This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2022.3151189, IEEE Transactions on Automatic Control

Dual Effect, Certainty Equivalence and Separation Revisited: A Counterexample and a Relaxed Characterization for Optimality

Milan S. Derpich and Serdar Yüksel

***Abstract*—We study the optimality of control policies admitting certainty equivalence or separation (of estimation and control) in discrete-time stochastic control, with two main contributions: (i) We first revisit the influential theorem given in the seminal 1974 paper by Bar-Shalom and Tse which studies the equivalence between certainty equivalence (CE) and no-dual-effect (NDE) properties in discrete-time stochastic control problems involving a linear dynamic system with a possibly non-linear measurement function. We show that there is a subtle error in Bar-Shalom and Tse's proof of the claim that CE implies NDE. Moreover, we prove that the claim does not hold by providing a counterexample. (ii) As our second and primary contribution, we introduce an alternative and a more relaxed notion of dual-freeness and establish that this new notion is sufficient to guarantee the separation of estimation and control and CE in the same control problem considered by Bar-Shalom and Tse.**

I. Introduction

In the classical theory of partially observed Linear Quadratic Gaussian problems under quadratic criteria [1]–[3], a celebrated result in the field is that the optimal control policy has a separation structure, where the optimal control policy (which turns out to be linear) has the same linear gain matrix that one would obtain for the corresponding fully observed problem but the state would be replaced with the conditional expectation of the state given the information at the controller. This phenomenon is known as the *separation of estimation and control* (where in its generality, this separation principle is said to hold when an optimal control exists in a subset of admissible policies where the control depends on the information only through the conditional expectation of the state given the information available), and for this particular case, a more special version of it, known as the *certainty equivalence* principle (precise definitions to follow).

This result significantly eases the study of optimal partially observed stochastic control problems; in general such problems require the usage of non-linear filtering, which is typically an infinite dimensional non-linear stochastic dynamical system taking values from a space of probability measures [4], [5], and hence are computationally challenging.

Milan S. Derpich is with the Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile. Email: milan.derpich@usm.cl. S. Yüksel is with the Department of Mathematics and Statistics, Queen's University, Kingston, Ontario, Canada, K7L 3N6. Email: yuksel@mast.queensu.ca. This research was partially supported by FONDECYT project 1171059, CONICYT basal fund FB0008, CONICYT postdoctoral fellowship 74180079, and the Natural Sciences and Engineering Research Council of Canada (NSERC).

The celebrated paper [6] due to Bar-Shalom and Tse establishes conditions under which, in a given control problem, certainty equivalence becomes equivalent to having the "so called" *no-dual-effect* (NDE) property. As defined in [6], the latter property is said to hold when the updated conditional distribution of the plant state given the plant output measurements does not depend on past control actions (the precise definitions of certainty equivalence and NDE property given in [6] are reproduced in Section II below). Besides revealing a fundamental aspect of a class of control problems, the equivalence between NDE and CE is useful when one of these properties is seemingly more difficult to verify than the other.

The problems on separation, certainty equivalence and related concepts (such as neutrality and active probing) are at the heart of partially observable optimal stochastic control problems, and hence have received a large interest in the classical stochastic control literature, as we briefly review in the following. In the economics theory literature also, there have been many contributions on this subject and the term *certainty-equivalence* appears first in this literature, to our knowledge [7] [8] [9]. Patchell and Jacobs [10] revisited the concepts of separability (the property that the plant state conditional estimate is a sufficient statistic), neutrality [11] (where dual effect is not present) and certainty equivalence (where the stochastic parameters are replaced by their conditional expectations). As noted by Patchell and Jacobs, *control laws that include active experimentation have been called 'dual' control laws*; active experimentation or probing is performed to reduce the uncertainty regarding a dynamical system. De Water and Willems [12] point to potentially ill-posed aspects of prior definitions for certainty-equivalence (in view of various interpretations for taking expectations) and develop a refined characterization on what it means to be certainty-equivalent, for both discrete-time and continuous-time systems. In their formulation, a control policy is said to be certainty equivalent when the conditional expectation of an *optimal control (policy realization) corresponding to the system with no uncertainty* conditioned on the information up to a given time is the actual control applied at that time. They also established that the LQG problem with linear Gaussian measurements admits optimal policies which are certainty-equivalent both in discrete-time and continuous-time [12]. For discrete-time linear system and measurement models that are not necessarily Gaussian (known as LQ, but not necessarily LQG models), we note that separation also applies when the noise processes are not Gaussian, though of course the estimations will no longer be linear; see [1, Lemma 5.2.1].

We also note that, in general continuous-time setups, the analysis can be subtle due to the fact that the control policy (only restricted to be measurable in general) may lead to issues on the existence of strong solutions for a given controlled stochastic differential equation. Lindquist [13] provides a general separation theorem provided that the control laws are among those which lead to the existence of a solution to the controlled stochastic differential system, generalizing previous analyses where only control laws of the Lipschitz type were considered, e.g. by Kushner [14] where Lipschitz continuity was in the conditional estimate and Wonham [15] where Lipschitz property holds in the control when viewed as a map from the normed linear space of continuous functions on measurement history to control actions to allow for the existence of strong solutions. To avoid such technical issues on strong solutions, relaxed solution concepts were introduced and studied in the literature based on the measure transformation technique due to Girsanov [16]–[18] which allows the control to be a function of an independent Brownian process. This approach requires absolute continuity conditions on the measurement process which may not be always applicable, though in the continuous-time literature this is the typical setup. An alternative approach is presented in a recent article by Georgiou and Lindquist [19], which provides a new characterization for when separation holds for the cases with linear Gaussian measurements (though with a more relaxed system noise).

A further related paper is [20] (see also [21]), which studied the separation problem in the context of networked LQG systems and established the joint optimality of a separated structure of optimal coding and control policies through the approach which is further refined in this paper. Part of the analysis (and the motivation) here builds on the approach developed in [20, Theorem 3.1], which however only considered sample path equivalence and lacked the generality with an argument tailored only for the networked control problem, and, more importantly, did not establish the relation with [6].

Contributions. (i.) In Section II we review the definition of NDE property provided by [6], highlighting one of its crucial but implicit properties, namely, that the no-dual effect of the control actions needs to hold even for arbitrary, unrestriced control actions. As discussed therein, this turns out to be a very demanding requirement. **(ii.)** After presenting in Section III the theorem of [6] which asserts the equivalence between CE and NDE, we show in Section IV-A that the proof of the claim that CE implies NDE given in [6] is unfortunately incorrect. Indeed, such an implication turns out to be false in general, as we demonstrate by providing a simple counterexample in Section IV-B. **(iii.)** The above mentioned findings motivate the need to consider an alternative and relaxed (or refined) notion of no dual effect which is less demanding and yet sufficient to imply CE or separation in the same family of control problems considered in [6]. We present such a refined definition of NDE property in Section V and show in Section VI that in the considered class of optimal stochastic control problems, this characterization is sufficient to guarantee that separation or certainty-equivalence holds.

II. Revisiting Bar-Shalom and Tse's Definitions

In [6], certainty equivalence and the no-dual-effect property are defined for the following general setup. The state of the dynamical system to be controlled is given by the recursion

$$x_{k+1} = f_k(x_k, u_k, v_k), \quad k = 0, 1, \ldots, N-1, \tag{1}$$

where u_k and v_k are the controller action and process noise at time k, respectively, and the distribution of the random initial state x_0 is known and given. The controller has access to the measurements

$$y_k = h_k(x_k, w_k), \quad k = 0, 1, \ldots, N-1, \tag{2}$$

where w_k is the measurement noise at time k. The $x^N = \{x_0, \ldots, x_N\}$, u^{N-1}, w^{N-1}, v^{N-1}, y^{N-1} sequences are random and vector-valued (with appropriate dimensions), and hence assumed to be measurable functions on $\mathcal{F}$, where $(\Omega, \mathcal{F}, P)$ is a probability space. In the above, f_k, h_k, for each k, are assumed to be measurable with no further restrictions.

The controller applies a policy $\{\gamma_0, \ldots, \gamma_{N-1}\}$ from a set of admissible control policies $\Pi^N = \times_{k=0}^{N-1} \Pi_k$, where, for each k, Π_k is the set of all functions of appropriate range measurable on the σ-field over Ω generated by $\mathcal{I}_k = \{y^k, u^{k-1}\}$ (the information available at the controller at time k), such that

$$u_k = \gamma_k(\mathcal{I}_t) = \gamma(y^k, u^{k-1}), \quad k = 0, \ldots, N-1. \tag{3}$$

A given closed-loop system together with a cost function and a class of admissible control policies is referred to as a *control problem.*

A. Certainty Equivalence

As expressed in [6, eqs. (2.20)–(2.22)], a control problem possesses the **certainty equivalence** (CE) property if the closed-loop optimal control policy has the same form as the deterministic optimal control policy under perfect state observation and in the absence of process noise. More precisely, if in the absence of process noise the optimal closed-loop control policy is

$$u_k^{\mathrm{CE}} = \phi_k(x_k), \tag{4}$$

and CE holds, then the optimal closed-loop control policy for the noisy and not necessarily fully observed system is

$$u_k^{\mathrm{CLO}} = \phi_k\left(\mathrm{E}[x_k | y^k, u^{k-1}]\right), \quad \forall k. \tag{5}$$

B. The No Dual Effect Property as Defined in [6]

The definition of the no dual effect property given in [6, eqs. (2.13)–(2.15)] can be stated as follows.[1]

[1] In [6, eq. (2.13)], instead of $M^r_{k|k}$ the authors define and utilize $M^r_{k|k,i} \triangleq \mathrm{E}[(x_{k,i} - \mathrm{E}[x_{k,i}|y^k, u^{k-1}])^r | y^k, u^{k-1}]$, where $x_{k,i}$ denotes the i-th entry in the vector x_k, stating in footnote 4 that joint moments should also be included. Our notation using tensor products fulfills this need.

Definition 1 (No Dual Effect Property as defined in [6])**.** *In a closed loop system, the controller is said to have* ***no dual effect of order*** r *(*$r \geq 2$*) if* [2]

$$\begin{aligned}&\mathrm{E}\Big[M^r_{k|k} \;\Big|\; y^j(\omega, u^{j-1}), u^{k-1}\Big]\\&= \mathrm{E}\Big[M^r_{k|k} \;\Big|\; y^j(\omega, 0^{j-1}), u^{k-1} = 0^{k-1}\Big],\\&\qquad \omega\text{-a.s., } \forall u^{k-1}, \forall j \leq k,\end{aligned} \tag{6a}$$

where

$$M^r_{k|k}(y^k(\omega, u^{k-1}), u^{k-1}) \triangleq \mathrm{E}\Big[(\tilde{w}_k)^{\{r\}} \Big| y^k(\omega, u^{k-1}), u^{k-1}\Big] \tag{6b}$$

$$\tilde{w}_k \triangleq x_k - E[x_k | y^k(\omega, u^{k-1}), u^{k-1}] \tag{6c}$$

and $a^{\{r\}}$ *denotes the* r*-th order tensor product of vector* a*.*

If $r = 2$, then (6) yields the second-order no-dual-effect condition

$$\begin{aligned}&\mathrm{E}\big[\tilde{w}_k \tilde{w}_k' \big| y^j(\omega, u^{j-1}), u^{k-1}\big]\\&= \mathrm{E}\big[\tilde{w}_k \tilde{w}_k' \big| y^j(\omega, 0^{j-1}), u^{k-1} = 0^{k-1}\big],\\&\qquad \omega\text{-a.s., } \forall u^{k-1}, \forall j \leq k.\end{aligned} \tag{7}$$

Remark 1. *[On the open-loop vs. closed-loop interpretation of control actions in Definition 1] The notation* $y^j(\omega, u^{k-1})$ *makes explicit the dependence of* y^j *on the underlying sample-space element* ω *and the control sequence* u^{k-1}*. The facts that* u^{k-1} *in* (6a) *does not depend on* ω *and that* (6a) *is to hold for all sequences* u^{k-1}*, imply that in* (6) u^{k-1} *is to be interpreted as an arbitrary control action sequence, not necessarily functionally dependent on* ω *or* y^{k-1}*. In other words,* (6) *is to hold even in open loop, wherein the control actions* u^{N-1} *are generated without feedback. The arbitrary and unrestricted nature of* u^{N-1} *in the above definition makes the NDE condition very demanding.*

Remark 2. *[On the closed-loop interpretation of control actions in Definition 1] If one were to interpret Definition 1 under a closed-loop, we naturally would have a more restrictive setup compared with an open-loop restriction. Since our analysis and counterexample to be presented is under an open-loop interpretation, this is also applicable to a closed loop interpretation. An additional, perhaps more direct, reason why the NDE condition defined above under a closed-loop interpretation is hard to satisfy, even for the 2nd-order NDE, is because equality in* (6a) *is to hold even when* $j = 0$*, i.e., without conditioning on the measurements. In this case, and if the control actions* u^{k-1} *are a function of the measurements* y^{k-1} *(i.e., in closed loop), then it is natural to expect*

$$E[\tilde{w}_k \tilde{w}_k' | u^{k-1}(y^{k-1}(\omega, u^{k-2}))]$$

to differ from

$$E[\tilde{w}_k \tilde{w}_k' | u^{k-1} = 0^{k-1}]$$

(for instance, the former is, in general, a random variable, while the latter is a deterministic quantity). As well-known, a notable exception is the LQG setup where conditional covariance matrices do not depend on the measurement or past control realizations.

III. The Equivalence Between CE and NDE Established in [6]

The connection between CE and NDE properties is studied in [6, § III], for the particular case in which (1) is a linear dynamical system with the same measurement function as before,

$$x_{k+1} = F_k x_k + G_k u_k + v_k, \quad k = 0, \dots, N-1 \tag{8a}$$

$$y_k = h_k(x_k, w_k), \quad k = 0, \dots, N-1 \tag{8b}$$

where each v_k has covariance V_k,

$$\mathrm{E}[v_k] = 0, \quad k = 0, \dots, N-1 \tag{9a}$$

$$\mathrm{E}[v_k v_j] = V_k \delta_{k,j}, \quad k = 0, \dots, N-1 \tag{9b}$$

$$w_0^{N-1} \perp\!\!\!\perp v_0^{N-1}, \tag{9c}$$

where $\perp\!\!\!\perp$ denotes probabilistic independence and δ_{jk} is the Kronecker delta function (notice that $\perp\!\!\!\perp$ differs from the symbol $\perp$ usually employed to denote the weaker notion of *uncorrelation* or *orthogonality*).

The cost to be minimized is quadratic, given by

$$J_0 \triangleq \mathrm{E}\left[x_N' Q_N x_N + \sum\nolimits_{i=0}^{N-1} x_i' Q_i x_i + u_i' R_i u_i\right] \tag{10}$$

with $Q_i \geq 0$ and $R_i > 0$ being matrices of appropriate dimensions. The control problem is the minimization of J_0 over all control policies $\{\gamma_0, \dots, \gamma_{N-1}\} \in \Pi^N$ (see (3)).

We can now re-state the theorem of [6] which, for this setting, establishes an equivalence between 2-nd order NDE and the CE property.

Theorem 1 (From [6, p. 98])**.** *The optimal stochastic control for the system with linear dynamics* (8a)*, process and measurement noises* (9)*, measurement equation* (8b) *and cost function* (10) *has the CE property for all* $Q_i \geq 0$*,* $R_i > 0$ *if and only if the control has no dual effect of second order, i.e., the updated covariance* $\Sigma_{k|k}$ *is not a function of the past control sequence* u^{k-1}*, for all* k*.*

In the latter statement, $\Sigma_{k|k}$ is the conditional covariance matrix defined as

$$\Sigma_{k|k} \triangleq \mathrm{E}\big[\tilde{w}_k \tilde{w}_k' \;\big|\; y^k, u^{k-1}\big]. \tag{11}$$

Remark 3. *The last condition stated in Theorem 1, which reads "the updated covariance matrix* $\Sigma_{k|k}$ *is not a function of the past control sequence* u^{k-1}*, for all* k*", translates into*

$$\begin{aligned}&\mathrm{E}\big[\tilde{w}_k \tilde{w}_k' \;\big|\; y^k(\omega, u^{k-1}), u^{k-1}\big]\\&= \mathrm{E}\big[\tilde{w}_k \tilde{w}_k' \;\big|\; y^k(\omega, 0^{k-1}), u^{k-1} = 0^{k-1}\big],\\&\qquad \omega\text{-a.s.}, \forall u^{k-1}\end{aligned} \tag{12}$$

which satisfies (7) *for* $j = k$*. However, in order to comply with the definition of NDE,* (12) *should hold also when, in its right-hand-side,* y^k *is replaced by* y^j*, for all* $j \leq k$ *(that is,*

[2]In [6, eq. (2.14)] (the original version of (6a)), the right-hand-side is actually $\mathrm{E}[M^r_{k|k} \;\big|\; y^j(\omega, 0^{k-1})]$. However, we believe the authors of [6] meant the more rigorously written form given by (6a), which is a weaker requirement than the one obtained by a faithful and rigid reading of [6, eq. (2.14)].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2022.3151189, IEEE Transactions on Automatic Control

while keeping the superscript $k-1$ of $u^{k-1}$and replacing y^k by y^j). ▲

In the next section we will review part of the proof of this theorem given in [6] and point out the main issue in it.

IV. Analysis of the Characterization by Bar-Shalom and Tse [6] and a Counterexample

A. On the Proof that CE Implies NDE Given in [6]

The proof of Theorem 1 provided in [6] relies on the well known fact that the minimum of (10) is given by the recursion

$$J_k^*(\mathcal{I}_k) \triangleq \min_{u_k} \mathrm{E}\left[x_k' Q_k x_k + u_k' R_k u_k + J_{k+1}^*(y^{k+1}, u^k)|\mathcal{I}_k\right], \quad k = N-1, \dots, 0 \tag{13a}$$

$$J_N^*(\mathcal{I}_N) \triangleq \mathrm{E}[x_N' Q_N x_N|\mathcal{I}_N], \tag{13b}$$

(recall that $\mathcal{I}_k = \{y^k, u^{k-1}\}$). For this dynamical system, the optimal deterministic control (in the absence of process noise and with perfect state information) is

$$u_k^{\mathrm{CE}} = -(R_k + G_k' P_{k+1} G_k)^{-1} G_k' P_{k+1} F_k x_k, \tag{14}$$

for $k = 0, 1, \dots, N-1$, and thus from (14), CE holds if and only if the optimal control action u_k^* satisfies

$$u_k^* = \tilde{u}_k \triangleq -(R_k + G_k' P_{k+1} G_k)^{-1} G_k' P_{k+1} F_k \, \mathrm{E}[x_k|\mathcal{I}_k], \tag{15}$$

for $k = 0, 1, 2, \dots, N-1$, where P_{k+1} is a matrix defined recursively as [6, eq. (311)]

$$P_k \triangleq Q_k + F_k'[P_{k+1} - P_{k+1} G_k (R_k + G_k' P_{k+1} G_k)^{-1} G_k' P_{k+1}] F_k, \quad k = 1, 2, \dots, N-1 \tag{16}$$

$$P_N \triangleq Q_N. \tag{17}$$

As argued in [6], if CE (i.e. (15)) holds, then the minimum cost to go at time k can be written in the form [6, eq. (3.13)]

$$J_{k+1}^* = \mathrm{E}\left[x_{k+1}' P_{k+1} x_{k+1}|\mathcal{I}_{k+1}\right] + \beta_{k+1}, \tag{18}$$

where, with $\beta_N = 0$, β_k is a sequence of scalar random variables defined recursively as (see [6, eqs. (3.12), (3.14)])

$$\beta_k \triangleq \mathrm{tr}\{V_k P_{k+1}\} + \mathrm{tr}\{(G_k' P_{k+1} F_k)'(R_k + G_k' P_{k+1} G_k)^{-1}(G_k' P_{k+1} F_k) \times \mathrm{E}[\tilde{w}_k \tilde{w}_k'|\mathcal{I}_k]\} + \mathrm{E}[\beta_{k+1}|\mathcal{I}_k] \tag{19}$$

and $\tilde{w}_k \triangleq x_k - \mathrm{E}[x_k|\mathcal{I}_k]$ (as in (6c)).

The "necessity" part in the proof of Theorem 1 given in [6, p. 498], claiming that (CE holds for every $R > 0$, $Q \geq 0$)⇒NDE, relies on the following argument:

Suppose that CE holds for every $R_k > 0$, $k = 1, \dots, N-1$, $Q_k \geq 0$, $k = 1, \dots, N$. Then, by definition, the u_k which minimizes[3]

$$J_k \triangleq \mathrm{E}\left[x_k' Q_k x_k + u_k' R_k u_k + J_{k+1}^*|\mathcal{I}_k\right] \tag{20}$$

(i.e., the RHS of (13a)), say u_k^*, is equal to $\tilde{u}_k$ (defined in (15)), for all $R_k > 0$. But $\tilde{u}_k$ is known to be the control that minimizes

$$\mathrm{E}\left[x_k' Q_k x_k + u_k' R_k u_k + \mathrm{E}\left[x_{k+1}' P_{k+1} x_{k+1}|\mathcal{I}_k\right]|\mathcal{I}_k\right]$$

(see [6, eq. (3.9)] and the discussion therein). For this to hold $\forall R_k > 0$, one of the following two options must hold:

1) β_{k+1} is not a function of u_k.
2) β_{k+1} is a function of u_k which is minimized by $\tilde{u}_k$, *for all* $R_k > 0$.

Here, the authors of [6] point out **in the paragraph at the end of the second column on [6, page 498]** that β_{k+1} does not depend on R_k while $\tilde{u}_k$ does, deducing from this that option 2) cannot happen and, accordingly, that β_{k+1} is not a function of u_k. Since β_{k+1} depends linearly on the expectations $\mathrm{E}[\tilde{w}_j \tilde{w}_j'|y^{k+1}, u^k]$, $j \geq k+1$ (see (19)), it is concluded that "$\Sigma_{k|k}$ is independent of u^{k-1}, i.e., it is necessary that the controls have no dual effect of second order" [6, p. 499]. From the above reasoning, what one actually concludes is that $\mathrm{E}[\tilde{w}_j \tilde{w}_j'|y^{k+1}(\omega, u^k), u^k]$ does not change with u^k, ω-a.s., for every $j \geq k+1$. Clearly, this does not coincide with the NDE property of second order defined by the authors in (7) (that is, in [6, eq. (2.14)]). See also Remark 3 above. Thus, we state explicitly below the main issue with the proof:

By looking at (15)*, it is clear that $\tilde{u}_k$ depends on R_k if and only if $E[x_k|y^k, u^{k-1}] \neq 0$. Hence, if $E[x_k|y^k, u^{k-1}] = 0$, which implies $\tilde{u}_k = 0$, it is possible that β_{k+1} depends on u_k and reaches its minimum precisely at $u_k = \tilde{u}_k = 0$ (i.e., option 2 above can indeed happen). In the next section, we present a simple setup in which this is the case, providing a counterexample to the claim that (CE holds for every $R > 0$, $Q \geq 0$)⇒NDE in Theorem 1 (even when regarding* (12) *as the definition for 2nd-order NDE instead of* (7)*).*

B. A Counterexample for the Claim that CE⇒NDE

Consider a special case of the linear dynamics (8a) in which x_k and u_k are scalars related by

$$x_{k+1} = F x_k + G u_k + v_k, \qquad k = 1, 2, \cdots \tag{21}$$

where $x_1 \sim \mathcal{N}(0,1)$ and v_k is i.i.d. with $v_k \sim \mathcal{N}(0,1)$. Suppose that the measurement system $h_k(x_k, w_k)$ is given by

$$y_k = h_k(x_k, w_k) = \begin{cases} 0 & , k \text{ is odd} \\ \mathrm{sgn}(x_k) & , k \text{ is even} \end{cases} \qquad k = 1, 2, \dots. \tag{22}$$

Let $N = 3$. At time $k = 1$, the sample $y_1 = 0$ deterministically, and provides no information about x_1. Hence

$$\mathrm{E}[x_1|y_1] = 0 \text{ and } x_1 \perp\!\!\!\perp y_1. \tag{23}$$

Thus, from (15), the certainty-equivalent control is

$$\tilde{u}_1 = 0, \quad \forall R_1 > 0, \ \forall P_2. \tag{24}$$

From (19),

$$\beta_3 \equiv 0 \tag{25}$$

[3] This equation corresponds to [6, eq. (3.15)]. However, in the latter the conditioning is on y_k, u^{k-1}, which we believe to be a typo. In any case, this conditioning does not change the forthcoming analysis.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2022.3151189, IEEE Transactions on Automatic Control

$$\beta_2(y^2, u_1) = Q_3 + \min_{u_2}(u_2 - \tilde{u}_2)^2(R_2 + GQ_3G) \tag{26}$$
$$+ \frac{G^2Q_3^2F^2}{R_2 + G^2Q_3}\,\mathrm{E}\big[(x_2 - \mathrm{E}[x_2|y^2, u_1])^2 \;\big|\; y^2, u_1\big] \tag{27}$$
$$= Q_3 + \frac{G^2Q_3^2F^2}{R_2 + G^2Q_3}\,\mathrm{E}\big[(x_2 - \mathrm{E}[x_2|y^2, u_1])^2 \;\big|\; y^2, u_1\big] \tag{28}$$

with the optimal control action given by

$$u_2^* = \tilde{u}_2. \tag{29}$$

Likewise,

$$\beta_1(y_1) = P_2 + \frac{G^2P_2^2F^2}{R_1 + G^2P_2}\,\mathrm{E}\big[(x_1 - \mathrm{E}[x_1|y_1])^2 \;\big|\; y_1\big] \tag{30}$$
$$+ \min_{u_1}\big\{(u_1 - \tilde{u}_1)^2(R_1 + GP_2G) + \mathrm{E}[\beta_2(y^2, u_1)|y_1]\big\} \tag{31}$$
$$\overset{(23)}{=} P_2 + \frac{G^2P_2^2F^2}{R_1 + G^2P_2}\mathrm{E}[(x_1)^2] \tag{32}$$
$$+ \min_{u_1}\big\{(u_1 - \tilde{u}_1)^2(R_1 + GP_2G) + \mathrm{E}[\beta_2(y^2, u_1)|y_1]\big\} \tag{33}$$
$$\overset{(24)}{=} P_2 + \frac{G^2P_2^2F^2}{R_1 + G^2P_2}\mathrm{E}[(x_1)^2] \tag{34}$$
$$+ \min_{u_1}\big\{(u_1)^2(R_1 + GP_2G) + \mathrm{E}[\beta_2(y^2, u_1)|y_1]\big\} \tag{35}$$
$$\overset{(28)}{=} P_2 + \frac{G^2P_2^2F^2}{R_1 + G^2P_2}\mathrm{E}[(x_1)^2] + Q_3 + \tag{36}$$
$$\min_{u_1}\Bigg((u_1)^2(R_1 + GP_2G) \tag{37}$$
$$+ \frac{G^2Q_3^2F^2}{R_2 + G^2Q_3}\,\mathrm{E}\big[(x_2 - \mathrm{E}[x_2|y_2, u_1])^2\big]\Bigg) \tag{38}$$

Finding the value of u_1 that minimizes

$$\mathrm{E}\big[(x_2 - \mathrm{E}[x_2|y_2, u_1])^2\big]$$

is equivalent to finding the *minimum mean squared error* (MMSE) two-cell quantizer for a zero-mean Gaussian variable, where the single threshold of the quantizer is $-Gu_1$. This optimal single threshold is known to be unique and equal to zero [22, footnote 1]. Thus,

$$\mathrm{E}\big[(x_2 - \mathrm{E}[x_2|y^2, u_1])^2\big] = \mathrm{E}\big[(x_2 - \mathrm{E}[x_2|y_2, u_1])^2\big]$$

depends on u_1 and its minimizer (and that of the rightmost term in (38)) is $u_1^* = 0 = \tilde{u}_1$.

Therefore, the linear dynamics system (21) and its measurement system (22) satisfy the CE condition for all $R_1, R_2 > 0$ and $Q_1, Q_2, Q_3 \geq 0$. However, this system does not have the NDE property, since, as already mentioned, $\mathrm{E}\big[(x_2 - \mathrm{E}[x_2|y^2, u_1])^2|y^2, u_1\big]$ depends on u_1.

V. Refined Characterizations of the NDE Property

We now propose alternative definitions of the NDE property which are less restrictive than [6] and yet sufficient to guarantee that CE holds in the same setup considered by [6] (as is shown in Section VI). We refer to them as *NDE control policy* (NDECP) properties. The first definition is sample-path based, which means that at each time k, the estimation error $\tilde{\omega}_k(\omega)$ remains unaltered (as a random variable) when one or more control policies prior to k are changed within a certain set. The second definition is second-order (or covariance based), requiring the same control-policy invariance criteria for $\mathrm{E}[\tilde{w}_k(\omega)\tilde{w}_k(\omega)']$ instead of $\tilde{w}_k(\omega)$.

A. The NDECP Property: Sample-Path Characterizations

To state our definitions in a precise manner, it is convenient to note that, for each $k = 1, \ldots, N$, the estimation error $\tilde{w}_k$ is a random variable that may functionally depend (at most) on the control policies $\gamma_0, \ldots, \gamma_{k-1}$, and thus we write it as $\tilde{w}_k(\omega, \gamma_0, \ldots, \gamma_{k-1})$.

Definition 2 (NDECP)**.** *A given control problem satisfies the refined NDECP Property if there exists a set* $\tilde{\Pi}^N = \tilde{\Pi}_0 \times \cdots \times \tilde{\Pi}_{N-1}$ *of policies containing an optimal control policy* $\{\gamma_j^\star\}_{j=0}^{N-1}$ ***and the CE control policy*** *such that for all* $\{\dot{\gamma}_j\}_{j=0}^{N-1}, \{\ddot{\gamma}_j\}_{j=0}^{N-1} \in \tilde{\Pi}^N$ *and* $t = 1, \ldots, N-1$,

$$\tilde{w}_t(\omega, \dot{\gamma}_0, \ldots, \dot{\gamma}_k, \gamma_{k+1}^\star, \ldots, \gamma_{t-1}^\star) = \tilde{w}_t(\omega, \ddot{\gamma}_0, \ldots, \ddot{\gamma}_k, \gamma_{k+1}^\star, \ldots, \gamma_{t-1}^\star), \quad \omega\text{-a.s.}, \tag{39}$$

for $k < t$.

Thus, the refined NDE property holds if perturbing all the policies $\{\gamma_0, \ldots, \gamma_k\} \in \tilde{\Pi}^{k+1}$ has no effect on future estimation errors, provided future policies are optimal.

We note that we can always take $\tilde{\Pi}^N$ to be Π^N, the set of all admissible policies, in the definition above. In some applications, one can relax the policies further without any loss in performance, this is why the definition includes such a refinement. We introduce next two alternative formulations of the NDECP property which we show to be equivalent to Definition 2. The value of providing these formulations is that one may be easier to verify than the other; the equivalence of these is shown in the Appendix.

Definition 3 (NDECP: "Mild" Formulation)**.** *A given control problem satisfies the weak NDECP property if it meets all the conditions of Definition 2 but replacing* (39) *by*

$$\tilde{w}_t(\omega, \gamma_0, \ldots, \gamma_{k-1}, \gamma_k, \gamma_{k+1}^\star, \ldots, \gamma_{t-1}^\star) = \tilde{w}_t(\omega, \gamma_0, \ldots, \gamma_{k-1}, \ddot{\gamma}_k, \gamma_{k+1}^\star, \ldots, \gamma_{t-1}^\star), \quad \omega\text{-a.s.} \tag{40}$$

In words, the mild NDECP property holds when perturbing the current policy, say $\gamma_k \in \tilde{\Pi}_k$, has no effect on future estimation errors, provided future policies are optimal and past policies belong to $\tilde{\Pi}^k$ and remain fixed.

Definition 4 (NDECP: "One-Step-Ahead" Formulation)**.** *A given control problem satisfies the one-step mild NDECP property if it meets all the conditions of Definition 2 but replacing* (39) *by*

$$\tilde{w}_t(\omega, \gamma_0, \ldots, \gamma_{t-1}) = \tilde{w}_t(\omega, \ddot{\gamma}_0, \ldots, \ddot{\gamma}_{t-1}), \quad \omega\text{-a.s.} \tag{41}$$

Unlike the previous definitions, the one-step mild NDECP property holds when perturbing all policies $\{\gamma_0, \ldots, \gamma_{t-1}\} \in \tilde{\Pi}^t$ does not change the next estimation error. As noted earlier in the paper, [20] already utilizes the condition that

$\mathrm{E}[x_{t+1}|\mathcal{I}_{t+1}]-\mathrm{E}[x_{t+1}|\mathcal{I}_t]$ does not depend on control policies $\gamma_0,\ldots,\gamma_t$ (provided that the control policies after time t are those that are restricted to be optimal) to develop a separation result on optimal control policies for networked control applications. However, the analysis in [20] is restricted to a networked control problem.

B. The NDECP Property: 2nd-Order Characterizations

As in Definition 1, it is also possible to define a NDECP property of order 2, as follows:

Definition 5 (2nd-Order NDECP Property)**.** *A given control problem satisfies the 2nd-order NDECP property if it meets the conditions of Definition 2 but replacing* (39) *by*

$$\begin{aligned}&\mathrm{E}\big[\tilde{w}_t(\omega,\gamma_0,\ldots,\gamma_k,\gamma^\star_{k+1},\ldots,\gamma^\star_{t-1})^{\{2\}}\big]\\&=\mathrm{E}\big[\tilde{w}_t(\omega,\ddot{\gamma}_0,\ldots,\ddot{\gamma}_k,\gamma^\star_{k+1},\ldots,\gamma^\star_{t-1})^{\{2\}}\big]\end{aligned}\tag{42}$$

where for column vector a we use the notation $a^{\{2\}}:=aa'$.

Definition 6 (2nd-Order NDECP: "Mild" Formulation)**.** *A given control problem satisfies the 2nd-order NDECP property if it meets the conditions of Definition 2 but replacing* (39) *by*

$$\begin{aligned}&\mathrm{E}\Big[\tilde{w}_t(\omega,\gamma_0,\ldots,\gamma_{k-1},\gamma_k,\gamma^\star_{k+1},\ldots,\gamma^\star_{t-1})^{\{2\}}\Big]\\&=\mathrm{E}\Big[\tilde{w}_t(\omega,\gamma_0,\ldots,\gamma_{k-1},\ddot{\gamma}_k,\gamma^\star_{k+1},\ldots,\gamma^\star_{t-1})^{\{2\}}\Big]\end{aligned}\tag{43}$$

Definition 7 (2nd-Order NDECP: "One-Step-Ahead" Formulation)**.** *A given control problem satisfies the 2nd-order NDECP property if it meets the conditions of Definition 2 but replacing* (39) *by*

$$\mathrm{E}\Big[\tilde{w}_t(\omega,\gamma_0,\ldots,\gamma_{t-1})^{\{2\}}\Big]=\mathrm{E}\Big[\tilde{w}_t(\omega,\ddot{\gamma}_0,\ldots,\ddot{\gamma}_{t-1})^{\{2\}}\Big]\tag{44}$$

The equivalence between Definitions 5, 6 and 7 follows readily from the equivalence between definitions 2, 3 and 4 shown in the Appendix.

C. Comparison Between the NDE and NDECP Properties

All the NDECP definitions introduced above depart from the NDE of [6] (stated in Definition 1) in the following two important aspects:

1) The alternative definitions involve an independence condition which is required to hold only for a set of policies $\tilde{\Pi}^N$ containing, at least, the optimal control policy and the CE policy. Instead, as discussed in Remark 1, the NDE notion of [6] requires such independence over *all* admissible control policies (see also Remark 2). On its own, this difference makes the NDE condition so restrictive that it is not difficult to find control problems where, because of it, the NDECP holds but the NDE doesn't, as the example in Section V-D illustrates.
2) The alternative definitions require non-dependence on control policies prior to time k without conditioning (in particular, not conditioning on the measurements $y^j,\ j\le k$, thus differing from the definition of NDE stated in Definition 1). Such conditioning makes the NDE of [6] harder to satisfy. To see why, consider the random variables $s(\omega,\gamma)$, $u=u(\omega,\gamma)$ and $y(\omega,u)$, where γ belongs to some set Π, and notice that for every $\gamma,\dot{\gamma}\in\Pi$

$$\begin{aligned}&\mathrm{E}[s(\omega,\gamma)|y(\omega,u),u(\omega,\gamma)]\\&\qquad=\mathrm{E}[s(\omega,\dot{\gamma})|y(\omega,u),u(\omega,\dot{\gamma})]\ \ \omega\text{-a.s.}\\&\Longrightarrow\mathrm{E}[s(\omega,\gamma)]=\mathrm{E}[s(\omega,\dot{\gamma})]\,,\end{aligned}\tag{45}$$

but the reverse implication is not necessarily true.

D. Example: Control Problem Satisfying NDECP but not NDE

Example 1. *Let*

$$x_{k+1}=ax_k+bu_k+v_k$$

$$y_k=x_k1_{\{|x_k|\le M\}}+(x_k+w_k)1_{\{|x_k|>M\}}$$

where v_k and w_k are i.i.d. $[-B,B]$-valued (with $B>0$) uniformly distributed random variables and M is a constant to be given below. Suppose that x_0 is uniformly distributed between $[-A,A]$ (with $A>0$). Consider the expected cost:

$$E\left[Px_N^2+\sum_{k=0}^{N-1}qx_k^2+ru_k^2\right],$$

with $q>0,r>0$. If the controller had access to the state for all time stages, then the optimal controller would be linear and the performance of this controller would be no worse than the performance of the controller which only had access to the measurement process given above [23], [24, p.457]. Now, suppose that $P=p$ is taken to be the fixed point solution of the Riccati equation for this problem (whose existence follows from the controllability and the scalar nature of the problem), so that the resulting controller (with full state information) is stationary. Accordingly, with p solving the Riccati equation: $p=q+a^2p-\frac{a^2p^{b2}}{b^2p+r}$, with $u_t=-\frac{bpax_t}{b^2p+r}=:kx_t$, we know that such a control is stabilizing with $(a+bk)=:\rho$, $|\rho|<1$. Then, the closed loop system under this optimal control is given with

$$x_{k+1}=ax_k+bu_k+v_k=\rho x_k+v_k.$$

Now, if $x_0\in[-A,A]$ and with $|v_k|\le B$, we are guaranteed that $x_k\in[-M,M]$ where we take $A+B\frac{1}{1-|\rho|}=:M$. Accordingly, the optimal control (with full state information) would be realizable with the information structure given above since the event $|x_k|\le M$ would always be active. In particular, in the definition of NDECP in Definition 2, if the policies considered were the ones keeping the state in $[-M,M]$. On the other hand, if one were to apply an arbitrary control, say one larger than $\frac{2M}{b}$, then $(x_k+w_k)1_{\{|x_k|>M\}}$ would be active, and the separation property would no longer be applicable as the control policy would affect the estimation error. The message is that requiring that the no-dual effect property holds for every possible action realization is too restrictive.

VI. The NDECP Implies CE

In this section we show that, despite being less restrictive than the 2nd-order NDE property, the 2nd-order NDECP is sufficient to guarantee that the CE property holds, for the same control problem considered in [6, Section III].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2022.3151189, IEEE Transactions on Automatic Control

Theorem 2. *Suppose that the control problem given by the dynamical system* (8) *with disturbances* (9) *and cost function* (10) *satisfies Definition 7. Then the CE property holds (and thus the optimal control policy is linear).*

Proof. Denote the optimal control policy as $\gamma_k^\star$, $k = 0, 1, \ldots, N-1$. The optimal *mean* cost-to-go from time k onward, defined as $\bar{J}_k^\star \triangleq \mathrm{E}[J_k^*(\mathcal{I}_k)]$ (see (13)), is given recursively by

$$\bar{J}_k^\star(\gamma_0, \ldots, \gamma_{k-1}) \tag{46a}$$

$$= \min_{\gamma_k \in \tilde{\Pi}_k} \mathrm{E}\big[x_k' Q_k x_k + u_k' R_k u_k + \bar{J}_{k+1}^\star(\gamma_0, \ldots, \gamma_k)\big], \quad k = 0, \ldots, N-1 \tag{46b}$$

$$\bar{J}_N^\star(\gamma_0, \ldots, \gamma_{N-1}) \triangleq \mathrm{E}[x_N' Q_N x_N] \tag{46c}$$

Suppose that $\bar{J}_{k+1}^\star$ has the form

$$\bar{J}_{k+1}^\star(\gamma_0, \ldots, \gamma_k) \tag{47a}$$

$$= \mathrm{E}\big[x_{k+1}' P_{k+1} x_{k+1}\big] + \alpha_{k+1}(\gamma_0, \ldots, \gamma_k), \tag{47b}$$

where the matrix $P_{k+1} \geq 0$ does not depend on γ^k or u^k and α_{k+1} satisfies

$$\alpha_{k+1}(\gamma_0, \ldots, \gamma_k) \text{ doesn't depend on } \{\gamma_0, \ldots, \gamma_k\} \in \tilde{\Pi}^{k+1} \tag{47c}$$

(The validity of this supposition will be demonstrated near the end of this proof.) For $k = N-1$ it is easy to see from (46c) that (47) is true with

$$P_N = Q_N, \qquad \alpha_N = 0. \tag{48}$$

Now, assuming that (47) is true for $k+1$, one has that

$$\bar{J}_k^\star(\gamma_0, \ldots, \gamma_{k-1}) \tag{49}$$

$$\stackrel{(a)}{=} \min_{\gamma_k \in \tilde{\Pi}_k} \mathrm{E}\big[x_k' Q_k x_k + u_k' R_k u_k + \bar{J}_{k+1}^\star\big] \tag{50}$$

$$\stackrel{(47a)}{=} \min_{\gamma_k \in \tilde{\Pi}_k} \mathrm{E}\big[x_k' Q_k x_k + u_k' R_k u_k + \mathrm{E}\big[x_{k+1}' P_{k+1} x_{k+1}\big] + \alpha_{k+1}\big] \tag{51}$$

$$\stackrel{(b)}{=} \min_{\gamma_k \in \tilde{\Pi}_k} \mathrm{E}\big[x_k' Q_k x_k + u_k' R_k u_k + x_{k+1}' P_{k+1} x_{k+1}\big] + \alpha_{k+1} \tag{52}$$

$$\begin{aligned}
&\stackrel{(8a)}{=} \min_{\gamma_k \in \tilde{\Pi}_k} \Big\{ E[x_k' Q_k x_k + u_k' R_k u_k \\
&+ (F_k x_k + G_k u_k + v_k)' P_{k+1} (F_k x_k + G_k u_k + v_k)] \Big\} + \alpha_{k+1} \\
&\stackrel{(9)}{=} \min_{\gamma_k \in \tilde{\Pi}_k} \Big\{ \mathrm{E}[u_k'(R_k + G_k' P_{k+1} G_k) u_k + 2u_k' G_k' P_{k+1} F_k x_k] \Big\} \\
&\qquad + \mathrm{E}[x_k'(Q_k + F_k' P_{k+1} F_k) x_k] + \mathrm{tr}\{V_k P_{k+1}\} + \alpha_{k+1} \\
&\stackrel{(c)}{=} \min_{\gamma_k \in \tilde{\Pi}_k} \Big\{ E[E[u_k'(R_k + G_k' P_{k+1} G_k) u_k \\
&+ \quad 2u_k' G_k' P_{k+1} F_k \,\mathrm{E}[x_k|\mathcal{I}_k] \,|\mathcal{I}_k]] \Big\} \\
&\qquad + \mathrm{E}[x_k'(Q_k + F_k' P_{k+1} F_k) x_k] + \mathrm{tr}\{V_k P_{k+1}\} + \alpha_{k+1} \\
&\stackrel{(d)}{\geq} E\Big[E\Big[\min_{u_k} \Big\{ u_k'(R_k + G_k' P_{k+1} G_k) u_k + \\
&\qquad 2u_k' G_k' P_{k+1} F_k \,\mathrm{E}[x_k|\mathcal{I}_k] \Big\} \Big| \mathcal{I}_k \Big]\Big] \\
&\qquad + \mathrm{E}[x_k'(Q_k + F_k' P_{k+1} F_k) x_k] + \mathrm{tr}\{V_k P_{k+1}\} + \alpha_{k+1}
\end{aligned} \tag{53}$$

where (a) holds because $\gamma_k^\star \in \tilde{\Pi}_k$ and $\tilde{\Pi}$ contains an optimal policy by hypothesis, (b) is due to applying iterated expectations so that $\mathrm{E}[\mathrm{E}[\cdot]] = \mathrm{E}[\cdot]$ and to (47c), (c) follows from applying iterated expectations, and (d) holds because $u_k = \gamma_k(\mathcal{I}_k)$ is a measurable map with argument $\mathcal{I}_k$. It is easy to verify (e.g., by completing squares) that the minimizer of the quadratic form in (53) over $\gamma_k \in \tilde{\Pi}_k$ is

$$u_k^\star = \gamma_k^\star(\mathcal{I}_k) = -(R_k + G_k' P_{k+1} G_k)^{-1} G_k' P_{k+1} F_k \,\mathrm{E}[x_k|\mathcal{I}_k], \tag{54}$$

which corresponds to the CE control policy. Since latter belongs to $\tilde{\Pi}_k$, it follows that $\gamma_k^\star$ is also the solution to (53). With this we have shown that if (47) holds at time $k+1$, then the optimal control policy at time k is linear and coincides with the certainty-equivalent control law. We now show that the above also implies that (47) is true at time k. For this purpose, substitute (54) into the conditional expectation in (53) to obtain

$$E\Big[\min_{\gamma_k \in \tilde{\Pi}_k} \Big\{ u_k'(R_k + G_k' P_{k+1} G_k) u_k + \quad 2u_k' G_k' P_{k+1} F_k \,\mathrm{E}[x_k|\mathcal{I}_k] \Big\} \Big| \mathcal{I}_k \Big] \tag{55}$$

$$= E\Big[-\mathrm{E}[x_k|\mathcal{I}_k]' \, (G_k' P_{k+1} F_k)' (R_k + G_k' P_{k+1} G_k)^{-1} \times (G_k' P_{k+1} F_k) \,\mathrm{E}[x_k|\mathcal{I}_k] \Big| \mathcal{I}_k \Big] \tag{56}$$

$$\begin{aligned}
&= E[\tilde{w}_k' (G_k' P_{k+1} F_k)' (R_k + G_k' P_{k+1} G_k)^{-1} (G_k' P_{k+1} F_k) \tilde{w}_k | \mathcal{I}_k] \\
&- E[x_k' (G_k' P_{k+1} F_k)' (R_k + G_k' P_{k+1} G_k)^{-1} (G_k' P_{k+1} F_k) x_k | \mathcal{I}_k],
\end{aligned} \tag{57}$$

where we have utilized the orthogonality principle and $\tilde{w}_k = x_k - \mathrm{E}[x_k|\mathcal{I}_k]$ (as defined in (6c)). Substituting (57) in (53) we obtain

$$\begin{aligned}
\bar{J}_k^* &= E[E[\tilde{w}_k' (G_k' P_{k+1} F_k)' (R_k + G_k' P_{k+1} G_k)^{-1} \\
&\quad \times (G_k' P_{k+1} F_k) \tilde{w}_k | \mathcal{I}_k]] && (58) \\
&\quad - E[E[x_k' (G_k' P_{k+1} F_k)' (R_k + G_k' P_{k+1} G_k)^{-1} \\
&\quad \times (G_k' P_{k+1} F_k) x_k | \mathcal{I}_k]] && (59) \\
&\quad + \mathrm{E}[x_k'(Q_k + F_k' P_{k+1} F_k) x_k] + \mathrm{tr}\{V_k P_{k+1}\} + \alpha_{k+1} \\
&= \mathrm{E}[x_k' P_k x_k] + \alpha_k && (60)
\end{aligned}$$

with

$$\begin{aligned}
&\alpha_k \triangleq \mathrm{tr}\{V_k P_{k+1}\} + \alpha_{k+1} + \\
&\mathrm{tr}\{(G_k' P_{k+1} F_k)' (R_k + G_k' P_{k+1} G_k)^{-1} (G_k' P_{k+1} F_k) \mathrm{E}[\tilde{w}_k \tilde{w}_k']\}
\end{aligned} \tag{61}$$

$$P_k \triangleq Q_k + F_k'[P_{k+1} - P_{k+1} G_k (R_k + G_k' P_{k+1} G_k)^{-1} G_k' P_{k+1}] F_k \tag{62}$$

The theorem statement assumes that the NDECP property of Definition 7 holds, and thus $\mathrm{E}[\tilde{w}_k \tilde{w}_k']$ does not depend on $\{\gamma_0, \ldots, \gamma_{k-1}\} \in \tilde{\Pi}^k$. Since we have supposed (47) to be

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2022.3151189, IEEE Transactions on Automatic Control

true for $k+1$, it follows that $\mathrm{E}[\alpha_{k+1}]$ does not depend on $\{\gamma_0,\dots,\gamma_{k-1}\}\in\tilde{\Pi}^k$. We then have from (61) that (47) is true for k as well. Therefore, by induction, (47) holds for every $k\in\{1,2,\dots,N-1\}$, implying from the first part of this proof that (54) does too, completing the proof. □

Remark 4. *Since the NDECP definitions 5, 6 and 7 are equivalent, it follows that Theorem 2 also holds if the control problem satisfies any of these versions of the NDECP. The same is true for the sample-based NDECP mutually equivalent definitions 2, 4 and 6, since a control problem satisfying any of these will also satisfy Definition 7.*

Appendix

Equivalence between Definitions 2, 3 and 4

The result is a direct consequence of the following lemma.

Lemma 1. *Let $\{s_i(\omega,\gamma_0,\dots,\gamma_{i-1})\}_{i=1}^N$ be a collection of random variables possibly depending on the sequence of parameters (policies) $\{\gamma_0,\dots,\gamma_{N-1}\}\in\Pi^N=\Pi_0\times\cdots\times\Pi_{N-1}$. Let $\{\gamma_0^\star,\dots,\gamma_{N-1}^\star\}\in\Pi^N$ be a specific choice of these parameters. Then the following statements are equivalent*

(i)

$$s_t(\omega,\gamma_0,\dots,\gamma_k,\gamma_{k+1}^\star,\dots,\gamma_{t-1}^\star) \text{ does not depend on } \{\gamma_0,\dots,\gamma_k\}\in\Pi^{k+1},\quad 0\le k<t\le N,\ \omega\in\Omega. \tag{63}$$

(ii)

$$s_t(\omega,\gamma_0,\dots,\gamma_k,\gamma_{k+1}^\star,\dots,\gamma_{t-1}^\star) \text{ does not depend on } \gamma_k\in\Pi_k,\quad 0\le k<t\le N,\ \omega\in\Omega \tag{64}$$

(iii)

$$s_t(\omega,\gamma_0,\dots,\gamma_{t-1}) \text{ does not depend on } \{\gamma_0,\dots,\gamma_{t-1}\}\in\Pi^t, \tag{65}$$

$$0<t\le N,\ \omega\in\Omega \tag{66}$$

Proof. (iii)⇒(ii): Immediate since (64) is a special case of (65).

(ii)⇒(iii): For $k=t-1$, (64) becomes

$$s_t(\omega,\gamma_0,\dots,\gamma_{t-1}) \text{ does not depend on } \gamma_{t-1}\in\Pi_{t-1},\quad 0<t\le N,\ \omega\in\Omega \tag{67}$$

Thus, for every pair of sequences $\{\dot\gamma_i\}_{i=0}^{t-1},\{\ddot\gamma_i\}_{i=0}^{t-1}\in\Pi^t$, we have

$$\begin{aligned}
&s_t(\omega,\dot\gamma_0,\dots,\dot\gamma_{t-2},\dot\gamma_{t-1}) && (68)\\
&\overset{(67)}{=} s_t(\omega,\dot\gamma_0,\dot\gamma_1,\dots,\dot\gamma_{t-3},\dot\gamma_{t-2},\gamma_{t-1}^\star),\ \omega\in\Omega && (69)\\
&\overset{(64)}{=} s_t(\omega,\dot\gamma_0,\dot\gamma_1,\dots,\dot\gamma_{t-3},\gamma_{t-2}^\star,\gamma_{t-1}^\star),\ \omega\in\Omega && (70)\\
&\ \vdots\\
&\overset{(64)}{=} s_t(\omega,\dot\gamma_0,\gamma_1^\star,\dots,\gamma_{t-3}^\star,\gamma_{t-2}^\star,\gamma_{t-1}^\star),\ \omega\in\Omega, && (71)\\
&\overset{(64)}{=} s_t(\omega,\ddot\gamma_0,\gamma_1^\star,\dots,\gamma_{t-3}^\star,\gamma_{t-2}^\star,\gamma_{t-1}^\star),\ \omega\in\Omega, && (72)\\
&\ \vdots\\
&\overset{(64)}{=} s_t(\omega,\ddot\gamma_0,\ddot\gamma,\dots,\ddot\gamma_{t-3},\ddot\gamma_{t-2},\gamma_{t-1}^\star),\ \omega\in\Omega, && (73)\\
&\overset{(67)}{=} s_t(\omega,\ddot\gamma_0,\ddot\gamma,\dots,\ddot\gamma_{t-3},\ddot\gamma_{t-2},\ddot\gamma_{t-1}),\ \omega\in\Omega, && (74)
\end{aligned}$$

which means (ii)⇒(iii) and thus (ii)⇔(iii).

(i)⇒(iii): Immediate since (65) is obtained from (63) by choosing $t=i$.

(iii)⇒(i): Immediate since (63) is a special case of (65). This completes the proof. □

References

[1] D. Bertsekas, *Dynamic Programming and Optimal Control Vol. 1*. Athena Scientific, 2000.

[2] H. Kushner, *Introduction to Stochastic Control Theory*. New York: Holt, Rinehart and Winston, 1972.

[3] P. E. Caines, *Linear Stochastic Systems*. New York, NY: John Wiley & Sons, 1988.

[4] A. Yushkevich, “Reduction of a controlled Markov model with incomplete data to a problem with complete information in the case of Borel state and control spaces,” *Theory Prob. Appl.*, vol. 21, pp. 153–158, 1976.

[5] D. Rhenius, “Incomplete information in Markovian decision models,” *Ann. Statist.*, vol. 2, pp. 1327–1334, 1974.

[6] Y. Bar-Shalom and E. Tse, “Dual effect, certainty equivalence, and separation in stochastic control,” *IEEE Transactions on Automatic Control*, vol. AC-19, no. 5, pp. 494–500, Oct. 1974.

[7] H. Theil, “A note on certainty equivalence in dynamic planning,” *Econometrica: Journal of the Econometric Society*, pp. 346–349, 1957.

[8] H. A. Simon, “Dynamic programming under uncertainty with a quadratic criterion function,” *Econometrica, Journal of the Econometric Society*, pp. 74–81, 1956.

[9] H. Theil, “Econometric models and welfare maximization,” in *Henri Theil’s Contributions to Economics and Econometrics*. Springer, 1992, pp. 1055–1075.

[10] J. W. Patchell and O. L. R. Jacobs, “Separability, neutrality and certainty equivalence,” *International Journal of Control*, vol. 13, no. 2, pp. 337–342, 1971.

[11] A. A. Feldbaum, “Dual control theory,” *Automation and Remote Control*, vol. 21, pp. 1033–1039, May 1961.

[12] H. V. de Water and J. Willems, “The certainty equivalence property in stochastic control theory,” *IEEE Transactions on Automatic Control*, vol. 26, no. 5, pp. 1080–1087, 1981.

[13] A. Lindquist, “On feedback control of linear stochastic systems,” *SIAM Journal on Control*, vol. 11, no. 2, pp. 323–343, 1973.

[14] H. Kushner, *Stochastic stability and control*. New York: Academic Press, 1967.

[15] W. M. Wonham, “On the separation theorem of stochastic control,” *SIAM Journal on Control*, vol. 6, no. 2, pp. 312–326, 1968.

[16] V. E. Beneš, “Existence of optimal stochastic control laws,” *SIAM Journal on Control*, vol. 9, no. 3, pp. 446–472, 1971.

[17] M. H. A. Davis and P. Varaiya, “Information states for linear stochastic systems,” *Journal of Mathematical Analysis and Applications*, vol. 37, no. 2, pp. 384–402, 1972.

[18] W. Fleming and E. Pardoux, “Optimal control for partially observed diffusions,” *SIAM J. Control Optim.*, vol. 20, no. 2, pp. 261–285, 1982.

[19] T. T. Georgiou and A. Lindquist, “The separation principle in stochastic control, redux,” *IEEE Transactions on Automatic Control*, vol. 58, no. 10, pp. 2481–2494, 2013.

[20] S. Yüksel, “Jointly optimal LQG quantization and control policies for multi-dimensional linear Gaussian sources,” *IEEE Transactions on Automatic Control*, vol. 59, pp. 1612–1617, June 2014.

[21] ——, “A note on the separation of optimal quantization and control policies in networked control,” *SIAM Journal on Control and Optimization*, vol. 57, no. 1, pp. 773–782, 2019.

[22] J. Max, “Quantizing for minimum distortion,” *IRE Trans. Inf. Theory*, pp. 7–12, 3 1960.

[23] D. Blackwell, “Memoryless strategies in finite-stage dynamic programming,” *Annals of Mathematical Statistics*, vol. 35, pp. 863–865, 1964.

[24] S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. New York: Springer, 2013.