

# Distributed Dynamic Programming for Discrete-Time Stochastic Control, and Idempotent Algorithms

William M. McEneaney\*

March 20, 2009

## 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. Instead, it is sufficient that certain solution forms are retained through application of the associated semigroup operator. Here, we see that idempotent methods may be used to solve some classes of stochastic control problems. The key is the use of the idempotent distributive property. This allows one to apply the curse-of-dimensionality-free idempotent approach. We demonstrate this approach for a class of nonlinear, discrete-time stochastic control problems.

## 1 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. Here, max-plus methods are appropriate for problems

---

\*Dept. of Mechanical and Aerospace Engineering, University of California San Diego, San Diego, CA 92093-0411, USA. [mceneaney@ucsd.edu](mailto:mceneaney@ucsd.edu) Research partially supported by AFOSR grant FA9550-06-1-0238, NSF grant DMS-0307229 and Northrop-Grumman.

with maximizing controllers and vice-versa. These methods include max-plus basis-expansion approaches [1], [2], [6], [7], [11], [15], [18], as well as the more recently developed curse-of-dimensionality-free methods [11], [16], [17]. However, stochastic control problems have eluded idempotent methods.

In a recent application, a min-plus approach was useful for a problem in sensing control [20], [21]. In that example, the state process was an observation-conditioned probability vector, where the stochasticity of the state process was due to the stochastic observation inputs. This led to the realization that idempotent methods were indeed applicable to stochastic control problems. The key tool that had been missing previously was simply the idempotent distributive property.

It was also necessary to realize that idempotent linearity of the associated semigroup was not required for applicability of curse-of-dimensionality-free types of idempotent methods. Instead, it is sufficient that certain solution forms are retained through application of the semigroup operator, i.e., the dynamic programming principle operator. We will see that, under certain conditions, pointwise minima of affine and quadratic forms will pass through this operator, when we have a minimizing control problem. As the operator contains an expectation component, this will require 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 will be propagated backwards with a representation as a pointwise minimum of quadratic or affine forms.

This approach will have some similarities to the idempotent curse-of-dimensionality-free methods developed for deterministic control problems [11], [13], [14], [16]. However, the need to use the idempotent distributive property makes this approach very different from those.

We will use a very limited class of control problems to introduce this new approach. We consider discrete-time, finite-time horizon control problems. The dynamics will contain a possibly continuum-valued stochastic input. They will also contain both a possibly continuum-valued control, and a control taking values in a finite set. The running cost will also depend on these inputs. We may think of the dynamics and the running cost as being indexed by the values of the finite control input. In the actual computational algorithms which will be obtained, each indexed dynamics will be linear, and each indexed running cost will be either quadratic or affine.

A similar approach was taken with the curse-of-dimensionality-free methods for deterministic control. In that case, it was shown that any semiconvex [respectively, convex] Hamiltonian could be arbitrarily well-approximated as a pointwise maximum of quadratic [respectively, affine] forms, and so this class is not so restrictive.

Note that the purpose of this paper is the announcement of this newly discovered class of algorithms. The general theory is indicated in an abstract, albeit discrete-time, setting, resulting in what we refer to here as the idempotent distributed dynamic programming principle (IDDPP). The basic algorithms will be described for two relatively simple classes of problems, where the value function will be given in terms of pointwise minima of quadratic and affine forms. Instantiation of these basic algorithms in computationally efficient forms requires approximation steps. A study of such approximations will not be included here, but the reader may refer to [20], [21] for some initial approximation methods relevant to this class of algorithms.

The layout of the paper is as follows. In Section 2, a finite time-horizon, discrete-time stochastic control problem is defined, and the standard dynamic programming principle (DPP) is given in this context. In Section 3, a continuum version of the min-plus distributive property is proved. This is the result that will allow us to obtain the main result. This main result is the IDDPP, which is obtained for this class of problems in Section 4. In Sections 5 and 6, the general IDDPP is reduced to computationally amenable forms for specific problem formulations.

## 2 Problem Definition and Dynamic Program

We begin by defining the specific class of problems which will be addressed here. In this initial foray into the domain of idempotent methods for stochastic control problems, we restrict ourselves to a rather narrow problem class. We will consider only discrete-time, finite time-horizon problems. Let the dynamics take the form

$$\xi_{t+1} = f(\xi_t, u_t, \mu_t, w_t) \tag{1}$$

$$\xi_s = x \in \mathbb{R}^n \tag{2}$$

where  $f$  is measurable, with more assumptions on it to follow. The  $u_t$  and  $\mu_t$  will be control inputs, and we find it helpful to differentiate these into the continuum-valued,  $u_t$ , and discrete-valued,  $\mu_t$ , components (see [11], [17] for

motivation). The  $w_t$  will be the stochastic disturbance inputs. The initial time is  $s$ , and terminal time is  $T$ , and we specifically write the time period as  $]s, T[ \doteq \{s, s+1, \dots, T\}$ .

We will use the term probability triple to denote a triple consisting of a (topological) sample space, the collection of Borel sets on the sample space, and a probability measure on the Borel sets; the generic notation will be  $(\Omega, \mathcal{B}, P)$  with various sub- and superscripts distinguishing such triples. We suppose that the range of each random  $w_t$  is separable metric space  $(W, d_w)$  (where, for simplicity, we skip the additional generality of letting the range space depend on time). We also suppose that each  $w_t$  is a random variable with respect to the time-indexed probability triple  $(\Omega_t^w, \mathcal{B}_t^w, P_t^w)$ , that is  $w_t : \Omega_t^w \rightarrow W$  is measurable with respect to  $\mathcal{B}_t^w$ . For each  $t \in ]s, T[$ , let  $\overline{\Omega}_t^w \doteq \prod_{r=s}^t \Omega_r^w$  where the product notation, indicating outer product here, will be used both for standard product and outer product; the meaning will be clear from context. Let  $\overline{\mathcal{B}}_t^w$  be the Borel sets on  $\overline{\Omega}_t^w$  (which we note are generated by the rectangular outer products of the  $\mathcal{B}_r^w$ , c.f., [8]). Let  $\overline{P}_t^w$  be the resulting probability measure on  $\overline{\mathcal{B}}_t^w$ .

We suppose each  $u_t : \overline{\Omega}_t^w \rightarrow U \subseteq \mathbb{R}^k$  and each  $\mu_t : \overline{\Omega}_t^w \rightarrow \mathcal{M} = ]1, 2 \dots M[$  is measurable with respect to Borel sets  $\overline{\mathcal{B}}_t^w$ . We suppose  $d_m(\cdot, \cdot)$  is a metric on  $\mathcal{M}$ . More particularly, we let the sets of allowable controls at time  $t$  be denoted by

$$\begin{aligned} \mathcal{U}_t &\doteq \{u_t : \overline{\Omega}_t^w \rightarrow U \mid \overline{\mathcal{B}}_{t-1}^w \times \{\Omega_t^w, \emptyset\} \text{ measurable}\}, \\ \mathcal{M}'_t &\doteq \{\mu_t : \overline{\Omega}_t^w \rightarrow \mathcal{M} \mid \overline{\mathcal{B}}_{t-1}^w \times \{\Omega_t^w, \emptyset\} \text{ measurable}\}, \end{aligned}$$

where the special case  $t = s$  is clear (one drops the  $\overline{\mathcal{B}}_{t-1}^w$  term). Let the allowable sets of control processes be denoted by  $\mathcal{U}_{]s, T-1[} \doteq \prod_{t=s}^{T-1} \mathcal{U}_t$  and  $\mathcal{M}'_{]s, T-1[} \doteq \prod_{t=s}^{T-1} \mathcal{M}'_t$ . Let generic elements of  $U$ ,  $\mathcal{M}$  and  $W$  be denoted by  $u$ ,  $m$  and  $w$ , respectively.

The payoff (to be minimized) will be

$$J(s, x, u, \mu) \doteq \mathbf{E} \left\{ \left[ \sum_{t=s}^{T-1} l(t, \xi_t, u_t, \mu_t, w_t) \right] + \Psi(\xi_T) \right\}, \quad (3)$$

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. It is important to note that the  $\mu_t$  process may serve as a means of generating

dynamics and payoff models as well as (or alternatively to) serving as a controller. That is, one may define a Hamiltonian which is a pointwise maximum of simpler (typically quadratic) forms, as a means for approximating a more general Hamiltonian, and then solve the problem with this approximating Hamiltonian as a means of obtaining an approximate solution of the original problem. This is the main reason for having the  $\mu$ . and  $u$ . processes separately denoted. See [11] for discussion of a similar approach in the context of deterministic control.

The value function for this control problem is

$$V_s(x) \doteq \inf_{u \in \mathcal{U}_{[s, T-1]}} \inf_{\mu \in \mathcal{M}'_{[s, T-1]}} J(s, x, u., \mu.). \quad (5)$$

We will assume that given  $\varepsilon > 0$ ,  $R < \infty$  and  $\bar{w} \in W$ , there exists  $\delta > 0$  such that

$$|f(x, u, m, w) - f(\bar{x}, u, m, \bar{w})| < \varepsilon \quad (A.1)$$

for all  $u \in U$ ,  $m \in \mathcal{M}$ ,  $w \in B_\delta(\bar{w})$ , and  $x, \bar{x} \in B_R(0)$  such that  $|x - \bar{x}| < \delta$ .

We assume that given  $R < \infty$ , there exists  $\hat{D}_R < \infty$  such that

$$|f(x, u, m, w)| \leq \hat{D}_R \quad (A.2)$$

for all  $x \in B_R(0)$ ,  $u \in U$ ,  $m \in \mathcal{M}$  and  $w \in W$ .

Regarding the running cost, we assume that given  $R < \infty$ , there exists  $C_R < \infty$  such that

$$|l(t, x, u, m, w)| \leq C_R \quad (A.3)$$

for all  $t \in ]s, T - 1[$ ,  $x \in B_R(0)$ ,  $u \in U$ ,  $m \in \mathcal{M}$  and  $w \in W$ .

We also assume that given  $\varepsilon > 0$ ,  $R < \infty$  and  $\bar{w} \in W$ , there exists  $\delta > 0$  such that

$$|l(t, x, u, m, w) - l(t, \bar{x}, u, m, \bar{w})| < \varepsilon \quad (A.4)$$

for all  $t \in ]s, T - 1[$ ,  $u \in U$ ,  $m \in \mathcal{M}$ ,  $w \in B_\delta(\bar{w})$ , and  $x, \bar{x} \in B_R(0)$  such that  $|x - \bar{x}| < \delta$ .

Regarding the terminal cost, we assume that given  $R < \infty$ , there exists  $C_R^T < \infty$  such that

$$|g_T(x, z_T)| \leq C_R^T \quad (A.5)$$

for all  $z_T \in Z_T$  and  $x \in B_R(0)$ .

Lastly, we assume that given  $\varepsilon > 0$  and  $\bar{R} < \infty$ , there exists  $\bar{\delta} > 0$  such that

$$|g_T(x, z_T) - g_T(\bar{x}, z_T)| < \varepsilon \quad (\text{A.6})$$

for all  $z_T \in Z_T$  and  $x, \bar{x} \in B_{\bar{R}}(0)$  such that  $|x - \bar{x}| < \bar{\delta}$ .

It is useful to note that, under these assumptions, given initial  $\xi_s = x$ , the state process is bounded independent of sample path. Further, the value function is finite for all  $s \in ]0, T[, x \in \mathbb{R}^n$ .

In this context, the dynamic programming principle (DPP) takes the following form. As it is entirely standard, a proof is not included.

**Theorem 2.1** *Let the value function  $V_t(x)$  be given by (1)–(5) for any  $t \in ]s + 1, T[$  and  $x \in \mathbb{R}^n$ . For any  $t \in ]s + 1, T[$ ,*

$$V_{t-1}(x) = \inf_{u \in U} \min_{m \in \mathcal{M}} \mathbf{E} \{l(t-1, x, u, m, w_{t-1}) + V_t(f(x, u, m, w_{t-1}))\} \quad \forall x \in \mathbb{R}^n. \quad (6)$$

For any Borel set,  $\mathcal{A} \subseteq W$ , let  $p_t^w(\mathcal{A}) \doteq P_t^w(w_t^{-1}(\mathcal{A}))$ , and note that this defines a probability measure on  $(W, \mathcal{B}^W)$ , where  $\mathcal{B}^W$  denotes the Borel sets on  $W$ . Then (6) may alternatively be written as

$$V_{t-1}(x) = \inf_{u \in U} \min_{m \in \mathcal{M}} \int_W [l(t-1, x, u, m, w) + V_t(f(x, u, m, w))] dp_{t-1}^w(w), \quad (7)$$

for all  $x \in \mathbb{R}^n$ .

The goal in this paper is the elucidation of an idempotent numerical approach to stochastic control problems. The approach will rely on the inheritance of certain functional forms through the expectation operator. Here, we will use quadratic forms as the functional forms because these form a min-plus basis for the min-plus vector space (i.e., moduloid) of semiconcave functions [11]. The key to an idempotent approach to stochastic control problems will be the idempotent distributive property (specifically the min-plus distributive property here). In the case of finite sums and products, this takes the standard form; here it will take a slightly different form which will nonetheless be familiar to control theorists. Lastly, there will be technical issues that must be sorted through in order to actually apply this distributive property in the context of the above DPP. In the next section, we present a general result about the min-plus distributive property. In the section following that, we work through the results necessary to apply that general result in this context.

### 3 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  $\mathbb{R}^+ \doteq \mathbb{R} \cup \{+\infty\}$  given by

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

c.f., [3], [9], [11]. 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  $\mathcal{J} = ]1, J[$ ,

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

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

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

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. The particular assumptions used here might not be necessary, but are sufficient for our needs.

**Theorem 3.1** *Let  $(Z, d_z)$  be a separable metric space. Recall that  $(W, d_w)$  is a separable metric space with Borel sets  $\mathcal{B}^W$ . Let  $p$  be a finite measure on  $(W, \mathcal{B}^W)$ , and let  $\bar{D} \doteq p(W)$ . Let  $h : W \times Z \rightarrow \mathbb{R}$  be Borel measurable, and suppose*

$$\left| \int_W \inf_{z \in Z} h(w, z) dp(w) \right| < \infty.$$

Also assume that, given  $\varepsilon > 0$  and  $\bar{w} \in W$ , there exists  $\delta > 0$  such that

$$|h(w, z) - h(\bar{w}, z)| < \varepsilon \quad \forall z \in Z, w \in B_\delta(\bar{w}).$$

Then,

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

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

PROOF. Let  $\tilde{z}_0 \in \tilde{Z}$ . Then,  $h(w, \tilde{z}_0(w)) \geq \inf_{z \in Z} h(w, z)$  for all  $w \in W$ , and so

$$\int_W h(w, \tilde{z}_0(w)) dp(w) \geq \int_W \inf_{z \in Z} h(w, z) dp(w).$$

Since this is true for all  $\tilde{z}_0 \in \tilde{Z}$ , one has

$$\inf_{\tilde{z} \in \tilde{Z}} \int_W h(w, \tilde{z}(w)) dp(w) \geq \int_W \inf_{z \in Z} h(w, z) dp(w). \quad (8)$$

We now turn to the reverse inequality. By the separability of  $W$ , there exists a countable, dense subset  $\{w_i \mid i \in \mathcal{I}\}$  in  $W$ . Fix  $\varepsilon > 0$ . Given  $i \in \mathcal{I}$ , by assumption, there exists  $\delta_i > 0$  such that

$$|h(w, z) - h(w_i, z)| < \varepsilon \quad \forall w \in B_{\delta_i}(w_i), \forall z \in Z. \quad (9)$$

Let  $\mathcal{D}_1 \doteq B_{\delta_1}(w_1)$ . Suppose one has obtained  $\mathcal{D}_j$  for all  $j \leq i$ . Then let

$$\mathcal{D}_{i+1} \doteq B_{\delta_{i+1}}(w_{i+1}) \setminus \left[ \bigcup_{j=1}^i \mathcal{D}_j \right]. \quad (10)$$

Let  $\tilde{\mathcal{I}} \doteq \{i \in \mathcal{I} \mid \mathcal{D}_i \neq \emptyset\}$ . By the density of the  $w_i$ , one sees that

$$W = \bigcup_{i \in \tilde{\mathcal{I}}} \mathcal{D}_i, \quad \text{and } \mathcal{D}_i \cap \mathcal{D}_j = \emptyset \quad \forall i \neq j. \quad (11)$$

For  $i \in \tilde{\mathcal{I}}$ , let  $z_i \in Z$  be such that

$$h(w_i, z_i) \leq \inf_{z \in Z} [h(w_i, z)] + \varepsilon. \quad (12)$$

Define  $\tilde{z}^\varepsilon : W \rightarrow Z$  by

$$\tilde{z}^\varepsilon(w) = z_i \quad \text{if } w \in \mathcal{D}_i, \quad (13)$$



and note that, by (11), this is well-defined.

Note that  $\mathcal{D}_1 = B_{\delta_1}(w_1)$  is measurable. If the  $\mathcal{D}_j$  are measurable for all  $j \leq i$ , then by (10),  $\mathcal{D}_{i+1}$  is measurable. By induction,  $\mathcal{D}_i$  is measurable for all  $i \in \tilde{\mathcal{I}}$ . Then, by (13),  $\tilde{z}^\varepsilon$  is measurable, which by the assumed measurability of  $h$ , implies that  $h(\cdot, \tilde{z}^\varepsilon(\cdot))$  is a (Borel) measurable map from  $W$  to  $\mathbb{R}$ .

Now fix any  $\bar{w} \in W$ . Let  $\bar{i}$  be the unique element of  $\tilde{\mathcal{I}}$  such that  $\bar{w} \in \mathcal{D}_{\bar{i}}$ . Let  $\bar{z} \in Z$  be such that

$$\inf_{z \in Z} h(\bar{w}, z) \geq h(\bar{w}, \bar{z}) - \varepsilon.$$

and then, since  $\bar{w} \in \mathcal{D}_{\bar{i}}$ ,  $d_w(\bar{w}, w_{\bar{i}}) < \delta_i$ , and so by (9),

$$\geq h(w_{\bar{i}}, \bar{z}) - 2\varepsilon,$$

which by (12), (13),

$$\geq h(w_{\bar{i}}, \tilde{z}^\varepsilon(w_{\bar{i}})) - 3\varepsilon$$

which by (9) again,

$$\geq h(\bar{w}, \tilde{z}^\varepsilon(w_{\bar{i}})) - 4\varepsilon,$$

which by (13),

$$= h(\bar{w}, \tilde{z}^\varepsilon(\bar{w})) - 4\varepsilon.$$

Since  $\bar{w} \in W$  was arbitrary, one has

$$h(w, \tilde{z}^\varepsilon(w)) \leq \inf_{z \in Z} h(w, z) + 4\varepsilon \quad \forall w \in W. \quad (14)$$

Consequently,

$$\int_W h(w, \tilde{z}^\varepsilon(w)) dp(w) \leq \int_W \left[ \inf_{z \in Z} h(w, z) + 4\varepsilon \right] dp(w),$$

which by assumption,

$$\leq \int_W \inf_{z \in Z} h(w, z) dp(w) + 4\bar{D}\varepsilon.$$

This immediately implies

$$\inf_{\tilde{z} \in \tilde{Z}} \int_W h(w, \tilde{z}(w)) dp(w) \leq \int_W \inf_{z \in Z} h(w, z) dp(w) + 4\bar{D}\varepsilon.$$

Since  $\varepsilon > 0$  was arbitrary, one has

$$\inf_{\tilde{z} \in \tilde{Z}} \int_W h(w, \tilde{z}(w)) dp(w) \leq \int_W \inf_{z \in Z} h(w, z) dp(w). \quad (15)$$

Combining (8) and (15) yields the result.  $\square$

## 4 Distributed Dynamic Programming

We now use the above infinite-version of the distributive property in the context of the dynamic program of Section 2. Suppose  $V_t$  has the form

$$V_t(x) = \inf_{z_t \in Z_t} g_t(x, z_t), \quad (16)$$

where  $(Z_t, d_{z_t})$  is a separable metric space. Note that  $V_T(x) = \Psi(x)$  has this form (see (4)). Then the dynamic programming principle of (7) becomes

$$\begin{aligned} V_{t-1}(x) &= \inf_{u \in U} \min_{m \in \mathcal{M}} \int_W \left[ l(t-1, x, u, m, w) + \inf_{z_t \in Z_t} g_t(f(x, u, m, w), z_t) \right] dp_{t-1}^w(w) \\ &= \inf_{u \in U} \min_{m \in \mathcal{M}} \int_W \inf_{z_t \in Z_t} [l(t-1, x, u, m, w) + g_t(f(x, u, m, w), z_t)] dp_{t-1}^w(w) \end{aligned} \quad (17)$$

for all  $x \in \mathbb{R}^n$ .

We will use the infinite distributive property of Theorem 3.1 to move the infimum over  $Z_t$  outside the integral. Then, letting

$$\tilde{Z}_t \doteq \{ \tilde{z}_t : W \rightarrow Z_t \mid \text{Borel measurable} \},$$

one will have

$$V_{t-1}(x) = \inf_{z_{t-1} \in Z_{t-1}} g_{t-1}(x, z_{t-1})$$

where

$$Z_{t-1} = U \times \mathcal{M} \times \tilde{Z}_t,$$

and

$$g_{t-1}(x, z_{t-1}) = \int_W l(t-1, x, u, m, w) + g_t(f(x, u, m, w), z_t) dp_{t-1}^w(w). \quad (18)$$

Consequently, the general form as an infimum will be inherited from  $V_t$  to  $V_{t-1}$ . Thus, one could propagate backward in this way indefinitely, neglecting practical issues of course. This what is will be referred to as idempotent distributed dynamic programming principle (IDDPP).

In order for this concept to work, it must be shown that the conditions necessary for application of Theorem 3.1 will hold for  $V_{t-1}$  if they held for  $V_t$ . This will be the task of this section. First, under suitable conditions, we demonstrate applicability of Theorem 3.1 for a single step.

**Lemma 4.1** *Suppose that  $g_t : \mathbb{R}^n \times Z_t \rightarrow \mathbb{R}$  is Borel measurable, and that given  $\varepsilon > 0$  and  $\bar{R} < \infty$ , there exists  $\bar{\delta} > 0$  such that*

$$|g_t(x, z_t) - g_t(\bar{x}, z_t)| < \varepsilon \quad (19)$$

for all  $z_t \in Z_t$  and  $x, \bar{x} \in B_{\bar{R}}(0)$  such that  $|x - \bar{x}| < \bar{\delta}$ . Then, given  $\varepsilon > 0$ ,  $R < \infty$  and  $\bar{w} \in W$ , there exists  $\delta > 0$  such that

$$|g_t(f(x, u, m, w), z_t) - g_t(f(x, u, m, \bar{w}), z_t)| < \varepsilon$$

for all  $z_t \in Z_t$ ,  $x \in B_R(0)$ ,  $u \in U$ ,  $m \in \mathcal{M}$  and  $w \in B_\delta(\bar{w})$ .

PROOF. By (A.2), given  $R < \infty$ ,

$$|f(x, u, m, w)|, |f(x, u, m, \bar{w})| \leq \widehat{D}_R \quad (20)$$

for all  $x \in B_R(0)$ ,  $u \in U$ ,  $m \in \mathcal{M}$  and  $w, \bar{w} \in W$ . Further, by (A.1), given  $\bar{\delta} > 0$ ,  $R < \infty$  and  $\bar{w} \in W$ , there exists  $\delta > 0$  such that

$$|f(x, u, m, w) - f(x, u, m, \bar{w})| < \bar{\delta} \quad (21)$$

for all  $x \in B_R(0)$ ,  $u \in U$ ,  $m \in \mathcal{M}$  and  $w \in B_\delta(\bar{w})$ . Combining (20), (21) and (19) (with  $\bar{R} = \widehat{D}_R$ ), one obtains the result.  $\square$

**Theorem 4.2** *Suppose  $V_t$  is given by (16), where  $g_t$  is Borel measurable and satisfies (19), and  $(Z_t, d_{z_t})$  is a separable metric space. Also assume that given  $R < \infty$ , there exists  $\widetilde{D}_{t,R} < \infty$  such that*

$$|g_t(x, z_t)| \leq \widetilde{D}_{t,R} \quad \forall x \in B_R(0), z_t \in Z_t. \quad (22)$$

Then,

$$V_{t-1}(x) = \inf_{z_{t-1} \in Z_{t-1}} g_{t-1}(x, z_{t-1}), \quad (23)$$

where

$$Z_{t-1} = U \times \mathcal{M} \times \widetilde{Z}_t, \quad (24)$$

$$\widetilde{Z}_t = \{\tilde{z}_t : W \rightarrow Z_t \mid \text{Borel measurable}\}, \quad (25)$$

and

$$g_{t-1}(x, z_{t-1}) = \int_W l(t-1, x, u, m, w) + g_t(f(x, u, m, w), \tilde{z}_t(w)) dp_{t-1}^w(w). \quad (26)$$

PROOF. By (7)

$$V_{t-1}(x) = \inf_{u \in U} \min_{m \in \mathcal{M}} \int_W \inf_{z_t \in Z_t} [l(t-1, x, u, m, w) + g_t(f(x, u, m, w), z_t)] dp_{t-1}^w(w). \quad (27)$$

Letting

$$\hat{g}(t, x, u, m, w, z_t) \doteq l(t-1, x, u, m, w) + g_t(f(x, u, m, w), z_t), \quad (28)$$

we see that we must show that  $\hat{g}(t, x, u, m, \cdot, \cdot) : W \times Z_t \rightarrow \mathbb{R}$  satisfies the conditions on  $h(\cdot, \cdot)$  from Theorem 3.1. (Note that  $p_{t-1}^w(W) = 1 < \infty$ .) By assumption, it is Borel measurable. Now,

$$\begin{aligned} & \int_W \inf_{z_t \in Z_t} [\hat{g}(t, x, u, m, w, z_t)] dp_{t-1}^w(w) \quad (29) \\ &= \int_W l(t-1, x, u, m, w) dp_{t-1}^w(w) + \int_W \inf_{z_t \in Z_t} [g_t(f(x, u, m, w), z_t)] dp_{t-1}^w(w). \end{aligned}$$

Fix any  $x \in \mathbb{R}^n$ ,  $u \in U$  and  $m \in \mathcal{M}$ , and let  $R \geq |x|$ . By (A.3),

$$\int_W |l(t-1, x, u, m, w)| dp_{t-1}^w(w) \leq \int_W C_R dp_{t-1}^w(w) = C_R < \infty. \quad (30)$$

By (A.2),  $|f(x, u, m, w)| \leq \hat{D}_R$ . Using this in (22), one finds

$$|g_t(f(x, u, m, w), z_t)| \leq \tilde{D}_{t, \hat{D}_R},$$

where we note that this bound is independent of  $w \in W$  and  $z_t \in Z_t$ . Consequently,

$$\int_W \left| \inf_{z_t \in Z_t} g_t(f(x, u, m, w), z_t) \right| dp_{t-1}^w(w) \leq \tilde{D}_{t, \hat{D}_R} < \infty. \quad (31)$$

Combining (29), (30) and (31), one sees that

$$\left| \int_W \inf_{z_t \in Z_t} [\hat{g}(t, x, u, m, w, z_t)] dp_{t-1}^w(w) \right| \leq C_R + D_{t, \hat{D}_R} < \infty,$$

which is one of the conditions from Theorem 3.1.

It only remains to prove that the continuity condition on  $h(\cdot, \cdot)$  of Theorem 3.1 holds for  $\hat{g}(t, x, u, m, \cdot, \cdot)$ . Recall that  $x, u, m$  are fixed, and that  $R \geq |x|$ . Fix  $\bar{w} \in W$ . Note that

$$\begin{aligned} & |\hat{g}(t, x, u, m, w, z_t) - \hat{g}(t, x, u, m, \bar{w}, z_t)| \\ & \leq |l(t-1, x, u, m, w) - l(t-1, x, u, m, \bar{w})| \\ & \quad + |g_t(f(x, u, m, w), z_t) - g_t(f(x, u, m, \bar{w}), z_t)|. \end{aligned} \quad (32)$$

Applying Lemma 4.1 and Assumption (A.4) to the right-hand side of (32) implies that given  $\varepsilon > 0$ , there exists  $\delta > 0$  such that

$$|\hat{g}(t, x, u, m, w, z_t) - \hat{g}(t, x, u, m, \bar{w}, z_t)| < \varepsilon,$$

for all  $z_t \in Z_t$  and  $w \in B_\delta(\bar{w})$ , which is the desired continuity condition. Consequently, we may apply Theorem 3.1 with  $h(\cdot, \cdot)$  replaced by  $\hat{g}(t, x, u, m, \cdot, \cdot)$ , and so

$$\begin{aligned} & \int_W \inf_{z_t \in Z_t} [l(t-1, x, u, m, w) + g_t(f(x, u, m, w), z_t)] dp_{t-1}^w(w) \\ & \inf_{\tilde{z}_t \in \tilde{Z}_t} \int_W l(t-1, x, u, m, w) + g_t(f(x, u, m, w), z_t) dp_{t-1}^w(w). \end{aligned}$$

Substituting this into (27) yields the result.  $\square$

Now, recalling that

$$V_T(x) = \Psi(x) = \inf_{z_T \in Z_T} \{g_T(x, z_T)\}, \quad (33)$$

by (A.5) and (A.6),  $V_T$  has a form satisfying the conditions of Theorem 4.2. Although Theorem 4.2 allows us to proceed one step with the IDDP (idempotent distributed dynamic programming principle), from say  $V_t$  to  $V_{t-1}$ , it does not imply that the resulting  $V_{t-1}$  will be of a form meeting the conditions of Theorem 4.2. We now proceed to show that the conditions are inherited. This will allow us to repeatedly apply the IDDP indefinitely.

**Theorem 4.3** *Suppose the conditions of Theorem 4.2 hold, and let  $V_{t-1}$  be given by (23)–(26). Then, given any  $R < \infty$ , there exists  $\tilde{D}_{t-1, R} < \infty$  such that*

$$|g_{t-1}(x, z_{t-1})| \leq \tilde{D}_{t-1, R} \quad \forall x \in B_R(0), z_{t-1} \in Z_{t-1}.$$

Also, given  $\varepsilon > 0$  and  $\bar{R} < \infty$ , there exists  $\delta > 0$  such that

$$|g_{t-1}(x, z_{t-1}) - g_{t-1}(\bar{x}, z_{t-1})| < \varepsilon$$

for all  $z_{t-1} \in Z_{t-1}$  and  $x, \bar{x} \in B_{\bar{R}}(0)$  such that  $|x - \bar{x}| < \delta$ . lastly,  $(Z_{t-1}, d_{z_{t-1}})$  is a separable metric space, where

$$\begin{aligned} d_{z_{t-1}}(z_{t-1}, z_{t-1}') &= d_{z_{t-1}}((\mathbf{u}, \mathbf{m}, \tilde{z}_t), (\mathbf{u}', \mathbf{m}', \tilde{z}_t')) \\ &= |\mathbf{u} - \mathbf{u}'| + d_m(\mathbf{m}, \mathbf{m}') + \int_W d_{z_{t-1}}(\tilde{z}_t, \tilde{z}_t') dp_{t-1}^w(\mathbf{w}). \end{aligned}$$

PROOF. The last assertion is standard (c.f. [22]), and we do not include a proof.

Now, fix any  $R < \infty$ , and let  $x \in B_R(0)$ . Let  $z_{t-1} = (\mathbf{u}, \mathbf{m}, \tilde{z}_t)$ . From (26),

$$\begin{aligned} |g_{t-1}(x, z_{t-1})| &\leq \int_W |l(t-1, x, \mathbf{u}, \mathbf{m}, \mathbf{w})| dp_{t-1}^w(\mathbf{w}) \\ &\quad + \int_W |g_t(f(x, \mathbf{u}, \mathbf{m}, \mathbf{w}), \tilde{z}_t(\mathbf{w}))| dp_{t-1}^w(\mathbf{w}), \end{aligned}$$

which by (A.3) and the fact that  $P_{t-1}^w(W) = 1$ ,

$$\leq C_R + \int_W |g_t(f(x, \mathbf{u}, \mathbf{m}, \mathbf{w}), \tilde{z}_t(\mathbf{w}))| dp_{t-1}^w(\mathbf{w}). \quad (34)$$

However, by (A.2),

$$|f(x, \mathbf{u}, \mathbf{m}, \mathbf{w})| \leq \hat{D}_R \quad \forall \mathbf{w} \in W, \quad (35)$$

and so by (22),

$$|g_t(f(x, \mathbf{u}, \mathbf{m}, \mathbf{w}), \tilde{z}_t(\mathbf{w}))| \leq \tilde{D}_{t, \hat{D}_R} \quad \forall \mathbf{w} \in W. \quad (36)$$

Substituting (36) into (34) implies

$$|g_{t-1}(x, z_{t-1})| \leq C_R + \tilde{D}_{t, \hat{D}_R} < \infty,$$

which yields the first assertion with  $\tilde{D}_{t-1, R} \doteq C_R + \tilde{D}_{t, \hat{D}_R}$ .

All that remains is to show that  $g_{t-1}$  satisfies the continuity assertion. Let  $R < \infty$  and  $x, \bar{x} \in B_R(0)$ . Again let  $z_{t-1} = (u, m, \tilde{z}_t)$ . One has

$$\begin{aligned} & |g_{t-1}(x, z_{t-1}) - g_{t-1}(\bar{x}, z_{t-1})| \\ & \leq \int_W |l(t-1, x, u, m, w) - l(t-1, \bar{x}, u, m, w)| dp_{t-1}^w(w) \\ & \quad + \int_W |g_t(f(x, u, m, w), \tilde{z}_t(w)) - g_t(f(\bar{x}, u, m, w), \tilde{z}_t(w))| dp_{t-1}^w(w), \end{aligned}$$

and by (A.4), there exists  $\delta_1 > 0$  such that

$$< \frac{\varepsilon}{2} + \int_W |g_t(f(x, u, m, w), \tilde{z}_t(w)) - g_t(f(\bar{x}, u, m, w), \tilde{z}_t(w))| dp_{t-1}^w(w), \quad (37)$$

if  $|x - \bar{x}| < \delta_1$ . As in (35),

$$|f(x, u, m, w)|, |f(\bar{x}, u, m, w)| \leq \widehat{D}_R \quad \forall w \in W. \quad (38)$$

By (19) (also assumed in Theorem 4.2), there exists  $\bar{\delta} > 0$  such that

$$|g_t(y, \tilde{z}_t(w)) - g_t(\bar{y}, \tilde{z}_t(w))| < \frac{\varepsilon}{2} \quad (39)$$

for all  $y, \bar{y} \in B_{\widehat{D}_R}(0)$  such that  $|y - \bar{y}| \leq \bar{\delta}$ . Also, by (A.1), there exists  $\delta_2 > 0$  such that

$$|f(x, u, m, w) - f(\bar{x}, u, m, w)| < \bar{\delta} \quad (40)$$

if  $|x - \bar{x}| < \delta_2$ . By (38)–(40), there exists  $\delta_2 > 0$  such that

$$|g_t(f(x, u, m, w), \tilde{z}_t(w)) - g_t(f(\bar{x}, u, m, w), \tilde{z}_t(w))| < \frac{\varepsilon}{2} \quad \forall w \in W \quad (41)$$

if  $|x - \bar{x}| < \delta_2$ . Substituting (41) into (37), one sees that

$$|g_{t-1}(x, z_{t-1}) - g_{t-1}(\bar{x}, z_{t-1})| < \frac{\varepsilon}{2} + \int_W \frac{\varepsilon}{2} dp_{t-1}^w(w) = \varepsilon$$

if  $|x - \bar{x}| < \delta \doteq \min\{\delta_1, \delta_2\}$ , and so the continuity assertion is proved.  $\square$

In summary, we have the following IDDP-based algorithm for the computation of  $V(\cdot)$ :

$$V_T(x) = \inf_{z_T \in Z_T} \{g_T(x, z_T)\}.$$

For  $t \in ]s+1, T[$ , one has

$$V_{t-1}(x) = \inf_{z_{t-1} \in Z_{t-1}} g_{t-1}(x, z_{t-1}), \quad (42)$$

where

$$Z_{t-1} = U \times \mathcal{M} \times \tilde{Z}_t, \quad (43)$$

$$\tilde{Z}_t = \{ \tilde{z}_t : W \rightarrow Z_t \mid \text{Borel measurable} \}, \quad (44)$$

and

$$g_{t-1}(x, z_{t-1}) = \int_W l(t-1, x, u, m, w) + g_t(f(x, u, m, w), \tilde{z}_t(w)) dp_{t-1}^w(w). \quad (45)$$

## 5 Quadratic Forms

A particularly useful class of forms to examine is that of quadratic cost criteria with affine-in-state dynamics. As an aside, it is useful to note that quadratic Hamiltonians form a min-plus basis for the space of semiconvex Hamiltonians, and that this (in the max-plus/semiconvex case) was used for deterministic control problems [11], [13], [14], [16], [17]. Further, we will see that in such cases, the  $g_t$  will all be quadratics, and so backward propagation via the IDDP reduces to propagation of the coefficients in the quadratic forms. Consequently, such problems are not subject to the curse-of-dimensionality.

We consider the case where

$$\begin{aligned} f(x, u, m, w) &\doteq A^m(u, w)x + B^m(u, w), \\ l(t, x, u, m, w) &= \frac{1}{2} [(x - \bar{x}_t^m(u, w))^T \bar{Q}_t^m(u, w)(x - \bar{x}_t^m(u, w)) + \bar{c}_t^m(u, w)], \end{aligned}$$

and

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

We will assume that given  $\varepsilon > 0$  and  $\bar{w} \in W$ , there exists  $\delta > 0$  such that

$$\begin{aligned} |A^m(u, w) - A^m(u, \bar{w})| &< \varepsilon \quad \forall m \in \mathcal{M}, u \in U, w \in B_\delta(\bar{w}), \\ |B^m(u, w) - B^m(u, \bar{w})| &< \varepsilon \quad \forall m \in \mathcal{M}, u \in U, w \in B_\delta(\bar{w}). \end{aligned} \quad (B.1)$$

We also assume that there exists  $D < \infty$  such that

$$|A^m(u, w)|, |B^m(u, w)| \leq D \quad \forall m \in \mathcal{M}, u \in U, w \in W. \quad (B.2)$$

We assume that given  $\varepsilon > 0$  and  $\bar{w} \in W$ , there exists  $\delta > 0$  such that



$$\begin{aligned}
|\bar{x}_t^m(\mathbf{u}, \mathbf{w}) - \bar{x}_t^m(\mathbf{u}, \bar{\mathbf{w}})| &< \varepsilon, \\
|\bar{Q}_t^m(\mathbf{u}, \mathbf{w}) - \bar{Q}_t^m(\mathbf{u}, \bar{\mathbf{w}})| &< \varepsilon, \\
|\bar{c}_t^m(\mathbf{u}, \mathbf{w}) - \bar{c}_t^m(\mathbf{u}, \bar{\mathbf{w}})| &< \varepsilon,
\end{aligned} \tag{B.3}$$

for all  $t \in ]s, T - 1[$ ,  $m \in \mathcal{M}$ ,  $\mathbf{u} \in U$ ,  $\mathbf{w} \in B_\delta(\bar{\mathbf{w}})$ . We assume that (with a possibly larger  $D < \infty$ ),

$$|\bar{x}_t^m(\mathbf{u}, \mathbf{w})|, |\bar{Q}_t^m(\mathbf{u}, \mathbf{w})|, |\bar{c}_t^m(\mathbf{u}, \mathbf{w})| < D \tag{B.4}$$

for all  $t \in ]s, T - 1[$ ,  $m \in \mathcal{M}$ ,  $\mathbf{u} \in U$ ,  $\mathbf{w} \in W$ . We assume that there exists  $\bar{C}_T < \infty$  such that

$$|x_T(z_T)|, |Q_T(z_T)|, |c_T(z_T)| < \bar{C}_T \tag{B.5}$$

for all  $z_T \in Z_T$ .

Lastly, we assume that for any  $z_T \in Z_T$ ,  $t \in ]s, T - 1[$ ,  $m \in \mathcal{M}$ ,  $\mathbf{u} \in U$  and  $\mathbf{w} \in W$ ,

$$\begin{aligned}
Q_T(z_T) &\text{ is positive definite (alternatively, negative definite),} \\
A^m(\mathbf{u}, \mathbf{w}) &\text{ is nonsingular, and } Q_t^m(\mathbf{u}, \mathbf{w}) \text{ is positive semidefinite} \\
&\text{(alternatively, negative semidefinite).}
\end{aligned} \tag{B.6}$$

It is easy to see that given Assumptions (B.1)–(B.5), Assumptions (A.1)–(A.6) hold. Consequently, the IDDP of the previous section may be applied to solve the problem. However, the computation of  $g_{t-1}$  from  $g_t$  is greatly simplified. In that regard, one may note that the last assumption, (B.6), guarantees that one may actually implement the computational scheme below (see (49)); alternate assumptions could guarantee this as well. Suppose

$$g_t(x, z_t) = \frac{1}{2} [(x - x_t(z_t))^T Q_t(z_t)(x - x_t(z_t)) + c_t(z_t)], \tag{46}$$

which is certainly the case for  $t = T$ . Then, by (26),

$$\begin{aligned}
g_{t-1}(x, z_{t-1}) &= g_{t-1}(x, (\mathbf{u}, \mathbf{m}, \tilde{z}_t)) \\
&= \int_W \frac{1}{2} [(x - \bar{x}_{t-1}^m(\mathbf{u}, \mathbf{w}))^T \bar{Q}_{t-1}^m(\mathbf{u}, \mathbf{w})(x - \bar{x}_{t-1}^m(\mathbf{u}, \mathbf{w})) + \bar{c}_{t-1}^m(\mathbf{u}, \mathbf{w})] dp_{t-1}^w(\mathbf{w}) \\
&\quad + \int_W \frac{1}{2} \left\{ [A^m(\mathbf{u}, \mathbf{w})x + B^m(\mathbf{u}, \mathbf{w}) - x_t(\tilde{z}_t(\mathbf{w}))]^T Q_t(\tilde{z}_t(\mathbf{w})) \right. \\
&\quad \quad \left. \cdot [A^m(\mathbf{u}, \mathbf{w})x + B^m(\mathbf{u}, \mathbf{w}) - x_t(\tilde{z}_t(\mathbf{w}))] + c_t(\tilde{z}_t(\mathbf{w})) \right\} dp_{t-1}^w(\mathbf{w}) \\
&= \int_W \frac{1}{2} [(x - x_{t-1}(z_{t-1}))^T Q_{t-1}(z_{t-1})(x - x_{t-1}(z_{t-1}))
\end{aligned}$$

$$+c_{t-1}(z_{t-1})] dp_{t-1}^w(\mathbf{w}), \quad (47)$$

where

$$Q_{t-1}(z_{t-1}) = \int_W \bar{Q}_{t-1}^m(\mathbf{u}, \mathbf{w}) + [A^m(\mathbf{u}, \mathbf{w})]^T Q_t(\tilde{z}_t(\mathbf{w})) A^m(\mathbf{u}, \mathbf{w}) dp_{t-1}^w(\mathbf{w}), \quad (48)$$

$$x_{t-1}(z_{t-1}) = Q_{t-1}^{-1}(z_{t-1}) \int_W \left[ \bar{Q}_{t-1}^m(\mathbf{u}, \mathbf{w}) \bar{x}_{t-1}^m(\mathbf{u}, \mathbf{w}) + [A^m(\mathbf{u}, \mathbf{w})]^T Q_t(\tilde{z}_t(\mathbf{w})) \cdot (x_t(\tilde{z}_t(\mathbf{w})) - B^m(\mathbf{u}, \mathbf{w})) \right] dp_{t-1}^w(\mathbf{w}), \quad (49)$$

(the existence of the inverse being guaranteed by Assumption (B.6)),

$$\begin{aligned} c_{t-1}(z_{t-1}) = & \int_W (\bar{x}_{t-1}^m(\mathbf{u}, \mathbf{w}))^T \bar{Q}_{t-1}^m(\mathbf{u}, \mathbf{w}) \bar{x}_{t-1}^m(\mathbf{u}, \mathbf{w}) \\ & + [x_t(\tilde{z}_t(\mathbf{w})) - B^m(\mathbf{u}, \mathbf{w})]^T Q_t(\tilde{z}_t(\mathbf{w})) [x_t(\tilde{z}_t(\mathbf{w})) - B^m(\mathbf{u}, \mathbf{w})] \\ & + \bar{c}_{t-1}^m(\mathbf{u}, \mathbf{w}) + c_t(\tilde{z}_t(\mathbf{w})) dp_{t-1}^w(\mathbf{w}) \\ & - [x_{t-1}(z_{t-1})]^T Q_{t-1}(z_{t-1}) x_{t-1}(z_{t-1}). \end{aligned} \quad (50)$$

In general, the functions  $Q_t(\cdot)$ ,  $x_t(\cdot)$ ,  $c_t(\cdot)$  will be over very large spaces,  $(Z_t, d_{z_t})$ . Discrete approximations will be needed for computation. In other works on idempotent methods in deterministic control and estimation, basic concepts were first laid out, and then the, very technical, error analysis followed in later papers. We propose the same approach here, as the error analysis will be very long, both obscuring the main point and delaying announcement. However, in the special case where  $U$ ,  $W$  and  $Z_T$  are all finite, we can immediately obtain an explicit algorithm. This will allow us to briefly discuss the curse-of-complexity, as it will appear in such stochastic control problems. (Note that although these methods eliminate the curse-of-dimensionality, they are subject to curse-of-complexity difficulties [11], [16], [17]; methods for attenuation of the curse-of-complexity are discussed in [12], [13], [21].

Let  $U = ]1, N^u[$ ,  $Z_T = ]1, N_T[$ , and  $W = ]1, N^w[$ . Suppose measure  $P_t^w$  is independent of  $t$ , and introduce the notation  $\tilde{p}_w \doteq P_t^w(\{\mathbf{w}\})$ . Suppose  $Z_t = ]1, N_t[$ , which is true for  $t = T$ . Then,

$$\tilde{Z}_t = \{ \tilde{z} = \{z_w\} \mid z_w \in Z_t \forall w \in W \},$$

i.e., the set of sequences of length  $N^w$  of elements of  $Z_t$ . The cardinality of  $\tilde{Z}_t$  is  $\#\tilde{Z}_t = (\#Z_t)^{\#W} = (N_t)^{N^w}$ . For example, if  $W = \{1, 2\}$  and  $Z_t = \{1, 2, 3\}$ , then  $\tilde{Z}_t = \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3)\}$ .

As always, one has  $Z_{t-1} = U \times \mathcal{M} \times \tilde{Z}_t$ . However, we now see that, abusing notation, we may take  $Z_{t-1} = ]1, N_{t-1}[$  where  $N_{t-1} = N^u M(N_t)^{N^w}$ , and the  $z_{t-1} \in ]1, N_{t-1}[$  are given by  $z_{t-1} = \mathcal{H}(u, m, \{z_w\})$  where  $\mathcal{H}$  is a bijection from  $U \times \mathcal{M} \times \tilde{Z}_t$  to  $]1, N_{t-1}[$ .

In this special case,  $g_{t-1}$  is given by (47), where the coefficients in (48)–(50) may be obtained as follows. Let  $(u, m, \{z_w\}) = \mathcal{H}^{-1}(z_{t-1})$ . Then,

$$Q_{t-1}(z_{t-1}) = \sum_{w \in W} [\bar{Q}_t^m(u, w) + (A^m(u, w))^T Q_t(z_w) A^m(u, w)] \tilde{P}_w, \quad (51)$$

$$x_{t-1}(z_{t-1}) = Q_{t-1}^{-1}(z_{t-1}) \sum_{w \in W} \left[ \bar{Q}_{t-1}^m(u, w) \bar{x}_{t-1}^m(u, w) + [A^m(u, w)]^T Q_t(\tilde{z}_t(w)) (x_t(\tilde{z}_t(w)) - B^m(u, w)) \right] \tilde{P}_w, \quad (52)$$

$$\begin{aligned} c_{t-1}(z_{t-1}) &= \sum_{w \in W} \left[ (\bar{x}_{t-1}^m(u, w))^T \bar{Q}_{t-1}^m(u, w) \bar{x}_{t-1}^m(u, w) + [x_t(\tilde{z}_t(w)) - B^m(u, w)]^T Q_t(\tilde{z}_t(w)) [x_t(\tilde{z}_t(w)) - B^m(u, w)] \right. \\ &\quad \left. + \bar{c}_{t-1}^m(u, w) + c_t(\tilde{z}(w)) \tilde{P}_w \right] \\ &\quad - [x_{t-1}(z_{t-1})]^T Q_{t-1}(z_{t-1}) x_{t-1}(z_{t-1}). \end{aligned} \quad (53)$$

The curse-of-complexity now manifests itself in the very rapid growth of  $N_t$  as one propagates backward in time. Successful application of the approach requires judicious pruning of the set of constituent quadratics at each iteration. In a stochastic control example with linear payoff, and the state lying in a simplex, this can be reasonably achieved via a linear program [12], [21]. In a deterministic setting with quadratic forms similar to those here, very effective pruning has been achieved via convex programming and linear matrix inequalities [13].

## 6 Affine Forms

Affine forms are certainly a degenerate case of quadratic forms, but the form in which we write the quadratics in the previous section does not allow one to obtain the affine form by setting the quadratic multiplier to zero. (Nonetheless, we prefer to write the quadratics in the form of the previous section as it has proved efficient for max-plus methods for deterministic control, c.f. [11], [16], [18].) Since one cannot immediately obtain the affine-constituents

case from the quadratic-constituents case of the previous section, we sketch the affine case here.

We suppose

$$\begin{aligned} f(x, \mathbf{u}, \mathbf{m}, \mathbf{w}) &= A^{\mathbf{m}}(\mathbf{u}, \mathbf{w})x + B^{\mathbf{m}}(\mathbf{u}, \mathbf{w}), \\ l(t, x, \mathbf{u}, \mathbf{m}, \mathbf{w}) &= \bar{b}_t^{\mathbf{m}}(\mathbf{u}, \mathbf{w}) \cdot x + \bar{c}_t^{\mathbf{m}}(\mathbf{u}, \mathbf{w}), \end{aligned}$$

and

$$g_T(x, z_T) = b_T(z_T) \cdot x + c_T(z_T).$$

In this case, we still assume (B.1)–(B.2) from the previous section. However, we replace (B.3)–(B.5) with the following assumptions.

We assume that given  $\varepsilon > 0$  and  $\bar{\mathbf{w}} \in W$ , there exists  $\delta > 0$  such that

$$|\bar{b}_t^{\mathbf{m}}(\mathbf{u}, \mathbf{w}) - \bar{b}_t^{\mathbf{m}}(\mathbf{u}, \bar{\mathbf{w}})|, |\bar{c}_t^{\mathbf{m}}(\mathbf{u}, \mathbf{w}) - \bar{c}_t^{\mathbf{m}}(\mathbf{u}, \bar{\mathbf{w}})| < \varepsilon, \quad (C.3)$$

for all  $t \in ]s, T - 1[$ ,  $\mathbf{m} \in \mathcal{M}$ ,  $\mathbf{u} \in U$ ,  $\mathbf{w} \in B_\delta(\bar{\mathbf{w}})$ . We also assume that (with a possibly larger  $D < \infty$  than in (B.2)),

$$|\bar{b}_t^{\mathbf{m}}(\mathbf{u}, \mathbf{w})|, |\bar{c}_t^{\mathbf{m}}(\mathbf{u}, \mathbf{w})| < D \quad (C.4)$$

for all  $t \in ]s, T - 1[$ ,  $\mathbf{m} \in \mathcal{M}$ ,  $\mathbf{u} \in U$ ,  $\mathbf{w} \in W$ . Lastly, we assume that there exists  $\bar{C}_T < \infty$  such that

$$|b_T(z_T)|, |c_T(z_T)| < \bar{C}_T \quad (C.5)$$

for all  $z_T \in Z_T$ .

As in the quadratic case, it is not difficult to show that (A.1)–(A.6) hold under (B.1), (B.2) and (C.3)–(C.5). Consequently, all the results through Section 4 continue to hold, and as in the previous section, the computations greatly simplify. In particular, given that  $g_t(x, z_t) = b_t(z_t) \cdot x + c_t(z_t)$ , one has

$$\begin{aligned} g_{t-1}(x, z_{t-1}) &= g_{t-1}(x, \mathbf{u}, \mathbf{m}, \tilde{z}_t) \\ &= \int_W \bar{b}_{t-1}^{\mathbf{m}}(\mathbf{u}, \mathbf{w}) \cdot x + \bar{c}_{t-1}^{\mathbf{m}} dp_{t-1}^w(\mathbf{w}) \\ &\quad + \int_W b_t(\tilde{z}_t(\mathbf{w})) \cdot [A^{\mathbf{m}}(\mathbf{u}, \mathbf{w})x + B^{\mathbf{m}}(\mathbf{u}, \mathbf{w})] + c_t(\tilde{z}_t(\mathbf{w})) dp_{t-1}^w(\mathbf{w}) \\ &= b_{t-1}(z_{t-1}) \cdot x + c_{t-1}(z_{t-1}), \end{aligned}$$

where

$$b_{t-1}(z_{t-1}) = \int_W \bar{b}_{t-1}^m(u, w) + (A^m(u, w))^T b_t(\tilde{z}_t(w)) dp_{t-1}^w(w)$$

and

$$c_{t-1}(z_{t-1}) = \int_W \bar{c}_{t-1}^m(u, w) + c_t(\tilde{z}_t(w)) + (B^m(u, w))^T b_t(\tilde{z}_t(w)) dp_{t-1}^w(w).$$

For the finite discrete case, and with the same notation as in the previous section, the above reduce to

$$b_{t-1}(z_{t-1}) = \sum_{w \in W} [\bar{b}_{t-1}^m(u, w) + (A^m(u, w))^T b_t(\tilde{z}_t(w))] \tilde{P}_w \quad (54)$$

and

$$c_{t-1}(z_{t-1}) = \sum_{w \in W} [\bar{c}_{t-1}^m(u, w) + c_t(\tilde{z}_t(w)) + (B^m(u, w))^T b_t(\tilde{z}_t(w))] \tilde{P}_w. \quad (55)$$

We should remark that the affine case is likely mainly of interest when the state space is restricted to a compact set as in the observation control problem of [20], [21]. This is due to the fact that both the running and terminal costs are in the form of pointwise minima of affine functionals, and hence concave, while the controller is a minimizing controller. This is not an obstacle in the quadratic case, as one can approximate any semiconcave function as a pointwise minimum of quadratics, including convex functions.

## 7 Simple Example

This paper is a first introduction to the use of idempotent methods for stochastic control problems. There are many directions in which the theory must be developed in order to render it useful for a reasonably large problem class. For example, if one is solving a continuous-time problem, then one must develop the requisite error analysis for discrete-time based computations. Closer to the main point of this paper, there is a very high curse-of-complexity for this approach. Certain pruning techniques have been developed for deterministic problems [13], and similar concepts may be useful here. In fact, such would be expected to be a requirement for useful application of these concepts. Nonetheless, we provide a simple example problem to illustrate that even with the very little machinery we have so far developed, the method may still be applied.

We will solve a problem using affine forms over a probability simplex. In fact, since we will be working over a probability simplex, linear forms (rather

than affine) are sufficient. We will use one additional tool, which was developed previously. In particular, as the algorithm proceeds, we will drop linear functionals (indexed by the coefficients  $b_t(z_t)$ ) which do not contribute at all to the solution. (We also drop duplicates and near-duplicates, but that is trivial.) One can determine whether a linear functional (the test linear functional) is completely redundant by solution of a linear program. One simply thinks of the pointwise minimum of the other linear functionals as forming part of the boundary of a convex polytope. Other portions of the polytope are defined by the boundary of the probability simplex. Then, one maximizes a linear functional defined by the coefficients of the test functional. The sign of this maximum value indicates whether the test linear functional contributes at all. Note that if it does not contribute at all, then its progeny at the next step will not contribute either. This contribution-check is discussed in more detail in [12].

We solved a simple example over the probability simplex in  $\mathbb{R}^4$ , i.e., over

$$S^4 \doteq \left\{ x \in \mathbb{R}^4 \mid x_i \in [0, 1] \ \forall i \in ]1, 4[, \sum_{i=1}^4 x_i = 1 \right\}.$$

The  $B^m$ ,  $\bar{b}_t^m$ ,  $\bar{c}_t^m$  and  $c_T$  were all zero. We took  $M = 1$ ,  $U = \{1, 2\}$ ,  $W = \{1, 2\}$  and  $Z_T = \{1, 2, 3, 4\}$ . We also took  $T = 8$ . We generated the problem data randomly. For the results depicted below, the problem data were as follows.

$$A(1, 1) = \begin{bmatrix} 0.2861 & 0.0598 & 0.3270 & 0.3270 \\ 0.0322 & 0.9327 & 0.0029 & 0.0322 \\ 0.0227 & 0.0138 & 0.9408 & 0.0227 \\ 0.0708 & 0.0708 & 0.0106 & 0.8477 \end{bmatrix},$$

$$A(2, 1) = \begin{bmatrix} 0.0737 & 0.0617 & 0.4323 & 0.4323 \\ 0.3855 & 0.0495 & 0.1796 & 0.3855 \\ 0.3626 & 0.1528 & 0.1219 & 0.3626 \\ 0.3282 & 0.3282 & 0.2266 & 0.1171 \end{bmatrix},$$

$$A(1, 2) = \begin{bmatrix} 0.4551 & 0.0147 & 0.2651 & 0.2651 \\ 0.0605 & 0.8511 & 0.0279 & 0.0605 \\ 0.2978 & 0.0851 & 0.3193 & 0.2978 \\ 0.0579 & 0.0579 & 0.0164 & 0.8678 \end{bmatrix},$$

$$A(2,2) = \begin{bmatrix} 0.3751 & 0.1715 & 0.2267 & 0.2267 \\ 0.3817 & 0.0466 & 0.1900 & 0.3817 \\ 0.0230 & 0.0049 & 0.9492 & 0.0230 \\ 0.0537 & 0.0537 & 0.0038 & 0.8888 \end{bmatrix},$$

$$b_T(1) = \begin{bmatrix} 1.5962 \\ 2.4124 \\ 2.5733 \\ 1.8181 \end{bmatrix}, \quad b_T(2) = \begin{bmatrix} 2.2257 \\ 1.7037 \\ 2.6366 \\ 2.4170 \end{bmatrix},$$

$$b_T(3) = \begin{bmatrix} 1.6777 \\ 1.4002 \\ 1.2280 \\ 2.1969 \end{bmatrix}, \quad b_T(4) = \begin{bmatrix} 1.4048 \\ 2.1111 \\ 2.2505 \\ 1.3557 \end{bmatrix},$$

$$P(W = 1) = 0.0653 \quad \text{and} \quad P(W = 2) = 0.9347.$$

The value function at time,  $t = 1$ , was obtained. Since  $S^4$  is a three-dimensional region, we plot the solution only on a typical plane, the plane  $x_4 = 0.2$ ,  $x_3 = 1 - (x_1 + x_2 + 0.2)$ , in Figure 1. This example should indicate that, even without serious numerical pruning technology, the method is still feasible. More reasonable complexity-reduction will be applied in later iterations of the method, along with relevant error analyses; the main focus of this paper is simply a demonstration of the existence of this new class of methods for stochastic control.

## 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<sup>th</sup> International Symposium on Mathematical Theory of Networks and Systems (2004).
- [3] F.L. Baccelli, G. Cohen, G.J. Olsder and J.-P. Quadrat, *Synchronization and Linearity*, John Wiley, New York, 1992.

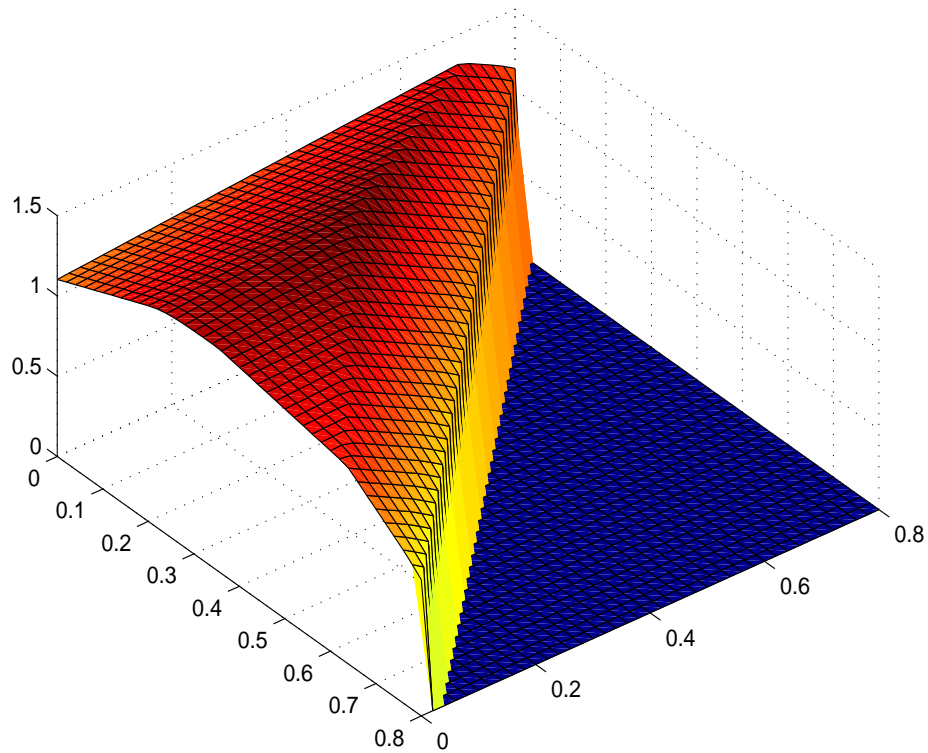


Figure 1: Value function along plane in  $S^4$ .

- [4] R.A. Cuninghame-Green, *Minimax Algebra*, Lecture Notes in Economics and Mathematical Systems 166, Springer, New York, 1979.
- [5] G. Cohen, S. Gaubert and J.-P. Quadrat, *Duality and Separation Theorems in Idempotent Semimodules*, Linear Algebra and Applications, 379 (2004), 395–422.
- [6] G. Collins and W.M. McEneaney, “Min-Plus Eigenvector Methods for Nonlinear  $H_\infty$  Problems with Active Control”, Optimal Control, Stabi-



- lization and Nonsmooth Analysis, LNCIS, Vol. 301 Queiroz, Marcio S. de; Malisoff, Michael; Wolenski, Peter (Eds.), Springer (2004), 101–120.
- [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] G.B. Folland, *Real Analysis, Modern Techniques and Their Applications*, John Wiley, New York, 1984.
- [9] V.N. Kolokoltsov and V.P. Maslov, *Idempotent Analysis and Its Applications*, Kluwer, 1997.
- [10] G.L. Litvinov, V.P. Maslov and G.B. Shpiz, *Idempotent Functional Analysis: An Algebraic Approach*, Mathematical Notes, 69 (2001), 696–729.
- [11] W.M. McEneaney, *Max-Plus Methods for Nonlinear Control and Estimation*, Birkhauser, Boston, 2006.
- [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.), AMS, (to appear).
- [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 and J.L. Kluberg, *Convergence Rate for a Curse-of-Dimensionality-Free Method for HJB PDEs Represented as Maxima fo Quadratic Forms*, SIAM J. Control and Optim., (submitted).
- [15] W.M. McEneaney, *Max-Plus Summation of Fenchel-Transformed Semigroups for Solution of Nonlinear Bellman Equations*, Systems and Control Letters, 56 (2007), 255-264.
- [16] 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.

- [17] W.M. McEneaney, “Curse-of-Dimensionality Free Method for Bellman PDEs with Hamiltonian Written as Maximum of Quadratic Forms” Proc. 44th IEEE Conf. on Decision and Control (2005).
- [18] W.M. McEneaney, *Max-Plus Eigenvector Representations for Solution of Nonlinear  $H_\infty$  Problems: Error Analysis*, SIAM J. Control and Optim., 43 (2004), 379–412.
- [19] W.M. McEneaney and P.M. Dower, *A max-plus affine power method for approximation of a class of mixed  $l_\infty/l_2$  value functions*, Proc. 42nd IEEE Conf. on Dec. and Control, Maui (2003), 2573-2578.
- [20] William M. McEneaney, Ali Oran and Andrew Cavender, *Value-Based Control of the Observation-Decision Process*, American Control Conf., Seattle, (2008).
- [21] William M. McEneaney, Ali Oran and Andrew Cavender, *Value-Based Tasking Controllers for Sensing Assets*, AIAA Guidance, Navigation and Control Conf., Honolulu, (2008).
- [22] R.L. Schilling, *Measures, Integrals and Martingales*, Cambridge, New York, 2005.