Cancer Therapy Optimization based on Unsupervised Learning and Multiple Model Adaptive Control

Francisco Manuel Ferreira Teles

Abstract—During the last years several clinical decision support systems have been developed, some of which clearly improved the results. However, this kind of decision computing process has not been quite explored in cancer treatment. In this work a control system that designs an optimal therapy based on adaptive control methods, allowing the eradication of a metastatic renal cell carcinoma as quickly and efficiently as possible and with lower associated toxicity, is developed. In order to do so, a new mathematical model describing the growth of this kind of tumor is developed, taking into account the effects of two of the most promising therapies: anti-angiogenesis and immunotherapy. Additionally, models describing pharmacodynamical aspects of the organism are also included. The therapy is designed through a multiple model adaptive control. Together with a system of selection and aggregation of key classes of models, it allows to deal with the uncertainty associated with the patient, namely his intra and inter variability. The results demonstrate that the built system presents robustness in terms of stability and performance. The reference following errors for the simulations are around 3 %, which allows a tumor eradication in less than a year and a half with mild and moderate toxicity levels. Therefore, the tool developed allows to create new perspectives in the creation of decision support systems for cancer therapy, thus enhancing the medical choices.

Index Terms—Cancer Therapy Design, Tumor Growth Model, Anti-angiogenesis, Immunotherapy, Multiple Model Adaptive Control, Model Clustering.

I. INTRODUCTION

A. Motivation and State-of-the-Art

Cancer is one of the most predominant causes of death with 14 million new cases globally, claiming 8.2 million lives yearly, as of 2012 [1]. If cancer-specific and sex-specific trends continue, an increase in the incidence of all-cancer cases by around 60%, will have reached 22.2 million cases in 2030, as predicted in [2]. The last fifteen years have seen remarkable progress in prevention, diagnosis and treatment of cancer, aided by technological advances and innovative research [3]. Nevertheless, one of the biggest problems remains: how to find the right treatment that can destroy a tumor in an efficient and effective way in the shortest period possible and with the least impact on patient’s life.

Automated decision makers or helpers are system-engineering tools that tend to make optimal decisions, taking into account the criteria and the parameterization of the problem. These tools have been widely used [4] to achieve considerable advancements in terms of quality and efficiency of products and services. In the health care sector this digital transformation is happening slowly [4], although it has been shown to yield added value for health organizations [5]. Therapy decisions are arduous not only because of the uncertainty associated with its effectiveness but also due to the possible negative outcomes that can arise for patients. Therefore, there are emotional factors and cognitive barriers that are limiting the implementation of Clinical Decision Support Systems (CDSSs) for cancer therapy design [6]. Nevertheless, some support systems have been implemented, such as FabAct [7] and LISA [8], systems that select the most appropriate drugs and dosages, respectively, for patients receiving chemotherapy and children with lymphoblastic leukemia.

Apart from the clinical experiments mentioned above, the scientific community has been creating and developing algorithms for optimizing cancer therapies. High costs, long trial times, difficulty in multiple hypothesis testing and ethical barriers make the feasibility of clinical trials, for testing different approaches and finding the optimal, an expensive and ineffective solution [9]. Therefore, mathematical modeling is an economical and powerful way of testing hypothesis and confirming therapies effects. Not only does it simulate the behavior of such complex systems, like tumor growth, pharmacokinetics (PK), pharmacodynamics (PD) and drug resistance (DR), but also and most importantly, their interactions.

Model based dynamics optimization or optimal control has been used in order to solve tumor eradication problems, using distinct techniques. For example, the optimal control study from Swan et al. [10] is critical for the understanding of the early modeling approaches for chemotherapy. Aside from that, Schättler et al. [11] applied geometrical optimal control to challenging problems. Subsequently, the same techniques were applied to a problem where a balance between cell kill and toxicity level was considered [12]. In [13], solutions are found using optimal control with different cost functions and also receding horizon control. Several other techniques are used such as model predictive control [14], or model reference adaptive control [15].

To the best of the author’s knowledge, there is no study approaching therapy optimization problem with Multiple Model Adaptive Control (MMAC), despite the fact that this control technique is used in several areas as, for example, for the control of blood pressure [16], neuromuscular blockade [17], [18], [19], or aircraft flight [20]. MMAC can also be used in combination with other techniques, such as, for example, with...
A tumor growth model is developed, using metastic renal cell carcinoma (mRCC) clinical results, through a combination of two different treatments: anti-angiogenesis and immunotherapy. To formulate the patient model, aspects as PK, PD, DR, and toxicity are also incorporated.

The uncertainty around patient characteristics, his reaction to therapy, and the inter and intra-patient variability is going to be taken into account by using MMAC, a control technique in which an optimal therapy is being constantly chosen by selecting, for each time $t$, from a large variety of models the one that, simultaneously best fits the unknown patient model and has the lowest toxicity level. The optimal control problem for each possible linearized model is solved using the Linear Quadratic Gaussian (LQG).

Finally, to allow the selection of any positive number of models, unsupervised learning in the form of clustering is going to be used to agglomerate patient models that are similar. This objective implies a constant number of controllers regardless the number of patient models.

The outline of the paper is as follows. The patient model is formulated in Section II. In Section III, a framework is developed for the control of a fully known patient. Section IV approaches the problem of the uncertainty around the patient, using the MMAC algorithm. In Section V, the addition of model clustering is explained. Numerical simulations are shown in Section VI and discussed in Section VII.

II. PATIENT MODEL

In order to develop a control algorithm to optimize cancer therapy, a patient model has to be formulated. This model is a mathematical formulation that incorporates the most important pharmacological aspects of a patient. Briefly, it can be summarized as in Figure 1, being each block, and their equations, described in the following subsections.

A. Tumor Growth Model

Anti-angiogenesis is a therapy that acts by blocking specific proteins, like VEGF-A, preventing the stimulation of endothelial cells, and thus the creation of new blood vessels in the tumor. This restriction limits the number of nutrients and oxygen in the tumor, resulting in a tumor shrinkage. On the other hand, immunotherapy can act as a boost to trigger the protective defenses of the immune system, or by inhibiting specific immune checkpoint pathways.

In order to develop a dynamic model that describes the dependency of the tumor behavior on the previous therapies, a fusion between Hahnfeldt et al. [22] and Ledzewicz et al. [23] models was performed, resulting in the following equations

\[ \dot{V} = -\lambda V \log \left( \frac{V}{K} \right) - \gamma VI, \]
\[ \dot{K} = -\mu K + bV - dKV^\frac{3}{2} - \eta Kg(t), \]
\[ \dot{I} = \mu_1 (V - \beta V^2)I - \delta I + \alpha + \kappa I_i(t). \]

The original constant terms were also modified in order to obtain the system following an approximation of a mRCC cancer clinical behavior, according to the clinical results described in [24], [25], resulting in: $\lambda = 0.025$ day$^{-1}$, $\gamma = 0.1$ day$^{-1}$, $\mu = 0.2$ day$^{-1}$, $b = 8.7$ day$^{-1}$, $d = 8.73 \times 10^{-3}$ day$^{-2}$ m$m^{-2}$, $\mu_1 = 3 \times 10^{-5}$, $\beta = 5 \times 10^{-4}$, $\delta = 0.37451$, and $\alpha = 0.15$. Furthermore, an interval of acceptable values for therapy effect sensibilities $\eta = [1.35, 11.25]$ kg/mg/day and $\kappa = [0.1, 0.5]$ kg/mg was also defined, as well as a range for tumor’s initial volume between 0.3 and 10 cm$^3$. It is worth noticing that the tumor volume $V$ represents the aggregated sum of tumors (main tumor and its metastases). Aside from that, $g(t)$ and $i(t)$ represent anti-angiogenesis and immunotherapy, respectively, as an effect, i.e. are between 0 and 1, where 0 means that no therapy is being applied, and 1 the maximum effect is achieved.

B. Pharmacokinetics

Pharmacokinetics (PK) is the one being responsible for determining the fate of substances administered to the organism and thus drug-concentration time profiles in the body.

This process can be incorporated in a two-compartmental model, which quantitatively describes the pharmacokinetic behavior of a drug in the organism [26]. The main compartment represents the blood and the heavily vascular organs supplied, being thus the compartment where drug administration and elimination are considered, while the peripheral compartment describes the rest of the body.

Since all the transfer process are of first-order, drug quantity in the central ($C_p$) and in the peripheral compartment ($P$) can be represented by the differential equations

\[ \dot{C}_p = g + k_{21}P - k_{12}C_p - k_{10}C_p, \]
\[ \dot{P} = k_{12}C_p - k_{21}P, \]

where $k_{12}$, $k_{21}$ are transfer rates, $k_{10}$ is the elimination rate, and $g$ is the drug administration rate.

Integrating equation (4) and dividing by the compartment volume results in the profile of the central compartment drug concentration $C_p$ [mg/kg/ml], given by

\[ C_p = \frac{D(\alpha - k_{21})}{V_c(\alpha - \beta)} e^{-\alpha t} + \frac{D(k_{21} - \beta)}{V_c(\alpha - \beta)} e^{-\beta t}, \]
where $D$ is the dose given to the patient in $mg/kg$, and the constants $\alpha$ and $\beta$ are the hybrid rates of distribution and elimination processes, respectively, that depend on $k_{12}, k_{21}$, and $k_{10}$.

The drugs considered in this work are bevacizumab, for anti-angiogenesis, and atezolizumab, for immunotherapy. For the first drug, the PK parameters estimated in [27] were used, resulting in $k_{12} = 0.223 \text{ day}^{-1}, k_{21} = 0.215 \text{ day}^{-1}, k_{10} = 0.0779 \text{ day}^{-1}$, and $V_c = 2600 \text{ ml}$. For the second drug, the two compartment model values were calculated from the experimental parameters presented in [28], [29], leading to $k_{12} = 0.3 \text{ day}^{-1}, k_{21} = 0.2455 \text{ day}^{-1}, k_{10} = 0.0643 \text{ day}^{-1}$, and $V_c = 3110 \text{ ml}$.

It is worth noticing that from equation (6) results a drug concentration in $mg/kg/ml$. However, drug doses are usually described in $mg/kg$. Therefore, in order to make a more straightforward analysis and to render an explicit comparison between tumor growth and PK models possible, the output of $C_p$ is multiplied by the volume $V_c$, being $C_p$ given in $[mg/kg]$.

### C. Pharmacodynamics

Pharmacodynamics (PD) allows the description of the relationship between drug effect, here described as $u$, and drug concentration $C_p$, which contributes to the understanding of a drug response and its effectiveness. It has been widely represented by the Hill Equation [30], a four parameter equation of a nonlinear relationship between concentration and drug effect, given by

$$u = u_{max}\frac{C_p^{\alpha}}{C_{50}^{\alpha} + C_p^{\alpha}}, \quad (7)$$

where $u_{max}$ is the maximum effect, typically unitary. $C_{50}$ is the drug concentration for which 50 \% of maximum effect is obtained, and $\alpha$ is the Hill coefficient determining the steepness of the resulting sigmoid. The $C_{50}$ parameter for bevacizumab and atezolizumab was calculated from the values in [31], [32], resulting $C_{50}_{\text{bev}} = 11.4274 \text{ mg/kg}$, and $C_{50}_{\text{ate}} = 7.1903 \text{ mg/kg}$. For simplicity, the constant $\alpha$ was considered unitary.

### D. Drug Resistance

It is necessary to assure that the correct level of drug concentration is presented in the target, taking into consideration the way the organism reacts to drugs, particularly how tumor cells resist to it. For this reason, a model for drug resistance was considered for both drugs, taking advantage of the capacity of malignant cells to proliferate into more resistant cells when low drug concentration is present in the plasma. This means that when $C_p$ is smaller than a threshold, drug resistance is acquired and a larger dose to achieve the same effect is necessary. Acquired resistance can be incorporated in the model by increasing the $C_{50}$ parameter from PD [33], since it will directly decrease the drug effect.

The drug resistance model can be defined as

$$C_{50}(t) = f(t)C_{50}^{\text{base}}, \quad (8)$$

where $C_{50}^{\text{base}}$ is the previously defined initial value and $f(t)$, which is a function that increases $C_{50}$ if the drug concentration $C_p$ is below the threshold $L_r$, is given by

$$f(t) = 1 + K_r \int_0^t \max[0, L_r - C_p(\tau)]d\tau. \quad (9)$$

The parameter $K_r$ is a positive constant that measures the capacity of the malignant cells to develop resistance against a certain drug.

### E. Toxicity

Clinical toxicity levels cannot be explicitly measured and there are no mathematical models for their evaluation. However, following the Common Terminology Criteria for Adverse Events (CTCAE) [34], toxicity in clinical trials can be graded as mild (grade 1), moderate (grade 2), severe (grade 3), life-threatening (grade 4) or death (grade 5).

By evaluating the concentration of drug induced in the body during therapy, toxicity levels $T_g$ for anti-angiogenesis and $T_i$ for immunotherapy, can be estimated by using a function yielding a certain small value until a certain amount of concentration is achieved. After this threshold, the function grows exponentially [33]. The toxicity curves considered in this work are illustrated in Figure 2. These curves were selected so that the grade 5 of toxicity was reached when the drug concentration in the organism was over 5 \% above the maximum dose allowed. These doses are 15 and 20 $mg/kg$, respectively, for bevacizumab [35] and atezolizumab [36].

Denoting $T_x$ as the total toxicity in the organism, it can be measured by averaging between $T_g$ and $T_i$, resulting in

$$T_x = \frac{T_g + T_i}{2}. \quad (10)$$

It is noticeable that toxicity does not have direct influence on any of the other models in Figure 1, since it does not influence the drug effect, being only an organism response to therapy. However, those are toxicity levels that constitute an important measurement that is going to be used in Section III.

### III. CONTROLLER DESIGN FOR A PATIENT MODEL

Considering that the patient model with state $x = [VKI]$ is fully understood, that is, all his parameters are known and do not vary over time, a feedback controller with proportional gains can be implemented, as illustrated in Figure 3. The roles of the Controller and Pharmacodynamical Simulation (PS) blocks are explained in the following subsections.
where $K_s$ is the feedback gain with dimension $2 \times 3$, being $G$ and $I$ the ideal effects of the anti-angiogenesis and the immunotherapy, respectively.

In order to compute the state estimation $\hat{x}$, an additional term has to be added to the state estimation differential equation. This term is proportional to the estimation error, and the multiplication by a matrix $L$ ensures the asymptotic convergence of the state estimations to the real state [37]. In this work, the observer matrix $L$ is calculated by using the Kalman estimator design for continuous-time systems. The covariance matrices were defined as $Q_o = 1000I$, where $I$ is the identity matrix, and $R_o = 0.1$.

The Linear Quadratic Regulator (LQR) controller [37] consists of a state feedback control law whose gains are selected in order to minimize the infinite horizon quadratic cost

$$J = \int x^TQx + U_i^TR_iU_i \, dt \quad i = 1, \ldots, 3 ,$$

where $Q > 0$ and $R_i > 0$ are matrices that can be tuned.

The selected value for the first were $Q = C^TC$. The three different input signals $U_i$ are performed by using different values of $R_i$, since decreasing $R_i$ means increasing the power of the manipulated variable $U_i$. The vector $U_i$ represents anti-angiogenesis and immunotherapy, hence a balance between both drugs in therapy can be made. The chosen matrices were

$$R_1 = \begin{bmatrix} 100 & 0 \\ 0 & 1000 \end{bmatrix}, R_2 = \begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix}, R_3 = \begin{bmatrix} 1000 & 0 \\ 0 & 100 \end{bmatrix}.$$  

Since $U_i = [G_i I_i]^T$, the effect of the immunotherapy in $U_1$ is smaller than in the same therapy in $U_2$, whilst the effect of the anti-angiogenesis in $U_3$ is lower than in $U_2$.

The input vector is given by $U_i = -K_i\hat{x}$, $i = 1, \ldots, 3$, where $K_i$ is the feedback gain with dimension $[2 \times 3]$, computed by

$$K_i = (B_i^TS_iB + R_i)^{-1}B_i^TS_iA \quad i = 1, \ldots, 3 ,$$

where $S_i$ is the only definite solution of the Riccati equation.

### A. Controller Implementation

In order to develop a feedback control system, the tumor growth model was linearized around the only real equilibrium point in the absence of therapy. Aside from that, the controllability and observability matrices were computed for confirming that the linearized system is fully controllable and observable.

In this work, the control action is executed depending on the process error, instead of directly in the feedback loop. The linear feedback controller used is constituted by an observer that estimates the state $\hat{x}$ for each model. The product between each state and three constant vectors $K_i$ with $i = 1, \ldots, 3$ is made in order to obtain three different therapies. Those therapies are described by a desired input effect vector $U = [U_1 U_2 U_3]$, where $U_i = [G_i I_i]^T$ with $i = 1, \ldots, 3$, being $G$ and $I$ the ideal effects of the anti-angiogenesis and the immunotherapy, respectively.

As side from what was mentioned, an input saturation was also included in order to guarantee that both drug effects are always between 0 and 1.

### B. Pharmacodynamical Simulation

The PS block has two main roles: choosing the least toxic therapy $U_T$ from the set of therapies $U$, and transforming it into a drug dose. To accomplish this roles, this block has to simulate pharmacodynamical processes of the patient before making any decision. Thus, one can consider the PS as a software that aims to approximate patient behavior, being possible to have some discrepancy with reality, being the uncertainty between the parameters of the PS and of the patient model given by $\Delta\theta_{PS}$ (in percentage) – see Section VI.

The PS’s composition and its interconnection with the Patient Model are shown in Figure 4. The process of selecting the least toxic therapy is executed by firstly transforming the effect vector $E$ into a concentration vector $C_{PS}^*$, through the inverse of the PD, denoted $PD^{-1}$. The parameter $C_{50}$ necessary for this transformation is calculated through an approximation of the DR model, named $DR'$. After the concentration calculation, it is possible to measure the toxicity levels, as described in Subsection II-E. Therefore, the least toxic drug concentration is the one chosen, by the block $T_{PS}$, to be administered to the patient.

In practice, it is desired that the concentration $C_p$ of the Patient Model follows the reference $C_{PS}^*$, selected by the PS block, since it is the least toxic drug concentration designed by the controller. It is achieved by the PK Control block that controls the PK model by using a linear feedback project. However, since PK information cannot be directly accessed and $C_p$ tends to be equal or almost the same as $C_{PS}^*$, one can assume PK Control as a unitary gain.
IV. MULTIPLE MODEL ADAPTIVE CONTROL

In oncology the uncertainty in the patient’s behavior is significant not only at the beginning of therapy, when patient’s characteristics are unknown, but mainly in its course, since the organism’s response is time-variant [39].

For models with large uncertainty, linear control theories, as pole placement or even LQR, cannot regulate the system in a satisfactory way, without an auxiliary mechanism. This problem led to the necessity of designing an adaptive control system sufficiently "smart" to operate in dynamic environments with a low degree of accuracy.

Multiple Model Adaptive Control (MMAC) is a modular technique, where the plant supervision is separated from the control. This technique overcomes the problem previously described, yielding to the control system a faster adaptation, improving also the transient performance [40].

Assuming that the plant dynamics \( P \) is defined by a set of parameters \( p \), one can consider \( p \in S \), being \( S \) a closed and bounded set of parameter space with finite dimension. The objective of MMAC is to control the plant by switching between a finite set of controllable and observable models \( M = \bigcup_{i=1}^{N} M_{i} \), being each model \( M_{i} \) characterized by a set of parameters \( \hat{p}_{i} \), that also respect \( \hat{p}_{i} \in S \), defined according to a convenient metrics.

It is considered that there exists a set of controllers \( C = \bigcup_{i=1}^{N} C_{i} \), where \( C_{i} \) is a local controller designed in order to have a good closed-loop performance in \( S_{i} \in S \), being \( S_{i} \) the neighborhood of the model parameter vector \( \hat{p}_{i} \).

Since the plant \( P \) is unknown, finding the model \( M_{i} \) that best fits the patient’s dynamics is a role executed by the supervisor, that has as inputs the least toxic therapy effect \( U_{T} \) and the plant output \( y \). Its output is the signal \( i^{*} \), responsible for activating the controller \( C_{i^{*}} \), which belongs to the controller bank, that is associated with the patient model \( M_{i^{*}} \). This process is illustrated in the simplified MMAC architecture in Figure 5 and it is fully detailed in the next subsections.

A. Supervisor

The supervisor decision making strategy is basically to choose for the output signal \( i^{*} \), a value corresponding to the index \( i \) of the model \( M_{i} \), whose performance index \( J_{i} \) is the smallest. To do so, the supervisor is constituted by an observer bank, a filter bank and a decision logic.

The observer bank is the responsible for computing the output estimations \( \hat{y}_{i} \) for each of the \( N \) considered models, based on \( U_{T} \), since \( u \) is not measurable, and on \( y \). Those estimations are performed by the equations (13) and (14), but with \( R_{o} = 10^{6} \) and \( Q_{o} = 10^{-5} I \).

Based on the models output estimations \( \hat{y}_{i} \), a prediction error \( \bar{e}_{i} = y - \hat{y}_{i} \) for each model \( M_{i} \) can be measured. A filtration process can be established in order to build a signal that quantitatively compares the plant with the model’s dynamics. Therefore, the prediction error \( \bar{e}_{i} \), after computed, is squared in order to be always positive and to present a favorable shape. The filtered error signal \( \pi_{i} \) is calculated by filtering the squared signal \( \bar{e}_{i}^{2} \) in order to smooth transactions and to remove high frequency components. A first-order low pass filter \( F_{i} \), with cutoff frequency of \( f_{c} = 30 \, MHz \) is used.

After obtaining the filtered signal \( \pi_{i} \), a local controller can be chosen by using a decision logic, which in this work is a three-step process composed by a performance index, a hysteresis condition and a dwell time criterion.

The performance index \( J_{i} \) is a cost function that evaluates which model should be chosen not only by looking for the current instant \( t \), but also by balancing it with the past information. In this way, a reliable and accurate model identifier can be built by using instantaneous and long-term empirical observations [41]. This feature is incorporated in the equation

\[
J_{i}(t) = \alpha \pi_{i}(t) + \beta \sum_{k=0}^{K} e^{-\lambda k} \pi_{i}(t - k) ,
\]

where \( K \) is the memory length, \( \alpha \) and \( \beta \) determine the importance of the current and past data, respectively, and \( \lambda \) is the forgetting factor. The values that demonstrated better results are \( K = 300 \, days, \alpha = 1, \beta = 1, \) and \( \lambda = 0.65 \).

To prevent excessive and inappropriate arbitrarily fast switches, a hysteresis algorithm [42] is used during the minimization of the performance index. Instead of simply minimizing at each instant the performance index, a hysteresis dead zone is considered. Thus, switching takes place only if

\[
\min_{i} J_{i}(t) < \delta J_{i}^{last}(t) ,
\]

where \( J_{i}^{last}(t) \) is the last performance index and \( \delta \) is the hysteresis constant, whose selected value was \( \delta = 0.65 \).

Chattering is avoided by using a dwell time condition [43]. After a switch the system continues selecting the same model for a minimum amount of time \( \tau_{D} \), denoted by dwell time. By dwelling in each model for a certain interval, at least during that time, the same controller remains connected with the plant which prevents possible instabilities that can arise due to swift switching. Besides that, in [44], despite using a slightly different structure, additional stability conditions are guaranteed for three different scenarios:

- **Scenario 1** System without noise, disturbances, and unmodelled dynamics;
- **Scenario 2** System with bounded noise and disturbances and without unmodelled dynamics;
- **Scenario 3** System with noise, disturbances, and unmodelled dynamics.

Hespansa et al. [44] assures that for Scenario 1 all the signals belonging to the Supervisor are bounded for every initial condition and, thus, the model switching stops in finite time, i.e. the system tends to choose the model that best fits...
the plant in a finite interval. For Scenario 2, by choosing a large value for $\tau_D$, the stability margin of the global system tends to the smallest stability margin of the local controllers. In this way, in [44] it is guaranteed that for linear systems, a uniformly exponentially stability can be considered. The situation is significantly different for Scenario 3 because the analysis becomes more complex. However, the authors assure that the control algorithm is robust for plant dynamics that are not being taken into account in the model space.

The value that showed better results in terms of stability, having thus few switch errors, was $\tau_D = 30$ days.

In the Figure 6, a computer diagram is illustrated, summarizing all the sequential steps performed by the decision logic.

### B. Controller Bank

The controller bank is a set of different controllers $C_1, ..., C_N$, each one designed as described in Subsection III-A by considering different patient models $M_i$. The optimal control is then chosen by switching the activated controller $C_i^*$ according to the switching signal $i^*$. Aside from this set, it is also constituted by a bumpless transfer and anti-windup system, and a mean filter. It is important to point out that although each local controller $C_i$ generates a set of three different signals, in order to simplify the explanations, from now on the signals are going to be considered unitary.

When changes in the plant are detected, the current controller is switched to a new one with a much more satisfactory performance. However, this ideal control process can be deteriorated by the switching since a degradation of the transient performance can be expected [45]. By swapping between one controller to another, abrupt changes in the manipulated variable $U$ can occur resulting in doses that largely exceed the maximum values allowed. However, bumpless transfer between different controllers can prevent this problem, for example, by inserting after the switch an integrator common to all the controllers [46]. Therefore, if an abrupt change happens in the manipulated variable, the integrator prevents it from drastically jump to higher values.

The addition of this integrator in the system has another implication since the local controllers design was performed without it. Therefore, the local controllers have to be redesigned by considering an augmented state space composed by the state $x$ and the state of the integrator $x_I$, resulting in

$$\begin{bmatrix} \dot{x} \\ \dot{U}_b \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ U_b \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} U_s ,$$  

where $A$, $B$ and $C$ are the previously defined state space matrices, and $U_s$ and $U_b$ are the drug effects before and after the bumpless transfer system, respectively.

The system composed by equations (18) and (19) is controllable but not observable. However, since $U_b$ is computed in software, direct measurements are available, being only necessary to estimate the state $x$ as it is currently implemented, but with the expansion of the matrices $K$ and $Q$.

An actuator saturation has to be considered after the integrator, since the drug effect $U$ must respect $U \in [0, 1]$. The actuator limits can be somehow achieved by the control variable leading the system to an open loop since the actuator is going to keep its maximum value independently of the system output. Besides that, when an integral action is used the control variable remains integrated leading to larger values. This process is called windup [38] and can cause large transients if the system is not correctly designed.

From the different anti-windup solutions, the scheme that is used in this work was introduced by Astrom and Wittenmark [47] and is shown in Figure 7. An extra feedback signal is added that measures the error $e_s$ between the desired and the true control signal. The signal $e_s$ is then fed back to the integrator, being the integrator input given by

$$U_f = U_s - \frac{1}{T_w} e_s .$$

When there is no saturation, the error $e_s$ is zero having no effect in the operation. However, when it is not null, the extra feedback path tries to drive the integrator input to zero, resetting the integrator so that the controller output is at the saturation limit, which prevents it from winding up. A value of $T_w = 10$ is used, since it represents the best tested scenario.

Despite the inclusion of the bumpless transfer, there may still exist situations in which the system is not able to prevent the occurrence of peaks. Example of this situation is a switch between two models that are very distinctive. The bumpless transfer is capable of smoothing these peaks but not of removing them. In this way, a cumulative moving average filter was included, which calculates for each time instant $t_n$, the mean of all the iterations between the first instant $t_0$ and the previous one $t_{n-1}$. In this way, the drug effect for the time instant $n$, $U^{(n)}$, is given by

$$U^{(n)} = \frac{U_{k}^{(n)}}{n} + n U^{(n-1)} ,$$

where $U_{k}^{(n)}$ is the output of the bumpless transfer and anti-windup system, and $U^{(n-1)}$ is the previous drug effect $U$.
V. MODEL CLUSTERING

The choice of the number of models, their selection and distribution in the parameter space $S$ are design issues crucial for achieving a good performance. In this work, it is essentially intended to use an algorithm where the controller bank is smaller than the model bank, i.e. where $N_c < N$, being $N_c$ the number of local controllers. Following the work of Oliveira et al. [48], it is expected to select a set of local controllers $C_j^c$, with $j = 1,..., N_c$, that best cover all the models $M_i$, where $i = 1,..., N$. For that, clustering algorithms have to be implemented in order to agglomerate similar models into a single cluster $c_j$. The centroids of the clusters $c_j$ form a new set of models $M_j^c$, from which the local controllers $C_j^c$ are going to be designed.

A. Model Sensitivity Analysis

The sensitivity analysis is used to choose from the set of 11 parameters of the patient model, the 2 most sensitive parameters, since clustering is performed based on them. Following Coito et al. work [49], the patient model can be described as $\dot{x} = f(x,u,\theta)$, where $\theta \in \mathbb{R}^{11}$ is the parameter vector. Deriving $\dot{x}$ with respect to the parameter $\theta_i$, where $i = 1,...,11$, and changing the order of derivatives, sensitivity $S_i$ can be represented as

$$\frac{d}{dt} S_i = \frac{\partial f}{\partial x} S_i + \frac{\partial f}{\partial \theta_i}. \quad (22)$$

After integrating equation (22) and simulating it, the parameters $\mu_1$ and $\beta$ were the ones that showed the greatest influence.

B. Vinnicombe Metric

An important factor regarding the classification and agglomeration of models is the choice of a good metric, i.e. a measurement of the closeness between linear time-invariant models. The Vinnicombe metric [50] was chosen to be used in this work, since an evaluation of the closed loop response is performed for comparing two distinct models. More precisely, if the Vinnicombe metric shows that two models are distinct, then a controller that gives satisfactory results for one model will behave poorly or even destabilize the other model.

Considering two models $M_i$ and $M_j$, that are described by a transfer function with the same number of inputs and outputs, the Vinnicombe metric $\delta_v(M_i,M_j)$ is then given by

$$\delta_v = \| (I + M_j M_j^{-1})^{-\frac{1}{2}} (M_j - M_i) (I + M_i M_j^{-1})^{-\frac{1}{2}} \|_{\infty}, \quad (23)$$

where $I$ is the identity matrix. The Vinnicombe metric can take values between $0 \leq \delta_v(M_i,M_j) \leq 1$. If $\delta_v(M_i,M_j) = 0$ the models are equal. On the other hand, $\delta_v(M_i,M_j) = 1$ means that both models are very distinctive. In this way, values closer to 0 indicate that a controller designed for $M_i$ can also stabilize $M_j$, having additionally similar closed loops gains.

C. Clustering Methods

After creating a model data set, constituted by $N$ models whose parameters were generated by a log-normal distribution, two clustering methods were applied: $k$-means [51] and complete linkage [52], which are described in the dissertation.

D. Clustering Results

For a test with $N = 100$ models and $N_c = 6$ clusters, the number of models $M_i$ in each cluster $c_i$, the mean of distances $d_c$ between $M_i$ and its clusters, and the worst case of the same distance $d_c$ were computed for both algorithms. The biggest cluster in the $k$-means had 24 models, whilst the larger one in complete linkage had 69. On the other hand, the smallest one for the first method had 7 clusters, whilst for the second it had only 1. The mean of $d_c$ was for the complete linkage 147.4% lower then for the $k$-means, being the worst $d_c$ also 91.8% lower for the first method. Aside from that, by increasing the number of clusters to 12, the complete linkage remains as the best method, however, with smaller differences.

At the first glance, the complete linkage seems to be the method with better performance in the tests made, having substantial differences compared with the $k$-means. However, the $k$-means seems also to agglomerate the models in a more uniform way, possibly yielding a better representation of the sample. This is mainly because the complete linkage tends to make partitions in areas where the models are further apart, creating also clusters with a large number of elements in areas where they are closer to each other. In this way, only the results of the control algorithm can say which algorithm is the best for the developed control framework.

VI. RESULTS AND SIMULATIONS

Several simulations were computed for a model data set that is distinct from the test set used in Section V for testing the clustering process. Comparison metrics were used in order to evaluate the quality of the results, which include the number of true and false switches, the time delay $\tau$ between the reference model and the true switch, and also information related with the reference following, described by the Mean Absolute Percentage Error (MAPE), that is given by

$$MAPE = \frac{100}{N} \sum_{i=1}^{N} \frac{|R_i - V_i|}{R_i},$$

where $R$ is the reference and $V$ the tumor volume output of the system.

This test simulates a scenario of a patient in which his dynamics, described by the equations summarized in Section II, is time varying. That is, the parameters $\mu$ and $\beta$ vary in a non linear way over time, as illustrated in Figure 8 by a solid black curve for the two different clustering algorithms with 6 and 12 clusters. This variation occurs in the following way: it takes 100 days from the model $M_A$ to $M_B$, 50 days from the model $M_B$ to $M_C$, and finally 350 days from $M_C$ to $M_D$.

The results of the MMAC algorithm’s simulations performed for the four methods are described in Table I.

All the four simulations present satisfactory results, with levels of MAPE around 3% and low toxicity levels, despite the fact that the trajectory in the parameter space happens in a sparse area in terms of models. What directly stands out from the simulations is the large number of false switches for the $12$ clusters’ algorithms with 5 and 8 false switches for $k$-means and complete linkage, respectively. This case happens because the computed centroids are closer to each other, which means that a small variation in the plant dynamics can result in a big oscillation in the selected model, until the stabilization of the control system. Aside from that, most of the false switches
Section III-B. Increasing $\Delta \theta_{PS}$ basically represents an increase in the parameters of the DR’ system, belonging to the PS, namely in the parameters $L_i^r$, $K_i^r$, and $C_i^{\text{base}}$. Increasing $L_i^r$ means that the drug resistance threshold is higher, but the same variation in the other two parameters is a warning for the control system that the mutant cells are more drug resistant. In this way, the control system is going to react by administrating a higher drug concentration in order to achieve the same desired effect, as illustrated in
Figure 11 where the drug concentration is higher for larger values of $\Delta \theta_{PS}$. Consequently, the toxicity maximum levels achieve values of $T_x = 4$, for $\Delta \theta_{PS} = 150 \%$, and $T_x = 49.6$, for $\Delta \theta_{PS} = 250 \%$, that are graded, respectively, as life-threatening and death – Section II-E.

However, this higher cell resistance is not real, since it is only an error between the PS’ parameters and the real ones from the patient. In this way, some repercussions in the system behavior are expected, more specifically, worse following errors for larger values of $\Delta \theta_{PS}$, which happens for $\Delta \theta_{PS} = 150$ and 250 %. However, for $\Delta \theta_{PS} = 50 \%$ the case is distinct, since its computed MAPE is smaller than the one without uncertainty – Table I.

This decrease in the MAPE for small values of $\Delta \theta_{PS}$ may be due to the existence of an offset in the system. Having a small positive value of $\Delta \theta_{PS}$ means administering a slightly higher drug concentration in the patient, which can dissipate the offset and, consequently, improve the reference following.

C. Noise and Disturbance Variation

The addition of uncertainty in the system is also simulated by adding disturbance $d$ and noise $n$, respectively, to the input and output. The system is going to be tested for three different situations: $R_1$ with $n = 0.1$ and $d = 10^{-4}$, $R_2$ with $n = 1$ and $d = 10^{-3}$, and $R_3$ with $n = 10^3$ and $d = 1$. As expected, by increasing the $n$ and $d$, the quality of the control system deteriorates. That is verified with the increase in the MAPE and in the number of false switches for the disturbed situations when compared with $R_0$ (no noise or disturbance) – Table I.

For the situation $R_1$, although the MAPE is slightly higher than in $R_0$, its curves for the selected model are similar. This event indicates that the values of noise and disturbance of the situation $R_1$ are acceptable for keeping a good performance.

On the other hand, for the other two disturbed situations, this conclusion cannot be expressed. For the situation $R_2$, there is an increase in the MAPE and, additionally, the switch between the models $M_2^*$ and $M_3^*$ is not correctly handled – Figure 12. For the situation $R_3$, in addition to the achieved higher MAPE, the MMAC system does not react to any changes, except at the first instants. This is exclusively due to the hysteresis condition. With the addition of noise and disturbance of large amplitude to the system, an “offset” with large varying amplitude is added to all the signals of the filtered prediction error $\pi_i$, that had small amplitudes. In this way, a set of signals with high order amplitudes that are closer to each other is created. For this reason, once the first model is selected, it is preserved until the end of the therapy since any possible switch “stagnates” in the hysteresis condition.

VII. CONCLUSION

The simulations shown in this abstract and the others described in the dissertation showed very satisfactory results, not only in terms of treatment efficacy, but also due to the low toxicity levels achieved. In all the situations, without the addition of error in the system, the tumor is eradicated in less than a year and a half, and the toxicity levels are graded between mild and moderate.

The clustering algorithms have impact on the behavior of the control system, since the positioning and the number considered for the key models, i.e. the centroids, influence the way the system performs. Particularly, with more clusters, and consequently more controllers, the system performance tends to degrade, since the system does not have the capacity of reacting with the necessary speed. Therefore, one can conclude that the addition of the model clustering to the MMAC brought performance improvement, since clustering’s main goal was exactly to reduce the numbers of controllers to operate.

Aside from that, the algorithm that showed the best results – $k$-means with 6 clusters – does not create clusters with a single model, which indicates that the centroids should be positioned in dense areas of models.

It was also concluded that the dwell time is important in some situations of chattering that can occur but, at the same time, can harm the system when a fast reaction is necessary.

Finally, improvements in the system can be settled, for example, considering the clusters with a single model as outliers and not as an important class of patients. Apart from that, a more variable controller bank can be also established, for example, by using an identification system or a neuronal network to adapt the set of models.

REFERENCES


