CONVERGENCE RATE FOR A CURSE-OF-DIMENSIONALITY-FREE METHOD FOR HJB PDES REPRESENTED AS MAXIMA OF QUADRATIC FORMS *

WILLIAM M. MCENEANEY †

Abstract. In previous work of the author and others, 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. Although max-plus basis expansion and max-plus finite-element methods provide computational-speed advantages, they still generally suffer from the curse-of-dimensionality. Here we consider HJB PDEs where the Hamiltonian takes the form of a (pointwise) maximum of linear/quadratic forms. The approach to solution will be rather general, but in order to ground the work, we consider only constituent Hamiltonians corresponding to long-run average-cost-per-unit-time optimal control problems for the development. We consider a previously obtained 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 linear/quadratic Hamiltonians. The dual-space semigroup is particularly useful due to its form as a max-plus integral operator with kernel obtained from the originating semigroup. One employs repeated application of the dual-space semigroup to obtain the solution. Here, we consider constituent Hamiltonians which contain linear and constant terms as well as purely quadratic terms. This greatly expands the allowable class of problems. However, there are solution existence issues, and the error bounds are more difficult.

Key words. partial differential equations, curse-of-dimensionality, dynamic programming, max-plus algebra, Legendre transform, Fenchel transform, semiconvexity, Hamilton-Jacobi-Bellman equations, idempotent analysis.

AMS subject classifications. 49LXX, 93C10, 35B37, 35F20, 65N99, 47D99

1. Introduction.

The use of dynamic programming (DP) to solve nonlinear control problems leads to the familiar DPE (dynamic programming equation). In the case of problems in continuous space/time governed by finite-dimensional "deterministic" (or max-plus stochastic) dynamics, the DPE takes the form of a Hamilton-Jacobi-Bellman partial differential equation (HJB PDE). In the infinite time-horizon case, this is a PDE over a region in a space whose dimension is the dimension of the state variable in the control problem. We remark that the solutions are generally nonsmooth, and the theory of viscosity solutions yields the appropriate solution definition (c.f., [4], [10], [11], [12], [21]).

The difficulty lies in computing the solution of the HJB PDE. The most intuitive, and commonly applied, approaches are grid-based (c.f., [4], [6], [14], [15], [16], [17], [21], [26] among many others), and are subject to the curse-of-dimensionality (whereby the computational cost growth is very roughly on the order of $(2D)^n$ where D is the required number of grid points per dimension, and more importantly, n is the space dimension.

A recent development is the discovery of the curse-of-dimensionality-free methods exploiting semiconvex dual operators and max-plus linearity [31], [32], [34]. This approach has, so far, only been developed for problems over the entire space. (For other max-plus-based methods developed for larger classes of problems, see [1], [2], [20], [34], [36], [37].) In particular, it deals with HJB PDEs of the form

$$0 = -\widetilde{H}(x, \operatorname{grad} V) \ \ \forall x \in I\!\!R^n \setminus \{0\}, \quad V(0) = 0 \tag{1.1}$$

where

$$\widetilde{H}(x, \operatorname{grad} V) = \max_{m \in \mathcal{M}} \{H^m(x, \operatorname{grad} V)\} \tag{1.2}$$

where $\mathcal{M} = \{1, 2 \dots M\}$, and the H^m have computationally simpler forms. In particular, this has been developed for long-run average-cost-per-unit-time problems where the H^m are quadratic forms.

*Research partially supported by NSF grant DMS-0307229 and AFOSR grant FA9550-06-1-0238.
†Dept. of Mechanical and Aerospace Eng., University of California San Diego, (wmceneaney@ucsd.edu)

In [32], [34], the method was developed and the curse-of-dimensionality-free nature was made clear. The H^m had the form

$$H^m(x,p) = (A^m x)'p + \tfrac{1}{2}x'C^m x + \tfrac{1}{2}p'\Sigma^m p, \tag{1.3}$$

where the A^m, C^m and Σ^m were $n \times n$ matrices meeting certain conditions which guaranteed existence and uniqueness of a solution within a certain class of functions. We refer to the H^m as the constituent Hamiltonians. In [31], the convergence rate for the algorithm was obtained. In particular, it was shown that there were two parameters, τ and $T = N\tau$ such that the errors go to zero as $T = N\tau \to \infty$ and $\tau \downarrow 0$. Further, a relation between the relative T and τ rates was indicated. The errors in the solution approximation are bounded in a form $0 \le \widetilde{V} - V^a \le \varepsilon(1+|x|^2)$ where $\widetilde{V}$ is the true solution and V^a is the computed approximation. Additionally, we had $T = N\tau \propto \varepsilon^{-1}$ and $\tau \propto \varepsilon^2$, and so $N \propto \varepsilon^{-3}$. The computational cost growth with n is only on the order of n^3 (due to some matrix inverses). However, the approach is subject to a curse-of-complexity, where the computational cost can grow like M^N. Various strategies are outlined in [31] for greatly attenuating this growth. In [30], a convex programming approach to attenuating the growth was employed, and a problem over $I\!R^6$ with $M = 6$ was solved in under an hour on a typical desktop machine (with the more general H^m of (1.4) below). Nevertheless, this curse-of-complexity attenuation remains an active research area.

It quickly became clear that the problem class given by (1.1), (1.2), (1.3) is quite restrictive. In particular, the solutions are quadratic along lines through the origin. On the other hand, it is well-known that any semiconvex function can be represented as a supremum of quadratic forms, and consequently, approximated (although we do not use a specific metric here) by a maximum of quadratic forms. This obviously holds true for Hamiltonians as well. However, the forms (1.3) are insufficient; we must expand to H^m of the form

$$H^m(x,p) = \tfrac{1}{2}x'D^m x + \tfrac{1}{2}p'\Sigma^m p + (A^m x)'p + (l_1^m)'x + (l_2^m)'p + \alpha^m, \tag{1.4}$$

where $l_1^m, l_2^m \in I\!R^n$ and $\alpha^m \in I\!R$. Although our interest here is on (numerical) solution of (1.1) with a Hamiltonian given by (1.2), (1.4) for its own sake, the additional motivation that these will approximate the very general class of HJB PDEs with semiconvex Hamiltonians, and so will eventually lead to approximate solution of such, provides further motivation. (Initial results on the Hamiltonian approximation question appear in [29].)

2. Problem class. In the general case, the consituent Hamiltonians, the H^m, still have corresponding HJB PDE problems which take the form

$$0 = -H^m(x, \operatorname{grad} V) \;\; \forall x \in I\!R^n \setminus \{0\}, \quad V(0) = 0. \tag{2.1}$$

These constituent Hamiltonians are associated, at least formally, with the following control problems. The dynamics of the problems are affine, and are given by

$$\dot{\xi}^m = A^m \xi^m + l_2^m + \sigma^m w, \quad \xi_0^m = x \in I\!R^n \tag{2.2}$$

where the nature of the (newly introduced) σ^m is specified just below. Let $w \in \mathcal{W} \doteq L_2^{loc}([0,\infty); I\!R^k)$, and we recall that $L_2^{loc}([0,\infty); I\!R^k) = \{w : [0,\infty) \to I\!R^k \;:\; \int_0^T |w_t|^2\,dt < \infty \;\forall T < \infty\}$. The cost functionals are

$$\begin{aligned} J^m(x,T;w) &\doteq \int_0^T L^m(\xi_t^m) - \frac{\gamma^2}{2}|w_t|^2\,dt, \\ &\doteq \int_0^T \tfrac{1}{2}(\xi_t^m)'D^m\xi_t^m + (l_1^m)'\xi_t^m + \alpha^m - \frac{\gamma^2}{2}|w_t|^2\,dt \end{aligned} \tag{2.3}$$

The associated value functions (which might not be finite) would be

$$V^m(x) = \sup_{w\in\mathcal{W}} \sup_{T<\infty} J^m(x,T;w). \tag{2.4}$$

Lastly, σ^m and γ are such that $\Sigma^m = \frac{1}{\gamma^2}\sigma^m(\sigma^m)'$.

We remark that a generalization of the second term in the integrand of the cost functional (in both cases) to $\frac{1}{2}w'C^m w$ with C^m symmetric, positive definite is not needed since this is equivalent to a change in σ^m in the dynamics (2.2).

Prior to introduction of the assumptions, some discussion is in order. In the purely quadratic case (with constituent Hamiltonians given by (1.3)) considered in [31], [32], [34], each of the constituent Hamiltonians is associated with a linear/quadratic problem with finite value, and solvable via a Riccati equation. However, in this more-general quadratic case, some of the constituent Hamiltonians might be associated with problems which might not be reasonable. For example, the nominal stable point of the dynamics might correspond to a state with negative running cost. Nonetheless, they may make a significant contribution to the originating Hamiltonian, $\widetilde{H}$. To motivate the assumptions for this rather general problem class, we return to the concept of $\widetilde{H}$ being constructed so as to resemble some given nonlinear control problem which has a (finite) solution. That is, we think of $\widetilde{H}$ as being chosen to resemble some given $\widetilde{\widetilde{H}}$. We suppose that problem

$$0 = -\widetilde{\widetilde{H}}(x, \operatorname{grad} V), \qquad V(0) = 0 \tag{2.5}$$

has finite value, and that we are choosing $\widetilde{H}$ to approximate $\widetilde{\widetilde{H}}$.

We make the following block of assumptions which we will refer to as Assumption Block $(A.m)$ throughout.

$(A.m)$

Assume there exists unique viscosity solution, $\widetilde{\widetilde{V}}$, to (2.5) in Q_K for some $K \in (0,\infty)$, where $Q_K = \{\phi : \mathbb{R}^n \to \mathbb{R} \,|\, \phi \text{ is semiconvex, and } 0 \le \Phi(x) \le (K/2)|x|^2 \ \forall x \in \mathbb{R}^n\}$.
Assume that $\widetilde{H}(x,p) = \max_{m\in\mathcal{M}} H^m(x,p) \le \widetilde{\widetilde{H}}(x,p)$ for all $x, p \in \mathbb{R}^n$.
Assume there exists $c_A \in (0,\infty)$ such that $x'A^m x \le -c_A|x|^2$ for all $x \in \mathbb{R}^n$ and all $m \in \mathcal{M}$.
Assume $H^1(x,p)$ has coefficients satisfying the following: $l_1^1 = l_2^1 = 0$; $\alpha^1 = 0$; D^1 is positive definite, symmetric; and $\gamma^2/c_\sigma^2 > c_D/c_A^2$, where c_D is such that $x'D^1x \le c_D|x|^2$ $\forall x \in \mathbb{R}^n$ and $c_\sigma \doteq |\sigma^1|$.
Assume that system (2.2) is controllable in the sense that given $x, y \in \mathbb{R}^n$ and $T > 0$, there exist processes $w \in \mathcal{W}$ and μ measurable with range in $\mathcal{M}$, such that $\xi_T = y$ when $\xi_0 = x$ and one applies controls w, μ.

Note that the last assumption, the controllability assumption, is satisfied if there exists at least one $m \in \mathcal{M}$ such that $\sigma^m(\sigma^m)'$ (which is $n \times n$) has n positive eigenvalues. We will be making an additional assumption, and that is:

$(A.w)$

Assume there exist $c_1, c_2 < \infty$ such that for any ε–optimal pair, $\mu^\varepsilon, w^\varepsilon$ for the $\widetilde{H}$ problem, one has

$$\|w^\varepsilon\|^2_{L_2[0,T]} \le c_1 + c_2|x|^2$$

for all $\varepsilon \in (0,1]$, all $T < \infty$ and all $x \in \mathbb{R}^n$.

Note that the behavior specified in $(A.w)$ is proved in the purely quadratic case (c.f., [34], [40]) under reasonable assumptions on the constituent-Hamiltonian matrices, but in this more general case, we assume it instead.

There are two more assumptions that were used in the error estimates. The first of these assumptions is that

$(A.\sigma)$

$$\sigma^m = \sigma \quad \forall\, m \in \mathcal{M}.$$

We do not use this assumption anywhere other than in the technical estimates of Section 6 (and, indirectly, in the Combined Errors section, i.e. Section 8, that follows). The author was not able to show that the assumption was necessary, but was unable to obtain the estimates of that section without it. Considering this, we leave the m-dependence on σ in the other sections.

In the introduction, we note that one can approximate any semiconvex Hamiltonian arbitrarily well by a pointwise maximum of quadratic "constituent" Hamiltonians. The restriction to m-independent σ does not prevent approximation of any semiconvex Hamiltonian with bounded semiconvexity constant, as the quadratic growth terms in the approximating functions may always be taken to be constant (c.f. [34]). However, although this is not a theoretical limitation, it does pose a practical limitation, and so one would prefer not to need this seemingly technical assumption. It should be noted that we have not found this assumption to be necessary in our, admittedly small, set of tests; one should note that in the six-dimensional example in [30], σ did depend on m.

There is one final technical assumption, this one is on the behavior of ε–optimal trajectories of the system, in the general class of coercivity-type assumptions. Since this assumption requires objects which have not been defined yet, and is used only in Section 7, we delay presentation of the assumption to the top of that section. There are no more assumptions.

In order to show that the conditions are not vacuous (in fact, they seem to be rather general relative to problems that might have finite value), we include a simple, one-dimensional example satisfying all the assumptions. Let $M = 2$ and

$$H^1(x,p) = \frac{x^2}{2} + \frac{p^2}{8} - xp,$$
$$H^2(x,p) = x^2 + \frac{p^2}{8} - xp - 1.$$

Given that $\Sigma^m = 1/4$ in both cases, we may take $\sigma = 1$ and $\gamma^2 = 4$ (so that $\sigma^2/(2\gamma^2) = 1/8$). It is obvious that Assumptions $(A.m)$ and $(A.\sigma)$ are satisfied. We demonstrate that the other two are met. Note that

$$\int_0^T L^{\mu_t^\varepsilon}(\xi_t^\varepsilon) - \frac{\gamma^2}{2}|w_t^\varepsilon|^2\,dt \le \int_0^T (\xi_t^\varepsilon)^2\,dt - 2\|w^\varepsilon\|^2_{L_2(0,T)},$$

which by (11.1) (in the proof of Lemma 4.3 in Appendix A),

$$\le \frac{x^2}{c_A} + \frac{c_\sigma^2}{c_A^2}\|w^\varepsilon\|^2_{L_2(0,T)} - 2\|w^\varepsilon\|^2_{L_2(0,T)} = x^2 - \|w^\varepsilon\|^2_{L_2(0,T)},$$

for all $T > 0$ and $x \in I\!R^n$. Comparing with $w \equiv 0$, where the payoff is then nonnegative, we see that if w^ε is ε–optimal with $\varepsilon \in (0,1]$, one must have $\|w^\varepsilon\|^2_{L_2(0,T)} \le 1 + x^2$. This yields $(A.w)$. Lastly, by this bound on w^ε and Lemma 4.3, noting that $l_2^1 = l_2^2 = 0$, one sees that $\int_0^T |\xi_t^\varepsilon|^2\,dt \le 2 + 3x^2$ for all T, x. This implies that for large time, the trajectory stays near the origin (in an L_2-sense), where H^1 dominates. This yields $(A.\xi)$ (of Section 7) with $c_3 = 1/4$, and we do not include the details.

3. Overview of required development. The curse-of-dimensionality-free approach is as follows. The solution to (1.1), (1.2), (1.4) is given by

$$\lim_{T\to\infty} \widetilde{S}_T[V_0]$$

where $\widetilde{S}_T$ is the associated semigroup, and V_0 is some initialization (where $V_0 \equiv 0$ is an acceptable initialization). Fixing some time-discretization, τ, and considering $T = N\tau$ (for T such that N is an integer), one may approximate $\widetilde{S}_T$ by $\left[\bar{S}_\tau\right]^N$ where the exponent indicates repeated composition and $\bar{S}_\tau[\phi](x) \doteq \max_{m\in\mathcal{M}} S_\tau^m[\phi](x)$ where S_τ^m is the semigroup associated with constituent HJB PDE $0 = H^m(x, \operatorname{grad} V)$.

As noted above, the errors are due to the finiteness of $T = N\tau$ and the time-discretization, τ. Now, the key to the curse-of-dimensionality-free aspect of the computations is in the means of computing the $\bar{S}_\tau$ operator actions. The computational speed is achieved through the use of max-plus linearity and semiconvex duality. We will not include those details which have been discussed elsewhere, but refer the reader to [31], [32], [34] for such.

Instead, we briefly outline the concept. Suppose V_0 is quadratic (or a max-plus sum of quadratics). It has a semiconvex dual (defined in Section 9), a_0, which is also a quadratic. In the case of quadratic H^m, the semiconvex dual operators of the S_τ^m take the form of max-plus integral operators

$$\widehat{\mathcal{B}}_\tau^m[a](x) = \mathcal{B}_\tau^m(x,\cdot) \odot a(\cdot) = \int_{I\!R^n}^{\oplus} \mathcal{B}_\tau^m(x,y) \otimes a(y)\, dy = \max_{y \in I\!R^n} \{\mathcal{B}_\tau^m(x,y) + a(y)\}$$

where the $\mathcal{B}_\tau^m$ are quadratic forms. Further, $\overline{\mathcal{B}}_\tau$, the kernel of the semiconvex dual of $\bar{S}_\tau$, is the max-plus sum of the $\mathcal{B}_\tau^m$, i.e.,

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

The dual of $V_1 = \bar{S}_\tau[V_0]$ is

$$\overline{a}_1(y) = \bigoplus_{m_1 \in \mathcal{M}} \mathcal{B}_\tau^{m_1}(y,\cdot) \odot a_0(\cdot) \doteq \bigoplus_{m_1 \in \mathcal{M}} \widehat{a}_1^{m_1}(y)$$

where each $\widehat{a}_1^{m_1}$ is a quadratic function (computed by analytical operations modulo a matrix inverse). The dual of $V_2 = (\bar{S}_\tau)^2[V_0]$, $\overline{a}_2$, has the form

$$\overline{a}_2(y) = \bigoplus_{m_2 \in \mathcal{M}} \bigoplus_{m_1 \in \mathcal{M}} \widehat{a}_2^{m_1,m_2}(y)$$

where

$$\widehat{a}_2^{m_1,m_2}(y) = \mathcal{B}_\tau^{m_2}(y,\cdot) \odot \widehat{a}_1^{m_1}(\cdot),$$

and we see that each $\widehat{a}_2^{m_1,m_2}$ is also a quadratic; and similarly obtained. The dual of $V_N = (\bar{S}_\tau)^N[V_0]$, $\overline{a}_N$, is obtained similarly. Details on the algorithm implementation and on methods for attenuating the curse-of-complexity appear in [30], [31]. We will briefly describe the minor modifications in the algorithm implementation due to the presence of the $l_1^m, l_2^m = 0, \alpha^m$ in Section 10.

There are four topics that must be addressed. First, we must obtain conditions under which we can guarantee the solutions we seek to compute actually exist, and this is done in Section 4. The second, and main, topic is the proof that $\widetilde{V} - (\bar{S}_\tau)^N[V_0]$ goes to zero as $N \to \infty$ and $\tau \downarrow 0$. While doing that, we will also obtain specific error bounds as functions of these parameters. In particular, under certain assumptions, we show that given $\varepsilon > 0$, there exist N, τ such that the error is less than $\varepsilon(1 + |x|^2)$ over $I\!R^n$. This appears in Sections 5–8. Next, in Section 9, we briefly demonstrate that the curse-of-dimensionality-free method outlined above may be used to compute $(\bar{S}_\tau)^N[V_0]$. Lastly, as noted just above, we indicate the minor changes in the algorithm (the basic algorithm being discussed more fully in [32], [34]) necessary for the more general H^m considered here, and this appears in Section 10.

4. Existence and semigroup limits. In this general-quadratic constituent Hamiltonian problem class, there are several technical issues, including specifically the question of existence of the solution. Recall that the HJB PDE problem of interest is

$$\begin{aligned} &0 = -\widetilde{H}(x, \operatorname{grad} V) \doteq -\max_{m \in \mathcal{M}} H^m(x, \operatorname{grad} V) \qquad x \in I\!R^n \setminus \{0\}, \\ &V(0) = 0. \end{aligned} \tag{4.1}$$

The corresponding value function is

$$\widetilde{V}(x) = \sup_{w\in\mathcal{W}} \sup_{\mu\in\mathcal{D}_\infty} \widetilde{J}(x,w,\mu) \doteq \sup_{w\in\mathcal{W}} \sup_{\mu\in\mathcal{D}_\infty} \sup_{T<\infty} \int_0^T L^{\mu_t}(\xi_t) - \frac{\gamma^2}{2}|w_t|^2\,dt \tag{4.2}$$

where

$$L^{\mu_t}(x) = \tfrac{1}{2}x'D^{\mu_t}x + (l_1^{\mu_t})'x + \alpha^{\mu_t},$$
$$\mathcal{D}_\infty = \{\mu : [0,\infty) \to \mathcal{M} \ : \ \text{measurable}\},$$

and ξ satisfies

$$\dot{\xi} = A^{\mu_t}\xi + l_2^{\mu_t} + \sigma^{\mu_t}w_t, \quad \xi_0 = x. \tag{4.3}$$

Let's first prove the finiteness of $\widetilde{V}(x)$.

THEOREM 4.1. *For all $x \in I\!R^n$, $0 \le \widetilde{V}(x) \le \widetilde{\widetilde{V}}(x)$. Consequently, $\widetilde{V}(x)$ lies in Q_K.*

Proof. If one had $\widetilde{\widetilde{V}}(\xi_t)$ absolutely continuous with respect to time, then the following proof would be greatly reduced. The central theme of the proof, neglecting this technical issue, is relatively straightforward once one uses the integral of the Hamiltonian as a tool. To handle the potential nonsmoothness of $\widetilde{\widetilde{V}}(\xi_t^\varepsilon)$, we consider a mollified seqeunce of functions, $V^\delta \in C^\infty$, approximating $\widetilde{\widetilde{V}}$. For $\delta > 0$, let $g^\delta : I\!R^n \to [0,\infty)$ be a one-parameter family of mollifiers, i.e., smooth functions such that $g^\delta(x) = 0$ for all $x \notin B_\delta(0)$ and $\int_{I\!R^n} g^\delta(x)\,dx = 1$ for all $\delta > 0$. In particular, we take $g^\delta(x) = \delta^{-n}\bar{g}(x/\delta)$ for some appropriate $\bar{g}$. Let

$$V^\delta(x) \doteq [g^\delta * \widetilde{\widetilde{V}}](x) = \int_{I\!R^n} g^\delta(x-y)\widetilde{\widetilde{V}}(y)\,dy.$$

We introduce the finite time-horizon value function

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

with ξ satisfying (4.3). Fix $w^\varepsilon \in \mathcal{W}, \mu^\varepsilon \in \mathcal{D}_T$ which are ε-optimal for $\widetilde{W}(x,T)$, and let ξ^ε be the corresponding trajectory with initial state ${\xi^\varepsilon}_0 = x$. We have

$$\begin{aligned}
\widetilde{W}(x,T) &\le \int_0^T L^{\mu_t^\varepsilon}(\xi_t^\varepsilon) - \frac{\gamma^2}{2}|w_t^\varepsilon|^2\,dt + \varepsilon \\
&= \int_0^T L^{\mu_t^\varepsilon}(\xi_t^\varepsilon) - \frac{\gamma^2}{2}|w_t^\varepsilon|^2 + [A^{\mu_t^\varepsilon}\xi_t^\varepsilon + l_2^{\mu_t^\varepsilon} + \sigma^{\mu_t^\varepsilon}w_t^\varepsilon]\cdot \operatorname{grad} V^\delta(\xi_t^\varepsilon) \\
&\qquad -[A^{\mu_t^\varepsilon}\xi_t^\varepsilon + l_2^{\mu_t^\varepsilon} + \sigma^{\mu_t^\varepsilon}w_t^\varepsilon]\cdot \operatorname{grad} V^\delta(\xi_t^\varepsilon)\,dt + \varepsilon,
\end{aligned}$$

which by the definition of $\widetilde{H}$

$$\le \int_0^T \widetilde{H}(\xi_t^\varepsilon, \operatorname{grad} V^\delta(\xi_t^\varepsilon))\,dt - \int_0^T [A^{\mu_t^\varepsilon}\xi_t^\varepsilon + l_2^{\mu_t^\varepsilon} + \sigma^{\mu_t^\varepsilon}w_t^\varepsilon]\cdot \operatorname{grad} V^\delta(\xi_t^\varepsilon)\,dt + \varepsilon$$

where, for clarity, we note that $\operatorname{grad} V^\delta$ indicates the gradient of V^δ rather than the δ-mollification of some $\operatorname{grad} V$, and by $(A.m)$

$$\begin{aligned}
&\le \int_0^T \widetilde{\widetilde{H}}(\xi_t^\varepsilon, \operatorname{grad} V^\delta(\xi_t^\varepsilon))\,dt - \int_0^T \dot{\xi}_t^\varepsilon \cdot \operatorname{grad} V^\delta(\xi_t^\varepsilon)\,dt + \varepsilon \\
&= \int_0^T \widetilde{\widetilde{H}}(\xi_t^\varepsilon, \operatorname{grad} V^\delta(\xi_t^\varepsilon))\,dt + V^\delta(x) - V^\delta(\xi_T^\varepsilon) + \varepsilon. \qquad (4.4)
\end{aligned}$$

(Note that if V^δ could be replaced by $\widetilde{\widetilde{V}}$, then the right-hand side would be less than $\widetilde{\widetilde{V}}(x)+\varepsilon$ by noting $\widetilde{\widetilde{V}}(\xi_T^\varepsilon)\geq 0$, and this would lead us to our goal below. Additional effort is required due to the fact that we needed to work with V^δ instead.)

Recalling that $\widetilde{\widetilde{V}}$ is assumed semiconvex, and consequently locally Lipschitz (c.f., [19]), it is easy to see that given $\bar{R}<\infty$ and $\varepsilon\in(0,1]$, there exists $\bar{\delta}>0$ such that

$$|V^\delta(y)-\widetilde{\widetilde{V}}(y)|\leq\varepsilon,\quad\forall\, y\in B_{\bar{R}}. \tag{4.5}$$

We now interject two lemmas. From [32], [34], we see that $V^1\in Q_K$ for $K=K_\delta\doteq(c_A\gamma^2/c_\sigma^2)-\delta$ for $\delta>0$ sufficiently small. Using this and Assumption $(A.w)$, one obtains the following lemmas essentially exactly as in [40] (see also [34]). For completeness, we include sketches of the proofs in Appendix A.

LEMMA 4.2. *For any $t<\infty$,*

$$|\xi_t^\varepsilon|^2\leq e^{-c_At}\,|x|^2+\frac{2}{c_A^2}\sup_m|l_2^m|^2+\frac{2c_\sigma^2}{c_A}\|w^\varepsilon\|^2_{L_2(0,t)}.$$

LEMMA 4.3. *For any $t<\infty$,*

$$\int_0^t|\xi_r^\varepsilon|^2\,dr\leq\frac{1}{c_A}|x|^2+\frac{2}{c_A^2}\sup_m|l_2^m|^2\,t+\frac{2c_\sigma^2}{c_A^2}\|w^\varepsilon\|^2_{L_2(0,t)}.$$

From Lemma 4.2, we see that there exists $\bar{R}<\infty$ such that $\xi_T^\varepsilon\leq\bar{R}$ for all $x\in\bar{B}_R$, $\varepsilon\in(0,1]$ and $T<\infty$. Combining this with (4.5), we see that

$$|V^\delta(\xi_T^\varepsilon)-\widetilde{\widetilde{V}}(\xi_T^\varepsilon)|\leq\varepsilon\quad\forall\, x\in\bar{B}_R,\varepsilon\in(0,1]\text{ and }T<\infty. \tag{4.6}$$

Combining (4.4), (4.5) and (4.6), one finds that for $\delta\leq\bar{\delta}$

$$\widetilde{W}(x,T)\leq\int_0^T\widetilde{\widetilde{H}}(\xi_t^\varepsilon,\operatorname{grad}V^\delta(\xi_t^\varepsilon))\,dt+\widetilde{\widetilde{V}}(x)-\widetilde{\widetilde{V}}(\xi_T^\varepsilon)+3\varepsilon$$

for all $x\in\bar{B}_R$, $\varepsilon\in(0,1]$ and $T<\infty$. Since $\widetilde{\widetilde{V}}\geq 0$, this yields

$$\widetilde{W}(x,T)\leq\int_0^T\widetilde{\widetilde{H}}(\xi_t^\varepsilon,\operatorname{grad}V^\delta(\xi_t^\varepsilon))\,dt+\widetilde{\widetilde{V}}(x)+3\varepsilon. \tag{4.7}$$

We now interject one more lemma. The proof of this lemma is given in Appendix B.

LEMMA 4.4. *Given $\bar{R}<\infty$, $\bar{\varepsilon}\in(0,1]$, sequence $\delta_k\downarrow 0$ and finite set $\{y^\lambda\}_{\lambda=1}^\Lambda\subset\bar{B}_{\bar{R}}$, there exists a subsequence, $\{\delta_{k_\kappa}\}$ such that for all κ sufficiently large*

$$\inf_{z\in\partial\widetilde{\widetilde{V}}(y^\lambda)}\left|\operatorname{grad}V^{\delta_{k_\kappa}}(y^\lambda)-z\right|<\bar{\varepsilon}\quad\forall\,\lambda\in\{1,2,\ldots\Lambda\}.$$

Recall that, since $\widetilde{\widetilde{V}}$ is a viscosity solution of $0=-\widetilde{\widetilde{H}}(x,\operatorname{grad}V)$ with boundary condition $V(0)=0$,

$$\widetilde{\widetilde{H}}(x,z)\leq 0\quad\forall\, z\in\partial\widetilde{\widetilde{V}}(x),x\neq 0. \tag{4.8}$$

Note that, by the semiconvexity of $\tilde{\tilde{V}}$ (which implies local Lipschitz behavior, c.f., [19]), there exists $\bar{R} < \infty$ such that $|\operatorname{grad} V^\delta(y)| < \bar{R}$ for all $y \in \bar{B}_R$ and all $\delta \in (0,1)$. By the continuity of $\tilde{\tilde{H}}$, given $R, \bar{R} < \infty$, there exists $\varepsilon_2 > 0$ such that

$$\left|\tilde{\tilde{H}}(y_1, z_1) - \tilde{\tilde{H}}(y_2, z_2)\right| < \frac{\varepsilon}{2T} \tag{4.9}$$

for all $y_1, y_2 \in \bar{B}_R$ and $z_1, z_2 \in \bar{B}_{\bar{R}}$ such that $|y_1 - y_2| < \varepsilon_2$ and $|z_1 - z_2| < \varepsilon_2$. Also, by the compactness of $\bar{B}_R$ and $\bar{B}_{\bar{R}}$, there exists finite set $\overline{Y} = \{y^\lambda\}_{\lambda=1}^{\Lambda} \subset \bar{B}_R$ and sequence $\delta_k \downarrow 0$ such that

$$\min_{\lambda \in \{1,2,\ldots\Lambda\}} \left[|y - y^\lambda| + \left|\operatorname{grad} V^{\delta_k}(y) - \operatorname{grad} V^{\delta_k}(y^\lambda)\right|\right] < \varepsilon_2 \quad \forall\, y \in \bar{B}_R,\ \forall k \in \mathbf{N}. \tag{4.10}$$

By (4.9) and (4.10), we may choose $\overline{Y}$ such that given any $y \in \bar{B}_R$, there exists $y^\lambda \in \overline{Y}$ such that

$$\left|\tilde{\tilde{H}}(y, \operatorname{grad} V^{\delta_k}(y)) - \tilde{\tilde{H}}(y^\lambda, \operatorname{grad} V^{\delta_k}(y^\lambda))\right| < \frac{\varepsilon}{2T} \quad \forall\, k \in \mathbf{N}. \tag{4.11}$$

Now choose subsequence δ_{k_κ} such that Lemma 4.4 holds with $\bar{\varepsilon} = \varepsilon_2$ and the finite set $\overline{Y}$. Then, for any $y^\lambda \in \overline{Y}$, there exists $z^\lambda \in \partial\tilde{\tilde{V}}(y^\lambda)$ such that

$$\left|\operatorname{grad} V^{\delta_{k_\kappa}}(y^\lambda) - z^\lambda\right| < \varepsilon_2$$

for any κ, which, by (4.9), implies

$$\left|\tilde{\tilde{H}}(y^\lambda, \operatorname{grad} V^{\delta_{k_\kappa}}(y^\lambda)) - \tilde{\tilde{H}}(y^\lambda, z^\lambda)\right| < \frac{\varepsilon}{2T} \quad \forall\, \kappa \in \mathbf{N}. \tag{4.12}$$

Combining (4.11) and (4.12), one sees that given $y \in \bar{B}_R$, there exists $\bar{\delta} = \delta_{k_\kappa}$ (actually any element of the subsequence), $y^\lambda \in \overline{Y}$ and $z^\lambda \in \partial\tilde{\tilde{V}}(y^\lambda)$ such that

$$\left|\tilde{\tilde{H}}(y, \operatorname{grad} V^{\bar{\delta}}(y)) - \tilde{\tilde{H}}(y^\lambda, z^\lambda)\right| < \frac{\varepsilon}{T}. \tag{4.13}$$

Then, by (4.8) and (4.13), for all $y \in \bar{B}_R$

$$\tilde{\tilde{H}}(y, \operatorname{grad} V^{\bar{\delta}}(y)) < \frac{\varepsilon}{T}. \tag{4.14}$$

Employing (4.14) in (4.7), one finds that

$$\widetilde{W}(x, T) \le \tilde{\tilde{V}}(x) + 4\varepsilon.$$

Since this is true for all $T \in (0, \infty)$ and all $\varepsilon \in (0, 1]$,

$$\widetilde{V}(x) = \sup_{T<\infty} \widetilde{W}(x, T) \le \tilde{\tilde{V}}(x). \tag{4.15}$$

On the other hand,

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

and by taking $\mu \equiv 1$ and $w \equiv 0$

$$\geq \int_0^T L^1(\xi_t)\,dt \geq 0.$$

Consequently,

$$\widetilde{V}(x) = \sup_{T<\infty} \widetilde{W}(x,T) \geq 0.$$

□

In analogy with the results of [31], we would like to know whether $\widetilde{S}_T[V](\cdot)$ converges to $\widetilde{V}(\cdot)$ as T goes to infinity. Unfortunately we haven't been able to establish this for arbitrary V in Q_K but only for the V under $\widetilde{V}$. We will however see that this is sufficient to prove the convergence of the algorithm.

THEOREM 4.5. *$\widetilde{V}$ is the unique continuous solution of $V = \widetilde{S}_T[V]$ in the class Q_K for any $T > 0$. $\widetilde{V}$ is also the unique viscosity solution of (4.1) in Q_K. Further, given any $V \in Q_K$ such that $0 \leq V \leq \widetilde{V}$, $\lim_{T\to\infty} \widetilde{S}_T[V](x) = \widetilde{V}(x)$ (uniformly on compact sets).*

Proof. The first two assertions of the theorem are classical results of the theory of viscosity solutions, and proofs can be found, respectively, in [34] and [40]. We turn to the last assertion.

$$\begin{aligned}\widetilde{V}(x) &= \sup_{w\in\mathcal{W}} \sup_{\mu\in\mathcal{D}_\infty} \sup_{T<\infty} \int_0^T L^{\mu_t}(\xi_t) - \frac{\gamma^2}{2}|w_t|^2\,dt \\ &= \sup_{T<\infty} \widetilde{S}_T[0](x).\end{aligned}$$

However,

$$\begin{aligned}\widetilde{S}_{T+\Delta_T}[0](x) &= \widetilde{S}_T[\widetilde{S}_{\Delta_T}[0](\cdot)](x) \\ &= \sup_{w\in\mathcal{W}} \sup_{\mu\in\mathcal{D}_T} \left\{ \int_0^T L^{\mu_t}(\xi_t) - \frac{\gamma^2}{2}|w_t|^2\,dt \right. \\ &\qquad \left. + \sup_{w^1\in\mathcal{W}} \sup_{\mu^1\in\mathcal{D}_{\Delta_T}} \int_T^{T+\Delta_T} L^{\mu^1_t}(\xi^1_t) - \frac{\gamma^2}{2}|w^1_t|^2\,dt \right\}\end{aligned}$$

where ξ^1 is driven by $w^1_{\cdot-T}, \mu^1_{\cdot-T}$ with initial condition $\xi^1_T = \xi_T$. By taking $w^1 \equiv 0$ and $\mu^1 \equiv 1$ on $[0, \Delta_T]$, we see

$$\begin{aligned}&\geq \sup_{w\in\mathcal{W}} \sup_{\mu\in\mathcal{D}_T} \int_0^T L^{\mu_t}(\xi_t) - \frac{\gamma^2}{2}|w_t|^2\,dt \\ &= \widetilde{S}_T[0](x),\end{aligned}$$

i.e., $\widetilde{S}_T[0](x)$ is a monotonically increasing function of T. Consequently, we have,

$$\lim_{T\to\infty} \widetilde{S}_T[0](x) = \widetilde{V}(x).$$

Further, for all V such that $0 \leq V \leq \widetilde{V}$ we find,

$$\widetilde{V}(x) = \lim_{T\to\infty} \widetilde{S}_T[0](x) \leq \lim_{T\to\infty} \widetilde{S}_T[V](x) \leq \lim_{T\to\infty} \widetilde{S}_T[\widetilde{V}](x) = \widetilde{V}(x),$$

which immediately yields the assertion. □

Now we would like to compute $\widetilde{V}(x)$ with a discrete-time approximation. The idea is of course to apply an approximation of $\widetilde{S}_T$ to a function $V \leq \widetilde{V}$ and to increase time towards infinity. We introduce the approximating, discrete-time operator

$$\begin{aligned}\bar{S}_\tau[\phi](x) &= \sup_{w\in\mathcal{W}} \max_{m\in\mathcal{M}} \left[\int_0^\tau L^m(\xi^m_t) - \frac{\gamma^2}{2}|w_t|^2\,dt + \phi(\xi^m_\tau) \right](x) \\ &= \max_{m\in\mathcal{M}} S^m_\tau[\phi](x)\end{aligned} \tag{4.16}$$

where ξ^m satisfies (2.2).

With τ acting as a time-discretization step-size, let

$$\mathcal{D}_\infty^\tau = \Big\{ \mu : [0,\infty) \to \mathcal{M} \mid \text{for each } n \in \mathbf{N} \cup \{0\}, \text{ there exists } m_n \in \mathcal{M} \\ \text{such that } \mu(t) = m_n \ \forall\, t \in [n\tau, (n+1)\tau) \Big\}, \tag{4.17}$$

and for $T = \bar{n}\tau$ with $\bar{n} \in \mathbf{N}$ define $\mathcal{D}_T^\tau$ 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

$$\bar{\bar{S}}_T^\tau[\phi](x) = \max_{\{m_k\}_{k=0}^{\bar{n}-1} \in \mathcal{M}^{\bar{n}}} \left\{ \prod_{k=0}^{\bar{n}-1} S_\tau^{m_k} \right\} [\phi](x) = (\bar{S}_\tau)^{\bar{n}}[\phi](0) \tag{4.18}$$

where the $\prod$ notation indicates operator composition, and the superscript in the last expression indicates repeated application of $\bar{S}_\tau$, $\bar{n}$ times. We will be approximating $\widetilde{V} = \lim_{T\to\infty} \widetilde{S}_T[\phi]$ by $\widetilde{S}_T[\phi]$ for some large T. We will approximate $\widetilde{S}_T[\phi]$ by $\bar{\bar{S}}_T^\tau[\phi]$ for some small τ.

Let $\overline{V} \doteq \lim_{N\to\infty} \bar{\bar{S}}_{N\tau}^\tau[0](x)$, where 0 denotes the zero-function.

THEOREM 4.6. *$\overline{V}$ satisfies*

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

Further, $0 \le V^1 \le \overline{V} \le \widetilde{V}$, and consequently, $\overline{V} \in Q_K$.

Neglecting trivial changes, the proof is identical to that provided in [32]. Now we need to know that by applying $\bar{\bar{S}}_{N\tau}^\tau$ to an initial V with N going to infinity, we converge to the solution $\overline{V}$.

THEOREM 4.7. *Given any $V \in Q_K$ such that $V \le \overline{V}$ we have, $\lim_{N\to\infty} \bar{\bar{S}}_{N\tau}^\tau[V](x) = \overline{V}(x)$ for all $x \in I\!R^n$ (uniformly on compact sets).*

Proof. By definition,

$$\overline{V}(x) = \lim_{N\to\infty} \bar{\bar{S}}_{N\tau}^\tau[0](x)$$

which by the monotonicity of $\bar{\bar{S}}_{N\tau}^\tau[\cdot]$ and our assumption,

$$\le \lim_{N\to\infty} \bar{\bar{S}}_{N\tau}^\tau[V](x) \le \lim_{N\to\infty} \bar{\bar{S}}_{N\tau}^\tau[\overline{V}](x)$$

which by Theorem 4.6,

$$= \overline{V}(x).$$

□

5. Error bounds and convergence. As indicated at the top of Section 3, there are two error sources with this curse-of-dimensionality-free approach to solution of the HJB PDE. For concreteness, we suppose that the algorithm is initialized with $\overline{V}^0 = V^1$ (where V^1 is given by (2.4) with $m = 1$). Letting $T \doteq N\tau$ where N is the number of iterations, the first error source is $\widetilde{S}_T\left[\overline{V}^0\right](x) - \bar{\bar{S}}_T^\tau\left[\overline{V}^0\right](x)$. This is the error due to the time-discretization of the μ process over finite time-horizon T. The second error source is $\widetilde{V}(x) - \widetilde{S}_T\left[\overline{V}^0\right](x)$. This is the error due to approximating the infinite time-horizon problem by the finite-time horizon problem with horizon T. The total error is obviously the sum of these. We begin the error analysis with the former of the two error sources in Section 6. Then, in Section 7, we consider the latter, and in Section 8, the two are combined.

6. Errors from time-discretization. We now obtain a bound on the difference between the action of $\bar{\bar{S}}^{\tau}_T$ and that of $\widetilde{S}_T$ depending on the time-dscretization step-size, τ. Assumption $(A.\sigma)$ which was given in Section 2 will be used in obtaining this bound. Again, the author was not able to determine whether the assumption is necessary, but needed it for technical reasons in the long proof to follow.

THEOREM 6.1. *There exists $\bar{K}_{13} < \infty$ such that for sufficiently small $\tau > 0$,*

$$0 \le \widetilde{S}_T[V^1](x) - \bar{\bar{S}}^{\tau}_T[V^1](x) \le \bar{K}_{13}(M+1)^4(1+|x|^2)(1+T)\tau$$

for all $x \in I\!R^n$ and $T \in (0,\infty)$.

Prior to proving this result, we will obtain a number of supporting results. Let μ^ε, w^ε be ε–optimal for $\widetilde{S}_T[V^1](x)$, and again let ξ^ε be the corresponding trajectory. The main issue here is the size of the error induced by the substitution of some $\overline{\mu}^\varepsilon \in \mathcal{D}^\tau_T$ as an approximation of $\mu^\varepsilon \in \mathcal{D}_T$. Let $\overline{\xi}^\varepsilon$ be the trajectory driven by $\overline{\mu}^\varepsilon$ and w^ε (again with initial condition $\overline{\xi}^\varepsilon_0 = x$). First, we indicate how we construct $\overline{\mu}^\varepsilon$ from μ^ε, and obtain a number of results about this approximation.

For any given $\tau > 0$, we build $\overline{\mu}^\varepsilon_r$ from μ^ε_r over $[0,T]$ in the following manner. Fix $\tau > 0$. Let N^t be the largest integer such that $N^t\tau \le t$. For any Lebesgue measurable subset of $I\!R$, $\mathcal{I}$, let $\mathcal{L}(\mathcal{I})$ be the measure of $\mathcal{I}$. For $t \in [0,T]$, $m \in \mathcal{M}$, let

$$\begin{aligned} \mathcal{I}^m_t &= \{r \in [0,t) \,|\, \mu^\varepsilon_r = m\}, & \overline{\mathcal{I}}^m_t &= \{r \in [0,t) \,|\, \overline{\mu}^\varepsilon_r = m\} \\ \lambda^m_t &= \mathcal{L}(\mathcal{I}^m_t), & \bar{\lambda}^m_t &= \mathcal{L}(\overline{\mathcal{I}}^m_t). \end{aligned} \tag{6.1}$$

At the end of any time step, $n\tau$, we pick one of the $m \in \mathcal{M}$ with the largest (positive) error so far committed and we correct it, i.e., let

$$\bar{m} \in \operatorname*{argmax}_{m\in\mathcal{M}} \{\lambda^m_{n\tau} - \bar{\lambda}^m_{(n-1)\tau}\}, \tag{6.2}$$

and we set

$$\overline{\mu}^\varepsilon_t = \bar{m} \quad \forall\, t \in [(n-1)\tau, n\tau). \tag{6.3}$$

Finally, we simply set $\overline{\mu}^\varepsilon_r \in \operatorname{argmax}_{m\in\mathcal{M}}\{\lambda^m_t - \bar{\lambda}^m_{N^t\tau}\}$ for all $r \in [N^t\tau, t]$. Obviously, $\forall n \in \{1,\cdots,N^t\}$ the sum of the errors in measure is null, that is,

$$\sum_{m\in\mathcal{M}} (\lambda^m_{n\tau} - \bar{\lambda}^m_{n\tau}) = \sum_{m\in\mathcal{M}} \lambda^m_{n\tau} - \sum_{m\in\mathcal{M}} \bar{\lambda}^m_{n\tau} = n\tau - n\tau = 0.$$

With this construction, we find the following two results. As the proofs are identical to those in [31], we do not include them.

LEMMA 6.2. *For any $t \in [0,T]$, and any $m \in \mathcal{M}$, one has $\lambda^m_t - \bar{\lambda}^m_t \ge -\tau$.*

LEMMA 6.3. *For any $t \in [0,T]$, and any $m \in \mathcal{M}$, one has $\lambda^m_t - \bar{\lambda}^m_t \le (M-1)\tau$.*

We now develop some more-delicate machinery, which will allow us to make fine estimates of the difference between ξ^ε_t and $\overline{\xi}^\varepsilon_t$. For each m we divide $\mathcal{I}^m_t$ into pieces $\widetilde{\mathcal{I}}^{m,t}_k$ of length τ as follows. Let $\widehat{K}^t_m = \max\{k \in \mathbf{N}\cup\{0\} \,|\, \exists$ integer $n \le t/\tau$ s.t. $\bar{\lambda}^m_{n\tau} = k\tau\}$. Then, for $k \le \widehat{K}^t_m$, let $n^{m,t}_k \doteq \min\{n \in \mathbf{N}\cup\{0\} \,|\, \bar{\lambda}^m_{n\tau} = k\tau\}$ and $\widetilde{\overline{\mathcal{I}}}^{m,t}_k = [(n^{m,t}_k - 1)\tau, n^{m,t}_k\tau]$. Let $K^t_m \doteq]1, \widehat{K}^t_m[$ where for any integers $m \le n$, $]m,n[$ denotes $\{m, m+1, \ldots, n\}$. Loosely speaking, we will now let $\widetilde{\mathcal{I}}^{m,t}_k$ denote the subset of $\mathcal{I}^m_t$ of measure τ, corresponding to $\widetilde{\overline{\mathcal{I}}}^{m,t}_k$. More specifically, we define $\widetilde{\mathcal{I}}^{m,t}_k$ as follows. Introduce the functions $\Phi^{m,t}_k(r)$ which are monotonically increasing (hence measurable) functions (that will match $\widetilde{\overline{\mathcal{I}}}_k^{m,t} = [(n_k^{m,t}-1)\tau; n_k^{m,t}\tau]$ with $\widetilde{\mathcal{I}}_k^{m,t}$) given by

$$\begin{aligned}\Phi_k^{m,t}(r) &= \inf\left\{\rho\in[0,t]\,|\,\lambda_\rho^m = (k-1)\tau + [r-(n_k^{m,t}-1)\tau]\right\}\\ &= \inf\left\{\rho\in[0,t]\,|\,\lambda_\rho^m = r + (k-n_k^{m,t})\tau\right\},\end{aligned}$$

where, in particular, we take $\Phi_k^{m,t}(r)=t$ if there does not exist $\rho\in[0,t]$ such that $\lambda_\rho^m = r+(k-n_k^{m,t})\tau$. We note that the $\Phi_k^{m,t}(r)$ are translations by part. Then, (neglecting the point $r=t$ which has measure zero anyway) $\widetilde{\mathcal{I}}_k^{m,t} = \Phi_k^{m,t}(\widetilde{\overline{\mathcal{I}}}_k^{m,t})$.

We also define $\widetilde{\overline{\mathcal{I}}}_f^{m,t}$ as the last part of $\overline{\mathcal{I}}_t^m$, with length $\mathcal{L}(\widetilde{\overline{\mathcal{I}}}_f^{m,t}) \le \tau$, and $\widetilde{\mathcal{I}}_f^{m,t}$ as the last part of $\mathcal{I}_t^m$ not corresponding to an interval of length τ of $\overline{\mathcal{I}}_t^m$. That is, $\widetilde{\overline{\mathcal{I}}}_f^{m,t} = \bar{\mathcal{I}}_t^m \setminus \bigcup_{k\in K_m^t}\widetilde{\overline{\mathcal{I}}}_k^{m,t}$ and $\widetilde{\mathcal{I}}_f^{m,t} = \mathcal{I}_t^m\setminus\bigcup_{k\in K_m^t}\widetilde{\mathcal{I}}_k^{m,t}$.

Define

$$\Phi_k^{m,t,+}(r) = \begin{cases}\Phi_k^{m,t}(r) & \text{if } \Phi_k^{m,t}(r)\ge r,\\ r & \text{otherwise,}\end{cases}, \qquad \Phi_k^{m,t,-}(r) = \begin{cases}\Phi_k^{m,t}(r) & \text{if } \Phi_k^{m,t}(r)\le r,\\ r & \text{otherwise.}\end{cases}$$

We need to evaluate the distance $|\Phi_k^{m,t}(r)-r|$ for r in $[(n_k^{m,t}-1)\tau, n_k^{m,t}\tau]$. The following two lemmas also appear in [31], and so we do not include the proofs.

LEMMA 6.4. *For all $m\in\mathcal{M}$, $\Phi_k^{m,t,+}(r) < n_{k+2}^{m,t}\tau + (r-n_k^{m,t}\tau)$ for all $r\in[(n_k^{m,t}-1)\tau, n_k^{m,t}\tau]$ and $k\in]1,\widehat{K}_m^t-2[$, and $\Phi_k^{m,t,-}(r) > n_{k-M}^{m,t}\tau + (r-n_k^{m,t}\tau)$ for all $r\in[(n_k^{m,t}-1)\tau, n_k^{m,t}\tau]$ and $k\in]M+1,\widehat{K}_m^t[$.*

LEMMA 6.5. *Suppose $f(\cdot)$ is nonnegative and integrable over $[0,T]$. For any $t\in[0,T]$ and any $m\in\mathcal{M}$,*

$$\sum_{k\in K_m^t}\left[\int_{(n_k^{m,t}-1)\tau}^{n_k^{m,t}\tau}\int_r^{\Phi_k^{m,t}(r)} f(\rho)\,d\rho\,dr\right] \le \min\{M,2\}\tau\int_0^t f(r)\,dr.$$

Now we demonstrate some results indicating simple forms for the matrizants corresponding to the differential equations defining our dynamics. Consider first the linear, homogeneous systems

$$\dot{\hat{\xi}}_t = A^{\mu_t^\varepsilon}\hat{\xi}_t, \tag{6.4}$$

$$\dot{\hat{\bar{\xi}}}_t = A^{\overline{\mu}_t^\varepsilon}\hat{\bar{\xi}}_t. \tag{6.5}$$

By the simple form of $\overline{\mu}^\varepsilon$ as constant on segments $[k\tau,(k+1)\tau)$, we have that for any $x\in I\!\!R^n$, the solution of (6.5) with $\hat{\bar{\xi}}_0 = x$ is

$$\hat{\bar{\xi}}_t = \exp\left[\int_0^t A^{\overline{\mu}_r^\varepsilon}\,dr\right]x. \tag{6.6}$$

Let the solution of (6.4) with $\hat{\xi}_0 = x$ be $\hat{\xi}_t$. One has

$$\hat{\xi}_t = \Psi(t,0)x \tag{6.7}$$

where Ψ is the matrizant of the system. In Appendix C, we prove:

THEOREM 6.6. $\Psi(t,0) = \exp\left[\int_0^t A^{\mu_r^\varepsilon}\,dr\right]$ *for all $t\in[0,\infty)$, that is,*

$$\hat{\xi}_t = \exp\left[\int_0^t A^{\mu_r^\varepsilon}\,dr\right]x$$

for all $0 \le t \le T < \infty$ *and all* $x \in \mathbb{R}^n$*.*

Now, fix $\delta > 0$ (used in the definition of K_δ – see the discussion just above Lemma 4.2). Fix any $T < \infty$ and $x \in \mathbb{R}^n$. Let $\varepsilon = (\hat{\varepsilon}/2)(1+|x|^2)$. We have

$$\widetilde{S}_T[V^1](x) - \left[\int_0^T L^{\mu_t^\varepsilon}(\xi^\varepsilon{}_t) - \frac{\gamma^2}{2}|w_t^\varepsilon|^2\,dt + V^1(\xi_T^\varepsilon)\right] \le \varepsilon = \frac{\hat{\varepsilon}}{2}(1+|x|^2) \tag{6.8}$$

where ξ^ε satisfies (4.3) with inputs w^ε, μ^ε.

By (6.6), Theorem 6.6 and standard results for nonhomogeneous linear systems, one has

$$\xi_t^\varepsilon = \exp\left[\int_0^t A^{\mu_r^\varepsilon}\,dr\right]x + \int_0^t \exp\left[\int_r^t A^{\mu_\rho^\varepsilon}\,d\rho\right] l_2^{\mu_r^\varepsilon}\,dr + \int_0^t \exp\left[\int_r^t A^{\mu_\rho^\varepsilon}\,d\rho\right]\sigma^{\mu_r^\varepsilon} w_r^\varepsilon\,dr \tag{6.9}$$

$$\overline{\xi}_t^\varepsilon = \exp\left[\int_0^t A^{\overline{\mu}_r^\varepsilon}\,dr\right]x + \int_0^t \exp\left[\int_r^t A^{\overline{\mu}_\rho^\varepsilon}\,d\rho\right] l_2^{\overline{\mu}_r^\varepsilon}\,dr + \int_0^t \exp\left[\int_r^t A^{\overline{\mu}_\rho^\varepsilon}\,d\rho\right]\sigma^{\overline{\mu}_r^\varepsilon} w_r^\varepsilon\,dr. \tag{6.10}$$

THEOREM 6.7. $|\xi_t^\varepsilon - \overline{\xi}_t^\varepsilon| \le \bar{K}_4(M+1)^2\sqrt{1+|x|^2}\,\tau$ *for all* $0 \le t \le T < \infty$*, for proper choice of* $\bar{K}_4$ *independent of* $x \in \mathbb{R}^n$*,* $\tau \in (0,1]$ *and* $\varepsilon \in (0,1]$ *(i.e.,* $\hat{\varepsilon} \in (0, 2/(1+|x|^2)]$*).*

Proof. First, using $(A.\sigma)$ in (6.9) and (6.10), one has

$$\begin{aligned}|\xi_t^\varepsilon - \overline{\xi}_t^\varepsilon| \le{}& \left|\exp\Big[\int_0^t A^{\mu_r^\varepsilon}\,dr\Big] - \exp\Big[\int_0^t A^{\overline{\mu}_r^\varepsilon}\,dr\Big]\right||x| \\ &+ \left|\int_0^t \exp\left[\int_r^t A^{\mu_\rho^\varepsilon}\,d\rho\right] l_2^{\mu_r^\varepsilon}\,dr - \int_0^t \exp\left[\int_r^t A^{\overline{\mu}_\rho^\varepsilon}\,d\rho\right] l_2^{\overline{\mu}_r^\varepsilon}\,dr\right| \\ &+ \left\{\int_0^t \left|\exp\Big[\int_r^t A^{\mu_\rho^\varepsilon}\,d\rho\Big]\sigma - \exp\Big[\int_r^t A^{\overline{\mu}_\rho^\varepsilon}\,d\rho\Big]\sigma\right|^2 dr\right\}^{1/2} \|w^\varepsilon\|_{L_2(0,t)}.\end{aligned} \tag{6.11}$$

Consider the first term on the right in (6.11). Using Lemmas 6.2 and 6.3, one finds

$$\left|\int_0^t A^{\overline{\mu}_r^\varepsilon}\,dr - \int_0^t A^{\mu_r^\varepsilon}\,dr\right| = \left|\sum_m \lambda_t^m A^m - \sum_m \bar{\lambda}_t^m A^m\right| \le \sum_m |\lambda_t^m - \bar{\lambda}_t^m||A^m| \le M^2\overline{A}\tau,$$

where $\overline{A} \doteq \max_{m\in\mathcal{M}}|A^m|$. Using this bound and the Mean Value Theorem (on e^x), we find that, for $\tau \in (0,1]$,

$$\left|1 - \exp\Big[\int_0^t A^{\overline{\mu}_r^\varepsilon}\,dr - \int_0^t A^{\mu_r^\varepsilon}\,dr\Big]\right| = \left|e^0 - \exp\Big[\int_0^t A^{\overline{\mu}_r^\varepsilon}\,dr - \int_0^t A^{\mu_r^\varepsilon}\,dr\Big]\right| \tag{6.12}$$

$$\le \left|\int_0^t A^{\overline{\mu}_r^\varepsilon}\,dr - \int_0^t A^{\mu_r^\varepsilon}\,dr\right| \exp\left\{\max_{t\in[0,T]}\left|\int_0^t A^{\overline{\mu}_r^\varepsilon}\,dr - \int_0^t A^{\mu_r^\varepsilon}\,dr\right|\right\} \le \beta_\tau^1 \tag{6.13}$$

where

$$\beta_\tau^1 = M^2\max_m|A^m|\left[\exp\{M^2\max_m|A^m|\}\right]\tau \doteq M^2\overline{A}\left[\exp\{M^2\overline{A}\}\right]\tau \tag{6.14}$$

independent of $t \ge 0$.

Let $y \in \mathbb{R}^n$. Define $F_t = \exp\big[\int_s^t A^{\mu_r^\varepsilon}\,dr\big]$. Then, using Assumption block $(A.m)$,

$$\frac{d}{dt}\left[y'F_t'F_t y\right] = 2y'\left[F_t'A^{\mu_t^\varepsilon}F_t\right]y = 2(F_t y)'A^{\mu_t^\varepsilon}(F_t y) \le -2c_A\big[y'F_t'F_t y\big].$$

Solving this ordinary differential inequality, one finds $y'F_t'F_ty \le |y|^2 e^{-2c_A(t-s)}$. Since this is true for all $y \in I\!R^n$, we have

$$\left|\exp\Big[\int_s^t A^{\mu_r^\varepsilon}\,dr\Big]\right| \le e^{-c_A(t-s)} \quad \forall\, t \ge s \ge 0. \tag{6.15}$$

By (6.13) and (6.15)

$$\left|\exp\Big[\int_s^t A^{\mu_r^\varepsilon}\,dr\Big] - \exp\Big[\int_s^t A^{\overline{\mu}_r^\varepsilon}\,dr\Big]\right| \le \beta_\tau^1 e^{-c_A(t-s)} \quad \forall\, t \ge s \ge 0. \tag{6.16}$$

We now turn to the third term on the right-hand side of (6.11). Note that

$$\begin{aligned}&\left\{\int_0^t \left|\exp\Big[\int_r^t A^{\mu_\rho^\varepsilon}\,d\rho\Big]\sigma - \exp\Big[\int_r^t A^{\overline{\mu}_\rho^\varepsilon}\,d\rho\Big]\sigma\right|^2 dr\right\}^{1/2}\\ &\le \left\{\int_0^t \left|\exp\Big[\int_r^t A^{\mu_\rho^\varepsilon}\,d\rho\Big] - \exp\Big[\int_r^t A^{\overline{\mu}_\rho^\varepsilon}\,d\rho\Big]\right|^2 |\sigma|^2\,dr\right\}^{1/2}\end{aligned}$$

which, by using (6.16),

$$\le \beta_\tau^1 c_\sigma (2c_A)^{-1/2}. \tag{6.17}$$

Combining this with Assumption $(A.w)$, we have

$$\left\{\int_0^t \left|\exp\Big[\int_r^t A^{\mu_\rho^\varepsilon}\,d\rho\Big]\sigma - \exp\Big[\int_r^t A^{\overline{\mu}_\rho^\varepsilon}\,d\rho\Big]\sigma\right|^2 dr\right\}^{1/2} \|w^\varepsilon\|_{L_2(0,t)} \le \beta_\tau^1 c_\sigma (2c_A)^{-1/2}\sqrt{c_1 + c_2|x|^2}. \tag{6.18}$$

The second term on the right-hand side of (6.11) requires substantially more work. Clearly,

$$\begin{aligned}\left|\int_0^t e^{\int_r^t A^{\mu_\rho^\varepsilon}\,d\rho} l_2^{\mu_r^\varepsilon}\,dr - \int_0^t e^{\int_r^t A^{\overline{\mu}_\rho^\varepsilon}\,d\rho} l_2^{\overline{\mu}_r^\varepsilon}\,dr\right| &\le \left|\int_0^t e^{\int_r^t A^{\mu_\rho^\varepsilon}\,d\rho}(l_2^{\mu_r^\varepsilon} - l_2^{\overline{\mu}_r^\varepsilon})\,dr\right|\\ &\quad + \left|\int_0^t (e^{\int_r^t A^{\mu_\rho^\varepsilon}\,d\rho} - e^{\int_r^t A^{\overline{\mu}_\rho^\varepsilon}\,d\rho}) l_2^{\overline{\mu}_r^\varepsilon}\,dr\right|.\end{aligned} \tag{6.19}$$

First we note that, using (6.16), the second term on the right-hand side of (6.19) satisfies the bound

$$\left|\int_0^t (e^{\int_r^t A^{\mu_\rho^\varepsilon}\,d\rho} - e^{\int_r^t A^{\overline{\mu}_\rho^\varepsilon}\,d\rho}) l_2^{\overline{\mu}_r^\varepsilon}\,dr\right| \le \max_{m\in\mathcal{M}}\{|l_2^m|\}\frac{\beta_\tau^1}{c_A}. \tag{6.20}$$

Now we turn to the first term on the right-hand side of (6.19). The approach will be similar to that used for some estimates in [31]. First, we introduce a shorthand notation which will also be helpful conceptually. For any given $t \in [0,T]$, let

$$\zeta_r \doteq e^{\int_r^t A^{\mu_\rho^\varepsilon}\,d\rho}, \quad \text{and} \quad \bar{\zeta}_r \doteq e^{\int_r^t A^{\overline{\mu}_\rho^\varepsilon}\,d\rho}. \tag{6.21}$$

Note that ζ and $\bar{\zeta}$ satisfy terminal value problems

$$\dot{\zeta} = A^{\mu_r^\varepsilon}\zeta, \qquad \zeta_t = I \tag{6.22}$$

and

$$\dot{\bar{\zeta}} = A^{\mu_r^\varepsilon}\bar{\zeta}, \qquad \bar{\zeta}_t = I. \tag{6.23}$$

Using these, one sees that the first term on the right-hand side of (6.19) takes the form $|\mathcal{A}_t|$ where

$$\mathcal{A}_t \doteq \int_0^t \zeta_r(l_2^{\mu_r^\varepsilon} - l_2^{\overline{\mu}_r^\varepsilon})\,dr = \sum_{m\in\mathcal{M}}\left\{\int_{\mathcal{I}_t^m} \zeta_r l_2^m\,dr - \int_{\overline{\mathcal{I}}_t^m} \zeta_r l_2^m\,dr\right\}. \tag{6.24}$$

Using the earlier definitions, one finds that (6.24) becomes

$$\mathcal{A}_t = \sum_{m\in\mathcal{M}} \mathcal{A}_t^m \tag{6.25}$$

where each

$$\begin{aligned}
\mathcal{A}_t^m = &\left\{ \sum_{k\in K_m^t} \left[\int_{\widetilde{\mathcal{I}}_k^{m,t}} \zeta_r l_2^m\, dr - \int_{(n_k^{m,t}-1)\tau}^{n_k^{m,t}\tau} \zeta_r l_2^m\, dr \right] \right\} \\
&+ \int_{\widetilde{\mathcal{I}}_f^{m,t}} \zeta_r l_2^m\, dr - \int_{\widetilde{\overline{\mathcal{I}}}_f^{m,\,t}} \zeta_r l_2^m\, dr \\
= &\left\{ l_2^m \sum_{k\in K_m^t} \left[\int_{(n_k^{m,t}-1)\tau}^{n_k^{m,t}\tau} \zeta_{\Phi_k^{m,t}(r)} - \zeta_r\, dr \right] \right\} \\
&+ l_2^m \left[\int_{\widetilde{\mathcal{I}}_f^{m,t}} \zeta_r\, dr - \int_{\widetilde{\overline{\mathcal{I}}}_f^{m,\,t}} \zeta_r\, dr \right] \\
\doteq &\left\{ l_2^m \sum_{k\in K_m^t} \left[\int_{(n_k^{m,t}-1)\tau}^{n_k^{m,t}\tau} \zeta_{\Phi_k^{m,t}(r)} - \zeta_r\, dr \right] \right\} + \mathcal{E}_t^m.
\end{aligned} \tag{6.26}$$

For ease of notation on the next few lines, in the case of $b < a$ and any integrable $f(t)$, we let $\int_a^b f(t)\,dt$ denote $\int_b^a f(t)\,dt$. Note that

$$\left|\zeta_{\Phi_k^{m,t}(r)} - \zeta_r\right| = \left| \int_r^{\Phi_k^{m,t}(r)} \dot{\zeta}_\rho\, d\rho \right|$$

which by (6.21)

$$= \left| \int_r^{\Phi_k^{m,t}(r)} A^{\mu_\rho^\varepsilon} \zeta_\rho\, d\rho \right| \le \overline{A} \int_r^{\Phi_k^{m,t}(r)} |\zeta_\rho|\, d\rho. \tag{6.27}$$

However, by (6.21) and (6.15), $|\zeta_\rho| \le e^{-c_A(t-\rho)}$. Employing this in (6.27) yields

$$\left|\zeta_{\Phi_k^{m,t}(r)} - \zeta_r\right| \le \overline{A} e^{-c_A t} \left| \int_r^{\Phi_k^{m,t}(r)} e^{c_A \rho}\, d\rho \right|. \tag{6.28}$$

Substituting (6.28) into (6.26) yields

$$|\mathcal{A}_t^m| \le |l_2^m| \overline{A} e^{-c_A t} \left[\sum_{k\in K_m^t} \int_{(n_k^{m,t}-1)\tau}^{n_k^{m,t}\tau} \left| \int_r^{\Phi_k^{m,t}(r)} e^{c_A \rho}\, d\rho \right| dr \right] + \mathcal{E}_t^m. \tag{6.29}$$

Employing Lemma 6.5 on the sum on the right-hand side of (6.29), one finds

$$\begin{aligned}
|\mathcal{A}_t^m| &\le |l_2^m| \overline{A} e^{-c_A t} \min\{M, 2\} \tau \int_0^t e^{c_a r}\, dr + \mathcal{E}_t^m \\
&\le \bar{K}_1 (M+1)^2 \tau + \mathcal{E}_t^m
\end{aligned} \tag{6.30}$$

for proper choice of $\bar{K}_1 < \infty$ independent of $0 \le t \le T < \infty$ and $m \in \mathcal{M}$.

Now, let $N_t = \max\{n \in \mathbf{N}\cup\{0\}\,|\,n\tau \le t\}$. Then, note that by Lemma 6.3 one has $\lambda^m_{N_t\tau} \le \bar\lambda^m_{N_t\tau} + (M-1)\tau$, and so

$$\lambda^m_t - \bar\lambda^m_{N_t\tau} \le M\tau$$

which implies $\mathcal{L}(\widetilde{\mathcal{I}}^{m,t}_f) \le M\tau$. Also, by the definition of $\widetilde{\overline{\mathcal{I}}}^{m,\,t}_f$, $\mathcal{L}(\widetilde{\overline{\mathcal{I}}}^{m,\,t}_f) \le t - N_t\tau \le \tau$. Then, noting that $|\zeta_r| \le 1$ for all $0 \le r \le t$, one sees that

$$\begin{aligned} |\mathcal{E}^m_t| &= \left| l^m_2 \left[\int_{\widetilde{\mathcal{I}}^{m,t}_f} \zeta_r\,dr - \int_{\widetilde{\overline{\mathcal{I}}}^{m,\,t}_f} \zeta_r\,dr \right] \right| \\ &\le \max_{m\in\mathcal{M}} |l^m_2|(M+1)\tau \end{aligned} \tag{6.31}$$

for all $m \in \mathcal{M}$. Combining (6.25), (6.26), (6.30) and (6.31), one finds that there exists $\bar K_2 < \infty$ such that

$$|\mathcal{A}_t| \le \bar K_2\tau \tag{6.32}$$

for all $0 \le t \le T < \infty$.

Then, combining (6.19), (6.20), the definition of $\mathcal{A}_t$ and (6.32), one finds that there exists $\bar K_3 < \infty$ such that

$$\left| \int_0^t e^{\int_r^t A^{\mu^\varepsilon_\rho}\,d\rho} l_2^{\mu^\varepsilon_r}\,dr - \int_0^t e^{\int_r^t A^{\overline{\mu}^\varepsilon_\rho}\,d\rho} l_2^{\overline{\mu}^\varepsilon_r}\,dr \right| \le \bar K_3\tau \tag{6.33}$$

for all $0 \le t \le T < \infty$. Finally, by combining (6.11), (6.16), (6.18), (6.33) and Assumption $(A.w)$, one has

$$|\xi^\varepsilon_t - \overline{\xi}^\varepsilon_t| \le \beta^1_\tau|x| + \bar K_3\tau + \beta^1_\tau c_\sigma(2c_A)^{-1/2}\sqrt{c_1 + c_2|x|^2} \tag{6.34}$$

which yields the desired result. □

Now that we have a bound on $|\xi^\varepsilon_t - \overline{\xi}^\varepsilon_t|$, we turn to the main problem of the section, that of proving Theorem 6.1. Note that

$$\begin{aligned} &\int_0^T L^{\mu^\varepsilon_t}(\xi^\varepsilon_t) - \frac{\gamma^2}{2}|w^\varepsilon_t|^2\,dt + V^1(\xi^\varepsilon_T) - \left[\int_0^T L^{\overline{\mu}^\varepsilon_t}(\overline{\xi}^\varepsilon_t) - \frac{\gamma^2}{2}|w^\varepsilon_t|^2\,dt + V^1(\overline{\xi}^\varepsilon_T) \right] \\ &= \tfrac12 \int_0^T \xi^\varepsilon_t D^{\mu^\varepsilon_t}\xi^\varepsilon_t - \overline{\xi}^\varepsilon_t D^{\overline{\mu}^\varepsilon_t}\overline{\xi}^\varepsilon_t\,dt + \int_0^T (l_1^{\mu^\varepsilon_t})'\xi^\varepsilon_t - (l_1^{\overline{\mu}^\varepsilon_t})'\,\overline{\xi}^\varepsilon_t\,dt \\ &\quad + \int_0^T \alpha^{\mu^\varepsilon_t} - \alpha^{\overline{\mu}^\varepsilon_t}\,dt + 1/2(\xi^\varepsilon_T)'P^1\xi^\varepsilon_T - 1/2(\overline{\xi}^\varepsilon_T)'P^1\overline{\xi}^\varepsilon_T \end{aligned} \tag{6.35}$$

where we use the fact that $V^1(x) = \frac12 x'P^1x$ for appropriate positive definite P^1. Note that the first integral term on the right-hand side in (6.35) is

$$\begin{aligned} &\int_0^T (\xi^\varepsilon_t)'\left(D^{\mu^\varepsilon_t} - D^{\overline{\mu}^\varepsilon_t}\right)\xi^\varepsilon_t + (\xi^\varepsilon_t)'D^{\overline{\mu}^\varepsilon_t}\xi^\varepsilon_t - (\overline{\xi}^\varepsilon_t)'D^{\overline{\mu}^\varepsilon_t}\overline{\xi}^\varepsilon_t\;dt \\ &= \int_0^T (\xi^\varepsilon_t)'\left(D^{\mu^\varepsilon_t} - D^{\overline{\mu}^\varepsilon_t}\right)\xi^\varepsilon_t\;dt + \int_0^T (\xi^\varepsilon_t + \overline{\xi}^\varepsilon_t)'D^{\overline{\mu}^\varepsilon_t}(\xi^\varepsilon_t - \overline{\xi}^\varepsilon_t)\;dt. \end{aligned} \tag{6.36}$$

Now, with $\bar c_D \doteq \max_{m\in\mathcal{M}}|D^m|$,

$$\left| \int_0^T (\xi^\varepsilon_t + \overline{\xi}^\varepsilon_t)'D^{\overline{\mu}^\varepsilon_t}(\xi^\varepsilon_t - \overline{\xi}^\varepsilon_t)\;dt \right| \le \int_0^T \left| (\overline{\xi}^\varepsilon_t - \xi^\varepsilon_t + 2\xi^\varepsilon_t)'D^{\overline{\mu}^\varepsilon_t}(\xi^\varepsilon_t - \overline{\xi}^\varepsilon_t) \right|\,dt$$

$$\leq \bar{c}_D \int_0^T |\xi_t^\varepsilon - \overline{\xi}_t^\varepsilon|^2 + 2|\xi_t^\varepsilon||\xi_t^\varepsilon - \overline{\xi}_t^\varepsilon|\,dt$$

which by Theorem 6.7

$$\leq \bar{c}_D \left\{ \bar{K}_4^2(M+1)^4(1+|x|^2)\tau^2 T + 2\bar{K}_4(M+1)^2\sqrt{1+|x|^2}\,\tau \int_0^T |\xi_t^\varepsilon|\,dt \right\}$$

which, by using a Holder inequality, Lemma 4.3 and Assumption $(A.w)$,

$$\leq \bar{c}_D \Bigg\{ \bar{K}_4^2(M+1)^4(1+|x|^2)\tau^2 T + 2\bar{K}_4(M+1)^2\sqrt{1+|x|^2}\,\tau\sqrt{T}\Bigg[\sqrt{\frac{2}{c_A}}|x| + \frac{2}{c_A}\max_{m\in\mathcal{M}}|l_2^m|\sqrt{T} + \frac{\sqrt{2}c_\sigma}{c_A}\sqrt{c_1+c_2|x|^2}\Bigg]\Bigg\}$$

$$\leq \bar{K}_5(M+1)^4(1+|x|^2)(1+T)\tau \tag{6.37}$$

where $\bar{K}_5$ is independent of $x \in I\!R^n$, $T \in [0,\infty)$, $\tau \in (0,1]$ and $\varepsilon \in (0,1]$.

We have obtained a bound on the second integral term on the right-hand side of (6.36). Now we turn to the first integral term on the right-hand side. This section of the proof, from here through (6.42), is nearly identical to the analogous estimate in the proof of Theorem 6.1 of [31]. It is also very similar, in technique, to the approach used to bound $|\xi_t^\varepsilon - \overline{\xi}_t^\varepsilon|$ just above (but with T replacing t). Due to this redundancy, only the main steps will be indicated.

One first finds that

$$\left|\int_0^T (\xi_t^\varepsilon)'\left(D^{\mu_t^\varepsilon} - D^{\overline{\mu}_t^\varepsilon}\right)\xi_t^\varepsilon\,dt\right| \leq \sum_{m\in\mathcal{M}} \Bigg\{ \sum_{k\in K_m^T} \left| \int_{(n_k^{m,T}-1)\tau}^{n_k^{m,T}\tau} (\xi^\varepsilon_{\Phi_k^{m,T}(t)})'D^m\xi^\varepsilon_{\Phi_k^{m,T}(t)} - (\xi_t^\varepsilon)'D^m\xi_t^\varepsilon\,dt \right| + \left| \int_{\widetilde{\mathcal{I}}_f^{m,T}} (\xi_t^\varepsilon)'D^m\xi_t^\varepsilon\,dt - \int_{\overline{\mathcal{I}}_f^{m,T}} (\xi_t^\varepsilon)'D^m\xi_t^\varepsilon\,dt \right| \Bigg\}. \tag{6.38}$$

However, note that

$$\left|(\xi^\varepsilon_{\Phi_k^{m,T}(t)})'D^m\xi^\varepsilon_{\Phi_k^{m,T}(t)} - (\xi_t^\varepsilon)'D^m\xi_t^\varepsilon\right| = \left|\int_t^{\Phi_k^{m,T}(t)} 2(\xi_r^\varepsilon)'D^m[A^{\mu_r^\varepsilon}\xi_r^\varepsilon + l_2^{\mu_r^\varepsilon} + \sigma w_r^\varepsilon]\,dr\right|$$

where the $l_2^{\mu_r^\varepsilon}$ term did not appear in the analogous estimate in [31]. Using Assumption Block $(A.m)$, this is

$$\leq \int_t^{\Phi_k^{m,T}(t)} \bar{K}_6\left[|\xi_r^\varepsilon|^2 + |w_r^\varepsilon|^2 + 1\right]\,dr \tag{6.39}$$

where

$$\bar{K}_6 = \bar{c}_D \max\left\{2\overline{A} + 2, c_\sigma^2, \max_{m\in\mathcal{M}}|l_2^m|^2\right\},$$

and we note that the only new term, as compared with the analogous inequality in the proof of Theorem 6.1 of [31], is the "+1" term in the integrand. Combining (6.38) and (6.39), one has

$$\left|\int_0^T (\xi_t^\varepsilon)'\left(D^{\mu_t^\varepsilon} - D^{\overline{\mu}_t^\varepsilon}\right)\xi_t^\varepsilon\,dt\right|$$

$$\leq \sum_{m\in\mathcal{M}} \Bigg\{ \sum_{k\in K_m^T} \left[\int_{(n_k^{m,T}-1)\tau}^{n_k^{m,T}\tau} \int_t^{\Phi_k^{m,T}(t)} \bar{K}_6\left(|\xi_r^\varepsilon|^2 + |w_r^\varepsilon|^2 + 1\right)\,dr \right]$$

$$+\left|\int_{\widetilde{\mathcal{I}}_f^{m,T}}(\xi_t^\varepsilon)'D^m\xi_t^\varepsilon\,dt-\int_{\widetilde{\overline{\mathcal{I}}}_f^{m,\,T}}(\xi_t^\varepsilon)'D^m\xi_t^\varepsilon\,dt\right|\Bigg\},$$

which by Lemma 6.5,

$$\leq M\min\{M,2\}\bar{K}_6\int_0^T\left(|\xi_r^\varepsilon|^2+|w_r^\varepsilon|^2+1\right)\,dr\,\tau$$

$$+\sum_{m\in\mathcal{M}}\left|\int_{\widetilde{\mathcal{I}}_f^{m,T}}(\xi_t^\varepsilon)'D^m\xi_t^\varepsilon\,dt-\int_{\widetilde{\overline{\mathcal{I}}}_f^{m,\,T}}(\xi_t^\varepsilon)'D^m\xi_t^\varepsilon\,dt\right|,$$

which by Assumption $(A.w)$ and Lemma 4.3,

$$\leq\bar{K}_7(M+1)^2(1+|x|^2)(1+T)\tau+\sum_{m\in\mathcal{M}}\left|\int_{\widetilde{\mathcal{I}}_f^{m,T}}(\xi_t^\varepsilon)'D^m\xi_t^\varepsilon\,dt-\int_{\widetilde{\overline{\mathcal{I}}}_f^{m,\,T}}(\xi_t^\varepsilon)'D^m\xi_t^\varepsilon\,dt\right| \tag{6.40}$$

for proper choice of $\bar{K}_7$ independent of $x\in I\!\!R^n$, $T\in[0,\infty)$ and $\tau\in[0,1)$. The last term in (6.40) is handled in the same way as the corresponding term in the proof of Theorem 6.1 of [31], and we do not repeat the details here. One obtains

$$\left|\int_0^T(\xi_t^\varepsilon)'\left(D^{\mu_t^\varepsilon}-D^{\overline{\mu}_t^\varepsilon}\right)\xi_t^\varepsilon\,dt\right|\leq\bar{K}_8(M+1)^2(1+|x|^2)(1+T)\tau, \tag{6.41}$$

for proper choice of $\bar{K}_8$ independent of $x\in I\!\!R^n$, $T\in[0,\infty)$ and $\tau\in[0,1)$. Combining (6.36), (6.37) and (6.41) yields

$$\int_0^T(\xi_t^\varepsilon)'D^{\mu_t^\varepsilon}\xi_t^\varepsilon-(\overline{\xi}_t^\varepsilon)'D^{\overline{\mu}_t^\varepsilon}\overline{\xi}_t^\varepsilon\;dt\leq\bar{K}_9(M+1)^4(1+|x|^2)(1+T)\tau, \tag{6.42}$$

for proper choice of $\bar{K}_9$ independent of $x\in I\!\!R^n$, $T\in[0,\infty)$ and $\tau\in[0,1)$.

We now have, in (6.42), a bound on the first integral term on the right-hand side of (6.35). Next, we turn to the second integral term on the right-hand side. Note that

$$\left|\int_0^T(l_1^{\mu_t^\varepsilon})'\xi_t^\varepsilon-(l_1^{\overline{\mu}_t^\varepsilon})'\,\overline{\xi}_t^\varepsilon\,dt\right|$$

$$\leq\left|\int_0^T(l_1^{\mu_t^\varepsilon}-l_1^{\overline{\mu}_t^\varepsilon})'\xi_t^\varepsilon\,dt\right|+\left|\int_0^T(l_1^{\overline{\mu}_t^\varepsilon})'(\xi_t^\varepsilon-\overline{\xi}_t^\varepsilon)\,dt\right|$$

which using Theorem 6.7,

$$\leq\left|\int_0^T(l_1^{\mu_t^\varepsilon}-l_1^{\overline{\mu}_t^\varepsilon})'\xi_t^\varepsilon\,dt\right|+\max_{m\in\mathcal{M}}|l_1^m|\bar{K}_4(M+1)^2\sqrt{1+|x|^2}\,T\tau. \tag{6.43}$$

The first term on the right-hand side of (6.43) may be treated in a similar manner to that used for $\int_0^T(\xi_t^\varepsilon)'(D^{\mu_t^\varepsilon}-D^{\overline{\mu}_t^\varepsilon})\xi_t^\varepsilon\,dt$ just above. That is,

$$\left|\int_0^T\left(l_1^{\mu_t^\varepsilon}-l_1^{\overline{\mu}_t^\varepsilon}\right)'\xi_t^\varepsilon\,dt\right|\leq\sum_m\Bigg\{\sum_{k\in K_m^T}\left|\int_{(n_k^{m,T}-1)\tau}^{n_k^{m,T}\tau}(l_1^m)'(\xi^\varepsilon_{\Phi_k^{m,T}(t)}-\xi_t^\varepsilon)\,dt\right|$$

$$+\left|\int_{\widetilde{\mathcal{I}}_f^{m,T}}(l_1^m)'\xi_t^\varepsilon\,dt-\int_{\widetilde{\overline{\mathcal{I}}}_f^{m,\,T}}(l_1^m)'\xi_t^\varepsilon\,dt\right|\Bigg\},$$

and by very similar steps, one finds that

$$\leq \bar{K}_{10}(M+1)^2(1+|x|^2)(1+T)\tau \tag{6.44}$$

for proper choice of $\bar{K}_{10}$ independent of $x \in I\!R^n$, $T \in [0,\infty)$, $\tau \in (0,1]$ and $\varepsilon \in (0,1]$. Combining (6.43) and (6.44), one obtains

$$\left| \int_0^T (l_1^{\mu_t^\varepsilon})' \xi_t^\varepsilon - (l_1^{\overline{\mu}_t^\varepsilon})' \,\overline{\xi}_t^\varepsilon \, dt \right| \leq \bar{K}_{11}(M+1)^2(1+|x|^2)(1+T)\tau \tag{6.45}$$

independent of $x \in I\!R^n$, $T \in [0,\infty)$, $\tau \in (0,1]$ and $\varepsilon \in (0,1]$.

We now have bounds on the first and second integral terms on the right-hand side of (6.35). For the third integral term, one easily finds (using Lemmas 6.2 and 6.3)

$$\left| \int_0^T \alpha^{\mu_t^\varepsilon} - \alpha^{\overline{\mu}_t^\varepsilon} \, dt \right| = \left| \sum_m \lambda_t^m \alpha^m - \sum_m \bar{\lambda}_t^m \alpha^m \right| \leq M \left(\sum_m |\alpha^m| \right) \tau. \tag{6.46}$$

Lastly,

$$\begin{aligned} &|(\xi_T^\varepsilon)' P^1 \xi_T^\varepsilon - (\overline{\xi}_T^\varepsilon)' P^1 \overline{\xi}_T^\varepsilon| \\ &\leq |P^1| \left[|\xi_T^\varepsilon - \overline{\xi}_T^\varepsilon|^2 + 2|\xi_T^\varepsilon|\, |\xi_T^\varepsilon - \overline{\xi}_T^\varepsilon| \right] \end{aligned}$$

which, using Theorem 6.7, is

$$\leq |P^1| \left[\bar{K}_4^2 (M+1)^4 (1+|x|^2)\tau^2 + 2\bar{K}_4(M+1)^2 \sqrt{1+|x|^2}\, \tau |\xi_T^\varepsilon| \right],$$

and by using Lemma 4.2 and Assumption $(A.w)$, this is

$$\begin{aligned} &\leq |P^1| \Bigg[\bar{K}_4^2 (M+1)^4 (1+|x|^2)\tau^2 \\ &\quad + 2\bar{K}_4(M+1)^2 \sqrt{1+|x|^2}\, \tau \left(|x| + \frac{2}{c_A} \max_{m \in \mathcal{M}} |l_2^m| + \frac{c_\sigma}{\sqrt{c_A}} \sqrt{c_1 + c_2|x|^2} \right) \Bigg] \end{aligned}$$

which for proper choice of $\bar{K}_{12}$,

$$\leq \bar{K}_{12}(M+1)^4(1+|x|^2)(1+T)\tau \tag{6.47}$$

Combining (6.35), (6.42), (6.45), (6.46) and (6.47), one has

$$\begin{aligned} &\left| \int_0^T L^{\mu_t^\varepsilon}(\xi_t^\varepsilon) - \frac{\gamma^2}{2}|w_t^\varepsilon|^2 \, dt + V^1(\xi_T^\varepsilon) - \left[\int_0^T L^{\overline{\mu}_t^\varepsilon}(\overline{\xi}_t^\varepsilon) - \frac{\gamma^2}{2}|w_t^\varepsilon|^2 \, dt + V^1(\overline{\xi}_T^\varepsilon) \right] \right| \\ &\leq \bar{K}_{13}(M+1)^4(1+|x|^2)(1+T)\tau \end{aligned} \tag{6.48}$$

for proper choice of $\bar{K}_{13}$ independent of $x \in I\!R^n$, $T \in [0,\infty)$, $\tau \in (0,1]$ and $\varepsilon \in (0,1]$.

Combining (4.16), (4.18), (6.8) and (6.48), one has

$$\widetilde{S}_T[V^1](x) - \bar{\bar{S}}_T^\tau[V^1](x) \leq \varepsilon + \bar{K}_{13}(M+1)^4(1+|x|^2)(1+T)\tau$$

Since it is true for all $\varepsilon \geq 0$ and since $\bar{K}_{13}$ is independent of $\varepsilon \in (0,1]$, we finally obtain

$$\widetilde{S}_T[V^1](x) - \bar{\bar{S}}_T^\tau[V^1](x) \leq \bar{K}_{13}(M+1)^4(1+|x|^2)(1+T)\tau,$$

which completes the proof of Theorem 6.1.

Combining Theorems 4.5 and 6.1, we see that for sufficiently large $T < \infty$ and sufficiently small $\tau > 0$, $\bar{\bar{S}}_T^\tau[V]$ approximates $\widetilde{V}$ arbitrarily well. However, this does not give us a specific error bound or requisite relative rates for T, τ. For this we will also need to consider the errors due to time truncation, and we now address such.

7. Finite-time truncation errors. We now know that by taking T large and τ small, we can approximate $\widetilde{V}(x)$ arbitrarily closely by $\bar{\bar{S}}_T^\tau[V^1](x)$, where the latter is obtained from the numerical algorithm (see Section 10). However, we would like to know how large $T = N\tau$ must be, and how small τ must be to achieve a desired accuracy. Further, we would like to know at what rate should one take $N \to \infty$ relative to the rate for $\tau \downarrow 0$ in order to obtain good convergence.

Given $T = N\tau$, Theorem 6.1 indicates an error bound on $\left|\bar{\bar{S}}_T^\tau[V^1] - \widetilde{S}_T[V^1]\right|$ as a function of τ (as well as T). Now we would like an estimate of $\left|\widetilde{S}_T[V^1] - \widetilde{V}\right|$ as a function of T. Unfortunately, this estimate will require another assumption on the behavior of ε–optimal trajectories of the system, in the general class of coercivity-type assumptions. (One would prefer an assumption that only involves the instantaneous structure of $\widetilde{H}$ or $\widetilde{\widetilde{H}}$, but we have, so far, been unable to obtain such.)

Assume there exist $\underline{T}, c_3 \in (0,\infty)$ such that for all $x \in I\!\!R^n$, all $\varepsilon \in (0,1]$, and all $\mu^\varepsilon, w^\varepsilon$ which are ε–optimal for $\widetilde{V}(x)$ (i.e., such that $\widetilde{J}(x,\mu^\varepsilon,w^\varepsilon) \geq \widetilde{V}(x) - \varepsilon$), one has

$$(A.\xi) \qquad \int_0^T L^{\mu_t^\varepsilon}(\xi_t^\varepsilon)\,dt \geq c_3 \int_0^T |\xi_t^\varepsilon|^2\ dt \quad \forall T \geq \underline{T}$$

where $\dot{\xi}_t^\varepsilon = A^{\mu_t^\varepsilon}\xi_t^\varepsilon + l_2^{\mu_t^\varepsilon} + \sigma^{\mu_t^\varepsilon} w_t^\varepsilon$, $\xi_0^\varepsilon = x$.

This assumption is used only in this section, and indirectly, in the Combined Errors section that follows. It is also worth noting that Assumption $(A.m)$ is not needed in this section.

LEMMA 7.1. *Assume $(A.\xi)$ (as well as the assumptions of Section 2). Then, there exist $\bar{c}_1, \bar{c}_2 < \infty$ such that for any $T \in [\underline{T},\infty)$, $\varepsilon \in (0,1]$, $x \in I\!\!R^n$ and ε–optimal $\mu^\varepsilon, w^\varepsilon$, one has*

$$\int_0^T |\xi_t^\varepsilon|^2\ dt \leq \bar{c}_1 + \bar{c}_2|x|^2$$

where

$$\dot{\xi}_t^\varepsilon = f^{\mu_t^\varepsilon}(\xi_t^\varepsilon, w_t^\varepsilon) \doteq A^{\mu_t^\varepsilon}\xi_t^\varepsilon + l_2^{\mu_t^\varepsilon} + \sigma^{\mu_t^\varepsilon} w_t^\varepsilon \ \textit{with}\ \xi_0^\varepsilon = x. \tag{7.1}$$

Proof. Let $x \in I\!\!R^n$ and $T \in [\underline{T},\infty)$. Let $R \geq |x|$. Let $\varepsilon, \bar{\varepsilon} \in (0,1]$ and $\mu^\varepsilon, w^\varepsilon$ be ε–optimal for $\widetilde{V}(x)$. By Assumption $(A.w)$ and Lemma 4.2, there exists $\bar{R} = \bar{R}(R) < \infty$ such that $|\xi_t^\varepsilon| \leq \bar{R}$ for all $t \in [0,\infty)$. Now, as in the proof of Theorem 4.1, we approximate $\widetilde{V}$ by a one-parameter family $V^\delta \in C^\infty$ where $V^\delta(x) = [g^\delta * \widetilde{V}](x) = \int_{I\!\!R^n} g^\delta(x-y)\widetilde{V}(y)\,dy$ with g^δ as described there. In particular, using the same analysis as that which yielded (4.5) and (4.14), there exists $\delta = \delta(T,\bar{R},\bar{\varepsilon}) > 0$ such that

$$\left|\widetilde{V}(y) - V^\delta(y)\right| < \bar{\varepsilon} \tag{7.2}$$

and

$$\widetilde{H}(y, \operatorname{grad} V^\delta(y)) < \frac{\bar{\varepsilon}}{T} \tag{7.3}$$

for all $y \in B_{\bar{R}}(0)$.

Let $\delta = \delta(T,\bar{R},\bar{\varepsilon})$ be such that (7.2),(7.3) hold. Then, by Assumption $(A.\xi)$,

$$\begin{aligned} &c_3 \int_0^T |\xi_t^\varepsilon|^2\,dt - \frac{\gamma^2}{2}\int_0^T |w_t^\varepsilon|^2\,dt \\ &\leq \int_0^T L^{\mu_t^\varepsilon}(\xi_t^\varepsilon) - \frac{\gamma^2}{2}|w_t^\varepsilon|^2\,dt, \end{aligned}$$

and using the fact that $V^\delta \in C^1$,

$$= \int_0^T L^{\mu_t^\varepsilon}(\xi_t^\varepsilon) - \frac{\gamma^2}{2}|w_t^\varepsilon|^2 + f^{\mu_t^\varepsilon}(\xi_t^\varepsilon, w_t^\varepsilon) \cdot \operatorname{grad} V^\delta(\xi_t^\varepsilon)\, dt - \int_0^T f^{\mu_t^\varepsilon}(\xi_t^\varepsilon, w_t^\varepsilon) \cdot \operatorname{grad} V^\delta(\xi_t^\varepsilon)\, dt,$$

which, using the definition of $\widetilde{H}$ and (7.1),

$$\leq \int_0^T \widetilde{H}(\xi_t^\varepsilon, \operatorname{grad} V^\delta(\xi_t^\varepsilon))\, dt - \int_0^T \dot{\xi}_t^\varepsilon \cdot \operatorname{grad} V^\delta(\xi_t^\varepsilon)\, dt,$$

which by (7.3),

$$\leq V^\delta(x) - V^\delta(\xi_T^\varepsilon) + \bar{\varepsilon}.$$

which by (7.2),

$$\leq \widetilde{V}(x) - \widetilde{V}(\xi_T^\varepsilon) + 3\bar{\varepsilon},$$

and since $\widetilde{V} \in Q_K$ by Theorem 4.1,

$$\leq \frac{K}{2}|x|^2 + 3\bar{\varepsilon}.$$

Rearranging this last statement, we have

$$c_3 \int_0^T |\xi_t^\varepsilon|^2\, dt \leq \frac{\gamma^2}{2}\|w^\varepsilon\|^2_{L_2(0,T)} + \frac{K}{2}|x|^2 + 3\bar{\varepsilon},$$

which, using Assumption $(A.w)$, implies that

$$\int_0^T |\xi_t^\varepsilon|^2\, dt \leq \bar{c}_1 + \bar{c}_2|x|^2 + 3\bar{\varepsilon},$$

where $\bar{c}_1 = \frac{\gamma^2 c_1}{2c_3}$ and $\bar{c}_2 = \frac{\gamma^2 c_2 + K}{2c_3}$. Since this is true for all $\bar{\varepsilon} > 0$, one has the asserted result. ∎

LEMMA 7.2. *Assume $(A.\xi)$ (as well as the assumptions of Section 2). Then, given $\overline{T} \in [\underline{T}, \infty)$, there exists $T \in [\overline{T}/2, \overline{T}]$ such that*

$$|\xi_T^\varepsilon|^2 \leq \frac{2}{\overline{T}}\left[\bar{c}_1 + \bar{c}_2|x|^2\right].$$

Proof. Suppose not. Then

$$\int_0^{\overline{T}} |\xi_t^\varepsilon|^2\, dt \geq \int_{\overline{T}/2}^{\overline{T}} |\xi_t^\varepsilon|^2\, dt > \int_{\overline{T}/2}^{\overline{T}} \frac{2}{\overline{T}}\left[\bar{c}_1 + \bar{c}_2|x|^2\right]\, dt = \bar{c}_1 + \bar{c}_2|x|^2$$

which contradicts Lemma 7.1. ∎

We now proceed to obtain the bound on $\left|\widetilde{S}_T[V^1] - \widetilde{V}\right|$. Let $x \in I\!R^n$, $\overline{T} \in [\underline{T}, \infty)$, $\varepsilon \in (0,1]$, and $\mu^\varepsilon, w^\varepsilon$ be ε–optimal for $\widetilde{V}(x)$ (with corresponding trajectory ξ^ε). For any $T \in [0, \infty)$,

$$\begin{aligned} \widetilde{V}(x) &\leq \int_0^\infty L^{\mu_t^\varepsilon}(\xi_t^\varepsilon) - \frac{\gamma^2}{2}|w_t^\varepsilon|^2\, dt + \varepsilon \\ &= \int_0^T L^{\mu_t^\varepsilon}(\xi_t^\varepsilon) - \frac{\gamma^2}{2}|w_t^\varepsilon|^2\, dt + \int_T^\infty L^{\mu_t^\varepsilon}(\xi_t^\varepsilon) - \frac{\gamma^2}{2}|w_t^\varepsilon|^2\, dt + \varepsilon \\ &\leq \int_0^T L^{\mu_t^\varepsilon}(\xi_t^\varepsilon) - \frac{\gamma^2}{2}|w_t^\varepsilon|^2\, dt + \widetilde{V}(\xi_T^\varepsilon) + \varepsilon. \end{aligned} \tag{7.4}$$

Let $T \in [\overline{T}/2, \overline{T}]$ be as asserted in Lemma 7.2. Using the fact that $\widetilde{V} \in Q_K$ (by Theorem 4.1), (7.4) implies

$$\widetilde{V}(x) \leq \widetilde{S}_T[0](x) + \frac{K}{2}|\xi_T^\varepsilon|^2 + \varepsilon$$

which by Lemma 7.2,

$$\leq \widetilde{S}_T[0](x) + \frac{K}{2\overline{T}}\left(\bar{c}_1 + \bar{c}_2|x|^2\right) + \varepsilon.$$

Since this holds for all $\varepsilon \in (0, 1]$, we have

$$\widetilde{S}_T[0](x) \geq \widetilde{V}(x) - \frac{K}{2\overline{T}}\left(\bar{c}_1 + \bar{c}_2|x|^2\right). \tag{7.5}$$

Suppose $V \in Q_K$. Then $V \geq 0$, and so, by the monotonicity of $\widetilde{S}_T[\cdot]$, (7.5) implies

$$\widetilde{V}(x) - \frac{K}{2\overline{T}}\left(\bar{c}_1 + \bar{c}_2|x|^2\right) \leq \widetilde{S}_T[V](x). \tag{7.6}$$

Suppose also that $V \leq \widetilde{V}$. Then, the monotonicity of $\widetilde{S}_T[\cdot]$ and Theorem 4.5 imply

$$\widetilde{S}_T[V] \leq \widetilde{S}_T[\widetilde{V}] = \widetilde{V}. \tag{7.7}$$

Combining (7.6) and (7.7), we have:

THEOREM 7.3. *Assume (A.ξ) (as well as the assumptions of Section 2). Let $V \in Q_K$ with $V \leq \widetilde{V}$. Then, there exists $\bar{K}_\delta < \infty$ such that for any $\overline{T} \in [\underline{T}, \infty)$,*

$$\widetilde{V}(x) - \frac{\bar{K}_\delta}{\overline{T}}(1 + |x|^2) \leq \widetilde{S}_{\overline{T}}[V](x) \leq \widetilde{V}(x).$$

8. Combined errors. We can now obtain an explicit estimate of the convergence of $\bar{\bar{S}}_T^\tau[V^1]$ to $\widetilde{V}$ as $\tau \downarrow 0$ and $T \to \infty$. Indeed, combining the results of Theorem 6.1 and Theorem 7.3, we have that for all $x \in I\!R^n$, all $T \in [\underline{T}, \infty)$, and all sufficiently small $\tau > 0$,

$$\begin{aligned}\bar{\bar{S}}_T^\tau[V^1](x) \leq \widetilde{V}(x) &= \widetilde{S}_T[\widetilde{V}](x)\\ &\leq \widetilde{S}_T[V^1](x) + \frac{\bar{K}_\delta}{T}(1 + |x|^2)\\ &\leq \bar{\bar{S}}_T^\tau[V^1](x) + \frac{\bar{K}_\delta}{T}(1 + |x|^2) + \bar{K}_{13}(M + 1)^4(1 + T)(1 + |x|^2)\tau.\end{aligned} \tag{8.1}$$

If we want $0 \leq \widetilde{V}(x) - \bar{\bar{S}}_T^\tau[V^1](x) \leq 2\varepsilon(1 + |x|^2)$, we can, for example, choose $T = \bar{K}_\delta/\varepsilon$ and $\tau \leq \varepsilon^2/[\bar{K}_{13}(M + 1)^4(1 + \bar{K}_\delta)]$. Consequently, one can get an approximation of the order ε, with $N\tau = T \propto \varepsilon^{-1}$ and $\tau \propto \varepsilon^2$, which implies $N \propto \varepsilon^{-3}$. Also, if the last term in (8.1) could be sharpened by elimination of the $(1+T)$ term (which we believe might be true for a reasonable problem class, but have been unable to prove), then one would have $N \propto \varepsilon^{-2}$ rather than $N \propto \varepsilon^{-3}$.

9. Propagation via dual operators. It remains to demonstrate that one can equivalently replace the propagation by semigroup operator $\bar{S}_\tau$ with the propagation in the semiconvex dual space by the corresponding max-plus integral operator with kernel $\overline{\mathcal{B}}_\tau(\cdot, \cdot)$ (where $\overline{\mathcal{B}}_\tau(x, y) = \bigoplus_{m \in \mathcal{M}} \mathcal{B}_\tau^m(x, y)$).

Let $\mathcal{S}_\beta$ be the set of $\phi : I\!R^n \to I\!R$ such that $\phi(x) + \frac{1}{2}x'\beta x$ is convex. It will be implicit throughout that we consider only $\mathcal{S}_\beta$ spaces where the β are symmetric and definite (either positive definite or negative definite).

We briefly review semiconvex duality. Proofs of the next results may be found in [20], [34]. We will employ certain transform kernel functions, $\psi : I\!R^n \times I\!R^n \to I\!R$ which take the form

$$\psi(x,z) = \tfrac{1}{2}(x-z)'C(x-z)$$

with nonsingular, symmetric C satisfying $C + \beta < 0$ (i.e., $C+\beta$ negative definite). The following semiconvex duality result requires only a small modification of convex duality and Legendre/Fenchel transform results; see Section 3 of [43], and more generally, [44].

THEOREM 9.1. *Let $\phi \in \mathcal{S}_\beta$. Let C and ψ be as above. Then, for all $x \in I\!R^n$,*

$$\phi(x) = \max_{z\in I\!R^n} [\psi(x,z) + a(z)] \tag{9.1}$$

$$= \int_{I\!R^n}^{\oplus} \psi(x,z) \otimes a(z)\,dz = \psi(x,\cdot) \odot a(\cdot) \tag{9.2}$$

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

$$a(z) = -\max_{x\in I\!R^n} [\psi(x,z) - \phi(x)] \tag{9.3}$$

$$= -\int_{I\!R^n}^{\oplus} \psi(x,z) \otimes [-\phi(x)]\,dx = -\left\{\psi(\cdot,z) \odot [-\phi(\cdot)]\right\} \tag{9.4}$$

which using the notation of [8]

$$= \left\{\psi(\cdot,z) \odot [\phi^-(\cdot)]\right\}^- . \tag{9.5}$$

We will refer to a as the *semiconvex dual* of ϕ (with respect to ψ).

THEOREM 9.2. *Let $\phi \in \mathcal{S}_\beta \subset \mathcal{S}_{-C}$ with semiconvex dual denoted by a. There exist $\eta, \overline{\tau} > 0$ such that $S^m_\tau[\phi] \in \mathcal{S}_{-(C+\eta I\tau)}$ for all $\tau \in [0,\overline{\tau}]$. Further, $S^m_\tau[\phi](x) = \psi(x,\cdot) \odot \hat{a}^1_m(\cdot)$ where $\hat{a}^1_m(y) = \mathcal{B}^m_\tau(y,\cdot) \odot a(\cdot)$, a is the semiconvex dual of ϕ and*

$$\mathcal{B}^m_\tau(y,z) = -\psi(\cdot,y) \odot [-S^m_\tau[\psi(\cdot,z)](\cdot)]. \tag{9.6}$$

Proof. The proof of the first assertion is very similar to the proof of Theorem 3.7 in [32], where it is demonstrated that $\widetilde{S}_\tau[\psi(\cdot,z)] \in \mathcal{S}_{-(C+\eta I\tau)}$. One difference is that in S^m_τ, the value of m is fixed rather than the case in $\widetilde{S}_\tau$ where one is optimizing over μ-processes, and so there is some simplification in that respect here.

A second difference is that $\psi(\cdot,z)$ is now replaced by $\phi \in \mathcal{S}_\beta$. However, $\phi \in \mathcal{S}_\beta$ implies $\phi(y) + \frac{1}{2}y'\beta y$ is convex and so for any $y, \delta_y \in I\!R^n$,

$$\phi(y-\delta_y) + \tfrac{1}{2}(y-\delta_y)'\beta(y-\delta_y) - 2[\phi(y) + \tfrac{1}{2}y'\beta y] + \phi(y+\delta_y) + \tfrac{1}{2}(y+\delta_y)'\beta(y+\delta_y) \geq 0$$

which implies

$$\begin{aligned}
&\phi(y-\delta_y) - 2\phi(y) + \phi(y+\delta_y)\\
&\geq -\tfrac{1}{2}\left[(y-\delta_y)'\beta(y-\delta_y) - 2y'\beta y + \tfrac{1}{2}(y+\delta_y)'\beta(y+\delta_y)\right]\\
&= -\delta_y'\beta\delta_y\\
&> \delta_y' C\delta_y = \psi(y-\delta_y,z) - 2\psi(y,z) + \psi(y+\delta_y,z)
\end{aligned} \tag{9.7}$$

for any $z \in I\!R^n$. This inequality will enable us to substitute $\psi(\cdot,z)$ for ϕ below.

Lastly, the running cost and dynamics have also been modified with the addition of linear and constant terms in the cost and a constant term in the dynamics. However, these will have little impact on the proof. We will begin the proof, and soon reach a point where we may refer to the remainder of the proof of Theorem 3.7 in [32].

Fix any $x, \nu \in I\!R^n$ with $|\nu| = 1$ and any $\delta > 0$. Fix $\tau > 0$ (to be specified below), and let $\varepsilon > 0$. Let w^ε be ε–optimal in $S_\tau^m[\phi](x)$. Let

$$\widehat{\mathcal{I}}^\phi(x,\tau,w^\varepsilon) \doteq \int_0^\tau L^m(\xi_t^\varepsilon) - \frac{\gamma^2}{2}|w_t^\varepsilon|^2\,dt + \phi(\xi_\tau^\varepsilon) \tag{9.8}$$

where ξ^ε satisfies (2.2) with input w^ε. Then,

$$\begin{aligned} &S_\tau^m[\phi](x-\delta\nu) - 2S_\tau^m[\phi](x) + S_\tau^m[\phi](x+\delta\nu) \\ &\geq \widehat{\mathcal{I}}^\phi(x-\delta\nu,\tau,w^\varepsilon) - 2\widehat{\mathcal{I}}^\phi(x,\tau,w^\varepsilon) + \widehat{\mathcal{I}}^\phi(x+\delta\nu,\tau,w^\varepsilon) - 2\varepsilon. \end{aligned} \tag{9.9}$$

Let $\xi^{\varepsilon,-\delta}, \xi^{\varepsilon,0}, \xi^{\varepsilon,\delta}$ satisfy (2.2) with input w^ε and initial conditions $\xi_0^{\varepsilon,-\delta} = x - \delta\nu$, $\xi_0^{\varepsilon,0} = x$, and $\xi_0^{\varepsilon,\delta} = x + \delta\nu$. Note that

$$\xi_t^{\varepsilon,\delta} - \xi_t^{\varepsilon,0} = \xi_t^{\varepsilon,0} - \xi_t^{\varepsilon,-\delta} \quad \forall\, t \in [0,\tau], \tag{9.10}$$

and we denote this difference as Δ_t^+. Substituting definition (9.8) in (9.9) yields

$$\begin{aligned} &S_\tau^m[\phi](x-\delta\nu) - 2S_\tau^m[\phi](x) + S_\tau^m[\phi](x+\delta\nu) \\ &\geq \int_0^\tau \tfrac{1}{2}(\xi_t^{\varepsilon,-\delta})' D^m \xi_t^{\varepsilon,-\delta} - (\xi_t^{\varepsilon,0})' D^m \xi_t^{\varepsilon,0} + \tfrac{1}{2}(\xi_t^{\varepsilon,\delta})' D^m \xi_t^{\varepsilon,\delta} \\ &\qquad + (l_2^m)'\xi_t^{\varepsilon,-\delta} - 2(l_2^m)'\xi_t^{\varepsilon,0} + (l_2^m)'\xi_t^{\varepsilon,\delta}\,dt \\ &\quad + \phi(\xi_\tau^{\varepsilon,-\delta}) - 2\phi(\xi_\tau^{\varepsilon,0}) + \phi(\xi_\tau^{\varepsilon,\delta}) - 2\varepsilon, \end{aligned}$$

and upon using (9.7) and (9.10), we see that this is

$$> \int_0^\tau (\Delta_t^+)' D^m \Delta_t^+\,dt + (\Delta_\tau^+)' C \Delta_\tau^+ - 2\varepsilon,$$

and we specifically note that the linear and constant terms in the payoff are irrelevant to this second difference bound.

Now, note that $\dot{\Delta}_t^+ = A^m(\xi_t^{\varepsilon,\delta} - \xi_t^{\varepsilon,0}) + l_2^m - l_2^m = A^m\Delta_t^+$ with $\Delta_0^+ = \delta\nu$. Consequently, the constant term in the dynamics is irrelevant to the second-difference, and in particular,

$$\Delta_t^+ = \exp[A^m t]\delta\nu.$$

The remainder of the proof of the first assertion is identical to the corresponding portion of the proof of Theorem 3.7 of [32], and we do not repeat it here.

Lastly, we turn to the second assertion. This is essentially the same as the proof of Proposition 3.10 in [32] (with S_τ^m replacing $\widetilde{S}_\tau$ there), and we do not repeat the steps here. □

The above theorem implies that one can propagate in the semiconvex dual space via the max-plus integral operation with kernel $\mathcal{B}_\tau^m$ in place of propagation via S_τ^m. Next we use that result to show that we may also replace propagation via $\bar{S}_\tau$ with propagation in the semiconvex dual space.

THEOREM 9.3. *Let $\phi \in \mathcal{S}_\beta \subset \mathcal{S}_{-C}$ with semiconvex dual denoted by a. For $\tau \in [0,\overline{\tau}]$, $\bar{S}_\tau[\phi](x) = \psi(x,\cdot) \odot a^1(\cdot)$ where $a^1(y) = \overline{\mathcal{B}}_\tau(y,\cdot) \odot a(\cdot)$, a is the semiconvex dual of ϕ, and*

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

Proof. Note that

$$\bar{S}_\tau[\phi](x) = \bigoplus_{m\in\mathcal{M}} S_\tau^m[\phi](x)$$

which by Theorem 9.2

$$\begin{aligned}
&= \bigoplus_{m\in\mathcal{M}} \psi(x,\cdot)\odot \hat{a}_m^1(\cdot) = \bigoplus_{m\in\mathcal{M}} \int_{I\!R^n}^{\oplus} \psi(x,y)\otimes \hat{a}_m^1(y)\,dy \\
&= \bigoplus_{m\in\mathcal{M}} \int_{I\!R^n}^{\oplus} \psi(x,y)\otimes \left[\int_{I\!R^n}^{\oplus} \mathcal{B}_\tau^m(y,z)\otimes a(z)\,dz\right] dy \\
&= \int_{I\!R^n}^{\oplus} \psi(x,y)\otimes \int_{I\!R^n}^{\oplus} \left[\bigoplus_{m\in\mathcal{M}} \mathcal{B}_\tau^m(y,z)\otimes a(z)\right] dy\,dz.
\end{aligned}$$

□

10. Modification of the algorithm. The algorithm was fully discussed in [32]. Also, some important comments appear in [30]. However, we now have the additional terms l_1^m, l_2^m and α^m. The only change this induces in the actual implementation of the algorithm is that there are now some additional terms in the quadratic functions $\mathcal{B}_\tau^m(y,z)$. We now indicate the minor modifications necessary for the generalization.

The computation of the $\mathcal{B}_\tau^m$ is performed by solving some differential Riccati equations and linear differential equations for each of the m systems. In particular, note that one may let $S_\tau^m[\psi(\cdot,z)](x)$ take the form

$$S_\tau^m[\psi(\cdot,z)](x) = \tfrac{1}{2}\Big[(x-\Lambda_\tau^m z)' P_\tau^m (x-\Lambda_\tau^m z) + z' R_\tau^m z + 2(L_\tau^m)' x + 2(\mathcal{K}_\tau^m)' z + b_\tau^m\Big].$$

The P_τ^m, Λ_τ^m and R_τ^m are $n\times n$ matrices, the L_τ^m and $\mathcal{K}_\tau^m$ are vectors of length n, and b_τ^m is a scalar. At $t=0$, these time-dependent coefficients satisfy $P_0^m = C$, $\Lambda_0^m = I$, $R_0^m = 0$, $L_0^m = 0$, $\mathcal{K}_0^m = 0$ and $b_0^m = 0$, because they correspond to ψ. They evolve according to

$$\begin{aligned}
\dot{P}^m &= (A^m)' P^m + P^m A^m + D^m + P^m \Sigma^m P^m \\
\dot{\Lambda}^m &= \left[(P^m)^{-1} D^m + A^m\right]\Lambda^m \\
\dot{R}^m &= (\Lambda^m)' D^m \Lambda^m, \\
\dot{L}^m &= [P^m\Sigma^m + (A^m)']\Lambda^m + l_1^m + P^m l_2^m, \\
\dot{\mathcal{K}}^m &= (\Lambda^m)'[P^m l_2^m + P^m \Sigma^m L^m], \\
\dot{b}^m &= \alpha^m + (L^m)'\Sigma^m L^m + 2(L^m)' l_2^m.
\end{aligned}$$

The $P_\tau^m, \Lambda_\tau^m, R_\tau^m, L_\tau^m, \mathcal{K}_\tau^m, b_\tau^m$ may be computed from these ordinary differential equations via a Runge-Kutta algorithm (or other technique) with initial time $t=0$ and terminal time $t=\tau$. We remark that each of these need only be computed once.

Next, noting that each $\mathcal{B}_\tau^m$ is given by (9.6), one has

$$\mathcal{B}_\tau^m(x,z) = \tfrac{1}{2}\left[x' M_{1,1}^m x + 2x' M_{1,2}^m z + z' M_{2,2}^m z + 2(\lambda_1^m)' x + 2(\lambda_2^m)' z + \beta^m\right]$$

where with shorthand notation $D_\tau \doteq (P_\tau^m - C)$,

$$\begin{aligned}
M_{1,1}^m &= -C D_\tau^{-1} P_\tau^m \\
M_{1,2}^m &= C D_\tau^{-1} P_\tau^m \Lambda_\tau^m
\end{aligned}$$

$$
\begin{aligned}
M^m_{2,2} &= R^m_\tau - (\Lambda^m_\tau)' C D_\tau^{-1} P^m_\tau \Lambda^m_\tau,\\
\lambda^m_1 &= -C D_\tau^{-1} L^m_\tau,\\
\lambda^m_2 &= \Lambda^m_\tau P^m_\tau D_\tau^{-1} L^m_\tau + \mathcal{K}^m_\tau,\\
\beta^m &= b^m_\tau - (L^m_\tau)' D_\tau^{-1} L^m_\tau.
\end{aligned}
$$

Note that each of these need only be computed once as well. Once one has these $\mathcal{B}^m_\tau$, the remainder of the algorithm is identical to what is presented in [32], with the exception that the initialization of the algorithm is with $\overline{V}^0 = V^1$.

Acknowledgments. The author thanks L. Jonathan Kluberg for his contribution to certain estimates appearing in Section 6.

11. Appendix A. The proofs of Lemmas 4.2 and 4.3 are nearly identical to the corresponding development in [40]. For the purposes of completeness and self-containment, we sketch the proofs of each.

Proof of Lemma 4.2: Using (4.3), one has

$$
\frac{d}{dt}|\xi^\varepsilon_t|^2 = 2\left\{(\xi^\varepsilon_t)' A^{\mu^\varepsilon_t}\xi^\varepsilon_t + (\xi^\varepsilon_t)'\left[l_2^{\mu^\varepsilon_t} + \sigma^{\mu^\varepsilon_t} w^\varepsilon_t\right]\right\},
$$

which by Assumption $(A.m)$,

$$
\leq 2\left\{-c_A|\xi^\varepsilon_t|^2 + (\xi^\varepsilon_t)'\left[l_2^{\mu^\varepsilon_t} + \sigma^{\mu^\varepsilon_t} w^\varepsilon_t\right]\right\},
$$

which, using the general inequality $a\cdot b \leq \frac{|a|^2}{2} + \frac{|b|^2}{2}$,

$$
\begin{aligned}
&\leq 2\left\{-\frac{c_A}{2}|\xi^\varepsilon_t|^2 + \frac{\left|l_2^{\mu^\varepsilon_t} + \sigma^{\mu^\varepsilon_t} w^\varepsilon_t\right|^2}{2c_A}\right\}\\
&\leq -c_A|\xi^\varepsilon_t|^2 + \frac{2\max_m |l^m_2|^2}{c_A} + \frac{2c_\sigma^2|w^\varepsilon_t|^2}{c_A}.
\end{aligned}
$$

Solving this ordinary differential inequality with $|\xi^\varepsilon_0|^2 = |x|^2$, one finds

$$
|\xi^\varepsilon_t|^2 \leq e^{-c_A t}|x|^2 + e^{-c_A t}\frac{2\max_m |l^m_2|^2}{c_A}\int_0^t e^{c_A r}\,dr + \frac{2c_\sigma^2}{c_A}\int_0^t e^{c_A(r-t)}|w^\varepsilon_r|^2\,dr,
$$

which yields the result.

Proof of Lemma 4.3: This proof is similar in form to the proof just above. Let $Q_t \doteq \int_0^t |\xi^\varepsilon_r|^2\,dr$. Then, following similar steps,

$$
\begin{aligned}
\dot{Q}_t &= |\xi^\varepsilon_t|^2 = |x|^2 + \int_0^t 2(\xi^\varepsilon_r)'\left[A^{\mu^\varepsilon_r}\xi^\varepsilon_r + l_2^{\mu^\varepsilon_r} + \sigma^{\mu^\varepsilon_r} w^\varepsilon_r\right]dr\\
&\leq |x|^2 + 2\left\{-c_A\int_0^t |\xi^\varepsilon_r|^2\,dr + \int_0^t (\xi^\varepsilon_r)'\left[l_2^{\mu^\varepsilon_r} + \sigma^{\mu^\varepsilon_r} w^\varepsilon_r\right]dr\right\}\\
&\leq |x|^2 - c_A Q + \frac{2}{c_A}\int_0^t \max_m |l^m_2|^2\,dr + \frac{2c_\sigma^2}{c_A}\int_0^t |w^\varepsilon_r|^2\,dr.
\end{aligned}
$$

Solving the ordinary differential inequality, one finds

$$
Q_t \leq \int_0^t e^{c_A(r-t)}|x|^2\,dr + \frac{2}{c_A}\int_0^t e^{c_A(r-t)}\int_0^r \max_m |l^m_2|^2\,d\rho\,dr
$$

$$+\frac{2c_\sigma^2}{c_A}\int_0^t e^{c_A(r-t)}\int_0^r |w_\rho^\varepsilon|^2\,d\rho\,dr$$

$$\le \frac{1}{c_A}|x|^2 + \frac{2\max_m |l_2^m|^2}{c_A^2}t + \frac{2c_\sigma^2}{c_A^2}\|w^\varepsilon\|_{L_2(0,t)}^2 \tag{11.1}$$

12. Appendix B. We prove Lemma 4.4. First, note that for any semiconvex function, the subdifferential and the Clarke generalized gradient are identical. (To see this, consider any semiconvex function, $f : I\!R^n \to I\!R$ and $x \in I\!R^n$. Note that $\tilde{f}(y) \doteq f(y) + c_\rho|y-x|^2$ is convex over a ball, $B_\rho = B_\rho(0)$, for appropriate c_ρ, and has the same subdifferential as f at x in B_ρ. The subdifferential of a convex function is identical to the Clarke generalized gradient [7].) Let $\mathcal{A}$ be the set of points where grad $\tilde{\tilde{V}}(x)$ exists (whose complement has measure zero by Rademacher's Theorem, c.f., [47]). For any set $S \subseteq I\!R^n$, let $\langle S\rangle$ denote the convex hull of S. Then note, from Theorem 2.5.1 of [7], that

$$\partial\tilde{\tilde{V}}(x) = \left\langle\left\{\lim \operatorname{grad}\tilde{\tilde{V}}(x_i)\,\middle|\, x_i \to x,\ x_i \in \mathcal{A},\ \lim\operatorname{grad}\tilde{\tilde{V}}(x_i)\exists\right\}\right\rangle, \tag{12.1}$$

where this indicates the convex hull of the set of limits points of grad $\tilde{\tilde{V}}(x_i)$ for sequences of points in $\mathcal{A}$ approaching x such that the limit exists.

Now we proceed to the proof. Fix any $\rho \in (0,\infty)$. By the local Lipschitz behavior of semiconvex functions (c.f., [19]), there exists $R < \infty$ such that

$$\operatorname{grad}\tilde{\tilde{V}}(x) \in B_R \quad \forall\, x \in B_{\rho+1}\cap\mathcal{A}. \tag{12.2}$$

Fix any $x_0 \in B_\rho$. Let $N \in \mathbf{N}$ (the set of natural numbers), and let $\Delta_N \doteq \{-N, -(N-1), \dots N-1\}$. Let $\Delta_N^n \doteq [\Delta_N]^n$ where the superscript n on the right-hand side denotes outer product, that is Δ_N^n is the set of vectors of length n of elements from Δ_N. The cardinality of Δ_N^n is $M_N \doteq (2N)^n$. For any $i \in \Delta_N^n$, let

$$S_i = \frac{R}{N}i + [0, R/N]^n, \tag{12.3}$$

where again the n superscript here indicates outer product. That is, the right-hand side is the cube with corner at iR/N and side-length R/N. Note that these cover the cube of half-length R, centered at the origin, and that that cube is chosen to cover B_R, which contains the gradients. Let $\delta \in (0,1]$. For each $i \in \Delta_N^n$, let

$$\mathcal{D}_i^\delta \doteq \{x \in B_\delta(x_0)\cap\mathcal{A}, \mid \operatorname{grad}\tilde{\tilde{V}}(x) \in S_i\}. \tag{12.4}$$

Also, let $L \in \mathbf{N}$ (with L to be chosen below), and let $\Delta_L^+ \doteq \{0, 1/L, 2/L, \dots 1\}$ Then, let $\vec{w}$ denote a vector of elements in Δ_L^+, with entries, w_i, indexed by $i \in \Delta_N^n$.

$$\Lambda_{L,N} \doteq \left\{\vec{w}\,\middle|\, w_i \in \Delta_L^+\ \ \forall i \in \Delta_N^n,\ \sum_{i\in\Delta_N^n} w_i = 1\right\}.$$

Fix any $\varepsilon \in (0,1]$. Let $\lceil a$ denote the smallest integer greater than or equal to a. It is easy to see that for any $L \ge \bar{L} \doteq \frac{M_N}{\varepsilon} = \lceil\frac{(2N)^n}{\varepsilon}$,

$$\min_{\vec{w}\in\Lambda_{L,N}} \sum_{i\in\Delta_N^n} |w_i - \mu_i| \le \varepsilon \quad \forall\,\vec{\mu} \in I\!R^{M_N} \text{ s.t. } \sum_{i\in\Delta_N^n}\mu_i = 1 \text{ and } \mu_i \in [0,1]\ \forall i \in \Delta_N^n. \tag{12.5}$$

For any such μ, define the mapping $\vec{w}^0$ by

$$\vec{w}^0(\vec{\mu}) \in \operatorname*{argmin}_{\vec{w}\in\Lambda_{L,N}} \sum_{i\in\Delta_N^n} |w_i - \mu_i|. \tag{12.6}$$

Define

$$\bar{z}^\delta = \bar{z}^\delta(x_0) \doteq \int_{B_\delta\cap\mathcal{A}_{x_0}} g^\delta(y)\,\mathrm{grad}\,\tilde{\tilde{V}}(x_0-y)\,dy \tag{12.7}$$

where $\mathcal{A}_{x_0}$ is the set of points, y, such that $\mathrm{grad}\,\tilde{\tilde{V}}(x_0-y)$ exists (and we note as before that $m(\mathcal{A}_{x_0}^c)=0$), and g^δ is the mollifier given at the top of the proof of Theorem 4.1. Note that

$$\bar{z}^\delta = [\mathrm{grad}\,\tilde{\tilde{V}}]^\delta(x_0), \tag{12.8}$$

i.e., $\bar{z}^\delta$ is the δ-mollification of $\mathrm{grad}\,\tilde{\tilde{V}}$ evaluated at x_0, and we note that this exists for all $x_0 \in I\!R^n$. Here we interject a lemma:

Lemma 12.1. *$\bar{z}^\delta(x) = \mathrm{grad}[\tilde{\tilde{V}}^\delta](x)$ for all $x \in I\!R^n$ and all $\delta > 0$.*

Proof. Let $u \in I\!R^n$, $|u|=1$. Let the directional derivative of any function, f, at point x_0 in direction u be denoted by $f_u(x_0)$. We have

$$\begin{aligned} u\cdot\mathrm{grad}[\tilde{\tilde{V}}^\delta](x_0) &= [\tilde{\tilde{V}}^\delta]_u(x_0)\\ &= \lim_{h\to 0}\frac{\tilde{\tilde{V}}^\delta(x_0+hu)-\tilde{\tilde{V}}^\delta(x_0)}{h},\end{aligned}$$

which by the definition of $\tilde{\tilde{V}}^\delta$

$$= \lim_{h\to 0}\int_{B_\delta\cap\mathcal{A}_{x_0}} \frac{g^\delta(y)[\tilde{\tilde{V}}(x_0+hu-y)-\tilde{\tilde{V}}(x_0-y)]}{h}\,dy$$

and by the Bounded Convergence Theorem,

$$= \int_{B_\delta\cap\mathcal{A}_{x_0}} g^\delta(y)\lim_{h\to 0}\frac{\tilde{\tilde{V}}(x_0+hu-y)-\tilde{\tilde{V}}(x_0-y)}{h}\,dy$$

which, since the gradient exists for $y \in \mathcal{A}_{x_0}$,

$$\begin{aligned} &= \int_{B_\delta\cap\mathcal{A}_{x_0}} g^\delta(y)\,\mathrm{grad}\,\tilde{\tilde{V}}(x_0-y)\cdot u\,dy\\ &= u\cdot\int_{B_\delta\cap\mathcal{A}_{x_0}} g^\delta(y)\,\mathrm{grad}\,\tilde{\tilde{V}}(x_0-y)\,dy = u\cdot[\mathrm{grad}\,\tilde{\tilde{V}}]^\delta(x_0).\end{aligned}$$

Since this holds for all $|u|=1$, $x_0\in I\!R^n$ and $\delta>0$, one has the desired result. ☐

Now, let $\mu_i^\delta = \int_{x_0-\mathcal{D}_i^\delta} g^\delta(y)\,dy$, and then let $\vec{\mu}^\delta$ denote the vector of elements μ_i^δ indexed by $i\in\Delta_N^n$. Next, let $\vec{w}^\delta \doteq \vec{w}^0(\vec{\mu}^\delta)$ where mapping $\vec{w}^0$ is given by (12.6), and let the elements of $\vec{w}^\delta$ be w_i^δ. For each $i \in \Delta_N^n$ such that $w_i^\delta > 0$, choose any $x_i^\delta \in \mathcal{D}_i^\delta \cap \mathcal{A}$. Define

$$\hat{z}^\delta \doteq \sum_{i\in\Delta_N^n} w_i^\delta\,\mathrm{grad}\,\tilde{\tilde{V}}(x_i^\delta). \tag{12.9}$$

(We will use the fact that this is a finite convex combination with rational weights with denominators L.) It is helpful to note that $x_0 - y \in \mathcal{D}_i^\delta$ is equivalent to $y \in x_0 - \mathcal{D}_i^\delta$. One now has

$$\begin{aligned}
|\bar{z}^\delta - \hat{z}^\delta| &= \left| \int_{B_\delta(0)\cap(x_0-\mathcal{A}_{x_0})} g^\delta(y) \operatorname{grad}\tilde{\tilde{V}}(x_0-y)\,dy - \sum_{i\in\Delta_N^n} w_i^\delta \operatorname{grad}\tilde{\tilde{V}}(x_i^\delta) \right| \\
&\le \sum_{i\in\Delta_N^n} \left| \int_{[x_0-\mathcal{D}_i^\delta]\cap\mathcal{A}_{x_0}} g^\delta(y) \operatorname{grad}\tilde{\tilde{V}}(x_0-y)\,dy - w_i^\delta \operatorname{grad}\tilde{\tilde{V}}(x_i^\delta) \right| \\
&\le \sum_{i\in\Delta_N^n} \left\{ \left| \int_{[x_0-\mathcal{D}_i^\delta]\cap\mathcal{A}_{x_0}} g^\delta(y)\,dy - w_i^\delta \right| \left| \operatorname{grad}\tilde{\tilde{V}}(x_i^\delta) \right| \right. \\
&\quad \left. + \int_{[x_0-\mathcal{D}_i^\delta]\cap\mathcal{A}_{x_0}} g^\delta(y) \left| \operatorname{grad}\tilde{\tilde{V}}(x_i^\delta) - \operatorname{grad}\tilde{\tilde{V}}(x_0-y) \right| dy \right\},
\end{aligned}$$

which by (12.2), (12.5), (12.6) and the definition of $\vec{w}^\delta$,

$$\le \varepsilon R + \sum_{i\in\Delta_N^n} \int_{[x_0-\mathcal{D}_i^\delta]\cap\mathcal{A}_{x_0}} g^\delta(y) \left| \operatorname{grad}\tilde{\tilde{V}}(x_i^\delta) - \operatorname{grad}\tilde{\tilde{V}}(x_0-y) \right| dy,$$

which by (12.3) and the definition of the x_i^δ,

$$\begin{aligned}
&\le \varepsilon R + \frac{\sqrt{n}R}{N} \sum_{i\in\Delta_N^n} \int_{[x_0-\mathcal{D}_i^\delta]\cap\mathcal{A}_{x_0}} g^\delta(y)\,dy \\
&= \varepsilon R + \frac{\sqrt{n}R}{N} \int_{B_\delta(0)\cap(x_0-\mathcal{A}_{x_0})} g^\delta(y)\,dy = \varepsilon R + \frac{\sqrt{n}R}{N}
\end{aligned}$$

which for $N \ge \overline{N} = \lceil(\sqrt{n}/\varepsilon)$,

$$\le 2R\varepsilon. \tag{12.10}$$

Note that (12.10) is true independent of $\delta \le 1$, $\vec{w}^\delta$, $\vec{x}^\delta$ chosen as $x_i^\delta \in \mathcal{D}_i^\delta \cap \mathcal{A}$ and $\vec{w}^\delta = \vec{w}^0(\vec{\mu}^\delta)$ for all $L \ge \bar{L}$.

Note that the weights w_i^δ depend on δ. We would like to have constant weights, and so will choose all weights to be $1/L$, with possible duplication of the x_i^δ points. Consider any mapping $\bar{i} : \{0,1,\ldots L\} \to \Delta_N^n$, and let $N_i(j) \doteq \#\{j \in]0,L[\,|\, \bar{i}(j) = i\}$, where $]l,k[\doteq \{l, l+1, l+2, \ldots, k\}$ for any $l,k \in \mathbf{N}$ with $l \le k$. Now choose the mapping $\bar{i}$ such that

$$\frac{N_i(j)}{L} = w_i^\delta \quad \forall\, i \in \Delta_N^n.$$

Let $\hat{x}_j^\delta \doteq x_{\bar{i}(j)}^\delta$ for all $j \in \{0,1,\ldots L\}$. Then,

$$\sum_{j\in]0,L[} \frac{1}{L} \operatorname{grad}\tilde{\tilde{V}}(\hat{x}_j^\delta) = \hat{z}^\delta. \tag{12.11}$$

Also, by (12.10),

$$\left| \sum_{j\in]0,L[} \frac{1}{L} \operatorname{grad}\tilde{\tilde{V}}(\hat{x}_j^\delta) - \bar{z}^\delta \right| \le 2R\varepsilon. \tag{12.12}$$

Further, note that $\operatorname{grad}\tilde{\tilde{V}}(\hat{x}_j^\delta) \in \bar{B}_R$ for all j. Consider sequence $\delta_k \downarrow 0$. By the compactness of $\overline{B}_R$, there exists a subsequence, $\{\delta_{k_\kappa}\}$ and $\{z_j \,|\, j \in]0,L[\} \subset \bar{B}_R$ such that

$$\operatorname{grad}\tilde{\tilde{V}}(\hat{x}_j^{\delta_{k_\kappa}}) \to z_j \quad \forall\, j \in]0,L[.$$

Consequently,

$$\sum_{j\in]0,L[}\frac{1}{L}\operatorname{grad}\widetilde{\widetilde{V}}(\hat{x}_j^{\delta_{k_\kappa}})\to\sum_{j\in]0,L[}\frac{1}{L}z_j\doteq\tilde{z}$$

$$\in\langle\{\lim\operatorname{grad}\widetilde{\widetilde{V}}(x_i)\,|\,x_i\to x,\,x_i\in\mathcal{A}\}\rangle,$$

which by (12.1)

$$=\partial\widetilde{\widetilde{V}}(x_0). \tag{12.13}$$

We can now put the above together so as to obtain the desired result. Suppose we are given any $\rho\in(0,\infty)$ and any $x_0\in B_\rho$. Let $R=R(\rho)\in(0,\infty)$ be as in (12.2). Let $\delta\in(0,1)$. Fix any $\bar{\varepsilon}>0$, and let $\varepsilon=\frac{\bar{\varepsilon}}{4R}$. Let $N=\bar{N}=\lceil(\sqrt{n}/\varepsilon)$, and let $L=\bar{L}=\lceil\frac{(2N)^n}{\varepsilon}$. Then, by Lemma 12.1, (12.11) and (12.12),

$$\left|\operatorname{grad}[\widetilde{\widetilde{V}}^{\delta}](x_0)-\sum_{j\in]0,L[}\frac{1}{\bar{L}}\operatorname{grad}\widetilde{\widetilde{V}}(\hat{x}_j^{\delta})\right|=|\bar{z}^{\delta}-\hat{z}^{\delta}|\le 2R\varepsilon=\bar{\varepsilon}/2, \tag{12.14}$$

where this is independent of $\delta\in(0,1)$ and sequence $\{x_i^\delta\}$ such that $x_i^\delta\in\mathcal{D}_i^\delta$ for all $i\in\Delta_N^n$. However, by (12.13), there exists $\tilde{z}\in\partial\widetilde{\widetilde{V}}(x_0)$ and $K<\infty$ such that for $\kappa\ge K$ (i.e., $\delta_{k_\kappa}\le\delta_{k_K}$), one has

$$|\hat{z}^{\delta_{k_\kappa}}-\tilde{z}|<\bar{\varepsilon}/2. \tag{12.15}$$

Combining (12.14) and (12.15), we see that given $\bar{\varepsilon}>0$, there exists $\tilde{z}\in\partial\widetilde{\widetilde{V}}(x_0)$ and $K<\infty$ such that for $\kappa\ge K$, and letting $\delta\le\delta_{k_\kappa}$,

$$\left|\operatorname{grad}[\widetilde{\widetilde{V}}^{\delta}](x_0)-\tilde{z}\right|<\bar{\varepsilon}.$$

Now, suppose that instead of a single point, x_0, we had a *finite* set, say $X=\left\{x_0^\lambda\,|\,\lambda\in\{0,1,\dots\Lambda\}\right\}$. Then, by taking further subsequences, we obtain set $Z=\left\{\tilde{z}^\lambda\,|\,\lambda\in\{0,1,\dots\Lambda\}\right\}$ and subsequence $\delta_{k_\kappa}\downarrow 0$ such that for all $\delta\le\delta_{k_\kappa}$ with $\kappa\ge K$ for some $K<\infty$,

$$\left|\operatorname{grad}[\widetilde{\widetilde{V}}^{\delta}](x_0^\lambda)-\tilde{z}^\lambda\right|<\bar{\varepsilon}$$

for all $\lambda\in\{0,1,\dots\Lambda\}$, which completes the proof of Lemma 4.4.

13. Appendix C. In this appendix, we prove Theorem 6.6. In this proof, we will be letting $\tau\downarrow 0$ in homogeneous system (6.5). This limit is taken at this point merely to prove our theorem (Th. 6.6) regarding the matrizant for the homogeneous system (6.4); it is not related to the argument for the proof of the main result of Section 6, where we take a similar limit for the overall convergence result. To obtain our result, we need the following intermediary theorem. The theorem is inequality (65) of [31], and so no proof is included; to specialize to this case, one simply takes $\sigma^m=0$ for all m, in which case one can take $w\equiv 0$ as optimal.

THEOREM 13.1. *There exists $K_0<\infty$ such that*

$$|\hat{\xi}_t-\hat{\bar{\xi}}_t|^2\le K_0(1+|x|^2)\tau\quad\forall\,t\in[0,\infty),\ \forall x\in I\!R^n.$$

Next, we include the following restatement of (6.16), which we note does not depend on anything not proven before Theorem 6.6.

LEMMA 13.2.

$$\left|\exp\left[\int_s^t A^{\mu_r}\,dr\right] - \exp\left[\int_s^t A^{\overline{\mu}_r}\,dr\right]\right| \le M^2\overline{A}e^{M^2\overline{A}-c_A(t-s)}\tau$$

for all $0 \le s \le t < \infty$*, where* $\overline{A} = \max_{m\in\mathcal{M}} |A^m|$.

Let

$$\hat{\xi}_t^0 \doteq \exp\left[\int_0^t A^{\mu_r}\,dr\right] x \quad \forall\, t \in [0,\infty),\, \forall x \in I\!R^n. \tag{13.1}$$

LEMMA 13.3. *There exists* $\widehat{K}_0 < \infty$ *such that*

$$|\hat{\xi}_t^0 - \hat{\overline{\xi}}_t| \le \widehat{K}_0|x|\tau \quad \forall\, t \in [0,\infty),\, \forall x \in I\!R^n.$$

Proof. This follows immediately from (6.6), (13.1) and Lemma 13.2 □

Now we have the material necessary to prove Theorem 6.6, and this follows.

Proof.(proof of Th. 6.6) Fix any $\tau > 0$. Combining Theorem 13.1 and Lemma 13.3, we see that for any $t \in [0,\infty)$ and any $x \in I\!R^n$,

$$|\hat{\xi}_t - \hat{\xi}_t^0| \le K_0\sqrt{1+|x|^2}\sqrt{\tau} + \widehat{K}_0|x|\tau.$$

Since this is true for all $\tau > 0$, one has $\hat{\xi}_t = \hat{\xi}_t^0$ for all $t \in [0,\infty)$ and all $x \in I\!R^n$. □

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, *The max-plus finite element method for optimal control problems: further approximation results,* Proc. joint 44th IEEE Conf. on Decision and Control and European Control Conf. 2005 (2005).

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

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

[5] F. Camilli, M. Falcone, P. Lanucara and A. Seghini, *A Domain decomposition method for Bellman equations,* Contemp. Math., 180 (1994), 477-483.

[6] E. Carlini, M. Falcone and R. Ferretti, *An Efficient algorithm for Hamilton-Jacobi equations in high dimensions,* Computing and Visualization Sci., 7 (2004), pp. 15–29.

[7] F.H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley, 1983.

[8] G. Cohen, S. Gaubert and J.-P. Quadrat, *Duality and separation theorems in idempotent semimodules*, Linear Algebra and Applications, 379 (2004), pp. 395–422.

[9] G. Collins and W.M. McEneaney, *Min-plus eigenvector methods for nonlinear H_∞ problems with active control,* Optimal Control, Stabilization and Nonsmooth Analysis, LNCIS, Vol. 301, M.S. de Queiroz; M. Malisoff; P. Wolenski (Eds.), Springer-Verlag (2004), pp. 101–120.

[10] M. G. Crandall, H. Ishii and P.-L. Lions, *Uniqueness of viscosity solutions of Hamilton–Jacobi equations revisited,* J. Math. Society Japan, 39 (1987), pp. 581–595.

[11] M. G. Crandall and P.-L. Lions, *Viscosity solutions of Hamilton–Jacobi equations,* Trans. of the Amer. Math. Soc., 277 (1983), pp. 1–42.

[12] M. G. Crandall and P.-L. Lions, *On existence and uniqueness of solutions of Hamilton–Jacobi equations,* Nonlinear Anal. Theory and Appl., 10 (1986), pp. 353–370.

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

[14] 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.

[15] P. Dupuis and A. Szpiro, *Convergence of the optimal feedback policies in a numerical method for a class of deterministic optimal control problems,* SIAM J. Control and Optim., 40 (2001), pp. 393–420.

[16] P. Dupuis and A. Szpiro, *Second order numerical methods for first order Hamilton-Jacobi equations,* SIAM J. on Num. Analysis, 40 (2002), pp. 113–118.
[17] M. Falcone and R. Ferretti, *Convergence analysis for a class of high-order semi-lagrangian advection schemes,* SIAM J. Num. Analysis, 35 (1998), pp. 909–940.
[18] W.H. Fleming, *Max-plus stochastic processes*, Applied Math. and Optim., 49 (2004), pp. 159–181.
[19] W.H. Fleming, *Functions of Several Variables*, Springer-Verlag, New York, 1977.
[20] 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), pp. 683–710.
[21] W.H. Fleming and H.M. Soner, *Controlled Markov Process and Viscosity Solutions, 2nd Ed.* Springer-Verlag, New York, 2006.
[22] F. John, *Partial Differential Equations*, Springer-Verlag, New York, 1978.
[23] J.W. Helton and M.R. James, *Extending H_∞ Control to Nonlinear Systems, Control of Nonlinear Systems to Achieve Performance Objectives*, Advances in Design and Control, SIAM, Philadelphia, 1999.
[24] M.R. James and W.M. McEneaney, *Max-plus approximation methods in partially observed H_∞ control,* 38th IEEE Conf. on Decision and Control, (1999), pp. 3011–3016.
[25] V.N. Kolokoltsov and V.P. Maslov, *Idempotent Analysis and Its Applications,* Kluwer, Dordrecht, 1997.
[26] H.J. Kushner and P. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer-Verlag, New York, 1992.
[27] G.L. Litvinov, V.P. Maslov and G.B. Shpiz, *Idempotent functional analysis: An algebraic approach,* Mathematical Notes, 69 (2001), pp. 696–729.
[28] V.P. Maslov, *On a new principle of superposition for optimization problems,* Russian Math. Surveys, 42 (1987), pp. 43–54.
[29] W.M. McEneaney and A. Deshpande, *Payoff Suboptimality and Errors in Value Induced by Approximation of the Hamiltonian,* Proc. IEEE CDC 2008.
[30] W.M. McEneaney, A. Deshpande and S. Gaubert, *Curse-of-Complexity Attenuation in the Curse-of-Dimensionality-Free Method for HJB PDEs,* Proc. ACC 2008.
[31] W.M. McEneaney and L.J. Kluberg, "Convergence rate for a curse-of-dimensionality-free method for a class of HJB PDEs", Submitted to SIAM J. Control and Optim.
[32] W.M. McEneaney, "A curse-of-dimensionality-free numerical method for solution of certain HJB PDEs", SIAM J. on Control and Optim., Vol. 46 (2007), 1239–1276.
[33] W.M. McEneaney, "Curse-of-dimensionality free method for Bellman PDEs with semiconvex Hamiltonians" Proc. 45th IEEE Conf. on Decision and Control, San Diego (2006).
[34] W.M. McEneaney, *Max-Plus Methods for Nonlinear Control and Estimation*, Birkhauser Systems and Control Series, Boston, 2006.
[35] W.M. McEneaney, "A Curse-of-dimensionality-free numerical method for a class of HJB PDEs", Proc. 16th IFAC World Congress (2005).
[36] W.M. McEneaney, *Max-plus summation of Fenchel-transformed semigroups for solution of nonlinear Bellman equations,* Systems and Control Letters, 56 (2007), pp. 255–264.
[37] W.M. McEneaney and P.M. Dower, *A max-plus affine power method for approximation of a class of mixed l_∞/l_2 value functions,* Proc. 42nd IEEE Conf. on Dec. and Control, Maui (2003), pp. 2573–2578.
[38] W.M. McEneaney, *Error analysis of a max-plus algorithm for a first-order HJB equation,* Stochastic Theory and Control, Proceedings of a Workshop held in Lawrence, Kansas, October 18-20, 2001, Lecture Notes in Control and Information Sciences, Springer-Verlag, B.Pasik-Duncan (Ed.) (2002), pp. 335–352.
[39] W.M. McEneaney and M. Horton, *Max-plus eigenvector representations for nonlinear H_∞ value functions,* 37th IEEE Conf. on Decision and Control (1998), pp. 3506–3511.
[40] W.M. McEneaney, *A Uniqueness result for the Isaacs equation corresponding to nonlinear H_∞ control,* Math. Controls, Signals and Systems, 11 (1998), pp. 303–334.
[41] W. M. McEneaney and M. V. Day, *Characteristic characterization of viscosity supersolutions corresponding to nonlinear H_∞ control,* 13^{th} World Congress of International Federation of Automatic Control, (1996), pp. 401–406.
[42] A.A. Melikyan, *Generalized Characteristics of First Order PDEs: Applications in Optimal Control and Differential Games*, Birkhauser, Boston, 1998.
[43] R.T. Rockafellar, *Conjugate Duality and Optimization*, Regional conference Series in Applied Mathematics 16, SIAM (1974).
[44] R.T. Rockafellar and R.J. Wets, *Variational Analysis*, Springer-Verlag, New York, 1997.
[45] P. Soravia, *H_∞ control of nonlinear systems: differential games and viscosity solutions,* SIAM J. Control and Optim., 34 (1996), pp. 1071–1097.
[46] A.J. van der Schaft, *L_2-gain analysis of nonlinear systems and nonlinear state feedback H_∞ control,* IEEE Trans. Auto. Control, 37 (1992), pp. 770–784.
[47] W.P. Ziemer, *Weakly Differentiable Functions,* Springer-Verlag, 1989.