2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Optimal Stopping Problems in Low-Dimensional Feature Spaces: Lossless Conditions and Approximations

M.J.T.C. van Zutphen, W.P.M.H. Heemels and D.J. Antunes

***Abstract*— Optimal control problems can be solved by dynamic programming. However, this method suffers from the curse of dimensionality. To resolve this, simplified versions of the original problem are often constructed in lower-dimensional feature spaces, leading to approximate policies. Yet, the connections between the original and the approximate policy and costs are rarely formalized. This paper addresses this challenge for optimal stopping problems. We start by providing conditions for lossless feature representations. This means that from an optimal policy obtained in feature space, an optimal policy in the original space can be constructed. Then, we search for modified versions of the original problem that (i) admit a lossless feature representation of far lower dimension; and (ii) provide upper and lower bounds on the optimal cost of the original problem. We can then use policies obtained in feature space using these modified problems to provide approximate policies for the original problem that are guaranteed to perform better than or equal to this aforementioned cost upper bound. We apply our tools in a high-dimensional precision farming intervention problem, where our tools allow for a dramatic decrease in complexity with only a small increase in the cost.**

I. Introduction

The method of *dynamic programming* is an invaluable tool for finding optimal policies in a wide range of optimal control and decision problems. However, for high-dimensional problems, the method is computationally intractable due to the curse of dimensionality. To mitigate this, approximate methods are often pursued.

Especially relevant to the present paper are approximate approaches that construct a feature space of lower dimension on which it is feasible to apply exact dynamic programming, providing a tractable *approximate* policy to the original problem. A prime method to accomplish this is state aggregation [1], [2], [3]. Other tools such as certainty equivalent control [4] and Belief Compression for POMDPs [5] can also be framed in this context. Moreover, a wide range of application-specific related approaches can be found in the literature, for instance in diesel engine control [6], resource allocation [7], missile defense problems [8], and vision-based control [9].

However, in these approximate approaches, the connection between the original and the approximate policy and costs is rarely formalized. The present paper aims to address this challenge for a specific class of dynamic programming problems, namely that of *optimal stopping problems*. The goal in these problems is to minimize a given cumulative cost function. This is done by finding a policy either to stop and incur a stopping cost or continue incurring the stage cost. These problems arise in a wide range of fields, from the pricing of financial derivatives [10] to connection scheduling in ad-hoc networks [11]. Moreover, average cost optimal state reset problems, that appear in event-triggered control and precision farming [12], can also be framed as optimal stopping problems.

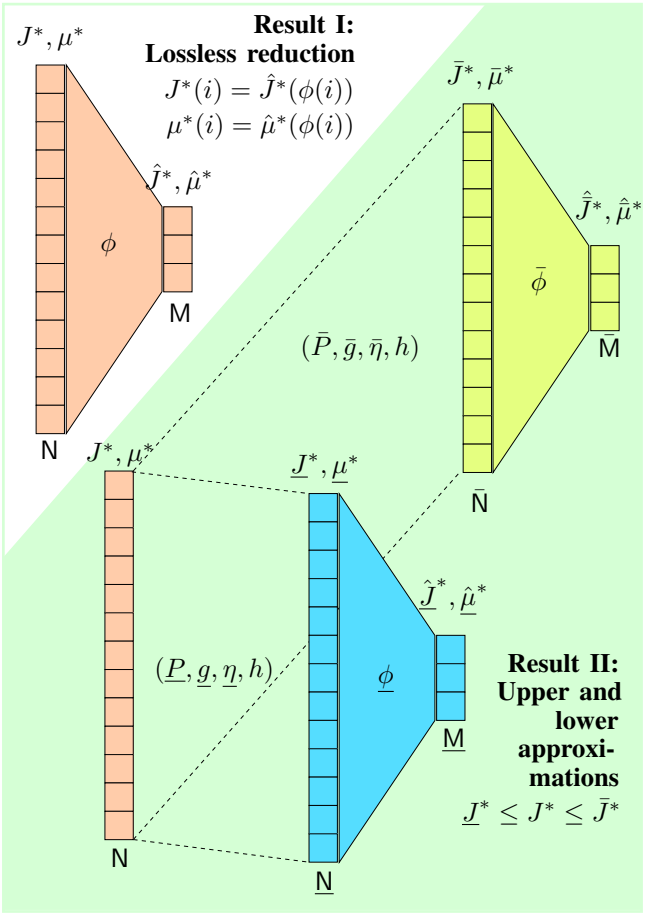

Fig. 1. Visualization of the main results on optimal stopping problems, (I) enabling lossless state reductions; (II) construction of highly-reducible upper- and lower approximations on the original problem.

In this paper, we provide two main theoretical contributions and an application of these ideas to a case study. The first contribution is a set of conditions on the structure of the optimal stopping problem that allows for a *lossless* feature representation of reduced dimension. This means that the optimal value function and the optimal policy in the original space can be obtained from the optimal value

The authors are with the Control Systems Technology Group, Department of Mechanical Engineering, Eindhoven University of Technology, The Netherlands. E-mails:{m.j.t.c.v.zutphen, m.heemels,d.antunes}@tue.nl. This research is part of the research program SYNERGIA (project number 17626), which is partly financed by the Dutch Research Council (NWO).

function and optimal policy in the feature space. As a second contribution, we develop a technique for creating modified versions of the original problem that (i) admit a lossless feature representation of far lower dimension; and (ii) provide upper and lower bounds on the optimal cost of the original problem. We can then use the optimal policies for the modified problems in feature space as approximate policies for the original problem with guaranteed upper and lower bounds on the associated cost (see Fig. 1).

Our results and methods are not only stated and applicable to the standard cumulative cost case, but also to the average cost optimal stopping problems with state resets. We apply our new method on an average cost problem arising in precision farming intervention problems, where our tools are shown to enable a dramatic decrease in complexity, at only a small cost increase.

The remainder of this paper is organized as follows. Section II introduces optimal stopping problems and states the problem. Conditions for a lossless reduction of the state space are presented in Section III. Approximate solutions that leverage these lossless conditions are discussed in Section IV. An example of a problem on which these techniques can be applied, together with the corresponding results, is provided in Section V. Lastly, Section VI provides conclusions and future work directions.

II. Problem formulation

We start by introducing the (standard) *cumulative cost* optimal stopping problem in Section II-A and the *average cost* optimal stopping problem in Section II-B. In Section II-C we define ϕ-lossless maps, upper- and lower approximations for the aforementioned problem classes, and provide the problem statement.

A. Cumulative cost optimal stopping problem

Consider a standard, cumulative cost, optimal stopping problem [13] with full state information on a discrete finite space $x_t \in \mathsf{N} := \{1, 2, \ldots, n\}$, for $t \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$ that aims to find a policy determining the stopping time $\tau \in \mathbb{N}$, which minimizes the expected cost function $J : \mathsf{N} \to \mathbb{R}$ given by

$$J(x_0) = \mathbb{E}[\sum_{t=0}^{\tau-1} g(x_t) + \eta(x_\tau)], \tag{1}$$

with initial condition $x_0 \in \mathsf{N}$, stage cost function $g : \mathsf{N} \to \mathbb{R}$, and stopping cost $\eta : \mathsf{N} \to \mathbb{R}$ subject to stationary stochastic dynamics

$$p_{ij} = \text{Prob}[x_{t+1} = j | x_t = i], \tag{2}$$

for $i, j \in \mathsf{N}$. We define $P := [p_{ij}]_{i,j \in \mathsf{N}}$. For simplicity, assume that the stopping time is restricted by $\tau \leq h$ for a given, possibly arbitrarily large $h \in \mathbb{N}$. The stopping time is technically a measurable function of the filtration associated with the process x_t, $t \in \mathbb{N}_0$ [14]. Equivalently, we can define $u_t \in \{0, 1\}$ to be such that $u_t = 1$, if $\tau = t$, and $u_t = 0$, otherwise, and search for a policy $u_t = \mu_t(x_t)$; then $\tau = \min\{\ t \in \mathbb{N}_0 \mid u_t = 1\}$.

The optimal solution to this problem can be obtained using the exact stochastic dynamic programming algorithm [14]; start with

$$J_h^*(i) = \eta(i), \tag{3a}$$

for all $i \in \mathsf{N}$. Subsequently, for each $t \in \{h-1, h-2, \ldots, 0\}$, and $i \in \mathsf{N}$, compute

$$J_t^*(i) = \min\{g(i) + \textstyle\sum_{j=1}^{n} p_{ij} J_{t+1}^*(j),\ \eta(i)\}, \tag{3b}$$

where the minimum of (1), denoted by $J^*(x_0)$, is obtained as $J^*(x_0) = J_0^*(x_0)$. Then, an associated optimal policy is

$$\mu_t^*(x_t) = \begin{cases} 1, & \text{if } J_t^*(x_t) = \eta(x_t), \\ 0, & \text{otherwise,} \end{cases} \tag{3c}$$

for every $t \in \{0, 1, \ldots, h-1\}$. We call (μ_t^*, J_t^*) a solution pair of the optimal stopping problem (P, g, η, h).

B. Average cost optimal stopping problems

Consider the average cost

$$\tilde{J}(x_0) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\left[(1 - u_t)\tilde{g}(x_t) + u_t \eta(x_t)\right], \tag{4}$$

subject to the same stationary stochastic dynamics (2), when $u_t = 0$, and state reset $x_{t+1} = x_0$, if $u_t = 1$, with stage cost function $\tilde{g} : \mathsf{N} \to \mathbb{R}$, and restriction $\forall k \in \mathbb{N}_0\ \exists \ell \in [k, k+h]$ such that $u_\ell = 1$, i.e., stopping time $\tau \leq h$. It is well known that the problem of minimizing this average expected cost (4) can equivalently be formulated as the average cost optimal stopping problem [12], [14]

$$\tilde{J}(x_0) = \frac{1}{\mathbb{E}[\tau]} \mathbb{E}[\textstyle\sum_{t=0}^{\tau-1} \tilde{g}(x_t) + \eta(x_\tau)]. \tag{5}$$

The optimal solution to this problem can be obtained using the numerical dynamic programming root-finding problem

$$\min_{\tau} \mathbb{E}[\textstyle\sum_{t=0}^{\tau-1} (\tilde{g}(x_t) - \beta) + \eta(x_\tau)] = 0, \tag{6}$$

with $\beta \geq 0$. Solutions to the minimization of the left-hand side of (6) for a given value of β are obtained by executing (3) with the adapted stage cost value $g(x_t) = \tilde{g}(x_t) - \beta$. The unique β that satisfies equation (6) can subsequently be found by, e.g., a binary search as the solution is monotonic in β.

C. Problem statement

As the solutions to both optimal stopping problems (1), (5) rely on exact dynamic programming, they become intractable for problems with a large state-space cardinality $n = |\mathsf{N}|$. Unfortunately, this is the case in many problems, e.g., those resulting from the discretization of problems in $\mathbb{R}^{\tilde{n}}$ with large $\tilde{n}$. In this paper, the aim is to achieve tractable solutions for these problems by means of a proxy problem in a lower-dimensional feature space, which we explain in more detail below.

Suppose that we group states according to a surjective map (projection)

$$\phi : \mathsf{N} \to \mathsf{M}, \tag{7}$$

where $\mathsf{M} := \{1, 2, \ldots, m\}$ with $m \leq n$. The map ϕ defines an equivalence relation. States $i, j \in \mathsf{N}$ are equivalent, denoted by $i \sim j$, if $\phi(i) = \phi(j)$. This map can induce an optimal control (proxy) problem in this lower-dimensional

space (see Fig. 2) with state $\hat{x}_t \in \mathsf{M}$, as we will explain in more detail below. We declare such map ϕ to be lossless if the solution pair $(\mu_t^*(x_t), J_t^*(x_t))$ can be recovered from a lower-dimensional solution pair $(\hat{\mu}_t^*(\hat{x}_t), \hat{J}_t^*(\hat{x}_t))$ of the problem in feature space taking value in M.

Definition 1 (ϕ-lossless) *A stopping time problem on state space $\mathsf{N} = \{1, 2, \ldots, n\}$ characterized by (P, g, η, h) is said to be ϕ-lossless for a given $\phi : \mathsf{N} \to \mathsf{M}$ with $m < n$, if there exists a reduced order problem $(\hat{P}, \hat{g}, \hat{\eta}, h)$ on state space $\mathsf{M} = \{1, 2, \ldots, m\}$, with optimal cost- and policy functions $\hat{J}_t^* : \mathsf{M} \to \mathbb{R}$, $\hat{\mu}_t^* : \mathsf{M} \to \{0, 1\}$, such that*

$$J_t^*(x_t) = \hat{J}_t^*(\phi(x_t)), \quad \textit{and} \quad \mu_t^*(x_t) = \hat{\mu}_t^*(\phi(x_t)), \tag{8}$$

for every $t \in \{0, 1, \ldots, h\}$ and $x_t \in \mathsf{N}$. ■

Note that for ϕ-lossless problems, we can work on the quotient space $\mathsf{N}/\sim$ of this equivalence relation, i.e., work with m states rather than n states, without losing optimality.

To tackle problems that are not ϕ-lossless for any ϕ, or that are ϕ-lossless for ϕ's where $m \approx n$, it is useful to define the notions of upper and lower approximations. Upper (lower) approximation refers to an abstract optimal stopping problem defined by the tuple $(\bar{P}, \bar{g}, \bar{\eta}, h)$ $((\underline{P}, \underline{g}, \underline{\eta}, h))$ and an associated cost function. Here, *abstract* refers to the fact that the rows of $\bar{P}$ ($\underline{P}$), in contrast to the rows of a standard optimal stopping transition probability matrix P, are not required to satisfy the discrete probability distribution property of summing to one. The abstract optimal stopping formulation requires a redefinition of the expected value operator present in e.g., (1) and (4). Let $\mathbf{x} = (x_1, x_2, \ldots, x_h) \in \mathsf{N}^h$ be a state sequence and let the set of all unique state sequences be indexed by ℓ, as $\mathsf{N}^h = \{\mathbf{x}^\ell = (x_1^\ell, x_2^\ell, \ldots, x_h^\ell) | \ell \in \{1, 2, \ldots, n^h\}\}$. Let the function $q : \{1, 2, \ldots, n^h\} \times \mathbb{R}^{n \times n} \to \mathbb{R}$ be defined as

$$q(\ell, P) = \Pi_{k \in \{1,2,\ldots,h-1\}} \; p_{x_k^\ell x_{k+1}^\ell}. \tag{9}$$

We use this to redefine the $\mathbb{E}_P[\,\cdot\,]$-operator as

$$\mathbb{E}_P[V(\mathbf{x})] = \sum\nolimits_{\ell=1}^{n^h} q(\ell, P) V(\mathbf{x}^\ell), \tag{10}$$

for $V : \mathsf{N}^h \to \mathbb{R}$. For transition probability matrices P that satisfy the conditions of a probability distribution, this definition coincides with the canonical definition of expected value $\mathbb{E}[\,\cdot\,]$, but for P-matrices which do not correspond to probability distributions, it should be interpreted as a distinct mathematical operator.

Evaluating the cost (1) using dynamic programming under this redefinition of the $\mathbb{E}[\,\cdot\,]$-operator, for a given stopping policy μ, and for any tuple (P, g, η, h) results in

$$J_t^\mu(i) = \begin{cases} g(i) + \sum_{j=1}^n p_{ij} J_{t+1}^\mu(j), & \text{if } \mu_t(i) = 0, \\ \eta(i), & \text{if } \mu_t(i) = 1. \end{cases} \tag{11}$$

Let us denote the cost (1) realized under policy μ_t and tuple (P, g, η, h) for $t \in \mathbb{N}_0$ by $J^\mu(x_0) = J_0^\mu(x_0)$, and a cost realized by an upper-bounding (lower-bounding) problem under the same policy by $\bar{J}^\mu(x_0) = \bar{J}_0^\mu(x_0)$ ($\underline{J}^\mu(x_0) = \underline{J}_0^\mu(x_0)$), where $\bar{J}_0^\mu(x_0)$ ($\underline{J}_0^\mu(x_0)$) is obtained by replacing p_{ij} by $\bar{p}_{ij}$ ($\underline{p}_{ij}$) in (9), (11).

Definition 2 (Upper- and lower approximations) *An abstract optimal stopping problem characterized by $(\bar{P}, \bar{g}, \bar{\eta}, h)$ ($(\underline{P}, \underline{g}, \underline{\eta}, h)$) is said to be an upper (lower) approximation of the optimal stopping problem characterized by (P, g, η, h), if for any stopping policy $\mu_t : \mathsf{N} \to \{0, 1\}$, $t \in \{0, \ldots, h-1\}$,*

$$J^\mu(x_0) \leq \bar{J}^\mu(x_0) \;\; (\underline{J}^\mu(x_0) \leq J^\mu(x_0)), \qquad \forall x_0 \in \mathsf{N}.$$

■

In this paper we are interested in two problems:

(i) Providing conditions on (P, g, η, h) under which cumulative and average cost optimal stopping problems are ϕ-lossless.

(ii) For problems that are either not ϕ-lossless or are ϕ-lossless with a ϕ where $m \approx n$, providing a procedure that modifies (P, g, η, h) to form upper and lower approximations that are ϕ-lossless with $m \ll n$.

III. Main result I: Lossless condition

The first main result provides sufficient conditions for the considered stopping time problems to be ϕ-lossless.

A. Conditions for lossless state aggregation

Let us define m equivalence sets $\mathcal{A}_j := \{i \in \mathsf{N} \mid \phi(i) = j\}$, $j \in \mathsf{M}$.

Theorem 1 (Lossless reduction) *If the three conditions*

$$\textit{(1.i)} \quad g(i) = g(r), \qquad \forall i, r \in \mathsf{N}, \; \textit{s.t. } i \sim r,$$

$$\textit{(1.ii)} \quad \sum_{\ell \in \mathcal{A}_j} p_{i\ell} = \sum_{\ell \in \mathcal{A}_j} p_{r\ell}, \; \forall i, r \in \mathsf{N}, \; \textit{s.t. } i \sim r, \forall j \in \mathsf{M},$$

$$\textit{(1.iii)} \quad \eta(i) = \eta(r), \qquad \forall i, r \in \mathsf{N}, \; \textit{s.t. } i \sim r,$$

are met, then the optimal stopping problem defined by (P, g, η, h) admits a ϕ-lossless feature-based state-space reduction. ■

This means we can aggregate all states from each class $\mathcal{A}_j$ for $j \in \mathsf{M}$ into a single representative state when running the DP algorithm, reducing the number of operations from order n to order m. This holds for both the cumulative cost (1) and average cost (4) optimal stopping problems.

B. Construction of the state aggregation

Provided a set of equivalence groups $\mathcal{A}_j$ for $j \in \mathsf{M}$ for problem (P, g, η, h) of state space size n, we can reduce the problem losslessly to $(\hat{P}, \hat{g}, \hat{\eta}, h)$ with a state space of size m through

$$\hat{P} = \Pi_r P \Pi_c, \quad \hat{g} = \Pi_r g, \quad \hat{\eta} = \Pi_r \eta, \tag{12}$$

Fig. 2. Solving the original problem (P, g, η, h) in feature space M by lossless feature mapping ϕ generating proxy problem $(\hat{P}, \hat{g}, \hat{\eta}, h)$ with state space cardinality $m < n$. An optimal solution pair (μ^*, J^*) is losslessly recovered from $(\hat{\mu}^*, \hat{J}^*)$.

where $\Pi_r \in \{0,1\}^{m\times n}$ and $\Pi_c \in \{0,1\}^{n\times m}$ are permutation matrices constructed by stacking row $r_j \in \{0,1\}^n$ and column $c_j \in \{0,1\}^n$ vectors for $j \in \mathsf{M}$ as

$$\Pi_r = \begin{bmatrix} r_1^\top & r_2^\top & \cdots & r_m^\top \end{bmatrix}^\top, \quad \Pi_c = \begin{bmatrix} c_1 & c_2 & \cdots & c_m \end{bmatrix},$$

with

$$r_j(i) = \begin{cases} 1, & \text{if } i = e_j, \\ 0, & \text{otherwise,} \end{cases} \quad \text{and} \quad c_j(i) = \begin{cases} 1, & \text{if } i \in \mathcal{A}_j, \\ 0, & \text{otherwise,} \end{cases}$$

for m arbitrary preselected elements $e_j \in \mathcal{A}_j$, and $i \in \mathsf{N}$, $j \in \mathsf{M}$.

In short, this procedure merges transition probabilities of each equivalence class into their respective feature space state $j \in \mathsf{M}$ through the post-multiplication with Π_c, while also removing all remaining redundant states through a pre-multiplication by Π_r.

IV. Main result II: Approximating policies

The second main result consists of sufficient conditions for modified problems to form an upper bound (or lower bound) on the cost of an optimal stopping problem. These conditions are leveraged below to create approximations endowed with ϕ-lossless structure, solvable on strongly reduced state spaces.

A. Conditions for upper and lower approximations

Theorem 2 (Upper and lower approximations) *If, with respect to optimal stopping problem (P, g, η, h), a proxy problem $(\underline{P}, \underline{g}, \underline{\eta}, h)$ ($(\bar{P}, \bar{g}, \bar{\eta}, h)$) satisfies*

(2.i) $\quad 0 \le \underline{g}(i) \le g(i)$ $(0 \le g(i) \le \bar{g}(i)), \quad \forall i \in \mathsf{N},$

(2.ii) $\quad 0 \le \underline{p}_{ij} \le p_{ij}$ $(0 \le p_{ij} \le \bar{p}_{ij}), \ \forall i, j \in \mathsf{N},$

(2.iii) $\quad 0 \le \underline{\eta}(i) \le \eta(i)$ $(0 \le \eta(i) \le \bar{\eta}(i)), \quad \forall i \in \mathsf{N},$

then the proxy problem is a lower (upper) approximation w.r.t. the optimal stopping problem. ■

The optimal policy $\bar{\mu}^*$ for an upper approximation has a proven upper-bound on the cost when applied to the original problem $J^{\mu^*} \le J^{\bar{\mu}^*} \le \bar{J}^{\bar{\mu}^*}$, where μ^* is the optimal policy for the original problem. Comparing the optimal cost of the upper approximation $\bar{J}^{\bar{\mu}^*}$ to the optimal cost of a lower approximation $\underline{J}^{\underline{\mu}^*}$ additionally provides an upper-bound on the difference between the cost realized using the approximate policy $\bar{\mu}^*$ and the true optimal cost as

$$\underline{J}^{\underline{\mu}^*} \le J^{\mu^*} \le J^{\bar{\mu}^*} \le \bar{J}^{\bar{\mu}^*}, \tag{13}$$

using Theorem 2 and the definition of optimality. Here, J should be interpreted as the cost w.r.t. either (1) or (4).

Note that Theorem 2 requires the stage cost $g(i)$, $i \in \mathsf{N}$, of the original problem to be non-negative, which is not necessarily the case for the reformulation (6) of the average cost problem (4). Yet, Theorem 2 does turn out to hold for the average cost problem (4) in the case of non-negative stage cost $\tilde{g}(i)$. Such average cost problem is characterized by the tuple $(P, \tilde{g}, \eta, h)$. Upper (lower) approximations can be defined as in Definition 2 provided that we now interpret $\bar{J}^\mu(x_0)$ ($\underline{J}^\mu(x_0)$) as cost (4), or equivalently (5), for policy μ). While such cost can no longer be written through (11), it is still well defined as we can write the numerator in (5) as a function of the $\bar{p}_{ij}$ ($\underline{p}_{ij}$) through (9), and the denominator as $\mathbb{E}_P[\tau(\mathbf{x})] = \sum_{\ell=1}^{n^h} q(\ell, P)\tau(\mathbf{x}^\ell)$.

Theorem 3 (Average cost approximation) *Consider the average cost (4), tuple $(P, \tilde{g}, \eta, h)$, and a proxy problem $(\underline{P}, \underline{g}, \underline{\eta}, h)$ ($(\bar{P}, \bar{g}, \bar{\eta}, h)$) that satisfies conditions (2.i), (2.ii), and (2.iii) with $g = \tilde{g}$. Then the modified problem is a lower (upper) approximation w.r.t. the optimal stopping problem.* ■

Theorem 3 allows us to construct upper- and lower approximations for average cost problems and cumulative cost problems alike.

B. Construction of the modified problems

In Algorithm 1, we present a general procedure that uses Theorem 2 to obtain an upper approximation for cumulative- and average cost optimal stopping problems with arbitrarily low $\bar{m} < m \le n$. For the method to be applicable, the original problem (P, g, η, h) must satisfy conditions (2.*i*), (2.*ii*) and (2.*iii*) of Theorem 2. Problems that satisfy these conditions include the many cumulative cost problems where $g(i) = 0$ and $\eta(i) > 0$ for all $i \in \mathsf{N}$, such as moment-of-purchase problems and the many average cost problems that generally satisfy $g(i) \ge 0$ and $\eta(i) > 0$ for all $i \in \mathsf{N}$.

Algorithm 1 Upper approximation

1. Select an arbitrary number $\bar{m} \in \mathbb{N}$ of *seed states* $\bar{\mathcal{S}} = \{i_1, i_2, \ldots, i_{\bar{m}}\} \subseteq \mathsf{N}$, each required to have a unique stage cost value, $g(i_j) \ne g(i_r)$ for $j, r \in \mathsf{M}$, $j \ne r$. It is further required that the state i with the largest stage cost $g(i)$ is in $\bar{\mathcal{S}}$.
2. Create stage cost function $\bar{g}$ such that $\bar{g}(i) := \min\{g(j) \mid g(i) \le g(j),\ j \in \bar{\mathcal{S}}\}$, for $i \in \mathsf{N}$, so that the equivalence classes emerge as $\bar{\mathcal{A}}_j := \{r \in \mathsf{N} \mid \bar{g}(r) = \bar{g}(i_j)\}$, for $j \in \bar{\mathsf{M}}$.
3. Set stopping costs $\bar{\eta}$ as $\bar{\eta}(i) = \max\{\eta(r) \mid r \in \bar{\mathcal{A}}_{j(i)}\}$, where $j(i) \in \bar{\mathsf{M}}$ is such that $i \in \mathcal{A}_{j(i)}$.
4. Define $\varepsilon_{ir} := \mathrm{Prob}[x_{t+1} \in \bar{\mathcal{A}}_r | x_t = i] = \Sigma_{\ell \in \bar{\mathcal{A}}_r} p_{i\ell}$, and find $M_{jr} = \max_{i \in \bar{\mathcal{A}}_j} \varepsilon_{ir}$ for all $j, r \in \bar{\mathsf{M}}$.
5. Compute the elements of $\bar{P}$ for $i, \ell \in \mathsf{N}$ as

$$\bar{p}_{i\ell} := \begin{cases} M_{jr}\frac{p_{i\ell}}{\varepsilon_{ir}} & \text{if } \varepsilon_{ir} > 0, \\ M_{jr}\frac{1}{|\bar{\mathcal{A}}_r|} & \text{if } \varepsilon_{ir} = 0, \end{cases}$$

where $j = \{k \in \bar{\mathsf{M}} | i \in \bar{\mathcal{A}}_k\}$, $r = \{k \in \bar{\mathsf{M}} | \ell \in \bar{\mathcal{A}}_k\}$.

The algorithm thus produces a ϕ-lossless upper approximation to the original problem with a significantly reduced number of equivalence classes $\bar{m} \ll m$ by increasing the stage- and stopping costs, and transition likelihoods of certain states to make previously dissimilar states equivalent. To initiate, in step 1, we somewhat arbitrarily select a set of $\bar{m}$ states from N such that each has a unique stage cost value $g(i)$ and one of the states with maximum stage cost is included. This is done to have a basis of seed states $\bar{\mathcal{S}}$ to which all other states are subsequently made equivalent.

In step 2, we construct $\bar{g}(i)$ by simply raising each of the original stage cost values $g(i)$ for $i \in \mathsf{N} \setminus \bar{\mathcal{S}}$ until they equal the stage cost of one of the seed states i_j for $j \in \bar{\mathsf{M}}$. This is

done to ensure $\bar{g}$ satisfies (2.*i*) for all emerging equivalence groups, which are made explicit as $\bar{\mathcal{A}}_j$, $j \in \bar{\mathsf{M}}$. Similarly, in step 3, we construct stopping cost function $\bar{\eta}(i)$ to always equal the maximum value $\eta(i)$ takes on the elements of each equivalence group for all equivalent states $i \in \bar{\mathcal{A}}_j$. This ensures the satisfaction of condition (2.*iii*).

Lastly, we find what the largest chance of transitioning from group $j \in \bar{\mathsf{M}}$ to any state in equivalence group $r \in \bar{\mathsf{M}}$ is (step 4) to define the elements $\bar{P} = [\bar{p}_{i\ell}]$ for $i, \ell \in \mathsf{N}$ such that the resulting transition sums $\bar{\varepsilon}_{ir}$ to each equivalence group will be equal for all equivalent states in $\bar{\mathcal{A}}_{j(i)}$ (step 5), note that $j(i) \in \bar{\mathsf{M}}$ selects the equivalence class index to which $i \in \mathsf{N}$ belongs.

Note that Algorithm 1 can be repurposed to produce lower- instead of upper approximations. This is done by (i) ensuring the smallest rather than the largest stage cost value state is included in $\bar{\mathcal{S}}$ in step 1, (ii) reversing the inequality in step 2, and (iii) replacing the max function with a min function in steps 3, and 4. Note also that for problems where more information about the states is stored in the terminal cost $\eta(i)$ than in the stage cost $g(i)$, e.g., when $g(i) = 0$ for all $i \in \mathsf{N}$, the algorithm should be modified by swapping g for η and vice versa in all steps.

V. Example application, precision farming

To demonstrate the efficacy of the tools developed above, an example problem in precision farming is employed. The problem is introduced and then solved both approximately and exactly when the size of its state space allows for it. Above this threshold, it is shown that the approximate methods can generate solutions when the exact methods have become intractable, providing approximate solutions with proven upper- and lower bounds on the cost.

Take an agricultural field, discretized spatially along a single dimension into $N \in \mathbb{N}$ subfields. We describe the state of each subfield f_i for $i \in \{1, 2, \dots, N\}$ at timestep $t \in \mathbb{N}_0$ using a binary value

$$f_i(t) \in \{0, 1\},$$

where $f_i(t) = 0$ corresponds to *"no weeds present"* and $f_i(t) = 1$ to *"subfield infected with weeds"*. These values are stored in field state vector $f(t) = \begin{bmatrix} f_1(t) & f_2(t) & \cdots & f_N(t) \end{bmatrix}^\top$.

Since $f \in \{0, 1\}^N$, the exact number n of unique field state values f can take is $|\{0, 1\}^N| = 2^N$. This allows the state of the entire field at every time step $t \in \mathbb{N}_0$ to be uniquely represented by a single integer state label as $x_t \in \mathsf{N} := \{1, 2, \dots, n\}$ for $n = 2^N$. This labeling can be defined through the bijective mapping $\pi : \{0, 1\}^N \to \mathsf{N}$.

Uninfected subfields $f_i(t) = 0$ for $i \in \{1, 2, \dots, N\}$, $t \in \mathbb{N}_0$, can become infected by seeds transmitted through the air or by infection from an infected neighboring subfield. The probabilities of either of these happening to an uninfected subfield at any timestep t are assumed

$$0 < p_{\text{air}} < p_{\text{neighbor}} < 1.$$

Thus, since subfields i for $i \in \{1, 2, \dots, N\}$, at any timestep t have $n_{\text{in},i}(t) \in \{0, 1, 2\}$ infected neighbors, their chance of infection

$$p_i(t) = \text{Prob}[f_i(t+1) = 1 | f_i(t), n_{\text{in},i}(t)],$$

at timestep t is found as $p_i(t) = 1$, if $f_i(t) = 1$, and

$$p_i(t) = \begin{cases} p_{\text{air}}, & \text{if } n_{\text{in},i}(t) = 0, \\ 1 - (1 - p_{\text{air}})(1 - p_{\text{neighbor}}), & \text{if } n_{\text{in},i}(t) = 1, \\ 1 - (1 - p_{\text{air}})(1 - p_{\text{neighbor}})^2, & \text{if } n_{\text{in},i}(t) = 2, \end{cases} \tag{14}$$

when $f_i(t) = 0$. Using π and this expression (14) we can create the transition probability matrix P that contains the probabilities $P = [p_{ij}]$ of the system transitioning from state $i \in \mathsf{N}$ to state $j \in \mathsf{N}$ in one timestep.

The cost of a field i being infected $f_i(t) = 1$ is $d \in \mathbb{R}_{>0}$ per timestep. Thus, the stage cost of each state $x_t \in \mathsf{N}$ is found as

$$g(x_t) = d \mathbb{1}_n^\top \pi^{-1}(x_t),$$

where $\pi^{-1}(x_t)$ is the inverse mapping $\pi^{-1} : \mathsf{N} \to \{0, 1\}^N$ of π and $\mathbb{1}_n$ is an $n \times 1$-vector of ones. The stopping cost is equal for all states and is defined as

$$\eta(x_t) = \delta, \quad \text{for all } x_t \in \mathsf{N},$$

where $\delta \in \mathbb{R}_{>0}$, as it is assumed there is a set cost associated with treating the entire field — so all subfields — with herbicides, which resets all subfield states to zero. Minimization of the average cost (4) is considered.

A. Solution

As the number of subfields N grows, the cardinality of the state space $n = 2^N$ grows exponentially. For a field that is subdivided in, e.g., $N = 30$ subfields, the related state space has a cardinality of $n > 1 \cdot 10^9$. Since P is an $n \times n$-matrix, this matrix alone would have more than $1 \cdot 10^{18}$ elements. This particular problem in its original form (P, g, η, h) turns out to be ϕ-lossless with $m \approx n/2$. This reducibility stems from mirror image states in $\{0, 1\}^N$. For $N = 2$, e.g., we have that $[0\ 1]^\top$ is the mirror image of $[1\ 0]^\top$, and thus enjoys the same current and expected future degrees of infection. Although this is a significant reduction, its lossless proxy problem is still far from tractable for $N \gg 10$.

Note that the original stage cost $g(i)$ takes only $N + 1$ unique values for $i \in \mathsf{N}$. Algorithm 1 applied to this example with $\bar{m} = N + 1$ thus generates upper- and lower approximations on the original problem with $\bar{g} = \underline{g} := g$ and $\bar{\eta} = \underline{\eta} := \eta$.

B. Results

We consider this example situation for cost- and transition probability parameters $(\delta, d, p_{\text{air}}, p_{\text{nei}}) = (10, 1, 0.1, 0.5)$, and a range of field discretizations, $N \in \{1, 2, \dots, 18\}$. The resulting optimal average cost is compared to the results from the related upper and lower approximations of the same problem. This is visualized in Fig. 3. Above $N = 14$, solving the original problem directly became too demanding without special hardware and only the approximate problems have been solved.

In Fig. 4, the relationships between the number of subfields N and the size of the corresponding state spaces

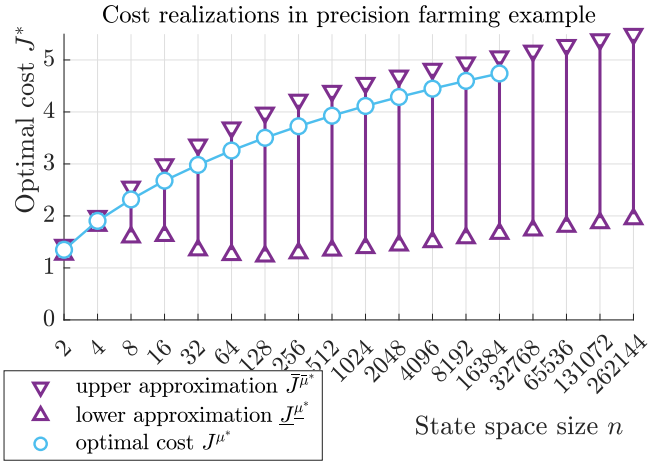

Fig. 3. The optimal cost of the example problem for $N = 1, 2, \ldots, 14$, as well as the costs associated with the upper- and lower approximations of this original problem for $N = 1, 2, \ldots, 18$.

$n = \bar{n} = \underline{n}$, m, $\bar{m} = \underline{m}$, are visualized. At $N = 132$, the number of values $n \times n$ comprising P already exceeds the estimated amount of atoms in the universe [15] by a factor of almost three. For this same scenario, the reduced P-matrices of the approximate problems ($\hat{\bar{P}}$ and $\hat{\underline{P}}$) require only $\bar{m} \times \bar{m} = \underline{m} \times \underline{m} = 19044$ elements.

As described in Section IV, the straightforward procedure of obtaining $\hat{\bar{P}}$ and $\hat{\underline{P}}$ starts by producing P. This matrix P is then adapted using Algorithm 1 to yield the $\bar{P}$ and $\underline{P}$ matrices. Lastly, the $\bar{P}$ and $\underline{P}$ matrices are reduced using Eq. (12) to form $\hat{\bar{P}}$ and $\hat{\underline{P}}$ (see also Fig. 1). As the size of P grows exponentially (as displayed in Fig. 4), executing the method this way is extremely memory intensive as it requires storing three matrices of size $n \times n$. This is why the computations that were required to obtain the results displayed in Fig. 3 have been modified to produce $\hat{\bar{P}}$ and $\hat{\underline{P}}$ directly. This modification can be obtained straightforwardly by executing Algorithm 1 and Eq. 12 for each element of $\hat{\bar{P}}$ and $\hat{\underline{P}}$ instead of for the entire matrices at once.

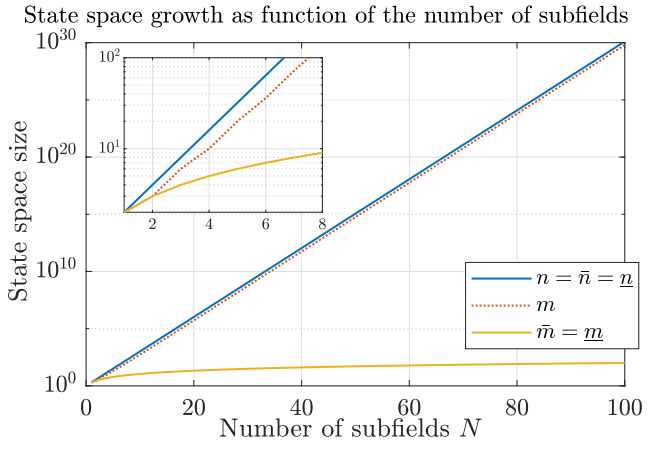

Fig. 4. The growth of the size of the original state space n, upper/lower approximation state space $\bar{n} = \underline{n}$, losslessly reduced state space m and losslessly reduced upper/lower approximation state space $\bar{m} = \underline{m}$ as a function of the number of subfields N.

VI. Conclusions, Future Work and Discussion

We introduced the notion of ϕ-losslessness for stopping time problems, provided sufficient conditions that can be used to verify this property for a given problem, and provided a method that generates the associated reduced problem. Secondly, we defined upper (lower) approximations of far lower dimension that not only form a guaranteed upper (lower) bound on the optimal cost of the original problem, but also yield an approximate policy with a guaranteed upperbound on its performance.

The natural extension of this work is to generalize the same techniques to general Markov decision problems. Additional research directions that emerge from the development of these techniques are (i) the investigation of how Algorithm 1 can be modified to produce the tightest possible upper/lower bounds, and (ii) the search for fast algorithms for finding ϕ, i.e., the existing equivalence classes in a given problem.

References

[1] D. P. Bertsekas, "Aggregation," in *Dynamic Programming and Optimal Control*. Athena Scientific, 2005, pp. 319–325.

[2] M. N. Katehakis and L. C. Smit, "A successive lumping procedure for a class of markov chains," *Probability in the Engineering and Informational Sciences*, vol. 26, pp. 483 – 508, 2012.

[3] P. J. Schweitzer and K. W. Kindle, "Iterative aggregation for solving undiscounted semi-markovian reward processes," *Stochastic Models*, vol. 2, pp. 1–41, 1986.

[4] D. P. Bertsekas, "Certainty equivalent and adaptive control," in *Dynamic Programming and Optimal Control*. Athena Scientific, 2005, pp. 283–300.

[5] N. Roy, G. Gordon, and S. Thrun, "Finding approximate POMDP solutions through belief compression," vol. 23, no. 1, 2005.

[6] Z. Huang and J. Ma, "Solving the curse of dimensionality utilizing action-dependent heuristic dynamic programming," *2011 IEEE International Conference on Computer Science and Automation Engineering*, vol. 2, pp. 289–292, 2011.

[7] A. Forootani, R. Iervolino, M. Tipaldi, and J. Neilson, "Approximate dynamic programming for stochastic resource allocation problems," *IEEE/CAA Journal of Automatica Sinica*, vol. 7, pp. 975–990, 2020.

[8] D. P. Bertsekas, M. L. Homer, D. A. Logan, S. D. Patek, and N. R. Sandell, "Missile defense and interceptor allocation by neuro-dynamic programming," *IEEE Trans. Syst. Man Cybern. Part A*, vol. 30, pp. 42–51, 2000.

[9] S. Kim, H. Seo, S. Choi, and H. J. Kim, "Vision-guided aerial manipulation using a multirotor with a robotic arm," *IEEE/ASME Transactions on Mechatronics*, vol. 21, pp. 1912–1923, 2016.

[10] J. Tsitsiklis and B. van Roy, "Optimal stopping of markov processes: Hilbert space theory, approximation algorithms, and an application to pricing high-dimensional financial derivatives," *IEEE Transactions on Automatic Control*, vol. 44, no. 10, pp. 1840–1851, 1999.

[11] D. Zheng, W. Ge, and J. Zhang, "Distributed opportunistic scheduling for ad hoc networks with random access: An optimal stopping approach," *IEEE Transactions on Information Theory*, vol. 55, no. 1, pp. 205–222, 2009.

[12] D. Antunes, R. Beumer, M. van de Molengraft, and W. Heemels, "Adaptive sampling and actuation for POMDPs: Application to precision agriculture," 2022, preprint on webpage at https://heemels.tue.nl/content/papers/AntBeu_CDC22a.pdf.

[13] T. S. Ferguson, *Optimal Stopping and Applications*. UCLA, 2008, ch. 1, pp. 1–8.

[14] D. P. Bertsekas, "Optimal stopping problems," in *Dynamic Programming and Optimal Control*. Athena Scientific, 2005, pp. 176–186.

[15] A. S. Eddington, "The constants of nature," in *The World of Mathematics*, J. R. Newman, Ed., vol. 2. Simon & Schuster, p. 1074–1093.