

# ULRR

## Some novel aspects of the positive linear observer problem: differential privacy and optimal l1 sensitivity

Item Type	Article
Authors	McGlinchey, Aisling;Mason, Oliver
Citation	Journal of the Franklin Institute;357, pp. 13923–13940
Publisher	Elsevier
Download date	2026-04-15 12:54:51
Item License	<a href="https://creativecommons.org/licenses/by-nc-sa/1.0/">https://creativecommons.org/licenses/by-nc-sa/1.0/</a>
Link to Item	<a href="https://hdl.handle.net/10344/9528">https://hdl.handle.net/10344/9528</a>



# Some novel aspects of the positive linear observer problem: Differential privacy and optimal $l_1$ sensitivity

Aisling McGlinchey<sup>a,b</sup>, Oliver Mason<sup>a,b,\*</sup>

<sup>a</sup>Department of Mathematics and Statistics/Hamilton Institute, Maynooth University, Co. Kildare, Ireland

<sup>b</sup>Lero, the Science Foundation Ireland Research Centre for Software, Ireland

Received 26 March 2020; received in revised form 10 August 2020; accepted 4 October 2020

Available online 10 October 2020

---

## Abstract

We present several results concerning the  $l_1$  sensitivity, a crucial parameter for differential privacy, of a positive linear observer. Specifically, for compartmental systems we derive explicit analytic expressions for positive observers that minimize a bound for the  $l_1$  sensitivity. Results are given for single-output systems and classes of multiple-output systems. For single-output general positive systems, we characterize the optimal  $l_1$  sensitivity bound of a positive observer with given convergence rate. We also make some initial observations on sensitivity for more general classes of positive observers.

© 2020 The Authors. Published by Elsevier Ltd on behalf of The Franklin Institute.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

---

## 1. Introduction

Motivated by applications in areas such as transportation, population dynamics, and communications, there has been a lot of interest in the theory of positive systems over recent decades. The linear time invariant (LTI) system

$$\begin{aligned}x(t+1) &= Ax(t) \\ y(t) &= Cx(t),\end{aligned}\tag{1}$$

---

\* Corresponding author at: Department of Mathematics and Statistics/Hamilton Institute, Maynooth University, Co. Kildare, Ireland.

E-mail address: [oliver.mason@mu.ie](mailto:oliver.mason@mu.ie) (O. Mason).

is positive if  $A \in \mathbb{R}_+^{n \times n}$ ,  $C \in \mathbb{R}_+^{p \times n}$  are both nonnegative matrices [1]. A positive (Luenberger) observer for Eq. (1) is defined by

$$\hat{x}(t+1) = (A - LC)\hat{x}(t) + Ly(t) \quad (2)$$

where the matrix  $L$  is such that  $A - LC$  and  $LC$  are nonnegative and the spectral radius of  $A - LC$  satisfies  $\rho(A - LC) < 1$  [2,3]. Fundamental results concerning existence of positive observers have been presented in references such as [2–6]. Recently researchers have sought to develop results for observers with specific forms [7], for nonlinear positive systems [8,9], systems with missing information [10], and to extend the theory of observers to the related class of max-plus linear systems [11]. On the other hand, there are fundamental and important problems concerning the limitations/difficulties of using positive observers in performance optimization and state feedback design; see for instance the work of [12,13]. Another interesting and important line of work related to positive observers concerns the development of methods that combine state estimation with fault detection and/or disturbance signal estimation; for recent work in this direction see [14,15].

*Compartmental systems* are an important subclass of positive systems arising in applications such as transportation and population dynamics [16,17]. The seminal reference on observer design for this system class is [6], while recent results on the dual problem of stabilizing a continuous time compartmental system can be found in [18]. Applications of compartmental systems such as transportation networks or smart buildings use personal data and there is a need to add privacy protections to such systems to address public concerns over personal privacy. In the recent past, differential privacy (DP) [19] has been applied to a variety of control problems. For example, the design of differentially private Kalman Filters was considered in [20] while differentially private observer design for general, nonlinear systems was first considered in [21]; motivating examples from epidemiology and social network analysis for the addition of differential privacy to observers can be found in [21].

To make an observer of the form (2) differentially private, we add appropriate noise  $N$  to the observer state  $\hat{x}$  so that changes to the measured signal  $y$  caused by a small number of individuals are difficult to detect. In this way the privacy of the individuals concerned is protected. The ‘randomized mapping’ [22] from the signal  $y$  to  $\hat{x} + N$  is called a mechanism. A mechanism can satisfy *strict* or *non-strict* differential privacy and can be constructed using noise from various distributions such as the Gaussian or Laplace distributions. We focus on strict differential privacy and Laplace mechanisms for a variety of reasons. First, the Laplace mechanism is the most widely used DP mechanism in practice. Second, the fundamental results in [20] underpinning mechanism design for control systems concern the Gaussian and Laplace distributions. As our interest in this paper is for strict DP, the Laplace is the natural choice as the results of [20] do not support the use of Gaussian mechanisms for strict DP.

The crucial step in constructing a DP Laplace mechanism is deciding the magnitude of noise to add to  $\hat{x}$ ; equivalently determining the variance of the associated Laplace distribution. This is determined by a key parameter of the system (2) known as its  $l_1$  sensitivity [21]. In general, it is very difficult to calculate exactly the sensitivity of a system. However, in [23] a bound for the  $l_1$  sensitivity of Eq. (2) was derived which can be used to determine how much noise is needed for differential privacy. This reference was the first paper to address the problem of DP positive observer design; a corresponding bound for the case of  $l_2$  sensitivity and Gaussian mechanisms for non-strict DP was presented in [24]. In the current paper, we consider the problem of finding an observer gain  $L$  that minimizes the bound in [23]; such

a gain is optimal in that it will lead to a mechanism with minimal noise (variance) which is thus the most accurate possible.

The structure of the paper is as follows. In Section 2, we recall the relevant background to our work. In Section 3, we present our main results on optimal positive observers for compartmental systems; results are given for the single output and multiple output cases. In Section 4, we consider the problem of minimising the  $l_1$  sensitivity bound for a single output system where the norm of the observer system matrix is specified. We briefly note in Section 5 that it is possible to improve on the optimal sensitivity for classical observers by considering the more general class [4]. Finally, in Section 6, we present our concluding remarks and some ideas for future work.

## 2. Background and technical preliminaries

$\mathbb{R}^n$  denotes the vector space of  $n$ -tuples of real numbers and  $\mathbb{R}^{m \times n}$  the space of  $m \times n$  matrices with real entries. For  $x \in \mathbb{R}^n$ :  $x \geq 0$  means that  $x_i \geq 0$  for  $1 \leq i \leq n$ .  $\mathbb{R}_+^n$  denotes the nonnegative orthant  $\mathbb{R}_+^n := \{x \in \mathbb{R}^n \mid x \geq 0\}$ . Similarly, for  $A \in \mathbb{R}^{m \times n}$ :  $A \geq 0$  means that  $a_{ij} \geq 0$  for  $1 \leq i \leq m$  and  $1 \leq j \leq n$ . We will sometimes use the notation  $[A]_{ij}$  to denote the  $i, j$  entry of a matrix  $A$ .  $\mathbb{R}_+^{m \times n}$  denotes the cone of nonnegative matrices.  $A \geq B$  denotes that  $A - B \geq 0$  is a nonnegative matrix.  $A^T$  denotes the transpose of  $A$ .

For a vector  $c \in \mathbb{R}_+^n$  the support of  $c$ ,  $\text{supp}(c)$  is given by  $\text{supp}(c) = \{j : c_j \neq 0\}$ . Throughout, we use  $\|x\|$  to denote the  $l_1$  norm of  $x \in \mathbb{R}^n$  and for  $T \in \mathbb{R}^{m \times n}$ ,  $\|T\|$  denotes the  $l_1$  induced norm of  $T$ . It is well known [25] that this is given by

$$\|T\| = \max_j \sum_{i=1}^m |t_{ij}|. \tag{3}$$

Our main results concern the case where the system (1) is compartmental [16–18]. A matrix  $A$  in  $\mathbb{R}_+^{n \times n}$  is compartmental if  $\sum_{i=1}^n a_{ij} \leq 1$  for all  $1 \leq j \leq n$ . If the system matrix  $A$  in Eq. (1) is compartmental, then Eq. (1) is a compartmental system.

### Differential privacy and linear observers

We very briefly recall the most relevant facts concerning differential privacy in the context of this paper. For reasons of space, we omit much of the fundamental details on probability; for more details and background, see [20].

For differential privacy, we first must choose a similarity relation on the space of signals  $y(\cdot)$ . Following [21], given two constants  $K > 0$ ,  $0 < \alpha < 1$ , the signals  $y, y'$  are said to be similar,  $y \sim y'$  if there is some  $t_0 \geq 0$  with

$$\begin{aligned} y(t) &= y'(t) & 0 \leq t < t_0 \\ \|y(t) - y'(t)\| &\leq K\alpha^{t-t_0} & t \geq t_0. \end{aligned} \tag{4}$$

The similarity relation captures changes in the signal due to the behaviour of a small number of individuals at time  $t_0$ . The parameters  $K$  and  $\alpha$  need to be chosen in advance and depend on the application context and level of protection required.  $K$  describes how large an initial change is permitted while  $\alpha$  describes the speed with which this change decays to zero.

The sensitivity of the system (2) is the key parameter that determines the magnitude of the noise required in order to achieve differential privacy. We use  $\hat{x}_y$  to denote the observer signal corresponding to the output  $y$ .

**Definition 2.1.** The  $l_1$  sensitivity of the observer (2) is given by:

$$\Delta := \sup\{\|\hat{x}_y - \hat{x}_{y'}\| : y \sim y'\}. \tag{5}$$

A real-valued Laplace random variable with mean 0 and scale parameter  $b > 0$  has probability density function (pdf)  $f(x) = \frac{1}{2b}e^{-\frac{|x|}{b}}$ . Given  $\varepsilon > 0$ , an  $\varepsilon$  differentially private mechanism for the observer (2) can be constructed by adding noise from a Laplace distribution to each component of  $\hat{x}(t)$  at each time  $t \geq 0$ . The key fact is the following.

**Proposition 2.1** [20]. *Let Eq. (2) have sensitivity  $\Delta$  and  $\varepsilon > 0$  be given. Choose  $b > \frac{\Delta}{\varepsilon}$ , and define the mechanism  $\hat{X}_y$  by adding a Laplace random variable with mean 0 and scale parameter  $b$ , to each component of  $\hat{x}_y(t)$  for all  $t$ . Then  $\hat{X}_y$  is  $\varepsilon$  differentially private.*

Proposition 2.1 highlights the important role played by the sensitivity  $\Delta$  of the system in determining the amount of noise required for differential privacy. Exact computation of the sensitivity can be difficult and complex. However, a differentially private mechanism for Eq. (2) can be constructed using the Laplace distribution provided we have an upper bound for the  $l_1$  sensitivity of Eq. (2).

Recall that Eq. (2) is a positive observer for the positive system (1) if and only if the following conditions are satisfied [3]: (i)  $LC \geq 0$ ; (ii)  $A - LC \geq 0$ ; (iii)  $A - LC$  is Schur-stable,  $\rho(A - LC) < 1$ . In order to ensure that an observer of the form (2) has a finite  $l_1$  sensitivity with respect to the similarity definition (5), we require that  $\|A - LC\| < 1$ . Thus we are considering a more restricted form of the positive observer problem. In [23] the following bound for the  $l_1$  sensitivity of a linear observer was derived.

**Theorem 2.1.** *Consider the observer (2) with  $\|A - LC\| < 1$  and let  $K > 0, 0 < \alpha < 1$  be given. The sensitivity  $\Delta$  of Eq. (2) with respect to the similarity relation (4) satisfies the following bound:*

$$\Delta \leq \frac{K}{1 - \alpha} \left( \frac{\|L\|}{1 - \|A - LC\|} \right) \tag{6}$$

To minimize the amount of noise added to the observer state  $\hat{x}$ , we want an observer gain  $L$  that minimizes the upper bound in Theorem 2.1. As  $K$  and  $\alpha$  are fixed parameters, we focus on minimizing the function

$$\Phi(L) := \frac{\|L\|}{1 - \|A - LC\|}, \tag{7}$$

for  $L$  satisfying  $0 \leq LC \leq A, \|A - LC\| < 1$ .

It is worth noting that the function  $\Phi$  is essentially the  $l_1$  norm of the zero-initial-state observer (2). Hence, it quantifies the sensitivity of this system to perturbations in the signal  $y$  in terms of the  $l_1$  norm on the associated sequence spaces. An interesting direction for future work would be to exploit this observation to investigate whether results on robust observer design, and the dual problem of robust state feedback, can be related to the problems of differential privacy discussed here. In later examples, we shall help clarify the (somewhat complicated) relationship between  $\Phi(L)$  and the norm of the observer gain  $\|L\|$ . In particular, we show that for some system classes, reducing  $\|L\|$  can increase  $\Phi(L)$  but the opposite may also occur.

### 3. Compartmental systems

We now assume Eq. (1) is a compartmental system and seek to minimize the function  $\Phi$ , in Eq. (7), over observer gain matrices  $L$  that define a positive observer for Eq. (1). First note that if  $\|A\| < 1$ , we can choose  $L = 0$  giving  $\Phi(L) = 0$ ; as this case is trivial for the problem we are studying here, we assume that  $\|A\| = 1$  (remember  $A$  is compartmental).

#### 3.1. Single output systems: $p = 1$

We first consider the case where  $p = 1$  so that  $C = c^T$  for some nonnegative column vector  $c$ . Given a compartmental matrix  $A \in \mathbb{R}_+^{n \times n}$  with  $\|A\| = 1$  and a column vector  $c \in \mathbb{R}_+^n$ , we say that  $(A, c)$  is a *feasible pair* if there exists some nonnegative  $l \in \mathbb{R}_+^n$  with  $\|A - lc^T\| < 1$ ,  $A - lc^T \geq 0$ . Note that when  $p = 1$ , the conditions  $lc^T \geq 0, l \geq 0$  are equivalent [3]. We denote the set of all such  $l$  by  $\mathcal{F}_{A,c}$ . We first characterize feasible pairs for the compartmental case.

**Lemma 3.1.** *Let a compartmental matrix  $A \in \mathbb{R}_+^{n \times n}$  with  $\|A\| = 1$ , and  $c \in \mathbb{R}_+^n$  be given. Let  $\mathcal{J} := \{j : \sum_{i=1}^n a_{ij} = 1\}$ . Then  $(A, c)$  is a feasible pair if and only if the following conditions are satisfied.*

- (i)  $\mathcal{J} \subseteq \text{supp}(c)$ .
- (ii) There exists some  $i$  with  $1 \leq i \leq n$  such that  $a_{ij} > 0$  for all  $j \in \text{supp}(c)$ .

**Proof.** First assume that (i) and (ii) hold. Choose some  $k$  such that  $a_{kj} > 0$  for all  $j \in \text{supp}(c)$  and set  $x = \min\{\frac{a_{kj}}{c_j} : j \in \text{supp}(c)\}$ . Clearly,  $x > 0$ . Now define  $l \in \mathbb{R}_+^n$  by

$$l_i = \begin{cases} x & i = k \\ 0 & i \neq k. \end{cases} \tag{8}$$

It follows immediately that for  $i \neq k, a_{ij} - l_i c_j = a_{ij} \geq 0$  for all  $j$ . Moreover, for  $i = k, a_{kj} - l_k c_j = a_{kj} \geq 0$  for  $j \notin \text{supp}(c)$ . For  $j \in \text{supp}(c)$ , the definition of  $x$  implies that  $a_{kj} - l_k c_j = a_{kj} - xc_j \geq 0$  and so  $A - lc^T \geq 0$ . Furthermore,  $x > 0$  and  $\mathcal{J} \subseteq \text{supp}(c)$ ; these conditions imply that for  $j \in \mathcal{J}$ ,

$$\sum_{i=1}^n (a_{ij} - l_i c_j) = \left( \sum_{i=1}^n a_{ij} \right) - xc_j = 1 - xc_j < 1.$$

From the definition of  $\mathcal{J}$ , we conclude that  $\|A - lc^T\| < 1$  and hence  $(A, c)$  is a feasible pair.

Conversely, assume that  $(A, c)$  is a feasible pair and let  $l \geq 0$  be such that  $A - lc^T \geq 0$  and  $\|A - lc^T\| < 1$ . As  $\|A\| = 1$ , it follows that  $\mathcal{J}$  is non-empty and that  $l \neq 0$ . Let  $j \in \mathcal{J}$  be given. Then  $\sum_{i=1}^n a_{ij} = 1$  and, as  $\|A - lc^T\| < 1$ , we must have  $\sum_{i=1}^n (a_{ij} - l_i c_j) < 1$ . This implies that  $c_j > 0$  and hence  $j \in \text{supp}(c)$ . This proves (i). As  $l \neq 0$ , we can choose some  $i$  with  $l_i > 0$ . Then as  $a_{ij} - l_i c_j \geq 0$  for all  $i, j$  we must have  $a_{ij} > 0$  for all  $j \in \text{supp}(c)$ . This proves (ii) and completes the proof of the Lemma.  $\square$

When  $p = 1$ , the gain matrix  $L$  is simply a column vector  $l \in \mathbb{R}_+^n$ ; hence, for this subsection we slightly alter our notation for for the function  $\Phi$  in Eq. (7) and write  $\Phi(l)$  for  $l$  in  $\mathcal{F}_{A,c}$ .

In our next result, we give an explicit analytic characterisation of this minimum value of  $\Phi(l)$  for  $l \in \mathcal{F}_{A,c}$ .

**Theorem 3.1.** Let a compartmental matrix  $A \in \mathbb{R}_+^{n \times n}$  with  $\|A\| = 1$ , and  $c \in \mathbb{R}_+^n$  be given such that the set  $\mathcal{F}_{A,c}$  is non-empty. Let  $\mathcal{J} := \{j : \sum_{i=1}^n a_{ij} = 1\}$ . There exists  $\hat{l} \in \mathcal{F}_{A,c}$  such that

$$\Phi(l) \geq \Phi(\hat{l}) = \frac{1}{\min_{j \in \mathcal{J}} c_j}. \tag{9}$$

**Proof.** We first note that for any  $l \in \mathcal{F}_{A,c}$ ,

$$\begin{aligned} \|A - lc^T\| &= \max_{1 \leq j \leq n} \sum_{i=1}^n (a_{ij} - l_i c_j) = \max_{1 \leq j \leq n} \left[ \left( \sum_{i=1}^n a_{ij} \right) - \|l\| c_j \right] \\ &\geq \max_{j \in \mathcal{J}} \left[ \left( \sum_{i=1}^n a_{ij} \right) - \|l\| c_j \right] = 1 - \|l\| \min_{j \in \mathcal{J}} c_j. \end{aligned}$$

This implies that  $1 - \|A - lc^T\| \leq \|l\| \min_{j \in \mathcal{J}} c_j$  and hence that

$$\Phi(l) \geq \frac{1}{\min_{j \in \mathcal{J}} c_j}.$$

Thus to complete the proof, we need to show that there exists some  $\hat{l}$  in  $\mathcal{F}_{A,c}$  with  $\Phi(\hat{l}) = \frac{1}{\min_{j \in \mathcal{J}} c_j}$ .

Denote by  $\mathcal{J}^c$  the set  $\{1, \dots, n\} \setminus \mathcal{J}$  and, for each  $k \in \mathcal{J}^c$ , let  $\alpha_k = \sum_{i=1}^n a_{ik}$  denote the corresponding column sum. Let  $j_0$  be any index such that  $c_{j_0} = \min\{c_j : j \in \mathcal{J}\}$  and let

$$M = \min \left\{ \frac{1 - \alpha_k}{c_{j_0} - c_k} : k \in \mathcal{J}^c, c_k < c_{j_0} \right\}.$$

Then as  $\alpha_k < 1$  for all  $k \in \mathcal{J}^c$ ,  $M > 0$ . By assumption,  $\mathcal{F}_{A,c}$  is non-empty. Thus, by Lemma 3.1,  $\mathcal{J} \subseteq \text{supp}(c)$  and we can choose some  $i_0$  such that  $a_{i_0 j} > 0$  for all  $j \in \text{supp}(c)$ . Define the vector  $\hat{l} \in \mathbb{R}_+^n$  by:

$$\hat{l}_i = \begin{cases} 0 & i \neq i_0 \\ \min \left( \{M\} \cup \left\{ \frac{a_{i_0 j}}{c_j} : j \in \text{supp}(c) \right\} \right) & i = i_0. \end{cases}$$

We note the following readily verifiable facts.

- (i)  $\hat{l} \geq 0$ ;  $\|\hat{l}\| \leq M$ .
- (ii)  $a_{ij} - \hat{l}_i c_j \geq 0$  for all  $1 \leq i, j \leq n$ .
- (iii)  $\sum_{i=1}^n (a_{ij} - \hat{l}_i c_j) < 1$  for  $1 \leq j \leq n$ . This follows from the facts that  $A$  is compartmental,  $\hat{l} \neq 0$ , and  $\mathcal{J} \subseteq \text{supp}(c)$ .

To finish the proof, we will show that for  $\hat{l}$  constructed above,  $\|A - \hat{l}c^T\| = \frac{1}{c_{j_0}}$ . First note that for any  $k \in \mathcal{J}$ , by the choice of  $j_0$ ,  $1 - \|\hat{l}\|c_{j_0} \geq 1 - \|\hat{l}\|c_k$ .

Now consider  $k \in \mathcal{J}^c$ . Then  $\alpha_k < 1$  so if  $c_k \geq c_{j_0}$ ,  $1 - \|\hat{l}\|c_{j_0} > \alpha_k - \|\hat{l}\|c_k$ . Finally, if  $k \in \mathcal{J}^c$ , and  $c_k < c_{j_0}$ ,

$$\begin{aligned} \|\hat{l}\| \leq M &\Rightarrow \|\hat{l}\| \leq \frac{1 - \alpha_k}{c_{j_0} - c_k} \\ &\Rightarrow \alpha_k - \|\hat{l}\|c_k \leq 1 - \|\hat{l}\|c_{j_0}. \end{aligned}$$

Putting the previous calculations together, we see that

$$\|A - \hat{l}c^T\| = \max_{1 \leq j \leq n} \sum_{i=1}^n a_{ij} - \|\hat{l}\|c_j = 1 - \|\hat{l}\|c_{j_0}.$$

It now follows immediately that

$$\Phi(\hat{l}) = \frac{1}{c_{j_0}} = \frac{1}{\min_{j \in \mathcal{J}} c_j},$$

which completes the proof.  $\square$

*Constructing the optimal  $\hat{l}$ :* It is relatively straightforward to construct the optimal observer gain by following the steps in the proof of [Theorem 3.1](#).

1. First find the minimum value of the entries of  $c$  corresponding to indices in  $\mathcal{J}$ ; choose one such entry  $c_{j_0}$ .
2. Compute  $M$  which is the minimum of the quotients  $(1 - \alpha_k)/(c_{j_0} - c_k)$  over indices  $k$  in  $\mathcal{J}^c$  for which the denominator is positive. If no such index exists, set  $M = \infty$ .
3. Choose a row  $i_0$  of  $A$  with  $a_{i_0j} > 0$  for all  $j$  with  $c_j > 0$  and set  $b = \min\{\frac{a_{i_0j}}{c_j} : c_j > 0\}$ .
4. Finally, define  $\hat{l}$  by setting its  $i_0$ th entry equal to the minimum of  $b$  and  $M$  and all other entries zero.

**Example 3.1.** Let

$$A = \begin{pmatrix} 1/3 & 0 & 1/2 \\ 0 & 1/2 & 1/4 \\ 1/3 & 0 & 1/4 \end{pmatrix}, \quad c^T = (0 \quad 1 \quad 3/4).$$

Clearly this is a feasible pair and  $\mathcal{J} = \{3\}$  so  $c_{j_0} = c_3 = 3/4$ . Next we compute  $M$  which is given by  $(1/3)/(3/4) = 4/9$ . Clearly,  $i_0 = 2$  and  $b = 1/3$ . Putting this together we see that  $\hat{l} = (0, 1/3, 0)^T$  and it can be easily verified that  $\Phi(\hat{l}) = \frac{4}{3}$  which is the optimal value established in [Theorem 3.1](#).

**Remark:** [Theorem 3.1](#) gives an explicit, analytic expression for the minimum of the sensitivity bound for a single-output compartmental system:  $\frac{K}{(1-\alpha) \min_{j \in \mathcal{J}} c_j}$ . Further, it provides a constructive way of obtaining the gain vector  $\hat{l}$  that achieves this minimum.

### 3.2. Extension to multiple output systems ( $p \geq 2$ )

We next consider the more general case where  $p \geq 2$ . Given a compartmental matrix  $A \in \mathbb{R}_+^{n \times n}$  with  $\|A\| = 1$  and  $C \in \mathbb{R}_+^{p \times n}$ , if there exists some  $L$  satisfying

$$0 \leq LC \leq A; \quad \|A - LC\| < 1, \tag{10}$$

we say that  $(A, C)$  is a feasible pair and denote the set of all such  $L$  by  $\mathcal{F}_{A,C}$ . We wish to minimize the function  $\Phi(L)$  given by [Eq. \(7\)](#) over  $L \in \mathcal{F}_{A,C}$ .

[Lemma 3.1](#) was important for our construction of the optimal observer gain vector  $l$  in [Theorem 3.1](#). It is tempting to conjecture the following natural generalisation of this result. For  $1 \leq k \leq p$ , let  $c^{(k)}$  denote the  $k$ th row of  $C$ . As above, let  $\mathcal{J}$  denote the set  $\{j : \sum_{i=1}^n a_{ij} = 1\}$ . A natural conjecture generalising [Lemma 3.1](#) is that  $(A, C)$  is a feasible pair if and only if the following two conditions are satisfied.

(F1)  $\mathcal{J} \subseteq \bigcup_{k=1}^p \text{supp}(c^{(k)})$ .

(F2) For each  $k$  in  $\{1, \dots, p\}$ , there exists some  $i_k$  such that  $a_{i_k, j} > 0$  for all  $j \in \text{supp}(c^{(k)})$ .

Unfortunately, this conjecture is not true as is shown by the following example.

**Example 3.2.** Consider the matrices:

$$A = \begin{pmatrix} 2/3 & 0 & 0 \\ 0 & 1/2 & 3/4 \\ 1/3 & 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Then clearly  $A$  is compartmental and  $\|A\| = 1$ . Moreover,  $(A, C)$  is feasible as the matrix

$$L = \begin{pmatrix} 1/3 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

clearly satisfies Eq. (10). However, while it is true that  $\mathcal{J} \subseteq \bigcup_{k=1}^p \text{supp}(c^{(k)})$ , there is no  $i_k$  satisfying condition (F2) for  $k = 2$ .

We will revisit the previous example after our next result which extends Theorem 3.1 to multiple output compartmental systems satisfying conditions (F1) and (F2) above.

**Theorem 3.2.** Let a compartmental  $A \in \mathbb{R}_+^{n \times n}$  with  $\|A\| = 1$  and  $C \in \mathbb{R}_+^{p \times n}$  with  $p \geq 2$  be given. Assume that the pair  $(A, C)$  satisfies conditions (F1), (F2) above. Let  $\mathcal{J} = \{j : \sum_{i=1}^n a_{ij} = 1\}$  and for  $1 \leq j \leq n$ , let  $\gamma_j = \sum_{i=1}^p c_{ij}$ . Then the pair  $(A, C)$  is feasible and moreover there exists  $\hat{L} \in \mathcal{F}_{A,C}$  such that for all  $L$  in  $\mathcal{F}_{A,C}$ :

$$\Phi(L) \geq \Phi(\hat{L}) = \frac{1}{\min_{j \in \mathcal{J}} \gamma_j}, \tag{11}$$

where  $\Phi$  is given by Eq. (7).

**Proof.** Let  $L \in \mathcal{F}_{A,C}$  be given. Then, as  $A - LC \geq 0$ :

$$\begin{aligned} \|A - LC\| &= \max_{1 \leq j \leq n} \sum_{i=1}^n [A - LC]_{ij} \\ &= \max_{1 \leq j \leq n} \left( \sum_{i=1}^n a_{ij} - \sum_{i=1}^n [LC]_{ij} \right) \\ &\geq \max_{j \in \mathcal{J}} \left( \sum_{i=1}^n a_{ij} - \sum_{i=1}^n [LC]_{ij} \right). \end{aligned} \tag{12}$$

Note that while we are not assuming  $L \geq 0$  here, we do have  $C \geq 0$  and hence for  $1 \leq i, j \leq n$ :

$$[LC]_{ij} = \sum_{k=1}^p l_{ik} c_{kj} \leq \sum_{k=1}^p |l_{ik}| c_{kj}.$$

It follows that for any  $j \in \mathcal{J}$ :

$$\sum_{i=1}^n [LC]_{ij} \leq \sum_{i=1}^n \sum_{k=1}^p |l_{ik}| c_{kj} = \sum_{k=1}^p \left( \sum_{i=1}^n |l_{ik}| \right) c_{kj} \leq \|L\| \sum_{k=1}^p c_{kj} = \|L\| \gamma_j.$$

Combining this with Eq. (12), we see that

$$\|A - LC\| \geq \max_{j \in \mathcal{J}} (1 - \|L\| \gamma_j) = 1 - \|L\| \min_{j \in \mathcal{J}} \gamma_j. \tag{13}$$

It now follows immediately that for any  $L \in \mathcal{F}_{A,C}$ :

$$\Phi(L) = \frac{\|L\|}{1 - \|A - LC\|} \geq \frac{1}{\min_{j \in \mathcal{J}} \gamma_j}.$$

It remains for us to show that there exists some  $\hat{L}$  in  $\mathcal{F}_{A,C}$  such that  $\Phi(\hat{L})$  attains this lower bound.

To begin, write  $\mathcal{J}^c$  for the complement of  $\mathcal{J}$ ,  $\mathcal{J}^c = \{1, \dots, n\} \setminus \mathcal{J}$ . Also, for  $1 \leq j \leq n$ , let  $\alpha_j = \sum_{i=1}^n a_{ij}$ ; thus  $\alpha_j < 1$  for all  $j \in \mathcal{J}^c$ . It is easy to see that by choosing  $x > 0$  sufficiently small, we can ensure that

$$1 - (\min_{j \in \mathcal{J}} \gamma_j)x > \alpha_k - \gamma_k x, \quad \forall k \in \mathcal{J}^c. \tag{14}$$

Now, assumption (F2) implies that for  $1 \leq k \leq p$ , there is some (not necessarily unique)  $i_k$  in  $\{1, \dots, n\}$  such that  $a_{i_k j} > 0$  for all  $j$  with  $c_{kj} > 0$ . We use this fact to construct  $\hat{L}$ .

First, choose  $i_1$  such that  $c_{1j} > 0$  implies  $a_{i_1 j} > 0$ . Then, for some  $x > 0$  (to be determined later) set  $\hat{l}_{i_1 1} = x$ ,  $\hat{l}_{s1} = 0$  for  $s \neq i_1$ . Repeat this, choosing  $i_k$  for  $k = 2, 3, \dots, p$  and in each case setting  $\hat{l}_{i_k k} = x$  and  $\hat{l}_{sk} = 0$  otherwise. Note that all of the  $i_k$  selected need not necessarily be distinct. However, it can be seen from the construction of  $\hat{L}$  that:

- (i) each column of  $\hat{L}$  has exactly one non-zero entry which is equal to  $x$ ;
- (ii)  $\hat{l}_{sq} > 0$ ,  $c_{qj} > 0$  implies that  $a_{sj} > 0$  (as in this case  $s = i_q$ ).

From (i), it follows that  $\sum_{i=1}^n \hat{l}_{ij} = x$  for  $1 \leq j \leq p$  and that  $\|\hat{L}\| = x$ . From point (ii), it follows that by choosing  $x$  sufficiently small (and positive) we can ensure that  $A - \hat{L}C \geq 0$ . Clearly as  $\hat{L} \geq 0$ ,  $\hat{L}C \geq 0$ . Finally, note that if  $j \in \mathcal{J}$ , there is some  $k \in \{1, \dots, p\}$  such that  $c_{kj} > 0$ . Hence, by the construction of  $\hat{L}$ ,  $\hat{l}_{i_k k} = x > 0$ . This implies that  $\sum_{i=1}^n [A - \hat{L}C]_{ij} < 1$  for any such  $j$  and by the definition of  $\mathcal{J}$ ,  $\|A - \hat{L}C\| < 1$ . Thus  $\hat{L} \in \mathcal{F}_{A,C}$ . Finally, note that

$$\begin{aligned} \|A - \hat{L}C\| &= \max_{1 \leq j \leq n} \sum_{i=1}^n \left( a_{ij} - \sum_{s=1}^p \hat{l}_{is} c_{sj} \right) = \max_{1 \leq j \leq n} \left( \alpha_j - \sum_{i=1}^n \sum_{s=1}^p \hat{l}_{is} c_{sj} \right) \\ &= \max_{1 \leq j \leq n} \left( \alpha_j - \sum_{s=1}^p \left( \sum_{i=1}^n \hat{l}_{is} \right) c_{sj} \right) = \max_{1 \leq j \leq n} \left( \alpha_j - \sum_{s=1}^p x c_{sj} \right) \\ &= \max_{1 \leq j \leq n} (\alpha_j - x \gamma_j). \end{aligned}$$

Now, if we choose  $x > 0$  sufficiently small so that Eq. (14) holds, then we can ensure that  $\max_{1 \leq j \leq n} (\alpha_j - x \gamma_j) = 1 - x \min_{j \in \mathcal{J}} \gamma_j$ . As  $\|\hat{L}\| = x$  by construction, it follows that for such a choice of  $x > 0$ :

$$\Phi(\hat{L}) = \frac{x}{x \min_{j \in \mathcal{J}} \gamma_j} = \frac{1}{\min_{j \in \mathcal{J}} \gamma_j}.$$

This completes the proof.  $\square$

**Remarks:**

- (i) **Theorem 3.2** explicitly characterizes the minimum value of  $\Phi(L)$  for a multiple output compartmental system satisfying (F1), (F2). Further, the proof is constructive as it demonstrates how to construct an optimal observer gain matrix  $L$ .
- (ii) For  $p \geq 2$ , it is in general *not necessary* for the gain matrix  $L$  of a positive observer (10) to be nonnegative. However the optimal  $L$  constructed in **Theorem 3.2** is indeed nonnegative. The problem of determining classes of positive linear systems for which there exists a nonnegative optimal observer gain is an interesting one for further research. It is certainly not going to be true for arbitrary positive systems as [3] contains examples of systems (1) for which there exists a positive observer with  $LC \geq 0$  but *no positive observer with  $L \geq 0$* .

*Constructing an optimal observer gain  $\hat{L}$ :* It is possible to construct an optimal gain  $\hat{L}$  for a system satisfying (F1), (F2) in the following way.

1. Set  $\Gamma_1 = \min_{j \in \mathcal{J}} \gamma_j$ .
2. Choose  $b_1 = \min\{\frac{1-\alpha_k}{\Gamma_1-\gamma_k} : k \in \mathcal{J}^c, \Gamma_1 - \gamma_k > 0\}$ . If the set is empty, set  $b_1 = \infty$ .
3.  $k = 1$ : First choose  $i_1$  so that  $c_{1j} > 0 \Rightarrow a_{i_1j} > 0$  (this is possible by (F2)). Define the first column of  $L$  by setting  $l_{i_11} = b_1$  and all other entries zero.
4.  $k = 2, \dots, p$ : Repeat the previous step, choosing a corresponding  $i_k$  for  $k \in \{2, \dots, p\}$ .
5. Form the product  $LC$ .
6. Find the minimum value  $b_2$  of  $\frac{a_{ij}}{[LC]_{ij}}$  over pairs  $(i, j)$  with  $[LC]_{ij} > 0$ .
7. Choose any  $x > 0$  that is less than  $\min\{b_1, b_2\}$ . The matrix  $\hat{L} = xL$  is optimal.

**Example 3.3.** Let

$$A = \begin{pmatrix} 1/2 & 0 & 1/4 & 0 \\ 0 & 1/3 & 1/4 & 1/3 \\ 1/3 & 1/4 & 0 & 1/4 \\ 1/6 & 0 & 1/2 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1/2 & 1 & 0 \\ 1 & 1/4 & 0 & 1/3 \end{pmatrix}.$$

It is straightforward to see that (F1) and (F2) are satisfied and  $\mathcal{J} = \{1, 3\}$  so the optimal observer sensitivity is given by 1. Following the algorithm steps to construct  $\hat{L}$ ,  $\Gamma_1 = 1$  and  $b_1 = \min\{4/3, 5/8\} = 5/8$ . Next we note that  $i_1 = 2, i_2 = 3$  so

$$L = \begin{pmatrix} 0 & 0 \\ 5/8 & 0 \\ 0 & 5/8 \\ 0 & 0 \end{pmatrix}, \quad LC = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 5/16 & 5/8 & 0 \\ 5/8 & 5/32 & 0 & 5/24 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

A simple calculation shows that  $b_2 = 2/5$  so we can (for example) choose  $x = 1/5$  to give

$$\hat{L} = \begin{pmatrix} 0 & 0 \\ 1/8 & 0 \\ 0 & 1/8 \\ 0 & 0 \end{pmatrix}.$$

Our next example shows that the conclusion of **Theorem 3.2** does not necessarily hold without the assumptions (F1), (F2) on the matrix pair  $(A, C)$ .

**Example 3.4.** Consider again the matrices  $A, C$  given in Example 3.2. We saw earlier that the pair  $(A, C)$ , while feasible, does not satisfy condition (F2). Now suppose that  $L \in \mathbb{R}^{3 \times 2}$  is such that  $LC \geq 0$  and  $A - LC \geq 0$ . A direct calculation shows that

$$LC = \begin{pmatrix} l_{11} + l_{12} & 0 & l_{12} \\ l_{21} + l_{22} & 0 & l_{22} \\ l_{31} + l_{32} & 0 & l_{32} \end{pmatrix}.$$

The conditions  $LC \geq 0, A - LC \geq 0$  now imply that

$$l_{12} = l_{32} = 0; \quad l_{22} \geq 0; \quad l_{21} = -l_{22}.$$

Thus

$$L = \begin{pmatrix} l_{11} & 0 \\ -l_{22} & l_{22} \\ l_{31} & 0 \end{pmatrix}, \quad A - LC = \begin{pmatrix} 2/3 - l_{11} & 0 & 0 \\ 0 & 1/2 & 3/4 - l_{22} \\ 1/3 - l_{31} & 0 & 0 \end{pmatrix}.$$

It now follows that  $\|L\| = l_{11} + l_{22} + l_{31}$  and that  $\|A - LC\| \geq 1 - (l_{11} + l_{31}) \geq 1 - \|L\|$ . This implies that

$$\frac{\|L\|}{1 - \|A - LC\|} \geq 1.$$

Thus, for any  $L$  with  $0 \leq LC \leq A$ , we must have  $\Phi(L) \geq 1$ . However, for this pair  $(A, C)$ ,  $\mathcal{J} = \{1\}$  and  $\min_{j \in \mathcal{J}} \gamma_j = 2$ . Thus in this case, there exists no  $L$  in  $\mathcal{F}_{A,C}$  with

$$\Phi(L) = \frac{1}{\min_{j \in \mathcal{J}} \gamma_j} = \frac{1}{2}.$$

#### 4. Trade-offs for $l_1$ sensitivity

When designing an observer (2), the speed of convergence to the true state, as determined by  $\|A - LC\|$ , is an important consideration. In general, there will be a conflict or trade-off between obtaining the lowest possible value for the sensitivity bound (or equivalently the function  $\Phi$ ) and minimising the norm  $\|A - LC\|$ . In this section, we consider this trade-off for a general positive LTI system, not necessarily compartmental, with a single output.

Consider the system (1), where  $A$  is a nonnegative matrix, not necessarily compartmental, and  $C$  is a (non-zero) nonnegative row vector, which we will write as  $c^T$  for some  $c \in \mathbb{R}_+^n \setminus \{0\}$ . Thus, we adopt the notation of Section 3.1 and consider the interplay between the objective functions  $\|A - lc^T\|$  and  $\Phi(l)$  where  $l \in \mathbb{R}_+^n$  is constrained to lie in the feasibility region  $\mathcal{F}_{A,c}$ . Formally, we will address the following question in detail.

**Problem 4.1.** Given  $A \in \mathbb{R}_+^{n \times n}$ ,  $c \in \mathbb{R}_+^n$  and  $\eta \in [0, 1)$ , find the minimum value of  $\Phi(l)$  subject to

$$l \geq 0; \quad A - lc^T \geq 0; \quad \|A - lc^T\| = \eta. \tag{15}$$

We first characterise the possible values of  $\eta = \|A - lc^T\|$  for  $l$  satisfying the first two conditions of Eq. (15).

**Lemma 4.1.** Let  $A \in \mathbb{R}_+^{n \times n}$ ,  $c \in \mathbb{R}_+^n$  be given and suppose that  $l \in \mathbb{R}_+^n$  is such that  $A - lc^T \geq 0$ . Define the vector  $\hat{l}$  for  $1 \leq i \leq n$  by

$$\hat{l}_i = \min_{k \in \text{supp}(c)} \frac{a_{ik}}{c_k}. \tag{16}$$

Then

$$\|A - \hat{l}c^T\| \leq \|A - lc^T\| \leq \|A\| \tag{17}$$

**Proof.** As  $A - lc^T \geq 0$  and  $l \geq 0$ , it follows that for all  $1 \leq i, j \leq n$ ,  $a_{ij} - l_i c_j \geq 0$  which implies that  $l_i \leq \min_{k \in \text{supp}(c)} \frac{a_{ik}}{c_k}$  for  $1 \leq i \leq n$ . If we set  $\hat{l}_i = \min_{k \in \text{supp}(c)} \frac{a_{ik}}{c_k}$ , then  $l \leq \hat{l}$ . As  $c \geq 0$  and the  $l_1$  induced matrix norm is monotonic [25], it follows that:

$$A - \hat{l}c^T \leq A - lc^T \leq A \Rightarrow \|A - \hat{l}c^T\| \leq \|A - lc^T\| \leq \|A\|$$

**Remarks:**

(i) For the rest of this section, given  $A \in \mathbb{R}_+^{n \times n}$  and  $c \in \mathbb{R}_+^n$ , we shall use  $\eta_{min}$  and  $\eta_{max}$  to denote

$$\eta_{min} = \|A - \hat{l}c^T\|, \quad \eta_{max} = \|A\|. \tag{18}$$

- (ii) If  $\eta_{min} = \eta_{max}$ , then  $\|A - lc^T\| = \|A\|$  for any  $l$  satisfying the first two conditions of Eq. (15). There are two possibilities in this case:  $\|A\| \geq 1$  in which case the problem is not feasible;  $\|A\| < 1$  and  $l = 0$  gives a solution of Problem 4.1 with  $\Phi(l) = 0$ . Thus if  $\eta_{min} = \eta_{max}$ , the problem is either infeasible or trivial, so we shall assume for the rest of this section that  $\eta_{min} < \eta_{max}$ ; this implies that  $\hat{l} \neq 0$ .
- (iii) Let  $A \in \mathbb{R}_+^{n \times n}$ ,  $c \in \mathbb{R}_+^n$  be such that  $\eta_{min} < \eta_{max}$ . Then it is not difficult to see that for any  $\eta \in [\eta_{min}, \eta_{max}]$  there exists some  $l \in \mathbb{R}_+^n$  that satisfies Eq. (15). This is essentially an application of the Intermediate Value Theorem coupled with the continuity of the norm.
- (iv) We are interested in the function  $\Phi$  as a bound for the  $l_1$  sensitivity of a linear observer (2). This bound, given by Theorem 2.1, is only valid if  $\|A - lc^T\| < 1$  holds. For this reason, we will assume for the rest of this section that  $\eta_{min} < 1$ .

Given  $\eta \in [\eta_{min}, \eta_{max}]$ , we define the set  $\mathcal{F}(\eta)$  as follows:

$$\mathcal{F}(\eta) = \{l \in \mathbb{R}_+^n : A - lc^T \geq 0, \|A - lc^T\| = \eta\}. \tag{19}$$

In view of the remarks above,  $\mathcal{F}(\eta) \neq \{0\}$ .

Given,  $A \in \mathbb{R}_+^{n \times n}$ ,  $c \in \mathbb{R}_+^n$ , and  $\eta \in [\eta_{min}, \eta_{max}]$  we define  $M(\eta)$  as:

$$M(\eta) = \max_{j \in \text{supp}(c)} \left( \frac{\sum_{i=1}^n a_{ij} - \eta}{c_j} \right) \tag{20}$$

The following lemma notes that under the assumption in point (ii),  $M(\eta)$  is nonnegative.

**Lemma 4.2.** Let  $A \in \mathbb{R}_+^{n \times n}$ ,  $c \in \mathbb{R}_+^n \setminus \{0\}$  be given. Assume that  $\eta_{\min} < \eta_{\max}$ . Then for any  $\eta \in [\eta_{\min}, \eta_{\max}]$ ,  $M(\eta) \geq 0$ .

**Proof.** By assumption,  $\|A - \hat{l}c^T\| < \|A\|$ . This implies that for any  $j$  with  $\sum_{i=1}^n a_{ij} = \|A\|$ , we must have  $c_j > 0$ . Let  $\sum_{i=1}^n a_{ik} = \|A\|$  for some  $k$ . It follows that  $k \in \text{supp}(c)$  and that for  $\eta \in [\eta_{\min}, \eta_{\max}]$ :

$$M(\eta) = \max_{j \in \text{supp}(c)} \left( \frac{\sum_{i=1}^n a_{ij} - \eta}{c_j} \right) \geq \frac{\sum_{i=1}^n a_{ik} - \eta}{c_k} = \frac{\|A\| - \eta}{c_k} \geq 0.$$

□

The next result characterizes the norm of  $l \in \mathcal{F}(\eta)$  in terms of  $\eta$ .

**Lemma 4.3.** Let  $A \in \mathbb{R}_+^{n \times n}$ ,  $c \in \mathbb{R}_+^n \setminus \{0\}$  be given and assume that  $\eta_{\min} < \eta_{\max}$ . Let  $l \in \mathcal{F}(\eta)$  for some  $\eta \in [\eta_{\min}, \eta_{\max}]$ . Then

$$\|l\| \geq M(\eta).$$

**Proof.** As  $\|A - lc^T\| = \eta$ ,  $\max_j \sum_{i=1}^n (a_{ij} - l_i c_j) = \eta$  and hence, for all  $1 \leq j \leq n$ ,

$$\sum_{i=1}^n a_{ij} - c_j \|l\| \leq \eta \Rightarrow c_j \|l\| \geq \sum_{i=1}^n a_{ij} - \eta$$

From the definition of  $\text{supp}(c)$ , it follows immediately that:

$$\|l\| \geq \max_{j \in \text{supp}(c)} \left( \frac{\sum_{i=1}^n a_{ij} - \eta}{c_j} \right) = M(\eta)$$

□

**Remark:** The last result gives us a lower bound for  $\|l\|$  where  $l$  is in  $\mathcal{F}(\eta)$  and  $\eta \in [\eta_{\min}, \eta_{\max}]$ . We next show that this lower bound is in fact a minimum.

**Proposition 4.1.** Let  $A \in \mathbb{R}_+^{n \times n}$ ,  $c \in \mathbb{R}_+^n \setminus \{0\}$  be given. Assume that  $\eta_{\min} < \eta_{\max}$ . For any  $\eta \in [\eta_{\min}, \eta_{\max}]$  let  $M(\eta)$  be given by Eq. (20). Then:

$$\min\{\|l\| : l \in \mathcal{F}(\eta)\} = M(\eta).$$

**Proof.** From Lemma 4.3, it is enough to show that there exists  $l^* \in \mathcal{F}(\eta)$  with  $\|l^*\| = M(\eta)$ . As  $\eta_{\min} < \eta_{\max}$ , it follows that there exists  $l \in \mathcal{F}(\eta)$  such that  $\|l\| > 0$ ; choose such an  $l$ .

As  $\|A - lc^T\| = \eta$ , for all  $j \in \{1, \dots, n\}$ , we have that  $\sum_{i=1}^n a_{ij} - c_j \|l\| \leq \eta$ . Furthermore, as  $A \geq 0, c \geq 0, l \geq 0$ , it follows that

$$\{j : \sum_{i=1}^n a_{ij} > \eta\} \subseteq \text{supp}(c)$$

Hence for all  $j \notin \text{supp}(c)$ ,  $\sum_{i=1}^n a_{ij} \leq \eta$ .

Define

$$l^* = \left( \frac{M(\eta)}{\|l\|} \right) l.$$

This is possible as we have chosen  $l$  with  $\|l\| > 0$ . Moreover, as  $\eta_{\min} < \eta_{\max}$ ,  $M(\eta) \geq 0$  by Lemma 4.2 and  $\|l^*\| = M(\eta)$ .

It remains to show that  $l^* \in \mathcal{F}(\eta)$ . By Lemma 4.3,  $\frac{M(\eta)}{\|l^*\|} \leq 1$ , thus  $0 \leq l^* \leq l$  and  $A - l^*c^T \geq A - lc^T \geq 0$ . To finish we need to show that  $\|A - l^*c^T\| = \eta$ .

First of all, note that for all  $j \notin \text{supp}(c)$ ,

$$\sum_{i=1}^n a_{ij} - c_j \|l^*\| = \sum_{i=1}^n a_{ij} \leq \eta$$

Next note that the definition of  $M(\eta)$  implies that:

1.  $M(\eta) \geq \frac{\sum_{i=1}^n a_{ij} - \eta}{c_j}$  for all  $j \in \text{supp}(c)$ ;
2.  $M(\eta) = \frac{\sum_{i=1}^n a_{ij_0} - \eta}{c_{j_0}}$  for some  $j_0 \in \text{supp}(c)$ .

As  $\|l^*\| = M(\eta)$ , it follows that for all  $j \in \text{supp}(c)$ ,  $\sum_{i=1}^n a_{ij} - c_j \|l^*\| \leq \eta$  while for  $j_0$ ,  $\|A - l^*c^T\| = \sum_{i=1}^n a_{ij_0} - c_{j_0} \|l^*\| = \eta$ . Thus  $l^* \in \mathcal{F}(\eta)$  and  $\|l^*\| = M(\eta)$ . This completes the proof.  $\square$

**Remark:** It is now straightforward to apply Proposition 4.1 to answer Problem 4.1. With the application to differential privacy in mind, we make the assumption that  $\eta_{\min} < 1$  (in order to ensure the bound in Theorem 2.1 is valid). The next result follows immediately from the definition of  $\Phi$  in Eq. (7) and the set  $\mathcal{F}(\eta)$ .

**Corollary 4.1.** *Let  $A \in \mathbb{R}_+^{n \times n}$ ,  $c \in \mathbb{R}_+^n \setminus \{0\}$  be such that  $\eta_{\min} < \eta_{\max}$  and  $\eta_{\min} < 1$ . Set  $\eta_1 = \min\{\eta_{\max}, 1\}$ . Let  $\eta \in [\eta_{\min}, \eta_1)$  be given. Then*

$$\min_{l \in \mathcal{F}(\eta)} \Phi(l) = \left( \frac{M(\eta)}{1 - \eta} \right). \tag{21}$$

**Remark:** In the corollary, if  $\eta_{\max} < 1$  we can include the right endpoint  $\eta_1$  (allow  $\eta \in [\eta_{\min}, \eta_{\max}]$ ) but this makes no material difference to the argument or conclusion.

**Example 4.1.** Let

$$A = \begin{pmatrix} 3/4 & 0 \\ 1/2 & 3/4 \end{pmatrix}, \quad c^T = (1 \quad 1).$$

Then we can see that  $\eta_{\min} = 3/4$ ,  $\eta_{\max} = 5/4$ . Then for any  $\eta \in [3/4, 1)$ , we can calculate  $M(\eta) = 5/4 - \eta$  so by Corollary 4.1 the minimum value of  $\Phi(l)$  for  $\|A - lc^T\| = \eta$  is given by  $\frac{5/4 - \eta}{1 - \eta}$ .

For this example, it is worth noting that if we take  $l = (0, 5/16)^T$ ,  $\hat{l} = (0, 6/16)^T$  then we find  $\Phi(l) = 5$ ,  $\Phi(\hat{l}) = 3$  showing that  $\Phi(l) > \Phi(\hat{l})$  while  $\|l\| < \|\hat{l}\|$ . So reducing the norm of the gain increases the sensitivity bound in this case.

#### 4.1. Implications for globally optimising the sensitivity bound

We now give two simple applications of Corollary 4.1 to the problem of finding a global minimum of  $\Phi$  over  $l$  in  $\mathcal{F}_{A,c}$ .

**Proposition 4.2.** *Let  $A \in \mathbb{R}_+^{n \times n}$ ,  $c \in \mathbb{R}_+^n \setminus \{0\}$  be given. Assume that  $\eta_{\min} < \eta_{\max}$  and  $\eta_{\min} < 1$ . If  $\sum_{i=1}^n a_{ij} \geq 1$  for all  $j \in \text{supp}(c)$ , then the infimum of  $\Phi(l)$  for  $l$  in  $\mathcal{F}_{A,c}$  is  $\frac{M(\eta_{\min})}{1 - \eta_{\min}}$ .*

**Proof.** Set  $\eta_1 = \min\{\eta_{max}, 1\}$ . For a given  $\eta \in [\eta_{min}, \eta_1)$  (as above, if  $\eta_1 < 1$  we can include the right endpoint here), write

$$\frac{M(\eta)}{1 - \eta} = \max_{j \in \text{supp}(c)} \phi_j(\eta) \quad \text{where} \quad \phi_j(\eta) = \left( \frac{\sum_{i=1}^n a_{ij} - \eta}{c_j} \right) \frac{1}{1 - \eta}.$$

For each  $j \in \text{supp}(c)$ , as  $\sum_{i=1}^n a_{ij} \geq 1$ , it follows by differentiation that  $\phi_j$  is a non-decreasing function on  $[\eta_{min}, \eta_1)$ . Hence,  $\frac{M(\eta)}{1 - \eta}$  is a non-decreasing function and

$$\frac{M(\eta_{min})}{1 - \eta_{min}} \leq \frac{M(\eta)}{1 - \eta}$$

for all  $\eta \in [\eta_{min}, \eta_1)$ . The result now follows from [Corollary 4.1](#).  $\square$

Finally for this section, we prove a simple result for the case where the non-zero entries of  $c$  are all equal.

**Proposition 4.3.** *Let  $A \in \mathbb{R}_+^{n \times n}$ ,  $c \in \mathbb{R}_+^n \setminus \{0\}$  be given. Assume that  $\eta_{min} < \eta_{max}$  and  $\eta_{min} < 1$ . Suppose that there is some real number  $\lambda > 0$  such that  $c_j = \lambda$  for all  $j \in \text{supp}(c)$ . If  $\max_{j \in \text{supp}(c)} \sum_{i=1}^n a_{ij} > 1$ , then the infimum of  $\Phi(l)$  for  $l$  in  $\mathcal{F}_{A,c}$  is  $\frac{M(\eta_{min})}{1 - \eta_{min}}$ .*

**Proof.** Again, set  $\eta_1 = \min\{1, \eta_{max}\}$  and let  $\eta \in [\eta_{min}, \eta_1)$ . As  $c_j = \lambda$  for all  $j \in \text{supp}(c)$ :

$$\begin{aligned} \frac{M(\eta)}{1 - \eta} &= \frac{\max_{j \in \text{supp}(c)} \left( \frac{\sum_{i=1}^n a_{ij} - \eta}{c_j} \right)}{1 - \eta} \\ &= \frac{\max_{j \in \text{supp}(c)} (\sum_{i=1}^n a_{ij}) - \eta}{\lambda(1 - \eta)} \end{aligned}$$

If  $\max_{j \in \text{supp}(c)} \sum_{i=1}^n a_{ij} > 1$ , then  $\frac{M(\eta)}{1 - \eta}$  is an increasing function for all  $\eta$  in  $[\eta_{min}, \eta_1)$ . The result follows immediately.  $\square$

### 5. The $l_1$ sensitivity of positive observers with coordinate transformation

In [\[4\]](#), a more general form of positive observer was studied for continuous time systems. For the discrete time setting, the observer structure studied in [\[4\]](#) is of the form:

$$\begin{aligned} z(t + 1) &= Fz(t) + Gy(t) \\ \hat{x}(t) &= T^{-1}z(t). \end{aligned} \tag{22}$$

In order for this to define a positive observer for [Eq. \(1\)](#), the following conditions are sufficient:  $F \geq 0$ ,  $\rho(F) < 1$ ;  $TA - FT = GC$ ;  $T$  is inverse positive, meaning  $T^{-1} \geq 0$ .

By suitably adapting the calculation previously published in [\[23\]](#) for the observer [\(2\)](#), we can readily derive the following bound for the  $l_1$  sensitivity of [Eq. \(22\)](#). As before, we require that  $\|F\| < 1$  in order to ensure that the sensitivity bound is finite.

**Proposition 5.1.** *Consider the observer [\(22\)](#) with  $\|F\| < 1$ . Let  $K > 0$ ,  $0 < \alpha < 1$  be given. The sensitivity  $\Delta$  of [Eq. \(22\)](#) with respect to the similarity relation [\(4\)](#) satisfies the following bound:*

$$\Delta \leq \frac{K}{1 - \alpha} \left( \frac{\|T^{-1}\| \|G\|}{1 - \|F\|} \right). \tag{23}$$

**Remark:** It is easy to see that a classical Luenberger observer (2) corresponds to the choice  $T = I$ ,  $F = A - LC$ ,  $G = L$ , so Theorem 2.1 is a corollary of Proposition 5.1. It is immediate that an optimal observer of the form (22) cannot perform worse than the optimal observer of the classical type. However, note that characterising the set of possible observers satisfying Eq. (22) may be significantly more complicated than for the classical case. Hence, the problem of determining an optimal observer belonging to the more general class is likely to be far more challenging and is beyond the scope of the present paper. The point of the next example is simply to show that it is possible to improve significantly on the theoretical minimum for a classical observer by considering the more general type.

In Example 5.1, we consider an observer (22) for a single-output compartmental system. Thus, an observer is defined by a triple  $(F, T, g)$  in  $\mathbb{R}_+^{n \times n} \times \mathbb{R}^{n \times n} \times \mathbb{R}_+^n$  satisfying:

$$\|F\| < 1; \quad T^{-1} \geq 0; \quad TA - FT = gc^T. \tag{24}$$

We construct such an observer for which  $\frac{\|T^{-1}\| \|g\|}{1 - \|F\|}$  is strictly less than the theoretical minimum given by Theorem 3.1 for the classical observer.

**Example 5.1.** Consider:

$$A = \begin{pmatrix} 1/2 & 1/4 \\ 1/2 & 1/3 \end{pmatrix}, \quad c^T = (1/3 \quad 1/2).$$

Next choose:

$$T = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \quad F = \begin{pmatrix} 1/3 & 0 \\ 0 & 1/30 \end{pmatrix}$$

and  $g^T = (1/2 \ 1/10)$ . Then it is readily verified that  $TA - FT = gc^T$ , and  $T^{-1} \geq 0$ ,  $\|F\| < 1$  so Eq. (22) defines a positive observer. Moreover  $\frac{1}{\min_{j \in \mathcal{J}} c_j} = 3$  while, for this choice of  $F, T, g$ ,  $\frac{\|T^{-1}\| \|g\|}{1 - \|F\|} = \frac{9}{5}$ . Thus, the upper bound for the general observer (23) of  $\frac{9}{5}$  is significantly lower than the theoretical minimum value of 3 for a classical observer in this case.

## 6. Concluding remarks

Theorems 3.1 and 3.2 provide simple, usable expressions for the minimum value of the  $l_1$  sensitivity bound in Theorem 2.1 as well as indicating a constructive procedure for computing an optimal observer gain. In Section 4 we characterize the interplay between the  $l_1$  sensitivity bound and the rate of convergence for single-output positive LTI systems. We have provided several numerical examples to illustrate the results of the paper; Example 5.1 shows explicitly that the more general type of observer in [4] can offer significantly improved performance.

There are several interesting directions for future research. For instance, extending Theorem 3.2 to a more general class of multiple-output compartmental systems. Specifically, can we relax the assumptions (F1) and (F2) of Section 3.2? The optimal gain matrix in Theorem 3.2 is nonnegative even though, as we have noted, this does not need to be the case for general positive systems. This suggests the problem of identifying matrix pairs  $(A, C)$  with  $p \geq 2$  for which the minimum value of  $\Phi(L)$  for  $L \in \mathcal{F}_{A,C}$  occurs at a nonnegative matrix  $L$ . Another question is to consider arbitrary output matrices  $C$  for the multi-output case and the corresponding optimization. Finally, determining the minimum value of the bound for the more general observer class in Section 5 presents a significant challenge.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

The authors would like to thank the associate editor and anonymous reviewers for their constructive feedback which has helped to significantly improve the paper.

This work was supported by [Science Foundation Ireland](#) (Grant no. 13/RC/2094), and the European Regional Development Fund through the Southern & Eastern Regional Operational Programme to Lero – the Irish Software Research Centre ([www.lero.ie](http://www.lero.ie)).

## References

- [1] L. Farina, S. Rinaldi, *Positive Linear Systems: Theory and Applications*, Wiley, 2000.
- [2] M. Ait Rami, F. Tadeo, Positive observation problem for linear discrete positive systems, in: *Proceedings of the IEEE 45th Annual Conference on Decision and Control (CDC)*, 2006.
- [3] M. Ait Rami, F. Tadeo, U. Helmke, Positive observers for linear positive systems, and their implications, *Int. J. Control* 84 (4) (2011) 716–725.
- [4] J. Back, A. Astolfi, Design of positive linear observers for positive linear systems via coordinate transformations and positive realizations, *SIAM J. Control Optim.* 47 (1) (2008) 345–373.
- [5] H.M. Hardin, J.H. van Schuppen, Observers for linear positive systems, *Linear Algebra Appl.* 425 (2007) 571–607.
- [6] J.M. Van den Hof, Positive linear observers for linear compartmental systems, *SIAM J. Control Optim.* 36 (2) (1998) 590–608.
- [7] L. Liu, H.R. Karimi, X. Zhao, New approaches to positive observer design for discrete-time positive linear systems, *J. Frankl. Inst.* 355 (2018) 4336–4350.
- [8] B. Brian, J. Wang, Z. Qu, Nonlinear positive observer design for positive dynamical systems, in: *Proceedings of the American Control Conference*, 2010, pp. 6231–6237.
- [9] X. Zhao, T. Wu, X. Zheng, R. Li, Discussions on observer design of nonlinear positive systems via T-S fuzzy modeling, *Neurocomputing* 157 (2015) 70–75.
- [10] G. Wang, B. Li, Q. Zhang, C. Yang, Positive observers design for discrete-time positive systems with missing data in output, *Neurocomputing* 168 (2015) 427–434.
- [11] L. Hardouin, C.A. Maia, B. Cottenceau, M. Lhommeau, Observer design for max plus linear systems, *IEEE Trans. Autom. Control* 55 (2) (2010) 538–543.
- [12] M. Naghnaeian, P.G. Voulgaris, Performance optimization over positive  $l_\infty$  cones, *Automatica* 80 (2017) 177–188.
- [13] Z. Shu, J. Lam, H. Gao, B. Du, L. Wu, Positive observer and dynamic output-feedback controllers for interval positive linear systems, *IEEE Trans. Circuits Syst.* 55 (10) (2008) 3209–3222.
- [14] A. Oghbaee, B. Shafai, S. Nazari, Complete characterization of disturbance estimation and fault detection for positive systems, *IET Control Theory Appl.* 12 (7) (2018) 883–891.
- [15] B. Shafai, S. Nazari, A. Oghbaee, Positive unknown input observer design for positive linear systems, in: *Proceedings of the 19th International Conference on System Theory, Control and Computing (ICSTCC)*, 2015.
- [16] J.A. Jacquez, C.P. Simon, Qualitative theory of compartmental systems, *Sci. World J.* 35 (1) (1993) 43–79.
- [17] P. Van den Driessche, J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, *Math. Biosci.* 180 (1) (2002) 29–48.
- [18] M.E. Valcher, I. Zorzan, State-feedback stabilization of multi-input compartmental systems, *Syst. Control Lett.* 119 (2018) 81–91.
- [19] C. Dwork, A. Roth, The algorithmic foundations of differential privacy, *Theor. Comput. Sci.* 9 (3) (2014) 211–407.
- [20] J. Le Ny, G.J. Pappas, Differentially private filtering, *IEEE Trans. Autom. Control* 59 (2) (2014) 341–354.
- [21] J. Le Ny, Privacy-preserving nonlinear observer design using contraction analysis, in: *Proceedings of the IEEE 54th Annual Conference on Decision and Control (CDC)*, 2015, pp. 4499–4504.

- [22] C. Dwork, Differential privacy, in: Proceedings of the 33rd Annual International Colloquium on Automata, Languages and Programming, LNCS 4051, Springer-Verlag, 2006, pp. 1–12.
- [23] A. McGlinchey, O. Mason, Differential privacy and the  $l_1$  sensitivity of positive linear observers, IFAC World Congr. 50 (1) (2017) 3111–3116.
- [24] A. McGlinchey, O. Mason, Bounding the  $l_2$  sensitivity for positive linear observers, in: Proceedings of the European Control Conference, 2018, pp. 1214–1219.
- [25] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge Univ. Press, 2012.