# Model-Invariant Safety-Preserving Control

Mahdi Yousefi, Klaske van Heusden, Guy A. Dumont, Ian M. Mitchell and J. Mark Ansermino

*Abstract*— **Employing viability techniques to preserve safety in safety-critical control applications has recently attracted much attention. The techniques rely on a conservative approximation of the viability kernel based on a state-space model of the system. However, in the case where the system's model is uncertain the safety concern has yet to be addressed in literature. In this work, we are seeking a single control action preserving safety for a set-represented model uncertainty. We define the model-invariant viability kernel for a multi-model uncertainty description. Based on this kernel, we show that under some assumptions, the control action taken based on any model in the model set is capable of preserving safety for the entire set. We propose a model-invariant safety-preserving control input which is the same for all models in the set. We evaluate and discuss the performance of the proposed scheme by applying it to the closed-loop control of anesthesia in which safety is critical.**

## I. INTRODUCTION

Addressing the problem of constraint satisfaction for systems under closed-loop control is a major concern, especially in safety-critical applications. Such applications include control of anesthesia [1], aircraft envelop protection [2], and process control [3]. Constrained model predictive control (MPC) [4] is commonly employed in cases where control variables are constrained and bounded due to safety concerns. However, there is no guarantee that MPC controller provides a feasible control input to keep the system's states within the constraint set (safe region).

Safety-preserving control addresses the above mentioned problem by guaranteeing the existence of a control input for a subset of the state space that can keep the states within the safe region. The set of states for which there exists a safety-preserving control input is called the "viability kernel" [5]. Accordingly, the first step in this method is viability kernel approximation. Margellos et al. [6] employ set theory and propose a dynamic programming based algorithm to approximate the viability kernel based on recursive computation of reachable sets. Kaynama et al. [7] use ellipsoidal representations of sets and propose a more computationally efficient approach to under-approximate the viability kernel. Gao et al. [8] discuss the viability approximation in the presence of uncertainty. They discuss the case where the model includes stochastic disturbances based on an approximation of the discriminating kernel, which is the viability kernel in the presence of a stochastic disturbance.

Once the viability kernel is approximated, one needs to synthesize a control law to preserve safety. Kurzhanski et al. [9] address the control synthesis problem through set-valued techniques as well as dynamic programming methods. Lygeros et al. [10] introduce a framework to design controllers to satisfy reachability specifications. This work was the basis for Kaynama et al. [11] to propose a hybrid automaton by combining the safety-preserving control with an arbitrary controller (performance controller) satisfying the performance criteria. In the scheme proposed in [11], the safety-preserving controller lets the performance controller achieve the desired performance as long as the states are sufficiently inside the safe region. As the states approach the boundaries of the viability kernel, the safety-preserving control adjusts the control input to prevent the states from going beyond the safe region.

Although a number of articles have been reported on different aspects of safety-preserving control, there is scant literature on invariability of the control technique to model uncertainties. Girard [12] and Kaynama et al. [7] discuss the safety issue in uncertain linear systems. However, the type of uncertainty they consider is additive state uncertainty, not model uncertainty. Although model-uncertainty can be represented as state uncertainty, the result might be very conservative and may not lead to satisfactory performance. Abate et al. [13] and Summers et al. [14] propose safety-preserving control approaches for hybrid stochastic systems. They define a control policy to be safety-preserving, if it maximizes the probability that the trajectory starting from the stochastic viability kernel remains within a safe region. In this case the safety-preserving control policy is selected from a set of control policies specified by the stochastic hybrid model.

In this paper, we aim to reduce the conservatism in safety-preserving control due to model uncertainty by extending the framework to model-invariant safety-preserving control. The proposed solution is limited to a specific model structure commonly used in control in anesthesia. We define a control input to be "model-invariant safety preserving" if it is capable of keeping the states of a set of state-space models within the safe region. Initially, for a finite set of models, we define the "model-invariant viability kernel" as the intersection of the viability kernels of all models in the model set. Subsequently, under certain assumptions, we prove that the safety preserving control input generated based on any model in the model set maintains the states of all models inside the safe region. We show that there exists a single control action

Mahdi Yousefi, Klaske van Heusden and Guy A. Dumont are with the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver BC. {mahdiyou,klaskeh,guyd}@ece.ubc.ca.

Ian M. Mitchell is with the Department of Computer Science, The University of British Columbia, Vancouver BC. mitchell@cs.ubc.ca.

J. Mark Ansermino is with the Department of Anesthesiology Pharmacology and Therapeutics, The University of British Columbia, Vancouver BC. anserminos@yahoo.ca.

capable of preserving safety of the entire model set.

In this work, we evaluate performance of the proposed technique by applying the model-invariant safety-preserving control to closed-loop control of anesthesia [15] in which safety is of utmost importance. In this application, a controller [1] manipulates the infusion rate of propofol to maintain a patient at a desired level of depth of hypnosis (DoH) based on the wavelet-based index ($WAV_{CNS}$) [16] feedback. To describe the relation between the drug infusion rates and a patient's DoH, we employ a PK[1]PD[2] model. Yousefi et al. [17] applied a safety-preserving control scheme to anesthesia control. In this paper, we extend this work by discussing the case where the patient model is within a set of models and there exists no exact knowledge of the model.

This paper is outlined as follows: Section II summarizes relevant results from the viability theory and safety-preserving control. In section III we introduce the model-invariant safety-preserving control and discuss the problem of control synthesis. Simulation results showing the effectiveness of this approach are presented in Section IV. Section V concludes the paper.

## II. Safety-Preserving Control

### A. Viability Theory

Consider the following state-space model:

$$\mathcal{X}: \ \dot{x}(t) = Ax(t) + Bu(t), \tag{1}$$

where $x(t)$ is the state vector in $\mathbb{R}^n$, $u(t) \in \mathcal{U}$ is the input vector in $\mathbb{R}^m$, and $A$ and $B$ are matrices in $\mathbb{R}^{n\times n}$ and $\mathbb{R}^{n\times m}$, respectively. The set $\mathcal{U}$ is the convex subset of $\mathbb{R}^m$ which specifies constraints on $u(t)$. The problem that the safety-preserving control techniques address is to guarantee the existence of a control input that can maintain the states of the model $\mathcal{X}$ inside the safe region $\mathcal{K}$.

*Definition 1:* (Viability Kernel) The finite-horizon viability kernel of $\mathcal{K}$ for the system $\mathcal{X}$ is a subset of $\mathcal{K}$ characterizing all initial conditions for which there exists an admissible control input $u(t, x(t)) \in \mathcal{U}$ (safety-preserving control) that keeps the trajectory $(x^{\mathcal{X}}_{x_0,u})$ of $\mathcal{X}$ emanating from those states within $\mathcal{K}$ for all $t \in [0, \tau]$:

$$Viab_{[0,\tau]}(\mathcal{K},\mathcal{U},\mathcal{X}) = \{x_0 \in \mathcal{K} \mid \exists u(t, x(t)) \in \mathcal{U} \text{ s.t. } \forall t \in [0,\tau], x^{\mathcal{X}}_{x_0,u} \in \mathcal{K}\ \}. \tag{2}$$

Kaynama et. al [7] use the results from set theory and propose an efficient recursive approach to approximate the viability kernel using the maximal reachable set. They partition the operational time interval ($[0, \tau]$) into $p$ subintervals. Then $Viab_{[t_i,\tau]}(\mathcal{K},\mathcal{U},\mathcal{X})$ can be approximated by [7]:

$$Viab_{[t_i,\tau]}(\mathcal{K},\mathcal{U},\mathcal{X}) = \mathcal{K}_{\downarrow} \bigcap Reach^{\#}_{[t_{i+1},t_i]}(Viab_{[t_{i+1},\tau]}(\mathcal{K},\mathcal{U},\mathcal{X}),\mathcal{U},\mathcal{X}) \tag{3}$$

where $Reach^{\#}_{[t_{i+1},t_i]}(Viab_{[t_{i+1},\tau]}(\mathcal{K},\mathcal{U},\mathcal{X}),\mathcal{U},\mathcal{X})$ is the maximal reachable set of the model $\mathcal{X}$ at time $t_i$ starting from $Viab_{[t_{i+1},\tau]}(\mathcal{K},\mathcal{U},\mathcal{X})$ at time $t_{i+1}$. $Viab_{[0,\tau]}(\mathcal{K},\mathcal{U},\mathcal{X})$ is approximated by calculating (3) recursively, i.e. moving backward in time from $t_p = \tau$ to $t_0 = 0$. $\mathcal{K}_{\downarrow}$ is the eroded version of $\mathcal{K}$ used instead of $\mathcal{K}$ to handle the effect of the time discretization [7]. $Reach^{\#}_{[t_{i+1},t_i]}(Viab_{[t_{i+1},\tau]}(\mathcal{K},\mathcal{U},\mathcal{X}),\mathcal{U},\mathcal{X})$ can be approximated using the level set toolbox developed by Mitchell [18]. To do so, we need to specify the constraint sets in an appropriate form. In this paper, we use the ellipsoidal technique [19] implemented in [20] to represent the constraint sets and to conduct ellipsoidal calculations. According to the ellipsoidal technique, any convex set can be approximated as an ellipsoid or a union of ellipsoids defined as follows:

$$\mathcal{E}(q, Q) = \{x \in \mathbb{R}^n \mid (x - q)^T Q^{-1}(x - q) \leq 1\} \tag{4}$$

where $Q$ is a positive definite matrix called the shape matrix, and $q$ is the center of the ellipsoid. Kaynama et al. [7] discuss the effect of approximating constraint sets with ellipsoids on accuracy of viability kernel approximation compared with other techniques such as polytopic, support vector and Hamilton-Jacobi methods.

### B. Control Synthesis

Safety-preserving control can be combined with any arbitrary controller (performance controller) and build a hybrid automaton to satisfy performance criteria while keeping the system safe [11]. Accordingly, once the system's states approach the constraint boundaries, the safety-preserving control adjusts the control input provided by the performance controller to keep the states within the safe bound. The hybrid automaton consists of two modes, namely, performance mode and safety mode [11]:

*1) Performance Mode:* Starting from any point in $Viab_{[0,\tau]}(\mathcal{K},\mathcal{U},\mathcal{X})$, if

$$x(t) \in \mathring{Viab}_{[t,\tau]}(\mathcal{K},\mathcal{U},\mathcal{X}) \tag{5}$$

we let the performance controller choose any value in $\mathcal{U}$ which satisfies the performance criteria. We denote by $\mathring{Viab}_{[t,\tau]}(\mathcal{K},\mathcal{U},\mathcal{X})$ the interior of $Viab_{[t,\tau]}(\mathcal{K},\mathcal{U},\mathcal{X})$. Since the control input is constrained, the control input ($\hat{u}_c$) provided by the performance controller may result in saturation. The saturated input can be determined by the support vector of $\mathcal{U}$ in the direction of the input generated by the performance controller:

$$u_{prf} = \begin{cases} \hat{u}_c & \text{if } \ u_{cnt} \in \mathcal{U}; \\ q_{\mathcal{U}} + Q_{\mathcal{U}} < \hat{u}_c, Q_{\mathcal{U}}\hat{u}_c >^{-1/2} & \text{if } \ u_c \notin \mathcal{U}, \end{cases} \tag{6}$$

where $q_{\mathcal{U}}$ and $Q_{\mathcal{U}}$ are the centre and the shape matrix of the ellipsoid specifying the input constraint and $< \bar{a}, \bar{b} >$ denotes the inner product of vectors $\bar{a}$ and $\bar{b}$.

*2) Safety Mode:* If $x(t) \notin \mathring{Viab}_{[t,\tau]}(\mathcal{K},\mathcal{U},\mathcal{X})$, the following control law will keep the states in $\mathcal{K}$ [11]:

$$u_{safe}(t, x(t)) = q_{\mathcal{U}} - Q_{\mathcal{U}}B^T d < d, BQ_{\mathcal{U}}B^T d >^{-1/2}, \tag{7}$$

where $d = {Q^{(l)}_{\mathcal{K},t}}^{-1}(x(t) - q^{(l)}_{\mathcal{K},t})$. $Q^{(l)}_{\mathcal{K},t}$ and $q^{(l)}_{\mathcal{K},t}$ denote the shape matrix and the center of the $l^{th}$ ellipsoid specifying

[1]Pharmacokinetics describes the distribution of drugs in plasma.

[2]Pharmacodynamics relates drug concentration in plasma to clinical effects.

$Viab_{[t,\tau]}(\mathcal{K},\mathcal{U},\mathcal{X})$. According to (7), the safety preserving control law pushes the states inside $\mathcal{K}$ toward the origin of the viability kernel. Kaynama et al. [11] prove that the control policy illustrated in (8) is capable of keeping $x(t)$ ($t \in [0,\tau]$) within $\mathcal{K}$.

$$u_c(t,x(t)) = \begin{cases} u_{prf}(t,x(t)) & \text{if } x(t) \in \mathring{Viab}_{[t,\tau]}(\mathcal{K},\mathcal{U},\mathcal{X}); \\ u_{safe}(t,x(t)) & \text{if } x(t) \notin \mathring{Viab}_{[t,\tau]}(\mathcal{K},\mathcal{U},\mathcal{X}). \end{cases} \tag{8}$$

Choosing an arbitrary value for $u_{prf}(t,x(t))$ without considering the main control objective, which is preserving safety, may result in high frequency switching between the two modes (chatter). Kaynama et al. [11] address this problem by using the convex combination of the two control modes as follows:

$$u_c(t,x(t)) = (1-\beta_\alpha(\zeta))u_{prf}(t,x(t)) + \beta_\alpha(\zeta)u_{safe}(t,x(t)), \tag{9}$$

where

$$\beta_\alpha(\zeta) = \begin{cases} 1 & \text{if } \zeta \le l; \\ \dfrac{1}{1-\alpha}(\zeta-\alpha) & \text{if } \alpha \le \zeta < 1; \\ 0 & \text{if } \zeta < \alpha. \end{cases} \tag{10}$$

In the above equation, $\alpha \in [0,1)$ is a design variable and $\zeta$ denotes how deep $x(t)$ is in $Viab_{[t,\tau]}(\mathcal{K},\mathcal{U},\mathcal{X})$:

$$\zeta = < (x(t) - q^{(l)}_{\mathcal{K},t}), {Q^{(l)}_{\mathcal{K},t}}^{-1}(x(t) - q^{(l)}_{\mathcal{K},t}) > . \tag{11}$$

If the states are sufficiently inside $Viab_{[t,\tau]}(\mathcal{K},\mathcal{U},\mathcal{X})$, the safety control policy allows the performance controller to choose any value in $\mathcal{U}$ to achieve the desired closed-loop performance. As the states approach the boundaries of $Viab_{[t,\tau]}(\mathcal{K},\mathcal{U},\mathcal{X})$, the safety preserving control manipulates $u_{prf}(t,x(t))$ to make an smooth transition between the performance and safety modes.

## III. Model-Invariant Safety-Preserving Control

The main objective of this paper is to extend the results from safety-preserving control with a known system model to the case where the system's model lies within a finite set of state-space models. Here, the objective is to find a control input which keeps the states of all models in the model set $\mathcal{M}$ specified below inside $\mathcal{K}$:

$$\mathcal{M} = \{\mathcal{X}_i | \mathcal{X}_i : \dot{x}(t) = A_i x(t) + \alpha_i B u(t),\ \alpha_i > 0, \\ i = 1,\dots,p\}. \tag{12}$$

*Definition 2:* (Model-Invariant Viability Kernel) The finite-horizon model-invariant viability kernel of $\mathcal{K}$ for the model set $\mathcal{M}$ is a subset of $\mathcal{K}$ characterizing all initial conditions for which there exists an admissible control input $u(t,x(t)) \in \mathcal{U}$ (model-invariant safety-preserving control) that keeps the trajectory of all $\mathcal{X} \in \mathcal{M}$ emanating from those states within $\mathcal{K}$ for all $t \in [0,\tau]$:

$$Viab_{[0,\tau]}(\mathcal{K},\mathcal{U},\mathcal{M}) = \{x_0 \in \mathcal{K} \mid \exists u(t,x(t)) \in \mathcal{U} \text{ s.t.} \\ \forall t \in [0,\tau] \,\&\, \forall \mathcal{X} \in \mathcal{M},\ x^{\mathcal{X}}_{x_0,u} \in \mathcal{K}\ \}. \tag{13}$$

To ensure the model-invariant viability kernel does exist, we must have:

$$\mathcal{K} \subseteq \bigcap_{\mathcal{X}\in\mathcal{M}} \{\mathcal{W}_{\mathcal{X}}\}, \tag{14}$$

where $\mathcal{W}_{\mathcal{X}}$ is the response space of $\mathcal{X}$. Assume the set $\mathcal{M}$ consists of a limited number of members. Let's define $\mathcal{I}$ as an intersection of the viability kernels of all individual models in $\mathcal{M}$:

$$\mathcal{I} = \bigcap_{\mathcal{X}\in\mathcal{M}} \{Viab_{[0,\tau]}(\mathcal{K},\mathcal{U},\mathcal{X})\}. \tag{15}$$

$\mathcal{I}$ specifies a subset of the state space for which there exists a safety-preserving control input for every model in the model set. However, existence of a unique input keeping the states of all models is debatable. Due to the convexity of the constraint sets, the approximated viability kernel is also convex [5]. Thus, the intersection $\mathcal{I}$ is convex too. In *Proposition 1*, we will show that under certain assumptions on $\mathcal{M}$, we have:

$$\mathcal{I} \subset Viab_{[0,\tau]}(\mathcal{K},\mathcal{U},\mathcal{M}). \tag{16}$$

Let's define $\tilde{\mathcal{I}}$ as an under-approximation of $\mathcal{I}$ with a set of ellipsoids. We will show by using $\tilde{\mathcal{I}}$ as a subset of the model-invariant viability kernel, the safety-preserving control (7) calculated based on any model in $\mathcal{M}$ is capable of maintaining the states of the entire set within the safe region.

*Lemma 1:* Assuming $\dfrac{d}{dt}V(t,x(t))$ exists, the following control policy preserves safety for the state-space model $\mathcal{X}$:

$$u_{safe}(t,x(t)) = \arg\min_{u(t)}\{\frac{d}{dt}V(t,x(t))|\ u(t) \in \mathcal{U}\}, \tag{17}$$

In the above equation, $x(t)$ is characterized by a state-space model $\mathcal{X}$ and $V(t,x(t)) = Dist^2(t,x(t),\mathcal{Z}(t))$, assuming $\mathcal{Z}(t) = Viab_{[t,\tau]}(\mathcal{K},\mathcal{U},\mathcal{X})$. $Dist(t,x(t),\mathcal{Z}(t))$ is the Hausdorff distance measuring the distance of $x(t)$ from the set $\mathcal{Z}(t)$, which is defined by:

$$Dist(t,x(t),\mathcal{Z}(t)) = \min\{\|x(t) - z\| \mid z \in \mathcal{Z}(t)\}. \tag{18}$$

According to 18, if $x(t) \in \mathcal{Z}(t)$, then $Dist(t,x(t),\mathcal{Z}(t)) = 0$ and if $x(t) \notin \mathcal{Z}(t)$, $Dist(t,x(t),\mathcal{Z}(t)) > 0$.

*Proof:* [21], chapter 1, page 25. ■

*Lemma 2:* (17) is capable of preserving safety for any subset of the viability kernel of $\mathcal{X}$ ($\mathcal{S} \subset Viab_{[t,\tau]}(\mathcal{K},\mathcal{U},\mathcal{X})$) if $V(t,x(t))$ is formulated based on $\mathcal{S}$ rather that $Viab_{[t,\tau]}(\mathcal{K},\mathcal{U},\mathcal{X})$ [21].

*Proof:* [21], chapter 1, page 25. ■

*Proposition 1:* For the model set $\mathcal{M}$ characterized as (12), using the model-invariant viability kernel defined in *Definition 2*, safety-preserving control (17) formulated based on any model in $\mathcal{M}$ is capable of maintaining the states of all members of $\mathcal{M}$ within the safe region $\mathcal{K}$.

*Proof:* According to *Lemma 1*, for $\mathcal{X} \in \mathcal{M}$, $u_{safe}(t,x(t))$ defined in (17) is capable of keeping the states of $\mathcal{X}$ inside $\mathcal{K}$. Kurzhanski et al. [21] show that $Dist(t,x(t),\mathcal{Z}(t))$ can be expressed as

$$\begin{aligned} Dist(t,x(t),\mathcal{Z}(t)) &= \max\{ < l, x(t) > - \rho(l|\mathcal{Z}(t))|\ \|l\| \le 1\} \\ &= < l^0, x(t) > - \rho(l^0|\mathcal{Z}(t)), \end{aligned} \tag{19}$$

where $l^0$ ($l^0 \neq 0$, $\|l^0\| = 1$) is a unique maximizer for $< l, x(t) > -\rho(l|\mathcal{Z}(t))$ and

$$\rho(l|\mathcal{Z}(t)) = \max\{< l, z > \mid z \in \mathcal{Z}(t)\} \tag{20}$$

is the support function of $\mathcal{Z}(t)$ in direction $l$. Accordingly, one can show (17) is equivalent to

$$u_{safe}(t, x(t)) = \arg\min_u \{\frac{d}{dt} Dist(t, x(t), \mathcal{Z}(t))| \ u \in \mathcal{U}\}. \tag{21}$$

According to *Lemma 2*, (21) also preserves safety for any subset of the viability kernel. Thus, assuming $\mathcal{Z}(t) = \mathcal{I}(\subset Viab_{[t,\tau]}(\mathcal{K}, \mathcal{U}, \mathcal{X}))$ which is defined in (15), and following the analysis in [21] we have

$$\begin{aligned}\frac{d}{dt} Dist(t, x(t), \mathcal{Z}(t)) =& \frac{\partial}{\partial t} Dist(t, x(t), \mathcal{Z}(t)) \\ &+ < \frac{\partial}{\partial x} Dist(t, x(t), \mathcal{Z}(t)), \dot{x}(t) > \\ =& < l^0, \dot{x}(t) > -\frac{\partial}{\partial t}\rho(l^0|\mathcal{Z}(t)).\end{aligned} \tag{22}$$

Kurzhanski et al. [21] show that

$$\frac{\partial}{\partial t}\rho(l^0|\mathcal{Z}(t)) =< l^0, Ax(t) > +\rho(l^0|B\mathcal{U}). \tag{23}$$

By using any member of $\mathcal{M}$ in (23), we have

$$\frac{\partial}{\partial t}\rho(l^0|\mathcal{Z}(t)) =< l^0, A_i x(t) > +\rho(l^0|\alpha_i B\mathcal{U}). \tag{24}$$

Substituting (24) in (22) yields

$$\begin{aligned}\frac{d}{dt} Dist(t, x(t), \mathcal{Z}(t)) =& < l^0, A_i x(t) + \alpha_i B u > \\ &- < l^0, A_i x(t) > -\rho(l^0|\alpha_i B\mathcal{U}) \\ =& < l^0, \alpha_i B u > -\rho(l^0|\alpha_i B\mathcal{U}).\end{aligned} \tag{25}$$

Thus, we can represent (17) as

$$u_{safe}(t, x) = \arg\min_u \{< l^0, \alpha_i B u > \mid u \in \mathcal{U}\}. \tag{26}$$

Due to the convexity of the optimization problem in (26), the minimizer is independent of the term $\alpha_i$, i.e. no matter which value we choose for $\alpha_i$, (26) always has a unique minimizer. Consequently, $u_{safe}(t, x)$ defined in (26) is a unique control input which can keep the states of all members of $\mathcal{M}$ defined in (12), within the safe region and

$$\mathcal{I} \subset Viab_{[0,\tau]}(\mathcal{K}, \mathcal{U}, \mathcal{M}). \tag{27}$$

■

(7) is the solution to (17) assuming the viability kernel as well as the constraint sets are under-approximated by ellipsoids [11]. Accordingly, one can show that using any model in the model set specified in (12) and $\tilde{\mathcal{I}}$ as an under-approximation of the model-invariant viability kernel of $\mathcal{M}$, (7) results in the same $u_{safe}$. Although the condition (12) on the model set appears to be very strict, it is satisfied in certain physical uncertain systems especially for all single input model sets in which the input has a direct influence on only one state of the models. This definition includes applications such as closed-loop control of anesthesia.

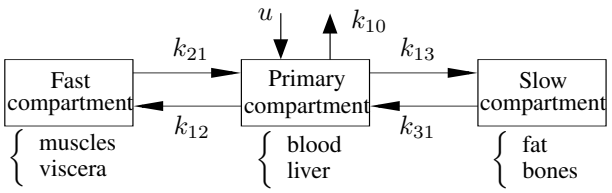


Fig. 1. Drug distribution in primary, fast and slow compartments (PK model)

According to the above proof, if we can approximate the model-invariant viability kernel for a set of models which can be represented as described in (12), it does not matter which model we use to calculate the safety-preserving control input, the input would be able to keep the states of the entire model set within the safe region.

## IV. Case Study

In this section, we apply model-invariant safety-preserving control to the closed-loop control of DoH. The effect of propofol anesthesia on DoH can be described by a Wiener model, generally containing a 3-compartment PK model [22], and a first-order model followed by a non-linearity to describe the PD model [23]. The PK and PD models can be combined and represented as a 4-state state-space model as follows:

$$\dot{x}(t) = Ax(t) + Bu(t - \tau_d), \ y = C_e. \tag{28}$$

$x \in \mathbb{R}^4$ includes drug concentrations in the plasma $C_p$ ($x_1$), fast peripheral ($x_2$), and slow peripheral ($x_3$) compartments as well as the effect-site concentration $C_e$ ($x_4$). As illustrated in fig. 1, the input $u$ is the anesthetic drug which is injected to the primary compartment and then it is distributed to the other compartments. Therefore, The matrix $B$ can be represented as:

$$B = \alpha \begin{bmatrix}1 & 0 & 0 & 0\end{bmatrix}^T. \tag{29}$$

The relation between the effect-site concentration and the anesthetic effect can be described by the non-linear Hill equation:

$$E(C_e) = \frac{C_e^\gamma}{EC_{50}^\gamma + C_e^\gamma}, \tag{30}$$

where $EC_{50}$ is the concentration results in $50\%$ of the overall effect and $\gamma$ is the cooperativity coefficient. During the maintenance phase of anesthesia, the control objective is to keep the drug effect at $E = 0.5$. Bibian et al. [23] identified and presented a set of 44 models corresponding individual patient responses. The first 15 models of this set, including PKPD models representing patients between ages 20 to 29 are used in this simulation example. Since the PKPD models can be represented in the form of (12), we can apply the model-invariant safety-preserving control to the model set. This paper focusses on state-space uncertainty. In the following, the anesthesia control problem is therefore simplified by assuming that all states can be measured, $EC_{50}$ is known and the system is delay free.

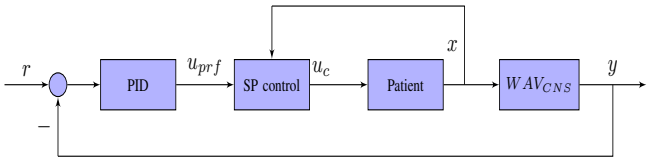


Fig. 2. Closed-loop diagram of the safety-preserving control of Anesthesia

To overcome the effect of the non-linearity in the PKPD model, we calculate the effect-site concentration by point-to-point inverse mapping of the drug effect:

$$E^{-1}(.): \ E(C_e) \rightarrow C_e. \tag{31}$$

Accordingly, if there is a constraint on the output (drug effect) it can be described as a constraint on $x_4$ (assuming $x_4 = C_e$ in (28)):

$$E(C_e) \in [a, b] \rightarrow C_e \in [E^{-1}(a), E^{-1}(b)]. \tag{32}$$

To satisfy the performance criteria in our simulations, we implement the robust PID controller used by van Heusden et al. in [1]. Fig. 2 depicts the block diagram of the safety-preserving closed-loop control implemented in this paper. The drug effect is measured by $\text{WAV}_{\text{CNS}}$ and is fed back to the PID controller. In this control scheme, initially, we let the PID controller bring the states inside the safe set. Once the predicted states reach the viability kernel, safety-preserving control is turned on and keeps the states within the constraints over the operational time. The constraints we define on the states and the input are:

$$0 \leq x_1, x_2, x_3 \leq 12, \tag{33}$$

$$1.52 \leq x_4 \leq 2.28, \tag{34}$$

$$0 \leq u \leq 600. \tag{35}$$

Fig. 3 illustrates the projection of the constraint set in $x_4$-$x_1$ plane (red region). In this figure, we focus on the plasma ($x_1$) and effect-site ($x_4$) concentrations which are the most important variables in closed-loop control of anesthesia. Fig. 3 represents the normalized version of the constraint set which is moved to the origin. The green region represents the model-invariant viability kernel under-approximated by a set of ellipsoids. This set is calculated by intersecting the viability kernels of 15 patient models between ages of 20 to 29. In this figure, we choose one of the patient models and compare the trajectory of the states when the safety-preserving control is applied and when it is not. Accordingly, in the case without the safety-preserving control, the trajectory (blue line) violates the constraint set for a short period of time. When the safety preserving control is turned on, it prevents the trajectory (red line) from going beyond the constraint set (red region.) Fig. 4 depicts DoH and the drug infusion rate in both cases for the same model. DoH of 100 is associated with the drug effect of 0 and DoH of 0 is associated with the drug effect of 1 (maximum effect). So, according to the constraint defined in (34) and the value for $EC_{50}$(=1.9), the objective is to keep DoH between 40 and 60. The upper plot in fig. 4 shows the capability of the safety-preserving control in

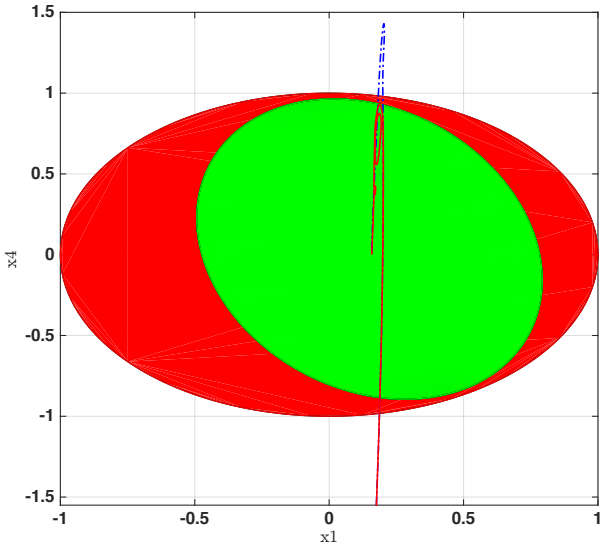


Fig. 3. $x_1$-$x_4$ plane, red region: constraint set ($\mathcal{K}$), green region: under-approximation of model-invariant viability kernel ($Viab(\mathcal{K}, \mathcal{U}, \mathcal{M})$), blue dashed line: the trajectory of the states with safety-preserving control off, red line: the trajectory of the states with safety-preserving control on. The trajectories are associated with the $4^{th}$ model in the model set.

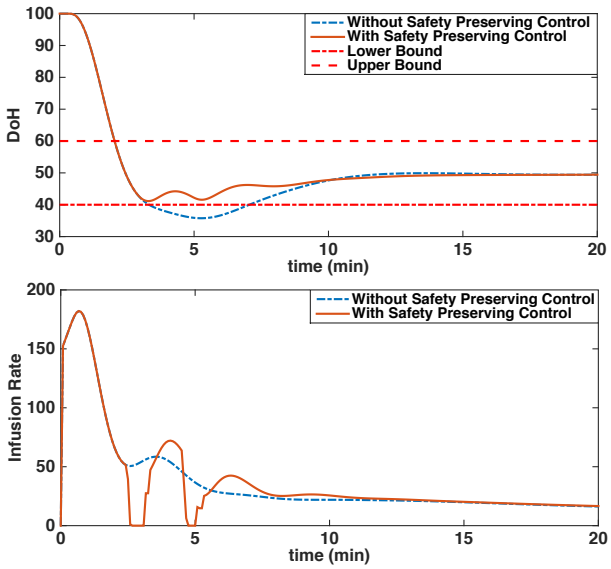


Fig. 4. The comparison of DoH and the infusion rate in the case when the safety-preserving control is on with the case when it is off. The result is associated with the $4^{th}$ model in the model set.

satisfying the constraints. Indeed, safety-preserving control brings DoH inside the limits once it approaches the lower bound. On the other hand, when DoH is within the constraint, safety- preserving control lets the PID controller operate in to satisfy the closed-loop performance. The lower plot shows how the safety-preserving control manipulates and corrects the control input provided by the PID controller to prevent the states from going outside the preset limits.

Fig. 5 depicts performance of the model-invariant safety-preserving control when it is applied to the set of PKPD models. We choose two random models from the model set to calculate the safety-preserving control input based on the model-invariant viability kernel depicted in fig. 3.

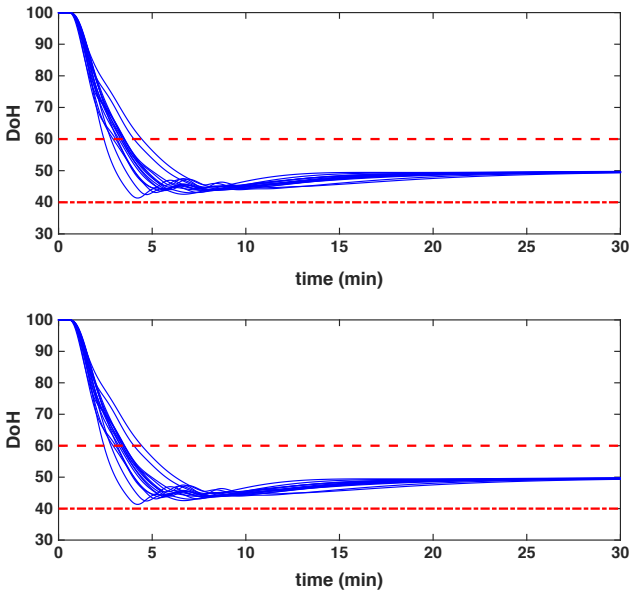


Fig. 5. Response of 15 PKPD models to MISP. In each plot, a single random model from the model set is used to control the entire population.

Fig. 5 shows that no matter which model we choose to calculate the safety-preserving control input, the proposed control approach is capable of maintaining the variables of the entire set within the safe bound.

## V. Conclusion and Future Work

In this paper, we introduced model-invariant safety-preserving control. This control approach is the extension of the safety-preserving control to the case where the model includes uncertainty. We have defined the model-invariant viability kernel as an intersection of the viability kernels of all model in the model set, followed by a proof for the model-invariant safety preserving control synthesis. We showed that using the model-invariant viability kernel, no matter which model from the model set we choose to calculate the safety-preserving control input, it is able to satisfy safety concerns. Finally, we illustrated the efficiency of the proposed technique by applying it to closed-loop control of anesthesia in which there is no unique model describing a patient's response to anesthetics.

Although we discussed model invariability of safety-preserving control, we only considered the case where the model set is finite. However, in an infinite model case or a case where the model set includes a large number of models, approximating the model-invariant viability kernel cannot be achieved by intersecting the viability kernels of all individual models. We also considered a specific uncertainty structure and do not discus the impact of time delays on safety. Moreover, we assumed states of the model are measurable. To make model-invariant safety-preserving control applicable to control of anethsia, further work on implementation of the proposed technique in the presence of output-feedback, infinite model sets, and time delays, will be needed.

## References


[1] K. Van Heusden, N. West, A. Umedaly, J. Mark. Ansermino, R. N. Merchant, and G. A. Dumont. Safety, constraints and anti-windup in closed-loop anesthesia. In *IFAC World Congress*, volume 19, pages 6569–6574, 2014.

[2] K. Margellos and J. Lygeros. Air traffic management with target windows: An approach using reachability. In *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*, pages 145–150, Dec 2009.

[3] J. Hasenauer, P. Rumschinski, S. Waldherr, S. Borchers, F. Allgöwer, and R. Findeisen. Guaranteed steady-state bounds for uncertain chemical processes. In *Proceedings of International Symposium on Advanced Control of Chemical Processes, ADCHEM'09, Istanbul, Turkey*, pages 674–679, 2009.

[4] E. F. Camacho and C. B. Alba. *Model predictive control*. Springer Science & Business Media, 2013.

[5] J. P. Aubin, A. M Bayen, and P. Saint-Pierre. *Viability theory: new directions*. Springer Science & Business Media, 2011.

[6] K. Margellos and J. Lygeros. Viable set computation for hybrid systems. *Nonlinear Analysis: Hybrid Systems*, 10:45–62, 2013.

[7] S. Kaynama, J. Maidens, M. Oishi, I. M. Mitchell, and G. A. Dumont. Computing the viability kernel using maximal reachable sets. In *Proceedings of the 15th ACM international conference on Hybrid Systems: Computation and Control*, pages 55–64. ACM, 2012.

[8] Y. Gao, J. Lygeros, and M. Quincampoix. On the reachability problem for uncertain hybrid systems. *Automatic Control, IEEE Transactions on*, 52(9):1572–1586, 2007.

[9] A. B. Kurzhanski and P. Varaiya. Ellipsoidal techniques for reachability analysis: internal approximation. *Systems & control letters*, 41(3):201–211, 2000.

[10] John Lygeros, Claire Tomlin, and Shankar Sastry. Controllers for reachability specifications for hybrid systems. *Automatica*, 35(3):349–370, 1999.

[11] S. Kaynama, I. M. Mitchell, M. Oishi, and G. A. Dumont. Scalable safety-preserving robust control synthesis for continuous-time linear systems. *Automatic Control, IEEE Transactions on*, 2015.

[12] A. Girard. Reachability of uncertain linear systems using zonotopes. In *Hybrid Systems: Computation and Control*, pages 291–305. Springer, 2005.

[13] A. Abate, M. Prandini, J. Lygeros, and S. Sastry. Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems. *Automatica*, 44(11):2724–2734, 2008.

[14] S. Summers, M. Kamgarpour, C. Tomlin, and J. Lygeros. Stochastic system controller synthesis for reachability specifications encoded by random sets. *Automatica*, 49(9):2906–2910, 2013.

[15] G. A. Dumont. Closed-loop control of anesthesia-a review. In *Biological and Medical Systems*, volume 8, pages 373–378, 2012.

[16] T. Zikov, S. Bibian, G. A. Dumont, M. Huzmezan, and C. R. Ries. Quantifying cortical activity during general anesthesia using wavelet analysis. *Biomedical Engineering, IEEE Transactions on*, 53(4):617–632, 2006.

[17] M. Yousefi, K. van Heusden, G. A. Dumont, and J. M. Ansermino. Safety-preserving closed-loop control of anesthesia. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pages 454–457. IEEE, 2015.

[18] I. M. Mitchell. A toolbox of level set methods. *Departmental of Computer Science, University of British Columbia, Vancouver, BC, Canada, http://www.cs.ubc.ca/˜mitchell/ToolboxLS/toolboxLS.pdf, Technical Report TR-2004-09*, 2004.

[19] A. B. Kurzhanski and P. Varaiya. Ellipsoidal techniques for reachability analysis. *preprint*, 1998.

[20] A. A. Kurzhanski and P. n. Varaiya. Ellipsoidal toolbox. *EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2006-46*, 2006.

[21] A. B. Kurzhanski and I. Vályi. *Ellipsoidal calculus for estimation and control*. Nelson Thornes, 1997.

[22] J. Schüttler and H. Ihmsen. Population pharmacokinetics of propofol: a multicenter study. *Anesthesiology*, 92(3):727–738, 2000.

[23] S. Bibian, G. A. Dumont, M. Huzmezan, and C. R. Ries. Patient variability and uncertainty quantification in clinical anesthesia: Part i–pkpd modeling and identification. In *IFAC Symposium on Modelling and Control in Biomedical Systems, Reims, France*, 2006.