European Journal of Operational Research 270 (2018) 260–281

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Decision Support

A linear-quadratic Gaussian approach to dynamic information acquisition☆

Thomas A. Weber[a,*], Viet Anh Nguyen[b]

[a] Chair of Operations, Economics and Strategy, École Polytechnique Fédérale de Lausanne, Station 5, Lausanne CH-1015, Switzerland
[b] École Polytechnique Fédérale de Lausanne, Station 5, Lausanne CH-1015, Switzerland

ARTICLE INFO

Article history:
Received 15 February 2017
Accepted 2 March 2018
Available online 8 March 2018

Keywords:
Dynamic emissions regulation
Information acquisition
Infinite-horizon optimal control
Linear-quadratic systems
Markov decision problems

ABSTRACT

We consider optimal information acquisition for the control of linear discrete-time random systems with noisy observations and apply the findings to the problem of dynamically implementing emissions-reduction targets. The optimal policy, which is provided in closed form, depends on a single composite parameter which determines the criticality of the system. For subcritical systems, it is optimal to perform "noise leveling," that is, to reduce the variance of the state uncertainty to an optimal level and keep it constant by a steady feed of information updates. For critical systems, the optimal policy is "noise attenuation," that is, to substantially decrease the variance once and never acquire information thereafter. Finally for supercritical systems, information acquisition is never in the best interest of the decision maker. In each case, an explicit expression of the value function is obtained. The criticality of the system, and therefore the tradeoff between spending resources on the control or on information to improve the control, is influenced by a "policy parameter" which determines the importance a decision maker places on uncertainty reduction. The dependence of the system performance on the policy parameter is illustrated using a practical climate-control problem where a regulator imposes state-contingent taxes to probabilistically attain emissions targets.

1. Introduction

The effective control of a stochastic system critically depends on sufficient information about its state. The optimal acquisition of state information balances the expected increase of the decision maker's value with the cost of the signal that is being acquired. The quality of the state information determines the precision with which, at any given point in time, the decision maker can condition the choice of the best available action on the actual system behavior. For example, when trying to implement greenhouse-gas emission-reduction targets a regulator can set taxes or quotas. The target for such stock pollutants are usually expressed in terms of aggregate emissions that should stay within a carbon budget. The latter is almost linearly related to the projected increase in average temperature (IPCC, 2014). Hence, while it is possible to steer aggregate emissions *in expectation* to a given target by dynamically setting carbon prices, the probability of the *actual* state being close to the target hinges on the quality of the acquired information about the emissions level. In this paper, we provide a closed-form solution to the combined control and information-acquisition problem for linear systems and quadratic costs with one-dimensional state. The application to emissions control is then discussed based on an established model by Hoel and Karp (2002) using recent global emissions data and targets (IEA, 2015).

The importance of combining the optimal control of a system with the estimation of its state was first recognized for engineering applications (Meier, 1965). Upon investigation, it was quickly realized that the estimation and optimization problems can be decoupled, in both discrete time (Striebel, 1965) and continuous time (Wonham, 1968), resulting in a "separation principle" (Davis, 1977; Fleming & Rishel, 1975). Yet, in virtually all of the extant work, the precision of the information about the state is taken as given.

☆ The authors would like to thank two anonymous referees and participants of the Moscow Summer Academy on Economic Growth and Governance of Natural Resources at Moscow State University (Russia), the 12th International Conference on Operations Research in Havana (Cuba), the 10th Conference of the Association of Asia-Pacific Operational Research Societies (APORS) in Kuching (Malaysia), the 2015 INFORMS Annual Meeting in Philadelphia, PA (USA), and the 10th Conference of the International Federation of Operational Research Societies (IFORS) in Barcelona (Spain) for their helpful comments and suggestions.

* Corresponding author.
E-mail addresses: thomas.weber@epfl.ch (T.A. Weber), viet-anh.nguyen@epfl.ch (V.A. Nguyen).

https://doi.org/10.1016/j.ejor.2018.03.003

That is, the decision maker remains unconcerned with the problem of acquiring an appropriate amount of information.[1] Here we consider the linear-quadratic control problem with costly information of varying precision as an archetypical case which can be solved completely. We show that the separation principle applies and that the best information-acquisition policy achieves an optimal noise level over and above the system noise.[2] The optimal information-acquisition policy is implemented by a threshold policy that at any time depends on the current variance of the decision maker's beliefs about the state of the system. The precise nature of this policy can be fully characterized by a "discriminant" of the problem.

In practice, a decision maker's incentives to acquire information involve more complex considerations than what is contained in the initially formulated combined control and information-acquisition problem. For example, it may be desirable to maximize the probability of the system's state to be close to a specified target state at a given date. For this, the decision maker can tune a "policy parameter" which describes the tradeoff between the control problem (referred to as *system-stabilization problem*) and the uncertainty-reduction problem (referred to as *information-acquisition problem*). The policy parameter modulates therefore the decision maker's information-acquisition effort. Our analysis addresses the comparative statics of the problem with respect to the policy parameter, thus illustrating the structural insights that can be obtained from a closed-form solution to the decision problem.

1.1. Literature

Blackwell (1951) showed that information is generally beneficial to decision makers and that more informative sources of information are of (weakly) greater value. Conversely, a pure increase in uncertainty about the state by adding noise to an information source corresponds to a "garbling" and therefore must (weakly) decrease the information value (DeGroot, 1962). The finding carries over to a Bayesian setting (Kihlstrom, 1984), and the decision maker's value for state observations with Gaussian noise is decreasing in the variance of those information sources. While there is a fairly rich work on the value of information in a (quasi-)static setting (LaValle, 1968; Lawrence, 1999), the literature on the sequential acquisition of information, after a promising start, experienced a long hiatus.[3] The seminal contribution by Wald (1947) provides a general approach to the information-acquisition problem, and then concentrates efforts on examining distributionally robust experimentation to improve a statistical decision function. Building on pieces of that initial framework, Moscarini and Smith (2001) perform an interesting analysis of (nonrobust) information acquisition in continuous time, by controlling the diffusion of a Brownian motion through the continuous acquisition of a somewhat peculiar, specially adapted sampling process so as to inform a binary decision. McCardle (1984) considers information acquisition in a discrete-time dynamic setting, where a firm gathers information to reduce uncertainty about a technology-adoption decision. The latter leads to an optimal stopping problem, where at an upper belief threshold the firm decides to adopt the technology (and stop information acquisition), at a lower belief threshold to reject the technology (and stop information acquisition), and otherwise to continue gathering information. The underlying problem of optimally stopping in a Markovian setting with costly information about an imperfectly observable state was discussed by Monahan (1980), and results about the convexity of policy regions are summarized by Lovejoy (1987). Similarly, Moore and Whinston (1986, 1987) discuss sequential information acquisition, followed by a final action.[4] By contrast, we are concerned here with problems where control interventions and information acquisition coexist from period to period. For example, in operations management, costly information about demand can help improve inventory-management decisions (DeCroix & Mookerjee, 1997). In a newsvendor setting with independent and identically distributed consumers, Milgrom and Roberts (1988) find that it is best to either survey none or all of them, i.e., to acquire either no or full information; the reason for this is a convexity in the value of additional information. In contrast to this, Fu and Zhu (2010), by using forecast-aggregation techniques, obtain a concave information value which generically leads to the optimality of intermediate levels of information acquisition. In a linear-quadratic setting, Bansal and Başar (1989) take an information-theoretic approach separating the measurement (or communications) task from the control task, in an iterative discrete-time framework. The authors consider a similar setting in continuous time (Başar & Bansal, 1994); see also Yüksel and Başar (2013, Ch. 11) for a summary of this decentralized encoder-decoder approach with ample additional references. Sims (2003) limits the flow of information by imposing a bound on the Shannon channel capacity, which then makes the amount of information collected over time subject to optimization. Provided the channel capacity is not too low, the best policy approaches an optimal signal-to-noise ratio in the long run, thus closing in on a stationary variance of the state given a stationary variance in the observational noise. Here we also consider a linear-quadratic problem setup, yet instead of imposing an exogenous limit on the information flow, we allow for a linear cost of the precision of the state observation.

The standard linear-quadratic Gaussian (LQG) optimal control problem consists in choosing the input for a linear system so as to maximize the expectation of a quadratic functional which depends on the realized trajectories of the state, the output, and the control (Anderson & Moore, 1971; Athans, 1972a). The linear-quadratic setup appears naturally in many managerial and policy-relevant contexts, such as inventory control (Holt, Modigliani, Muth, & Simon, 1960; West, 1986), error-correction mechanisms (Salmon, 1982), production smoothing and scheduling (Gallego, 1990; Naish, 1994), dynamic oligopoly (Kydland, 1975; Fudenberg and Tirole, 1986), monetary and fiscal policy (Benigno & Woodford, 2004; Pindyck & Roberts, 1974), forecasting of economic equilibria (Townsend, 1983), nonlinear pricing with learning (Bonatti, 2011), and dynamic regulation (Auray, Mariotti, & Moizeau, 2011; Friedman, 1981), to just name a few. The optimal value of the objective depends on the quality of the state observations. The corresponding linear-quadratic combined estimation and control problem was solved by Kalman (1960), resembling results by Thiele (1880) (see Lauritzen, 1981). Athans (1972b) provides an algorithm for optimally switching among a finite number of sensors in continuous time, effectively solving an offline sensor-selection

[1] A notable exception is the dynamic sensor-selection problem introduced by Athans (1972b) which, from an engineering standpoint, amounts to scheduling the best camera view onto the state. Due to the combinatorial nature of this problem, it leads only to an algorithmic solution without much structural insight; see also Section 1.1.

[2] The findings can be interpreted in terms of signal-to-noise ratios, but they appear most naturally in terms of the additionally tolerated observational noise.

[3] There is significant work on the exploration-versus-exploitation tradeoff, inherent in the multi-armed bandit problem, but here information acquisition is mixed with reward-oriented actions and they are difficult to disentangle; see Gittins, Glazebrook, and Weber (2011) for details.

[4] Applications in healthcare also have this feature, where a sequence of tests ("screening" actions) may be followed by a treatment; this is complicated by the fact that the disease progression or population characteristics may be nonstationary. Tsodikov and Yakovlev (1991) examine aperiodic cancer screening, while Maillart, Ivy, Ransom, and Diehl (2008) use a hidden Markov-chain approach. The latter is also used by Cipriano and Weber (2018) to determine when and how much sampling is needed before the decision to discontinue a public health screening for a population with declining hepatitis C prevalence.

problem within a Hamiltonian framework. At any given time, the selected sensor serves as the best-matched observational element of a Kalman filter compared to all the available information sources. The optimal sensor selection thus reflects the best tradeoff between information and the costly signal. The structural insights of this and subsequent related work in discrete time (e.g., by Gupta, Chung, Hassibi, & Murray, 2006 and Krishnamurthy, 2002) are quite limited because no closed-form solutions to the embedded dynamic programs or two-point boundary problems are available and *ad hoc* numerical methods are employed. In contrast to this, by focusing on one-dimensional systems and a natural structure of the available continuum of information sources we provide a complete closed-form solution to the combined linear-quadratic stochastic control and information-acquisition problem. Despite the linear-quadratic setup, the underlying optimization problems are nonlinear, driven by the nonlinearity of the Bayesian information update. Lindset, Lund, and Matsen (2009) examine a continuous-time finite-horizon version of the LQG problem with quadratic information-acquisition cost but do not provide an analytical solution. Their main structural insight is somewhat unsurprising: towards the end of the horizon it is best to acquire the least amount of information, since clearly the remaining time to act on it is very limited. In our discrete-time infinite-horizon setting, information acquisition may be persistent, effectively keeping the variance of the state estimate at an optimal steady-state level, after a finite time.

1.2. Outline

The remainder of this paper is organized as follows. Section 2 introduces the decision problem, consisting of a discrete-time LQG control problem with the option for the decision maker to acquire information in each period. This problem is then reduced to an equivalent decision problem with only half as many parameters. Section 3 provides a complete solution to the reduced-form decision problem, effectively decoupling the control problem and the information-acquisition problem. A mapping of the reduced-form solution to the solution of the original problem is given, as well as a detailed discussion of how the solution can be tuned using a "policy parameter" to achieve secondary policy objectives such as maximizing the probability of goal attainment. Section 4 applies the model to the global emissions-control problem, highlighting the need for information acquisition. Section 5 concludes.

2. Model

Given the real constants A and B, with $AB^2 > 0$, consider a noisy linear system,

$$\tilde{x}_{t+1} = A\tilde{x}_t + Bu_t + \tilde{\varepsilon}_t, \tag{1}$$

where $\tilde{x}_t$ (with realizations $x_t \in \mathcal{X}$) denotes the state, $u_t \in \mathcal{U}$ the control input, and $\tilde{\varepsilon}_t$ a zero-mean independent and identically distributed (i.i.d.) Gaussian noise term with positive variance $\bar{N}$, at the discrete time instants $t \in \mathcal{T} = \{0, 1, 2, \ldots\}$. We assume that the state space $\mathcal{X}$ and the control set $\mathcal{U}$ are unconstrained, so $\mathcal{X} = \mathcal{U} = \mathbb{R}$. The initial state $\tilde{x}_0$ follows a normal distribution with given mean $\bar{x}_0$ and positive variance $\bar{N}_0$, independent of the $\tilde{\varepsilon}_t$. The tuple $(\bar{x}_0, \bar{N}_0)$ characterizes the decision maker's prior belief about the state of the system at $t = 0$.[5]

Remark 1. The linear system is invariant with respect to translations in the state space. In applications it often represents the *deviation* of a system state x' from a deterministic target state $\bar{x}'$, so $x_t \equiv x'_t - \bar{x}'$.

[5] A summary of notation is provided in Appendix B.

At each time $t \in \mathcal{T}$, the decision maker observes the output z_t of the system which follows a linear law of motion,

$$\tilde{z}_{t+1} = C\tilde{x}_t + (\tilde{\eta}_t / \nu_t), \tag{2}$$

where C is a given positive constant, $\tilde{\eta}_t$ denotes a zero-mean i.i.d. Gaussian noise term with positive variance M, and $\nu_t \in \mathcal{V} = \mathbb{R}_+$ is a control input which describes the quality of the acquired state information in terms of its precision (ν_t^2/M). A value of $\nu_t = 0$ means that no new information about the state is observed at time $t+1$. In accordance with the decision maker's prior belief about x_0, an initial observation z_0 can be defined, without loss of generality, as the realization of a normal distribution with mean $C\bar{x}_0$ and variance $M_0 = C^2\bar{N}_0$.

Remark 2. In contrast to engineering systems, where the output is typically assumed to be a simultaneous reflection of the present state (Kailath, 1980), social systems naturally carry an observation delay. For example, a corporation issues a report usually for the past period, not the current period. Hence, for $t \geq 1$ the observation of the system output z_t includes a waiting period and thus depends on the previous state x_{t-1}.

2.1. Belief propagation

Based on the system dynamics in Eq. (1) and the output characteristics in Eq. (2), the decision maker uses the observations z_t to update his belief about the state $\tilde{x}_t$. The linearity of the system implies that the decision maker's posterior belief, conditional on the applied controls u_{t-1} and ν_{t-1} in the previous period and the current realization of z_t, is normally distributed with mean $\hat{x}_t$ and variance $\hat{V}_t$.

Proposition 1. *Let $(\hat{x}_0, \hat{V}_0) = (\bar{x}_0, \bar{N}_0)$. Then for any $t \geq 1$, the decision maker's belief conditional on the history $\mathcal{H}_t = \{(z_1, \ldots, z_t), (u_0, \ldots, u_{t-1}), (\nu_0, \ldots, \nu_{t-1})\}$ is normally distributed with mean $\hat{x}_t$ and variance $\hat{V}_t$, such that*

$$\hat{x}_t = A\hat{x}_{t-1} + Bu_{t-1} + \frac{AC\,\nu_{t-1}^2 \hat{V}_{t-1}}{M + C^2\nu_{t-1}^2\hat{V}_{t-1}}(z_t - C\hat{x}_{t-1}), \tag{3}$$

$$\hat{V}_t = \bar{N} + A^2\left(1 - \frac{C^2\,\nu_{t-1}^2\hat{V}_{t-1}}{M + C^2\,\nu_{t-1}^2\hat{V}_{t-1}}\right)\hat{V}_{t-1}. \tag{4}$$

Whenever $\nu_{t-1} = 0$, the update $(\hat{x}_t, \hat{V}_t)$ is affine in $(\hat{x}_{t-1}, \hat{V}_{t-1}, u_{t-1})$.

The updated state estimate $\hat{x}_t$ in Eq. (3) corresponds to the predicted state based on the evolution of the previous state estimate $\hat{x}_{t-1}$ in Eq. (1) and a correction term that is proportional to the difference between the actual output observation z_t and the predicted output observation $C\hat{x}_{t-1}$ in Eq. (2). The variance is bounded from below by the variance of the system noise $(\bar{N})$, and its (nonlinear) adjustment in Eq. (4) is increasing in the variance of the measurement noise (M/ν_{t-1}^2). For any given information quality ν_t, Eqs. (3) and (4) present a version of discrete-time Kalman filtering (Kalman, 1960) with delayed state observation. For perfect observation quality (i.e., when $\nu_t \to \infty$), the system variance remains at its minimum value, $\bar{N}$.

Remark 3. The law of motion (3) for the estimated mean $\hat{x}_t$ at time t can be rewritten in terms of the realization ω_{t-1} of a standard normal distribution at time $t-1$,

$$\hat{x}_t = A\hat{x}_{t-1} + Bu_{t-1} + \frac{AC\,\nu_{t-1}\hat{V}_{t-1}}{\sqrt{M + C^2\nu_{t-1}^2\hat{V}_{t-1}}}\,\omega_{t-1}, \tag{3'}$$

for all $t \geq 1$. Based on (3′) and (4) the decision maker can therefore also condition his belief updates on the history $\hat{\mathcal{H}}_t = \{(\omega_0, \ldots, \omega_{t-1}), (u_0, \ldots, u_{t-1}), (\nu_0, \ldots, \nu_{t-1})\}$ instead of $\mathcal{H}_t$. This transformation effectively normalizes the output variance to one.

2.2. Decision problem

The decision maker's goal is to steer the estimated state $\hat{x}_t$ of the system in Eq. (3) (or alternatively, Eq. (3′)) towards a target and at the same time reduce the uncertainty about target achievement in form of the state variance $\hat{V}_t$ in Eq. (4) as much as possible. The tradeoff between these primary objectives is described by weights for the cost of the estimated state deviation and variance in the objective function.[6] The decision maker disposes of a 'standard' system control u_t to regulate the uncertain state in the linear system in Eq. (1) and a 'nonstandard' information control v_t to regulate the information acquisition about the state in Eq. (2). The tradeoff between target achievement and uncertainty reduction determines the amount of resources used for the system control and the information control, respectively. Overall, the decision maker needs to find control trajectories $u = (u_0, u_1, u_2 \ldots)$ and $v = (v_0, v_1, v_2, \ldots)$ to solve the infinite-horizon optimal control problem

$$\begin{aligned}
&\hat{K}(\bar{x}_0,\bar{N}_0)\\
&= \inf \quad \sum_{t=0}^{\infty} \beta^t \, \mathbb{E}\Big[p_2 \hat{x}_t^2 + p_1 \hat{x}_t + p_0 + q\hat{V}_t + \gamma u_t^2 + \delta v_t^2 \Big| \bar{x}_0, \bar{N}_0 \Big],\\
&\quad \text{s.t.} \quad \hat{x}_{t+1} = A\hat{x}_t + Bu_t + \frac{AC\, v_t \hat{V}_t}{\sqrt{M + C^2 v_t^2 \hat{V}_t}}\, \omega_t, \quad \hat{x}_0 = \bar{x}_0,\\
&\qquad\quad \hat{V}_{t+1} = \bar{N} + A^2 \left(1 - \frac{C^2 v_t^2 \hat{V}_t}{M + C^2 v_t^2 \hat{V}_t} \right) \hat{V}_t, \quad \hat{V}_0 = \bar{N}_0,\\
&\qquad\quad (u_t, v_t) \in \mathcal{U} \times \mathcal{V}, \quad t \in \mathcal{T},
\end{aligned} \tag{P}$$

where $(\bar{x}_0, \bar{N}_0)$ describes the decision maker's prior belief about the distribution of the state at time $t = 0$, and $\beta \in (0, 1)$ is a given discount factor. The cost parameters p_0, p_1, p_2, q, γ, δ are assumed positive, except for p_0 and p_1 which may take nonpositive values. Given a target state $\bar{x}' = -p_1/(2p_2)$ as in Remark 1 and $p_0 = p_1^2/(4p_2)$, the objective function in (P) penalizes the target deviation: $p_2 \hat{x}_t^2 + p_1 \hat{x}_t + p_0 = p_2 \big(\hat{x}_t - \bar{x}'\big)^2$; without loss of generality, the constant p_0 can be set to zero.

Remark 4. The problem (P) is more general than the standard linear-quadratic Gaussian (LQG) optimal control problem with information acquisition of the form

$$\sum_{t=0}^{\infty} \beta^t \, \mathbb{E}\Big[p_2 \tilde{x}_t^2 + p_1 \tilde{x}_t + \gamma u_t^2 + \delta v_t^2 \Big| \mathcal{H}_0 \Big] \longrightarrow \min,$$

subject to (1) and (2), where we have set $\mathcal{H}_0 \triangleq \hat{\mathcal{H}}_0 \triangleq \{\bar{x}_0, N_0\}$. In the extant literature on the discrete-time LQG problem, information acquisition cannot be actively controlled, so $v_t \equiv$ const. By the law of iterated expectations,

$$\begin{aligned}
\mathbb{E}\Big[\mathbb{E}\Big[\tilde{x}_t^2 \Big| \hat{\mathcal{H}}_t \Big] \Big| \mathcal{H}_0 \Big] &= \mathbb{E}\Big[\mathbb{E}\Big[(\hat{x}_t + (\tilde{x}_t - \hat{x}_t))^2 \Big| \hat{\mathcal{H}}_t \Big] \Big| \mathcal{H}_0 \Big]\\
&= \mathbb{E}\Big[\mathbb{E}\Big[\hat{x}_t^2 + 2(\tilde{x}_t - \hat{x}_t) + (\tilde{x}_t - \hat{x}_t)^2 \Big| \hat{\mathcal{H}}_t \Big] \Big| \mathcal{H}_0 \Big]\\
&= \mathbb{E}\Big[\hat{x}_t^2 + \hat{V}_t \Big| \bar{x}_0, \bar{N}_0 \Big],
\end{aligned}$$

for all $t \in \mathcal{T}$, where by construction $\hat{x}_t = \mathbb{E}[\tilde{x}_t | \hat{\mathcal{H}}_t]$ and $\hat{V}_t = \mathbb{E}[\,(\tilde{x}_t - \hat{x}_t)^2 \big| \hat{\mathcal{H}}_t]$; see Section 2.1. Thus, the above LQG problem emerges as a special case of the decision problem (P) for $q = p_2$.

[6] The choice of the weight influences the characteristics of the solution, and can therefore be viewed as a policy problem; see Section 3.4 for details.

2.3. Parameter reduction

The decision problem (P) depends on the 12-dimensional parameter vector

$\theta \triangleq (A, B, C, p_0, p_1, p_2, q, M, \bar{N}, \beta, \gamma, \delta).$

While in a given practical situation each of these parameters can be directly interpreted, it turns out that half of them are not needed to fully characterize an optimal solution.

Proposition 2. *For any given initial belief $(\bar{x}_0, \bar{N}_0)$, the decision problem (P) is equivalent to the reduced-form decision problem*[7]

$$\begin{aligned}
&\hat{K}(\bar{x}_0, \quad \bar{N}_0) = \frac{p_0}{1-\beta} + (p_2 + q)\lambda^2\\
&\times \left(\inf \quad \sum_{t=0}^{\infty} \beta^t \, \mathbb{E}\Big[\big(1 - r\big)\hat{x}_t^2 + s\hat{x}_t + r\hat{V}_t + u_t^2 + v_t^2 \Big| \bar{x}_0', N_0' \Big] \right),\\
&\text{s.t.} \quad \hat{x}_{t+1} = a\hat{x}_t + bu_t + \frac{a\, v_t \hat{V}_t}{\sqrt{1 + v_t^2 \hat{V}_t}}\, \omega_t, \quad \hat{x}_0 = \bar{x}_0',\\
&\qquad \hat{V}_{t+1} = N + a^2 \left(1 - \frac{v_t^2 \hat{V}_t}{1 + v_t^2 \hat{V}_t} \right) \hat{V}_t, \quad \hat{V}_0 = N_0',\\
&\qquad (u_t, v_t) \in \mathcal{U} \times \mathcal{V}, \quad t \in \mathcal{T},
\end{aligned} \tag{P′}$$

with $a \triangleq A$, $b \triangleq B\sqrt{(p_2+q)/\gamma}$, $r \triangleq q/(p_2+q)$, $s \triangleq p_1/[(p_2+q)\lambda]$, and $N \triangleq \bar{N}/\lambda^2$, where $\lambda^2 \triangleq \sqrt{\delta M/(p_2+q)}/C$. The reduced-form initial belief $(\bar{x}_0', N_0')$ is such that $\bar{x}_0' \triangleq \bar{x}_0/\lambda$ and $N_0' \triangleq \bar{N}_0/\lambda^2$.

The solution to the reduced-form decision problem (P′) merely requires the 6-dimensional *reduced* parameter vector

$\theta' \triangleq (a, b, r, s, N, \beta).$

As in the original decision problem (P), the solution to (P′) does not depend on p_0. The 'policy parameter' $r \in (0, 1)$ determines the relative weight that is being put on minimizing the state deviation as opposed to minimizing the state variance; see Section 3.4 for details.

Remark 5. If the decision maker wants to stabilize the estimated mean of the random state $\tilde{x}_t$ around a non-zero target $\bar{x}'$ instead of the origin, then a translation of the form $\hat{x}_t = \hat{x}_t' - \bar{x}'$ leads to precisely the reduced-form problem (P′); the variance of the translated state remains unchanged; see also Remark 1.

Since the reduced-form problem contains all generically important information about the original problem, without any loss of generality one can restrict attention to analyzing the properties of (P′) instead of the original problem (P). The mapping of the original parameter vector θ to the reduced-form parameter vector θ' is provided in Proposition 2.

3. Optimal system-stabilization and information-acquisition policies

For any given initial belief $(\bar{x}_0, N_0)$ about the Gaussian distribution of the initial state $\tilde{x}_0$, an optimal solution $(v^*, u^*) = ((u_t^*)_{t\in\mathcal{T}}, (v_t^*)_{t\in\mathcal{T}})$ to the (reduced-form) decision problem (P′) attains by Proposition 2 the optimal cost $\hat{K}(\bar{x}_0, N_0)$ in (P). If we set

$$K(x, y) \triangleq \frac{1}{(p_2+q)\lambda^2} \left(\hat{K}(\lambda x, \lambda^2 y) - \frac{p_0}{1-\beta} \right),$$

[7] The transformed state-variance tuple $(\hat{x}_t', \hat{V}_t') = (\hat{x}_t/\lambda, \hat{V}_t/\lambda^2)$ has been replaced by the original notation, $(\hat{x}_t, \hat{V}_t)$. Similarly, the transformed control, $(u_t', v_t') = (\sqrt{\gamma'} u_t, \sqrt{\delta'} v_t)$, where $(\gamma', \delta') \triangleq (\gamma, \delta)\,(p_2+q)^{-1}\lambda^{-2}$, has been replaced again by (u_t, v_t). The control sets $\mathcal{U}$ and $\mathcal{V}$ in the original problem are transformed to $\mathcal{U}' = \sqrt{\gamma'}\,\mathcal{U}$ and $\mathcal{V}' = \sqrt{\delta'}\,\mathcal{V}$ in the reduced-form problem. Given our assumptions, it is $\mathcal{U} = \mathcal{U}' = \mathbb{R}$ and $\mathcal{V} = \mathcal{V}' = \mathbb{R}_+$.

for all $(x,y)\in\mathcal{X}\times\mathcal{Y}$ with $\mathcal{Y}\triangleq\mathbb{R}_+$, then by the dynamic-programming principle (Bellman, 1954); in particular, Bertsekas and Shreve (1978/1996, Proposition 9.12), the optimal cost K satisfies the Bellman equation[8]

$$K(x,y)=\min_{(u,v)\in\mathcal{U}\times\mathcal{V}}\ \{(1-r)x^2+sx+ry+u^2+v^2+\beta\,\mathbb{E}[K(\tilde{x}',y')|x,y]\},$$
$$\text{s.t.}\quad \tilde{x}'=ax+bu+\frac{avy\,\tilde{\omega}}{\sqrt{1+v^2y}},\qquad y'=N+a^2\left(1-\frac{v^2y}{1+v^2y}\right)y, \tag{P''}$$

for all $(x,y)\in\mathcal{X}\times\mathcal{Y}$, where the random variable $\tilde{\omega}$ follows a standard normal distribution. We first establish that the optimal cost K is separable, which allows us to decompose the problem (P′′) into a stabilization problem (which yields the system control u^*) and an information-acquisition problem (which yields the information control v^*). The key idea for the separation is to first assume and then verify a quadratic cost K_1 for the stabilization portion, as in the standard linear-quadratic control problem without information acquisition. This induces additive separability of the optimization problems and restricts the impact of the nonlinearity (in the belief-propagation with respect to the precision v) to the cost K_2 of information acquisition.

Proposition 3. *The Bellman equation (P′′) is additively separable, i.e.,*[9]

$$K(x,y)\equiv K_1(x)+K_2(y),$$

with

$$K_1(x)\triangleq\min_{u\in\mathcal{U}}\Big\{(1-r)x^2+sx+u^2+\beta\Big[P(ax+bu)^2+Q(ax+bu)+R\Big]\Big\}, \tag{SS}$$

given appropriate values of $P,Q,R\in\mathbb{R}$, *and*

$$K_2(y)\triangleq\min_{v\in\mathcal{V}}\left\{ry+v^2+\beta\left[\frac{Pa^2v^2y^2}{1+v^2y}+K_2(y')\Big|_{y'=N+a^2\left(1-\frac{v^2y}{1+v^2y}\right)y}\right]\right\} \tag{IA}$$

for all states $(x,y)\in\mathcal{X}\times\mathcal{Y}$.

Problem (SS) is called the *system-stabilization problem*, while problem (IA) is referred to as the *information-acquisition problem*. The solutions of these two problems, which together amount to the solution of the (reduced-form) decision problem (P′), are now discussed in turn.

3.1. System-stabilization problem

As a function of the current state $x\in\mathcal{X}$, an optimal stabilizing policy $u^*=\mu(x)$ can be found as the unique solution to the quadratic minimization problem on the right-hand side of the Bellman equation (SS).

Proposition 4 (System stabilization)**.**

(i) *The unique solution to (SS) is* $K_1(x)=Px^2+Qx+R$, *for all* $x\in\mathcal{X}$, *where* $P>0>R$ *and*

$$P\triangleq\frac{a^2+(1-r)b^2-1/\beta+\sqrt{(a^2+(1-r)b^2-1/\beta)^2+4(1-r)b^2/\beta}}{2b^2},$$
$$Q\triangleq\frac{(Pb^2+1/\beta)s}{Pb^2-a+1/\beta},$$
$$R\triangleq-\frac{\beta}{1-\beta}\frac{(Q/2)^2b^2}{Pb^2+1/\beta}. \tag{5}$$

(ii) *The optimal system control* $u^*=\mu(x)$ *is given by*

$$\mu(x)=-\frac{b}{Pb^2+1/\beta}\left(Pa\,x+\frac{Q}{2}\right), \tag{6}$$

for all $x\in\mathcal{X}$.

The optimal stabilizing policy in Eq. (6) implements an affine state-feedback law. While the coefficient P is independent of both the precision cost δ for acquired information and the output variance M, the coefficients Q and R depend on the product δM through s (see Proposition 2). Thus, the optimal design of the state feedback in Proposition 4 takes into account the characteristics of the information-acquisition technology, including observation noise, as specified in the original problem (P). Because of the separability of the decision problem (P′) in (SS) and (IA) in Proposition 3, the optimal information-acquisition policy will implement a state feedback that only takes into account the variance of the state estimate. The benefit of information acquisition for the state feedback therefore comes not from changing the design of the feedback law for the system stabilization but from providing a more reliable input (i.e., a better state estimate) for the state feedback.

Remark 6. As is well known for linear-quadratic dynamic optimization problems, the optimal policy in Eq. (6) satisfies the "certainty equivalence principle" (Simon, 1956; Theil, 1957), in the sense that the state-feedback law is entirely independent of the stochasticity of the problem; see also Bertsekas (1995).

3.2. Information-acquisition problem

For any given *prior variance* y, an optimal information-acquisition policy $v^*=\nu(y)$ solves the nonlinear minimization problem on the right-hand side of the Bellman equation (IA) and

---

[8] If, without any loss of generality, we choose λ, p_0 so $(p_2+q)\lambda^2=1$ and $p_0=s^2/(4(1-r))$, then the objective function in problem (P′) becomes

$$\hat{K}(\bar{x}_0,N_0)=\inf\sum_{t=0}^{\infty}\beta^t\mathbb{E}\Big[g(\hat{x}_t,\hat{V}_t,u_t,v_t)\Big|\bar{x}_0,N_0\Big],$$

where $\bar{x}=-s/(2(1-r))$ is a target state (as noted after the formulation of the original decision problem (P)), so that

$$g(\hat{x}_t,\hat{V}_t,u_t,v_t)\triangleq(1-r)\big(\hat{x}_t-\bar{x}\big)^2+r\hat{V}_t+u_t^2+v_t^2$$

is a (stationary) per-period cost function, defined for all $(\hat{x}_t,\hat{V}_t,u_t,v_t)\in\mathcal{X}\times\mathcal{Y}\times\mathcal{U}\times\mathcal{V}$. Importantly, the per-period cost is nonnegative on its domain, irrespective of the sample path $(\omega_0,\omega_1,\ldots)$. Hence, in the terminology of Strauch (1966) who considered the maximization of a discounted sum of nonpositive per-period rewards, problem (P′) is a "negative dynamic programming problem." Thus, by Bertsekas and Shreve (1978/1996, Proposition 9.12), (taken from Schäl, 1975, Theorem 5.2.2) a stationary state-feedback policy $(\mu,\nu):\mathcal{X}\times\mathcal{Y}\to\mathcal{U}\times\mathcal{V}$ is optimal (i.e., it solves problem (P′)) if and only if the Bellman equation (P′′) is satisfied, for $(u,v)=(\mu,\nu)(x,y)$. Alternately, given any $\varepsilon>0$ (e.g., $\varepsilon=\alpha^*-N$; see Section 3.2), there exists an information control $v_t=\hat{v}_\varepsilon$ with some finite constant $0\le\hat{v}_\varepsilon<\infty$ such that $\hat{V}_t\le N+\varepsilon$ for all $t\ge0$, resulting in a total discounted cost of $\hat{v}_\varepsilon^2/(1-\beta)$. For this constant information-acquisition policy and a standard stabilizing control $u_t=\mu(x_t)$ with $\mu(\cdot)$ as in Eq. (6) below, the resulting total cost $\check{K}_\varepsilon(\bar{x}_0,N_0)$ remains finite and constitutes an upper bound for the optimal cost $\hat{K}(\bar{x}_0,N_0)$. It is therefore possible to restrict attention to the compact subset $\bar{\mathcal{U}}\times\bar{\mathcal{V}}_\varepsilon$ of controls, so the per-period cost remains uniformly bounded, in the sense that $0\le g(\hat{x}_t,\hat{V}_t,u_t,v_t)\le m_\varepsilon$, for all $(u_t,v_t)\in\bar{\mathcal{U}}\times\bar{\mathcal{V}}_\varepsilon$, i.e., on any "reasonably controlled" state trajectory $(\hat{x}_t,\hat{V}_t)$ starting at $(\bar{x}_0,N_0)$, where $m_\varepsilon\triangleq\check{K}_\varepsilon(\bar{x}_0,N_0)$. Then Theorem 4.2.3 and part (a) of Proposition 4.3.1 in Hernández-Lerma and Lasserre (1996) together imply that the optimal cost $\hat{K}(\bar{x}_0,N_0)$ can be obtained as unique pointwise solution to the Bellman equation in (P′′).

[9] The functional forms of K_1 (see Proposition 4) and K_2 (see Lemma 2, Corollary 1, Corollary 3, and Corollary 4) are specified below; see also Proposition 9.

thus determines the optimal information-acquisition cost $K_2(y)$. The problem of solving this recursive relation can be simplified by thinking of the optimization in terms of varying the *posterior variance* y' in (P'') instead of varying the information control ν. For any $y \in \mathcal{Y}$, there is a one-to-one relationship between the (nonnegative) information control and the posterior variance,

$$\nu^2 = \frac{N + a^2 y - y'}{y(y' - N)} = \frac{a^2}{y' - N} - \frac{1}{y} \in \mathbb{R}_+, \tag{7}$$

and consequently the decision maker's posterior variance y' is restricted to a bounded interval,

$$y' \in (N, N + a^2 y].$$

The posterior variance N, corresponding to noiseless observation of the system output, cannot be attained at finite cost, so that an optimal solution must lie strictly inside (a compact subset of) the feasible interval. With this in mind, the information-acquisition problem (IA) can be rewritten in the form

$$K_2(y) = ry - \frac{1}{y} + \beta P(N + a^2 y) + \min_{y' \in (N, N + a^2 y]} \left\{ \frac{a^2}{y' - N} + \beta \left[K_2(y') - P y' \right] \right\}, \tag{IA$'$}$$

for any $y \in \mathcal{Y}$. To solve the Bellman equation (IA$'$) one needs to distinguish the case of information acquisition, where $\nu > 0$ (or equivalently, $y' \in (N, N + a^2 y)$) and the case of no information acquisition, where $\nu = 0$ (and thus $y' = N + a^2 y$). Only in the second case does y' actually depend on y, since the minimand in (IA$'$) is independent of y. In other words, *the posterior variance depends on the prior variance only if no information is acquired.* Conversely, if information is in fact acquired, then the decision maker selects the amount of information so as to reach a target posterior variance, whereby this chosen target does *not* depend on the current variance.[10]

Remark 7. The additive separability of the information control in Eq. (7) with respect to y and y' is critical for our relatively simple solution to the information-acquisition problem. In particular, the cost-separability in (IA$'$) does not extend to cases where the information-acquisition cost is not affine in ν^2. It also does not generalize to the multi-dimensional case of a vector-valued source (with scalar information control) where (suboptimal) results may be obtained by restricting attention to specific classes of state feedback (e.g., linear).

3.2.1. Threshold optimality

Consider as target variance α^* a solution to the minimization problem in (IA$'$). Given a continuously differentiable optimal cost $K_2(\cdot)$, the first- (and second-) order necessary optimality conditions of the minimization problem yield the (y-)*conditional target variance*,

$$\alpha(y) = \sup \mathcal{A}_y, \tag{8}$$

for all $y \in \mathcal{Y}$, where the (y-)*conditional cost-improvement* set,

$$\mathcal{A}_y \triangleq \left\{ y' \in (N, N + a^2 y] : -\frac{a^2}{(y' - N)^2} + \beta[K_2'(y') - P] < 0 \right\}, \tag{9}$$

is the set of all posterior variances with negative cost gradient, i.e., those posterior variances which may be improved upon with the help of a larger conditional target variance. Because K_2 is continuously differentiable, its derivative K_2' is bounded on any compact subinterval of $(N, N + a^2 y]$, so that $\mathcal{A}_y$ must be nonempty. Thus, the conditional target variance in Eq. (8) is well-defined. Furthermore, if $\alpha(y) = N + a^2 y$, then by Eq. (7) no information acquisition is optimal, i.e.,

$$\alpha(y) = N + a^2 y \quad \Leftrightarrow \quad \nu^* = \nu(y) = 0.$$

Because for any $\hat{y}, y \in \mathcal{Y}$ with $\hat{y} > y$ it is $\mathcal{A}_y \subset \mathcal{A}_{\hat{y}}$, necessarily also

$$\hat{y} > y \quad \Rightarrow \quad \alpha(\hat{y}) \geq \alpha(y),$$

that is, the y-conditional target variance is nondecreasing in the prior variance y. For $y \to \infty$, one obtains the (*unconditional*) *cost-improvement set* $\mathcal{A}$ as the union of the $\mathcal{A}_y$ for all $y \in \mathcal{Y}$,

$$\mathcal{A} \triangleq \bigcup_{y \in \mathcal{Y}} \mathcal{A}_y = \left\{ y' \in (N, \infty) : -\frac{a^2}{(y' - N)^2} + \beta[K_2'(y') - P] < 0 \right\}. \tag{10}$$

The corresponding (*unconditional*) *target variance* is defined as the limit of the y-conditional target variance for $y \to \infty$,[11]

$$\alpha^* \triangleq \lim_{y \to \infty} \alpha(y) = \sup \mathcal{A}. \tag{11}$$

The target variance always exists as an element of $(N, \infty]$.[12] If α^* is finite, then necessarily $\alpha(y) = \alpha^*$ for all $y \geq (\alpha^* - N)/a^2$. Provided the minimization problem in (IA$'$) is convex (to be established below), one further obtains that $\alpha(y) = N + a^2 y$ for $y < (\alpha^* - N)/a^2$, whence

$$\alpha(y) = \min\{\alpha^*, N + a^2 y\},$$

for all $y \in \mathcal{Y}$. Corresponding to the y-conditional variance threshold, by Eq. (7) one can derive an equivalent information-acquisition policy of the form

$$\nu(y) = \sqrt{\frac{a^2}{\alpha(y) - N} - \frac{1}{y}}, \quad y > 0,$$

and $\nu(0) = 0$. The last relation defines a "(y^*, α^*)-threshold policy," relative to the target variance α^* and the variance threshold

$$y^* \triangleq \frac{\alpha^* - N}{a^2} < \infty,$$

which can be written in the form

$$\nu(y) = \sqrt{\left[\frac{a^2}{\alpha^* - N} - \frac{1}{y} \right]_+} = \sqrt{\left[\frac{1}{y^*} - \frac{1}{y} \right]_+}, \quad y \in \mathcal{Y}, \tag{12}$$

where $\nu(0) \triangleq \lim_{y \to 0^+} \nu(y) = 0$, by continuous completion. If $\alpha^* = \infty$, then $y^* = \infty$, and for any finite y it is therefore best to not acquire information, so $\nu(y) = 0$ for all $y \in \mathcal{Y}$. Note that the situation with infinite variance threshold is already contained as limiting case in Eq. (12); the latter can therefore be used as a general representation for the (optimal) information-acquisition policy, including the case where there is no information collection at all (for $y^* = \infty$).

Proposition 5. *Provided that $\mathcal{A}$ is convex, the (y^*, α^*)-threshold policy in Eq. (12) is optimal.*

In what follows, we first describe the "autonomous" behavior of the variance when no information is collected. We then focus on the basic shapes of the optimal cost K_2 based on whether $y^* = \infty$ (zero information acquisition) or $y^* < \infty$ (positive information acquisition). Then we establish the convexity of $\mathcal{A}$ and the optimal variance threshold (or equivalently, the optimal target variance), depending on the problem parameters.

---

10 This is similar to the (s, S)-policy in inventory control (Arrow, Harris, & Marschak, 1951; Scarf, 1960).

11 For any increasing sequence $(y_n)_{n=1}^\infty \subset \mathcal{Y}$ with $y_n \to \infty$, as $n \to \infty$, the monotonic sequence $(\alpha(y_n))_{n=1}^\infty$ converges towards the smallest upper bound, $\lim_{n \to \infty} \alpha(y_n) = \alpha^* = \sup \mathcal{A} = \lim_{n \to \infty} (\sup \mathcal{A}_{y_n})$; see, e.g., Rudin (1976, Theorem 3.14).

12 For details on the (affine) extension of the real numbers, $\bar{\mathbb{R}} = [-\infty, +\infty]$, see, e.g., Aubin (1977, Section 1.3).

3.2.2. Autonomous variance trajectory

When the decision maker acquires no information, then the posterior variance $y' = N + a^2 y$ at the end of any given period is determined recursively by the prior variance y at the beginning of that period. Thus, using the geometric-series formula, after $n > 0$ periods without information acquisition the system variance becomes

$$y^{(n)} = \begin{cases} N + ny, & \text{if } a = 1, \\ \bar{y} + a^{2n}(y - \bar{y}), & \text{otherwise,} \end{cases} \tag{13}$$

where (for $a \neq 1$) we set [13]

$$\bar{y} \triangleq \frac{N}{1 - a^2}. \tag{14}$$

The stability properties of the system variance depend on the magnitude of a.

Lemma 1. *Let $y \in \mathcal{Y}$ be a given prior variance at a certain time $t \geq 0$, and assume that there is no (further) information acquisition. (i) For $a < 1$, the variance $\bar{y} > N$ is a globally asymptotically stable steady state, which is approached monotonically at the exponential rate $\ln(1/a^2)$: $\lim_{n\to\infty} y^{(n)} = \bar{y}$. (ii) For $a \geq 1$, the system variance diverges monotonically: $y^{(n)} \uparrow \infty$ as $n \to \infty$, at the exponential rate $\ln(a^2)$ for $a > 1$, or linearly for $a = 1$.*

The preceding result establishes an important dichotomy for the autonomous behavior of the variance, in terms of the parameter a, namely whether the system is *expanding* ($a \geq 1$) or *contracting* ($a < 1$). The noise of a contracting system, if left alone, is such that its variance converges towards the steady state $\bar{y}$ in Eq. (14), which is increasing in a. By contrast, the noise of an expanding system keeps on increasing. When left alone, i.e., without information acquisition, its variance becomes eventually larger than any finite bound.

3.2.3. Optimal cost without information acquisition

Since the law of motion for the variance in problem (P′′) becomes affine when no information is collected, the corresponding cost of the information-acquisition problem (IA′) can be obtained in closed form.[14]

Lemma 2. *Let $a < 1/\sqrt{\beta}$. If after reaching the variance y, no further information is collected, then the optimal cost in the information-acquisition problem (IA′) is*[15]

$$K_2(y) = \frac{r}{1 - \beta a^2}\left(\frac{\beta N}{1 - \beta} + y\right). \tag{15}$$

The "no-information cost" is positively affine in the prior variance y and in the variance N of the system noise. It is also linearly increasing in the policy parameter r, as well as increasing in the system parameter a and the discount factor $\beta \in (0, 1)$. For any admissible parameters, this cost stays finite.

3.2.4. Optimal cost with information acquisition

At any prior variance y, for which information acquisition is optimal, the envelope theorem (see, e.g., Dixit, 1990) applied to (IA′) yields

$$K_2'(y) = r + \beta P a^2 + \frac{1}{y^2} > 0, \tag{16}$$

implying that the optimal information-acquisition cost increases in y. This in turn means that if it is optimal to acquire information for a given prior variance y, then it is also optimal to acquire information for any prior variance that exceeds y.

Lemma 3. *If there is a variance threshold $y^* > N$ such that information acquisition is optimal for $y \geq y^*$, then the optimal cost in the information-acquisition problem (IA′) is*

$$K_2(y) = k_0 + k_1 y - \frac{1}{y}, \quad y \geq y^*,$$

where k_0 and k_1 are nonnegative constants.

The optimal cost with information acquisition is increasing and concave in the prior variance y; the constants k_0, k_1 are specified in Eq. (38) of Appendix A.

3.2.5. Information acquisition for expanding systems (a ≥ 1)

Let $a \geq 1$. Given information acquisition, by substituting K_2 from Lemma 3 into the definition of the cost-improvement set $\mathcal{A}$ in Eq. (10) the target variance α^* in Eq. (11) becomes

$$\alpha^* = \sup\left\{ y' > N : \frac{\beta}{(y')^2} - \frac{a^2}{(y' - N)^2} - \Delta < 0 \right\}, \tag{17}$$

where $\Delta \triangleq \beta((1 - \beta a^2)P - r)$ acts as a "discriminant" of the information-acquisition problem. As is shown below (see Proposition 9), whenever $\Delta \geq 0$ no information collection is optimal (i.e., $\alpha^* = \infty$). For $\Delta < 0$, the target variance becomes finite and we therefore consider the corresponding (y^*, α^*)-threshold policy in Eq. (12) with $y^* = (\alpha^* - N)/a^2$.

Lemma 4. *For an expanding system ($a \geq 1$), the cost-improvement set $\mathcal{A}$ is convex.*

The condition that $a \geq 1$ in the preceding result is tight in the sense that for any $a < 1$, there exists a discount factor $\beta \in (0, 1)$ and a prior variance $y \in \mathcal{Y}$, both large enough, so that the convexity of the minimand in (IA′) is violated.

Proposition 6. *(i) If $\Delta < 0$, then the (y_1^*, α_1^*)-threshold policy $\nu(\cdot)$ is optimal, where $y_1^* \triangleq (\alpha_1^* - N)/a^2$ is the unique optimal variance threshold, and the optimal target variance α_1^* is the unique solution of*[16]

$$\frac{\beta}{(\alpha_1^*)^2} - \frac{a^2}{(\alpha_1^* - N)^2} = \Delta. \tag{18}$$

(ii) If $\Delta \geq 0$, then no information acquisition is optimal, i.e., $y_1^* = \alpha_1^* = \infty$ and $\nu(y) \equiv 0$.

If the time-t variance y_t exceeds the variance threshold y_1^*, then it is optimal to set $\nu_t^* = \nu(y_t)$ to reduce the variance to α_1^*. Thereafter, the decision maker acquires a constant amount of information in each period $t + n$, by setting $\nu_{t+n}^* = \nu(\alpha_1^*)$, resulting in $y_{t+n} = \alpha_1^*$ for all $n > 0$. Note that in the special case where $y_t = y_1^*$, then—even though y_{t+1} is larger than y_t—information acquisition is positive, since otherwise the time-$(t+1)$ variance would exceed the target variance (because in that case $y_{t+1} = N + a^2 y_t > \alpha_1^*$).

Corollary 1. *For $y \geq y_1^*$, the optimal information-acquisition cost is*

$$K_2(y) = \left(r + \beta P a^2\right)y - \frac{1}{y} + \frac{1}{1 - \beta}\left(\beta P N - \Delta\alpha_1^* + \frac{a^2}{\alpha_1^* - N} - \frac{\beta}{\alpha_1^*}\right). \tag{19}$$

[13] The constant $\bar{y}$ takes on negative values if (and only if) $a > 1$; it can therefore be interpreted as a system variance only if $a < 1$. For a derivation of Eq. (13), see the proof of Lemma 1.

[14] No information collection may be optimal only if $a < 1/\sqrt{\beta}$, which is therefore included in Lemma 2; for details see Proposition 9 below.

[15] In the proof of Lemma 2, we use the implicit assumption that no information acquisition is optimal in a (one-sided) neighborhood of y.

[16] As solution to a quartic equation, the value of α_1^* can be obtained in closed form; see, e.g., Shmakov (2011) and the references therein. In Appendix A, Eq. (51) provides an explicit expression using Cardan's formula.

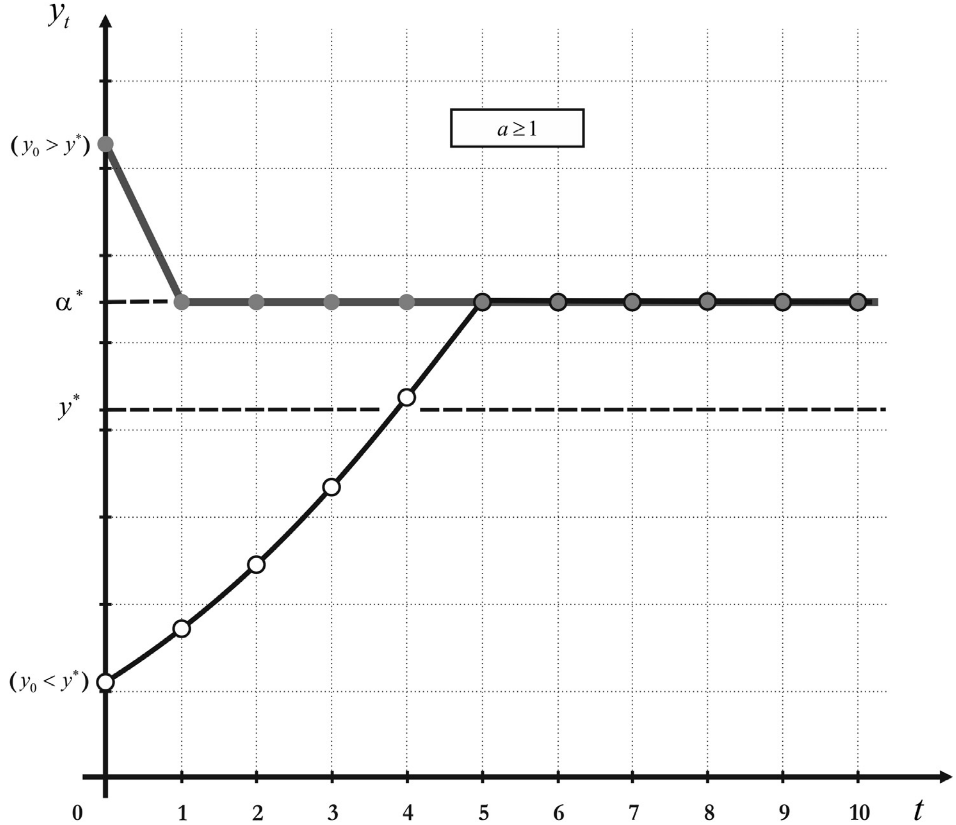

Fig. 1. Optimal variance trajectories for an expanding system.

For $y < y_1^*$, the optimal information-acquisition cost is[17]

$$K_2(y) = r \sum_{n=0}^{n^*-1} \beta^n y^{(n)} + \beta^{n^*} K_2(y^{(n^*)}), \tag{20}$$

where $y^{(n)}$, for $n > 0$, is given in Eq. (13), $y^{(0)} \triangleq y \geq y^{(n^*)}$, and

$$n^* \triangleq \inf\{n > 0 : y^{(n)} > y_1^*\} = \begin{cases} \lceil (y_1^* - N)/y \rceil, & \text{if } a = 1, \\ \left\lceil \frac{1}{\ln(a^2)} \ln\left(\frac{(1-a^2)y_1^* - N}{(1-a^2)y - N}\right) \right\rceil, & \text{otherwise.} \end{cases}$$

Despite the fact that the cost is convex in v, the threshold policy is such that it is optimal for the decision maker to decrease the system variance directly to the target variance. Thus, even if y is very large, a full decrease of the variance to α_1^* in a single period is better than multiple partial decreases. The intuition for this is that first, the information-acquisition cost is in fact linear in the precision v^2 of the observed signal. Second, the reduction in variance has not only the advantage of avoiding large future information-acquisition costs, it also decreases the cost of variance for the system, as implied by the policy parameter r. The policy parameter needs to be sufficiently large (see Proposition 11 below) for the discriminant Δ to become negative, which in turn is necessary and sufficient for the control of expanding systems to benefit from information acquisition (according to the threshold policy in Proposition 6) to be optimal. The choice of r becomes irrelevant when a is sufficiently large relative to the discount factor (so that there is always net present growth of the system in future periods).

Corollary 2. *If $a \geq 1/\sqrt{\beta}$, then the (y_1^*, α_1^*)-threshold policy is optimal.*

Fig. 1 shows a typical variance trajectory. Starting at $y > y_1^*$, one can directly reduce the system variance to α_1^*; thereafter the variance is kept at α_1^*. On the other hand, with an initial value $y < y_1^*$, there is no information acquisition until the variance surpasses y_1^*. The variance is kept at α_1^* by acquiring information of precision $(v^*)^2 = \frac{a^2}{\alpha_1^* - N} - \frac{1}{\alpha_1^*}$ in each subsequent period.

3.2.6. Information acquisition for contracting systems (a < 1)

Let $a < 1$. As noted in Section 3.2.3, the variance of a contracting system tends towards a steady state $\bar{y} > N$, in the absence of information acquisition. The optimal information-acquisition policy depends on whether the (unconditional) target variance α^* lies above or below $\bar{y}$. The reason for this comes from difference in direction of the autonomous variance movement after the optimal target has been attained. While for $\alpha^* < \bar{y}$, the variance grows without information, it shrinks autonomously when $\alpha^* > \bar{y}$. Hence, in the former situation, noise leveling is optimal (just as it is for expanding systems), whereas in the latter situation it is best to perform at most a one-time noise attenuation, and then let the variance decrease (towards the steady state) on its own.

The magnitude of the target variance depends on the magnitude of the discriminant Δ,[18] and we denote by $\hat{\Delta}$ the *subcritical*

[17] The β-discounted partial sum of variances can be written more simply as follows:

$$\sum_{n=0}^{n^*-1} \beta^n y^{(n)} = \begin{cases} \left(\frac{1-\beta^{n^*}}{1-\beta}\right) y + \frac{N\beta}{(1-\beta)^2}\left[(1-\beta^{n^*}) - (1-\beta) n^* \beta^{n^*}\right], & \text{if } a = 1, \\ \left(\frac{1-\beta^{n^*}}{1-\beta}\right)\bar{y} + n^*(y - \bar{y}), & \text{if } a = 1/\sqrt{\beta}, \\ \left(\frac{1-\beta^{n^*}}{1-\beta}\right)\bar{y} + \left(\frac{1-(\beta a^2)^{n^*}}{1-\beta a^2}\right)(y - \bar{y}), & \text{otherwise.} \end{cases}$$

[18] The dependence of α^* on Δ is monotonic: $\Delta_1 < \Delta_2$ implies that $\mathcal{A}|_{\Delta=\Delta_1} \subset \mathcal{A}|_{\Delta=\Delta_2}$, which yields $\alpha^*|_{\Delta=\Delta_1} \leq \alpha^*|_{\Delta=\Delta_2}$. Moreover, it is straightforward to show that for any $\alpha_0 \in (N, \infty]$, there exists a finite $\Delta_0 \leq 0$ such that $\alpha^*|_{\Delta=\Delta_0} = \alpha_0$.

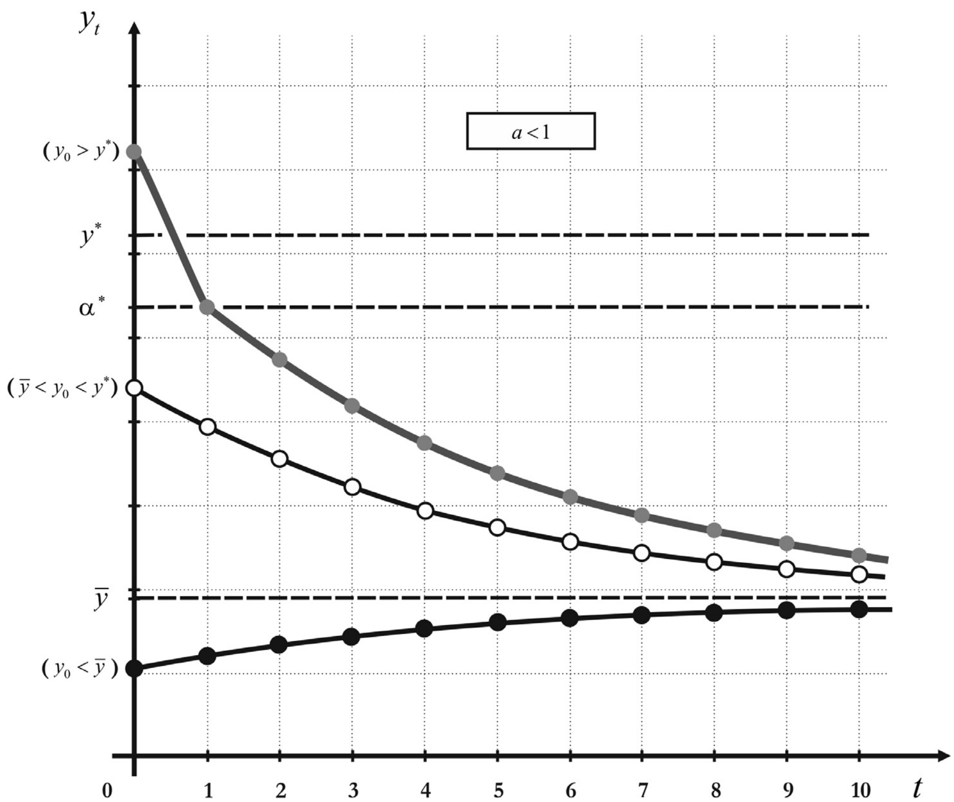

Fig. 2. Variance trajectories for a contracting system (with $y^* > \bar{y}$).

discriminant threshold, such that $\alpha^*|_{\Delta=\hat{\Delta}} = \bar{y}$. The next results establish the optimal information-acquisition policies for "subcritical systems" (where $\Delta < \hat{\Delta}$) and "critical systems" (where $\hat{\Delta} \leq \Delta < 0$), respectively.

Subcritical Systems ($\Delta < \hat{\Delta}$). In expanding systems, information collection always leads to noise leveling, and information collection takes place if the discriminant Δ is below the critical value of zero. In contracting systems, noise leveling is obtained if the discriminant lies below the subcritical discriminant threshold,

$$\hat{\Delta} \triangleq -\frac{1-\beta a^2}{(a\bar{y})^2}, \quad (<0) \tag{21}$$

which, consistent with Eq. (17) applied to the case where $a<1$, is such that $\alpha^*|_{\Delta=\hat{\Delta}} = \bar{y}$.

Lemma 5. *For a contracting system ($a<1$), the $\bar{y}$-conditional cost-improvement set $\mathcal{A}_{\bar{y}}$ is convex.*

If the target variance α^* lies in $(N, \bar{y}]$, then after reaching it, the variance will again increase (towards $\bar{y}$), thus producing a noise-leveling policy akin to the optimal information-collection policy for expanding systems.

Proposition 7. *If $\Delta < \hat{\Delta}$, then the (y_2^*, α_2^*)-threshold policy $\nu(\cdot)$ is optimal, where $y_2^* \triangleq (\alpha_2^* - N)/a^2 \in (N, \bar{y}]$ is the unique optimal variance threshold, and the optimal target variance $\alpha_2^* < \bar{y}$ is the unique solution of*

$$\frac{\beta}{(\alpha_2^*)^2} - \frac{a^2}{(\alpha_2^* - N)^2} = \Delta. \tag{22}$$

The intuition for the optimality of the (y_2^*, α_2^*)-threshold policy is simple. By construction, the target variance of a subcritical system lies in $\mathcal{A}_{\bar{y}}$. Since this target variance is in fact unconditional, by Eq. (10) the $\bar{y}$-conditional cost-improvement set $\mathcal{A}_{\bar{y}}$ must be equal to the unconditional cost-improvement set $\mathcal{A}$, which is therefore also convex. The optimality of the threshold policy is then obtained as a consequence of Proposition 5 at the outset of Section 3.2. Note also that Eq. (22) justifies *ex post* the application of Eq. (17) in the case of subcritical systems to obtain the subcritical discriminant threshold in Eq. (21).

Corollary 3. *For a subcritical system ($\Delta < \hat{\Delta}$), the information-acquisition cost $K_2(y)$ of the (y_2^*, α_2^*)-threshold policy is provided by Eq. (19) for $y \geq y_2^*$ (with α_1^* replaced by α_2^*), and by Eq. (20) for $y < y_2^*$.*

The optimal information-acquisition cost for noise-leveling policies is always given by Eqs. (19) and (20), for both expanding and subcritical contracting systems.

Critical Systems ($\hat{\Delta} \leq \Delta < 0$). When the target variance α_3^* exceeds the steady state $\bar{y}$ of a contracting system, information can be collected at most once. The reason is that the variance threshold y_3^* to trigger information acquisition can never be reached again, once the target variance has been reached.

Proposition 8. *If $\hat{\Delta} \leq \Delta < 0$, then the (y_3^*, α_3^*)-threshold policy $\nu(\cdot)$ is optimal, where $y_3^* \triangleq (\alpha_3^* - N)/a^2 = \sqrt{(1-\beta a^2)/(-\Delta a^2)}$ is the unique variance threshold, and the optimal target variance $\alpha_3^* \geq \bar{y}$ is given by*

$$\alpha_3^* = N + a\sqrt{\frac{1-\beta a^2}{(-\Delta)}}. \tag{23}$$

Fig. 2 depicts a typical variance trajectory under the noise-attenuation policy in Proposition 8. Starting at $y \geq y_3^*$, the decision maker acquires information to achieve the target variance α_3^*. Thereafter, as in the case without information acquisition (when $y < y_3^*$), the variance converges autonomously to $\bar{y}$. As before, the optimal threshold policy implies the optimal information-acquisition cost for the decision maker.

Corollary 4. *Let $\hat{\Delta} \le \Delta < 0$. For $y \ge y_3^*$, the optimal information-acquisition cost is*

$$K_2(y) = \left(r + \beta P a^2\right)y - \frac{1}{y} + \frac{1}{1-\beta a^2}\left(\frac{\beta r N}{1-\beta} - \Delta(\alpha_3^* - N)\right) + \frac{a^2}{\alpha_3^* - N}. \tag{24}$$

For $y < y_3^*$, the optimal information-acquisition cost is

$$K_2(y) = \frac{r}{1-\beta a^2}\left(y + \frac{\beta N}{1-\beta}\right). \tag{25}$$

In the remainder of this section, we first summarize the solution to the joint system-stabilization and information-acquisition problem, and then discuss the choice of the policy parameter r with respect to the reduced problem (P′′) and the original decision problem (P).

3.3. Optimal synthesis: noise leveling vs. noise attenuation

When presented with an infinite-horizon optimal control problem (P) in its general form, given a parameter vector $\theta = (A, B, C, p_0, p_1, p_2, q, M, \bar{N}, \beta, \gamma, \delta)$, the first step towards its solution is to find the reduced parameter vector $\theta' = (a, b, r, s, N, \beta)$, using the transformations in Proposition 2, for the equivalent reduced-form decision problem (P′′). The reduced-form decision problem can be decomposed into a system-stabilization problem (SS) and an information-acquisition problem (IA), which can be solved separately. The former yields the optimal system control u_t^* and the latter the optimal information control v_t^*, for all $t \ge 0$.

Proposition 9. *(i) The system-stabilization problem leads to the optimal state-feedback law $\mu(x)$ in Eq. (6), valid for any (transformed) state $x \in \mathcal{X}$, so that $u_t^* = \mu(x_t)$ for all $t \ge 0$, where the initial state x_0 is given. (ii) The solution to the information-acquisition problem (IA) consists of a feedback law $\nu(y)$ for the (transformed) variance $y \in \mathcal{Y}$, so that $v_t^* = \nu(y_t)$, for all $t \ge 0$, where the initial variance y_0 is given. The optimal information-acquisition policy is always a (y^*, α^*)-threshold policy described by Eq. (12), which can be found as follows:*[19]

1. *Determine whether the system is expanding ($a \ge 1$) or contracting ($a < 1$), and compute the discriminant for the information-acquisition problem, $\Delta \triangleq \beta((1 - \beta a^2)P - r)$, where P is given in Eq. (5).*
2. *If the system is expanding, then the (y_1^*, α_1^*)-threshold policy is optimal, as specified in Proposition 6. (For $\Delta \ge 0$, the optimal policy amounts to no information collection.)*
3. *If the system is contracting, then determine the criticality of the system by comparing the Δ to the subcritical discriminant threshold $\hat{\Delta} \triangleq -(1 - \beta a^2)/(a\bar{y})^2$, where $\bar{y} \triangleq N/(1 - a^2)$.*

(3a.) *If $\Delta < \hat{\Delta}$, then the (y_2^*, α_2^*)-threshold policy $\nu(\cdot)$ is optimal, as specified in Proposition 7.*

(3b.) *If $\hat{\Delta} \le \Delta < 0$, then the (y_3^*, α_3^*)-threshold policy $\nu(\cdot)$ is optimal, as specified in Proposition 6.*

(3c.) *If $\Delta \ge 0$, then no information acquisition is optimal, so $\nu(y) \equiv 0$.*

Fig. 3 shows the three regimes for the optimal information-acquisition policy in (r, a)-space. The intercept at $a = 1/\sqrt{\beta}$ is consistent with Corollary 2. Note also that all the curves $\Delta = \hat{\Delta}$, $\Delta = 0$, and $a = 1$ intersect at the same point, $(\bar{r}, 1)$, where $\bar{r}$ is implicitly determined by a zero discriminant at $a = 1$, i.e., by the condition $\Delta|_{(r,a)=(\bar{r},1)} = 0$.

Equipped with the optimal state-variance feedback law $(\mu, \nu) : \mathcal{X} \times \mathcal{Y} \to \mathcal{U} \times \mathcal{V}$ for the reduced-form decision problem (P′′), given any initial value (x_0, y_0), the optimal feedback law $(\hat{\mu}, \hat{\nu})$ for the original problem (P) is obtained by setting

$$(\hat{\mu}(\hat{x}), \hat{\nu}(\hat{V})) \triangleq \left(\frac{\mu(\hat{x}/\lambda)}{\sqrt{\gamma'}}, \frac{\nu(\hat{V}/\lambda^2)}{\sqrt{\delta'}}\right), \tag{26}$$

for all $(\hat{x}, \hat{V}) \in \mathcal{X} \times \mathcal{Y}$, where $\lambda \triangleq \sqrt[4]{\delta M/(p_2 + q)}/\sqrt{C} > 0$ is a scaling factor and γ', δ' are as in footnote 7. The state-variance feedback law $(\hat{\mu}, \hat{\nu})$ solves (P) for any given initial value $(\hat{x}_0, \hat{V}_0) = (\bar{x}_0, \bar{N}_0)$. The optimal cost function $\hat{K}$ for (P) is found via an affine transformation of the reduced cost $K = K_1 + K_2$ as in Proposition 2, so

$$\hat{K}(\hat{x}, \hat{V}) = \frac{p_0}{1-\beta} + (p_2 + q)\lambda^2 \cdot \left(K_1(\hat{x}/\lambda) + K_2(\hat{V}/\lambda^2)\right), \tag{27}$$

for all $(\hat{x}, \hat{V}) \in \mathcal{X} \times \mathcal{Y}$, where $K_1(\cdot)$ is specified in Proposition 4 and $K_2(\cdot)$ is given in Corollary 1 (for expanding systems), Corollary 3 (for subcritical systems), Corollary 4 (for critical systems), and Lemma 2 (for supercritical systems, with $\Delta \ge 0$).

3.4. Choosing the policy parameter

For a given parameter vector θ in the original decision problem (P), the parameter $q > 0$ describes the marginal cost of the state variance $\hat{V}_t$ at time t. When transforming (P) to its reduced form (P′) using Proposition 2, all components of the reduced parameter vector θ', with the exception of the system parameter a and the discount factor β, depend on $q = p_2 r/(1 - r)$, or equivalently, on the "policy parameter" $r = q/(p_2 + q) \in (0, 1)$.

Remark 8. To be clear, both q and r are policy parameters, the former in the original decision problem (P) and the latter in the equivalent reduced-form problem (P′′). For a given value of p_2, the relation between $q \in (0, \infty)$ and $r \in (0, 1)$ is one-to-one. The analysis of the comparative statics in the reduced policy parameter r is more convenient than in q because one obtains multiplicative separability (e.g., $p_2 + q = (1 - r)^{-1} p_2$). On the other hand, any optimal value for r implies a unique optimal value for q as well.

Let $\hat{\theta}'$ be the reduced parameter vector θ' with the components a, r, β deleted. Then, as a function of r it is

$$\hat{\theta}'(r) = (b, s, N) = \left((1-r)^{-1/2} b_0, (1-r)^{3/4} s_0, (1-r)^{-1/2} n_0\right),$$

where $\hat{\theta}'_0 = (b_0, s_0, n_0) \triangleq (B\sqrt{p_2/\gamma}, p_1/(p_2\lambda_0), \bar{N}/\lambda_0^2)$ is the reduced parameter vector for $r \to 0^+$, with $\lambda_0^2 \triangleq \sqrt{\delta M/p_2}/C$. We now examine the impact of r on the optimal feedback law in Eq. (26) and on the optimal cost in Eq. (27). The latter can be written in the form

$$\hat{K}(\hat{x}, \hat{V}) = \frac{p_0}{1-\beta} + \hat{K}_1(\hat{x}) + \hat{K}_2(\hat{V}),$$

for all $(\hat{x}, \hat{V}) \in \mathcal{X} \times \mathcal{Y}$, where $\hat{K}_1(\hat{x}) \triangleq (\frac{p_2\lambda^2}{1-r})K_1(\hat{x}/\lambda)$ and $\hat{K}_2(\hat{V}) \triangleq (\frac{p_2\lambda^2}{1-r})K_2(\hat{V}/\lambda^2)$.

3.4.1. Impact of r on system-stabilization policy

To understand the impact of the policy parameter on the optimal policy and the behavior of the regulated system in (P) we first examine how r affects the components P, Q, R of the optimal system-stabilization policy in Proposition 4. By setting $(P_0, Q_0, R_0) \triangleq (P, Q, R)|_{r \to 0^+}$, one obtains

$$(P, Q, R) = \left((1-r)P_0, (1-r)^{3/4} Q_0, (1-r)^{1/2} R_0\right).$$

[19] We refer to the information-acquisition policy under 2. and 3a. as *noise leveling*, and to the information-acquisition policy under 3b. as *noise attenuation. A closed-form solution to the quartic Eqs. (18) and (22), needed for determining the optimal thresholds of the noise-leveling policies, is provided in Appendix A.*

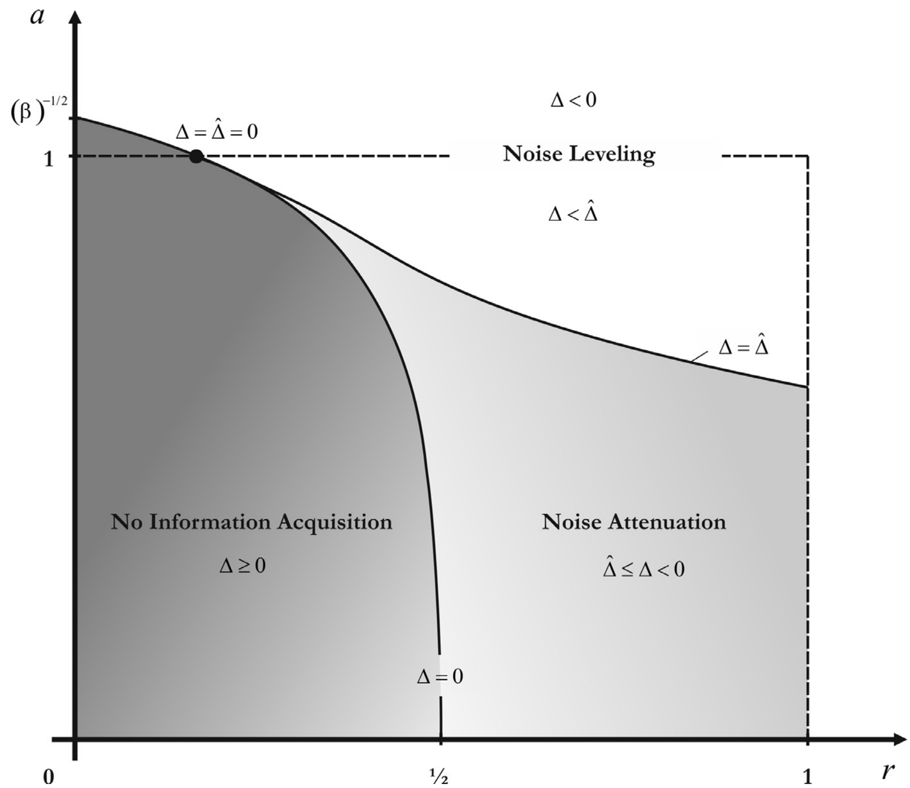

Fig. 3. Regimes of the optimal information-acquisition policy in (r, a)-space.

This dependence on r of the state feedback in the reduced problem compensates for the scaling required to obtain the solution of the original problem.

Proposition 10. *The optimal state-feedback law for (P) is of the form*[20]

$$\hat{\mu}(\hat{x}) = -\frac{(b_0/\sqrt{\gamma_0'})}{P_0 b_0^2 + 1/\beta}\left(P_0 a\left(\frac{\hat{x}}{\lambda_0}\right) + \frac{Q_0}{2}\right), \tag{28}$$

for all $\hat{x} \in \mathcal{X}$. It remains unaffected by the choice of the policy parameter r (and q), i.e., $d\mu(\hat{x})/dr \equiv 0$.

The invariance of the optimal system-stabilization policy with respect to r implies that the monetary expenditure for reaching the target state, i.e., the pure control cost ($\gamma(u_t^*)^2$ with $u_t^* = \hat{\mu}(\hat{x}_t)$) is independent of the policy parameter. Thus, the system stabilization cost in the overall objective function decreases linearly in r.

Corollary 5. *The optimal system-stabilization cost for (P) is $\hat{K}_1(\hat{x}) \equiv p_2\left[P_0\hat{x}^2 + \lambda_0 Q_0 \hat{x} + \lambda_0^2 R_0\right]$.*

The state-feedback law in Eq. (28) is nonlinear and its absolute value generically nonmonotonic in p_2.

3.4.2. Impact of r *on information-acquisition policy*

As noted in Section 3.3, depending on the value of the discriminant $\Delta = (1-r)\Delta_0 - r\beta$, where $\Delta_0 \triangleq \Delta|_{r\to 0^+} = (1-\beta a^2)\beta P_0$, the optimal information-acquisition policy for (P) is to acquire no information at all ($\Delta \geq 0$), to perform noise attenuation ($\hat{\Delta} \leq \Delta < 0$), or to perform noise leveling ($\Delta < \hat{\Delta}$). Both the discriminant Δ and the subcritical discriminant threshold $\hat{\Delta} = (1-r)\hat{\Delta}_0$, with $\hat{\Delta}_0 \triangleq \hat{\Delta}|_{r\to 0^+} = -(1-\beta a^2)/(a^2\bar{y}_0^2)$ (and $\bar{y}_0$ as given in Proposition 11 below), depend on the policy parameter r.

Proposition 11. *The optimal variance-feedback law for (P) is of the form*

$$\hat{\nu}(\hat{V}) = \lambda_0^2\sqrt{\frac{p_2}{\delta}\left[\frac{1}{\hat{V}^*(r)} - \frac{1}{\hat{V}}\right]_+}, \tag{29}$$

where $\hat{V}^*(r) \triangleq \lambda^2 y^*(r) = \lambda_0^2(1-r)^{1/2} y^*(r)$ is the variance threshold and

$$y^*(r) = \begin{cases} \infty, & \text{if } 0 < r \leq \underline{r}, \quad \text{('no information acquisition')} \\ y_1^*(r), & \text{if } \underline{r} < r < 1 \text{ and } a \geq 1, \quad \text{('noise leveling')} \\ y_2^*(r), & \text{if } \hat{r} < r < 1 \text{ and } a < 1, \quad \text{('noise leveling')} \\ y_3^*(r), & \text{if } \underline{r} < r \leq \hat{r} \text{ and } a < 1, \quad \text{('noise attenuation')} \end{cases} \tag{30}$$

is the corresponding threshold for (P′′), with the critical values

$$\underline{r} \triangleq \frac{\Delta_0}{\beta + |\Delta_0|} = \frac{(1-\beta a^2)P_0}{1 + |1-\beta a^2|P_0},$$

and

$$\underline{\hat{r}} \triangleq \frac{\Delta_0 - \hat{\Delta}_0}{\beta + |\Delta_0 - \hat{\Delta}_0|} = \frac{(1-\beta a^2)(P_0 + (a\bar{y}_0)^{-2})}{1 + |1-\beta a^2|(P_0 + (a\bar{y}_0)^{-2})},$$

for the policy parameter $r \in (0, 1)$, with $\bar{y}_0 \triangleq \bar{y}|_{r\to 0^+} = N_0/(1-a^2)$ for $a < 1$.

The dependence of the feedback law in Eq. (29) on the policy parameter r varies with the regime of the optimal information-acquisition policy.

Proposition 12. *The optimal information-acquisition policy $\hat{\nu}(\hat{V})$ is decreasing in r for noise leveling and increasing in r for noise attenuation, and otherwise constant (zero).*

[20] $P_0 = (2b_0^2)^{-1}\left(a^2 + b_0^2 - 1/\beta + \sqrt{(a^2 + b_0^2 - 1/\beta)^2 + 4b_0^2/\beta}\right)$, $Q_0 = (P_0 b_0^2 - a + 1/\beta)^{-1}(P_0 b_0^2 + 1/\beta)s_0$, and $R_0 = -(\beta/(1-\beta))\left(P_0 b_0^2 + 1/\beta\right)^{-1}(Q_0/2)^2 b_0^2$; moreover, $\gamma_0' = \gamma/(p_2\lambda_0^2)$, and $\hat{\theta}_0' = (b_0, s_0, n_0)$ with λ_0 as introduced earlier.

The piecewise monotonic behavior of information acquisition with respect to the policy parameter means that maximal one-time information acquisition (with respect to r) occurs for contracting systems when the decision maker achieves noise leveling by just doing noise attenuation.

Corollary 6. *The optimal variance-feedback law* $\hat{\nu}(\hat{V})$ *is maximal for* $r = \max\{0, \underline{r}, \hat{\underline{r}}\}$.

The overall information-acquisition cost is not necessarily monotonic in the policy parameter.

Lemma 6. *The change of the optimal information-acquisition cost with respect to the policy parameter is*

$$\frac{d\hat{K}_2(\hat{V})/dr}{\hat{K}_2(\hat{V})} = \left[\frac{1+\epsilon_2(y)}{2(1-r)} + \frac{\partial K_2(y)/\partial r}{K_2(y)}\right]\Bigg|_{y=\hat{V}/\lambda^2} \tag{31}$$

where $\epsilon_2(y) \triangleq yK_2'(y)/K_2(y) \geq 0$ *is the cost elasticity with respect to the variance.*

For large values of r, the first term in Eq. (31) dominates the second term: increasing the decision maker's attention on variance reduction also increases the overall cost of information acquisition.

3.4.3. Implications for the choice of r

The policy parameter r (or q) can be used to accomplish additional objectives in the decision problem (P). One can think of a social planner who by choosing r can adapt the optimal policy and cost.[21] For example, given the tradeoff between speed of target attainment and accuracy, it can play an important role in the credible achievement of timed policy goals, such as the reduction targets implied by the Intended Nationally Determined Contributions (INDC) from individual nations at the 2015 Climate Conference in Paris (COP21), discussed in the next section. There we show that an intermediate value of the policy r can maximize the probability of reaching a given target at a fixed finite time horizon.

4. Application: dynamic emissions control and measurement

As an application of the model, consider the problem of dynamic pollution control using closed-loop taxation (Hoel & Karp, 2002). Without a price for carbon, i.e., under business-as-usual (BAU) conditions, a representative firm's random emissions level at time $t \geq 0$ is $\bar{e} + \tilde{\theta}_t$, where $\bar{e} > 0$ denotes the stationary BAU emissions level and $\tilde{\theta}_t$ is an i.i.d. zero-mean normal random variable with homoscedastic variance $\sigma^2 > 0$. Given a time-t pollution tax $\tau_t \geq 0$ and a quadratic abatement cost $c(\bar{e} + \theta_t - e_t)^2/2$ for emission levels below the realized BAU emissions level $\bar{e} + \theta_t$ at time t (where the marginal abatement cost $c > 0$ is constant), the firm's total cost of emissions becomes

$$\mathrm{TC}(e_t) = \frac{c(\bar{e} + \theta_t - e_t)^2}{2} + \tau_t e_t.$$

[21] Selecting r so as to minimize the decision maker's cost may seem to perfectly align the social planner's and the decision maker's objectives, but in fact gives rise to an "ostrich bias" because an optimal r renders the lowest cost easiest to achieve, i.e., the one that perhaps ignores most of the difficulty. It also ignores side-benefits such as the one discussed in Section 4. A somewhat more aggressive and robust approach the social planner could take, quite different from aligning objectives with the decision maker, would be to *maximize* the competitive ratio $\rho(r) \triangleq \inf_{\hat{r}\in(0,1)} \left\{\bar{K}^*(\hat{r})/\bar{K}(u^*(r), \nu^*(r), \hat{r})\right\} \in [0, 1]$ (see, e.g., Goel, Myerson, & Weber, 2009), so the solution performs reasonably well for *any* r, and the policy parameter may therefore allowed to be ex-ante random with unknown distribution. For a given initial condition, we mean by $\bar{K}^*(r) > 0$ the optimal cost given r (assumed positive) and by $\bar{K}(u^*(r), \nu^*(r), \hat{r})$ the (generally *sub*optimal) cost, given the objective function with policy parameter $\hat{r}$, evaluated for a policy that is optimal for the policy parameter r. Given $p_0 = 0$, this analysis can be carried out for the reduced cost in the reduced-form decision problem (P″).

The representative firm's optimal emissions level e_t minimizes its total cost of emissions, so

$$e_t = \bar{e} + \theta_t - (\tau_t/c), \tag{32}$$

for all $t \geq 0$. For simplicity, let us denote the regulator's control variable by $u_t \triangleq \tau_t/c$. The regulator's goal is to reduce the current random emissions stock S_0 to a stationary target level $\bar{S}$ ($< S_0$) in the long run. In other words, the regulatory problem is to steer the random excess pollution stock $\tilde{s}_t \triangleq S_t - \bar{S}$ to zero, where S_t denotes the random pollution stock at time $t \geq 0$. The excess pollution stock evolves according to

$$\tilde{s}_{t+1} = A\tilde{s}_t + B\tilde{e}_t = A\tilde{s}_t - Bu_t + B\bar{e} + \tilde{\varepsilon}_t, \quad t \in \mathcal{T},$$

where $\tilde{\varepsilon}_t \triangleq B\tilde{\theta}_t$. The system coefficient $A \in (0, 1)$ is such that without further emissions the excess pollution stock would decrease exponentially at the rate $\ln(1/A)$, which is the fastest trajectory towards the target state. The control coefficient $B = 1$ indicates that all considered emissions are atmospheric.

The decision maker, referred to here as "policy maker," minimizes the sum of the expected abatement cost and the expectation of the (excess) environmental damage, $D(\tilde{s}_t) \triangleq d\tilde{s}_t^2/2$, where $d > 0$, solving the problem

$$\begin{aligned} \min_u \quad & \sum_{t=0}^{\infty} \beta^t \mathbb{E}\left[\frac{c(\bar{e} + \tilde{\theta}_t - e_t)^2}{2} + \frac{d\tilde{s}_t^2}{2}\middle|\bar{s}_0, \bar{N}_0\right], \\ \text{s.t.} \quad & \tilde{s}_{t+1} = A\tilde{s}_t - Bu_t + B\bar{e} + \tilde{\varepsilon}_t, \quad t \in \mathcal{T}, \end{aligned} \tag{33}$$

where $\bar{s}_0$, N_0 denote the mean and the variance of the excess pollution stock at time $t = 0$, respectively, and where $\beta \in (0, 1)$ is a given discount factor. Let $\tilde{x}_t \triangleq \kappa - \tilde{s}_t$ be the global emissions abatement (including the "natural" abatement level $\bar{S}$), where $\kappa = \frac{B\bar{e}}{1-A}$ is the BAU emissions steady state under zero taxes. Using a positive affine transformation of the objective function and substituting the firms' optimal emissions level e_t, the regulator's decision problem is equivalent to

$$\begin{aligned} \min_u \quad & \sum_{t=0}^{\infty} \beta^t \mathbb{E}\left[\frac{d\tilde{x}_t^2}{2} - d\kappa\tilde{x}_t + \frac{cu_t^2}{2}\middle|\bar{x}_0, \bar{N}_0\right], \\ \text{s.t.} \quad & \tilde{x}_{t+1} = A\tilde{x}_t + Bu_t + \tilde{\varepsilon}_t, \quad t \in \mathcal{T}, \end{aligned} \tag{34}$$

where $\tilde{\varepsilon}_t \sim \mathcal{N}(0, \bar{N})$ i.i.d., and $\tilde{x}_0 \sim \mathcal{N}(\bar{x}_0, \bar{N}_0)$, with $\bar{x}_0 \triangleq \kappa - \bar{s}_0$.

Information about the uncertain time-t stock of the pollutant, S_t, can be acquired at a cost. Because of the affine relationship between the pollution stock and the global emissions abatement, this is equivalent to obtaining an informative signal about $\tilde{x}$. Thus, through a costly pollution-stock measurement the policy maker receives a signal $\tilde{z}_{t+1} = \tilde{x}_t + (\tilde{\eta}_{t+1}/\nu_t)$, where $\tilde{\eta}_{t+1}$ is (without loss of generality) normally distributed with mean 0 and variance $M > 0$. This implies that the signal is unbiased and of precision ν_t^2/M. As in earlier sections, the cost of acquiring the informative signal is assumed to be linear in its precision. As in Section 2.2, the regulator's emissions-control problem with information acquisition is therefore

$$\begin{aligned} &\hat{K}(\bar{x}_0, \bar{N}_0) \\ &\quad = \min_{(u,\nu)} \sum_{t=0}^{\infty} \beta^t\, \mathbb{E}\left[\frac{d\hat{x}_t^2}{2} - d\kappa\hat{x}_t + q\hat{V}_t + \frac{cu_t^2}{2} + \delta\nu_t^2\middle|\bar{x}_0, \bar{N}_0\right], \\ &\quad \text{s.t.} \quad \hat{x}_{t+1} = A\hat{x}_t + Bu_t + \frac{AC\,\nu_t\hat{V}_t}{\sqrt{M + C^2\nu_t^2\hat{V}_t}}\,\omega_t, \quad \hat{x}_0 = \bar{x}_0, \\ &\quad\qquad \hat{V}_{t+1} = \bar{N} + A^2\left(1 - \frac{C^2\nu_t^2\hat{V}_t}{M + C^2\nu_t^2\hat{V}_t}\right)\hat{V}_t, \quad \hat{V}_0 = \bar{N}_0, \\ &\quad\qquad (u_t, \nu_t) \in \mathcal{U} \times \mathcal{V}, \quad t \in \mathcal{T}, \end{aligned}$$

where $(\bar{x}_0, \bar{N}_0)$ describes the decision maker's prior belief about the distribution of the state at time $t = 0$. Thus, the problem has

Table 1
Parameters for the global dynamic emissions-control and measurement problem.

Symbol	Value	Unit	Source
A	0.97	–	(IPCC, 2014)
B	1	–	(Atmospheric emissions)
C	0.75	–	(Model)
p_0	0	\$	(Model)
p_1	$-3{,}986.4 \times 10^9$ $(= -d\kappa)$	\$ /($GtCO_2$)	(Model)
p_2	1.65×10^9 $(= d/2)$	\$ /$(GtCO_2)^2$	(Model)
q	policy parameter $(= p_2 r/(1-r) > 0)$	–	(Model)
M	2.25×10^4	$(GtCO_2)^2$	(Estimated)
$\bar{N}_0 = \bar{N}$	12.96 $(= B^2\sigma^2)$	$(GtCO_2)^2$	(Model)
β	0.9524	–	(Arrow et al., 2004)
γ	15.15×10^9 $(= c/2)$	\$ /$(GtCO_2)^2$	(Model)
δ	5×10^6	\$	(Estimated)
c	30.3×10^9	\$ /$(GtCO_2)^2$	(Weber & Neuhoff, 2010)
d	3.3×10^9	\$ /$(GtCO_2)^2$	(Weber & Neuhoff, 2010)
$\bar{e}$	36.23	$GtCO_2$	(IEA, 2015; IPCC, 2014; Weber & Neuhoff, 2010)
S_0	820	$GtCO_2$	(IPCC, 2014)
$\bar{S}$	1080	$GtCO_2$	(IEA, 2015; IPCC, 2014)
$\bar{x}_0$	1468 $(= \kappa - (S_0 - \bar{S}))$	$GtCO_2$	(Model)
κ	1208 $(= B\bar{e}/(1-A))$	$GtCO_2$	(IEA, 2015; IPCC, 2014; Weber & Neuhoff, 2010)
σ	3.6	$GtCO_2$	(IPCC, 2014)

been rewritten in the form of the general decision problem (P), with parameter vector $\theta = (A, B, C, p_0, p_1, p_2, q, M, \bar{N}, \beta, \gamma, \delta)$, where $(p_0, p_1, p_2) = (0, -d\kappa, d/2)$, $\gamma = c/2$, and $\bar{N} = B^2\sigma^2$.

4.1. Model identification

The Intergovernmental Panel on Climate Change (IPCC, 2014) estimates that cumulative anthropogenic carbon-dioxide (CO_2) emissions into the atmosphere between 1750 and 2011 were 2040 ± 310 metric gigatons ($GtCO_2$). About half of the cumulative emissions occurred over the last 40 years of this period, and 40% of the cumulative emissions have remained in the atmosphere by 2011.[22] In simulation models with horizon up to the end of the 21st century, the temperature increase depends almost linearly on cumulative emissions (ibid., p. 9). Based on these estimates, the International Energy Agency notes that for the expected average temperature increase, compared with preindustrial (1870-)levels, to stay below 2 degrees Celsius with probability of 50%, the remaining “carbon budget” would allow for approximately 1000 ± 150 $GtCO_2$ of additional emissions from the year 2014 onwards (IEA, 2015, p. 18; numbers slightly rounded for convenience).

To estimate the system coefficient A, we use the IPCC (2014, p. 4) CO_2 emissions data $\hat{e}_\tau$ from $\tau_0 = 1750$ to $\tau_1 = 2010$ for the linear system $\hat{S}_{\tau+1} = A\hat{S}_\tau + B\hat{e}_\tau$ for $B = 1$ (atmospheric emissions) and $\tau \in \{\tau_0, \tau_1\}$. The annual carbon-dioxide outputs sum to the IPCC cumulative emissions estimate of about 2040 $GtCO_2$. For $A = 0.97$, the atmospheric carbon-dioxide stock $\hat{S}_{\tau_1}$ at the end of the observation period corresponds to 40% of the cumulative emissions estimate (i.e., $S_0 = \hat{S}_{\tau_1} = 820\,GtCO_2$), consistent with IPCC’s 40% atmospheric persistence estimate. With the aid of the simulated emissions data by the IEA (2015, p. 39) we find that the target emissions stock level, which exhausts the 2°C global-warming carbon budget in the year 2040, is $\bar{S} = 1080$ $GtCO_2$. The selected discount factor of $\beta = 0.9524$ corresponds to an annual “social rate of interest on consumption” of 5% (Arrow et al., 2004).

To determine the BAU emissions level $\bar{e}$, note first that about 11% of the global energy-related emissions (about 90% of the total carbon-dioxide output) was subject to emissions-trading schemes at an average price of about $p =$ \$ 7/tCO_2 (IEA 2015, p. 23). On the other hand, 13% of energy-related emissions receive consumption subsidies. Using the marginal cost estimate $c = 30.3 \times 10^9$ \$ $\times (GtCO_2)^{-2}$ by Weber and Neuhoff (2010), we separate the non-energy related and the non-priced emissions (assuming both as inelastic), so

$$\bar{e} = (5 + (89\%) \times 31 + (11\%) \times 31)\ GtCO_2 + (p/c) \approx 36.23\ GtCO_2,$$

where the current effect of carbon pricing, $p/c = (7/30.3)$ $GtCO_2 \approx 0.23$ $GtCO_2$, is almost negligible, leading to a current emissions reduction of less than 1%. For our analysis, we neglect the time trend, consistent with the relative stagnation of energy-related carbon-dioxide emissions forecasts in the INDC scenario (IEA, 2015, p. 62).[23] With this, we obtain the BAU emissions steady state under zero taxes ($\kappa = 1208$ $GtCO_2$). As in the IPCC emissions estimates (IPCC, 2014, p. 5) we assume a 10% noise level in the emissions, corresponding to a standard deviation of the macroeconomic uncertainty of $\sigma = 3.6$ $GtCO_2$.

In 2011, the atmospheric concentration of CO_2 is estimated at 430 ± 90 parts per million (ppm), corresponding to a coefficient of variation of about 20%. Given a quasi-linear relationship between atmospheric CO_2 concentration and cumulative emissions, for the cumulative excess emission to be observable at 15% via the concentration (at 20% standard deviation), the observation coefficient is $C = 15/20 = 0.75$. Therefore the observed signal z_t of the global emissions abatement $\tilde{x}_t$ must be in the order of 10^3 $GtCO_2$.[24] As a result, $M \approx (0.15 \times 1000)^2 (GtCO_2)^2$. Assuming, for simplicity, that atmospheric concentration estimates are based on measurements by the NOAA Mauna Loa, Hawai’i, observatory alone and that the annual operating budget for a comparable weather station is \$ 5 million, the use of m such (spatially i.i.d.) measurements increases the cost linearly and decreases the observational variance M proportionally; the information-acquisition cost becomes $\delta =$ \$ 5×10^6, where $m = v_t^2$ denotes the employed number of observatories with i.i.d. measurements; see Table 1.

4.2. Practical considerations

To limit global warming to 2°C with probability 50% by 2040, the remaining emissions budget of about 1,000 $GtCO_2$ must not

[22] Absorption takes place via storage in plants and soils; oceans have absorbed about 30% of total emissions (ibid., p. 4).

[23] The simple model in this paper incorporates neither economic growth dynamics, nor innovation and substitution of traditional energy sources by renewables, thus implicitly assuming that one effect cancels the other.

[24] For example, $\bar{x}_0 = \kappa - (S_0 - \bar{S}) = 1468$ $GtCO_2$, so $\mathbb{E}[\tilde{z}_0] = (3/4)\mathbb{E}[\tilde{x}_0] = 1101$ $GtCO_2$.

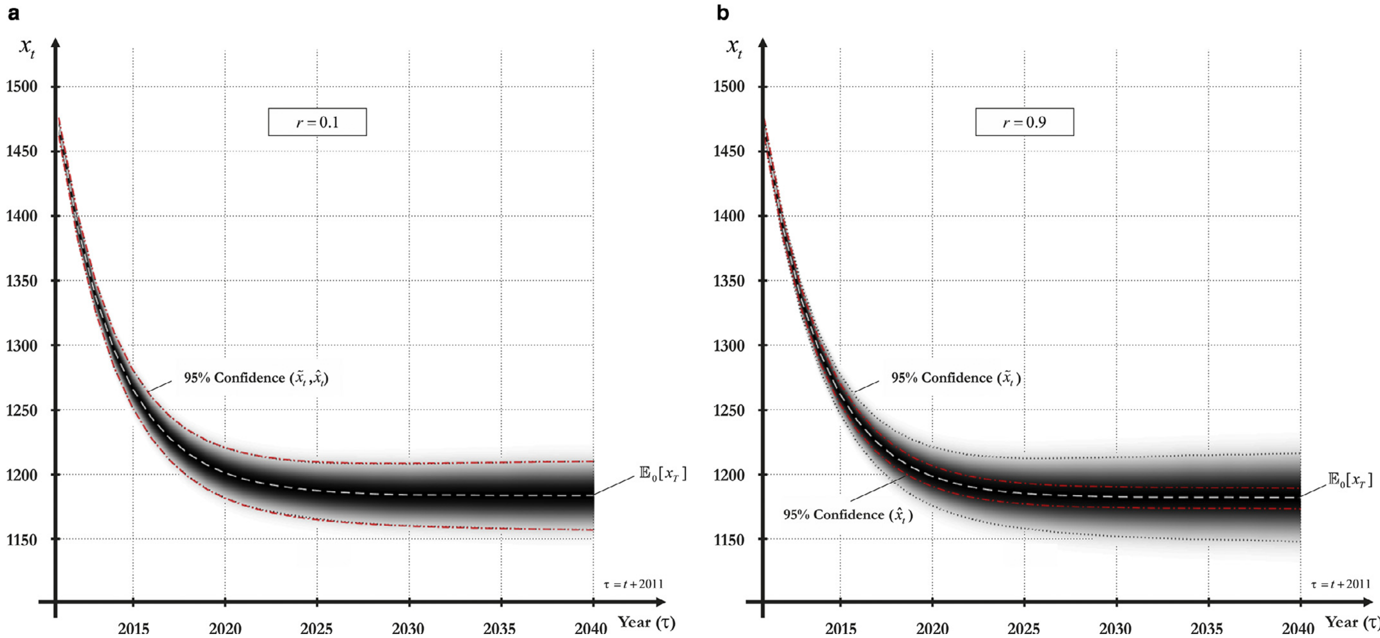

Fig. 4. Uncertain state evolution for (a) $r = 0.1$ and (b) $r = 0.9$, with an uncertain initial value $\tilde{x}_0$.

be spent before the end of this horizon (IPCC, 2014). The parameter values in Section 4.1 were chosen as if the global emissions-control and measurement problem was formulated in infinite horizon. Naturally, because of the many uncertainties regarding the distant future (see Fig. 4), it is unreasonable to expect the infinite-horizon solution of the decision problem (P) to be of great use even after, say, 10 to 20 years. Yet, it is always possible to resolve the problem in the future, leading to a receding-horizon policy (see, e.g., Bemporad & Morari, 1999; Weber, 1997).[25] To heed the IPCC recommendations, the decision maker can use the policy parameter r (or q) to maximize the probability of being in the target interval with cumulative emissions (since 1750) of 3040 ± 150 $GtCO_2$. In our model (which takes into account atmospheric self-cleansing), this corresponds to $S_T \in [1025, 1135]$ $GtCO_2$, where $T = 2040 - 2011$ (with 2011 as the starting date). Thus, as mentioned in Section 3.4.3, the policy maker may select a policy-parameter value $r = r^*$ (or equivalently, $q^* = p_2 r^*/(1 - r^*)$) so the estimated global emissions abatement $\hat{x}_T$, at $t = T$, lies in some interval. For example, to achieve global carbon-abatement goals, a decision maker may set the policy parameter to maximize the likelihood that $\mathbb{E}_T[S_T]$ (which is a random variable at time zero) lies in the interval $[\underline{S}_T, \bar{S}_T]$, by solving[26]

$$r^* \in \arg\max_{r \in [0,1]} \mathbb{P}(\mathbb{E}_T[S_T] \in [\underline{S}_T, \bar{S}_T]) = \arg\max_{r \in [0,1]} \mathbb{P}(\hat{x}_T \in [\underline{x}_T, \bar{x}_T]),$$

where $\underline{x}_T \triangleq \kappa - (\bar{S}_T - \bar{S})$ and $\bar{x}_T \triangleq \kappa - (\underline{S}_T - \bar{S})$ are the corresponding thresholds in the state space. Note that conditional on the information at time $t = 0$, the distribution of the random variable $\hat{x}_T$ is Gaussian with mean $\mathbb{E}_0[\hat{x}_T]$ (computed using iterated expectations, and using the optimal state-feedback law in Proposition 10) and deterministic variance $\hat{V}_T$ (computed using the optimal variance-feedback law in Proposition 11). Fig. 5 shows the target-achievement probability as a function of $r \in (0, 1)$ for a given interval $[\underline{x}_T, \bar{x}_T]$ of time-T values of $\hat{x}_T$ with an interior maximum at $r^* \approx 0.38$. The solution of the climate-control problem (formulated as decision problem (P)) further yields that for values of the policy parameter r below $\underline{r} \approx 0.2438$, it is best to not acquire any information. With a subcritical system (when $r \geq \underline{\hat{r}} \approx 0.2440$), noise leveling is optimal. Specifically, for $r = 0.4$, one obtains a stationary information-acquisition policy, equivalent to using about $m =$ 1000 uncorrelated measurement stations ($\upsilon^* \approx 31.83$), suggesting a worldwide annual information-acquisition cost of about \$5 billion.[27] In terms of taxation, the model finds an optimal carbon tax of slightly exceeding \$1000/tCO_2 ($u^* \approx 35.75$). While this is significantly higher than extant predictions of actual year-2040 carbon prices of about \$140/tCO_2 (IEA, 2015, p. 33), it also indicates an upper bound for what reasonable carbon prices might eventually look like, in the absence of political considerations. At such elevated carbon prices, the attainment of current IPCC thresholds poses no difficulty. By virtue of the separation principle (see Section 1) a constrained solution, with limits on the emissions price, can be obtained by capping the system-stabilization control, without any effect on the optimal information-acquisition policy.

5. Conclusion

The closed-form solution to the decision problem (P) in Proposition 9 was obtained by considering an equivalent reduced-form decision problem (P′′) with only half as many parameters. In line with the separation principle of stochastic control, the solution to the system-stabilization and information-acquisition subproblems (SS) and (IA) can be obtained separately. The optimal system-stabilization policy satisfies the certainty-equivalence principle, and it consists of an affine state feedback. The solution to the highly nonlinear information-acquisition problem is a threshold policy, the nature of which is completely characterized by the value of the discriminant Δ (see Section 3.2.5 or Proposition 9). For $\Delta \geq 0$, it is optimal to not acquire any information, i.e., the variance threshold is infinite. For $\Delta < 0$, information acquisition is optimal. In particular, for contracting systems (which have a natural steady

[25] In response to significant outside shocks to the state, e.g., the recent forest fires in Canada, which can be in the order of several $GtCO_2$ (Lamberty, Peckham, Ahl, & Gower, 2007), it is possible to resolve the decision problem with an updated (shifted) belief distribution (amounting to a translated value of the initial state).

[26] For optimization purposes, we allow for values in the compact closure [0, 1] of the domain of r.

[27] Given a world population of about 7 billion, the cost for determining the carbon concentration would therefore amount to about one measurement station for every 7 million people.

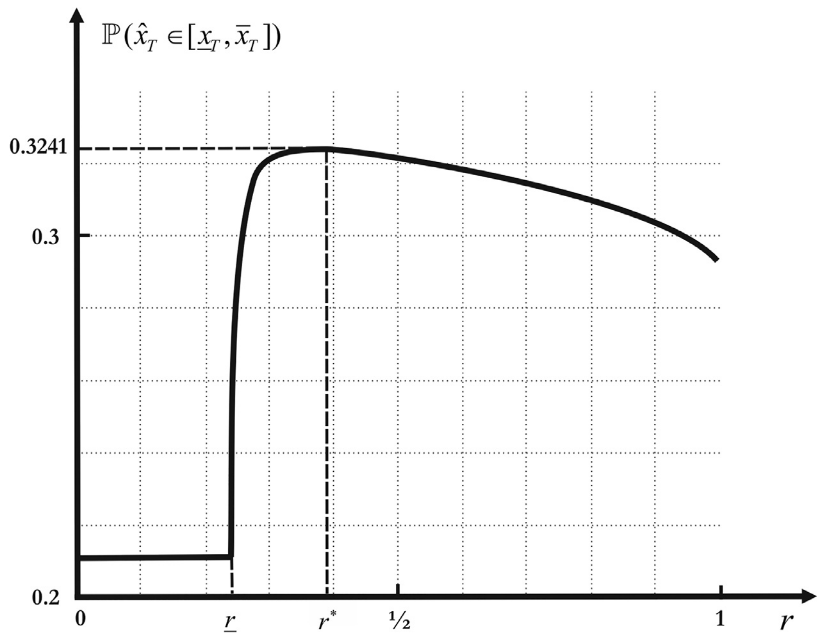

Fig. 5. Target-achievement probability for $[\underline{x}_T, \bar{x}_T] = [1172, 1180]$ $GtCO_2$ as a function of r.

state), when Δ is sufficiently small, the decision maker should pursue "noise leveling," that is, to acquire a large amount of information so as to decrease uncertainty to an optimal variance level. The latter is then held constant by acquiring only small amounts of information in subsequent periods. For intermediate (negative) values of Δ ("critical systems"), the variance steady-state lies below the target variance, so that beyond a one-time "noise attenuation" upon reaching the variance threshold, no further informational action by the decision maker is required. Lastly, for expanding systems and negative values of Δ, noise leveling is always optimal. Overall, the best information-acquisition policy is reminiscent of the (s, S)-threshold policy in inventory control: upon reaching a critical level of uncertainty, the decision maker orders enough information to reach and maintain (or autonomously improve upon) an optimal noise level. The particularity of this solution is that the variance threshold y^* (corresponding to the reordering point s) and the variance target α^* (corresponding to the order-up-to level S) are related by a one-period autonomous variance evolution (see Section 3.2.2), i.e., $y^{(1)}\big|_{y=\alpha^*} = y^*$. This produces the optimality of noise-leveling for all "subcritical" systems.

The decision maker can use the policy parameter q in (P), or equivalently r in (P′′) to "color" the intervention, without losing optimality. While the system-stabilization policy is invariant with respect to the weight on the variance in the objective function, the information-acquisition policy tends to increase in the policy parameter. As illustrated by the application of our results to the dynamic emissions-control problem in Section 4, the policy parameter can help the decision maker (e.g., a regulator) to pursue secondary objectives, such as maximizing the probability of being close to an intermediary target state. In a policy setting, the optimal information-acquisition policy may provide a justification for information acquisition (in terms of noise leveling or noise attenuation) or no information acquisition. It therefore limits the required amount of information from above and from below, thus possibly forestalling an "analysis paralysis" [28] caused by a perceived need to gather more and more information before starting to implement a costly stabilization policy. On the other hand, it can also provide the justification for substantial expenditures on information, especially when initial uncertainty is high.

[28] The phenomenon dates back at least to Aesop's fable of the "Fox and the Cat."

Appendix A. Proofs

Proof of Proposition 1. Consider an instant at time period $t \geq 0$, just after the decision maker has determined the controls u_t and v_t, but the realization of $\tilde{z}_{t+1}$ has not yet been observed. Conditioning on $(\mathcal{H}_t, u_t, v_t)$, one obtains the expected values of next period's state and system output:

$$\mathbb{E}[\tilde{x}_{t+1}|\mathcal{H}_t, u_t, v_t] = A\hat{x}_t + Bu_t,$$
$$\mathbb{E}[\tilde{z}_{t+1}|\mathcal{H}_t, u_t, v_t] = C\hat{x}_t.$$

As alluded to in Remark 1, consider now the deviation of the state from its current expected value, $\tilde{\varphi}_t = \tilde{x}_t - \hat{x}_t$, which follows a zero-mean normal distribution with variance $\hat{V}_t$. Correspondingly, the random next-period deviations, $\tilde{\xi}_{t+1}$ and $\tilde{\zeta}_{t+1}$, conditional on $\{\mathcal{H}_t, u_t, v_t\}$, follow the same law of motion as the original system in Eqs. (1) and (2),

$$\tilde{\xi}_{t+1} \triangleq \tilde{x}_{t+1} - \mathbb{E}[\tilde{x}_{t+1}|\mathcal{H}_t, u_t, v_t] = A\tilde{\varphi}_t + \tilde{\varepsilon}_{t+1},$$
$$\tilde{\zeta}_{t+1} \triangleq \tilde{z}_{t+1} - \mathbb{E}[\tilde{z}_{t+1}|\mathcal{H}_t, u_t, v_t] = C\tilde{\varphi}_t + (\tilde{\eta}_{t+1}/v_t).$$

The random vector $(\tilde{\xi}_{t+1}, \tilde{\zeta}_{t+1})$ follows therefore a joint two-dimensional normal distribution with zero mean and (conditional) covariance matrix

$$\Sigma = \begin{bmatrix} \bar{N} + A^2\hat{V}_t & AC\hat{V}_t \\ AC\hat{V}_t & C^2\hat{V}_t + M/v_t^2 \end{bmatrix}.$$

Hence, conditional on $\{\mathcal{H}_t, u_t, v_t, \zeta_{t+1}\}$, the time-$(t+1)$ state deviation $\tilde{\xi}_{t+1}$, is normally distributed with mean $\frac{AC\hat{V}_t\,\zeta_{t+1}}{C^2\hat{V}_t + M/v_t^2}$ and variance $\bar{N} + A^2(1 - \frac{C^2\hat{V}_t}{C^2\hat{V}_t + M/v_t^2})\hat{V}_t$. The time-$(t+1)$ update for the state and variance of the system, conditional on $(\mathcal{H}_t, u_t, v_t, z_{t+1})$, is therefore

$$\hat{x}_{t+1} = A\hat{x}_t + Bu_t + \frac{ACv_t^2\hat{V}_t}{M + C^2v_t^2\hat{V}_t}\left(z_{t+1} - C\hat{x}_t\right),$$

$$\hat{V}_{t+1} = \bar{N} + A^2\left(1 - \frac{C^2v_t^2\hat{V}_t}{M + C^2v_t^2\hat{V}_t}\right)\hat{V}_t,$$

where we have used the fact that $\zeta_{t+1} = z_{t+1} - C\hat{x}_t$. □

Proof of Proposition 2. Using a positive scaling factor λ, consider the decision problem (P) in the transformed state variable $\hat{x}' = \frac{\hat{x}}{\lambda}$ and the transformed variance $\hat{V}' = \hat{V}/\lambda^2$, so

$$\hat{K}(\bar{x}_0, \bar{N}_0) = \frac{p_0}{1-\beta}$$

$$+\inf \quad \sum_{t=0}^{\infty} \beta^t \mathbb{E}\Big[p_2\lambda^2(\hat{x}'_t)^2 + p_1\lambda\hat{x}'_t + q\lambda^2\hat{V}'_t + \gamma u_t^2 + \delta v_t^2 \Big| \hat{\mathcal{H}}_t \Big],$$

$$\text{s.t.} \quad \hat{x}'_{t+1} = A\hat{x}'_t + \frac{B}{\lambda}u_t + \frac{\lambda A C v_t \hat{V}'_t}{\sqrt{M + \lambda^2 C^2 v_t^2 \hat{V}'_t}}\omega_t, \quad \hat{x}'_0 = \frac{\bar{x}_0}{\lambda},$$

$$\hat{V}'_{t+1} = \frac{N}{\lambda^2} + A^2\left(1 - \frac{\lambda^2 C^2 v_t^2 \hat{V}'_t}{M + \lambda^2 C^2 v_t^2 \hat{V}'_t}\right)\hat{V}'_t, \quad \hat{V}'_0 = \frac{\bar{N}_0}{\lambda^2}.$$

We normalize the cost parameters, dividing them by $(p_2+q)\lambda^2$, whence

$$\hat{K}(\bar{x}_0, \bar{N}_0) = \frac{p_0}{1-\beta} + (p_2+q)\lambda^2 \times \left(\inf \quad \sum_{t=0}^{\infty} \beta^t \mathbb{E}\Big[(1-r)(\hat{x}'_t)^2 + s\hat{x}'_t + r\hat{V}'_t + \gamma' u_t^2 + \delta' v_t^2 \Big| \hat{\mathcal{H}}_t \Big] \right),$$

$$\text{s.t.} \quad \hat{x}'_{t+1} = A\hat{x}'_t + \frac{B}{\lambda}u_t + \frac{A v_t \hat{V}'_t}{\sqrt{M' + v_t^2 \hat{V}'_t}}\omega_t, \quad \hat{x}'_0 = \frac{\bar{x}_0}{\lambda},$$

$$\hat{V}'_{t+1} = \frac{N}{\lambda^2} + A^2\left(1 - \frac{v_t^2 \hat{V}'_t}{M' + v_t^2 \hat{V}'_t}\right)\hat{V}'_t, \quad \hat{V}'_0 = \frac{\bar{N}_0}{\lambda^2},$$

with $r = \frac{q}{p_2+q}$, $s = \frac{p_1}{(p_2+q)\lambda}$, $\gamma' = \frac{\gamma}{(p_2+q)\lambda^2}$, $\delta' = \frac{\delta}{(p_2+q)\lambda^2}$, and $M' = \frac{M}{\lambda^2 C^2}$. Setting $(u'_t)^2 = \gamma' u_t^2$ and $(v'_t)^2 = \delta' v_t^2$ then yields

$$\hat{K}(\bar{x}_0, \bar{N}_0) = \frac{p_0}{1-\beta} + (p_2+q)\lambda^2 \times \left(\inf \quad \sum_{t=0}^{\infty} \beta^t \mathbb{E}\Big[(1-r)(\hat{x}'_t)^2 + s\hat{x}'_t + r\hat{V}'_t + (u'_t)^2 + (v'_t)^2 \Big| \hat{\mathcal{H}}_t \Big] \right),$$

$$\text{s.t.} \quad \hat{x}'_{t+1} = A\hat{x}'_t + \frac{B}{\lambda\sqrt{\gamma'}}u'_t + \frac{A v'_t \hat{V}'_t}{\sqrt{\delta' M' + (v'_t)^2 \hat{V}'_t}}\omega_t, \quad \hat{x}'_0 = \frac{\bar{x}_0}{\lambda}, \tag{35}$$

$$\hat{V}'_{t+1} = \frac{N}{\lambda^2} + A^2\left(1 - \frac{(v'_t)^2 \hat{V}'_t}{\delta' M' + (v'_t)^2 \hat{V}'_t}\right)\hat{V}'_t, \quad \hat{V}'_0 = \frac{\bar{N}_0}{\lambda^2}.$$

Thus, by choosing λ such that

$$\lambda^2 = \sqrt{\frac{\delta M}{(p_2+q)C^2}},$$

one obtains $\delta' M' = 1$; moreover, the transformed decision problem (35) becomes

$$\hat{K}(\bar{x}_0, \bar{N}_0) = \frac{p_0}{1-\beta} + (p_2+q)\lambda^2 \times \left(\inf \quad \sum_{t=0}^{\infty} \beta^t \mathbb{E}\Big[(1-r)(\hat{x}'_t)^2 + s\hat{x}'_t + r\hat{V}'_t + (u'_t)^2 + (v'_t)^2 \Big| \hat{\mathcal{H}}_t \Big] \right),$$

$$\text{s.t.} \quad \hat{x}'_{t+1} = A\hat{x}'_t + \frac{B}{\lambda\sqrt{\gamma'}}u'_t + \frac{A v'_t \hat{V}'_t}{\sqrt{1 + (v'_t)^2 \hat{V}'_t}}\omega_t, \quad \hat{x}'_0 = \frac{\bar{x}_0}{\lambda},$$

$$\hat{V}'_{t+1} = \frac{N}{\lambda^2} + A^2\left(1 - \frac{(v'_t)^2 \hat{V}'_t}{1 + (v'_t)^2 \hat{V}_t}\right)\hat{V}'_t, \quad \hat{V}'_0 = \frac{\bar{N}_0}{\lambda^2}.$$

The result is then obtained by carefully converting the problem parameters. □

Proof of Proposition 3. If $K_1(x) = Px^2 + Qx + R$, then by virtue of the first constraint in (P′′) it is

$$\mathbb{E}\Big[K_1(\tilde{x}')\Big| x, y\Big] = P\left[(ax + b'u)^2 + \frac{a^2 v^2 y^2}{1 + v^2 y}\right] + Q(ax + b'u) + R.$$

Moreover, the minimization problem with respect to (u, v) in the Bellman equation (P′′) decomposes into the two given Bellman equations, (SS) and (IA). It remains to be shown that the first of these admits a quadratic solution for K_1.[29] The first-order condition for the minimization on the right-hand side yields that the optimal control u^* must be a linear function of x, which in turn implies that the cost K_1 can be written as a polynomial with a degree of at most two, so $K_1(x) = Px^2 + Qx + R$ can be made consistent with the solution to the first Bellman equation as claimed, which concludes our proof. □

Proof of Proposition 4. Suppose that the optimal cost function K in problem (P′′) is separable as in Proposition 3, with $K_1(x) = Px^2 + Qx + R$, for all $x \in \mathcal{X}$. Then

$$\mathbb{E}[K(\tilde{x}', y')] = \mathbb{E}[K_1(\tilde{x}')] + K_2(y') = \mathbb{E}[P(\tilde{x}')^2 + Q\tilde{x}' + R] + K_2(y')$$
$$= P(ax + bu)^2 + Q(ax + bu) + R + P\frac{a^2 v^2 y^2}{1 + v^2 y} + K_2(y').$$

Substituting the last expression in the Bellman equation of (P′′) yields

$$K(x, y) = \min_{(u,v)} \Big\{ (1-r)x^2 + sx + ry + u^2 + v^2 + \beta\left(P\cdot(ax+bu)^2 + Q\cdot(ax+bu) + R + P\frac{a^2v^2y^2}{1+v^2y} + K_2(y')\right)\Big\},$$

$$\text{s.t.} \quad y' = N + a^2 y - \frac{a^2 v^2 y^2}{1 + v^2 y}.$$

[29] The solution of the Bellman equation is unique; see also footnote 8.

The problem is separable in the system control u and the information control v. Minimization with respect to u,

$$Px^2 + Qx + R = \min_u\{(1-r)x^2 + sx + u^2 + \beta P\cdot(ax+bu)^2 + \beta Q\cdot(ax+bu) + \beta R\},$$

yields the optimal system control,

$$u^* = -\frac{\beta(2Pabx + Qb)}{2(1+\beta Pb^2)}.$$

Using this control (which implements an affine state-feedback law), one therefore obtains the optimal cost K_1 as a quadratic function of x. The values of P, Q, R are obtained by comparing the corresponding coefficients for x^2, x^1, and x^0, so necessarily

$$P = 1 - r + \frac{\beta a^2 P}{1+\beta b^2 P},$$

$$Q = s + \frac{\beta a Q}{1+\beta b^2 P},$$

$$R = \beta R - \frac{\beta^2 b^2 P^2/4}{1+\beta b^2 P}.$$

The preceding system of nonlinear equations has a unique solution for (P, Q, R), which is provided in Eq. (5). Minimizing the separable cost with respect to the information control v gives

$$K_2(y) = \min_{v\ge 0}\left\{(ry+v^2) + \beta\left(P\frac{a^2v^2y^2}{1+v^2y} + K_2(y')\right)\right\},$$

$$\text{s.t. } y' = N + a^2y - \frac{a^2v^2y^2}{1+v^2y}.$$

The solution to the preceding problem depends solely on y, not on x. Hence, the optimal cost K in problem (P'') is indeed separable in x and y. □

Proof of Proposition 5. Let $\mathcal{A}$ be convex. We distinguish two cases, depending on whether the target variance α^* is infinite or finite.

Case 1: α^* is infinite. Since by Eq. (11) the target variance is $\alpha^* = \infty$, the convex cost-improvment set must be of the form $\mathcal{A} = (N, \infty]$. Moreover, the variance threshold is $y^* = \infty$. The corresponding (∞, ∞)-threshold policy in Eq. (12) is equivalent to the information-acquisition policy $\nu(y)\equiv 0$, i.e., no information acquisition for any prior variance $y \in \mathcal{Y}$. This policy is optimal, because by the definition of the cost-improvement set $\mathcal{A}$, the cost in the information-acquisition problem (IA') is strictly decreasing in the posterior variance y' in (N, ∞). Thus, for any finite prior variance $y \in \mathcal{Y}$, the optimal solution to (IA') is $\nu(y) = 0$.

Case 2: α^* is finite. Given that $\alpha^* < \infty$, by the convexity of $\mathcal{A}$, it is necessarily $\mathcal{A} = (N, \alpha^*)$. As in Case 1, the optimum of the information-acquisition problem (IA') is attained at the posterior variance $y' = N + a^2y$ if $y < y^* = (\alpha^* - N)/a^2 = y^*$, and at the posterior variance $y' = \alpha^*$ if $y \ge y^*$. Therefore the information-acquisition policy $\nu(y)$, defined by the (y^*, α^*)-threshold policy in Eq. (12) must be optimal for all $y \in \mathcal{Y}$.

Both cases together establish the optimality of the threshold policy. □

Proof of Lemma 1. Without information acquisition, using the recursion $y' = N + a^2y$, the variance after $n \ge 1$ periods starting from y becomes, using the geometric-series formula,

$$y^{(n)} = N\sum_{k=0}^{n-1} a^{2k} + a^{2n}y = N\frac{1-a^{2n}}{1-a^2} + a^{2n}y = \bar{y} + a^{2n}(y-\bar{y}), \quad (36)$$

provided that $a \ne 1$. For $a = 1$, it is $y^{(n)} = N + ny$. (i) For $a < 1$, $y^{(2n)} - \bar{y} = a^{2n}(y-\bar{y}) \to 0$ as $n \to \infty$, which implies global asymptotic stability for all $y \in \mathcal{Y}$. More specifically, since $|y^{(n)} - \bar{y}| \le |y-\bar{y}|\cdot\exp[-n\ln(1/a^2)]$, for all $n \ge 1$, the convergence is exponential with rate $\ln(1/a^2)$; the convergence is also monotonic. (ii) For $a \ge 1$, it is $y' - y \ge N > 0$, so the variance increases monotonically beyond any given bound $\bar{B} > 0$ in at most $\lceil \bar{B}/N \rceil$ periods, i.e., $y^{(n)} \uparrow \infty$ as $n \to \infty$. For $a = 1$, the increase is linear. For $a > 0$, the increase is by Eq. (36) exponential,

$$y^{(n)} = \left(y + \frac{N}{a^2-1}\right)\exp[n\ln a^2] - \frac{N}{a^2-1}, \quad n > 0,$$

at the rate $\ln(a^2)$, which completes our proof. □

Proof of Lemma 2. Without information collection at the prior variance y, i.e., when $\nu(y) = 0$, the optimal posterior variance in the information-acquisition problem (IA') is $y' = N + a^2y$, and the Bellman equation takes on the form

$$K_2(y) = ry + \beta K_2(N + a^2y). \quad (37)$$

Using the affine ansatz $\bar{K}_2(y) \equiv \bar{k}_0 + \bar{k}_1 y$, for suitable constants $\bar{k}_0, \bar{k}_1$, Eq. (37) becomes

$$\bar{k}_0 + \bar{k}_1 y = ry + \beta\left(\bar{k}_0 + \bar{k}_1(N + a^2y)\right) = \beta(\bar{k}_0 + \bar{k}_1 N) + (r + \bar{k}_1\beta a^2)y.$$

This relation needs to hold in a (either one-sided or two-sided) neighborhood of y, which implies—by comparing the coefficients for the different powers of y—that

$$\bar{k}_0 = \frac{r}{1-\beta a^2}\cdot\frac{\beta N}{1-\beta} \quad \text{and} \quad \bar{k}_1 = \frac{r}{1-\beta a^2},$$

thus establishing Eq. (15), completing the proof. □

Proof of Lemma 3. Provided there is a variance threshold y^* such that information acquisition is optimal for $y \ge y^*$, integration of the expression for the gradient of K_2 in (16), from y^* to y, yields

$$K_2(y) = K_2(y^*) + \left(r + \beta Pa^2\right)(y - y^*) + \left(\frac{1}{y^*} - \frac{1}{y}\right) = k_0 + k_1 y - \frac{1}{y}, \quad y \ge y^*,$$

where

$$k_0 \triangleq K_2(y^*) - \left(r + \beta Pa^2\right)y^* + \frac{1}{y^*} \quad \text{and} \quad k_1 \triangleq r + \beta Pa^2 \quad (38)$$

are constants. Since necessarily $K_2(y) \ge 0$ for all $y \ge y^*$, the constant k_0 must be nonnegative. The constant k_1 is nonnegative because the reduced parameter vector $\theta' \ge 0$ by assumption. □

Proof of Lemma 4. For any given y, the Lagrangian for the constrained minimization in (IA') is [30]

$$\mathcal{L}(y', \lambda; y) = ry - \frac{1}{y} + \beta P(N + a^2y) + \frac{a^2}{y'-N} + \beta\left[K_2(y') - Py'\right] - \lambda(N + a^2y - y'),$$

where $\lambda \ge 0$ denotes the shadow price associated with the constraint that y' cannot exceed the posterior variance $N + a^2y$, which is attained without information acquisition. Because the minimand diverges to $+\infty$ when y' tends to N (from the right), $y' = N$ can never be optimal, so that $y' > N$ is automatically satisfied. The relevant first-order necessary optimality condition is

$$-\frac{a^2}{(y'-N)^2} + \beta\left[K_2'(y') - P\right] + \lambda = 0, \quad (39)$$

[30] The minimization on the right-hand side of (IA') is with respect to the posterior variance y'. The extra terms relating to the given prior variance y are included to be able to apply the envelope theorem to the entire right-hand side of (IA'), as will become clear below.

with complementary-slackness condition

$$\lambda(N + a^2 y - y') = 0. \tag{40}$$

Let $(y', \lambda) = (\nu(y), \ell(y))$ be a solution to the system (39) and (40). Applying the envelope theorem to the right-hand side of (IA′) gives

$$K_2'(y) = \left.\frac{\partial \mathcal{L}(y', \lambda; y)}{\partial y}\right|_{(y',\lambda)=(\nu(y),\ell(y))} = r + \frac{1}{y^2} + \beta P a^2 - a^2 \ell(y).$$

Substituting the last relation into Eq. (39) then yields

$$-\frac{a^2}{(y'-N)^2} + \beta\left[(y')^{-2} - a^2\ell(y')\right] - \Delta + \lambda = 0. \tag{41}$$

where $\Delta = \beta(P(1-\beta a^2) - r)$, as defined in the text. The left-hand side of the last equation is increasing in y' if the corresponding gradient is positive, i.e.,

$$2\left(\frac{a^2}{(y'-N)^3} - \frac{\beta}{(y')^3}\right) - \beta a^2 \ell'(y') > 0. \tag{42}$$

Since $\ell'(y') \le 0$ almost everywhere (provided the remaining terms are positive; see Remark 9 below), the last inequality holds as long as

$$\frac{a^2}{\beta} > \left(1 - \frac{N}{y'}\right)^3,$$

which—in turn—is true for all $y' \in (N, N + a^2 y]$ if and only if

$$\frac{a^2}{\beta} > \left(1 - \frac{N}{N + a^2 y}\right)^3 = \left(\frac{a^2 y}{N + a^2 y}\right)^3.$$

The term on the right-hand side of the last equation is increasing in y and tends to 1 as $y \to \infty$, so the inequality holds for all $y \in \mathcal{Y}$ if and only if

$$a^2 > \beta.$$

Finally, the preceding inequality holds for all $\beta \in (0, 1)$ if and only if

$$a \ge 1.$$

By the definition of the y-conditional cost-improvement set $\mathcal{A}_y$ in Eq. (9), the last condition therefore guarantees that for all $\hat{y}' \in \mathcal{A}_y$:

$$N < y' < \hat{y}' \quad \Rightarrow \quad y' \in \mathcal{A}_y.$$

This implies that $\mathcal{A}_y = (N, \alpha(y))$, where $\alpha(y)$ is the y-conditional target variance in Eq. (8). By taking the limit for $y \to \infty$, one therefore obtains that the unconditional cost-improvement set is convex and of the form $\mathcal{A} = (N, \alpha^*)$, which completes the proof. □

Remark 9. The following monotonicity property of a convex optimization problem is used in the proof of Lemma 4: *Let $\phi : \mathbb{R} \to \mathbb{R}$ be a twice differentiable, strictly convex function and $\alpha \in \mathbb{R}$ a given constant. Then the problem*

$$\min_{\xi \in (-\infty, \alpha]} \phi(\xi)$$

has a nonnegative Lagrange multiplier, $\lambda(\alpha) = \max\{0, -\phi'(\alpha)\}$, which is nonincreasing in α. The proof of this auxiliary result follows directly from the Kuhn-Tucker conditions in the standard Lagrangian framework.

Proof of Proposition 6. (i) Using the Lagrangian framework established in the Proof of Lemma 4, consider $(y', \lambda) = (\nu(y), \ell(y))$ as a solution to the necessary optimality conditions in Eqs. (40) and (41), where necessarily $\ell(y) \ge 0$. If $\ell(y) > 0$, then the complementary-slackness condition (40) requires that $y' = N + a^2 y$, which implies $\nu(y) = 0$, i.e., no information collection is undertaken. If $\ell(y) = 0$, then the optimal solution $\alpha_1^* = y'$ does not depend on y and satisfies

$$\begin{cases} \dfrac{\beta}{(\alpha_1^*)^2} - \dfrac{a^2}{(\alpha_1^* - N)^2} = \Delta + \beta a^2 \ell(\alpha_1^*), \\ \alpha_1^* \le N + a^2 y. \end{cases} \tag{43}$$

For $a \ge 1$, Lemma 4 together with Proposition 5 implies that α_1^*, as unique solution to Eq. (18), is an optimal target variance. The left-hand side of Eq. (18) is monotonic (by Lemma 4), and for $y' \to \infty$:

$$\frac{\beta}{(y')^2} - \frac{a^2}{(y'-N)^2} \uparrow 0.$$

As a result, the condition $\Delta < 0$ is necessary and sufficient to ensure that target variance α_1^* is *finite*, and the (y_1^*, α_1^*)-threshold policy is by Proposition 5 optimal, where $y_1^* = (\alpha_1^* - N)/a^2$ is the optimal threshold. Note that because $a \ge 1$, it is $\alpha_1^* > y_1^*$, so $\ell(\alpha_1^*) = 0$ holds necessarily. (ii) In part (i), it was established that $\Delta < 0$ is a necessary and sufficient condition for $\alpha_1^* < \infty$. Thus, for $\Delta \ge 0$, both the target variance α_1^* and the associated variance threshold y_1^* become infinite, so the (∞, ∞)-threshold policy is optimal: with $\nu(y) \equiv 0$, no information acquisition takes place. □

Proof of Corollary 1. By Proposition 6 the (y_1^*, α_1^*)-threshold policy is optimal for the information-acquisition problem (IA′). Thus, for $y \ge y_1^*$, the optimal cost is of the form given by Lemma 3 (and its proof), so

$$K_2(y) = k_0 + \left(r + \beta P a^2\right) y - \frac{1}{y}, \tag{44}$$

where k_0 is a nonnegative constant. Since $\alpha_1^* \ge y_1^*$, the cost $K_2(\alpha_1^*)$ is of the same form,

$$K_2(\alpha_1^*) = k_0 + \left(r + \beta P a^2\right)\alpha_1^* - \frac{1}{\alpha_1^*}. \tag{45}$$

On the other hand, the target variance also solves the Bellman equation in (IA′), whence

$$K_2(\alpha_1^*) = \beta P N + \left(r + \beta P a^2\right)\alpha_1^* - \frac{1}{\alpha_1^*} + \frac{a^2}{\alpha_1^* - N} - \beta P \alpha_1^* + \beta K_2(\alpha_1^*). \tag{46}$$

Eqs. (45) and (46) together yield that

$$k_0 = \frac{1}{1-\beta}\left(\beta P N - \Delta \alpha_1^* + \frac{a^2}{\alpha_1^* - N} - \frac{\beta}{\alpha_1^*}\right),$$

which in conjunction with Eq. (44) establishes the optimal information-acquisition cost in Eq. (19), for all $y \ge y_1^*$.

For $y < y_1^*$, the decision maker does not collect information, and the solution to the minimization problem in (IA′) is $y' = N + a^2 y$. The optimal information-acquisition cost therefore simplifies to

$$K_2(y) = r y + \beta K_2(N + a^2 y). \tag{47}$$

By Eq. (13) in Section 3.2.2, after n periods without information collection the variance y autonomously evolves to $y^{(n)} > y$, and $y^{(n)} \uparrow \infty$ as $n \to \infty$ by virtue of Lemma 1. Thus, the minimum number of periods required to exceed the variance threshold is finite (see the proof of Lemma 1 for details),

$$n^* \triangleq \inf\{n > 0 : y^{(n)} > y_1^*\} \le \left\lceil \frac{y_1^*}{N} \right\rceil.$$

By a straightforward recursion, the optimal information-acquisition cost therefore takes on the form in Eq. (20), which completes our proof. □

Proof of Corollary 2. This result follows from Proposition 6 (i), since for $a \ge 1/\sqrt{\beta}$, the discriminant $\Delta < 0$. □

Proof of Lemma 5. If $\beta < a^2 < 1$, then $\mathcal{A}$ (and therefore in particular also $A_{\bar{y}}$) is convex, by the proof of Lemma 4. To obtain the result for the case where $a^2 < \beta$, one can use an approach analogous to the proof of Lemma 4. The left-hand side of Eq. (42),

$$2\left(\frac{a^2}{(y'-N)^3} - \frac{\beta}{(y')^3}\right) - \beta a^2 \ell'(y') > 0,$$

given that ℓ' is nonpositive as before (by Remark 9), is positive if

$$y' < N/(1-\chi),$$

where $\chi \triangleq (a^2/\beta)^{1/3}$, provided that $\chi < 1$. Hence, for $a^2 < \beta$, the y-conditional cost-improvement set $\mathcal{A}_y$ is convex for all

$$N + a^2 y < \frac{N}{1-a^2} < \frac{N}{1-\chi},$$

since $a^2 < \chi$. The preceding inequality is equivalent to $y < \bar{y}$, which completes the proof. □

Proof of Proposition 7. By Lemma 5, the cost-improvement set $\mathcal{A}_{\bar{y}}$ is convex, where $\bar{y} = N/(1-a^2) < \infty$ is by virtue of Lemma 1 the exponentially stable steady state for the variance, in the absence of information acquisition. The result now follows in a way that is analogous to the proof of Proposition 6, provided the discriminant Δ is subcritical, to ensure that $\alpha_2^* \in \mathcal{A}_{\bar{y}}$:

$$\hat{\Delta} \triangleq \frac{\beta}{\bar{y}^2} - \frac{a^2}{(\bar{y}-N)^2} \geq \Delta.$$

Using the definition of $\bar{y}$, one obtains that $\hat{\Delta} = -(1-\beta a^2)/(a\bar{y})^2$ (<0). □

Proof of Corollary 3. The proof of this result is analogous to the proof of Corollary 1. □

Proof of Proposition 8. Let $a \in (0, 1)$. By the construction in the main text, for critical systems with $\hat{\Delta} \leq \Delta < 0$, the (finite, unconditional) target variance must exceed the finite steady state $\bar{y}$ at least weakly. Therefore, once the target variance $y' = \alpha_3^*$ has been attained, there cannot be any further information acquisition. By Lemma 2 the optimal information-acquisition cost for variances after attaining the target variance must be affine, so the information-acquisition problem (IA′) becomes a Bellman equation of the form

$$K_2(y) = \beta PN + (r + \beta PN)y - \frac{1}{y} + \min_{y' \in (N, N+a^2 y]} \left\{ \frac{a^2}{y'-N} - \beta P y' + \frac{\beta r}{1-\beta a^2}\left(\frac{\beta N}{1-\beta} + y\right) \right\}. \tag{48}$$

Provided that $y \geq y_3^*$, the optimal variance target α_3^* satisfies the first-order necessary optimality condition,

$$-\frac{a^2}{(\alpha_3^* - N)^2} = \frac{\Delta}{1-\beta a^2}. \tag{49}$$

The unique solution of Eq. (49) on the interval (N, ∞) is equal to the target variance α_3^* given by Eq. (23). Correspondingly one obtains the variance threshold $y_3^* = (\alpha_3^* - N)/a^2 = \sqrt{(1-\beta a^2)/(-\Delta a^2)}$. Since the cost-improvement set $\mathcal{A} = (N, \alpha_3^*)$ is convex, the optimality of the (y_3^*, α_3^*)-threshold policy follows by virtue of Proposition 5, which concludes our proof. □

Proof of Corollary 4. For $y \geq y_3^*$, the optimal information-acquisition cost $K_2(y)$ follows from Eq. (48) in the proof of Proposition 8 by substituting the expression of the optimal target variance α_3^* from Eq. (23). For $y < y_3^*$, because there is no further information acquisition, the optimal cost $K_2(y)$ is given by Eq. (15) in Lemma 2. Thus, one obtains Eq. (24) for $y \geq y_3^*$ and Eq. (25) for $y < y_3^*$, respectively. □

Proof of Proposition 9. This result follows from Proposition 1 (for part (i)), and from Props. 6–8 (for part (ii)). The fact that $\Delta \geq 0$ implies that $\nu(y) \equiv 0$ is optimal follows immediately from the discussion in Section 3.2.1 because then the cost-improvement set $\mathcal{A} = (N, \infty)$, so the target variance is infinite, which is equivalent to no information collection. □

Proof of Proposition 10. Since $Pb^2 = P_0 b_0^2$ (independent of r) and $b = b_0(1-r)^{1/2}$, Proposition 4 implies

$$\mu(x) - \mu(0) = -\frac{P_0 b_0 \, ax}{P_0 b_0^2 + (1/\beta)} (1-r)^{1/2} = (\mu_0(x) - \mu_0(0))(1-r)^{1/2},$$

where $\mu_0(x) \triangleq \lim_{r \to 0^+} \mu(x)$ for all $x \in \mathcal{X}$ and $\mu_0(0) = \mu(0)$.[31] By Eq. (26) the optimal state-feedback law for the decision problem (P) is of the form

$$\hat{\mu}(\hat{x}) = \frac{\mu(\hat{x}/\lambda)}{\sqrt{\gamma'}} = \frac{\mu(\hat{x}/\lambda)}{\sqrt{\gamma_0'}}(1-r)^{-1/4},$$

for all $\hat{x} \in \mathcal{X}$, where $\gamma_0' \triangleq \gamma'|_{r \to 0^+} = \gamma/(p_2 \lambda_0^2)$. Hence,

$$\hat{\mu}(\hat{x}) - \hat{\mu}(0) = \frac{\mu_0(\hat{x}/\lambda) - \mu_0(0)}{\sqrt{\gamma_0'}}(1-r)^{1/4}.$$

Differentiation with respect to r yields

$$\begin{aligned}\frac{d(\hat{\mu}(\hat{x}) - \hat{\mu}(0))}{dr} &= -\frac{(1-r)^{-3/4}}{4\sqrt{\gamma_0'}}(\mu_0(\hat{x}/\lambda) - \mu_0(0)) + \frac{(1-r)^{1/4}}{4\sqrt{\gamma_0'}} \frac{\mu_0'(\hat{x}/\lambda)}{\lambda_0}(1-r)^{-5/4} \\ &= \frac{\mu_0(\hat{x}/\lambda) - \mu_0(0)}{4\sqrt{\gamma_0'}}\left[\frac{(\hat{x}/\lambda)\mu_0'(\hat{x}/\lambda)}{\mu_0(\hat{x}/\lambda) - \mu_0(0)} - 1\right] \times (1-r)^{-3/4}.\end{aligned}$$

Since $x\mu_0'(x)/(\mu_0(x) - \mu_0(0)) \equiv 1$, we therefore obtain that

$$\frac{d(\hat{\mu}(\hat{x}) - \hat{\mu}(0))}{dr} = 0,$$

for all $\hat{x} \in \mathcal{X}$. On the other hand,

$$\hat{\mu}(0) = \frac{\mu(0)}{\sqrt{\gamma'}} = -\frac{b_0 Q_0}{2\sqrt{\gamma_0'}(P_0 b_0^2 + 1/\beta)} = \mu_0(0),$$

independent of r. This implies that the feedback law $\hat{\mu}(\hat{x})$ is independent of r (and independent of q):

$$\hat{\mu}(\hat{x}) = -\frac{(b_0/\sqrt{\gamma_0'})}{P_0 b_0^2 + 1/\beta}\left(P_0 a\left(\frac{\hat{x}}{\lambda_0}\right) + \frac{Q_0}{2}\right),$$

for all $\hat{x} \in \mathcal{X}$. □

Proof of Corollary 5. As pointed out in the main text, $(P, Q, R) = ((1-r)P_0, Q_0(1-r)^{3/4}, R_0(1-r)^{1/2})$, so $\hat{K}_1(\hat{x}) = (p_2 + q)\lambda^2 K_1(\hat{x}/\lambda) = \left(\frac{p_2\lambda^2}{1-r}\right)K_1(\hat{x}/\lambda)$, and by Proposition 4 therefore

$$\hat{K}_1(\hat{x}) = \left(\frac{p_2 \lambda^2}{1-r}\right)K_1(\hat{x}/\lambda) = p_2 \lambda_0^2 K_{10}(\hat{x}/\lambda_0) = p_2\left[P_0 \hat{x}^2 + \lambda_0 Q_0 \hat{x} + \lambda_0^2 R_0\right],$$

[31] In particular, $\lim_{r \to 1^-} \mu(x) = \mu(0)$, for all $x \in \mathcal{X}$. That is, if the decision maker cares only about uncertainty reduction, not about target achievement, then the feedback gain in the reduced-form problem (P′′) is constant.

where we have set $K_{10}(x) \triangleq \lim_{r\to 0^+} K_1(x) = P_0x^2 + Q_0x + R_0$, for all $x \in \mathcal{X}$. Note also that

$$\frac{d\hat{K}_1(\hat{x})}{dr} = 0,$$

for all $\hat{x} \in \mathcal{X}$. That is, the optimal system-stabilization cost is invariant with respect to r. □

Proof of Proposition 11. By Eqs. (12) and (26) the optimal variance-feedback law for the decision problem (P) is

$$\hat{\nu}(\hat{V}) = \frac{\nu(\hat{V}/\lambda^2)}{\sqrt{\delta'}} = \frac{\nu(\hat{V}/\lambda^2)}{\sqrt{\delta'_0}}(1-r)^{-1/4} = \frac{\lambda_0}{\sqrt{\delta'_0}}\sqrt{\left[\frac{1}{\lambda^2 y^*} - \frac{1}{\hat{V}}\right]_+}$$
$$= \lambda_0^2\sqrt{\frac{p_2}{\delta}\left[\frac{1}{\hat{V}^*} - \frac{1}{\hat{V}}\right]_+},$$

where $\delta'_0 \triangleq \delta'|_{r\to 0^+} = \delta/(p_2\lambda_0^2)$ and $\hat{V}^* \triangleq \lambda^2 y^*$. The (unreduced) variance threshold $\hat{V}^*$ depends on r just as the reduced variance threshold y^* depends on r, so $\hat{V}^* = \hat{V}^*(r)$ and $y^* = y^*(r)$. Consider now the discriminant $\Delta = \beta((1-\beta a^2)P - r) = (1-r)\Delta_0 - r\beta$, where $\Delta_0 \triangleq \Delta|_{r\to 0^+} = (1-\beta a^2)\beta P_0$, so that $d\Delta/dr = -\Delta_0 - \beta$. Since

$$\lim_{r\to 1^-} \Delta = -\beta < 0,$$

there is always information acquisition, provided that r (or q) is large enough. It is $\Delta < 0$ if and only if

$$r > \frac{\Delta_0}{\beta + |\Delta_0|} = \frac{(1-\beta a^2)P_0}{1 + |1-\beta a^2|P_0} = \underline{r}.$$

The last inequality holds in the interesting case where $\beta a^2 < 1$. It fully characterizes information acquisition for expanding systems ($a \geq 1$) and leads to noise leveling. For contracting systems ($a < 1$), noise attenuation may be optimal when the last inequality holds. Note first that $\hat{\Delta} = (1-r)\hat{\Delta}_0$, where $\hat{\Delta}_0 \triangleq \hat{\Delta}|_{r\to 0^+} = -(1-\beta a^2)/(a^2\bar{y}_0^2)$. Thus, $\hat{\Delta} \leq \Delta$ is equivalent to $(1-r)\hat{\Delta}_0 \leq (1-r)\Delta_0 - r\beta$, which in turn means that there is noise attenuation if and only if $\underline{r} < r \leq \hat{\underline{r}}$, where

$$\hat{\underline{r}} = \frac{\Delta_0 + (-\hat{\Delta}_0)}{\beta + |\Delta_0 + (-\hat{\Delta}_0)|} = \frac{(1-\beta a^2)(P_0 + (a\bar{y}_0)^{-2})}{1 + |1-\beta a^2|(P_0 + (a\bar{y}_0)^{-2})}.$$

On the other hand, noise leveling is optimal for contracting systems if $\hat{\underline{r}} < r < 1$. □

Proof of Proposition 12. Differentiation of the threshold information-acquisition policy in Eq. (29) with respect to the policy parameter r yields

$$\frac{d\hat{\nu}(\hat{V})}{dr} = -\frac{\lambda_0^2}{2(\hat{V}^*)^2}\sqrt{\frac{p_2}{\delta}}\left(\left[\frac{1}{\hat{V}^*(r)} - \frac{1}{\hat{V}}\right]_+\right)^{-1/2}\frac{d\hat{V}^*(r)}{dr}$$
$$= -\frac{a^2\sqrt{p_2/\delta}}{2(\hat{\alpha}^*(r) - n_0)^2}\left(\left[\frac{1}{\hat{V}^*(r)} - \frac{1}{\hat{V}}\right]_+\right)^{-1/2}\frac{d\hat{\alpha}^*(r)}{dr}, \quad (50)$$

where $\hat{\alpha}^*(r) \triangleq (1-r)^{1/2}\alpha^*(r)$ and $\alpha^*(r) = N + a^2y^*(r)$.

Consider first noise-leveling policies. With the substitution $\hat{y} \triangleq (1-r)^{1/2}y'$ the cost-improvement set, transformed to the $\hat{y}$-domain, becomes

$$\hat{\mathcal{A}}(r) = \left\{\hat{y} > n_0 : \frac{\beta}{\hat{y}^2} - \frac{a^2}{(\hat{y}-n_0)^2} < \Delta_0 - \frac{r\beta}{1-r}\right\}.$$

Since $0 < r < \hat{r} < 1$ implies that $\hat{\mathcal{A}}(\hat{r}) \subseteq \hat{\mathcal{A}}(r)$, the transformed variance target $\hat{\alpha}^*(r) = \sup \hat{\mathcal{A}}(r)$ is necessarily (weakly) decreasing in the policy parameter r. Thus, $\hat{\alpha}^*(r) = (1-r)^{1/2}\alpha^*(r)$ is decreasing in r, which by Eq. (50) implies that $\hat{\nu}(\hat{V})$ is decreasing in r.

Consider now noise-attenuation policies, relevant for contracting systems ($a < 1$), where—by construction—the target variance is

$$\alpha^*(r) = \frac{n_0}{(1-r)^{1/2}} + \frac{a\sqrt{1-\beta a^2}}{(-\Delta)^{1/2}}.$$

Thus,

$$\frac{d\hat{\alpha}^*(\hat{V})}{dr} = \frac{d((1-r)^{1/2}\alpha^*(r))}{dr}$$
$$= -\frac{\beta a\sqrt{1-\beta a^2}}{2(1-r)^2}\left(-\Delta_0 + \frac{\beta r}{1-r}\right)^{-3/2}$$
$$= -\frac{\beta a\sqrt{1-\beta a^2}}{2(1-r)^{1/2}}(-\Delta)^{-3/2} < 0,$$

so that, by virtue of Eq. (50), the information-acquisition policy $\hat{\nu}(\hat{V})$ must be decreasing in r. □

Proof of Corollary 6. For expanding systems ($a \geq 1$), by Propositions 11 and 12 the optimal variance feedback is largest for $r \to \max\{0, \underline{r}\}^+$. For contracting systems ($a < 1$), by Propositions 11 and 12 the optimal variance feedback is largest for $r = \hat{\underline{r}}$ ($> \max\{0, \underline{r}\}$). □

Proof of Lemma 6. The derivative of the optimal information-acquisition cost with respect to r is[32]

$$\frac{d\hat{K}_2(\hat{V})}{dr} = \frac{d}{dr}\left[\left(\frac{p_2\lambda^2}{1-r}\right)K_2(\hat{V}/\lambda^2)\right]$$
$$= \left(\frac{p_2\lambda^2}{2(1-r)^2}\right)K_2(\hat{V}/\lambda^2) + \left(\frac{p_2\lambda^2}{1-r}\right)\left(K'_2(\hat{V}/\lambda^2)\left(\frac{\hat{V}/\lambda^2}{2(1-r)}\right) + \left.\frac{\partial K_2(y)}{\partial r}\right|_{y=\hat{V}/\lambda^2}\right)$$
$$= \left(\frac{p_2\lambda^2}{1-r}\right)\left(\frac{K_2(\hat{V}/\lambda^2)}{2(1-r)}\left[\frac{(\hat{V}/\lambda^2)K'_2(\hat{V}/\lambda^2)}{K_2(\hat{V}/\lambda^2)} + 1\right] + \left.\frac{\partial K_2(y)}{\partial r}\right|_{y=\hat{V}/\lambda^2}\right).$$

Hence, the behavior of the information-acquisition cost with respect to the policy parameter r depends on its elasticity, $\epsilon_2(y) \triangleq yK'_2(y)/K_2(y)$, with respect to the variance. □

Solution to the quartic equation

If $\mathbf{1}_{\{a\geq 1\}}\Delta < 0$ or $\Delta - \hat{\Delta} < 0$, i.e., if $\min\{\mathbf{1}_{\{a\geq 1\}}\Delta, \Delta - \hat{\Delta}\} < 0$, then the optimal target variance is [33]

$$\alpha^* = \frac{N + \sqrt{L_1} + \sqrt{L_2 - L_3}}{2} \in \left\{y' > N : \frac{\beta}{(y')^2} - \frac{a^2}{(y'-N)^2} = \Delta\right\}, \quad (51)$$

[32] As before, we use the fact that $\lambda^2 = \lambda_0^2(1-r)^{1/2}$, where $\lambda_0^2 = \sqrt{\delta M/p_2}/C$ is independent of r.

[33] The expression singles out a root from the solutions of (a cubic resolvent for) the quartic equation using Cardan's formula. All four (generally complex-valued) roots are given by $(N + \sqrt{L_1} \pm \sqrt{L_2 - L_3})/2$ and $(N - \sqrt{L_1} \pm \sqrt{L_2 + L_3})/2$.

where

$$L_1 \triangleq \frac{N^2}{3} + \frac{2(\beta - a^2) + L_4^{1/3} + (\beta - a^2 - \Delta N^2)^2 L_4^{-1/3}}{3\Delta},$$

$$L_2 \triangleq \frac{2N^2}{3} + \frac{4(\beta - a^2) - L_4^{1/3} - (\beta - a^2 - \Delta N^2)^2 L_4^{-1/3}}{3\Delta},$$

$$L_3 \triangleq \frac{2N(\beta + a^2)}{\Delta\sqrt{L_1}},$$

$$L_4 \triangleq L_5 + 27\Delta N^2 a^2 \beta - 6\sqrt{3\beta\Delta L_5} Na,$$

$$L_5 \triangleq \Delta^3 N^6 + 3\,(\beta - a^2)(\beta a^2 - \Delta^2 N^4) + 3\,(a^4 + 7\beta a^2 + \beta^2)\,\Delta N^2 - (\beta^3 - a^6).$$

Appendix B. Notation

Symbol	*Description*	*Domain/Definition*
A, a	System coefficient (original, reduced)	$\mathbb{R}_{++}$
B, b	Control coefficient (original, reduced)	$\mathbb{R}\backslash\{0\}$
C	Measurement/observation coefficient	$\mathbb{R}_{++}$
$K(\cdot)$	Optimal cost, $K = K_1 + K_2$	$K : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$
$K_1(\cdot)$	Optimal cost for estimated mean (x)	$K_1 : \mathcal{X} \to \mathbb{R}_+$
$K_2(\cdot)$	Optimal cost for estimated variance (y)	$K_2 : \mathcal{Y} \to \mathbb{R}_+$
M	Variance of the measurement/observation noise	$\mathbb{R}_+$
N	Reduced noise variance of the linear system ($N = \bar{N}/\lambda^2$)	$\mathbb{R}_+$
$\bar{N}$	Inherent noise variance of the linear system	$\mathbb{R}_+$
P	Optimal quadratic cost coefficient for mean cost function	$\mathbb{R}_+$
Q	Optimal linear cost coefficient for mean cost function	$\mathbb{R}_+$
R	Optimal constant coefficient for mean cost function	$\mathbb{R}_+$
$\hat{V}_t$	Estimated variance at time t	$\mathcal{Y}$
k_0	Optimal constant coefficient of $K_2(\cdot)$	$\mathbb{R}_{++}$
k_1	Optimal linear coefficient of $K_2(\cdot)$	$\mathbb{R}_{++}$
n	Period index (to denote $t+n$, given some t)	$\mathbb{N}$
p	Price (e.g., tax) for carbon emissions	$\mathbb{R}$
p_0	Constant cost coefficient (original)	$\mathbb{R}$
p_1	Linear cost coefficient (original)	$\mathbb{R}$
p_2	Quadratic cost coefficient (original)	$\mathbb{R}_{++}$
q	Cost coefficient of the estimated variance	$\mathbb{R}_{++}$
r	Policy parameter	$(0, 1)$
s	Reduced linear cost coefficient	$\mathbb{R}$
t	Discrete time	$\mathcal{T}$
u_t	System control at time t	$\mathcal{U}$
v_t	Information control at time t	$\mathcal{V}$
x	Prior estimated mean	$\mathcal{X}$
$\hat{x}_t$	Estimated mean at time t	$\mathcal{X}$
y	Prior variance	$\mathcal{Y}$
y'	Posterior variance	$\mathcal{Y}$
y^*	Variance threshold	$(N, \infty]$
$\bar{y}$	Asymptotic variance (steady state), for $a \in (0, 1)$	(N, ∞)
$y^{(n)}$	Variance without information at time $t+n$, given y at time t	$\mathcal{Y}$
z_t	System output	$\mathbb{R}$
$\alpha(y)$	y-conditional target variance	$(N, \infty]$
α^*	Unconditional target variance	$(N, \infty]$
β	Discount factor	$(0, 1)$
γ	Control-cost coefficient	$\mathbb{R}_+$
δ	Information-acquisition-cost coefficient	$\mathbb{R}_+$
Δ	Discriminant of the information-acquisition problem	$\mathbb{R}$
$\hat{\Delta}$	Subcritical discriminant threshold, for $a \in (0, 1)$	$\mathbb{R}_{--}$
$\tilde{\varepsilon}_t$	Gaussian system noise (i.i.d.) at time t	$\mathcal{N}(0, \bar{N})$
$\tilde{\eta}_t$	Gaussian measurement/observation noise (i.i.d.) at time t	$\mathcal{N}(0, M)$

(*continued on next column*)

Symbol	*Description*	*Domain/Definition*
θ, θ'	Parameter vector (original, reduced)	$\mathbb{R}^{12}, \mathbb{R}^6$
λ	Scaling factor	$\mathbb{R}_{++}$
$\hat{\mu}(\cdot), \mu(\cdot)$	Optimal system-stabilization policy (original, reduced)	$\hat{\mu}, \mu : \mathcal{X} \to \mathbb{R}$
$\hat{\nu}(\cdot), \nu(\cdot)$	Optimal information-acquisition policy (original, reduced)	$\hat{\nu}, \nu : \mathcal{Y} \to \mathbb{R}_+$
$\mathcal{A}_y$	y-conditional cost-improvement set	Eq. (9)
$\mathcal{A}$	Unconditional cost-improvement set	Eq. (10)
$\mathcal{T}$	Discrete time domain	$\mathbb{N}$
$\mathcal{U}$	Domain for system control u	$\mathbb{R}$
$\mathcal{V}$	Domain for information control v	$\mathbb{R}_+$
$\mathcal{X}$	State space	$\mathbb{R}$
$\mathcal{Y}$	Variance space	$\mathbb{R}_+$

References

Anderson, B. D. O., & Moore, J. B. (1971). *Linear optimal control*. Englewood Cliffs, NJ: Prentice-Hall.

Arrow, K., Dasgupta, P., Goulder, L., Daily, G., Ehrlich, P., Heal, G., et al. (2004). Are we consuming too much? *Journal of Economic Perspectives, 18*(3), 147–172.

Arrow, K. J., Harris, T., & Marschak, J. (1951). Optimal inventory policy. *Econometrica, 19*(3), 250–272.

Athans, M. (1972a). The discrete time linear-quadratic-Gaussian stochastic control problem. *Annals of Economic and Social Measurement, 1*(4), 449–491.

Athans, M. (1972b). On the determination of optimal costly measurement strategies for linear stochastic systems. *Automatica, 8*(4), 397–412.

Aubin, J. P. (1977). *Applied abstract analysis*. New York, NY: Wiley.

Auray, S., Mariotti, T., & Moizeau, F. (2011). Dynamic regulation and quality. *Rand Journal of Economics, 42*(2), 246–265.

Başar, T., & Bansal, R. (1994). Optimum design of measurement channels and control policies for linear-quadratic stochastic systems. *European Journal of Operational Research, 73*(2), 226–236.

Bansal, R., & Başar, T. (1989). Simultaneous design of measurement and control strategies for stochastic systems with feedback. *Automatica, 25*(5), 679–694.

Bellman, R. E. (1954). The theory of dynamic programming. *Bulletin of the American Mathematical Society, 60*(6), 503–516.

Bemporad, A., & Morari, M. (1999). Robust model predictive control: A survey. In A. Garulli, A. Tesi, & A. Vicino (Eds.), *Robustness in identification and control*. In *Lecture notes in control and information systems* (pp. 207–226). NY: Springer.

Benigno, P., & Woodford, M. (2004). Optimal monetary and fiscal policy: A linear-quadratic approach*: 18. NBER macroeconomics annual 2013* (pp. 271–333).

Bertsekas, D. P. (1995). *Dynamic programming and optimal control*. Belmont, MA: Athena Scientific.

Bertsekas, D. P., & Shreve, S. E. (1978/1996). *Stochastic optimal control: The discrete-time case*. Belmont, MA: Athena Scientific. [Originally published by Academic Press in 1978.]

Blackwell, D. (1951). Comparison of experiments. In *Proceedings of the second Berkeley symposium on mathematical statistics and probability* (pp. 93–102). Berkeley, CA: University of California Press.

Bonatti, A. (2011). Menu pricing and learning. *American Economic Journal: Microeconomics, 3*(3), 124–163.

Bond-Lamberty, B., Peckham, S. D., Ahl, D. E., & Gower, S. T. (2007). Fire as the dominant driver of central Canadian boreal forest carbon balance. *Nature, 450*(1), 89–92.

Cipriano, L., & Weber, T. A. (2018). Population-level intervention and information collection in dynamic healthcare policy. *Health Care Management Science, 21*(4). doi:10.1007/s10729-017-9415-5. forthcoming.

Davis, M. H. A. (1977). *Linear estimation and stochastic control*. London, UK: Chapman and Hall/CRC.

DeCroix, G. A., & Mookerjee, V. S. (1997). Purchasing demand information in a stochastic-demand inventory system. *European Journal of Operational Research, 102*(1), 36–57.

DeGroot, M. (1962). Uncertainty, information, and sequential experiments. *Annals of Mathematical Statistics, 33*(2), 404–419.

Dixit, A. K. (1990). *Optimization in economic theory* (Second Edition). Oxford, UK: Oxford University Press.

Fleming, W. H., & Rishel, R. W. (1975). *Deterministic and stochastic optimal control*. Berlin, Germany: Springer.

Friedman, R. (1981). The reliability of policy recommendations and forecasts from linear econometric models. *International Economic Review, 22*(2), 415–428.

Fu, Q., & Zhu, K. (2010). Endogenous information acquisition in supply chain management. *European Journal of Operational Research, 201*(2), 454–462.

Fudenberg, D., & Tirole, J. (1986). *Dynamic models of oligopoly*. Oxon, UK: Routledge.

Gallego, G. (1990). Scheduling the production of several items with random demands in a single facility. *Management Science, 36*(12), 1579–1592.

Gittins, J., Glazebrook, K., & Weber, R. (2011). *Multi-armed bandit allocation indices* (Second Edition). New York, NY: Wiley.

Goel, A., Myerson, A., & Weber, T. A. (2009). Fair welfare maximization. *Economic Theory, 41*(3), 465–494.

Gupta, V., Chung, T. H., Hassibi, B., & Murray, R. M. (2006). On a stochastic sensor selection algorithm with applications in sensor scheduling and sensor coverage. *Automatica, 42*(2), 251–260.

Hernández-Lerma, O., & Lasserre, J. B. (1996). *Discrete-time Markov control processes: Basic optimality criteria*. New York, NY: Springer.

Hoel, M., & Karp, L. (2002). Taxes versus quotas for a stock pollutant. *Resource and Energy Economics, 24*(4), 367–384.

Holt, C. C., Modigliani, F., Muth, J., & Simon, H. (1960). *Planning production, inventories and work force*. Englewood Cliffs, NJ: Prentice-Hall.

IEA (2015). Energy and climate change. In *World energy outlook special report*. International Energy Agency, Paris, France.

IPCC (2014). *Climate change 2014: Synthesis report*. Geneva, Switzerland: Intergovernmental Panel on Climate Change.

Kailath, T. (1980). *Linear systems*. Englewood Cliffs, NJ: Prentice-Hall.

Kalman, R. E. (1960). A new approach to filtering and prediction problems. *Transactions of the ASME – Journal of Basic Engineering, D82*(1), 35–45.

Kihlstrom, R. E. (1984). A Bayesian exposition of Blackwell's theorem on the comparison of experiments. In M. Boyer, & R. E. Kihlstrom (Eds.), *Bayesian models in economic theory* (pp. 13–31). Amsterdam, NL: North-Holland.

Krishnamurthy, V. (2002). Algorithms for optimal scheduling and management of hidden Markov model sensors. *IEEE Transactions on Signal Processing, 50*(6), 1382–1397.

Kydland, F. (1975). Noncooperative and dominant player solutions in discrete dynamic games. *International Economic Review, 16*(2), 321–335.

Lauritzen, S. L. (1981). Time series analysis in 1880: A discussion of contributions made by T.N. Thiele. *International Statistical Review, 49*(3), 319–331.

LaValle, I. (1968). On cash equivalents and information evaluation in decisions under uncertainty. *Journal of the American Statistical Association, 63*(321), 252–290.

Lawrence, D. B. (1999). *The economic value of information*. New York, NY: Springer.

Lindset, S., Lund, A. C., & Matsen, E. (2009). Optimal information acquisition for a linear quadratic control problem. *European Journal of Operational Research, 199*(2), 435–441.

Lovejoy, W. S. (1987). On the convexity of policy regions in partially observed systems. *Operations Research, 35*(4), 619–621.

Maillart, L. M., Ivy, J. S., Ransom, S., & Diehl, K. (2008). Assessing dynamic breast cancer screening policies. *Operations Research, 56*(6), 1411–1427.

McCardle, K. F. (1984). Information acquisition and the adoption of new technology. *Management Science, 31*(11), 1372–1389.

Meier, L. (1965). *Combined optimum control and estimation theory*. Stanford Research Institute Nasa contractor report, contract nas 2-2457.

Milgrom, P., & Roberts, J. (1988). Communication and inventory as substitutes in organizing production. *Scandinavian Journal of Economics, 90*(3), 275–289.

Monahan, G. E. (1980). Optimal stopping in a partially observable Markov process with costly information acquisition. *Operations Research, 28*(6), 1319–1334.

Moore, J. C., & Whinston, A. B. (1986). A model of decision-making with sequential information-acquisition (part 1). *Decision Support Systems, 2*(4), 285–307.

Moore, J. C., & Whinston, A. B. (1987). A model of decision-making with sequential information-acquisition (part 2). *Decision Support Systems, 3*(1), 47–72.

Moscarini, G., & Smith, L. (2001). The optimal level of experimentation. *Econometrica, 69*(6), 1629–1644.

Naish, H. F. (1994). Production smoothing in the linear-quadratic inventory model. *Economic Journal, 104*(425), 864–875.

Pindyck, R. S., & Roberts, S. M. (1974). Optimal policies for monetary control. *Annals of Economic and Social Measurement, 3*(1), 207–237.

Rudin, W. (1976). *Principles of mathematical analysis* (Third Edition). New York, NY: McGraw-Hill.

Salmon, M. (1982). Error correction mechanisms. *Economic Journal, 92*(367), 615–629.

Scarf, H. E. (1960). The optimality of (s, s) policies in the dynamic inventory problem. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), *Mathematical methods in the social sciences*. Stanford, CA: Stanford University Press. Chapter 11

Schäl, M. (1975). Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete (Probability Theory and Related Fields), 32*(3), 179–196.

Shmakov, S. L. (2011). A universal method of solving quartic equations. *International Journal of Pure and Applied Mathematics, 71*(2), 251–259.

Simon, H. (1956). Dynamic programming under uncertainty with a quadratic criterion function. *Econometrica, 24*(1), 74–81.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics, 50*(3), 665–690.

Strauch, R. E. (1966). Negative dynamic programming. *Annals of Mathematical Statistics, 37*(4), 871–890.

Striebel, C. (1965). Sufficient statistics on the optimum control of stochastic systems. *Journal of Mathematical Analysis and Applications, 12*(3), 576–592.

Theil, H. (1957). A note on certainty equivalence in dynamic planning. *Econometrica, 25*(2), 346–349.

Thiele, T. N. (1880). Om anvendelse af mindste kvadraters methode i nogle tilfælde, kvor en komplikation af visse slags uensartede tilfældige fejlkilder giver fejlene en 'systematisk' karakter. *Det kongelige danske Videnskabernes Selskabs Skrifter, 5*. Sur la compensation de quelques erreurs quasi-systématiques par la méthode des moindres carrés, C.A. Reitzel, Copenhagen, Denmark; reprinted in 2013 by Forgotten Books, London, UK.

Townsend, R. M. (1983). Forecasting the forecasts of others. *Journal of Political Economy, 91*(4), 546–588.

Tsodikov, A. D., & Yakovlev, A. Y. (1991). On the optimal policies of cancer screening. *Mathematical Biosciences, 107*(1), 21–45.

Wald, A. (1947). Foundations of a general theory of sequential decision functions. *Econometrica, 15*(4), 279–313.

Weber, T. A. (1997). Constrained predictive control for corporate policy. *Technical report LIDS-TH 2398*. Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology, Cambridge, MA.

Weber, T. A., & Neuhoff, K. (2010). Carbon markets and technological innovation. *Journal of Environmental Economics and Management, 60*(2), 115–132.

West, K. D. (1986). A variance bound test of the linear-quadratic inventory model. *Journal of Political Economy, 94*(2), 374–401.

Wonham, W. M. (1968). On the separation theorem of stochastic control. *SIAM Journal on Control, 6*(2), 312–326.

Yüksel, S., & Başar, T. (2013). *Stochastic networked control systems*. New York, NY: Springer.