Idempotent Algorithms for Continuous-Time Stochastic Control

Hidehiro Kaise William M. McEneaney

***Abstract*— Previously, idempotent methods have been found to be extremely fast for solution of dynamic programming equations associated with deterministic control problems. The original methods exploited the idempotent (e.g., max-plus) linearity of the associated semigroup operator. However, it is now known that the curse-of-dimensionality-free idempotent methods do not require this linearity, and may be used to solve some classes of stochastic control problems. The key is the use of the idempotent distributive property. This was previously demonstrated for a class of discrete-time stochastic control problems. Here, we extend this approach to a class of continuous-time stochastic control problems**

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], [8], [13], as well as the more recently developed curse-of-dimensionality-free methods [13], [19]. It has recently been discovered that idempotent methods are applicable to stochastic control and games. The methods are related to the the above curse-of-dimensionality-free methods for deterministic control. In particular, a min-plus based method was found for stochastic control problems [15], [20], and a min-max method was discovered for games [16].

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 $\mathbb{R}^m$) 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

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

Dept. of Mechanical and Aerospace Engineering, University of California San Diego, San Diego, CA 92093-0411, USA. wmceneaney@ucsd.edu Research partially supported by NSF grant DMS-0307229.

ISBN 978-963-311-370-7

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

$$d\xi_s = f(\xi_s, u_s)\, ds + \sigma(\xi_s, u_s)\, dB_s, \qquad (1)$$
$$\xi_t = x \in I\!R^n \qquad (2)$$

where f is measurable, with more assumptions on it to follow. The u_s will be control inputs taking values in $U \subset I\!R^k$. (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. [13] for motivation.) Also, $\{B_\cdot, \mathcal{F}_\cdot\}$ is an m-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 $n \times m$ 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

$$J(t, x, u_\cdot) \doteq \mathbf{E}\left\{ \int_t^T l(\xi_s, u_s)\, ds + \Psi(\xi_T) \right\} \qquad (3)$$

where

$$\Psi(x) \doteq \inf_{z_T \in Z_T} \{ g_T(x, z_T) \}, \qquad (4)$$

where l and the g_T are measurable, and (Z_T, d_{z_T}) is a separable metric space. The value function is

$$V(t, x) = \inf_{u \in \mathcal{U}_t} J(t, x, u_\cdot), \qquad (5)$$

where $\mathcal{U}_t$ is the set of $\mathcal{F}_t$-progressively measurable controls, taking values in U, 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, $y \in I\!R^n$, and any $u \in U$,

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

It is seen that $V(t,x)$ can be characterized as the viscosity solution of the HJB PDE associated with (1) and (2) (see [9] for such discussion). Indeed, $V(t,x)$ satisfies the dynamic programming principle (DPP)

$$V(t,x) = \inf_{u \in \mathcal{U}_t} \mathbf{E}\left[\int_t^s l(\xi_r, u_r) dr + V(s, \xi_s) \right].$$

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

$$\begin{aligned} &\frac{\partial V}{\partial t} + \mathcal{H}(x, D_x V(t,x), D_x^2 V(t,x)) = 0, \qquad (6) \\ &V(T,x) = \Psi(x), \end{aligned}$$

where

$$\mathcal{H}(x, D_x V(t,x), D_x^2 V(t,x)) = \inf_{u \in U} \left\{ \frac{1}{2} \mathrm{tr}(\sigma\sigma^T(x,u) D_x^2 V(t,x)) + f(x,u) \cdot D_x V(t,x) + l(x,u) \right\}.$$

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

$$F_{t,s}\phi(x) = \inf_{u \in U} \{ l(x,u)(s-t) + \mathbf{E}[\phi(x + f(x,u)(s-t) + \sigma(x,u)(B_s - B_t))] \}. \qquad (7)$$

(See [3] for general problems 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:

$$V_N(t,x) = \begin{cases} \Psi(x), & t = T, \\ F_{t,t_{i+1}} V_N(t_{i+1}, \cdot)(x), & t_i \le t < t_{i+1}, \end{cases}$$

where $F_{t,t_{i+1}} V_N(t_{i+1}, \cdot)(x)$ is $F_{t,t_{i+1}}\phi(x)$ with $\phi(\cdot) = V_N(t_{i+1}, \cdot)$. We note that the infinitesimal generator of $\{F_{t,s}\}$ is $\mathcal{H}$: For any smooth function $\varphi(t,x)$,

$$\frac{F_{t,t+h}\varphi(t+h, \cdot)(x) - \varphi(t,x)}{h} \to \frac{\partial \varphi}{\partial t} + \mathcal{H}(x, D_x\varphi(t,x), D_x^2\varphi(t,x)) \ (h \to 0+).$$

By using arguments similar to the stability of viscosity solutions and combining the uniqueness of viscosity solutions, we can show that $V_N(t,x)$ converges to $V(t,x)$ as $N \to \infty$ uniformly on each compact set of $[0,T] \times I\!R^n$. (See [7] for the uniqueness of viscosity solutions under (A1) and (A2)).

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 semifield on $I\!R^+ \doteq I\!R \cup \{+\infty\}$ given by

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

c.f., [4], [11], [13]. The distributive property is, of course,

$$(a_{1,1} \oplus a_{1,2}) \otimes (a_{2,1} \oplus a_{2,2}) = a_{1,1} \otimes a_{2,1} \oplus a_{1,1} \otimes a_{2,2} \oplus a_{1,2} \otimes a_{2,1} \oplus a_{1,2} \otimes a_{2,2}.$$

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

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

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 [15]. However, the assumptions on 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. Recall that (W, d_w) is 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

$$\int_W h(w,\bar{z})\, dP(w) < \infty \tag{8}$$

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

$$\int_{(\overline{B}_R(0))^c} \inf_{z\in Z} h(w,z)\, dP(w) \geq -\varepsilon. \tag{9}$$

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 Z is countable or $h(w,z)$ is continuous on z for each $w \in W$. 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 $\inf_{z\in Z} h(w,z)$ on w, we need to check $\{w \in W; \inf_{z\in Z} h(w,z) \geq \alpha\}$ is measurable for any α. To see this, we note that

$$\{w \in W; \inf_{z\in Z} h(w,z) \geq \alpha\} = \bigcap_{z\in Z} \{w \in W; h(w,z) \geq \alpha\}. \tag{10}$$

In case of countable Z, (10) is measurable because (10) is a countable intersection of $\{w \in W; h(w,z) \geq \alpha\}$. For general Z, we shall show that for some countable $\tilde{Z} \subset Z$,

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

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

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

The opposite inclusion is obvious.

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

$$\inf_{\tilde{z}\in\widetilde{Z}} \left\{ \int_w h(w,\tilde{z}_0(w))\, dP(w) \right\} \geq \int_w \inf_{z\in Z} [h(w,z)]\, dP(w). \tag{11}$$

We now proceed to prove the reverse.

Let $\varepsilon > 0$. By (8) and the Dominated Convergence Theorem, there exists $\bar{z} \in Z$ and $R_1 < \infty$ such that

$$\int_{[\overline{B}_{R_1}(0)]^c} h(w, \bar{z})\, dP(w) < \varepsilon. \tag{12}$$

Further, by (9), there exists $R_2 < \infty$ such that

$$\int_{[\overline{B}_{R_1}(0)]^c} \inf_{z \in Z} [h(w, z)]\, dP(w) \geq -\varepsilon. \tag{13}$$

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

$$|h(w, z) - h(\bar{w}, z)| < \varepsilon \tag{14}$$

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 $\{w_i\}_{i \in \mathbb{N}} \subseteq \overline{B}_R(0)$ such that $\bigcup_{i \in \mathbb{N}} B_\delta(w_i) \supseteq \overline{B}_R(0)$. For each $i \in \mathbb{N}$, let $z_i \in Z$ be such that

$$h(w_i, z_i) \leq \inf_{z \in Z} h(w_i, z) + \varepsilon. \tag{15}$$

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

$$h(w, z_i) > \inf_{z \in Z} h(w, z) + 4\varepsilon. \tag{16}$$

Then,

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

which by (14),

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

whcih by (12),

$$> \inf_{z \in Z} h(w, z) + 3\varepsilon. \tag{17}$$

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

$$h(w, z_w^\varepsilon) \leq \inf_{z \in Z} h(w, z) + \varepsilon. \tag{18}$$

Combining (17) and (18), one has

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

which by (14) again,

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

which contradicts (15). Therefore,

$$h(w, z_i) \leq \inf_{z \in Z} h(w, z) + 4\varepsilon, \tag{19}$$

for all $w \in B_\delta(w_i)$ and all $i \in \mathbb{N}$.

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

$$\tilde{z}^\varepsilon(w) = \begin{cases} z_k & \text{if } w \in D_k, \\ \bar{z} & \text{if } w \in [\bigcup_{i \in \mathbb{N}} B_\delta(w_i)]^c. \end{cases} \tag{20}$$

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

$$\int_W h(w, \tilde{z}^\varepsilon(w)\, dP(w) = \int_{\bigcup_{i \in \mathbb{N}} B_\delta(w_i)} h(w, \tilde{z}^\varepsilon(w))\, dP(w) + \int_{[\bigcup_{i \in \mathbb{N}} B_\delta(w_i)]^c} h(w, \tilde{z}^\varepsilon(w))\, dP(w),$$

which by (19) and (20),

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

which by (12) and the assumption $P(W) = \overline{D} < \infty$,

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

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

which by (13),

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

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 use the above infinite-version of the distributive property in the context of the dynamic program of Section II. This will yield what we refer to as an idempotent distributed dynamic programming principle (IDDPP). In this development, it is necessary to show that the conditions allowing application of Theorem 3.1 are inherited through the backward IDDPP operator. This will be included, and is similar in overall structure to the analogous result in [15].

V. Quadratic Forms and an Example

In the presentation, we will include the development in the particular case where the g_T and l take quadratic forms. In this case, the computations reduce to a simple form. In order to keep the computations tractable, a projection down to the optimal lower-dimensional min-plus subspace is performed at each step. This is performed by an approach similar to that discussed in [17], with a partial ordering on the quadratic forms which allows such. Bounds on the additional errors introduced by the projection operation will be included. A simple example will also be given.

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 determinsitic 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] R.A. Cuninghame-Green, *Minimax Algebra*, Lecture Notes in Economics and Mathematical Systems 166, Springer, New York, 1979.

[6] G. Cohen, S. Gaubert and J.-P. Quadrat, *Duality and Separation Theorems in Idempotent Semimodules*, Linear Algebra and Applications, 379 (2004), 395–422.

[7] 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.

[8] 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.

[9] W.H. Fleming and H.M. Soner, *Controlled Markov Processes and Viscosity Solutions*, 2nd ed., Springer-Verlag, New York, 2006.

[10] G.B. Folland, *Real Analysis, Modern Techniques and Their Applications*, John Wiley, New York, 1984.

[11] V.N. Kolokoltsov and V.P. Maslov, *Idempotent Analysis and Its Applications,* Kluwer, 1997.

[12] G.L. Litvinov, V.P. Maslov and G.B. Shpiz, *Idempotent Functional Analysis: An Algebraic Approach*, Mathematical Notes, 69 (2001), 696–729.

[13] W.M. McEneaney, *Max-Plus Methods for Nonlinear Control and Estimation*, Birkhauser, Boston, 2006.

[14] W.M. McEneaney, *Idempotent Algorithms for Discrete-Time Stochastic Control*, Automatica, (to appear).

[15] W.M. McEneaney, *Idempotent Algorithms for Discrete-Time Stochastic Control through Distributed Dynamic Programming,* Proc. 48th IEEE CDC (2009).

[16] W.M. McEneaney, *Idempotent Method for Dynamic Games and Complexity Reduction in Min-Max Expansions,* Proc. 48th IEEE CDC (2009).

[17] 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.

[18] 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).

[19] 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.

[20] W.M. McEneaney, A. Oran and A. Cavender, *Value-Based Control of the Observation-Decision Process,* Proc. ACC, Seattle, (2008).

[21] R.L. Schilling, *Measures, Integrals and Martingales,* Cambridge, New York, 2005.