



**HAL**  
open science

## Identifiability and observability of the SIR model with quarantine

Frédéric Hamelin, Abderrahman Iggidr, Alain Rapaport, Gauthier Sallet, Max O Souza

► **To cite this version:**

Frédéric Hamelin, Abderrahman Iggidr, Alain Rapaport, Gauthier Sallet, Max O Souza. Identifiability and observability of the SIR model with quarantine. 2021. hal-03160829v1

**HAL Id: hal-03160829**

**<https://hal.inrae.fr/hal-03160829v1>**

Preprint submitted on 5 Mar 2021 (v1), last revised 20 Mar 2023 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Identifiability and observability of the SIR model with quarantine

Frédéric Hamelin<sup>\*</sup>, Abderrahman Iggidr<sup>†</sup>, Alain Rapaport<sup>‡</sup>, Gauthier Sallet<sup>§</sup> and Max O. Souza<sup>¶</sup>

March 5, 2021

## Abstract

We analyze the identifiability and observability of the well-known SIR epidemic model with an additional compartment  $Q$  of the sub-population of infected individuals that are placed in quarantine (SIQR model), considering that the flow of individuals placed in quarantine and the size of the quarantine population are known at any time. Then, we focus on the problem of identification of the model parameters, with the synthesis of an observer.

## 1 Introduction

Many papers in epidemiology proposing a mathematical model using dynamical systems, face the problem of parameter estimation. In general, some parameters are given, extracted from the literature, while the remaining unknown parameters are estimated by fitting the model to some observed data, usually by means of an optimization algorithm based on least squares or maximum likelihood methods. Nevertheless, relatively few studies about the intrinsic property of a model to admit a unique set of parameters values for a given choice of measured variables. On the other hand, this a question that is well known in automatic control. Investigation of identifiability in mathematical epidemiology is relatively recent [11, 8, 13, 10, 6, 1, 2]. Indeed, to the best of our knowledge, the first paper considering the problem in a epidemic model studies identifiability of an intra-host model of HIV, and it has been published in 2003 in an Automatic Control journal [13]. It is also surprising that the observability and identifiability of the original Kermack-Mckendrick model has not been more studied, since this system has been widely used to model an outbreak of an infection. The observability and identifiability of the classical model SIR, with demography and constant population, has been first studied in 2005 [2].

The global health crisis of COVID-19 outbreak has led to a spectacular resurgence of interest in this type of models, but with specificities related to the detection and isolation of infected individuals [9]. This is why we revisit here the issues of identification and observability for an extended ‘SIQR’ model [4] for which such an analysis had not yet been performed (to the best of our knowledge). Once the question of identification has been settled, we also tackle the task of proposing a practical strategy for reconstructing the unique values of the parameters.

## 2 The models

Inspired by [4, 7], we consider the classical SIR model (see for instance [5]), where  $S, I, R$  denote the size of the populations of respectively susceptible, infected and recovered individuals, with an additive compartment where  $Q$  denotes the size of the population of identified and isolated infectious individuals that have been removed from the infected population and placed in quarantine:

$$\begin{cases} \dot{S} &= -\beta S \frac{I}{N-Q} \\ \dot{I} &= \beta S \frac{I}{N-Q} - (\rho + \alpha)I \\ \dot{Q} &= \alpha I - \rho Q \\ \dot{R} &= \rho I + \rho Q \end{cases} \quad (1)$$

---

<sup>\*</sup>F. Hamelin is with Institut Agro, Rennes, France [frederic.hamelin@agrocampus-ouest.fr](mailto:frederic.hamelin@agrocampus-ouest.fr)

<sup>†</sup>A. Iggidr is with Université de Lorraine, CNRS, Inria, IECL, F-57000 Metz, France [Abderrahman.Iggidr@inria.fr](mailto:Abderrahman.Iggidr@inria.fr)

<sup>‡</sup>A. Rapaport is with MISTEA, Univ. Montpellier, INRAE, Institut Agro, Montpellier, France [alain.rapaport@inrae.fr](mailto:alain.rapaport@inrae.fr)

<sup>§</sup>G. Sallet is with Université de Lorraine, CNRS, IECL, F-57000 Metz, France [gauthier.sallet@univ-lorraine.fr](mailto:gauthier.sallet@univ-lorraine.fr)

<sup>¶</sup>M.O. Souza is with Instituto de Matemática e Estatística, Universidade Federal Fluminense, Niterói - RJ, 24210-201, Brasil [maxsouza@id.uff.br](mailto:maxsouza@id.uff.br)

When the size of the total population  $N$  is large and the size of the population placed in quarantine remains small compared to  $N$  during the considered interval of time, one can consider a simplified model:

$$\begin{cases} \dot{S} &= -\beta S \frac{I}{N} \\ \dot{I} &= \beta S \frac{I}{N} - (\rho + \alpha)I \\ \dot{Q} &= \alpha I - \rho Q \\ \dot{R} &= \rho I + \rho Q \end{cases} . \quad (2)$$

Note that for both models, one has

$$S(t) + I(t) + Q(t) + R(t) = N, \quad \forall t \geq 0.$$

These models have three parameters: the infectivity parameter  $\beta$ , the recovery rate  $\rho$ , that we assume to be identical for the infected populations placed in quarantine or not, and the rate of placement in quarantine  $\alpha$ . These parameters are unknown but we assume the following hypothesis.

**Assumption 1.** *The reproduction number  $\mathcal{R}_0$  verifies*

$$\mathcal{R}_0 := \frac{\beta}{\rho + \alpha} > 1.$$

This assumption implies that the epidemic can spread in the population i.e. at initial time with  $S(0) = N - I(0)$  close to  $N$  one has  $I(0) > 0$ .

### 3 The identification problem

We assume that

- the flow  $\alpha I(t)$  of infected people placed in quarantine is known at any time  $t \geq 0$
- the size  $Q(t)$  of the population placed in quarantine is perfectly known at any time  $t \geq 0$
- the size  $N$  of the total population is known
- at initial time 0, one has  $S(0) = N - \varepsilon$ ,  $I(0) = \varepsilon$ ,  $Q(0) = 0$ ,  $R(0) = 0$  with  $\varepsilon \in (0, N)$ .

We consider then the observation function

$$y(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} := \begin{bmatrix} \alpha I(t) \\ Q(t) \end{bmatrix} \quad (3)$$

and follow the usual definitions of identifiability and observability of systems [12, 3]. However, note that when  $Q = 0$ , the system is not infinitesimally identifiable: the knowledge of the outputs and all its derivative do not allow to determine formally  $\rho$ . At  $I = 0$ , the system is not identifiable neither. We adopt the following definition of identifiability for these models.

**Definition 1.** *Given  $N > 0$  and  $\varepsilon \in (0, N)$ , we shall say that system (1) resp. (2) is identifiable for the observation (3) if there exists  $t > 0$  such that the map*

$$\begin{bmatrix} \alpha \\ \beta \\ \rho \end{bmatrix} \in (\mathbb{R}_+^*)^3 \mapsto y(\cdot) \in \mathcal{C}^\infty([0, t], \mathbb{R}_+^2)$$

is injective, where  $(S(\cdot), I(\cdot), Q(\cdot), R(\cdot))$  is solution of the Cauchy problem for the differential system (1) resp. (2) with  $S(0) = N - \varepsilon$ ,  $I(0) = \varepsilon$ ,  $Q(0) = 0$  and  $R(0) = 0$ . If moreover the map

$$\begin{bmatrix} \alpha \\ \beta \\ \rho \\ \varepsilon \end{bmatrix} \in (\mathbb{R}_+^*)^3 \times (0, N) \mapsto y(\cdot) \in \mathcal{C}^\infty([0, t], \mathbb{R}_+^2)$$

is injective, then the system (1) resp. (2) is identifiable and observable for the observation (3).

## 4 Analysis of the first model

**Proposition 1.** *System (1) is identifiable and observable for the observation (3), in the sense of Definition 1.*

*Proof.* It consists in showing that parameters and unmeasured variables  $S$  and  $I$  can be expressed as functions of the successive derivatives of the output vector  $y$ . As the variable  $I$  cannot reach 0 in finite time, we shall assume  $I \neq 0$  in the following.

Note first that with  $Q(0) = 0$  one has  $\dot{Q}(0) > 0$  and then  $y_2(t) = Q(t) > 0$  for any  $t > 0$ . The dynamics of  $Q$  gives directly the expression of the parameter  $\rho$  as:

$$\rho = \frac{y_1(t) - \dot{y}_2(t)}{y_2(t)}, \quad t > 0. \quad (4)$$

Posit  $h_1 := \frac{\dot{y}_1}{y_1}$ . One has from the dynamics of  $I$

$$h_1 = \frac{\beta S}{N - Q} - \alpha - \rho. \quad (5)$$

and then

$$(N - Q)\dot{h}_1 = -\frac{\beta^2 S}{N - Q}I + \frac{\beta S}{N - Q}\dot{Q}. \quad (6)$$

Using the equality  $\frac{\beta S}{N - Q} = \alpha + h_1 + \rho$  from (5), one obtains

$$h_2 := (N - y_2)\dot{h}_1 = (h_1 + \alpha + \rho)(-\beta I + \dot{Q}). \quad (7)$$

Let us write the derivative of  $h_2$ :

$$\begin{aligned} \dot{h}_2 &= \dot{h}_1(-\beta I + \dot{Q}) \\ &+ (h_1 + \alpha + \rho) \left[ -\beta \frac{\beta S}{N - Q}I + \beta(\alpha + \rho)I + \ddot{Q} \right] \end{aligned}$$

which can be also expressed as

$$\begin{aligned} \dot{h}_2 &= \dot{h}_1(-\beta I + \dot{Q}) \\ &+ (h_1 + \alpha + \rho) [-\beta I(h_1 + \alpha + \rho) + \beta(\alpha + \rho)I + \ddot{Q}] \\ &= \dot{h}_1(-\beta I + \dot{Q}) + (h_1 + \alpha + \rho) [-h_1\beta I + \ddot{Q}]. \end{aligned}$$

Then, using relation (7), one obtains the expression

$$\dot{h}_2 = \dot{h}_1(-\beta I + \dot{Q}) + \frac{h_2}{(-\beta I + \dot{Q})} [h_1(-\beta I + \dot{Q}) - h_1\dot{Q} + \ddot{Q}]$$

which implies

$$\begin{aligned} (-\beta I + \dot{Q})\dot{h}_2 &= \dot{h}_1(-\beta I + \dot{Q})^2 \\ &+ h_2 h_1(-\beta I + \dot{Q}) + h_2(-h_1\dot{Q} + \ddot{Q}) \end{aligned}$$

or equivalently the equation

$$\dot{h}_1(-\beta I + \dot{Q})^2 + (h_2 h_1 - \dot{h}_2)(-\beta I + \dot{Q}) + h_2(-h_1\dot{Q} + \ddot{Q}) = 0$$

to be fulfilled.

Observe that this last equation is a second order polynomial in the variable  $X = -\beta I + \dot{Q}$ . From (6) and  $\mathcal{R}_0 > 1$  one has

$$\{h_1\}_{t=0} = \frac{\beta \varepsilon}{N}(-\beta + \alpha) \left(1 - \frac{\varepsilon}{N}\right) < 0 \quad (8)$$

and this allows us to show that one also has

$$\{h_2(-h_1\dot{Q} + \ddot{Q})\}_{t=0} = \alpha \beta \rho \varepsilon^2 (\beta - \alpha) \left(1 - \frac{\varepsilon}{N}\right) > 0 \quad (9)$$

Indeed, one has

$$\begin{aligned} \ddot{Q} &= \alpha \beta \frac{SI}{N-Q} - \alpha(\rho + \alpha)I - \rho \dot{Q} \\ \Rightarrow \{\ddot{Q}\}_{t=0} &= \alpha \varepsilon \left(\beta \left(1 - \frac{\varepsilon}{N}\right) - \alpha - 2\rho\right) \end{aligned}$$

With  $h_1(0) = \beta \left(1 - \frac{\varepsilon}{N}\right) - \rho - \alpha$ , one obtains

$$\begin{aligned} \{(-h_1\dot{Q} + \ddot{Q})\}_{t=0} &= \alpha \varepsilon \left(\beta \left(1 - \frac{\varepsilon}{N}\right) - \alpha - 2\rho\right) \\ &\quad - \left(\beta \left(1 - \frac{\varepsilon}{N}\right) - \rho - \alpha\right) \alpha \varepsilon \\ &= -\alpha \rho \varepsilon < 0 \end{aligned}$$

and with  $h_2(0) = \beta \varepsilon (-\beta + \alpha) \left(1 - \frac{\varepsilon}{N}\right)$ , one gets

$$\{(h_2(-h_1\dot{Q} + \ddot{Q}))\}_{t=0} = \alpha \beta \rho \varepsilon^2 (\beta - \alpha) \left(1 - \frac{\varepsilon}{N}\right) > 0.$$

Observe also that one has  $X(0) = -\beta I(0) + \dot{Q}(0) = \varepsilon(-\beta + \alpha) < 0$ . Therefore, by continuity w.r.t.  $t$ , we obtain that for  $t > 0$  small enough,  $X$  is the unique negative solution of

$$\dot{h}_1 X^2 + (h_2 h_1 - \dot{h}_2)X + h_2(-h_1\dot{Q} + \ddot{Q}) = 0.$$

that is

$$X = \frac{-2(h_2 h_1 - \dot{h}_2) - \sqrt{(h_2 h_1 - \dot{h}_2)^2 + 4\dot{h}_1(h_1 \dot{y}_2 - \dot{y}_2)}}{2\dot{h}_1}.$$

The parameter  $\alpha$  can be then obtained from equation (6)

$$\alpha = \frac{(N - y_2)\dot{h}_1}{X} - h_1 - \rho$$

where  $\rho$  is given by (4). The initial condition  $\varepsilon$  is simply reconstructed by  $\varepsilon = y_1(0)/\alpha$  and finally one obtains the parameter  $\beta = \alpha - X(0)/\varepsilon$ .  $\square$

## 5 Analysis of the simplified model

**Proposition 2.** *System (2) is identifiable and observable for the observation (3), in the sense of Definition 1.*

*Proof.* As for model (1), one can determine the parameter  $\rho$  from any positive time as

$$\rho = \frac{y_1(t) - \dot{y}_2(t)}{y_2(t)}, \quad t > 0.$$

Then from the dynamics of  $I$  one can write

$$\frac{\beta S(t)}{N} - \alpha = h_1(t) := \frac{\dot{y}_1(t)}{y_1(t)} + \rho, \quad t > 0 \quad (10)$$

where  $h_1$  is a known function. Differentiating  $h_1$  with respect to the time gives

$$\dot{h}_1(t) = -\beta^2 S(t) \frac{I(t)}{N^2} = -\frac{\beta I(t)}{N} (h_1(t) + \alpha) \quad (11)$$

and differentiating twice

$$\ddot{h}_1(t) = -\frac{\beta}{N} \left( \frac{\beta S(t)}{N} - \rho - \alpha \right) I(t) (h_1(t) + \alpha) - \frac{\beta I(t)}{N} \dot{h}_1(t).$$

With the expression (10), we rewrite this last equation as follows

$$\ddot{h}_1(t) = -\frac{\beta I(t)}{N}(h_1(t) - \rho)(h_1(t) + \alpha) - \frac{\beta I(t)}{N}\dot{h}_1(t)$$

and with expression (11)

$$\ddot{h}_1(t) = -\dot{h}_1(t)(h_1(t) - \rho) - \frac{\beta I(t)}{N}\dot{h}_1(t).$$

One obtains then the following expression for  $\beta I(t)$

$$\beta I(t) = -N \left( \frac{\ddot{h}_1(t)}{\dot{h}_1(t)} + h_1(t) - \rho \right)$$

(note that this expression is well defined because  $h_1(t) + \alpha > 0$  for any  $t$  and thus  $\dot{h}_1(t) < 0$ ).

Finally, from (11) one reconstructs the parameter

$$\alpha = -\frac{N\dot{h}_1(t)}{\beta I(t)} - h_1(t)$$

and then the parameter

$$\beta = \alpha \frac{\beta I(t)}{y_1(t)}$$

At last, the initial condition is recovered as  $\varepsilon = y_1(0)/\alpha$ . □

## 6 Parameter estimation

The former analyses have shown that models (1) and (2) are not infinitesimally identifiable at time 0 when the initial state is  $(N - \varepsilon, \varepsilon, 0, 0)$ . One has to wait a short time  $t > 0$  to have  $Q(t) > 0$  and formally identify parameters. Thus, we expect a weak accuracy of the parameters estimation at the very beginning, that should improve with time while the state get away from this initial state and new measurements come. This is why we have opted for a dynamical estimation with the help of observers. The use of observers, although usually dedicated to state estimation (rather to parameters estimation) possesses the advantage to tune the speed of error decay. Moreover, the choice of a speed-accuracy compromise can be balanced thru simulations with synthetic data corrupted by noise, before facing real data.

Note also that for large times, the solutions of (1) and (2) converge asymptotically to non-observable non-identifiable states when  $I$  and  $Q$  are both null. Consequently, we do not look precisely for results about asymptotic convergence of the error (as this is usually done in observers theory), but rather for an exponential decay of the error estimation during initial transients.

In this section, we shall consider the additional hypothesis

**Assumption 2.** *One has*

$$\alpha \leq \rho$$

that means that the rate of placement in quarantine is not larger than the recovery rate, which is often the case in epidemic regimes.

We shall denote the elementary symmetric polynomials, where  $X$  is a vector in  $\mathbb{R}^n$ , as

$$\sigma_k^n(X) := \sum_{1 \leq i_1 < \dots < i_k \leq n} \left( \prod_{j=1}^k X_{i_j} \right), \quad i = 1 \dots n$$

**Proposition 3.** Let  $\lambda$  and  $\mu$  be two positive vectors in  $\mathbb{R}^4$  and  $\mathbb{R}^3$  respectively. For  $t > 0$ , consider the dynamical system

$$\begin{cases} \dot{\hat{z}}_1 = \hat{\delta} - \hat{\rho} - K_1(\hat{z}_1 - \log(y_1(t))), \\ \dot{\hat{z}}_2 = \frac{y_1(t)}{y_2(t)} - \hat{\rho} - K_2(\hat{z}_1 - \log(y_1(t))), \\ \dot{\hat{\delta}} = -K_3(\hat{z}_1 - \log(y_1(t))), \\ \dot{\hat{\rho}} = -K_4(\hat{z}_1 - \log(y_1(t))) - (\hat{z}_2 - \log(y_2(t))), \\ \dot{\hat{y}}_1 = \hat{v}y_1(t) - K_5y_1(t)(\hat{y}_1 - y_1(t)), \\ \dot{\hat{v}} = -\hat{k}\frac{y_1(t)}{N} - K_6y_1(t)(\hat{y}_1 - y_1(t)), \\ \dot{\hat{k}} = -K_7Ny_1(t)(\hat{y}_1 - y_1(t)) \end{cases} \quad (12)$$

with the gains vector

$$K = \begin{bmatrix} \sigma_1^4(\lambda) \\ \sigma_1^4(\lambda) + \sigma_3^4(\lambda) \\ -\sigma_4^4(\lambda) \\ -\sigma_2^4(\lambda) - \sigma_4^4(\lambda) - 1 \\ \sigma_1^3(\mu) \\ \sigma_2^3(\mu) \\ -\sigma_3^3(\mu) \end{bmatrix}$$

Then, the output vector

$$\begin{bmatrix} \hat{\rho}(t) \\ \hat{\beta}(t) = \frac{1}{2} \left( \hat{k}(t) - \sqrt{\max(\hat{k}(t)^2 - 4\hat{\delta}(t)\hat{k}(t), 0)} \right) \\ \hat{\alpha}(t) = \hat{\beta}(t) - \hat{\delta}(t) \end{bmatrix} \quad (13)$$

is an estimator of  $[\rho, \beta, \alpha]^\top$ , whose exponential decay of the error can be made as fast as desired keeping the number

$$l = \min_{i,j} \min(\lambda_i, \mu_j).$$

large, as long as  $S(t)/N$  remains close to 1. Moreover, the state  $I$  is estimated with

$$\hat{I}(t) = \frac{y_1(t)}{\hat{\alpha}(t)} \quad (14)$$

*Proof.* Posit

$$\delta := \beta - \alpha. \quad (15)$$

As far as  $S/N$  remains close to 1, the size of the population  $I$  is small compared to  $S$ , and the dynamics of the outputs  $y_1 = \alpha I$  and  $y_2 = Q$  can be approximated by the linear dynamics

$$\begin{cases} \dot{y}_1 = \delta y_1 - \rho y_1 \\ \dot{y}_2 = y_1 - \rho y_2 \end{cases} \quad (16)$$

For  $t > 0$  and  $i = 1, 2$ , consider the new outputs  $z_i(t) = \log(y_i(t))$ , whose dynamics is given by the system

$$\begin{cases} \dot{z}_1 = \delta - \rho \\ \dot{z}_2 = \exp(z_1 - z_2) - \rho \end{cases} \quad (17)$$

For this sub-system with unknown parameters  $\delta$  and  $\rho$ , we consider the following candidate observer in  $\mathbb{R}^4$ :

$$\begin{cases} \dot{\hat{z}}_1 = \hat{\delta} - \hat{\rho} - K_1(\hat{z}_1 - z_1), \\ \dot{\hat{z}}_2 = \exp(z_1 - z_2) - \hat{\rho} - K_2(\hat{z}_1 - z_1), \\ \dot{\hat{\delta}} = -K_3(\hat{z}_1 - z_1) \\ \dot{\hat{\rho}} = -K_4(\hat{z}_1 - z_1) - (\hat{z}_2 - z_2), \end{cases} \quad (18)$$

The dynamics of the error  $e_1 = (\hat{z}_1, \hat{z}_2, \hat{\delta}, \hat{\rho})^\top - (z_1, z_2, \delta, \rho)^\top$  is given by the linear time invariant system  $\dot{e}_1 = M_1 e_1$  with

$$M_1 = \begin{pmatrix} -K_1 & 0 & 1 & -1 \\ -K_2 & 0 & 0 & -1 \\ -K_3 & 0 & 0 & 0 \\ -K_4 & -1 & 0 & 0 \end{pmatrix}$$

A calculation shows that with the choice

$$\begin{cases} K_1 = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4, \\ K_2 = \sum \lambda_i + \sum \lambda_i \lambda_j \lambda_k, \\ K_3 = -\lambda_1 \lambda_2 \lambda_3 \lambda_4 K_{42} \\ K_4 = -(\sum \lambda_i \lambda_j + \lambda_1 \lambda_2 \lambda_3 \lambda_4 + 1), \end{cases}$$

the spectrum of  $M_1$  is  $\{-\lambda_i, i = 1 \dots 4\}$ . This shows that the first four equations of system (12) gives the reconstruction of the parameters  $\delta$  and  $\rho$  with an exponential decay of the error larger than  $\min_i \lambda_i$ .

Posit now

$$k := \frac{\beta^2}{\alpha}. \quad (19)$$

and consider the variable

$$v(t) := \beta \frac{S(t)}{N} - \rho - \alpha \quad (20)$$

As far as  $S(T)/N$  remains close to 1, one can write the approximation

$$\dot{v} = -\frac{\beta^2}{\alpha} \frac{S(t)}{N^2} y_1(t) \simeq -\frac{\beta^2}{\alpha N} y_1(t)$$

Then, this amounts to approximate the dynamics of (1) or (2) in the  $(y_1, v, k)$  coordinates by the following dynamical system

$$\begin{cases} \dot{y}_1 = v y_1 \\ \dot{v} = -\frac{k}{N} y_1 \end{cases}$$

where  $k$  is an unknown parameter, for which we consider the following candidate observer in  $\mathbb{R}^3$

$$\begin{cases} \dot{\hat{y}}_1 = \hat{v} y_1 - K_5 y_1 (\hat{y}_1 - y_1) \\ \dot{\hat{v}} = -\hat{k} \frac{y_1}{N} - K_6 y_1 (\hat{y}_1 - y_1) \\ \dot{\hat{k}} = -K_7 N y_1 (\hat{y}_1 - y_1) \end{cases} \quad (21)$$

whose dynamics of the error  $e_2 = (\hat{y}_1, \hat{v}, \hat{k})^\top - (y_1, v, k)^\top$  is given by the non-autonomous linear system

$$\dot{e}_2 = y_1(t) M_2 e_2 \quad (22)$$

with

$$M_2 = \begin{pmatrix} -K_5 & 1 & 0 \\ -K_6 & 0 & -\frac{1}{N} \\ -K_7 N & 0 & 0 \end{pmatrix}$$

One can easily check that for the choice

$$\begin{cases} K_5 = \mu_1 + \mu_2 + \mu_3, \\ K_6 = \mu_1 \mu_2 + \mu_1 \mu_3 + \mu_2 \mu_3, \\ K_7 = -\mu_1 \mu_2 \mu_3, \end{cases}$$



the spectrum of  $M_2$  is  $\{-\mu_i, i = 1 \dots 3\}$ . Then, from (22), we obtain the upper bound on the error decrease

$$|\hat{k}(t) - k| \leq \exp\left(-(\min_j \mu_j) \int_0^T y_1(\tau) d\tau\right) \|e_2(0)\|$$

whose exponential decay can be made as large as desired with large  $l$ .

Finally, from the reconstruction of parameters  $\delta, k$  by observers (18), (21) and expressions (15) and (19), the original parameters  $\alpha, \beta$  are recovered as roots of

$$\beta^2 - k\beta + k\delta = 0 \Rightarrow \beta = \frac{k \pm \sqrt{k^2 - 4k\delta}}{2} \quad (23)$$

Note first that Assumption 1 implies  $\beta > \alpha$  and thus  $k^2 - 4k\delta > 0$ . Moreover, one has  $k > \beta(1 + \frac{\rho}{\alpha})$ , and by Assumption 2 one has  $k > 2\beta$ , which implies that only the smaller root of (23) is valid, leading to the expression (13) of the estimator. Note that this expression preserves the exponential decay of the error obtained for  $\delta, k$  and  $\rho$ .  $\square$

Let us make some comments about this observer. It consists in reconstructing functions of the parameters  $\delta$  and  $k$  and not directly the parameters  $\alpha, \beta$ . There is an apparent redundancy of variables  $\hat{z}_1$  and  $\hat{y}_1$  in dynamics (12), which reconstruct  $\log y_1$  and  $y_1$ . Indeed, this allows to decouple the observer into two sub-systems of dimensions 4 and 3, which avoids the use of two large correction gains compared to a full order observer. Finally, outputs of these two sub-systems are coupled in expression (13) to reconstruct the original parameters.

## 7 Numerical illustrations

The proposed observer has been tested with synthetic data for a population size  $N = 10^5$  with parameter values  $\alpha = 0.07$ ,  $\beta = 0.4$ ,  $\rho = 0.1$  and initial condition  $I(0) = 10$ ,  $Q(0) = 5$ ,  $R(0) = 0$  over a time horizon of 10 days (see Figure 1). The gains have been computed for the choice of vectors  $\lambda = [1; 1.5; 2; 2.5]$  and  $\mu = [1/(13 \cdot 10^3); 1/(15 \cdot 10^3); 1/(19 \cdot 10^3)]$ . Note that vector  $\mu$  has been chosen quite small to avoid too large gains when multiplied by  $N$  in the observer equations.

Then, we have simulated a measurement noise with a centered Gaussian law of variance equal to 5% of the signal (see Figure 2). Because the expression (14) of the estimation of the state  $I$  is not filtered, we have applied a moving average smoothing to the estimation  $\hat{I}$  (see Figure 3).

Finally, these simulations show that the method allows to reconstruct the parameter values in few days in a quite accurate manner. The estimation of the size of the infected population  $I$  allows then the use of the model for predictions of the epidemics.

## 8 Conclusion

This work shows that although the identifiability of the SIR-Q models has singularity points where measured variables are null, it is possible to design an observer with exponential decay of the estimation error during the first stage of the epidemics, and recover parameters in few days. Further investigations will concern real data of COVID epidemics provided by various territories.

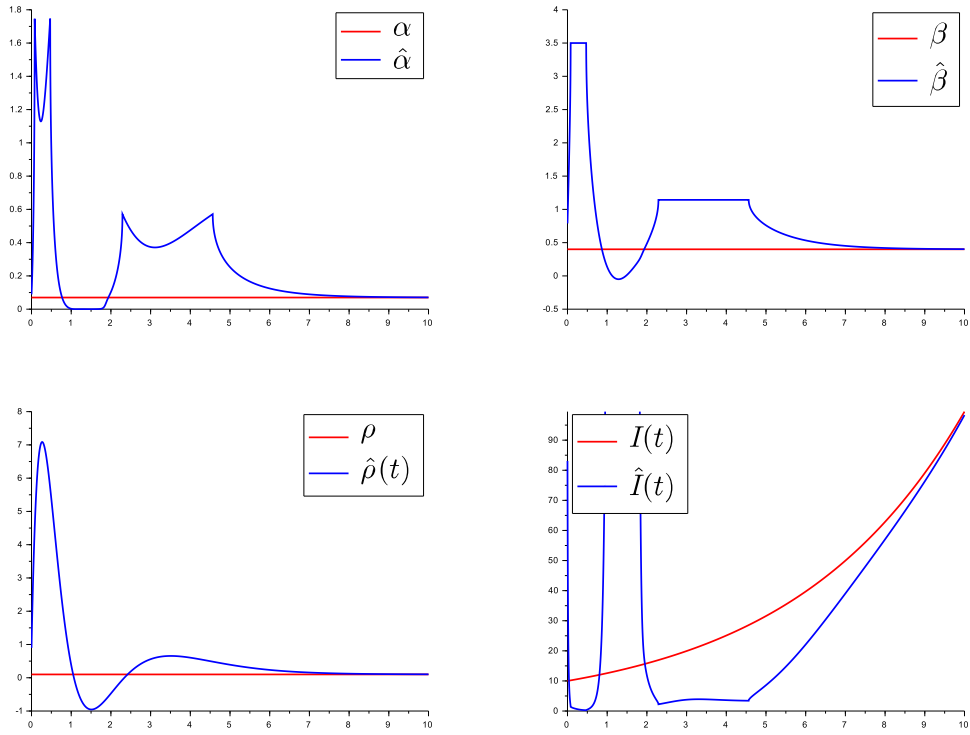


Figure 1: Observer simulation without measurement noise

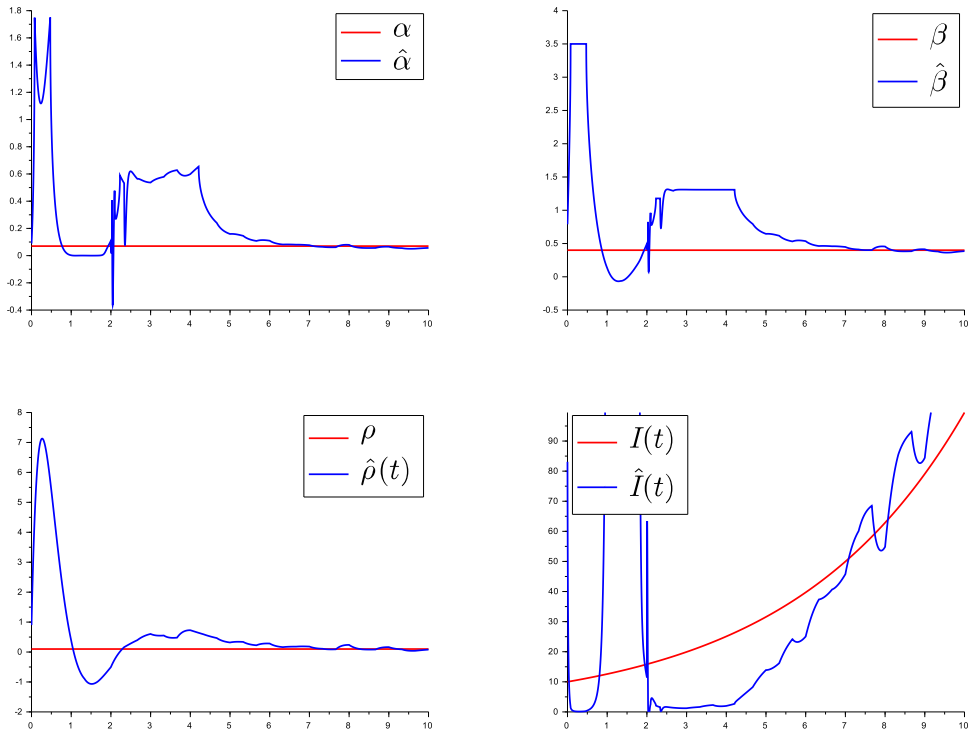


Figure 2: Observer simulation with measurement noise

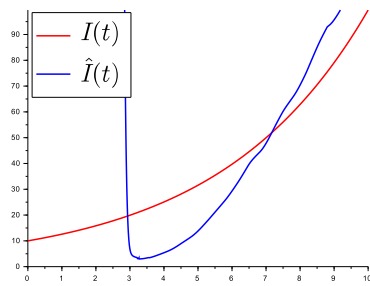


Figure 3: Smoothed estimation of the infected population in presence of measurement noise

## References

- [1] M. C. Eisenberg, S. L. Robertson, and J. H. Tien. Identifiability and estimation of multiple transmission pathways in Cholera and waterborne disease. *J Theor Biol*, 324:84–102, May 2013.
- [2] N. D. Evans, L. J. White, M. J. Chapman, K. R. Godfrey, and M. J. Chappell. The structural identifiability of the susceptible infected recovered model with seasonal forcing. *Math. Biosci.*, 194(2):175–197, 2005.
- [3] J. P. Gauthier and I. A. K. Kupka. Observability for systems with more outputs than inputs and asymptotic observers. *Math. Z.*, 223(1):47–78, 1996.
- [4] H. Hethcote, M. Zhien, and L. Shengbing. Effects of quarantine in six endemic models for infectious diseases. *Math. Biosci.*, 180:141–160, 2002. John A. Jacquez memorial volume.
- [5] M. Y. Li. *An introduction to mathematical modeling of infectious diseases*, volume 2 of *Mathematics of Planet Earth*. Springer, Cham, 2018.
- [6] H. Miao, X. Xia, A. S. Perelson, and H. Wu. On identifiability of nonlinear ODE models and applications in viral dynamics. *SIAM Rev.*, 53(1):3–39, 2011.
- [7] M. Nuño, C. Castillo-Chavez, Z. Feng, and M. Martcheva. Mathematical models of influenza: the role of cross-immunity, quarantine and age-structure. In *Mathematical epidemiology*, volume 1945 of *Lecture Notes in Math.*, pages 349–364. Springer, Berlin, 2008.
- [8] A. Perasso, B. Laroche, Y. Chitour, and S. Touzeau. Identifiability analysis of an epidemiological model in a structured population. *J. Math. Anal. Appl.*, 374(1):154–165, 2011.
- [9] C. Roda, B. Varughese, D. Han, and M. Li. Why is it difficult to accurately predict the covid-19 epidemic? *Infect. Dis. Model.*, (5).
- [10] M. P. Saccomani. An effective automatic procedure for testing parameter identifiability of HIV/AIDS models. *Bull. Math. Biol.*, 73(8):1734–1753, 2011.
- [11] N. Tuncer, H. Gulbudak, V. L. Cannataro, and M. Martcheva. Structural and practical identifiability issues of immunoparasitological vector-host models with application to Rift Valley Fever. *Bull. Math. Biol.*, 78(9):1796–1827, 2016.
- [12] E. Walter and L. Pronzato. *Identification of parametric models*. Communications and Control Engineering Series. Springer-Verlag, Berlin; Masson, Paris, 1997. From experimental data, Translated from the 1994 French original and revised by the authors, with the help of John Norton.
- [13] X. Xia and C. H. Moog. Identifiability of nonlinear systems with application to HIV/AIDS models. *IEEE Trans. Automat. Control*, 48(2):330–336, 2003.