Responses to reviewer 21

Corrections to your comments are highlighted in colors ■ and ■. In the revised version, we have made the following changes to the symbols and terminologies used in the manuscript.

Original version	Revised version
Controller,disturbance	Player I, Player II
$a \in \mathcal{A}$, $d \in \mathcal{D}$	$a \in \mathcal{A}$, $b \in \mathcal{B}$
$\mathscr{A}$, $\mathscr{D}$	$\mathscr{A}_0$, $\mathscr{B}_0$
Given time horizon $\bar{T}$	T or $T_1, ..., T_M$
Given admissible cost $\bar{J}$	J or $J_1, ..., J_M$
$\phi_{t_0}^{t_1}(., s_0, a, d)$	$\phi_{s,t}^{a,b}(.)$
$\min\limits_{a(.)} \max\limits_{d(.)} t_f$	$\sup\limits_{\beta(.)} \inf\limits_{a(.)} t_f$ (Emphasize that Player II plays nonanticipative strategies)
$J_{t_0}^{t_1}(s_0, a, d)$	$\mathcal{J}_{t_0}^{t_1}(s, a, b)$
$\min\limits_{a(.)} \max\limits_{d(.)} J_{t_0}^{t_1}(s_0, a, d)$	$\sup\limits_{\beta(.)} \inf\limits_{a(.)} \mathcal{J}_{t_0}^{t_1}(s, a, \beta[a])$
Given final time t_f	$\bar{T}$ (The symbol of the free final time t_f remains unchanged)
$\widehat{W}_{t_f}^c(.)$	$W_{\bar{T}}^c(., 0)$, $\left(W_{\bar{T}}^c(s, \tau) = \sup\limits_{\beta(.)} \inf\limits_{a(.)} \mathcal{J}_{\tau}^{\bar{T}}(s, a, \beta[a]) \right)$

1. First, a huge limitation is that based on Theorem 1, verification of $W^c(s)$ only works when $\tilde{J}$ is smaller than $\gamma * t_f$. The implication of this is that CBRT can be verified only for small cost constraints, therefore, the limitation is significant.

Response: This may be a misunderstanding caused by the symbols we use. We have used a new set of symbols in the revised version to avoid these ambiguities. In the original version, the final time t_f in Eq. (6) and (10) is free, not given, it is dependent on the initial state. The t_f in Eq. (15) and $\widehat{W}_{t_f}^c$ is not free, but given (In the modified version, the given t_f is replaced with $\bar{T}$, which can be regarded as a parameter of our algorithm, and $\widehat{W}_{t_f}^c(.)$ is replaced with $W_{\bar{T}}^c(., 0)$). Given a larger $\bar{T}$, the function$W_{\bar{T}}^c(., 0)$ can be equivalent to $W^c(.)$ in a larger region, which of course means that more recursions are required (This problem also exists in the level set method, where the HJ PDE needs to be solved over a larger time interval in order to compute the BRT of a longer time horizon). The choice of $\bar{T}$ depends on the time horizon T (or admissible cost J) of the BRT (or CBRT) to be computed, see Eq. (25) of the revision. Take Example 1 in the manuscript as an example, the following animation shows how the function $W_{\bar{T}}^c(., 0)$ varies with $\bar{T}$. It can be seen that as $\bar{T}$ increases, the range that makes function $W_{\bar{T}}^c(., 0)$ and function $W^c(.)$ equal becomes larger and larger, and the time horizon of the BRT that can be represented becomes larger and larger. Opening this pdf file with software "Adobe Acrobat Pro DC" will play this animation.

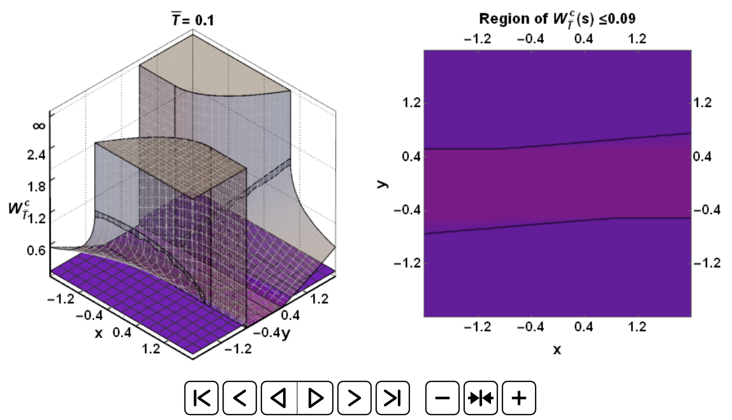


Figure 1: The variation of function $W_{\bar{T}}^c(., 0)$ with $\bar{T}$.

In the animation, the gray surface represents the analytic expression of function $W^c(.)$, i.e.:

$$W^c(s) = \begin{cases} -\sqrt{x^2+2x-2y+2}+x+1 & x^2+2x-2y+2 \geq 0 \wedge y > 0.5 \wedge x \geq -1 \\ -\sqrt{x^2-2x+2y+2}-x+1 & x^2-2x+2y+2 \geq 0 \wedge y < -0.5 \wedge x \leq 1 \\ 0 & |y| \leq 0.5 \\ \infty & \left(x^2+2x-2y+2 < 0 \vee x < -1\right) \wedge y \geq 0.5 \\ \infty & \left(x^2-2x+2y+2 < 0 \vee x > 1\right) \wedge y \leq -0.5 \end{cases} \tag{1}$$

2. Moreover, the definition of t_f is not provided in the paper (it is strongly recommended to provide its definition before equation (6)); in many parts of the paper where it is used, t_f implies different definitions and therefore, it is ambiguous whether the theory is valid. Specifically, let's assume that t_f is a first hitting time to the target set K because if it's just a general terminal time, $s(t_f) \in K$ in equation (10) is not a valid constraint for the backward reachable tube problem (it will rather be a backward reachable set problem). If t_f is a first hitting time, it should be a state (s_0) dependent function. However, it is clear that in the proof of theorem 1, t_f is used as a general terminal time. (Also side note: τ before eq (18) shouldn't be "a" $\tau \in [0, t_f) s.t. \phi(\;) \in K$, it should be a minimal time s.t. $\phi(\;) \in K$.)

Response: Some other reviewers have also raised this issue, and we apologize for any ambiguity this has caused. In the original version, the final time t_f in Eq. (6) and (10) is free, not given, it is dependent on the initial state. The t_f in Eq. (15) and $\widehat{W}^c_{t_f}$ is not free, but given (In the modified version, the given t_f is replaced with $\bar{T}$, and $\widehat{W}^c_{t_f}(.)$ is replaced with $W^c_{\bar{T}}(., 0)$). In Theorem 1, t_f is indeed given (a general terminal time) and it can be regarded as a parameter of our algorithm, so it is replaced with $\bar{T}$ in the revision. In the revision, we have clarified the definitions of both t_f and $\bar{T}$.

3. Finally, how you introduced the control signal and the disturbance signal in equation (3) is very wrong. The paper [23] does *not* propose the value function defined in equation (3), it does not concern a differential game formulation. The correct citation should be [25] in the references. It is crucial to note the difference between the definition (3) in this paper (which is wrong) and the definition (16) in [25]. Specifically, the viscosity solution of equation (5) does not provide the value function (3); the order of inf and sup is flipped, and the disturbance should be a strategy rather than a control signal itself. For more details, please refer to [25] or Evans, Souganidis, Differential games and representation formulas for solutions of Hamilton-Jacobi-Isaacs equations, 1984. Because of this, all the statements with respect to the disturbance are wrong. Please take a careful look at the relevant literature and address the statements and proofs accordingly.

Response: Yes, we take too much for granted about Eq. (3). Therefore, in the revision, we removed Eq. (3).

The order of these two symbols ("$\min, \max$") is related to the information pattern of the game. And the choice of information pattern is related to the nature of the target set. In our paper, the target set is treated as a safe area in the state space. For example, the target set of the aircraft in Example 2. This target set is the region near the equilibrium point in the aircraft state space. In general, the state of the aircraft is always kept in the neighborhood of equilibrium. Therefore, in our paper, the BRT has the meaning of "safe set". In this aspect, literature [1] (references [25] of the original manuscript) and this paper are opposite, they consider the target set as a dangerous area.

In this manuscript, the information pattern is: Player II (disturbance) plays a nonanticipative strategy. This information pattern has the following benefits: The purpose of Player I (controller) is to drive the system state to the target set, while the purpose of Player II is to avoid or delay the entry of the system state into the target set. The player who takes nonanticipative strategy has an advantage over its opponent [1]. Player II taking nonanticipative strategy can result in an underapproximation of BRT. An BRT is a safe set, a favorable set. Therefore, to take the worst case into account, we prefer to underapproximate it rather than overapproximate it.

In literature [1], the player attempting to drive the system state to the target set plays nonanticipative strategies. This is to overapproximate BRT rather than underapproximate BRT, because in [1], the BRT is an unsafe set.

In the revision, we added a description of the information pattern and detailed the reasons for choosing this information pattern.

4. In fact, the techniques you used to verify $W^c(s_0)$ are nothing new. How you convert this problem to $\hat{W}^c(s_0)$ (as in eq (12) - (15)) is already done similarly in [25] (Appendix A) to derive the corresponding HJI-VI (your eq (5)). Also, "the recursive formula" (23) that you propose in this paper is just a dynamic programming principle, which is very similar to Lemma 1 in reference [23]. Therefore, I don't see any significant novelty in this numerical method and one of your main contributions (enumeration (2) in the introduction) does not seem valid at all.

Response: Yes, the method of computing the value function in this paper is not new, but the value function itself is very different from the value function in the level set method. The HJ PDE in the level set method can be written in the following backward recursive form:

$$\begin{aligned} &V(s,T) = l(s) \\ &V(s,t-\Delta t) = \min\left[V(s,t), \min_{a\in\mathcal{A}}\max_{b\in\mathcal{B}} V\left(F(s,a,b),t\right)\right] \end{aligned} \tag{2}$$

where $s(t+\Delta t) = F(s(t),a(t),b(t))$ is the time discrete form of $\dot{s} = f(s,a,b)$. Our recursive form is:

$$\begin{aligned} &W^c_{\bar{T}}(s,\bar{T}) = 0 \\ &W^c_{\bar{T}}(s,t-\Delta t) = \min_{a\in\mathcal{A}}\max_{b\in\mathcal{B}}\left[\widehat{C}(s,a,b) + W^c_{\bar{T}}(\widehat{F}(s,a,b),t)\right] \end{aligned} \tag{3}$$

As can be seen, the two recursive forms are quite different. Such a value function brings some benefits, one of which is the reduction of storage space consumption. In the level set method, the BRT of time horizon T is represented as:

$$R_K(T) = \{s_0 \in \mathbb{R}^n | V(s_0,0) \leq 0\} \tag{4}$$

This means that to save $R_K(T)$, one needs to save value function $V(.,0)$. In this case, the storage space consumed is proportional to $\prod_{i=1}^n N_i$, where N_i is the number of grid points in the ith dimension of the Cartesian grid. For $T_1,...,T_M \in [0,\infty)$, the representations of the BRTs of these time horizons are:

$$\begin{aligned} &R_K(T_1) = \{s_0 \in \mathbb{R}^n | V(s_0,T-T_1) \leq 0\} \\ &... \\ &R_K(T_M) = \{s_0 \in \mathbb{R}^n | V(s_0,T-T_M) \leq 0\} \end{aligned} \tag{5}$$

Since the value functions $V(.,T-T_1),...,V(.,T-T_M)$ are each different, the storage space consumption required to save these BRTs is proportional to $M\prod_{i=1}^n N_i$, which is M times as much as the original. In the proposed method, only $W(.,\bar{T})$ ($\bar{T} > \max(T_1,...,T_M)$) needs to be saved to save $R_K(T_1),...,R_K(T_M)$. These BRTs are represented as:

$$\begin{aligned} &R_K(T_1) = \left\{s_0 \in \mathbb{R}^n | W^c_{\bar{T}}(s_0,0) < T_1\right\} \\ &... \\ &R_K(T_M) = \left\{s_0 \in \mathbb{R}^n | W^c_{\bar{T}}(s_0,0) < T_M\right\} \end{aligned} \tag{6}$$

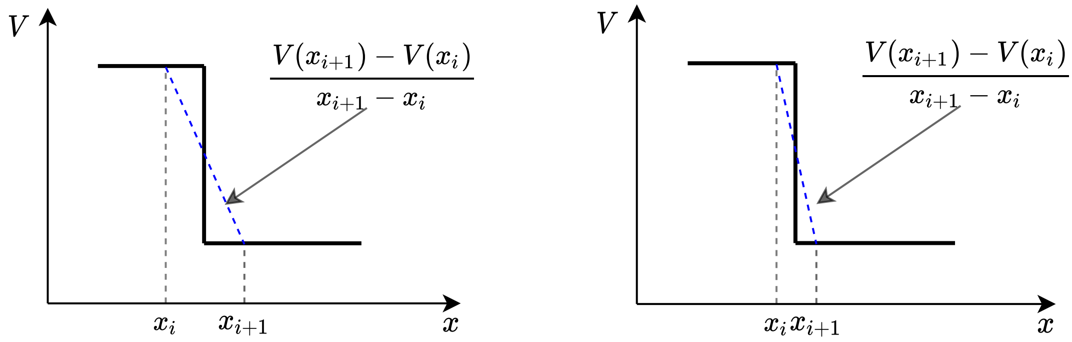
The storage space consumed does not vary with M and is always proportional to $\prod_{i=1}^n N_i$, which is one part in M of that of the level set method. We visually display the superiority of this storage approach in the first example of the revised manuscript.

Another advantage is that the CBRT can be computed by setting different running cost functions. Therefore, we modified enumeration (2) in the introduction to focus the description on the properties of the value function itself rather than on the method of computing the value function.

5. In fact, your argument that your interpolation-based method is better than solving the HJI-PDE is only supported by Table 1 and the paragraph starting with "It is worth pointing out that" in Section 4.1. However, I think for a fair comparison, you should provide the result from the level set toolbox for the

same settings and for the same different Delta ts (0.01 0.04). Do those settings have worse numerical errors? Without providing this together, it is hard to believe that your interpolation-based method gives better results than the level set toolbox. As a matter of fact, I think they are in principle doing the same thing, so I don't see any new contribution from this second part of the paper.

Response: The value function in this article has some disadvantages in addition to the benefits described above. The biggest disadvantage is that it is often discontinuous, In this case the left or right derivative cannot be defined and the difference between the grid points does not converge. Take a one-dimensional problem as an example:



(a) (b)

Figure 2:

As the grid size decreases, the difference between grid points near discontinuous points becomes larger. Therefore, this paper cannot represent the value function by the viscosity solution of HJ PDE as in the level set method. In this paper, we take a recursion and interpolation-based method because this method does not require the computation of spatial derivatives and also avoids the computation of dissipation function.

In the level set toolbox, the time step size is not specified by user, but is determined by CFL condition, which helps to avoid numerical instability. In the revision, we replaced the time-discrete form from forward Euler method to fourth-order Runge-Kutta method. The results show that the error decreases as the grid size decreases, and the convergence rate is comparable to that of the level set method. see Fig. 4 in the revision.

References

[1] I. M. Mitchell, A. M. Bayen, and C. J. Tomlin. A time-dependent hamilton-jacobi formulation of reachable sets for continuous dynamic games. *IEEE Transactions on Automatic Control*, 50(7):947–957, 2005.