CYBER-PHYSICAL CLOUD COMPUTING LAB
UNIVERSITY OF CALIFORNIA, BERKELEY

REAL-TIME STATE ESTIMATION OF LINEAR PROCESSES OVER A SHARED AWGN CHANNEL WITHOUT FEEDBACK

Ching-Ling Huang and Raja Sengupta

Working Papers
CPCC-WP-2012-08-02

August 2012

Real-Time State Estimation of Linear Processes over a Shared AWGN Channel without Feedback

Ching-Ling Huang, Raja Sengupta

Abstract

We formulate the problem of estimating the state in real-time of multiple continuous-time linear dynamical systems over a shared Additive White Gaussian Noise (AWGN) channel without channel feedback and characterize the optimal encoders and decoder for minimum mean-squared error (MMSE) estimation. The first analyses are for a a single linear system being observed over an AWGN channel with a power constraint. One optimal encoder is a scaled innovation encoder and the decoder a scaling of the channel output. We then study the same problem for multiple linear systems communicating over a shared AWGN channel and characterize the optimal encoders and decoder. Finally the encoders and decoder are characterized in closed form for two identical linear sources with some correlation. We bound the optimal costs without the correlation assumption.

I. Introduction

This paper is about remotely observing multiple linear dynamical systems or diffusions, over a shared channel, to estimate their states in real-time as illustrated in figure 2. The shared channel is an Additive White Gaussian Noise (AWGN) channel. Each dynamical system uses an encoder to transform its state for communication. The channel outputs the superposition of its inputs corrupted by noise. This output is received by one single decoder that produces estimates of the states of all the dynamical systems. The encoders and decoder are causal and continuous. If X^t is the state of all of the dynamical systems at time t and $\hat{X}^t$ the output of the decoder at the same time, then the encoders and decoders are to be designed to make $\hat{X}^t$ as close as possible to X^t using only information available up to time t. Thus we label the formulation real-time. More precisely the performance measure is the expected value of $(X^t - \hat{X}^t)^2$, i.e., we characterize the Minimum Mean Square Error (MMSE) encoders and decoder.

C. L. Huang and R. Sengupta are with the Systems Program, Dept. Civil and Environmental Engineering, University of California at Berkeley. Contact: sengupta@ce.berkeley.edu

It has been noted (e.g., by [30], [23]) that Shannon's classical information theorems [25] cannot be applied directly to real-time observation or state estimation problems. For example, the Shannon capacity of a channel is not realizable when the encoded sequence is short. The encoded block needs to be long enough to achieve decoding typicality and thus channel capacity [25]. Since real-time estimation cannot tolerate long encoding delays, there is a newer literature on the problem of remotely observing or estimating the state of a dynamical system over a noisy channel, e.g., see [2], [27], [24], [26], [34], [12], [28]. Most of This literature has focused on observing a single dynamical system. This paper adds some results to the single dynamical system case and breaks new ground on optimally estimating the states of multiple dynamical systems sending signals through a noisy channel .

The literature can be separated into a part focused on stability and another focused on optimality. The stablity literature, e.g., see [23], [27], [24], [26], [34], focuses on the necessary and sufficient conditions for asymptotically stable real-time state estimation. Tatikonda [27], Sahai [23], and Mitter treat the case of the unstable Linear Time-Invariant (LTI) source and derive real-time information measures for estimating the state of this type of source over an unreliable channel, namely 1) the sequential rate distortion [27] for the bit rate at the source encoder required to achieve stable tracking; and 2) the anytime capacity [23] which the channel input bit rate cannot exceed so that the channel bit errors can be corrected exponentially fast thus enabling use of the channel for stable real-time state estimation. Others such as Seiler and Sengupta [24] and Sinopoli etal [26] have investigated stability in the context of a packet delivering channel. Such a channel accepts packets but loses each packet with some probability. A received packet can communicate a real number. Seiler and Sengupta use this channel model to study real-time observation and control of a discrete-time, continuous-state Markovian Jump Linear Process (MJLP) and derive a necessary and sufficient condition as a linear matrix inequality (LMI) for stability. Sinopoli et. al. [26]use a time-varying Kalman filter as the decoder and derive a bound on the channel error probability limiting stable real-time estimation.

The optimal observation literature e.g., see [30], [12], [28], focuses on finding the optimal encoder-decoder pair. This paper is also focused on optimality. Like us, Gupta and Murray [12], as well as Xu and Hespanha [34] have proposed designs for remote estimation of multiple dynamical systems, but assume communication is based on packets containing real numbers. Gupta and Murray [12] derive an optimal MMSE encoder for a discrete time lossy channel similar to [24], [26], [34]. A packet is either lost or delivered in its entirety. Since the packet bears a real number it can convey the accumulated source state innovation. The arrival of the latest packet gives all the past source innovation and washes away all previous channel errors .This assumption is required for the encoder structure in [12]. The results in this paper show the optimal encoder structure to be different without the assumption. Xu and Hespanha [34] derive the minimum required packet rate to achieve a stable mean-squared error (MSE) when tracking an unstable LTI source using the error-dependent transmission. The packet channel [24], [26], [34] is widely analyzed in control theory as evidenced by the literature survey on Networked Control Systems [13]. We study the same performance measure for the continuous time continuous state AWGN channel of information theory. In the control literature on real-time observation over information-theoretic channel models, e.g., [23], [27], [2], [30], [28], a packet communicating a real number is ill-posed since it would bear infinite information.

The literature on optimal observation over information-theoretic channel models is more closely related to this paper. Walrand and Varaiya [30] derived an optimal information structure for the encoder. A discrete-time discrete-state Markov source is observed over an unreliable discrete- time channel but with a perfect feedback channel from decoder to encoder, i.e., the encoder knows exactly what has been received by the decoder. Their cost function is defined as a measure over the true state and estimated state. The measurability of the encoder in the current source state is shown to be sufficient for optimal real-time state estimation. This formulation is extended by Teneketzis [28] with the channel feedback removed. The optimal encoder in [28] turns out to be a function of current source state and the probability measure over the decoder memory (as a sufficient statistic at the encoder for the decoders knowledge of the source). The optimal encoder can be computed by dynamic programming. This paper is also on the channel without feedback and builds on these results in two ways. We model the dynamical system as a continuous-time continuous state diffusion with an encoder communicating continuously over a continuous-time AWGN channel as conventional in information-theory [25]. This allows us to go beyond the algorithmic characterization by Teneketzis [28], to closed-form characterizations of the optimal encoders and decoders using results in non-linear filtering theory. Secondly, we go beyond the observation of a single dynamical system to observing multiple dynamical systems over a shared channel.

Our single dynamical system results may also be compared to results by Charalambous and

Farhadi [2]. Like this paper, they investigate an Ito diffusion process observed by a remotely located estimator over an AWGN channel. Their first result concerns the AWGN channel without feedback. The plant has to be remotely estimated and stabilized by control. The result asserts that if a stabilizing controller, linear encoder, and decoder exist then an inequality involving the eigenvalues of the plant and the information-theoretic channel capacity must hold. Corollary 4 in this paper shows that for unstable plants and some stable plants no stable encoder-decoder pair exists. The stable plants without stable encoder decoder pairs are intuitively those with high gain. See Corollary 4 for the precise characterization. Moreover, since we do not assume linearity, this negative result holds for non-linear encoders as well. This suggests the premises of theorem 3.1 in [2] can only be satisfied for low-gain plants. Charalambous and Farhadi then characterize the optimal linear encoder and decoder for a scalar diffusion over an AWGN channel with perfect feedback. Our single dynamical system results do the same but for an AWGN channel without feedback. Their optimal encoder turns out to be a scaled version of the difference between source state and decoder output. Our optimal encoder turns out be a scaled innovation encoder. This show the presence or absence of channel feedback is significant as it changes the optimal encoder-decoder structure. Charalambous and Farhadi then generalize to the vector diffusion while we generalize to the multiple access channel. Their results generalize naturally to the vector case, while our single source results do not. We need to assume some correlation over the sources. This again suggests that estimating the state of a vector-valued dynamical system is fundamentally different from estimating the state of multiple dynamical systems. Finally, the proof techniques are different. Charalambous and Farhadi reach their minimum mean-square error (MMSE) designs by maximizing mutual information. The connection to MMSE estimation is by Tyrone Duncan [5]. We work directly on the MMSE estimation problem using results in filtering theory.

This paper is focused on the real-time estimation problem. We do not discuss corresponding real-time control problems. Our system architecture assumes there is a noisy channel between plant or sensor and controller, but not between controller and actuator. The same assumption is made in all the literature reviewed above. When the communication channel exists between plant (source) and estimator (decoder) and not between actuator and plant, the encoders and decoders are restricted to linear functions, and the quadratic cost criterion is used to optimize the control loop, the estimation and control problems are separable without loss of optimality by Assertion 7, section E of [33]. In this setting, the focus on the estimation problem is without loss of generality and hence there are several papers focused on the estimation problem alone [31], [34], [23], [30], [28]. Certainty equivalent control is optimal. The loss of generality when these assumptions are relaxed is also known [32], [29].

We arrived at the theoretical formulation in this paper through work on the design of inter-vehicle communications for collision warning [14], [15], [16], [7]. Cars use WiFi like channels to broadcast their GPS position, speed, and heading. Each car receives this information from cars around it and estimates the motion state of its neighboring cars. This motion state then drives algorithms that select certain proximate cars as collision threats and warn the driver if necessary. The estimates must be real-time since they are used for collision warning. Thus each car has a decoder like the one in figure 2. The signals received by this decoder are the superposition of the signals sent by the encoders in the neighboring cars as in the figure. The similarities end there. In this paper an AWGN channel replaces WiFi and Ito diffusions replace the cars, to enable theoretical insight. The inter-vehicle communication problem also led us to the value of analyzing channels without feedback. The common broadcast wireless communication protocols have no feedback [4] and broadcast is widely used in many sensor and vehicular networks. Perfect channel feedback is a suitable model for CSMA/CD on wired ethernets, the dominant type of local area network before the rise of WiFi.

Finally, we use a continuous-time continuous state model like [2], though modern computer networks are digital. Our aim was to find out if the optimal encoder and decoder design problem could be addressed using the mathematics of optimal filtering familiar to control engineers (see for example Kailath 1971 [9]) , and this paper is the result. To enable this we move away from the band-limited channel in Shannon's work to the non-bandlimited channel used by Wiener to advance communication theory. Our channel is power constrained.

In this paper, we focus on finding the optimal encoder-decoder design for real-time state estimation. We start with a continuous-time, continuous-state source and formulate the MMSE real-time state estimation problem. Like [28], we study the formulation with no channel feedback. For a single dynamical system being estimated in real-time over a scalar AWGN channel, we prove a scaled innovation encoder is optimal. The scaling serves to utilize all available power. The optimal decoder simply scales the observation. The main result is Theorem 6. Theorem 10 is the main result for the general case. We characterize the optimal decoder and show the innovation encoders are optimal, though we are unable to obtain them in closed form. Next we study two dynamical systems sharing the channel. The scaled innovation encoders and scaled observation decoder are once again optimal, but this time the theorem needs a correlation condition on the two systems. The main result is Theorem 11. Theorem 13 provides bounds on the optimal costs without the correlation condition. We obtain our results by leveraging non-linear filtering theory [9], [21]. The paper is organized as follows. The problem of estimating one linear dynamical system over a scalar AWGN channel is formulated in Section II along with our results on the MMSE optimal encoder-decoder pair and the optimal costs. The multiple-access AWGN channel formulation is presented in Section III followed by the problem of estimating two identical linear systems. The results are also in this section. Section IV summarizes the paper. Some of the results in this paper appear in [17] without proof.

It may be noted that the cost function is the integral of mean-square estimation error over the time interval $[0, 1.]$ The structure of the optimal encoders and decoders in Theorems 6,10, 11, and 13, shows they would remain optimal if the upper limit of integration were changed. For example, the same encoders and decoders in an interval $[0, t]$ minimize the integral of mean square estimation error from 0 to t.

II. One-to-One Channel Formulation

Figure 1 shows the analyzed framework. We consider a finite time horizon problem with time $t \in [0, 1]$. Let C be the Banach space of all continuous functions $z : [0, 1] \to \mathbf{R}$ with norm $\|z\| = \max\{|z(t)| : 0 \leq t \leq 1\}$, where $|r|$ is the Euclidean norm of $r \in \mathbf{R}$. Let Γ_t be the smallest σ-field of subsets of C which contains all sets of the form $\{z|z(\tau) \in \beta\}$ where $\tau \in [0, t]$ and β is a Borel subset of $\mathbf{R}$. Let $\Gamma = \Gamma_1$. Let B be the set of the Borel measurable subsets of $[0, 1]$.

The source is a specific type of *Itô process* (e.g., see [9], [18], [20]) with the differential structure in eq. (1). More specifically, the source process x^t is a scalar linear continuous-time process described by the stochastic differential equation:

$$dx^t = a^t x^t dt + dw^t, x^0 \sim \mathbf{N}(0, \Lambda^0) \tag{1}$$

where $t \in [0, 1]$ is the time, $x^t \in \mathbf{R}$ is the process state, $a^t \in \mathbf{R}, a^t > 0$ is the amplification factor, and $w^0 = 0, w^t \in \mathbf{R}$ is a Wiener process [20], [6]. that steers the source process. For $0 \leq s < t \leq 1$, $w^t - w^s \sim \mathbf{N}(0, V(t-s))$ for $V > 0$. The dw^t in eq. (1) is defined as

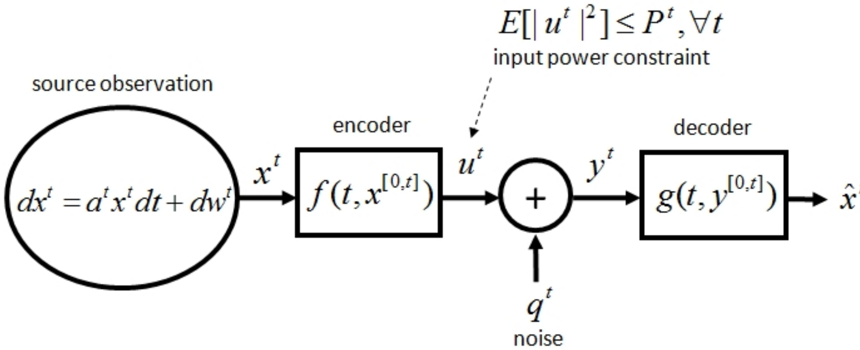

Fig. 1. Analyzed one-to-one channel formulation: the real-time tracking of a scalar linear process over an AWGN channel *without* feedback

$dw^t \equiv w^{t+dt} - w^t$ and thus $dw^t \sim \mathbf{N}(0, V dt)$. The trajectory of a Wiener process is known to be continuous but not differentiable *almost everywhere* [20], [6]. However, a generalized sense of derivative (defined with the integration by parts) for Wiener process can be shown to be a white Gaussian noise [20]. In other words, a Wiener process can be viewed as the limiting behavior of a random walk consisting of small independent increments. In our formulation, this Wiener process w^t is equipped with a generalized sense of derivative $\frac{d}{dt}w^t$, denoted as $v^t \equiv \frac{d}{dt}w^t$ for convenience. This white Gaussian process v^t has $E[v^t] = 0$ and $E[v^t v^s] = V\delta(t-s)$, for $t, s \in (0, 1]$. Throughout this paper, this process v^t is referred as the *innovation* of the source process. The initial condition x^0 is assumed to be zero-mean and Gaussian distributed with variance $\Lambda^0 > 0$. In addition, the Wiener process $w^{[0,1]}$ and the innovation process $v^{(0,1]}$ are assumed to be independent of the initial state x^0.

The process x^t in eq. (1) is a Gaussian process [6].. The probability measure on the state trajectory $x^{[0,t]}$ is induced by the probability measure on the initial state x^0 and the Wiener measure on the trajectory $w^{[0,t]}$. The probability measure on $x^{[0,t]}$ is *absolutely continuous* with respect to the Wiener measure on the space C of all continuous functions from $[0, 1]$ into $\mathbf{R}$ (e.g., see Lemma 2 and Corollary 2 of [5]). The source process x^t is observed perfectly by the causal measurable encoder function at time t, denoted as f, which produces u^t as the input to the AWGN channel at time t. Let $f : [0, 1] \times C \to \mathbf{R}$ be causal, i.e., f is not only $B \otimes \Gamma$ measurable but also measurable with respect to Γ_t for each fixed $t \in [0, 1)$. The encoder f can be parameterized as:

$$u^t = f(t, x^{[0,t]}) \tag{2}$$

where $x^{[0,t]}$ represents the trajectory of the source process as the full history up to time t. The channel input $u^t \in \mathbf{R}$ is chosen to deliver information to the decoder so that the decoder can produce an estimate of the state of the source process. The encoder function f can be time-varying and hence the parameter t. With the given trajectory of $x^{[0,t]}$, the trajectory of $w^{[0,t]}$ can be perfectly reconstructed for each t (e.g., see the discussion on pp. 355-358 and Theorem 1 of [5]). Therefore, $f(t, \cdot)$ is measurable with respect to the σ-algebra generated by the trajectory of $x^{[0,t]}$ and the trajectory of $w^{[0,t]}$ (and thus $v^{(0,t]}$ according to the generalized sense of derivative in [6]).

The AWGN channel output is the superposition of channel input u^t and the channel noise q^t:

$$y^t = u^t + q^t \tag{3}$$

where $y^t \in \mathbf{R}$ is the channel output at time t, and $q^t \in \mathbf{R}$ is a white Gaussian noise process independent of the source initial state x^0, the Wiener process $w^{[0,1]}$, and the innovation process $v^{(0,1]}$. The noise process q^t has $E[q^t] = 0$ and $E[q^t q^s] = Q^t \delta(t-s)$, $Q^t > 0$, for $t, s \in [0,1]$. We choose the formulation in eq. (3) instead of a stochastic differential form (e.g., eq. (1)) to match with most definitions of an AWGN channel [11], [10]. The channel output y^t is observed by the decoder, denoted as g, which produces a real-time estimate of current state of the source process x^t. Let $g : [0,1] \times C \to \mathbf{R}$ be causal, i.e., g is not only $B \otimes \Gamma$ measurable but also measurable with respect to Γ_t for each fixed $t \in [0,1)$. The decoder g can be parameterized as:

$$\hat{x}^t = g(t, y^{[0,t]}) \tag{4}$$

where $\hat{x}^t \in \mathbf{R}$ is the estimate of x^t and $y^{[0,t]}$ is the full history of observed channel outputs up to time t. In this real-time tracking formulation, causality is imposed on the encoder f and decoder g. Once this estimate $\hat{x}^t$ is produced at time t, it is final and can not be improved later based on future channel outputs $y^{(t,1]}$. This decoder function g can also be time-varying and hence the parameter t.

The cost function $J_{(1)}$ is defined over a finite time horizon $t \in [0,1]$:

$$J_{(1)} \equiv \int_0^1 E[|x^t - \hat{x}^t|^2] dt \tag{5}$$

of which the expectation is taken with respect to the probability measure of the initial condition x^0, the Wiener process $w^{[0,1]}$, and the white Gaussian channel noise process $q^{[0,1]}$. For a given encoder-decoder pair, the real-time MSE $E[|x^t - \hat{x}^t|^2]$ can be calculated for each t. This cost function $J_{(1)} \in \mathbf{R}, J_{(1)} \geq 0$ is essentially the Lebesgue integral of the real-time MSE over the unit time interval $[0, 1]$.

We define a finite time horizon optimization problem based on the cost function $J_{(1)}$ in eq. (5) to find an encoder-decoder pair, denoted as $\{f, g\}$, minimizing the cost function $J_{(1)}$:

$$\{f^*, g^*\} = \underset{E[|u^t|^2] \leq P^t, \forall t \in [0,1]}{\operatorname{argmin}} J_{(1)} \tag{6}$$

where $\{f^*, g^*\}$ denotes the optimal encoder-decoder pair and $P^t < \infty$ is the finite power constraint on the channel input at each time t. The total transmission energy allowed is also finite over the same time horizon $t \in [0, 1]$, i.e., $\int_0^1 P^t dt < \infty$. This power constraint P^t avoids the degenerate solution in which the encoder amplifies the channel input signal to an arbitrarily large value to *overwhelm* the channel noise q^t and the decoder can then restore the signal with an arbitrarily small error.

A. Preliminaries for the One-to-One Channel Formulation

Let $\bar{g} : [0, 1] \times C \to \mathbf{R}$ be causal, i.e., $\bar{g}$ is not only $B \otimes \Gamma$ measurable but also measurable with respect to Γ_t for each fixed $t \in [0, 1)$. We define a differential decoder for $\hat{x}^t$ (the real-time estimate of the source process) by the stochastic differential equation:

$$d\hat{x}^t = a^t \hat{x}^t dt + \bar{g}(t, y^{[0,t]})dt, \hat{x}^0 = \bar{g}(0, y^0) \tag{7}$$

where a^t is the same amplification factor in eq. (1) and the function $\bar{g}(t, y^{[0,t]})$ steers the evolution of $\hat{x}^t$. The similarity between eq. (1) and eq. (7) acknowledges the rationale that the decoder output $\hat{x}^t$ in eq. (7) incorporates the knowledge of the model of the source process x^t.

Lemma 1: For the real-time tracking problem in (6), there is no loss of optimality (in the MMSE sense for each $t \in [0, 1]$) by assuming the differential decoder in eq. (7).

Proof: The form of the differential decoder in eq. (7) is MMSE optimal according to the nonlinear filtering analysis by Clark [3], Frost and Kailath [9], and Lo [21] (e.g., see Theorem 3 in [3], Theorem 3-5 in [9], and Theorem 1 in [21]. The finite channel input energy and Gaussian channel noise are the key elements guaranteeing the optimality of the differential decoder form in eq. (7). See specifically the discussion of Gauss-Markov Process and Linear Case on pp. 222 in [9] . ■

Theorem 2: Let $\Sigma^t \equiv E[(x^t - \hat{x}^t)^2]$ denote the tracking MSE at time t. For the optimal differential decoder (defined in eq. (7)), denoted as $\bar{g}^*(t, y^{[0,t]})$, the MMSE at time t, denoted as Σ^{*t}, can be described by the following differential equation. For $t \in (0, 1]$,

$$d\Sigma^{*t} = 2a^t\Sigma^{*t}dt + E[(dw^t - \bar{g}^*(t, y^{[0,t]})dt)^2] \tag{8}$$

with the initial condition, $\Sigma^{*0} = E[(x^0 - \bar{g}^*(0, y^0))^2]$.

Proof: Let $e^t \equiv x^t - \hat{x}^t$. We get the differential equation of e^t based on eq. (1) and eq. (7):

$$de^t = a^te^tdt + dw^t - \bar{g}(t, y^{[0,t]})dt. \tag{9}$$

With eq. (9) and the definition $\Sigma^t = E[(e^t)^2]$, the differential of the MMSE Σ^{*t} for the optimal differential decoder $\bar{g}^*$ can be derived as follows:

$d\Sigma^{*t} = \Sigma^{*(t+dt)} - \Sigma^{*t}$

$= E[(e^{*t} + de^{*t})^2] - E[(e^{*t})^2]$

$= 2E[e^{*t}de^{*t}] + E[(de^{*t})^2]$

$= 2a^tE[(e^{*t})^2]dt + 2E[e^{*t}(dw^t - \bar{g}^*(t, y^{[0,t]})dt)] + E[(a^te^{*t}dt + dw^t - \bar{g}^*(t, y^{[0,t]})dt)^2]$

The last step is derived by substituting $de^{*t} = a^te^{*t}dt + dw^t - \bar{g}^*(t, y^{[0,t]})dt$.

$= 2a^tE[(e^{*t})^2]dt - 2E[e^t\bar{g}^*(t, y^{[0,t]})dt] + (a^t)^2E[(e^{*t})^2](dt)^2 + E[(dw^t - \bar{g}^*(t, y^{[0,t]})dt)^2]$
$-2a^tE[e^{*t}\bar{g}^*(t, y^{[0,t]})dt]dt$

The last step is true because the Wiener process has independent increments and thus dw^t is independent of e^{*t}.

$= 2a^tE[(e^{*t})^2]dt + (a^t)^2E[(e^{*t})^2](dt)^2 + E[(dw^t - \bar{g}^*(t, y^{[0,t]})dt)^2]$

This is due to a general form of Orthogonality Principle: the optimal tracking error is orthogonal to any function of all the past observations, e.g., see pp. 268 in [22].

$= 2a^tE[(e^{*t})^2]dt + E[(dw^t - \bar{g}^*(t, y^{[0,t]})dt)^2]$

The last step is derived by neglecting the higher order term associated with $(dt)^2$.

which leads to the differential equation in (8). ■

Corollary 3: For the real-time tracking problem in (6), an MMSE optimal differential decoder $\bar{g}^*$ can be expressed, for $t \in (0, 1]$, as

$$\bar{g}^*(t, y^{[0,t]}) = E[v^t|y^{[0,t]}] \tag{10}$$

with the initial condition, $\bar{g}^*(0, y^0) = E[x^0|y^0]$.

Proof: It has been noted that conditional expectation given observations minimizes the MSE (e.g., see the discussion on pp. 218 in [9]). From the differential form in eq. (8), observe that the MMSE Σ^{*t} can be minimized if all the previous $E[(dw^\tau - \bar{g}^*(\tau, y^{[0,\tau]})d\tau)^2]$ are minimized for all $0 < \tau \leq t$ and $E[(x^0 - \bar{g}^*(0, y^0))^2]$ is minimized for the initial condition. The differential decoder function $\bar{g}(t, \cdot)$ has all the channel observations $y^{[0,t]}$ and can be chosen for each t independently. This $\bar{g}(t, \cdot)$ does not depend on previous decoder functions $\bar{g}(\tau, \cdot), 0 \leq \tau < t$.

Because the conditional expectation $E[dw^\tau|y^{[0,\tau]}]$ minimizes this expectation for each τ, we get $\bar{g}^*(\tau, y^{[0,\tau]})d\tau = E[dw^\tau|y^{[0,\tau]}]$ and thus $\bar{g}^*(\tau, y^{[0,\tau]}) = E[\frac{d}{d\tau}w^\tau|y^{[0,\tau]}]$. Since the innovation process v^t is defined as the generalized sense of derivative of w^t, i.e., $v^t \equiv \frac{d}{dt}w^t$, we get eq. (10). At the initial condition, $t = 0$, the decoder is the conditional expectation of the source initial state x^0 given the available channel observation y^0. ■

Corollary 4: For the real-time tracking problem in (6) and $a^t \geq 0.5$, for each t, it is impossible to have asymptotic stable MSE, i.e., $\lim_{t\to\infty} \Sigma^t \to \infty$ for any encoder-decoder pair.

Proof: First partition infinite time line into segments of unit time length as in our formulation. In each unit time segment, the same differential equation (8) for Σ^{*t} still applies but with different initial conditions. Define the variable Ψ_1^t by

$$d\Psi_1^t = 2a^t\Psi_1^t dt, \Psi_1^0 = \Sigma^{*0} \tag{11}$$

and thus $\Psi_1^t \leq \Sigma^{*t}, \forall t > 0$ by comparing eq. (11) and eq. (8) and the fact that $E[(dw^t - \bar{g}^*(t, y^{[0,t]})dt)^2] > 0$ for nontrivial channel noise q^t, $\forall t > 0$. Now define another variable Ψ_2^t by

$$d\Psi_2^t = \Psi_2^t dt, \Psi_2^0 = \Sigma^{*0} \tag{12}$$

and thus, if $2a^t \geq 1, \forall t$, $\Psi_2^t \leq \Psi_1^t \leq \Sigma^{*t}, \forall t > 0$. The solution of the differential equation (12) can be expressed as $\Psi_2^t = \exp(t)\Psi_2^0$. Since $\Psi_2^0 = E[(x^0 - \bar{g}^*(0, y^0))^2] > 0$ for nontrivial channel noise q^0,

$$\infty = \lim_{t\to\infty} \Psi_2^t \leq \lim_{t\to\infty} \Psi_1^t \leq \lim_{t\to\infty} \Sigma^{*t}.$$

Therefore, in our formulation *without* channel feedback, if $2a^t \geq 1, \forall t$, no encoder-decoder pair can achieve stable MSE tracking asymptotically. ■

B. Main Results for the One-to-One Channel Formulation

In this subsection, we state an optimal encoder-decoder pair for the one-to-one channel formulation (as shown in Figure 1) and the optimal real-time tracking performance. With the Gaussian distributed source innovation v^t, the encoder is shown to be a linear innovation encoder and the associated decoder is a differential decoder that steers the state estimate $\hat{x}^t$ according to eq. (7).

Lemma 5: For the real-time tracking problem in (6), one optimal form of the encoder function and the differential decoder function (defined in eq. (7)) can be expressed as follows: for each $t \in (0, 1]$, the encoder is a function of v^t, denoted as $\bar{f}^*(t, v^t)$, and the associated differential decoder is denoted as $\bar{g}^*(t, y^t) = E[v^t|y^t]$. For the initial condition, $t = 0$, one optimal form can be written as $\bar{f}^*(0, x^0)$ and $\bar{g}^*(0, y^0) = E[x^0|y^0]$.

Proof: Assume an optimal encoder-decoder pair at t: $f^*(t, \cdot)$ and $\bar{g}^*(t, \cdot)$. The optimal differential decoder form $\bar{g}^*$ from Corollary 3 is:

$\bar{g}^*(t, y^{[0,t]}) = E[v^t|y^{[0,t]}] = E[v^t|y^t]$

Since v^t is independent of x^0, $v^{(0,t)}$, past channel outputs $y^{[0,t)}$ do not contain any information about v^t due to causality.

$= E[v^t|q^t + u^t]$

$= E[v^t|q^t + f^*(t, \{x^0, v^{(0,t)}, v^t\})]$

The source process, encoder f^*, noisy channel output, and decoder $\bar{g}^*$ form a Markov chain: $\{x^0, v^{(0,t)}, v^t\} \rightarrow u^t \rightarrow y^t \rightarrow E[v^t|y^t]$.

$= E[v^t|q^t + \bar{f}^*(t, v^t)]$ for some function $\bar{f}^*(t, \cdot)$

Since v^t is independent of x^0, $v^{(0,t)}$, v^t is as *informative* as the set $\{x^0, v^{(0,t)}, v^t\}$ for the decoder to estimate v^t, e.g., see Theorem 3 and 4 in [1]. Thus one can design another function $\bar{f}^*(t, \cdot)$ to focus on delivering v^t without loss of optimality.

Therefore, without loss of optimality, the encoder at time $t \in (0, 1]$ can focus on delivering v^t to help the decoder better estimate the innovation v^t to steer the state estimate $\hat{x}^t$ according to eq. (7). The decoder can be a function of current channel observation y^t mainly due to causality. ∎

Theorem 6: For the real-time tracking problem in (6), one optimal encoder function $\bar{f}^*$ and its matched differential decoder $\bar{g}^*$ can be expressed, for $t \in (0, 1]$, as:

$$\bar{f}^*(t, v^t) = v^t \sqrt{\frac{P^t}{V}} \tag{13}$$

and

$$\bar{g}^*(t, y^t) = y^t \frac{\sqrt{VP^t}}{P^t + Q^t}. \tag{14}$$

At the initial condition, $t = 0$, the optimal form can be written as $\bar{f}^{*0}(x^0) = x^0\sqrt{\frac{P^0}{\Lambda^0}}$ and $\bar{g}^{*0}(y^0) = y^0\frac{\sqrt{\Lambda^0 P^0}}{P^0+Q^0}$.

Proof: Based on Lemma 5, the optimal encoder scales the Gaussian distributed innovation v^t to match with the channel input power constraint P^t (see pp. 561-562, 564 in Goblick [11] and pp. 1153 in Gastpar [10]), and this direct transmission can minimize the term $E[(dw^t - \bar{g}^*(t, y^{[0,t]})dt)^2]$ in eq. (8) for each $t \in (0, 1]$ and thus achieves MMSE optimal. The optimal differential decoder in eq. (14) is the conditional expectation of the source innovation v^t given y^t and the linear innovation encoder in eq. (13). The optimal encoder/decoder pair at $t = 0$ (i.e., the initial condition) follows similarly. ∎

Corollary 7: For the real-time tracking problem in (6), the optimal cost $J^*_{(1)}$ achievable is given by

$$J^*_{(1)} = \int_0^1 \Sigma^{*t} dt \tag{15}$$

where Σ^{*t} can be described by the differential equation:

$$\frac{d}{dt}\Sigma^{*t} = 2a^t\Sigma^{*t} + V(\frac{Q^t}{P^t + Q^t})^2, \Sigma^{*0} = \frac{\Lambda^0 Q^0}{P^0 + Q^0}. \tag{16}$$

Proof: Based on the optimal encoder and differential decoder in Theorem 6,

$E[(dw^t - \bar{g}^*(t, y^{[0,t]})dt)^2]$
$= E[(dw^t - (v^t\sqrt{\frac{P^t}{V}} + q^t)\frac{\sqrt{VP^t}}{P^t+Q^t}dt)^2]$
$= E[(dw^t - v^t dt\frac{P^t}{P^t+Q^t} - q^t dt\frac{\sqrt{VP^t}}{P^t+Q^t})^2]$
$= E[(dw^t\frac{Q^t}{P^t+Q^t} - q^t dt\frac{\sqrt{VP^t}}{P^t+Q^t})^2]$
$= E[(dw^t\frac{Q^t}{P^t+Q^t})^2] + E[(q^t dt\frac{\sqrt{VP^t}}{P^t+Q^t})^2]$
The last step is true because dw^t is independent of the channel noise q^t.
$= (\frac{Q^t}{P^t+Q^t})^2 E[(dw^t)^2] + \frac{VP^tQ^t}{(P^t+Q^t)^2}(dt)^2 = (\frac{Q^t}{P^t+Q^t})^2 V dt$
The last step is derived by neglecting the higher order term associated with $(dt)^2$ and then applying the property of the Wiener process $E[(dw^t)^2] = Vdt$.

Combining the above with eq. (8), we get the differential form in eq. (16). The derivation for the initial condition at $t = 0$ is as follows:

$E[(x^0 - \bar{g}^*(0, y^0))^2] = E[(x^0 - (x^0\sqrt{\frac{P^0}{\Lambda^0}} + q^0)\frac{\sqrt{\Lambda^0 P^0}}{P^0+Q^0})^2]$
$= E[(x^0\frac{Q^0}{P^0+Q^0} - q^0\frac{\sqrt{\Lambda^0 P^0}}{P^0+Q^0})^2]$
$= E[(x^0\frac{Q^0}{P^0+Q^0})^2] + E[(q^0\frac{\sqrt{\Lambda^0 P^0}}{P^0+Q^0})^2]$

The last step is true because the initial state x^0 is independent of the channel noise q^0.
$= (\frac{Q^0}{P^0+Q^0})^2\Lambda^0 + \frac{\Lambda^0 P^0}{(P^0+Q^0)^2}Q^0 = \frac{\Lambda^0 Q^0}{P^0+Q^0}$

which leads to the Σ^{*0} term in eq. (16). ■

The simple form of the optimal encoder-decoder pair in eq. (13) and eq. (14) is mainly due to the fact that the Gaussian distributed source innovation v^t and Gaussian channel noise q^t are *matched* (e.g., see the discussion on pp. 1152-1153 in Gastpar [10]). One can also observe from eq. (14) that, when the transmission power is zero (i.e., $P^t = 0$, no communication at all), the differential decoder will propagate the state estimate $\hat{x}^t$ in eq. (7) with $g^* = E[v^t] = 0$ and the MMSE is given by $\frac{d}{dt}\Sigma^{*t} = 2a^t\Sigma^{*t} + V$ as derived in [20]. The reduction from V to $V(\frac{Q^t}{P^t+Q^t})^2$ in eq. (16) can be viewed as the *benefit* when the encoder communicates v^t to the decoder.

III. Multiple-Access Channel Formulation

In this section, we extend the one-to-one channel formulation in Section II to a multiple-access channel formulation. There are two main differences: 1) $n \in \mathbf{N}$, $n \geq 2$ scalar linear processes are observed by n encoders individually which produce n channel inputs into a multiple-access scalar AWGN channel; and 2) a single decoder reads the channel output and produces a vector of estimates for all n source processes in real-time. We are looking for an optimal set of encoders and decoder that achieves Minimum Sum of MSE (MSMSE) for tracking all the source processes. Figure 2 illustrates this formulation for $n = 2$.

Similar to the *Itô process* defined in eq. (1), each source process $x_i^t, i = 1, ..., n$, is a scalar linear continuous-time process described by the stochastic differential equation:

$$dx_i^t = a_i^t x_i^t dt + dw_i^t, x_i^0 \sim \mathbf{N}(0, \Lambda_i^0) \tag{17}$$

where $t \in [0, 1]$ is the time, $x_i^t \in \mathbf{R}$ is the process state, $a_i^t \in \mathbf{R}$ is the amplification factor, and $w_i^t \in \mathbf{R}$ is a Wiener process that steers the i-th source process. For $0 \leq s < t \leq 1$, $w_i^t - w_i^s \sim \mathbf{N}(0, V_i(t - s))$ for $V_i > 0$. This Wiener process w_i^t is equipped with a generalized

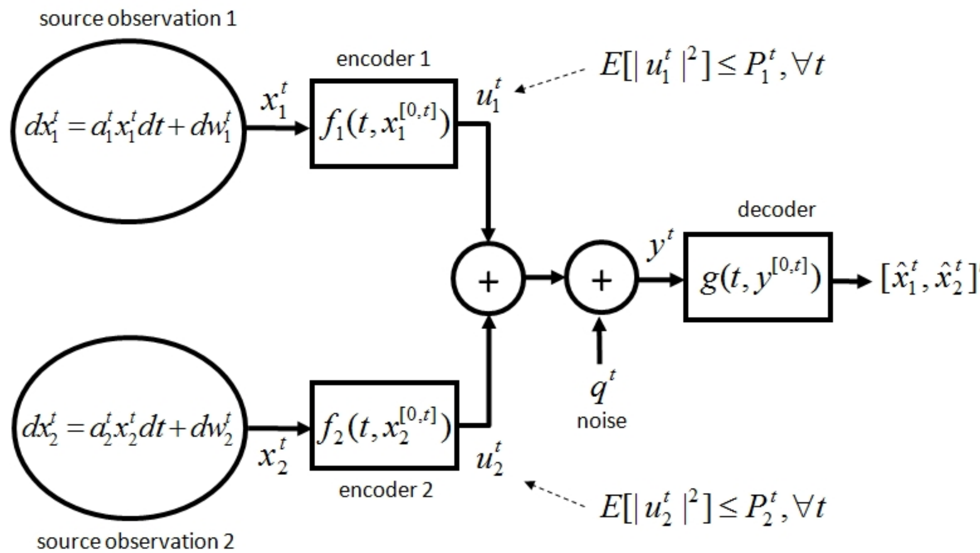

Fig. 2. Analyzed multiple-access channel formulation: the real-time tracking of $n = 2$ linear processes over an AWGN channel *without* feedback

sense of derivative $\frac{d}{dt}w_i^t$, denoted as the i-th innovation process $v_i^t \equiv \frac{d}{dt}w_i^t$. This white Gaussian process v_i^t has $E[v_i^t] = 0$ and $E[v_i^t v_i^s] = V_i \delta(t-s)$, for $t, s \in (0,1]$. Within the n sources, for the same time index $t \in (0,1]$ the v_i^t can be correlated with v_j^t for $i, j = 1, ..., n$ with the correlation coefficient ρ_{ij}^t, $i \neq j$, and $-1 \leq \rho_{ij}^t \leq 1$. The initial condition x_i^0 is assumed to be zero-mean and Gaussian distributed with variance $\Lambda_i^0 > 0$. Among n sources, the initial condition x_i^0 can be correlated with x_j^0 for $i, j = 1, ..., n$ with the correlation coefficient ρ_{ij}^0, $i \neq j$, and $-1 \leq \rho_{ij}^0 \leq 1$. The probability measure induced by each source x_i^t is similar to that of the single source x^t in Section II.

Each source process $x_i^t, i = 1, ..., n,$ is observed perfectly by the i-th causal measurable encoder function f_i which produces $u_i^t \in \mathbf{R}$ as the i-th input to the AWGN channel. Let $f_i : [0,1] \times C \to \mathbf{R}$ be causal, i.e., f_i is not only $B \otimes \Gamma$ measurable but also measurable with respect to Γ_t for each fixed $t \in [0,1)$. The encoder f_i can be parameterized as:

$$u_i^t = f_i(t, x_i^{[0,t]}) \tag{18}$$

where $x_i^{[0,t]}$ represents the trajectory of the i-th source process as the full history up to time t. Note that the information structure of our formulation only allows the i-th encoder to observe the i-th source process.

The AWGN channel output is the superposition of the sum of the channel inputs $\sum_{i=1}^{n} u_i^t$ and the channel noise q^t:

$$y^t = \sum_{i=1}^{n} u_i^t + q^t \tag{19}$$

where $y^t \in \mathbf{R}$ is the channel output at time t, and $q^t \in \mathbf{R}$ is a white Gaussian noise process independent of $x_i^0, i = 1, ..., n$, $w_i^{[0,1]}, i = 1, ..., n$, and $v_i^{(0,1]}, i = 1, ..., n$. The noise process q^t has $E[q^t] = 0$ and $E[q^t q^s] = Q^t \delta(t-s)$, $Q^t > 0$, for $t, s \in [0,1]$. The channel output y^t is observed by the decoder g to produce a vector of estimates of $x_i^t, i = 1, ..., n$. Let $g : [0,1] \times C \rightarrow \mathbf{R}_{n \times 1}$ be causal, i.e., g is not only $B \otimes \Gamma$ measurable but also measurable with respect to Γ_t for each fixed $t \in [0,1)$. The decoder g can be parameterized as:

$$[\hat{x}_1^t, ..., \hat{x}_n^t]' = g(t, y^{[0,t]}) \tag{20}$$

where each $\hat{x}_i^t \in \mathbf{R}$ is the estimate of x_i^t and $y^{[0,t]}$ is the full history of observed channel outputs up to time t.

The cost function $J_{(n)}$ is defined over a finite time horizon $t \in [0,1]$:

$$J_{(n)} \equiv \int_0^1 \sum_{i=1}^{n} E[|x_i^t - \hat{x}_i^t|^2] dt \tag{21}$$

where n is the number of source processes and the expectation is taken with respect to the probability measure of the initial condition $x_i^0, i = 1, ..., n$, the Wiener processes $w_i^{[0,1]}, i = 1, ..., n$, and the channel noise process $q^{[0,1]}$. With a given set of encoders-decoder, the MSE $E[|x_i^t - \hat{x}_i^t|^2], \forall i$ can be calculated for each t. This cost function $J_{(n)} \in \mathbf{R}$, $J_{(n)} \geq 0$ is essentially the Lebesgue integral of the sum of all the real-time tracking MSE over the unit time interval $[0,1]$. The subscript in $J_{(n)}$ denotes that this multiple-access formulation considers the tracking of n sources, which is different from the cost function in eq. (5).

We define an optimization problem based on the cost function $J_{(n)}$ in eq. (21). It is to find the set of encoders and decoder for $t \in [0,1]$, denoted as $\{f_1, ..., f_n, g\}$, such that the cost function $J_{(n)}$ is minimized:

$$\{f_1^*, ..., f_n^*, g^*\} = \underset{E[|u_i^t|^2] \leq P_i^t, \forall i, \forall t \in [0,1]}{\operatorname{argmin}} J_{(n)} \tag{22}$$

where f_i^* is the optimal encoder for the i-th source, g^* is the optimal decoder, and $P_i^t < \infty$ is the finite power constraint on u_i^t, the channel input producing the i-th encoder, at time $t \in [0,1]$, and the total energy is also finite: $\sum_{i=1}^{n}\int_0^1 P_i^t dt < \infty$. Again, the finite power and energy constraints avoid the degenerate solution mentioned in Section II.

A. Preliminaries for the Multiple-Access Channel Formulation

Let $\mathbf{x}^t \equiv [x_1^t, ..., x_n^t]'$ denote the vector of the source states and similarly $\mathbf{w}^t \equiv [w_1^t, ..., w_n^t]'$. The source processes can then be described by the stochastic differential equation:

$$d\mathbf{x}^t = \mathbf{a}^t \mathbf{x}^t dt + d\mathbf{w}^t \tag{23}$$

with $\mathbf{a}^t \equiv diag(a_1^t, ..., a_n^t)$ where $diag$ represents a diagonal matrix with the indicated diagonal elements.

Let $\bar{g} : [0,1] \times C \rightarrow \mathbf{R}_{n\times 1}$ be causal, i.e., $\bar{g}$ is not only $B \otimes \Gamma$ measurable but also measurable with respect to Γ_t for each fixed $t \in [0,1)$. We define a differential decoder for the real-time estimate $\hat{\mathbf{x}}^t \equiv [\hat{x}_1^t, ..., \hat{x}_n^t]'$ by the stochastic differential equation:

$$d\hat{\mathbf{x}}^t = \mathbf{a}^t \hat{\mathbf{x}}^t dt + \bar{g}(t, y^{[0,t]})dt, \hat{\mathbf{x}}^0 = \bar{g}(0, y^0) \tag{24}$$

where $\mathbf{a}^t$ is the same amplification factor in eq. (23) and the function $\bar{g}(t, y^{[0,t]})$ produces the $n \times 1$ vector that steers the evolution of $\hat{\mathbf{x}}^t$, the decoder output.

Lemma 8: Given the multiple-access real-time tracking formulation in (22), there is no loss of optimality (in the MSMSE sense for each $t \in [0,1]$) by assuming the differential decoder in eq. (24).

Proof: Our argument here is similar to the proof of Lemma 1. The differential form of eq. (24) is MSMSE optimal according to the nonlinear filtering analysis by Clark [3], Frost and Kailath [9], and Lo [21]. ∎

Theorem 9: Let $\mathbf{\Sigma}^t \equiv E[(\mathbf{x}^t - \hat{\mathbf{x}}^t)(\mathbf{x}^t - \hat{\mathbf{x}}^t)']$ denote the tracking error covariance matrix at time t. The optimal error covariance $\mathbf{\Sigma}^{*t}$ (in the MSMSE sense) can be described by the differential equation:

$$d\mathbf{\Sigma}^{*t} = 2\mathbf{a}^t \mathbf{\Sigma}^{*t} dt + E[\Psi^t (\Psi^t)'] \tag{25}$$

where $\Psi^t \equiv d\mathbf{w}^t - \bar{g}^*(t, y^{[0,t]})dt$. The vector $\mathbf{w}^t$ is the vector of Wiener processes as defined in eq. (23). The function $\bar{g}^*(t, y^{[0,t]})$ is the optimal differential decoder with the form defined in eq. (24).

Proof: Let $\mathbf{e}^t \equiv \mathbf{x}^t - \hat{\mathbf{x}}^t$ denote the error vector. We get the differential equation of $\mathbf{e}^t$ based on eq. (23) and eq. (24):

$$d\mathbf{e}^t = \mathbf{a}^t\mathbf{e}^t dt + d\mathbf{w}^t - \bar{g}(t, y^{[0,t]})dt. \tag{26}$$

With eq. (26) and the definition $\mathbf{\Sigma}^t = E[\mathbf{e}^t(\mathbf{e}^t)']$, the differential of $\mathbf{\Sigma}^{*t}$ with the optimal differential decoder $\bar{g}^*$ can be expressed as eq. (25) using the similar arguments in the proof of Theorem 2. ∎

Theorem 10: Let $tr(\cdot)$ denote the trace of the input matrix. Given the multiple-access real-time tracking formulation in (22), the optimal differential decoder $\bar{g}^*$ that achieves the Minimum Sum of MSE (MSMSE) $tr(\mathbf{\Sigma}^{*t})$ can be expressed, for $t \in (0, 1]$, as:

$$\bar{g}^*(t, y^t) = E[\mathbf{v}^t|y^t] \tag{27}$$

where the innovation vector $\mathbf{v}^t \equiv [v_1^t, ..., v_n^t]'$. One set of optimal form of encoders is the innovation encoders: for $t \in (0, 1]$, $\bar{f}_i^*(t, v_i^t), i = 1, ..., n$. For the initial condition, $t = 0$: one optimal decoder is $\bar{g}^*(0, y^0) = E[\mathbf{x}^0|y^0]$, and one optimal form of the encoders is $\bar{f}_i^*(0, x^0), i = 1, ..., n$.

Proof: Based on the differential form of $\mathbf{\Sigma}^{*t}$ in eq. (25) and similar arguments as in the proof of Corollary 3, to achieve MSMSE $tr(\mathbf{\Sigma}^{*t})$, $\bar{g}^*(t, y^{[0,t]})dt = E[d\mathbf{w}^t|y^{[0,t]}]$ and thus $\bar{g}^*(t, y^{[0,t]}) = E[\frac{d}{dt}\mathbf{w}^t|y^{[0,t]}]$. Since $v_i^t \equiv \frac{d}{dt}w_i^t, \forall i$, we get $\bar{g}^*(t, y^{[0,t]}) = E[\mathbf{v}^t|y^{[0,t]}]$. Based on the arguments in the proof of Lemma 5, $\bar{g}^*(t, y^{[0,t]}) = E[\mathbf{v}^t|y^t]$ due to causality. Since the i-th source innovation can only be observed by the i-th encoder, one optimal set of encoders is for each encoder to focus on delivering the innovation v_i^t at each time t as in the proof of Lemma 5. The initial condition at $t = 0$ follows similarly. ∎

Theorem 10 says that, at each t, the i-th encoder can focus on communicating v_i^t to the decoder. However, how to design optimal encoders to communicate v_i^t over a shared channel to the decoder is an on-going research [19], [8]. In the following, we study a simple case of tracking two identical linear sources and apply recent results in [19] to get the optimal encoders and the optimal real-time tracking performance.

B. Tracking Two Identical Linear Sources over a Shared AWGN Channel

In this subsection, we assume the same multiple access tracking formulation (as shown in Figure 2) with two identical linear sources and the same power constraint: $\forall t$, $a_1^t = a_2^t = a^t$,

$P_1^t = P_2^t = P^t$, $V_1 = V_2 = V$, and $\Lambda_1^0 = \Lambda_2^0 = \Lambda^0$. Theorem 11 characterizes the optimal encoders when the Signal-to-Noise Ratio (SNR) P^t/Q^t is within a range decided by the source correlation factor ρ^t (see eq. (28)). The optimal encoder of the i-th source uses the uncoded transmission of the source innovation $v_i^t, i = 1, 2$, i.e., a linear scaling of the innovation to match with the power P^t. The optimal MSMSE is given in Corollary 12.

Theorem 11: For the real-time tracking problem in (22) with two identical linear sources and, $\forall t \in [0, 1]$,

$$0 < \frac{P^t}{Q^t} < \frac{|\rho^t|}{1 - |\rho^t|^2}, \tag{28}$$

one pair of optimal encoder functions $\bar{f}_1^*$, $\bar{f}_2^*$, and the associated differential decoder $\bar{g}^*$ (as defined in eq. (24)) can be expressed, for $t \in (0, 1]$, as,

$$\bar{f}_i^*(t, v_i^t) = v_i^t \sqrt{\frac{P^t}{V}}, i = 1, 2, \tag{29}$$

and

$$\bar{g}^*(t, y^t) = y^t \frac{\sqrt{VP^t}(1 + |\rho^t|)^2}{2P^t(1 + |\rho^t|) + Q^t}[1, 1]'. \tag{30}$$

For the initial condition, $t = 0$, the optimal form can be written as $\bar{f}_i^{*0}(x_i^0) = x_i^0 \sqrt{\frac{P^0}{\Lambda^0}}, i = 1, 2$ and $\bar{g}^{*0}(y^0) = y^0 \frac{\sqrt{\Lambda^0 P^0}(1+|\rho^0|)^2}{2P^0(1+|\rho^0|)+Q^0}[1, 1]'$.

Proof: Based on Theorem 10, the optimal encoder can focus on transmitting the source innovation v_i^t for each t. With the analysis by Lapidoth and Tinguely [19] (see specifically pp. 2720-2722 and Corollary IV.1 and IV.3 in [19]), as long as the SNR P^t/Q^t satisfies eq. (28), the uncoded transmission of innovation in eq. (29) is optimal to minimize the distortion for each t. The optimal decoder in eq. (30) is the conditional expectation (see pp. 2729 in [19]) given the encoders in eq. (29). The optimal encoder/decoder pair at $t = 0$ (i.e., the initial condition) follows similarly. ■

Corollary 12: For the real-time tracking problem in (22) with two identical linear sources, and $\forall t \in [0, 1]$, SNR P^t/Q^t satisfies eq. (28), the optimal cost $J_{(2)}^*$ achievable is given by

$$J_{(2)}^* = 2\int_0^1 \Sigma^{*t} dt \tag{31}$$

where Σ^{*t} can be described by the differential equation:

$$\frac{d}{dt}\Sigma^{*t} = 2a^t\Sigma^{*t} + V\frac{P^t(1 - |\rho^t|^2) + Q^t}{2P^t(1 + |\rho^t|) + Q^t}, \Sigma^{*0} = \Lambda^0 \frac{P^0(1 - |\rho^0|^2) + Q^0}{2P^0(1 + |\rho^0|) + Q^0}. \tag{32}$$

Proof: Based on Theorem 11 and Theorem 9, the achieved MSMSE can be described as the differential form in eq. (32). The derivation of the distortion of v_i^t for each t can be found on pp. 2722 of [19]. The initial condition $t = 0$ follows similarly. ■

When the SNR P^t/Q^t does not satisfy the condition in eq. (28), the form of the optimal encoder is still under investigation, e.g., see [19], [8]. We provide the upper and lower bounds on the optimal real-time tracking performance.

Theorem 13: For the real-time tracking problem in (22) with two identical linear sources, when the SNR P^t/Q^t does not satisfy the condition in eq. (28), the optimal cost $J^*_{(2)}$ achievable is upper and lower bounded by

$$2\int_0^1 \underline{\Sigma}^{*t} dt \leq J^*_{(2)} \leq 2\int_0^1 \overline{\Sigma}^{*t} dt \tag{33}$$

where $\overline{\Sigma}^{*t}$ can be described by the differential equation:

$$\frac{d}{dt}\overline{\Sigma}^{*t} = 2a^t\overline{\Sigma}^{*t} + V\frac{P^t(1-|\rho^t|^2)+Q^t}{2P^t(1+|\rho^t|)+Q^t}, \overline{\Sigma}^{*0} = \Lambda^0\frac{P^0(1-|\rho^0|^2)+Q^0}{2P^0(1+|\rho^0|)+Q^0}, \tag{34}$$

and $\underline{\Sigma}^{*t}$ can be described by the differential equation:

$$\frac{d}{dt}\underline{\Sigma}^{*t} = 2a^t\underline{\Sigma}^{*t} + V\sqrt{\frac{Q^t(1-|\rho^t|)^2}{2P^t(1+|\rho^t|)+Q^t}}, \underline{\Sigma}^{*0} = \Lambda^0\sqrt{\frac{Q^0(1-|\rho^0|)^2}{2P^0(1+|\rho^0|)+Q^0}}. \tag{35}$$

Proof: This result is based on Theorem 10 and the analysis in [19] (see specifically Corollary IV.1 and IV.3 in [19]): The upper bound $\overline{\Sigma}^{*t}$ can be achieved by uncoded transmission for each t as in Theorem 11. The lower bound $\underline{\Sigma}^{*t}$ can be achieved by a joint encoder function that takes both source innovations as input. ■

It is not yet known theoretically how close a distributed encoder design (i.e., each encoder only observes the corresponding source innovation) can approach the lower bound. However, some distributed encoders have been proposed to approach this bound for high SNR [19], [8].

IV. Summary

We first analyzed the problem of MMSE estimation in real-time of the state of a scalar linear continuous-time system over a scalar AGWN channel *without* channel feedback. The optimality of the scaled innovation encoder is proven and its optimal costs are precisely characterized. We then formulated the problem of MMSE state estimation in real-time of multiple linear systems over a shared AWGN channel and prove the optimality of the innovation encoder in this more general setting. These results are strengthened for the special case of two linear systems by obtaining the encoders and decoder in closed form when a correlation condition holds for the two systems. We provide bounds on the optimal costs when the correlation condition does not hold.

REFERENCES

[1] Striebel C. Sufficient statistics in the optimal control of stochastic systems. *Journal of Math. Analysis and App*, 12:576–592, 1965.

[2] C.D. Charalambous, A. Farhadi, and S.Z. Denic. Control of continuous-time linear gaussian systems over additive gaussian wireless fading channels: A separation principle. *Automatic Control, IEEE Transactions on*, 53(4):1013 –1019, may 2008.

[3] J.M.C. Clark. Conditions for one-to-one correspondence between an observation process and its innovation. Technical Report 1, 1969.

[4] B.P. Crow, I. Widjaja, L.G. Kim, and P.T. Sakai. Ieee 802.11 wireless local area networks. *Communications Magazine, IEEE*, 35(9):116 –126, sep 1997.

[5] T. Duncan and Varaiya P. On the solutions of a stochastic control system. *SIAM Journal on Control*, 1971.

[6] R. Durrett. *Probability: Theory and Examples*. Brooks Cole, 2002.

[7] Y.P. Fallah, Ching-Ling Huang, R. Sengupta, and H. Krishnan. Analysis of information dissemination in vehicular ad-hoc networks with application to cooperative vehicle safety systems. *Vehicular Technology, IEEE Transactions on*, 60(1):233 –247, jan. 2011.

[8] P. Floor, A. Kim, N. Wernersson, T. Ramstad, M. Skoglund, and I. Balasingham. Zero-delay joint source-channel coding for a bivariate gaussian on a gaussian mac. *Communications, IEEE Transactions on*, PP(99):1 –12, 2012.

[9] P. Frost and T. Kailath. An innovations approach to least-squares estimation–part iii: Nonlinear estimation in white gaussian noise. *Automatic Control, IEEE Transactions on*, 16(3):217 – 226, jun 1971.

[10] M. Gastpar, B. Rimoldi, and M. Vetterli. To code, or not to code: lossy source-channel communication revisited. *Information Theory, IEEE Transactions on*, 49(5):1147 – 1158, may 2003.

[11] Jr. Goblick, T. Theoretical limitations on the transmission of data from analog sources. *Information Theory, IEEE Transactions on*, 11(4):558 – 567, oct 1965.

[12] V. Gupta, A.F. Dana, J.P. Hespanha, R.M. Murray, and B. Hassibi. Data transmission over networks for estimation and control. *Automatic Control, IEEE Transactions on*, 54(8):1807 –1819, aug. 2009.

[13] J.P. Hespanha, P. Naghshtabrizi, and Yonggang Xu. A survey of recent results in networked control systems. *Proceedings of the IEEE*, 95(1):138 –162, jan. 2007.

[14] Ching-Ling Huang, Y.P. Fallah, R. Sengupta, and H. Krishnan. Adaptive intervehicle communication control for cooperative safety systems. *Network, IEEE*, 24(1):6 –13, jan.-feb. 2010.

[15] Ching-Ling Huang, Y.P. Fallah, R. Sengupta, and H. Krishnan. Intervehicle transmission rate control for cooperative active safety system. *Intelligent Transportation Systems, IEEE Transactions on*, 12(3):645 –658, sept. 2011.

[16] Ching-Ling Huang, R. Sengupta, H. Krishnan, and Y.P. Fallah. Implementation and evaluation of scalable vehicle-to-vehicle safety communication control. *Communications Magazine, IEEE*, 49(11):134 –141, november 2011.

[17] Ching-Ling Huang, Raja Sengupta, and Jiangchuan Huang. Real-time tracking of linear processes over an awgn channel without feedback: The mmse optimality of the innovation encoder. In *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pages 6691 –6696, dec. 2011.

[18] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*. North-Holland Publishing Company, 1981.

[19] A. Lapidoth and S. Tinguely. Sending a bivariate gaussian over a gaussian mac. *Information Theory, IEEE Transactions on*, 56(6):2714 –2752, june 2010.

[20] R.S. Lipster. *Lecture Notes # 10 for the course Stochastic Process*.

[21] J.T. Lo. On optimal nonlinear estimation part i: Continuous observation. *Information Sciences*, 6(0):19 – 32, 1973.

[22] A. Papoulis and Pillai S.U. *Probability, Random Variables and Stochastic Processes*. McGrawHill, 2002.

[23] A. Sahai and S. Mitter. The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link part i: Scalar systems. *Information Theory, IEEE Transactions on*, 52(8):3369 –3395, aug. 2006.

[24] P. Seiler and Raja Sengupta. An h-infinity approach to networked control. *Automatic Control, IEEE Transactions on*, 50(3):356 – 364, march 2005.

[25] C. E. Shannon. A mathematical theory of communication. *SIGMOBILE Mob. Comput. Commun. Rev.*, 5(1):3–55, January 2001.

[26] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M.I. Jordan, and S.S. Sastry. Kalman filtering with intermittent observations. *Automatic Control, IEEE Transactions on*, 49(9):1453 – 1464, sept. 2004.

[27] S. Tatikonda, A. Sahai, and S. Mitter. Stochastic linear control over a communication channel. *Automatic Control, IEEE Transactions on*, 49(9):1549 – 1561, sept. 2004.

[28] D. Teneketzis. On the structure of optimal real-time encoders and decoders in noisy communication. *Information Theory, IEEE Transactions on*, 52(9):4017 –4035, sept. 2006.

[29] P. Varaiya and J. Walrand. On delayed sharing patterns. *Automatic Control, IEEE Transactions on*, 23(3):443– 445, june 1978.

[30] J. Walrand and P. Varaiya. Optimal causal coding - decoding problems. *Information Theory, IEEE Transactions on*, 29(6):814 – 820, nov 1983.

[31] S. W. Wing and R.W. Brockett. Systems with finite communication bandwidth constraints. i. state estimation problems. *Automatic Control, IEEE Transactions on*, 42(9):1294 –1299, sep 1997.

[32] H.S. Witsenhausen. A counter-example in stochastic optimum control. *SIAM Journal on Control*, 1968.

[33] H.S. Witsenhausen. Separation of estimation and control for discrete time systems. *Proceedings of the IEEE*, 1971.

[34] Yonggang Xu and J.P. Hespanha. Estimation under uncontrolled and controlled communications in networked control systems. In *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05. 44th IEEE Conference on*, pages 842 – 847, dec. 2005.