2019 IEEE 58th Conference on Decision and Control (CDC)
Palais des Congrès et des Expositions Nice Acropolis
Nice, France, December 11-13, 2019

# Stochastic Optimal Control with Markovian Lossy State Observations

Minyi Huang

***Abstract*— We consider optimal control of a diffusion process where the state is either observed exactly or completely lost at the controller, as described by a binary Markov chain. The random observation loss, coupled with the nonlinear dynamics, makes the conventional optimal control techniques difficult to apply. We introduce an abstract state space from the point of view of a hybrid system and next apply dynamic programming. For this purpose, we apply tools of differentiation of functions defined on a space of probability measures, which has no linear structure. Our approach is explicitly illustrated by a linear quadratic (LQ) control problem.**

***Index Terms*— Controlled diffusion, observation loss, hybrid system, dynamic programming, Wasserstein metric**

## I. Introduction

Consider the controlled diffusion process

$$dX_t = f(t,X_t,u_t)dt + \sigma(t,X_t,u_t)dw_t, \quad t \geq 0, \tag{1}$$

where $X_t \in \mathbb{R}^n$ is the state, $u_t \in \mathbb{R}^{n_1}$ the control, and $w_t$ an $\mathbb{R}^{n_2}$-valued standard Brownian motion. The initial state is $X_0 = x_0$. The state $X_t$ is sent to the controller via an unreliable observation link, the operational condition of which is described by a continuous-time two state Markov chain $\{\theta_t, t \geq 0\}$ with state space $\Theta = \{0,1\}$ and stationary transition probabilities. The two processes $w_t$ and $\theta_t$ are independent. Denote

$$Y_t = X_t 1_{\{\theta_t = 1\}}.$$

The observation at the controller consists of $(\theta_t, Y_t)$ at time $t$. We interpret $\theta_t = 0$ as observation link failure so that the state information is lost at the controller. Suppose for $\Delta t > 0$,

$$P(\theta_{\Delta t} = 1 | \theta_0 = 0) = q_0 \Delta t + o(\Delta t),$$
$$P(\theta_{\Delta t} = 0 | \theta_0 = 1) = q_1 \Delta t + o(\Delta t),$$

where $q_0 > 0$ is the recovery rate, and $q_1 > 0$ the failure rate.

The cost function to be minimized is

$$J(0,x_0,u(\cdot)) = E\int_0^T e^{-\rho t} L(t,X_t,u_t)dt. \tag{2}$$

For simplicity, the terminal cost is taken as 0. The general case can be treated without further difficulty.

Let $\psi$ stand for $f$ or $\sigma$. We make the following standing assumptions:

A1) $u \in U$ which is a nonempty closed subset of $\mathbb{R}^{n_1}$.

A2) $\psi(t,x,u)$ is a continuous function on $[0,T]\times\mathbb{R}^n\times U$, and is Lipschitz continuous in $x$ uniformly with respect to $(t,u)$. In addition, $|\psi| \leq C(1+|x|+|u|)$.

This work was supported by NSERC.

The author is with the School of Mathematics and Statistics, Carleton University, Ottawa, ON K1S 5B6, Canada (mhuang@math.carleton.ca).

A3) $L$ is nonnegative and continuous in $(t,x,u)$ and $L(t,x,u) \leq C(1+|x|^2+|u|^2)$, and for some $c_0 > 0$,

$$L(t,x,u) \geq c_0|u|^2 - C_0, \quad \forall\ (t,x,u) \in [0,T]\times\mathbb{R}^n\times U. \tag{3}$$

If $U$ is a compact set, (3) is automatically true.

**Admissible control:** Denote $\mathscr{F}_t^o = \sigma(\theta_s, Y_s, s \leq t)$, which is the $\sigma$-algebra generated by the observations up to time $t$. The admissible control set $\mathscr{U}_{0,T}$ consists of all process $u_t$ which is adapted to $\mathscr{F}_t^o$ such that $E\int_0^T |u_t|^2 dt < \infty$ and such that (1) has a unique strong solution.

### *A. Related Literature and Our Contributions*

We mention that, in the literature of networked systems, lossy observations or measurements have been considered in optimal filtering [13], [23], optimal control [12], [14], [17], [24], or stabilization [22], [25].

Our nonlinear model does not lead to a standard optimal control problem due to incomplete information of the state $X_t$. It is also different from control problems with partial observations [4], [10]. The application of conventional approaches, such as dynamic programming and the calculus of variations, is difficult. The key challenge is that during observation link failure, one faces an open loop control situation with random entry and exit times for that period.

To overcome the above difficulty we introduce an abstract state space on which a new control problem with complete state information is solved. This is done by specifying the distribution of the diffusion when it is not observed. Dynamic programming is applied when the domain of the value function alternates in two spaces. This is facilitated by differentiating functions defined on the space of probability measures, which does not have a linear structure. Related analysis has been developed in the analysis of abstract dynamical systems [2], [3]. This method has also been used in the recent literature of mean field control [5], [8], [11], [15]. Our model together with its analysis involves both discrete and continuous components in the states, and a switch in the observation pattern. This can be regarded as a particular form of stochastic hybrid systems [16]. The use of dynamic programming can be found in deterministic hybrid systems [7], [20] and stochastic hybrid systems [6]. Quadratic optimal control of linear hybrid systems with independent sojourn times in the modes is considered in [21]. LQ optimal control with a fixed number of Poisson observations is solved in [1].

The paper is organized as follows. Section II defines the value function in an extended state space and introduces preliminaries on differentiation of functions defined on a space of probability measures. The Hamilton-Jacobi-Bellman

(HJB) equations are presented in Section III, and their derivation is given in Section IV. An LQ example is solved in Section V for illustration. Section VI concludes the paper.

### B. Notation

Throughout the paper, the time variable $t$ (or $s$) often appears as a subscript (example: $X_t$, $u_t$, $P_t$, $g_t$) in a function or a process and should not be confused as a derivative. A partial derivative in $t$ is always denoted by $\partial_t$. For the space variable $x$ (or $y$ etc.) in a function $\phi$, the derivative or partial derivative may be denoted by $\phi_x$, $\phi_{xx}$, $\partial_x\phi$, etc. We use $\delta_x$ to denote the dirac measure at $x$. Let $\mathscr{P}_2(\mathbb{R}^n)$ be the set of Borel probability measures on $\mathbb{R}^n$ with finite second moments, and $C_B^2(\mathbb{R}^n;\mathbb{R})$ the set of functions $\phi \in C^2(\mathbb{R}^n;\mathbb{R})$ with bounded $\phi_{xx}$. For a function $\phi$ and probability measure $\mu$, we denote the integral $\int_{\mathbb{R}^n}\phi(x)\mu(dx)$ as $\langle\mu,\phi\rangle$ or $\langle\phi\rangle_\mu$. For two probability measures $\mu_1$ and $\mu_2$, $\langle\mu_1-\mu_2,\phi\rangle = \langle\mu_1,\phi\rangle-\langle\mu_2,\phi\rangle$. We use $C$ (or $C_1$, etc) to denote a generic constant. If $\phi\in C_B^2(\mathbb{R}^n;\mathbb{R})$, it can be checked that $|\phi_x(x)|\le C(1+|x|)$ and $|\phi(x)|\le C(1+|x|^2)$.

## II. The Extended State Space Model

We introduce an abstract state $X_t^* = (\theta_t, Z_t)$. The extended state space for $X_t^*$ is defined as

$$\mathfrak{X} = \mathfrak{X}^1 \cup \mathfrak{X}^0,$$

where $\mathfrak{X}^1 = \{1\}\times\mathbb{R}^n$ and $\mathfrak{X}^0 = \{0\}\times\mathscr{P}_2(\mathbb{R}^n)$. The discrete component (0 or 1) comes from the Markov chain $\theta_t$ and is called the mode. The component $Z_t$ provides information on the diffusion process by being equal to $X_t$ when it is observed or to its distribution otherwise. This enables us to convert the original optimal control problem (1)-(2) into a new one with complete state information so that dynamic programming is applicable in the new state space.

For illustration, suppose $\theta_t = 1$. Denote the stopping times $\beta_0 = \inf\{s>t|\theta_s=0\}$ and $\beta_1=\inf\{s>\beta_0|\theta_s=1\}$. By the property of the Markov chain, $\beta_0-t$ is an exponentially distributed random variable with mean $1/q_1$. The hybrid system state is $X_s^*=(\theta_s,X_s)=(1,X_s)\in\mathfrak{X}^1$ for $s\in[t,\beta_0]$, on which the observation link is active, and $X_s^*=(\theta_s,\mu_s)=(0,\mu_s)\in\mathfrak{X}^0$ for $s\in[\beta_0,\beta_1]$, where $\mu_s$ is the distribution of $X_s$ conditional on $(\beta_0, X_{\beta_0})$. Along a sample path, one can determine $\mu_{\beta_0}=\delta_{X_{\beta_0}}$ depending on the underlying sample $\omega\in\Omega$. In a long period, the trajectory of $X_t^*$ evolves back and forth between $\mathfrak{X}^0$ and $\mathfrak{X}^1$.

### A. The Value Function

We follow the method of dynamic programming [26] by defining a family of optimal control problems with different initial conditions although this is now done with switching observation patterns as described below.

Case i) Consider the initial pair $(t,x)$ and

$$dX_s^{t,x} = f(s,X_s^{t,x},u_s)ds+\sigma(s,X_s^{t,x},u_s)dw_s, \quad s\ge t, \qquad (4)$$

where $\theta_t=1$ and $X_t^{t,x}=x$.

Case ii) Consider the initial pair $(t,\mu)$ and

$$dX_s^{t,\mu} = f(s,X_s^{t,\mu},u_s)ds+\sigma(s,X_s^{t,\mu},u_s)dw_s, \quad s\ge t, \qquad (5)$$

where $\theta_t=0$ and the initial state is a random variable $X_t^{t,\mu}$ which is independent of $\{w_s, s\ge t\}$ and has distribution $\mu\in\mathscr{P}_2(\mathbb{R}^2)$. We give explanation on introducing the process in Case ii). With $\theta_t=0$, it is used to generate a physical process whose state value $X_s^{t,\mu}$ is not observed at least on a small interval after $t$. Its distribution $\mu_s$, however, can be determined with the initial condition $\mu_t=\mu$. Thus, $\mu_s$ is regarded as an abstract state evolving in an infinite dimensional space. We take $u_s=u$ and may describe $\mu_s$ by the following differential equation in a weak form: for any $\varphi\in C_B^2(\mathbb{R}^n;\mathbb{R})$,

$$\frac{d}{ds}\langle\mu_s,\varphi\rangle = \langle\mu_s, f^T(s,x,u)\varphi_x+\tfrac{1}{2}\mathrm{Tr}(\varphi_{xx}(x)(\sigma\sigma^T)(s,x,u))\rangle, \qquad (6)$$

where $s\ge t$, $\mu_t=\mu$ and the integration on the right hand side is with respect to $\mu_s(dx)$. This equation can be derived by applying Ito's formula to $\varphi(X_s^{t,\mu})$.

By restricting to the sub-interval $[t,T]$, on which the initial condition $X_t^*$ at $t$ is given, we may similarly define the admissible control set $\mathscr{U}_{t,T}$. Define two value functions as follows: For case i),

$$V^1(t,x)=\inf_{u\in\mathscr{U}_{t,T}} E[\int_t^T e^{-\rho(s-t)}L(s,X_s^{t,x},u_s)ds|\theta_t=1,X_t^{t,x}=x],$$

and for case ii),

$$V^0(t,\mu)=\inf_{u\in\mathscr{U}_{t,T}} E[\int_t^T e^{-\rho(s-t)}L(s,X_s^{t,\mu},u_s)ds|\theta_t=0,\mu_t=\mu],$$

where $\mu\in\mathscr{P}_2(\mathbb{R}^n)$. We give some prior estimates.

*Lemma 1:* For some fixed constant $C$, we have

$$0\le V^1(t,x)\le C(1+|x|^2), \quad 0\le V^0(t,\mu)\le C(1+\langle\mu,|x|^2\rangle).$$

*Proof:* Fix any $u\in U$, and we take $u_s\equiv u$. We next use $J(u(\cdot))$ defined on $[t,T]$ as an upper bound. The inequalities follow from standard SDE estimates and A3). ■

### B. Preliminary on Wasserstein Metric Space

For $\mu,\nu\in\mathscr{P}_2(\mathbb{R}^n)$, let $\Gamma(\mu,\nu)$ be the set of probability measures on $\mathbb{R}^{2n}$ which have $\mu$ and $\nu$ as the first and second marginals, respectively. The Wasserstein distance $W_2$ is defined by

$$W_2^2(\mu,\nu)=\inf_{\gamma\in\Gamma(\mu,\nu)}\int_{\mathbb{R}^{2n}}|x-y|^2\gamma(dx,dy).$$

$(\mathscr{P}_2(\mathbb{R}^n),W_2)$ is a complete and separable metric space [3].

We follow a method similar to [8] to differentiate a function defined on $\mathscr{P}_2(\mathbb{R}^n)$. This can be viewed as a local linearization of the function. More general definitions can be found in [3]. We say $\Phi:\mathscr{P}_2(\mathbb{R}^n)\to\mathbb{R}$ is differentiable at $\mu$ if there exists a Borel function $\phi:\mathbb{R}^n\to\mathbb{R}$ such that for any $\nu\in\mathscr{P}_2(\mathbb{R}^n)$, we have i) $\langle\nu,|\phi|\rangle<\infty$ and, ii)

$$\Phi(\nu)-\Phi(\mu)=\int\phi(y)(\nu-\mu)(dy)+o(W_2(\nu,\mu)).$$

We may write the derivative $\phi$ as $(\partial_\mu\Phi)(\mu,y)$ to explicitly indicate its dependence on $\mu$. Any function $\phi$ satisfying $|\phi(x)|\le C(1+|x|^2)$ ensures condition i), but this growth rate is not a necessary condition to guarantee i). Since for any constant $C$, $\int C(\mu-\nu)(dx)=0$, $\partial_\mu\Phi$ is defined up to an additive constant.

We say $\Phi$ is twice differentiable at $\mu$ if there exists a real valued Borel function $\psi(y_1,y_2)$ on $\mathbb{R}^{2n}$ such that we have i) for all $\nu_1,\nu_2\in\mathscr{P}_2(\mathbb{R}^n)$, $\iint|\psi(y_1,y_2)|\nu_1(dy_1)\nu_2(dy_2)<\infty$, and ii) for all $\nu\in\mathscr{P}_2(\mathbb{R}^n)$,

$$\begin{aligned}\Phi(\nu)-\Phi(\mu)=&\int\phi(y)(\nu-\mu)(dy)\\&+\frac{1}{2}\iint[\psi(y_1,y_2)(\nu-\mu)(dy_1)](\nu-\mu)(dy_2)\\&+o(W_2^2(\nu,\mu)).\end{aligned}\tag{7}$$

Since $\psi$ depends on $\mu$, we write $\psi=(\partial_{\mu\mu}\Phi)(\mu,y_1,y_2)$. Note that after the second differentiation, a new independent variable $y_2$ is introduced.

*Example 1:* Let $F_1(\mu)=\langle\mu,x^TQx\rangle$, where the matrix $Q\geq 0$. Then $\partial_\mu F_1=x^TQx$. ■

*Example 2:* Let $F_2(\mu)=\langle x\rangle_\mu^TQ\langle x\rangle_\mu$. Then we have

$$\begin{aligned}F_2(\nu)-F_2(\mu)=&2\langle x\rangle_\mu^TQ\langle\nu-\mu,x\rangle\\&+\langle\nu-\mu,x\rangle^TQ\langle\nu-\mu,x\rangle.\end{aligned}$$

Denote $\xi=\langle\nu-\mu,x\rangle$. We can show $|\xi|\leq W_2(\nu,\mu)$. It follows that $F_2(\nu)-F_2(\mu)=2\langle x\rangle_\mu^TQ\langle x,\nu-\mu\rangle+o(W_2(\nu,\mu))$. We obtain $\partial_\mu F_2=2\langle x\rangle_\mu^TQx$ as a linear function of $x$ for the given $\mu$. The second derivative is $\partial_{\mu\mu}F_2=2y_2^TQy_1$. ■

We give another example of higher nonlinearity.

*Example 3:* Suppose $F_3(\mu)=e^{\langle\mu,|x|^2\rangle}$. We have

$$F_3(\nu)-F_3(\mu)=e^{\langle\nu,|x|^2\rangle}-e^{\langle\mu,|x|^2\rangle}=e^{\langle\mu,|x|^2\rangle}\sum_{k=1}^\infty\frac{\Delta^k}{k!}.$$

where $\Delta=\langle\nu-\mu,|x|^2\rangle$. We have the bound

$$|\Delta|=|\langle\nu,|y|^2\rangle-\langle\mu,|x|^2\rangle|\leq W_2^2(\mu,\nu)+2W_2(\mu,\nu)\langle|x|^2\rangle_\mu^{1/2}.$$

So $\partial_\mu F_3=e^{\langle\mu,|x|^2\rangle}|x|^2$ and $\partial_{\mu\mu}F_3=e^{\langle\mu,|x|^2\rangle}|x|^2|y|^2$. ■

## III. Dynamic Programming

Define the matrix function $\Lambda(t,x,u)=\sigma(t,x,u)\sigma^T(t,x,u)$, which is from $[0,T]\times\mathbb{R}^n\times U$ to $\mathbb{R}^{n\times n}$.

*Lemma 2:* Suppose $u_s\equiv u\in U$ for $s\in[t,t_1]$ for some $t_1>t$, and let $\mu_s$ be the distribution of $X_s^{t,\mu}$ in (5) for $s\in[t,t_1]$. If $\phi\in C_B^2(\mathbb{R}^n;\mathbb{R})$, we have

$$\begin{aligned}&\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\langle\mu_{t+\varepsilon}-\mu,\phi\rangle\\=&\langle\mu,f^T(t,y,u)\phi_y+\tfrac{1}{2}\mathrm{Tr}(\phi_{yy}\Lambda(t,u,y))\rangle,\end{aligned}\tag{8}$$

where the right hand takes integration with respect to $\mu(dy)$.

*Proof:* We have $\langle\mu_{t+\varepsilon}-\mu,\phi\rangle=E[\phi(X_{t+\varepsilon}^{t,\mu})-\phi(X_t^{t,\mu})]$. By Ito's formula for (5), we obtain the limit in (8). ■

*Lemma 3:* For (5), suppose $\mu_t=\delta_z$, and $u_s=u$ on $[t,t+\varepsilon]$ for some fixed $\varepsilon>0$ in (5). Then $\langle\mu_{t+\varepsilon},|x|^2\rangle\leq C(1+|z|^2)$, where $C$ does not depend on $\varepsilon$.

*Proof:* We have $\langle\mu_{t+\varepsilon},|x|^2\rangle=E|X_{t+\varepsilon}^{t,\mu}|^2$, where $X_t^{t,\mu}=z$. The lemma follows from the second moment estimate for the SDE with the deterministic initial condition. ■

The following hypothesis H1) is the optimality principle.

H1) $U$ is compact. The dynamic programming principle holds for the value function. Specifically, for any $0\leq\varepsilon\leq T-t$, we have the relation

$$\begin{aligned}V^1(t,x)=&\min_{u(\cdot)}E[\int_t^{t+\varepsilon}e^{-\rho(s-t)}L(s,X_s^{t,x},u_s)ds\\&+e^{-\rho\varepsilon}V^{\theta_{t+\varepsilon}}(t+\varepsilon,Z_{t+\varepsilon})|\theta_t=1,X_t^{t,x}=x],\end{aligned}\tag{9}$$

$$\begin{aligned}V^0(t,\mu)=&\min_{u(\cdot)}E[\int_t^{t+\varepsilon}e^{-\rho(s-t)}L(s,X_s^{t,\mu},u_s)ds\\&+e^{-\rho\varepsilon}V^{\theta_{t+\varepsilon}}(t+\varepsilon,Z_{t+\varepsilon})|\theta_t=0,\mu_t=\mu],\end{aligned}\tag{10}$$

where $\mu\in\mathscr{P}_2(\mathbb{R}^n)$.

We further introduce some regularity assumptions on the value functions regarding differentiability.

H2) $\partial_tV^1$, $\partial_xV^1$, $\partial_{xx}V^1$ are continuous functions on $[0,T]\times\mathbb{R}^n$, and $\partial_{xx}V^1$ is bounded.

H3) $\partial_tV^0$ is a continuous function of $(t,\mu)$. $\partial_\mu V^0$ has continuous up to second order partial derivatives in $y$, and $\partial_{yy}\partial_\mu V^0$ is bounded.

*Remark 1:* H3) implies $|\partial_y(\partial_\mu V^0)(t,\mu,y)|\leq C(1+|y|)$.

H4) $\partial_{\mu\mu}V^0$ exists for all $\mu\in\mathscr{P}_2(\mathbb{R}^n)$. For $\psi(y_1,y_2)=(\partial_{\mu\mu}V^0)(t,\mu,y_1,y_2)$, the partial derivatives $\psi_{y_1}$, $\psi_{y_2}$, $\psi_{y_1y_2}$, $\psi_{y_1y_1}$, $\psi_{y_2y_2}$ exist and are continuous in $(y_1,y_2)$. Moreover, $|\psi_{y_1y_1}|+|\psi_{y_2y_2}|+|\psi_{y_1y_2}|\leq C$.

*Remark 2:* H4) implies $|\psi_{y_1}|+|\psi_{y_2}|\leq C(1+|y_1|+|y_2|)$ and $|\psi|\leq C(1+|y_1|^2+|y_2|^2)$.

*Lemma 4:* Let $\mu_s$, $s\in[t,t+\varepsilon]$, be determined by (5) with $u_s=u$. Then we have $W_2(\mu_{t+\varepsilon},\mu)=O(\varepsilon^{1/2})$.

*Proof:* We have $W_2^2(\mu_{t+\varepsilon},\mu)\leq E|X_{t+\varepsilon}^{t,\mu}-X_t^{t,\mu}|^2=O(\varepsilon)$, and the lemma follows. ■

*Lemma 5:* Denote $\nu=\mu_{t+\varepsilon}-\mu$. Suppose that H4) and the assumption in Lemma 4 hold. Then we have

$$\Big|\int(\partial_{\mu\mu}V^0)(t,\mu,y_1,y_2)\nu(dy_1)\nu(dy_2)\Big|=o(\varepsilon).\tag{11}$$

*Proof:* See appendix. ■

*Theorem 6:* Under H1)-H4), we have the dynamic programming equations

$$\begin{aligned}\rho V^1(t,x)=&\partial_tV^1(t,x)+q_1[V^0(t,\delta_x)-V^1(t,x)]\\&+\min_{u\in U}[f^T(t,x,u)V_x^1(t,x)+L(t,x,u)\\&+\tfrac{1}{2}\mathrm{Tr}(V_{xx}^1(t,x)\Lambda(t,x,u))],\end{aligned}\tag{12}$$

$$\begin{aligned}\rho V^0(t,\mu)=&\partial_tV^0(t,\mu)+q_0[\langle\mu,V^1(t,\cdot)\rangle-V^0(t,\mu)]\\&+\min_{u\in U}\Big[\langle\mu,L(t,\cdot,u)\rangle\\&+\int f^T(t,y,u)\partial_y(\partial_\mu V^0)(t,\mu,y)\mu(dy)\\&+\int\tfrac{1}{2}\mathrm{Tr}[(\partial_{yy}(\partial_\mu V^0)(t,\mu,y))\Lambda(t,y,u)]\mu(dy)\Big],\end{aligned}\tag{13}$$

where $\mu\in\mathscr{P}_2(\mathbb{R}^n)$, $V^1(T,x)=0$ and $V^0(T,\mu)=0$. ■

We call (12)-(13) the HJB equations, where (12) will reduce to a standard HJB equation with unknown $V^1$ if $q_1$ is taken as 0. The proof of Theorem 6 is given in Section IV. The proof needs certain regularity in terms of $\partial_{\mu\mu}V^0$ (see H4)) although this term does not appear in (13).

*Remark 3:* We assume compact control space $U$ in H1). This will simplify the derivation of the HJB equations. If $U$

is unbounded, the SDE estimates on short time intervals as developed in Section IV will be more difficult. This bounded control set does not cover the LQ case.

On the other hand, the two HJB equations are still well defined in the LQ case for which we expect them to be applicable. This case will be solved explicitly in Section V.

## IV. Derivation of the HJB Equations

### A. The Analysis of (12)

Given $\theta_t = 1$ and $X_t^{t,x} = x$, we have the following estimate $V^0(t+\varepsilon, \mu_{t+\varepsilon}) = V^0(t, \delta_x) + o(1)$. By the dynamic programming principle and local Taylor expansion of the function $V^1$, we can establish (12). See standard techniques in [26].

### B. The Analysis of (13)

Part 1. Denote the right hand side of (13) by $H_\mu$. Our method is to first show

$$\rho V^0(t,\mu) \le H_\mu. \tag{14}$$

Fix $t \in [0,T)$. Take $u_s \equiv u \in U$ for $s \in [t, t+\varepsilon]$, where $\varepsilon > 0$ is sufficiently small. Given $\theta_t = 0$, define $\tau_1 = \inf\{s > t | \theta_s = 1\}$, $\tau_2 = \inf\{s > \tau_1 | \theta_s = 0\}$. By H1), we have

$$\begin{aligned} V^0(t,\mu) \le & E\int_t^{t+\varepsilon} e^{-\rho(s-t)} L(s, X_s^{t,\mu}, u_s) ds \qquad (15)\\ & + e^{-\rho\varepsilon} E V^{\theta_{t+\varepsilon}}(t+\varepsilon, Z_{t+\varepsilon}) \\ & = \varepsilon E L(t, X_t^{t,\mu}, u) + o(\varepsilon) + e^{-\rho\varepsilon}(K_0 + K_1), \end{aligned}$$

where the expectation is calculated given the initial condition $\theta_t = 0$, $X_t$ having distribution $\mu$, and we denote

$$\begin{aligned} K_0 &= E[V^{\theta_{t+\varepsilon}}(t+\varepsilon, Z_{t+\varepsilon}) 1_{(\tau_1 > t+\varepsilon)}], \\ K_1 &= E[V^{\theta_{t+\varepsilon}}(t+\varepsilon, Z_{t+\varepsilon}) 1_{(\tau_1 \le t+\varepsilon)}]. \end{aligned}$$

It follows that

$$\begin{aligned} & EL(t, X_t^{t,\mu}, u) = \langle \mu, L(t, \cdot, u) \rangle, \\ & K_0 = E[V^0(t+\varepsilon, \mu_{t+\varepsilon}) 1_{(\tau_1 > t+\varepsilon)}] = V^0(t+\varepsilon, \mu_{t+\varepsilon}) e^{-q_0\varepsilon}. \end{aligned}$$

Since $\partial_{\mu\mu} V^0$ exists, we obtain

$$\begin{aligned} V^0(t+\varepsilon, \mu_{t+\varepsilon}) =& V^0(t,\mu) + V^0(t+\varepsilon, \mu_{t+\varepsilon}) - V^0(t, \mu_{t+\varepsilon}) \\ & + V^0(t, \mu_{t+\varepsilon}) - V^0(t,\mu) \\ =& V^0(t,\mu) + (\partial_t V^0(t,\mu) + o(1))\varepsilon \\ & + \langle \mu_{t+\varepsilon} - \mu, (\partial_\mu V^0)(t,\mu,\cdot) \rangle \\ & + \tfrac{1}{2} \int (\partial_{\mu\mu} V^0)(t,\mu,y_1,y_2) \nu(dy_1)\nu(dy_2) \\ & + o(W_2^2(\mu_{t+\varepsilon}, \mu)), \end{aligned} \tag{16}$$

where $\nu = \mu_{t+\varepsilon} - \mu$ and $\mu_t = \mu$. By (16), Lemmas 4 and 5,

$$\begin{aligned} K_0 =& V^0(t,\mu) + \partial_t V^0(t,\mu)\varepsilon \\ & + \langle \mu_{t+\varepsilon} - \mu, (\partial_\mu V^0)(t,\mu,y) \rangle - q_0 V^0(t,\mu)\varepsilon + o(\varepsilon). \end{aligned}$$

By Lemma 2,

$$\begin{aligned} & \langle \mu_{t+\varepsilon} - \mu, (\partial_\mu V^0)(t,\mu,y) \rangle \\ =& \Big[ \int f^T(t,y,u) \partial_y (\partial_\mu V^0)(t,\mu,y) \mu(dy) \\ & + \int \tfrac{1}{2} \mathrm{Tr}[(\partial_{yy}(\partial_\mu V^0)(t,\mu,y)) \Lambda(t,y,u)] \mu(dy) \Big] \varepsilon + o(\varepsilon). \end{aligned}$$

For the remaining part, we estimate $K_1$. We have

$$\begin{aligned} K_1 &= E[V^{\theta_{t+\varepsilon}}(t+\varepsilon, Z_{t+\varepsilon}) 1_{(\tau_1 \le t+\varepsilon, \tau_2 > t+\varepsilon)}] + K_3 \\ &= E[V^1(t+\varepsilon, X_{t+\varepsilon}^{t,\mu}) 1_{(\tau_1 \le t+\varepsilon, \tau_2 > t+\varepsilon)}] + K_3, \end{aligned}$$

where $K_3 = E[V^{\theta_{t+\varepsilon}}(t+\varepsilon, Z_{t+\varepsilon}) 1_{(\tau_1 \le t+\varepsilon, \tau_2 \le t+\varepsilon)}]$. Since $X_{t+\varepsilon}^{t,\mu}$ is determined by a constant input $u$, it is independent of the event $\{\tau_1 \le t+\varepsilon, \tau_2 > t+\varepsilon\}$. So we obtain

$$\begin{aligned} & E[V^1(t+\varepsilon, X_{t+\varepsilon}^{t,\mu}) 1_{(\tau_1 \le t+\varepsilon, \tau_2 > t+\varepsilon)}] \\ =& EV^1(t+\varepsilon, X_{t+\varepsilon}^{t,\mu}) q_0 \varepsilon + o(\varepsilon) \\ =& EV^1(t, X_t^{t,\mu}) q_0 \varepsilon + o(\varepsilon) = q_0 \varepsilon \langle \mu, V^1(t,\cdot) \rangle + o(\varepsilon). \end{aligned} \tag{17}$$

To estimate $K_3$, we need some care since it is the product of two terms which are not independent. Next we have

$$\begin{aligned} & |V^{\theta_{t+\varepsilon}}(t+\varepsilon, Z_{t+\varepsilon}) 1_{(\theta_{t+\varepsilon}=1)}| = |V^1(t+\varepsilon, X_{t+\varepsilon}^{t,\mu}) 1_{(\theta_{t+\varepsilon}=1)}| \\ \le & |V^1(t+\varepsilon, X_{t+\varepsilon}^{t,\mu})| \le C(1 + |X_{t+\varepsilon}^{t,\mu}|^2) \end{aligned} \tag{18}$$

by Lemma 1. We further have

$$|V^{\theta_{t+\varepsilon}}(t+\varepsilon, Z_{t+\varepsilon}) 1_{(\theta_{t+\varepsilon}=0)}| = |V^0(t+\varepsilon, \mu_{t+\varepsilon}) 1_{(\theta_{t+\varepsilon}=0)}|,$$

where $\mu_{t+\varepsilon}$ is random. By Lemma 1, we have

$$\begin{aligned} |V^0(t+\varepsilon, \mu_{t+\varepsilon})| &\le C(1 + \langle \mu_{t+\varepsilon}(\omega), |x|^2 \rangle) \\ &\le C(1 + \sup_{t \le s \le t+\varepsilon} |X_s^{t,\mu}|^2), \end{aligned} \tag{19}$$

where we have used the fact that $\mu_{t+\varepsilon}$ is determined from the distribution of $X_{t+\varepsilon}^{t,\mu}$ with a random initial condition of the form $\delta_{X_{\tau(\omega)}^{t,\mu}(\omega)}$, where $\tau$ is the last time when the diffusion is observed. Note that

$$E \sup_{t \le s \le t+\varepsilon} |X_s^{t,\mu}|^2 \le C, \quad E 1_{(\tau_1 \le t+\varepsilon, \tau_2 \le t+\varepsilon)} = O(\varepsilon^2).$$

It follows that

$$K_3 \le CE[(1 + \sup_{t \le s \le t+\varepsilon} |X_s^{t,\mu}|^2) 1_{(\tau_1 \le t+\varepsilon, \tau_2 \le t+\varepsilon)}] = O(\varepsilon^2),$$

where the process $\{X_s^{t,\mu}, t \le s \le t+\varepsilon\}$ is independent of the indicator random variable.

Since $u$ is arbitrary, we obtain (14).

Part 2. We show the other direction of the inequality, i.e.,

$$\rho V^0(t,\mu) \ge H_\mu. \tag{20}$$

Fix any small $\delta > 0$. For any $\varepsilon > 0$, we can find a control on $[t, t+\varepsilon]$ such that

$$\begin{aligned} V^0(t,\mu) + \delta\varepsilon \ge & E\int_t^{t+\varepsilon} e^{-\rho(s-t)} L(s, X_s^{t,\mu}, u_s) ds \\ & + e^{-\rho\varepsilon} E V^{\theta_{t+\varepsilon}}(t+\varepsilon, Z_{t+\varepsilon}). \end{aligned}$$

We have

$$\begin{aligned} E\int_t^{t+\varepsilon} e^{-\rho(s-t)} L(s, X_s^{t,\mu}, u_s) ds = & E\int_t^{t+\varepsilon} \langle \mu_t, L(s,\cdot,u_s) \rangle ds \\ & + o(\varepsilon). \end{aligned}$$

Denote $k_0 = E[V^{\theta_{t+\varepsilon}}(t+\varepsilon, Z_{t+\varepsilon}) 1_{(\tau_1 > t+\varepsilon)}]$, $k_1 = E[V^{\theta_{t+\varepsilon}}(t+\varepsilon, Z_{t+\varepsilon}) 1_{(\tau_1 \le t+\varepsilon)}]$. Next, we have

$$\begin{aligned} EV^{\theta_{t+\varepsilon}}(t+\varepsilon, Z_{t+\varepsilon}) =& k_0 + k_1 \\ =& EV^0(t+\varepsilon, \mu_{t+\varepsilon}) 1_{(\tau_1 > t+\varepsilon)} + k_1. \end{aligned}$$

We can further show (20). The remaining estimates are similar to those in Part 1, and we omit the detail. ∎

## V. Linear Quadratic Case

We consider the linear case of the SDE (1) in the form

$$dX_t = (AX_t + Bu_t)dt + Ddw_t, \quad t \in [0,T]. \tag{21}$$

The cost integrand is $(X_t - \phi_t)^T Q(X_t - \phi_t) + u_t^T R u_t$, where $\phi_t$ is a continuous function on $[0,T]$ and $Q \geq 0$, $R > 0$. For simplicity, we consider constant matrices $A$, $B$, $D$, $Q$ and $R$.

Under the lossy observation conditions, the separate principle holds for this LQ model and the optimal control law can be determined using either the state or its linear prediction when state loss occurs. This can be done by applying or adapting the methods in [24] for i.i.i. packet losses or in [17] for Markovian packet losses. We will not give the solution by this approach. Instead, this model serves to illustrate the general solution in Section III so as to give insights into the abstract dynamic programming. Our method further gives the optimal cost in a closed-form.

The HJB equation (12) is simplified into the form

$$\begin{aligned}\rho V^1(t,x) =& \partial_t V^1(t,x) + q_1[V^0(t,\delta_x) - V^1(t,x)] \\ &+ x^T A^T V_x^1 + (x-\phi_t)^T Q(x-\phi_t) + \tfrac{1}{2}\mathrm{Tr}(V_{xx}^1 DD^T) \\ &- \tfrac{1}{4} V_x^{1T} B^T R^{-1} B^T V_x^1, \end{aligned} \tag{22}$$

where the optimal control is $u = -\frac{1}{2}R^{-1}B^T V_x^1$. The second HJB equation (13) reduces to

$$\begin{aligned}&\rho V^0(t,\mu) = \partial_t V^0(t,\mu) + q_0[\langle \mu, V^1(t,\cdot)\rangle - V^0(t,\mu)] \\ &+ \langle \mu, (y-\phi_t)^T Q(y-\phi_t) + y^T A^T \partial_y \partial_\mu V^0(t,\mu,y)\rangle \\ &+ \tfrac{1}{2}\langle \mu, \mathrm{Tr}(\partial_{yy}\partial_\mu V^0(t,\mu,y)DD^T)\rangle \\ &- \tfrac{1}{4}\langle \mu, \partial_y\partial_\mu V^0(t,\mu,y)\rangle^T BR^{-1}B^T \langle \mu, \partial_y \partial_\mu V^0(t,\mu,y)\rangle, \end{aligned} \tag{23}$$

where the optimal control is

$$u = -\tfrac{1}{2}R^{-1}B^T \langle \mu, \partial_y \partial_\mu V^0(t,\mu,y)\rangle.$$

We introduce two ODE systems:

$$\begin{cases}(\rho+q_1)P = \dot{P} + A^T P + PA - PBR^{-1}B^T P + Q + q_1(K+G), \\ (\rho+q_1)S = \dot{S} + A^T S - PBR^{-1}B^T S + q_1 N - Q\phi_t, \\ (\rho+q_1)r = \dot{r} + q_1 g - S^T BR^{-1}B^T S + \phi_t^T Q\phi_t + \mathrm{Tr}(PDD^T), \end{cases} \tag{24}$$

where $P_T = 0$, $S_T = 0$ and $r_T = 0$; and

$$\begin{cases}\rho K = \dot{K} + q_0(P-K) + A^T K + KA + Q, \\ \rho G = \dot{G} - q_0 G + A^T G + GA - (K+G)BR^{-1}B^T(K+G), \\ \rho N = \dot{N} + q_0(S-N) + A^T N - (K+G)BR^{-1}B^T N - Q\phi_t, \\ \rho g = \dot{g} + q_0(r-g) - N^T BR^{-1}B^T N + \phi_t^T Q\phi_t + \mathrm{Tr}(KDD^T), \end{cases} \tag{25}$$

where $K_T = G_T = 0$, $N_T = 0$ and $g_T = 0$.

*Lemma 7:* There exists a unique solution to (24)-(25).

*Proof:* Denoting $\widehat{K} = K + G$, we obtain two coupled Riccati equations

$$\rho P = \dot{P} + A^T P + PA - PBR^{-1}B^T P + Q + q_1(\widehat{K} - P),$$
$$\rho \widehat{K} = \dot{\widehat{K}} + A^T \widehat{K} + \widehat{K}A - \widehat{K}BR^{-1}B^T\widehat{K} + Q + q_0(P - \widehat{K}),$$

where $P_T = \widehat{K}_T = 0$. This system has a unique solution $P(t) \geq 0$, $\widehat{K}(t) \geq 0$, $t \in [0,T]$ ([9], [27]). Then we uniquely solve $K$ from a linear ODE when $P$ is given. This further gives $G$. Next, we uniquely solve $(S,N)$, and finally $(r,g)$. ∎

By use of the ODE of $P - \widehat{K}$, we can further show $P = \widehat{K}$ on $[0,T]$. Therefore, $P$ is simply solved from

$$\rho P = \dot{P} + A^T P + PA - PBR^{-1}B^T P + Q, \quad P(T) = 0.$$

*Theorem 8:* The pair of functions

$$V^1(t,x) = x^T P_t x + 2x^T S_t + r_t, \tag{26}$$
$$V^0(t,\mu) = \langle \mu, x^T K_t x\rangle + \langle x\rangle_\mu^T G_t \langle x\rangle_\mu + 2\langle x\rangle_\mu^T N_t + g_t, \tag{27}$$

is a solution to (22)-(23).

*Proof:* By (26)-(27), we obtain

$$\begin{aligned}&V^0(t,\delta_x) = x^T(K_t + G_t)x + 2x^T N_t + g_t, \\ &\partial_t V^1 = x^T \dot{P}_t x + 2x^T \dot{S}_t + \dot{r}_t, \\ &V_x^1 = 2P_t x + 2S_t, \quad V_{xx}^1 = 2P_t.\end{aligned}$$

Next, we have

$$\begin{aligned}&(\partial_\mu V^0)(t,\mu,y) = y^T K_t y + 2y^T N_t + 2\langle x\rangle_\mu^T G_t y, \\ &\partial_y(\partial_\mu V^0)(t,\mu,y) = 2K_t y + 2N_t + 2G_t\langle x\rangle_\mu, \\ &\partial_{yy}(\partial_\mu V^0)(t,\mu,y) = 2K_t.\end{aligned}$$

By (24)-(25), we verify that $V^0$ and $V^1$ satisfy (22)-(23). ∎

*Remark 4:* Assumptions H2-H4) hold in this LQ model.

## VI. Concluding Remarks

This paper considers optimal control of a diffusion process with Markovian lossy state observations at the controller. We apply dynamic programming in a hybrid state space, where the continuous state evolves between the Euclidean space and an infinite dimensional space of probability measures. This relies on differentiation of functions defined on Wasserstein metric space without a linear structure. Our approach is illustrated by an LQ model with closed-form solutions.

For future work, it is of interest to develop existence results with unbounded control space and to generalize this modeling framework to other non-diffusion system models. The uniqueness analysis of the solution by restricting to a certain space is another interesting issue. The dynamic programming principle may be further studied by use of nonlinear semi-groups [18], [19]. A further generalization is to consider the case where the normal observation duration vanishes toward zero. This will give a model with observations arriving as a Poisson process [1].

## Appendix: Proof of Lemma 5.

We prove by a probabilistic approach. Denote the SDE

$$d\tilde{X}_s^{t,\mu} = f(s,\tilde{X}_s^{t,\mu},u)ds + \sigma(s,\tilde{X}_s^{t,\mu},u)d\tilde{w}_s, \tag{28}$$

where $t \leq s \leq T$. The initial condition $\tilde{X}_t^{t,\mu}$ is independent of $\{\tilde{w}_s, s \geq t\}$ and has the distribution $\mu$. In addition, $(\tilde{X}_t^{t,\mu}, \tilde{w}_s)$ and $(X_t^{t,\mu}, w_s)$ are independent. We may introduce a product probability space on which both $\tilde{X}_s^{t,\mu}$ and $X_s^{t,\mu}$ are defined.

We still use $E$ to denote expectation on the product probability space. In the remaining proof we simply write $X_s^{t,\mu}$ and $\tilde{X}_s^{t,\mu}$ as $X_s$ and $\tilde{X}_s$, respectively.

Denote $\phi(y_1,y_2) := (\partial_{\mu\mu}V^0)(t,\mu,y_1,y_2)$. Then

$$\int \phi(y_1,y_2)\nu(dy_1)\nu(dy_2) = E[\phi(X_{t+\varepsilon},\tilde{X}_{t+\varepsilon}) - \phi(X_t,\tilde{X}_{t+\varepsilon}) - \phi(X_{t+\varepsilon},\tilde{X}_t) + \phi(X_t,\tilde{X}_t)] =: C_\nu. \quad (29)$$

By applying Ito's formula, we have

$$E\phi(X_{t+\varepsilon},\tilde{X}_{t+\varepsilon}) = E\phi(X_t,\tilde{X}_t) + \int_t^{t+\varepsilon} Eh_1(s)ds, \quad (30)$$

where

$$\begin{aligned} h_1(s) =& \phi_{y_1}^T(X_s,\tilde{X}_s)f(s,X_s,u) + \phi_{y_2}^T(X_s,\tilde{X}_s)f(s,\tilde{X}_s,u) \\ &+ \tfrac{1}{2}\mathrm{Tr}[\phi_{y_1y_1}(X_s,\tilde{X}_s)(\sigma\sigma^T)(s,X_s,u)] \\ &+ \tfrac{1}{2}\mathrm{Tr}[\phi_{y_2y_2}(X_s,\tilde{X}_s)(\sigma\sigma^T)(s,\tilde{X}_s,u)]. \end{aligned}$$

Similarly, we have

$$E\phi(X_t,\tilde{X}_{t+\varepsilon}) = E\phi(X_t,\tilde{X}_t) + \int_t^{t+\varepsilon} Eh_2(s)ds, \quad (31)$$

$$E\phi(X_{t+\varepsilon},\tilde{X}_t) = E\phi(X_t,\tilde{X}_t) + \int_t^{t+\varepsilon} Eh_3(s)ds, \quad (32)$$

where the expressions of $h_2$ and $h_3$ are easily determined but omitted here. Subsequently, we have

$$C_\nu = E\int_t^{t+\varepsilon} [\xi_1(s) + \xi_2(s) + \mathrm{Tr}\xi_3(s) + \mathrm{Tr}\xi_4(s)]ds,$$

where

$$\begin{aligned} \xi_1(s) &= [\phi_{y_1}^T(X_s,\tilde{X}_s) - \phi_{y_1}^T(X_s,\tilde{X}_t)]f(s,X_s,u), \\ \xi_2(s) &= [\phi_{y_2}^T(X_s,\tilde{X}_s) - \phi_{y_2}^T(X_t,\tilde{X}_s)]f(s,\tilde{X}_s,u), \\ \xi_3(s) &= \tfrac{1}{2}[\phi_{y_1y_1}(X_s,\tilde{X}_s) - \phi_{y_1y_1}(X_s,\tilde{X}_t)](\sigma\sigma^T)(s,X_s,u), \\ \xi_4(s) &= \tfrac{1}{2}[\phi_{y_2y_2}(X_s,\tilde{X}_s) - \phi_{y_2y_2}(X_t,\tilde{X}_s)](\sigma\sigma^T)(s,\tilde{X}_s,u). \end{aligned}$$

Then by H4),

$$E\int_t^{t+\varepsilon} |\xi_1(s)|ds \le CE\int_t^{t+\varepsilon} |\tilde{X}_s - \tilde{X}_t|(1+|X_s|)ds \le C\varepsilon^{3/2}.$$

Similarly, $E\int_t^{t+\varepsilon}|\xi_2(s)|ds \le C\varepsilon^{3/2}$. Denote $m_{tT} = \sup_{t\le r\le T}|X_r|$. Then $Em_{tT}^2 < \infty$. Note that $|\xi_k(s)| \le C(1+m_{tT}^2)$ for $k = 3,4$. By dominated convergence, we can show that $E\xi_k(s)$ is a continuous function on $[t,T]$ and $\lim_{s\downarrow t} E\zeta_k(s) = 0$. It follows that

$$\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon} \mathrm{Tr}[E\xi_3(s) + E\xi_4(s)]ds = 0. \quad (33)$$

We obtain $|C_\nu| = o(\varepsilon)$. The lemma follows. ■

## REFERENCES


[1] M. Ades, P.E. Caines, and R.P. Malhamé. Stochastic optimal control under Poisson-distributed observations. *IEEE Transactions on Automatic Control*. vol. 45, no. 1, pp. 3-13, 2000.

[2] L. Ambrosio and W. Gangbo. Hamiltonian ODEs in the Wasserstein space of probability measures. *Commun. Pure Applied Math.*, vol. 61, no. 1, pp. 18-53, 2008.

[3] L. Ambrosio, N. Gigli, and G. Savare. *Gradient flows in metric spaces and the Wasserstein spaces of probability measures*, Lectures in Mathematics, ETH Zurich, Birkhauser, 2005.

[4] A. Bensoussan. *Stochastic Control of Partially Observable Systems*. Cambridge University Press, Cambridge, 1992.

[5] A. Bensoussan, J. Frehse, and P. Yam. *Mean Field Games and Mean Field Type Control Theory*. Springer, New York, 2013.

[6] A. Bensoussan and J.L. Menaldi. Stochastic hybrid control, *J. Math. Anal. Appl.*, vol. 249, pp. 21-288, 2000.

[7] M. S. Branicky, V.S. Borkar, and S.K. Mitter. A unified framework for hybrid control: Model and optimal control theory, *IEEE Trans. Autom. Control*, vol. 43, pp. 31-45, 1998.

[8] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions. The master equation and the convergence problem in mean field games, preprint, 2015.

[9] O.L.V. Costa, M.D. Fragoso, and M. G. Todorov. *Continuous-Time Markov Jump Linear Systems*, Springer-Verlag, Berlin, 2013.

[10] M.H.A. Davis and P. Varaiya. Dynamic programming conditions for partially observable stochastic systems, *SIAM J. Control*, vol. 11, no. 2, pp. 226-261, 1973.

[11] W. Gangbo and A. Swiech. Existence of a solution to an equation arising from the theory of mean field games. *J. Diff. Equations*, vol. 259, no. 11, pp. 6573-6643, 2015.

[12] V. Gupta, B. Hassibi, and R.M. Murray. Optimal LQG control across packet-dropping links. *Syst. Control Lett.*, vol. 56, pp. 439-446, 2007.

[13] M. Huang and S. Dey. Stability of Kalman filtering with Markovian packet losses. *Automatica*, vol. 43, pp. 598-607, 2007.

[14] O.C. Imer, S. Yuksel, and T. Basar. Optimal control of LTI systems over unreliable communication links. *Automatica*, vo. 42 no. 9, pp. 1429-1439, 2006.

[15] V. N. Kolokoltsov, J. Li, and W. Yang. Mean field games and nonlinear Markov processes, arXiv:1112.3744v2, 2011.

[16] J. Lygeros and M. Prandini. Stochastic hybrid systems: A powerful framework for complex, large scale applications. *European Journal of Control*, vol. 6, pp.583-594, 2010.

[17] Y. Mo, E. Garone, and B. Sinopoli. LQG control with Markovian packet loss. *Proc. European Control Conference*, Zurich, Switzerland, pp. 2380-2385, July 2013.

[18] M. Nisio. On a non-linear semi-group attached to stochastic optimal control, *Publ. RIMS, Kyoto Univ*, vol. 13, pp. 513-537, 1976.

[19] M. Nisio. *Stochastic Control Theory: Dynamic Programming Principle*, 2nd ed., Springer, Tokyo, 2015.

[20] A. Pakniyat and P. E. Caines. On the relation between the minimum principle and dynamic programming for classical and hybrid control systems. *IEEE Trans. Autom. Control*, vol. 62, no. 9, pp. 4347-4362, Sept. 2017.

[21] F. R. Pour Safaei, K. Rohb, S. R. Proulx, and J. P. Hespanha. Quadratic control of stochastic hybrid systems with renewal transitions. *Automatica*, vol. 50, pp. 2822-2834, 2014.

[22] D. E. Quevedo and D. Nesic. Robust stability of packetized predictive control of nonlinear systems with disturbances and Markovian packet losses. *Automatica*, vol. 48, no. 8, pp. 1803-1811, 2012.

[23] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry. Kalman filtering with intermittent observations. *IEEE Trans. Autom. Control*, vol. 49, pp. 1453-1464, 2004.

[24] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry. Foundations of control and estimation over lossy networks. *Proceedings of the IEEE*, vol. 96, no. 1, pp. 163-187, 2007.

[25] L. Xie and L. Xie. Stability analysis of networked sampled-data linear systems with Markovian packet losses *IEEE Transactions on Automatic Control*, vol. 54, no. 6, pp. 1375-1381, 2009.

[26] J. Yong and X.Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer, New York, 1999.

[27] W. M. Wonham. Random differential equations in control theory, in *Probabilistic Methods in Applied Mathematics* (Edited by A. T. Bharucha-Reid), 2(1970) pp. 131-212, Academic Press, New York.