

Risk-Sensitive Methods in Deception Games

Srinivas Sridharan

Dept. of Mechanical and Aerospace Engineering
University of California San Diego
Email: srsridharan@ucsd.edu

William M. McEneaney

Dept. of Mechanical and Aerospace Engineering
University of California San Diego
Email: wmceneaney@ucsd.edu

Abstract— We introduce the application of semi-field analysis to stochastic dynamic games of information. Such problems typically have a high curse-of-dimensionality; hence a method of managing this issue is developed by formulating the payoff function as a min-log sum whose structure is preserved. This technique is applied to a problem of deception and search with a risk-sensitive payoff function where the information is represented using a log-plus probability space.

I. INTRODUCTION

There has been a strong growth, in the recent past, of the application of max-plus and idempotent analysis methods to nonlinear optimal control problems. Such methods exploit the preservation of the form of the optimal cost function when it is propagated via the semi-group operator associated with the control problem. Some applications of these techniques to various problems are described in [6], [7], [11], [12].

Recently, the second author developed an idempotent method for deception games involving two players where the opponent is allowed to distort information in order to hinder the achievement of the desired objectives [9]. Such a game problem leads to a min-max value function since the opponent's strategy is assumed to be antagonistic. However, even given an opponent with a purely malevolent agenda, not all noise sources are necessarily antagonistically generated. This implies that a more suitable model would be that of a stochastic game. Of course, that added generality implies added theoretical and computational difficulties. One approach to attenuating these additional difficulties is through the use of a risk-sensitive stochastic model. This was seen to be superior to a purely stochastic model in an air tasking application in [8]. There, it is also noted that in the linear-quadratic case, the risk-sensitive stochastic control problem formulation yields the same dynamic programming equation and solution as a stochastic game formulation. Consequently, here we consider a risk-sensitive framework – obtained via a relaxation of the payoff used in the robust control approach (in which the worst case payoff is the quantity of interest). This change in the assumptions leads to an intriguing change in the problem framework and the mathematical structure used to describe it; specifically the algebra that describes this problem is no longer an idempotent semi-field.

We demonstrate that the structure of the optimal cost function is preserved in this case, thereby offering the potential to

manage the curse-of-dimensionality. However, just as in the idempotent case, the solution complexity in the risk-sensitive case grows exponentially with back-propagation step due to the use of the distributed dynamic programming (DDP) procedure. This will be handled through the use of a certain projection, taking the form of a pruning operation.

II. PROBLEM FORMULATION

We begin by formulating the deception game problem that leads in a natural way to the log-plus algebra framework. Consider a two player game (with Player 1 (P1) and Player 2 (P2)). P1 searches for the assets of P2 by choosing a set of observations. P2 is able to deceive P1 by influencing the measurements made by P1, albeit at a cost. After several repetitions of this measurement and deception actions, P1 uses the information state obtained from its various measurements to takes an action, at the terminal time, drawn from the set \mathcal{A} of possible final time actions (c.f Fig. 1). The true location of the asset of P2 is denoted by x which takes values from the set \mathcal{X} . In this article, we assume that the time steps are discrete, the action set \mathcal{A} is discrete and has finite cardinality A . i.e., it can be represented by the set $]1, A[$. Throughout, for integers $a \leq b$, $]a, b[$ denotes $\{a, a+1, \dots, b\}$. The asset location set (\mathcal{X}) of P2 is also assumed to be discrete and to have finite cardinality (and can therefore be represented without loss of generality by the set $]1, L[$). Now if the true asset location of P2 is x and P1 takes an action a , then we denote the loss incurred by P2 by $c(x, a)$. Hence P1 wishes to minimize (make more negative) the loss $c(x, a)$ and P2 attempts to maximize (make less negative) this loss. We represent the vector of values $c(x, a)$, of length L , across all possible states $x \in \mathcal{X}$ by $C(a)$.

We define q_x to be the cost to P2 to make P1 believe that the state is ' x '. Now, in the case when P2's objective is purely antagonistic and there are negligible stochastic observation errors (robust/worst case formulation), the costs are taken to be elements of a max-plus idempotent semi-field. The cost q is normalized via the condition $\max_{x \in \mathcal{X}} [q_x] = 0$ (where we note that ' 0 ' is the multiplicative identity element for the semi-field). In the case where not all noise sources are antagonistic, we use a different algebraic structure to represent and determine the risk-sensitive cost. The normalization

Research partially supported by NSF grant DMS-0808131 and AFOSR.

condition for the cost in this case is taken to be

$$\bigoplus_{x \in \mathcal{X}}^{\varepsilon} q_x = 0, \quad (1)$$

$$\text{where } \bigoplus_{x \in \mathcal{X}}^{\varepsilon} q_x := \varepsilon \log \left\{ \sum_{x \in \mathcal{X}} \exp \left[\frac{q_x}{\varepsilon} \right] \right\}.$$

Note that the max function is obtained as a limiting case of the log-plus function as $\varepsilon \rightarrow 0$. The product operation \otimes^{ε} for this new algebraic structure is taken to be the same (\otimes) as for the max-plus semi-field, namely the standard sum operation. It can be shown that the log-plus sum and product operations used above to represent the costs in this risk-sensitive case, define a log-plus semi-field. However unlike the max-plus case, this structure lacks the idempotent property. The max-plus algebra is thus obtained as a limiting case of the log-plus semi-field as $\varepsilon \rightarrow 0$. Furthermore, the expected payoff in the antagonistic case is given by

$$J(q, a) = \bigoplus_{x \in \mathcal{X}} [c(a, x) \otimes q_x].$$

Hence in the log-plus case the payoff has the form

$$J(q, a) = \bigoplus_{x \in \mathcal{X}}^{\varepsilon} [c(a, x) \otimes^{\varepsilon} q_x]. \quad (2)$$

III. ANALYSIS OF THE DECEPTION GAME AND DIMENSIONALITY REDUCTION

We now discuss some intriguing aspects of the log-plus structure and its interpretation. Firstly, if q_x takes up values in $\mathbb{R} \cup \{-\infty\}$ then the log-plus structure is a semi-field (but not a field) as there is no additive inverse operation in this case. However by extending the set of possible costs to the set $\mathbb{R} \cup \{-\infty\} \cup [\mathbb{R} + i\pi\varepsilon]$, the algebraic structure becomes a field where the log-plus additive inverse to any element $a \in \mathbb{R}$ is $a + i\pi\varepsilon$ (and vice-versa for $a + i\pi\varepsilon$). Secondly, the costs q can be viewed as a log-plus probability mass function taking up values q_x at any point $x \in \mathcal{X}$. This is intuitively supported by the fact that $q_x \in]I_a^{\varepsilon}, I_m^{\varepsilon}[$, where $I_a^{\varepsilon} := -\infty$ and $I_m^{\varepsilon} := 0$ are the additive and multiplicative identities respectively for the log-plus semi-field. The normalization constraint on the log-plus probability takes the form in Eq. (1). A more detailed introduction to a probabilistic interpretation of idempotent analysis can be found in [1], [2], [3], [10]. In addition, a risk-sensitive version for two different applications is described in [4], [5].

In the log-plus case the cost function Eq. (2) can be interpreted as a log-plus expectation of $c(a, x)$ distributed according to q_x . Hence the cost function (payoff) in Eq. (2) is

$$E^{\oplus^{\varepsilon}}[c(a, \xi)] := \bigoplus_{x \in \mathcal{X}}^{\varepsilon} [c(a, x) \otimes^{\varepsilon} q_x] := C(a) \odot^{\varepsilon} q,$$

where ξ is distributed according to q . This can be viewed as the value of the information q . The notation $E^{\oplus^{\varepsilon}}$ denotes the risk-sensitive expectation operator arising from the distribution q .

Thus the payoff that P1 wishes to minimize (the value of information q at the terminal time T which is the cost to P2) is

$$\phi(q) := \min_{a \in \mathcal{A}} J(q, a) = \bigwedge_{a \in \mathcal{A}} [C(a) \odot^{\varepsilon} q]. \quad (3)$$

Now if ' q ' is the log-plus probability distribution after observation at time t . Then the cost for any true state $x \in \mathcal{X}$ is

$$\begin{aligned} \hat{q}_x(t+1) &= p(y|x; \hat{u}) + q_x(t), \\ &= p(y|x; \hat{u}) \otimes^{\varepsilon} q_x(t). \end{aligned}$$

However this can be normalized such that Eq. (1) holds. This normalized cost (denoted \hat{q}) is obtained from the unnormalized \hat{q} using the expression

$$\begin{aligned} q_x(t+1) &= p^{\oplus^{\varepsilon}}(y|x; \hat{u}) \otimes^{\varepsilon} \hat{q}_x(t) - \\ &\quad \bigoplus_{\xi \in \mathcal{X}}^{\varepsilon} \left[p^{\oplus^{\varepsilon}}(y|\xi; \hat{u}) \otimes^{\varepsilon} \hat{q}_{\xi}(t) \right]. \end{aligned} \quad (4)$$

Note that given the above form of the cost, the normalization condition in Eq. (1) holds due to the following result.

Lemma 1 (Distributive property):

$$\bigoplus_{x \in \mathcal{X}}^{\varepsilon} [g_x \otimes^{\varepsilon} M] = M \otimes^{\varepsilon} \left[\bigoplus_{x \in \mathcal{X}}^{\varepsilon} [g_x] \right], \quad (5)$$

for any finite sequence, g ., and any constant, M .

Proof: From the definition of \oplus^{ε} we have

$$\begin{aligned} \bigoplus_{x \in \mathcal{X}}^{\varepsilon} [g_x \otimes^{\varepsilon} M] &= \varepsilon \log \left[\sum_{x \in \mathcal{X}} \exp \left\{ \frac{g_x + M}{\varepsilon} \right\} \right] \\ &= \varepsilon \log \left[\sum_{x \in \mathcal{X}} \exp(g_x/\varepsilon) \times \exp(M/\varepsilon) \right] \\ &= M + \left[\bigoplus_{x \in \mathcal{X}}^{\varepsilon} g_x \right] = M \otimes^{\varepsilon} \left[\bigoplus_{x \in \mathcal{X}}^{\varepsilon} g_x \right]. \end{aligned} \quad (6)$$

The fact that Eq. (4) satisfies the normalization condition can be seen by applying this result after setting

$$M = - \bigoplus_{\xi \in \mathcal{X}}^{\varepsilon} [p^{\oplus^{\varepsilon}}(y|\xi; \hat{u}) \otimes^{\varepsilon} \hat{q}_{\xi}(t)].$$

At the terminal time the optimal cost function has the form

$$V(T, q) = \phi(q) = \bigwedge_{a \in \mathcal{A}} [C(a) \odot^{\varepsilon} q]. \quad (7)$$

We now demonstrate that this form of the cost function is preserved under propagation by the dynamic programming propagation operator over each time step. This is a core result as it helps enable the efficient computation of the solution to this optimal control problem. This structure preservation approach has in fact, also been exploited to tackle a range of optimal control problems (c.f., [9], [11]) where such a idempotent field structures arise.

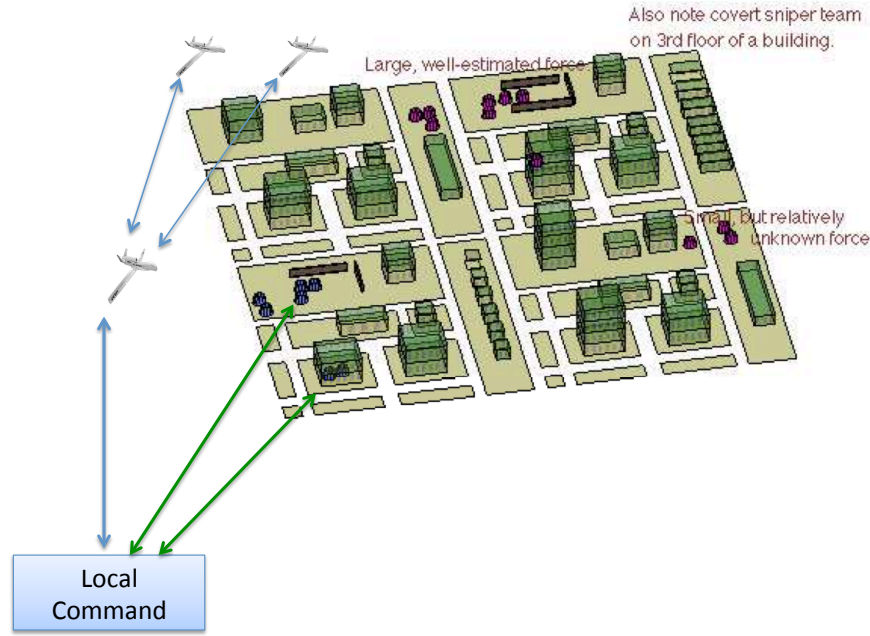


Fig. 1. A scenario of interest that demonstrates an instance of a deception game. Here player 1 (P1) could direct the UAVs depending on the choice of measurements to be made. Player 2 (P2) can influence the observations of P1 and P1 takes actions depending on the measurements obtained.

The DPP applied to Eq. (7) yields

$$\begin{aligned} V(t, q) &= \bigwedge_{u \in \mathcal{U}} E^{\oplus^\varepsilon} [V(t+1, q(t+1))], \\ &= \bigwedge_{u \in \mathcal{U}} \bigoplus_{y \in \mathcal{Y}}^\varepsilon [V(t+1, q(t+1)) \otimes^\varepsilon p_{t+1}^{\oplus^\varepsilon}(y)]. \end{aligned}$$

Here $p_t^{\oplus^\varepsilon}(y)$ is the max-plus probability of observing y at time t . The probability distribution $q_{(\cdot)}(t+1)$ is generated from $q_{(\cdot)}(t)$ using Eq. (4). In order to indicate the preservation of this structure, we suppose that at a time $t+1$ the form of the value function is given by

$$V(t+1, q) = \bigwedge_{z \in \mathcal{Z}_{t+1}} [d_{t+1}(z) \odot^\varepsilon q], \quad (8)$$

for some index set \mathcal{Z}_{t+1} of finite cardinality. This structure is intuitively inspired by the form of the value function Eq. (3) and the approach in [9]. Here \mathcal{Z}_{t+1} is an index set with associated coefficient vectors $d_{t+1}(z)$. Thus

$$\begin{aligned} V(t, q) &= \bigwedge_{u \in \mathcal{U}} \bigoplus_{y \in \mathcal{Y}}^\varepsilon \left\{ \bigwedge_{z \in \mathcal{Z}_{t+1}} \left[\left(d_{t+1}(z) \odot^\varepsilon q(t+1) \right) \right. \right. \\ &\quad \left. \left. \otimes^\varepsilon p_{t+1}^{\oplus^\varepsilon}(y) \right] \right\} \end{aligned} \quad (9)$$

Note that

$$p_{t+1}^{\oplus^\varepsilon}(y) = \bigoplus_{\xi \in \mathcal{X}}^\varepsilon p^{\oplus^\varepsilon}(y|\xi; \hat{u}) \otimes^\varepsilon q_\xi(t). \quad (10)$$

Applying (10) to (9) yields

$$\begin{aligned} V(t, q) &= \bigwedge_{u \in \mathcal{U}} \bigoplus_{y \in \mathcal{Y}}^\varepsilon \bigwedge_{z \in \mathcal{Z}_{t+1}}^\varepsilon \left[[d_{t+1}(z) \odot^\varepsilon q(t+1)] \otimes^\varepsilon \right. \\ &\quad \left. \bigoplus_{\xi \in \mathcal{X}}^\varepsilon \left(p^{\oplus^\varepsilon}(y|\xi; \hat{u}) \otimes^\varepsilon q_\xi \right) \right]. \end{aligned} \quad (11)$$

For consistency of notation we denote \oplus^ε by \vee^ε and \bigoplus^ε by \bigvee^ε . Hence Eq. (11) becomes

$$\begin{aligned} V(t, q) &= \bigwedge_{u \in \mathcal{U}} \bigvee_{y \in \mathcal{Y}}^\varepsilon \bigwedge_{z \in \mathcal{Z}_{t+1}}^\varepsilon \left[[d_{t+1}(z) \odot^\varepsilon q(t+1)] \otimes^\varepsilon \right. \\ &\quad \left. \bigoplus_{\xi \in \mathcal{X}}^\varepsilon \left[p^{\oplus^\varepsilon}(y|\xi; \hat{u}) \otimes^\varepsilon q_\xi \right] \right]. \end{aligned} \quad (12)$$

We can rewrite all components of Eq. (4) into a single equation of the form

$$q(t+1) = D^{y, \hat{u}} \otimes^\varepsilon q(t) \oslash^\varepsilon \left\{ \bigoplus_{\xi \in \mathcal{X}}^\varepsilon \left[p^{\oplus^\varepsilon}(y|\xi; \hat{u}) \otimes^\varepsilon q_\xi \right] \right\}, \quad (13)$$

where \oslash^ε denotes log-plus division. Here $D^{y, \hat{u}}$ is a diagonal matrix with diagonal terms

$$D_{x, x}^{y, \hat{u}} = p^{\oplus^\varepsilon}(y|x; \hat{u}),$$

and the remaining terms set to the log-plus additive identity $\{-\infty\}$. Substituting Eq. (13) into Eq. (12) and canceling

terms yields

$$V(t, q) = \bigwedge_{u \in \mathcal{U}} \bigvee_{y \in \mathcal{Y}} \left\{ \bigwedge_{z \in \mathcal{Z}_{t+1}} \left[\hat{d}_t(z, y, u) \odot^\varepsilon q \right] \right\}, \quad (14)$$

where $\hat{d}_t(z, y, u) := [D^{y,u}]^T \bullet d_{t+1}(z) = D^{y,u} \bullet d_{t+1}(z)$. Here the operation \bullet is interpreted as a log-plus matrix vector product.

In order to proceed with transforming Eq. (14) to the simpler structural form in Eq. (8) we now prove that

Lemma 2:

$$\bigvee_{y \in \mathcal{Y}} \bigwedge_{z \in \mathcal{Z}_{t+1}} [\hat{d}_t(z, y, u) \odot^\varepsilon q] = \bigwedge_{\tilde{z} \in \tilde{\mathcal{Z}}_t} [\tilde{d}_t(\tilde{z}, u) \odot^\varepsilon q], \quad (15)$$

where

$$\begin{aligned} \tilde{d}_t(\tilde{z}, u) &:= \bigoplus_{y \in \mathcal{Y}} \hat{d}_t(\tilde{z}_y, y, u), \\ \tilde{\mathcal{Z}}_t &:= \{\tilde{z} = \{\tilde{z}_y\}_{y \in \mathcal{Y}} | \tilde{z}_y \in \mathcal{Z}_{t+1}, \forall y \in \mathcal{Y}\}. \end{aligned} \quad (16)$$

Proof: The notation \tilde{z} in the statement of the result denotes a particular action strategy that returns an action \tilde{z}_y for any possible observation y (drawn from the set \mathcal{Y}). We first simplify the symbols used in the proof of this result as follows.

Noting that the result holds for any time t and choice of sensing action u , we define

$$\begin{aligned} \tilde{a}_z &:= \tilde{d}_t(\tilde{z}, u) \odot^\varepsilon q, \\ \hat{a}_{y,z} &:= \hat{d}_t(z, y, u) \odot^\varepsilon q. \end{aligned}$$

We can thus reformulate the statement in Eq. (15) into the form

$$\bigvee_{y \in \mathcal{Y}} \bigwedge_{z \in \mathcal{Z}_{t+1}} \hat{a}_{y,z} = \bigwedge_{\tilde{z} \in \tilde{\mathcal{Z}}_t} \tilde{a}_{\tilde{z}}. \quad (17)$$

Without loss of generality we assume that the observations \mathcal{Y} and the set of actions \mathcal{Z}_{t+1} take on values which are parameterized by a set of integers $\{1, 2, \dots, I\}$ and $\{1, 2, \dots, J\}$ respectively. This is valid due to the assumptions that the observations (and actions) are drawn from a discrete set of finite cardinality. The LHS of Eq. (17) is

$$\begin{aligned} \bigvee_{y \in \mathcal{Y}} \bigwedge_{z \in \mathcal{Z}_{t+1}} \hat{a}_{y,z} &= \\ &(\hat{a}_{11} \wedge \hat{a}_{12} \wedge \dots \wedge \hat{a}_{1J}) \vee^\varepsilon \dots \vee^\varepsilon (\hat{a}_{I1} \wedge \hat{a}_{I2} \wedge \dots \wedge \hat{a}_{IJ}). \end{aligned} \quad (18)$$

Here $\hat{a}_{1,1}$ denotes the value \hat{a}_{1,z_1} i.e., the value of \hat{a} corresponding to the output y_1 (indexed by 1) and the action z_1 (indexed by 1). Similarly $\hat{a}_{1,2}$ is the value of \hat{a} for the output y_1 with the action used being z_2 (indexed by 2). The RHS of Eq. (17) is of the form

$$\begin{aligned} \bigwedge_{\tilde{z} \in \tilde{\mathcal{Z}}} \tilde{a}_{\tilde{z}} &= \bigwedge_{\tilde{z} \in \tilde{\mathcal{Z}}} \left\{ \tilde{d}_t(\tilde{z}, u) \odot^\varepsilon q \right\} \\ &= \bigwedge_{\tilde{z} \in \tilde{\mathcal{Z}}} \left\{ \left[\bigoplus_{y \in \mathcal{Y}} \hat{d}_t(\tilde{z}_y, y, u) \right] \odot^\varepsilon q \right\}. \end{aligned}$$

Applying the distributive property from lemma 1 this is

$$= \bigwedge_{\tilde{z} \in \tilde{\mathcal{Z}}} \bigoplus_{y \in \mathcal{Y}}^\varepsilon \left[\hat{d}_t(\tilde{z}_y, y, u) \odot^\varepsilon q \right].$$

Using \vee^ε to denote \bigoplus^ε in the preceding equation, it follows that the RHS of Eq. (17) is

$$\bigwedge_{\tilde{z} \in \tilde{\mathcal{Z}}} \tilde{a}_{\tilde{z}} = \bigwedge_{\tilde{z} \in \tilde{\mathcal{Z}}} \bigvee_{y \in \mathcal{Y}}^\varepsilon \left[\hat{d}_t(\tilde{z}_y, y, u) \odot^\varepsilon q \right]. \quad (19)$$

This proof will proceed by demonstrating that the RHS Eq. (19) and LHS Eq. (18) are equal.

We start by considering the RHS. Now, for every strategy \tilde{z} , there is a product $b_1^J(\tilde{z})$ defined as

$$b_1^J(\tilde{z}) := [\hat{a}_{1,\tilde{z}_1} \vee^\varepsilon \hat{a}_{2,\tilde{z}_2} \dots \vee^\varepsilon \hat{a}_{J,\tilde{z}_J}]. \quad (20)$$

Rewriting Eq. (20) as

$$[\hat{a}_{1,\tilde{z}_1} \vee^\varepsilon \hat{a}_{2,\tilde{z}_2} \dots \vee^\varepsilon \hat{a}_{J,\tilde{z}_J}] = \hat{a}_{1,\tilde{z}_1} \vee^\varepsilon b_2^J(\tilde{z}),$$

where

$$b_2^J(\tilde{z}) := \hat{a}_{2,\tilde{z}_2} \vee^\varepsilon \hat{a}_{3,\tilde{z}_3} \dots \vee^\varepsilon \hat{a}_{J,\tilde{z}_J}, \quad (21)$$

and using this in Eq. (19) yields

$$\bigwedge_{\tilde{z} \in \tilde{\mathcal{Z}}} [\hat{a}_{1,\tilde{z}_1} \vee^\varepsilon \hat{a}_{2,\tilde{z}_2} \dots \vee^\varepsilon \hat{a}_{J,\tilde{z}_J}] = \bigwedge_{\substack{z \in \mathcal{Z}_{t+1} \\ b \in \mathcal{B}_2^J}} \left[\hat{a}_{1,z} \vee^\varepsilon b \right]. \quad (22)$$

Here \mathcal{B}_k^J is defined to be the set of log-plus products of the form $\hat{a}_{k,\tilde{z}_k} \vee^\varepsilon \hat{a}_{k+1,\tilde{z}_{k+1}} \dots \vee^\varepsilon \hat{a}_{J,\tilde{z}_J}$.

We note from the definition of \vee^ε that it is monotonic in each of its terms, i.e., for any $a \in \mathbb{R}^-$

$$\begin{aligned} a \vee^\varepsilon x &\leq a \vee^\varepsilon y, & \forall x \leq y, \\ \text{and } x \vee^\varepsilon a &\leq y \vee^\varepsilon a, & \forall x \leq y. \end{aligned} \quad (23)$$

Using Eqns. (22)-(23) in Eq. (19) yields

$$\bigwedge_{\tilde{z} \in \tilde{\mathcal{Z}}} \bigvee_{y \in \mathcal{Y}}^\varepsilon \hat{a}_{y,\tilde{z}_y} = \bigwedge_{b \in \mathcal{B}_2^J} \left\{ \left[\bigwedge_{z \in \mathcal{Z}_{t+1}} \hat{a}_{1,z} \right] \vee^\varepsilon b \right\}.$$

Denoting $\bigwedge_{z \in \mathcal{Z}_{t+1}} \hat{a}_{1,z}$ by \mathcal{A}_1 and using the distributive property demonstrated in lemma 1, we can rewrite the above as

$$\bigwedge_{b \in \mathcal{B}_2^J} (\mathcal{A}_1 \vee^\varepsilon b) = \mathcal{A}_1 \vee^\varepsilon \left[\bigwedge_{b \in \mathcal{B}_2^J} b \right]. \quad (24)$$

We can now rewrite Eq. (21) as follows

$$b_2^J(\tilde{z}) = \hat{a}_{2,\tilde{z}_2} \vee^\varepsilon b_3^J(\tilde{z}), \quad (25)$$

$$\text{where } b_3^J(\tilde{z}) := \hat{a}_{3,\tilde{z}_3} \vee^\varepsilon \hat{a}_{4,\tilde{z}_4} \dots \vee^\varepsilon \hat{a}_{J,\tilde{z}_J}.$$

Therefore using Eq. (25) and Eq. (24), we write Eq. (19) as

$$\bigwedge_{\tilde{z} \in \tilde{\mathcal{Z}}} \bigvee_{y \in \mathcal{Y}}^\varepsilon \hat{a}_{y,\tilde{z}_y} = \mathcal{A}_1 \vee^\varepsilon \left[\bigwedge_{b \in \mathcal{B}_2^J} b \right] = \mathcal{A}_1 \vee^\varepsilon \left[\bigwedge_{\substack{z \in \mathcal{Z}_{t+1} \\ b \in \mathcal{B}_3^J}} (\hat{a}_{2,z} \vee^\varepsilon b) \right], \quad (26)$$

which, using the monotonicity of the \vee^ε operator as before, becomes

$$= \mathcal{A}_1 \vee^\varepsilon \bigwedge_{b \in \mathcal{B}_3^I} \left\{ \left[\bigwedge_{z \in \mathcal{Z}_{t+1}} \hat{a}_{2,z} \right] \vee^\varepsilon b \right\},$$

and denoting $\bigwedge_{z \in \mathcal{Z}} \hat{a}_{2,z}$ by \mathcal{A}_2 this is

$$= \mathcal{A}_1 \vee^\varepsilon \mathcal{A}_2 \vee^\varepsilon \left[\bigwedge_{b \in \mathcal{B}_3^I} b \right]. \quad (27)$$

From Eqns. (26), (27) it is seen that by repeating this procedure we have that Eq. (19) can be written as

$$\begin{aligned} \mathcal{A}_1 \vee^\varepsilon \mathcal{A}_2 \vee^\varepsilon \dots \mathcal{A}_J &= \left(\bigwedge_{z \in \mathcal{Z}_{t+1}} \hat{a}_{1,z} \right) \vee^\varepsilon \left(\bigwedge_{z \in \mathcal{Z}_{t+1}} \hat{a}_{2,z} \right) \vee^\varepsilon \dots \left(\bigwedge_{z \in \mathcal{Z}_{t+1}} \hat{a}_{J,z} \right) \\ &= \bigvee_{y \in \mathcal{Y}} \bigwedge_{z \in \mathcal{Z}_{t+1}} \hat{a}_{y,z}. \end{aligned} \quad (28)$$

This Eq. (28) is equal to Eq. (18) and hence the statement of the result follows. ■

We now obtain the required result on the preservation of the structure of the payoff function.

Corollary 3: Given the payoff function Eq. (8) at time $t+1$, the payoff function in Eq. (14) has the form

$$V(t, q) = \bigwedge_{z \in \mathcal{Z}_t} [d_t(z) \odot^\varepsilon q], \quad (29)$$

Proof: Applying Lemma 2 to Eq. (14) we have

$$\begin{aligned} V(t, q) &= \bigwedge_{u \in \mathcal{U}} \bigvee_{y \in \mathcal{Y}} \left\{ \bigwedge_{z \in \mathcal{Z}_{t+1}} \left[\hat{d}_t(z, y, u) \odot^\varepsilon q \right] \right\} \\ &= \bigwedge_{u \in \mathcal{U}} \bigwedge_{\tilde{z} \in \mathcal{Z}_t} [\tilde{d}_t(\tilde{z}, u) \odot^\varepsilon q]. \end{aligned}$$

By identifying \mathcal{Z}_t with $\mathcal{U} \times \mathcal{Z}_t$ this is

$$= \bigwedge_{z \in \mathcal{Z}_t} [d_t(z) \odot^\varepsilon q], \quad (30)$$

where the $d_t(z)$ terms correspond to $\tilde{d}_t(\tilde{z}, u)$. ■

Thus, it is seen that the structure of the optimal payoff function is preserved. This feature of the problem helps manage the rate of growth of memory and computational time requirements, as the solution is propagated backwards from the terminal time, during the numerical solution procedure for the optimal control problem. This leads to the potential for efficiently obtaining the solution to the risk-sensitive deception game problem.

IV. CONCLUSIONS

In this article we describe a risk sensitive analogue to the idempotent field approach for the representation of deception games. It was proved that the value function has a invariant property with respect to the dynamic programming operator - a feature that enables us to avoid the curse of dimensionality. However there still remains a curse of complexity i.e., a growth in the number of control, sensing actions that cause a growth in the number of parameters to be stored while computing the value function. The avenues for future work are therefor to develop an approach to reduce this growth in

complexity by removing certain parameters (termed pruning) via ordering their contribution to achieving the optimal value function. Furthermore there exists a need to obtain error analysis for such pruning.

REFERENCES

- [1] M. Akian, Densities of idempotent measures and large deviations, *Transactions of the American Mathematical Society*, 351 (1999), 4515–4543.
- [2] P. Del Moral and M. Doisy, Maslov Idempotent Probability Calculus, I. *Theory of Probability and its Applications*, 43:562, 1999.
- [3] P. Del Moral and M. Doisy, Maslov Idempotent Probability Calculus, II. *Theory of Probability and its Applications*, 44:319, 2000.
- [4] W.H. Fleming, H. Kaise, and S.J. Sheu, Max-plus stochastic control and risk-sensitivity. *Applied Mathematics and Optimization*, 62(1):81–144, 2010.
- [5] W.M. McEneaney and C.D. Charalambous, Large deviations theory, induced log-plus and max-plus measures and their applications. *Mathematical Theory of Network and Systems (2000)*, 19–23.
- [6] W.M. McEneaney, *Max-plus methods for nonlinear control and estimation*. Birkhauser, 2006.
- [7] W.M. McEneaney, A. Deshpande, and S. Gaubert, Curse-of-complexity attenuation in the curse-of-dimensionality-free method for HJB PDEs. In *American Control Conference (2008)*, 4684–4690.
- [8] W.M. McEneaney, B.G. Fitzpatrick and I.G. Lauko, “Stochastic Game Approach to Air Operations”, *IEEE Trans. Aerospace and Electronic Systems*, 40 (2004), 1191–1216.
- [9] W.M. McEneaney, Idempotent Method for Deception Games and Complexity Attenuation. In *Proc. 18th IFAC Conf. World Congress, Milano*, 2011 (to appear).
- [10] T. Puhalskii, *Large Deviations and Idempotent Probability*, Chapman & Hall/CRC, 2001.
- [11] S. Sridharan, M. Gu, M.R. James, and W.M. McEneaney, An efficient computational method for the optimal control of higher dimensional quantum systems. In *Proceedings of the, 49th IEEE Conference on Decision and Control (2010)*, 2996–3001.
- [12] S. Sridharan, M. Gu, Matthew R. James, and W.M. McEneaney. Reduced-complexity numerical method for optimal gate synthesis. *Physical Review A (Atomic, Molecular, and Optical Physics)*, 82(4), 2010.