Idempotent Expansions for Continuous-Time Stochastic Control

Hidehiro Kaise
Graduate School of Information Science
Nagoya University
Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan
Email: kaise@is.nagoya-u.ac.jp

William M. McEneaney
Department of Mechanical and Aerospace Engineering
University of California San Diego
San Diego, CA 92093-0411, USA
Email: wmceneaney@ucsd.edu

I. Introduction

It is now well-known that many classes of deterministic control problems may be solved by max-plus or min-plus (more generally, idempotent) numerical methods. These methods include max-plus basis-expansion approaches [1], [2], [6], [9], as well as the more recently developed curse-of-dimensionality-free methods [9], [14]. It has recently been discovered that idempotent methods are applicable to stochastic control and games. The methods are related to the above curse-of-dimensionality-free methods for deterministic control. In particular, a min-plus based method was found for stochastic control problems [10], [15], and a min-max method was discovered for games [11].

The first such methods for stochastic control were developed only for discrete-time problems. The key tools enabling their development were the idempotent distributive property and the fact that certain solution forms are retained through application of the semigroup operator (i.e., the dynamic programming principle operator). In particular, under certain conditions, pointwise minima of affine and quadratic forms pass through this operator. As the operator contains an expectation component, this requires application of the idempotent distributive property. In the case of finite sums and products, this property looks like our standard-algebra distributive property; in the infinitesimal case, it is familiar to control theorists through notions of strategies, non-anticipative mappings and/or progressively measurable controls. Using this technology, the value function can be propagated backwards with a representation as a pointwise minimum of quadratic or affine forms.

Here, we will remove the severe restriction to discrete-time problems. This extension requires overcoming significant technical hurdles. First, note that as these methods are related to the max-plus curse-of-dimensionality-free methods of deterministic control, there will be a discretization over time, but not over space. We will first define a parameterized set of operators, approximating the dynamic programming operator. We obtain the solutions to the problem of backward propagation by repeated application of the approximating operators. These solutions are parameterized by the time-discretization step size. Using techniques from the theory of viscosity solutions, we show that the solutions converge to the viscosity solution of the Hamilton-Jacobi-Bellman partial differential equation (HJB PDE) associated with the original problem.

The problem is now reduced to backward propagation by these approximating operators. The min-plus distributive property is employed. A generalization of this distributive property, applicable to continuum versions will be obtained. This will allow interchange of expectation over normal random variables (and other random variables with range in $I\!R^m$) with infimum operators. At each time-step, the solution will be represented as an infimum over a set of quadratic forms. Use of the min-plus distributive property will allow us to maintain that solution form as one propagates backward in time. Backward propagation is reduced to simple standard-sense linear algebraic operations for the coefficients in the representation. We also demonstrate that the assumptions on the representation which allow one to propagate backward one step are inherited by the representation at the next step. The difficulty with the approach is an extreme curse-of-complexity, wherein the number of terms in the min-plus expansion grows very rapidly as one propagates. The complexity growth will be attenuated via projection onto a lower dimensional min-plus subspace at each time step. At each step, one desires to project onto the optimal subspace relative to the solution approximation. That is, the subspace is not set a priori. In the discrete-time case, it has been demonstrated that for some problem classes, this approach is substantially superior to grid-based methods. Simple numerical examples with continuous-time dynamics will be examined with this new approach.

II. Problem Definition and Dynamic Program

We begin by defining the specific class of problems which will be addressed here. Let the dynamics take the form

$$d\xi_s = f(\xi_s, u_s, \mu_s)\, ds + \sigma(\xi_s, u_s, \mu_s)\, dB_s, \tag{1}$$

$$\xi_t = x \in I\!R^n \tag{2}$$

where f is measurable, with more assumptions on it to follow. The u_s and μ_s will be control inputs taking values in $U \subset I\!R^p$ and $\mathcal{M} =]1, M[= \{1, 2, \cdots, M\}$, respectively. In practice, we often find it useful to allow both a continuum-valued control component and a finite set-valued component, where the latter is used to allow approximation of more general nonlinear Hamiltonians, c.f. [9] for motivation. Also, $\{B_\cdot, \mathcal{F}_\cdot\}$ is an l-dimensional Brownian motion on the probability space $(\Omega, \mathcal{F}, P)$, where $\mathcal{F}_0$ contains all the P-negligible elements of $\mathcal{F}$ and σ is an $n \times l$ matrix-valued diffusion coefficient. We will be examining a finite time-horizon formulation, with terminal time, T, and will take initial time $t \in [0, T]$.

The payoff (to be minimized) will be

$$J(t,x,u.,\mu.) \doteq \mathbf{E}\left\{\int_t^T l(\xi_s,u_s,\mu_s)\,ds + \Psi(\xi_T)\right\} \quad (3)$$

where

$$\Psi(x) \doteq \inf_{z_T \in Z'_T} \{g_T(x, z_T)\}, \quad (4)$$

where l and the g_T are measurable, and $(Z'_T, d_{Z'_T})$ is a separable metric space. The value function is

$$V(t,x) = \inf_{u.\in\mathcal{U}_t, \mu.\in\widetilde{\mathcal{M}}_t} J(t,x,u.,\mu.), \quad (5)$$

where $\mathcal{U}_t$ (*resp.* $\widetilde{\mathcal{M}}_t$) is the set of $\mathcal{F}_t$-progressively measurable controls, taking values in U (*resp.* $\mathcal{M}$), such that there exists a strong solution to (1), (2).

We will assume that the given data in the dynamics and the payoff satisfy the following conditions:

(A1) U is compact.

(A2) There exist L, $K > 0$ such that for any x, $x' \in I\!R^n$, u, $u' \in U$, m, $m' \in \mathcal{M}$,

$$\begin{aligned}
&|f(x,u,m) - f(x',u',m')| + \|\sigma(x,u,m) - \sigma(x',u',m')\| \\
&\quad \le L|x-x'| + L(1+|x|+|x'|)(|u-u'|+|m-m'|), \\
&|l(x,u,m) - l(x',u',m')| \le L(1+|x|+|x'|)|x-x'| \\
&\quad +L(1+|x|^2+|x'|^2)(|u-u'|+|m-m'|), \\
&|f(x,u,m)| + \|\sigma(x,u,m)\| \le K(1+|x|), \\
&|l(x,u,m)| \le K(1+|x|^2).
\end{aligned}$$

(A3) There exist $\hat{L}$, $\hat{K} > 0$ such that for any x, $x' \in I\!R^n$,

$$\begin{aligned}
&|\Psi(x) - \Psi(x')| \le L(1+|x|+|x'|)|x-x'|, \\
&|\Psi(x)| \le K(1+|x|^2).
\end{aligned}$$

It is seen that $V(t,x)$ can be characterized as the viscosity solution of the HJB PDE associated with (1), (2), (3) (see [7] for such discussion). Indeed, $V(t,x)$ satisfies the dynamic programming principle (DPP)

$$V(t,x) = \inf_{u.\in\mathcal{U}_t, \mu.\in\widetilde{\mathcal{M}}_t} \mathbf{E}\left[\int_t^s l(\xi_r,u_r,\mu_r)dr + V(s,\xi_s)\right], \quad 0 \le t < s \le T,\ x \in I\!R^n.$$

By using the notion of viscosity solutions, it can be shown that $V(t,x)$ is a viscosity solution of

$$\begin{aligned}
&\frac{\partial V}{\partial t} + \mathcal{H}(x, D_xV(t,x), D_x^2V(t,x)) = 0 \text{ in } (0,T)\times I\!R^n, \\
&V(T,x) = \Psi(x),
\end{aligned} \quad (6)$$

where

$$\begin{aligned}
&\mathcal{H}(x, D_xV(t,x), D_x^2V(t,x)) \\
&= \inf_{u\in U}\min_{m\in\mathcal{M}}\left\{\frac{1}{2}\mathrm{tr}(\sigma\sigma^T(x,u,m)D_x^2V(t,x)) \right. \\
&\qquad \left. + f(x,u,m)\cdot D_xV(t,x) + l(x,u,m)\right\}.
\end{aligned}$$

Since $V(t,x)$ is quadratically growing on x, i.e., there exists $K > 0$ such that

$$|V(t,x)| \le K(1+|x|^2),\ (t,x) \in [0,T]\times I\!R^n, \quad (7)$$

$V(t,x)$ is the unique viscosity solution in such class (c.f. [5]).

To approximate the viscosity solution of (6) by discrete-time stochastic control problems, we introduce a family of parameterized operators $\{F_{t,s}\}_{t<s}$ defined by

$$\begin{aligned}
F_{t,s}\phi(x) = &\inf_{u\in U, m\in\mathcal{M}}\{l(x,u,m)(s-t) \\
&+\mathbf{E}[\phi(x+f(x,u,m)(s-t)+\sigma(x,u,m)(B_s-B_t))]\}.
\end{aligned} \quad (8)$$

See [3] for general viscosity techniques for approximations of second order PDEs under strong assumptions. Let $\pi_N = \{t_0 = 0 < t_1 < \cdots < t_N = T\}$ be a partition of $[0,T]$ with the step size $t_{i+1} - t_i = T/N$ $(i = 0, 1, \cdots, N-1)$. We define a discrete-time value function $V^N(t,x)$ ($(t,x) \in [0,T] \times I\!R^n$) associated with π_N recursively backward in time:

$$V^N(t,x) = \begin{cases} \Psi(x), & t = T, x \in I\!R^n, \\ F_{t,t_{i+1}}V^N(t_{i+1},\cdot)(x), & t_i \le t < t_{i+1}, x \in I\!R^n \end{cases} \quad (9)$$

where $F_{t,t_{i+1}}V^N(t_{i+1},\cdot)(x)$ is $F_{t,t_{i+1}}\phi(x)$ with $\phi(\cdot) = V^N(t_{i+1},\cdot)$. Under (A1)–(A3), we can obtain the uniform estimates of $V^N(t,x)$.

Proposition 2.1: Suppose that (A1)–(A3) hold. There exists $K > 0$, which does not depend on partition π_N, such that for x,$x' \in I\!R^n$, $t, s \in [0,T]$, $N = 1, 2, \cdots$,

$$\begin{aligned}
&|V^N(t,x) - V^N(t,x')| \le K(1+|x|+|x'|)|x-x'|, \\
&|V^N(t,x) - V^N(s,x)| \le K(1+|x|^2)|t-s|^{1/2}.
\end{aligned}$$

This proposition can be shown in a straightforward way. Since the full argument is tedious, we omit the proof.

As a corollary of Proposition 2.1, we have

Corollary 2.2: There exists a subsequence $\{V^{N_k}(t,x)\}$ and a continuous function $W(t,x)$ on $[0,T]\times I\!R^n$ such that $V^{N_k}(t,x)$ converges to $W(t,x)$ uniformly on each compact set of $[0,T]\times I\!R^n$ as $N_k \to \infty$.

To relate the limit with the unique viscosity solution of (6), we note that the infinitesimal generator of $\{F_{t,s}\}$ is $\mathcal{H}$. More precisely, we can show that for any smooth function $\varphi(t,x)$ with bounded $\partial^2\varphi/\partial t^2$, $\partial^2\varphi/\partial x_i\partial t$, $\partial^3\varphi/\partial x_i\partial x_j\partial x_k$, $\partial^3\varphi/\partial x_i\partial x_j\partial t$ $(i,j,k = 1,2,\cdots,n)$,

$$\begin{aligned}
&\frac{F_{t,t+\Delta}\varphi(t+\Delta,\cdot)(x) - \varphi(t,x)}{\Delta} \\
&\to \frac{\partial\varphi}{\partial t} + \mathcal{H}(x, D_x\varphi(t,x), D_x^2\varphi(t,x))\ (\Delta \to 0+).
\end{aligned} \quad (10)$$

The convergence is uniform on each compact set of $[0,T]\times \mathbb{R}^n$.

By using arguments similar to those regarding stability of viscosity solutions and combining them with uniqueness results for viscosity solutions, we can relate the discrete-time stochastic control value with the viscosity solution of (6).

Theorem 2.3: Under (A1)–(A3), $V^N(t,x)$ converges to a viscosity solution of (6) as $N\to\infty$ uniformly on each compact set of $[0,T]\times\mathbb{R}^n$. The limit of $V^N(t,x)$ is the unique viscosity solution, $V(t,x)$, among the class of solutions satisfying (7).

The following proof is in the class of viscosity solution proofs. Readers mainly interested in the developments in the vein of min-plus analysis, might reasonably choose to skip this argument on a first reading.

Proof: Let $W(t,x)$ be a limit of $V^N(t,x)$ in Corollary 2.2. We use the full sequence of $V^N(t,x)$ for simplicity of notation. We will only prove $W(t,x)$ is a viscosity subsolution of (6). The supersolution part can be proved in a similar way. Let $(\hat t,\hat x)\in(0,T)\times\mathbb{R}^n$ be a maximum point of $W(t,x)-\varphi(t,x)$ on $B_\delta(\hat t,\hat x)$ with $W(\hat t,\hat x)=\varphi(\hat t,\hat x)$. We may suppose that $(\hat t,\hat x)$ is a strict local maximum point. Note that from Proposition 2.1 with (A2) and (A3), there exists $K>0$ independent of N such that

$$\begin{aligned}|V^N(t,x)|&\le K(1+|t-\hat t|^2+|x-\hat x|^2),\\ |W(t,x)|&\le K(1+|t-\hat t|^2+|x-\hat x|^2).\end{aligned}$$

For a given $\bar K>K$, by modifying $\varphi(t,x)$ outside of a neighborhood of $(\hat t,\hat x)$, we can have a smooth $\psi(t,x)$ satisfying the following conditions:

$$\begin{gathered}W(t,x)<\psi(t,x)\text{ if }(t,x)\ne(\hat t,\hat x), W(\hat t,\hat x)=\psi(\hat t,\hat x),\\ \psi(t,x)=\varphi(t,x)\text{ on }B_{\delta/2}(\hat t,\hat x),\\ \psi(t,x)=\bar K(1+|t-\hat t|^2+|x-\hat x|^2)\text{ on }B_\delta(\hat t,\hat x)^c.\end{gathered}$$

Take a maximum point $(\hat t_N,\hat x_N)$ of $V^N(t,x)-\psi(t,x)$ on $\overline{B}_\delta(\hat t,\hat x)$. Since $V^N(t,x)$ converges to $W(t,x)$ uniformly on $\overline{B}_\delta(\hat t,\hat x)$ and $(\hat t,\hat x)$ is a global maximum point of $W(t,x)-\psi(t,x)$,

$$(\hat t_N,\hat x_N)\to(\hat t,\hat x)\ (N\to\infty).$$

Since $V^N(t,x)\le K(1+|t-\hat t|^2+|x-\hat x|^2)$ and $\bar K>K$, we have for $(t,x)\notin B_\delta(\hat t,\hat x)$,

$$\begin{aligned}V^N(t,x)-\psi(t,x)&\le K(1+|t-\hat t|^2+|x-\hat x|^2)\\ &\quad-\bar K(1+|t-\hat t|^2+|x-\hat x|^2)\\ &\le-(\bar K-K)(1+\delta^2)<0.\end{aligned}$$

Then, since $\max_{\overline{B}_\delta(\hat t,\hat x)}(V^N-\psi)\to\max_{\overline{B}_\delta(\hat t,\hat x)}(W-\psi)=0$, $(\hat t_N,\hat x_N)$ is a global maximum point of $V^N(t,x)-\psi(t,x)$ for large N.

Modify the notation in π_N to $\pi_N=\{t_0^{(N)}=0<t_1^{(N)}<\cdots<t_N^{(N)}=T\}$ so as to indicate the dependence on N. Since $\hat t_N\to\hat t\in(0,T)$, there exists i such that

$$t_i^{(N)}\le\hat t_N<t_{i+1}^{(N)}.$$

From (9),

$$V^N(\hat t_N,\hat x_N)=F_{\hat t_N,t_{i+1}^{(N)}}V^N(t_{i+1}^{(N)},\cdot)(\hat x_N).$$

Using the property that $F_{t,s}(\phi+c)=F_{t,s}\phi+c$ for any scalar c, we have

$$0=F_{\hat t_N,t_{i+1}^{(N)}}(V^N(t_{i+1}^{(N)},\cdot)-V^N(\hat t_N,\hat x_N))(\hat x_N).$$

Since $F_{t,s}$ is monotone and $(\hat t_N,\hat x_N)$ is a global maximum point of $V^N(t,x)-\psi(t,x)$, we can see that

$$\begin{aligned}0&\le F_{\hat t_N,t_{i+1}^{(N)}}(\psi(t_{i+1}^{(N)},\cdot)-\psi(\hat t_N,\hat x_N))(\hat x_N)\\ &=F_{\hat t_N,t_{i+1}^{(N)}}\psi(t_{i+1}^{(N)},\cdot)(\hat x_N)-\psi(\hat t_N,\hat x_N).\end{aligned}$$

Thus we have

$$0\le\frac{1}{t_{i+1}^{(N)}-\hat t_N}\Big\{F_{\hat t_N,t_{i+1}^{(N)}}\psi(t_{i+1}^{(N)},\cdot)(\hat x_N)-\psi(\hat t_N,\hat x_N)\Big\}.$$

Note that $\partial^2\psi/\partial t^2$, $\partial^2\psi/\partial t\partial x_i$, $\partial^3\psi/\partial x_i\partial x_j\partial x_k$ and $\partial^3\psi/\partial t\partial x_i\partial x_j$ are bounded. Therefore, if we take the limit as $N\to\infty$, we have from (10)

$$0\le\frac{\partial\psi}{\partial t}(\hat t,\hat x)+\mathcal{H}(\hat x,D_x\psi(\hat t,\hat x),D_x^2\psi(\hat t,\hat x)).$$

Since $\varphi(t,x)=\psi(t,x)$ on $B_{\delta/2}(\hat t,\hat x)$,

$$0\le\frac{\partial\varphi}{\partial t}(\hat t,\hat x)+\mathcal{H}(\hat x,D_x\varphi(\hat t,\hat x),D_x^2\varphi(\hat t,\hat x)).$$

Hence $W(t,x)$ is a viscosity subsolution of (6).

Lastly, we note that $W(t,x)$ is the unique viscosity solution of (6) satisfying (7) by the comparison theorem of [5, Theorem 2.1]. Therefore, $V^N(t,x)$ converges to the unique viscosity solution of (6). ∎

III. Min-Plus Distributive Property

We will use an infinite version of the min-plus distributive property to move a certain infimum from inside an expectation operator to outside. It will be familiar to control and game theorists who often work with notions of non-anticipative mappings and strategies.

Recall that the min-plus algebra is the commutative semi-field on $\mathbb{R}^+\doteq\mathbb{R}\cup\{+\infty\}$ given by

$$a\oplus b\doteq\min\{a,b\},\qquad a\otimes b\doteq a+b,$$

c.f., [4], [8], [9]. The distributive property is, of course,

$$\begin{aligned}(a_{1,1}\oplus a_{1,2})\otimes(a_{2,1}\oplus a_{2,2})=\;&a_{1,1}\otimes a_{2,1}\oplus a_{1,1}\otimes a_{2,2}\\ &\oplus a_{1,2}\otimes a_{2,1}\oplus a_{1,2}\otimes a_{2,2}.\end{aligned}$$

By induction, one finds that for finite index sets $\mathcal{I}=]1,I[=\{1,2,\cdots,I\}$ and $\mathcal{J}=]1,J[=\{1,2,\cdots,J\}$,

$$\bigotimes_{i\in\mathcal{I}}\left[\bigoplus_{j\in\mathcal{J}}a_{i,j}\right]=\bigoplus_{\{j_i\}_{i\in\mathcal{I}}\in\mathcal{J}^I}\left[\bigotimes_{i\in\mathcal{I}}a_{i,j_i}\right],$$

where $\mathcal{J}^I=\prod_{i\in\mathcal{I}}J$, the set of ordered sequences of length I of elements of $\mathcal{J}$. Alternatively, we may write this as

$$\sum_{i\in\mathcal{I}}\left[\min_{j\in\mathcal{J}}a_{i,j}\right]=\min_{\{j_i\}_{i\in\mathcal{I}}\in\mathcal{J}^I}\left[\sum_{i\in\mathcal{I}}a_{i,j_i}\right].$$

In this latter form, one naturally thinks of the sequences $\{j_i\}_{i\in\mathcal{I}}$ as mappings from $\mathcal{I}$ to $\mathcal{J}$, i.e., as mappings or strategies.

When we move to the infinite version of the distributive property, some technicalities arise. One version of such appeared in [10]. However, the assumptions in that result are too restrictive for the class of problems we are considering. Instead, we generalize that result to:

Theorem 3.1: Let (Z, d_Z) be a separable metric space and (W, d_W) be a separable Banach space with Borel sets $\mathcal{B}^W$. Let p be a finite measure on $(W, \mathcal{B}^W)$, and let $\overline{D} \doteq p(W)$. Let $h : W \times Z \to I\!\!R$ be Borel measurable. Suppose there exists $\bar{z} \in Z$ such that

$$\int_W h(w, \bar{z})\, dP(w) < \infty \tag{11}$$

and suppose for given $\varepsilon > 0$, there exists $R < \infty$ such that

$$\int_{(\overline{B}_R(0))^c} \inf_{z\in Z} h(w, z)\, dP(w) \geq -\varepsilon. \tag{12}$$

Also, suppose that given $\varepsilon > 0$ and $R < \infty$, there exists $\delta > 0$ such that $|h(w, z) - h(\bar{w}, z)| < \varepsilon$ for all $z \in Z$ and all $w, \bar{w} \in \overline{B}_R(0)$ such that $d_W(w, \bar{w}) < \delta$. Lastly, we suppose that either Z is countable or $h(w, z)$ is continuous on z for each $w \in W$ (where of course, the former supposition can be embedded within the latter, but that is less illuminating). Then,

$$\int_W \inf_{z\in Z} h(w, z)\, dP(w) = \inf_{\tilde{z}\in\widetilde{Z}} \int_W h(w, \tilde{z}(w))\, dP(w),$$

where $\widetilde{Z} \doteq \{\tilde{z} : W \to Z \mid \text{Borel measurable }\}$.

Proof: For the measurability of $\inf_{z\in Z} h(w, z)$, note that for $\alpha \in I\!\!R$,

$$\{w \in W; \inf_{z\in Z} h(w, z) \geq \alpha\} = \bigcap_{z\in Z} \{w \in W; h(w, z) \geq \alpha\}.$$

If Z is countable, the measurability is immediate. For general Z, we shall show that for some countable $Z' \subset Z$,

$$\bigcap_{z\in Z} \{w \in W; h(w, z) \geq \alpha\} = \bigcap_{z\in Z'} \{w \in W; h(w, z) \geq \alpha\}.$$

Take a countable dense set Z' of Z. Let $w \in W$ satisfy $h(w, z) \geq \alpha$ for any $z \in Z'$. Suppose that $h(w, \hat{z}) < \alpha$ for some $\hat{z} \in Z$. Since $h(w, z)$ is continuous on z and Z' is dense, there exists $\bar{z} \in Z'$ such that $h(w, \bar{z}) < \alpha$, which is a contradiction. Therefore we have

$$\bigcap_{z\in Z'} \{w \in W; h(w, z) \geq \alpha\} \subseteq \bigcap_{z\in Z} \{w \in W; h(w, z) \geq \alpha\}.$$

The opposite inclusion is obvious.

Now, for any $\tilde{z}_0 \in \widetilde{Z}$, $\int_W h(w, \tilde{z}_0(w))\, dP(w) \geq \int_W \inf_{z\in Z}[h(w, z)]\, dP(w)$, and so

$$\inf_{\tilde{z}\in\widetilde{Z}} \left\{ \int_W h(w, \tilde{z}_0(w))\, dP(w) \right\} \geq \int_W \inf_{z\in Z} [h(w, z)]\, dP(w). \tag{13}$$

We now proceed to prove the reverse.

Let $\varepsilon > 0$. By (11) and the Dominated Convergence Theorem, there exists $R_1 < \infty$ such that

$$\int_{[\overline{B}_{R_1}(0)]^c} h(w, \bar{z})\, dP(w) < \varepsilon. \tag{14}$$

Further, by (12), there exists $R_2 < \infty$ such that

$$\int_{[\overline{B}_{R_2}(0)]^c} \inf_{z\in Z} [h(w, z)]\, dP(w) \geq -\varepsilon. \tag{15}$$

Let $R = \max\{R_1, R_2\}$. By assumption, there exists $\delta = \delta(R, \varepsilon) > 0$ such that

$$|h(w, z) - h(\bar{w}, z)| < \varepsilon \tag{16}$$

for all $z \in Z$ and all $w, \bar{w} \in \overline{B}_R(0)$ such that $d_W(\bar{w}, w) < \delta$.

By the separability of W, there exists $\{w_i\}_{i\in I\!\!N} \subseteq \overline{B}_R(0)$ such that $\bigcup_{i\in I\!\!N} B_\delta(w_i) \supseteq \overline{B}_R(0)$. For each $i \in I\!\!N$, let $z_i \in Z$ be such that

$$h(w_i, z_i) \leq \inf_{z\in Z} h(w_i, z) + \varepsilon. \tag{17}$$

We next follow a standard continuity-type argument. Let $w \in B_\delta(w_i)$, and suppose

$$h(w, z_i) > \inf_{z\in Z} h(w, z) + 4\varepsilon. \tag{18}$$

Then,

$$h(w_i, z_i) \geq h(w, z_i) - |h(w_i, z_i) - h(w, z_i)|,$$

which by (16),

$$> h(w, z_i) - \varepsilon,$$

which by (18),

$$> \inf_{z\in Z} h(w, z) + 3\varepsilon. \tag{19}$$

Let $z_w^\varepsilon \in Z$ be such that

$$h(w, z_w^\varepsilon) \leq \inf_{z\in Z} h(w, z) + \varepsilon. \tag{20}$$

Combining (19) and (20), one has

$$h(w_i, z_i) > h(w, z_w^\varepsilon) + 2\varepsilon,$$

which by (16) again,

$$> h(w_i, z_w^\varepsilon) + \varepsilon \geq \inf_{z\in Z} h(w_i, z) + \varepsilon,$$

which contradicts (17). Therefore,

$$h(w, z_i) \leq \inf_{z\in Z} h(w, z) + 4\varepsilon, \tag{21}$$

for all $w \in B_\delta(w_i)$ and all $i \in I\!\!N$.

Now let $D_1 = B_\delta(w_1) \cap \overline{B}_R(0)$ and, for all $k > 1$, $D_k = (B_\delta(w_k) \cap \overline{B}_R(0)) \setminus \bigcup_{i<k} D_i$. Note that $\{D_k\}_{k\in I\!\!N}$ is disjoint and $\overline{B}_R(0) = \bigcup_{i\in I\!\!N} D_i$. Define $\tilde{z}^\varepsilon : W \to Z$ given by

$$\tilde{z}^\varepsilon(w) = \begin{cases} z_k & \text{if } w \in D_k, \\ \bar{z} & \text{if } w \in [\bigcup_{i\in I\!\!N} D_i]^c = \overline{B}_R(0)^c. \end{cases} \tag{22}$$

Then, $\tilde{z}^\varepsilon$ is well-defined and measurable. Further,

$$\int_W h(w, \tilde{z}^\varepsilon(w)\, dP(w)$$
$$= \int_{\bigcup_{i\in I\!\!N} D_i} h(w, \tilde{z}^\varepsilon(w))\, dP(w) + \int_{[\bigcup_{i\in I\!\!N} D_i]^c} h(w, \tilde{z}^\varepsilon(w))\, dP(w),$$

which by (21) and (22),

$$\leq \int_{\bigcup_{i\in \mathbf{N}} D_i} \inf_{z\in Z}[h(w,z)+4\varepsilon]\,dP(w)+\int_{[\bigcup_{i\in \mathbf{N}} D_i]^c} h(w,\bar{z})\,dP(w),$$

which by (14) and the assumption that $P(W)=\overline{D}<\infty$,

$$\leq (4\overline{D}+1)\varepsilon + \int_{\bigcup_{i\in \mathbf{N}} D_i} \inf_{z\in Z}[h(w,z)]\,dP(w)$$

$$= (4\overline{D}+1)\varepsilon + \int_W \inf_{z\in Z}[h(w,z)]\,dP(w) - \int_{[\bigcup_{i\in \mathbf{N}} D_i]^c} \inf_{z\in Z}[h(w,z)]\,dP(w),$$

which by (15),

$$\leq (4\overline{D}+2)\varepsilon + \int_W \inf_{z\in Z}[h(w,z)]\,dP(w).$$

Since this is true for all $\varepsilon>0$,

$$\inf_{\tilde{z}\in\widetilde{Z}}\left\{\int_W h(w,\tilde{z}(w)\,dP(w)\right\} \leq \int_W \inf_{z\in Z}[h(w,z)]\,dP(w).$$

■

IV. Distributed Dynamic Programming

We will use the above infinite-version of the min-plus distributive property in conjunction with the dynamic programming principle of Section II. This will yield what we refer to as an idempotent distributed dynamic programming principle (IDDPP), which is the basis of the numerical approach we take.

Recall our discrete-time value function, $V^N(t_k,x)$ given by (9) for $t_k\in\pi_N$ and $x\in I\!R^n$. Suppose that at time, t_{k+1}, one has representation

$$V^N(t_{k+1},x)=\inf_{z\in Z_{k+1}} g^N_{k+1}(x,z), \tag{23}$$

where $(Z_{k+1},d_{Z_{k+1}})$ is a separable metric space. Letting $g^N_N(x,z)=g_T(x,z)$ and $Z_N=Z'_T$, we see that $V^N(t_N,x)=V^N(T,x)=\Psi(x)$ has this form. Then the dynamic program of (8), (9) with $\Delta=T/N$ becomes

$$V^N(t_k,x)=\inf_{u\in U}\min_{m\in\mathcal{M}}\Big\{l(x,u,m)\Delta+\mathbf{E}\Big[\inf_{z\in Z_{k+1}} g^N_{k+1}\big(x + f(x,u,m)\Delta+\sigma(x,u,m)w,z\big)\Big]\Big\}$$
$$=\inf_{u\in U}\min_{m\in\mathcal{M}}\int_W \inf_{z\in Z_{k+1}}\Big[l(x,u,m)\Delta+g^N_{k+1}\big(x + f^\Delta(x,u,m,w),z\big)\Big]\,dP_\Delta(w), \tag{24}$$

where $f^\Delta(x,u,m,w)=f(x,u,m)\Delta+\sigma(x,u,m)w$, P_Δ is the measure corresponding to a normal random variable over $I\!R^l$ with mean zero and covariance ΔI, and $W=I\!R^l$.

We will use the min-plus distributive property of Theorem 3.1 to move the infimum over Z_{k+1} outside the integral. Letting

$$\widetilde{Z}_{k+1}\doteq\{\tilde{z}_{k+1}: W\to Z_{k+1}\,|\,\text{Borel measurable}\},$$

we will have

$$V^N(t_k,x)=\inf_{z\in Z_k} g^N_k(x,z), \tag{25}$$

where $Z_k=U\times\mathcal{M}\times\widetilde{Z}_{k+1}$ and for $x\in I\!R^n$ and $z\in Z_k$,

$$g^N_k(x,z)=\int_W l(x,u,m)\Delta + g^N_{k+1}\big(x+f^\Delta(x,u,m,w),\tilde{z}_{k+1}(w)\big)\,dP_\Delta(w). \tag{26}$$

Consequently, the general form of (23) will be inherited from $V^N(t_{k+1},\cdot)$ to $V^N(t_k,\cdot)$, and one can propagate backward in this manner indefinitely. This is what we referred to above as the IDDPP.

In order to make this program rigorous, we have to verify two results. The first is to find a sufficient condition on $g^N_{k+1}(x,z)$ under which we can apply Theorem 3.1 at (24).

Proposition 4.1: In addition to (A1) and (A2), we suppose that Z_{k+1} and $g^N_{k+1}(x,z)$ satisfy the following:

(i)$_{k+1}$ $(Z_{k+1},d_{Z_{k+1}})$ is a bounded and closed subset of a separable Banach space $\mathcal{X}_{k+1}$ where metric $d_{Z_{k+1}}$ is induced by norm $\|\cdot\|_{\mathcal{X}_{k+1}}$ of $\mathcal{X}_{k+1}$.

(ii)$_{k+1}$ There exists $C>0$ such that for any $x,x'\in I\!R^n$, $z\in Z_{k+1}$,

$$|g^N_{k+1}(x,z)|\leq C(1+|x|^2),$$
$$|g^N_{k+1}(x,z)-g^N_{k+1}(x',z)|\leq C(1+|x|+|x'|)|x-x'|.$$

(iii)$_{k+1}$ There exists $C>0$ such that for any $z,z'\in Z_{k+1}$, $x\in I\!R^n$,

$$|g^N_{k+1}(x,z)-g^N_{k+1}(x,z')|\leq C(1+|x|^2)d_{Z_{k+1}}(z,z').$$

Then (25) holds.

Secondly, in order to repeatedly apply Theorem 3.1, we need to show that properties (i)$_{k+1}$–(iii)$_{k+1}$ on g^N_{k+1} are inherited by the g^N_k given by (26).

Proposition 4.2: Suppose that $(Z_{k+1},d_{Z_{k+1}})$ and $g^N_{k+1}: I\!R^n\times Z_{k+1}\to I\!R$ satisfy (i)$_{k+1}$–(iii)$_{k+1}$. Let $\mathcal{X}_k=I\!R^p\times I\!R\times L^1(W,\mathcal{B}^W,(1+|w|^2)P_\Delta(dw);\mathcal{X}_{k+1})$ be a product of the Banach spaces with the norm

$$\|z\|_{\mathcal{X}_k}=|u|+|m|+\int_W \|\tilde{z}_{k+1}(w)\|_{\mathcal{X}_{k+1}}(1+|w|^2)P_\Delta(dw)$$

where $z=(u,m,\tilde{z}_{k+1})\in\mathcal{X}_k$. Under (A1) and (A2), $Z_k=U\times\mathcal{M}\times\widetilde{Z}_{k+1}\subset\mathcal{X}_k$ and $g^N_k: I\!R^n\times Z_k\to I\!R$ given by (26) satisfy (i)$_k$–(iii)$_k$.

Finally, we obtain the IDDPP.

Theorem 4.3: In addition to (A1) and (A2), suppose that $(Z'_T,d_{Z'_T})$ and $g_T: I\!R^n\times Z'_T\to I\!R$ satisfy the following:

(i-T) $(Z'_T,d_{Z'_T})$ is a bounded and closed subset of a separable Banach space $\mathcal{X}'_T$ where metric $d_{Z'_T}$ is induced by the norm of $\mathcal{X}'_T$.

(ii-T) There exists $C>0$ such that for any $x,x'\in I\!R^n$, $z\in Z'_T$,

$$|g_T(x,z)|\leq C(1+|x|),$$
$$|g_T(x,z)-g_T(x',z)|\leq C(1+|x|+|x'|)|x-x'|.$$

(iii-T) There exists $C>0$ such that for any $z,z'\in Z'_T$, $x\in I\!R^n$,

$$|g_T(x,z)-g_T(x,z')|\leq C(1+|x|^2)d_{Z'_T}(z,z').$$

Letting $Z_N=Z'_T$ and $g^N_N(x,z)=g_T(x,z)$, (25) with (26) holds for $k=N-1,N-2,\cdots,0$.

V. QUADRATIC FORMS

We will give an example where a quadratic structure is retained in course of the IDDPP. Consider the following particular case:

$$f(x,u,m) = A^m x + b^m(u), \;\; \sigma(x,u,m) = \sigma^m(u),$$
$$l(x,u,m) = \frac{1}{2}(x-\overline{x}^m)^T \overline{Q}^m (x-\overline{x}^m) + \frac{1}{2}\overline{c}^m(u),$$

where, for each $m \in \mathcal{M}$, A^m is an $n\times n$ matrix, $\overline{Q}^m$ is an $n\times n$ positive-definite symmetric matrix, $\overline{x}^m \in I\!R^n$ and $b^m(\cdot)$, $\sigma^m(\cdot)$, $\overline{c}^m(\cdot)$ are $I\!R^n$-valued, $n\times l$ matrix-valued, $I\!R$-valued Lipschitz continuous functions on a compact set U of $I\!R^p$, respectively.

Let $(Z'_T, d_{Z'_T})$ be a bounded and closed subset of a separable Banach space. We suppose that $g_T : I\!R^n \times Z'_T \to I\!R$ is a quadratic form on x:

$$g_T(x,z) = \frac{1}{2}(x - x_T(z))^T Q_T(z)(x - x_T(z)) + \frac{1}{2}c_T(z),$$

where $Q_T(\cdot)$, $x_T(\cdot)$, and $c_T(\cdot)$ are $n\times n$ nonnegative-definite symmetric matrix-valued, $I\!R^n$-valued, and $I\!R$-valued bounded Lipschitz continuous functions on Z'_T, respectively, i.e., there exists $L>0$ such that $z,z' \in Z'_T$,

$$\begin{aligned}\|Q_T(z)\| + |x_T(z)| + |c_T(z)| &\le L,\\ \|Q_T(z) - Q_T(z')\| + |x_T(z) - x_T(z')| + |c_T(z) - c_T(z')| &\\ \le L d_{Z'_T}(z,z').&\end{aligned}$$

Under these assumptions, we can verify the conditions of Theorem 4.3, and we have $V^N(t_k,x)$ described by quadratic $g_k^N(x,z)$:

$$V^N(t_k,x) = \inf_{z\in Z_k} g_k^N(x,z),$$
$$g_k^N(x,z) = \frac{1}{2}(x - x_k(z))^T Q_k(z)(x - x_k(z)) + \frac{1}{2}c_k(z),$$

where for $k=N$, we let $Z_N = Z'_T$, $Q_N = Q_T$, $x_N = x_T$, $c_N = c_T$. For $k = N-1,\cdots,0$, Z_k, Q_k, x_k, c_k are recursively determined backward in k:

$$Z_k = U \times \mathcal{M} \times \widetilde{Z}_{k+1} \tag{27}$$

and for $z = (u,m,\tilde{z}_{k+1}) \in Z_k$,

$$\begin{aligned}Q_k(z) &= \overline{Q}^m + (I + A^m\Delta)^T \int_W Q_{k+1}(\tilde{z}_{k+1}(w))dP_\Delta(w)\\ &\quad\times(I + A^m\Delta),\\ x_k(z) &= -Q_k(z)^{-1}\Big\{ -\overline{Q}^m\overline{x}^m + (I+A^m\Delta)^T\\ &\quad\times\int_W Q_{k+1}(\tilde{z}_{k+1}(w))\big(b^m(u)\Delta + \sigma^m(u)w\\ &\quad -x_{k+1}(\tilde{z}_{k+1}(w))\big)dP_\Delta(w)\Big\},\\ c_k(z) &= x_{k+1}(z)^T Q_k(z) x_{k+1}(z) + (\overline{x}^m)^T\overline{Q}^m(z)\overline{x}^m + \overline{c}^m(u)\\ &\quad+\int_W\big(b^m(u)\Delta + \sigma^m(u)w - x_{k+1}(\tilde{z}_{k+1}(w))\big)^T Q_{k+1}(\tilde{z}_{k+1}(w))\\ &\quad\times\big(b^m(u)\Delta + \sigma^m(u)w - x_{k+1}(\tilde{z}_{k+1}(w))\big)dP_\Delta(w)\\ &\quad+\int_W c_{k+1}(\tilde{z}_{k+1}(w))dP_\Delta(w).\end{aligned}$$

Here we note that $Q_k(z)$ is positive and therefore $Q_k(z)^{-1}$ exists because $\overline{Q}^m$ is positive and $Q_{k+1}(\tilde{z}_{k+1}(w))$ is non-negative. More generally, in practice, we use quadratics in the form $g_k^N(x,z) = \frac{1}{2}x^T Q_k(z)x + b_k^T(z)x + c_k(z)$, which avoids the inverse, but for reasons of space we do not include the details.

The key to this class of methods lies in the repeated projection of the solution down onto a low-dimensional (min-plus) subspace. Importantly, the subspace is chosen at each step so as to minimize the error induced by this projection. One sees in (27) that after one step of the IDDPP, the set Z_k will have the cardinality of the continuum even in the case where Z_{k+1} is finite. Consequently, the projection down to a finite-dimensional subspace is a critical step. We will use an approach analogous to that in [12], and very briefly discuss this below.

Each quadratic form is defined by the triple of its coefficients, and we let $\tau = \tau_k^z \doteq \big(Q_k(z), b_k(z), c_k(z)\big)$. Let $\widetilde{T}$ denote the set of all possible such triples. One first defines a relaxed partial order on $\widetilde{T}$ by $\alpha \preceq \tau$ if

$$\int_{I\!R^n} \mathcal{G}[\alpha](x)\,d\pi(x) \le \int_{I\!R^n} \mathcal{G}[\tau](x)\,d\pi(x) \quad \forall\,\pi\in\Pi,$$

where $\mathcal{G}[\tau](\cdot)$ denotes the quadratic function induced by coefficients τ and Π is a set of probability measures on $I\!R^n$. With this ordering, we reduce the optimal projection problem to minimization of a decreasing, convex functional over a cornice structure [12]. In particular, the appropriate cornice structure is the upward cone (according to the partial order) of the convex hull of $\widehat{T}_k \doteq \{\tau_k(z)\,|\,z\in Z_k\} \subset \widetilde{T}$. One can show that the optimal projection onto an $\widehat{N}$-dimensional min-plus subspace reduces to selection of a finite subset of $\widehat{T}_k$. One further shows that this selection problem can be reduced to optimization of a submodular functional on a domain of sets. One may then employ, for example, a greedy algorithm for (suboptimal) selection relative to the submodular criterion, and we note that there exists extensive literature regarding error bounds for such algorithms in relation to submodular criteria.

Lastly, we note that there are two sources of error. The first is that induced by the time-discretization (see Section II). The second is due to the projection operation. Both need to be estimated in order to develop a clear sense of the approach. Based on the results in the discrete-time stochastic [10], [12], [15] and continuous-time deterministic [9], [13] cases, we have positive expectations here.

REFERENCES

[1] M. Akian, S. Gaubert and A. Lakhoua, *The max-plus finite element method for solving deterministic optimal control problems: basic properties and convergence analysis*, SIAM J. Control and Optim., **47** (2008), 817–848.

[2] M. Akian, S. Gaubert and A. Lakhoua, *A max-plus finite element method for solving finite horizon deterministic optimal control problems*, Proc. 16^{th} International Symposium on Mathematical Theory of Networks and Systems (2004).

[3] G. Barles and P.E. Souganidis, *Convergence of approximation schemes for fully nonlinear second order equations*, Asymptotic Anal., 4 (1991), 271–283.

[4] F.L. Baccelli, G. Cohen, G.J. Olsder and J.-P. Quadrat, *Synchronization and Linearity,* John Wiley, New York, 1992.
[5] F. Da Lio and O. Ley, *Uniqueness results for second-order Bellman-Isaacs equations under quadratic growth assumptions and applications*, SIAM J. Control Optim., 40 (2006), 74–106.
[6] W.H. Fleming and W.M. McEneaney, *A max-plus based algorithm for an HJB equation of nonlinear filtering,* SIAM J. Control and Optim., 38 (2000), 683–710.
[7] W.H. Fleming and H.M. Soner, *Controlled Markov Processes and Viscosity Solutions*, 2nd ed., Springer-Verlag, New York, 2006.
[8] V.N. Kolokoltsov and V.P. Maslov, *Idempotent Analysis and Its Applications,* Kluwer, 1997.
[9] W.M. McEneaney, *Max-Plus Methods for Nonlinear Control and Estimation*, Birkhauser, Boston, 2006.
[10] W.M. McEneaney, *Idempotent Algorithms for Discrete-Time Stochastic Control through Distributed Dynamic Programming,* Proc. 48th IEEE CDC (2009).
[11] W.M. McEneaney, *Idempotent Method for Dynamic Games and Complexity Reduction in Min-Max Expansions*, Proc. 48th IEEE CDC (2009).
[12] W.M. McEneaney, *Complexity Reduction, Cornices and Pruning,* Proc. of the International Conference on Tropical and Idempotent Mathematics, G.L. Litvinov and S.N. Sergeev (Eds.), Contemporary Math. 495, Amer. Math. Soc. (2009), 293–303.
[13] W.M. McEneaney, A. Deshpande, S. Gaubert, *Curse-of-Complexity Attenuation in the Curse-of-Dimensionality-Free Method for HJB PDEs,* Proc. ACC 2008, Seattle (2008).
[14] W.M. McEneaney, “A Curse-of-Dimensionality-Free Numerical Method for Solution of Certain HJB PDEs”, SIAM J. on Control and Optim., 46 (2007) 1239-1276.
[15] W.M. McEneaney, A. Oran and A. Cavender, *Value-Based Control of the Observation-Decision Process,* Proc. ACC, Seattle, (2008).