Value of Communication in Goal-Oriented Semantic Communications: A Pareto Analysis

Jiping Luo, Bowen Li, and Nikolaos Pappas, *Senior Member, IEEE*

***Abstract*—Emerging cyber-physical systems increasingly operate under stringent communication constraints that preclude the reliable transmission of their extensive machine-type data streams. Since raw measurements often contain correlated or redundant components, effective operation depends not on transmitting all available data but on selecting the information that contributes to achieving the objectives of the system. Beyond accuracy, goal-oriented semantic communication assesses the *value of information* and aims to generate and transmit only what is relevant and at the right time. Motivated by this perspective, this work studies the *value of communication* through the canonical setting of remote estimation of Markov sources, where a value-of-information measure quantifies the relevance of information. We investigate how optimal estimation performance varies with the available communication budget and determine the marginal performance gain attributable to additional communication. Our approach is based on a *Pareto analysis* that characterizes the complete set of policies that achieve optimal trade-offs between estimation performance and communication cost. The value of communication is defined as the absolute slope of the resulting Pareto frontier. Although computing this frontier is non-trivial, we demonstrate that in our setting it admits a notably tractable structure: it is strictly decreasing, convex, and piecewise linear, and its slope is governed by a finite collection of constants. Moreover, each Pareto-optimal operating point is realizable as a convex combination of two stationary deterministic policies, enabling practical implementation. Leveraging these structural insights, we introduce SPLIT, an efficient and provably optimal algorithm for constructing the complete Pareto frontier.**

***Index Terms*—Age and value of information, value of communication, goal-oriented semantic communications, Pareto analysis, remote estimation.**

I. Introduction

Virtually all modern wireless systems are built upon fundamental principles of reliable communications first developed in Shannon's seminal work [1]. Under this classical view, "reliability" primarily refers to the *accuracy* of signal reconstruction. However, in emerging cyber-physical systems that generate massive volumes of machine-type data, the timeliness and contextual relevance of information often outweigh mere signal fidelity. It has become increasingly important that the information conveyed is *fresh*, *significant*, and aligned with the application context and decision-making goals. Thus motivated, goal-oriented semantic communication aims to exploit the *value (semantics) of information* to generate and transmit only what is relevant and at the right time [2]–[4].

The authors are with the Department of Computer and Information Science, Linköping University, Linköping 58183, Sweden (e-mail: jiping.luo@liu.se; bowen.li@liu.se; nikolaos.pappas@liu.se). This work has been supported by the Swedish Research Council (VR), ELLIIT, the European Union (ETHER, 101096526, 6G-LEADER, 101192080, and ROBUST-6G, 101139068), and the Graduate School in Computer Science (CUGS).

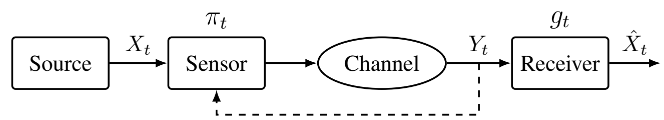


Fig. 1. Remote state estimation of a Markov source.

The value of information depends on the application context and goals. In causal systems, this value can be inferred from the history of all past system realizations [4], in contrast to distortion measures [5], which are typically history-independent. The notion of *age* of information (AoI) is the first concrete step in this direction, built on the premise that information holds the greatest value when it is fresh [6], [7]. Yet, this assumption deserves re-examination, as the freshest measurement may be inaccurate or uninformative. More recently, a series of semantics-aware metrics has been introduced that go beyond accuracy and age, capturing aspects such as context-awareness [3], [8], [9], the lasting impact of consecutive errors [10]–[12], and the urgency of lasting impact [4], [13].

Despite this progress, the *value of communication* has yet to be satisfactorily characterized. In the existing literature, one relatively well-studied question is the following: Given a measure of the value of information and a communication budget (or cost), what is the best achievable performance and the optimal communication timing (i.e., the *value of timing* problem)? Analytical tools from Markov decision theory, such as Lagrangian and constrained formulations, play a central role in addressing this problem; see, e.g., [4], [10], [14]. In practice, however, the communication budget is a parameter that may vary over time and requires careful calibration by the system designer. This motivates the problem of value of communication: Given a desired performance goal and an operating budget range, *how sensitive is the optimal performance to changes in the communication budget?* In other words, what is the marginal benefit of communication, and when is it worth allocating additional transmissions?

In this work, we characterize the value of communication through the remote state estimation of Markov sources, a fundamental problem in information and control theory. As depicted in Figure 1, a sensor observes a Markov source X and decides when to send its current observation to a receiver over a wireless channel. The receiver employs the maximum a posteriori (MAP) estimation policy g that reconstructs the source based on all received messages. Our goal is to devise a communication policy π that achieves an optimal trade-off between the long-run average estimation and communication costs. Unlike the classical notion that estimation quality is synonymous with accuracy, we filter out less informative measurements by quantifying their value through the urgency of the lasting impact metric [4].

Our main contributions are as follows:

(1) We approach the problem through the lens of *Pareto analysis*. In contrast to the Lagrangian and constrained formulations, which yield a single "best" policy, the Pareto formulation seeks to characterize the entire set of optimal tradeoff policies (i.e., the Pareto frontier) that balance the conflicting estimation and communication costs. *The value of communication is defined as the absolute slope of this frontier, quantifying the marginal improvement in estimation performance per unit increase in communication.*

(2) However, the *exact* computation of the Pareto frontier for multi-objective Markov decision problems (MDPs) is a nontrivial task. As will become evident, while the Lagrangian and constrained formulations can identify *some* Pareto points, no practical algorithms currently exist to construct the complete frontier with formal guarantees. In this work, we shall address this challenge for the semantics-aware remote estimation problem and, more broadly, for a class of bi-objective MDPs that share similar tradeoffs. We show that the frontier in our study has a simple and exploitable structure: it is *strictly decreasing*, *convex*, and *piecewise linear*. Each corner point (where the slope changes) on this frontier corresponds to a pure (i.e., stationary and deterministic) policy, and the value of communication takes values in a *finite* set of constants (more precisely, Lagrange multipliers). Moreover, every Pareto-optimal policy is a randomized mixture (i.e., convex combination) of two pure policies, and only finitely many such policies are needed to construct the entire frontier. Importantly, these mixture policies are easy to implement: when the operating budget changes, one selects the appropriate pair of pure policies from the lookup table and adjusts the mixture coefficient accordingly.

(3) We then leverage the structural properties of the Pareto frontier to develop an efficient, *provably* optimal algorithm called Structured Pareto Line Intersection Tracing (SPLIT). SPLIT constructs the *exact* frontier by identifying all its corner points, with a time complexity that is *linear* in the number of corner points. Surprisingly, this complexity is lower than that of classical methods that solve the constrained problem for a single budget. In contrast, existing frontier approximation methods entail substantially higher complexity while providing only coarse representations.

Organization: The paper is organized as follows. Section II summarizes some related work. Sections III and IV present the system model and problem formulations. Section V establishes the structural properties of the Pareto frontier, building on which Section VI develops a fast and exact frontier construction algorithm. Numerical results and conclusions are provided in Sections VII and VIII, respectively.

II. Literature Overview

A. Goal-Oriented Semantic Communications

Information *accuracy* has long been the primary performance metric in information and control theory [1], [5], [15]. The rate-distortion theorem, however, makes clear that many emerging cyber-physical systems cannot reliably communicate the massive volumes of machine-type data they generate. From a data significance perspective, they need not do so, as raw measurements can be correlated, redundant, or irrelevant to achieving the system's goals.

Timeliness has emerged as a crucial requirement in applications such as intelligent transportation, industrial automation, and swarm robotics [16]–[18]. The AoI was introduced in [6] to quantify the timeliness (or freshness) of the information that a remote receiver has about a source. Reviews of AoI and its applications can be found in [19]–[21]. It is worth noting that the impact of aged information on networked control systems (NCSs) has been independently and extensively studied from a control-theoretic perspective [22], [23]. In particular, a Kalman filter produces estimates of a linear Gaussian process and pushes them into the data queue for transmission. For control stability, it is unnecessary to transmit every queued packet, since the newest estimate is as accurate as the older ones and therefore best reflects the current state of the system (see, e.g., [24]–[27]). This canonical example illustrates the basic idea of AoI: when the newest measurement best reflects the system status, outdated information can be discarded to ensure timely operation at the destination. However, this assumption does not always hold. Recent studies have shown that age-based policies are often suboptimal for monitoring general information sources, such as Wiener processes [14], Ornstein-Uhlenbeck processes [28], and Markov chains [8], [9], [11].

Several AoI variants have been proposed to balance timeliness and accuracy. The effective age [29] penalizes the cumulative error incurred in the absence of updates. Similar in spirit, the age of incorrect information (AoII) [10] and the cost of memory error [11] quantify the duration for which the system remains in erroneous states. Subsequent work analyzes these metrics in systems with random channel delays [30], multiple access [31], and continuous-time Markov sources [32], [33], among others. There are also variants such as version age designed to prevent unnecessary growth in age for specific tasks, including [34]–[39]. The intuition is that the age need not grow when the source remains unchanged.

The aforementioned metrics, however, treat all source states equally and do not account for their *contextual relevance* or goal-oriented usefulness. Information that is both fresh and accurate can still be uninformative if it does not meaningfully contribute to the goal. Early efforts to address this limitation include the context-aware age metric [8], which models the information source as a Markov chain with normal and alarm states. This metric penalizes the staleness of information about urgent states more heavily, highlighting that delayed updates may have unequal consequences depending on the context. Another relevant metric is the cost of actuation error (CAE) [40], which captures the fact that the cost of estimation error depends not only on the physical discrepancy but also on the potential control risks to system performance. Follow-up studies in [11], [41], [42] analyzed CAE for monitoring Markov sources. The work in [9], [43] extended these results to resource-constrained multi-source systems, a setting particularly relevant for multimodal systems.

An important semantic attribute motivated by many NCSs is the *lasting impact* in consecutive estimation errors; that is, the longer an error persists, the more severe its consequences can become. The AoII and the cost of memory error capture this notion, though they are context-agnostic. Motivated by this, [12], [44] introduced the age of missed alarm (AoMA) and the age of false alarm (AoFA) to quantify, respectively, the persistence costs in the more urgent missed alarm error and the less important false alarm error. Moreover, the severity of an estimation error depends on both its context and duration. The significance-aware age of consecutive error (AoCE) [4], [45] thus measures the *urgency of lasting impact* through a set of context-aware nonlinear functions. Optimal communication policies have been derived through the lens of Markov decision theory. Prior work has established that the optimal policy under a Lagrangian formulation exhibits a simple switching structure (i.e., multiple thresholds) whose transmission thresholds depend on the instantaneous estimation error [4], [12]. Furthermore, [13] shows that when a MAP estimator is used, this switching structure also depends on the AoI. The optimal policy for the constrained formulation is a randomized mixture of two such switching policies.

B. Multi-Objective MDPs

There is a rich literature on single-objective MDPs, including the Lagrangian formulation [46] and the constrained formulation [47]. These formulations play a central role in goal-oriented semantic communications (see, e.g., [4], [9], [13]). To tackle the value of communication problem, we need to resort to multi-objective MDPs, which involve multiple, potentially dependent and conflicting objectives. However, the *exact* characterization of the Pareto frontier, even for two objectives, is a non-trivial task [48]–[50].

There are two main approaches to multi-objective MDPs: the *scalarization* method and the *evolutionary* method. However, only approximate or partial solutions can be obtained. In the scalarization approach, the multiple objectives are combined into a single objective using a weighted sum (i.e., the Lagrangian formulation) [51], [52] or other functional mapping [53], allowing standard single-objective MDP techniques to be applied. However, it yields only a *single* Pareto point per weight configuration, requiring multiple runs to explore different parts of the frontier [54]. Moreover, as we shall show, traversing all weights is neither practical nor meaningful, since it can identify only finitely many points on the frontier.

Evolutionary methods, in contrast, *approximate* the Pareto frontier by exploring a population of policies in a single run. For example, the vector-valued dynamic programming [55] computes a set of non-stationary Pareto-optimal policies; however, it is computationally infeasible to enumerate all Pareto-optimal policies. This algorithm was later modified in [56] to accelerate computation by restricting the search to stationary deterministic policies. Similarly, the convex hull value iteration [52] computes a set of deterministic policies that form the convex hull of the Pareto frontier. The Pareto Q-learning [57] maintains a vector Q-table to identify a set of deterministic policies. Nevertheless, existing evolutionary algorithms do not provide formal guarantees regarding the optimality or quality of the solutions they produce.

In this work, we address a bi-objective MDP that encompasses a broad class of problems in communication and control systems. In particular, it exhibits a tradeoff in which relaxing one objective (e.g., energy consumption) never degrades the best achievable performance of the other objective (e.g., estimation accuracy). We derive several structural properties of the Pareto frontier and develop an efficient algorithm, SPLIT, to compute the exact frontier.

III. Semantics-Aware Communication Model

A. System Model

Consider a remote estimation system consisting of an information source, a sensor (transmitter), a wireless channel, and a remote estimator (receiver), as depicted in Figure 1. The information source $\{X_t\}_{t\geq 1}$ is modeled as an irreducible Markov chain taking values in a finite alphabet $\mathcal{X}$. The chain's transition probability matrix is given by

$$Q = \left(Q_{i,j}, i, j \in \mathcal{X}\right),\ Q_{i,j} = \Pr[X_{t+1} = j | X_t = i].$$

Each state of the chain represents either a quantized level of a physical process or, as is our main focus, an abstract status of a dynamical system (e.g., operating modes, component failures, or anomalies). As such, the chain is *prioritized*, with states differing in their relative importance to system performance.

At each time t, the sensor observes the source and decides whether to transmit its current observation to the receiver. Denote by $U_t \in \mathcal{U}$ the sensor's decision at time t, where $U_t = 0$ means no transmission and $U_t = 1$ means a decision to transmit. We consider a practical network in which the sensor can transmit only intermittently over a packet-drop wireless channel. At each time t, the transmitted packet is either successfully decoded by the receiver or erased due to channel unreliability. Let H_t denote the erasure indicator, where $H_t = 1$ represents a successful reception and $H_t = 0$ represents an erasure event. The sequence $\{H_t\}$ is modeled as an independent and identically distributed (i.i.d.) Bernoulli process with

$$\Pr[H_t = 1] = p_s,\ \Pr[H_t = 0] = 1 - p_s = p_f.$$

The receiver generates a real-time estimate $\hat{X}_t \in \mathcal{X}$ based on all received messages $Y_{1:t}$. At each time t, the received message is given by

$$Y_t = \begin{cases} X_t, & U_t H_t = 1, \\ \mathcal{E}, & U_t H_t = 0, \end{cases} \tag{1}$$

where $\mathcal{E}$ denotes that no message was received, which occurs either when no transmission is attempted or when a transmission fails. The sensor is informed of the channel outcomes through an error-free feedback link.

B. Decision Rules

The time ordering of the relevant variables and decision rules is depicted in Figure 2.

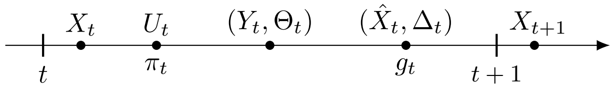


Fig. 2. Time ordering of the relevant variables and decision rules.

The sensor has causal access to the current and past source states $X_{1:t}$, the past decisions $U_{1:t-1}$, and the delivered messages $Y_{1:t-1}$. For each t, the sensor's decision rule is

$$\pi_t : \mathcal{X}^t \times \mathcal{U}^{t-1} \times \mathcal{Y}^{t-1} \to \text{Prob}(\mathcal{U}),$$

where $\text{Prob}(\mathcal{U})$ denotes the set of all probability distributions over the action space $\mathcal{U}$. Then we write

$$U_t \sim \pi_t(\cdot|X_{1:t}, U_{1:t-1}, Y_{1:t-1}). \tag{2}$$

The collection $\pi = (\pi_1, \pi_2, \ldots)$ is called a *communication policy*. Let Π denote the set of admissible policies. A policy is stationary if it applies the same decision rule at all times, i.e., $\pi_t = \pi$ for all t. Of particular interest are stationary deterministic policies that select an action with certainty, i.e.,

$$U_t = \pi(X_{1:t}, U_{1:t-1}, Y_{1:t-1}). \tag{3}$$

The receiver employs the MAP estimation rule to reconstruct the source state, and this rule is known to the sensor. Upon receiving Y_t at time t, an estimate $\hat{X}_t$ is generated by the rule [58, Ch. 3.5]

$$\hat{X}_t = g_t(Y_{1:t}) = \arg\max_{x \in \mathcal{X}} b_t(x), \tag{4}$$

where

$$b_t(x) := \Pr[X_t = x | Y_{1:t}]$$

denotes the receiver's posterior belief that the source is in state x at time t.

Let Θ_t denote the AoI at the receiver, defined as

$$\Theta_t := t - \max\{\tau \leq t : Y_\tau \neq \mathcal{E}\}. \tag{5}$$

In other words, Θ_t measures how fresh the receiver's knowledge is about an information source. The receiver's belief can then be expressed as

$$b_t(x) = \begin{cases} 1, & \Theta_t = 0, Y_t = x, \\ 0, & \Theta_t = 0, Y_t \neq x, \\ Q^{\Theta_t}_{Z_t, x}, & \Theta_t > 0, Y_t = \mathcal{E}, \end{cases} \tag{6}$$

where $Z_t = X_{t-\Theta_t}$ is the most recently received update at the receiver, and $Q^{\Theta_t}_{Z_t,x} = \Pr[X_t = x | X_{t-\Theta_t} = Z_t]$. The MAP estimate can then be written as a function of the AoI as

$$\hat{X}_t = g_t(Y_{1:t}) = g(Z_t, \Theta_t). \tag{7}$$

It reveals that the usefulness of the received data depends on both its content and its age.

For comparison, we also consider the widely used zero-order hold (ZOH) estimator [4], which holds the most recently received update until a new one arrives; that is

$$\hat{X}_t = g_{\text{ZOH}}(Z_t, \Theta_t) = Z_t.$$

The ZOH estimator serves as an important analytical tool in semantics-aware communication systems and is relevant when the receiver lacks prior knowledge of source statistics, which is not the case in this study.

C. Performance Criterion

In classical estimation systems, the system's performance is measured by *average distortion*, i.e.,

$$\mathcal{J}_{\text{classic}}(\pi) := \limsup_{T \to \infty} \mathbb{E}^\pi \left[\frac{1}{T} \sum_{t=1}^{T} d(X_t, \hat{X}_t) \right], \tag{8}$$

where $d : \mathcal{X} \times \mathcal{X} \to [0, \infty)$ is a bounded distortion measure, such as Hamming distortion or mean-square error (MSE), that quantifies the discrepancy between the source and its reconstruction. The expectation in (8) is taken with respect to a probability measure that is determined by the distributions of the source, the channel, and the choice of decision rules.

In semantics-aware systems, however, the contextual relevance and goal-oriented usefulness of information often outweigh mere signal fidelity. We now provide a formal definition of the *value of information*. Let

$$I_t = (X_{1:t}, U_{1:t}, Y_{1:t}, \hat{X}_{1:t})$$

denote the history of system realizations up to time t, and let $\mathcal{I}_t$ denote the set of all possible realizations.

Definition 1: The value of information conveyed by the current state X_t is extracted from the history I_t by

$$\Delta_t = \psi_t(I_t) = \psi_t(X_{1:t}, U_{1:t}, Y_{1:t}, \hat{X}_{1:t}),$$

where $\psi_t : \mathcal{I}_t \to \mathcal{D}$ is an extraction function at time t, and $\mathcal{D} \subseteq [0, \infty)$ is the value domain of the semantic attribute Δ_t.

In this work, we instantiate the semantic attribute as the *urgency of lasting impact* in consecutive errors [4]. That is, the cost of an estimation error depends on both its contextual significance and its duration. Intuitively, the more urgent an error and the longer it persists, the more severe its consequences can become. The error holding time is recursively defined as

$$\Delta_t = \begin{cases} \Delta_{t-1} + 1, & X_t \neq \hat{X}_t, (X_t, \hat{X}_t) = (X_{t-1}, \hat{X}_{t-1}), \\ 1, & X_t \neq \hat{X}_t, (X_t, \hat{X}_t) \neq (X_{t-1}, \hat{X}_{t-1}), \\ 0, & X_t = \hat{X}_t, \end{cases} \tag{9}$$

where $\Delta_t \in \mathbb{N}$, and $\mathbb{N}$ denotes the set of natural numbers.

We then define the urgency of lasting impact as

$$c(X_t, \hat{X}_t, \Delta_t) := \bar{d}(X_t, \hat{X}_t) \rho_{X_t, \hat{X}_t}(\Delta_t), \tag{10}$$

which incorporates both the instantaneous error $(X_t, \hat{X}_t)$ and the *history-dependent* value Δ_t. Here, $\bar{d} : \mathcal{X} \times \mathcal{X} \to [0, \infty)$ is a *context-aware* distortion measure capturing the contextual significance (or potential control risks) of an error [9], and $\rho_{i,j} : \mathbb{N} \to [0, \infty)$, $i, j \in \mathcal{X}$, $i \neq j$, are weakly increasing functions representing the persistence cost for the error (i, j).

We note that the analysis presented in this paper is readily applicable to many existing semantics-aware metrics, including those mentioned in Section II-A.

Then, the performance of the semantics-aware communication system is measured by

$$\mathcal{J}(\pi) := \limsup_{T\to\infty} \mathbb{E}^{\pi}\left[\frac{1}{T}\sum_{t=1}^{T} c(X_t, \hat{X}_t, \Delta_t)\right]. \tag{11}$$

Since Δ_t can grow indefinitely due to channel unreliability, the average cost $\mathcal{J}(\pi)$ may become unbounded regardless of the choice of policy π. To avoid such pathological cases, we impose the following asymptotic growth condition on the functions $\rho_{i,j}$ [4, Theorem 2]:

$$\lim_{\delta\to\infty} \frac{\rho_{i,j}(\delta+1)}{\rho_{i,j}(\delta)} < \frac{1}{Q_{i,i}p_f}, \quad i \neq j. \tag{12}$$

Similarly, the transmission frequency under π is defined as

$$F(\pi) := \limsup_{T\to\infty} \mathbb{E}^{\pi}\left[\frac{1}{T}\sum_{t=1}^{T} U_t\right]. \tag{13}$$

IV. Problem Formulations

A. Pareto Formulation

The goal is to achieve an optimal tradeoff between the average estimation cost $\mathcal{J}$ and the average communication cost F. To this end, we formulate a bi-objective Pareto optimization problem as

$$\inf_{\pi\in\Pi} \big\{\mathcal{J}(\pi), F(\pi)\big\}, \tag{14}$$

where Π denotes the set of admissible policies.

Problem (14) is a bi-objective MDP. There is generally no single "best" policy in the Pareto sense. Instead, the aim is to characterize the complete Pareto frontier (i.e., the tradeoff curve) that contains all Pareto-optimal solutions and, more importantly, provides insights into the *value of communication*. Such insights are particularly useful, for example, in quantifying the marginal benefit of allowing additional transmissions or in adapting to time-varying communication budget.

Definition 2: A policy $\pi^* \in \Pi$ is said to be Pareto-optimal if there exists no other policy $\pi \in \Pi$ such that

$$\mathcal{J}(\pi) \leq \mathcal{J}(\pi^*) \text{ and } F(\pi) \leq F(\pi^*)$$

with at least one of the above inequalities being *strict*. Let $\mathcal{P} \subseteq \Pi$ denote the set of all Pareto-optimal policies. The Pareto frontier $\mathcal{C}$ is defined in the objective space $(F, \mathcal{J})$ as

$$\mathcal{C} = \big\{\big(F(\pi), \mathcal{J}(\pi)\big) : \pi \in \mathcal{P}\big\}. \tag{15}$$

Definition 3: The value of communication represents the sensitivity of the estimation performance with respect to changes in the communication budget. It is defined as the absolute value of the right-hand derivative of $\mathcal{J}(F)$; that is

$$\eta(F) := \lim_{h\to 0^+} \frac{\big|\mathcal{J}(F+h) - \mathcal{J}(F)\big|}{h}. \tag{16}$$

The existence of this limit will be shown to hold.

A large value of η implies that allowing more transmissions leads to a significant reduction in estimation cost, whereas a small η implies that further transmissions provide only marginal improvements in estimation performance.

In optimal decision-making, it is of primary importance to identify a *sufficient statistic* to constrain the policy search space. Let

$$I'_t = (X_{1:t}, U_{1:t-1}, Y_{1:t-1})$$

denote the information available to the sensor at the beginning of time t. Note that $\hat{X}_{1:t-1}$, and hence $\Theta_{1:t-1}$ and $\Delta_{1:t-1}$, are known at the sensor through the MAP estimation rule. A sufficient statistic $S_t \subseteq I_t$ is the minimal information the sensor needs to retain without losing optimality. The next lemma identifies a sufficient statistic for this study (see, e.g., [4], [13]).

Lemma 1: A sufficient statistic for the sensor is

$$S_t = (X_t, Z_{t-1}, \Theta_{t-1}, \Delta_{t-1}). \tag{17}$$

Therefore, without loss of optimality, it suffices to consider communication policies of the form $U_t \sim \pi_t(\cdot|S_t)$.

Let $\mathcal{S} = \mathcal{X}^2 \times \mathbb{N}^2$ denote the state space of $\{S_t\}$.

The following formulations serve as the basis for characterizing the Pareto frontier. While they have been relatively well studied in semantic communication systems, their connections to Pareto optimization remain largely unexplored.

B. Constrained Formulation

This formulation seeks to minimize the average cost $\mathcal{J}(\pi)$ subject to a hard constraint on the transmission frequency $F(\pi)$. Let

$$\Pi_\beta := \big\{\pi \in \Pi : F(\pi) \leq \beta\big\} \tag{18}$$

denote the set of feasible policies for a fixed budget $\beta \in [0, 1]$. This constrained optimization problem is formulated as

$$\inf_{\pi\in\Pi_\beta} \mathcal{J}(\pi). \tag{19}$$

Definition 4: A policy $\pi^*_\beta \in \Pi_\beta$ is constrained-optimal for a given budget β if it attains the minimum in (19). Let $\Pi^*_\beta \subseteq \Pi_\beta$ denote the set of constrained-optimal policies for that β. For notational simplicity, we shall write

$$\mathcal{J}^*_\beta = \mathcal{J}(\pi^*_\beta) \text{ and } F^*_\beta = F(\pi^*_\beta).$$

Problem (19) is a Constrained MDP (CMDP) that is computationally challenging. A formal approach to CMDPs is the Lagrange multiplier method, which relaxes the constraint through a weighted sum of the two objectives.

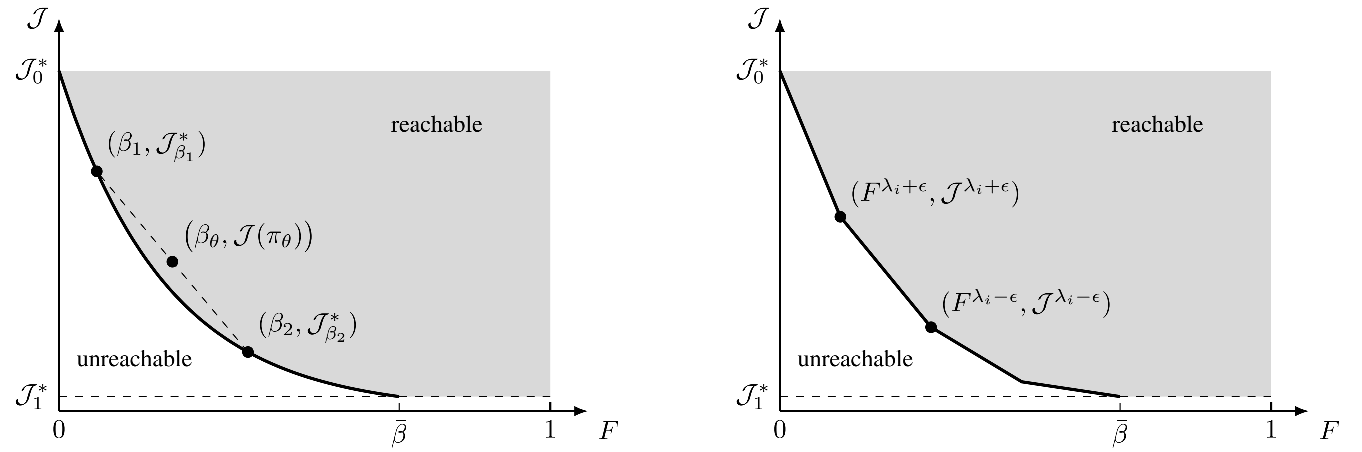
C. Lagrangian Formulation

Let $\lambda > 0$ denote the Lagrange multiplier, which can be interpreted as the resource utilization cost associated with each transmission attempt. The corresponding Lagrangian formulation yields a standard (unconstrained) MDP defined as

$$\inf_{\pi\in\Pi} \mathcal{L}^\lambda(\pi) = \mathcal{J}(\pi) + \lambda F(\pi). \tag{20}$$

Definition 5: A policy $\pi^\lambda \in \Pi$ is called λ-optimal if it attains the minimum in (20) for a given λ. Let $\Pi^\lambda \subseteq \Pi$ denote the set of λ-optimal policies for that λ. We shall write

$$\mathcal{J}^\lambda = \mathcal{J}(\pi^\lambda), \; F^\lambda = F(\pi^\lambda),$$



(a) Monotonicity and convexity (Theorem 1).

(b) Piecewise linearity (Theorem 2).

Fig. 3. Illustration of the Pareto frontier (solid line) obtained from (a) the constrained formulation and (b) the Lagrangian formulation. The shaded region represents the set of attainable objective values.

$$\mathcal{L}^\lambda = \mathcal{L}^\lambda(\pi^\lambda) = \mathcal{J}^\lambda + \lambda F^\lambda.$$

The following questions relating to the Pareto frontier are of interest:

1) How are these formulations related?
2) What is the structure of the Pareto frontier? What are the characteristics of Pareto-optimal policies, and, more importantly, are they easy to implement?
3) Can the value of communication be expressed explicitly?
4) Can the frontier be computed exactly and efficiently?

We shall answer the first three questions in Section V and the last question in Section VI.

V. Pareto Analysis

This section characterizes the structure of the Pareto frontier and quantifies the value of communication in the considered system model. Two key results, Theorem 1 and Theorem 2, show that every constrained-optimal policy (within a certain budget range) and every λ-optimal policy are Pareto-optimal. Importantly, the Pareto frontier is *strictly decreasing*, *convex*, and *piecewise linear*, and the value of communication is fully characterized by a *finite* set of Lagrange multipliers. These findings reveal the connections among the constrained, Lagrangian, and Pareto formulations and will later prove useful for efficient algorithm design.

A. Monotonicity and Convexity of the Pareto Frontier

We start with the constrained formulation. Let $\bar{\beta}$ be the minimum "sufficient" budget, defined as

$$\bar{\beta} = \inf\big\{\beta \in [0,1] : \mathcal{J}^*_\beta = \mathcal{J}^*_1\big\}. \tag{21}$$

In other words, $\bar{\beta}$ is the smallest budget beyond which additional communication provides no further improvement (value) in estimation performance.

The next theorem states that every constrained-optimal policy π^*_β, $\beta \in [0, \bar{\beta}]$, is Pareto-optimal, and that varying β traces the entire Pareto frontier.

Theorem 1: For any $\beta \in [0, \bar{\beta}]$, every constrained-optimal policy π^*_β is Pareto-optimal. The Pareto frontier is characterized by

$$\mathcal{C} = \big\{\big(\beta, \mathcal{J}^*_\beta\big) : \beta \in [0, \bar{\beta}]\big\}. \tag{22}$$

The set of Pareto-optimal policies is given by

$$\mathcal{P} = \big\{\pi^*_\beta \in \Pi^*_\beta : \beta \in [0, \bar{\beta}]\big\}. \tag{23}$$

To establish this result, we first introduce two key lemmas concerning the *monotonicity* and *convexity* of $\mathcal{J}^*_\beta$, as illustrated in Figure 3a. These properties play a central role in our analysis and algorithmic design that follow.

Lemma 2: $\mathcal{J}^*_\beta$ is weakly decreasing and convex in $\beta \in [0, 1]$.

Proof: The proof proceeds in two parts.

Part (i): Monotonicity. Since allowing more transmissions never worsens the best achievable estimation performance, we have for any $\beta_1 < \beta_2$, $\Pi_{\beta_1} \subseteq \Pi_{\beta_2}$ and

$$\mathcal{J}^*_{\beta_1} = \inf_{\pi \in \Pi_{\beta_1}} \mathcal{J}(\pi) \geq \inf_{\pi \in \Pi_{\beta_2}} \mathcal{J}(\pi) = \mathcal{J}^*_{\beta_2}.$$

This shows that $\mathcal{J}^*_\beta$ is weakly decreasing in $\beta \in [0, 1]$.

Part (ii): Convexity. Convexity requires that for any $\beta_1 < \beta_2$, $\beta_1, \beta_2 \in [0, 1]$, and any $\theta \in [0, 1]$, the constrained-optimal policy $\pi^*_{\beta_\theta}$ at the interpolated budget

$$\beta_\theta = \theta\beta_1 + (1-\theta)\beta_2$$

satisfies the inequality

$$\mathcal{J}^*_{\beta_\theta} \leq \theta\mathcal{J}^*_{\beta_1} + (1-\theta)\mathcal{J}^*_{\beta_2}.$$

We establish convexity by constructing a feasible policy $\pi_\theta \in \Pi_\theta$ such that $\mathcal{J}(\pi_\theta) = \theta\mathcal{J}^*_{\beta_1} + (1-\theta)\mathcal{J}^*_{\beta_2}$. The result then follows directly from the constrained-optimality of $\pi^*_{\beta_\theta}$ at β_θ. See Figure 3a for an illustration. We construct π_θ as follows. Let $\pi^*_{\beta_1}$ and $\pi^*_{\beta_2}$ be the constrained-optimal policies at budgets β_1 and β_2. In effect, they are stationary [47, Theorem 4.1]. Consider their convex combination

$$\pi_\theta = \theta\pi^*_{\beta_1} + (1-\theta)\pi^*_{\beta_2}, \tag{24}$$

where

$$\pi_\theta(u|s) = \theta\pi^*_{\beta_1}(u|s) + (1-\theta)\pi^*_{\beta_2}(u|s)$$

is the probability of choosing action $u \in \mathcal{U}$ in state $s \in \mathcal{S}$. By construction, π_θ satisfies $\sum_{u\in\mathcal{U}} \pi_\theta(u|s) = 1$ for all s. Hence, π_θ can be understood as a *randomized mixture policy* that selects the policy $\pi^*_{\beta_1}$ with probability θ and $\pi^*_{\beta_2}$ with probability $1-\theta$. Formally, we write

$$\pi_\theta = \begin{cases} \pi^*_{\beta_1}, & \text{w.p. } \theta, \\ \pi^*_{\beta_2}, & \text{w.p. } 1-\theta. \end{cases} \tag{25}$$

Randomized mixture policies will prove useful in several parts of this study.

By construction, the estimation cost of π_θ satisfies

$$\begin{aligned}
\mathcal{J}(\pi_\theta) &= \limsup_{T\to\infty} \mathbb{E}^{\pi_\theta}\left[\frac{1}{T}\sum_{t=1}^{T} c(X_t, \hat{X}_t, \Delta_t)\right] \\
&= \theta \limsup_{T\to\infty} \mathbb{E}^{\pi^*_{\beta_1}}\left[\frac{1}{T}\sum_{t=1}^{T} c(X_t, \hat{X}_t, \Delta_t)\right] \\
&\quad + (1-\theta) \limsup_{T\to\infty} \mathbb{E}^{\pi^*_{\beta_2}}\left[\frac{1}{T}\sum_{t=1}^{T} c(X_t, \hat{X}_t, \Delta_t)\right] \\
&= \theta\mathcal{J}(\pi^*_{\beta_1}) + (1-\theta)\mathcal{J}(\pi^*_{\beta_2}) \\
&= \theta\mathcal{J}^*_{\beta_1} + (1-\theta)\mathcal{J}^*_{\beta_2},
\end{aligned} \tag{26}$$

where the last equality follows because $\pi^*_{\beta_1}$ and $\pi^*_{\beta_2}$ are the constrained-optimal policies at budgets β_1 and β_2.

Similarly, the communication cost of π_θ satisfies

$$\begin{aligned}
F(\pi_\theta) &= \theta F(\pi^*_{\beta_1}) + (1-\theta)F(\pi^*_{\beta_2}) \\
&= \theta F^*_{\beta_1} + (1-\theta)F^*_{\beta_2} \\
&\le \theta\beta_1 + (1-\theta)\beta_2 = \beta_\theta,
\end{aligned} \tag{27}$$

where the inequality holds because every constrained-optimal policy π^*_β satisfies the constraint $F^*_\beta \le \beta$.

From (27), π_θ is feasible for the budget β_θ, i.e., $\pi_\theta \in \Pi_{\beta_\theta}$. Since $\mathcal{J}^*_{\beta_\theta}$ is the minimum average cost under budget β_θ, it follows that for any $\theta \in [0,1]$,

$$\mathcal{J}^*_{\beta_\theta} \le \mathcal{J}(\pi_\theta) = \theta\mathcal{J}^*_{\beta_1} + (1-\theta)\mathcal{J}^*_{\beta_2}, \tag{28}$$

which establishes the convexity of $\mathcal{J}^*_\beta$ in $\beta \in [0,1]$. ■

With Lemma 2, we can establish a stronger result as follows.

Lemma 3: $\mathcal{J}^*_\beta$ is strictly decreasing in $\beta \in [0,\bar{\beta}]$.

Proof: Suppose, to the contrary, that there exist $0 \le \beta_1 < \beta_2 < \bar{\beta}$ such that

$$\mathcal{J}^*_{\beta_1} = \mathcal{J}^*_{\beta_2}.$$

From Lemma 2, we know that $\mathcal{J}^*_\beta$ is convex and nonincreasing in $\beta \in [0,1]$. Convexity implies that for any $\beta_3 \in (\beta_2, \bar{\beta}]$,

$$\mathcal{J}^*_{\beta_2} \le \tilde{\theta}\mathcal{J}^*_{\beta_1} + (1-\tilde{\theta})\mathcal{J}^*_{\beta_3}, \tag{29}$$

where $\tilde{\theta} = \frac{\beta_3-\beta_2}{\beta_3-\beta_1} < 1$. By assumption, $\mathcal{J}^*_{\beta_1} = \mathcal{J}^*_{\beta_2}$, this yields

$$\mathcal{J}^*_{\beta_2} \le \mathcal{J}^*_{\beta_3}.$$

Given that $\mathcal{J}^*_\beta$ is nonincreasing in β, we must have $\mathcal{J}^*_{\beta_2} \ge \mathcal{J}^*_{\beta_3}$. Thus, we obtain

$$\mathcal{J}^*_{\beta_3} = \mathcal{J}^*_{\beta_2} \text{ for all } \beta_3 \in (\beta_2, \bar{\beta}], \tag{30}$$

which contradicts the definition of $\bar{\beta}$ in (21) that $\bar{\beta}$ as the minimal sufficient budget with $\mathcal{J}^*_\beta = \mathcal{J}^*_1$. Therefore, $\mathcal{J}^*_\beta$ is strictly decreasing in $\beta \in [0,\bar{\beta}]$. ■

With Lemma 3, we are now ready to establish Theorem 1.

Lemma 4: For any $\beta \in [0,\bar{\beta}]$, $\pi^*_\beta \in \mathcal{P}$ and $\big(\beta, \mathcal{J}^*_\beta\big) \in \mathcal{C}$.

Proof: We first prove that $\pi^*_\beta \in \mathcal{P}$. Suppose, to the contrary, that π^*_β, $\beta \in [0,\bar{\beta}]$, is not Pareto-optimal. Then there exists a feasible policy $\tilde{\pi} \in \Pi_\beta$ such that $\big(F(\tilde{\pi}), \mathcal{J}(\tilde{\pi})\big)$ weakly dominates $(F^*_\beta, \mathcal{J}^*_\beta)$, i.e.,

$$F(\tilde{\pi}) \le F^*_\beta \text{ and } \mathcal{J}(\tilde{\pi}) \le \mathcal{J}^*_\beta,$$

with at least one strict inequality.

Case (i): $F(\tilde{\pi}) \le F^*_\beta$ and $\mathcal{J}(\tilde{\pi}) < \mathcal{J}^*_\beta$. This contradicts the fact that π^*_β is constrained-optimal for Problem (19).

Case (ii): $F(\tilde{\pi}) < F^*_\beta$ and $\mathcal{J}(\tilde{\pi}) \le \mathcal{J}^*_\beta$. Let $\beta' = F(\tilde{\pi})$. By optimality at β', we have

$$\mathcal{J}^*_{\beta'} \le \mathcal{J}(\tilde{\pi}) \le \mathcal{J}^*_\beta. \tag{31}$$

However, by Lemma 3, $\mathcal{J}^*_\beta$ is strictly decreasing in $\beta \in [0,\bar{\beta}]$. Therefore, we must have

$$\mathcal{J}^*_{\beta'} > \mathcal{J}^*_\beta, \tag{32}$$

which leads to a contradiction. Hence, no such $\tilde{\pi}$ exists, and $\pi^*_\beta \in \mathcal{P}$ and $(F^*_\beta, \mathcal{J}^*_\beta) \in \mathcal{C}$.

It remains to show that $F^*_\beta = \beta$ and $(\beta, \mathcal{J}^*_\beta) \in \mathcal{C}$. Suppose, to the contrary, that $F^*_\beta < \beta$, i.e., the optimal performance is achieved with a smaller budget. Then for a budget $\beta'' = F^*_\beta < \beta$, Lemma 3 implies that $\mathcal{J}^*_{\beta''} > \mathcal{J}^*_\beta$, which contradicts optimality. Therefore, $F^*_\beta = \beta$, and the lemma follows. ■

Lemma 4 concludes the proof of Theorem 1.

Remark 1: Note, however, that the exact computation of the Pareto frontier based on Theorem 1 is numerically intractable, as it requires solving infinitely many constrained optimization problems across all possible budgets. Convexity suggests that convex combinations of some Pareto points can approximate the frontier (see the randomized mixture policy in (24)). Nevertheless, this approach is conservative, inefficient, and provides only a coarse characterization of the value of communication. The next section presents more structural properties of the frontier that provide further insights into efficient algorithm design.

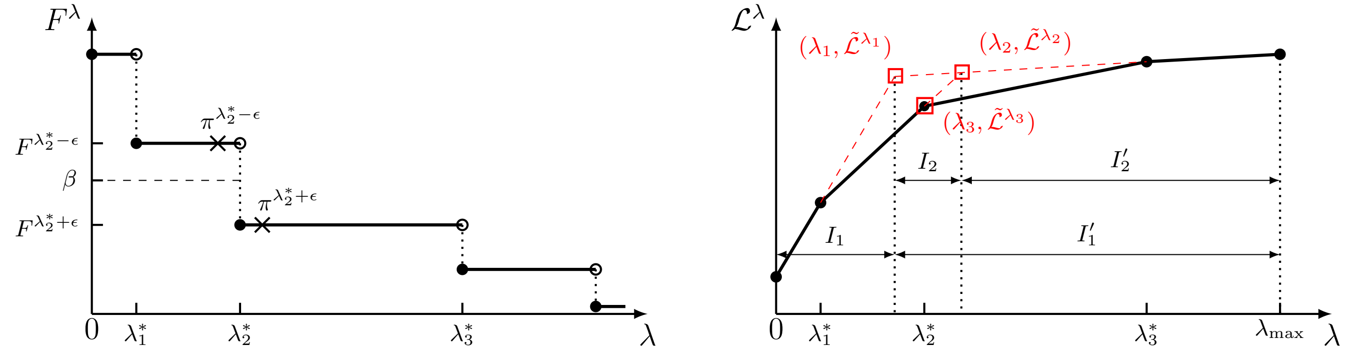
B. Piecewise Linearity of Pareto Frontier

This section examines the Lagrangian formulation. A central result is Theorem 2, which shows that the Pareto frontier is differentiable almost everywhere and, somewhat surprisingly, *piecewise linear*. Each corner point on this frontier corresponds to a stationary deterministic λ-optimal policy, and the value of communication is then characterized by a *finite* set of Lagrange multipliers. These findings will soon prove useful in computing the Pareto frontier in Section VI.

To establish the theorem, we begin by presenting three key propositions concerning λ-optimal policies.

Proposition 1: For any $\lambda > 0$, every λ-optimal policy π^λ is Pareto-optimal. The Pareto frontier is characterized by

$$\mathcal{C} = \big\{(F^\lambda, \mathcal{J}^\lambda) : \pi^\lambda \in \Pi^\lambda, \lambda > 0\big\}, \tag{33}$$



(a) F^λ is weakly decreasing and piecewise constant.

(b) $\mathcal{L}^\lambda$ is weakly increasing, continuous, piecewise linear, and concave.

Fig. 4. Illustration of the cost functions (solid lines) in the Lagrangian formulation and the search for breakpoints.

where Π^λ denotes the set of λ-optimal policies for a given λ. The set of Pareto-optimal policies is given by

$$\mathcal{P} = \{\pi^\lambda \in \Pi^\lambda : \lambda > 0\}. \tag{34}$$

Proof: Suppose, to the contrary, that π^λ is not Pareto-optimal. Then there exists a policy $\hat{\pi} \in \Pi$ such that

$$\mathcal{J}(\hat{\pi}) \leq \mathcal{J}^\lambda \text{ and } F(\hat{\pi}) \leq F^\lambda$$

with at least one strict inequality. This yields

$$\mathcal{L}(\hat{\pi}) = \mathcal{J}(\hat{\pi}) + \lambda F(\hat{\pi}) < \mathcal{L}^\lambda,$$

which contradicts λ-optimality. This proves the result. ■

The following result shows how to find *stationary* and *deterministic* λ-optimal policies.

Proposition 2 ([4, Theorem 1]): For any $\lambda > 0$, there exists a bounded value function V satisfying Bellman's equation

$$\mathcal{L}^\lambda + V^*(s) = \min_u \{l(s,u) + \mathbb{E}[V^*(s')|s,u]\},\ s \in \mathcal{S}, \tag{35}$$

where $l(s,u) = c(s) + \lambda u$. Define a sequence $\{V^n(s)\}$ with $V^0(s) = 0$ and, for $n = 1, 2, \ldots$,

$$Q^n(s,u) = l(s,u) + \mathbb{E}\big[V^{n-1}(s')|s,u\big], \tag{36a}$$
$$\tilde{V}^n(s) = \min_u\{Q^n(s,u)\}, \tag{36b}$$
$$V^n(s) = \tilde{V}^n(s) - \tilde{V}^n(s_{\text{ref}}), \tag{36c}$$

where s_{ref} is an arbitrarily chosen reference state. Then, the sequence $\{V^n\}$ converges to V^* as $n \to \infty$, and any policy that attains the minimum in (35) is λ-optimal.

Remark 2: Note that it is incorrect to conclude that solving Bellman's equation (35) for all $\lambda > 0$ allows one to trace the entire Pareto frontier. The reasons are twofold:

- For a given λ, there may exist infinitely many λ-optimal policies, including randomized or nonstationary ones, whereas (35) yields only a single stationary deterministic policy.
- The functions $\mathcal{J}^\lambda$ and F^λ are discrete-valued (see Proposition 3 below), offering only countably many isolated Pareto-optimal points.

Proposition 3: F^λ ($\mathcal{J}^\lambda$) is piecewise constant and weakly decreasing (increasing) in λ, and $\mathcal{L}^\lambda$ is continuous, weakly increasing, piecewise-linear, and concave (see Figure 4).

Proof: A proof of this result can be pieced together from [9, Proposition 3], [59, Lemmas 3.1-3.2], and [60, Lemmas 3.4-3.5]. For completeness and self-containment, we provide a consolidated proof in the Appendix. ■

We now show how to trace the entire Pareto frontier using these stationary deterministic λ-optimal policies.

Denote by Λ the ordered set of breakpoints on F^λ, where each element λ_i is the ith breakpoint satisfying

$$F^{\lambda_i - \epsilon} > F^{\lambda_i + \epsilon}$$

for any arbitrarily small $\epsilon > 0$.

Lemma 5: Λ is a finite set.

Proof: Let $\kappa = \min_i |F^{\lambda_{i+1}} - F^{\lambda_i}|$. From Proposition 3, we know that F^λ is discrete-valued. Then we have $\kappa > 0$ and

$$|\Lambda| \leq \frac{F^0 - F^\infty}{\kappa} \leq \frac{1}{\kappa} < \infty,$$

which proves the claim. ■

Theorem 2: For any breakpoint $\lambda_i \in \Lambda$ and any arbitrarily small constant $\epsilon > 0$, the randomized mixture policy

$$\pi_\theta^{\lambda_i} = \theta \pi^{\lambda_i - \epsilon} + (1-\theta)\pi^{\lambda_i + \epsilon},\ \theta \in [0,1] \tag{37}$$

is Pareto-optimal at the budget

$$\hat{\beta}_\theta = \theta F^{\lambda_i - \epsilon} + (1-\theta) F^{\lambda_i + \epsilon}.$$

Moreover, the Pareto frontier is piecewise linear and consists of line segments connecting the points $(F^{\lambda_i-\epsilon}, \mathcal{J}^{\lambda_i-\epsilon})$ and $(F^{\lambda_i+\epsilon}, \mathcal{J}^{\lambda_i+\epsilon})$ for all $\lambda_i \in \Lambda$. The value of communication on this line segment is

$$\eta = \lim_{\epsilon \to 0^+} \frac{\mathcal{J}^{\lambda_i+\epsilon} - \mathcal{J}^{\lambda_i-\epsilon}}{F^{\lambda_i-\epsilon} - F^{\lambda_i+\epsilon}} = \lambda_i. \tag{38}$$

Proof: We first show that $\pi_\theta^{\lambda_i}$, $\theta \in [0,1]$, is constrained-optimal (and thus by Theorem 1, Pareto-optimal) at budget $\hat{\beta}_\theta$, i.e., $\mathcal{J}(\pi_\theta^{\lambda_i}) = \mathcal{J}^*_{\hat{\beta}_\theta}$. Suppose, to the contrary, that $\pi_\theta^{\lambda_i}$ is not constrained-optimal. Then there exists a constrained-optimal policy $\pi^*_{\hat{\beta}_\theta}$ such that

$$\mathcal{J}^*_{\hat{\beta}_\theta} < \mathcal{J}(\pi_\theta^{\lambda_i}). \tag{39}$$

This leads to the following inequality

$$\mathcal{L}^{\lambda_i}(\pi^*_{\hat{\beta}_\theta}) = \mathcal{J}(\pi^*_{\hat{\beta}_\theta}) + \lambda_i F(\pi^*_{\hat{\beta}_\theta})$$

$$
\begin{aligned}
&\overset{\text{(a)}}{=} \mathcal{J}^*_{\hat{\beta}_\theta} + \lambda_i \hat{\beta}_\theta \\
&\overset{\text{(b)}}{<} \mathcal{J}(\pi_\theta^{\lambda_i}) + \lambda_i \hat{\beta}_\theta \\
&= \mathcal{J}\big(\theta\pi^{\lambda_i-\epsilon} + (1-\theta)\pi^{\lambda_i+\epsilon}\big) \\
&\quad + \lambda_i\big(\theta F^{\lambda_i-\epsilon} + (1-\theta)F^{\lambda_i+\epsilon}\big) \\
&\overset{\text{(c)}}{=} \theta\mathcal{J}^{\lambda_i-\epsilon} + (1-\theta)\mathcal{J}^{\lambda_i+\epsilon} \\
&\quad + \lambda_i\big(\theta F^{\lambda_i-\epsilon} + (1-\theta)F^{\lambda_i+\epsilon}\big) \\
&= \theta\big(\mathcal{J}^{\lambda_i-\epsilon} + (\lambda_i-\epsilon)F^{\lambda_i-\epsilon}\big) \\
&\quad + (1-\theta)\big(\mathcal{J}^{\lambda_i+\epsilon} + (\lambda_i+\epsilon)F^{\lambda_i+\epsilon}\big) + o(\epsilon) \\
&= \theta\mathcal{L}^{\lambda_i-\epsilon} + (1-\theta)\mathcal{L}^{\lambda_i+\epsilon} + o(\epsilon),
\end{aligned} \tag{40}
$$

where (a) follows from Lemma 4, which states that every constrained-optimal policy π^*_β, $\beta \in [0, \bar{\beta}]$, satisfies $F^*_\beta = \beta$; (b) follows from the contradiction assumption in (39); and (c) follows from the same argument as in (26). In the final expression, we have collected the residual terms into

$$
o(\epsilon) = \theta\epsilon F^{\lambda_i-\epsilon} - (1-\theta)\epsilon F^{\lambda_i+\epsilon}.
$$

Taking the limit as $\epsilon \to 0^+$ and using the continuity of $\mathcal{L}^\lambda$ (see Proposition 3), we obtain

$$
\begin{aligned}
\mathcal{L}^{\lambda_i}(\pi^*_{\hat{\beta}_\theta}) &< \lim_{\epsilon\to 0^+} \theta\mathcal{L}^{\lambda_i-\epsilon} + (1-\theta)\mathcal{L}^{\lambda_i+\epsilon} + o(\epsilon) \\
&= \mathcal{L}^{\lambda_i}.
\end{aligned} \tag{41}
$$

This contradicts the λ_i-optimality of π^{λ_i}. Therefore, $\pi_\theta^{\lambda_i}$ is constrained-optimal by construction.

Furthermore, by Proposition 1, both $\pi^{\lambda_i-\epsilon}$ and $\pi^{\lambda_i+\epsilon}$ are Pareto-optimal. Hence, the randomized mixture policy $\pi_\theta^{\lambda_i}$ forms a line segment between two Pareto-optimal points $(F^{\lambda_i-\epsilon}, \mathcal{J}^{\lambda_i-\epsilon})$ and $(F^{\lambda_i+\epsilon}, \mathcal{J}^{\lambda_i+\epsilon})$, as illustrated in Figure 3b. Because all points on this segment are Pareto-optimal, the Pareto frontier is piecewise linear, and the system exhibits a constant communication value along this segment.

It remains to show that the limit in (38) exists and equals the Lagrange multiplier λ_i. The proof relies on the fundamental inequality that reads

$$
\begin{aligned}
2\epsilon F^{\lambda_i-\epsilon} &= \mathcal{J}(\pi^{\lambda_i-\epsilon}) + (\lambda_i+\epsilon)F(\pi^{\lambda_i-\epsilon}) \\
&\quad - \big(\mathcal{J}(\pi^{\lambda_i-\epsilon}) + (\lambda_i-\epsilon)F(\pi^{\lambda_i-\epsilon})\big) \\
&= \mathcal{L}^{\lambda_i+\epsilon}(\pi^{\lambda_i-\epsilon}) - \mathcal{L}^{\lambda_i-\epsilon} \\
&\overset{\text{(d)}}{\geq} \mathcal{L}^{\lambda_i+\epsilon} - \mathcal{L}^{\lambda_i-\epsilon} \\
&\overset{\text{(e)}}{\geq} \mathcal{L}^{\lambda_i+\epsilon} - \mathcal{L}^{\lambda_i-\epsilon}(\pi^{\lambda_i+\epsilon}) \\
&= \mathcal{J}(\pi^{\lambda_i+\epsilon}) + (\lambda_i+\epsilon)F(\pi^{\lambda_i+\epsilon}) \\
&\quad - \big(\mathcal{J}(\pi^{\lambda_i+\epsilon}) + (\lambda_i-\epsilon)F(\pi^{\lambda_i+\epsilon})\big) \\
&= 2\epsilon F^{\lambda_i+\epsilon} \geq 0
\end{aligned} \tag{42}
$$

for any $\lambda_i > 0$ and $\epsilon > 0$, where (d) holds because the $(\lambda_i-\epsilon)$-optimal policy, $\pi^{\lambda_i-\epsilon}$, is not necessarily optimal at multiplier $\lambda_i+\epsilon$, and similarly (e) holds because $\pi^{\lambda_i+\epsilon}$ is not necessarily optimal at multiplier $\lambda_i-\epsilon$. Removing redundant terms gives

$$
2\epsilon F^{\lambda_i-\epsilon} \geq \mathcal{L}^{\lambda_i+\epsilon} - \mathcal{L}^{\lambda_i-\epsilon} \geq 2\epsilon F^{\lambda_i+\epsilon} \geq 0. \tag{43}
$$

Substituting

$$
\mathcal{L}^{\lambda_i+\epsilon} = \mathcal{J}^{\lambda_i+\epsilon} + (\lambda_i+\epsilon)F^{\lambda_i+\epsilon},
$$

$$
\mathcal{L}^{\lambda_i-\epsilon} = \mathcal{J}^{\lambda_i-\epsilon} + (\lambda_i-\epsilon)F^{\lambda_i-\epsilon}
$$

in (43) gives

$$
\begin{aligned}
(\lambda_i+\epsilon)\big(F^{\lambda_i-\epsilon} - F^{\lambda_i+\epsilon}\big) &\geq \mathcal{J}^{\lambda_i+\epsilon} - \mathcal{J}^{\lambda_i-\epsilon} \\
&\geq (\lambda_i-\epsilon)\big(F^{\lambda_i-\epsilon} - F^{\lambda_i+\epsilon}\big).
\end{aligned}
$$

Dividing all terms by $F^{\lambda_i-\epsilon} - F^{\lambda_i+\epsilon} > 0$, we obtain

$$
\lambda_i + \epsilon \geq \frac{\mathcal{J}^{\lambda_i+\epsilon} - \mathcal{J}^{\lambda_i-\epsilon}}{F^{\lambda_i-\epsilon} - F^{\lambda_i+\epsilon}} \geq \lambda_i - \epsilon. \tag{44}
$$

Taking a limit as $\epsilon \to 0^+$ yields the result in (38). ■

Remark 3: The Lagrange multiplier serves not only as a measure of resource consumption or a tool for solving constrained problems, but, more importantly, as a quantitative measure of the value of communication. This underscores the role of Pareto optimization in semantic communications, a facet that has not been sufficiently recognized in the literature.

Remark 4: The randomized mixture policy (37) offers a provably low-complexity representation of the Pareto frontier. It suffices to compute a finite set of breakpoints and store the associated deterministic λ-optimal policy pairs. Notably, this mixture policy is easy to implement: when the operating budget changes, one selects the appropriate policy pairs from the lookup table and adjusts the interpolation coefficient θ.

Remark 5: Our analysis is based on the principle that *allowing more transmissions never worsens the best achievable performance*[1]. This tradeoff is common in communication and control systems. Therefore, the structural results are readily applicable to a broader class of bi-objective MDPs that share similar tradeoffs.

VI. Computing the Pareto Frontier

We now leverage the structural properties established in Section V to develop the Structured Pareto Line Intersection Tracing (SPLIT) algorithm, which computes the exact Pareto frontier $\mathcal{C}_\mathcal{B}$ over any budget interval $\mathcal{B} = [\beta_\text{L}, \beta_\text{H}]$.

SPLIT identifies a set $\Lambda_\mathcal{B}$ of breakpoints of F^λ over $\mathcal{B}$, i.e.,

$$
\Lambda_\mathcal{B} = \big\{\lambda_k^* \in \Lambda : \beta_\text{L} \leq F^{\lambda_k^*+\epsilon} < F^{\lambda_k^*-\epsilon} \leq \beta_\text{H}\big\}.
$$

Let $\Pi_\mathcal{B}$ denote the set of all policy pairs corresponding to the breakpoints in $\Lambda_\mathcal{B}$, where

$$
\Pi_\mathcal{B} = \big\{(\pi^{\lambda_k^*-\epsilon}, \pi^{\lambda_k^*+\epsilon}) : \lambda_k^* \in \Lambda_\mathcal{B}\big\}.
$$

Figure 4a and Figure 3b illustrate how to construct the frontier segment associated with a given breakpoint λ_k^*:

a. Compute two stationary deterministic policies, $\pi^{\lambda_k^*-\epsilon}$ and $\pi^{\lambda_k^*+\epsilon}$, by solving Bellman's equation (35) for the multipliers $\lambda_k^* - \epsilon$ and $\lambda_k^* + \epsilon$, respectively.
b. Then, by Theorem 2, the randomized mixture policy

$$
\pi_\theta^{\lambda_k^*} = \theta\pi^{\lambda_k^*-\epsilon} + (1-\theta)\pi^{\lambda_k^*+\epsilon}, \ \theta \in [0,1]
$$

traces out the frontier for budgets $\beta \in [F^{\lambda_k^*+\epsilon}, F^{\lambda_k^*-\epsilon}]$.

The pseudocode of SPLIT is provided in Algorithm 1. It consists of two main components: (1) an Intersection Search

[1]Note that allowing more transmissions does not necessarily mean mandating them. In energy-harvesting systems [8] or wearing channels [27], frequent transmissions can impair the hardware and degrade overall performance.

Algorithm 1 `SPLIT`($\mathcal{B}$)

Input: Operating budgets $\mathcal{B} = [\beta_{\mathrm{L}}, \beta_{\mathrm{H}}]$
Output: Exact frontier $\mathcal{C}_{\mathcal{B}}$

$(\Lambda_{\mathcal{B}}, \Pi_{\mathcal{B}}) \leftarrow$ `Insec`$(\mathcal{B})$ ▷ Breakpoints
for each $\lambda_k^* \in \Lambda_{\mathcal{B}}$ **do** ▷ Frontier
  $\pi_\theta^{\lambda_k^*} \leftarrow \theta\, \pi^{\lambda_k^* - \epsilon} + (1-\theta)\, \pi^{\lambda_k^* + \epsilon},\ \theta \in [0,1]$
end for

Algorithm 2 `Insec`($\mathcal{B}$)

Input: $\mathcal{B} = [\beta_{\mathrm{L}}, \beta_{\mathrm{H}}]$
Output: $(\Lambda_{\mathcal{B}}, \Pi_{\mathcal{B}})$

INITIALIZE: $I_0 \leftarrow [0, \lambda_{\max}]$, $\Lambda_{\mathcal{B}} \leftarrow \emptyset$, $\Pi_{\mathcal{B}} \leftarrow \emptyset$
BREAKPOINTS: $(\Lambda_{\mathcal{B}}, \Pi_{\mathcal{B}}) \leftarrow$ RECURSE(I_0)
function RECURSE(I_n)
  $\lambda_n \leftarrow$ `Update`(I_n), $\pi^{\lambda_n} \leftarrow$ `SPI`(λ_n) ▷ Intersect
  if $\mathcal{L}^{\lambda_n} = \tilde{\mathcal{L}}^{\lambda_n}$ **and** $F^{\lambda_n} \in \mathcal{B}$ **then** ▷ Breakpoint
    $\pi^{\lambda_n - \epsilon} \leftarrow$ `SPI`$(\lambda_n - \epsilon)$
    $\pi^{\lambda_n + \epsilon} \leftarrow$ `SPI`$(\lambda_n + \epsilon)$
    $\Lambda_{\mathcal{B}} \leftarrow \Lambda_{\mathcal{B}} \cup \{\lambda_n\}$
    $\Pi_{\mathcal{B}} \leftarrow \Pi_{\mathcal{B}} \cup \{(\pi^{\lambda_n - \epsilon}, \pi^{\lambda_n + \epsilon})\}$
  else ▷ Split Interval
    $I_{n+1} \leftarrow [\lambda_n^-, \lambda_n]$, $I'_{n+1} \leftarrow [\lambda_n, \lambda_n^+]$
    RECURSE(I_{n+1}), RECURSE(I'_{n+1})
  end if
end function

Algorithm 3 `SPI`(λ)

Input: Lagrange multiplier λ
Output: λ-optimal policy π^λ

INITIALIZE: $V^0(s) \leftarrow 0$, $\pi_0^\lambda(s) \leftarrow \infty$, $s_{\mathrm{ref}} \leftarrow 0$
for $n = 1, 2, \ldots$ **do**
  $\mathcal{L}_n^\lambda + V^n(s) \leftarrow Q^n\big(s, \pi_n^\lambda(s)\big)\,,\ s \in \bar{\mathcal{S}}$ ▷ Evaluate
  for each (x, z, θ) with $x \neq g(z, \theta)$ **do**
    for $\delta = 1 \ldots \Delta_{\max}$ **do**
      $\pi_{n+1}^\lambda(s) \leftarrow \arg\min_u Q^n(s, u)$ ▷ Improve
      if $\pi_{n+1}^\lambda(s) = 1$ **then** ▷ Exploit
        $\pi_{n+1}^\lambda(s') \leftarrow 1$ for all $\delta' > \delta$
        break
      end if
    end for
  end for
  if $\pi_{n+1}^\lambda = \pi_n^\lambda$ **then** ▷ Terminate
    return $\pi^\lambda \leftarrow \pi_n^\lambda$
  end if
end for

(`Insec`) for identifying the set of breakpoints $\Lambda_{\mathcal{B}}$, given in Algorithm 2, and (2) a Structured Policy Iteration (`SPI`) for computing λ-optimal policies, given in Algorithm 3. These components are discussed in detail in the following sections.

A. Intersection Search

The `Insec` algorithm is built on the prior work of Luo and Pappas [9], which introduced an efficient multiplier update method for solving CMDPs. In this work, we develop this idea further to identify all breakpoints with *linear* complexity.

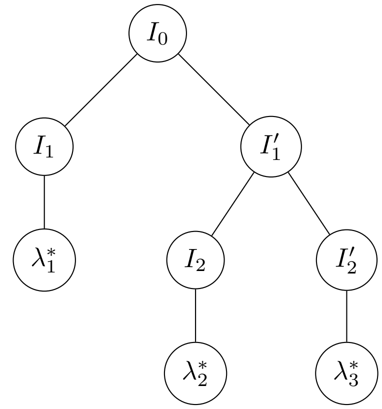


Fig. 5. Binary tree representation of `Insec` for the case in Figure 4.

We leverage two key properties of $\mathcal{L}^\lambda$: (i) $\mathcal{L}^\lambda$ is piecewise-linear and concave, and (ii) any breakpoint on F^λ corresponds to a corner point on $\mathcal{L}^\lambda$, as illustrated in Figure 4. The main idea is to recursively split the initial multiplier search interval I_0 into minimal subintervals, each of which contains exactly one breakpoint. The algorithm proceeds as follows:

a. *Initialization:* Choose a sufficiently large $\lambda_{\max}$ and set the initial interval $I_0 = [\lambda_0^-, \lambda_0^+] = [0, \lambda_{\max}]$.
b. *Intersection:* Compute the intersection point of two tangents to $\mathcal{L}^\lambda$: one passing through $(\lambda_n^-, \mathcal{L}^{\lambda_n^-})$ with slope $F^{\lambda_n^-}$, and another through $(\lambda_n^+, \mathcal{L}^{\lambda_n^+})$ with slope $F^{\lambda_n^+}$. The intersection point $(\lambda_n, \tilde{\mathcal{L}}^{\lambda_n})$, illustrated by the squares in Figure 4b, is given by
$$\lambda_n = \texttt{Update}(I_n) = \frac{\mathcal{J}^{\lambda_n^+} - \mathcal{J}^{\lambda_n^-}}{F^{\lambda_n^-} - F^{\lambda_n^+}}, \tag{45a}$$
$$\tilde{\mathcal{L}}^{\lambda^n} = \mathcal{J}^{\lambda_n^-} + \lambda_n F^{\lambda_n^-} \geq \mathcal{L}^{\lambda_n}. \tag{45b}$$
Apply SPI to solve Bellman's equation (35) at multiplier λ_n and obtain π^{λ_n}.
c. *Recursion:* If the intersection point $(\lambda_n, \tilde{\mathcal{L}}^{\lambda_n})$ lies on $\mathcal{L}^\lambda$, i.e., $\tilde{\mathcal{L}}^{\lambda_n} = \mathcal{L}^{\lambda_n}$, then λ_n is identified as a breakpoint (e.g., $(\lambda_3, \tilde{\mathcal{L}}^{\lambda_3})$ in Figure 4b). Otherwise, split the current interval into $I_{n+1} = [\lambda_n^-, \lambda_n]$ and $I'_{n+1} = [\lambda_n, \lambda_n^+]$, and repeat Step (b) for each subinterval.
d. *Termination:* The algorithm terminates when each split interval brackets exactly one breakpoint.

Theorem 3: The complexity of `Insec` is $\mathcal{O}(|\Lambda|)$.

Proof: We analyze the algorithm using the binary tree representation depicted in Figure 5, where:

- Each leaf node corresponds to a unique breakpoint;
- Each internal node corresponds to an intersection point that results in a split of the interval.

`Insec` maintains that each active interval brackets at least one breakpoint, guaranteed by the concavity of $\mathcal{L}^\lambda$. Each recursive call either identifies a breakpoint or splits the interval, strictly reducing the search space. In the corresponding binary tree, each internal node has either two internal children or one leaf child. The search for breakpoints is thus equivalent to traversing the tree until all leaves have been visited. Since the total number of breakpoints $|\Lambda|$ is finite, the tree depth is bounded, and the algorithm is guaranteed to converge in a finite number of calls.

The binary tree has $|\Lambda|$ leaves and $2|\Lambda|-1$ internal nodes. The number of recursive calls equals the number of internal nodes. Therefore, the complexity of `Insec` is $\mathcal{O}(|\Lambda|)$. ■

Remark 6: `Insec` identifies *all* breakpoints (i.e., optimality) with linear complexity (i.e., efficiency). According to the binary tree representation in Figure 5, `Insec` naturally enables parallel computation. In contrast, classical approaches to CMDPs, such as the bisection search [9], find only a *single* breakpoint with complexity $\mathcal{O}\big(\log(\lambda_{\max}/\epsilon)\big)$, which is higher than that of `Insec` when high precision is required.

B. Structured Policy Iteration

This section presents the `SPI` algorithm for computing stationary deterministic λ-optimal policies. Although conceptually simpler than the constrained and Pareto formulations, the Lagrangian approach remains computationally challenging for two main reasons: (1) the state space $\mathcal{S}$ is countably infinite, making classical dynamic programming methods inapplicable; and (2) most existing dynamic programming methods, such as relative value iteration (RVI) [46, Ch. 8.5], solve Bellman's equation (35) in an unstructured manner; that is, they update the recursions in (36) for all state-action pairs without exploiting the structure of the optimal policy.

These challenges have been effectively addressed by Luo and Pappas in [4], [12], [13]. We now summarize two key results from this line of work concerning the *asymptotic optimality* and *switching structure* of λ-optimal policies.

Proposition 4: Define $\tilde{\Delta}_t = \min\{\Delta_t, \Delta_{\max}\}$ and $\tilde{\Theta}_t = \min\{\Theta_t, \Theta_{\max}\}$. Then the Lagrangian problem defined on the truncated state space $\bar{\mathcal{S}} = \mathcal{X}^2 \times \{0,1,\dots,\Delta_{\max}\} \times \{0,1,\dots,\Theta_{\max}\}$ converges to the original problem exponentially fast as the truncation bounds $\Delta_{\max}$ and $\Theta_{\max}$ increase.

Then we shall feel safe to truncate the state space with appropriately chosen bounds.

Proposition 5: For any given $\lambda > 0$, the λ-optimal policy has a switching structure; that is, $\pi^{\lambda}(x,z,\theta,\delta)=1$ only when the δ exceeds a threshold $\delta^*_{x,z,\theta}$. Formally,

$$\pi^{\lambda}(x,z,\theta,\delta) = \begin{cases} 1, & x \neq g(z,\theta), \delta \geq \delta^*_{x,z,\theta}, \\ 0, & \text{otherwise.} \end{cases} \tag{46}$$

`SPI` leverages the switching structure to reduce computation overhead. The algorithm proceeds as follows:

a. *Initialization:* Select an initial switching policy π^{λ}_0 with thresholds $\delta_{x,z,\theta} = \infty$. Choose a reference state s_{ref} and sufficiently large truncation bounds $\Delta_{\max}$ and $\Theta_{\max}$. Then, for $n = 0,1,2,\dots$ do the following:
b. *Evaluation:* Find a scalar $\mathcal{L}^{\lambda}_n$ and a vector V^n by solving
$$\mathcal{L}^{\lambda}_n + V^n(s) = Q^n\big(s, \pi^{\lambda}_n(s)\big),\ s \in \bar{\mathcal{S}}$$
such that $V^n(s_{\text{ref}}) = 0$, where Q^n is defined in (36).
c. *Improvement:* For each (x,z,θ), $x \neq g(z,\theta)$, update
$$\pi^{\lambda}_{n+1}(s) = \operatorname*{arg\,min}_{u\in\mathcal{U}} Q^n\big(s,u\big)$$
in the order $\delta = 1,2,\dots,\Delta_{\max}$. When $\pi^{\lambda}_{n+1}(s)=1$, the optimal action for all subsequent states $s' = (x,z,\theta,\delta')$ with $\delta' > \delta$ is to transmit without further computation.
d. *Termination:* If $\pi^{\lambda}_{n+1} = \pi^{\lambda}_n$, the algorithm terminates with $\mathcal{L}^{\lambda} = \mathcal{L}^{\lambda}_n$ and $\pi^{\lambda} = \pi^{\lambda}_n$; otherwise increase $n = n+1$ and return to Step (b).

VII. Numerical Results

This section presents numerical results to corroborate the theoretical findings on the Pareto frontier. For comparison, we consider both slowly- and rapidly-evolving sources with $|\mathcal{X}| = 3$ states and transition matrices

$$Q_{\text{slow}} = \begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.3 & 0.6 & 0.1 \\ 0.2 & 0.1 & 0.7 \end{bmatrix},\ Q_{\text{rapid}} = \begin{bmatrix} 0.4 & 0.3 & 0.3 \\ 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.3 \end{bmatrix}.$$

Moreover, we examine both prioritized and non-prioritized sources. In the prioritized case, state 1 is designated as an alarm state, and the urgency of lasting impact is modeled as

$$g_{i,j}(\delta) = \begin{cases} 4.8e^{0.55\delta} + 1.2, & \text{missed alarms } (i=1, j\neq 1), \\ 2.4e^{0.35\delta} + 0.6, & \text{false alarms } (i\neq 1, j=1), \\ 1.2e^{0.15\delta} + 0.3, & \text{normal errors } (i\neq j, i,j\neq 1), \\ 0, & \text{synced states } (i=j). \end{cases}$$

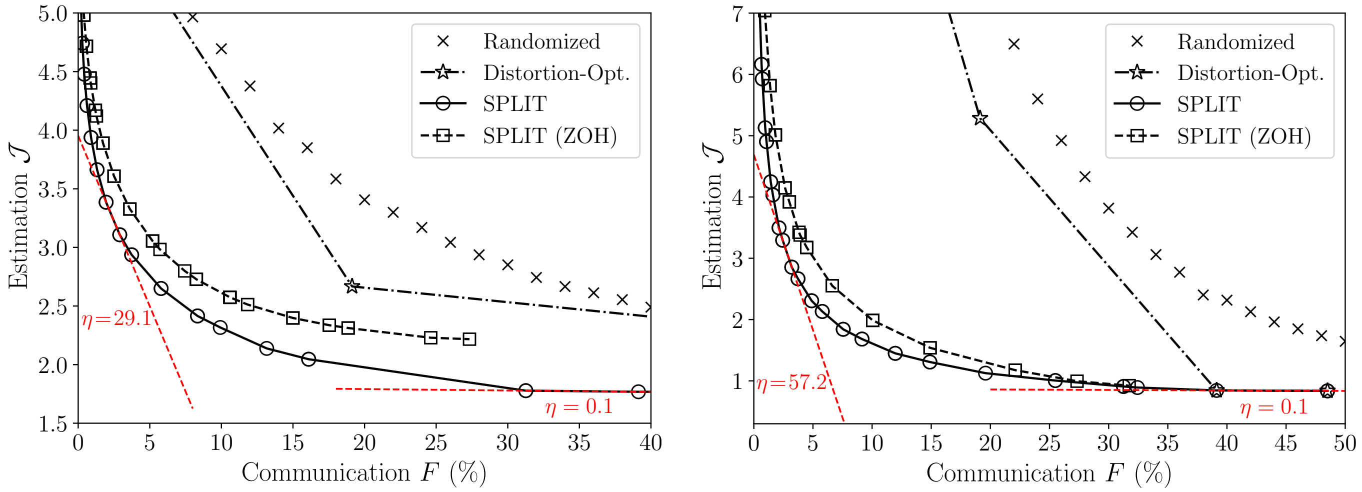
In the non-prioritized case, all errors are treated equally, i.e.,

$$g_{i,j}(\delta) = 1.2e^{0.55\delta} + 0.3, \quad i \neq j.$$

The remaining parameters are set as follows: packet erasure probability $p_f = 0.3$, operating budgets $\mathcal{B} = [0, 0.5]$, maximum multiplier $\lambda_{\max} = 10^5$, search tolerance $\epsilon = 10^{-3}$, and truncation bounds $\Delta_{\max} = 50$ and $\Theta_{\max} = 20$.

Figure 6 illustrates the Pareto frontiers of the slowly-evolving source obtained under different estimators. Each point on the frontier represents a Pareto-optimal operating point that balances estimation performance against communication cost. The frontier is piecewise-linear and convex, with squares (circles) marking the corner points obtained using SPLIT with the ZOH (MAP) estimator. Notably, the marginal value of communication diminishes when operating at relatively high frequencies (i.e., $F > 20\%$). Conversely, when communication is scarce (i.e., $F < 10\%$), careful calibration of the operating point becomes crucial. Moreover, the distortion-optimal and randomized policies (both using MAP) exhibit significant performance degradation, particularly in the low-frequency region. Interestingly, the frontier for the distortion-optimal policy is characterized by only a few breakpoints, resulting in low computation overhead but reduced performance granularity.

For a given communication budget (i.e., drawing a vertical line), the MAP estimator achieves better estimation performance than the ZOH estimator. For the prioritized source (Figure 6a), the two estimators achieve comparable performance only when communication is extremely scarce, in which case the receiver has little information about the source to make reliable predictions. Observe from Figure 6b that, for the non-prioritized source, the two estimators also achieve similar performance when operating at a relatively high frequency



(a) Slowly-evolving prioritized source.

(b) Slowly-evolving non-prioritized source.

Fig. 6. Pareto frontier of different estimators. Squares (circles) denote the corner points on the frontier obtained by SPLIT with the ZOH (MAP) estimator.

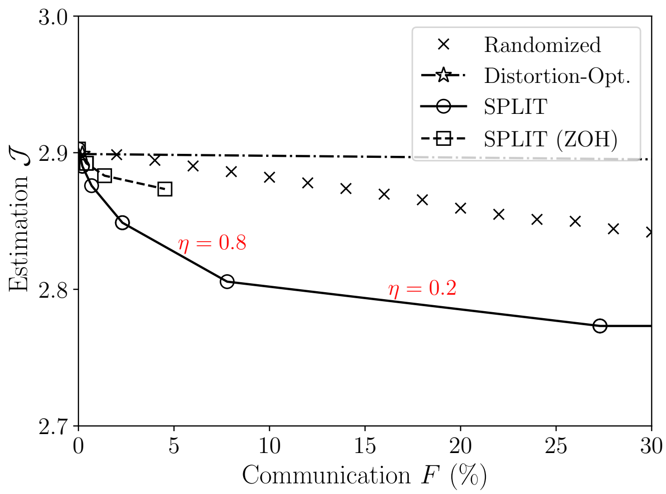


Fig. 7. Pareto frontier of the rapidly-evolving prioritized source.

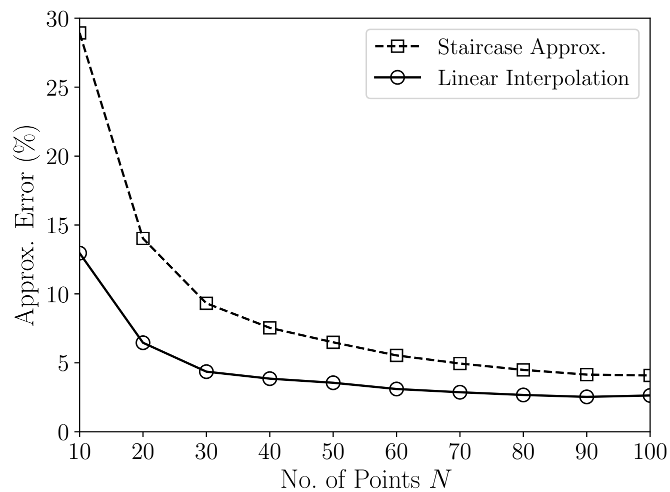


Fig. 8. Relative approximation error of different approximation methods.

($F > 25\%$). This is because the estimation errors across different states are equally important; hence, the benefit of incorporating prior information in the MAP estimator becomes negligible under frequent updates. The randomized and distortion-optimal policies, however, suffer noticeable performance degradation due to uninformative transmission timing.

On the other hand, for a given performance level (i.e., drawing a horizontal line), the optimal policy with the MAP estimator significantly reduces communication by exploiting the timing of informative updates. For instance, in Figure 6a, to achieve a target performance of $\mathcal{J} = 2.5$, the optimal policy transmits at a frequency of only $F = 8\%$. In contrast, the distortion-optimal and randomized policies consume 40% of the communication resources. This highlights the effectiveness of exploiting the value of information in such systems.

Figure 7 illustrates the frontiers of the rapidly-evolving prioritized source. In contrast to the slowly-evolving source, the frontier becomes noticeably flat, with a much smaller value of communication. In other words, additional transmissions provide limited marginal improvement because the prediction horizon is inherently short: rapid state transitions reduce the usefulness of both memory and timing information. In such cases, communication is not sufficient for accurate estimation, but it helps to prevent prolonged error periods.

We now discuss classical approaches for approximating the Pareto frontier. A common method is to traverse all Lagrange multipliers and solve the corresponding Bellman equation for all $\lambda > 0$. This approach is neither practical nor meaningful, since practical dynamic programming methods can identify only finitely many points on the frontier.

An alternative approach is to quantize the budget set $\mathcal{B}$ into N equal intervals and solve a total of N constrained optimization problems to obtain a coarse approximation of the frontier. The time complexity of this approach is

$$\mathcal{O}\Big(N \log_2\left(\lambda_{\max}/\epsilon\right) I_{\text{SPI}}\Big),$$

where I_{SPI} is the complexity of the `SPI`. We consider two approximation methods:

- *Staircase approximation:* Construct a step function from the computed Pareto points. While conservative, this approach prevents overestimation when the frontier shape is unknown.
- *Linear interpolation:* Connect the computed points to form a piecewise-linear convex frontier. This approach implicitly leverages the known convexity of the Pareto frontier to achieve a tighter approximation.

Figure 8 shows the relative approximation error of the above methods. Both require a considerably large N to achieve the desired accuracy. In contrast, the proposed SPLIT algorithm constructs the exact frontier with complexity $\mathcal{O}\big(|\Lambda| I_{\text{SPI}}\big)$, far outperforming the approximation methods in terms of both complexity and optimality.

VIII. Conclusion

This paper investigates the value of communication in goal-oriented semantic communications by examining the canonical problem of remote estimation of Markov sources. The problem is formulated as a bi-objective MDP, whose optimal solutions constitute the Pareto frontier in the communication-estimation objective plane. We then define the value of communication as the absolute slope of this frontier.

Our analysis reveals that the frontier is strictly decreasing, convex, and piecewise linear. Each corner point on the frontier corresponds to a stationary deterministic Pareto-optimal policy, and each linear segment is formed by a convex combination of two such neighboring policies. Moreover, the value of communication is governed by a finite collection of Lagrange multipliers. Building on these findings, we develop a fast and exact algorithm, SPLIT, to compute the entire frontier. The algorithm is provably optimal, with time complexity linear in the number of corner points. Importantly, the frontier is easy to implement: when the operating budget changes, one selects the appropriate policy pairs from the lookup table and adjusts the mixture coefficient.

The structural results developed in this paper can be applied to most existing semantics-aware metrics, as well as to a broader class of bi-objective MDPs characterized by a fundamental tradeoff: relaxing one objective never compromises the best achievable performance of the other.

Appendix

Proposition 3 is established using the following lemmas.

Lemma 6 (Monotonicity): $\mathcal{L}^\lambda$ and J^λ are weakly increasing in λ, while F^λ is weakly decreasing in λ.

Proof: For any $\lambda > 0$ and $\epsilon > 0$, let π^λ and $\pi^{\lambda+\epsilon}$ denote the λ-optimal and $(\lambda+\epsilon)$-optimal policies. Then, we have

$$\begin{aligned}\mathcal{L}^{\lambda+\epsilon} - \mathcal{L}^\lambda &= \mathcal{L}^{\lambda+\epsilon}(\pi^{\lambda+\epsilon}) - \mathcal{L}^\lambda(\pi^\lambda) \\ &\geq \mathcal{L}^{\lambda+\epsilon}(\pi^{\lambda+\epsilon}) - \mathcal{L}^\lambda(\pi^{\lambda+\epsilon}) \\ &= \big(\mathcal{J}(\pi^{\lambda+\epsilon}) + (\lambda+\epsilon)F(\pi^{\lambda+\epsilon})\big) \\ &\quad - \big(\mathcal{J}(\pi^{\lambda+\epsilon}) + \lambda F(\pi^{\lambda+\epsilon})\big) \\ &= \epsilon F^{\lambda+\epsilon} \geq 0. \end{aligned} \tag{47}$$

Hence, $\mathcal{L}^\lambda$ is weakly increasing in λ. Similarly,

$$\begin{aligned}\mathcal{L}^{\lambda+\epsilon} - \mathcal{L}^\lambda &\leq \mathcal{L}^{\lambda+\epsilon}(\pi^\lambda) - \mathcal{L}^\lambda(\pi^\lambda) \\ &= \big(\mathcal{J}(\pi^\lambda) + (\lambda+\epsilon)F(\pi^\lambda)\big) \\ &\quad - \big(\mathcal{J}(\pi^\lambda) + \lambda F(\pi^\lambda)\big) \\ &= \epsilon F^\lambda. \end{aligned} \tag{48}$$

Combining (47) and (48), we obtain

$$0 \leq \epsilon F^{\lambda+\epsilon} \leq \mathcal{L}^{\lambda+\epsilon} - \mathcal{L}^\lambda \leq \epsilon F^\lambda, \tag{49}$$

which implies that F^λ is weakly decreasing in λ.

It remains to show that $\mathcal{J}^\lambda$ is weakly increasing. Suppose, to the contrary, that $\mathcal{J}^\lambda$ is decreasing in λ. Then we must have

$$\mathcal{L}^\lambda = \mathcal{J}^\lambda + \lambda F^\lambda > \mathcal{J}^{\lambda+\epsilon} + \lambda F^{\lambda+\epsilon} = \mathcal{L}^\lambda(\pi^{\lambda+\epsilon}), \tag{50}$$

which contradicts the λ-optimality of π^λ. Therefore, $\mathcal{J}^\lambda$ is weakly increasing in λ. ■

Lemma 7 (Continuity and concavity): $\mathcal{L}^\lambda$ is concave and continuous on $(0, \infty)$.

Proof: For any $0 < \lambda_1 < \lambda_2$ and $0 \leq \theta \leq 1$, let $\lambda_\theta = \theta\lambda_1 + (1-\theta)\lambda_2$. Then we have

$$\begin{aligned}\mathcal{L}^{\lambda_\theta} &= \mathcal{J}^{\lambda_\theta} + \lambda_\theta F^{\lambda_\theta} \\ &= \mathcal{J}^{\lambda_\theta} + \big(\theta\lambda_1 + (1-\theta)\lambda_2\big)F^{\lambda_\theta} \\ &= \theta\big(\mathcal{J}^{\lambda_\theta} + \lambda_1 F^{\lambda_\theta}\big) + (1-\theta)\big(\mathcal{J}^{\lambda_\theta} + \lambda_2 F^{\lambda_\theta}\big) \\ &\geq \theta\mathcal{L}^{\lambda_1} + (1-\theta)\mathcal{L}^{\lambda_2}. \end{aligned} \tag{51}$$

This establishes the concavity of $\mathcal{L}^\lambda$ on $(0, \infty)$.

From (49) we obtain that for any $0 < \lambda_1, \lambda_2 < \infty$,

$$\begin{aligned}\big|\mathcal{L}^{\lambda_1} - \mathcal{L}^{\lambda_2}\big| &\leq \max\{F^{\lambda_1}, F^{\lambda_1}\}\big|\lambda_1 - \lambda_2\big| \\ &\leq \big|\lambda_1 - \lambda_2\big|.\end{aligned}$$

Hence $\mathcal{L}^\lambda$ is Lipschitz and therefore continuous on $(0, \infty)$. ■

Lemma 8 (Piecewise linearity): $\mathcal{L}^\lambda$ is piecewise linear, and $\mathcal{J}^\lambda$ and F^λ are piecewise constant.

Proof: We prove this by induction on the value iteration recursion in (36). Initialize

$$Q^0(s,u) = l^\lambda(s,u) = c(s) + \lambda u.$$

Clearly, $Q^0(s,u)$ is linear in λ. Since the pointwise minimum of linear functions is piecewise linear, both $\tilde{V}^0(s) = \min_u\{Q^0(s,u)\}$ and $V^0(s) = \tilde{V}^0(s) - \tilde{V}^0(s_{\text{ref}})$ are piecewise linear in λ for all s and u.

Assume as the induction hypothesis that $Q^n(s,u)$ is piecewise linear for all $s \in \mathcal{S}$. Because a weighted sum of piecewise linear functions remains piecewise linear, it follows that

$$Q^{n+1}(s,u) = l^\lambda(s,u) + \mathbb{E}\big[V^n(s')|s,u\big]$$

is piecewise linear in λ for all s and u. Since this property is preserved as $n \to \infty$, the optimal value functions $V^*(s)$ and $Q^*(s,u)$ are piecewise linear. Consequently,

$$\mathcal{L}^\lambda = \min_u Q^*(s,u) - V^*(s)$$

is piecewise linear in λ.

From (49), the right-hand derivative of $\mathcal{L}^\lambda$ satisfies

$$\lim_{\epsilon\to 0^+} F^{\lambda+\epsilon} \leq \lim_{\epsilon\to 0^+} \frac{\mathcal{J}^{\lambda+\epsilon} - \mathcal{J}^\lambda}{\epsilon} \leq F^\lambda. \tag{52}$$

Thus, the derivative of $\mathcal{L}^\lambda$ exists almost everywhere and equals F^λ. Since $\mathcal{L}^\lambda$ is piecewise linear, its derivative must be constant on each linear segment. Therefore, both F^λ and $\mathcal{J}^\lambda$ are piecewise constant. ■

REFERENCES

[1] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[2] R. A. Howard, "Information value theory," *IEEE Trans. Syst. Sci. Cybern.*, vol. 2, no. 1, pp. 22–26, Aug. 1966.

[3] M. Kountouris and N. Pappas, "Semantics-empowered communication for networked intelligent systems," *IEEE Commun. Mag.*, vol. 59, no. 6, pp. 96–102, June 2021.

[4] J. Luo and N. Pappas, "On the cost of consecutive estimation error: Significance-aware non-linear aging," *IEEE Trans. Inf. Theory*, vol. 71, no. 10, pp. 7976–7989, Oct. 2025.

[5] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: Wiley, 1968.

[6] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 2731–2735.

[7] A. Kosta, N. Pappas, V. Angelakis *et al.*, "Age of information: A new concept, metric, and tool," *Found. Trends Netw.*, vol. 12, no. 3, pp. 162–259, Nov. 2017.

[8] G. Stamatakis, N. Pappas, and A. Traganitis, "Control of status updates for energy harvesting devices that monitor processes with alarms," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2019, pp. 1–6.

[9] J. Luo and N. Pappas, "Semantic-aware remote estimation of multiple Markov sources under constraints," *IEEE Trans. Commun.*, vol. 73, no. 11, pp. 11 093–11 105, Nov. 2025.

[10] A. Maatouk, S. Kriouile, M. Assaad, and A. Ephremides, "The age of incorrect information: A new performance metric for status updates," *IEEE/ACM Trans. Netw.*, vol. 28, no. 5, pp. 2215–2228, Oct. 2020.

[11] M. Salimnejad, M. Kountouris, and N. Pappas, "State-aware real-time tracking and remote reconstruction of a Markov source," *J. Commun. Netw.*, vol. 25, no. 5, pp. 657–669, Oct 2023.

[12] J. Luo and N. Pappas, "Exploiting data significance in remote estimation of discrete-state Markov sources," *arXiv preprint arXiv:2406.18270*, 2024.

[13] J. Luo and N. Pappas, "On the role of age and semantics of information in remote estimation of Markov sources," *arXiv preprint arXiv:2507.18514*, 2025.

[14] Y. Sun, Y. Polyanskiy, and E. Uysal, "Sampling of the Wiener process for remote estimation over a channel with random delay," *IEEE Trans. Inf. Theory*, vol. 66, no. 2, pp. 1118–1135, Feb. 2020.

[15] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry, "Foundations of control and estimation over lossy networks," *Proc. IEEE*, vol. 95, no. 1, pp. 163–187, Jan. 2007.

[16] J. Luo, T. Zhang, R. Hao, D. Li, C. Chen, Z. Na, and Q. Zhang, "Real-time cooperative vehicle coordination at unsignalized road intersections," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 5, pp. 5390–5405, May 2023.

[17] S. Vitturi, F. Tramarin, and L. Seno, "Industrial wireless networks: The significance of timeliness in communication systems," *IEEE Ind. Electron. Mag.*, vol. 7, no. 2, pp. 40–51, June 2013.

[18] J. Gielis, A. Shankar, and A. Prorok, "A critical review of communications in multi-robot systems," *Curr. Robot. Rep.*, vol. 3, no. 4, pp. 213–225, Dec. 2022.

[19] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of information: An introduction and survey," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1183–1210, May 2021.

[20] Y. Sun, I. Kadota, R. Talak, and E. Modiano, "Age of information: A new metric for information freshness," *Synth. Lectures Commun. Netw.*, vol. 12, no. 2, pp. 1–224, Dec. 2019.

[21] N. Pappas, M. A. Abd-Elmagid, B. Zhou, W. Saad, and H. S. Dhillon, *Age of Information: Foundations and Applications*. Cambridge, UK: Cambridge Univ. Press, 2023.

[22] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1453–1464, Sept. 2004.

[23] L. Schenato, "Optimal estimation in networked control systems subject to random delay and packet drop," *IEEE Trans. Autom. Control*, vol. 53, no. 5, pp. 1311–1317, June 2008.

[24] L. Shi, P. Cheng, and J. Chen, "Sensor data scheduling for optimal state estimation with communication energy constraint," *Automatica*, vol. 47, no. 8, pp. 1693–1698, Aug. 2011.

[25] A. S. Leong, S. Dey, and D. E. Quevedo, "Sensor scheduling in variance based event triggered estimation with packet drops," *IEEE Trans. Autom. Control*, vol. 62, no. 4, pp. 1880–1895, Apr. 2017.

[26] W. Liu, D. E. Quevedo, Y. Li, K. H. Johansson, and B. Vucetic, "Remote state estimation with smart sensors over Markov fading channels," *IEEE Trans. Autom. Control*, vol. 67, no. 6, pp. 2743–2757, June 2022.

[27] J. Luo, G. Stamatakis, O. Simeone, and N. Pappas, "Remote state estimation over a wearing channel: Information freshness vs. channel aging," *arXiv preprint arXiv:2501.17473*, 2025.

[28] T. Z. Ornee and Y. Sun, "Sampling and remote estimation for the Ornstein-Uhlenbeck process through queues: Age of information and beyond," *IEEE/ACM Trans. Netw.*, vol. 29, no. 5, pp. 1962–1975, Oct. 2021.

[29] C. Kam, S. Kompella, G. D. Nguyen, J. E. Wieselthier, and A. Ephremides, "Towards an effective age of information: Remote estimation of a Markov source," in *Proc. IEEE INFOCOM WKSHPS*, Apr. 2018, pp. 367–372.

[30] Y. Chen and A. Ephremides, "Minimizing age of incorrect information over a channel with random delay," *IEEE/ACM Trans. Netw.*, vol. 32, no. 4, pp. 2752–2764, Aug. 2024.

[31] A. Munari, "Monitoring IoT sources over random access channels: Age of incorrect information and missed detection probability," in *Proc. IEEE Int. Conf. Commun. (ICC)*, June 2024, pp. 207–213.

[32] I. Cosandal, N. Akar, and S. Ulukus, "Modeling AoII in push-and pull-based sampling of continuous time Markov chains," in *Proc. IEEE INFOCOM WKSHPS*, May 2024, pp. 01–06.

[33] I. Cosandal, N. Akar, and S. Ulukus, "Multi-threshold AoII-optimum sampling policies for continuous-time Markov chain information sources," *IEEE Trans. Inf. Theory*, vol. 71, no. 9, pp. 6968–6988, Sept. 2025.

[34] M. Bastopcu and S. Ulukus, "Information freshness in cache updating systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 1861–1874, Mar. 2021.

[35] R. D. Yates, "The age of gossip in networks," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2021, pp. 2984–2989.

[36] B. Buyukates, M. Bastopcu, and S. Ulukus, "Version age of information in clustered gossip networks," *IEEE J. Sel. Areas Inf. Theory*, vol. 3, no. 1, pp. 85–97, Mar. 2022.

[37] E. Delfani and N. Pappas, "Version age-optimal cached status updates in a gossiping network with energy harvesting sensor," *IEEE Trans. Commun.*, vol. 73, no. 4, pp. 2344–2360, Apr. 2025.

[38] M. Salimnejad, M. Kountouris, A. Ephremides, and N. Pappas, "Age of information versions: a semantic view of Markov source monitoring," *IEEE Trans. Commun.*, 2025.

[39] J. P. Champati, M. Skoglund, M. Jansson, and J. Gross, "Detecting state transitions of a Markov source: Sampling frequency and age trade-off," *IEEE Trans. Commun.*, vol. 70, no. 5, pp. 3081–3095, May 2022.

[40] N. Pappas and M. Kountouris, "Goal-oriented communication for real-time tracking in autonomous systems," in *Proc. IEEE Int. Conf. Autom. Syst. (ICAS)*, Aug. 2021, pp. 1–5.

[41] M. Salimnejad, M. Kountouris, and N. Pappas, "Real-time reconstruction of Markov sources and remote actuation over wireless channels," *IEEE Trans. Commun.*, vol. 72, no. 5, pp. 2701–2715, May 2024.

[42] A. Zakeri, M. Moltafet, and M. Codreanu, "Semantic-aware sampling and transmission in real-time tracking systems: A pomdp approach," *IEEE Trans. Commun.*, vol. 73, no. 7, pp. 4898–4913, Jul. 2025.

[43] J. Luo and N. Pappas, "Goal-oriented estimation of multiple Markov sources in resource-constrained systems," in *Proc. IEEE 35th Int. Symp. Pers. Indoor Mob. Radio Commun. (PIMRC)*, Sept. 2024, pp. 1–6.

[44] J. Luo and N. Pappas, "Minimizing the age of missed and false alarms in remote estimation of Markov sources," in *Proc. 25th ACM MobiHoc*, Oct. 2024, pp. 381–386.

[45] J. Luo and N. Pappas, "Revisiting estimation quality: Significance-aware age of consecutive error," in *Proc. 23rd Int. Symp. Model. Optim. Mobile, Ad Hoc, Wireless Netw. (WiOpt)*, May 2025, pp. 1–8.

[46] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Hoboken, NJ, USA: Wiley, 1994.

[47] E. Altman, *Constrained Markov Decision Processes*. Boca Raton, FL, USA: CRC Press, 1999.

[48] D. M. Roijers, P. Vamplew, S. Whiteson, and R. Dazeley, "A survey of multi-objective sequential decision-making," *J. Artif. Intell. Res.*, vol. 48, pp. 67–113, Oct. 2013.

[49] K. Chatterjee, R. Majumdar, and T. A. Henzinger, "Markov decision processes with multiple objectives," in *Proc. 23rd Annu. Symp. Theor. Aspects Comput. Sci. (STACS)*, Feb. 2006, pp. 325–336.

[50] K. Chatterjee, "Markov decision processes with multiple long-run average objectives," in *Proc. 27th Int. Conf. Found. Softw. Technol. Theor. Comput. Sci. (FSTTCS)*, Dec. 2007, pp. 473–484.

[51] Z. Gábor, Z. Kalmár, and C. Szepesvári, "Multi-criteria reinforcement learning." in *Proc. 15th int. conf. mach. learn. (ICML)*, July 1998, pp. 197–205.

[52] L. Barrett and S. Narayanan, “Learning all optimal policies with multiple criteria,” in *Proc. 25th int. conf. Mach. learn. (ICML)*, July 2008, pp. 41–47.
[53] K. Van Moffaert, M. M. Drugan, and A. Nowé, “Scalarized multi-objective reinforcement learning: Novel design techniques,” in *Proc. 2013 IEEE Symp. Adapt. Dyn. Program. Reinforc. Learn. (ADPRL)*, Apr. 2013, pp. 191–199.
[54] P. Vamplew, R. Dazeley, A. Berry, R. Issabekov, and E. Dekker, “Empirical evaluation methods for multiobjective reinforcement learning algorithms,” *Mach. Learn.*, vol. 84, no. 1, pp. 51–80, July 2011.
[55] D. White, “Multi-objective infinite-horizon discounted Markov decision processes,” *J. Math. Anal. Appl.*, vol. 89, no. 2, pp. 639–647, Oct. 1982.
[56] M. A. Wiering and E. D. de Jong, “Computing optimal stationary policies for multi-objective Markov decision processes,” in *Proc. IEEE Symp. Approx. Dyn. Program. Reinforc. Learn.*, Apr. 2007, pp. 158–165.
[57] K. V. Moffaert and A. Nowé, “Multi-objective reinforcement learning using sets of pareto dominating policies,” *J. Mach. Learn. Res.*, vol. 15, no. 107, pp. 3663–3692, Nov. 2014.
[58] V. Krishnamurthy, *Partially Observed Markov Decision Processes: From Filtering to Controlled Sensing*. Cambridge, UK: Cambridge Univ. Press, 2016.
[59] F. J. Beutler and K. W. Ross, “Optimal policies for controlled Markov chains with a constraint,” *J. Math. Anal. Appl.*, vol. 112, no. 1, pp. 236–252, 1985.
[60] L. I. Sennott, “Constrained average cost Markov decision chains,” *Probab. Eng. Inf. Sci.*, vol. 7, no. 1, pp. 69–83, 1993.