to be $\vec{a}_2 = \mathsf{P}_{A_2}(\vec{a}_1)$. Since $A_2(\vec{a}_1) \leq \vec{a}_1$, it follows from part 1 of Lemma 1.4 that $\vec{a}_2 \leq \vec{a}_1$. At the same time, if $i \in \mathcal{C}(A_2)$, then by part 5 of Lemma 1.4, $(\mathsf{P}_{A_2})_{ii} = 0$ and so $(\vec{a}_2)_i \geq (\vec{a}_1)_i$. Hence, $(\vec{a}_2)_i = (\vec{a}_1)_i$ for all $i \in \mathcal{C}(A_2)$.

We can now carry on and generate a sequence $(A_s, \vec{a}_s)$ for $s = 1, 2, \cdots$, such that the following properties hold:

$$\begin{array}{lrcl}
1) & A_s \in S & \text{and} & a_s \in \mathbb{R}^n \\
2) & A_s(\vec{a}_{s-1}) & = & F(\vec{a}_{s-1}) \\
3) & A_s(\vec{a}_s) & = & \vec{a}_s \\
4) & \vec{a}_s & \leq & \vec{a}_{s-1} \\
5) & (\vec{a}_s)_i & = & (\vec{a}_{s-1})_i \text{ for all } i \in \mathcal{C}(A_s) \\
6) & \mathcal{C}(A_s) & \subseteq & \mathcal{C}(A_{s-1})\,.
\end{array}$$

Evidently, since S is finite, we must have $A_k = A_l$ for some $k < l$. By property 6, $\mathcal{C}(A_k) = \cdots = \mathcal{C}(A_l)$. Hence, by property 5, $(\vec{a}_k)_i = \cdots = (\vec{a}_l)_i$ for all $i \in \mathcal{C}(A_l)$. It follows from property 3 that $\vec{a}_k$ and $\vec{a}_l$ are fixed points of A_l which agree on $\mathcal{C}(A_l)$. Hence, by part 4 of Lemma 1.4, $\vec{a}_k = \vec{a}_l$. By property 4, $\vec{a}_k = \cdots = \vec{a}_l$. In particular, $\vec{a}_l = \vec{a}_{l-1}$. By property 2, $\vec{a}_l = A_l(\vec{a}_l) = A_l(\vec{a}_{l-1}) = F(\vec{a}_{l-1}) = F(\vec{a}_l)$. It follows that $\vec{a}_l$ is a fixed point of F. This completes the proof.

□

Corollary 2.2 ([19, §3]) *If $F \in \mathsf{MM}(n,n)$ satisfies the Duality Conjecture, then the fixed point result (1) holds.*

Proof Evident.

□

Corollary 2.3 ([6]) *Suppose that $F \in \mathsf{MM}(n,n)$ has a max-representation S such that each $H \in S$ has a fixed point. Then F has a fixed point.*

Proof Since $\chi(H) = \mathsf{t}\chi(H)$ for each $H \in S$, the result follows immediately from the Theorem.

□

Corollary 2.4 ([19, Theorem 3.1]) *Let $F \in \mathsf{MM}(n,n)$ and let S, T be as in Theorem 2.2. F has a fixed point, with $F(\vec{x}) = \vec{x} + h$, if, and only if,*

$$\bigvee_{G \in T} \mathsf{b}\chi(G) = h = \bigwedge_{H \in S} \mathsf{t}\chi(H)\,.$$

Proof If F has a fixed point, this is just Proposition 2.1. If the formula holds, then it follows from (16) that condition 2 and condition 3 of Theorem 2.2 hold and hence that F has a fixed point.

□

2.3 Algorithmic issues

Finding fixed points of min-max functions is an important problem in applications. For instance, the clock schedule verification problem mentioned in §1.2 is equivalent to finding a fixed point of a min-max function associated to a digital circuit. The particular form of the min-max functions which arise in this application leads to efficient algorithms for finding fixed points. For general min-max functions the situation is less clear. Although the methods of the previous section are constructive in nature, they do not give rise to an efficient general algorithm.

The problem stems from the fact that a min-max function is typically presented in the form $F = \bigwedge_{H \in S} H$ where S is a subset of max-only functions which is not necessarily rectangular. In order to make use of the method in Theorem 2.2, it is necessary to find

$A \in \mathsf{Rec}(S)$, such that A has a fixed point and $\chi(A)$ is minimal; this is the starting point for the iteration. Searching all of $\mathsf{Rec}(S)$ to find such a function is prohibitively expensive. However, it is sometimes the case that all functions $H \in \mathsf{Rec}(S)$ have fixed points. This occurs, for instance, when S consists of functions for which the corresponding max-only matrices have no $-\infty$ entries. In this case it is easy to see, using Proposition 1.3, that each function $H \in \mathsf{Rec}(S)$ satisfies $\chi(H) = h$, for some $h \in \mathbb{R}$. Hence, by Proposition 1.4, each H has a fixed point. This situation does arise in applications. We can adapt the method of Theorem 2.2 to give a tractable algorithm in this case.

It will be convenient to extend the spectral projector notation P_A (see (12)) to general matrices: if $\mathsf{t}\mu(A) \neq 0$, let $\tilde{A} = -\mathsf{t}\mu(A) + A$, so that $\mathsf{t}\mu(\tilde{A}) = 0$, and define $\mathsf{P}_A = \mathsf{P}_{\tilde{A}}$.

Suppose that a min-max function F is given in the form:

$$F_i(\vec{x}) = \bigwedge_{u \in U(i)} A_{iu}\vec{x} \ , \tag{18}$$

where $U(1), \cdots, U(n)$ are finite sets and A_{iu} are row vectors with entries in $\mathbb{R}_{\max}$. Borrowing the vocabulary of optimal control, we say that a policy is a map $\pi : \{1, \cdots, n\} \to \bigcup_{1 \leq i \leq n} U(i)$, such that $\pi(i) \in U(i)$, for all $1 \leq i \leq n$. The corresponding policy matrix $A[\pi]$ is defined by $A[\pi]_i = A_{i\pi(i)}$. By construction, the set of policy matrices $A[\pi]$ is rectangular.

The fixed point algorithm takes as input a min-max function of the form (18) each of whose policy matrices has a fixed point. Equivalently, by Proposition 1.4, for each policy matrix, π, there exists $h_\pi \in \mathbb{R}$ such that $\chi(A[\pi]) = h_\pi$. The algorithm produces as output $\vec{x} \in \mathbb{R}^n$ and $h \in \mathbb{R}$ such that $F(\vec{x}) = \vec{x} + h$. The steps are as follows.

1. *Initialisation.* Select an arbitrary policy π_1. Set $s = 1$ and let $A_1 = A[\pi_1]$. Find $\vec{x}_1 \in \mathbb{R}^n$ and $h_1 \in \mathbb{R}$, such that $A_1\vec{x}_1 = \vec{x}_1 + h_1$.
2. If $F(\vec{x}_s) = \vec{x}_s + h_s$, then stop.
3. *Policy improvement.* Define π_{s+1} by
$$\forall 1 \leq i \leq n, \qquad \bigwedge_{u \in U(i)} A_{iu}\vec{x}_s = A_{i\pi_{s+1}(i)}\vec{x}_s \ .$$
The choice should be conservative, in the sense that $\pi_{s+1}(i) = \pi_s(i)$ whenever possible. Let $A_{s+1} = A[\pi_{s+1}]$.
4. *Value determination.*
   (a) If $\mu(A_{s+1}) < h_s$, then, take any fixed point $\vec{x}_{s+1}$ of A_{s+1}.
   (b) If $\mu(A_{s+1}) = h_s$, then, take the particular fixed point $\vec{x}_{s+1} = P_{A_{s+1}}\vec{x}_s$.
5. Increment s by one and go to step 2.

The cycle time vector $\mu(A)$ of a max-plus matrix A can be computed by Karp's algorithm, [1, 27] while Gondran and Minoux give algorithms in [16, Chapter 3, §4] which can be adapted for computing the spectral projector P_A.

The proof that the algorithm terminates is a straightforward generalisation of the method of Theorem 2.2 and is left as an exercise to the reader. The following example illustrates how the algorithm works in practice.

Consider the min-max function:

$$\begin{aligned}
F_1(x_1, x_2, x_3) &= (x_1 \vee x_2 \vee x_3 + 1) \wedge (x_1 - 1 \vee x_2 + 1 \vee x_3 + 1) \\
F_2(x_1, x_2, x_3) &= (x_1 - 1 \vee x_2 + 2 \vee x_3 + 1) \wedge (x_1 \vee x_2 + 1 \vee x_3) \\
F_3(x_1, x_2, x_3) &= (x_1 + 1 \vee x_2 \vee x_3 + 2) \wedge (x_1 \vee x_2 + 1 \vee x_3 + 2)
\end{aligned}$$

Alternatively, F can be written in the form (18), with

$$U(1) = U(2) = U(3) = \{1, 2\}, \quad \begin{array}{ll} A_{11} = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} & A_{12} = \begin{pmatrix} -1 & 1 & 1 \end{pmatrix} \\ A_{21} = \begin{pmatrix} -1 & 2 & 1 \end{pmatrix} & A_{22} = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \\ A_{31} = \begin{pmatrix} 1 & 0 & 2 \end{pmatrix} & A_{32} = \begin{pmatrix} 0 & 1 & 2 \end{pmatrix} \end{array}$$

Note that each policy matrix of F has all its entries finite and so, as discussed above, F satisfies the conditions required by the algorithm.

Initialisation. Select $\pi_1(1) = 1, \pi_1(2) = 1, \pi_1(3) = 1$ so that

$$A_1 = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix} .$$

Since $\mu(A_1) = 2$, we set $h_1 = 2$ and we choose some fixed point of A_1, for instance,

$$\vec{x}_1 = \begin{pmatrix} -2 & -1 & -1 \end{pmatrix}^T .$$

Policy improvement. We have $F(\vec{x}_1) = A_2 \vec{x}_1$ where

$$A_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix} .$$

Here, $\pi_2(1) = 1$, $\pi_2(2) = 2$ and $\pi_2(3) = 1$.

Value determination. We have $h_2 = h_1 = 2$. Accordingly, we select the particular fixed point

$$\vec{x}_2 = \begin{pmatrix} -2 & -3 & -1 \end{pmatrix}^T = \mathsf{P}_{A_2} \vec{x}_1 ,$$

where

$$\mathsf{P}_{A_2} = \begin{pmatrix} -2 & -3 & -1 \\ -3 & -4 & -2 \\ -1 & -2 & 0 \end{pmatrix} .$$

Since $F(\vec{x}_2) = \vec{x}_2 + h_2$, the algorithm terminates.

The complexity of one iteration of the algorithm is $O(n^3 + (\sum_{1 \leq i \leq n} |U(i)|)n)$, where n is the dimension of the ambient space, and $|X|$ denotes the cardinality of the set X. Indeed, Karp's algorithm for computing the cycle time vector $\mu(A)$ of a matrix A has time complexity $O(n^3)$ as do the algorithms of [16]. It follows that one value determination step costs $O(n^3)$ time. Clearly, one policy improvement step requires $(\sum_{1\leq i\leq n}|U(i)|)$ scalar products, which can be done in time $O((\sum_{1\leq i\leq n}|U(i)|)n)$.

It seems difficult to bound accurately the number of iterations of the algorithm. Experiments suggest that its average value is well below n, at least when $|U(i)|$ is $O(n)$, for all i. The situation seems very similar to that of conventional policy improvement algorithms, which are known to be excellent in practice although no polynomial bound is known in general for their execution time.

2.4 Derivation of Olsder's Theorem

Olsder has proved a fixed point theorem for certain separated min-max functions, [33, Theorem 2.1]. Although this applies only to a restricted class, it was the first result to be proved on min-max functions beyond the $\mathbb{R}_{\max}$ linear setting. We now show that it follows from Theorem 2.2.

Let $F \in \mathsf{MM}(n,n)$ be a separated function. We can assume, without loss of generality, that F has the following form

$$F_i = \begin{cases} x_1 + K_{i1} \vee \cdots \vee x_n + K_{in} & \text{if } 1 \leq i \leq s \\ x_1 + K_{i1} \wedge \cdots \wedge x_n + K_{in} & \text{if } s+1 \leq i \leq n \end{cases}$$

where $1 \leq s < n$ and K_{ij} is an $n \times n$ matrix of elements satisfying $K_{ij} \in \mathbb{R}_{\max}$ for $1 \leq i \leq s$ and $K_{ij} \in \mathbb{R}_{\min}$ for $s+1 \leq i \leq n$. Let $t = n - s$. Recall the notation introduced at the end of §1.4: if B is a matrix over $\mathbb{R}_{\min}$, then $\eta(B)$ denotes the vector of minimum upstream cycle means and $\chi(B) = \eta(B)$.

K is neither a matrix over $\mathbb{R}_{\max}$ nor over $\mathbb{R}_{\min}$ but it is convenient to break it into blocks which are. Let A, C be matrices over $\mathbb{R}_{\max}$ of size $s \times s$ and $s \times t$, respectively, corresponding to the top left and top right blocks of K and let D, B be matrices over $\mathbb{R}_{\min}$ of size $t \times s$ and $t \times t$, respectively, corresponding to the bottom left and bottom right blocks of K:

$$\begin{array}{rcllll} A_{ij} & = & K_{ij} & i \in \{1, \cdots, s\} & j \in \{1, \cdots, s\} \\ C_{i(j-s)} & = & K_{ij} & i \in \{1, \cdots, s\} & j \in \{s+1, \cdots, s+t\} \\ D_{(i-s)j} & = & K_{ij} & i \in \{s+1, \cdots, s+t\} & j \in \{1, \cdots, s\} \\ B_{(i-s)(j-s)} & = & K_{ij} & i \in \{s+1, \cdots, s+t\} & j \in \{s+1, \cdots, s+t\}\,. \end{array}$$

Suppose that F has a fixed point: $F(\vec{x}) = \vec{x} + h$. Let $\vec{y} \in \mathbb{R}^s$ be the vector obtained from $\vec{x}$ by truncating the last t components: $y_i = x_i$ for $1 \leq i \leq s$. Evidently, $A(\vec{y}) \leq \vec{y} + h$, so that $\mu(A) \leq h$. Equivalently, in scalar terms, $\mathsf{t}\mu(A) \leq h$. A dual argument shows that $h \leq \mathsf{b}\eta(B)$. Hence, a necessary condition for F to have a fixed point is that $\mathsf{t}\mu(A) \leq \mathsf{b}\eta(B)$.

Olsder's result is by way of a converse to this but requires more assumptions on the structure of A, B, C and D. To discuss it, we need to review some further material from matrix theory. For more details, see [1].

Let A be an $n \times n$ matrix over $\mathbb{R}_{\max}$. A is said to be irreducible if there does not exist any permutation matrix P such that $P^{\mathsf{t}}AP$ is in upper triangular block form. (This is identical to the notion of irreducibility for nonnegative matrices, [30, §1.2].) An equivalent condition is that $\mathcal{G}(A)$ is strongly connected. That is, if i and j are any two nodes in $\mathcal{G}(A)$, then they are upstream from each other: $i \Leftarrow j$ and $j \Leftarrow i$. If $i \leftarrow j$ in $\mathcal{G}(A)$ then it is easy to see that $\mu_i(A) \geq \mu_j(A)$. It follows that if A is irreducible then $\mu(A) = \mathsf{t}\mu(A)$. (By Proposition 1.4 we see that A has an eigenvector lying in $\mathbb{R}^n$. This is a max-plus version of the Perron-Frobenius Theorem, [1, Theorem 3.23].) If $U \subseteq \{1, \cdots, n\}$ is a subset of nodes, all of which are upstream from each other—$i \Leftarrow j$ for all $i, j \in U$—then we shall say that U is upstream (respectively, downstream) from some node $k \in \{1, \cdots, n\}$, if there is some $i \in U$ such that $k \Leftarrow i$ (respectively, $i \Leftarrow k$).

Theorem 2.3 ([33, Theorem 2.1]) *Suppose that $F \in \mathsf{MM}(n, n)$ is separated. Using the notation above, suppose further that A and B are irreducible and that both C and D have at least one finite entry. Then F has a fixed point if, and only if, $\mathsf{t}\mu(A) \leq \mathsf{b}\eta(B)$.*

Proof We begin with some initial constructions and observations. Define rectangular max and min-representations of F, $S, T \subseteq \mathsf{MM}(n, n)$, as follows.

$$\begin{aligned} S &= \{F_1\} \times \cdots \times \{F_s\} \times \textstyle\prod_{s+1 \leq i \leq s+t} \{x_j + K_{ij} \mid K_{ij} \neq +\infty\} \\ T &= \textstyle\prod_{1 \leq i \leq s} \{x_j + K_{ij} \mid K_{ij} \neq -\infty\} \times \{F_{s+1}\} \times \cdots \times \{F_{s+t}\}\,. \end{aligned}$$

Let $\vec{\mu} = \bigwedge_{H \in S} \mu(H)$ and $\vec{\eta} = \bigvee_{G \in T} \eta(G)$. We know from (16) that $\vec{\eta} \leq \vec{\mu}$.

For any $H \in S$, $H_i = F_i$ for $i \in \{1, \cdots, s\}$. Hence the top left block of H is equal to A: $H_{ij} = A_{ij}$ for $i, j \in \{1, \cdots, s\}$. Since A is irreducible by hypothesis, it follows that, for any $i, j \in \{1, \cdots, s\}$, $\mu_i(H) = \mu_j(H)$. Furthermore, $\mathsf{t}\mu(A) \leq \mu_i(H)$ for all $H \in S$. Hence,

$$\mu_i = \bigwedge_{H \in S} \mu_i(H) = \bigwedge_{H \in S} \mu_j(H) = \mu_j\,.$$

It follows that $\mu_1 = \cdots = \mu_s$. Let us call the common value μ. Evidently, $\mathsf{t}\mu(A) \leq \mu$. Dually, $\eta_{s+1} = \cdots = \eta_{s+t} = \eta$ and $\mathsf{b}\eta(B) \geq \eta$.

Consider the min-plus matrix B. It has a rectangular max-representation, $R \subseteq \mathsf{MM}(t, t)$, of the form

$$R = \prod_{1 \leq i \leq t} \{x_j + B_{ij} \mid B_{ij} \neq +\infty\}\,.$$

Each element of R is a simple function. Since B is irreducible, $\eta(B) = \mathsf{b}\eta(B)$. By Lemma 2.4 there exists a simple function $U \in R$ such that $\eta(U) = \mathsf{b}\eta(B)$. Now construct an element $K \in S$ by choosing $K_j = U_{j-s}$ for $j \in \{s{+}1, \cdots, s{+}t\}$. It is clear that $\mu_j(K) = \mathsf{b}\eta(B)$ for $j \in \{s{+}1, \cdots, s{+}t\}$. By hypothesis, $C_{ik} \neq -\infty$ for some $i \in \{1, \cdots, s\}$ and $k \in \{s{+}1, \cdots, s{+}t\}$. It follows that $i \leftarrow k$ in $\mathcal{G}(K)$. It is not difficult to see that $\mu_i(K) = \mathsf{t}\mu(A) \vee \mathsf{b}\eta(B)$. Hence,

$$\mu_i(K) = \begin{cases} \mathsf{t}\mu(A) \vee \mathsf{b}\eta(B) & \text{if } i \in \{1, \cdots, s\} \\ \mathsf{b}\eta(B) & \text{if } i \in \{s+1, \cdots, s+t\}\,. \end{cases}$$

It follows that $\mu \leq \mathsf{t}\mu(A) \vee \mathsf{b}\eta(B)$ and $\mu_i \leq \mathsf{b}\eta(B)$ for $i \in \{s+1, \cdots, s+t\}$. By a dual construction we can show that $\eta \geq \mathsf{b}\eta(B) \wedge \mathsf{t}\mu(A)$ and $\eta_i \geq \mathsf{t}\mu(A)$ for $i \in \{1, \cdots, s\}$. We can summarise what we have shown in the table below.

$$\begin{array}{ccccccc}
\mathsf{t}\mu(A) & \leq & \eta_1 & \leq & \mu & \leq & \mathsf{t}\mu(A) \vee \mathsf{b}\eta(B) \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\mathsf{t}\mu(A) & \leq & \eta_s & \leq & \mu & \leq & \mathsf{t}\mu(A) \vee \mathsf{b}\eta(B) \\
\mathsf{b}\eta(B) \wedge \mathsf{t}\mu(A) & \leq & \eta & \leq & \mu_{s+1} & \leq & \mathsf{b}\eta(B) \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\mathsf{b}\eta(B) \wedge \mathsf{t}\mu(A) & \leq & \eta & \leq & \mu_{s+t} & \leq & \mathsf{b}\eta(B)
\end{array}$$

We can now embark on the proof proper. If F has a fixed point then we saw above, in the preamble to Theorem 2.3, that $\mathsf{t}\mu(A) \leq \mathsf{b}\eta(B)$. Now suppose that $\mathsf{t}\mu(A) \leq \mathsf{b}\eta(B)$. We shall show that F must have a fixed point.

Choose $j \in \{s+1, \cdots, s+t\}$. We claim that $\mu_j \leq \mu$. To see this, choose $H \in S$ such that $\mu_1(H) = \cdots = \mu_s(H) = \mu$. By construction of S, H is simple in the components $s+1, \cdots, s+t$. Hence the node j must have an unique edge leading to it in $\mathcal{G}(H)$: say, $j \leftarrow k$. If $k \in \{s+1, \cdots, s+t\}$ then it has a similar property and we can proceed in this way until one of two mutually exclusive possibilities occur. Either the path remains entirely among the nodes in the range $\{s+1, \cdots, s+t\}$ or it contains a node $i \in \{1, \cdots, s\}$. In the latter case, $\mu_j(H) = \mu_i(H) = \mu$ since the path out of j is unique until it reaches a node in $\{1, \cdots, s\}$. Hence, $\mu_j \leq \mu$.

In the former case, j is not downstream from $\{1, \cdots, s\}$ in $\mathcal{G}(H)$. Suppose, to begin with, that there is no node in the range $\{s+1, \cdots, s+t\}$ which is downstream from $\{1, \cdots, s\}$. Because C has at least one real entry, some node in this range is upstream from $\{1, \cdots, s\}$. Since every circuit of $\mathcal{G}(H)$ in the range $\{s+1, \cdots, s+t\}$ must also be a circuit in $\mathcal{G}(B)$, it follows that $\mu \geq \mathsf{b}\eta(B)$. From the table, we see that $\mu_j \leq \mathsf{b}\eta(B)$ and so $\mu_j \leq \mu$.

We may now assume that there exists $u \in \{s+1, \cdots, s+t\}$ downstream from $\{1, \cdots, s\}$ in $\mathcal{G}(H)$. Since B is irreducible, there exists a path in $\mathcal{G}(B)$ from u to j:

$$j = u_1 \leftarrow u_2 \leftarrow \cdots \leftarrow u_m = u\,, \tag{19}$$

where $1 < m$, $\{u_1, \cdots, u_m\} \subseteq \{s+1, \cdots, s+t\}$. We may assume furthermore, without loss of generality, that $u_1, \cdots, u_{m-1}$ are not downstream from $\{1, \cdots, s\}$ in $\mathcal{G}(H)$. It follows that $\mu_{u_i}(H) \geq \mathsf{b}\eta(B)$ for $1 \leq i < m$. Define $H' \in S$ by altering H as follows:

$$\begin{array}{rcll}
(H')_i & = & H_i & \text{if } i \notin \{u_1, \cdots, u_{m-1}\} \\
(H')_{u_i} & = & B_{(u_i - s)(u_{i+1} - s)} & \text{if } 1 \leq i \leq m-1
\end{array}$$

By construction, (19) is also a path in $\mathcal{G}(H')$ and j is downstream from $\{1, \cdots, s\}$ in $\mathcal{G}(H')$. It follows that $\mu_j(H') = \mu_1(H')$. The only difference between $\mathcal{G}(H')$ and $\mathcal{G}(H)$ is at the nodes $u_1, \cdots, u_{m-1}$ which may have different edges leading to them.

Suppose that $\mu_1(H') > \mu$. This can only happen if one of the edges on (19) has created a new circuit upstream from 1 in $\mathcal{G}(H')$. Let u_r be the first node on (19) which is upstream from $\{1, \cdots, s\}$ in $\mathcal{G}(H')$. We may assume that $1 \leq r < m$, for if $r = m$, then, contrary to what was just said, no edge of (19) can have caused the change. It must now be the case that u_r was also upstream from $\{1, \cdots, s\}$ in $\mathcal{G}(H)$. Hence, $\mu = \mu_1(H) \geq \mu_{u_r}(H)$. But, as we saw above, $\mu_{u_r}(H) \geq \mathsf{b}\eta(B)$. It follows once again that $\mu_j \leq \mu$. Hence, we may assume that $\mu_1(H') = \mu$. But then $\mu_j \leq \mu_j(H') = \mu_1(H') = \mu$. In either case, $\mu_j \leq \mu$, which establishes the claim.

Now suppose that for some $j \in \{s+1, \cdots, s+t\}$, it is the case that $\mu_j < \mu$. Choose $H \in S$ such that $\mu_j(H) = \mu_j$. If p is the path in $\mathcal{G}(H)$ leading to node j then p cannot start from any node $i \in \{1, \cdots, s\}$. For if it did, $\mu \leq \mu_i(H) = \mu_j(H) < \mu$, which is nonsense. Hence p must terminate in a circuit g, which must also be a circuit in $\mathcal{G}(B)$. Hence $\mathsf{m}(g) \geq \mathsf{b}\eta(B)$. But evidently, $\mu_j(H) = \mathsf{m}(g)$, since g is the only circuit upstream from j. Hence $\mu_j(H) \geq \mathsf{b}\eta(B)$, from which it follows that $\mathsf{b}\eta(B) \geq \mu > \mu_j(H) \geq \mathsf{b}\eta(B)$, which is also nonsense. It follows that for all $j \in \{s+1, \cdots, s+t\}$, $\mu_j = \mu$. We have shown that $\bigwedge_{H \in S} \chi(H) = \mu$. By Theorem 2.2, F has a fixed point. This completes the proof.

□

The above proof is straightforward in comparison with Olsder's original argument and fits within the general framework established by the Duality Conjecture.

3 Conclusion

If H is a max-only function which does not have a fixed point (so that $\chi(H) \neq h$ for any $h \in \mathbb{R}$), it can nevertheless be shown that there exists $\vec{x} \in \mathbb{R}^n$ such that $H^k(\vec{x}) = \vec{x} + k\chi(H)$ for all $1 \leq k$. We can think of this as a generalised fixed point appropriate for the situation in which $\chi(H) \neq h$. It is shown in [13] that the policy improvement methods of §2.2 can be adapted to prove that any min-max function has a generalised fixed point. This immediately yields a proof of the Duality Conjecture. An alternative proof, which works for a larger class of topical functions, is given in [12]. It remains an open problem to show that the algorithm given in §2.3 for a subclass of min-max functions can be extended to the more general topical functions considered in [12].

References

[1] F. Baccelli, G. Cohen, G. J. Olsder, and J.-P. Quadrat. *Synchronization and Linearity.* Wiley Series in Probability and Mathematical Statistics. John Wiley, 1992.

[2] F. Baccelli and J. Mairesse. Ergodic theorems for stochastic operators and discrete event systems. Appears in [22].

[3] S. M. Burns. *Performance Analysis and Optimization of Asynchronous Circuits.* PhD thesis, California Institute of Technology, 1990.

[4] Z.-Q. Cao, K. H. Kim, and F. W. Roush. *Incline Algebra and Applications.* Mathematics and its Applications. Ellis-Horwood, 1984.

[5] J. Cochet-Terrasson. Modélisation et Méthodes Mathématiques en Économie. Rapport de Stage de DEA. Université de Paris I, 1996.

[6] J. Cochet-Terrasson and Stéphane Gaubert. A min-max analogue of the Howard algorithm. Privately circulated draft, 1996. Preliminary version appears as an Appendix to [5].

[7] G. Cohen and J.-P. Quadrat, editors. *11th International Conference on Analysis and Optimization of Systems.* Springer LNCIS 199, 1994.

[8] M. G. Crandall and L. Tartar. Some relations between nonexpansive and order preserving maps. *Proceedings of the AMS*, 78(3):385–390, 1980.

[9] R. A. Cuninghame-Green. Describing industrial processes with interference and approximating their steady-state behaviour. *Operational Research Quarterly*, 13(1):95–100, 1962.

[10] R. A. Cuninghame-Green. *Minimax Algebra*, volume 166 of *Lecture Notes in Economics and Mathematical Systems.* Springer-Verlag, 1979.

[11] S. Gaubert. On the Burnside problem for semigroups of matrices in the (max,+) algebra. *Semigroup Forum*, 52:271–292, 1996.

[12] S. Gaubert and J. Gunawardena. A non-linear hierarchy for discrete event dynamical systems. Proceedings of WODES'98, IEE, Cagliari, Italy, August 1998.

[13] S. Gaubert and J. Gunawardena. The duality theorem for min-max functions. *Comptes Rendus Acad. Sci.*, 326:43–48, 1998.

[14] S. Gaubert and M. Plus. Methods and applications of (max,+) linear algebra. In *Proceedings STACS'97.* Springer LNCS 1200, 1997.

[15] P. Glasserman and D. D. Yao. *Monotone Structure in Discrete Event Systems.* Wiley Series in Probability and Mathematical Statistics. John Wiley, 1994.

[16] M. Gondran and M. Minoux. *Graphs and Algorithms.* Wiley-Interscience Series in Discrete Mathematics. John Wiley, 1984.

[17] J. Gunawardena. An introduction to idempotency. Appears in [22].

[18] J. Gunawardena. Timing analysis of digital circuits and the theory of min-max functions. In *Digest of Technical Papers of the ACM International Workshop on Timing Issues in the Specification and Synthesis of Digital Systems.* ACM, 1993.

[19] J. Gunawardena. Cycle times and fixed points of min-max functions. In G. Cohen and J.-P. Quadrat, editors, *11th International Conference on Analysis and Optimization of Systems*, pages 266–272. Springer LNCIS 199, 1994.

[20] J. Gunawardena. A dynamic approach to timed behaviour. In B. Jonsson and J. Parrow, editors, *CONCUR'94: Concurrency Theory*, pages 178–193. Springer LNCS 836, 1994.

[21] J. Gunawardena. Min-max functions. *Discrete Event Dynamic Systems*, 4:377–406, 1994.

[22] J. Gunawardena, editor. *Idempotency.* Publications of the Isaac Newton Institute. Cambridge University Press, 1998.

[23] J. Gunawardena and M. Keane. On the existence of cycle times for some nonexpansive maps. Technical Report HPL-BRIMS-95-003, Hewlett-Packard Labs, 1995.

[24] J. Gunawardena, M. Keane, and C. Sparrow. In preparation, 1998.

[25] Y. C. Ho, editor. *Special issue on Dynamics of Discrete Event Systems*. Proceedings of the IEEE, 77(1), January 1989.

[26] P. T. Johnstone. *Stone Spaces*, volume 3 of *Studies in Advanced Mathematics*. Cambridge University Press, 1982.

[27] R. M. Karp. A characterization of the minimum cycle mean in a digraph. *Discrete Mathematics*, 23:309–311, 1978.

[28] V. N. Kolokoltsov. On linear, additive and homogeneous operators in idempotent analysis. Appears in [29].

[29] V. P. Maslov and S. N. Samborskiĭ, editors. *Idempotent Analysis*, volume 13 of *Advances in Soviet Mathematics*. American Mathematical Society, 1992.

[30] H. Minc. *Nonnegative Matrices*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley, 1988.

[31] R. D. Nussbaum. Hilbert's projective metric and iterated nonlinear maps. *Memoirs of the AMS*, 75(391), 1990.

[32] R. D. Nussbaum. Convergence of iterates of a nonlinear operator arising in statistical mechanics. *Nonlinearity*, 4:1223–1240, 1991.

[33] G. J. Olsder. Eigenvalues of dynamic max-min systems. *Discrete Event Dynamic Systems*, 1:177–207, 1991.

[34] J.-E. Pin. Tropical semirings. Appears in [22].

[35] M. L. Puterman. Markov decision processes. In *Handbook on Operations Research and Management Science*, pages 331–434. North Holland, 1990.

[36] S. Rajsbaum and M. Sidi. On the performance of synchronized programs in distributed networks with random processing times and transmission delays. *IEEE Transactions on Parallel and Distributed Systems*, 5:939–950, 1994.

[37] K. A. Sakallah, T. N. Mudge, and O. A. Olukotun. $checkT_c$ and $minT_c$: timing verification and optimal clocking of synchronous digital circuits. In *Digest of Technical Papers of the IEEE International Conference on Computer-Aided Design of Integrated Circuits*, pages 552–555. IEEE Computer Society, 1990.

[38] I. Simon. On semigroups of matrices over the tropical semiring. *Theoretical Informatics and Applications*, 28:277–294, 1994.

[39] T. Szymanski and N. Shenoy. Verifying clock schedules. In *Digest of Technical Papers of the IEEE International Conference on Computer-Aided Design of Integrated Circuits*, pages 124–131. IEEE Computer Society, 1992.

[40] J. M. Vincent. Some ergodic results on stochastic iterative discrete event systems. *Discrete Event Dynamic Systems*, 7:209–233, 1997.

[41] P. Whittle. *Optimization over Time*. John Wiley and Sons, 1986.

[42] U. Zimmermann. *Linear and Combinatorial Optimization in Ordered Algebraic Structures*, volume 10 of *Annals of Discrete Mathematics*. North-Holland, 1981.