

4

A Note on Stochastic Dissipativeness

Vivek S. Borkar *Sanjoy K. Mitter*

Abstract

In this paper we present a stochastic version of Willems' ideas on Dissipativity and generalize the dissipation inequality to Markov Diffusion Processes. We show the relevance of these ideas by examining the problem of Ergodic Control of partially observed diffusions.

4.1 Introduction

In [8, 9], Willems introduced the notion of a dissipative dynamical system with associated ‘supply rate’ and ‘storage function,’ with a view to building a Lyapunov-like theory of input-output stability for deterministic control systems. In this article, we extend these notions to stochastic systems, specifically to controlled diffusions. This makes contact with the ergodic control problem for controlled diffusions ([2], Chapter VI) and offers additional insight into the latter. In particular, it allows us to obtain a “martingale dynamic programming principle” for ergodic control under partial observations in the spirit of Davis and Varaiya [6]. So far this has been done only in special cases using a vanishing discount limit, see Borkar [3, 4, 5].

The next section introduces the notation and key definitions. Section 3 considers the links with ergodic control with complete or partial observations.

4.2 Notation and Definitions

Our controlled diffusion will be a $d \geq 1$ dimensional process

$$X(\cdot) = [X_1(\cdot), \dots, X_d(\cdot)]^T$$

satisfying the stochastic differential equation

$$X(t) = X_0 + \int_0^t m(X(s), u(s)) ds + \int_0^t \sigma(X(s)) dW(s) \quad , \quad t \geq 0 \quad . \quad (4.1)$$

Here,

- (i) for a prescribed compact metric ‘control’ space U ,

$$m(\cdot, \cdot) = [m_1(\cdot, \cdot), \dots, m_d(\cdot, \cdot)]^T : \mathbf{R}^d \times U \rightarrow \mathbf{R}^d$$

is continuous and Lipschitz in its first argument uniformly with regards to the second argument,

- (ii) $\sigma(\cdot) = [[\sigma_{ij}(\cdot)]]_{1 \leq i, j \leq d} : \mathbf{R}^d \rightarrow \mathbf{R}^{d \times d}$ is Lipschitz,
 (iii) X_0 is an \mathbf{R}^d -valued random variable with a prescribed law π_0 ,
 (iv) $W(\cdot) = [W_1(\cdot), \dots, W_d(\cdot)]^T$ is a d -dimensional standard Brownian motion independent of X_0 and
 (v) $u(\cdot) : [0, \infty) \rightarrow U$ is a control process with measurable sample paths satisfying the nonanticipativity condition: for $t \geq s$, $W(t) - W(s)$ is independent of $u(\tau)$, $W(\tau)$, $\tau \leq s$ and X_0 .

We shall consider a weak formulation of (4.1), i.e., we look for $(X(\cdot), u(\cdot), W(\cdot), X_0)$ on *some* probability space so that (4.1) holds. See [2], Chapter 1, for an exposition of the weak formulation. In particular, letting \mathcal{F}_t denote the right-continuous completion of $\sigma(X(s), s \leq t)$ for $t \geq 0$, it is shown in Theorem 2.2, pp. 18-19, [2] that it suffices to consider $u(\cdot)$ adapted to $\langle \mathcal{F}_t \rangle$, i.e., $u(\cdot)$ of the

form $u(t) = f_t(X([0, t]))$ where $X([0, t])$ denotes the restriction of $X(\cdot)$ to $[0, t]$ and $f_t : C([0, t]; R^d) \rightarrow U$ are measurable maps. If in addition $u(t) = v(X(t))$, $t \geq 0$, for a measurable $v : R^d \rightarrow U$, we call $u(\cdot)$ (or, by abuse of terminology, the map $v(\cdot)$ itself) a Markov control.

We shall also be interested in the partially observed case ([2], Chapter V). Here one has an associated observation process $Y(\cdot)$ taking values in R^m ($m \geq 1$), given by:

$$Y(t) = \int_0^t h(X(s)) ds + W'(t) \quad , \quad t \geq 0 \quad . \tag{4.2}$$

where $h : R^d \rightarrow R^m$ is continuous and $W'(\cdot)$ an m -dimensional standard Brownian motion independent of $W(\cdot), X_0$. Let $\langle \mathcal{G}_t \rangle$ denote the right-continuous completion of $\sigma(Y(s), s \leq t)$ for $t \geq 0$. We say that $u(\cdot)$ is strict sense admissible if it is adapted to $\langle \mathcal{G}_t \rangle$.

For the completely observed control problem where we observe $X(\cdot)$ directly and $u(t) = f_t(X([0, t]))$ for $t \geq 0$, we define

DEFINITION 4.1

A measurable function $V : R^d \rightarrow R$ is said to be a storage function associated with a supply rate function $g \in C(R^d \times U)$ if it is bounded from below and $V(X(t)) + \int_0^t g(X(s), u(s)) ds$, $t \geq 0$, is an $\langle \mathcal{F}_t \rangle$ -super martingale for all $(X(\cdot), u(\cdot))$ satisfying (4.1) as above. □

The storage function need not be unique. For example, we get another by adding a constant. For $g, (X(\cdot), u(\cdot))$ as above, let

$$V_c(x) = \sup_{u(\cdot)} \sup_{\tau} E \left[\int_0^{\tau} g(X(s), u(s)) ds / X_0 = x \right] \quad , \quad x \in R^d \quad ,$$

where the first supremum is over all bounded $\langle \mathcal{F}_t \rangle$ -stopping times and the second supremum is over all $\langle \mathcal{F}_t \rangle$ -adapted $u(\cdot)$. Since $\tau = 0$ is a stopping time, $V_c(\cdot) \geq 0$.

LEMMA 4.1

If $V_c(x) < \infty$ for all x , it is the least nonnegative storage function associated with g . □

Proof In the following, τ denotes an $\langle \mathcal{F}_t \rangle$ -stopping time. For $t \geq 0$,

$$\begin{aligned} V_c(x) &\geq \sup_{u(\cdot)} \sup_{\tau \geq t} E \left[\int_0^{\tau} g(X(s), u(s)) ds / X_0 = x \right] \\ &= \sup_{u([0, t])} E \left[\int_0^t g(X(s), u(s)) ds \right. \\ &\quad \left. + \sup_{u(t+\cdot)} \sup_{\tau \geq t} E \left[\int_t^{\tau} g(X(s), u(s)) ds / X(t) \right] / X_0 = x \right] \quad , \end{aligned}$$

where the equality follows by a standard dynamic programming argument. Thus

$$V_c(x) \geq \sup_{u(\cdot)} E \left[\int_0^t g(X(s), u(s)) ds + V_c(X(t)) / X_0 = x \right] . \quad (4.3)$$

Now suppose $s \leq \tau \leq T < \infty$, where $T > 0$ is deterministic and s, τ are $\langle \mathcal{F}_t \rangle$ -stopping times. Then

$$\begin{aligned} & E \left[\int_0^\tau g(X(s), u(s)) ds + V_c(X(\tau)) / \mathcal{F}_s \right] \\ &= \int_0^s g(X(s), u(s)) ds + E \left[\int_s^\tau g(X(s), u(s)) ds + V_c(X(\tau)) / \mathcal{F}_s \right] . \end{aligned}$$

By Theorem 1.6, p. 13 of [2], the regular condition law of $(X_{s+\cdot}, u(s+\cdot))$ given \mathcal{F}_s is again the law of a pair $(\tilde{X}(\cdot), \tilde{u}(\cdot))$ satisfying (4.1) with initial condition $X(s)$, a.s. Therefore by (4.3), the above is less than or equal to

$$\int_0^s g(X(s), u(s)) ds + E \left[V_c(X(s)) / \mathcal{F}_s \right] = \int_0^s g(X(s), u(s)) ds + V_c(X(s)) .$$

It follows that

$$\int_0^t g(X(s), u(s)) ds + V_c(X(t)) , \quad t \geq 0 ,$$

is an $\langle \mathcal{F}_t \rangle$ -supermartingale. Thus $V_c(\cdot)$ is a storage function. If $F(\cdot)$ is another nonnegative storage function, we have, by the optional sampling theorem

$$\begin{aligned} F(x) &\geq E \left[\int_0^\tau g(X(s), u(s)) ds + F(X(\tau)) / X_0 = x \right] \\ &\geq E \left[\int_0^\tau g(X(s), u(s)) ds / X_0 = x \right] . \end{aligned}$$

for any bounded $\langle \mathcal{F}_t \rangle$ -stopping time τ . Therefore

$$F(x) \geq \sup_{u(\cdot)} \sup_{\tau} E \left[\int_0^\tau g(X(s), u(s)) ds / X_0 = x \right] = V_c(x) .$$

This completes the proof. □

LEMMA 4.2

If $V_c(\cdot) < \infty$, it can also be defined by

$$V_c(x) = \sup_{u(\cdot)} \sup_{t \geq 0} E \left[\int_0^t g(X(s), u(s)) ds / X_0 = x \right] .$$

□

Proof Let $\bar{V}_c(x)$ denote the R.H.S. above. Then clearly $V_c(x) \geq \bar{V}_c(x)$. On the other hand, an argument similar to that of the preceding lemma shows that for $t \geq r \geq 0$,

$$V(X(r)) + \int_0^r g(X(s), u(s)) ds \geq E \left[V(X(t)) + \int_0^t g(X(s), u(s)) ds / \mathcal{F}_r \right] ,$$

implying that

$$\int_0^t g(X(s), u(s)) ds + V(X(t)) , \quad t \geq 0 ,$$

is an $\langle \mathcal{F}_t \rangle$ -supermartingale. Thus $\bar{V}_c(\cdot)$ is a storage function. Clearly, $\bar{V}_c(\cdot) \geq 0$. Thus by the preceding lemma, $\bar{V}_c(\cdot) \geq V_c(\cdot)$ and hence $\bar{V}_c(\cdot) = V_c(\cdot)$. \square

For the partially observed control problem, the correct ‘state’ is $\pi_t \triangleq$ the regular conditional law of $X(t)$ given $\zeta_t \triangleq$ the right-continuous completion of $\sigma(Y(s), u(s), s \leq t)$, $t \geq 0$. Note that $\zeta_t = \mathcal{G}_t$ for strict sense admissible $u(\cdot)$, but we shall allow the so called wide-sense admissible $u(\cdot)$ of [7]. (See, also, [2], Chapter V.) Thus, in general, $\mathcal{G}_t \subset \zeta_t$. Let $\mathcal{P}(R^d)$ = the Polish space of probability measures on R^d with Prohorov topology. Viewing $\langle \pi_t \rangle$ as a $\mathcal{P}(R^d)$ -valued process, its evolution is given by the nonlinear filter, defined as follows: For $f : R^d \rightarrow R$ that are twice continuously differentiable with compact supports,

$$\pi_t(f) = \pi_0(f) + \int_0^t \pi_s(Lf(\cdot, u(s))) ds + \int_0^t \langle \pi_s(hf) - \pi_s(h)\pi_s(f), d\tilde{Y}(s) \rangle ,$$

$t \geq 0 , \quad (4.4)$

where:

- (i) $v(f) \triangleq \int f dv$ for $f \in C_b(R^d)$, $v \in \mathcal{P}(R^d)$,
- (ii) $Lf(x, u) = \frac{1}{2} \sum_{ijk} \sigma_{ik}(x)\sigma_{jk}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum_i m_i(x, u) \frac{\partial f}{\partial x_i}(x)$ for twice continuously differentiable $f : R^d \rightarrow R$,
- (iii) $\tilde{Y}(t) = Y(t) - \int_0^t \pi_s(h) ds$, $t \geq 0$, is an m -dimensional standard Brownian motion.

See [2], Chapter V, for a discussion of wellposedness and related issues for (4.4). In particular, for a given pair $(\tilde{Y}(\cdot), u(\cdot))$ of a Brownian motion and a wide sense admissible control, (4.4) has a unique solution if

- (i) $\sigma(\cdot)$ is nondegenerate, i.e., the least eigenvalue of $\sigma(\cdot)\sigma(\cdot)^T$ is uniformly bounded away from zero, and
- (ii) $h(\cdot)$ is twice continuously differentiable, bounded, with bounded first and second partial derivatives. (This can be relaxed — see [7].)

The partially observed control problem then is equivalent to the completely observed control problem of controlling the $\mathcal{P}(R^d)$ -valued process $\{\pi_t\}$ governed by (4.4), with wide sense admissible $u(\cdot)$. By analogy with Definition 4.1 above, we have

DEFINITION 4.2

A measurable function $\bar{V} : \mathcal{P}(R^d) \rightarrow R$ is said to be a storage function associated with the supply rate function $g \in C(\mathcal{P}(R^d) \times U)$ if it is bounded from below and

$$\bar{V}(\pi_t) + \int_0^t g(\pi_s, u(s)) ds, \quad t \geq 0$$

is a $\langle \zeta_t \rangle$ -supermartingale for all $\{\pi_t, u(t)\}_{t \geq 0}$ as above. \square

Define $V_p : \mathcal{P}(R^d) \rightarrow R$ by

$$V_p(\pi) = \sup_{u(\cdot)} \sup_{\tau} E \left[\int_0^{\tau} g(\pi_s, u(s)) ds / \pi_0 = \pi \right],$$

where the first supremum is over all bounded $\langle \zeta_t \rangle$ -stopping times τ and the second supremum is over all wide sense admissible $u(\cdot)$. Then the following can be proved exactly as in Lemmas 4.1 and 4.2.

LEMMA 4.3

If $V_p(\cdot) < \infty$, it is the least nonnegative storage function associated with supply rate g and permits the alternative definition:

$$V_p(\pi) = \sup_{u(\cdot)} \sup_{t \geq 0} E \left[\int_0^t g(\pi_s, u(s)) ds / \pi_0 = \pi \right]$$

where the outer supremum is over all wide sense admissible controls. \square

4.3 Connections to Ergodic Control

Let $k \in C_b(R^d \times U)$. The ergodic control problem seeks to maximize over admissible $u(\cdot)$ the reward

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \int_0^t E \left[k(X(s), u(s)) \right] ds. \quad (4.5)$$

Likewise, the ergodic control problem under partial observations is to maximize over all wide sense admissible $u(\cdot)$ the above reward, rewritten as

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \int_0^t E \left[\hat{k}(\pi_s, u(s)) \right] ds,$$

where $\hat{k}(\mu, u) = \mu(k(\cdot, u))$, $\mu \in \mathcal{P}(R^d)$, $u \in U$. Under suitable conditions, this problem can be shown to have an optimal stationary solution ([2], Chapter VI, [1]) with $u(\cdot)$ a Markov control. If $\sigma(\cdot)$ is nondegenerate (i.e., the least eigenvalue of $\sigma(\cdot)\sigma(\cdot)^T$ is uniformly bounded away from zero), then one can in fact have, under suitable hypotheses, a Markov control that is optimal for any initial law ([2], Chapter VI). Here our interest is in the ‘martingale dynamic programming principle’ elucidated in [6], [2], Chapter III, among other places, albeit for cost criteria other than ergodic. Let β (resp. $\hat{\beta}$) denote the optimal costs for the completely observed (resp., partially observed) ergodic control problem.

DEFINITION 4.3

A measurable map $\psi : R^d \rightarrow R$ is said to be a value function for the completely observed ergodic control problem if for all $(X(\cdot), u(\cdot))$ satisfying (4.1), the process

$$\psi(X(t)) + \int_0^t [k(X(s), u(s)) - \beta] ds \quad , \quad t \geq 0 \quad ,$$

is an $\langle \mathcal{F}_t \rangle$ -supermartingale and is a martingale if and only if $(X(\cdot), u(\cdot))$ is an optimal pair. □

DEFINITION 4.4

A measurable map $\bar{\psi} : \mathcal{P}(R^d) \rightarrow R$ is said to be a value function for the partially observed ergodic control problem if for all $(\pi_t, u(t))$, $t \geq 0$, as in (4.4), the process

$$\bar{\psi}(\pi_t) + \int_0^t [\pi_s(k(\cdot), u(s)) - \hat{\beta}] ds \quad , \quad t \geq 0 \quad ,$$

is a $\langle \zeta_t \rangle$ -supermartingale, and is a $\langle \zeta_t \rangle$ -martingale if and only if $\{\pi_t, u(t)\}$, $t \geq 0$, is an optimal pair. □

Proving a martingale dynamic programming principle in either case amounts to exhibiting a ψ (resp. $\bar{\psi}$) satisfying the above. The following lemmas establish a link with the developments of the preceding section.

LEMMA 4.4

- (a) If $\psi \geq 0$ is as in Definition (4.3), then ψ is a storage function for $g(\cdot, \cdot) = k(\cdot, \cdot) - \beta$.
 - (b) If $\bar{\psi} \geq 0$ is as in Definition (4.4), then $\bar{\psi}$ is a storage function for $g(\cdot, \cdot) = \hat{k}(\cdot, \cdot) - \hat{\beta}$.
-

This is immediate from the definitions. Note that if ψ is a value function, so is $\psi + c$ for any scalar c . Thus, in particular, it follows that there is a nonnegative value function whenever there is one that is bounded from below. This is the case for nondegenerate diffusions with ‘near-monotone’ $k(\cdot, \cdot)$, i.e., $k(\cdot, \cdot)$ satisfying

$$\liminf_{\|\cdot\| \rightarrow \infty} \inf_u k(u, u) > \beta \quad .$$

See [2], Chapter VI for details.

Going in the other direction, we have

LEMMA 4.5

- (a) If $V_c(\cdot) < \infty$ for $g(\cdot, \cdot) = k(\cdot, \cdot) - \beta$, then $V_c(X(t)) + \int_0^t (k(X(s), u(s)) - \beta) ds$, $t \geq 0$, is an $\langle \mathcal{F}_t \rangle$ -supermartingale for all $(X(\cdot), u(\cdot))$ as in (4.1). Furthermore, if $(X(\cdot), u(\cdot))$ is a stationary optimal solution and $V_c(X(t))$ is integrable under this stationary law, then the above process is in fact a martingale.

- (b) If $V_p(\cdot) < \infty$ for $g(\cdot, \cdot) = \hat{k}(\cdot, \cdot) - \hat{\beta}$, then $V_p(\pi_t) + \int_0^t (\hat{k}(\pi_s, u(s)) - \hat{\beta}) ds, t \geq 0$, is a $\langle \zeta_t \rangle$ -supermartingale for all $(\pi_t, u(t)), t \geq 0$, as in (4.4). Furthermore, if $(\pi_t, u(t)), t \geq 0$, is a stationary optimal solution and $V_p(\pi_t)$ is integrable under this stationary law, then the above process is in fact a martingale. \square

Proof We prove only (a), the proof of (b) being similar. The first claim is immediate. For stationary optimal $(X(\cdot), u(\cdot))$,

$$E[V(X(0))] \geq \int_0^t [E[k(X(s), u(s))] - \beta] ds + E[V(X(t))] .$$

Hence

$$0 \geq E[k(X(t), u(t))] - \beta .$$

But since $(X(\cdot), u(\cdot))$ are stationary, the corresponding reward (4.5) in fact equals $E[k(X(t), u(t))]$. Since it is optimal, this equals β , so equality must hold throughout, which is possible only if

$$V(X(t)) + \int_0^t [k(X(s), u(s)) - \beta] ds , \quad t \geq 0$$

is in fact an $\langle \mathcal{F}_t \rangle$ -martingale. \square

What we have established is the fact that storage functions are candidate value functions and vice versa, at least for the situations where the latter are known to be bounded from below. In cases where this is possible, we thus have an explicit stochastic representation for the value function of ergodic control. While an explicit stochastic representation, albeit a different one, was available for the completely observed control problem (see [2], p. 161), its counterpart for partial observations was not available. In the foregoing, however, there is little difference in the way we handle complete or partial observations.

Recall also that the usual approach for arriving at value functions for ergodic control is to consider the vanishing discount limit of suitably renormalized value functions for the associated infinite horizon discounted cost control problems. This limit is often difficult to justify and has been done under suitable hypotheses for completely observed ergodic control in [2], Chapter VI, and under rather restrictive conditions for partially observed ergodic control in [3, 4, 5]. Use of the storage function approach allows us to directly define a candidate value function V_c or V_p as above. The task then is to show that they are finite for the problem at hand. For linear stochastic differential equations describing the controlled process and noisy linear observations this can be done and a theory analogous to that of Willems [8, 9] can be developed. It is worth noting that V_c defined above for $g(\cdot, \cdot) = k(\cdot, \cdot) - \gamma$ would certainly be finite for $\gamma > \beta$ and $+\infty$ for $\gamma < \beta$, thus $\gamma = \beta$ is the ‘critical’ case.

It would be interesting to investigate the relationship of these ideas to that of Rantzer [10].

Acknowledgments

Vivek S. Borkar's research has been partially supported by a grant for 'Nonlinear Studies' from Indian Space Research Organization and Defence Research and Development Organization, Govt. of India, administered through Indian Institute of Science, Bangalore.

Sanjoy K. Mitter's research has been supported by the NSF-KDI Grant ECS-9873451 and by the Army Research Office under the MURI Grant: Data Fusion in Large Arrays of Microsensors DAAD19-00-1-0466.

4.4 References

- [1] A. G. Bhatt, V. S. Borkar: Occupation measures for controlled Markov processes: characterization and optimality. *The Annals of Probability*, vol. 24, 1996, p. 1531-1562.
- [2] V. S. Borkar: *Optimal Control of Diffusion Processes*. Pitman Research Notes in Mathematics No. 203, Longman Scientific and Technical, Harlow, England.
- [3] V. S. Borkar: The value function in ergodic control of diffusion processes with partial observations. *Stochastics and Stochastic Reports*, vol. 67, 1999, p. 255-266.
- [4] V. S. Borkar: The value function in ergodic control of diffusion processes with partial observations II. *Applicationes Mathematicae*, vol. 27, 2000, p. 455-464 (Correction note in *ibid.*, vol. 28, 2001, p. 245-246).
- [5] V. S. Borkar: Dynamic programming for ergodic control with partial observations. To appear in 'Stochastic Processes and Their Applications'.
- [6] M. H. A. Davis, P. P. Varaiya: Dynamic programming conditions for partially observable stochastic systems. *SIAM Journal of Control and Optimization*, Vol. 11, 1973, p. 226-261.
- [7] W. H. Fleming, E. Pardoux: Optimal control of partially observed diffusions. *SIAM Journal of Control and Optimization*, vol. 20, 1982, p. 261-285.
- [8] J.C. Willems: Dissipative Dynamical Systems, Part I: General Theory. *Archive for Rational Mechanics and Analysis*, Vol. 45, No. 5, 1972, p. 321-351 (Springer-Verlag, 1972).
- [9] J.C. Willems: Dissipative Dynamical Systems, Part II: Linear Systems with Quadratic Supply Rates. *Archive for Rational Mechanics and Analysis*, Vol. 45, No. 5, 1972, p. 352-393 (Springer-Verlag, 1972).
- [10] A. Rantzer: A Dual to Lyapunov's Stability Theorem, *Systems and Control Letters*, Vol. 42, No. 3, 2001, p. 161-168.