**Noname manuscript No.**
(will be inserted by the editor)

# Optimal Geometric Control Applied to the Protein Misfolding Cyclic Amplification Process

**Monique Chyba · Jean-Michel Coron · Pierre Gabriel · Alain Jacquemard · Geoff Patterson · Gautier Picot · Peipei Shang**

Received: date / Accepted: date

**Abstract** Protein Misfolding Cyclic Amplification is a procedure used to accelerate the prion-replication process involved during the incubation period of transmissible spongiform encephalopathies. This technique could be used to design an efficient diagnosis test detecting the abnormally-shaped protein responsible of the decease before the affected person is at an advanced stage of the illness. In this paper, we investigate the open problem to determine what is the optimal strategy

J.-M. Coron supported and M. Chyba, P. Shang partially supported by ERC advanced grant 266907 (CPDENL) of the 7th Research Framework Programme (FP7). M. Chyba, G. Patterson partially supported by the National Science Foundation (NSF) Division of Graduate Education, award #0841223 and the NSF Division of Mathematical Sciences, award #1109937. P. Shang partially supported by the National Natural Science Foundation of China (No. 11301387)

M. Chyba, G. Patterson, G. Picot
Mathematics Department
University of Hawai'i
Honolulu, HI 96822, USA
E-mail: chyba@hawaii.edu, gautier@math.hawaii.edu, gpatters.uh@gmail.com

J.-M. Coron
Université Pierre et Marie Curie-Paris 6
UMR 7598 Laboratoire Jacques-Louis Lions
75005 Paris, France
E-mail: coron@ann.jussieu.fr

P. Gabriel
Université de Versailles Saint-Quentin
Laboratoire de Mathématiques de Versailles, CNRS UMR 8100
45 Avenue des Etats-Unis, 78035 Versailles, France
E-mail: pierre.gabriel@uvsq.fr

A. Jacquemard
Institut de Mathématiques de Bourgogne UMR 5584
Université de Bourgogne
9 avenue Alain Savary, 21078 Dijon, France
E-mail: Alain.Jacquemard@u-bourgogne.fr

P. Shang
Department of Mathematics, Tongji University
Shanghai 200092, China
E-mail: peipeishang@hotmail.com

to produce maximum replication in a fixed time. Primarily, we expand on prior attempt to answer this question in general, and provide results under some specific assumptions.

**Keywords** Transmissible spongiform encephalopathies · Nucleated polymerization · Optimal Control · Singular Extremals

**Mathematics Subject Classification (2000)** MSC 93C95 · MSC 70Q05 · MSC 70E60

## 1 Introduction

Transmissible spongiform encephalopathies (TSEs) are a group of progressive diseases affecting the brains and nervous systems of many animals, including humans. The disorders cause impairment of brain function, including memory changes, personality changes and problems with movement that worsen over time. Some examples include scrapie in sheep, bovine spongiform encephalopathies in cows ("mad cow disease"), and Creutzfeld-Jakob disease in humans [11]. Transmission of TSEs can occur when healthy animals consume tainted tissues from others with the disease. The mad cow disease epidemic occurred because cattles were being fed the processed remains of other cattle, a practice now banned in many countries. TSEs are characterized by their long incubation periods, lack of immune response, and invisibility to detection as viruses [30]. It is postulated that TSEs are invisible to detection as viruses because, in fact, they are not caused by viruses, but instead by abnormally shaped proteins – the so called, *prion* protein [10]. This hypothesis explains many of the features of the infectious agents of TSEs, except for their ability to replicate.

Prions do not contain DNA or RNA, which are the commonly accepted basis for replication. Ongoing research seeks to explain the replication mechanism for prions. The leading theory is nucleated polymerization, in which healthy prions are converted to infectious ones [21]. In this paper, we use standard notation for the normal prion $\mathrm{PrP}^{\mathrm{c}}$ (prion protein cellular) and the abnormal prion $\mathrm{PrP}^{\mathrm{Sc}}$ (prion protein scrapie). The biological processes associated with nucleated polymerization theory and the replication of $\mathrm{PrP}^{\mathrm{Sc}}$ polymers can be summarized as: lengthening (by addition of $\mathrm{PrP}^{\mathrm{c}}$ monomers); splitting (into two smaller polymer lengths); degradation (by metabolic processes). The polymers of $\mathrm{PrP}^{\mathrm{Sc}}$ can replicate by attaching units of $\mathrm{PrP}^{\mathrm{c}}$ in a string-like fashion to its ends (i.e. lengthening). After a $\mathrm{PrP}^{\mathrm{c}}$ monomer attaches to the $\mathrm{PrP}^{\mathrm{Sc}}$ polymer it is converted to the infectious form. When the $\mathrm{PrP}^{\mathrm{Sc}}$ polymer is long enough to wrap into a helical shape, it forms stabilizing bonds that constitute the polymer strings. These strings can grow to lengths of thousands of monomer units. The $\mathrm{PrP}^{\mathrm{Sc}}$ polymers can also split into two polymers of smaller lengths, each of which are then capable of further lengthening. If after splitting, a smaller polymer is below the threshold to maintain the stabilizing bonds, it degrades immediately into the normal $\mathrm{PrP}^{\mathrm{c}}$ monomers. It is the collection of long $\mathrm{PrP}^{\mathrm{Sc}}$ polymers which effectively interrupt normal nervous system processes.

Since incubation takes place over very long period of time, detecting the protein is often not possible until the affected person is at an advanced stage of a related disease. Protein Misfolding cyclic amplification is a technique to simulate

an accelerated replication process for prions in a laboratory environment [32]. This amplification process for instance permitted detection of abnormal prions in blood sample of affected hamsters [9]. The idea is to create a cyclic scheme that mimic polymerase chain reaction by alternating incubation phases to allow lengthening of the abnormal prion with sonication phases to break the polymers into smaller ones. In [31], the incubation phase is set to be more than 30 times the duration of the sonication phase. In control theory we refer to such strategy as bang-bang, the sonication pulse alternates between a minimum value (corresponding to no sonication) and a maximum value. There is an extensive number of papers published on the subject including[15,17,27], in which the authors discuss strategies to improve PCMA. Those include more sophisticated equipment, the use of different tissues for detection, variation of concentration of $PrP^c$ in the substitute, introduction of Teflon beads to increase the conversion of $PrP^c$ into $PrP^{Sc}$ , and increasing the number of cycles. In all those paper the procedure of the PCMA is always identical to the bang-bang strategy described above. However, bang-bang approaches are not necessarily the most efficient ones in particular to be applied to human blood samples and a more refine strategy is necessary. It is well known that the so-called singular arcs play a major role in optimal synthesis [4]. For instance, it has been shown that optimal control strategies contain singular arcs to provide, in some situations, viable options for chemoterapy treatments, see for instance [22–24].

It is an open problem to determine what is the optimal strategy to produce accelerate replication in a given time, and in this paper we expand on the first attempt made in [12] to answer this question in general. In particular, techniques form geometric control are used to study the singular flow in dimensions greater than two.

## 2 Statement of the problem as a Mayer Problem

In [26], the authors introduce a mathematical model of nucleated polymerization that is now the most widely used. In this paper, we consider a finite version of Masel's model that contains an infinite number of coupled ordinary differential equations. Note that there also exists a continuous version of this model, see [18]. These models capture the competition between the two processes of nucleated polymerization: lengthening and splitting. An in-depth study of the correlation between those two phenomena based on these models can be found in [16]. Here, we focus on a mathematical model to describe the protein misfolded cyclic amplification (PMCA) technique that was introduced in [16]. The model reflects the cyclic process that includes incubation phases, corresponding to the lengthening of the polymers, with sonication phases corresponding to the splitting of the $PrP^{Sc}$ . The two major assumptions for the model used here are first that there is an infinite population of $PrP^c$ throughout the experiment to induce the polymerization process and second that the growth and fragmentation constants of the polymers are not impacted by the sonication.

A detailed introduction to the model can be found in [12,16], we therefore restrict ourselves here to a brief introduction. The $n$-compartment approximation

of the continuous version of this model is given by

$$\dot{x}_i(t) + r(u(t))(\tau_i x_i(t) - \tau_{i-1}x_{i-1}(t)) = u(t)\left(2\sum_{j=i+1}^{n}\beta_j\kappa_{i,j}x_j(t) - \beta_i x_i(t)\right) \quad (1)$$

for $1 \leq i \leq n$, where $x_i(t)$ represents the density of polymer of size $i$ at a given time $t$, and the measurable bounded function $u(.)$ called the control represents the sonication intensity applied throughout the entire process. Practical implementation forces the control to be bounded: $u(t) \in [u_{min}, u_{max}]$ with $u_{max} \geq u_{min} > 0$, and the function $r(.)$ is decreasing to reflect the negative effect of ultrasound on the polymerization. The constants $\tau_i, \beta_i$ and $\kappa_i$ represents respectively the growth rate, the global fragmentation and the fragmentation kernel of polymers of size $i$.

Equation (1) can be seen as a (first order) size discretization of the partial differential equation

$$\partial_t f(t,y) + r(u(t))\partial_y\big(\tau(y)f(t,y)\big) = u(t)\Big(2\int_y^\infty \beta(z)\kappa(y,z)f(t,z)\,dz - \beta(y)f(t,y)\Big), \quad (2)$$

where $y > 0$ represents the length of the polymers. Most of our study will be made on the $n$-compartment approximation, but we will briefly come back to the continuous model (2) at the end of the paper.

System (1) can be rewritten as a bi-input control system,

$$\dot{x}(t) = \Big(u(t)F + v(t)G\Big)x(t), \qquad x_i(0) > 0 \quad (3)$$

where $v(t) = r(u(t))$, and $F, G$ represent respectively the fragmentation and growth matrices:

$$F := \begin{bmatrix} 0 & & & \\ -\beta_2 & (2\kappa_{ij}\beta_j)_{i<j} & & \\ & \ddots & & \\ 0 & & & \\ & & & -\beta_n \end{bmatrix}, G := \begin{bmatrix} -\tau_1 & & & \\ \tau_1 & -\tau_2 & 0 & \\ & \ddots & \ddots & \\ & 0 & \tau_{n-2} & -\tau_{n-1} \\ & & & \tau_{n-1} \quad 0 \end{bmatrix}, \quad (4)$$

with the following assumptions:

$$\tau_i > 0,\ \beta_{i+1} > 0 \text{ for all } i \in [1, n-1], \quad (5)$$

and

$$\sum_{i=1}^{j-1}\kappa_{ij} = 1, \qquad \sum_{i=1}^{j-1} i\kappa_{ij} = \frac{j}{2}, \quad (6)$$

which ensure that the breakage of a polymer produces two smaller ones with a half mean size.

The problem is to maximize the final density of $\mathrm{PrP}^{\mathrm{Sc}}$ represented by

$$\sum_{i=1}^{n} i x_i(T)$$

where $T$ is the fixed final time.

It is relevant, from the biological point of view, to suppose that the function $r$ is a decreasing convex positive function belonging to $\mathcal{C}^2([u_{\min}, u_{\max}])$. Thus, we assume that, for all $u \in [u_{min}, u_{max}]$,

$$r'(u) > 0, \tag{7}$$

$$r''(u) > 0, \tag{8}$$

$$r(u) - ur'(u) > 0. \tag{9}$$

The pair $(u(t), v(t))$ belongs by definition to the graph of the function $r$, denoted $\Omega$ and defined by

$$\Omega = \{(u, r(u)),\ u_{min} \leq u \leq u_{max}\}.$$

This set is not convex in $\mathbf{R}^2$, since $r$ is a strictly convex function. However, to guarantee the existence of an optimal control, we must consider a convex control set, see [25]. We consequently relax the problem by assuming that $(u(t), v(t))$ belongs to the convex hull of $\Omega$, denoted $H(\Omega)$. It can then be proved, see [12], that, for almost every $t \in [0, T]$, the optimal control $(u^*(t), v^*(t))$ belongs to the line $l$ that links $(u_{min}, r(u_{min}))$ to $(u_{max}, r(u_{max}))$ in the $uv-$plane. Therefore, in the sequel we assume $v = \xi + u\theta > 0$, where the negative slope $\theta$ and the positive $v$-intercept $\xi$ of the line $l$ are given by

$$\begin{aligned} \theta &= \frac{r(u_{\max}) - r(u_{\min})}{u_{\max} - u_{\min}} < 0 \\ \xi &= r(u_{\max}) - u_{\max}\frac{r(u_{\max}) - r(u_{\min})}{u_{\max} - u_{\min}} > 0 \end{aligned} \tag{10}$$

Using a time reparametrization, our system can be written as a bi-linear control system

$$\dot{x}(t) = Ax(t) + u(t)Bx(t) \tag{11}$$

where $A = G, B = F + \theta G$, $x_i(0) > 0$. To summarize, our goal is to solve the following problem

$$\dot{x}(t) = Gx(t) + u(t)(F + \theta G)x(t),\ \ x_i(0) = x_i^0 > 0 \tag{12}$$

$$\max_{u \in \mathcal{U}}\ \sum_{i=1}^{n} i x_i(T), \tag{13}$$

where $\mathcal{U}$ is the set of measurable functions that satisfy $u(t) \in [u_{min}, u_{max}]$ for almost every $t$. This is a standard Mayer problem in optimal control.

*Remark 1* Using that with our reparametrization we have $1 + u\theta > 0$, we conlude that system (11) is of the form $\dot{x}(t) = M_u(t)x(t)$, where $M_u = G + u(F + \theta G)$, with negative diagonal entries and non negative off diagonal entries. As a consequence the first orthant of the state space is positively invariant. In particular, if $x_i(0) > 0$ then $x_i(t) > 0$ for all $t \in [0, T]$.

The set of points where the drift $Gx$ and the control vector field $(F + \theta G)x$ are collinear plays an important role to analyze the dynamics of the system. The following lemma states that for our problem this set is empty on the first orthant.

**Lemma 1** *The vector fields $Gx$ and $(F+\theta G)x$ are linearly independent for every $x$ such that $x_i > 0$.*

*Proof* If $Gx$ and $(F+\theta G)x$ are collinear, there exists $\lambda \in \mathbb{R}\backslash\{0\}$ such that $\lambda Gx = (F+\theta G)x$ which is equivalent to $(\lambda-\theta)Gx = Fx$. Renaming $\lambda - \theta = \gamma$, we search for the collinear set of $F$ and $G$. Given the structure of the matrices $F$ and $G$ and the constraints on the coefficients we conclude that if $Fx$ and $Gx$ are collinear then $0 = \sum_{i=2}^{n} \beta_i x_i$. Since $x_i > 0, \beta_i > 0$ for all $i$, we reach a contradiction.

Notice that since $u_{\min} > 0$, the study of the uncontrolled system ($u = 0$) is irrelevant.

The current experimental procedure for PMCA consist of alternate phases of 30 min. incubation (corresponding to no sonication, $u = u_{min}$) and 40 sec. pulse of sonication set (constant pulse at 60% of max performance, $u = 0.6 * u_{max}$), see [31]. This procedure does not take into account that there exists a special value of the control that provides the most efficient increase in the number of polymers nor is based on an analytic analysis to determine the best duration for which the various values of the control should be applied. Using tools of geometric control, we can prove in particular that singular arcs provides optimal strategies with an increased performance. We recall here some results from [12] and expand on their analysis.

As in [12], for the relaxed optimal control problem, to a fixed pair of control $(u, v) \in H(\Omega)$ we associate the Perron eigenvalue (the single dominant eigenvalue) $\lambda_p(u, v)$ of the matrix $uF+vG$. Its existence is guaranteed by the Perron-Frobenius theorem [33] (the matrix is irreducible and has non-negative extra-diagonal values). View as a function $\lambda_p(u)$ of the control defined on a compact set, there exists a maximum value $\bar{\lambda}_p$ corresponding to a control that we denote $(\bar{u}, \bar{v})$, i.e. $\bar{\lambda}_p = \lambda(\bar{u}, \bar{v})$. In [12], it is proved that $(\bar{u}, \bar{v}) \in \Sigma$. Under our reparametrization, we have by construction that $\bar{v} = 1 + \bar{u}\theta$. To summarize, we associate to problem (12) the control $\bar{u}$ defined as producing the maximum of the function:

$$\begin{aligned} \lambda : [u_{\min}, u_{\max}] &\to \mathbb{R}^+\backslash\{0\} \\ u &\to \lambda_u. \end{aligned}$$

where $\lambda_u$ is the Perron eigenvalue of the matrix $M_u = G + u(F + \theta G)$ introduced in Remark 1.

For every $u \in [u_{\min}, u_{\max}]$, let's denote $\xi_u$ (seen as a column vector) and $\phi_u$ (seen as a row vector) the corresponding right and left eigenvectors such that

$$M_u \xi_u = \lambda_u \xi_u \tag{14}$$

$$\phi_u M_u = \lambda_u \phi_u \tag{15}$$

normalized by the constraints $||\xi_u||_1 := \sum_{i=1}^{n} |\xi_i| = 1$ and $\phi_u \xi_u = 1$.

A remarkable fact shown in the next section is that $\bar{u}$ is a so-called singular with corresponding singular trajectory $\xi_{\bar{u}}$.

## 3 Maximum Principle

The Pontryagin Maximum Principle, see [4, 29], gives necessary conditions for a solution of a control problem to be optimal. Let $u^*(.)$ be an optimal control for

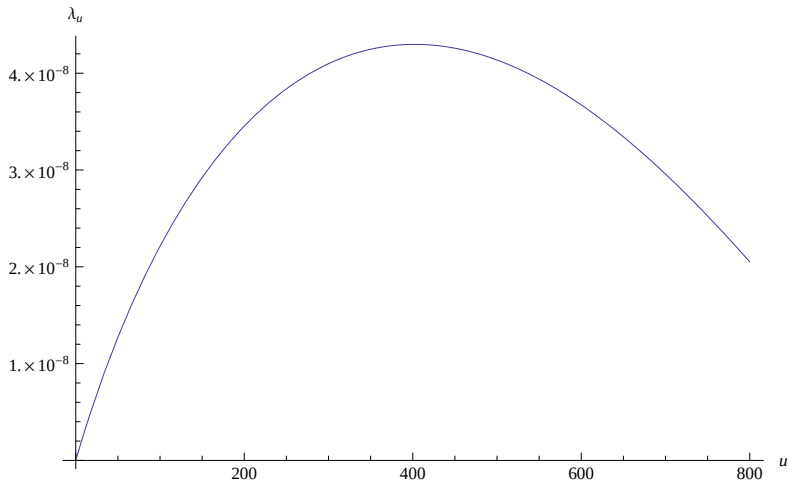



**Fig. 1** Graph of the mapping $u \to \lambda_u$, for $\tau_1 = 1.10^{-7}$, $\tau_2 = 3.10^{-7}$, $\beta_2 = 2.10^{-10}$, $\beta_3 = 3.10^{-10}$, $\theta = -1.10^{-3}$ over the interval $[u_{\min}, u_{\max}] = [1, 800]$. For this set of parameters, we find a maximum perron eigenvalue $\bar{\lambda}_p = 4.29787.10^{-8}$ corresponding to a control $\bar{u} = 402.265$.

Mayer problem (12) defined on $[0,T]$, and $x^*(.)$ be the corresponding trajectory. Then there exists an absolutely continuous row vector $p^* : [0,T] \longrightarrow \mathbb{R}^n \backslash \{0\}$ such that, for almost every $t \in [0,T]$

$$\dot{p^*}(t) = -p^*(t)\Big(G + u(t)(F + \theta G)\Big). \tag{16}$$

Moreover, if we introduce the Hamiltonian function

$$H(x,p,u) = pGx + up(F + \theta G)x, \tag{17}$$

the maximization condition

$$H(x^*(t), p^*(t), u^*(t)) = \max_{u \in [u_{\min}, u_{\max}]} H(x^*(t), p^*(t), u), \tag{18}$$

holds for almost every time $t \in [0,T]$ and the Hamiltonian is constant. Finally, the transversality condition implies that, up to a normalization,

$$p^*(T) = (1 \ \ldots \ n). \tag{19}$$

Similarly to Remark 1, we deduce from (19) that $p_i^*(t) > 0$ for all $t < T$, since $p_i(T) > 0$. Thus, both $x^*(t)$ and $p^*(t)$ are strictly positive for $t \in [0,T]$. A 3-tuple $(x(.), p(.), u(.))$ solution of (11), (16), (18) and (19) is called an *extremal* of the problem.

Introduce the function $\Phi(t) = p(t)(F + \theta G)x(t)$, then the maximization condition (18) implies the following. If $u^*(.)$ is an optimal control, we have

$$u^*(t) := \begin{cases} u_{min} & \text{if } \Phi(t) < 0 \\ u_{max} & \text{if } \Phi(t) > 0 \\ \in [u_{\min}, u_{\max}] & \text{if } \Phi(t) = 0. \end{cases}$$

The function $\Phi(.)$ is called the switching function and an isolated zero $t$ of the switching function is called a switching time. If $\Phi(t) = 0$ on a nonempty time interval, a further analysis must be performed to deduce the value of $u^*(t)$ (see below). A bang extremal defined on $[t_1, t_2]$, $t_2 > t_1$, corresponds to a constant maximum or minimum control, i.e. the sign of the switching function is constant over the entire interval (either strictly positive or strictly negative). An extremal is said to be singular on $[t_1, t_2]$, $t_2 > t_1$, if $\Phi(.)$ identical zero on that interval. The maximum principle implies that an optimal solution $x^*(.)$ is the projection of a concatenation of bang and singular extremals. The main difficulty is to analyze the structure of the optimal control, it is well-known that very complex situation such as the Füller Phenomenon can happen [20,35].

An important remark for our problem and direct consequence from the maximum principle is that along an extremal the inner product between the adjoint and the state variable is constant.

**Lemma 1** *Let $(x(.), p(.), u(.))$ be an extremal defined on $[0, T]$. We have that the quantity $p(t)x(t)$ is constant, in particular our cost $\sum_{i=1}^{n} ix_i(T)$ is equal to $p(0)x(0)$.*

*Proof* By direct differentiation of $p(t)x(t)$.

Lemma 1 asserts that the cost is completely determined at time 0 since $p(T)x(T) = p(0)x(0)$. Our problem is consequently equivalent to find an extremal $(x(.), p(.), u(.))$ which maximizes the quantity $p(0)x(0)$.

Next, we state some facts that will be useful to understand the optimal structure of the control, see [12] for a proof.

**Lemma 2** *Let $(x(.), p(.), u(.))$ be an extremal defined on $[0, T]$. The following holds:*

1. $p(T)Fx(T) = 0$,
2. $p(t)Gx(t) > 0$ *for every* $t \in [0, T]$, *and* $p(T)Gx(T) = \sum_{i=1}^{n-1} \tau_i x_i(T)$.

A direct consequence from Lemma 2 is the following.

**Corollary 1** *The Hamiltonian $H$ is strictly positive along an extremal $(x(.), p(.), u(.))$.*

*Proof* From Lemma 2, we obtain

$$H(x(T), p(T), u(T)) = (1 + \theta u(T))p(T)Gx(T) > 0.$$

Since the Hamiltonian is constant along an extremal, the result follows.

The following lemma provides a general result about the structure of the optimal arc at the end.

**Lemma 3** *The optimal strategy must finish by a bang arc $u_{min}$.*

*Proof* By lemma 2, we have $p(T)Fx(T) = 0$, therefore the switching function at the final time is given by $\Phi(T) = \theta p(T)Gx(T)$ which, by Lemma 2, is strictly negative. As a consequence of the maximization condition there exists an time interval $[\tau, T]$ with $\tau < T$ such that $u(t) = u_{min}$ on that interval.

### 3.1 Singular Arcs

To compute singular controls we proceed as follows. Along a singular arc, we have $\Phi(t) = 0$ on a nonzero time interval $[t_1, t_2]$ which is equivalent to

$$p(t)(F + \theta G)x(t) = \Phi(t) = 0 \tag{20}$$

on that same interval. Differentiating this equation, and using $[G, G] = 0$, we obtain

$$p(t)[G, F]x(t) = \dot{\Phi}(t) = 0, \tag{21}$$

where $[G, F] = GF - FG$. Differentiating (21) with respect to time, we get

$$p(t)[[G, F], G]x(t) + u(t)p(t)([[G, F], F] + \theta[[G, F], G])x(t) = \ddot{\Phi}(t) = 0 \tag{22}$$

almost everywhere on $[t_1, t_2]$. Outside the surface $S = \{(x, p);\ p([[G, F], F] + \theta[[G, F], G])x = 0\}$, the singular control is said to be of order two and is given by

$$u_{sing}(t) = -\frac{p(t)[[G, F], G]x(t)}{p(t)([[G, F], F] + \theta[[G, F], G])x(t)}. \tag{23}$$

We define the switching surface by $\Sigma = \{(x, p);\ p(F + \theta G)x = p[G, F]x = 0\}$. The generalized Legendre-Clebsh condition is a second order necessary condition for optimality in the Mayer problem [19]. We have the following.

**Lemma 4** *The generalized Legendre Clebsh condition* $\frac{\partial}{\partial u}\frac{d^2}{dt^2}\frac{\partial H}{\partial u} \leq 0$ *is equivalent to* $\frac{\partial}{\partial u}\ddot{\phi} \leq 0$ *for a single-input affine control system. It is satisfied for our problem if*

$$p(t)([[G, F], F] + \theta[[G, F], G])x(t) \leq 0 \tag{24}$$

*along the singular arc.*

In section 2, we introduce the function $\lambda(u)$ where $\lambda(u)$ is the Perron eigenvalue for our problem associated to a constant control $u$. Multiplying (14) on the left by $\phi_u$ and using $\phi_u \xi_u = 1$, we have that $\lambda(u) = \lambda_u = \phi_u M_u \xi_u$. By differentiating this last equation, we can show that

$$\lambda'(u) = \phi_u(F + \theta G)\xi_u. \tag{25}$$

Since $\Phi(t) = p(t)(F+\theta G)x(t)$, the above equation explicits the correlation between the critical points of the function $\lambda(.)$ and the singular extremals. Indeed, by definition $\bar{u}$ as defined in section 2 is a critical point of $\lambda(.)$: $\lambda'(\bar{u}) = 0$. Therefore, if we consider the trajectories

$$\bar{x}(t) = R\xi_{\bar{u}}e^{\lambda_{\bar{u}}t}, \qquad \bar{p}(t) = S\phi_{\bar{u}}e^{-\lambda_{\bar{u}}t}, \tag{26}$$

where $R, S > 0$ are two constants that satisfy $RS = \sum_{i=1}^{n} i x_i(T)$, they are clearly solutions of the equations (11), (16) and moreover, we have, for every $t$

$$\Phi(t) = \bar{p}(F+\theta G)\bar{x} = RS\phi_{\bar{u}}e^{-\lambda_{\bar{u}}t}(F+\theta G)\xi_{\bar{u}}e^{\lambda_{\bar{u}}t} = RS\phi_{\bar{u}}(F+\theta G)\xi_{\bar{u}} = RS\lambda'(\bar{u}) = 0. \tag{27}$$

Our result is summarized in the following theorem.

**Theorem 1** *The constant control* $\bar{u}$ *providing the maximum Perron eigenvalue is a singular control.*

We define $\bar{u}$ the Perron-singular control and $(\bar{x}, \bar{p})$ as defined in (26) the Perron singular extremal. It has been proved in [12] that in two dimensions, it is the only singular control. However the singular flow is much more complicated in higher dimension.

Remark that from (14) and $\phi_u \xi_u = 1$, the Perron-singular eigenvalue is given by

$$\lambda_{\bar{u}} = \phi_{\bar{u}} G \xi_{\bar{u}}. \tag{28}$$

**Proposition 1** *If a Perron-singular arc belongs to the optimal solution, then we have $H = \lambda_{\bar{u}} RS$ along the optimal trajectory. However, it cannot contain a Perron-singular arc if the parameters satisfy the following constraints:*

$$\frac{i}{\tau_i} > \frac{1 + u_{\min}\theta}{\lambda_{\bar{u}}} \qquad \text{for all } i = 1, \cdots, n-1. \tag{29}$$

*Proof* Along the Perron-singular extremal, the Hamiltonian is given by $H(\bar{x}, \bar{p}, \bar{u}) = \bar{p}G\bar{x} + \bar{u}\bar{p}(F + \theta G)\bar{x}$. However $\bar{p}(F + \theta G)\bar{x}$ is zero since it is a singular arc, therefore $H(\bar{x}, \bar{p}, \bar{u}) = \bar{p}G\bar{x}$. Since we have $\bar{p}M_{\bar{u}}\bar{x} = \bar{p}G\bar{x}$, equation (14) implies that $H(\bar{x}, \bar{p}, \bar{u}) = \lambda_{\bar{u}}\bar{p}\bar{x} = \lambda_{\bar{u}} RS$. Using the fact that $p(t)x(t)$ is constant we have $\lambda_{\bar{u}} RS = \lambda_{\bar{u}} p(T)x(T) = \lambda_{\bar{u}} \sum_{i=1}^{n} i x_i$. Since we know that the last arc of an optimal solution corresponds to a $u_{\min}$ control, if there is a singular-Perron arc the Hamiltonian satisfy

$$\lambda_{\bar{u}} \sum_{i=1}^{n} i x_i(T) = (1 + u_{\min}\theta) \sum_{i=1}^{n-1} \tau_i x_i(T). \tag{30}$$

The above equation can be written as $n\lambda_{\bar{u}} x_n(T) + \sum_{i=1}^{n-1} (\lambda_{\bar{u}} i - \tau_i(1 + u_{\min}\theta)) x_i(T) = 0$. Since $\lambda_u > 0$ and $x_i(T) > 0$ the result is proved.

## 4 Conjugate Points

In this section, we give the definition of the first conjugate time and provide algorithms of computation.

**Lemma 5** *The generalized Legendre Clebsh condition in Lemma 4 along the extremal $(\bar{x}(\cdot), \bar{p}(\cdot), \bar{u})$ defined by* (26) *is*

$$\phi_{\bar{u}}([[G, F], F] + \theta[[G, F], G])\xi_{\bar{u}} \leq 0, \tag{31}$$

*moreover,* (31) *can be reformulated as*

$$\sum_{i=2}^{n} (\lambda_i - \lambda_{\bar{u}})(\phi_i(F + \theta G)\xi_{\bar{u}})(\phi_{\bar{u}}(F + \theta G)X_i) \geq 0, \tag{32}$$

*where $(X_i, \phi_i, \lambda_i)$ is a basis of eigenvectors of $G + \bar{u}(F + \theta G)$ defined by*

$$(G + \bar{u}(F + \theta G))X_i = \lambda_i X_i, \quad \phi_i(G + \bar{u}(F + \theta G)) = \lambda_i \phi_i, \quad \phi_i X_j = \delta_{i,j}.$$

*Proof* Direct computation gives

$$\phi_{\bar{u}} G(F+\theta G)^2 \xi_{\bar{u}} = \phi_{\bar{u}} (F+\theta G)^2 G \xi_{\bar{u}}. \tag{33}$$

By (33), (31) can be reduced to

$$\begin{aligned} &\phi_{\bar{u}}([[G,F],F]+\theta[[G,F],G])\xi_{\bar{u}} \\ =&2\phi_{\bar{u}}(F+\theta G)[F,G]\xi_{\bar{u}} \leq 0. \end{aligned} \tag{34}$$

Condition (34) can also be formulated as

$$\phi_{\bar{u}}(F+\theta G)\Big((G+\bar{u}(F+\theta G))(F+\theta G)-(F+\theta G)(G+\bar{u}(F+\theta G))\Big)\xi_{\bar{u}} \geq 0,$$

which writes

$$\phi_{\bar{u}}(F+\theta G)\Big(G+\bar{u}(F+\theta G)\Big)(F+\theta G)\xi_{\bar{u}} - \lambda_{\bar{u}}\phi_{\bar{u}}(F+\theta G)^2\xi_{\bar{u}} \geq 0. \tag{35}$$

If we decompose the vector $(F+\theta G)\xi_{\bar{u}}$ on the basis $(X_i)$, we get

$$(F+\theta G)\xi_{\bar{u}} = \sum_{i=1}^{n} (\phi_i (F+\theta G)\xi_{\bar{u}}) X_i,$$

then (35) can be rewritten as

$$\begin{aligned} &\phi_{\bar{u}}(F+\theta G)\Big(G+\bar{u}(F+\theta G)\Big)(F+\theta G)\xi_{\bar{u}} - \lambda_{\bar{u}}\phi_{\bar{u}}(F+\theta G)^2\xi_{\bar{u}} \\ =&\sum_{i=2}^{n}(\lambda_i-\lambda_{\bar{u}})(\phi_i(F+\theta G)\xi_{\bar{u}})(\phi_{\bar{u}}(F+\theta G)X_i) \geq 0. \end{aligned} \tag{36}$$

*Remark 2* We remark here that (36) also appears in the variational analysis of the Floquet eigenvalue around $\bar{u}$ (see [8] for more details). If it is negative, then the existence of periodic controls providing a growth rate greater than $\lambda_{\bar{u}}$ is ensured.

Based on Lemma 5 and Remark 2, we have the following theorem (see [3], note that in our case, the Goh condition is empty.)

**Theorem 2** *There exists $\varepsilon > 0$ so that the trajectory $\bar{x}$ defined by* (26) *is locally optimal in $L^\infty$ topology on $[0,\varepsilon]$.*

For any $t \in [0,T]$, we define the following *end-point mapping* of system (12) as

$$\begin{aligned} E : \mathbb{R}^+ \times \mathcal{U} &\longrightarrow \mathbb{R}^n \\ (t,u) &\longrightarrow x(t,u), \end{aligned} \tag{37}$$

where $s \to x(s,u)$ is the trajectory solution of (12), associated to the control $u(.)$ such that $x_i(0,u) = x_i(0)$.

For every $t \in [0,T]$, denote by $Q_t$ the intrinsic second order derivative of the end-point mapping on $[0,t]$.

**Definition 1 (The first conjugate time)** Define the first conjugate time $t_c$ along a singular extremal $(x(\cdot),p(\cdot),u_{sing}(x,p))$ as the supremum of times $t$ so that $Q_t$ is positive definite.

From the definition of a first conjugate time, it is clear that a singular trajectory $x(\cdot)$ is locally optimal in $L^\infty$ topology on $[0, t_c)$.

Our goal in this paper is to determine the existence or not of a conjugate point along a Perron-singular extremal. Using formula (23), we introduce $H_r(x,p) = H_r(x,p,u_{sing}(x,p))$. Let the vector field $\mathbf{H}_r$ be defined by

$$\mathbf{H}_r = \Big(\frac{\partial H_r}{\partial p}, -\frac{\partial H_r}{\partial x}\Big)^T,$$

and let $z(t) := (x(t), p(t))$ be a trajectory of $\mathbf{H}_r$ defined on $[0,T]$, i.e., $\dot{z}(t) = \mathbf{H}_r(z(t))$, for every $t \in [0,T]$. The differential equation on $[0,T]$

$$\dot{\delta} z(t) = d\mathbf{H}_r(z(t))\delta z(t)$$

is called Jacobi equation along $z(\cdot)$, or variational system along $z(\cdot)$.

A Jacobi field $J(t)$ is a nontrivial solution of the Jacobi equation along $z(\cdot)$. It is said to be vertical at time $t$ whenever $d\pi(z(t)).J(t) = 0$, where $\pi$ is the canonical projection. In local coordinates, set $J(t) := (\delta x(t), \delta p(t))$, then $J(t)$ is vertical at the time $t$ whenever $\delta x(t) = 0$.

**Definition 2 (The first geometric conjugate time)** A time $t_c$ is said to be geometrically conjugate if there exists a Jacobi field that is vertical at times 0 and $t_c$.

From results in [3] and Lemma 5, we claim that the first geometric conjugate time coincides with the first conjugate time along the Perron-singular extremal $(\bar{x}(\cdot), \bar{p}(\cdot), \bar{u})$ defined by (26). The test to determine a conjugate point along the Perron-singular extremal is detailed in [3]. We recall it briefly. Consider the Jacobi fields solutions of (39) with initial data as a basis

$$(\delta x_1(0), \cdots, \delta x_n(0), \delta p_1(0), \cdots, \delta p_n(0))$$

satisfying the following conditions

$$\begin{aligned}
&\langle \delta p_i(0), p(0)\rangle = 0,\\
&\langle \delta p_i(0), (F+\theta G)x(0)\rangle + \langle \delta p_i(0), (F+\theta G)\delta x_i(0)\rangle = 0,\\
&\langle \delta p_i(0), [G, F+\theta G]x(0)\rangle + \langle \delta p_i(0), [G, F+\theta G]\delta x_i(0)\rangle = 0,\\
&\delta x_i(0) \in \mathbb{R}(F+\theta G)x(0),
\end{aligned}$$

where $x(0) = R\xi_{\bar{u}}$ and $p(0) = S\phi_{\bar{u}}$. We compute the $n-2$ associated Jacobi fields. We then compute numerically the conjugate time, i.e., times for which the following condition is satisfied

$$\text{rank } (\delta x_1(t), \cdots, \delta x_{n-2}(t), (F+\theta G)x(t)) \leqslant n-2. \tag{38}$$

For our problem, it is a straightforward calculation (but tedious) to show that We have

$$\dot{\delta z} = \begin{pmatrix} \dot{\delta x} \\ \dot{\delta p} \end{pmatrix} = d\mathbf{H}_r \begin{pmatrix} \delta x \\ \delta p \end{pmatrix} = \begin{pmatrix} \Delta_1 & e^{2\lambda_{\bar{u}} t}\Delta_2 \\ -e^{-2\lambda_{\bar{u}} t}\Delta_3 & -\Delta_4 \end{pmatrix} \begin{pmatrix} \delta x \\ \delta p \end{pmatrix}, \tag{39}$$

where

$$\Delta_1 = (G + \bar{u}(F + \theta G)) - \frac{(F + \theta G)\xi_{\bar{u}}\phi_{\bar{u}}([[G, F], G] + \bar{u}[[G, F], F + \theta G])}{\phi_{\bar{u}}[[G, F], F + \theta G]\xi_{\bar{u}}} \quad (40)$$
$$- \frac{([[G, F], G] + \bar{u}[[G, F], F + \theta G])\xi_{\bar{u}}\phi_{\bar{u}}(F + \theta G)}{\phi_{\bar{u}}[[G, F], F + \theta G]\xi_{\bar{u}}},$$

$$\Delta_2 = -\frac{(F + \theta G)\xi_{\bar{u}}\xi_{\bar{u}}^T([[G, F], G] + \bar{u}[[G, F], F + \theta G])^T}{\phi_{\bar{u}}[[G, F], F + \theta G]\xi_{\bar{u}}} \quad (41)$$
$$- \frac{([[G, F], G] + \bar{u}[[G, F], F + \theta G])\xi_{\bar{u}}\xi_{\bar{u}}^T(F + \theta G)^T}{\phi_{\bar{u}}[[G, F], F + \theta G]\xi_{\bar{u}}},$$

$$\Delta_3 = -\frac{(F + \theta G)^T\phi_{\bar{u}}^T\phi_{\bar{u}}([[G, F], G] + \bar{u}[[G, F], F + \theta G])}{\phi_{\bar{u}}[[G, F], F + \theta G]\xi_{\bar{u}}} \quad (42)$$
$$- \frac{([[G, F], G] + \bar{u}[[G, F], F + \theta G])^T\phi_{\bar{u}}^T\phi_{\bar{u}}(F + \theta G)}{\phi_{\bar{u}}[[G, F], F + \theta G]\xi_{\bar{u}}},$$

$$\Delta_4 = \Delta_1^T. \quad (43)$$

Using the change of variables $\tilde{\delta z} = (\tilde{\delta x}, \tilde{\delta p}) := (e^{-\lambda_{\bar{u}}t}\delta x, e^{\lambda_{\bar{u}}t}\delta p)$, we obtain

$$\dot{\tilde{\delta z}} = \begin{pmatrix} \dot{\tilde{\delta x}} \\ \dot{\tilde{\delta p}} \end{pmatrix} = \begin{pmatrix} -\lambda_{\bar{u}}I + \Delta_1 & \Delta_2 \\ -\Delta_3 & \lambda_{\bar{u}}I - \Delta_4 \end{pmatrix} \begin{pmatrix} \tilde{\delta x} \\ \tilde{\delta p} \end{pmatrix}. \quad (44)$$

Note that numerical softwares such as Hampath are now used to determine the existence of conjugate points. Hampath is precisely based on an indirect method and the algorithm described above with the variational equation [6, 13].

## 5 Two dimensional Case

The two dimensional case corresponds to $n = 2$, and the optimal control problem is

$$\begin{cases} \dot{x}(t) = Gx(t) + u(t)\Big(F + \theta G\Big)x(t), \\ \max\limits_{u \in \mathcal{U}} \; x_1(T) + 2x_2(T), \\ x_1(0) > 0, x_2(0) > 0 \end{cases} \quad (45)$$

where $\mathcal{U}$ is the set of measurable functions that satisfy $u(t) \in [u_{min}, u_{max}]$ almost everywhere. In this case, the matrices $F$ and $G$ are given by

$$F = \begin{bmatrix} 0 & 2\beta \\ 0 & -\beta \end{bmatrix}, \; G = \begin{bmatrix} -\tau & 0 \\ \tau & 0 \end{bmatrix} \quad (46)$$

with $\beta > 0$, $\tau > 0$. In [12], the authors provide an optimal synthesis, and we summarize their results here. In dimension two, it can be shown that the only singular control is the one determined as the critical point of the application $\lambda$ defined in section 2:

$$\lambda : [u_{\min}, u_{\max}] \to \mathbb{R}^+ \backslash \{0\}$$
$$u \to \lambda_u.$$

In other words, a control $\bar{u}$ is singular if and only if it maximizes the Perron-eigenvalue of the matrix $G + u(F + \theta G)$ over the set of $u \in [u_{\min}, u_{\max}]$. Such control $\bar{u}$ is constant and its expression can be computed explicitly in terms of the parameters $\tau$ and $\beta$. It follows that a singular arc belongs to the line defined in the $x_1, x_2$-plane by the right Perron eigenvector $\bar{\xi}$ associated to $\lambda_{\bar{u}}$. These observations lead to the optimal synthesis of the problem depicted in Fig. 2. In [12], it is proved that for a final time $T$ sufficiently large, the unique optimal control $u^*$ is piecewise constant of the form $u_{bang}u_{sing}u_{min}$. More precisely, we have that the positive quadrant of the $x_1, x_2$-plane is partitioned into two regions by the line $l$ generated by $\bar{\xi}$. If the initial condition $x(0)$ belongs to the region above the line the optimal control starts with a sonication impulse $u_{max}$ until the trajectory reaches the line $l$. If $x(0)$ is in the region below the line, the optimal process start with an incubation period associated to a control $u_{min}$ until the trajectory reaches the line $l$. Once the trajectory is such that $(x_1(t), x_2(t))$ is proportional to $\bar{\xi}$ the control switch to the singular control $\bar{u}$ and stay on this arc as long as possible. The value $\bar{u}$ maximizing the Perron eigenvalue guarantees an exponential growth of maximal rate $\lambda_{\bar{u}}$ of the quantity $x_1 + 2x_2$ which is to be maximized. Such a stategy is said to exhibit the so-called turnpike properties, see [34] for instance. By Lemma 3, the last arc must be an incubation period corresponding to a control $u_{min}$ to meet the transversality condition. Notice that the duration of the last arc is function of the parameters $\tau$ and $\beta$ and is fixed for all optimal solutions.

On Fig. 2, we represent an example of an optimal solution (in blue) for a given set of parameters. The optimal solution is compared to the one following a PMCA procedure. Table 1 displays the performance comparison for the set of parameters used in Fig. 2. A spectacular improvement of the order of a $10^6$ magnitude can be observed between the optimal solution and the experimental one over a 48 hours duration period. In 48 hours, the experimental trajectory yields $J(T) = 2.824085876741727e + 025$. In only 39.7068 hours, the optimal trajectory achieves the same $J$ value, and in 48 hours the optimal trajectory gives $J(T) = 9.076925755681900e + 031$.

| Control regime | $J(T)$ |
|---|---|
| $u^*$ | $9.0769 \cdot 10^{31}$ |
| $u_{exp}$ | $2.8241 \cdot 10^{25}$ |
| $u_{min}$ | $6.2635 \cdot 10^{23}$ |
| $u_{max}$ | $4.3950 \cdot 10^{8}$ |

**Table 1** The final costs for the two-dimensional example trajectories shown in Figure 2. Also, for comparison, the final costs for trajectories using the same parameters but with the control frozen to $u_{min}$ and $u_{max}$ are shown. For this set of parameters, the optimal control provides over a $10^6$ magnitude improvement of the final cost when compared to the traditional experimental setting.

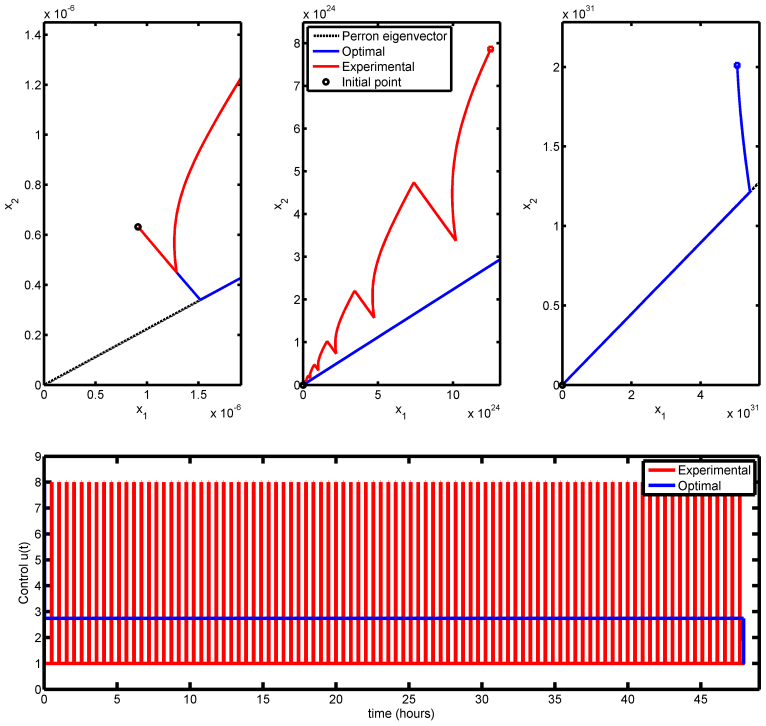



**Fig. 2** A two-dimensional example illustrating several properties of the optimal synthesis, for a specific set of parameters ($\theta = -0.1$, $\tau = 10^{-3}$, $\beta = 10^{-3}$, $T = 48$ hours, $[u_{min}, u_{max}] = [1,\ 8]$). The plot on the left first shows the partition of the state space by the right-eigenvector $\bar{\xi}$ in the $x_1$-$x_2$ plane. The area above $\bar{\xi}$ corresponds to initial points $x(0)$ such that the associated optimal control is of the form $u_{max}$-$u_{sing}$-$u_{min}$. Similarly, the area below corresponds to points $x(0)$ such that the associated optimal control is of the form $u_{min}$-$u_{sing}$-$u_{min}$. The top three plots are actually the same plot viewed at different zoom levels. A complete optimal trajectory is displayed in blue from an arbitrary initial point $x(0) = (0.9133 \cdot 10^{-6}, 0.6323 \cdot 10^{-6})$. It is defined on the time interval $[0, 48]$ hours and corresponds to the control $u^*(t) = \{8 : 0 \leq t \leq 83.3 \text{ sec}, \quad 2.74 : 83.3 \text{ sec} < t \leq 47.93 \text{ hrs}, \quad 1 : 47.93 \text{ hrs} < t \leq 48 \text{ hrs}\}$. For comparison, in red we display the experimental trajectory defined over the same time interval, which corresponds to a $u_{min}$ control with 94 $u_{max}$ pulses. The duration of the $u_{max}$ pulses are 45 seconds, and the $u_{min}$ portions are 30 minutes each. The bottom plot shows the corresponding controls over time.

## 6 The 3D case

In dimension $n = 3$, the optimal control system can be written under the form:

$$\begin{cases} \dot{x}(t) = Gx(t) + u(t)(F + \theta G)x(t), \quad x_i(0) > 0 \\ \max\limits_{u \in \mathcal{U}} \ x_1(T) + 2x_2(T) + 3x_3(T) \end{cases}$$

where $\theta < 0$, and $G$ and $F$ are the constant matrices

$$G = \begin{bmatrix} -\tau_1 & 0 & 0 \\ \tau_1 & -\tau_2 & 0 \\ 0 & \tau_2 & 0 \end{bmatrix}, \ F = \begin{bmatrix} 0 & 2\beta_2 & \beta_3 \\ 0 & -\beta_2 & \beta_3 \\ 0 & 0 & -\beta_3 \end{bmatrix}$$

$\tau_i > 0, \beta_i > 0$. We have

$$M_u = G + u(F + \theta G) = \begin{bmatrix} a & 2d & c \\ -a & -d-b & c \\ 0 & b & -c \end{bmatrix}, \tag{47}$$

where $a = -\tau_1(1+u\theta)$, $b = \tau_2(1+u\theta)$, $c = u\beta_3$ and $d = u\beta_2$.

We define $\alpha = a - b - c - d$ and $\gamma = ab + ac - ad - cd$. Notice that $\alpha < 0$. We have that the characteristic polynomial of $M_u$ is given by

$$p(\lambda) = \lambda^3 - \alpha\lambda^2 - \gamma\lambda + ac(b+d). \tag{48}$$

There exists at least one positive real value since the coefficient in front of $\lambda^3$ is positive and the $y-$intercept is negative. By differentiating the characteristic polynomial, we see the critical points are $\lambda = \frac{2\alpha}{6} \pm \frac{\sqrt{4\alpha^2 - 12\gamma}}{6}$ which is smaller than 0. Therefore, for each fixed value of $u$ there is only one real positive eigenvalue which is given by:

$$\lambda_u = \frac{1}{3}\big(\alpha - [\frac{1}{2}(\Delta - 4(\alpha^2+3\gamma)^2)]^{1/3} - [\frac{1}{2}(\Delta + 4(\alpha^2+3\gamma)^2)]^{1/3}\big) \tag{49}$$

where $\Delta = -2\alpha^3 - 9\alpha\gamma + 27ac(b+d)$. To find the Perron Eigenvalue we must differentiate $\lambda_u$ with respect to $u$ and compute the zero of that polynomial that belongs between $[u_{\min}, u_{\max}]$.

Using the fact that $\dot{p} = -pM_u$, we obtain the adjoint system:

$$\dot{p}_1 = -\tau_1(1+\theta u)(p_2 - p_1) \tag{50}$$

$$\dot{p}_2 = -\beta_2 u(2p_1 - p_2) - \tau_2(1+\theta u)(p_3 - p_2) \tag{51}$$

$$\dot{p}_3 = -\beta_3 u(p_1 + p_2 - p_3) \tag{52}$$

where $p = (p_1, p_2, p_3)$ represents the adding vector, and the transversality condition imposes $(p_1(T), p_2(T), p_3(T)) = (1, 2, 3)$.

### 6.1 Singular Flow

By theorem 1, we know that $\bar{u}$ is a singular control. But contrary to the case in dimension two the singular flow is more complex. In dimension three, the adjoint vector can be eliminated and the singular control is determined as a feedback from the state only. First let us introduce the following determinants:

$$D'(x) = \det\Big((F+\theta G)x, [G,F]x, [[G,F],G]x\Big) \tag{53}$$

$$D(x) = \det\Big((F+\theta G)x, [G,F]x, [[G,F],F]x\Big) \tag{54}$$

Notice that, since $G$ and $F + \theta G$ are constant matrices, both $D'(x)$ and $D(x)$ are homogeneous degree 3 polynomial functions in the variables $x_1$, $x_2$ and $x_3$. Moreover, $DD' < 0$. Indeed, since we must have $u_{sing} > 0$ it implies from equation $(1+u\theta)D' + uD = 0$ and the fact that $1 + u\theta > 0$, that the two determinants must have opposite sign.

**Proposition 2** *Singular trajectories of order 1 are solutions of* $\dot{x}(t) = X_s(x(t))$ *where* $X_s$ *is given by:*

$$X_s(x) = Gx - \frac{D'(x)}{D(x) + \theta D'(x)}(Fx + \theta Gx), \tag{55}$$

*with the feedback control* $u_{sing} = -\frac{D'(x)}{D(x)+\theta D'(x)}$.

As it was done in [5], the vector field $X_s(x)$ can be desingularized using the time reparametrization $ds = \frac{dt}{D(x(t))+\theta D'(x(t))}$ to produce a smooth vector field

$$X^r_{sing}(x) = D(x)Gx - D'(x)Fx. \tag{56}$$

Notice that $X^r_{sing}$ is homogeneous of order 4.

As in [5], the geometric classification of the singular flow is equivalent to the classification of the pair $(X_{sing}, D+\theta D')$. In the three dimensional case, the surface $D+\theta D' = 0$ plays an important role. Indeed, outside the surface $D' = 0$ it is where the control blow-up and is not admissible anymore, while the set $S' = \{x; D(x) = D'(x) = 0\}$ represents the points where the singular trajectory might possibly cross the surface $D + \theta D' = 0$ with an admissible control.

To compute the Lie Bracket as well as $D$ and $D'$, we introduce $\alpha = \tau_1 - \tau_2$, $\gamma = 2\tau_1 - \tau_2$, $\sigma = \beta_2 - \beta_3$, $\psi = \beta_2\tau_1$, $\rho = \tau_1 + \tau_2$ and $\mu = 4\beta_2\beta_3 - \beta_3^2 - 2\beta_2^2$. Then, the Lie Brackets are given by:

$$[G, F] = \begin{pmatrix} -2\beta_2\tau_1 & -2\beta_2\alpha - \beta_3\tau_2 & -\beta_3\tau_1 \\ \beta_2\tau_1 & 2\beta_2\tau_1 - \beta_3\tau_2 & \beta_3\alpha \\ 0 & -\sigma\tau_2 & \beta_3\tau_2 \end{pmatrix} \tag{57}$$

$$[[G, F], G] = \begin{pmatrix} -\tau_1\left(2\beta_2\alpha + \beta_3\tau_2\right) - \left(2\beta_2\alpha^2 + \beta_3\gamma\tau_2\right) & -\beta_3\tau_1^2 \\ \tau_1\left(3\tau_1\beta_2 + \tau_2\sigma\right) & \gamma\beta_3\tau_2 + 2\beta_2\tau_1\alpha \quad \beta_3\left(\tau_1^2 + \tau_1\tau_2 - \tau_2^2\right) \\ \left(\beta_3 - 2\beta_2\right)\tau_1\tau_2 & -\tau_2\left(\beta_2\gamma - \beta_3\tau_2\right) \quad -\beta_3\alpha\tau_2 \end{pmatrix} \tag{58}$$

$$[[G, F], F] = \begin{pmatrix} -\psi\beta_2 & \tau_2\mu - 6\psi\beta_2 & \beta_3\left(\beta_3\left(\tau_1 - 2\tau_2\right) + \beta_2\left(-6\tau_1 + 4\tau_2\right)\right) \\ \psi\beta_2 & 2\psi\beta_2 + \sigma\beta_3\tau_2 & -\beta_3\left(\beta_2\left(-4\tau_1 + \tau_2\right) + \beta_3\rho\right) \\ 0 & \sigma^2\tau_2 & -\beta_3\sigma\tau_2 \end{pmatrix} \tag{59}$$

The determinants $D$ and $D'$ are homogenous of order 3, the coefficients for the polynomials $D$ and $D'$ are given below.

Coefficients of the determinant $D = \sum_{i+j+k=3} D_{ijk}x_1^i x_2^j x_3^k$:

$$\begin{aligned}
D_{300} &= 0 \\
D_{030} &= \beta_2\tau_2\left(2\theta\beta_3^2\tau_1\tau_2 + \beta_2\beta_3\left(\beta_3\gamma + \theta\tau_2\left(-8\tau_1 + \tau_2\right)\right) + \beta_2^2\left(8\theta\tau_1^2 - \beta_3\gamma\right)\right) \\
D_{003} &= -\beta_2\beta_3^3\gamma\left(\tau_1 + 2\tau_2\right) \\
D_{210} &= \theta\beta_2\left(2\beta_2^2 - 3\beta_2\beta_3 + \beta_3^2\right)\tau_1^2\tau_2 \\
D_{201} &= \theta\beta_2\beta_3\left(-2\beta_2 + \beta_3\right)\tau_1^2\tau_2 \\
D_{111} &= -\beta_3\tau_1(4\beta_2^3\tau_1 + \beta_2\left(-2\beta_2^2 + \theta\left(3\beta_2 + \beta_3\right)\tau_1\right)\tau_2 \\
&\quad + \theta\left(2\beta_2^2 - 7\beta_2\beta_3 + 2\beta_3^2\right)\tau_2^2) \\
D_{120} &= 2\theta\psi\left(2\beta_2 - \beta_3\right)\tau_2\left(2\psi - \beta_3\tau_2\right)
\end{aligned}$$

$$
\begin{aligned}
D_{102} &= -\beta_3^2\tau_1\left(-\sigma\psi+\beta_2\left(2\beta_2+\beta_3+2\theta\tau_1\right)\tau_2-2\theta\sigma\tau_2^2\right)\\
D_{012} &= \beta_3^2\left(-8\psi^2+\psi\left(-\sigma+2\theta\tau_1\right)\tau_2-2\left(\beta_2\left(\beta_2-2\beta_3\right)+\theta\sigma\tau_1\right)\tau_2^2\right)\\
D_{021} &= \beta_3\Big(2\theta\beta_3^2\tau_1\tau_2^2+\beta_2^3\tau_1\left(-8\tau_1+7\tau_2\right)+2\beta_2\beta_3\tau_2\left(-2\theta\tau_1\tau_2+\beta_3\rho\right)\\
&\quad -\beta_2^2\tau_2\left(\theta\tau_1\left(-8\tau_1+\tau_2\right)+\beta_3\left(\tau_1+4\tau_2\right)\right)\Big)
\end{aligned}
$$

Coefficients of the determinant $D'=\sum_{i+j+k=3} D'_{ijk}x_1^i x_2^j x_3^k$:

$$
\begin{aligned}
D'_{300} &= \theta\beta_2\left(-2\beta_2+\beta_3\right)\tau_1^3\tau_2\\
D'_{030} &= -2\tau_2(\psi^2\left(\beta_2+\beta_3\right)+\psi\left(\beta_2^2-5\beta_2\beta_3+\beta_3^2-\theta\sigma\tau_1\right)\tau_2\\
&\quad +\left(\beta_2\beta_3^2+\theta\sigma^2\tau_1\right)\tau_2^2)\\
D'_{003} &= -\beta_3^3\tau_1\tau_2\gamma\\
D'_{210} &= \theta\beta_2\tau_1^2\tau_2\left(\beta_3\rho+\beta_2\left(-7\tau_1+2\tau_2\right)\right)\\
D'_{201} &= \beta_3\tau_1^2\left(4\beta_2^2\tau_1-2\beta_2^2\tau_2+\theta\left(2\beta_2-\beta_3\right)\tau_2^2\right)\\
D'_{111} &= \beta_3\tau_1(2\theta\beta_3\tau_2^3+\beta_2^2\left(4\tau_1^2-15\tau_1\tau_2+2\tau_2^2\right)+\beta_2\tau_2\left(3\beta_3\rho-2\theta\left(\tau_1^2+\tau_2^2\right)\right))\\
D'_{120} &= \tau_1\tau_2(2\theta\beta_3^2\tau_2^2-\theta\beta_2\beta_3\tau_2\left(5\tau_1+3\tau_2\right)\\
&\quad +\beta_2^2\left(-2\tau_1\left(\beta_3+2\theta\tau_1\right)+\left(\beta_3+8\theta\tau_1\right)\tau_2\right))\\
D'_{102} &= \beta_2\beta_3^2\tau_1\left(\tau_1-2\tau_2\right)\gamma\\
D'_{012} &= -2\beta_2\beta_3^2\tau_2\gamma^2\\
D'_{021} &= -\beta_3\tau_2\left(9\beta_2^2\tau_1\alpha+\theta\beta_3\tau_1\tau_2\rho+\beta_2\left(\beta_3\left(\tau_1-2\tau_2\right)^2-\theta\tau_1\tau_2\rho\right)\right)
\end{aligned}
$$

In this paper, we focus on a specific case: $\tau$ is constant and $\beta$ comes from a linear function [26]. This is equivalent to $\tau_1=\tau_2$ and $\beta_2=2\beta, \beta_3=3\beta$. Using a time reparametrization we can assume that $\tau=1$. We have:

$$
G=\begin{pmatrix}-1&0&0\\1&-1&0\\0&1&0\end{pmatrix},\quad F=\beta\begin{pmatrix}0&4&3\\0&-2&3\\0&0&-3\end{pmatrix} \tag{60}
$$

**Proposition 3** *The Lie brackets are given by:*

$$
[G,F]=\beta\begin{pmatrix}-4&-3&-3\\2&1&0\\0&1&3\end{pmatrix}, \tag{61}
$$

$$
[[G,F],F]=\beta^2\begin{pmatrix}-8&-17&-21\\4&5&0\\0&1&3\end{pmatrix},\ [[G,F],G]=\beta\begin{pmatrix}-3&-3&-3\\5&3&3\\-1&1&0\end{pmatrix} \tag{62}
$$

A straightforward computation shows that we have the following coefficients for $D$ and $D'$.

$$\begin{array}{ll} D_{300} = 0 & D'_{300} = -2\beta^2\theta \\ D_{030} = 4\beta^3(3\beta + 4\theta) & D'_{030} = -2\beta^2(4\beta + 3\theta) \\ D_{003} = -162\beta^4 & D'_{003} = -27\beta^3 \\ D_{210} = -2\beta^3\theta & D'_{210} = -8\beta^2\theta \\ D_{201} = -6\beta^3\theta & D'_{201} = 3\beta^2(8\beta + \theta) \\ D_{111} = -6\beta^3(8\beta + \theta) & D'_{111} = -6\beta^2\theta \\ D_{120} = 4\beta^3\theta & D'_{120} = -2\beta^2(6\beta + 7\theta) \\ D_{102} = -18\beta^3(8\beta + 3\theta) & D'_{102} = -18\beta^3 \\ D_{012} = 18\beta^3(-7\beta + 3\theta) & D'_{012} = -36\beta^3 \\ D_{021} = 6\beta^3(2\beta + 11\theta) & D'_{021} = -6\beta^2(3\beta + \theta) \end{array}$$

It is an easy verification that

$$D = 2(x_2 + 3x_3)\hat{D}, \tag{63}$$

with $\hat{D} = \theta x_1^2 + (24\beta x_3 - 2\theta x_2 + 9\theta x_3)x_1 + 12\beta x_3 x_2 - 6\beta x_2^2 + 27\beta x_3^2 - 8\theta x_2^2 - 9\theta x_2 x_3$ and

$$\begin{aligned} D' = \beta^2[&2\theta x_1^3 + (8\theta x_2 - (24\beta + 3\tau)x_3)x_1^2 \\ &+((14\theta + 12\beta)x_2^2 + 6\theta x_2 x_3 + 18\beta x_3^2)x_1 \\ &+(6\theta + 8\beta)x_2^3 + (6\theta + 18\beta)x_3 x_2^2] + 36\beta x_3^2 x_1 x_2 + 27\beta x_3^3 \end{aligned} \tag{64}$$

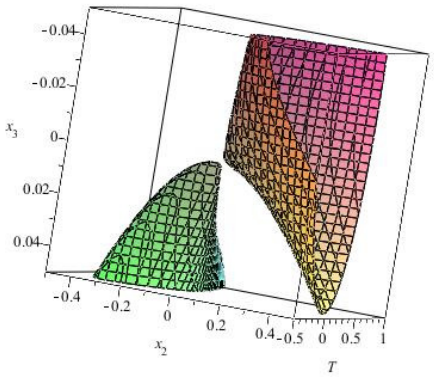
*6.1.1 The set $S = \{x; D(x) + \theta D'(x) = 0\}$*

It is important to determine the states $x \in S$ such that $D(x) \neq 0$ (and therefore $D'(x) \neq 0$) since they represent the states where the control explodes which can lead that a saturation phenomenon in the neighborhood since our control is bounded.

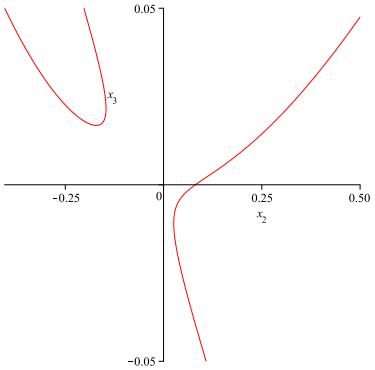
We can assume by homogeneity that $x_1 = 1$. To study the singularities of the set $S$ we proceed as follows. First, we divide $D+\theta D' = 0$ by $\beta^2$, then we evaluate at $x_1 = 1$. This produce an homogeneous of degree 3 polynomial, that we denote by $Pol_{D+\theta D'}$, in $x_2, x_3$ that is also homogeneous of degree 2 in $\beta, \theta$ (we will set $\beta = 1$ for the figures). To study the singularities, a Gröbner basis with graded reverse lex order on $(\theta, \beta, x_2, x_3)$ is computed for $\{Pol_{D+\theta D'}, \frac{\partial Pol_{D+\theta D'}}{\partial x_2}, \frac{\partial Pol_{D+\theta D'}}{\partial x_3}\}$. The basis contains 37 polynomials. The last polynomial in the basis just involves the indeterminates $\theta$ and $\beta$. Furthermore, it is factorizable. It then follows that there are no singularities in the positive quadrant (we omit the lengthly but straightforward calculations here).

**Proposition 4** *The set $S = \{x; D(x) + \theta D'(x) = 0\} \cap \{x_1 = 1\}$ does not contain any singularity in the positive quadrant.*

This set is represented on Fig. 3 for various values of $\theta$ with a curve. It unfolds on the 3-dimensional state space $x$ as an homogeneous cone, with a vertex at the origin, whose generator is the curve corresponding to $x_1 = 1$.




(a) Here $\theta$ varies between -0.5 and 1. Both the negative and positive quadrants have been represented to observe that there is indeed only one singularity at $x_2 = x_3 = 0$.




(b) $D + \theta D' = 0$ for $\theta = -0.1$

**Fig. 3** The set $S = \{x; D(x) + \theta D'(x) = 0\}$ for $x_1 = 1$, $x_2 \in [-0.5, 0.5]$, $x_3 \in [-0.05, 0.05]$, $\beta = 1$.

*6.1.2 The sets* $S' = \{x; D(x) = D'(x) = 0\}$

This set represents the states where the control can cross the set $S$. From formula (63), we have two cases to study: $x_2 + 3x_3 = 0$ and $\hat{D} = 0$.

**Case 1:** $x_2 + 3x_3 = 0$**.** Since we have $x_i > 0$ along the entire trajectory, the condition $x_2 = -3x_3$ is not relevant for our problem.

**Case 2:** $\hat{D} = 0$**.** We solve $\hat{D} = 0$ for $\theta$ which gives

$$\theta = \frac{3\beta(-2x_2^2 + 4x_2x_3 + 8x_1x_3 + 9x_3^2)}{-8x_2^2 - 2x_1x_2 - 9x_2x_3 + 9x_1x_3 + x_1^2}. \tag{65}$$

Outside the set of parameters where both the numerator and denominator vanish (since $\theta$ is bounded), we can plug in this above expression for $\theta$ in $D'$ to obtain

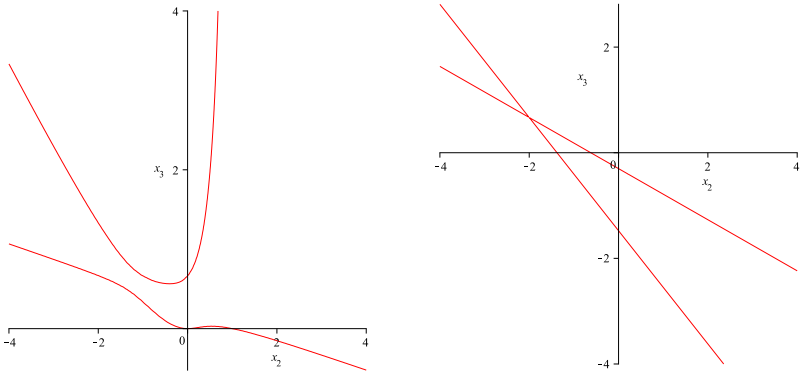
$$D' = \frac{\beta^3}{Den}(27x_3^2 + 48x_1x_3 + 12x_1^2 + 42x_2x_3 + 28x_1x_2 + 14x_2^2)(2x_2^3 - 2x_1x_2^2 + 12x_2^2x_3 + 9x_2x_3^2 - 9x_1x_3^2 + 6x_1^2x_3) \tag{66}$$

where $Den = -8x_2^2 - 2x_1x_2 - 9x_2x_3 + 9x_1x_3 + x_1^2$. We therefore have two conditions:

$$Pol1 = 2x_2^3 - 2x_1x_2^2 + 12x_2^2x_3 + 9x_2x_3^2 - 9x_1x_3^2 + 6x_1^2x_3 = 0, \tag{67}$$

or

$$Pol2 = 27x_3^2 + 48x_1x_3 + 12x_1^2 + 42x_2x_3 + 28x_1x_2 + 14x_2^2 = 0. \tag{68}$$




(a) Pol1 for $x_2$ and $x_3$ between -4 and 4.

(b) Pol2 for $x_2$ and $x_3$ between -4 and 4. There is a singularity at $(-2, -\frac{2}{3})$

**Fig. 4** Representation of the two polynomials $Pol1$ and $Pol2$ in a neighborhood of the origin for $x_1 = 1$.

The second condition corresponds to an union of two transversal lines intersecting at a singular point $x_2 = -2, x_3 = -\frac{2}{3}$, which is not relevant for our constraints on the state. The sets obtained by the two conditions are pictured in Fig. 4.

### 6.2 Rays

Our goal is to investigate the behavior of a singular trajectory associated with such a singular control $u_{sing}$ in the neighborhood of the half-ray $L = \mathbb{R}^+ \xi_{\bar{u}}$, in order to derive qualitative informations regarding the structure of the optimal control. In particular, an interesting problem is to determine whether or not any singular arc in the neighborhood of $L$ converges asymptotically to $L$. It turns out that this problem can be reduced to study the stability of the equilibrium of a linear problem. Indeed, as stated in proposition 2, the projection on the state space of the singular extremals of order 1 are the integral curves of the following vector field:

$$X_s(x) = Gx - \frac{D'(x)}{D(x) + \theta D'(x)}(Fx + \theta Gx)$$

which, using the time reparameterization $ds = \frac{dt}{D(x(t)) + \theta D'(x(t))}$, writes

$$X^r_{sing}(x) = D(x)Gx - D'(x)Fx.$$

Denote $\mathcal{B} = \{e_1, e_2, e_3\}$ the canonical basis, $\mathcal{B}' = \{\xi_{\bar{u}}, \xi_2, \xi_3\}$ the Graham-Schmidt orthonormalization of $\{\xi_{\bar{u}}, e_2, e_3\}$, $y = (y_1, y_2, y_3)$ the new coordinates in $\mathcal{B}'$ and $P$ the change-of-basis matrix from $\mathcal{B}$ to $\mathcal{B}'$. In the new set of coordinates, the dynamics $\dot{x} = X^r_{sing}(x)$ becomes

$$\dot{y} = P^{-1}\Big(D(Py)G - D'(Py)F\Big)Py. \tag{69}$$

and $L$ becomes to the positive $y_1$-axis. The transformation $P$ being linear, $D(Py)$ and $D'(Py)$ are two homogeneous degree 3 polynomial functions in the variables $y_1,\ y_2$ and $y_3$. Thus $\dot{y}_1,\ \dot{y}_2$ and $\dot{y}_3$ are homogeneous degree 4 polynomial functions in $y_1,\ y_2$ and $y_3$. Denote

$$\dot{y}_1 = \sum_{i+j+k=4} p_1^{ijk} y_1^i y_2^j y_3^k \tag{70}$$

$$\dot{y}_2 = \sum_{i+j+k=4} p_2^{ijk} y_1^i y_2^j y_3^k \tag{71}$$

$$\dot{y}_3 = \sum_{i+j+k=4} p_3^{ijk} y_1^i y_2^j y_3^k \tag{72}$$

where the coefficients $p_1^{ijk},\ p_2^{ijk}$ and $p_3^{ijk} \in \mathbb{R}$. Projecting on the tangent plan to the unit sphere at $(1,0,0)$, the dynamics reduces to 2-dimensional control system. Denote $z_2 = \frac{y_2}{y_1}$ and $z_3 = \frac{y_3}{y_1}$. We have

$$\begin{aligned}
\dot{z}_2 &= \frac{y_1\dot{y}_2 - y_2\dot{y}_1}{y_1^2} \\
&= \frac{\sum_{i+j+k=4} p_2^{ijk} y_1^{i+1} y_2^j y_3^k - \sum_{i+j+k=4} p_1^{ijk} y_1^i y_2^{j+1} y_3^k}{y_1^2} \\
&= \frac{\sum_{i+j+k=4} (p_2^{ijk} z_2^j z_3^k - p_1^{ijk} z_2^{j+1} z_3^k) y_1^5}{y_1^2}.
\end{aligned}$$

Similarly, we obtain

$$\dot{z}_3 = \frac{\sum_{i+j+k=4} (p_3^{ijk} z_2^j z_3^k - p_1^{ijk} z_2^j z_3^{k+1}) y_1^5}{y_1^2}. \tag{73}$$

Hence, using the time reparameterization $dr = y_1^3 ds$, we obtain the 2-dimensional system

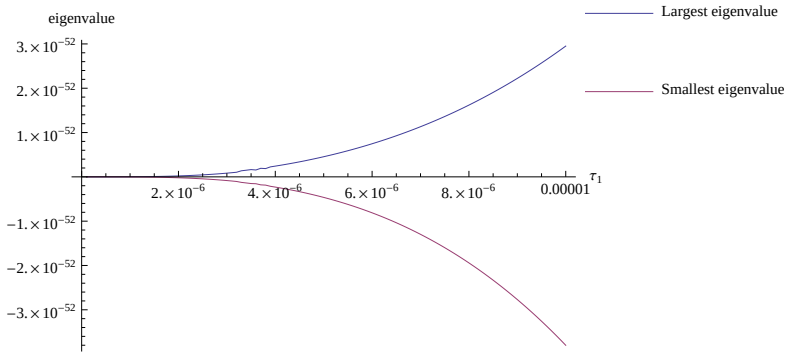
$$\begin{pmatrix} \dot{z}_2 \\ \dot{z}_3 \end{pmatrix} = Q \begin{pmatrix} z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} \sum_{i+j+k=4} p_2^{ijk} z_2^j z_3^k - p_1^{ijk} z_2^{j+1} z_3^k \\ \sum_{i+j+k=4} p_3^{ijk} z_2^j z_3^k - p_1^{ijk} z_2^j z_3^{k+1} \end{pmatrix}. \tag{74}$$

In this tangent plane,$L$ reduces to the equilibrium point $(0,0)$. We then can study its Lyapunov stability, see [1], by computing the eigenvalues the linearized system

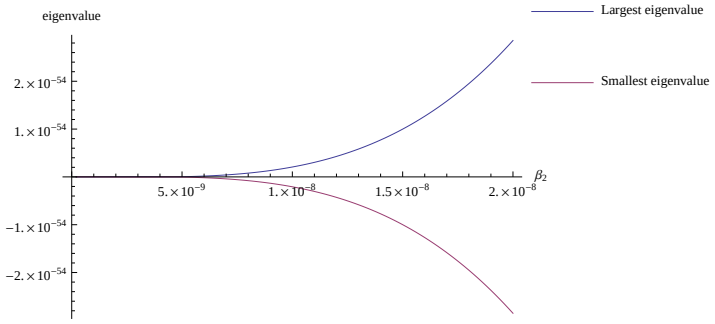
$$\begin{pmatrix} \dot{z}_2 \\ \dot{z}_3 \end{pmatrix} = J_{(0,0)} Q \begin{pmatrix} z_2 \\ z_3 \end{pmatrix} \tag{75}$$

where $J_{(0,0)}Q$ is the Jacobian matrix of the application $Q$ evaluated at (0,0).

Fig. 5 depicts the results of numerical computations that we carried out to study the Lyapunov stability of the equilibrium point $(0,0)$ of the linearized system (75) for different sets of parameters $\tau_1,\ \tau_2,\ \beta_2$ and $\beta_3$ . We investigated two different




(a) $1.10^{-7} \leq \tau_1 \leq 1.10^{-5},\ \tau_2 = 3\tau_1,\ \beta_2 = 2.10^{-10},\ \beta_3 = 3.10^{-10}$




(b) $2.10^{-10} \leq \beta_2 \leq 2.10^{-8},\ \beta_3 = \frac{3}{2}\beta_2, \tau_1 = 10^{-7},\ \tau_2 = 3.10^{-7}$

**Fig. 5** These two figures show the graphs of both the largest (in blue) and smallest (in red) eigenvalues of the Jacobian Matrix $J_{(0,0)}Q$ from the linearized system (75) when two of the four parameters $\tau_1,\ \tau_2,\ \beta_2$ and $\beta_3$ are held constant and the two other varie along given intervals. For these figures, we fixed $\theta = -1.10^{-3},\ \sigma = 1,\ u_{\min} = 1$ and $u_{\max} = 800$ so that the condition $\sigma + u\theta > 0$ is satisfied on the interval $[u_{\min}, u_{\max}]$.

cases. In the first case, the parameter $\tau_1$ is considered as a variable on the interval $[1.10^{-7}, 1.10^{-8}]$ and the parameter $\tau_2$ is set equal to $3\tau_1$. The two other parameters $\beta_2$ and $\beta_3$ are set equal to the constant values $2.10^{-10}$ and $3.10^{-10}$. In the second case, the parameter $\beta_2$ is considered as a variable on the interval $[2.10^{-10}, 2.10^{-8}]$, $\beta_3$ is set equal to $\frac{3}{2}\beta_2$ and $\tau_1$ and $\tau_2$ are set equal to constant values $1.10^{-7}$ and $3.10^{-7}$. In both cases, we used $\theta = -1.10^{-3},\ \sigma = 1,\ \ u_{\min} = 1$ and $u_{\max} = 800$. These values were chosen to guarantee that the Perron eigenvalue $\lambda_{\bar{u}}$ remained small enough so that our numerical computations could be performed and also to satisfy the constraint $\sigma + u\theta > 0$. For each combination $(\tau_1, \tau_2, \beta_2, \beta_3)$, we computed, by means of the method described previously, the corresponding matrix $J_{(0,0)}Q$ from the system (75) and its two real eigenvalues. We then plotted the resulting curves representing the evolution of both the eigenvalues as the variable parameter runs along the interval $I$. In the four different cases, computations showed that the

Jacobian matrix $J_{(0,0)}Q$ has always one positive and one negative eigenvalue. As a result, $(0,0)$ is an unstable equilibrium point and all singular arcs do not converge asymptotically to the half-ray $L$. It is observed that the eigenvalues of $J_{(0,0)}Q$ are opposite of each other (i.e. $tr(J(0,0)Q)$ is zero).

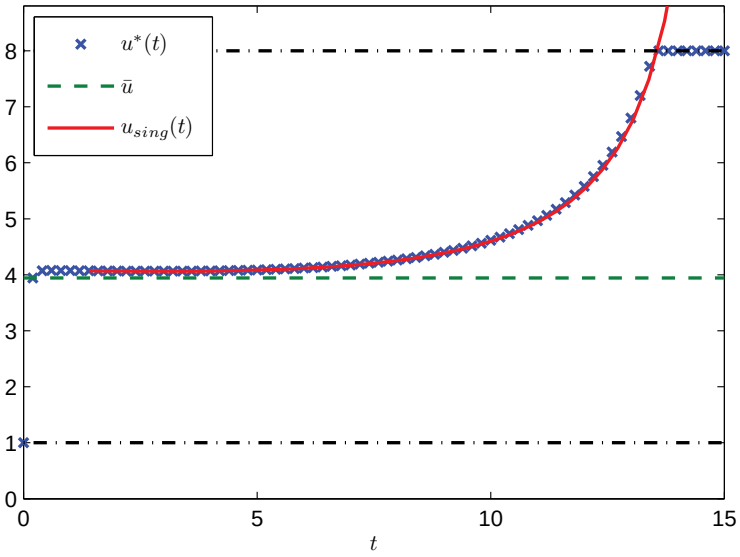



**Fig. 6** In this figure we have plotted the parameter $\bar{u}$ which maximizes the Perron eigenvalue, the optimal control $u^*(t)$, and the singular control $u_{sing}(t)$ obtained from the optimal trajectory by using Formula (55). The numerical values of the parameters are $u_{min} = 1,\ u_{max} = 8,\ \tau_1 = 0.01,\ \tau_2 = 10,\ \beta_2 = 0.1,\ \beta_3 = 0.9,\ \theta = 0$, and $x(0) = (3\ 4\ 4)^T$.

The behavior of the singular control plotted in Fig. 6 supports the instability result of the half-ray $L$. Indeed if $L$ was attractive, Formula (55) indicates that the singular control $u_{sing}(t)$ should converge to $\bar{u}$. But we see in Fig. 6 that the singular control, which coincides well with the optimal control $u^*(t)$ until it saturates, moves away from $\bar{u}$ as times passes.

## 6.3 Numerical Simulations

### *6.3.1 Comparison between strategies including the Perron singular control and the PMCA procedure*

Here, we compare numerically the performance of a constant Perron-singular control to the experimental strategy of 45 second pulses of $u_{max}$ followed by 30 minute periods of $u_{min}$ over at 48 hour period.

Figure 7 displays a comparison similar to the one done for the two-dimensional case. Once again, the comparison trajectory is obtained by applying the constant

Perron-singular control $\bar{u}$ over the entire trajectory but notice that in three dimensions it has not been proved to be the optimal strategy. It produces for this state of parameters an improvement of a magnitude of $10^{13}$, see Table 2. In 48 hours, the experimental trajectory yields $J(T) = 1.195303141025706e + 013$. In only 27.8278 hours, the $\bar{u}$ trajectory achieves the same $J$ value, and in 48 hours the $\bar{u}$ trajectory gives $J(T) = 4.290058715348198e + 026$.

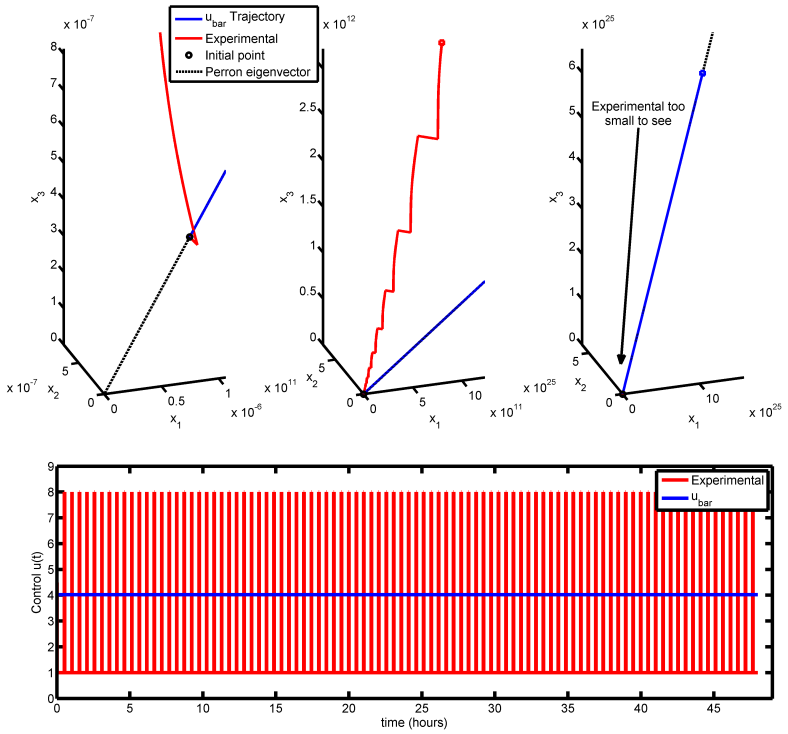



**Fig. 7** A three-dimensional example comparing the experimental control (red) versus the control frozen to $u = \bar{u}$ (blue), for a specific set of parameters: ($\theta = -0.1$, $\tau_1 = 10^{-3}$, $\tau_2 = 3 \cdot 10^{-3}$, $\beta_2 = 2 \cdot 10^{-4}$, $\beta_3 = 3 \cdot 10^{-4}$, $T = 48$ hours, $[u_{min}, u_{max}] = [1,\ 8]$). The initial point $x(0)$ is chosen to be on the eigenvector $\xi_{\bar{u}}$, and integrated for $T = 48$ hours. The top three plots are the same plot viewed at different zoom levels, showing the state trajectory. Again, the experimental control is defined as 45 second pulses of $u_{max}$ with 30 minute $u_{min}$ rest periods in between. The bottom plot shows the coressponding controls over time.

### *6.3.2 Conjugate points*

The goal of this section is to determine numerically whether the Perron-singular extremal posseses a conjugate point. Since the property could be dependent on the choice of parameters we proceed to an extended search by varying the physical parameters $\tau_i$ and $\beta_i$ as follows. We choose to fix $[u_{min}, u_{max}] = [1,\ 8]$, and $\theta = -0.1$. We also fix our simulation time to 48 hours which is consistent with the duration used by the experimentalists [15].

| Control regime | $J(T)$ |
|---|---|
| $\bar{u}$ | $4.2901 \cdot 10^{26}$ |
| $u_{exp}$ | $1.1953 \cdot 10^{13}$ |
| $u_{min}$ | $1.4061 \cdot 10^{11}$ |
| $u_{max}$ | $6.1708 \cdot 10^{9}$ |

**Table 2** The final costs for the three-dimensional example trajectories shown in Fig. 7. Also, for comparison, the final costs for trajectories using the same parameters but with the control frozen to $u_{min}$ and $u_{max}$ are shown. For this set of parameters, the $\bar{u}$-control provides over a $10^{13}$ magnitude improvement of the final cost when compared to the traditional experiment setting.

Using a Monte Carlo method, we randomally select 1000 sets of parameter values such that $\tau_1, \tau_2, \beta_2, \beta_3 \in [10^{-10}, 10^{-6}]$, and so that $\tau_1$ and $\tau_2$ are on the same order of magnitude: $\text{floor}(\log_{10}(\tau_1)) = \text{floor}(\log_{10}(\tau_2))$, and similarly for $\beta_2$ and $\beta_3$: $\text{floor}(\log_{10}(\beta_2)) = \text{floor}(\log_{10}(\beta_3))$.

A search for suitable parameters suggested that these assumptions achieve two nice properties:

1. the integrated values do not blow up to infinity before the integration reaches $T = 48$ hours, and thereby avoids numerical difficulties for the computer,
2. the Perron-singular control satisfies $u_{min} < \bar{u} < u_{max}$ (with strict inequality).

For each randomly selected parameter set, $\bar{u}$ is determined (to five decimal places of accuracy), as well as the corresponding $\lambda_{\bar{u}}$. The initial state conditions are chosen to be $\xi_{\bar{u}} \cdot 10^{-7}$, and we test for the first geometric conjugate time as described above, with the control fixed $u = \bar{u}$.

To compute the conjugate points, a similar algorithm to Hampath was used and coded in Matlab.

For all but two of the 1000 sets of parameters, the Perron-singular arc does not contain a conjugate point for parameters over a duration of 48 hours. For the set $\tau_1 = 0.948002359333351 * 10^{-6}, \tau_2 = 0.059641623585366 * 10^{-6}$, $\beta_2 = 0.026871234524279 * 10^{-6}$ and $\beta_3 = 0.098668004398028 * 10^{-6}$ we have $\bar{u} = 5.43549$ and a conjugate time is found at 133920 seconds = 37.2 hours. Along the second set of parameters with a conjugate time, we have $t_c$ =91450 seconds = 25.4 hours.

The results in this section are inconclusive for several reasons. First, the values for the physical parameters had to be chosen to allow an integration over a 48 hours period of time for both the original system and the variational equation. This allowed us to determine the existence or not of the conjugate point. However, for these choices it was also observed that no improvement on the final cost is produced with a Perron-singular control over the experimental procedure. For the values of parameters previously used for comparison, see section 6.3.1, that display significant improvement we have that the state variables of the variational equation blow-up quickly and the conjugate points cannot be computed. This suggests that the model needs to be altered for instance by incorporating a component to guarantee some sort of threshold in the evolution of the state and adjoint variables. The second issue in this section is the existence of conjugate points for two sets of parameters over the pool of 1000 candidates. It is doubtful that this could be attributed to numerical issues, but it has been observed that on those two cases a very slight modification in the parameters (taking $0.99\beta_3$ instead of $\beta_3$) produced a Perron-singular without a conjugate point. This should be further analyzed.

## 7 The continuous model

In this last section, we give some tracks to tackle the case of the continuous model (2). For the sake of simplicity, we consider the uniform fragmentation kernel $\kappa(y,z)=\frac{1}{z}$, meaning that a polymer can break anywhere along its length. We define the corresponding fragmentation operator $\mathcal{F}$ by

$$(\mathcal{F}f)(y)=2\int_y^\infty \frac{\beta(z)}{z}f(z)\,dz-\beta(y)f(y)$$

and the growth operator $\partial_y\tau$ by

$$(\partial_y\tau f)(y)=\partial_y\big(\tau(y)f(y)\big).$$

With this notation, the continuous version of the relaxed problem (11) writes

$$\partial_t f=-\partial_y\tau f+u(t)(\mathcal{F}-\theta\partial_y\tau)f. \tag{76}$$

Considering the duality bracket

$$\langle f,\varphi\rangle=\int_0^\infty f(y)\varphi(y)\,dy,$$

we define the dual fragmentation operator $\mathcal{F}^*$ by

$$\langle\mathcal{F}f,\varphi\rangle=\langle f,\mathcal{F}^*\varphi\rangle$$

where $f$ and $\varphi$ are any functions such that the left and right hand sides are well defined. Using the Fubini theorem, we can explicitely compute

$$(\mathcal{F}^*\varphi)(y)=2\frac{\beta(y)}{y}\int_0^y\varphi(z)\,dz-\beta(y)\varphi(y).$$

The continuous eigenfunctions associated with the optimal control $\bar u$ are defined by the Perron spectral problem (see [14,7])

$$\bar\lambda_P\Xi=-\partial_y\tau\Xi+\bar u(\mathcal{F}-\theta\partial_y\tau)\Xi,\quad \Xi(y)>0,\quad \int_0^\infty\Xi(y)\,dy=1,$$

$$\bar\lambda_P\Phi=\tau\partial_y\Phi+\bar u(\mathcal{F}^*+\theta\tau\partial_y)\Phi,\quad \Phi(y)>0,\quad \int_0^\infty\Phi(y)\Xi(y)\,dy=1.$$

With these definitions, the analogue of the Legendre condition (31) writes

$$\langle(\mathcal{F}\partial_y\tau-\partial_y\tau\mathcal{F})\Xi,(\mathcal{F}^*+\theta\tau\partial_y)\Phi\rangle>0. \tag{77}$$

We compute

$$\begin{aligned}\mathcal{F}\partial_y\tau\Xi(y)&=2\int_y^\infty\frac{\beta(z)}{z}\partial_z(\tau(z)\Xi(z))\,dz-\beta(y)\partial_y(\tau(y)\Xi(y))\\&=-2\frac{\beta(y)\tau(y)}{y}\Xi(y)-2\int_y^\infty\frac{z\beta'(z)-\beta(z)}{z^2}\tau(z)\Xi(z)\,dz\\&\qquad-\beta(y)\partial_y(\tau(y)\Xi(y))\end{aligned}$$

and

$$\begin{aligned}\partial_y\tau\mathcal{F}\Xi(y) &= 2\partial_y\Big(\tau(y)\int_y^\infty \frac{\beta(z)}{z}\Xi(z)\,dz\Big) - \partial_y(\tau(y)\beta(y)\Xi(y))\\ &= 2\tau'(y)\int_y^\infty \frac{\beta(z)}{z}\Xi(z)\,dz - 2\frac{\tau(y)\beta(y)}{y}\Xi(y)\\ &\qquad - \beta'(y)\tau(y)\Xi(y) - \beta(y)\partial_y(\tau(y)\Xi(y)).\end{aligned}$$

We get

$$\begin{aligned}&\langle(\mathcal{F}\partial_y\tau - \partial_y\tau\mathcal{F})\Xi, (\mathcal{F}^* + \theta\tau\partial_y)\Phi\rangle = \\ &\quad\int_0^\infty \Big[\beta(y)\Big(\frac{2}{y}\int_0^y \Phi(z)\,dz - \Phi(y)\Big) + \theta\tau(y)\Phi'(y)\Big]\\ &\qquad\times\Big[2\int_y^\infty \Big(\frac{\tau'(y)}{\tau(z)} + \frac{\beta'(z)}{\beta(z)} - \frac{1}{z}\Big)\frac{\tau(z)\beta(z)}{z}\Xi(z)\,dz - \beta'(y)\tau(y)\Xi(y)\Big]\,dy.\end{aligned} \tag{78}$$

For general coefficients, the sign of this quantity is not clear since it requires fine informations on the eigenfunctions $\Xi$ and $\Phi$. The profiles of these eigenfunctions have been studied in [2], but only asymptotically when $y$ tends to zero or infinity. Here more precise informations are necessary to deduce the bracket, and new investigations still have to be carried out.

Nevertheless we can give some words about the simple case, but biologically relevant [18], which corresponds to a constant polymerization rate $\tau(y) = \tau$ and a linear fragmentation rate $\beta(y) = \beta y$.

**Proposition 5** *For coefficients $\tau(y) = \tau$ and $\beta(y) = \beta y$, condition* (77) *is fulfilled.*

*Proof* For the coefficients considered, the dual eigenfunction $\Phi$ is explicitely given (see [14]) by

$$\Phi(y) = \frac{1}{2}\Big(1 + \sqrt{\frac{\bar{u}\beta}{(1+\theta\bar{u})\tau}}\,y\Big).$$

With this expression we can compute for any $y > 0$,

$$\Big[\beta(y)\Big(\frac{2}{y}\int_0^y \Phi(z)\,dz - \Phi(y)\Big) + \theta\tau(y)\Phi'(y)\Big] = \frac{\beta y}{2} + \frac{1}{2}\theta\tau\sqrt{\frac{\bar{u}\beta}{(1+\theta\bar{u})\tau}}$$

and

$$\Big[2\int_y^\infty \Big(\frac{\tau'(y)}{\tau(z)} + \frac{\beta'(z)}{\beta(z)} - \frac{1}{z}\Big)\frac{\tau(z)\beta(z)}{z}\Xi(z)\,dz - \beta'(y)\tau(y)\Xi(y)\Big] = -\beta\tau\Xi(y).$$

Finally we get for (78), using the normalizations $\int\Xi = \int\Phi\Xi = 1$,

$$\begin{aligned}\langle(\mathcal{F}\partial_y\tau - \partial_y\tau\mathcal{F})\Xi, (\mathcal{F}^* + \theta\tau\partial_y)\Phi\rangle &= -\frac{\theta\beta\tau}{2}\sqrt{\frac{\bar{u}\beta\tau}{1+\theta\bar{u}}} + \frac{\beta^2\tau}{2}\int_0^\infty y\,\Xi(y)\,dy\\ &= -\frac{\theta\beta\tau}{2}\sqrt{\frac{\bar{u}\beta\tau}{1+\theta\bar{u}}} + \frac{\beta\tau}{2}\sqrt{\frac{(1+\theta\bar{u})\beta\tau}{\bar{u}}}\\ &= -\frac{\beta\tau}{2}\sqrt{\frac{\bar{u}\beta\tau}{1+\theta\bar{u}}}\Big(\theta - \frac{1+\theta\bar{u}}{\bar{u}}\Big)\\ &= \frac{\beta\tau}{2\bar{u}}\sqrt{\frac{\bar{u}\beta\tau}{1+\theta\bar{u}}} > 0.\end{aligned}$$

Proposition 5 indicates that the Perron strategy is locally optimal for small times. This conclusion can even be strengthened by reducing Equation (76) to a system of two ordinary differential equations, as it is used in [18] to perform a stability analysis. Denoting

$$N(t) := \int_0^\infty f(t,y)\,dy \qquad \text{and} \qquad M(t) := \int_0^\infty y f(t,y)\,dy$$

we obtain by testing Equation (76) against 1 and $y$, the closed 2-dimensional system

$$\begin{cases} \dot{N} = u(t)\beta M, \\ \dot{M} = (1+\theta u(t))\tau N. \end{cases} \tag{79}$$

The optimal control problem for the continuous model is equivalent to the question of maximizing the quantity $M(T)$ subjected to the dynamics (79). But we know [12] that in dimension 2, the best strategy consists in the control which maximizes the Perron eigenvalue.

## References


1. V.I. Arnold, *Mathematical methods of classical mechanics*, Springer, New York (1989)
2. D. Balagué, J. Cañizo and P. Gabriel, *Fine asymptotics of profiles and relaxation to equilibrium for growth-fragmentation equations with variable drift rates*, Kinet. Relat. Models. **6**(2), 219-243 (2013)
3. B. Bonnard, J.-B. Caillau and E. Trélat, *Second order optimality conditions in the smooth case and applications in optimal control*, ESAIM Control Optim. and Calc. Var., **13**, 207–236 (2007)
4. B. Bonnard and M. Chyba, "Singular trajectories and their role in control theory", Springer-Verlag, Berlin, (2003)
5. B. Bonnard, M. Chyba, A. Jacquemard and J. Marriott, *Algebraic geometric classification of the singular flow in the contrast imaging problem in nuclear magnetic resonance*, Mathematical Control and Related Fields-AIMS, Special issue in the honor of Bernard Bonnard. Part II., **3(4)**, 397-432 (2013).
6. J.-B. Caillau, O. Cots & J. Gergaud, *Differential continuation for regular optimal control problems*, Optim. Methods Softw., **27** (2012), no 2, 177–196.
7. V. Calvez, M. Doumic and P. Gabriel *Self-similarity in a general aggregation-fragmentation problem. Application to fitness analysis*, J. Math. Pures Appl. (9) **98**, 1-27 (2012)
8. V. Calvez, P. Gabriel and S. Gaubert *Non-linear eigenvalue problems arising from growth maximization of positive linear dynamical systems*, submitted (2014), arXiv:1404.1868
9. J. Castilla and P. Saa and C. Soto, *Detection of prions in blood*, Nat. Med. **11**, 982-985 (2005)
10. F.E. Cohen and S.B. Prusinier, *Pathologic conformations of prion proteins.*, Annu. Rev. Biochem **67**, 793-819 (1998)
11. J. Collinge, *Prion diseases of humans and animals: their causes and molecular basis.* Annu. Rev. Neurosci. **24**, 519-555 (2001)
12. J.-M. Coron, P. Gabriel and P. Shang, *Optimization of an Amplification Protocol for Misfolded Proteins by using Relaxed Control* J. Math. Biol., Published online: 25 February 2014. DOI: 10.1007/s00285-014-0768-9
13. O. Cots, *Contrôle optimal géométrique : méthodes homotopiques et applications*. Phd thesis, Institut Mathématiques de Bourgogne, Dijon, France, 2012.
14. M. Doumic and P. Gabriel, *Eigenelements of a general aggregation-fragmentation model* Math. Models Methods Appl. Sci. **20**, 757–783 (2010)
15. N. Fernandez-Borges, J. Castilla, *PMCA. A Decade of In Vitro Prion Replication* Current Chemical Biology **4**, 200-207 (2010)
16. P. Gabriel, *Equations de transport-fragmentation et applications aux maladies à prions [Transport-fragmentation equations and applications to prion deceases]* PhD thesis, Paris, (2011)

17. N. Gonzalez-Montalban, N. Makarava, VG. Ostapchenko, R. Savtchenk, I. Alexeeva and et al. *Highly Efficient Protein Misfolding Cyclic Amplification* PLoS Pathog 7(2): e1001277. doi:10.1371/journal.ppat.1001277
18. M.L. Greer, L. Pujo-Menjouet and G. F. Webb, *A mathematical analysis of prion proliferation* J. Theoret. Biol., **242**(3), 598-606 (2006)
19. A. J. Krener, *The high order maximal principle and its application to singular extremals*, SIAM J. Control Optim., **15** (1977), no. 2, 256–293.
20. I. Kupka, *The ubiquity of Fuller's phenomenon* Nonlinear controllabillity and optimal control, Monogr. Textbooks Pure Apll. Math. **133** Dekker, New York, 313-350 (1990)
21. J. T. Jarrett and P. T. Lansbury, *Seeding one-dimensional crystallization of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie?* Cell **73** (6), 1055-1058 (1993)
22. U. Ledzewicz and H. Schättler, *Analysis of a cell-cycle specific model for cancer chemotherapy* Journal of Biological Systems, **10** (03), 183-206 (2002)
23. U. Ledzewicz and H. Schättler, *Analysis of models for evolving drug resistance in cancer chemotherapy* Dyn. Contin. Discrete Impuls. Syst. Ser. A Math. Anal., **13B** (suppl.) (03), 291-304 (2006)
24. U. Ledzewicz and H. Schättler, *Drug resistance in cancer chemotherapy as an optimal control problem* Discrete. Contin. Discrete Dyn. Syst. Ser. B, **6** (1), 129-150 (2002)
25. E.B. Lee and L. Markus, "Fondations of optimal control theory" Robert E. Krieger Publishing Co. Inc., Melbourne, FL, second edition (1986).
26. J. Masel, V.A.A Jansen and M.A. Nowak, *Quantifying the kinetic parameters of prion replication* Biophysical Chemistry, **77** (2-3), 139-152 (1999)
27. C.E. Mays, W. Titlow, T. Seward, G.C. Telling, C. Ryou, *Enhancement of protein misfolding cyclic amplification by using concentrated cellular prion protein source* Biochemical and Biophysical Research Communications **388**, 306310 (2009)
28. D. Mumford, J. Fogarty and F. Kirwan, *Geometric Invariant Theory,* Ergebnisse der Mathematik und ihrer Grenzgebiete (2) [Results in Mathematics and Related Areas (2)], **34**, Springer-Verlag, Berlin, 1994.
29. L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mishchenko, "The Mathematical Theory of Optimal Processes", John Wiley & Sons, New York, 1962.
30. R. Roos, D.C. Gajdusek and C.J., Jr. Gibbs, *The clinical characteristics of transmissible Creutzfeldt-Jakob disease.* Brain **96**, 1-20 (1973)
31. P. Saa and J. Castilla and C. Soto, *Ultra-efficient Replication of Infectious Prions by Automated Protein Misfolding Cyclic Amplification* Biological Chemistry **281** (16), 35245-3525 (2006)
32. G.P. Saborio and B. Permanne and C. Soto, *Sensitive detection of pathological prion protein by cyclic amplification of protein misfolding* Nature **411**, 810-813 (2001)
33. D. Serre, "Matrices", volume 16 of Graduate Texts in Mathematics. Springer-Verlag., New York, (2002) Theory and applications, Translated from the 2001 French original.
34. J. Zaslavski "Turnpike properties in the calculus of variations and optimal control", volume 80 of Nonconvex Optimization and its applications. Springer, New York, 2006.
35. M.I. Zelikin and V.F Borisov, "Theory of Chattering Control with Applications to Astronautics, Robotics, Economics and Engineering", Birkhäuser, 1994.