

A verification theorem for threshold-indexability of real-state discounted restless bandits

José Niño-Mora*

Department of Statistics

Carlos III University of Madrid

28903 Getafe (Madrid), Spain

jose.nino@uc3m.es, <http://orcid.org/0000-0002-2172-3983>

Published in *Mathematics of Operations Research*, vol. 45, 465–496, 2020

DOI: [10.1287/moor.2019.0998](https://doi.org/10.1287/moor.2019.0998)

Online companion available at

<https://pubsonline.informs.org/doi/suppl/10.1287/moor.2019.0998>

Abstract

The Whittle index, which characterizes optimal policies for controlling certain single restless bandit projects (a Markov decision process with two actions: active and passive) is the basis for a widely used heuristic index policy for the intractable restless multiarmed bandit problem. Yet two roadblocks need to be overcome to apply such a policy: the individual projects in the model at hand must be shown to be indexable, so that they possess a Whittle index; and the index must be evaluated. Such roadblocks can be especially vexing when project state spaces are real intervals, as in recent sensor scheduling applications. This paper presents sufficient conditions for indexability (relative to a generalized Whittle index) of general real-state discrete-time restless bandits under the discounted criterion, which are not based on elucidating properties of the optimal value function and do not require proving beforehand optimality of threshold policies as in prevailing approaches. The main contribution is a verification theorem establishing that, if project performance metrics under threshold policies and an explicitly defined marginal productivity (MP) index satisfy three conditions, then the project is indexable with its generalized Whittle index being given by the MP index, and threshold policies are optimal for dynamic project control.

Keywords: Markov decision processes; discounted criterion; discrete time; Whittle index; index policies; indexability; threshold policies

Contents

1 Introduction.	3
1.1 RMABP and Whittle’s index policy.	3
1.2 Indexability.	3

*This work has been developed over a number of years since 2008, with support from the Spanish Ministry of Education and Science [Grant MTM2007-63140], Ministry of Science and Innovation [Grant MTM2010-20808], and Ministry of Economics and Competitiveness [Grant ECO2015-66593-P]. The author has presented parts of this work in a 2016 seminar at Xerox Research Centre (Grenoble, France) and at the 4th Euro-NGI Conference on Next Generation Internet Networks (NGI 2008) (Krakow, Poland), the 3rd Euro-NF Conference on Network Control and Optimization (NETCOOP 2009) (Eindhoven, Netherlands), the 21st International Symposium on Mathematical Programming (ISMP 2012) (Berlin, Germany), the 23rd International Symposium on Mathematical Programming (ISMP 2018) (Bordeaux, France), and the 2013 and 2017 Conferences of the Spanish Royal Mathematical Society. A previous version was posted on arXiv and submitted in 2015.

1.3	Real-state projects: prevailing approaches to indexability based on optimal thresholds.	4
1.4	PCL-based sufficient indexability conditions.	5
1.5	Goal and main results.	6
1.6	Organization of the paper.	6
2	Preliminaries and formulation of the main result.	7
2.1	Notation.	7
2.2	Project model, performance metrics, and λ -price problems.	7
2.3	Indexability.	9
2.4	Thresholdability and threshold-indexability: weak and strong versions.	10
2.5	Marginal metrics and the MP index.	11
2.6	Main result: a verification theorem for threshold-indexability.	12
3	Motivation of the PCL-indexability conditions.	14
4	Indexability of strongly thresholdable projects.	16
5	LP reformulation of the λ-price problem starting from $X_0 \sim \nu_0$.	17
6	Decomposition of performance metrics.	18
7	Analysis of metrics as functions of the threshold.	20
7.1	Càdlàg property.	20
7.2	Monotonicity of resource metric.	22
7.3	Analysis of discontinuities.	22
7.4	Bounded variation and MP index as Radon–Nikodym derivative.	24
7.5	Relations between marginal project metrics.	25
8	Further relations between MP metrics.	26
9	Proving Theorem 1.	27
10	How to establish condition (PCLI3).	28
10.1	Piecewise differentiable $F(x, \cdot)$ and $G(x, \cdot)$	28
10.2	Piecewise constant $G(x, \cdot)$	29
11	Metrics and index computation.	30
12	Examples.	31
12.1	Optimal web crawling.	32
12.2	Optimal dynamic transmission over a noisy channel.	33
13	Concluding remarks.	35

1 Introduction.

1.1 RMABP and Whittle’s index policy.

In the *restless multiarmed bandit problem* (RMABP), introduced by Whittle (1988), a decision maker aims to maximize the expected reward earned from a finite set of projects. These are modeled as *restless bandits* (throughout, the terms *project* and *bandit* are used interchangeably), i.e., infinite-horizon *Markov decision processes* (MDPs) —see Puterman (1994)— that can be operated in two modes, *active* and *passive*. In the RMABP, a fixed number of projects are selected to be active at each time. Note that projects are *restless* in that they can change state while passive.

The RMABP extends the classic *MABP*, where projects are *nonrestless* and one project is engaged at each time, which is a rare example of a multidimensional MDP for which a tractable policy is available. Thus, one can define for each project a scalar function of its state, the *Gittins index*, so that the resulting *priority-index policy*, which engages at each time a project of largest index, is optimal. See, e.g., Gittins (1979), Whittle (1980), and Katehakis and Veinott (1987). Cowan and Katehakis (2015) have extended the scope of such a landmark result to a broader MABP setting, which allows real-state projects and non-uniform discounting, among other features. In contrast, index policies are generally suboptimal for the RMABP, as this is known to be PSPACE-hard. See Papadimitriou and Tsitsiklis (1999).

However, Whittle (1988) extended in part the index-based solution approach to the RMABP. He defined an index for restless bandits —the *Whittle index*— and proposed as a heuristic the resulting priority-index policy, which engages at each time the required number of projects with larger indices. He conjectured the asymptotic optimality of such a policy as both the number of active projects and their total number grow to infinity in fixed ratio, which has been proved by Weber and Weiss (1990), Ouyang et al. (2016) and Verloop (2016) under certain conditions.

Being an intuitive practical heuristic, researchers have deployed Whittle’s policy over the last two decades in diverse applications, providing a growing body of evidence that its performance is often nearly optimal. Examples include, e.g., machine maintenance (Whittle (1996, Ch. 14.6)), scheduling multiclass queues (Ansell et al. (2003), Veatch and Wein (1996), Niño-Mora (2006)), admission and routing to queues (Niño-Mora (2002, 2012, 2019)), projects with response delays (Caro and Yoo (2010)), dynamic spectrum access (Liu and Zhao (2010), Ouyang et al. (2015)), demand response (Taylor and Mathieu (2014)), and web crawling (Avrachenkov and Borkar (2018)).

1.2 Indexability.

Yet, to apply Whittle’s index policy to a given RMABP model researchers need to overcome two roadblocks: (1) establish *indexability* (existence of the index) for the projects in the model at hand; and (2) devise an efficient index-computing scheme.

Indeed, unlike the Gittins index, which is defined for any classic bandit, the Whittle index is defined only for a limited class of restless bandits, called *indexable*. Whittle (1988, p. 292) pointed out that “*Indexability cannot be taken for granted ... One would very much like to have simple sufficient conditions for indexability; at the moment, none are known.*” Liu and Zhao (2010, p. 5547) further noted that “*indexability ... is often difficult to establish ... and computing Whittle’s index can be complex.*”

We next outline the concepts of indexability and Whittle index as considered herein, which apply to a *single* project. Consider a project evolving in discrete time that consumes a generic resource. At the start of each time period, the controller observes the project state and then chooses an action (active or passive). Then the project (i) consumes an amount of resource and yields a reward, discounted over time; and (ii) moves to the next state according to a Markov transition law.

Action choice is governed by a policy π from the class Π of *nonanticipative policies*. Suppose that the resource is charged at a unit price λ , and let $V_\lambda(x, \pi)$ be the resulting net project value starting from state x under π , giving the expected total discounted value of rewards earned minus resource charges incurred. Now, consider the parametric collection of λ -price problems $\{P_\lambda: \lambda \in \mathbb{R}\}$, with

$$P_\lambda: \quad \text{find } \pi_\lambda^* \in \Pi \text{ such that } V_\lambda(x, \pi_\lambda^*) = V_\lambda^*(x) \triangleq \sup_{\pi \in \Pi} V_\lambda(x, \pi) \text{ for every state } x. \quad (1)$$

Thus, a P_λ -optimal policy achieves an optimal tradeoff between rewards and resource usage costs.

We shall call the project *indexable* if there exists a scalar map $\lambda^*(x)$ on X (the project's *Whittle index*) characterizing P_λ -optimal policies for any price λ , as follows: when the project is in state x , it is P_λ -optimal to take the active (resp. passive) action if and only if $\lambda^*(x) \geq \lambda$ (resp. $\lambda^*(x) \leq \lambda$).

1.3 Real-state projects: prevailing approaches to indexability based on optimal thresholds.

While most work on the RMABP has considered discrete-state projects, models with real-state projects have drawn growing attention since the last decade, motivated by *partially observed MDP* (POMDP) formulations of sensor scheduling applications. See, e.g., [Krishnamurthy \(2016, Ch. 8\)](#) and the surveys [Washburn \(2008\)](#) and [Kuhn and Nazarathy \(2017\)](#).

In such applications, a set of sensors are to be dynamically allocated to sense a larger set of objects, e.g., moving targets (in multitarget tracking, where the sensors are radars) or communication channels (in multichannel access, where sensors probe channel transmission availability). Due to the limited number of sensors and, possibly, their providing noisy measurements, true object states are unknown at decision times. Instead, under a POMDP model sensor allocation decisions are based on the *belief state* of each object, which gives the posterior state distribution and is updated by a Bayesian filter depending on whether the object is sensed. Belief states can take a continuum of values, being scalar in certain models. Thus, identifying objects with projects, the active and passive actions with sensing and not sensing an object, and using belief states as project states, when the latter are scalar we obtain special cases of the RMABP with real-state projects.

[La Scala and Moran \(2006\)](#) first formulated the problem of dynamic radar scheduling to minimize the track error variance of two moving targets as an RMABP, where projects follow deterministic Kalman filter dynamics. They proved optimality of the greedy policy in the symmetric case.

More recent work has addressed indexability in POMDP models with real-state projects, including, e.g., multisite tracking ([Le Ny et al. \(2008\)](#)), dynamic multichannel access ([Liu and Zhao \(2010\)](#) and [Ouyang et al. \(2015\)](#)), continuous-time ([Le Ny et al. \(2011\)](#)) and discrete-time ([Dance and Silander \(2015\)](#)) multitarget tracking, demand response ([Taylor and Mathieu \(2014\)](#)), web crawling ([Avrachenkov and Borkar \(2018\)](#)), and hidden Markov bandits ([Meshram et al. \(2018\)](#)).

Such work follows a common pattern, based on elucidating and exploiting properties of the optimal value function $V_\lambda^*(x)$ of λ -price problems (1) for the model at hand through *dynamic programming* (DP) arguments, to prove optimality of *threshold policies*. Such policies, also known as *control limit policies*, make the project active when its state is (at or) above a given threshold, and make it passive otherwise. They are widely used in applications due to their simplicity and ease of implementation.

A typical analysis proceeds in two stages. The first is to identify and exploit properties of $V_\lambda^*(x)$ to prove existence of an optimal threshold policy for each λ -price problem, with the optimal threshold $z^*(\lambda)$ being unique for a certain interval of λ 's. In the second stage, it is claimed that to establish indexability it suffices to show that $z^*(\lambda)$ is *increasing* (or, in some papers, *nondecreasing*) in λ . The Whittle index $\lambda^*(x)$ is then obtained by inverting $z^*(\lambda)$, i.e., solving the equation $z^*(\lambda) = x$ in λ .

Yet such analyses, while insightful and useful, are often incomplete, due to: (1) as proving optimality of threshold policies may be elusive, it is sometimes only conjectured (as in [Le Ny et al. \(2011\)](#) and [Dance and Silander \(2015\)](#)), or the arguments provided are sketchy and not fully rigorous; and (2) showing *increasingness* (in the strict sense) of $z^*(\lambda)$ does not suffice to establish indexability; *continuity* of $z^*(\lambda)$ needs also be proven (see Proposition 3).

Indeed, proving optimality of threshold policies can be far from trivial. It can be formulated in terms of $V_\lambda^{(a,*)}(x)$, the optimal value with initial action a . Thus, it needs to be shown that, for λ in a certain interval, the equation $\Delta_{a=0}^{a=1} V_\lambda^{(a,*)}(x) \triangleq V_\lambda^{(1,*)}(x) - V_\lambda^{(0,*)}(x) = 0$ in state variable x has a unique root $z^*(\lambda)$, with $\Delta_{a=0}^{a=1} V_\lambda^{(a,*)}(x) < 0$ for $x < z^*(\lambda)$ and $\Delta_{a=0}^{a=1} V_\lambda^{(a,*)}(x) > 0$ for $x > z^*(\lambda)$.

Such a result follows if the difference $V_\lambda^{(1,*)}(x) - V_\lambda^{(0,*)}(x)$ is concave in x and changes sign at the endpoints of the state space, as in [Le Ny et al. \(2008\)](#) and [Liu and Zhao \(2010\)](#), where $V_\lambda^{(1,*)}(x)$ is affine and $V_\lambda^{(0,*)}(x)$ convex. Yet, in the model extension in [Ouyang et al. \(2015\)](#), while the $V_\lambda^{(a,*)}(x)$ are convex in x , it is no longer

ensured that $\Delta_{a=0}^{a=1}V_{\lambda}^{(a,*)}(x)$ is concave, as $V_{\lambda}^{(1,*)}(x)$ may not be affine. Hence, existence of a $z^*(\lambda)$ as above cannot be concluded from $\Delta_{a=0}^{a=1}V_{\lambda}^{(a,*)}(x)$ changing sign.

The most widely used approach for proving optimality of threshold policies draws on the theory of *supermodularity*. See, e.g., [Topkis \(1998\)](#), [Heyman and Sobel \(1984, Ch. 8.2\)](#), [Puterman \(1994, Ch. 6.11.2\)](#), [Altman and Stidham \(1995\)](#), and [Krishnamurthy \(2016, Ch. 9\)](#). Thus, if $V_{\lambda}^{(a,*)}(x)$ is *strictly supermodular* in (x, a) , so $\Delta_{a=0}^{a=1}V_{\lambda}^{(a,*)}(x)$ is increasing in x , being negative for small x and positive for large x , existence of a unique threshold $z^*(\lambda)$ satisfying the required properties follows.

Yet, in the setting of a Kalman filter target tracking model, [Dance and Silander \(2019\)](#) identify an instance demonstrating that efforts to prove optimality of threshold policies in such a model through the above approaches are doomed to failure. Using the notation in Section 2, their instance has rewards $r(x, a) = -x$, passive dynamics $X_{t+1} = X_t + 1$, active dynamics $X_{t+1} = 1/(\alpha + 1/(X_t + 1))$ with $\alpha = 0.1$, and discount factor $\beta = 0.95$. Figure 1 in their paper shows that, for $\lambda = 0.7647$, $\Delta_{a=0}^{a=1}V_{\lambda}^{(a,*)}(x)$ is neither concave nor monotonic over $x \in [0, 0.25]$. The author has independently verified such a result, which highlights a limitation of the scope of prevailing approaches to optimality of threshold policies and indexability, raising the need for different approaches with a wider scope.

1.4 PCL-based sufficient indexability conditions.

An alternative to prevailing approaches to indexability for discrete-state restless bandits was introduced and developed in [Niño-Mora \(2001, 2002, 2006, 2007\)](#) based on two sufficient conditions, which can be verified analytically or numerically, along with an index algorithm. The validity of such conditions stems from satisfaction by project metrics of *partial conservation laws* (PCLs), whence we shall refer to *PCL-indexability conditions*. PCLs extend earlier types of conservation laws in stochastic scheduling, which have been used to prove optimality of index policies through achievable performance region analyses. See, e.g., [Coffman and Mitrani \(1980\)](#), [Shanthikumar and Yao \(1992\)](#), [?](#), and [Niño-Mora \(2011\)](#). Yet, the proof techniques used in the aforementioned work rely critically on the discreteness of the project state space, so their extension to real-state projects is nontrivial.

[Niño-Mora \(2008\)](#) first outlined, without proving its validity, an extension to real-state projects of such PCL-indexability conditions, which roughly correspond to (PCLI1) and (PCLI2) in this paper (see Definition 7). Yet condition (PCLI3) herein was not considered, nor was the continuity requirement in (PCLI2). That paper further illustrated application of the conditions in the model in [Le Ny et al. \(2008\)](#) and [Liu and Zhao \(2010\)](#), sketching an incomplete analysis.

Later work exploring use of such an approach includes [Niño-Mora \(2009, 2016\)](#) and the PhD thesis of [Villar \(2012\)](#). The latter, supervised by the author, deepened the study of the models in [Niño-Mora and Villar \(2009, 2011\)](#), including the multitarget Kalman filter tracking model in [La Scala and Moran \(2006\)](#). While such work did not establish validity of the conditions, it presented partial analyses as well as numerical evidence supporting the practical value of the proposed approach: the Whittle index could be efficiently computed and the index policy was found to be nearly optimal.

In the first arXiv version of this paper, [Niño-Mora \(2015\)](#) modified the PCL-indexability conditions in [Niño-Mora \(2008\)](#) by adding a continuity requirement in (PCLI2), as well as the new condition (PCLI3) in Definition 7. [Niño-Mora \(2015\)](#) gave a proof that conditions (PCLI1–PCLI3) imply indexability *and* optimality of threshold policies for the λ -price problems. The proof relied on analysis of an infinite-dimensional *linear programming* (LP) formulation of such problems.

Based on [Niño-Mora \(2015\)](#), [Dance and Silander \(2019\)](#) have applied the PCL-indexability conditions herein to establish for the first time indexability, along with optimality of threshold policies, of an extension of the aforementioned Kalman filter target tracking model. To prove (PCLI1–PCLI3) they deploy a novel analysis of action trajectories under threshold policies drawing on the theory of combinatorics on words. In a preliminary arXiv version of such work, further outline a short proof of the validity of such conditions for their model, which motivated the author to develop the simpler proof presented here for the general case that avoids the infinite-dimensional LP analysis in [Niño-Mora \(2015\)](#).

1.5 Goal and main results.

The major goal of this paper is to develop a rigorous PCL-based approach to establish indexability and evaluate the Whittle index of real-state discounted restless bandits in a form consistent with optimality of threshold policies for their λ -price problems.

The main result is a *verification theorem* (Theorem 1) which, under three *PCL-indexability conditions*, ensures in one fell swoop *both* indexability of a real-state restless bandit *and* optimality of threshold policies for its λ -price problems (1). This differs from prevailing approaches (see Section 1.3), where first optimality of threshold policies is proved, and then indexability is established.

Under the proposed conditions, the Whittle index is given by a *marginal productivity* (MP) index that is explicitly defined in terms of project metrics under threshold policies. This is in contrast both to the Whittle index definition, which characterizes it implicitly, and to the way in which the index is determined in prevailing approaches, by inverting an optimal threshold function.

The conditions are formulated as properties of *project metrics* under threshold policies. Thus, we consider the *reward* and *resource (usage)* metrics denoted respectively by $F(x, z)$ and $G(x, z)$, with x the initial state and z shorthand for the z -policy (see Section 2.4). We also consider *marginal reward* and *marginal resource* metrics, denoted respectively by $f(x, z) \triangleq F(x, \langle 1, z \rangle) - F(x, \langle 0, z \rangle)$ and $g(x, z) \triangleq G(x, \langle 1, z \rangle) - G(x, \langle 0, z \rangle)$, where policy $\langle a, z \rangle$ first takes action a and then follows the z -policy. We further consider the *marginal productivity* (MP) metric $m(x, z) \triangleq f(x, z)/g(x, z)$, which is defined when $g(x, z) \neq 0$, and the *MP index*, given by $m(x) \triangleq m(x, x)$ for every state x .

The *PCL-indexability conditions* are: (PCLI1) the marginal resource metric $g(x, z)$ is positive for any state x and threshold z ; (PCLI2) the MP index $m(x)$ is monotone nondecreasing and continuous; and (PCLI3) the reward and resource metrics and the MP index are linked by

$$F(x, z_2) - F(x, z_1) = \int_{(z_1, z_2]} m(z) G(x, dz), \quad -\infty < z_1 < z_2 < \infty, \quad (2)$$

so, for each x , and viewed as functions of the threshold z , $F(x, z)$ is an indefinite *Lebesgue–Stieltjes* (LS) *integral* (see, e.g., Carter and van Brunt (2000)) of the MP index $m(z)$ with respect to $G(x, z)$. Proposition 4 shows that (PCLI3) characterizes the index as a *Radon–Nikodym derivative*.

Theorem 1 ensures that a project satisfying conditions (PCLI1–PCLI3) is indexable with $m(x)$ being its Whittle index. Furthermore, as the latter is nondecreasing and continuous, the optimality of threshold policies for the project’s λ -price problems (1) follows. See Lemma 2.

This result reduces the task of establishing a project’s indexability to that of proving certain properties of project metrics *under threshold policies*, and provides an explicit expression for the index. Such an approach is in contrast to prevailing approaches, based on exploiting properties of the optimal value function, which require establishing beforehand optimality of threshold policies.

Lemma 2 and Proposition 2 show that conditions (PCLI2, PCLI3) are necessary for *threshold-indexability* (see Definition 4), and that the Whittle index matches the MP index under (PCLI1). See Proposition 1(c). Proposition 7 gives an index-computing recursion.

Note that the PCL-indexability conditions for discrete-state projects referred to in Section 1.4 roughly correspond to (PCLI1, PCLI2), as in such a setting they imply (PCLI3). In the present real-state setting, since (PCLI3) may be hard to verify, we give Propositions 5 and 6, which ensure that (PCLI3) holds under (PCLI1, PCLI2) and some additional simpler conditions.

1.6 Organization of the paper.

The remainder of the paper is organized as follows. Section 2 sets up notation and presents concepts to formulate the main result in Section 2.6. Section 3 motivates the PCL-indexability conditions. Section 4 presents results for the indexability analysis of *strongly thresholdable* projects (see Definition 3), which are relevant for complementing the approaches reviewed in Section 1.3. Section 5 reformulates the λ -price problem P_λ as an LP problem over a space of measures. Section 6 presents a key result on decomposition of project metrics. Section 7 establishes required properties of project metrics as functions of the threshold, including that they are *càdlàg* (“continue à droite, limite à gauche,” i.e., right-continuous with left limits), and characterizes

the MP index as a Radon–Nikodym derivative. Section 8 derives further required relations between project metrics. Section 9 gives the proof of Theorem 1. Section 10 presents tools to establish that condition (PCLI3) holds. Section 11 addresses index and metrics computation. Section 12 illustrates application of the proposed approach in two models. Section 13 concludes.

An online companion containing two appendices provides relevant ancillary material. Section A motivates the interest of indexability through a general resource-constrained RMABP. Section B presents geometric and economic interpretations of the MP index under PCL-indexability, showing in particular that it can be characterized as a *resource shadow price*.

2 Preliminaries and formulation of the main result.

We next set up notation, describe the single-project model, and present required concepts to formulate the main result.

2.1 Notation.

We denote by $\mathsf{X} \triangleq \{x \in \mathbb{R} : \ell \leq x \leq u\}$ the project state space, which is a closed (possibly unbounded) interval of \mathbb{R} with endpoints $-\infty \leq \ell < u \leq \infty$, and by $\mathsf{K} \triangleq \mathsf{X} \times \{0, 1\}$ its state-action space. For $S \in \mathcal{B}(\mathsf{X})$, the Borel σ -algebra of X , we write the complement of S in X as $S^c \triangleq \mathsf{X} \setminus S$, and denote by $1_S(\cdot)$ its indicator function. All subsets of \mathbb{R} and functions referred to are assumed to be Borel measurable. We write functions as, e.g., $J(\cdot)$, and denote by $|J|(\cdot)$ the corresponding absolute value, so $|J|(x) = |J(x)|$. We denote by $\mathbb{C}(\mathsf{X})$ the space of real-valued continuous functions on X .

Throughout, x and y denote project states, and we use $z \in \overline{\mathbb{R}} \triangleq \mathbb{R} \cup \{-\infty, \infty\}$ to denote thresholds. For a given threshold z , we use shorthands such as, e.g., $\{x > z\}$ to mean $\{x \in \mathsf{X} : x > z\}$.

We write as $\int_B J d\alpha$ or $\int_B J(x) \alpha(dx)$ the LS integral (see Carter and van Brunt (2000)) of a function $J(\cdot)$ with respect to a function $\alpha(\cdot)$ over B . When B is the domain of $J(\cdot)$ we write $\int J d\alpha$.

We use the terms “increasing” and “decreasing” in the strict sense, and “iff” to mean “if and only if.”

2.2 Project model, performance metrics, and λ -price problems.

We consider a general restless bandit model of optimal dynamic resource allocation to a project, whose state X_t moves in discrete time over an infinite horizon across the state space X .

At the start of period $t = 0, 1, \dots$, a controller observes the project state X_t and then chooses *action* $A_t \in \{0, 1\}$, where $A_t = 1$ (resp. $A_t = 0$) codes the project being *active* (resp. *passive*). When $(X_t, A_t) = (x, a)$: (i) the project consumes $c(x, a)$ units of a generic resource and yields a reward $r(x, a)$, discounted with factor $\beta \in [0, 1]$; and (ii) its state moves to X_{t+1} according to a Markov transition kernel $\kappa^a(\cdot, \cdot)$, so $\mathbb{P}\{X_{t+1} \in B \mid (X_t, A_t) = (x, a)\} = \kappa^a(x, B)$ for $B \in \mathcal{B}(\mathsf{X})$.

We make the following standing assumption.

Assumption 1. For each action $a \in \{0, 1\}$:

- (i) $r(\cdot, a), c(\cdot, a) \in \mathbb{C}(\mathsf{X})$, with $c(\cdot, a)$ satisfying $0 \leq c(x, 0) < c(x, 1)$ for $x \in \mathsf{X}$;
- (ii) there is a function $w : \mathsf{X} \rightarrow [1, \infty)$ and $M > 0, \beta \leq \gamma < 1$ such that, for any state x ,
 - (ii.a) $\max\{|r|(x, a), c(x, a)\} \leq M w(x)$; and
 - (ii.b) $\beta \tilde{w}(x, a) \leq \gamma w(x)$, where $\tilde{w}(x, a) \triangleq \int w(y) \kappa^a(x, dy)$;

Remark 1. (i) Assumption 1(ii) corresponds to the *weighted sup-norm* framework for discounted MDPs (see, e.g., Lippman (1975); Wessels (1977) and Hernández-Lerma and Lasserre (1999, assumption 8.3.2)), which allows functions $r(\cdot, a)$ and $c(\cdot, a)$ to be unbounded.

- (ii) If $\tilde{w}(x, a) \leq w(x)$, Assumption 1(ii.b) holds for any β by taking $\gamma \triangleq \beta$.

Actions are selected through a *policy* $\pi = (\pi_t)_{t=0}^\infty$ from the class Π of *nonanticipative policies*, which at time t selects action $A_t = a$ with probability $\pi_t(a | \mathcal{H}_t)$, a measurable function of the history $\mathcal{H}_t \triangleq ((X_s, A_s)_{s=0}^{t-1}; X_t)$. We shall further refer to the class Π^{SR} of *stationary randomized policies*, where $\pi_t(a | \mathcal{H}_t) = q(a | X_t)$ for some $q(\cdot | \cdot) \geq 0$ with $\sum_a q(a | x) = 1$, and to the class Π^{SD} of *stationary deterministic policies*, where $A_t = 1_B(X_t)$ for some *active region* $B \in \mathcal{B}(X)$, in which case we shall refer to the *B-policy*, writing $\pi = B$.

For a policy $\pi \in \Pi$ and an initial-state probability measure ν_0 on $\mathcal{B}(X)$ (written $X_0 \sim \nu_0$), let $\mathbb{P}_{\nu_0}^\pi$ be the probability measure determined (by Ionescu-Tulcea's extension theorem; see [Hernández-Lerma and Lasserre \(1996, proposition C.10\)](#)) on the space K^∞ of state-action paths with the product σ -algebra, with expectation $\mathbb{E}_{\nu_0}^\pi$. We write \mathbb{P}_x^π and \mathbb{E}_x^π when $\nu_0 = \delta_x$, the Dirac measure at x .

We evaluate a policy π starting from an initial state drawn from ν_0 by the *reward* and *resource (usage) performance metrics*

$$F(\nu_0, \pi) \triangleq \mathbb{E}_{\nu_0}^\pi \left[\sum_{t=0}^{\infty} \beta^t r(X_t, A_t) \right] \quad \text{and} \quad G(\nu_0, \pi) \triangleq \mathbb{E}_{\nu_0}^\pi \left[\sum_{t=0}^{\infty} \beta^t c(X_t, A_t) \right], \quad (3)$$

which we write as $F(x, \pi)$ and $G(x, \pi)$ when $\nu_0 = \delta_x$, and as $F(\nu_0, B)$ and $G(\nu_0, B)$ when $\pi = B$.

If resource usage is charged at a price λ per unit, denote by $V_\lambda(\nu_0, \pi) \triangleq F(\nu_0, \pi) - \lambda G(\nu_0, \pi)$ the (*net*) *project value*, and by $V_\lambda^*(\nu_0) \triangleq \sup_{\pi \in \Pi} V_\lambda(\nu_0, \pi)$ the *optimal project value*.

For any resource price λ consider the λ -*price problem*

$$P_\lambda: \quad \text{find } \pi_\lambda^* \in \Pi \text{ such that } V_\lambda(x, \pi_\lambda^*) = V_\lambda^*(x) \text{ for every } x \in X. \quad (4)$$

We call a policy π_λ^* solving (4) P_λ -*optimal*. Consider the parametric problem collection

$$\mathcal{P} \triangleq \{P_\lambda: \lambda \in \mathbb{R}\}. \quad (5)$$

We next review results from the weighted sup-norm framework that are needed in the sequel. Let $\mathbb{B}_w(K)$ and $\mathbb{B}_w(X)$ be the Banach spaces (see [Hernández-Lerma and Lasserre \(1999, proposition 7.2.1\)](#)) of w -*bounded* functions $u: K \rightarrow \mathbb{R}$ and $v: X \rightarrow \mathbb{R}$, respectively, having finite w -*norms*

$$\|u\|_w \triangleq \sup_{(x,a) \in K} \frac{|u|(x,a)}{w(x)} \quad \text{and} \quad \|v\|_w \triangleq \sup_{x \in X} \frac{|v|(x)}{w(x)}, \quad (6)$$

and write as $\mathbb{P}_w(X)$ the space of probability measures ν_0 on $\mathcal{B}(X)$ with finite w -norm

$$\|\nu_0\|_w \triangleq \int w d\nu_0 = \mathbb{E}_{\nu_0}[w(X_0)]. \quad (7)$$

Let further $\mathbb{C}_w(K) \triangleq \mathbb{C}(K) \cap \mathbb{B}_w(K)$ and $\mathbb{C}_w(X) \triangleq \mathbb{C}(X) \cap \mathbb{B}_w(X)$ denote the corresponding subspaces of w -bounded continuous functions on K and X .

Henceforth, $F(\cdot, \pi)$, $G(\cdot, \pi)$ and $V_\lambda(\cdot, \pi)$ denote the reward, resource and net value project metrics viewed as functions of the initial state x for a fixed policy π .

Remark 2. Under Assumption 1 (cf. [Hernández-Lerma and Lasserre \(1999, theorem 8.3.6\)](#)):

(i) For any action a , $r(\cdot, a), c(\cdot, a) \in \mathbb{C}_w(X)$; $r(\cdot, \cdot), c(\cdot, \cdot) \in \mathbb{C}_w(K)$; and $\kappa^a(x, \cdot) \in \mathbb{P}_w(X)$.

(ii) For any $\pi \in \Pi$ and $\lambda \in \mathbb{R}$, $F(\cdot, \pi), G(\cdot, \pi), V_\lambda(\cdot, \pi) \in \mathbb{B}_w(X)$; and, with $M_\gamma \triangleq M/(1 - \gamma)$,

$$\max\{\|F(\cdot, \pi)\|_w, \|G(\cdot, \pi)\|_w\} \leq M_\gamma \quad \text{and} \quad \|V_\lambda(\cdot, \pi)\|_w \leq (1 + |\lambda|)M_\gamma. \quad (8)$$

(iii) For $\pi \in \Pi$, $\nu_0 \in \mathbb{P}_w(X)$ and $\lambda \in \mathbb{R}$, $F(\nu_0, \pi)$, $G(\nu_0, \pi)$ and $V_\lambda(\nu_0, \pi)$ are well defined and finite, being equal to $\int F(x, \pi) \nu_0(dx)$, $\int G(x, \pi) \nu_0(dx)$ and $\int V_\lambda(x, \pi) \nu_0(dx)$, respectively, and satisfy

$$\max\{|F|(\nu_0, \pi), |G|(\nu_0, \pi)\} \leq M_\gamma \|\nu_0\|_w \quad \text{and} \quad |V_\lambda|(\nu_0, \pi) \leq (1 + |\lambda|)M_\gamma \|\nu_0\|_w. \quad (9)$$

- (iv) For any active region $B \in \mathcal{B}(\mathsf{X})$, $F(\cdot, B)$ and $G(\cdot, B)$ are characterized (cf. [Hernández-Lerma and Lasserre \(1999, remark 8.3.10\)](#)) as the unique solutions in $\mathbb{B}_w(\mathsf{X})$ to the fixed-point equations

$$\begin{aligned} F(x, B) &= r(x, 1_B(x)) + \beta \int F(y, B) \kappa^{1_B(x)}(x, dy), \quad x \in \mathsf{X}, \\ G(x, B) &= c(x, 1_B(x)) + \beta \int G(y, B) \kappa^{1_B(x)}(x, dy), \quad x \in \mathsf{X}. \end{aligned} \tag{10}$$

Later (see [Section 6](#)) we will find it convenient to use the reformulation of [\(10\)](#) in terms of the bounded linear operator $\mathcal{L}^*: \mathbb{B}_w(\mathsf{X}) \rightarrow \mathbb{B}_w(\mathsf{K})$ mapping $v(\cdot)$ to

$$\mathcal{L}^*v(x, a) \triangleq v(x) - \beta \int v(y) \kappa^a(x, dy), \quad (x, a) \in \mathsf{K}, \tag{11}$$

namely,

$$\begin{aligned} \mathcal{L}^*F(\cdot, B)(x, 1_B(x)) &= r(x, 1_B(x)), \quad x \in \mathsf{X}, \\ \mathcal{L}^*G(\cdot, B)(x, 1_B(x)) &= c(x, 1_B(x)), \quad x \in \mathsf{X}. \end{aligned} \tag{12}$$

- (v) For any price $\lambda \in \mathbb{R}$, the *Bellman operator* $\mathcal{T}_\lambda: \mathbb{B}_w(\mathsf{X}) \rightarrow \mathbb{B}_w(\mathsf{X})$ defined by

$$\mathcal{T}_\lambda v(x) \triangleq \max_{a \in \{0,1\}} \mathcal{T}_\lambda^a v(x), \quad \text{where } \mathcal{T}_\lambda^a v(x) \triangleq r(x, a) - \lambda c(x, a) + \beta \int v(y) \kappa^a(x, dy), \tag{13}$$

is a *contraction mapping* with modulus γ (see [Assumption 1\(ii\)](#)) and the *Bellman equation (BE)*

$$v = \mathcal{T}_\lambda v, \tag{14}$$

has the *optimal value function* $V_\lambda^*(\cdot)$ as its unique fixed point in $\mathbb{B}_w(\mathsf{X})$, which satisfies

$$\|V_\lambda^*\|_w \leq (1 + |\lambda|)M_\gamma. \tag{15}$$

- (vi) For any price $\lambda \in \mathbb{R}$ there is a P_λ -optimal stationary deterministic policy; a policy in Π^{SD} is P_λ -optimal iff it selects in each state x an action a with $\mathcal{T}_\lambda^a V_\lambda^*(x) \geq \mathcal{T}_\lambda^{1-a} V_\lambda^*(x)$.

We denote the optimal net project value starting from $X_0 = x$ with initial action $A_0 = a$ by

$$V_\lambda^{(a,*)}(x) \triangleq \mathcal{T}_\lambda^a V_\lambda^*(x) = r(x, a) - \lambda c(x, a) + \beta \int V_\lambda^*(y) \kappa^a(x, dy). \tag{16}$$

We say that *action a is P_λ -optimal in state x* if $V_\lambda^{(a,*)}(x) \geq V_\lambda^{(1-a,*)}(x)$. It is convenient to reformulate such a definition in terms of the sign of $\Delta_{a=0}^{a=1} V_\lambda^{(a,*)}(x) \triangleq V_\lambda^{(1,*)}(x) - V_\lambda^{(0,*)}(x)$, the *marginal value of engaging the project in state x* . Thus, action $a = 1$ is P_λ -optimal in x if $\Delta_{a=0}^{a=1} V_\lambda^{(a,*)}(x) \geq 0$; action $a = 0$ is P_λ -optimal in x if $\Delta_{a=0}^{a=1} V_\lambda^{(a,*)}(x) \leq 0$; and both actions are P_λ -optimal in x if

$$\Delta_{a=0}^{a=1} V_\lambda^{(a,*)}(x) = 0. \tag{17}$$

2.3 Indexability.

We will address the parametric problem collection \mathcal{P} in [\(4\)](#) through the concept of *indexability*, extended by [Whittle \(1988\)](#) from classic to restless bandits with resource consumption $c(x, a) \triangleq a$, and further extended by [Niño-Mora \(2002\)](#) to general $c(x, a)$.

For a given resource price λ , the P_λ -optimal *active* and *passive regions* are given by

$$S_\lambda^{*,1} \triangleq \{x \in \mathsf{X} : \Delta_{a=0}^{a=1} V_\lambda^{(a,*)}(x) \geq 0\} \quad \text{and} \quad S_\lambda^{*,0} \triangleq \{x \in \mathsf{X} : \Delta_{a=0}^{a=1} V_\lambda^{(a,*)}(x) \leq 0\}, \tag{18}$$

respectively. Under indexability, such regions are characterized by an *index* attached to project states. Note that the definition below refers to the *sign function* $\text{sgn}: \mathbb{R} \rightarrow \{-1, 0, 1\}$.

Definition 1 (Indexability). The project is *indexable* if there is a map $\lambda^*: \mathsf{X} \rightarrow \mathbb{R}$ with

$$\operatorname{sgn} \Delta_{a=0}^{a=1} V_{\lambda}^{(a,*)}(x) = \operatorname{sgn}(\lambda^*(x) - \lambda), \quad \text{for } x \in \mathsf{X} \text{ and } \lambda \in \mathbb{R}. \quad (19)$$

We call $\lambda^*(\cdot)$ the project's *Whittle index* (*Gittins index* if the project is nonrestless).

Remark 3.

- (i) [Whittle \(1988\)](#) defined indexability when $c(x, a) \triangleq a$ through an equivalent problem collection $\mathscr{P}' \triangleq \{P'_{\lambda} : \lambda \in \mathbb{R}\}$ differing from \mathscr{P} (see (4)) in the net rewards, which are $r(x, a) + \lambda(1 - a)$ for P'_{λ} (so λ represents a *subsidy for passivity*) and $r(x, a) - \lambda a$ for P_{λ} (so λ is a *price of activity*). Yet, their equivalence is not ensured for projects with general $c(x, a)$, where net rewards for P'_{λ} and P_{λ} would be $r(x, a) + \lambda c(x, 1 - a)$ and $r(x, a) - \lambda c(x, a)$, respectively. In such a general case, [Niño-Mora \(2002\)](#) introduced a definition of indexability through the problem collection \mathscr{P} as herein. The resulting index is a *generalized Whittle index*, or a *generalized Gittins index* if projects are nonrestless.
- (ii) [Whittle \(1988, p. 291\)](#) defined indexability as a property of the family of P'_{λ} -optimal passive sets $D(\lambda)$, in that $D(\lambda)$ “increases monotonically from \emptyset to X as ν (his notation for our λ) increases from $-\infty$ to $+\infty$.” He defined the index as the root (assumed to be unique) in λ of (17).
- (iii) The project is indexable with index $\lambda^*(\cdot)$ if the P_{λ} -optimal action in each state x is nonincreasing in λ with both actions being P_{λ} -optimal only for $\lambda = \lambda^*(x)$; thus, for $\lambda < \lambda^*(x)$ (resp. $\lambda > \lambda^*(x)$), $a = 1$ (resp. $a = 0$) is the only P_{λ} -optimal action (cf. [Niño-Mora \(2014, section 2.3\)](#)); equivalently, for each x , equation (17) in λ has the unique root $\lambda^*(x)$, and $\Delta_{a=0}^{a=1} V_{\lambda}^{(a,*)}(x) > 0$ for $\lambda < \lambda^*(x)$ and $\Delta_{a=0}^{a=1} V_{\lambda}^{(a,*)}(x) < 0$ for $\lambda > \lambda^*(x)$.

2.4 Thresholdability and threshold-indexability: weak and strong versions.

We address indexability in the case that it holds along with optimality of threshold policies. For any threshold z we define the z -policy, with active and passive regions $S^1(z) \triangleq \{x \in \mathsf{X} : x > z\}$ and $S^0(z) \triangleq \{x \in \mathsf{X} : x \leq z\}$; and the z^- -policy, with active and passive regions $S^1(z^-) \triangleq \{x \in \mathsf{X} : x \geq z\}$ and $S^0(z^-) \triangleq \{x \in \mathsf{X} : x < z\}$. We write their metrics as $F(\nu_0, z)$, $G(\nu_0, z)$, $F(\nu_0, z^-)$ and $G(\nu_0, z^-)$.

The notation $F(\nu_0, \cdot)$ and $G(\nu_0, \cdot)$ refers to metrics $F(\nu_0, z)$ and $G(\nu_0, z)$ as functions of the threshold z on \mathbb{R} . The ambiguity of the notation $F(\nu_0, z^-)$ and $G(\nu_0, z^-)$ is resolved in Lemma 10(a), ensuring that the latter metrics are also the left limits at z of $F(\nu_0, \cdot)$ and $G(\nu_0, \cdot)$, respectively.

To formalize the idea of optimality of threshold policies, we distinguish between weak and strong versions of such a concept. The following weak version corresponds to the case that P_{λ} -optimal actions are partially characterized by a threshold. Note that the term *thresholdability* has been used in a similar sense in the literature. See, e.g., [Ouyang et al. \(2015\)](#).

Definition 2 (Thresholdability). We call the project *thresholdable* if there is a map $z^*: \mathbb{R} \rightarrow \overline{\mathbb{R}}$ such that, for every project state $x \in \mathsf{X}$ and price $\lambda \in \mathbb{R}$:

$$\begin{aligned} \text{if } x \leq z^*(\lambda) \text{ then } \Delta_{a=0}^{a=1} V_{\lambda}^{(a,*)}(x) &\leq 0; & \text{and} \\ \text{if } x \geq z^*(\lambda) \text{ then } \Delta_{a=0}^{a=1} V_{\lambda}^{(a,*)}(x) &\geq 0. \end{aligned} \quad (20)$$

We call any such $z^*(\cdot)$ an *optimal threshold map*.

Remark 4. (i) If the project is thresholdable with $z^*(\cdot)$ satisfying (20), the optimal value function is $V_{\lambda}^*(x) = V_{\lambda}(x, z^*(\lambda)) = V_{\lambda}(x, z^*(\lambda)^-)$ for every x and λ .

(ii) The project is thresholdable if for every price $\lambda \in \mathbb{R}$ there is a threshold $z \in \overline{\mathbb{R}}$ (not necessarily unique) such that both the z -policy and the z^- -policy are P_{λ} -optimal.

(iii) A sufficient condition for thresholdability is that, for every price $\lambda \in \mathbb{R}$, $V_{\lambda}^{(a,*)}(x)$ is *supermodular* in (x, a) , i.e., $\Delta_{a=0}^{a=1} V_{\lambda}^{(a,*)}(x)$ is nondecreasing in x .

- (iv) Suppose a project is thresholdable. Could we use such information to establish its indexability, using only intrinsic properties of its optimal threshold maps $z^*(\cdot)$? Clearly the answer is negative, since (20) does not fully characterize when each action is optimal, whereas (19) does.

Motivated by Remark 4(iv), we will further consider the following strong version of the thresholdability property, under which the P_λ -optimal actions are fully characterized by a threshold.

Definition 3 (Strong thresholdability). We call the project *strongly thresholdable* if there is a map $z^*: \mathbb{R} \rightarrow \overline{\mathbb{R}}$ such that

$$\text{sgn } \Delta_{a=0}^{a=1} V_\lambda^{(a,*)}(x) = \text{sgn}(x - z^*(\lambda)), \quad \text{for } x \in \mathbb{X} \text{ and } \lambda \in \mathbb{R}. \quad (21)$$

Remark 5.

- (i) There can only be one optimal threshold map $z^*(\cdot)$ satisfying (21).
- (ii) The project is strongly thresholdable with optimal threshold map $z^*(\cdot)$ if the P_λ -optimal action is non-decreasing in the state x , with both actions being P_λ -optimal only for $x = z^*(\lambda)$; thus, for $x < z^*(\lambda)$ (resp. $x > z^*(\lambda)$), $a = 0$ (resp. $a = 1$) is the only P_λ -optimal action (cf. Remark 3(iii)); equivalently, for each λ , equation (17) in x has $z^*(\lambda)$ as its unique root and, further, $\Delta_{a=0}^{a=1} V_\lambda^{(a,*)}(x) < 0$ for $x < z^*(\lambda)$ and $\Delta_{a=0}^{a=1} V_\lambda^{(a,*)}(x) > 0$ for $x > z^*(\lambda)$; or, for every λ , the P_λ -optimal active and passive regions are given by

$$S_\lambda^{*,1} = \{x \in \mathbb{X}: x \geq z^*(\lambda)\} \quad \text{and} \quad S_\lambda^{*,0} = \{x \in \mathbb{X}: x \leq z^*(\lambda)\}. \quad (22)$$

The following definition formalizes the concept of indexability holding consistently with optimality of threshold policies.

Definition 4 (Threshold-indexability). We call the project *threshold-indexable* if it is thresholdable and indexable.

Our main aim is to identify sufficient conditions for threshold-indexability. We will further consider the corresponding strong version of the latter property.

Definition 5 (Strong threshold-indexability). We call the project *strongly threshold-indexable* if it is strongly thresholdable and indexable.

2.5 Marginal metrics and the MP index.

As key tools in our indexability analyses, we will use the concepts of *marginal metrics* and *marginal productivity* (MP) index, introduced in Niño-Mora (2002) for the analysis of discrete-state projects.

For an action a and policy π , let $\langle a, \pi \rangle$ be the policy that takes action a in period $t = 0$ and adopts π afterwards. We measure the changes in reward earned and in resource expended, respectively, resulting from a change from $\langle 0, \pi \rangle$ to $\langle 1, \pi \rangle$ starting from x , by the *marginal reward metric*

$$f(x, \pi) \triangleq \Delta_{a=0}^{a=1} F(x, \langle a, \pi \rangle) = F(x, \langle 1, \pi \rangle) - F(x, \langle 0, \pi \rangle) \quad (23)$$

and the *marginal resource (usage) metric*

$$g(x, \pi) \triangleq \Delta_{a=0}^{a=1} G(x, \langle a, \pi \rangle). \quad (24)$$

Remark 6. (i) It follows from Remark 2(ii) and (23, 24) that, for any policy $\pi \in \Pi$,

$$f(\cdot, \pi), g(\cdot, \pi) \in \mathbb{B}_w(\mathbb{X}), \quad \text{with } \max\{\|f(\cdot, \pi)\|_w, \|g(\cdot, \pi)\|_w\} \leq 2M_\gamma. \quad (25)$$

(ii) For a given active region B , the functions $f(\cdot, B)$ and $g(\cdot, B)$ can be evaluated by

$$f(x, B) = \begin{cases} F(x, B) - r(x, 0) - \beta \int F(y, B) \kappa^0(x, dy), & x \in B \\ r(x, 1) + \beta \int F(y, B) \kappa^1(x, dy) - F(x, B), & x \in B^c, \end{cases}$$

$$g(x, B) = \begin{cases} G(x, B) - c(x, 0) - \beta \int G(y, B) \kappa^0(x, dy), & x \in B \\ c(x, 1) + \beta \int G(y, B) \kappa^1(x, dy) - G(x, B), & x \in B^c; \end{cases}$$

or, in terms of the linear operator \mathcal{L}^* in (11),

$$f(x, B) = \begin{cases} \mathcal{L}^* F(\cdot, B)(x, 0) - r(x, 0), & x \in B \\ r(x, 1) - \mathcal{L}^* F(\cdot, B)(x, 1), & x \in B^c, \end{cases} \quad (26)$$

$$g(x, B) = \begin{cases} \mathcal{L}^* G(\cdot, B)(x, 0) - c(x, 0), & x \in B \\ c(x, 1) - \mathcal{L}^* G(\cdot, B)(x, 1), & x \in B^c. \end{cases} \quad (27)$$

(iii) For a given project state x and policy π , $f(x, \pi)$ and $g(x, \pi)$ can be represented as

$$f(x, \pi) = \Delta_{a=0}^{a=1} \{r(x, a) + \beta F(\kappa^a(x, \cdot), \pi)\} \quad \text{and} \quad g(x, \pi) = \Delta_{a=0}^{a=1} \{c(x, a) + \beta G(\kappa^a(x, \cdot), \pi)\}. \quad (28)$$

If $g(x, B) \neq 0$ for a given x and B , we measure the MP of engaging the project in state x at time $t = 0$, given that the B -policy is adopted thereafter, by the *MP metric*

$$m(x, B) \triangleq \frac{f(x, B)}{g(x, B)}. \quad (29)$$

We shall further write as $f(x, z)$, $g(x, z)$, $m(x, z)$ and $f(x, z^-)$, $g(x, z^-)$, $m(x, z^-)$ the marginal metrics for the z -policy and the z^- -policy, respectively (see Section 2.4).

We next use the MP metric to define the *MP index*.

Definition 6 (MP index). Suppose $g(x, x) \neq 0$ for any state x . The project's *MP index* is the map $m: \mathsf{X} \rightarrow \mathbb{R}$ given by

$$m(x) \triangleq m(x, x) = \frac{f(x, x)}{g(x, x)}, \quad x \in \mathsf{X}. \quad (30)$$

Note the abuse of notation in (30), since m is used both for the MP index and the MP metric.

2.6 Main result: a verification theorem for threshold-indexability.

We next present our main result, giving sufficient conditions for threshold-indexability. The conditions correspond to the concept of *PCL-indexability*, which is extended here to the real-state setting.

Definition 7 (PCL-indexability). We call the project *PCL-indexable* (with respect to threshold policies) if the following *PCL-indexability conditions* hold:

- (PCLI1) the marginal resource metric satisfies $g(x, z) > 0$ for every $x \in \mathsf{X}$ and $z \in \overline{\mathbb{R}}$;
- (PCLI2) the MP index $m(\cdot)$ is nondecreasing and continuous on X ;
- (PCLI3) for each state x , the metrics $F(x, \cdot)$, $G(x, \cdot)$ and $m(\cdot)$ are related by (2).

Remark 7. (i) Conditions (PCLI1, PCLI2) match those given for discrete-state projects in Niño-Mora (2001, 2002, 2006), except for the continuity in (PCLI2).

- (ii) Under (PCLI1), the MP metric $m(x, z)$ (see (29)) is well defined for every x and z .
- (iii) Under (PCLI1, PCLI2), we extend the MP index domain to $\overline{\mathbb{R}}$ by $m(z) \triangleq m(\ell)$ for $z < \ell$ if $\ell > -\infty$, $m(z) \triangleq m(u)$ for $z > u$ if $u < \infty$, $m(-\infty) \triangleq \lim_{z \rightarrow -\infty} m(z)$, and $m(\infty) \triangleq \lim_{x \rightarrow \infty} m(x)$.
- (iv) Under (PCLI1, PCLI2), it follows immediately that the range $m(\mathbf{X})$ of $m(\cdot)$ is an interval.
- (v) (PCLI3) requires that, for each state x , $F(x, \cdot)$ be an indefinite LS integral of the MP index $m(\cdot)$ with respect to $G(x, \cdot)$. The existence of such integrals is justified in Lemma 12.
- (vi) For discrete-state projects, (PCLI1) implies (PCLI3). See Niño-Mora (2002, theorem 6.4(b)) for the finite-state case and Niño-Mora (2006, lemma 5.5.(a)) for the denumerable-state case.
- (vii) For any initial-state distribution $\nu_0 \in \mathbb{P}_w(\mathbf{X})$, (PCLI3) yields (cf. Remark 2(iii))

$$F(\nu_0, z_2) - F(\nu_0, z_1) = \int_{(z_1, z_2]} m(z) G(\nu_0, dz), \quad -\infty < z_1 < z_2 < \infty. \quad (31)$$

For a project satisfying (PCLI1, PCLI2), define for any price $\lambda \in \mathbb{R}$ the *threshold set*

$$Z_\lambda \triangleq \begin{cases} \{z \in \overline{\mathbb{R}}: m(z) = \lambda\} & \text{if } \{x \in \mathbf{X}: m(x) = \lambda\} \neq \emptyset \\ \{z \in \overline{\mathbb{R}}: z < \ell\} & \text{if } \ell > -\infty \text{ and } \lambda < m(x) \text{ for every } x \in \mathbf{X} \\ \{-\infty\} & \text{if } \ell = -\infty \text{ and } \lambda < m(x) \text{ for every } x \in \mathbf{X} \\ \{z \in \overline{\mathbb{R}}: z > u\} & \text{if } u < \infty \text{ and } \lambda > m(x) \text{ for every } x \in \mathbf{X} \\ \{\infty\} & \text{if } u = \infty \text{ and } \lambda > m(x) \text{ for every } x \in \mathbf{X}. \end{cases} \quad (32)$$

Under (PCLI1, PCLI2), we shall refer to the *generalized inverse* (cf. Falkner and Teschl (2012, p. 413)) of the MP index. We shall call a threshold map $\zeta: \mathbb{R} \rightarrow \overline{\mathbb{R}}$ a *generalized inverse* of $m(\cdot)$ if

$$\zeta(\lambda) \in Z_\lambda, \quad \lambda \in \mathbb{R}. \quad (33)$$

The following result gives two key properties of such generalized inverses of the MP index.

Lemma 1. *Assume (PCLI1, PCLI2) hold and let $\zeta(\cdot)$ be a generalized inverse of $m(\cdot)$. For any project state x and resource price λ the following holds:*

- (a) *if $x \leq \zeta(\lambda)$ then $m(x) \leq \lambda$; and if $x \geq \zeta(\lambda)$ then $m(x) \geq \lambda$;*
- (b) *$\zeta(\cdot)$ is increasing on the interval $m(\mathbf{X})$.*

Proof. (a) This part follows straightforwardly from (32) and (33).

(b) Let $\zeta(\cdot)$ be a generalized inverse of $m(\cdot)$ and suppose $\zeta(\cdot)$ were not increasing on $m(\mathbf{X})$, which is an interval by (PCLI2). Then there would exist $m(x_1) = \lambda_1 < \lambda_2 = m(x_2)$ with $x_1, x_2 \in \mathbf{X}$ and $\zeta(\lambda_1) \geq \zeta(\lambda_2)$. Now, from $\zeta(\lambda_k) \in Z_{\lambda_k}$, (32), (33) and nondecreasingness of $m(\cdot)$ it would follow that $\lambda_1 = m(\zeta(\lambda_1)) \geq m(\zeta(\lambda_2)) = \lambda_2$, a contradiction. Hence $\zeta(\cdot)$ must be increasing on $m(\mathbf{X})$. \square

We next state the verification theorem.

Theorem 1. *Under (PCLI1–PCLI3), the project is threshold-indexable with Whittle index $m(\cdot)$, and its optimal threshold maps are the generalized inverses $\zeta(\cdot)$ of $m(\cdot)$.*

The proof of Theorem 1, given in Section 9, builds on a number of implications of PCL-indexability conditions (PCLI1–PCLI3), which are developed in Section 5–8.

3 Motivation of the PCL-indexability conditions.

This section presents results motivating the PCL-indexability conditions.

We start by characterizing thresholdability for an indexable project, in its weak and strong versions (see Definitions 2 and 3), in terms of properties of the Whittle index. The following result shows that condition (PCLI2) is necessary for threshold-indexability, and further characterizes optimal threshold maps as generalized inverses of the Whittle index (cf. Theorem 1).

Lemma 2. *Suppose the project is indexable with Whittle index $\lambda^*(\cdot)$. Then it is thresholdable iff $\lambda^*(\cdot)$ is non-decreasing and continuous, in which case its optimal threshold maps $z^*(\cdot)$ are the generalized inverses of $\lambda^*(\cdot)$, which are increasing on $\lambda^*(\mathbf{X})$.*

Proof. Suppose the project is thresholdable. To obtain a contradiction, suppose $\lambda^*(\cdot)$ were not nondecreasing. Then there would be states $x < y$ with $\lambda^*(y) < \lambda^*(x)$ and, for any price $\lambda^*(y) < \lambda < \lambda^*(x)$, indexability would imply (see Remark 3(iii)) that $a = 0$ and $a = 1$ are the only P_λ -optimal actions in y and x , respectively, which would rule out P_λ -optimality of a threshold policy. Hence $\lambda^*(\cdot)$ must be nondecreasing. It further follows that $\lambda^*(\cdot)$ can only have jump discontinuities, so the left and right limits $\lambda^*(x^-)$ and $\lambda^*(x^+)$ exist at states $x > \ell$ and $x < u$, respectively.

To show that $\lambda^*(\cdot)$ is left-continuous, suppose $\lambda^*(y^-) < \lambda^*(y)$ for some $y > \ell$. Then for any $\lambda^*(y^-) < \lambda < \lambda^*(y)$ and $x_1 < y < x_2$, it would be $\lambda^*(x_1) \leq \lambda^*(y^-) < \lambda < \lambda^*(y) < \lambda^*(x_2)$ by nondecreasingness of $\lambda^*(\cdot)$. Hence, indexability (cf. Remark 3(iii)) would imply that the only P_λ -optimal threshold policy would be the y^- -policy, contradicting thresholdability (cf. Remark 4(i)).

As for right-continuity of $\lambda^*(\cdot)$, if there were a state $y < u$ with $\lambda^*(y) < \lambda^*(y^+)$, then for any $\lambda^*(y) < \lambda < \lambda^*(y^+)$ and states $x_1 < y < x_2$ it would be $\lambda^*(x_1) \leq \lambda^*(y) < \lambda < \lambda^*(y^+) \leq \lambda^*(x_2)$. Hence by indexability the y -policy would be the only P_λ -optimal threshold policy, again a contradiction.

As for the reverse implication, suppose $\lambda^*(\cdot)$ is nondecreasing and continuous, and fix $\lambda \in \mathbb{R}$. Then, it is immediate from the assumptions that for the following choices of z both the z -policy and the z^- -policy are P_λ -optimal: if $\lambda^*(x) < \lambda$ for every state x , take any $z > u$ if $u < \infty$, and $z = \infty$ if $u = \infty$; if $\lambda^*(x) > \lambda$ for every state x , take any $z < \ell$ if $\ell > -\infty$, and $z = -\infty$ if $\ell = -\infty$; and, otherwise, take as z any state with $\lambda^*(z) = \lambda$, whose existence is ensured by (PCLI2). This shows both that the project is thresholdable (see Remark 4(i)), and that its optimal threshold maps $z^*(\cdot)$ are the generalized inverses of $\lambda^*(\cdot)$. It further follows as in the proof of Lemma 1(b) that any such $z^*(\cdot)$ must be increasing on $\lambda^*(\mathbf{X})$. \square

The next result characterizes via marginal metrics the optimal regions $S_\lambda^{*,a}$ in (18). Let $\lambda \in \mathbb{R}$.

Lemma 3. *Let policy π^* be P_λ -optimal. Then*

$$S_\lambda^{*,1} = \{x \in \mathbf{X} : f(x, \pi^*) - \lambda g(x, \pi^*) \geq 0\} \quad \text{and} \quad S_\lambda^{*,0} = \{x \in \mathbf{X} : f(x, \pi^*) - \lambda g(x, \pi^*) \leq 0\}.$$

Proof. Since π^* is P_λ -optimal, we can reformulate (18) as

$$S_\lambda^{*,a} = \{x \in \mathbf{X} : V_\lambda(x, \langle a, \pi^* \rangle) \geq V_\lambda(x, \langle 1-a, \pi^* \rangle)\}, \quad a \in \{0, 1\}. \quad (34)$$

Now, it immediately follows that

$$\Delta_{a=0}^{a=1} V_\lambda(x, \langle a, \pi^* \rangle) = f(x, \pi^*) - \lambda g(x, \pi^*), \quad (35)$$

which, along with (34), yields the result. \square

The next result justifies using the MP index $m(\cdot)$ as a candidate for the project's Whittle index.

Proposition 1. *Let the project be threshold-indexable with Whittle index $\lambda^*(\cdot)$. Then*

- (a) $f(x, x) - \lambda^*(x)g(x, x) = 0 = f(x, x^-) - \lambda^*(x)g(x, x^-)$ for every state x ;
- (b) $\lambda^*(x) = m(x)$ for any state x with $g(x, x) \neq 0$;

(c) under (PCLI1), $\lambda^*(\cdot) = m(\cdot)$.

Proof. (a) Since the project is threshold-indexable, $\lambda^*(\cdot)$ is nondecreasing by Lemma 2. It follows from this and Definition 1 that both the x -policy and the x^- -policy are P_λ -optimal for $\lambda = \lambda^*(x)$. Hence, Lemma 3 implies that $f(x, x) - \lambda^*(x)g(x, x) = f(x, x^-) - \lambda^*(x)g(x, x^-) = 0$.

(b) This part follows from the first identity in (a) and (30).

(c) This part follows from (b) and (PCLI1). \square

The following two lemmas will be used to establish in Proposition 2 the necessity of (PCLI3) for threshold-indexability. Note that, for any x , the map $\lambda \mapsto V_\lambda^*(x)$ is convex on \mathbb{R} , as $V_\lambda^*(x)$ is the supremum of linear maps in λ . Hence, it is differentiable almost everywhere and has a nonempty set of *subgradients* (see, e.g., Niculescu and Persson (2006, §1.5)) at every λ , which we denote by $\partial V_\lambda^*(x)$. The following result shows how to obtain subgradients of such a map. Let $x \in X$ and $\lambda \in \mathbb{R}$.

Lemma 4. *If π^* is a P_λ -optimal policy, then $-G(x, \pi^*) \in \partial V_\lambda^*(x)$.*

Proof. Let $\delta > 0$. Using that $V_\lambda^*(x) = F(x, \pi^*) - \lambda G(x, \pi^*)$, we have

$$V_\lambda^*(x) - \delta G(x, \pi^*) = F(x, \pi^*) - (\lambda + \delta)G(x, \pi^*) = V_{\lambda+\delta}(x, \pi^*) \leq V_{\lambda+\delta}^*(x)$$

and

$$V_\lambda^*(x) + \delta G(x, \pi^*) = F(x, \pi^*) - (\lambda - \delta)G(x, \pi^*) = V_{\lambda-\delta}(x, \pi^*) \leq V_{\lambda-\delta}^*(x),$$

whence

$$V_\lambda^*(x) - V_{\lambda-\delta}^*(x) \leq -\delta G(x, \pi^*) \leq V_{\lambda+\delta}^*(x) - V_\lambda^*(x).$$

This shows that $-G(x, \pi^*)$ is a subgradient of $V_\lambda^*(x)$. \square

Lemma 5. *Suppose the project is thresholdable and let $z^*(\cdot)$ be an optimal threshold map. Then for any state x and prices $\lambda_1 < \lambda_2$,*

$$(a) \quad V_{\lambda_1}^*(x) - V_{\lambda_2}^*(x) = \int_{(\lambda_1, \lambda_2]} G(x, z^*(\lambda)) d\lambda;$$

$$(b) \quad \Delta_{a=0}^{a=1} V_{\lambda_1}^{(a,*)}(x) - \Delta_{a=0}^{a=1} V_{\lambda_2}^{(a,*)}(x) = \int_{(\lambda_1, \lambda_2]} g(x, z^*(\lambda)) d\lambda.$$

Proof. (a) This follows from the integral representation of convex functions (see Niculescu and Persson (2006, proposition 1.6.1)) and Lemma 4, as $-G(x, z^*(\lambda)) \in \partial V_\lambda^*(x)$.

(b) We have

$$\begin{aligned} V_{\lambda_1}^{(a,*)}(x) - V_{\lambda_2}^{(a,*)}(x) &= (\lambda_2 - \lambda_1)c(x, a) + \beta \int [V_{\lambda_1}^*(y) - V_{\lambda_2}^*(y)] \kappa^a(x, dy) \\ &= (\lambda_2 - \lambda_1)c(x, a) + \beta \int \int_{(\lambda_1, \lambda_2]} G(y, z^*(\lambda)) d\lambda \kappa^a(x, dy) \\ &= \int_{(\lambda_1, \lambda_2]} \left[c(x, a) + \beta \int G(y, z^*(\lambda)) \kappa^a(x, dy) \right] d\lambda, \end{aligned}$$

where we have used that $V_\lambda^{(a,*)}(x) = r(x, a) - \lambda c(x, a) + \beta \int V_\lambda^*(y) \kappa^a(x, dy)$, part (a), and Fubini's theorem (using that $G(x, \pi) \geq 0$). Therefore we obtain, as required,

$$\Delta_{a=0}^{a=1} [V_{\lambda_1}^{(a,*)}(x) - V_{\lambda_2}^{(a,*)}(x)] = \int_{(\lambda_1, \lambda_2]} g(x, z^*(\lambda)) d\lambda.$$

\square

The following result shows that condition (PCLI3) in (2) is necessary for threshold-indexability. Its proof relies on $G(x, \cdot)$ being a right-continuous function of bounded variation on finite intervals, to ensure the validity of identities (36) and (37). This is proved in Lemmas 10 and 11.

Proposition 2. *Suppose the project is threshold-indexable with index $\lambda^*(\cdot)$. Then for any x ,*

$$F(x, z_2) - F(x, z_1) = \int_{(z_1, z_2]} \lambda^*(z) G(x, dz), \quad -\infty < z_1 < z_2 < \infty.$$

Proof. Assume with no loss of generality that $z_1, z_2 \in \mathsf{X}$. Let $z^*(\cdot)$ be an optimal threshold map. By Lemma 2, $z^*(\cdot)$ is a generalized inverse of the nondecreasing continuous map $\lambda^*(\cdot)$, and hence we can apply the change of variable formula in Falkner and Teschl (2012, Eq. (1)) to obtain

$$\int_{(z_1, z_2]} G(x, z) \lambda^*(dz) = \int_{(\lambda_1, \lambda_2]} G(x, z^*(\lambda)) d\lambda, \quad (36)$$

where $\lambda_k = \lambda^*(z_k)$ for $k = 1, 2$.

On the other hand, integration by parts (see Carter and van Brunt (2000, theorem 6.2.2)) gives

$$\int_{(z_1, z_2]} G(x, z) \lambda^*(dz) = \lambda_2 G(x, z_2) - \lambda_1 G(x, z_1) - \int_{(z_1, z_2]} \lambda^*(z) G(x, dz). \quad (37)$$

Furthermore, we can write

$$V_{\lambda_1}^*(x) - V_{\lambda_2}^*(x) = F(x, z_1) - \lambda_1 G(x, z_1) - F(x, z_2) + \lambda_2 G(x, z_2). \quad (38)$$

Now, we use in turn (38), Lemma 5(a), (37) and (36) to obtain

$$\begin{aligned} F(x, z_2) - F(x, z_1) &= \lambda_2 G(x, z_2) - \lambda_1 G(x, z_1) - V_{\lambda_1}^*(x) + V_{\lambda_2}^*(x) \\ &= \int_{(z_1, z_2]} G(x, z) \lambda^*(dz) + \int_{(z_1, z_2]} \lambda^*(z) G(x, dz) - \int_{(\lambda_1, \lambda_2]} G(x, z^*(\lambda)) d\lambda \\ &= \int_{(z_1, z_2]} \lambda^*(z) G(x, dz). \end{aligned}$$

□

4 Indexability of strongly thresholdable projects.

We next present results relevant to prevailing approaches to indexability (see Section 1.3), which rely on first establishing strong thresholdability.

Lemma 6. *An indexable project is strongly thresholdable iff its index $\lambda^*(\cdot)$ is increasing and continuous, in which case its optimal threshold map $z^*(\cdot)$ is increasing and continuous on $\lambda^*(\mathsf{X})$.*

Proof. Suppose the project is strongly thresholdable. Then, by Lemma 2, $\lambda^*(\cdot)$ is nondecreasing and continuous. To show that its Whittle index $\lambda^*(\cdot)$ must be increasing note that, from Definitions 1 and 3, its unique (see Remark 5(i)) optimal threshold map $z^*(\cdot)$ and Whittle index satisfy

$$\text{sgn}(x - z^*(\lambda)) = \text{sgn}(\lambda^*(x) - \lambda), \quad \text{for } x \in \mathsf{X} \text{ and } \lambda \in \mathbb{R}. \quad (39)$$

Take states $x_1 < x_2$. Then, writing $\lambda_k \triangleq \lambda^*(x_k)$, (39) yields $z^*(\lambda_k) = x_k$ for $k = 1, 2$. We now use again (39) to obtain, as required,

$$\text{sgn}(\lambda_1 - \lambda_2) = \text{sgn}(\lambda^*(x_1) - \lambda_2) = \text{sgn}(x_1 - z^*(\lambda_2)) = \text{sgn}(x_1 - x_2) < 0.$$

Consider now the reverse implication, so $\lambda^*(\cdot)$ is assumed increasing and continuous. Then it has a unique inverse $z^*(\cdot) \triangleq \lambda^{*, -1}(\cdot)$ on the interval $\lambda^*(\mathsf{X})$, which is also increasing and continuous, and satisfies the identity in (39) for $x \in \mathsf{X}$ and $\lambda \in \lambda^*(\mathsf{X})$. If $\lambda^*(\mathsf{X}) \neq \mathbb{R}$, it is immediate to extend the domain of such a $z^*(\cdot)$ to \mathbb{R} consistently with (39). This yields strong thresholdability. □

The next result shows that to establish indexability of a strongly thresholdable project it does not suffice to prove that its optimal threshold map is increasing. It needs also to be proved that the optimal threshold map is continuous. Note that $z^{*, -1}(\mathbf{X}) \triangleq \{\lambda: z^*(\lambda) \in \mathbf{X}\}$.

Proposition 3. *Assume the project is strongly thresholdable with optimal threshold map $z^*(\cdot)$. Then it is indexable iff $z^*(\cdot)$ is increasing and continuous on $z^{*, -1}(\mathbf{X})$, in which case its Whittle index is also increasing and continuous, being $\lambda^*(x) = z^{*, -1}(x)$ for $x \in \mathbf{X}$.*

Proof. If the project is indexable, the stated properties of $z^*(\cdot)$ follow by Lemma 2.

As for the reverse implication, suppose the project is strongly thresholdable with $z^*(\cdot)$ increasing and continuous on $z^{*, -1}(\mathbf{X})$. Then the restriction of $z^*(\cdot)$ on the domain $z^{*, -1}(\mathbf{X})$ has a unique inverse $\lambda^*(\cdot) \triangleq z^{*, -1}(\cdot)$, which is also increasing and continuous, and satisfies (39). This implies that the project is indexable with Whittle index $\lambda^*(\cdot)$, which completes the proof. \square

5 LP reformulation of the λ -price problem starting from $X_0 \sim \nu_0$.

This section lays down further groundwork for proving Theorem 1. It draws upon standard results to reformulate the problem

$$P_\lambda(\nu_0): \quad \underset{\pi \in \Pi}{\text{maximize}} \quad V_\lambda(\nu_0, \pi), \quad (40)$$

which is the λ -price problem P_λ in (4) starting from $X_0 \sim \nu_0 \in \mathbb{P}_w(\mathbf{X})$, as an LP problem.

For any policy $\pi \in \Pi$, consider the *state-action occupation measure* $\mu_{\nu_0}^\pi$ defined by

$$\mu_{\nu_0}^\pi(\Gamma) \triangleq \mathbb{E}_{\nu_0}^\pi \left[\sum_{t=0}^{\infty} \beta^t \mathbf{1}_\Gamma(X_t, A_t) \right] = \sum_{t=0}^{\infty} \beta^t \mathbb{P}_{\nu_0}^\pi \{(X_t, A_t) \in \Gamma\}, \quad \Gamma \in \mathcal{B}(\mathbf{K}). \quad (41)$$

Since the objective of problem $P_\lambda(\nu_0)$ can be represented in terms of $\mu_{\nu_0}^\pi$ as

$$V_\lambda(\nu_0, \pi) = \int (r - \lambda c) d\mu_{\nu_0}^\pi = \int \{r(y, a) - \lambda c(y, a)\} \mu_{\nu_0}^\pi(d(y, a)),$$

we can reformulate the dynamic optimization problem $P_\lambda(\nu_0)$ as the static optimization problem

$$L_\lambda(\nu_0): \quad \underset{\mu \in \mathcal{M}_{\nu_0}}{\text{maximize}} \quad \int (r - \lambda c) d\mu, \quad (42)$$

where $\mathcal{M}_{\nu_0} \triangleq \{\mu_{\nu_0}^\pi: \pi \in \Pi\}$, which is characterized by linear constraints (see, e.g., Heilmann (1979) and Hernández-Lerma and Lasserre (1996, §6.3)) as outlined below.

The $\mu_{\nu_0}^\pi$ belong (see Hernández-Lerma and Lasserre (1999, proposition 7.2.2)) to the Banach space $\mathbb{M}_w(\mathbf{K})$ of finite *signed measures* μ on $\mathcal{B}(\mathbf{K})$ with finite w -norm (with $|\mu|$ the *total variation* of μ)

$$\|\mu\|_w \triangleq \sup_{h \in \mathbb{B}_w(\mathbf{K}), \|h\|_w \leq 1} \left| \int h d\mu \right| = \int w d|\mu| = \int w(y) |\mu|(d(y, a)). \quad (43)$$

For any $\mu \in \mathbb{M}_w(\mathbf{K})$ consider the marginal signed measures on $\mathcal{B}(\mathbf{X})$ defined by $\tilde{\mu}^a(S) \triangleq \mu(S \times \{a\})$ for $a \in \{0, 1\}$, and $\tilde{\mu}(S) \triangleq \mu(S \times \{0, 1\}) = \sum_{a \in \{0, 1\}} \tilde{\mu}^a(S)$, which belong to the Banach space $\mathbb{M}_w(\mathbf{X})$ of finite signed measures ν on $\mathcal{B}(\mathbf{X})$ with finite w -norm (cf. (7))

$$\|\nu\|_w \triangleq \sup_{v \in \mathbb{B}_w(\mathbf{X}), \|v\|_w \leq 1} \left| \int v d\nu \right| = \int w d|\nu| = \int w(y) |\nu|(dy). \quad (44)$$

In particular, the *marginal action/state-occupation measures* for $\mu_{\nu_0}^\pi$, $\tilde{\mu}_{\nu_0}^{\pi, a}$ and $\tilde{\mu}_{\nu_0}^\pi$, are in $\mathbb{M}_w(\mathbf{X})$.

Both $(\mathbb{M}_w(\mathbb{K}), \mathbb{B}_w(\mathbb{K}))$ and $(\mathbb{M}_w(\mathbb{X}), \mathbb{B}_w(\mathbb{X}))$ are known to be *dual pairs of vector spaces* (see [Hernández-Lerma and Lasserre \(1999, §12.2.A\)](#)) with respect to the bilinear forms

$$\langle \mu, h \rangle \triangleq \int h d\mu = \int h(y, a) \mu(d(y, a)), \quad \mu \in \mathbb{M}_w(\mathbb{K}), h \in \mathbb{B}_w(\mathbb{K}), \quad (45)$$

$$\langle \nu, v \rangle \triangleq \int v d\nu = \int v(y) \nu(dy), \quad \nu \in \mathbb{M}_w(\mathbb{X}), v \in \mathbb{B}_w(\mathbb{X}). \quad (46)$$

Consider now the linear operators $\mathcal{L}^a: \mathbb{M}_w(\mathbb{K}) \rightarrow \mathbb{M}_w(\mathbb{X})$ for $a \in \{0, 1\}$, mapping μ to

$$\mathcal{L}^a \mu(S) \triangleq \tilde{\mu}^a(S) - \beta \int \kappa^a(y, S) \tilde{\mu}^a(dy), \quad S \in \mathcal{B}(\mathbb{X}), \quad (47)$$

and the linear operator $\mathcal{L} \triangleq \mathcal{L}^0 + \mathcal{L}^1$. Note that

$$\mathcal{L}\mu(S) \triangleq \tilde{\mu}(S) - \beta \int \kappa^a(y, S) \mu(d(y, a)), \quad S \in \mathcal{B}(\mathbb{X}). \quad (48)$$

Such operators are *bounded* (cf. [Hernández-Lerma and Lasserre \(1999, §7.2.B\)](#)).

It is known (cf. [Heilmann \(1977, theorem 8\)](#) and [Hernández-Lerma and Lasserre \(1996, theorem 6.3.7\)](#)) that \mathcal{M}_{ν_0} in (42) is a compact and convex region of $\mathbb{M}_w(\mathbb{K})$, which is spanned by stationary randomized policies and is characterized by linear constraints as

$$\mathcal{M}_{\nu_0} = \{\mu_{\nu_0}^\pi: \pi \in \Pi^{\text{SR}}\} = \{\mu \in \mathbb{M}_w^+(\mathbb{K}): \mathcal{L}\mu = \nu_0\}, \quad (49)$$

where $\mathbb{M}_w^+(\mathbb{K}) \triangleq \{\mu \in \mathbb{M}_w(\mathbb{K}): \mu \geq 0\}$ is the *cone of measures* in $\mathbb{M}_w(\mathbb{K})$.

We can thus explicitly formulate problem $L_\lambda(\nu_0)$ in (42) as the infinite-dimensional LP problem

$$\begin{aligned} L_\lambda(\nu_0): \quad & \text{maximize } \langle \mu, r - \lambda c \rangle \\ & \text{subject to: } \mathcal{L}\mu = \nu_0, \mu \in \mathbb{M}_w^+(\mathbb{K}). \end{aligned} \quad (50)$$

6 Decomposition of performance metrics.

We present in this section a key decomposition of project performance metrics, which will be used in subsequent analyses.

We need a preliminary result. Note that the linear operator \mathcal{L}^* in (11) is (cf. [Hernández-Lerma and Lasserre \(1996, p. 139\)](#)) the *adjoint* of \mathcal{L} in (48), as it satisfies the *basic adjoint relation*

$$\langle \mathcal{L}\mu, v \rangle = \langle \mu, \mathcal{L}^*v \rangle, \quad \mu \in \mathbb{M}_w(\mathbb{K}), v \in \mathbb{B}_w(\mathbb{X}), \quad (51)$$

or, in terms of the operators \mathcal{L}^a in (47),

$$\sum_{a \in \{0, 1\}} \langle \mathcal{L}^a \mu, v \rangle = \langle \mu, \mathcal{L}^*v \rangle. \quad (52)$$

The next result (which the author has not found in the literature) gives a decomposition of (52).

Lemma 7. *For any $a \in \{0, 1\}$, $\mu \in \mathbb{M}_w(\mathbb{K})$ and $v \in \mathbb{B}_w(\mathbb{X})$,*

$$\langle \mathcal{L}^a \mu, v \rangle = \langle \tilde{\mu}^a, \mathcal{L}^*v(\cdot, a) \rangle.$$

Proof. Let $a \in \{0, 1\}$ and $\mu \in \mathbb{M}_w(\mathbb{K})$. For any $S \in \mathcal{B}(\mathbb{X})$ we have

$$\begin{aligned} \langle \mathcal{L}^a \mu, 1_S \rangle &= \mathcal{L}^a \mu(S) = \int \left\{ 1_S(x) - \beta \int 1_S(y) \kappa^a(x, dy) \right\} \tilde{\mu}^a(dx) \\ &= \int \mathcal{L}^* 1_S(x, a) \tilde{\mu}^a(dx) = \langle \tilde{\mu}^a, \mathcal{L}^* 1_S(\cdot, a) \rangle, \end{aligned}$$

where we have used (46), (47), (11), and (45). By standard arguments, it follows from the above that $\langle \mathcal{L}^a \mu, v \rangle = \langle \tilde{\mu}^a, \mathcal{L}^*v(\cdot, a) \rangle$ for any $v \in \mathbb{B}_w(\mathbb{X})$. \square

We next give the main result of this section, consisting of two parts. Part (a) decomposes the reward metric $F(\nu_0, \pi)$ as the sum of $F(\nu_0, B)$ for a given active region B and a linear combination of $\tilde{\mu}_{\nu_0}^{\pi, a}$. Part (b) similarly decomposes the resource metric $G(\nu_0, \pi)$. This result extends to real-state projects corresponding results for the discrete-state case in Niño-Mora (2001, theorem 3), Niño-Mora (2002, proposition 6.1) and Niño-Mora (2006, lemma 5.4). Let $\nu_0 \in \mathbb{P}_w(\mathbf{X})$.

Lemma 8. *For any policy $\pi \in \Pi$ and active region B ,*

$$(a) \quad F(\nu_0, \pi) = F(\nu_0, B) - \int_B f(y, B) \tilde{\mu}_{\nu_0}^{\pi, 0}(dy) + \int_{B^c} f(y, B) \tilde{\mu}_{\nu_0}^{\pi, 1}(dy), \text{ viz.,}$$

$$F(\nu_0, \pi) = F(\nu_0, B) + \mathbf{E}_{\nu_0}^{\pi} \left[\sum_{t=0}^{\infty} \beta^t \{A_t - 1_B(X_t)\} f(X_t, B) \right]; \quad (53)$$

$$(b) \quad G(\nu_0, \pi) = G(\nu_0, B) - \int_B g(y, B) \tilde{\mu}_{\nu_0}^{\pi, 0}(dy) + \int_{B^c} g(y, B) \tilde{\mu}_{\nu_0}^{\pi, 1}(dy), \text{ viz.,}$$

$$G(\nu_0, \pi) = G(\nu_0, B) + \mathbf{E}_{\nu_0}^{\pi} \left[\sum_{t=0}^{\infty} \beta^t \{A_t - 1_B(X_t)\} g(X_t, B) \right]. \quad (54)$$

Proof. (a) Using in turn Lemma 7, (46), (26) and (12), we can write

$$\begin{aligned} \langle \mathcal{L}^0 \mu_{\nu_0}^{\pi}, F(\cdot, B) \rangle &= \langle \tilde{\mu}_{\nu_0}^{\pi, 0}, \mathcal{L}^* F(\cdot, B)(\cdot, 0) \rangle = \int \mathcal{L}^* F(\cdot, B)(y, 0) \tilde{\mu}_{\nu_0}^{\pi, 0}(dy) \\ &= \int_B \mathcal{L}^* F(\cdot, B)(y, 0) \tilde{\mu}_{\nu_0}^{\pi, 0}(dy) + \int_{B^c} \mathcal{L}^* F(\cdot, B)(y, 0) \tilde{\mu}_{\nu_0}^{\pi, 0}(dy) \\ &= \int_B \{r(y, 0) + f(y, B)\} \tilde{\mu}_{\nu_0}^{\pi, 0}(dy) + \int_{B^c} r(y, 0) \tilde{\mu}_{\nu_0}^{\pi, 0}(dy) \\ &= \int r(y, 0) \tilde{\mu}_{\nu_0}^{\pi, 0}(dy) + \int_B f(y, B) \tilde{\mu}_{\nu_0}^{\pi, 0}(dy) \end{aligned}$$

and

$$\begin{aligned} \langle \mathcal{L}^1 \mu_{\nu_0}^{\pi}, F(\cdot, B) \rangle &= \langle \tilde{\mu}_{\nu_0}^{\pi, 1}, \mathcal{L}^* F(\cdot, B)(\cdot, 1) \rangle = \int \mathcal{L}^* F(\cdot, B)(y, 1) \tilde{\mu}_{\nu_0}^{\pi, 1}(dy) \\ &= \int_B \mathcal{L}^* F(\cdot, B)(y, 1) \tilde{\mu}_{\nu_0}^{\pi, 1}(dy) + \int_{B^c} \mathcal{L}^* F(\cdot, B)(y, 1) \tilde{\mu}_{\nu_0}^{\pi, 1}(dy) \\ &= \int_B r(y, 1) \tilde{\mu}_{\nu_0}^{\pi, 1}(dy) + \int_{B^c} \{r(y, 1) - f(y, B)\} \tilde{\mu}_{\nu_0}^{\pi, 1}(dy) \\ &= \int r(y, 1) \tilde{\mu}_{\nu_0}^{\pi, 1}(dy) - \int_{B^c} f(y, B) \tilde{\mu}_{\nu_0}^{\pi, 1}(dy). \end{aligned}$$

Further, using that $\mathcal{L} = \sum_{a \in \{0,1\}} \mathcal{L}^a$, (49) and (46), we obtain

$$\sum_{a \in \{0,1\}} \langle \mathcal{L}^a \mu_{\nu_0}^{\pi}, F(\cdot, B) \rangle = \langle \mathcal{L} \mu_{\nu_0}^{\pi}, F(\cdot, B) \rangle = \langle \nu_0, F(\cdot, B) \rangle = F(\nu_0, B).$$

Combining the above identities with

$$F(\nu_0, \pi) = \int r d\mu_{\nu_0}^{\pi} = \sum_{a \in \{0,1\}} \int r(y, a) \tilde{\mu}_{\nu_0}^{\pi, a}(dy).$$

we readily obtain the stated decomposition identity

$$F(\nu_0, \pi) = F(\nu_0, B) - \int_B f(y, B) \tilde{\mu}_{\nu_0}^{\pi, 0}(dy) + \int_{B^c} f(y, B) \tilde{\mu}_{\nu_0}^{\pi, 1}(dy). \quad (55)$$

As for (53), it follows by reformulating (55) using that

$$\int_B f(y, B) \tilde{\mu}_{\nu_0}^{\pi,0}(dy) = \mathbb{E}_{\nu_0}^{\pi} \left[\sum_{t=0}^{\infty} \beta^t \mathbf{1}_{B \times \{0\}}(X_t, A_t) f(X_t, B) \right]$$

and

$$\int_{B^c} f(y, B) \tilde{\mu}_{\nu_0}^{\pi,1}(dy) = \mathbb{E}_{\nu_0}^{\pi} \left[\sum_{t=0}^{\infty} \beta^t \mathbf{1}_{B^c \times \{1\}}(X_t, A_t) f(X_t, B) \right]$$

as

$$F(\nu_0, \pi) + \mathbb{E}_{\nu_0}^{\pi} \left[\sum_{t=0}^{\infty} \beta^t \mathbf{1}_{B \times \{0\}}(X_t, A_t) f(X_t, B) \right] = F(\nu_0, B) + \mathbb{E}_{\nu_0}^{\pi} \left[\sum_{t=0}^{\infty} \beta^t \mathbf{1}_{B^c \times \{1\}}(X_t, A_t) f(X_t, B) \right],$$

and then reformulating the latter identity as (53) noting that, for any state x and action a ,

$$\mathbf{1}_{B^c \times \{1\}}(x, a) - \mathbf{1}_{B \times \{0\}}(x, a) = a - \mathbf{1}_B(x).$$

Part (b) follows along the same lines as part (a). \square

We will use Lemma 8 as a powerful tool to generate useful identities.

7 Analysis of metrics as functions of the threshold.

This section presents properties of metrics under threshold policies as functions of the threshold. Throughout, $\nu_0 \in \mathbb{P}_w(\mathbb{X})$ and $x \in \mathbb{X}$.

7.1 Càdlàg property.

Lemma 10 shows that $F(\nu_0, z)$, $G(\nu_0, z)$, $f(x, z)$ and $g(x, z)$ are càdlàg in $z \in \mathbb{R}$. We need the following result, referring to the extension to \mathbb{R} of the state-occupation measures $\tilde{\mu}_{\nu_0}^{\pi}$, viz., $\tilde{\mu}_{\nu_0}^{\pi}(S) \triangleq \tilde{\mu}_{\nu_0}^{\pi}(S \cap \mathbb{X})$ for $S \in \mathcal{B}(\mathbb{R})$. We abuse notation by letting, e.g., $[z_1, z_2]$ mean $\{z \in \mathbb{R} : z_1 \leq z \leq z_2\}$, even when $z_1 = -\infty$ or $z_2 = \infty$, and similarly with other intervals.

Lemma 9. *For any thresholds $-\infty \leq z_1 < z_2 \leq \infty$,*

- (a) $F(\nu_0, z_1) - F(\nu_0, z_2) = \int_{(z_1, z_2]} f(y, z_2) \tilde{\mu}_{\nu_0}^{z_1}(dy) = \int_{(z_1, z_2]} f(y, z_1) \tilde{\mu}_{\nu_0}^{z_2}(dy);$
- (b) $G(\nu_0, z_1) - G(\nu_0, z_2) = \int_{(z_1, z_2]} g(y, z_2) \tilde{\mu}_{\nu_0}^{z_1}(dy) = \int_{(z_1, z_2]} g(y, z_1) \tilde{\mu}_{\nu_0}^{z_2}(dy);$
- (c) $F(\nu_0, z_1) - F(\nu_0, z_2^-) = \int_{(z_1, z_2)} f(y, z_2^-) \tilde{\mu}_{\nu_0}^{z_1}(dy) = \int_{(z_1, z_2)} f(y, z_1) \tilde{\mu}_{\nu_0}^{z_2^-}(dy);$
- (d) $G(\nu_0, z_1) - G(\nu_0, z_2^-) = \int_{(z_1, z_2)} g(y, z_2^-) \tilde{\mu}_{\nu_0}^{z_1}(dy) = \int_{(z_1, z_2)} g(y, z_1) \tilde{\mu}_{\nu_0}^{z_2^-}(dy).$
- (e) $F(\nu_0, z_1^-) - F(\nu_0, z_2) = \int_{[z_1, z_2]} f(y, z_2) \tilde{\mu}_{\nu_0}^{z_1^-}(dy) = \int_{[z_1, z_2]} f(y, z_1^-) \tilde{\mu}_{\nu_0}^{z_2}(dy);$
- (f) $G(\nu_0, z_1^-) - G(\nu_0, z_2) = \int_{[z_1, z_2]} g(y, z_2) \tilde{\mu}_{\nu_0}^{z_1^-}(dy) = \int_{[z_1, z_2]} g(y, z_1^-) \tilde{\mu}_{\nu_0}^{z_2}(dy).$

Proof. (a, b) For the first identity take $\pi = z_1$ and $B = (z_2, u]$ in Lemma 8(a, b), using that $\tilde{\mu}_{\nu_0}^{z_1,0}(z_1, u] = \tilde{\mu}_{\nu_0}^{z_1,1}[\ell, z_1] = 0$ and $\tilde{\mu}_{\nu_0}^{z_1,1} = \tilde{\mu}_{\nu_0}^{z_1}$ on $\mathcal{B}(z_1, z_2]$. For the second identity take $\pi = z_2$ and $B = (z_1, u]$, using that $\tilde{\mu}_{\nu_0}^{z_2,0}(z_2, u] = \tilde{\mu}_{\nu_0}^{z_2,1}[\ell, z_1] = 0$ and $\tilde{\mu}_{\nu_0}^{z_2,0} = \tilde{\mu}_{\nu_0}^{z_2}$ on $\mathcal{B}(z_1, z_2]$.

(c, d) For the first identity take $\pi = z_1$ and $B = [z_2, u]$ in Lemma 8(a, b), using that $\tilde{\mu}_{\nu_0}^{z_1,0}[z_2, u] = \tilde{\mu}_{\nu_0}^{z_1,1}[\ell, z_1] = 0$ and $\tilde{\mu}_{\nu_0}^{z_1,1} = \tilde{\mu}_{\nu_0}^{z_1}$ on $\mathcal{B}(z_1, z_2)$. For the second identity take $\pi = z_2^-$ and $B = (z_1, u]$, using that $\tilde{\mu}_{\nu_0}^{z_2^-,0}[z_2, u] = \tilde{\mu}_{\nu_0}^{z_2^-,1}[\ell, z_1] = 0$ and $\tilde{\mu}_{\nu_0}^{z_2^-,0} = \tilde{\mu}_{\nu_0}^{z_2^-}$ on $\mathcal{B}(z_1, z_2)$.

(e, f) For the first identity take $\pi = z_1^-$ and $B = (z_2, u]$ in Lemma 8(a, b), using that $\tilde{\mu}_{\nu_0}^{z_1^-, 0}(z_2, u] = \tilde{\mu}_{\nu_0}^{z_1^-, 1}(\bar{X}_{z_1}^c) = 0$ and $\tilde{\mu}_{\nu_0}^{z_1^-, 1} = \tilde{\mu}_{\nu_0}^{z_1^-}$ on $\mathcal{B}[z_1, z_2]$. For the second identity take $\pi = z_2$ and $B = [z_1, u]$, using that $\tilde{\mu}_{\nu_0}^{z_2, 0}(z_2, u] = \tilde{\mu}_{\nu_0}^{z_2, 1}[\ell, z_1] = 0$ and $\tilde{\mu}_{\nu_0}^{z_2, 0} = \tilde{\mu}_{\nu_0}^{z_2}$ on $\mathcal{B}[z_1, z_2]$. \square

Recall from Section 2.4 that the notation $F(\nu_0, z^-)$ (for this and other metrics) refers to $F(\nu_0, \pi)$ under the z^- -policy. The next result ensures the consistency of such notation with the standard one where $F(\nu_0, z^-)$ denotes the left limit at z of function $F(\nu_0, \cdot)$. The result further shows the consistency of the notation $F(\nu_0, -\infty)$ and $F(\nu_0, \infty)$, where the second argument refers to the threshold policies $-\infty$ (“always active”) and ∞ (“never active”) with the standard notation where $F(\nu_0, -\infty)$ and $F(\nu_0, \infty)$ are the limits of $F(\nu_0, z)$ as $z \rightarrow -\infty$ and as $z \rightarrow \infty$, respectively.

Lemma 10. *The following holds:*

- (a) $F(\nu_0, \cdot), G(\nu_0, \cdot), f(x, \cdot)$ and $g(x, \cdot)$ are bounded càdlàg functions with left limits at $z \in \mathbb{R}$ given by $F(\nu_0, z^-), G(\nu_0, z^-), f(x, z^-)$ and $g(x, z^-)$;
- (b) $\lim_{z \rightarrow -\infty} (F(\nu_0, z), G(\nu_0, z), f(x, z), g(x, z)) = (F(\nu_0, -\infty), G(\nu_0, -\infty), f(x, -\infty), g(x, -\infty))$;
- (c) $\lim_{z \rightarrow \infty} (F(\nu_0, z), G(\nu_0, z), f(x, z), g(x, z)) = (F(\nu_0, \infty), G(\nu_0, \infty), f(x, \infty), g(x, \infty))$.

Proof. (a) We first show that $F(\nu_0, \cdot)$ is right-continuous at z . From Lemma 9(a) we obtain (with M_γ as in (8)), for $\delta > 0$,

$$|F(\nu_0, z + \delta) - F(\nu_0, z)| \leq \int_{(z, z+\delta]} |f|(y, z + \delta) \tilde{\mu}_{\nu_0}^z(dy) \leq 2M_\gamma \int_{(z, z+\delta]} w(y) \tilde{\mu}_{\nu_0}^z(dy) \rightarrow 0 \text{ as } \delta \searrow 0,$$

where the second inequality follows from (25), and the limit follows by the dominated convergence theorem, since $\tilde{\mu}_{\nu_0}^z \in \mathbb{M}_w(\mathbb{X})$ (see Section 5). The result for $G(\nu_0, \cdot)$ follows similarly. Furthermore,

$$\begin{aligned} \lim_{\delta \searrow 0} f(x, z + \delta) &= \lim_{\delta \searrow 0} \Delta_{a=0}^{a=1} \left\{ r(x, a) + \beta \int F(y, z + \delta) \kappa^a(x, dy) \right\} \\ &= \Delta_{a=0}^{a=1} \left\{ r(x, a) + \beta \int F(y, z) \kappa^a(x, dy) \right\} = f(x, z), \end{aligned}$$

where the interchange of limit and integral is justified by the dominated convergence theorem, using (8) and the right-continuity of $F(y, \cdot)$. The result for $g(x, \cdot)$ follows similarly.

Consider now the left limits. From Lemma 9(c) and arguing as above we obtain

$$|F(\nu_0, z - \delta) - F(\nu_0, z^-)| \leq \int_{(z-\delta, z)} |f|(y, z - \delta) \tilde{\mu}_{\nu_0}^{z^-}(dy) \leq 2M_\gamma \int_{(z-\delta, z)} w(y) \tilde{\mu}_{\nu_0}^{z^-}(dy) \rightarrow 0 \text{ as } \delta \searrow 0,$$

and similarly for the left limits of $G(\nu_0, \cdot), f(x, \cdot)$ and $g(x, \cdot)$ at z .

(b) The result is trivial if ℓ is finite, so suppose $\ell = -\infty$. For any finite z , taking $z_1 = -\infty$ and $z_2 = z$ in the first identity in Lemma 9(a) yields

$$F(\nu_0, -\infty) = F(\nu_0, z) + \int_{(-\infty, z]} f(y, z) \tilde{\mu}_{\nu_0}^{-\infty}(dy). \quad (56)$$

The result that $\lim_{z \rightarrow -\infty} F(\nu_0, z) = F(\nu_0, -\infty)$ now follows from (56) by the dominated convergence theorem, arguing as in part (a). The results for $G(\nu_0, -\infty), f(x, -\infty)$ and $g(x, -\infty)$ follow similarly.

(c) The result is trivial if u is finite, so suppose $u = \infty$. For any finite z , taking $z_1 = z$ and $z_2 = \infty$ in the second identity in Lemma 9(a) yields

$$F(\nu_0, \infty) = F(\nu_0, z) - \int_{(z, \infty)} f(y, z) \tilde{\mu}_{\nu_0}^{\infty}(dy). \quad (57)$$

The result that $\lim_{z \rightarrow \infty} F(\nu_0, z) = F(\nu_0, \infty)$ follows from (57) by the dominated convergence theorem, arguing as in part (a). The results for $G(\nu_0, \infty), f(x, \infty)$ and $g(x, \infty)$ follow similarly. \square

7.2 Monotonicity of resource metric.

The next result shows that, under (PCLI1), the resource metric $G(\nu_0, z)$ is, as intuition would suggest, monotone nonincreasing in the threshold z .

Lemma 11. *Let (PCLI1) hold. Then*

- (a) $G(\nu_0, z)$ is nonincreasing in z on \mathbb{R} ;
- (b) if ν_0 has full support X then $G(\nu_0, z)$ is decreasing in z on X .

Proof. (a) For finite $z_1 < z_2$, Lemma 9(b) and (PCLI1) yield $G(\nu_0, z_1) \geq G(\nu_0, z_2)$.
(b) Lemma 9(b) further implies that, for any finite thresholds $z_1 < z_2$ in X ,

$$G(\nu_0, z_1) - G(\nu_0, z_2) = \int_{(z_1, z_2]} g(y, z_2) \tilde{\mu}_{\nu_0}^{z_1}(dy) > 0, \quad (58)$$

where the inequality follows from (PCLI1) and $\tilde{\mu}_{\nu_0}^{z_1}(z_1, z_2] \geq \nu_0(z_1, z_2] > 0$. \square

We next establish existence and finiteness of the LS integrals in condition (PCLI3).

Lemma 12. *Under (PCLI1, PCLI2), $\int_{(z_1, z_2]} m(z) G(\nu_0, dz)$ exists and is finite for finite $z_1 < z_2$.*

Proof. By Lemmas 10 and 11(a), $G(\nu_0, \cdot)$ is bounded, càdlàg and nonincreasing. Since $m(\cdot)$ is continuous, $\int_{(z_1, z_2]} m(z) G(\nu_0, dz)$ exists and is finite by standard results on LS integration. \square

7.3 Analysis of discontinuities.

Being càdlàg, functions $F(\nu_0, \cdot)$, $G(\nu_0, \cdot)$, $f(x, \cdot)$ and $g(x, \cdot)$ have at most countably many discontinuities, which are of jump type. We analyze below such discontinuities. The following result gives simple formulae for their jumps. We denote by $\Delta_2 F(\nu_0, z) \triangleq F(\nu_0, z) - F(\nu_0, z^-)$ the jump of $F(\nu_0, \cdot)$ at z , and similarly for $\Delta_2 G(\nu_0, z)$, $\Delta_2 f(x, z)$ and $\Delta_2 g(x, z)$.

Lemma 13. *For any threshold $z \in \mathsf{X}$,*

- (a) $\Delta_2 F(\nu_0, z) = -f(z, z^-) \tilde{\mu}_{\nu_0}^z\{z\} = -f(z, z) \tilde{\mu}_{\nu_0}^{z^-}\{z\}$;
- (b) $\Delta_2 G(\nu_0, z) = -g(z, z^-) \tilde{\mu}_{\nu_0}^z\{z\} = -g(z, z) \tilde{\mu}_{\nu_0}^{z^-}\{z\}$.

Proof. (a) Taking $\pi = z$ and $B = \{y \in \mathsf{X} : y \geq z\}$ in Lemma 8(a) gives

$$F(\nu_0, z) + \int_{y \geq z} f(y, z^-) \tilde{\mu}_{\nu_0}^{z, 0}(dy) = F(\nu_0, z^-) + \int_{y < z} f(y, z^-) \tilde{\mu}_{\nu_0}^{z, 1}(dy),$$

so $\Delta_2 F(\nu_0, z) + f(z, z^-) \tilde{\mu}_{\nu_0}^z\{z\} = 0$, as $\tilde{\mu}_{\nu_0}^{z, 0}\{z\} = \tilde{\mu}_{\nu_0}^z\{z\}$ and $\tilde{\mu}_{\nu_0}^{z, 0}\{y > z\} = \tilde{\mu}_{\nu_0}^{z, 1}\{y < z\} = 0$.

Further, taking $\pi = z^-$ and $B = \{y \in \mathsf{X} : y > z\}$ in Lemma 8(a) gives

$$F(\nu_0, z^-) + \int_{y > z} f(y, z) \tilde{\mu}_{\nu_0}^{z^-, 0}(dy) = F(\nu_0, z) + \int_{y \leq z} f(y, z) \tilde{\mu}_{\nu_0}^{z^-, 1}(dy),$$

so $\Delta_2 F(\nu_0, z) + f(z, z) \tilde{\mu}_{\nu_0}^{z^-}\{z\} = 0$, as $\tilde{\mu}_{\nu_0}^{z^-, 1}\{z\} = \tilde{\mu}_{\nu_0}^{z^-}\{z\}$ and $\tilde{\mu}_{\nu_0}^{z^-, 0}\{y > z\} = \tilde{\mu}_{\nu_0}^{z^-, 1}\{y < z\} = 0$.

Part (b) follows along the same lines as part (a). \square

Recall that PCL-indexability condition (PCLI1) requires $g(x, z)$ to be positive for any state x and threshold z . The following result shows that it further implies $g(x, x^-) > 0$.

Lemma 14. *Under (PCLI1), $g(x, x^-) > 0$ and $\Delta_2 G(x, x) < 0$.*

Proof. Taking $\nu_0 = \delta_x$ and $z = x$ in Lemma 13(b) and using (PCLI1) and $\tilde{\mu}_x^{x^-}\{x\} \geq 1$ gives

$$g(x, x^-) \tilde{\mu}_x^{x^-}\{x\} = g(x, x) \tilde{\mu}_x^{x^-}\{x\} \geq g(x, x) > 0,$$

whence $g(x, x^-) > 0$ and $\Delta_2 G(x, x) < 0$. □

The next result characterizes the discontinuities of $G(\nu_0, \cdot)$ in terms of $T_z \triangleq \min\{t \geq 0: X_t = z\}$, the *first hitting time to z* of state process $\{X_t\}_{t=0}^\infty$ under the z -policy, with $T_z \triangleq \infty$ if z is never hit.

Lemma 15. *Under (PCLI1), $G(\nu_0, \cdot)$ is discontinuous at z iff $\mathbb{P}_{\nu_0}^z\{T_z < \infty\} > 0$.*

Proof. It follows from Lemmas 13(b) and 14 that $G(\nu_0, \cdot)$ is discontinuous at z iff $\tilde{\mu}_{\nu_0}^z\{z\} > 0$, which happens iff $\mathbb{P}_{\nu_0}^z\{T_z < \infty\} > 0$. □

The next result relates the jumps of $F(\nu_0, \cdot)$ and $G(\nu_0, \cdot)$, gives a corresponding result for $f(x, \cdot)$ and $g(x, \cdot)$, and ensures that the MP metric $m(x, \cdot)$ (see (29)) is continuous at $z = x$.

Lemma 16. *Let (PCLI1) hold. Then, for any threshold $z \in \mathbb{X}$,*

- (a) $\Delta_2 F(\nu_0, z) = m(z) \Delta_2 G(\nu_0, z)$;
- (b) $\Delta_2 f(x, z) = m(z) \Delta_2 g(x, z)$;
- (c) $m(x, \cdot)$ is continuous at x , viz., $m(x, x^-) = m(x, x) = m(x)$.

Proof. (a) Lemma 13(a, b) and $g(z, z) > 0$ yield

$$\Delta_2 F(\nu_0, z) = -f(z, z) \tilde{\mu}_{\nu_0}^{z^-}\{z\} = -m(z) g(z, z) \mu_{\nu_0}^{z^-}\{z\} = m(z) \Delta_2 G(\nu_0, z).$$

(b) Using in turn (23), part (a) and (24), gives

$$\Delta_2 f(x, z) = \beta \Delta_{a=0}^{a=1} \int \Delta_2 F(y, z) \kappa^a(x, dy) = m(z) \beta \Delta_{a=0}^{a=1} \int \Delta_2 G(y, z) \kappa^a(x, dy) = m(z) \Delta_2 g(x, z).$$

(c) Taking $z = x$ in part (b) and noting that $m(x) = f(x, x)/g(x, x)$, we obtain

$$f(x, x^-) - f(x, x) = m(x)[g(x, x^-) - g(x, x)] = m(x)g(x, x^-) - f(x, x).$$

whence $f(x, x^-) = m(x)g(x, x^-)$ and, by Lemma 14, $m(x) = m(x, x^-)$. □

Lemma 16(a) yields an alternate characterization of the MP index under (PCLI1) in terms of jumps $\Delta_2 F(x, x) = F(x, x) - F(x, x^-)$ and $\Delta_2 G(x, x) = G(x, x) - G(x, x^-)$ (cf. Lemma 14).

Corollary 1. *Under (PCLI1), $m(x) = \Delta_2 F(x, x)/\Delta_2 G(x, x)$.*

Further, Lemma 16(a, b) immediately implies the following result relating the continuity points of functions $F(x, \cdot)$ and $G(x, \cdot)$, and of $f(x, \cdot)$ and $g(x, \cdot)$.

Corollary 2. *Under (PCLI1),*

- (a) if $G(\nu_0, \cdot)$ is continuous at z , so is $F(\nu_0, \cdot)$;
- (b) if $g(x, \cdot)$ is continuous at z , so is $f(x, \cdot)$.

7.4 Bounded variation and MP index as Radon–Nikodym derivative.

We next show that the metrics of a PCL-indexable project, as functions of the threshold, belong to the linear space $\mathbb{V}(\mathbb{R})$ of functions of *bounded variation* on \mathbb{R} (see Carter and van Brunt (2000, §2.7)). We further build on that result to characterize the MP index as a Radon–Nikodym derivative.

We start with the resource metrics, for which condition (PCLI1) suffices.

Lemma 17. *Under (PCLI1), $G(\nu_0, \cdot), g(x, \cdot) \in \mathbb{V}(\mathbb{R})$.*

Proof. $G(\nu_0, \cdot) \in \mathbb{V}(\mathbb{R})$ follows from $G(\nu_0, \cdot)$ being bounded (see (8)) and nonincreasing on \mathbb{R} (see Lemma 11(a)). Since $g(x, z) = \Delta_{a=0}^{a=1}[c(x, a) + \beta G(\kappa^a(x, \cdot), z)]$, $g(x, \cdot)$ is a difference of bounded nonincreasing functions, whence (see Carter and van Brunt (2000, theorem 2.7.2)) $g(x, \cdot) \in \mathbb{V}(\mathbb{R})$. \square

Remark 8. Since the function $G(\nu_0, \cdot)$ is bounded, càdlàg and nonincreasing, by Carathéodory’s extension theorem it induces a unique *finite LS measure* $\nu_{G(\nu_0, \cdot)}$ on $\mathcal{B}(\mathbb{R})$ satisfying

$$\nu_{G(\nu_0, \cdot)}(z_1, z_2] = G(\nu_0, z_1) - G(\nu_0, z_2), \quad -\infty < z_1 < z_2 < \infty.$$

In light of Remark 8, we have by standard results that the LS integral

$$\int |m|(z) G(\nu_0, dz) = - \int |m| d\nu_{G(\nu_0, \cdot)}$$

is well defined. The following result shows that such an integral is finite, even when \mathbf{X} is unbounded.

Lemma 18. *Under PCL-indexability, $m(\cdot)$ is $\nu_{G(\nu_0, \cdot)}$ -integrable.*

Proof. Consider first the case that $m(\cdot) \geq 0$ over \mathbb{R} . Then,

$$\begin{aligned} \int |m| d\nu_{G(\nu_0, \cdot)} &= \int m \nu_{G(\nu_0, \cdot)} = \lim_{n \rightarrow \infty} \int_{(-n, n]} m \nu_{G(\nu_0, \cdot)} \\ &= \lim_{n \rightarrow \infty} \{F(\nu_0, -n) - F(\nu_0, n)\} = F(\nu_0, -\infty) - F(\nu_0, \infty) < \infty, \end{aligned}$$

where we have used the monotone convergence theorem, condition (PCLI3) and Lemma 10(b, c). A similar argument yields the result in the case that $m \leq 0$ over \mathbb{R} .

Otherwise, (PCLI2) implies that there is a state b with $m(b) = 0$ such that $|m| = -m$ on $(-\infty, b]$ and $|m| = m$ on (b, ∞) . Arguing along the same lines as above yields

$$\begin{aligned} \int |m| \nu_{G(\nu_0, \cdot)} &= \int_{(-\infty, b]} -m d\nu_{G(\nu_0, \cdot)} + \int_{(b, \infty)} m \nu_{G(\nu_0, \cdot)} = \lim_{n \rightarrow \infty} \int_{(-n, b]} -m d\nu_{G(\nu_0, \cdot)} + \lim_{n \rightarrow \infty} \int_{(b, n]} m \nu_{G(\nu_0, \cdot)} \\ &= \lim_{n \rightarrow \infty} \{F(\nu_0, b) - F(\nu_0, -n)\} + \lim_{n \rightarrow \infty} \{F(\nu_0, b) - F(\nu_0, n)\} \\ &= 2F(\nu_0, b) - F(\nu_0, -\infty) - F(\nu_0, \infty) < \infty. \end{aligned}$$

\square

The following result shows that, under PCL-indexability, the identity (31) —and, in particular condition (PCLI3)— extends to infinite intervals.

Lemma 19. *Under PCL-indexability,*

$$F(\nu_0, z_2) - F(\nu_0, z_1) = \int_{z_1 < z \leq z_2} m(z) G(\nu_0, dz), \quad -\infty \leq z_1 < z_2 \leq \infty. \quad (59)$$

Proof. The result follows from (31), Lemmas 10(b) and 18, and dominated convergence. \square

We next present the counterpart of Lemma 17 for reward measures.

Lemma 20. *Under PCL-indexability, $F(\nu_0, \cdot), f(x, \cdot) \in \mathbb{V}(\mathbb{R})$.*

Proof. Since $\nu_{G(\nu_0, \cdot)}$ (see Remark 8) is a finite measure, Lemma 18 ensures that $\tilde{\nu}_{\nu_0}(S) \triangleq \int_S m d\nu_{G(\nu_0, \cdot)}$ is a finite signed measure, and hence admits a Jordan decomposition $\tilde{\nu}_{\nu_0} = \tilde{\nu}_{\nu_0}^+ - \tilde{\nu}_{\nu_0}^-$ with $\tilde{\nu}_{\nu_0}^+$ and $\tilde{\nu}_{\nu_0}^-$ finite measures. On the other hand, (PCLI3) and Lemma 10(b) yield $\tilde{\nu}_{\nu_0}(-\infty, z] = F(\nu_0, -\infty) - F(\nu_0, z)$ for $z \in \mathbb{R}$, whence $F(\nu_0, z) = F(\nu_0, -\infty) + \tilde{\nu}_{\nu_0}^-(-\infty, z] - \tilde{\nu}_{\nu_0}^+(-\infty, z]$. Hence, $F(\nu_0, \cdot) \in \mathbb{V}(\mathbb{R})$, being a difference of bounded nondecreasing functions. Further, it follows that function $z \mapsto F(\kappa^a(\nu_0, \cdot), z)$ is in $\mathbb{V}(\mathbb{R})$. Thus, the identity $f(x, z) = \Delta_{a=0}^{a=1}[r(x, a) + \beta F(\kappa^a(x, \cdot), z)]$ represents $f(x, \cdot)$ as the difference of two functions in $\mathbb{V}(\mathbb{R})$, whence $f(x, \cdot) \in \mathbb{V}(\mathbb{R})$. \square

Remark 9. In light of Lemma 20 and $F(\nu_0, \cdot)$ being bounded càdlàg under PCL-indexability (see (9) and Lemma 10), Carathéodory's theorem for signed measures (cf. Doob (1994, section X.6)) ensures existence of a unique finite LS signed measure $\nu_{F(\nu_0, \cdot)}$ on $\mathcal{B}(\mathbb{R})$ satisfying

$$\nu_{F(\nu_0, \cdot)}(z_1, z_2] \triangleq F(\nu_0, z_1) - F(\nu_0, z_2), \quad -\infty < z_1 < z_2 < \infty.$$

The next result shows that the MP index is a Radon–Nikodym derivative of $\nu_{F(\nu_0, \cdot)}$ with respect to $\nu_{G(\nu_0, \cdot)}$. The notation \ll below means “is absolutely continuous with respect to”.

Proposition 4. *Under PCL-indexability, $\nu_{F(\nu_0, \cdot)} \ll \nu_{G(\nu_0, \cdot)}$, with $m(\cdot)$ being a Radon–Nikodym derivative of $\nu_{F(\nu_0, \cdot)}$ with respect to $\nu_{G(\nu_0, \cdot)}$, viz.,*

$$\nu_{F(\nu_0, \cdot)}(S) = \int_S m(x) \nu_{G(\nu_0, \cdot)}(dx), \quad S \in \mathcal{B}(\mathbb{R}). \quad (60)$$

Proof. The result that $\nu_{F(\nu_0, \cdot)}$ satisfies (60) follows from (2) and Remark 9, whereas $\nu_{F(\nu_0, \cdot)} \ll \nu_{G(\nu_0, \cdot)}$ follows directly from (60). \square

By standard results on differentiation of LS measures (cf. Doob (1994, section X.4)), the above ensures that $m(\cdot)$ is $\nu_{G(\nu_0, \cdot)}$ -a.s. the derivative of the signed measure $\nu_{F(\nu_0, \cdot)}$ with respect to the measure $\nu_{G(\nu_0, \cdot)}$, denoted by $d\nu_{F(\nu_0, \cdot)}/d\nu_{G(\nu_0, \cdot)}$, or of $F(\nu_0, \cdot)$ with respect to $G(\nu_0, \cdot)$, denoted by $dF(\nu_0, \cdot)/dG(\nu_0, \cdot)$.

7.5 Relations between marginal project metrics.

This section presents relations between marginal metrics under threshold policies as functions of the threshold. The following result gives counterparts of Lemmas 18 and 19 for marginal metrics $f(x, \cdot)$ and $g(x, \cdot)$. Note that the LS integrals below are well defined, since $g(x, \cdot)$ is a function of bounded variation on \mathbb{R} (see Lemma 17) and $m(\cdot)$ is continuous under PCL-indexability. The result refers to the finite signed measure $\nu_{g(x, \cdot)}$ determined by $g(x, \cdot)$ through $\nu_{g(x, \cdot)}(z_1, z_2] \triangleq g(x, z_1) - g(x, z_2)$ for finite $z_1 < z_2$.

Lemma 21. *Under PCL-indexability,*

- (a) $m(\cdot)$ is $\nu_{g(x, \cdot)}$ -integrable, i.e., the integral $\int |m| d|\nu_{g(x, \cdot)}|$ is finite;
- (b) $f(x, z_2) - f(x, z_1) = \int_{(z_1, z_2]} m(z) g(x, dz)$, for $-\infty \leq z_1 < z_2 \leq \infty$.

Proof. (a) This part follows from Lemma 18 using that, by (28), $\nu_{g(x, \cdot)} = \beta \Delta_{a=0}^{a=1} \nu_{G(\kappa^a(x, \cdot), \cdot)}$.

(b) Using part (a), Lemma 19 and (28), we obtain

$$\begin{aligned} \int_{(z_1, z_2]} m(z) g(x, dz) &= \beta \Delta_{a=0}^{a=1} \int_{(z_1, z_2]} m(z) G(\kappa^a(x, \cdot), dz) \\ &= \beta \Delta_{a=0}^{a=1} \{F(\kappa^a(x, \cdot), z_2) - F(\kappa^a(x, \cdot), z_1)\} = f(x, z_2) - f(x, z_1). \end{aligned}$$

\square

8 Further relations between MP metrics.

This section builds on the above to present further relations between MP metrics that will play a key role in the proof of the verification theorem. The next result shows that the MP index satisfies certain linear Volterra–Stieltjes integral equations.

Lemma 22. *Suppose the project is PCL-indexable and let $x \in X$ and $z \in \overline{\mathbb{R}}$. Then*

- (a) $m(x, z) - m(z) = \int_{z < b < x} \frac{g(x, b)}{g(x, z)} m(db)$, if $x > z$;
- (b) $m(x, z) - m(z) = - \int_{x \leq b \leq z} \frac{g(x, b)}{g(x, z)} m(db)$, if $x \leq z$;

Proof. (a) From Lemmas 10, 21, and the dominated convergence theorem, we have

$$f(x, x^-) - f(x, z) = \lim_{y \nearrow x} f(x, y) - f(x, z) = \lim_{y \nearrow x} \int_{(z, y]} m(b) g(x, db) = \int_{(z, x)} m(b) g(x, db).$$

Consider the case $z > -\infty$. Using the latter identity, integration by parts, and Lemma 16(c) gives

$$\begin{aligned} f(x, z) &= f(x, x^-) - \int_{(z, x)} m(b) g(x, db) \\ &= f(x, x^-) - [m(x)g(x, x^-) - m(z)g(x, z)] + \int_{(z, x)} g(x, b) m(db) \\ &= m(z)g(x, z) + \int_{(z, x)} g(x, b) m(db). \end{aligned} \tag{61}$$

The result now follows by dividing each term by $g(x, z)$.

In the case $z = -\infty$, the result follows from the limiting argument

$$\begin{aligned} f(x, -\infty) - m(-\infty)g(x, -\infty) &= \lim_{z' \rightarrow -\infty} f(x, z') - m(z')g(x, z') \\ &= \lim_{z' \rightarrow -\infty} \int_{(z', x)} g(x, b) m(db) = \int_{(-\infty, x)} g(x, b) m(db), \end{aligned}$$

where we have used Lemma 10(b), (PCLI1), (61) and (PCLI2).

(b) Consider the case $z < \infty$. Arguing as in part (a), we obtain

$$\begin{aligned} f(x, z) &= f(x, x^-) + \int_{[x, z]} m(b) g(x, db) \\ &= f(x, x^-) + m(z)g(x, z) - m(x)g(x, x^-) - \int_{[x, z]} g(x, b) m(db) \\ &= m(z)g(x, z) - \int_{[x, z]} g(x, b) m(db). \end{aligned} \tag{62}$$

In the case $z = \infty$, the result follows from the limiting argument

$$\begin{aligned} f(x, \infty) - m(\infty)g(x, \infty) &= \lim_{z' \rightarrow \infty} f(x, z') - m(z')g(x, z') \\ &= - \lim_{z' \rightarrow \infty} \int_{[x, z']} g(x, b) m(db) = - \int_{[x, \infty)} g(x, b) m(db), \end{aligned}$$

where we have used Lemma 10(c), (PCLI1), (62) and (PCLI2). This completes the proof. \square

The next result shows that $m(x, z) - m(z)$ and $m(x) - m(z)$ have the same sign.

Lemma 23. *Suppose the project is PCL-indexable. Then, for any state x and threshold z ,*

$$\operatorname{sgn}(m(x, z) - m(z)) = \operatorname{sgn}(m(x) - m(z)).$$

Proof. If $x > z$, we use Lemma 22(a) and (PCLI1, PCLI2) to obtain

$$\operatorname{sgn}(m(x, z) - m(z)) = \operatorname{sgn} \int_{z < b < x} g(x, b) m(db) = \operatorname{sgn}(m(x) - m(z)) \geq 0.$$

If $x \leq z$, we use Lemma 22(b) and (PCLI1, PCLI2) to obtain

$$\operatorname{sgn}(m(x, z) - m(z)) = -\operatorname{sgn} \int_{x \leq b \leq z} g(x, b) m(db) = \operatorname{sgn}(m(x) - m(z)) \leq 0.$$

□

We need a final preliminary result, showing that $m(x, z) - \lambda$ and $m(x) - \lambda$ have the same sign when threshold z is taken from the set Z_λ defined in (32).

Lemma 24. *Suppose the project is PCL-indexable. Then, for any $x \in \mathbf{X}$, $\lambda \in \mathbb{R}$ and $z \in Z_\lambda$,*

$$\operatorname{sgn}(m(x, z) - \lambda) = \operatorname{sgn}(m(x) - \lambda). \quad (63)$$

Proof. Let $x \in \mathbf{X}$, $\lambda \in \mathbb{R}$ and $z \in Z_\lambda$. We distinguish three cases. In the case $\{y \in \mathbf{X}: m(y) = \lambda\} \neq \emptyset$, we have $m(z) = \lambda$ (see (32)), and hence (63) holds by Lemma 23.

If $\lambda < m(y)$ for every y , then (see (32)) $z < \ell$ if $\ell > -\infty$, and $z = -\infty$ if $\ell = -\infty$, so $m(z) = m(\ell) \geq \lambda$, with strict inequality if $\ell > -\infty$ (as then $\ell \in \mathbf{X}$). We can now write, using Lemma 22(a),

$$m(x, z) - \lambda = m(x, z) - m(z) + m(z) - \lambda = \int_{(z, x)} \frac{g(x, b)}{g(x, z)} m(db) + m(\ell) - \lambda > 0, \quad (64)$$

where the inequality follows from (PCLI1) if $m(\ell) > \lambda$. If $m(\ell) = \lambda$ then $\ell = -\infty$, and by the assumptions the integral in the right-hand side of (64) is positive, yielding the inequality and (63).

If $\lambda > m(y)$ for every y , then (see (32)) $z > u$ if $u < \infty$, and $z = \infty$ if $u = \infty$, so $m(z) = m(u) \leq \lambda$, with strict inequality if u is finite (because then $u \in \mathbf{X}$). We can now write, using Lemma 22(b),

$$m(x, z) - \lambda = m(x, z) - m(z) + m(z) - \lambda = -\int_{b \geq x} \frac{g(x, b)}{g(x, z)} m(db) + m(u) - \lambda < 0, \quad (65)$$

where the inequality follows from (PCLI1) if $m(u) < \lambda$. If $m(u) = \lambda$ then $u = \infty$, and by the assumptions the integral in the right-hand side of (65) is positive, which yields (63). □

9 Proving Theorem 1.

We are now ready to prove Theorem 1.

Proof. Proof of Theorem 1. Let $\zeta(\cdot)$ be a generalized inverse of the MP index $m(\cdot)$. For any $x \in \mathbf{X}$ and $\lambda \in \mathbb{R}$, we have $\zeta(\cdot) \in Z_\lambda$ and, using (PCLI1) and Lemma 63,

$$\operatorname{sgn} \Delta_{a=0}^{a=1} V_\lambda(x, z) = \operatorname{sgn}(f(x, \zeta(\lambda)) - \lambda g(x, \zeta(\lambda))) = \operatorname{sgn}(m(x, \zeta(\lambda)) - \lambda) = \operatorname{sgn}(m(x) - \lambda). \quad (66)$$

The outermost identity in (66) and Lemma 1(a) imply that $\zeta(\cdot)$ satisfies (20), whence the project is thresholdable with (see Remark 4(i)) optimal value $V_\lambda^*(x) = V_\lambda(x, \zeta(\lambda))$. We thus obtain

$$\operatorname{sgn} \Delta_{a=0}^{a=1} V_\lambda^*(x) = \operatorname{sgn}(m(x) - \lambda), \quad (67)$$

which shows (cf. Definition 1) that the project is indexable with Whittle index $m(\cdot)$. □

10 How to establish condition (PCLI3).

This section presents conditions under which PCL-indexability condition (PCLI3) can be established by simpler means than direct verification.

10.1 Piecewise differentiable $F(x, \cdot)$ and $G(x, \cdot)$.

We start with the case that metrics $F(x, \cdot)$ and $G(x, \cdot)$ are piecewise differentiable. Denote by $F'_2(x, z)$ and $G'_2(x, z)$ their partial derivatives with respect to z . Below, \mathring{S} denotes the interior of a set S .

Assumption 2. For each state x there is a countable partition $\{S_i(x) : i \in N(x)\}$ of \mathbb{R} with each $S_i(x)$ being either a left-semiclosed interval or a singleton, such that:

- (i) $F(x, \cdot) = \sum_{i \in N(x)} F_i(x, \cdot) 1_{S_i(x)}(\cdot)$ and $G(x, \cdot) = \sum_{i \in N(x)} G_i(x, \cdot) 1_{S_i(x)}(\cdot)$, with $F_i(x, \cdot)$ and $G_i(x, \cdot)$ being both differentiable on $\mathring{S}_i(x)$.
- (ii) $F'_2(x, z) = m(z)G'_2(x, z)$ for $z \in \mathring{S}_i(x)$ and $i \in N(x)$.

We need two preliminary results. Denote by $\text{TV}_{F(x, \cdot)}(I)$ and $\text{TV}_{G(x, \cdot)}(I)$ the total variation of $F(x, \cdot)$ and $G(x, \cdot)$ on an interval I , and by $D(x) \triangleq \mathbb{R} \setminus \cup_{i \in N(x)} \mathring{S}_i(x)$ the set of interval endpoints for partition $\{S_i(x) : i \in N(x)\}$. Recall that $\mathbb{V}(I)$ denotes the functions of bounded variation on I .

Lemma 25. Let (PCLI1, PCLI2) and Assumption 2 hold. Let x be a state and I a finite open interval where both $F(x, \cdot)$ and $G(x, \cdot)$ are differentiable. Then $F(x, \cdot) \in \mathbb{V}(I)$ and

$$\int_I F(x, dz) = \int_I m(z) G(x, dz).$$

Proof. Since $G(x, \cdot)$ is bounded nonincreasing by Lemmas 10(a) and 11(a), and $m(\cdot)$ is continuous, by standard results $\int_I |m|(z) G(x, dz)$ is finite. Furthermore, using Assumption 2(ii),

$$\int_I |m|(z) G(x, dz) = \int_I |m|(z) G'_2(x, z) dz = - \int_I |F'_2|(x, z) dz = -\text{TV}_{F(x, \cdot)}(I)$$

(cf. Carter and van Brunt (2000, theorem 6.1.7)). Hence, $F(x, \cdot) \in \mathbb{V}(I)$ and

$$\int_I m(z) G(x, dz) = \int_I m(z) G'_2(x, z) dz = \int_I F'_2(x, z) dz = \int_I F(x, dz).$$

□

In the next result, $\mathbb{V}_{\text{loc}}(\mathbb{R})$ denotes the class of functions of locally bounded variation on \mathbb{R} , i.e., having bounded variation on every finite interval.

Lemma 26. Under (PCLI1, PCLI2) and Assumption 2, $F(x, \cdot) \in \mathbb{V}_{\text{loc}}(\mathbb{R})$.

Proof. Let $I = (z_1, z_2)$ with $-\infty < z_1 < z_2 < \infty$ and $K(I) \triangleq \max\{|m|(z) : z \in [z_1, z_2]\}$, which is finite by continuity of $m(\cdot)$. The total variation $\text{TV}_{F(x, \cdot)}(I)$ of $F(x, \cdot)$ on I is finite, since

$$\begin{aligned} \text{TV}_{F(x, \cdot)}(I) &= \sum_{z \in D(x) \cap I} |\Delta_2 F(x, z)| + \sum_{i: \mathring{S}_i(x) \cap I \neq \emptyset} \int_{\mathring{S}_i(x) \cap I} |F'_2|(x, z) dz \\ &= \sum_{z \in D(x) \cap I} |m|(z) |\Delta_2 G(x, z)| + \sum_{i: \mathring{S}_i(x) \cap I \neq \emptyset} \int_{\mathring{S}_i(x) \cap I} |m|(z) |G'_2|(x, z) dz \\ &\leq K(I) \left\{ \sum_{z \in D(x) \cap I} |\Delta_2 G(x, z)| + \int_{i: \mathring{S}_i(x) \cap I \neq \emptyset} |G'_2|(x, z) dz \right\} \\ &= K(I) \text{TV}_{G(x, \cdot)}(I) \leq K(I) \text{TV}_{G(x, \cdot)}(\mathbb{R}) < \infty, \end{aligned}$$

where we have used Assumption 2 and Lemmas 16(a), 25 and 17. □

We can now give the main result of this section.

Proposition 5. *Under (PCLI1, PCLI2) and Assumption 2, (PCLI3) holds.*

Proof. Let $-\infty < z_1 < z_2 < \infty$. We have

$$\begin{aligned}
F(x, z_2) - F(x, z_1) &= \int_{(z_1, z_2]} F(x, dz) = \Delta_2 F(x, z_2) + \int_{(z_1, z_2)} F(x, dz) \\
&= \sum_{z \in D(x) \cap (z_1, z_2]} \Delta_2 F(x, z) + \sum_{i: \dot{S}_i(x) \cap (z_1, z_2) \neq \emptyset} \int_{\dot{S}_i(x) \cap (z_1, z_2)} F'_2(x, z) dz \\
&= \sum_{z \in D(x) \cap (z_1, z_2]} m(z) \Delta_2 G(x, z) + \sum_{i: \dot{S}_i(x) \cap (z_1, z_2) \neq \emptyset} \int_{\dot{S}_i(x) \cap (z_1, z_2)} m(z) G'_2(x, z) dz \\
&= \int_{(z_1, z_2]} m(z) G(x, dz),
\end{aligned}$$

where we have used Lemma 26, Assumption 2 and Lemmas 16(a), 25 and 12. \square

10.2 Piecewise constant $G(x, \cdot)$.

We next consider the case that the resource metric $G(x, z)$ is piecewise constant in the threshold variable z , with a finite or denumerable number of pieces.

Assumption 3. *For each state x , $G(x, \cdot) = \sum_{i \in N(x)} G_i(x) 1_{S_i(x)}(\cdot)$, where $\{S_i(x) : i \in N(x)\}$ is a countable partition of \mathbb{R} with each $S_i(x)$ being either a left-semiclosed interval or a singleton.*

Note that Assumption 3 does not refer to the reward metric, so it is not evident that this case reduces to that in Section 10.1. Yet, we will show that Assumption 3 implies Assumption 2.

The following result can be used to verify Assumption 3 through analysis of state trajectories $\{X_t\}_{t=0}^\infty$ under threshold policies. Parts (a, c) characterize when $G(x, \cdot)$ is constant on an interval. Parts (b, d) ensure that $F(x, \cdot)$ is constant over the same intervals that $G(x, \cdot)$. For a Borel set S , denote by $T_S \triangleq \min\{t \geq 0 : X_t \in S\}$ the first hitting time to S , with $T_S \triangleq \infty$ if S is never hit.

Lemma 27. *Let (PCLI1) hold. For any state x and finite $z_1 < z_2$,*

- (a) $G(x, \cdot)$ is constant on $[z_1, z_2]$ iff $\mathbb{P}_x^{z_1}\{T_{(z_1, z_2]} < \infty\} = 0$;
- (b) if $G(x, \cdot)$ is constant on $[z_1, z_2]$, then so is $F(x, \cdot)$;
- (c) $G(x, \cdot)$ is constant on (z_1, z_2) iff $\mathbb{P}_x^{z_1}\{T_{(z_1, z_2)} < \infty\} = 0$;
- (d) if $G(x, \cdot)$ is constant on (z_1, z_2) , then so is $F(x, \cdot)$.

Proof. (a) Since $G(x, \cdot)$ is nonincreasing on \mathbb{R} by Lemma 11(a), it is constant on $[z_1, z_2]$ iff $G(x, z_1) = G(x, z_2)$, which occurs, by Lemma 9(b), iff

$$\int_{(z_1, z_2]} g(y, z_2) \tilde{\mu}_x^{z_1}(dy) = 0. \tag{68}$$

Now, under (PCLI1), (68) holds iff $\tilde{\mu}_x^{z_1}(z_1, z_2] = 0$, which occurs iff $\mathbb{P}_x^{z_1}\{T_{(z_1, z_2]} < \infty\} = 0$.

(b) If $G(x, \cdot)$ is constant on $(z_1, z_2]$ then $\tilde{\mu}_x^{z_1}(z_1, z_2] = 0$, which, along with Lemma 9(a), yields

$$F(x, z_1) - F(x, z) = \int_{(z_1, z]} f(y, z) \tilde{\mu}_x^{z_1}(dy) = 0, \quad z \in (z_1, z_2].$$

Hence, $F(x, \cdot)$ is also constant on $[z_1, z_2]$.

(c) Let $\{z^n\}_{n=0}^\infty$ be increasing with $z_1 < z^n < z_2$ and $\lim_{n \rightarrow \infty} z^n = z_2$. Then $G(x, \cdot)$ is constant on $[z_1, z_2]$ iff $G(x, \cdot)$ is constant on $[z_1, z^n]$ for each n , which by part (a) occurs iff $\tilde{\mu}_x^{z_1}(z_1, z^n] = 0$ for each n . By continuity from below, such will be the case iff $\tilde{\mu}_x^{z_1}(z_1, z_2) = 0$, i.e., iff $\mathbb{P}_x^{z_1}\{T_{(z_1, z_2)} < \infty\} = 0$.

(d) The proof of this part follows along the lines of that for part (b). \square

Note that, e.g., part (c) says that $G(x, \cdot)$ is constant on $[z_1, z_2)$ iff the project state never hits (z_1, z_2) under the z_1 -policy starting from x . Parts (b, d) ensure that $F(x, \cdot)$ is of the form $F(x, \cdot) = \sum_{i \in N(x)} F_i(x) 1_{S_i(x)}(\cdot)$.

Lemma 28. *Assumption 3 implies Assumption 2.*

Proof. The result follows immediately from Lemma 27(d). \square

Verification of Assumption 3 for a given model may be a nontrivial task, as it requires knowledge of $G(x, \cdot)$. Instead, we will draw on Lemma 27 to show that it suffices to verify the following assumption, which only involves analyzing project state trajectories under threshold policies.

Assumption 4. *For each x there is a countable set $D(x)$ giving the endpoints of a partition of \mathbb{R} into left-semiclosed intervals or singletons, such that $\mathbb{P}_x^z\{\{X_t\}_{t=0}^\infty \subseteq D(x)\} = 1$ for any z .*

Note that Assumption 4 ensures that, starting from $X_0 = x$, the project state trajectory $\{X_t\}_{t=0}^\infty$ lies in $D(x)$ under *any* threshold policy.

Lemma 29. *Assumption 4 implies Assumption 3.*

Proof. The result follows immediately from Lemma 27(c), using the points in $D(x)$ to define the endpoints of the partition $\{S_i(x) : i \in N(x)\}$. \square

We thus obtain the following result.

Proposition 6. *Under (PCLI1, PCLI2) and Assumption 3 or 4, (PCLI3) holds.*

Proof. The result follows directly from Lemmas 28 and 29 and Proposition 5. \square

11 Metrics and index computation.

This section gives recursions for approximately computing the project metrics and index, in cases they cannot be evaluated in closed form.

For a state x , threshold z and $k = 0, 1, \dots$, consider the k -horizon metrics

$$F_k(x, z) \triangleq \mathbb{E}_x^z \left[\sum_{t=0}^k \beta^t r(X_t, A_t) \right] \quad \text{and} \quad G_k(x, z) \triangleq \mathbb{E}_x^z \left[\sum_{t=0}^k \beta^t c(X_t, A_t) \right]. \quad (69)$$

The function sequences $\{F_k(\cdot, z)\}_{k=0}^\infty$ and $\{G_k(\cdot, z)\}_{k=0}^\infty$ are determined by the following *value iteration* recursions: $F_0(x, z) \triangleq r(x, 1_{\{x > z\}})$, $G_0(x, z) \triangleq c(x, 1_{\{x > z\}})$ and, for $k = 0, 1, \dots$,

$$F_{k+1}(x, z) \triangleq r(x, 1_{\{x > z\}}) + \beta \int F_k(y, z) \kappa^{1_{\{x > z\}}}(x, dy), \quad (70)$$

$$G_{k+1}(x, z) \triangleq c(x, 1_{\{x > z\}}) + \beta \int G_k(y, z) \kappa^{1_{\{x > z\}}}(x, dy). \quad (71)$$

Consider further the k -horizon marginal metrics $f_k(x, z)$ and $g_k(x, z)$, given by $f_0(x, z) = \Delta_{a=0}^{a=1} r(x, a)$, $g_0(x, z) = \Delta_{a=0}^{a=1} c(x, a)$ and, for $k \geq 1$,

$$f_k(x, z) = \Delta_{a=0}^{a=1} \{r(x, a) + \beta \int F_{k-1}(y, z) \kappa^a(x, dy)\}, \quad (72)$$

$$g_k(x, z) = \Delta_{a=0}^{a=1} \{c(x, a) + \beta \int G_{k-1}(y, z) \kappa^a(x, dy)\}. \quad (73)$$

Finally, consider the k -horizon MP metric $m_k(x, z) \triangleq f_k(x, z)/g_k(x, z)$ and the k -horizon MP index $m_k(x) \triangleq m_k(x, x)$, which are defined when $g_k(x, z) \neq 0$ and $g_k(x, x) \neq 0$, respectively.

The next result shows, in parts (a, c), that $F_k(\cdot, z)$, $G_k(\cdot, z)$, $f_k(\cdot, z)$ and $g_k(\cdot, z)$ converge linearly with rate γ in the w -norm $\|\cdot\|_w$ (see (6) and Remark 2(ii)). Parts (b, d) bound their w -norms. Note that M , γ and w are as in Assumption 1(ii), and M_γ as in (8).

Lemma 30. For any threshold z and integer $k \geq 0$,

- (a) $\|F_k(\cdot, z) - F(\cdot, z)\|_w \leq M_\gamma \gamma^k$ and $\|G_k(\cdot, z) - G(\cdot, z)\|_w \leq M_\gamma \gamma^k$;
- (b) $\|F_k(\cdot, z)\|_w \leq M_\gamma$ and $\|G_k(\cdot, z)\|_w \leq M_\gamma$;
- (c) $\|f_k(\cdot, z) - f(\cdot, z)\|_w \leq 2M_\gamma \gamma^k$ and $\|g_k(\cdot, z) - g(\cdot, z)\|_w \leq 2M_\gamma \gamma^k$;
- (d) $\|f_k(\cdot, z)\|_w \leq 2M_\gamma$ and $\|g_k(\cdot, z)\|_w \leq 2M_\gamma$;

Proof. We only consider reward metrics, as the results for resource metrics follow similarly.

(a) (70) is a contraction with modulus γ on $\mathbb{B}_w(\mathbf{X})$ (cf. Hernández-Lerma and Lasserre (1999, remark 8.3.10)), so (cf. Remark 2(iv)) Banach's fixed point theorem yields the result.

(b) We first show by induction on k that $|F_k|(x, z) \leq C_k w(x)$, where $C_0 \triangleq 0$ and $C_k \triangleq M(1 + \dots + \gamma^{k-1})$ for $k \geq 1$, so $C_{k+1} = M + \gamma C_k$. For $k = 0$, such a result trivially holds. Assume now $|F_k|(x, z) \leq C_k w(x)$ for some $k \geq 0$. Then, from Assumption 1 and (70) we obtain

$$\begin{aligned} |F_{k+1}|(x, z) &\leq |r|(x, 1_{\{x > z\}}) + \beta \int |F_k|(y, z) \kappa^{1_{\{x > z\}}}(x, dy) \leq Mw(x) + \beta \int C_k w(y) \kappa^{1_{\{x > z\}}}(x, dy) \\ &\leq Mw(x) + \gamma C_k w(x) = C_{k+1} w(x), \end{aligned}$$

which proves the induction step. Since $C_k < M_\gamma$, it follows that $|F_k|(x, z) \leq M_\gamma w(x)$.

(c) Using (23), (72), part (a) and Assumption 1(ii.b), we obtain

$$\begin{aligned} |f_k(x, z) - f(x, z)| &\leq \beta \int |F_{k-1}(y, z) - F(y, z)| \kappa^1(x, dy) + \beta \int |F_{k-1}(y, z) - F(y, z)| \kappa^0(x, dy) \\ &\leq \beta M_\gamma \gamma^{k-1} \int w(y) \kappa^1(x, dy) + \beta M_\gamma \gamma^{k-1} \int w(y) \kappa^0(x, dy) \leq 2M_\gamma \gamma^k w(x). \end{aligned}$$

(d) This part follows along the same lines as part (b). □

Letting $\underline{g}(x, z) \triangleq \inf_k g_k(x, z)$, consider the following strong version of condition (PCLII):

$$\underline{g}(x, z) > 0 \quad \text{for any state } x \text{ and threshold } z. \tag{74}$$

Since $g_k(x, z)$ converges to $g(x, z)$ (see Lemma 30(c)), condition (74) implies (PCLII).

The following result shows that the finite-horizon MP metric $m_k(x, z)$ and MP index $m_k(x)$ converge at least linearly with rate γ to $m(x, z)$ and $m(x)$, respectively.

Proposition 7. Let (74) hold. Then, for any state x , threshold z and integer $k \geq 0$,

- (a) $|m_k(x, z) - m(x, z)| \leq 2M_\gamma \gamma^k w(x)(1 + |m|(x, z))/\underline{g}(x, z)$;
- (b) $|m_k(x) - m(x)| \leq 2M_\gamma \gamma^k w(x)(1 + |m|(x))/\underline{g}(x, x)$.

Proof. (a) From

$$m_k(x, z) - m(x, z) = \frac{f_k(x, z) - f(x, z)}{g_k(x, z)} - \frac{g_k(x, z) - g(x, z)}{g_k(x, z)} m(x, z),$$

Lemma 30(c) and $g_k(x, z) \geq \underline{g}(x, z) > 0$, we obtain the stated inequality.

(b) This part follows by setting $z = x$ in part (a). □

12 Examples.

This section illustrates application of the proposed approach to two models.

12.1 Optimal web crawling.

This is the model analyzed in [Avrachenkov and Borkar \(2018\)](#) under the average-reward criterion, which we analyze below under the discounted criterion. The project has reward and resource consumption $r(x, a) \triangleq xa$ and $c(x, a) \triangleq Ca$ with $C > 0$, and the following dynamics. When $X_t = x$ and $A_t = 0$, $X_{t+1} = \ell + \alpha x$, for certain parameters $0 < \alpha < 1$ and $\ell = (1 - \alpha)b$ with $b > 0$; if $A_t = 1$, $X_{t+1} = \ell$. The state space is $\mathsf{X} \triangleq [\ell, u]$, with $u \triangleq \ell/(1 - \alpha)$.

The evaluation equations for the reward and resource metrics under the z -threshold policy are

$$F(x, z) = \begin{cases} \beta F(\ell + \alpha x, z) & \text{if } x \leq z \\ x + \beta F(\ell, z) & \text{if } x > z \end{cases} \quad \text{and} \quad G(x, z) = \begin{cases} \beta G(\ell + \alpha x, z) & \text{if } x \leq z \\ C + \beta G(\ell, z) & \text{if } x > z. \end{cases}$$

To solve such equations in closed form, let $h(x) \triangleq \ell + \alpha x$, and recursively define $h_0(x) \triangleq x$ and $h_t(x) \triangleq h(h_{t-1}(x))$ for $t \geq 1$, so $h_t(x) = u - (u - x)\alpha^t$. Let $\tau(x, z) \triangleq \min\{t \geq 0: h_t(x) > z\}$, with $\tau(x, z) \triangleq \infty$ if $h_t(x) \leq z$ for all t . Through elementary arguments we obtain

$$F(x, z) = \begin{cases} \beta^{\tau(x, z)}(h_{\tau(x, z)}(x) + \beta F(\ell, z)) & \text{if } x \leq z \\ x + \beta F(\ell, z) & \text{if } x > z \end{cases} \quad \text{with} \quad F(\ell, z) = \frac{\beta^{\tau(\ell, z)} h_{\tau(\ell, z)}(\ell)}{1 - \beta^{\tau(\ell, z)+1}},$$

$$G(x, z) = \begin{cases} \beta^{\tau(x, z)}(C + \beta G(\ell, z)) & \text{if } x \leq z \\ C + \beta G(\ell, z) & \text{if } x > z \end{cases} \quad \text{with} \quad G(\ell, z) = \frac{\beta^{\tau(\ell, z)} C}{1 - \beta^{\tau(\ell, z)+1}}.$$

Note that, since the u -threshold policy is never active, we have $F(x, u) \equiv 0$ and $G(x, u) \equiv 0$.

Further, the marginal metrics are given by $f(x, u) = x$, $g(x, u) = C > 0$, and

$$f(x, z) = x + \beta F(\ell, z) - \beta F(h(x), z) = x + \beta F(\ell, z) - \beta^{\tau(x, z)}(h_{\tau(x, z)}(x) + \beta F(\ell, z)),$$

$$g(x, z) = C + \beta G(\ell, z) - \beta G(h(x), z) = C + \beta G(\ell, z) - \beta^{\tau(x, z)}(C + \beta G(\ell, z)) \\ = (1 - \beta^{\tau(x, z)})(C + \beta G(\ell, z)) > 0.$$

As for the MP index $m(\cdot)$, since $\tau(x, x) = 1$ for $\ell \leq x < u$, we have, for such x and $t = 1, 2, \dots$,

$$m(x) \triangleq \frac{f(x, x)}{g(x, x)} = \frac{x - \beta h(x) + \beta(1 - \beta)F(\ell, x)}{(1 - \beta)(C + \beta G(\ell, x))} = \frac{x - \beta h(x) + \beta(1 - \beta)F_t(\ell)}{(1 - \beta)(C + \beta G_t(\ell))}, \quad h_{t-1}(\ell) \leq x < h_t(\ell),$$

with $F_t(\ell) \triangleq \beta^t h_t(\ell)/(1 - \beta^{t+1})$ and $G_t(\ell) \triangleq \beta^t C/(1 - \beta^{t+1})$. Further, $m(u) \triangleq f(u, u)/g(u, u) = u/C$.

Proposition 8. *The above optimal web crawling project model is PCL-indexable.*

Proof. It is shown above that $g(x, z) > 0$, hence condition (PCLI1) holds. The MP index $m(\cdot)$ is piecewise affine right-continuous, increasing within each piece, since

$$\frac{d}{dx} m(x) = \frac{1 - \alpha\beta}{(1 - \beta)(C + \beta G_t(\ell))} > 0, \quad h_{t-1}(\ell) < x < h_t(\ell), \quad t = 1, 2, \dots$$

As for continuity of $m(\cdot)$, we need to check that, for $t = 1, 2, \dots$, $m(h_t(\ell)^-) = m(h_t(\ell))$, i.e.,

$$\frac{h_t(\ell) - \beta h_{t+1}(\ell) + \beta(1 - \beta)F_t(\ell)}{(1 - \beta)(C + \beta G_t(\ell))} = \frac{h_t(\ell) - \beta h_{t+1}(\ell) + \beta(1 - \beta)F_{t+1}(\ell)}{(1 - \beta)(C + \beta G_{t+1}(\ell))},$$

which is straightforward to verify algebraically. We further need to check that $m(h_t(\ell)) \rightarrow m(u)$ as $t \rightarrow \infty$, which also follows immediately. This establishes condition (PCLI2).

For checking condition (PCLI3), we apply the results in Section 10.2, since the resource metric $G(x, z)$ is piecewise constant in the threshold z for fixed x , as it is immediate that, for $z < x$,

$$G(x, z) = \begin{cases} \frac{\beta^{t+1}}{1 - \beta^{t+1}}C, & x > z, h_{t-1}(\ell) \leq z < h_t(\ell), t \geq 1 \\ \frac{\beta^{s+t+1}}{1 - \beta^{t+1}}C, & x \leq z, h_{s-1}(x) \leq z < h_s(x), h_{t-1}(\ell) \leq z < h_t(\ell), 1 \leq s \leq t. \end{cases}$$

Thus, Assumption 3 holds, and hence Proposition 6 implies that (PCLI3) holds. \square

Hence, Theorem 1 gives that the model is threshold-indexable with index $m(\cdot)$. Note further that the index $m(\cdot)$ converges as $\beta \rightarrow 1$ to the limit

$$\bar{m}(x) = \frac{(t+1)(x - h(x)) + h_t(\ell)}{C}, \quad h_{t-1}(\ell) \leq x < h_t(\ell), t \geq 1,$$

which is precisely the average-reward Whittle index derived in Avrachenkov and Borkar (2018).

12.2 Optimal dynamic transmission over a noisy channel.

This is the model in Liu and Zhao (2010), which motivated the proposal of a previous incomplete version of the PCL-indexability conditions for real-state projects in Niño-Mora (2008). This section completes the analysis in Niño-Mora (2008), and further proves satisfaction of (PCLI3), which was not considered there.

A user dynamically attempts to send packets over a noisy channel. In each period t , the channel can be in either state 1 (good), in which an transmission attempt succeeds, or 0, in which it fails. State transitions are Markovian, with probabilities p and q for transitions $1 \rightarrow 0$ and $0 \rightarrow 1$, satisfying $\rho \triangleq 1 - p - q > 0$. The user cannot observe the channel state, and bases decisions on the *belief state* $X_t \in \mathbf{X} \triangleq [0, 1]$, the posterior probability that the channel state is 1. If $X_t = x$ and action $A_t = 1$ is taken (attempt to transmit) then X_{t+1} takes the values $q + \rho$ and q with probabilities x and $1 - x$. If action $A_t = 0$ is taken, $X_{t+1} = q + \rho x$. The reward and resource functions are $r(x, a) \triangleq ax$ and $c(x, a) \triangleq a$. The cost per transmission attempt is λ .

The λ -price problem P_λ in (4) is to find a transmission policy maximizing the expected total β -discounted value. To apply the present framework we need to evaluate the project metrics. Under the z -policy, the reward and resource metrics are determined by the functional equations

$$\begin{aligned} F(x, z) &= \beta F(q + \rho x, z)1_{\{x \leq z\}} + (x + \beta x F(q + \rho, z) + \beta(1 - x)F(q, z))1_{\{x > z\}}, \\ G(x, z) &= \beta G(q + \rho x, z)1_{\{x \leq z\}} + (1 + \beta x G(q + \rho, z) + \beta(1 - x)G(q, z))1_{\{x > z\}}, \end{aligned}$$

and the marginal metrics are given by

$$\begin{aligned} f(x, z) &= x - \beta(F(q + \rho x, z) - xF(q + \rho, z) - (1 - x)F(q, z)), \\ g(x, z) &= 1 - \beta(G(q + \rho x, z) - xG(q + \rho, z) - (1 - x)G(q, z)). \end{aligned}$$

Provided $g(x, x) \neq 0$, the MP index is

$$m(x) = \frac{x - \beta(F(q + \rho x, x) - xF(q + \rho, x) - (1 - x)F(q, x))}{1 - \beta(G(q + \rho x, x) - xG(q + \rho, x) - (1 - x)G(q, x))}.$$

The following analysis uses the map $h(x) \triangleq q + \rho x$ and iterates $h_t(x)$, with $h_0(x) = x$ and $h_t(x) = h(h_{t-1}(x))$ for $t \geq 1$, so $h_t(x) = h_\infty - (h_\infty - x)\rho^t \rightarrow h_\infty \triangleq q/(1 - \rho)$ as $t \rightarrow \infty$. It further uses the *backward iterates* $h_{-t}(z) = h_\infty - (h_\infty - z)\rho^{-t}$ for $t \geq 1$. Four cases need be considered.

Case I: $z < q$. In this case the belief state is above threshold at times $t \geq 1$. The metrics are

$$\begin{aligned} F(x, z) &= \frac{\beta(q + (1 - \beta)\rho x) + (1 - \beta)(1 - \beta\rho)x1_{\{x > z\}}}{(1 - \beta)(1 - \beta\rho)}, & f(x, z) &= x \\ G(x, z) &= \frac{\beta + (1 - \beta)1_{\{x > z\}}}{1 - \beta}, & g(x, z) &= 1. \end{aligned}$$

Thus, condition (PCLI1) holds for any x when $z < q$, and the MP index is $m(x) = x$ for $x < q$.

Case II: $q \leq z < h_\infty$. Consider the subcase $h_{t-1}(q) \leq z < h_t(q)$ for $t \geq 1$. The metrics can be given in closed form in terms of $G_t(q) \triangleq G(q, h_{t-1}(q))$, $G_t(q + \rho) \triangleq G(q + \rho, h_{t-1}(q))$, $F_t(q) \triangleq F(q, h_{t-1}(q))$ and $F_t(q + \rho) \triangleq F(q + \rho, h_{t-1}(q))$, which are readily evaluated. Thus, for $x > z$,

$$\begin{aligned} F(x, z) &= x + \beta(xF_t(q + \rho) + (1 - x)F_t(q)) \\ f(x, z) &= x - \beta(h(x) + \beta h(x)F_t(q + \rho) + \beta(1 - h(x))F_t(q) - xF_t(q + \rho) - (1 - x)F_t(q)) \\ G(x, z) &= 1 + \beta(xG_t(q + \rho) + (1 - x)G_t(q)) \\ g(x, z) &= 1 - \beta(1 + \beta h(x)G_t(q + \rho) + \beta(1 - h(x))G_t(q) - xG_t(q + \rho) - (1 - x)G_t(q)). \end{aligned}$$

For $x \leq z$, letting $s \leq t + 1$ be such that $h_{s-1}(x) \leq z < h_s(x)$, we have

$$\begin{aligned} F(x, z) &= \beta^s(h_s(x) + \beta h_s(x)F_t(q + \rho) + \beta(1 - h_s(x))F_t(q)) \\ f(x, z) &= x - \beta(\beta^{s-1}[h_s(x) + \beta h_s(x)F_t(q + \rho) + \beta(1 - h_s(x))F_t(q)] - xF_t(q + \rho) - (1 - x)F_t(q)) \\ G(x, z) &= \beta^s(1 + \beta h_s(x)G_t(q + \rho) + \beta(1 - h_s(x))G_t(q)) \\ g(x, z) &= 1 - \beta(\beta^{s-1}[1 + \beta h_s(x)G_t(q + \rho) + \beta(1 - h_s(x))G_t(q)] - xG_t(q + \rho) - (1 - x)G_t(q)). \end{aligned}$$

Hence, for $q \leq x < h_\infty$, and provided $g(x, x) > 0$, the MP index is

$$m(x) = \frac{x - \beta(h(x) + \beta h(x)F_t(q + \rho) + \beta(1 - h(x))F_t(q) - xF_t(q + \rho) - (1 - x)F_t(q))}{1 - \beta(1 + \beta h(x)G_t(q + \rho) + \beta(1 - h(x))G_t(q) - xG_t(q + \rho) - (1 - x)G_t(q))}.$$

Case III: $h_\infty \leq z < q + \rho$. In this case, the project metrics are given by

$$\begin{aligned} F(x, z) &= \frac{x}{1 - \beta(q + \rho)} 1_{\{x > z\}} \\ f(x, z) &= \frac{x}{1 - \beta(q + \rho)} 1_{\{h(x) \leq z\}} + \frac{(1 - \beta\rho)x - \beta q}{1 - \beta(q + \rho)} 1_{\{h(x) > z\}} \\ G(x, z) &= \frac{1 - \beta(q + \rho - x)}{1 - \beta(q + \rho)} 1_{\{x > z\}} \\ g(x, z) &= \frac{1 - \beta(q + \rho - x)}{1 - \beta(q + \rho)} 1_{\{h(x) \leq z\}} + \left(1 - \beta + \beta \frac{(1 - \beta\rho)x - \beta q}{1 - \beta(q + \rho)}\right) 1_{\{h(x) > z\}}. \end{aligned}$$

Hence, for $h_\infty \leq x < q + \rho$, as $h(x) \leq x$, the MP index is $m(x) = x/(1 - \beta(q + \rho - x))$.

Case IV: $z \geq q + \rho$. Finally, in this case we have $F(x, z) = x 1_{\{x > z\}}$, $f(x, z) = x$, $G(x, z) = 1_{\{x > z\}}$, and $g(x, z) = 1$, and therefore, for $x \geq q + \rho$, the MP index is $m(x) = x$.

Proposition 9. *The above optimal channel transmission model is PCL-indexable.*

Proof. Consider (PCL1). The stronger result $g(x, z) \geq 1 - \beta$ holds, which is trivial in cases I, III and IV. In case II, for $h_{t-1}(q) < z < h_t(q)$, $g(\cdot, z)$ is piecewise linear *càglàd* (left-continuous with right limits) with $t + 1$ increasing pieces, which follows from $G_t(q + \rho) > G_t(q)$ and

$$\frac{\partial}{\partial x} g(x, z) = \begin{cases} \beta(1 - (\beta\rho)^{t+1})(G_t(q + \rho) - G_t(q)), & \text{for } 0 < x < h_{-t}(z) \\ \beta(1 - (\beta\rho)^s)(G_t(q + \rho) - G_t(q)), & \text{for } h_{-s}(z) < x < h_{-s+1}(z), s = 2, \dots, t \\ \beta(1 - \beta\rho)(G_t(q + \rho) - G_t(q)), & \text{for } h_{-1}(z) < x < 1. \end{cases}$$

Hence, $g(x, z) > g(0^+, z)$ for $0 < x \leq h_{-t}(z)$, $g(x, z) > g(h_{-s}(z)^+, z)$ for $h_{-s}(z) < x \leq h_{-s+1}(z)$ with $2 \leq s \leq t$, and $g(x, z) > g(h_{-1}(z)^+, z)$ for $h_{-1}(z) < x \leq 1$. Now, it can be verified that $g(0^+, z) = 1$, and, using that $(\partial/\partial z)g(h_{-s}(z)^+, z) = \beta(\rho^{-s} - \beta^s)(G_t(q + \rho) - G_t(q)) > 0$, $g(h_{-s}(z)^+, z) > g(h_{t-s-1}(q)^+, h_{t-1}(q))$ for $s = 1, \dots, t$. One can then show through algebraic and calculus arguments that $g(h_{t-s-1}(q)^+, h_{t-1}(q)) \geq 1 - \beta$ for each such s , and thus establish satisfaction of (PCL1).

As for condition (PCLI2), it is also easily verified in cases I, III and IV, viz., for $x \in [0, q) \cup [h_\infty, 1]$. Its verification in case II involves algebraic and calculus arguments.

Regarding (PCLI3), we use the sufficient conditions given in Section 10.2. Thus, taking

$$D(x) \triangleq \{h_t(x) : t \geq 0\} \cup \{h_t(q) : t \geq 0\} \cup \{h_\infty, q, q + \rho\},$$

Assumption 4 holds, whence Lemma 29 and Proposition 6 imply (PCLI3). \square

Thus, Theorem 1 gives that the model is threshold-indexable with Whittle index $m(\cdot)$.

13 Concluding remarks.

This paper has presented PCL-based sufficient conditions for indexability along with optimality of threshold policies, for general restless bandits in a real-state discrete-time discounted setting, demonstrating their validity. This extends the approach previously developed by the author for discrete-state bandits. As for the value of the proposed approach relative to prevailing approaches based on first proving optimality of threshold policies (in a strong sense; see Section 2.4 and Section 4), the main example demonstrating it is the recent work in [Dance and Silander \(2019\)](#), which deploys the sufficient conditions presented here to establish threshold-indexability of an important model that has not yielded to prevailing approaches, the Kalman filter target tracking bandit. Furthermore, the proposed approach applies to the larger class of threshold-indexable projects, whereas the conventional approach applies to the more restricted class of strongly threshold-indexable projects. We further stress that two of the three sufficient conditions for threshold-indexability, viz. (PCLI2, PCLI3), are necessary (see Section 3).

Among the issues raised by this paper we highlight that further work is required to facilitate application of the proposed conditions to relevant model classes, by providing simpler means of checking their satisfaction than direct verification. Steps in that direction are already taken in this paper with condition (PCLI3). See Section 10. It would also be desirable to identify properties of model primitives implying PCL-indexability. Extending the results to the average-reward criterion is another goal for further research.

References

- E. Altman and S. Stidham, Jr. Optimality of monotonic policies for two-action Markovian decision processes, with applications to control of queues with delayed information. *Queueing Syst.*, 21:267–291, 1995.
- P. S. Ansell, K. D. Glazebrook, J. Niño Mora, and M. O’Keeffe. Whittle’s index policy for a multi-class queueing system with convex holding costs. *Math. Meth. Oper. Res.*, 57:21–39, 2003.
- K. E. Avrachenkov and V. S. Borkar. Whittle index policy for crawling ephemeral content. *IEEE Trans. Control Netw. Syst.*, 5:446–455, 2018.
- F. Caro and O. S. Yoo. Indexability of bandit problems with response delays. *Probab. Eng. Inform. Sci.*, 24: 349–374, 2010.
- M. Carter and B. van Brunt. *The Lebesgue–Stieltjes Integral: A Practical Introduction*. Springer, New York, NY, 2000.
- E. G. Coffman, Jr. and I. Mitrani. A characterization of waiting time performance realizable by single-server queues. *Oper. Res.*, 28:810–821, 1980.
- W. Cowan and M. N. Katehakis. Multi-armed bandits under general depreciation and commitment. *Probab. Eng. Inform. Sci.*, 29:51–76, 2015.
- C. R. Dance and T. Silander. When are Kalman-filter restless bandits indexable? In C. Cortes, D. D. Lee, M. Sugiyama, and R. Garnett, editors, *Proceedings of the 2015 Conference on Neural Information Processing Systems (NIPS 2015)*, Montreal, Canada, pages 1711–1719. MIT Press, Cambridge, MA, 2015.

- C. R. Dance and T. Silander. Optimal policies for observing time series and related restless bandit problems. *J. Mach. Learn. Res.*, 20, 2019. Article Number 35.
- J. L. Doob. *Measure Theory*. Springer, New York, NY, 1994.
- N. Falkner and G. Teschl. On the substitution rule for Lebesgue–Stieltjes integrals. *Expo. Math.*, 30:412–418, 2012.
- J. C. Gittins. Bandit processes and dynamic allocation indices. *J. Roy. Statist. Soc. Ser. B*, 41:148–177, 1979.
- W. -R. Heilmann. Linear programming of dynamic programs with unbounded rewards. *Meth. Oper. Res.*, 24: 94–105, 1977.
- W. -R. Heilmann. Solving a general discounted dynamic program by linear programming. *Probab. Theory Rel. Fields*, 48:339–346, 1979.
- O. Hernández-Lerma and J. B. Lasserre. *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. Springer, New York, NY, 1996.
- O. Hernández-Lerma and J. B. Lasserre. *Further Topics on Discrete-Time Markov Control Processes*. Springer, New York, NY, 1999.
- D. P. Heyman and M. J. Sobel. *Stochastic Models in Operations Research. Vol. II: Stochastic Optimization*. McGraw-Hill, New York, NY, 1984.
- M. N. Katehakis and A. F. Veinott, Jr. The multi-armed bandit problem: decomposition and computation. *Math. Oper. Res.*, 12:262–268, 1987.
- V. Krishnamurthy. *Partially Observed Markov Decision Processes: From Filtering to Controlled Sensing*. Cambridge University Press, Cambridge, UK, 2016.
- J. Kuhn and Y. Nazarathy. Wireless channel selection with restless bandits. In R. J. Boucherie and N. M. VanDijk, editors, *Markov Decision Processes in Practice*, chapter 7, pages 463–485. Springer, Cham, Switzerland, 2017.
- B. F. La Scala and B. Moran. Optimal target tracking with restless bandits. *Digit. Signal Process.*, 16:479–487, 2006.
- J. Le Ny, M. Dahleh, and E. Feron. Multi-UAV dynamic routing with partial observations using restless bandit allocation indices. In *Proceedings of the 2008 American Control Conference, Seattle, WA*, pages 4220–4225, New York, NY, 2008. IEEE.
- J. Le Ny, E. Feron, and M. Dahleh. Scheduling continuous-time Kalman filters. *IEEE Trans. Autom. Control*, 56:1381–1394, 2011.
- S. A. Lippman. On dynamic programming with unbounded rewards. *Management Sci.*, 21:1225–1233, 1975.
- K. Liu and Q. Zhao. Indexability of restless bandit problems and optimality of Whittle index for dynamic multichannel access. *IEEE Trans. Inform. Theory*, 56:5547–5567, 2010.
- R. Meshram, D. Manjunath, and A. Gopalan. On the Whittle index for restless multi-armed hidden Markov bandits. *IEEE Trans. Automat. Control*, 63:3046–3053, 2018.
- J. Niño-Mora. Restless bandits, partial conservation laws and indexability. *Adv. Appl. Probab.*, 33:76–98, 2001.
- J. Niño-Mora. Dynamic allocation indices for restless projects and queueing admission control: A polyhedral approach. *Math. Program.*, 93:361–413, 2002.
- J. Niño-Mora. Restless bandit marginal productivity indices, diminishing returns and optimal control of make-to-order/make-to-stock M/G/1 queues. *Math. Oper. Res.*, 31:50–84, 2006.

- J. Niño-Mora. Dynamic priority allocation via restless bandit marginal productivity indices. *TOP*, 15:161–198, 2007.
- J. Niño-Mora. An index policy for dynamic fading-channel allocation to heterogeneous mobile users with partial observations. In *Proceedings of the 4th Conference on Next Generation Internet Networks (NGI 2008) (Cracow, Poland)*, pages 231–238, New York, NY, 2008. IEEE.
- J. Niño-Mora. A restless bandit marginal productivity index for opportunistic spectrum access with sensing errors. In *Proceedings of the 3rd Euro-NF Conference on Network Control and Optimization (NETCOOP 2009), Eindhoven, Netherlands*, volume 5894 of *Lecture Notes in Computer Science*, pages 60–74, Berlin, Germany, 2009. Springer.
- J. Niño-Mora. Conservation laws and related applications. In J. J. Cochran, L. A. Cox, Jr., P. Keskinocak, J. P. Kharoufeh, and J. C. Smith, editors, *Wiley Encyclopedia of Operations Research and Management Science*. Wiley, New York, NY, 2011. DOI: 10.1002/9780470400531.eorms0186.
- J. Niño-Mora. Admission and routing of soft real-time jobs to multiclusters: Design and comparison of index policies. *Comput. Oper. Res.*, 39:3431–3444, 2012.
- J. Niño-Mora. A dynamic page-refresh index policy for web crawlers. In *Proceedings of the 21st International Conference on Analytical and Stochastic Modelling Techniques and Applications (ASMTA 2014), Budapest, Hungary*, volume 8499 of *Lecture Notes in Computer Science*, pages 46–60, Berlin, Germany, 2014. Springer.
- J. Niño-Mora. A verification theorem for indexability of discrete time real state discounted restless bandits. *arXiv:1512.04403v1*, 2015.
- J. Niño-Mora. Whittle’s index policy for multi-target tracking with jamming and nondetections. In *Proceedings of the 23rd International Conference on Analytical and Stochastic Modelling Techniques and Applications (ASMTA 2016), Cardiff, Wales*, volume 9845 of *Lecture Notes in Computer Science*, pages 210–222, Berlin, Germany, 2016. Springer.
- J. Niño-Mora. Resource allocation and routing in parallel multi-server queues with abandonments for cloud profit maximization. *Comput. Oper. Res.*, 103:221–236, 2019.
- J. Niño-Mora and S. S. Villar. Multitarget tracking via restless bandit marginal productivity indices and Kalman filter in discrete time. In *Proceedings of the Joint 48th IEEE Conference on Decision and Control (CDC) / 28th Chinese Control Conference (CCC), Shanghai, China*, pages 2905–2910, New York, NY, 2009. IEEE.
- J. Niño-Mora and S. S. Villar. Sensor scheduling for hunting elusive hiding targets via Whittle’s restless bandit index policy. In *Proceedings of the 5th International Conference on Network Games, Control and Optimization (NETGCOOP 2011), Paris France*, pages 124–131, New York, NY, 2011. IEEE.
- C. Niculescu and L. E. Persson. *Convex Functions and Their Applications: A Contemporary Approach*. Springer, New York, NY, 2006.
- W. Z. Ouyang, S. Murugesan, A. Eryilmaz, and N. B. Shroff. Exploiting channel memory for joint estimation and scheduling in downlink networks — a Whittle’s indexability analysis. *IEEE Trans. Inform. Theory*, 61: 1702–1719, 2015.
- W. Z. Ouyang, A. Eryilmaz, and N. B. Shroff. Downlink scheduling over Markovian fading channels. *IEEE/ACM Trans. Netw.*, 24:1801–1812, 2016.
- C. H. Papadimitriou and J. N. Tsitsiklis. The complexity of optimal queuing network control. *Math. Oper. Res.*, 24:293–305, 1999.
- M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley, New York, NY, 1994.

- J. G. Shanthikumar and D. D. Yao. Multiclass queueing systems: Polymatroidal structure and optimal scheduling control. *Oper. Res.*, 40:S293–S299, 1992.
- J. A. Taylor and J. L. Mathieu. Index policies for demand response. *IEEE Trans. Power Syst.*, 29:1287–1295, 2014.
- D. M. Topkis. *Supermodularity and Complementarity*. Princeton University Press, Princeton, NJ, 1998.
- M. H. Veatch and L. M. Wein. Scheduling a multiclass make-to-stock queue: index policies and hedging points. *Oper. Res.*, 44:634–647, 1996.
- I. M. Verloop. Asymptotically optimal priority policies for indexable and nonindexable restless bandits. *Ann. Appl. Probab.*, 26:1947–1995, 2016.
- S. Villar. *Restless Bandit Index Policies for Dynamic Sensor Scheduling Optimization*. PhD thesis, Carlos III University of Madrid, Spain, 2012.
- R. B. Washburn. Application of multi-armed bandits to sensor management. In *Foundations and Applications of Sensor Management*, chapter 7, pages 153–175. Springer, New York, NY, 2008.
- R. R. Weber and G. Weiss. On an index policy for restless bandits. *J. Appl. Probab.*, 27:637–648, 1990.
- J. Wessels. Markov programming by successive approximations with respect to weighted supremum norms. *J. Math. Anal. Appl.*, 58:326–335, 1977.
- P. Whittle. Multi-armed bandits and the Gittins index. *J. Roy. Statist. Soc. Ser. B*, 42:143–149, 1980.
- P. Whittle. Restless bandits: Activity allocation in a changing world. In J. Gani, editor, *A Celebration of Applied Probability*, volume 25A of *J. Appl. Probab.*, pages 287–298. Applied Probability Trust, Sheffield, UK, 1988.
- P. Whittle. *Optimal Control: Basics and Beyond*. Wiley, Chichester, UK, 1996.