MANAGEMENT SCIENCE
Vol. 00, No. 0, Xxxxx 0000, pp. 000–000
ISSN 0025-1909 | EISSN 1526-5501 | 00 | 0000 | 0001

INFORMS
DOI 10.1287/xxxx.0000.0000

Monitoring With Limited Information

We consider a system with an evolving state that can be stopped at any time by a decision maker (DM), yielding a state-dependent reward. The DM does not observe the state except for a limited number of monitoring times, which he must choose, in conjunction with a suitable stopping policy, to maximize his reward. Dealing with this type of stopping problems, which arise in a variety of applications from healthcare to finance, often requires excessive amounts of data for calibration purposes, and prohibitive computational resources. To overcome these challenges, we propose a robust optimization approach, whereby adaptive uncertainty sets capture the information acquired through monitoring. We consider two versions of the problem—static and dynamic—depending on how the monitoring times are chosen. We show that, under certain conditions, the same worst-case reward is achievable under either static or dynamic monitoring. This allows recovering the optimal dynamic monitoring policy by resolving static versions of the problem. We discuss cases when the static problem becomes tractable, and highlight conditions when monitoring at equi-distant times is optimal. Lastly, we showcase our framework in the context of a healthcare problem (monitoring heart transplant patients for Cardiac Allograft Vasculopathy), where we design optimal monitoring policies that substantially improve over the status quo treatment recommendations.

Key words: robust optimization, monitoring, optimal stopping problem

1. Introduction

Consider a system with a randomly evolving state that can be stopped at any time by a decision maker (DM), yielding a state-dependent reward. The DM does not observe the system's state except for a specific limited number of times when he chooses to monitor the system, perhaps at some cost. To maximize his reward, the DM needs to devise dynamic policies that prescribe when to monitor the system and whether to continue or stop, based on all acquired information.

Problems of this kind have received a lot of attention in the stochastic optimization literature: categorized as a form of optimal stopping problems, the commonly suggested solution approaches involve characterizing the uncertain evolution through probability distribution functions, and relying on dynamic programming techniques to derive policies that maximize the expected reward.

Many practical considerations that arise in application domains ranging from healthcare to finance, however, often render such approaches unsuitable. First, calibrating joint probability distribution functions for all possible time points in the planning horizon typically requires a prohibitively large amount of data that might not be available in practice, or relies on structural assumptions on state evolution, such as independence or lack of autocorrelation, that might not conform with the true dynamics of the system. Second, the system's complexity is often such that it requires a high-dimensional state representation, resulting in "curse of dimensionality" issues. Third, modeling the learning process that emerges from acquiring information through system monitoring usually requires augmenting this state beyond computational reach.

To see this in a concrete example, consider patients suffering from Cardiac Allograft Vasculopathy (CAV), who need to appropriately time their only viable treatment, namely heart transplant: performing the transplant too early reduces its success chances, while performing it too late comes with the risk of CAV progression and health deterioration. Although monitoring the disease status requires invasive and potentially expensive procedures, it is necessary for decision making, in the absence of credible ways to model and predict CAV progression. Indeed, very little data is available from prior studies of CAV patients, a problem that is compounded by the disease's pathology, which is known to be very complex (see §5). How should CAV patients and their physicians make monitoring and treatment decisions facing this considerable uncertainty?

Similarly, consider a lending institution that extends commercial loans backed by working assets, such as inventory. For a given loan, the lender can occasionally monitor the value of the pledged collateral through costly field visits and inspections, and decide whether to request early repayment or force collateral liquidation, to avert future possible losses. Making predictions about future collateral value, however, could be a very challenging task, as it would involve tracking the value of a considerable number of assets classes that are present in or correlated with the pledged assets. Furthermore, these assets may be illiquid and the lender may lack expertise in valuing them, such as with specialized types of inventory (CH 2014). How should asset-based lenders make monitoring and liquidation decisions when they need to track a multitude of highly uncertain assets?

To address such questions, we develop a new solution methodology that overcomes the challenges discussed above, and then apply it in a practical setting using real data. In particular, we rely on the robust optimization paradigm to develop a model where the DM's information concerning future state evolution is captured through multi-dimensional uncertainty sets. Rather than being exogenously specified, the uncertainty sets governing future state values depend on state values observed at past monitoring times. We study the DM's problem of finding a policy for choosing the monitoring times and for stopping the system so as to maximize his worst-case reward. Our framework alleviates some of the aforementioned practical challenges: uncertainty sets can be more readily calibrated from limited data, and the robust optimization problems that need to be solved for finding optimal policies cope more favorably with "curse of dimensionality" issues.

We make the following contributions.

- We consider two versions of the problem—static and dynamic—depending on whether the monitoring times are committed to upfront, or respectively adjusted depending on acquired information. We show that when the reward is monotonic and the uncertainty sets have a lattice structure, the same worst-case reward is achievable under either static or dynamic monitoring. This allows recovering an optimal dynamic monitoring policy by re-solving static versions of the monitoring problem, drastically simplifying the decision problem. In an extension, we generalize these results to settings where monitoring is costly, or where the DM can adjust the state values by extracting rewards or incurring a cost.
- We then discuss sufficient conditions under which the static monitoring and stopping problem can be efficiently solved. We focus on generalizations of box-type (Ben-Tal et al. 2009) and Central-Limit-Theorem-type (Bandi and Bertsimas 2012) uncertainty sets, wherein past measurements impose lower and upper bounds on future state values, giving rise to an "uncertainty envelope." We show that the curvature of this uncertainty envelope is critical. When the envelope is generated by bounding functions that are convex in the elapsed time, a single monitoring opportunity suffices for recovering the worst-case reward, and the optimal time can be found by solving a one-dimensional optimization problem. When the bounding functions are concave, all monitoring times are needed for reducing the uncertainty, and it is optimal for the DM to distribute these times uniformly over the horizon. We discuss conditions relying on complementarity properties (e.g., supermodularity) of the reward function and uncertainty envelopes that allow solving the problem through combinatorial algorithms.

- We leverage our approach to devise monitoring and transplantation policies for patients suffering from Cardiac Allograft Vasculopathy (CAV). We calibrate our model using data from health studies and published medical papers, and show how optimal monitoring and stopping policies can be found by solving mixed-integer linear programming problems. Simulation suggests that our policies generate life year gains that stochastically dominate those provided by existing medical guidelines, with a substantial increase in lower percentiles and a slight increase in median and higher percentiles.

1.1. Literature Review

Our assumption that information is captured through uncertainty sets relates this paper to the extensive literature in robust optimization (RO) and robust control (see, e.g., Ben-Tal and Nemirovski 2002, Bertsimas et al. 2011 and references therein for the former, and Dullerud and Paganini 2005, Zhou and Doyle 1998 for the latter). Typical models in RO consider uncertainty sets that are exogenous to the decision process, and focus on the computational tractability of the resulting optimization problems. Poss (2013) and Nohadani and Sharma (2016) consider decision-dependent uncertainty sets, and propose mixed-integer programming reformulations for the resulting non-convex optimization problems (also see Jaillet et al. 2016 for similar concepts applied in a model with satisficing objectives and constraints). Similarly, Zhang et al. (2016) study robust control problems where the decision maker *chooses* the uncertainty set from a family of structured sets, and incurs a penalty for “smaller” choices (in terms of radius, volume, etc.); since the dynamic decision problems are generally intractable, the paper resorts to policy approximations through affine decision rules, which are numerically shown to deliver good performance. Bertsimas and Vayanos (2017) consider a pricing problem where the demand function coefficients are unknown, and propose an approximation scheme that requires solving mixed-binary conic optimization problems. Different from these papers, our uncertainty sets depend on a monitoring policy chosen by the DM, which reduces future uncertainty under well defined information updating rules. Closest to our paper is Nohadani and Roy (2017), who also construct conic uncertainty sets where observation times (chosen statically at the beginning of the planning horizon) influence the feasible uncertainties in each period, and provide conditions when monitoring once at the midpoint is optimal for a two-stage linear model. In contrast, we consider both static and dynamic monitoring policies under a multi-period setting and nonlinear objectives, and provide conditions when both formulations yield the same value, and monitoring at equidistant times is optimal.

Our work is also related to a subset of papers in RO that discuss the optimality of static policies in dynamic decision problems. Ben-Tal et al. (2004) prove this for a linear program where the uncertain parameters affecting distinct constraints are disjoint, i.e., the uncertainty is "constraint-wise." Marandi and den Hertog (2017) extend this condition to two-stage problems where constraints are convex in decisions and concave in uncertainties, and Bertsimas et al. (2014) prove static optimality for two-stage linear programs when a particular transformation of the uncertainty set is convex. We extend this literature by proving optimality for static monitoring rules in a broad class of multi-period problems, provided the objectives possess certain monotonicity properties, and the uncertainty sets have ordering properties (e.g., they are lattices).

Our work is also related to the broad literature on stopping problems. These are typically formulated under a fully specified probability distribution for the stochastic processes of interest (see, e.g., Snell (1952), Taylor (1968), Karatzas and Shreve (2012), and references therein). Some of this literature also incorporates robustness by allowing for a set of possible distributions chosen by nature, and extending the martingale approach to characterize stopping in the (continuous-time) game between the DM and nature (see, e.g., Riedel (2009), Bayraktar and Yao (2014), Bayraktar et al. (2010), and references therein). In contrast, we consider all measures with a given support (i.e., the uncertainty set, as in classical RO), and restrict the stopping decision to a finite number of times, chosen by the decision maker. This provides a more sharp characterization of the optimal stopping policy, as well as computationally tractable procedures for finding it.

Finally, our applications of interest also relate our paper (albeit more loosely) to an extensive literature studying monitoring and stopping problems in healthcare or finance. For the former stream, we direct the interested reader to Shechter et al. (2008) (studying the optimal time to initiate HIV treatment), Maillart et al. (2008) and Ayer et al. (2012) (assessing breast cancer screening policies), Denton et al. (2009) (optimizing the start of statin therapy for diabetes patients), Lavieri et al. (2012) (optimizing the start of radiation therapy treatment for prostate cancer), and the numerous references therein. This line of work typically relies on models based on (partially observable) Markov Decision Processes, and unique probability distributions for transitions and rewards. Instead, we adopt a robust model where limited information is available; this allows a sharper characterization of optimal policies, which may not be possible when insisting on Bellman-optimal policies, as required under uniquely specified stochastic processes and risk-neutral objectives (also see Delage and Iancu 2015 and Iancu et al. 2013 for a discussion).

For the latter stream, we mention the extensive work on pricing exotic options, of which Bermudan options—which can only be exercised on a pre-determined set of dates—are perhaps closest to our framework (see Ibanez and Zapatero (2004) and Kolodko and Schoenmakers (2006) for Bermudan options, and Wilmott et al. (1994) and Karatzas et al. (1998) for a broader overview). In contrast to our work, stopping times are typically exogenous, and exact computations of the liquidation frontier are generally too complex, so that solution methods typically rely on discretized numerical integration and/or simulation-based approximations. Also related are papers in finance that deal with debt monitoring and repayment/amortization schedules (see, e.g., Leland (1994), Leland and Toft (1996), Gorton and Winton (2003), Rajan and Winton (1995), and references therein). Such papers typically consider stylized settings (e.g., two or three-periods, a single, uniquely defined stochastic process of interest, etc.), since the focus is characterizing the optimal capital structure, rather then optimizing monitoring schedules or liquidation policies.

1.2. Notation and Terminology

To distinguish vectors from scalars, we denote the former by bold fonts, e.g., $\boldsymbol{x} \in \mathbb{R}^n$. For a set of vectors $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ in $\mathbb{R}^d$, we use $\boldsymbol{x}^{\{k\}} \stackrel{\text{def}}{=} [\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_k] \in \mathbb{R}^{d\times(k+1)}$ to denote the matrix obtained by horizontal concatenation. Similarly, the operator ; is used to denote vertical concatenation of matrices with the same number of columns. For a set of matrices $\mathcal{U} \subset \mathbb{R}^{d\times k}$ and indices $1 \le i_1 \le i_2 \le d$ and $1 \le j_1 \le j_2 \le k$, we define

$$\Pi_{i_1:i_2,:}\mathcal{U} = \Big\{ Y \in \mathbb{R}^{(i_2-i_1+1)\times k} \,:\, \exists\, X \in \mathbb{R}^{(i_1-1)\times k},\, Z \in \mathbb{R}^{(d-i_2)\times k} \;:\; [X; Y; Z] \in \mathcal{U} \Big\},$$
$$\Pi_{:,j_1:j_2}\mathcal{U} = \Big\{ Y \in \mathbb{R}^{d\times(j_2-j_1+1)} \,:\, \exists\, X \in \mathbb{R}^{d\times(j_1-1)},\, Z \in \mathbb{R}^{d\times(k-j_2)} \;:\; [X, Y, Z] \in \mathcal{U} \Big\},$$

as the projections along a subset of rows and a subset of columns, respectively. To simplify notation, we also use $\mathcal{U}_j \stackrel{\text{def}}{=} \Pi_{:,j:j}\mathcal{U}$ to denote the latter projection along a single column.

We use the terms "increasing" and "decreasing" in a weak sense, and we refer to functions of multiple arguments as being increasing (decreasing) instead of isotone (antitone), which is the established lattice terminology (Topkis 1998). We compare vectors in $\mathbb{R}^d$

using the typical component-wise order, i.e., $\boldsymbol{x} \succeq \boldsymbol{y}$ for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$ if and only if $x_i \geq y_i$, $1 \leq i \leq d$. Similarly, we compare matrices by first viewing them as vectors, and applying the component-wise order above. For sets of matrices or vectors, "increasing" and "decreasing" are understood in the typical set order (Topkis 1998), unless explicitly stated otherwise.

2. Model

Consider a system whose state is characterized by a finite number of exogenous processes evolving over a finite time horizon. A decision maker (DM) starts with limited information on the future process values and can observe (or monitor) the system at a finite number of times. These times, which we refer to as *monitoring times*, are chosen by the DM. At every monitoring time, the DM observes the state of the system, updates his information about the future process values, and decides whether to stop the system or allow it to continue evolving until the next monitoring time. If at some monitoring time the DM chooses to stop the system, he collects a reward that depends on the stopping time and the state. If he decides to continue until the end, he collects the reward associated with the terminal state. We focus on a system with high uncertainty, where the DM's information on the future state evolution is described by uncertainty sets rather than a complete probability law. In this setting, the DM's problem is to find a policy for choosing the monitoring times and for stopping the system so as to maximize his worst-case reward.

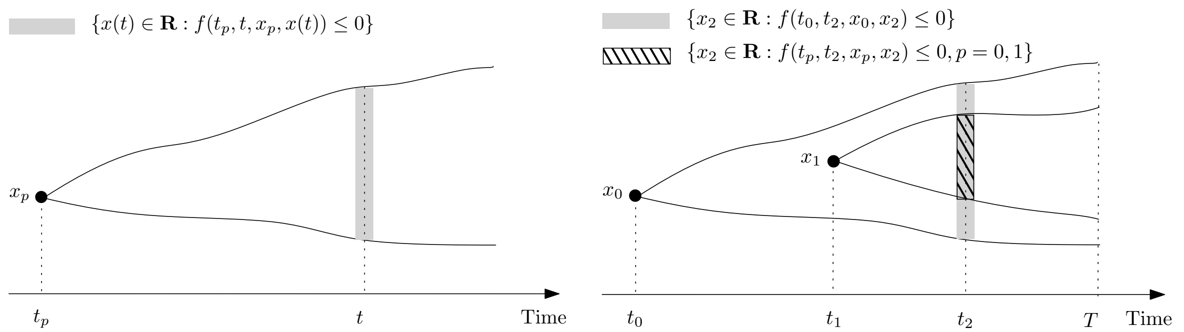
We next introduce notation and provide a formal model description.

2.1. State Evolution and Uncertainty Set

The system evolves over a continuous time frame $[0, T]$, and is characterized by d real-valued processes. Let $\boldsymbol{x}(t) \in \mathbb{R}^d$ be a vector whose components are the system's processes, for all t in $[0, T]$. We refer to $\boldsymbol{x}(t)$ as the state of the system at time t.

The DM has knowledge of the initial state, i.e., $\boldsymbol{x}(0)$, and can monitor the system at most n times throughout the planning horizon, in addition to the default monitoring times of 0 and T. Let $t_1, \ldots, t_n$ denote these n monitoring times with $0 \leq t_1 \leq \cdots \leq t_n \leq T$. To ease notation, we also let $t_0 \stackrel{\text{def}}{=} 0$ and $t_{n+1} \stackrel{\text{def}}{=} T$.

At each monitoring time t_p, $0 \leq p \leq n+1$, the DM observes the state $\boldsymbol{x}_p \stackrel{\text{def}}{=} \boldsymbol{x}(t_p)$ and updates his information on the future possible state values. More precisely, each observed state $\boldsymbol{x}_p$ imposes restrictions on the feasible set of values $\boldsymbol{x}(t)$ at every future time $t > t_p$. These restrictions are summarized through m constraints, written compactly as $f(t_p, t, \boldsymbol{x}_p, \boldsymbol{x}(t)) \leq 0$, where $f : \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^m$. For an illustration, see Figure 1(a).



(a) Consider the case $d = 1$. Each observation x_p imposes a restriction on the future process values $x(t)$ for all $t > t_p$, which is identified by $f(t_p, t, x_p, x(t))$.

(b) With two observations $\boldsymbol{x}^{\{1\}} = [x_0, x_1]$, the set of feasible state values x_2 at time t_2, given by the projection of $\mathcal{U}(\boldsymbol{t}^{\{2\}}, \boldsymbol{x}^{\{1\}})$ on x_2 (dashed), is the intersection of the projection of $\mathcal{U}(\boldsymbol{t}^{\{2\}}, \boldsymbol{x}^{\{0\}})$ on x_2 (gray) and the restriction imposed by the second observation x_1.

Figure 1 Illustration of how observed state values constrain future values for $d = 1$.

We assume that the DM reconciles new information (from new measurements) with old information (from existing measurements) by considering all the restrictions imposed. To make this precise and introduce some notation, consider the k-th monitoring time t_k. At that time, the DM has made k observations at monitoring times $t_1, \ldots, t_k$, in addition to the initial observation at t_0, resulting in an observation matrix $\boldsymbol{x}^{\{k\}} \stackrel{\text{def}}{=} [\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_k]$. Suppose the DM is considering some future monitoring times $t_{k+1}, \ldots, t_r$ satisfying $t_k \le t_{k+1} \le \cdots \le t_r \le t_{n+1}$, for $r > k$. Then, in view of observations $\boldsymbol{x}^{\{k\}}$, the process values at $t_{k+1}, \ldots, t_r$, together with the terminal process value at t_{n+1}, lie in the following uncertainty set:

$$\mathcal{U}\big(\boldsymbol{t}^{\{r\}}, \boldsymbol{x}^{\{k\}}\big) = \Big\{ [\boldsymbol{x}_{k+1}, \ldots, \boldsymbol{x}_r, \boldsymbol{x}_{n+1}] \in \mathbb{R}^{d \times (r-k+1)} : f(t_p, t_q, \boldsymbol{x}_p, \boldsymbol{x}_q) \le 0, \\ \forall p, q \in \{0, 1, \ldots, r, n+1\}, p < q \Big\}, \tag{1}$$

where $\boldsymbol{t}^{\{r\}} \stackrel{\text{def}}{=} [t_0, t_1, \ldots, t_r]$ denotes the vector of the first $r + 1$ monitoring times (including the initial time t_0). As such, information updating in our model corresponds to taking intersections of the corresponding uncertainty sets.[1]

To illustrate this, consider the example in Figure 1(b), where a system with a single process ($d = 1$) is monitored twice ($n = 2$), at t_1 and t_2. At time t_0, the DM starts with an

[1] Taking intersection also implies that uncertainty sets updated after new observations can only include values that were already considered plausible at time t_0. This allows the DM to behave in a dynamically consistent fashion, and ensures that having additional information is always preferable (Machina 1989, Hanany and Klibanoff 2009). Note that considering a large enough set of future plausible state values at time t_0 always suffices for this to hold.

initial information on the future values x_1, x_2 and $x_3 = x_{n+1}$, represented by $\mathcal{U}(\boldsymbol{t}^{\{2\}}, \boldsymbol{x}^{\{0\}}) = \{[x_1, x_2, x_3] \in \mathbb{R}^3 : f(t_0, t_q, x_0, x_q) \le 0, q = 1, 2, 3\}$. Then, with an observation x_1 at t_1, the updated set $\mathcal{U}(\boldsymbol{t}^{\{2\}}, \boldsymbol{x}^{\{1\}})$ becomes the intersection of $\{[x_2, x_3] \in \mathbb{R}^2 : f(t_0, t_q, \boldsymbol{x}_0, \boldsymbol{x}_q) \le 0, q = 2, 3\}$ and $\{[x_2, x_3] \in \mathbb{R}^2 : f(t_1, t_q, x_1, x_q) \le 0, q = 2, 3\}$, which are the restrictions imposed by x_0 and x_1, respectively. Figure 1(b) depicts the set of feasible values for x_2, which is the projection of $\mathcal{U}(\boldsymbol{t}^{\{2\}}, \boldsymbol{x}^{\{1\}})$ on the first dimension. In §2.4, we discuss additional assumptions concerning the structure of the uncertainty sets and the restrictions in f.

2.2. Decision Problem

We consider two models of the problem—*static* and *dynamic*—depending on how the monitoring times are chosen. In the static model, the DM selects and commits to all monitoring times $t_1, \ldots, t_n$ at time t_0. The stopping decision remains dynamic, adapted to available information; more formally, a stopping policy at the k-th monitoring time, Π_k^S, is a mapping from the σ-algebra induced by the choice of times $t_1, \ldots, t_{n+1}$ and the observation matrix $\boldsymbol{x}^{\{k\}}$ to the set {Stop, Continue}. The DM's *static monitoring problem* thus entails choosing the monitoring times $\boldsymbol{t}^{\{n\}}$ and a set of stopping policies, $\Pi^S \stackrel{\text{def}}{=} \{\Pi_0^S, \Pi_1^S, \ldots \Pi_n^S\}$.

In the dynamic model, the DM only needs to select and commit to the nearest future monitoring time. More precisely, at time t_k, the DM simultaneously chooses whether to stop the process at time t_k or not, and in the latter case, selects the time t_{k+1} when to next monitor the process. Therefore, in the *dynamic monitoring problem*, the DM needs to choose a monitoring and stopping policy, $\Pi^D = \{\Pi_0^D, \Pi_1^D, \ldots, \Pi_n^D\}$, where each Π_k^D is a mapping from the σ-algebra induced by $(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}})$ to $\{\text{Stop}, \text{Continue}\} \times [t_k, T]$.

In both models, the DM collects a reward $g(t, \boldsymbol{x})$ when stopping at time t with state $\boldsymbol{x}$.[2] His objective is to maximize his worst-case stopping reward over the possible monitoring and stopping policies. To ease the notational burden, we formalize these problems in §3.1.

2.3. Applications

Our framework captures several problems encountered in practice. We next discuss two such examples.

Healthcare. Consider a patient suffering from a chronic condition for which the preferred initial treatment is passive, by regularly monitoring disease progression through office

[2] That the reward is collected only when stopping is without loss of generality here, since we could always append an additional process to the state, accumulating all the intermediate (discounted) rewards.

visits and/or various tests. The act of monitoring is often invasive, requiring exposure to toxic agents (e.g., radioactive agents in imaging studies) or even micro-surgical procedures (e.g., collecting tissues for biopsy). In addition, testing can also be expensive, generating both direct and indirect costs for to the patient. These considerations limit the number of monitoring chances, and require physicians to make judicious use of them. It thus becomes critical to understand when to monitor and—depending on observed outcomes—whether to discontinue the passive treatment and switch to an active one, such as a disease modifying agent or a surgical procedure.

In this context, the state $\boldsymbol{x}(t)$ can capture the health measurements related to disease progression, which can be collected during an office visit. For instance, when monitoring heart-transplanted patients for CAV, $\boldsymbol{x}(t)$ could consist of the degree of angiographic lesion, the number of acute rejections, the CAV stage, and the age. The stopping reward $g(t, \boldsymbol{x})$ depends on the focus, but can generally capture the patient's total cumulative quality-adjusted life years (QALY) from switching to an active treatment.

Lending. Consider a lender issuing a loan with collateral consisting of several working assets, e.g., accounts receivable, inventory, equipment, etc. Since lenders often lack extensive in-house expertise for evaluating collateralized working assets, it is customary to contract with third party liquidation houses to obtain periodic appraisals (see, e.g., Foley et al. 2013 and CH 2014, p.16). Based on these appraisals and on certain advance rates (i.e., hair-cuts) applied to each collateral type,[3] the lender then calculates a borrowing base, which essentially corresponds to the liquidity-adjusted collateral value (CH 2014, p.15-20). This borrowing base is critical in determining whether the borrower is "over-advanced," i.e., whether the outstanding loan is not supported by the collateral value.

Relating to our model, the state $\boldsymbol{x}(t)$ can capture the value of the d assets pledged as collateral, which fluctuates over time. With $a_i \in [0, 1]$ denoting the advance rate assigned to the i-th collateral type, each monitoring opportunity then allows the lender to observe the current value of the borrowing base, $\boldsymbol{a}^T \boldsymbol{x}(t)$. The reward $g(t, \boldsymbol{x})$ can be modeled in several ways, depending on the remedial actions imposed by the lender when the borrower is over-advanced (e.g., requiring early repayment, renegotiating the contract, or even liquidating the assets), but it would typically be a function of t and $\boldsymbol{x}$ that increases in $\boldsymbol{x}$.

[3] In practice, advance rates for accounts receivable range from 70% to 90%, and advance rates for inventory are at most 65% of the book value or 80% of the net orderly liquidation value (see CH 2014, p.17-19).

2.4. Assumptions

The monitoring and stopping problem we described so far is generally intractable due to the curse of dimensionality. In view of the motivating applications, we introduce assumptions that make it amenable to analysis. The first concerns the reward function g.

ASSUMPTION 1. *$g(t,\boldsymbol{x})$ is monotonic in each x_j, $\forall j \in \{1,\ldots,d\}$.*

Reward monotonicity is consistent with several practical considerations. In chronic disease monitoring, the patient's reward is usually monotonic in the health state $\boldsymbol{x}$. For CAV patients, for instance, the expected QALY with re-transplantation decrease with age, the number of acute rejections and CAV stage (Johnson et al. 2007, Johnson 2007). In collateralized lending, $g(t,\boldsymbol{x})$ increases in $\boldsymbol{x}$ because the lender's payoff upon a remedial action typically grows with the values of collateralized assets (CH 2014).

Without loss, we henceforth assume that $g(t,\boldsymbol{x})$ is increasing in all states (this can always be achieved through a change of variables, by replacing x_j with $-x_j$ for any j such that g is decreasing in x_j).

The next assumption relates to the structure of the uncertainty set(s). Its requirements bear intuitive interpretations, and are compatible with several classical families of uncertainty sets considered in the literature. We first introduce and discuss these requirements, and then provide illustrative examples.

ASSUMPTION 2. *For any $0 \le k \le r \le n$ and given $\boldsymbol{t}^{\{r\}}$ and $\boldsymbol{x}^{\{k\}}$,*

(i) (Lattice) $\mathcal{U}\big(\boldsymbol{t}^{\{r\}},\boldsymbol{x}^{\{k\}}\big)$ is a lattice;

(ii) (Monotonicity) $\mathcal{U}\big(\boldsymbol{t}^{\{r\}},\boldsymbol{x}^{\{k\}}\big)$ is increasing in $\boldsymbol{x}^{\{k\}}$;

(iii) (Dynamic Consistency) $\Pi_{:,i:j}\mathcal{U}(\boldsymbol{t}^{\{r\}},\boldsymbol{x}^{\{k\}}) = \Pi_{:,i:j}\mathcal{U}(\boldsymbol{t}^{\{r'\}},\boldsymbol{x}^{\{k\}})$, $\forall i \le j \le r \le r' \le n+1$.

The lattice requirement is primarily of technical nature.[4] The monotonicity requirements are guided by practical considerations. At an intuitive level, these are akin to "better past performance" being "good news" for the future, and "worst past performance" being "bad news" for the future. Mathematically, "better past performance" means higher $\boldsymbol{x}^{\{k\}}$, which leads to larger future state values, i.e., good news for the future. Conversely, "worse past performance" means lower $\boldsymbol{x}^{\{k\}}$, which leads to lower future state values, i.e., bad news for the future. This makes sense in chronic disease monitoring: better (worse) medical history is

[4] Strictly speaking, our results only require $\mathcal{U}$ to be a meet-semilattice.

often indicative of good (bad) outcomes in the future. For CAV patients, for instance, higher CAV stages in the past lead to higher stages in the future (see §5). In collateralized lending, several empirical studies show that pledged asset values are positively autocorrelated over a short time horizon (e.g., Cutler et al. 1991). Finally, dynamic consistency corresponds to the natural requirement that committing to additional monitoring times in the future, $t_{r+1},\dots,t_{r'}$, does not affect the set of feasible state values at monitoring times before t_r.

In practice, requirements (i) and (ii) are relatively easy to check, as our examples will shortly illustrate. To facilitate checking (iii), we introduce the following sufficient condition.

Proposition 1. *Suppose that for any t_p and $\boldsymbol{x}_p$, and any t,t' satisfying $t_p<t<t'$,*

$$\{\boldsymbol{x}\in\mathbb{R}^d : f(t_p,t,\boldsymbol{x}_p,\boldsymbol{x})\le 0\}\subseteq\{\boldsymbol{x}\in\mathbb{R}^d : f(t_p,t',\boldsymbol{x}_p,\boldsymbol{x})\le 0\}. \tag{2}$$

Then, Assumption 2(iii) holds.

Proof It suffices to show that for any $\boldsymbol{x}^{\{k\}}$ and monitoring times $\boldsymbol{t}^{\{r\}}$ $(k<r\le n)$,

$$\Pi_{:,i:j}\mathcal{U}(\boldsymbol{t}^{\{r\}},\boldsymbol{x}^{\{k\}})=\Pi_{:,i:j}\mathcal{U}(\boldsymbol{t}^{\{k+j\}},\boldsymbol{x}^{\{k\}}),\ \forall\, 1\le i\le j\le r-k+1.$$

By definition, $\Pi_{:,i:j}\mathcal{U}(\boldsymbol{t}^{\{r\}},\boldsymbol{x}^{\{k\}})\subseteq\Pi_{:,i:j}\mathcal{U}(\boldsymbol{t}^{\{k+j\}},\boldsymbol{x}^{\{k\}})$, because introducing more monitoring times imposes more constraints, thus shrinking the uncertainty sets and their projections. To prove the opposite direction, we must show that if $\boldsymbol{B}\in\Pi_{:,i:j}\mathcal{U}(\boldsymbol{t}^{\{k+j\}},\boldsymbol{x}^{\{k\}})$, then $\boldsymbol{B}\in\Pi_{:,i:j}\mathcal{U}(\boldsymbol{t}^{\{r\}},\boldsymbol{x}^{\{k\}})$, i.e., $\exists\,\boldsymbol{A}\in\mathbb{R}^{d\times(i-1)},\ \boldsymbol{C}\in\mathbb{R}^{d\times(r-j)}$ such that $[\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}]\in\mathcal{U}(\boldsymbol{t}^{\{r\}},\boldsymbol{x}^{\{k\}})$. By definition, since $\boldsymbol{B}\in\Pi_{:,i:j}\mathcal{U}(\boldsymbol{t}^{\{k+j\}},\boldsymbol{x}^{\{k\}})$, there exists a matrix of possible future values $A\overset{\text{def}}{=}[\boldsymbol{x}_{k+1},\dots,\boldsymbol{x}_{k+i-1}]\in\mathbb{R}^{d\times(i-1)}$, such that $[\boldsymbol{A},\boldsymbol{B}]\in\mathcal{U}(\boldsymbol{t}^{\{k+j\}},\boldsymbol{x}^{\{k\}})$. Now, consider the matrix $[\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}]$, where $\boldsymbol{C}=(\Pi_{:,(j-i+1):(j-i+1)}\boldsymbol{B})\mathbf{1}^{1\times(r-j)}$, i.e., obtained by repeating the last element of $\boldsymbol{B}$ for $r-j$ times. Then, by Proposition, the last entry of $\boldsymbol{B}$, which is feasible for t_{k+j}, is feasible for $t_{i+j+1},\dots,t_r$ as well. Therefore, $[\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}]\in\mathcal{U}(\boldsymbol{t}^{\{r\}},\boldsymbol{x}^{\{k\}})$. ■

The sufficient condition (2) for dynamic consistency requires that the possible process values at time t are among the possible values at time $t'>t$. This is a very mild requirement if $\boldsymbol{x}$ corresponds to a “zero-mean” (noise) process, since it is natural to expect information for more distant future times to become “less accurate”/“more noisy” in such settings. But the assumption remains innocuous even when $\boldsymbol{x}$ exhibits predictable trends or seasonalities: one can decompose the process into a component satisfying (2) and a predictable time-varying component, whose effect can be captured through the time-dependency of the reward function g.

2.5. Examples of Uncertainty Sets

We next present several examples of uncertainty sets satisfying our requirements.

Example 1 (Lattice with Cross-Constraints). We consider a generalization of the classical box uncertainty sets to our dynamic setting, where bounds depend on observed information and monitoring times. More precisely, for $\beta \geq \alpha \geq 0$, $\mathcal{M} \subseteq \{1,\ldots,d\}^2$, $\ell : \mathbb{R}^2 \to \mathbb{R}_-$ decreasing in its second argument, and $u : \mathbb{R}^2 \to \mathbb{R}_+$ increasing in its second argument, consider the set:

$$\mathcal{U}(\boldsymbol{t}^{\{r\}}, \boldsymbol{x}^{\{k\}}) = \Big\{ [\boldsymbol{x}_{k+1},\ldots,\boldsymbol{x}_r,\boldsymbol{x}_{n+1}] \in \mathbb{R}^{d\times(r-k+1)} : \alpha \cdot x_p^m + \ell(t_p, t_q - t_p) \leq x_q^{m'} \leq \beta \cdot x_p^m + u(t_p, t_q - t_p), \forall (m,m') \in \mathcal{M}, \forall p \in \{0,1,\ldots,k\}, \forall q \in \{k+1,\ldots,r\} \Big\}. \tag{3}$$

Here, $\mathcal{M}$ specifies a group of processes that are dependent on each other. When $(m,m') \in \mathcal{M}$, an observation with value x_p^m for the m-th process at time t_p would impose restrictions on the future process value $x_q^{m'}$ for the m'-th process at time t_q, in the form of a lower bound $\alpha x_p^m + \ell(t_p, t_q - t_p)$ and an upper bound $\beta x_p^m + u(t_p, t_q - t_p)$. The bounds depends on the time of the past observation, t_p, and on the elapsed time between the observations, $t_q - t_p$. That $-\ell$ and u increase in the elapsed time (i.e., their second argument) reflects more variability in process values due to passage of time.

To see that Assumption 2(i) is satisfied, note that the uncertainty set in (3) is given by a finite collection of bimonotone linear inequalities,[5] and is therefore a lattice (Queyranne and Tardella 2006). Assumption 2(ii) is readily satisfied since both $\alpha\, x + \ell(t_p, t_q - t_p)$ and $\beta\, x + u(t_p, t_q - t_p)$ are increasing in x. Lastly, condition (2)—and thus Assumption 2(iii)—is also satisfied since $-\ell$ and u are increasing in their second argument.

Sets such as (3) are a natural extension of the typical box uncertainty model (Ben-Tal et al. 2009) to our dynamic setting. The example generalizes to bounding functions that depend on the pair of components (i.e., we have $\alpha^{m,m'}, \beta^{m,m'}, \ell^{m,m'}$ and $u^{m,m'}$), and constraints written more generally as $\tilde{\ell}(x_p^m, t_p, t_q - t_p) \leq x_q^{m'} \leq \tilde{u}(x_p^m, t_p, t_q - t_p)$, where $\tilde{\ell}$ is convex and increasing in its first argument, and $\tilde{u}$ is concave and increasing in its first argument. We omit more details for simplicity of exposition.

[5] A linear inequality in variables $\boldsymbol{x}$ is bimonotone if it can be written as $a_i x_i + a_j x_j \leq c$, for some i, j, with $a_i \cdot a_j \leq 0$.

Example 2 (CLT-Budgeted Uncertainty Sets). Consider the case where the d processes correspond to random walks with independent increments. In particular, consider $d=1$, and assume that increments corresponding to a time length Δt have mean $\mu\Delta t$ and standard deviation $\sigma\sqrt{\Delta t}$. As such, by expressing the difference between two observations x_q and x_p (for $q>p$) as a sum of intermediate independent increments, i.e., $x_q - x_p = \sum_{i=p+1}^{q}(x_i - x_{i-1})$, one can then rely on the ideas introduced by Bandi and Bertsimas (2012) to formulate the following Central-Limit-Theorem-type uncertainty set:

$$\mathcal{U}(\boldsymbol{t}^{\{r\}}, \boldsymbol{x}^{\{k\}}) = \Big\{[x_{k+1}, \dots, x_r, x_{n+1}] \in \mathbb{R}^{r-k+1} : -\Gamma \le \frac{x_q - x_p - (t_q - t_p)\mu}{\sqrt{t_q - t_p}\sigma} \le \Gamma, \\ \forall p, q \in \{0, 1, \dots, r, n+1\},\, p<q\Big\},$$

for some $\Gamma > 0$. Note that this set can be reformulated as a lattice uncertainty set with $\alpha = \beta = 1$, $\ell(t_p, t_q - t_p) = (t_q - t_p)\mu - \Gamma\sqrt{t_q - t_p}\sigma$, and $u(t_p, t_q - t_p) = (t_q - t_p)\mu + \Gamma\sqrt{t_q - t_p}\sigma$. Assumption 2(ii) is immediate, and by Proposition 1, Assumption 2(iii) is satisfied when the budget is sufficiently high ($\Gamma\sigma \ge 2\mu\sqrt{T}$).

3. Analysis

We begin by formulating the two problems—for dynamic and static monitoring—as dynamic programs.

3.1. Dynamic Programming Formulations

Dynamic Monitoring. Let $J_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}})$ denote the worst-case value-to-go at time t_k, with the first k observations $\boldsymbol{x}^{\{k\}}$ made at times $\boldsymbol{t}^{\{k\}}$. The Bellman recursions become:

$$J_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}}) = \max\Big(g\big(t_k, \boldsymbol{x}_k\big), \max_{t_{k+1} \in (t_k, T]} \min_{\boldsymbol{x}_{k+1} \in \mathcal{U}_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}})} J_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k+1\}})\Big), \tag{4}$$

where $J_{n+1}(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{n+1\}}) = g\big(t_{n+1}, \boldsymbol{x}_{n+1}\big)$. For an intermediate monitoring time t_k, $J_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}})$ is obtained as the maximum of the reward from stopping at t_k and the worst-case continuation value obtained by (optimally) picking the next monitoring time t_{k+1}. The DM's problem is then to find a set of monitoring and stopping policies $\Pi^D = \{\Pi_k^D\}_{k=0}^n$, where Π_k^D consists of a stopping policy at time t_k and a policy $\tau_k^D(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}})$ for choosing the next monitoring time t_{k+1}. More precisely,

$$\tau_k^D(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}}) = \operatorname*{arg\,max}_{t_{k+1} \in (t_k, T]} \min_{\boldsymbol{x}_{k+1} \in \mathcal{U}_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}})} J_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k+1\}}).$$

Let $J = J_0(t_0, \boldsymbol{x}^{\{0\}})$.

Static Monitoring. Suppose all monitoring times were chosen as $\boldsymbol{t}^{\{n+1\}}$. Let $V_k(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{k\}})$ denote the worst-case value-to-go function at time t_k with the first k observations $\boldsymbol{x}^{\{k\}}$ made. The Bellman equations under static monitoring become:

$$V_k(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{k\}}) = \max\Big(g\big(t_k, \boldsymbol{x}_k\big), \min_{\boldsymbol{x}_{k+1} \in \mathcal{U}_{k+1}(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{k\}})} V_{k+1}\big(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{k+1\}}\big)\Big), \tag{5}$$

where $V_{n+1}(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{n+1\}}) = g\big(t_{n+1}, \boldsymbol{x}_{n+1}\big)$. The DM's problem is to choose a set of stopping policies $\Pi^S = \{\Pi_k^S\}_{k=0}^n$ and a vector of monitoring times $\boldsymbol{t}^S \in \arg\max_{\boldsymbol{t}^{\{n+1\}}} V_0(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{0\}})$. Let $V = V_0(\boldsymbol{t}^S, \boldsymbol{x}^{\{0\}})$.

3.2. Optimal Policy Under Dynamic Monitoring

We will show that the optimal policy under dynamic monitoring can be recovered by dynamically (re)solving a sequence of static monitoring problems. The first step is to show that the optimal worst-case reward achieved under dynamic monitoring is the same as under static monitoring.

THEOREM 1. *Under Assumption 1 and Assumption 2, we have*

$$J = V, \qquad t_1^S \in \tau_1^D(t_0, \boldsymbol{x}_0).$$

Proof of Theorem 1. It is trivial that $J \geq V$, since any optimal solution for the static model is also feasible for the dynamic model. To show the opposite direction, define

$$\begin{aligned} J_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}}) &= \max\Big(g\big(t_k, \boldsymbol{x}_k\big),\, C_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}})\Big) \\ C_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}}) &\overset{\text{def}}{=} \max_{t_{k+1} \in [t_k, T]} G_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}}) \\ G_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}}) &\overset{\text{def}}{=} \min_{\boldsymbol{x}_{k+1} \in \mathcal{U}_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}})} J_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k+1\}}). \end{aligned}$$

The proof relies on the following auxiliary result, which we prove in Lemma 1: for any $\boldsymbol{t}^{\{k+1\}}$ and $\boldsymbol{x}^{\{k-1\}}$, with $\underline{\boldsymbol{x}}_k \overset{\text{def}}{=} \min\big\{\boldsymbol{x} : \boldsymbol{x} \in \mathcal{U}_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k-1\}})\big\}$, we have

$$\begin{aligned} \underline{\boldsymbol{x}}_k \in \arg\min_{\boldsymbol{x}_k \in \mathcal{U}_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k-1\}})} g(t_k, \boldsymbol{x}_k) \quad &\cap \arg\min_{\boldsymbol{x}_k \in \mathcal{U}_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k-1\}})} C_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}}) \\ &\cap \arg\min_{\boldsymbol{x}_k \in \mathcal{U}_k(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k-1\}})} G_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}}). \end{aligned} \tag{6}$$

We then have:

$$
\begin{aligned}
G_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k-1\}}) &= \min_{\boldsymbol{x}_k \in \mathcal{U}_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k-1\}})} J_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}}) \\
&\overset{(6)}{=} \max\Big(g\big(t_k, \underline{\boldsymbol{x}}_k\big),\, C_k(\boldsymbol{t}^{\{k\}}, [\boldsymbol{x}^{\{k-1\}}, \underline{\boldsymbol{x}}_k])\Big) \\
&= \max\Big(g\big(t_k, \underline{\boldsymbol{x}}_k\big),\, \max_{t_{k+1} \in [t_k, T]} G_{k+1}(\boldsymbol{t}^{\{k+1\}}, [\boldsymbol{x}^{\{k-1\}}, \underline{\boldsymbol{x}}_k])\Big) \\
\text{(since } \underline{\boldsymbol{x}}_k \text{ independent of } t_{k+1}) \quad &= \max_{t_{k+1} \in [t_k, T]} \max\Big(g\big(t_k, \underline{\boldsymbol{x}}_k\big),\, G_{k+1}(\boldsymbol{t}^{\{k+1\}}, [\boldsymbol{x}^{\{k-1\}}, \underline{\boldsymbol{x}}_k])\Big) \\
&\overset{(6)}{\le} \max_{t_{k+1} \in [t_k, T]} \min_{\boldsymbol{x}_k \in \mathcal{U}_k(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k-1\}})} \max\Big(g\big(t_k, \boldsymbol{x}_k\big),\, G_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}})\Big).
\end{aligned}
$$

The argument then follows by induction. ■

LEMMA 1. *Consider any $1 \le k \le n$, any $\boldsymbol{t}^{\{k+1\}}$ and $\boldsymbol{x}^{\{k-1\}}$, and let*

$$\underline{\boldsymbol{x}}_k \overset{\text{def}}{=} \min\big\{\boldsymbol{x} : \boldsymbol{x} \in \mathcal{U}_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k-1\}})\big\}.$$

Then,

$$\underline{\boldsymbol{x}}_k \in \operatorname*{arg\,min}_{\boldsymbol{x}_k \in \mathcal{U}_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k-1\}})} g(t_k, \boldsymbol{x}_k) \cap \operatorname*{arg\,min}_{\boldsymbol{x}_k \in \mathcal{U}_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k-1\}})} C_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}}) \cap \operatorname*{arg\,min}_{\boldsymbol{x}_k \in \mathcal{U}_k(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k-1\}})} G_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}}).$$

Proof Note that $\underline{\boldsymbol{x}}_k$ is well defined since $\mathcal{U}_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k-1\}})$ is a lattice, by Assumption 2(i). Since $g(t_k, \boldsymbol{x}_k)$ is increasing in $\boldsymbol{x}_k$ for any t_k, it suffices to prove that $C_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}})$ and $G_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}})$ are increasing functions, for any k and any $\boldsymbol{t}^{\{k+1\}}$. To the latter point, we claim that it suffices to show that $J_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k+1\}})$ is increasing in $\boldsymbol{x}^{\{k+1\}}$; when this holds, we readily have that:

$$G_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}}) \overset{\text{def}}{=} \min_{\boldsymbol{x}_{k+1} \in \mathcal{U}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}})} J_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k+1\}}) = J_{k+1}(\boldsymbol{t}^{\{k+1\}}, [\boldsymbol{x}^{\{k\}}, \underline{\boldsymbol{x}}_{k+1}])$$

is increasing in $\boldsymbol{x}^{\{k\}}$, since $\underline{\boldsymbol{x}}_{k+1}$ is increasing in $\boldsymbol{x}^{\{k\}}$, by Assumption 2(ii). And thus

$$C_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}}) \overset{\text{def}}{=} \max_{t_{k+1} \in [t_k, T]} G_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}})$$

is also increasing in $\boldsymbol{x}^{\{k\}}$, as a maximum of increasing functions.

To complete our proof, it thus suffices to show that $J_k(\boldsymbol{t}^{\{k\}}, \cdot)$ is increasing, for any k and any $\boldsymbol{t}^{\{k+1\}}$. We prove this by induction. For $k = n+1$, we have that $J_{n+1}(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{n+1\}}) \overset{\text{def}}{=} g(t_{n+1}, \boldsymbol{x}_{n+1})$, so $J_{n+1}(\boldsymbol{t}^{\{n+1\}}, \cdot)$ is increasing. Assume the property holds for J_{k+1}. Then,

$G_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}})$ and $C_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}})$ are both increasing in $\boldsymbol{x}^{\{k\}}$, by the argument above. And since $g(t_k, \boldsymbol{x}_k)$ is also increasing in $\boldsymbol{x}_k$ by Assumption 1, we have that

$$J_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}}) = \max\Big(g\big(t_k, \boldsymbol{x}_k\big), C_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}})\Big)$$

is also increasing in $\boldsymbol{x}^{\{k\}}$, completing our inductive step. ■

Although static monitoring is able to achieve the same worst-case reward as dynamic monitoring and to generate an optimal initial monitoring time t_1^S, the subsequent optimal static monitoring times $t_2^S, \ldots, t_n^S$ are not necessarily dynamically consistent. In particular, if the process values observed at t_1^S do not correspond to nature's worst-case actions, the DM may prefer to adjust the second monitoring time to a value that differs from t_2^S (and the same rationale applies at subsequent monitoring times). However, this issue can be addressed by re-solving a static monitoring problem over the remaining horizon, using the updated information. This intuition is formalized in the following result.

Corollary 1. *The following algorithm yields an optimal dynamic monitoring policy: For all $k = 0, 1, \ldots, n$:*

1. *At the k-th monitoring time t_k, having observed $\boldsymbol{x}_k$, find an optimal static monitoring policy over the remaining time horizon $[t_k, t_{n+1}]$, with $n-k$ monitoring chances, initial state $\boldsymbol{x}_k$, and observation matrix $\boldsymbol{x}^{\{k\}}$. Let $V(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}})$ denote the worst-case optimal reward, and $t_1^S(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}})$ denote the first monitoring time under this policy.*
2. *If $g(t_k, \boldsymbol{x}_k) \geq V(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}})$, then stop. Otherwise, continue and $\tau_k^D(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}}) \leftarrow t_1^S(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}})$.*

This result underscores the importance of solving the static monitoring problem, which we analyze next.

3.3. Optimal Policy Under Static Monitoring

Our approach is to first characterize nature's optimal actions and the DM's optimal stopping strategy for a given set of monitoring times. This will allow us to then reformulate (and simplify) the DM's problem to one of choosing only the monitoring times.

Proposition 2. *Consider a fixed set of monitoring times $\boldsymbol{t}^{\{n+1\}}$. Nature's optimal (worst-case) response when the DM continues in period t_k can be recovered recursively as:*

$$\underline{\boldsymbol{x}}_k(\boldsymbol{t}^{\{n+1\}}) \stackrel{\text{def}}{=} \min\Big\{\boldsymbol{\xi} \in \mathbb{R}^d : \boldsymbol{\xi} \in \mathcal{U}_k\big(\boldsymbol{t}^{\{n+1\}}, [\boldsymbol{x}_0, \underline{\boldsymbol{x}}_1(\boldsymbol{t}^{\{n+1\}}), \ldots, \underline{\boldsymbol{x}}_{k-1}(\boldsymbol{t}^{\{n+1\}})]\big)\Big\}, 1 \leq k \leq n+1. \tag{7}$$

Proof According to Lemma 1, the worst-case process value at time t_k, which we denote by $\underline{\boldsymbol{x}}_k(\boldsymbol{t}^{\{n+1\}}, \underline{\boldsymbol{x}}^{\{k-1\}})$, can be obtained by choosing the smallest element of the corresponding uncertainty set, i.e.,

$$\underline{\boldsymbol{x}}_k(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{k-1\}}) \stackrel{\text{def}}{=} \min\Big\{\boldsymbol{\xi} \in \mathbb{R}^d : \boldsymbol{\xi} \in \mathcal{U}_k\big(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{k-1\}}\big)\Big\}, \forall k \in \{1, \ldots, n+1\}.$$

The result follows by induction. ■

A given set of monitoring times thus induces a particular (predictable) worst-case path, and the DM's stopping problem reduces to choosing when to stop along that path. The DM's problem (5) can thus be re-formulated as:

$$V = \max_{\boldsymbol{t}^{\{n+1\}}} \max_{k \in \{1, \ldots, n+1\}} g(t_k, \underline{\boldsymbol{x}}_k(\boldsymbol{t}^{\{n+1\}})). \tag{8}$$

Our next result further simplifies this problem, proving that under an optimally chosen set of monitoring times and along the resulting worst-case path, it is optimal either to stop at the last monitoring time t_n or to continue until the end.

THEOREM 2. *The optimal value in (8) can be obtained as*

$$V = \max_{\boldsymbol{t}^{\{n+1\}}} \max_{k \in \{n, n+1\}} g\big(t_k, \underline{\boldsymbol{x}}_k(\boldsymbol{t}^{\{n+1\}})\big). \tag{9}$$

Proof We claim that there always exists an optimal choice of $\boldsymbol{t}^{\{n+1\}}$ such that the inner maximum in (8) is reached at t_n or $t_{n+1} = T$ (i.e., for $k = n$ or $n+1$, respectively). To see this, assume that the maximum occurs at $\bar{k} < n$, and introduce a new set of monitoring times $\boldsymbol{z}$ such that $\boldsymbol{z} = [t_0, t_1, \ldots, t_{\bar{k}-1}, t_{\bar{k}-1}, t_{\bar{k}}, t_{\bar{k}+2}, \ldots, t_n, t_{n+1}]$. Then,

$$V \geq \max_{k \in \{1, \ldots, n+1\}} g(z_k, \underline{\boldsymbol{x}}_k(\boldsymbol{z})) \stackrel{(8)}{\geq} g(\boldsymbol{z}_{\bar{k}+1}, \underline{\boldsymbol{x}}_{\bar{k}+1}(\boldsymbol{z})) = g(t_{\bar{k}}, \underline{\boldsymbol{x}}_{\bar{k}}(\boldsymbol{t}^{\{n+1\}})) = V.$$

The penultimate equality holds because $\mathcal{U}_{\bar{k}}\big(\boldsymbol{t}^{\{n+1\}}, \underline{\boldsymbol{x}}^{\{\bar{k}-1\}}\big) = \mathcal{U}_{\bar{k}+1}\big(\boldsymbol{z}, \underline{\boldsymbol{x}}^{\{\bar{k}\}}\big)$, by the construction of $\boldsymbol{z}$ and by the definition of $\mathcal{U}$ in (1). Thus, under the monitoring times $\boldsymbol{z}$, the same optimum is reached at $\bar{k}+1$. Repeating the argument inductively yields the result. ■

Theorem 2 implies that the static problem entails solving two optimization problems. Since $t_{n+1} \equiv T$ is fixed, this can be done provided we can find a solution for the problem:

$$\max_{\boldsymbol{t}^{\{n+1\}}} g\big(t_n, \underline{\boldsymbol{x}}_n(\boldsymbol{t}^{\{n+1\}})\big). \tag{10}$$

Solution approaches critically depend on the structure of the reward function g and the uncertainty sets. For our subsequent analysis, we focus on the class of lattice uncertainty sets with cross constraints introduced in Example 1, which subsumes many interesting sets considered in the literature (e.g., sets with box constraints and CLT-budgeted sets, per our discussion in §2.5). For simplicity, we take $\alpha = \beta = 1$, so that:

$$\begin{aligned} \mathcal{U}\big(\boldsymbol{t}^{\{r\}}, \boldsymbol{x}^{\{k\}}\big) = \Big\{ & [\boldsymbol{x}_{k+1}, \ldots, \boldsymbol{x}_r, \boldsymbol{x}_{n+1}] \in \mathbb{R}^{d\times(r-k+1)} : \\ & x_p^m + \ell(t_p, t_q - t_p) \le x_q^{m'} \le x_p^m + u(t_p, t_q - t_p), \\ & \forall\, (m, m') \in \mathcal{M},\, \forall p \in \{0, 1, \ldots, k\},\, \forall q \in \{k+1, \ldots, r\} \Big\}. \end{aligned} \tag{11}$$

To rule out degenerate cases where the uncertainty sets could be empty, we require that $\ell(\cdot, 0) = u(\cdot, 0) = 0$. Let $\underline{x}_0^{m'} \stackrel{\text{def}}{=} \max_{m:(m,m')\in\mathcal{M}} x_0^m, \forall\, m' \in \{1, \ldots, d\}$.

3.3.1. Stationary Bound Functions. We first discuss the case where ℓ and u only depend on their second argument, i.e., the elapsed time between monitoring opportunities; with a slight overload of notation, we write $\ell(t_q - t_p)$ and $u(t_q - t_p)$. This already captures many uncertainty sets of practical interest, such as the CLT-budgeted sets of Example 2.

THEOREM 3. *Under Assumption 1 and for the uncertainty set in (11),*

(i) if $\ell(\cdot)$ is convex, then

$$V = \max_{t\in[0,T]} g\big(t, \underline{\boldsymbol{x}}_0 + \ell(t)\mathbf{1}\big), \qquad t_1^S \in \arg\max_{t\in[0,T]} g\big(t, \underline{\boldsymbol{x}}_0 + \ell(t)\mathbf{1}\big), \tag{12}$$

and $t_2^S, \ldots, t_n^S$ can be chosen arbitrarily from $[t_1^S, T]$;

(ii) if $\ell(\cdot)$ is concave, then

$$V = \max\Big(\max_{t\in[0,T]} g\big(t, \boldsymbol{\phi}_n(t)\big),\, g\big(T, \boldsymbol{\phi}_{n+1}(T)\big) \Big), \tag{13}$$

where $\boldsymbol{\phi}_k(t) \stackrel{\text{def}}{=} \underline{\boldsymbol{x}}_0 + k\ell\big(t/k\big)\mathbf{1}$, $\forall\, k \in \{n, n+1\}$. Moreover, if $g(t, \boldsymbol{x})$ is jointly concave, then (13) reduces to solving a convex optimization problem.

Proof For the convex case in (i), please refer to the proof of Theorem 4 for a more general result. For the concave case in (ii), consider problem (10), $\max_{\boldsymbol{t}^{\{n+1\}}} g(t_n, \underline{\boldsymbol{x}}_n(\boldsymbol{t}^{\{n+1\}}))$. By Theorem 4, $\underline{\boldsymbol{x}}_n(\boldsymbol{t}^{\{n+1\}}) = \underline{\boldsymbol{x}}_0 + \sum_{i=1}^n \ell(t_i - t_{i-1})\mathbf{1}$ when ℓ is concave. Thus, for a fixed t_n, maximizing $g(t_n, \underline{\boldsymbol{x}}_n(\boldsymbol{t}^{\{n+1\}}))$ is equivalent to maximizing $\sum_{i=1}^n \ell(t_i - t_{i-1})$, since $g(t, \boldsymbol{x})$ is increasing in $\boldsymbol{x}$. Due to the concavity of $\ell(\delta)$, the maximum is achieved when $t_i - t_{i-1} = t_n/n$, $i = 1, \ldots, n$ (by Jensen's inequality). Therefore, problem (10) becomes equivalent to $\max_{t_n} g(t_n, \boldsymbol{\phi}_n(t_n))$. Similarly, $\max_{\boldsymbol{t}^{\{n+1\}}} g(t_{n+1}, \underline{\boldsymbol{x}}_{n+1}(\boldsymbol{t}^{\{n+1\}})) = g(T, \boldsymbol{\phi}_{n+1}(T))$. ■

According to Theorem 3, when the bound functions have well-defined curvatures, solving the static monitoring problem reduces to optimizing a one-dimensional function.

When ℓ is convex, it is worst-case optimal for the DM to stop at the first monitoring time, found by solving (12), which renders the choice of subsequent monitoring times irrelevant. Intuitively, this occurs because a convex lower bound function corresponds to a decreasing rate of uncertainty growth over time. This results in the initial estimates from time t_0 providing the least conservative lower bound estimate of the future process values, and makes any additional observations irrelevant for the DM's problem of "learning" the worst-case path. An example of such convex bound functions is the CLT-budgeted uncertainty set in Example 2, where $\ell(\delta) = \delta\mu - \Gamma\sqrt{\delta}\sigma$.

When ℓ is concave, it is worst-case optimal for the DM to evenly space the monitoring times, stopping either at the last monitoring time t_n (chosen to maximize $g(t, \phi_n(t))$) or at the end of the planning horizon, T. Intuitively, a concave lower bound induces an increasing rate of uncertainty growth over time, so that worst-case estimates of future process values quickly degrade. Thus, new observations can substantially reduce the uncertainty, and the optimal policy avails itself of all monitoring chances, distributing them uniformly over time so as to maximally reduce the uncertainty and improve worst-case outcomes.

3.3.2. Non-stationary Bound Functions. We now treat the more general case, where the bounds depend on both the monitoring time and the elapsed time, i.e., $\ell(t_p, t_q - t_p)$ and $u(t_p, t_q - t_p)$. As expected, additional conditions are required here to characterize the solutions in the static monitoring problem. Our first result parallels Theorem 3, highlighting the importance of the bounds' curvature.

THEOREM 4. *Under Assumption 1 and for the uncertainty set in (11),*

(i) if $\ell(t,\delta)$ is decreasing in t and convex in δ, then

$$V = \max_{t\in[0,T]} g\big(t, \underline{\boldsymbol{x}}_0 + \ell(t_0, t-t_0)\mathbf{1}\big), \qquad t_1^S \in \arg\max_{t\in[0,T]} g\big(t, \underline{\boldsymbol{x}}_0 + \ell(t_0, t-t_0)\mathbf{1}\big),$$

and $t_2^S, \ldots, t_n^S$ can be chosen arbitrarily from $[t_1^S, T]$;

(ii) if $\ell(t,\delta)$ is increasing in t and concave in δ, then

$$V = \max\Big(\max_{t_n\in[0,T]} g\big(t_n, \boldsymbol{x}_0 + \xi_n(t_n)\mathbf{1}\big),\, g\big(T, \boldsymbol{x}_0 + \xi_{n+1}(T)\mathbf{1}\big)\Big), \quad \textit{where} \tag{14a}$$

$$\xi_k(t_k) \stackrel{\text{def}}{=} \max_{0\le t_1\cdots\le t_{k-1}\le t_k} \left[\sum_{i=1}^{k} \ell(t_{i-1}, t_i - t_{i-1})\right], \forall\, k\in\{n, n+1\}. \tag{14b}$$

Furthermore, if ℓ is jointly concave, then (14b) is a convex optimization problem; and if $g(t,\boldsymbol{x})$ is jointly concave, then (14a) reduces to solving a convex optimization problem.

Proof For case (i), we first show that $\underline{\boldsymbol{x}}_k(\boldsymbol{t}^{\{n+1\}}) = \underline{\boldsymbol{x}}_0 + \ell(t_0, t_k - t_0)\mathbf{1}$. To that end, note that for any $j \in \{1,\ldots,d\}$ and $1 \le p < q \le n$:

$$x_0^j + \ell(t_0, t_p - t_0) + \ell(t_p, t_q - t_p) \le x_0^j + \ell(t_0, t_p - t_0) + \ell(t_0, t_q - t_p) \le x_0^j + \ell(t_0, t_q - t_0), \quad (15)$$

where the first inequality follows since $\ell(t,\delta)$ is decreasing in t, and the second inequality follows because ℓ is superadditive in δ (since ℓ is convex in δ and $\ell(\cdot,0)=0$). Therefore,

$$\begin{aligned}\underline{\boldsymbol{x}}_k(\boldsymbol{t}^{\{n+1\}}) &\stackrel{\text{def}}{=} \min\Big\{\boldsymbol{\xi} : \boldsymbol{\xi} \in \mathcal{U}_k\big(\boldsymbol{t}^{\{n+1\}}, [\boldsymbol{x}_0, \underline{\boldsymbol{x}}_1, \ldots, \underline{\boldsymbol{x}}_{k-1}]\big)\Big\} \\ &= \underline{\boldsymbol{x}}_0 + \left[\max_{\{0=k_1 \le \cdots \le k_r = k\} \in \{1,\ldots,k\}} \sum_{i=1}^{r} \ell(t_{k_i - 1}, t_{k_i} - t_{k_i - 1})\right]\mathbf{1} = \underline{\boldsymbol{x}}_0 + \ell(t_0, t_k - t_0)\mathbf{1}.\end{aligned}$$

For case (ii), note that when $\ell(t,\delta)$ is increasing in t and concave in δ, the reverse inequalities hold in (15). Thus, nature's optimal (worst-case) response is given by:

$$\underline{\boldsymbol{x}}_k(\boldsymbol{t}^{\{n+1\}}) = \underline{\boldsymbol{x}}_0 + \left[\sum_{i=1}^{k} \ell(t_{i-1}, t_i - t_{i-1})\right]\mathbf{1}.$$

Since $g(t,\boldsymbol{x})$ is jointly concave and increasing in $\boldsymbol{x}$, and $\ell(t,\delta)$ is jointly concave, the composition $g\big(t_n, \underline{\boldsymbol{x}}_n(\boldsymbol{t}^{\{n+1\}})\big)$ is concave in $\boldsymbol{t}^{\{n+1\}}$, so that problem (10) requires maximizing a concave function over the convex set $0 \le t_1 \le \cdots \le t_n \le T$. ■

Part (i) shows that when the bound $\ell(t,\delta)$ is convex in δ and also *decreasing in* t, it would again be worst-case optimal for the DM to stop at the first monitoring time, found by solving a one-dimensional optimization problem; as before, the choice of subsequent monitoring times is irrelevant for maximizing the worst-case reward. Intuitively, that bound functions would decrease in t if the forecasting technology improved over time, or the processes became "more predictable."

Part (ii) shows that when $\ell(t,\delta)$ is concave in δ and also *increasing in* t, it would be worst-case optimal for the DM to rely on all monitoring chances, which are picked so as to reduce the uncertainty as much as possible, per (14b). When ℓ and g satisfy additional (mild) requirements, the requisite problems have a convex structure, and are tractable.

In the absence of curvature information about ℓ, Problem (10) can also be viewed from a combinatorial optimization perspective. In particular, if the function $g\big(t_n, \underline{\boldsymbol{x}}_n(\boldsymbol{t}^{\{n+1\}})\big)$ is a supermodular function of $\boldsymbol{t}^{\{n+1\}}$, Problem (10) would involve maximizing a supermodular function over a lattice—a problem that can be tackled through tractable combinatorial algorithms (see, e.g., Fujishige (2005), Schrijver (2003), and references therein). Although it is difficult to exactly characterize when $g\big(t_n, \underline{\boldsymbol{x}}_n(\boldsymbol{t}^{\{n+1\}})\big)$ becomes supermodular, we provide a set of sufficient conditions in the following result.

PROPOSITION 3. *Under Assumption 1 and for the uncertainty set in (11), if:*
(i) $g(t,\boldsymbol{x})$ *is supermodular in* $\boldsymbol{x}$*, convex in every component* x_k*, and*
(ii) ℓ *is supermodular and decreasing in* (t_p, t_q) *on the lattice* $[0,T]^2$*,*
then $g\big(t_k, \underline{\boldsymbol{x}}_k(\boldsymbol{t}^{\{k\}})\big)$ *is a supermodular function of* $\boldsymbol{t}^{\{k-1\}}$ *for any fixed* t_k*, for* $k \in \{n, n+1\}$.

Proof Our proof relies on the following known result from lattice programming.

PROPOSITION 4 **(Topkis 1998)**. *If h is a convex, increasing (i.e., isotone) function, and f is supermodular and either isotone or antitone, then $h \circ f$ is supermodular.*

Note that every component of $\underline{\boldsymbol{x}}_n$ can be written as:

$$\underline{\boldsymbol{x}}_n^k = \underline{\boldsymbol{x}}_0^k + \max_{s\in\mathcal{S}} f_s(\boldsymbol{t}^{\{n\}})$$
$$f_s(\boldsymbol{t}^{\{n\}}) \stackrel{\text{def}}{=} \sum_{i=0}^{p-1} \ell(t_{s(i)}, t_{s(i+1)} - t_{s(i)}),$$

where $\mathcal{S}$ denotes the set of all ordered subsets of $\{0,\dots,n\}$ that include 0 and n. By assumption (ii), for every $s \in \mathcal{S}$, f_s is supermodular in $\boldsymbol{t}^{\{n\}}$ and decreasing (i.e., antitone) in $\boldsymbol{t}^{\{n\}}$. Since the max function is convex and increasing (i.e., isotone), we can invoke Proposition 4 to conclude that $\underline{\boldsymbol{x}}_n^k$ is supermodular in $\boldsymbol{t}^{\{n\}}$. Since $\underline{\boldsymbol{x}}_n^k$ is also decreasing in $\boldsymbol{t}^{\{n\}}$, and $g(t,\boldsymbol{x})$ is increasing and component-wise convex in $\boldsymbol{x}$, we can again invoke Proposition 4 to conclude that $g\big(t_n, \underline{\boldsymbol{x}}_n(\boldsymbol{t}^{\{n\}})\big)$ is supermodular in $\boldsymbol{t}^{\{n-1\}}$ for any fixed t_n. ■

To solve Problem (10), one should thus follow a two-stage approach.[6] In the second stage, for a fixed t_n, one can rely on Proposition 3 to find $t_1,\dots,t_{n-1}$, by maximizing the supermodular function $g\big(t_n, \underline{\boldsymbol{x}}_n(\boldsymbol{t}^{\{n\}})\big)$ over the lattice $0 \le t_1 \le t_2 \le \cdots \le t_{n-1} \le t_n$. In the first stage, one would solve a one-dimensional optimization problem over t_n.

In practice, this combinatorial approach could be quite useful, since t_n may be required to be integer or a multiple of some base planning period, such as a month, a week, a

[6] In case $g(t,\boldsymbol{x})$ is also increasing in t and (jointly) convex in $(t,\boldsymbol{x})$, then this two-step approach is not needed, since $g\big(t_n, \underline{\boldsymbol{x}}_n(\boldsymbol{t}^{\{n\}})\big)$ is actually supermodular in $\boldsymbol{t}^{\{n\}}$, by Proposition 4.

day, etc. Furthermore, some of the premises of Proposition 3 are reasonable. For instance, having rewards g that are supermodular—i.e., exhibiting complementarity in the states—is natural in many healthcare applications, where better performance in one biometric can relate to better (marginal) performance in the others. In CAV monitoring, for example, the number of acute rejections and CAV stage are complementary (John 2009). Component-wise convexity of g suggests increasing returns to scale in each component of $\boldsymbol{x}$, which could be expected particularly for low values of biometrics in healthcare applications. The supermodularity of ℓ implies that the uncertainty is also complementary.

Finally, when none of the approaches above is applicable, one may still be able to formulate the static monitoring problem as a mixed-integer linear program, and rely on modern MIP solvers to determine the monitoring times. This approach works particularly well when the primitives (i.e., the functions g and ℓ) are piece-wise linear. Since the resulting model heavily depends on the problem specifics, we omit a general-purpose formulation here; for a concrete example, please see our numerical study in §5.

4. Extensions

We explore several extensions of our model that may be relevant in practice.

4.1. Costly Monitoring

Consider our base model, but assume that each monitoring incurs a fixed cost $c \geq 0$. In particular, the reward from stopping at the k-th monitoring time t_k with a state of $\boldsymbol{x}_k$ becomes $g(t_k, \boldsymbol{x}_k) - kc$, for all $k = 1, \ldots, n$. Then, it can be readily checked that Theorem 1 continues to hold, so that the worst-case optimal rewards under dynamic and static monitoring are the same. Thus, it suffices again to focus on the static monitoring problem.

Let $\tilde{V}_k^c(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{k\}})$ be the worst-case value-to-go function for the static monitoring problem at time t_k, with the first k observations made. The Bellman equations become:

$$\tilde{V}_k^c(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{k\}}) = \max\left(g(t_k, \boldsymbol{x}_k) - ck, \min_{\boldsymbol{x}_{k+1} \in \tilde{\mathcal{U}}_{k+1}(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{k\}})} \tilde{V}_{k+1}^c(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{k+1\}}) \right),$$
$$\tilde{V}_{n+1}^c(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{n+1\}}) = g(t_{n+1}, \boldsymbol{x}_{n+1}) - c\,(n+1).$$

It can be readily checked that, under a given set of monitoring times $\boldsymbol{t}^{\{n+1\}}$, nature's optimal (worst-case) response $\underline{\boldsymbol{x}}_k(\boldsymbol{t}^{\{n+1\}})$ in period t_k is still given by Proposition 2. Thus, the DM's problem can again be reformulated as:

$$\tilde{V}_0^c(t_0, \boldsymbol{x}_0) = \max_{\boldsymbol{t}^{\{n+1\}}} \max_{k \in \{1, \ldots, n+1\}} \left[g\big(t_k, \underline{\boldsymbol{x}}_k(\boldsymbol{t}^{\{n+1\}})\big) - c\,k \right]. \tag{16}$$

A key difference from our earlier results lies in the DM's optimal stopping strategy. Recall that in our base model, it was worst-case optimal to either stop at the last monitoring time t_n or continue until the end of the horizon (see Theorem 2). This is no longer the case here, as an optimal policy may require stopping at an earlier time due to the monitoring cost c. Thus, the optimal k^* in (16) may be strictly smaller than n.

To solve (16), one can switch the order of the maximization operators. Since finding the optimal $\boldsymbol{t}^{\{n+1\}}$ for a fixed k requires solving problems that are structurally identical to Problem (10), our results in §3.3.1 and §3.3.2 can be directly leveraged. By iterating over k, one can then recover the optimal number of monitoring times.

Moreover, when the bounds are stationary, the problem of finding the optimal number of monitoring times is also tractable under mild conditions, as summarized next.

Proposition 5. *Under Assumption 1 and for the uncertainty set in (11) with stationary lower bounds $\ell(t_q - t_p)$,*

(i) if $\ell(\cdot)$ is convex, then a single monitoring time is sufficient for achieving the worst-case optimal reward;

(ii) if $\ell(\cdot)$ is concave, and $g(t,\boldsymbol{x})$ is jointly concave, then the optimal number of monitoring times can be done by solving convex optimization problems.

Proof. Part (i) follows directly from Theorem 3(i). For part (ii), recall from Theorem 3(ii) that finding the optimal stopping time under a fixed number of monitoring times n requires solving the problem $\max_{t\in[0,T]} g\big(t, \underline{\boldsymbol{x}}_0 + n\ell\big(t/n\big)\mathbf{1}\big)$. The function to be maximized is jointly concave in (t,n), since $g(t,\boldsymbol{x})$ is jointly concave and increasing in $\boldsymbol{x}$, and the functions $n\ell(t/n)$ are jointly concave in (t,n) since ℓ is concave. Thus, one can find an optimal t and n by first maximizing a concave function over a convex feasible set (considering n continuous), and then checking the nearest integers (possibly solving two additional one dimensional convex optimization problems to determine the corresponding t). ■

4.2. More General Decision Process

Some of our results concerning the monitoring policy also extend to a more general decision problem, where the DM, instead of simply stopping, can modify the processes by increasing the state values (an action we refer to as "injection") or decreasing them ("extraction"). This allows capturing several applications of interest. In chronic disease monitoring, "injections" could denote interventions that are costly or have immediate side-effects in the short run, but carry long-term benefits, while "extractions" could capture relaxing a strict treatment, leading to immediate relief but carrying potential long-term consequences. In collateralized lending, "injections" could denote the costly addition of new collateral, which improves the borrowing base, and "extractions" could denote immediate collateral liquidations, which generate cash but reduce the borrowing base.

To formalize this, consider our setup in §2, but assume that at the k-th monitoring time t_k, upon observing the state value $\boldsymbol{x}_k \stackrel{\text{def}}{=} \boldsymbol{x}(t_k)$, the DM decides an action $\boldsymbol{y}_k \in A(\boldsymbol{x}_k) \subseteq \mathbb{R}^d$, which results in an immediately updated state $\boldsymbol{z}_k \stackrel{\text{def}}{=} \boldsymbol{x}_k - \boldsymbol{y}_k$, and a net reward $r(t_k, \boldsymbol{x}_k, \boldsymbol{y}_k)$ accruing to the DM. When $\boldsymbol{y}_k \geq 0$ (< 0), the action can be thought of as extracting value from (injecting value into) the processes, in which case the corresponding net reward would typically be positive (respectively, negative). Not all actions are possible, and $A(\boldsymbol{x}_k)$ captures the feasible set when the initial state is $\boldsymbol{x}_k$.

Following the DM's action, the system subsequently evolves from time t_k to the next monitoring time t_{k+1}, where it takes a value of $\boldsymbol{x}_{k+1}$, chosen by nature from an uncertainty set. More precisely, for any $0 \leq k \leq r \leq n+1$, and given a fixed choice of monitoring times $\boldsymbol{t}^{\{r\}}$, observations $\boldsymbol{x}^{\{k\}}$ and post-action states $\boldsymbol{z}^{\{k\}}$ up to time t_k, the set of possible future values for $[\boldsymbol{x}_{k+1}, \ldots, \boldsymbol{x}_r, \boldsymbol{x}_{n+1}]$ is given by:

$$\tilde{\mathcal{U}}(\boldsymbol{t}^{\{r\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k\}}) \stackrel{\text{def}}{=} \Big\{ [\boldsymbol{x}_{k+1}, \ldots, \boldsymbol{x}_r] \in \mathbb{R}^{d \times (r-k+1)} : \tilde{f}(t_p, t_q, \boldsymbol{x}_p, \boldsymbol{z}_p, \boldsymbol{x}_q), \quad \forall\, p, q \in \{0, \ldots, r, n+1\},\, p < q \Big\}. \tag{17}$$

As before, we consider two versions of the DM's problem—*static* and *dynamic*—depending on whether the monitoring times are chosen at inception or throughout the problem horizon. The DM's objective is to determine the monitoring times $\boldsymbol{t}^{\{n+1\}}$ and the optimal actions $\boldsymbol{y}^{\{n+1\}}$ that maximize his cumulative reward up to time t_{n+1}.

Assumptions. We assume that rewards and action sets are monotonic in states.

Assumption 3. *The net reward $r(t, \boldsymbol{x}, \boldsymbol{y})$ is increasing in $\boldsymbol{x}$, and the action set $A(\boldsymbol{x})$ is increasing in $\boldsymbol{x}$ with respect to set inclusion, i.e., $\boldsymbol{x}_1 \leq \boldsymbol{x}_2 \Rightarrow A(\boldsymbol{x}_1) \subseteq A(\boldsymbol{x}_2)$.*

Several feasible sets satisfy our requirement; for instance, $A(\boldsymbol{x}) = \{\boldsymbol{y} : 0 \leq \boldsymbol{y} \leq \boldsymbol{x}\}$. Paralleling Assumption 2, we also require the uncertainty sets to be lattices, with suitable monotonicity and dynamic consistency properties.

Assumption 4. *For any $0 \leq k \leq r \leq n$, and given $\boldsymbol{t}^{\{r\}}, \boldsymbol{x}^{\{k\}}$ and $\boldsymbol{z}^{\{k\}}$,*

(i) *(Lattice)* $\tilde{\mathcal{U}}(\boldsymbol{t}^{\{r\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k\}})$ *is a lattice;*

(ii) *(Monotonicity)* $\tilde{\mathcal{U}}(\boldsymbol{t}^{\{r\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k\}})$ *is increasing in* $\boldsymbol{x}^{\{k\}}$ *and* $\boldsymbol{z}^{\{k\}}$*;*

(iii) *(Dynamic Consistency)* $\Pi_{:,i:j}\tilde{\mathcal{U}}(\boldsymbol{t}^{\{r\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k\}}) = \Pi_{:,i:j}\tilde{\mathcal{U}}(\boldsymbol{t}^{\{r'\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k\}})$, $\forall i \le j \le r \le r' \le n+1$.

These generalized sets allow future states to depend on historical state values both immediately before and immediately after the DM's actions. As before, we can prove that dynamic consistency is guaranteed when $\tilde{f}$ is monotonic in its second argument (details are omitted.)

Analysis. We first consider the dynamic problem. With $\tilde{J}_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k-1\}})$ denoting[7] the DM's value-to-go function at time t_k, the Bellman recursions become:

$$\begin{aligned}
\tilde{J}_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k-1\}}) &= \max_{\boldsymbol{y}_k \in A(\boldsymbol{x}_k)} \Bigg[r(t_k, \boldsymbol{x}_k, \boldsymbol{y}_k) + \\
&\qquad \max_{t_{k+1} \in [t_k, T]} \min_{\boldsymbol{x}_{k+1} \in \tilde{\mathcal{U}}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k\}})} \tilde{J}_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k+1\}}, \boldsymbol{z}^{\{k\}})) \Bigg], \\
\tilde{J}_{n+1}(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{n+1\}}, \boldsymbol{z}^{\{n\}}) &= \max_{\boldsymbol{y}_{n+1} \in A(\boldsymbol{x}_{n+1})} r(t_{n+1}, \boldsymbol{x}_{n+1}, \boldsymbol{y}_{n+1}).
\end{aligned}$$

Let $\tilde{J}_0 \stackrel{\text{def}}{=} \tilde{J}_0(t_0, \boldsymbol{x}_0)$.

In the static problem, the DM chooses $\boldsymbol{t}^{\{n+1\}}$ at time t_0. With $\tilde{V}_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k-1\}})$ denoting the value-to-go function at time t_k, the Bellman recursions become:

$$\begin{aligned}
\tilde{V}_k(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k-1\}}) &= \max_{\boldsymbol{y}_k \in A(\boldsymbol{x}_k)} \Bigg[r(t_k, \boldsymbol{x}_k, \boldsymbol{y}_k) + \min_{\boldsymbol{x}_{k+1} \in \tilde{\mathcal{U}}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k\}})} \tilde{V}_{k+1}\big(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{k+1\}}, \boldsymbol{z}^{\{k\}})\big) \Bigg], \\
\tilde{V}_{n+1}(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{n+1\}}, \boldsymbol{z}^{\{n\}}) &= \max_{\boldsymbol{y}_{n+1} \in A(\boldsymbol{x}_{n+1})} r(t_{n+1}, \boldsymbol{x}_{n+1}, \boldsymbol{y}_{n+1}),
\end{aligned}$$

and the optimal choice of monitoring times yields a value of $\tilde{V}_0 \stackrel{\text{def}}{=} \max_{\boldsymbol{t}^{\{n+1\}}} \tilde{V}_0(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{0\}})$.

In this context, we can confirm that an analogous result to Theorem 1 holds, and the dynamic problem yields the same worst-case optimal reward as the static problem.

THEOREM 5. *Under Assumption 3 and Assumption 4,* $\tilde{J}_0 = \tilde{V}_0$.

Proof. The Bellman recursion under dynamic monitoring can be written:

$$\begin{aligned}
\tilde{J}_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k-1\}}) &= \max_{\boldsymbol{y}_k \in A(\boldsymbol{x}_k)} \max_{t_{k+1} \in [t_k, T]} \Big[r(t_k, \boldsymbol{x}_k, \boldsymbol{y}_k) + \tilde{G}_k(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k\}}) \Big], \quad \text{where} \\
\tilde{G}_k(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k\}}) &\stackrel{\text{def}}{=} \min_{\boldsymbol{x}_{k+1} \in \tilde{\mathcal{U}}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k\}})} \tilde{J}_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k+1\}}, \boldsymbol{z}^{\{k\}}), \forall k \in \{1, \ldots, n\}.
\end{aligned}$$

[7] To simplify notation, we define $\boldsymbol{z}^{\{-1\}} \stackrel{\text{def}}{=} \emptyset$.

First, using induction, we prove that $\tilde{J}_k$ and $\tilde{G}_k$ are increasing in all arguments except time. By Assumption 3, this is true[8] for $\tilde{J}_{n+1}(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{n+1\}}, \boldsymbol{z}^{\{n\}})$. Assuming this is true at $k+1$, and using Assumption 4(i,ii), note that:

$$\begin{aligned}\operatorname{argmin}_{\boldsymbol{x}_{k+1}\in\tilde{\mathcal{U}}(\boldsymbol{t}^{\{k+1\}},\boldsymbol{x}^{\{k\}},\boldsymbol{z}^{\{k\}})}\tilde{J}_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k+1\}}, \boldsymbol{z}^{\{k\}}) &= \min_{\boldsymbol{x}_{k+1}\in\tilde{\mathcal{U}}(\boldsymbol{t}^{\{k+1\}},\boldsymbol{x}^{\{k\}},\boldsymbol{z}^{\{k\}})} \boldsymbol{x}_{k+1} \\ &\stackrel{\text{def}}{=} \underline{\boldsymbol{x}}_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k\}}),\end{aligned}$$

and $\underline{\boldsymbol{x}}_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k\}})$ is increasing in $\boldsymbol{x}^{\{k\}}$ and $\boldsymbol{z}^{\{k\}}$. Therefore,

$$\tilde{G}_k(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k\}}) = \tilde{J}_{k+1}\Big(\boldsymbol{t}^{\{k+1\}}, \big[\boldsymbol{x}^{\{k\}}, \underline{\boldsymbol{x}}_{k+1}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k\}})\big], \boldsymbol{z}^{\{k\}}\Big)$$

is increasing in $(\boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k\}})$. But then, note that the maximand in the problem:

$$\tilde{J}_k(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k-1\}}) = \max_{\boldsymbol{y}_k\in A(\boldsymbol{x}_k)} \max_{t_{k+1}\in[t_k,T]} \Big[r(t_k, \boldsymbol{x}_k, \boldsymbol{y}_k) + \tilde{G}_k\big(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}}, [\boldsymbol{z}^{\{k-1\}}, \boldsymbol{x}_k - \boldsymbol{y}_k]\big)\Big].$$

is increasing in $(\boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k-1\}})$, for any fixed value of $\boldsymbol{y}_k$ and t_{k+1}. And since the action set $A(\boldsymbol{x}_k)$ is increasing in $\boldsymbol{x}_k$ with respect to set inclusion by Assumption 3, this implies that $\tilde{J}_k$ is increasing in $(\boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k-1\}})$, which completes our induction.

Using these monotonicity properties, we then obtain:

$$\tilde{G}_{k-1}(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k-1\}}, \boldsymbol{z}^{\{k-1\}}) \tag{18}$$

$$\begin{aligned}&= \min_{\boldsymbol{x}_k\in\tilde{\mathcal{U}}(\boldsymbol{t}^{\{k\}},\boldsymbol{x}^{\{k-1\}},\boldsymbol{z}^{\{k-1\}})} \max_{\boldsymbol{y}_k\in A(\boldsymbol{x}_k)} \max_{t_{k+1}\in[t_k,T]} \Big[r(t_k, \boldsymbol{x}_k, \boldsymbol{y}_k) + \tilde{G}_k\Big(\boldsymbol{t}^{\{k+1\}}, [\boldsymbol{x}^{\{k-1\}}, \boldsymbol{x}_k], \boldsymbol{z}^{\{k\}}\Big)\Big] \\ &= \max_{t_{k+1}\in[t_k,T]} \max_{\boldsymbol{y}_k\in A(\boldsymbol{x}_k)} \Big[r(t_k, \underline{\boldsymbol{x}}_k, \boldsymbol{y}_k) + \tilde{G}_k\Big(\boldsymbol{t}^{\{k+1\}}, [\boldsymbol{x}^{\{k-1\}}, \underline{\boldsymbol{x}}_k], \boldsymbol{z}^{\{k\}}\Big)\Big] \\ &= \max_{t_{k+1}\in[t_k,T]} \min_{\boldsymbol{x}_k\in\tilde{\mathcal{U}}(\boldsymbol{t}^{\{k+1\}},\boldsymbol{x}^{\{k-1\}},\boldsymbol{z}^{\{k-1\}})} \max_{\boldsymbol{y}_k\in A(\boldsymbol{x}_k)} \Big[r(t_k, \boldsymbol{x}_k, \boldsymbol{y}_k) + \tilde{G}_k\Big(\boldsymbol{t}^{\{k+1\}}, [\boldsymbol{x}^{\{k-1\}}, \boldsymbol{x}_k], \boldsymbol{z}^{\{k\}}\Big)\Big].\end{aligned} \tag{19}$$

The second equality follows from the monotonicity of r and $\tilde{G}_k$ in $\boldsymbol{x}_k$; the last equality follows from the same monotonicity and the dynamic consistency Assumption 4(iii), which ensures that $\tilde{\mathcal{U}}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k-1\}}, \boldsymbol{z}^{\{k-1\}}) = \tilde{\mathcal{U}}(\boldsymbol{t}^{\{k\}}, \boldsymbol{x}^{\{k-1\}}, \boldsymbol{z}^{\{k-1\}})$, so that the nature's worst-case response $\underline{\boldsymbol{x}}_k$ is independent of the choice t_{k+1}. Therefore, we can interchange the order of $\max_{t_{k+1}}$ and $\min_{\boldsymbol{x}_k\in\tilde{\mathcal{U}}_k}$. Repeating the argument inductively, we obtain $\tilde{J}_0 = \tilde{V}_0$. ∎

[8] To see why this follows, consider $f(x) \stackrel{\text{def}}{=} \max_{y\in A(x)} g(x,y)$ where $g(\cdot, y)$ is increasing for any y, and let $y^*(x)$ denote a maximizer in the problem. Then, for $x_1 \le x_2$, we have $f(x_1) = g(x_1, y^*(x_1)) \le g(x_2, y^*(x_1)) \le \max_{y\in A(x_2)} g(x_2, y) = f(x_2)$, where the inequality in the second step relies on $y^*(x_1) \in A(x_2)$, which is guaranteed by Assumption 3.

This result again allows reconstructing the DM's optimal dynamic monitoring policy by (resolving) static versions of the monitoring problem. In fact, a further simplification is also possible here, as summarized in our next result.

Proposition 6. *Consider the static monitoring problem. The DM can make all the injection decisions at time t_0 and recover the same worst-case reward, i.e.,*

$$\tilde{V}_0 = \max_{\boldsymbol{t}^{\{n+1\}}} \max_{\boldsymbol{y}^{\{n+1\}} \in \mathbb{R}^{d\times(n+1)}} \min_{\substack{\boldsymbol{x}^{\{n+1\}}: \forall k \in \{1,\ldots,n+1\},\\ \boldsymbol{x}_k \in \tilde{\mathcal{U}}(\boldsymbol{t}^{\{n+1\}}, \boldsymbol{x}^{\{k-1\}}, \boldsymbol{z}^{\{k-1\}})}} \sum_{k=0}^{N} r(t_k, \boldsymbol{x}_k, \boldsymbol{y}_k).$$

Proof. Running through the same arguments as in the proof of Theorem 5, let $\boldsymbol{y}_k^*(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k\}}, \boldsymbol{z}^{\{k\}})$ denote an optimal policy for the DM in (19). It can be checked that the operators $\min_{\boldsymbol{x}_k \in \tilde{\mathcal{U}}(\boldsymbol{t}^{\{k+1\}}, \boldsymbol{x}^{\{k-1\}}, \boldsymbol{z}^{\{k-1\}})} \max_{\boldsymbol{y}_k \in A(\boldsymbol{x}_k)}$ in (19) can be interchanged under a choice $\boldsymbol{y}_k^*(\boldsymbol{t}^{\{k+1\}}, \underline{\boldsymbol{x}}^{\{k\}}, \boldsymbol{z}^{\{k-1\}})$, since this action remains feasible for any $\boldsymbol{x}_k$ by Assumption 3, and nature's worst-case response under knowledge of this action remains $\underline{\boldsymbol{x}}^{\{k\}}$. Repeating the argument by induction then yields the result. ■

In view of Proposition 6, for purposes of recovering the worst-case reward, the DM can restrict attention to static policies for *both* monitoring *and* extraction; this simplifies the problem, and allows reconstructing a dynamic policy by repeatedly finding static policies.

5. Robust Monitoring of Cardiac Allograft Vasculopathy

We leverage our approach to devise monitoring policies for patients suffering from Cardiac Allograft Vasculopathy (CAV). As discussed in the Introduction, CAV patients face disease progression that is highly uncertain and therefore could benefit from policies prescribed by our method. We use real data to calibrate our model and then evaluate the performance of our proposed monitoring policies vis-a-vis established guidelines in the medical community. Simulation suggests that our policies provide a QALY distribution that stochastically dominates that provided by existing guidelines, with a substantial increase in lower percentiles and a slight increase in median and higher percentiles. The results showcase the efficacy and robustness of our policies.

Background. Heart transplantation often represents the only viable treatment option for patients suffering from refractory or end-stage heart failure. Although its post-operative survival rates have increased over the last decades, with 5-year survival rates currently around 72.5% (Wilhelm 2015), heart transplantation could still result in serious long-term complications, such as CAV.

CAV is caused by a thickening and hardening of coronary arteries, which obstructs the blood circulation through the heart; this can cause various cardiac problems, from abnormal heart rhythms (arrhythmias) to heart attacks, heart failure, or even sudden cardiac death. CAV is the primary limiting factor for the long-term survival in heart transplantation, accounting for 17% of deaths by the third post-transplant year and 30% by the fifth year. When early stages are included, CAV affects up to 75% of heart-transplanted patients during the three years after transplantation (Ramzy et al. 2005).

The pathology of CAV is such that it presents two main challenges in managing the disease. First, CAV is a form of chronic rejection that lacks prominent symptoms, and it is particularly difficult to predict its progression. Consequently, to obtain information about CAV development, patients need to undergo periodic monitoring procedures, which involve invasive and often expensive examinations, such as coronary angiography and intravascular ultrasonography. Second, after detection, re-transplantation is the only viable treatment option, which itself introduces a tradeoff. On the one hand, re-transplantation in early stages is not advisable, because of the substantial evidence that its success is negatively correlated with the interval between transplants (Schnetzler et al. 1998, Johnson 2007, Saito et al. 2013, John et al. 1999, Vistarini et al. 2010). On the other hand, as the disease progresses, quality of life prognosis before and after re-transplantation worsens.

In order to manage CAV in an effective way, patients and their physicians need to make joint monitoring and re-transplantation timing decisions, relying on all available information and capturing all aforementioned tradeoffs. Unfortunately, the International Society for Heart and Lung Transplantation (ISHLT) provides crude recommendations for re-transplantation and monitoring, which are not personalized.

5.1. Data Description

There is limited data that can be used to guide CAV disease management decisions. To calibrate CAV development, we use a dataset from the Papworth Hospital that was compiled by monitoring 622 heart-transplanted patients on a yearly basis and recording their disease progression, starting from the transplant surgery.[9] The length of the monitoring horizon across all patients varies from a few days to 20 years, with an average of 3.85 years and a standard deviation of 3.34 years. At each monitoring time, the patient's CAV stage

[9] Data is publicly available in an R package called 'msm.'

is recorded; CAV progression is classified into three stages—stage 1 (Not Significant), stage 2 (Mild) and stage 3 (Severe)—depending on the observed degree of angiographic lesion and allograft dysfunction. The dataset comprises 2,846 monitoring observations.

To calibrate survival after re-transplantation, we obtained data from Tjang (2007), Copeland et al. (2011), Schnetzler et al. (1998), Novitzky et al. (1985); in particular, we obtained 37 observations of survivability outcomes after the second transplantation, for different intervals between the two transplants.

5.2. Robust Monitoring

Consider a CAV patient who had a heart transplantation at $t = 0$; the patient and his physician need to decide when to monitor disease progression and when to perform a re-transplantation throughout a horizon T. The patient's state at time t is given by the number of years that the patient spent in a CAV stage of i and above, denoted by $x^i(t)$, for $i \in \{2, 3\}$, and the patient's age, denoted by $age(t)$. Re-transplantation corresponds to "stopping," which yields a reward that we calculate as the total Quality-Adjusted Life Years (QALY), i.e.,

$$\text{reward at } t = (\text{QALY until } t) + (\text{QALY after re-transplant at } t).$$

If we let $Z^i(t)$ be the time spent in each CAV stage i, for $i \in \{1, 2, 3\}$, and $LY(t)$ be the expected life years after re-transplantation at t, the reward can be expressed as

$$\text{reward at } t = \sum_i QOL_i \cdot Z^i(t) + QOL_{re} \cdot LY(t),$$

where the factors QOL_i and QOL_{re} adjust for quality of life at stage i and after re-transplantation, respectively. Table 1 reports the quality of life factors we use, which we approximated based on the related literature (Long et al. 2014, Feingold et al. 2015, Grauhan et al. 2010, Montazeri 2009, Johnson 2007).

QOL_1	QOL_2	QOL_3	QOL_{re}
0.8583	0.7138	0.5774	0.6456

Table 1 Quality of Life Factors

The expected life years after re-transplantation could depend in principle on the patient's state. However, the substantial literature on the expected outcomes after re-transplantation provides consistent evidence that the interval between transplants is the only statistically significant predictor (Schnetzler et al. 1998, Saito et al. 2013, Vistarini et al. 2010, John et al. 1999). Therefore, we model expected life years after re-transplantation only as a function of t. Furthermore, many papers suggest that re-transplantations made within 1 year after the primary transplant result in worse survival rate compared to those performed after 1 year (Saito et al. 2013, Johnson et al. 2007, Goldraich et al. 2016). Based on this evidence, we consider the following piecewise linear model

$$LY(t) = \beta_0 + \beta_1 * t + \beta_2 * (t - b)^+ + \epsilon, \tag{20}$$

so that the survival could increase with t up to some point b, and then decrease. The model is fitted using regression based on our heart re-transplantation data (Figure 2):

$$\hat{\beta}_0 = 2.1635,\ \hat{\beta}_1 = 1.0356,\ \hat{\beta}_2 = -1.7727,\ \hat{b} = 5.060. \tag{21}$$

By combining all these results and by noting that the variables $\boldsymbol{Z}(t)$ relate to our state variables as $Z^1(t) = t - x^2(t)$, $Z^2(t) = x^2(t) - x^3(t)$, $Z^3(t) = x^3(t)$, we can formulate the reward function as

$$\begin{aligned} g(t, x^2, x^3, age) = {} & 0.8583\, t - 0.1445\, x^2(t) - 0.1364\, x^3(t) \\ & + 0.6456\,(2.1635 + 1.0356\, t - 1.7727\,(t - 5.CANWE060)^+), \end{aligned} \tag{22}$$

which is monotonic in x^2 and x^3 and therefore satisfies Assumption 1.

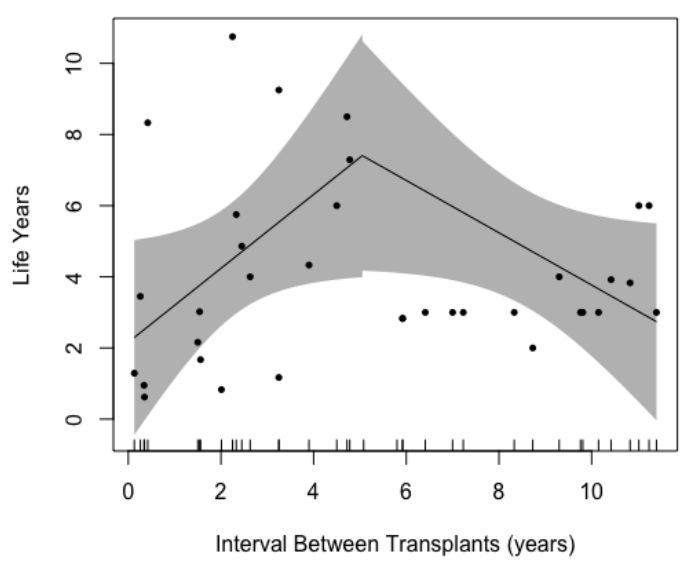


Figure 2 The fitted life year after re-transplantation as a function of the interval between transplants (n = 37).

Uncertainty Sets for CAV Progression

To model CAV progression, we start by calibrating a Cox model, which is one of the most popular models for survival analysis in healthcare. Because of limited data, such a model is likely to provide noisy estimates of the true progression probabilities. Moreover, the underlying assumption—typical in a Cox model— that a unit increase in a covariate has a constant multiplicative effect might not be reasonable for CAV progression, further making the goodness of fit questionable. Therefore, we only use the estimates of the Cox model to construct uncertainty sets.

In particular, consistent with a Cox model, we start by assuming that the time until a patient transitions from one CAV stage to another are exponentially distributed, with a mean that depends linearly on t and $age(t)$. More precisely, when a patient with $age(t)$ is observed to be in stage i at post-transplant time t, the time until he progresses to stage j, denoted $L_{ij}(t, age)$, is exponentially distributed with rate $\lambda_{ij}(t, age)$ such that

$$1/\lambda_{ij}(t, age) = \beta_0^{ij} + \beta_1^{ij} * t + \beta_2^{ij} * age(t), \quad \forall (i,j) \in \{(1,2),(1,3),(2,3)\}. \tag{23}$$

Table 2 reports the coefficients we obtained after fitting this model to our CAV monitoring data using maximum likelihood estimation.

(i,j)	β_0^{ij}	β_1^{ij}	β_2^{ij}
(1,2)	18.0826 (0.8900)	-0.6511 (0.0456)	-0.1820 (0.0177)
(1,3)	38.8900 (2.8844)	-1.1412 (0.1754)	-0.3991 (0.0573)
(2,3)	6.5479 (0.9671)	-0.1576 (0.0745)	-0.0492 (0.0178)

Table 2 Linear Coefficients for $1/\lambda_{ij}$ (standard errors).

Using the model for transition times, we then derive confidence intervals that we use to construct uncertainty sets for the amount of time spent in each disease state. Specifically, suppose a patient is monitored post-transplant at time t, resulting in an observation $(x^2(t), x^3(t), age(t))$. When the patient is again monitored after an additional Δt time units, the possible values for $x^2(t+\Delta t)$ can be described as follows.

1. If $x^2(t) > 0$, the patient is already in stage 2 or above at time t. Therefore, $x^2(t+\Delta t) = x^2(t) + \Delta t$ without any uncertainty.

2. If $x^2(t)=0$, the patient must have been in stage 1 at time t. Then, the distribution of $x^2(t+\Delta t)$ can be expressed as:

$$\mathbb{P}(x^2(t+\Delta t) \leq s | x^2(t)=0) = \mathbb{P}(L_{12}(t, age) \geq \Delta t - s) = e^{-\lambda_{12}(t,age)(\Delta t - s)}, \ \forall s \in [0, \Delta t].$$

Using the $(1-\rho)$-percentile of L_{12}, i.e., $s_\rho^{12} \stackrel{\text{def}}{=} \inf\{s : \mathbb{P}(L_{12}(t, age(t)) \leq s) \geq 1-\rho\}$, we can construct a $\rho\%$-confidence-level upper bound for $x^2(t+\Delta t)$ as:

$$(\Delta t - s_\rho^{12})^+ = \Big(\Delta t + \frac{\log(\rho)}{\lambda_{12}(t, age)}\Big)^+.$$

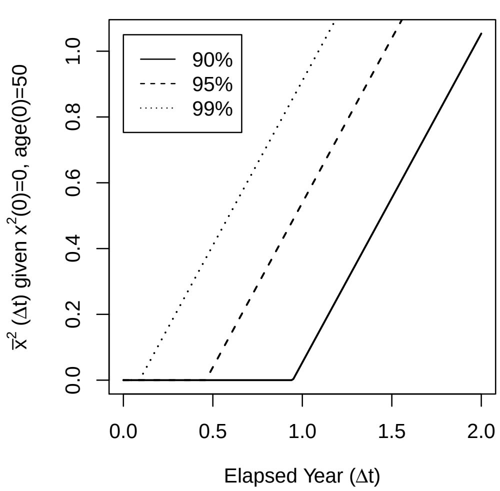
Similarly, we obtain $\rho\%$-confidence-level upper bounds for $x^2(t+\Delta t)$ and $x^3(t+\Delta t)$ given $x^2(t), x^3(t)$ denoted by $\bar{x}^2(t+\Delta t)|x^2(t)$ and $\bar{x}^3(t+\Delta t)|x^2(t), x^3(t)$ respectively, as follows:

$$\bar{x}^2(t+\Delta t)|x^2(t) = \begin{cases} \big(\Delta t + \frac{\log(\rho)}{\lambda_{12}(t,age)}\big)^+, & \text{if } x^2(t)=0 \\ x^2(t)+\Delta t, & \text{otherwise.} \end{cases} \tag{24a}$$

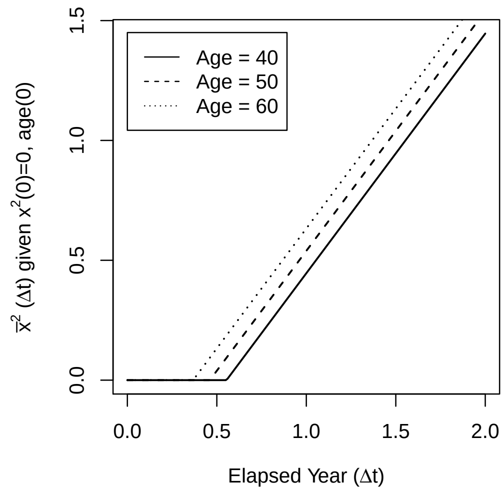
$$\bar{x}^3(t+\Delta t)|x^2(t), x^3(t) = \begin{cases} \big(\Delta t + \frac{\log(\rho)}{\lambda_{13}(t,age)}\big)^+, & \text{if } x^2(t)=0, x^3(t)=0 \\ \big(\Delta t + \frac{\log(\rho)}{\lambda_{23}(t,age)}\big)^+, & \text{if } x^2(t)>0, x^3(t)=0 \\ x^3(t)+\Delta t, & \text{otherwise.} \end{cases} \tag{24b}$$

This implies that, at $\rho\%$ confidence level, an observation $(x^2(t),\ x^3(t))$ at time t imposes restrictions on the future state values $(x^2(t+\Delta t),\ x^3(t+\Delta t))$ specified by $x^2(t+\Delta t) \leq \bar{x}^2(t+\Delta t)$ and $x^3(t+\Delta t) \leq \bar{x}^3(t+\Delta t)$. To complete our uncertainty sets, we also set the lower bounds on $\big(x^2(t+\Delta t),\ x^3(t+\Delta t)\big)$ to zero, without loss of generality. Since both the upper and the lower bounds are increasing in x^2, x^3, and Δt, and are concave and convex, respectively, in x_2 and x_3, the resulting uncertainty sets satisfy Assumption 2.[10]

[10] Several technical remarks may be in order. First, since the reward g is decreasing in x_2 and x_3 by (22), the lower bounds are irrelevant for determining the worst-case rewards. Our choice of zero thus only serves to ensure that the resulting uncertainty sets satisfy Assumption 2. Second, note that according to (24a) and (24b), the upper bounds $\bar{x}^2(t+\Delta t)|x^2(t)$ and $\bar{x}^3(t+\Delta t)|x^2(t), x^3(t)$ are discontinuous in $(x^2(t), x^3(t))$ as we approach the boundary $\{x^2(t)=0\} \cup \{x^3(t)=0\}$. However, since $\log \rho < 0$, the functions are upper-semicontinuous at the boundary, and are thus concave in $(x^2(t), x^3(t))$. To rigorously prove that our results in Theorem 1 would hold, we could introduce an intermediate region in (24a) for $0 < x^2(t) < \epsilon$ (for $\epsilon > 0$), where $\bar{x}^2(t+\Delta t)|x^2(t)$ is defined by linearly interpolating between the value at $x^2(t)=0$ and the value at $x^2(t)+\Delta t$ (and similarly for $\bar{x}^3(t+\Delta t)$). For any $\epsilon > 0$, one can check that Assumption 2 is satisfied, and Theorem 1 thus holds. Since the worst-case optimal values under both static and dynamic monitoring would be continuous functions of ϵ, and have equal values for any $\epsilon > 0$, their limiting values as $\epsilon \to 0$ must also correspond, so that Theorem 1 also holds for the uncertainty set given by the upper bounds in (24a) and (24b). Lastly, note that we do not need to check concavity (convexity) of the upper (lower) bounds in $age(t)$, since the latter can be taken without loss outside the state of the system (its value is simply t plus a constant denoting the patient's age at time 0).



(a) The worst-case values for $x^2(\Delta t)$ depend on the confidence level ρ (initial age = 50).



(b) The worst-case values for $x^2(\Delta t)$ depend on the patient's initial age ($\rho = 95\%$).

Figure 3 The worst-case values for $x^2(\Delta t)$ given the initial observation.

Naturally, the uncertainty sets we construct here depend on the confidence level ρ and the patient's initial age. Higher confidence level produces more conservative upper bounds, resulting in higher worst-case values for $x^2(t+\Delta t)$ and $x^3(t+\Delta t)$. Figure 3(a) illustrates how the upper bounds for $x^2(\Delta t)$ imposed by the initial observation $(x^2(0), x^3(0)) = (0, 0)$ depend on the confidence level ρ.[11] Second, the negative values of β_2^{ij} (Table 2) imply that older patients are more likely to have shorter transition times. Thus, the worst-case values for $x^2(t+\Delta t)$ and $x^3(t+\Delta t)$ are increasing in the patient's initial age (Figure 3(b)).

5.3. Monitoring Policy

To solve the dynamic monitoring problem using the algorithm described in Corollary 1, it suffices to solve a series of static monitoring problems. Based on our analysis, the static monitoring problem with n monitoring times $t_1, \ldots, t_n$ chosen at time 0 can be written as

$$\max_{\boldsymbol{t}^{\{n+1\}}} \max\Big(g\big(t_n, \bar{x}^2(t_n), \bar{x}^3(t_n), age(t_n)\big),\ g\big(t_{n+1}, \bar{x}^2(t_{n+1}), \bar{x}^3(t_{n+1}), age(t_{n+1})\big)\Big), \tag{25}$$

where $\bar{x}^2(t_k)$ and $\bar{x}^3(t_k)$ satisfy

$$\bar{x}^2(t_{k+1}) = \begin{cases} \big(t_{k+1} - t_k + \log(\rho)(\beta_0^{12} + \beta_1^{12}\, t_k + \beta_2^{12}\, age(t_k))\big)^+, & \text{if } \bar{x}^2(t_k) = 0, \\ \bar{x}^2(t_k) + t_{k+1} - t_k, & \text{otherwise,} \end{cases} \tag{26a}$$

[11] A CAV patient's stage is 1 at time 0, that is $x^2(0) = 0, x^3(0) = 0$.

$$\bar{x}^3(t_{k+1}) = \begin{cases} \left(t_{k+1} - t_k + \log(\rho)(\beta_0^{13} + \beta_1^{13}\, t_k + \beta_2^{13}\, age(t_k))\right)^+, & \text{if } \bar{x}^2(t_k) = 0, \bar{x}^3(t_k) = 0, \\ \left(t_{k+1} - t_k + \log(\rho)(\beta_0^{23} + \beta_1^{23}\, t_k + \beta_2^{23}\, age(t_k))\right)^+, & \text{if } \bar{x}^2(t_k) > 0, \bar{x}^3(t_k) = 0, \\ \bar{x}^3(t_k) + t_{k+1} - t_k, & \text{otherwise,} \end{cases} \tag{26b}$$

for all $k = 0, 1, \ldots, n$. Notice that the worst-case value for $x^2(t_{k+1})$ at time t_{k+1} is determined solely by the upper bound imposed by the latest observation at t_k. This relies on the fact that β_1^{ij} and β_2^{ij} are negative, so that the associated upper bounds decrease with t.

Note that the reward function and the bound functions defining the uncertainty sets are all piece-wise linear. Therefore, we can formulate Problem (25) as a mixed-integer optimization problem. We start by defining binary variables $\boldsymbol{y}$, $\boldsymbol{z}$ as follows:

$$y_k \stackrel{\text{def}}{=} \begin{cases} 1, & \text{if } \bar{x}^2(t_k) > 0, \\ 0, & \text{otherwise} \end{cases}, \quad z_k \stackrel{\text{def}}{=} \begin{cases} 1, & \text{if } \bar{x}^3(t_k) > 0, \\ 0, & \text{otherwise} \end{cases}, \quad \forall\, k \in \{0, \ldots, N\}. \tag{27}$$

Similarly, consider the binary variable u_n such that $u_n = 1$ if and only if $t_n > b = 5.060$. Then, for sufficiently large $M > 0$ and sufficiently small $\epsilon > 0$, maximizing $g(t_n, \bar{x}^2(t_n), \bar{x}^3(t_n), age(t_n))$ in Problem (25) can be reformulated as follows:

$$\max_{\substack{\bar{\boldsymbol{x}}^2, \bar{\boldsymbol{x}}^3, \boldsymbol{t}, s_n, \\ u_n, \boldsymbol{y}, \boldsymbol{z}}} \; 0.8583\, t_n - 0.1445\, \bar{x}_n^2 - 0.1364\, \bar{x}_n^3 + 0.6456\,(2.1635 + 1.0356\, t_n - 1.7727\, s_n)$$

$$\text{subject to } \bar{x}_0^2 = 0, \quad \bar{x}_0^3 = 0, \quad \bar{x}_N^2 \le T, \quad \bar{x}_N^3 \le T \tag{28a}$$

$$t_k \le t_{k+1}, \quad \bar{x}_k^2 \le \bar{x}_{k+1}^2, \quad \bar{x}_k^3 \le \bar{x}_{k+1}^3, \quad k \in \{0, 1, \ldots, n+1\} \tag{28b}$$

$$\bar{x}_k^2 \le M y_k, \quad \bar{x}_k^2 \ge \epsilon y_k, \quad \bar{x}_k^3 \le M z_k, \quad \bar{x}_k^3 \ge \epsilon z_k, \quad k \in \{0, 1, \ldots, n+1\} \tag{28c}$$

$$t_n \ge b\, u_n, \quad t_n \le b + M u_n \tag{28d}$$

$$s_n \ge -M u_n, \quad s_n \ge t_n - b - M(1 - u_n) \tag{28e}$$

$$s_n \le M u_n, \quad s_n \le t_n - b + M(1 - u_n) \tag{28f}$$

$$\bar{x}_{k+1}^2 \ge t_{k+1} - t_k + \log(\rho)(\beta_0^{12} + \beta_1^{12}\, t_k + \beta_2^{12}\, age(t_k)) - M y_k, \tag{28g}$$

$$\bar{x}_{k+1}^2 \ge \bar{x}_k^2 + t_{k+1} - t_k - M + M y_k, \tag{28h}$$

$$\bar{x}_{k+1}^3 \ge t_{k+1} - t_k + \log(\rho)(\beta_0^{13} + \beta_1^{13}\, t_k + \beta_2^{13}\, age(t_k)) - M y_k - M z_k, \tag{28i}$$

$$\bar{x}_{k+1}^3 \ge t_{k+1} - t_k + \log(\rho)(\beta_0^{23} + \beta_1^{23}\, t_k + \beta_2^{23}\, age(t_k)) - M + M y_k - M z_k, \tag{28j}$$

$$\bar{x}_{k+1}^3 \ge \bar{x}_k^3 + t_{k+1} - t_k - 2M + M y_k + M z_k, \quad k \in \{0, 1, \ldots, n\}. \tag{28k}$$

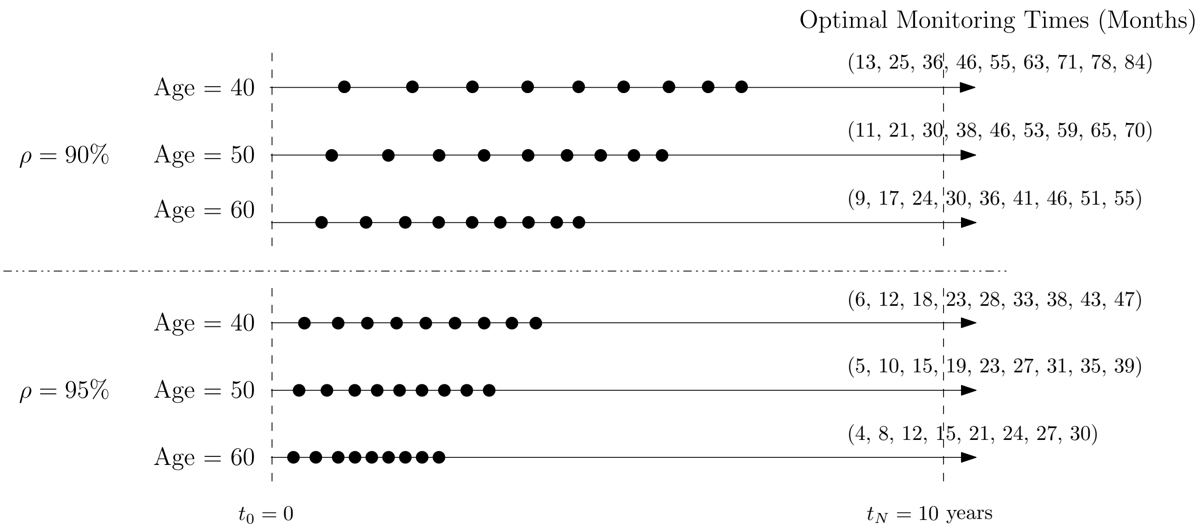


Figure 4 **Monitoring times (in months) for static monitoring prescribed by our framework.**

The inequalities in (28c) define the binary variables $\boldsymbol{y}, \boldsymbol{z}$, the inequalities in (28d) define u_n, and the inequalities in (28e)-(28f) define s_n as $(t_n - b)^+$. Then, the inequalities (28g)-(28h) and (28i)-(28k) correspond to (26a) and (26b), respectively.

The problem of maximizing the term $g(t_{n+1}, \bar{x}^2(t_{n+1}), \bar{x}^3(t_{n+1}), age(t_{n+1}))$ in (25) can similarly be formulated as an IP.

5.4. Results

We consider CAV management for a patient following heart transplantation. According to ISHLT, it is recommended that the patient undergoes annual or bi-annual examinations of coronary arteries through coronary angiography and intravascular ultrasonography to monitor progression in the first 3-5 years after transplant (Labarrere et al. 2017, Costanzo et al. 2010), and perform re-transplantation once CAV reaches stage 3 (Johnson 2007). We compare these recommended policies with the ones prescribed by our framework.

Static Monitoring. Given that ISHLT essentially prescribes a static monitoring policy, we begin by determining the optimal static monitoring policy in our framework. We consider a time horizon of 10 years; since ISHLT prescribes yearly examinations, we thus derive a static monitoring policy that allows for $n = 9$ intermediate monitoring times. Figure 4 illustrates our policy, depending on the patient's age and conservativeness (as measured by the desired confidence level ρ). Notably, our policy departs from equi-sized monitoring times and rather suggests a frequency that increases as time goes by, at a rate that also increases with a patient's age and conservativeness.

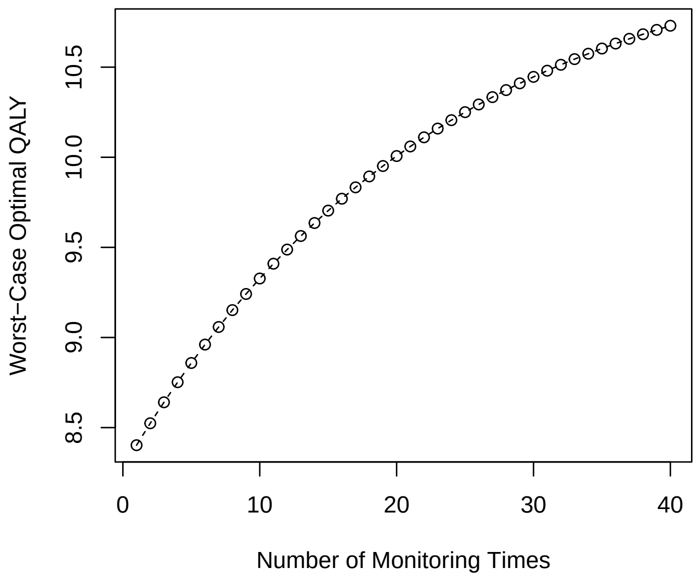


Figure 5 **Worst-case optimal QALYs as a function of the number of monitoring times ($T = 10$ years, age = 50, $\rho = 95$%).**

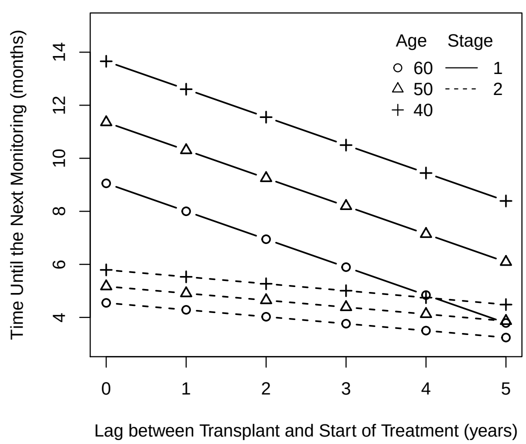
Our framework can also be used to explore the value that added monitoring opportunities would bring, therefore enabling patients to navigate this important tradeoff. As an illustrative example, Figure 5 depicts the worst-case QALYs for a 50 year old patient as a function of monitoring times. This type of analysis could help patients choose the appropriate number of monitoring times, by comparing the marginal QALY improvements of additional monitoring with the associated costs.

Dynamic Monitoring. Unlike ISHLT guidelines, our method can also be adjusted to account for information gained during monitoring when recommending the next examination. Table 3 reports our framework's recommended next monitoring time for a patient who just had a heart transplant, depending on age, CAV stage and conservativeness (we again assume $n = 9$ for a fair comparison). For example, when a 50-year-old patient is diagnosed with CAV stage 1, the first monitoring time would be 6 months later, but when diagnosed with stage 2, he would be monitored 3 months later (for $\rho = 95\%$).

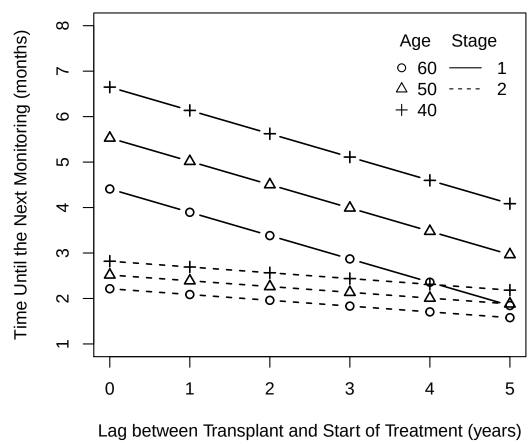
If treatment began with some lag following the transplant, Figure 6 illustrates when to schedule the next monitoring time, depending on the lag, the age and the CAV stage at the time of treatment. For example, a 60-year old patient who began treatment one year after transplant at CAV stage 1, should be monitored 8 months later (for $\rho = 90\%$). The results suggest that the longer the delay, the sooner the patient should be monitored next.

$\rho = 90\%$			$\rho = 95\%$		
Age	Stage 1	Stage 2	Age	Stage 1	Stage 2
40	14	6	40	7	3
50	11	5	50	6	3
60	9	5	60	4	2

Table 3 The worst-case optimal first monitoring time (in months).



(a) When $\rho = 90\%$



(b) When $\rho = 95\%$

Figure 6 The time until the next monitoring (in months) as a function of the lag between transplant and the start of treatment (in years).

Simulation. To compare the performance of our policies with that of the ISHLT guidelines, we simulate CAV progression using the transition probabilities we estimated in the model calibration. In particular, transition times are randomly generated according to an exponential distribution ($T = 10$ years, initial age = 50, number of iterations = 5×10^3). We simulate both our static and dynamic monitoring policies.

The results are summarized in Table 4 and Figure 7. Both the static and dynamic robust policies outperform the benchmark policy regardless of the confidence level, by yielding a QALY distribution that stochastically dominates that of the ISHLT policy. The average QALY under the robust policies are slightly better than under the benchmark policy, and the improvements become more noticeable for lower percentiles. To assess the robustness of our policies to the distributional assumptions of transition times, we also considered the cases where the true distributions (used in simulation) were different from the exponential distribution (used by the model). The results are included in the appendix, and are consistent with the ones reported here.

Policy	Min	25% Quantile	Median	75% Quantile	Max	Mean
Benchmark	4.0449	8.9784	9.9097	10.6657	11.0120	9.5069
Static RO ($\rho = 95\%$)	8.2711	9.2823	10.2498	10.8847	11.0120	10.0727
Dynamic RO ($\rho = 95\%$)	8.5748	9.5362	10.3140	10.8910	11.0120	10.1657
Static RO ($\rho = 90\%$)	8.3389	9.3209	9.9657	10.6961	11.0120	9.9732
Dynamic RO ($\rho = 90\%$)	8.4966	9.3770	10.0130	10.7070	11.0120	10.0132

Table 4 QALY under different monitoring policies
($T = 10$ years, initial age = 50, number of monitoring = 9, number of iterations = 5×10^3).

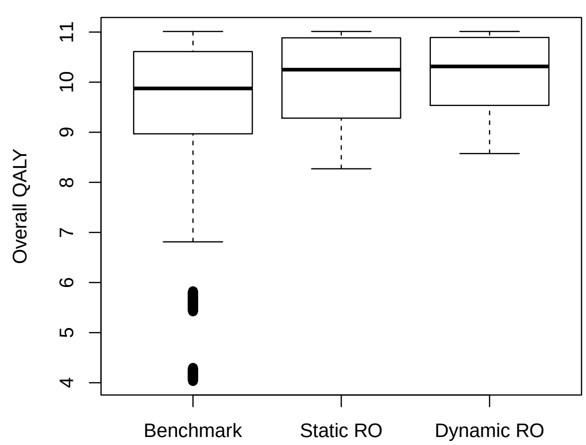


Figure 7 Boxplot for simulated QALYs ($\rho = 95$% under the benchmark, and the static and dynamic RO policies, Number of Simulations $= 5 \times 10^3$, Number of Monitoring Times $= 10$, $T = 10$ years).

Appendix A: Unknown Transition Times Distribution

Figure 8 illustrates the performance of our robust optimal policies when the distribution of the transition times are different from the one assumed by the model. We consider four distributions—Weibull, log-normal, gamma and log-logistic—that are widely used in parametric models for survival analysis. The distributions are fitted with the CAV monitoring data (§5.1) and used to simulate CAV progression. The box plots in Figure 8 imply that the performance of our robust optimal policies is not sensitive to this distributional misspecification.

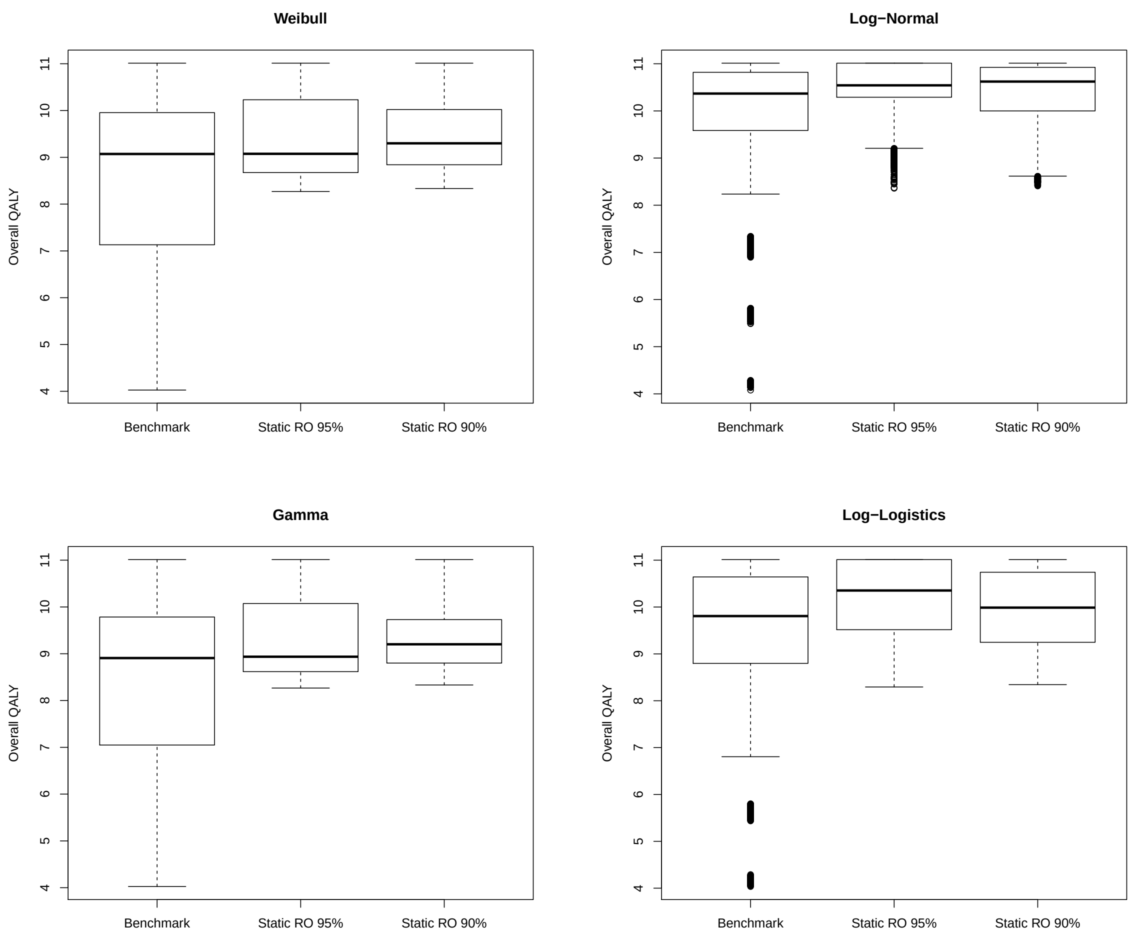


Figure 8 Boxplot for Non-Worst-Case Simulations under Different Distributional Assumptions (Number of Simulations $= 5 \times 10^3$, Number of Monitoring Times $= 10$, $T = 10$ years).

References

Ayer, Turgay, Oguzhan Alagoz, Natasha K Stout. 2012. OR forum – a POMDP approach to personalize mammography screening decisions. *Operations Research* **60**(5) 1019–1034.

Bandi, Chaithanya, Dimitris Bertsimas. 2012. Tractable stochastic analysis in high dimensions via robust optimization. *Mathematical Programming* **134** 23–70.

Bayraktar, E., I. Karatzas, S. Yao. 2010. Optimal stopping for dynamic convex risk measures. *Illinois Journal of Mathematics* **54**(3) 1025–1067.

Bayraktar, Erhan, Song Yao. 2014. On the robust optimal stopping problem. *SIAM Journal on Control and Optimization* **52**(5) 3135–3175.

Ben-Tal, Aharon, Laurent El Ghaoui, Arkadi Nemirovski. 2009. *Robust optimization*. Princeton University Press.

Ben-Tal, Aharon, A. Goryashko, E. Guslitzer, Arkadi Nemirovski. 2004. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming* **99**(2) 351–376.

Ben-Tal, Aharon, Arkadi Nemirovski. 2002. Robust optimization – methodology and applications. *Mathematical Programming* **92**(3) 453–480.

Bertsimas, D., V. Goyal, B. Lu. 2014. A tight characterization of the performance of static solutions in two-stage adjustable robust linear optimization. *Mathematical Programming* **150**(2) 281–319.

Bertsimas, D., P. Vayanos. 2017. Data-driven learning in dynamic pricing using adaptive optimization. Working paper.

Bertsimas, Dimitris, David B Brown, Constantine Caramanis. 2011. Theory and applications of robust optimization. *SIAM review* **53**(3) 464–501.

CH. 2014. *Accounts Receivable and Inventory Financing*. Comptroller of the currency, administrator of national banks.

Copeland, Hannah, Romana Coelho-Anderson, Nicole Mineburg, Michael McCarthy, Jack G Copeland. 2011. Elective cardiac retransplantation: a viable option that can be repeated. *The Journal of thoracic and cardiovascular surgery* **141**(3) 822–827.

Costanzo, Maria Rosa, Anne Dipchand, Randall Starling, Allen Anderson, Michael Chan, Shashank Desai, Savitri Fedson, Patrick Fisher, Gonzalo Gonzales-Stawinski, Luigi Martinelli, et al. 2010. The International Society of Heart and Lung Transplantation Guidelines for the care of heart transplant recipients.

Cutler, David M, James M Poterba, Lawrence H Summers. 1991. Speculative dynamics. *The Review of Economic Studies* **58**(3) 529–546.

Delage, E., D. A. Iancu. 2015. Robust multistage decision making. *Tutorials in Operations Research* 20–46.

Denton, Brian T, Murat Kurt, Nilay D Shah, Sandra C Bryant, Steven A Smith. 2009. Optimizing the start time of statin therapy for patients with diabetes. *Medical Decision Making* **29**(3) 351–367.

Dullerud, G. E., F. Paganini. 2005. *A Course in Robust Control Theory*. Springer.

Feingold, B, SA Webber, CL Bryce, SY Park, HE Tomko, SC West, SA Hart, WT Mahle, KJ Smith. 2015. Cost-effectiveness of pediatric heart transplantation across a positive crossmatch for high waitlist urgency candidates. *American Journal of Transplantation* **15**(11) 2978–2985.

Foley, C.F., A. Raman, N.C. Craig. 2013. Inventory-based lending industry note.

Fujishige, Satoru. 2005. *Submodular Functions and Optimization*, *Annals of Discrete Mathematics*, vol. 58. 2nd ed. Elsevier.

Goldraich, LA, Josef Stehlik, AY Kucheryavaya, LB Edwards, HJ Ross. 2016. Retransplant and medical therapy for cardiac allograft vasculopathy: International society for Heart and Lung Transplantation Registry analysis. *American Journal of Transplantation* **16**(1) 301–309.

Gorton, Gary, Andrew Winton. 2003. Financial intermediation. *Handbook of the Economics of Finance* **1** 431–552.

Grauhan, O, H Huang Chang, W Albert, N Hiemann, R Meyer, H Lehmkuhl, M Dandel, C Knosalla, M Hübler, M Pasic, et al. 2010. Impact of donor-transmitted coronary atherosclerosis on quality of life (QOL) and quality-adjusted life years (QALY) after heart transplantation. *Applied Cardiopulmonary Pathophysiology* **14** 58–65.

Hanany, Eran, Peter Klibanoff. 2009. Updating ambiguity averse preferences. *The B.E. Journal of Theoretical Economics* **9**(1).

Iancu, Dan A, Mayank Sharma, Maxim Sviridenko. 2013. Supermodularity and affine policies in dynamic robust optimization. *Operations Research* **61**(4) 941–956.

Ibanez, Alfredo, Fernando Zapatero. 2004. Monte Carlo valuation of American options through computation of the optimal exercise frontier. *Journal of Financial and Quantitative Analysis* **39**(2) 253–275.

Jaillet, P., S.D. Jena, T.S. Ng, M. Sim. 2016. Satisficing awakens: Models to mitigate uncertainty. Working paper.

John, Ranjit, Jonathan M Chen, Alan Weinberg, Mehmet C Oz, Donna Mancini, Silviu Itescu, Mark E Galantowicz, Craig R Smith, Eric A Rose, Niloo M Edwards. 1999. Long-term survival after cardiac retransplantation: a twenty-year single-center experience. *The Journal of thoracic and cardiovascular surgery* **117**(3) 543–555.

Johnson, Maryl R. 2007. When is retransplantation a viable option? *Heart failure clinics* **3**(1) 97–105.

Johnson, MR, Keith D Aaronson, CE Canter, JK Kirklin, DM Mancini, MR Mehra, B Radovancevic, DO Taylor, SA Webber. 2007. Heart retransplantation. *American Journal of Transplantation* **7**(9) 2075–2081.

Karatzas, Ioannis, Steven Shreve. 2012. *Brownian motion and stochastic calculus*, vol. 113. Springer Science & Business Media.

Karatzas, Ioannis, Steven E Shreve, I Karatzas, Steven E Shreve. 1998. *Methods of mathematical finance*, vol. 39. Springer.

Kolodko, Anastasia, John Schoenmakers. 2006. Iterative construction of the optimal Bermudan stopping time. *Finance and Stochastics* **10**(1) 27–49.

Labarrere, CA, BR Jaeger, GS Kassab. 2017. Cardiac allograft vasculopathy: Microvascular arteriolar capillaries (capioles) and survival. *Frontiers in bioscience (Elite edition)* **9** 110–128.

Lavieri, Mariel, Martin Puterman, Scott Tyldesley, William Morris. 2012. When to treat prostate cancer patients based on their PSA dynamics. *IIE Transactions on Healthcare Systems Engineering* **2**(1) 62–77.

Leland, Hayne E. 1994. Corporate debt value, bond covenants, and optimal capital structure. *The journal of finance* **49**(4) 1213–1252.

Leland, Hayne E, Klaus Bjerre Toft. 1996. Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads. *The Journal of Finance* **51**(3) 987–1019.

Long, Elisa F, Gary W Swain, Abeel A Mangi. 2014. Comparative survival and cost effectiveness of advanced therapies for end-stage heart failure. *Circulation: Heart Failure* .

Machina, Mark. 1989. Dynamic consistency and non-expected utility models of choice under uncertainty. *Journal of Economic Literature* **27** 1622–1668.

Maillart, Lisa M, Julie Simmons Ivy, Scott Ransom, Kathleen Diehl. 2008. Assessing dynamic breast cancer screening policies. *Operations Research* **56**(6) 1411–1427.

Marandi, A., Dick den Hertog. 2017. When are static and adjustable robust optimization problems with constraint-wise uncertainty equivalent? *Mathematical Programming, Series A* .

Montazeri, Ali. 2009. Quality of life data as prognostic indicators of survival in cancer patients: an overview of the literature from 1982 to 2008. *Health and quality of life outcomes* **7**(1) 102.

Nohadani, O., A. Roy. 2017. Robust optimization with time-dependent uncertainty in radiation therapy. *IISE Transactions on Healthcare Systems Engineering* **7**(2) 81–92.

Nohadani, O., K. Sharma. 2016. Optimization under decision-dependent uncertainty. Working paper.

Novitzky, D, DKC Cooper, RP Lanza, CN Barnard. 1985. Further cardiac transplant procedures in patients with heterotopic heart transplants. *The Annals of thoracic surgery* **39**(2) 149–154.

Poss, M. 2013. Robust combinatorial optimization with variable budgeted uncertainty. *4OR: A Quarterly Journal of Operations Research* **11**(1) 75–92.

Queyranne, M., F. Tardella. 2006. Bimonotone linear inequalities and sublattices of r^n. *Linear Algebra and Its Applications* **413** 100–120.

Rajan, Raghuram, Andrew Winton. 1995. Covenants and collateral as incentives to monitor. *The Journal of Finance* **50**(4) 1113–1146.

Ramzy, Danny, Vivek Rao, Julie Brahm, Santiago Miriuka, Diego Delgado, Heather J Ross. 2005. Cardiac allograft vasculopathy: a review. *Canadian journal of surgery* **48**(4) 319.

Riedel, Frank. 2009. Optimal stopping with multiple priors. *Econometrica* **77**(3) 857–908.

Saito, Aya, Richard J Novick, Bob Kiaii, F Neil McKenzie, Mackenzie Quantz, Peter Pflugfelder, Grant Fisher, Michael WA Chu. 2013. Early and late outcomes after cardiac retransplantation. *Canadian Journal of Surgery* **56**(1) 21.

Schnetzler, Bruno, Alain Pavie, Richard Dorent, Anne-Claude Camproux, Philippe Leger, Annick Delcourt, Iradj Gandjbakhch. 1998. Heart retransplantation: a 23-year single-center clinical experience. *The Annals of thoracic surgery* **65**(4) 978–983.

Schrijver, Alexander. 2003. *Combinatorial Optimization: Polyhedra and Efficiency*. 1st ed. Springer.

Shechter, Steven M, Matthew D Bailey, Andrew J Schaefer, Mark S Roberts. 2008. The optimal time to initiate HIV therapy under ordered health states. *Operations Research* **56**(1) 20–33.

Snell, James Laurie. 1952. Applications of martingale system theorems. *Transactions of the American Mathematical Society* **73**(2) 293–312.

Taylor, Howard M. 1968. Optimal stopping in a Markov process. *The Annals of Mathematical Statistics* 1333–1344.

Tjang, Yanto Sandy. 2007. Cardiac retransplantation. Ph.D. Thesis.

Topkis, Donald M. 1998. *Supermodularity and Complementarity*. Princeton University Press.

Vistarini, Nicola, Carlo Pellegrini, Marco Aiello, Alessia Alloni, Cristian Monterosso, Barbara Cattadori, Carmine Tinelli, Andrea M D'Armini, Mario Vigano. 2010. Should we perform heart retransplantation in early graft failure? *Transplant International* **23**(1) 47–53.

Wilhelm, Markus J. 2015. Long-term outcome following heart transplantation: current perspective. *Journal of thoracic disease* **7**(3) 549.

Wilmott, P., J. Dewynne, S. Howison. 1994. *Option Pricing: Mathematical Models and Computation*. Oxford Financial Press.

Zhang, X., M. Kamgarpour, A. Georghiou, P. Goulart, J. Lygeros. 2016. Robust optimal control with adjustable uncertainty sets. *Automatica* **75** 249–259.

Zhou, K., J. C. Doyle. 1998. *Essentials of Robust Control*. Prentice Hall.