Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Curse-of-Dimensionality Free Method for Bellman PDEs with Hamiltonian Written as Maximum of Quadratic Forms

William M. McEneaney

Abstract— **Max-plus methods have been explored for solution of first-order, nonlinear Hamilton-Jacobi-Bellman partial differential equations (HJB PDEs) and corresponding nonlinear control problems. These methods exploit the max-plus linearity of the associated semigroups. Although these methods provide advantages, they still suffer from the curse-of-dimensionality. Here we consider HJB PDEs where the Hamiltonian takes the form of a (pointwise) maximum of quadratic forms. We obtain a numerical method not subject to the curse-of-dimensionality. The method is based on construction of the dual-space semigroup corresponding to the HJB PDE. This dual-space semigroup is constructed from the dual-space semigroups corresponding to the constituent quadratic Hamiltonians. The actual computations in the algorithm involve repeatedly computing coefficients of quadratics which are obtained as the maxima of two other quadratics.**

I. Introduction

One approach to nonlinear control is through Dynamic Programming (DP). With DP, solution of the control problem "reduces" to solution of the corresponding partial differential equation (PDE). In the case of Deterministic Optimal Control or Deterministic Games (such as H_∞ control) where one player's feedback is prespecified, the PDE is a Hamilton-Jacobi-Bellman (HJB) PDE. The difficulty is that one must solve the HJB PDE.

Various approaches have been taken to solution of the HJB PDE. The most common methods by far all fall into the class of grid-based methods. (cf. [3], [4], [8] among many others). These require that one generate a grid over some bounded region of the state-space. Suppose the region over which one constructs the grid is rectangular. Suppose one uses 100 grid points per dimension. If the state dimension is n, then one has 100^n grid points. Thus the computations grow exponentially in state-space dimension n.

In recent years, an entirely new class of numerical methods for HJB PDEs has emerged (c.f. [7], [14], [13], [1]). These methods exploit the max-plus linearity of the associated semigroup. They employ a max-plus basis function expansion of the solution, and the numerical methods obtain the coefficients in the basis expansion. Much of the work has concentrated on the (harder) steady-state HJB PDE class. With the max-plus methods, the number of basis functions required still typically grows exponentially with space dimension. For instance, one might use 25 basis functions per space dimension to cover a rectangular region well. Consequently, one still has the curse-of-dimensionality. Even with that max-plus approach, one cannot expect to solve problems of more than dimension 4 or 5 on current machines.

This paper discusses an approach to certain nonlinear HJB PDEs which is not subject to the curse-of-dimensionality. In fact, the computational growth in state-space dimension is on the order of n^3. However, there is exponential computational growth in a certain measure of complexity of the Hamiltonian. Under this measure, the minimal complexity Hamiltonian is the quadratic Hamiltonian – corresponding to solution by a Riccati equation. If the Hamiltonian is given as a point-wise maximum of M quadratic Hamiltonians, then one could say the complexity of the Hamiltonian is M. It should be remarked here that any semiconvex Hamiltonian can be represented as a supremum of quadratics, and therefore approximated arbitrarily well by a maximum of a finite number of quadratics.

The approach has been applied on some simple nonlinear problems. A few simple examples comprised of 3 linear/quadratic components were solved in 10-20 seconds over $\mathbf{R}^3$ and 10-45 seconds over $\mathbf{R}^4$. For these particular problems, the solution was obtained over the entire space with the resulting errors in the *gradients* growing linearly in $|x|$. (See Section VI for specific examples.) These speeds are of course unprecedented. This code was not optimized. Further, the computational growth in going from $n = 4$ up to say $n = 6$ would be on the order of $6^3/4^3 \simeq 4$ as opposed to say more than 10^4 for a finite element method.

We will consider HJB PDEs given as

$$\widetilde{H}(x, \nabla V) = \max_{m \in \{1,2,\ldots,M\}} \{H^m(x, \nabla V)\} \qquad (1)$$

with boundary data $V(0) = 0$ (V being zero at the origin). In order to make the problem tractable, we will concentrate on a single class of HJB PDEs of form (1). However, the theory can obviously be expanded to a much larger class.

II. Review of Theory

We first outline the theory underlying the algorithm. As indicated above, we suppose the individual H^m are quadratic Hamiltonians. Consequently, consider a finite set of linear systems

$$\dot{\xi}^m = A^m \xi^m + \sigma^m w, \quad \xi^m_0 = x \in I\!R^n. \qquad (2)$$

Let $w \in \mathcal{W} \doteq L_2^{loc}([0,\infty); I\!R^m)$. Let the cost functionals and value functions be

$$J^m(x, T; w) \doteq \int_0^T \tfrac{1}{2} \xi^m_t D^m \xi^m_t - \frac{\gamma^2}{2} |w_t|^2 \, dt, \qquad (3)$$

$$V^m(x) = \lim_{T \to \infty} \sup_{w \in \mathcal{W}} J^m(x, T; w). \qquad (4)$$

Research supported by NSF grant DMS-0307229.

Dept. of Mech. and Aero. Eng. and Dept. of Math., University of California San Diego, La Jolla, CA 92093-0112, USA, wmceneaney@ucsd.edu

0-7803-9568-9/05/$20.00 ©2005 IEEE

Obviously J^m and V^m require some assumptions in order to guarantee their existence.

Assume that there exists $c_A \in (0,\infty)$ such that $x^T A^m x \le -c_A|x|^2$ for all $x \in I\!R^n$ and $m \in \mathcal{M}$. Assume that there exists $c_\sigma < \infty$ such that $|\sigma^m| \le c_\sigma \ \forall m \in \mathcal{M}$. Assume that all D^m are positive definite, symmetric, and let c_D be such that $x^T D^m x \le c_D|x|^2$ for all $x \in I\!R^n$ and $m \in \mathcal{M}$. Lastly, assume that $\gamma^2/c_\sigma^2 > c_D/c_A^2$. (A1)

These assumptions guarantee the existence of the V^m as locally bounded functions which are zero at the origin (cf. [16]).

The corresponding HJB PDEs are

$$\begin{aligned} 0 &= H^m(x,\nabla V) \\ &= \tfrac{1}{2}x^T D^m x + (A^m x)^T \nabla V + \tfrac{1}{2}\nabla V^T \Sigma^m \nabla V \\ V(0) &= 0 \end{aligned} \quad (5)$$

where $\Sigma^m \doteq \frac{1}{\gamma^2}\sigma^m(\sigma^m)^T$. Let $\mathcal{G}_\delta$ be the subset of $C(I\!R^n)$ such that $0 \le V(x) \le \frac{c_A(\gamma-\delta)^2}{c_\sigma^2}|x|^2$. For $m \in \mathcal{M}$, let P^m satisfy the algebraic Riccati equations

$$0 = (A^m)^T P^m + P^m A^m + D^m + P^m \Sigma^m P^m. \quad (6)$$

Standard results imply that each value function (4) is the unique classical solution of its corresponding HJB PDE (5) in the class $\mathcal{G}_\delta$ for sufficiently small $\delta > 0$. Further, $V^m(x) = \frac{1}{2}x^T P^m x$ where P^m is the smallest symmetric, positive definite solution of (6). In particular, there exists symmetric, positive definite $\overline{C}$ such that $V^m(x) - \frac{1}{2}x^T\overline{C}x$ is convex for all $m \in \mathcal{M}$.

The method we will use to obtain these value functions/HJB PDE solutions will be through the associated semigroups. For each m define the semigroup

$$S_T^m[\phi] \doteq \sup_{w\in\mathcal{W}} \int_0^T \tfrac{1}{2}(\xi_t^m)^T D^m \xi_t^m - \frac{\gamma^2}{2}|w_t|^2\,dt + \phi(\xi_T^m)$$

where ξ^m satisfies (2). By [16], the domain of S_T^m includes $\mathcal{G}_\delta$ for all $\delta > 0$.

Theorem 2.1: Fix any $T > 0$. Each value function, V^m, is the unique smooth solution of $V = S_T^m[V]$ in the class $\mathcal{G}_\delta$ for sufficiently small $\delta > 0$. Further, given any $V \in \mathcal{G}_\delta$, $\lim_{T\to\infty} S_T^m[V](x) = V^m(x)$ (uniformly on compact sets).

Recall that the HJB PDE of interest is (1) with H^m given by (5). The corresponding value function is

$$\widetilde{V}(x) = \sup_{w\in\mathcal{W}} \sup_{\mu\in\mathcal{D}_\infty} \sup_{T<\infty} \int_0^T l^{\mu_t}(\xi_t) - \frac{\gamma^2}{2}|w_t|^2\,dt \quad (7)$$

where

$$\begin{aligned} l^{\mu_t}(x) &= \tfrac{1}{2}x^T D^{\mu_t} x, \\ \mathcal{D}_\infty &= \{\mu : [0,\infty) \to \mathcal{M} \ : \ \text{measurable}\}, \end{aligned}$$

and ξ satisfies

$$\dot{\xi} = A^{\mu_t}\xi + \sigma^{\mu_t} w_t, \quad \xi_0 = x. \quad (8)$$

Define the semigroup

$$\widetilde{S}_T[\phi] = \sup_{w\in\mathcal{W}} \sup_{\mu\in\mathcal{D}_T} \int_0^T l^{\mu_t}(\xi_t) - \frac{\gamma^2}{2}|w_t|^2\,dt + \phi(\xi_T)$$

where $\mathcal{D}_T = \{\mu : [0,T) \to \mathcal{M} \ : \ \text{measurable}\}$.

Theorem 2.2: Fix any $T > 0$. Value function $\widetilde{V}$ is the unique continuous solution of $V = \widetilde{S}_T[V]$ in the class $\mathcal{G}_\delta$ for sufficiently small $\delta > 0$. Further, given any $V \in \mathcal{G}_\delta$, $\lim_{T\to\infty} \widetilde{S}_T[V](x) = \widetilde{V}(x)$ (uniformly on compact sets). Lastly, there exists $c_V > 0$ such that $\widetilde{V}(x) - \frac{1}{2}c_V|x|^2$ is convex.

III. Max-Plus Dual Operators

We use $\oplus, \otimes$ to indicate max-plus addition and multiplication; max-plus integration (supremization) is indicated by an $\oplus$ superscript on the integral sign. Let $I\!R^- = I\!R \cup \{-\infty\}$. Recall that a function, $\phi : I\!R^n \to I\!R^-$ is semiconvex if given any $R \in (0,\infty)$ there exists $\beta_R \in I\!R$ such that $\phi(x) + \frac{\beta_R}{2}|x|^2$ is convex over $\overline{B}_R(0) = \{x \in I\!R^n : |x| \le R\}$. We say ϕ is uniformly semiconvex with constant β if $\phi(x) + \frac{\beta}{2}|x|^2$ is convex over $I\!R^n$. Let $\mathcal{S}_\beta = \mathcal{S}_\beta(I\!R^n)$ be the set of functions mapping $I\!R^n$ into $I\!R^-$ which are uniformly semiconvex with constant β. Note that $\mathcal{S}_\beta$ is a max-plus vector space (also known as a moduloid) [2], [7], [9], [14]. Combining this notation with the above results, we see that we have

Theorem 3.1: There exists $\overline{\beta} \in I\!R$ such that given any $\beta > \overline{\beta}$, $\widetilde{V} \in \mathcal{S}_\beta$ and $V^m \in \mathcal{S}_\beta$ for all $m \in \mathcal{M}$. Further, one may take $\beta < 0$ (i.e. $\widetilde{V}, V^m$ convex).

The following semiconvex duality result [7], [14] requires only a small modification of convex duality and Legendre/Fenchel transform results (c.f. [17]).

Theorem 3.2: Let $\phi \in \mathcal{S}_\beta$. Let C be a symmetric matrix such that $C + \beta I > 0$ (i.e. $C + \beta I$ positive definite) with either $C > 0$ or $C < 0$. Define $\psi : I\!R^n \times I\!R^n \to I\!R$ by $\psi(x,z) = -\frac{1}{2}(x-z)^T C(x-z)$. Then, for all $x \in I\!R^n$,

$$\begin{aligned} \phi(x) &= \max_{z\in I\!R^n} [\psi(x,z) + a(z)] \\ &\doteq \int_{I\!R^n}^{\oplus} \psi(x,z) \otimes a(z)\,dz \doteq \psi(x,\cdot) \odot a(\cdot) \end{aligned} \quad (9)$$

where for all $z \in I\!R^n$

$$\begin{aligned} a(z) &= -\int_{I\!R^n}^{\oplus} \psi(x,z) \otimes [-\phi(x)]\,dx \\ &= -\{\psi(\cdot,z) \odot [-\phi(\cdot)]\} \doteq \{\psi(\cdot,z) \odot [\phi^-(\cdot)]\}^- . \end{aligned} \quad (10)$$

We will refer to a as the *semiconvex dual* of ϕ.

Semiconcavity is the obvious analogue of semiconvexity. Let $\mathcal{S}_\beta^-$ be the set of functions mapping $I\!R^n$ into $I\!R \cup \{+\infty\}$ which are uniformly semiconcave with constant β ($\phi(x) - (\beta/2)|x|^2$ concave over all of $I\!R^n$).

Lemma 3.3: Let $\phi \in \mathcal{S}_\beta$, and let a be the semiconvex dual of ϕ. Then $a \in \mathcal{S}_\beta^-$. Further, suppose $b \in \mathcal{S}_\beta^-$ is such that $\phi = \psi(x,\cdot) \odot b(\cdot)$. Then $b = a$.

For simplicity, we will henceforth specialize to the case where $\psi(x,z) \doteq (c/2)|x-z|^2$. It will be critical to the method that $\widetilde{S}_\tau[\psi(\cdot,z)] \in \mathcal{S}_{-(c+\varepsilon)}$ for some $\varepsilon > 0$. This is the subject of the next theorem.

Theorem 3.4: We may choose $c > 0$ such that $\widetilde{V}, V^m \in \mathcal{S}_{-c}$, and such that there exists $\overline{\tau} > 0$ and $\eta > 0$ such that ,

$$\widetilde{S}_\tau[\psi(\cdot,z)], S_\tau^m[\psi(\cdot,z)] \in \mathcal{S}_{-(c+\eta\tau)} \qquad \forall\, \tau \in [0,\overline{\tau}].$$

Henceforth, we suppose c, τ, η chosen so that the results of Theorem 3.4 hold. Now for each $z \in I\!R^n$, $\widetilde{S}_\tau[\psi(\cdot, z)] \in \mathcal{S}_{-(c+\eta\tau)}$. Therefore, by Theorem 3.2

$$\widetilde{S}_\tau[\psi(\cdot, z)](x) = \psi(x, \cdot) \odot \widetilde{\mathcal{B}}_\tau(\cdot, z) \tag{11}$$

where for all $y \in I\!R^n$

$$\widetilde{\mathcal{B}}_\tau(y, z) = \left\{\psi(\cdot, y) \odot [\widetilde{S}_\tau[\psi(\cdot, z)](\cdot)]^-\right\}^- \tag{12}$$

It is handy to define the max-plus linear operator with "kernel" $\widetilde{\mathcal{B}}_\tau$ as $\widehat{\widetilde{\mathcal{B}}}_\tau[a](z) \doteq \widetilde{\mathcal{B}}_\tau(z, \cdot) \odot a(\cdot)$ for all $a \in \mathcal{S}_{-c}$. Note that (11), (12) introduce the dual-space operator kernel $\widetilde{\mathcal{B}}_\tau$ which propagates the dual equivalently to propagation in the original space by $\widetilde{S}_\tau$.

Proposition 3.5: Let $\phi \in \mathcal{S}_{-c}$ with semiconvex dual denoted by a. Define $\phi^1 = \widetilde{S}_\tau[\phi]$. Then $\phi^1 \in \mathcal{S}_{-(c+\eta\tau)}$, and $\phi^1(x) = \psi(x, \cdot) \odot a^1(\cdot)$ where $a^1(x) = \widetilde{\mathcal{B}}_\tau(x, \cdot) \odot a(\cdot)$.

Theorem 3.6: Let $V \in \mathcal{S}_{-c}$, and let a be its semiconvex dual (with respect to ψ). Then $V = \widetilde{S}_\tau[V]$ if and only if

$$\begin{aligned} a(z) &= \int_{I\!R^n}^{\oplus} \widetilde{\mathcal{B}}_\tau(z, y) \otimes a(y)\, dy \\ &= \widetilde{\mathcal{B}}_\tau(z, \cdot) \odot a(\cdot) = \widehat{\widetilde{\mathcal{B}}}_\tau[a](z) \qquad \forall\, z \in I\!R^n. \end{aligned}$$

Corollary 3.7: The value function $\widetilde{V}$ is given by $\widetilde{V}(x) = \psi(x, \cdot) \odot \widetilde{a}(\cdot)$ where $\widetilde{a}$ is the unique solution of $\widetilde{a}(y) = \widetilde{\mathcal{B}}_\tau(y, \cdot) \odot \widetilde{a}(\cdot)\ \forall y \in I\!R^n$ or equivalently, $\widetilde{a} = \widehat{\widetilde{\mathcal{B}}}_\tau[\widetilde{a}]$.

Similarly, for each $m \in \mathcal{M}$ and $z \in I\!R^n$, $S^m_\tau[\psi(\cdot, z)] \in \mathcal{S}_{-(c+\eta\tau)}$ and

$$S^m_\tau[\psi(\cdot, z)](x) = \psi(x, \cdot) \odot \mathcal{B}^m_\tau(\cdot, z) \qquad \forall\, x \in I\!R^n$$

where

$$\mathcal{B}^m_\tau(y, z) = \left\{\psi(\cdot, y) \odot \left[S^m_\tau[\psi(\cdot, z)]\right]^-(\cdot)\right\}^-.$$

As before, it will be handy to define the max-plus linear operator with "kernel" $\mathcal{B}^m_\tau$ as $\widehat{\mathcal{B}}^m_\tau[a](z) \doteq \mathcal{B}^m_\tau(z, \cdot) \odot a(\cdot)$ for all $a \in \mathcal{S}_{-c}$. Further, one also obtains analogous results (by similar proofs). In particular, one has the following

Theorem 3.8: Let $V \in \mathcal{S}_{-c}$, and let a be its semiconvex dual (with respect to ψ). Then $V = S^m_\tau[V]$ if and only if $a(z) = \mathcal{B}^m_\tau(z, \cdot) \odot a(\cdot)\ \forall z \in I\!R^n$.

Corollary 3.9: Each value function V^m is given by $V^m(x) = \psi(x, \cdot) \odot a^m(\cdot)$ where each a^m is the unique solution of $a^m(y) = \mathcal{B}^m_\tau(y, \cdot) \odot a^m(\cdot)\ \forall y \in I\!R^n$.

IV. Discrete Time Approximation

The method developed here will not involve any discretization over space nor any spatially distributed basis functions. Of course this is obvious since otherwise one could not avoid the curse-of-dimensionality. The discretization will be over time, where approximate μ processes will be constant over the length of each time-step. We define the operator $\bar{S}_\tau$ on $\mathcal{G}_\delta$ by

$$\begin{aligned} \bar{S}_\tau[\phi](x) &= \sup_{w \in \mathcal{W}} \max_{m \in \mathcal{M}} \Big[\int_0^\tau l^m(\xi^m_t) - \frac{\gamma^2}{2}|w_t|^2\, dt \\ &\qquad + \phi(\xi^m_\tau)\Big](x) \\ &= \max_{m \in \mathcal{M}} S^m_\tau[\phi](x) \end{aligned}$$

where ξ^m satisfies (2). Let

$$\overline{\mathcal{B}}_\tau(y, z) \doteq \max_{m \in \mathcal{M}} \mathcal{B}^m_\tau(y, z) = \bigoplus_{m \in \mathcal{M}} \mathcal{B}^m_\tau(y, z).$$

The corresponding max-plus linear operator is

$$\widehat{\overline{\mathcal{B}}}_\tau = \bigoplus_{m \in \mathcal{M}} \widehat{\mathcal{B}}^m_\tau.$$

Lemma 4.1: For all $z \in I\!R^n$, $\bar{S}_\tau[\psi(\cdot, z)] \in \mathcal{S}_{-(c+\eta\tau)}$. Further, $\bar{S}_\tau[\psi(\cdot, z)](x) = \psi(x, \cdot) \odot \overline{\mathcal{B}}_\tau(\cdot, z)$.

One has $S^m_\tau \le \bar{S}_\tau \le \widetilde{S}_\tau$ for all $m \in \mathcal{M}$. With τ acting as a time-discretization step-size, let

$$\begin{aligned} \mathcal{D}^\tau_\infty = \Big\{&\mu : [0, \infty) \to \mathcal{M} \mid \text{for each } n \in \mathbf{N} \cup \{0\}, \\ &\text{there exists } m_n \in \mathcal{M} \text{ such that} \\ &\mu(t) = m_n\ \forall\ t \in [n\tau, (n+1)\tau)\Big\}, \end{aligned}$$

and for $T = \bar{n}\tau$ with $\bar{n} \in \mathbf{N}$ define $\mathcal{D}^\tau_T$ similarly but with domain $[0, T)$ rather than $[0, \infty)$. Let $\mathcal{M}^{\bar{n}}$ denote the outer product of $\mathcal{M}$, $\bar{n}$ times. Let $T = \bar{n}\tau$, and define

$$\bar{\bar{S}}^\tau_T[\phi](x) = \max_{\{m_k\}_{k=0}^{\bar{n}-1} \in \mathcal{M}^{\bar{n}}} \left\{\prod_{k=0}^{\bar{n}-1} S^{m_k}_\tau\right\}[\phi](x)$$

where the $\prod$ indicates operator composition.

We will be approximating $\widetilde{V}$ by solving $V = \bar{S}_\tau[V]$ via its dual problem $a = \widehat{\overline{\mathcal{B}}}_\tau[a]$ for small τ. Consequently, we will need to show that there exists a solution to $V = \bar{S}_\tau[V]$, that the solution is unique, and that it can be found by solving the dual problem. We begin with existence.

Theorem 4.2: Let

$$\overline{V}(x) \doteq \lim_{N \to \infty} \bar{\bar{S}}^\tau_{N\tau}[0](x) \tag{13}$$

for all $x \in I\!R^n$ where 0 here represents the zero-function. Then, $\overline{V}$ satisfies

$$V = \bar{S}_\tau[V], \quad V(0) = 0. \tag{14}$$

Further, $0 \le V^m \le \overline{V} \le \widetilde{V}$ for all $m \in \mathcal{M}$, and consequently, $\overline{V} \in \mathcal{G}_\delta$.

Similar techniques to those used for V^m and $\widetilde{V}$ will prove uniqueness for (14) within $\mathcal{G}_\delta$.

Theorem 4.3: $\overline{V}$ is the unique solution of (14) within the class $\mathcal{G}_\delta$ for sufficiently small $\delta > 0$. Further, given any $V \in \mathcal{G}_\delta$, $\lim_{N \to \infty} \bar{\bar{S}}^\tau_{N\tau}[V](x) = \overline{V}(x)$ for all $x \in I\!R^n$ (uniformly on compact sets).

Henceforth, we let $\delta > 0$ be sufficiently small such that $V^m, \widetilde{V}, \overline{V} \in \mathcal{G}_\delta$ for all $m \in \mathcal{M}$.

Theorem 4.4: Let $V \in \mathcal{S}_{-c}$, and let a be its semiconvex dual. Then $V = \bar{S}_\tau[V]$ if and only if $a(y) = \overline{\mathcal{B}}_\tau(y, \cdot) \odot a(\odot)$ $\forall y \in I\!R^n$.

Corollary 4.5: Value function $\overline{V}$ given by (13) is in $\mathcal{S}_{-c}$, and has representation $\overline{V}(x) = \psi(x, \cdot) \odot \overline{a}(\cdot)$ where $\overline{a}$ is the unique solution of $\overline{a} = \widehat{\overline{\mathcal{B}}}_\tau[\overline{a}]$.

The following result on propagation of the semiconvex dual will also come in handy.

Proposition 4.6: Let $\phi \in \mathcal{S}_{-c}$ with semiconvex dual denoted by a. Define $\phi^1 = \bar{S}_\tau[\phi]$. Then $\phi^1 \in \mathcal{S}_{-(c+\eta\tau)}$,

and $\phi^1(x) = \psi(x,\cdot) \odot a^1(\cdot)$ where $a^1(y) = \overline{\mathcal{B}}_\tau(y,\cdot) \odot a(\cdot)$ $\forall y \in I\!R^n$.

The next result indicates that one may approximate $\widetilde{V}$, the solution of $V = \widetilde{S}_\tau[V]$, to as accurate a level as one desires by solving $V = \bar{S}_\tau[V]$ for sufficiently small τ. Recall that if $V = \bar{S}_\tau[V]$, then it satisfies $V = \bar{\bar{S}}^\tau_{N\tau}[V]$ for all $N > 0$ (while V satisfies $V = \widetilde{S}_{N\tau}[V]$), and so this is essentially equivalent to introducing a discrete-time $\overline{\mu} \in \mathcal{D}^\tau_{N\tau}$ approximation to the μ process in $\widetilde{S}_{N\tau}$.

Theorem 4.7: Given $\bar{\varepsilon} > 0$ and $R < \infty$, there exists $\tau > 0$ such that

$$\widetilde{V}(x) - \bar{\varepsilon} \le \overline{V}(x) \le \widetilde{V}(x) \qquad \forall\, x \in \overline{B}_R(0).$$

V. Algorithm

Due to space limitations, we only outline the algorithm.

From Theorem 4.2, $\overline{V} = \lim_{N\to\infty} \bar{\bar{S}}^\tau_{N\tau}[0]$. Let $\overline{V}^0 \equiv 0$. Then $\overline{V}^0 \in \mathcal{S}_{-c}$ of course. Given $\overline{V}^k$, let

$$\overline{V}^{k+1} \doteq \bar{S}_\tau[\overline{V}^k]$$

so that $\overline{V}^k = \bar{\bar{S}}^\tau_{k\tau}[0]$ for all $k \ge 1$.

Let $\overline{a}^k$ be the semiconvex dual of $\overline{V}^k$ for all k. Since $\overline{V}^0 \equiv 0$, one easily finds that $\overline{a}^0(y) = 0$ for all $y \in I\!R^n$. Note also that by Proposition 4.6,

$$\overline{a}^{k+1} = \overline{\mathcal{B}}_\tau(x,\cdot) \odot \overline{a}^k(\cdot) = \widehat{\overline{\mathcal{B}}}_\tau[\overline{a}^k]$$

for all $n \ge 0$.

Recall that

$$\overline{\mathcal{B}}_\tau(x,\cdot) \odot \overline{a}^k(\cdot) = \bigoplus_{m\in\mathcal{M}} \left[\mathcal{B}^m_\tau(x,\cdot) \odot \overline{a}^k(\cdot)\right]. \tag{15}$$

By (15),

$$\begin{aligned} &\overline{a}^1(x) = \bigoplus_{m\in\mathcal{M}} \widehat{a}^1_m(x) \\ &\text{where} \\ &\widehat{a}^1_m(x) \doteq \mathcal{B}^m_\tau(x,\cdot) \odot \overline{a}^0(\cdot) \qquad \forall\, m. \end{aligned} \tag{16}$$

Using (15) and (16),

$$\begin{aligned} &\overline{a}^2(x) = \bigoplus_{\{m_1,m_2\}\in\mathcal{M}^2} \widehat{a}^2_{\{m_1,m_2\}}(x) \\ &\text{where} \\ &\widehat{a}^2_{\{m_1,m_2\}}(x) \doteq \mathcal{B}^{m_2}_\tau(x,\cdot) \odot \widehat{a}^1_{m_1}(\cdot) \qquad \forall\, m_1, m_2 \end{aligned} \tag{17}$$

and $\mathcal{M}^2$ represents the outer product $\mathcal{M} \times \mathcal{M}$. Proceeding with this, one finds that in general,

$$\begin{aligned} &\overline{a}^k(x) = \bigoplus_{\{m_i\}_{i=1}^k \in \mathcal{M}^k} \widehat{a}^k_{\{m_i\}_{i=1}^k}(x) \\ &\text{where} \\ &\widehat{a}^k_{\{m_i\}_{i=1}^k}(x) \doteq \mathcal{B}^{m_k}_\tau(x,\cdot) \odot \widehat{a}^{k-1}_{\{m_i\}_{i=1}^{k-1}}(\cdot) \end{aligned} \tag{18}$$

Of course one can obtain $\overline{V}^n$ from its dual as

$$\begin{aligned} \overline{V}^k(x) &= \max_{y\in I\!R^n} [\psi(x,y) + \overline{a}^k(y)] \\ &\doteq \max_{\{m_i\}_{i=1}^k \in \mathcal{M}^k} \widehat{V}^k_{\{m_i\}_{i=1}^k}(x) \end{aligned} \tag{19}$$

where

$$\widehat{V}^k_{\{m_i\}_{i=1}^k} = \int^{\oplus}_{I\!R^n} \psi(x,y) \otimes \widehat{a}^k_{\{m_i\}_{i=1}^k}(y)\, dy. \tag{20}$$

The algorithm will consist of the forward propagation of the $\widehat{a}^k_{\{m_i\}_{i=1}^k}$ (according to (18)) from $k = 0$ to some termination step $k = N$, followed by construction of the value as $\widehat{V}^k_{\{m_i\}_{i=1}^k}$ (according to (20)).

It is important to note that the computation of each $\widehat{a}^k_{\{m_i\}_{i=1}^k}$ is analytical. We will indicate the actual analytical computations.

By the linear/quadratic nature of the m-indexed systems, we find that the $S^m_\tau[\psi(\cdot,z)]$ take the form

$$S^m_\tau[\psi(\cdot,z)](x) = \tfrac{1}{2}(x - \Lambda^m_\tau z)^T P^m_\tau (x - \Lambda^m_\tau) + \tfrac{1}{2} z^T R^m_\tau z$$

where the time-dependent $n \times n$ matrices P^m_t, Λ^m_t and R^m_t are obtained from solution of Riccati equations, and we note that each of the $P^m_\tau, \Lambda^m_\tau, R^m_\tau$ need only be computed once.

Next one computes each quadratic function $\mathcal{B}^m_\tau(x,z)$ (one time only) as follows. One has

$$\mathcal{B}^m_\tau = -\max_{y\in I\!R^n} \{\psi(y,x) - S^m_\tau[\psi(\cdot,z)](y)\}$$

This takes the form

$$\begin{aligned} \mathcal{B}^m_\tau(x,z) = \tfrac{1}{2}\Big[&x^T M^m_{1,1} x + x^T M^m_{1,2} z + z^T (M^m_{1,2})^T x \\ &+ z^T M^m_{2,2} z\Big] \end{aligned}$$

where each of the $M^m_{i,j}$ have analytical forms (involving matrix inverses) which, due to space limitations, we do not include. Also note that all the matrices in the definition of $\mathcal{B}^m_\tau$ may be precomputed.

Now let us write the (quadratic) $\widehat{a}^k_{\{m_i\}_{i=1}^k}$ in the form

$$\begin{aligned} \widehat{a}^k_{\{m_i\}_{i=1}^k}(x) = \ &\tfrac{1}{2}(x - \widehat{z}^k_{\{m_i\}_{i=1}^k})^T \widehat{Q}^k_{\{m_i\}_{i=1}^k} (x - \widehat{z}^k_{\{m_i\}_{i=1}^k}) \\ &+ \widehat{r}^k_{\{m_i\}_{i=1}^k}. \end{aligned}$$

Then, for each m_{k+1}, we may obtain the coefficients from

$$\begin{aligned} \widehat{Q}^{k+1}_{\{m_i\}_{i=1}^{k+1}} &= M^{m_{k+1}}_{1,1} - M^{m_{k+1}}_{1,2} \widehat{D} \left(M^{m_{k+1}}_{1,2}\right)^T \\ \widehat{z}^{k+1}_{\{m_i\}_{i=1}^{k+1}} &= -\left(\widehat{Q}^{k+1}_{\{m_i\}_{i=1}^{k+1}}\right)^{-1} M^{m_{k+1}}_{1,2} \widehat{E} \\ \widehat{r}^{k+1}_{\{m_i\}_{i=1}^{k+1}} &= \widehat{r}^k_{\{m_i\}_{i=1}^k} + \tfrac{1}{2}\widehat{E}^T M^m_{2,2} \widehat{z}^k_{\{m_i\}_{i=1}^k} \\ &\quad -\tfrac{1}{2}\left(\widehat{z}^{k+1}_{\{m_i\}_{i=1}^{k+1}}\right)^T \widehat{Q}^{k+1}_{\{m_i\}_{i=1}^{k+1}} \widehat{z}^{k+1}_{\{m_i\}_{i=1}^{k+1}} \\ \widehat{D} &= \left(M^{m_{k+1}}_{2,2} + \widehat{Q}^k_{\{m_i\}_{i=1}^k}\right)^{-1} \\ \widehat{E} &= \widehat{D}\widehat{Q}^k_{\{m_i\}_{i=1}^k} \widehat{z}^k_{\{m_i\}_{i=1}^k}. \end{aligned} \tag{21}$$

Thus we have the analytical expression for the propagation of each (quadratic) $\widehat{a}^k_{\{m_i\}_{i=1}^k}$ function. Specifically, we see that the propagation of each $\widehat{a}^k_{\{m_i\}_{i=1}^k}$ amounts to a set of matrix multiplications (and an inverse).

At each step, k, the semiconvex dual of $\overline{V}^k$, $\overline{a}^k$, is represented as the finite set of functions

$$\widehat{\mathcal{A}}_k \doteq \left\{\widehat{a}^k_{\{m_i\}_{i=1}^k} \,|\, m_i \in \mathcal{M}\ \forall i \in \{1,2,\ldots,k\}\right\}.$$

where this is equivalently represented as the set of triples

$$\widehat{\mathcal{Q}}_k \doteq \left\{ \left(\widehat{Q}^k_{\{m_i\}_{i=1}^k}, \widehat{z}^k_{\{m_i\}_{i=1}^k}, \widehat{r}^k_{\{m_i\}_{i=1}^k} \right) \right\}.$$

At any desired stopping time, one can recover a representation of $\overline{V}^k$ as

$$\widehat{\mathcal{V}}_k \doteq \left\{ \widehat{V}^k_{\{m_i\}_{i=1}^k} \,|\, m_i \in \mathcal{M} \ \forall i \in \{1,2,\ldots,k\} \right\}$$

where these $\widehat{V}^k_{\{m_i\}_{i=1}^k}$ are also quadratics. In fact, recall

$$\begin{aligned}\overline{V}^k(x) &= \max_{\{m_i\}_{i=1}^k} \tfrac{1}{2}(x - \widehat{x}^k_{\{m_i\}_{i=1}^k})^T \widehat{P}^k_{\{m_i\}_{i=1}^k}(x - \widehat{x}^k_{\{m_i\}_{i=1}^k}) \\ &\quad + \widehat{\rho}^k_{\{m_i\}_{i=1}^k} \\ &\doteq \bigoplus_{\{m_i\}_{i=1}^k} \widehat{V}^k_{\{m_i\}_{i=1}^k}(x)\end{aligned}$$

where the coefficients $\widehat{P}^k_{\{m_i\}_{i=1}^k}, \widehat{x}^k_{\{m_i\}_{i=1}^k}, \widehat{\rho}^k_{\{m_i\}_{i=1}^k}$ have analytical forms which we do not include due to space limitations. Thus, $\overline{V}^k$ has the representation as the set of triples

$$\mathcal{P}_k \doteq \left\{ \left(\widehat{P}^k_{\{m_i\}_{i=1}^k}, \widehat{x}^k_{\{m_i\}_{i=1}^k}, \widehat{\rho}^k_{\{m_i\}_{i=1}^k} \right) \right\}. \qquad (22)$$

We note that the triples which comprise $\mathcal{P}_k$ are analytically obtained from the triples $(\widehat{Q}^k_{\{m_i\}_{i=1}^k}, \widehat{z}^k_{\{m_i\}_{i=1}^k}, \widehat{r}^k_{\{m_i\}_{i=1}^k})$ by matrix multiplications and an inverse. The transference from $(\widehat{Q}^k_{\{m_i\}_{i=1}^k}, \widehat{z}^k_{\{m_i\}_{i=1}^k}, \widehat{r}^k_{\{m_i\}_{i=1}^k})$ to $(\widehat{P}^k_{\{m_i\}_{i=1}^k}, \widehat{x}^k_{\{m_i\}_{i=1}^k}, \widehat{\rho}^k_{\{m_i\}_{i=1}^k})$ need only be done once which is at the termination of the algorithm propagation. We note that (22) is the representation of our approximate solution of the original control problem/HJB PDE.

The errors are due to our approximation of $\widetilde{V}$ by $\overline{V}$, and to the approximation of $\overline{V}$ by the prelimit $\overline{V}^N$ for stopping time $k = N$. Neither of these errors are related to the space dimension. The errors in $|\widetilde{V} - \overline{V}|$ are dependent on the step size τ. The errors in $|\overline{V}^N - \overline{V}| = |\bar{\bar{S}}^\tau_{N\tau}[0] - \overline{V}|$ are due to premature termination in the limit $\overline{V} = \lim_{N\to\infty} \bar{\bar{S}}^\tau_{N\tau}[0]$. The computation of each triple $(\widehat{P}^k_{\{m_i\}_{i=1}^k}, \widehat{x}^k_{\{m_i\}_{i=1}^k}, \widehat{\rho}^k_{\{m_i\}_{i=1}^k})$ grows like the cube of the space dimension (due to the matrix operations). Thus one avoids the curse-of-dimensionality. Of course if one then chooses to compute $\overline{V}^N(x)$ for all x on some grid over say a rectangular region in $\mathbb{R}^n$, then by definition one has exponential growth in this computation as the space dimension increases. The point is that one does not need to compute $\overline{V}^N \simeq \widetilde{V}$ at each such point.

However, the curse-of-dimensionality is replaced by another type of rapid computational cost growth. Here, we refer to this as the curse-of-complexity. If $\#\mathcal{M} = 1$, then all the computations of our algorithm (excepting the solution of the Riccati equation) are unnecessary, and we *informally* refer to this as complexity one. When there are $M = \#\mathcal{M}$ such quadratics in the Hamiltonian, $\widetilde{H}$, we say it has complexity M. Note that

$$\# \left\{ \widehat{V}^k_{\{m_i\}_{i=1}^k} \,|\, m_i \in \mathcal{M} \ \forall i \in \{1,2,\ldots,k\} \right\} \sim M^N.$$

For large N, this is indeed a large number. (Pruning and proper initialization can reduce this problem, but we do not include these techniques.) Nevertheless, for small values of M, we obtain a very rapid solution of such nonlinear HJB PDEs, as will be indicated in the examples to follow. Further, the computational cost growth in space dimension n is limited to cubic growth. We emphasize that the existence of an algorithm avoiding the curse-of-dimensionality is significant regardless of the practical issues.

VI. Examples

A number of examples have so far been tested. In these tests, the computational speeds were very great. (Again, some practical issues involving pruning and initialization are not discussed here due to space limitations.) This is due to the fact that $M = \#\mathcal{M}$ was small. The algorithm as described above was coded in MATLAB (with a very simple pruning technique and initialization). The quoted computational times were obtained with a standard 2001 PC. The times correspond to the time to compute $\overline{V}_N \doteq \bar{\bar{S}}^\tau_{N\tau}[0]$. The plots below require one to compute the value function and/or gradients pointwise on planes in the state space. These plotting computations are not included in the timing.

Consider a four-dimensional example with constituent Hamiltonians, H^m, whose A^m are

$$A^1 = \begin{bmatrix} -1.0 & 0.5 & 0.0 & 0.1 \\ 0.1 & -1.0 & 0.2 & 0.0 \\ 0.2 & 0.0 & -1.5 & 0.1 \\ 0.0 & -0.1 & 0.0 & -1.5 \end{bmatrix},$$

$$A^2 = (A^1)^T,$$

$$A^3 = \begin{bmatrix} -1.0 & 0.5 & 0.0 & 0.1 \\ 0.1 & -1.0 & 0.2 & 0.0 \\ 0.2 & 0.0 & -1.6 & -0.1 \\ 0.0 & -0.05 & 0.1 & -1.5 \end{bmatrix}.$$

The D^m and Σ^m were simply

$$D^1 = D^2 = D^3 = \begin{bmatrix} 1.5 & 0.2 & 0.1 & 0.0 \\ 0.2 & 1.5 & 0.0 & 0.1 \\ 0.1 & 0.0 & 1.5 & 0.0 \\ 0.0 & 0.1 & 0.0 & 1.5 \end{bmatrix},$$

and

$$\Sigma^1 = \Sigma^2 = \Sigma^3 = \begin{bmatrix} 0.2 & -0.01 & 0.02 & 0.01 \\ -0.01 & 0.2 & 0.0 & 0.0 \\ 0.02 & 0.0 & 0.25 & 0.0 \\ 0.01 & 0.0 & 0.0 & 0.25 \end{bmatrix}.$$

The results of this four-dimensional example appear in Figures 1–4. In this case, the results have been plotted over the region of the affine plane $x_3 = 3$, $x_4 = -0.5$ given by $x_1 \in [-10, 10]$ and $x_2 \in [-10, 10]$. The backsubstitution error has been scaled by dividing by $|x|^2 + 10^{-5}$. The computations required approximately 40 seconds.

In order to indicate that the form of HJB PDE solutions obtained by this approach are not limited to the type of shapes appearing in the previous example, we include an additional partial derivative plot from another example, and this is depicted in Figure 5. There is not space here to give the full details of the example, but we note that it includes

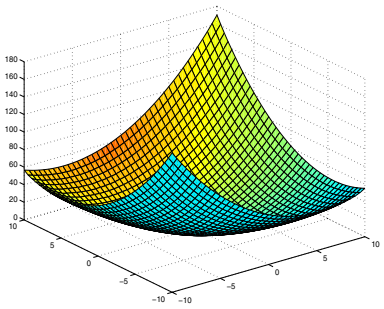

Fig. 1. Value function (4-D case)

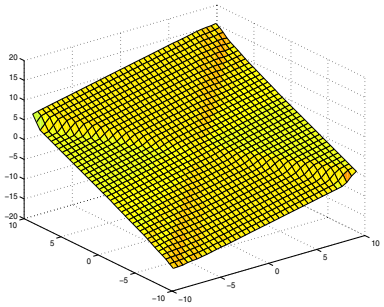

Fig. 2. Partial with respect to x_1 (4-D case)

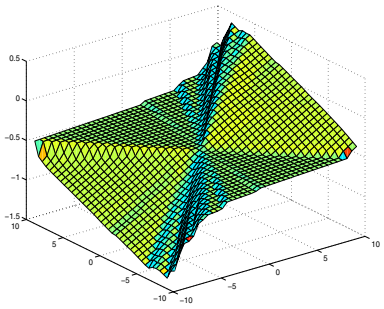

Fig. 3. Partial with respect to x_4 (4-D case)

constant and linear terms in the H^m, yielding a system where the behavior changes when the state exceeds a certain bound.

REFERENCES

[1] M. Akian, S. Gaubert and A. Lakhoua, "A max-plus finite element method for solving finite horizon determinsitic optimal control problems", Proc. MTNS 2004 (to appear).

[2] F.L. Baccelli, G. Cohen, G.J. Olsder and J.-P. Quadrat, *Synchronization and Linearity*, John Wiley (1992).

[3] M. Bardi and I. Capuzzo-Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, Birkhauser, Boston, 1997.

[4] F. Camilli, M. Falcone, P. Lanucara and A. Seghini, "A Domain Decomposition Method for Bellman Equations," Domain Decomposition Methods in Scientific and Engineering Computation, D.E. Keyes and J. Xu (eds.), Contemporary Math. **180**, AMS (1994), pp. 477-483.

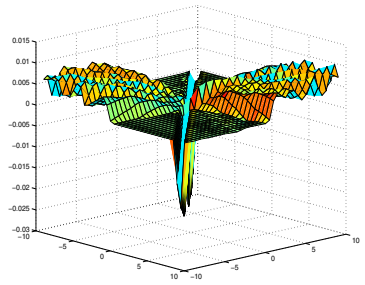

Fig. 4. Scaled backsubstitution error (4-D case)

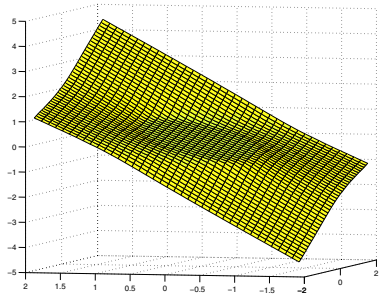

Fig. 5. Partial with respect to x_1, extra example

[5] R.A. Cuninghame-Green, *Minimax Algebra*, Lecture Notes in Economics and Mathematical Systems 166, Springer (1979).

[6] P. Dupuis and M. Boué, "Markov chain approximations for deterministic control problems with affine dynamics and quadratic cost in the control", SIAM J. on Numerical Analysis, **36** (1999), pp. 667–695.

[7] 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.

[8] H.J. Kushner and P. Dupuis, *Numerical methods for Stochastic Control Problems in Continuous Time*, Springer Verlag, New York, 1992.

[9] G.L. Litvinov, V.P. Maslov and G.B. Shpiz, "Idempotent Functional Analysis: An Algebraic Approach", Mathematical Notes, **69**, No. 5 (2001), 696–729.

[10] V.P. Maslov, "On a new principle of superposition for optimization problems", Russian Math. Surveys, **42** (1987) 43–54.

[11] W.M. McEneaney, "Legendre Transforms on Max-Plus Spaces as a Tool for Nonlinear Control Problems", Proc. MTNS 2004.

[12] W.M. McEneaney, "Max-Plus Summation of Fenchel-Transformed Semigroups for Solution of Nonlinear Bellman Equations", Submitted to Systems and Control Letters.

[13] W.M. McEneaney, "Max–Plus Eigenvector Representations for Solution of Nonlinear H_∞ Problems: Error Analysis", SIAM J. Control and Optim. **43**, (2004) 379–412.

[14] W.M. McEneaney, "Max–Plus Eigenvector Representations for Solution of Nonlinear H_∞ Problems: Basic Concepts", IEEE Trans. Auto. Control **48**, (2003) 1150–1163.

[15] W.M. McEneaney and M. Horton, "Max–plus eigenvector representations for nonlinear H_∞ value functions", 37th IEEE Conf. on Decision and Control (1998), 3506–3511.

[16] W.M. McEneaney, "A Uniqueness result for the Isaacs equation corresponding to nonlinear H_∞ control", Math. Controls, Signals and Systems, **11** (1998), 303–334.

[17] R.T. Rockafellar and R.J. Wets, *Variational Analysis*, Springer–Verlag (1997).

[18] P. Soravia, "H_∞ control of nonlinear systems: differential games and viscosity solutions", SIAM J. Control and Optim., **34** (1996), 1071–1097.