Multitarget Tracking via Restless Bandit Marginal Productivity Indices and Kalman Filter in Discrete Time

José Niño-Mora and Sofía S. Villar

Abstract—This paper designs, evaluates, and tests a tractable priority-index policy for scheduling target updates in a discrete-time multitarget tracking model, which aims to be close to optimal relative to a discounted or average performance objective for tracking error variance and measurement costs. The policy is to be used by $M$ phased-array radars who coordinate to track the positions of $N$ targets moving according to independent scalar Gaussian–Markov linear dynamics, which allows use of the Kalman filter for track estimation. The paper exploits the natural problem formulation as a multiarmed restless bandit problem (MARBP) with real-state projects subject to deterministic dynamics by deploying Whittle’s (1988) index policy for the MARBP. The challenging issues of indexability (existence of the index) and index evaluation are resolved by applying a method recently introduced by the first author for the analysis of real-state restless bandits. Preliminary computational results are reported demonstrating the tractability of index evaluation and comparing the MP index policy against myopic policies advocated in previous work.

I. INTRODUCTION

A. Motivation and Background

In contrast to conventional track–while-scan radars, which track targets while the radar’s antenna mechanically rotates at a constant rate, modern phased-array radars are capable of electronically steering the radar’s beam to point toward desired directions. The operating flexibility of such an agile beam-steering capability raises the research challenge of tracking-performance optimization in radar surveillance, via appropriate design of the scheduling control policy adopted for dynamic prioritization of target track updates.

In this paper we seek to design a tractable and near-optimal policy for dynamic scheduling of track updates, in a highly idealized system model for dynamic tracking of a fixed number $N$ of well-separated moving targets, using $M$ radars which are controlled by a central coordinator. We ignore other relevant issues such as target detection, waveform selection, and control of the pulse repetition interval (PRI) durations. We refer the reader to [1] for a survey of the substantial literature in the area. For early work on the subject of track-update scheduling, see, e.g., [2], [3], [4].

The model we consider here is based on and extends that formulated in [5], in which targets and target track measurements follow scalar, linear Gauss–Markov dynamics, and target track error variances (TEVs) are updated via Kalman filter’s equations. The objective to optimize in [5] is the sum of the targets’ track error variances over a finite horizon of discrete time periods. The authors propose a greedy scheduling policy, which updates at each time a target of largest TEV, thus taking a target’s current TEV as its priority index. They further claim such a greedy-index policy to be optimal in for the case of two symmetric targets.

In [6], the authors extend the results in [5] on optimality of the greedy-index scheduling policy for tracking two symmetric targets to more general linear dynamical systems under the same finite-horizon total TEV performance objective. They further remark that such a problem falls within the framework of the multiarmed restless bandit problem (MARBP) introduced by Whittle in [7], although they do not use the indexation approach proposed there.

The MARBP is a powerful modeling framework which concerns the optimal dynamic scheduling of a collection of $N$ stochastic projects, $M$ of which can be engaged at a time. Each project (or bandit) is modeled as a binary-action (active or passive) Markov decision process (MDP). The goal is to find a scheduling policy that maximizes the expected total discounted (ETD) or the long-run expected time-average reward earned over an infinite horizon. In the special classic case where passive projects do not change state and $M = 1$, the optimal policy is given by the Gittins index [8] policy: an index $\lambda(n), s(n)$ is attached to each project $n$ as a function of its state $s(n)$, and then a project of largest index is engaged at each time. In [9], a Gittins index policy is proposed for multitarget tracking in a classic (nonrestless) multiarmed bandit model, where projects represent targets, assuming that a target’s motion from one PRI to the next is negligible. As pointed out in [6], however, such an assumption is typically inadequate, calling for use of the MARBP model where passive projects change state.

Indeed, Whittle used multitarget tracking as one of his motivating applications for introducing the MARBP, in the example of $M$ aircraft trying to track the positions of $N$ enemy submarines, where, as he put it, “the bandits are restless in the most literal sense.”

Although index policies are generally suboptimal for the MARBP, Whittle introduced in [7] a heuristic index policy based on a particular index for restless bandits, which emerges from a Lagrangian relaxation and decomposition approach that also yields a bound on the optimal problem value. The Whittle index, which has been extended in [10] into the more general concept of marginal productivity (MP) index—named after its economic interpretation—raises substantial research challenges as (i) indexability (i.e., existence
of the index) needs to be established for the model at hand; and (ii) the index needs to be evaluated in a tractable fashion.

The first author has developed a methodology for resolving such issues on restless bandit indexation in the discrete-state case, in a stream of work starting in [11]–[13], which is reviewed in [14]. More recently he has announced in [15] extensions to real-state restless bandits, which are the cornerstone of the present paper’s approach, and are also deployed in an opportunistic spectrum access model in [16].

B. Goals and contributions

This paper extends such a line of work by investigating an MARBP formulation of dynamic tracking of multiple asymmetric targets with scalar linear Gauss–Markov dynamics, which incorporates both tracking-error and measurement (energy) costs, the main goal being to obtain a tractable index policy that performs well based on restless bandit indexation.

The paper deploys the methodology for real-state restless bandit indexation announced in [15] to establish indexability and evaluate the MP index in an efficient fashion for the model of concern. Further, the paper presents some evidence that the resulting MP index policy outperforms conventional policies advocated in previous work, in instances with asymmetric targets.

C. Organization of the paper

The remainder of the paper is organized as follows. Section II describes the multitarget tracking model and the MARBP formulation. Section III discusses the restless bandit indexation methodology for real-state restless bandits introduced in [15] as it applies to the design of index policies for multitarget tracking. Section IV discusses how to deploy such a methodology in the present model to verify indexability and approximately evaluate the Whittle’s MP index. Section V reports the results of a preliminary computational study on index evaluation and on the relative quality of the Whittle’s MP index policy. No proofs are given due to space constraints. They will be included, along with extensive computational experiments, in the full journal version of this work.

II. MULTITARGET TRACKING AND RESTLESS BANDITS

A. Multitarget Tracking Kalman Filter Model

We consider a system consisting of $M$ radars labeled by $m \in M \triangleq \{1, \ldots, M\}$, which cooperate to track $N \geq M$ moving targets labeled by $n \in \mathbb{N} \triangleq \{1, \ldots, N\}$. All radars are synchronized to operate over time slots $t = 0, 1, \ldots$, where a time slot corresponds to a PRI.

The system is controlled by a central coordinator, who at each slot $t$ can update the tracks of at most $M$ targets by steering toward them the beams of as many radars to measure their positions. For simplicity we assume that targets move in one dimension. Let $x_{t}^{(n)}$ be the (unobservable) position of target $n$ in the real line $\mathbb{R}$ at the beginning of slot $t$. If a radar measures target $n$’s position in slot $t$, a noisy measurement $y_{t}^{(n)}$ is obtained. We assume that there are no clutter or false measurements, and that the probability of target detection is unity. Decisions on which target tracks to update at each time are formulated by binary action processes $a_{t}^{(n)} \in \{0, 1\}$, where $a_{t}^{(n)} = 1$ if target $n$ is measured in slot $t$.

The targets move over $\mathbb{R}$ following independent linear Gauss–Markov dynamics

$$x_{t}^{(n)} = F^{(n)} x_{t-1}^{(n)} + \omega^{(n)}_{t}, \quad t \geq 0,$$

where the position-noise process $\omega^{(n)}_{t}$ is an i.i.d. zero-mean Gaussian white noise with variance $q^{(n)}$, and $F^{(n)} > 0$.

At a slot $t$ in which target $n$ is measured, the measurement $y_{t}^{(n)}$ is generated by the linear Gauss–Markov dynamics

$$y_{t}^{(n)} = H^{(n)} x_{t}^{(n)} + \nu^{(n)}_{t},$$

which do not depend on the radar being used, where the measurement-noise process $\nu^{(n)}_{t}$ is an i.i.d. zero-mean Gaussian white noise with variance $r^{(n)}$, and $H^{(n)} > 0$.

Although our approach applies to arbitrary parameters $F^{(n)}$ and $H^{(n)}$, for simplicity of exposition we focus the subsequent discussion on the case $F^{(n)} = 1$ and $H^{(n)} = 1$.

If an initial estimate of the position and of the tracking error variance (TEV), denoted by $\hat{x}_{0}^{(n)}$ and $p_{0}^{(n)}$, respectively, are given for each target $n$, then the optimal minimum-variance predicted estimates are given by the Kalman filter. The TEV $p_{t}^{(n)}$, which describes the uncertainty in target $n$’s track at the beginning of slot $t$, is recursively updated by the Kalman equations

$$p_{t}^{(n)} = \begin{cases} q^{(n)} + p_{t-1}^{(n)}, & \text{if } a_{t}^{(n)} = 0 \\ q^{(n)} + p_{t-1}^{(n)} + \frac{1}{1 + q^{(n)}/r^{(n)} + p_{t-1}^{(n)}/r^{(n)}}, & \text{if } a_{t}^{(n)} = 1. \end{cases}$$

We shall take the state of each target $n$ to be its scaled TEV (STEV) $s_{t}^{(n)} \triangleq p_{t}^{(n)}/r^{(n)}$, which follows the dynamics

$$s_{t}^{(n)} = \begin{cases} F_{0}^{(n)}(s_{t-1}^{(n)}) + \phi^{0,1}_{t}^{(n)}(s_{t-1}^{(n)}), & \text{if } a_{t}^{(n)} = 0 \\ F_{1}^{(n)}(s_{t-1}^{(n)}) + \phi^{1,1}_{t}^{(n)}(s_{t-1}^{(n)}), & \text{if } a_{t}^{(n)} = 1, \end{cases}$$

where

$$\phi^{0,1}_{t}^{(n)}(s) = \theta^{(n)} + x, \quad \phi^{1,1}_{t}^{(n)}(s) = \frac{\theta^{(n)} + x}{1 + \theta^{(n)} + x}$$

and $\theta^{(n)} \triangleq q^{(n)}/r^{(n)}$ is the position to measurement noise variance ratio for target $n$. Such a STEV state moves over the state space $S \triangleq [0, \infty)$.

Note that, for any initial state $s_{0}^{(n)} = s$, the $t$th iterate of $\phi^{1,1}_{t}^{(n)}(s)$, $\phi_{t}^{1,1}(s)$, which is generated by $\phi_{0}^{1,1}(s) \triangleq s$ and $\phi_{t}^{1,1}(s) \triangleq \phi_{t-1}^{1,1}(s) \phi_{t-1}^{1,1}(s)$, converges to the limit

$$\lim_{t \to \infty} \phi_{t}^{1,1}(s) = \phi_{1}^{1,1}(s) = \frac{1}{2} \left( \sqrt{\theta^{(n)}(4 + \theta^{(n)})} - \theta^{(n)} \right),$$

which is the unique nonnegative root of $\phi_{1}^{1,1}(s) = s$ and is an attractive fixed point. Note further that $\phi_{1}^{1,1} \leq \theta^{(n)}$ iff $\theta^{(n)} \geq 1/2$, which will be the case if radar’s measurements on target $n$ are precise enough.

We assume henceforth that $\theta^{(n)} \geq 1/2$ for each target $n$. Note that in such a case the subset of states $S^{(n)} \triangleq [\phi_{1}^{1,1}, \infty)$ is absorbing for target $n$. 2906
To take actions $a_t^{(n)}$, the coordinator adopts a scheduling policy $\pi$, which is drawn from the class $\Pi$ of admissible policies that are nonanticipative (based on the history of states and actions) and measure at most $M$ targets per slot,

$$\sum_{n \in N} a_t^{(n)} \leq M, \quad t = 0, 1, 2, \ldots \tag{4}$$

We assume that a radar which updates the target $n$’s track in a time slot incurs a measurement cost $h^{(n)} \geq 0$, representing the cost of radar energy expended for the track’s update. Further, we take the tracking-error cost at slot $t$ to be $d^{(n)}(t+1) = d^{(n)}(t+1)$, where $d^{(n)} > 0$ is a constant that may differ by target. The flexibility furnished by the $d^{(n)}$ will be of use if the relative importance of tracking precision differs across targets. Hence, the one-slot cost incurred by taking action $a$ on target $n$ when it occupies STEV state $s$ is $C^{(n)}(s, a) \subseteq d^{(n)}(s) \beta^{(n)}(s) + h^{(n)} a$.

B. Multiarmed Restless Bandit Formulation

Consider the following dynamic optimization problems: (1) find a discount-optimal policy:

$$\min_{\pi \in \Pi} E^{n}_{s} \left[ \sum_{t=0}^{\infty} \sum_{n \in N} \beta^t C^{(n)}(s_t^{(n)}, a_t^{(n)}) \right], \tag{5}$$

which minimizes the expected total discounted (ETD) cost, where $0 < \beta < 1$ is the discount factor, $s_0 = s = (s^{(n)})$ is the initial joint STEV state, and $E^{n}_{s} \boldsymbol{[\cdot]}$ denotes expectation under policy $\pi$ conditional on $s_0 = s$; and (2) find an average-optimal policy,

$$\min_{\pi \in \Pi} \lim_{T \to \infty} \frac{1}{T} E^{n}_{s} \left[ \sum_{t=0}^{T} \sum_{n \in N} C^{(n)}(s_t^{(n)}, a_t^{(n)}) \right], \tag{6}$$

which minimizes the expected long-run average cost.

Problems (5) and (6) are multiarmed restless bandit problems with real-state projects, which are generally intractable.

C. Index policies

Instead of attempting to solve such problems optimally, we shall pursue the more practical goals of designing and computing a good heuristic policy of priority-index type. Such policies attach an index $\lambda^{(n)}(s^{(n)})$ to each target $n$ as a function of its STEV state $s^{(n)}$, depending on target parameters. At time $t$, the resulting index policy selects at most $M$ targets to measure, using $\lambda^{(n)}(s_t^{(n)})$ as a priority index for measuring target $n$ (where a larger index value means a higher priority), among those targets, if any, for which the index exceeds the measurement cost, i.e., $\lambda^{(n)}(s_t^{(n)}) > h^{(n)}$, breaking ties arbitrarily.

In the sequel, we shall focus for concreteness on discounted-cost problem (5), although our approach also applies to average-cost problem (6).

III. RESTLESS BANDIT INDEXATION

A. Relaxed Problem, Lagrangian Relaxation, Decomposition

Along the lines introduced in [7] for the equality-constrained case, we first construct a relaxation of (5), relaxing the requirement (4) that at most $M$ targets are measured at each time by the averaged version that the ETD number of measured targets does not exceed $M/(1 - \beta)$, i.e.,

$$E^{\pi}_{s} \left[ \sum_{t=0}^{\infty} \sum_{n \in N} \beta^t a_t^{(n)} \right] \leq M \frac{1}{1 - \beta}. \tag{7}$$

Denoting by $\Pi^{NA}$ the class of nonanticipative scheduling policies (which can measure any number of targets at any time), the relaxed primal problem is

$$\min_{(\pi, s) \in \Pi^{NA}} E^{\pi}_{s} \left[ \sum_{t=0}^{\infty} \sum_{n \in N} \beta^t C^{(n)}(s_t^{(n)}, a_t^{(n)}) \right]. \tag{8}$$

Note that the optimal value (cost) $V^R(s)$ of (8) gives a lower bound on the optimal value $V^*(s)$ of (5).

To address (8) we deploy a Lagrangian approach, attaching a multiplier $\lambda \geq 0$ to (7). The resulting problem

$$\min_{\pi \in \Pi^{NA}} E^{\pi}_{s} \left[ \sum_{t=0}^{\infty} \sum_{n \in N} \beta^t \left( C^{(n)}(s_t^{(n)}, a_t^{(n)}) + \lambda a_t^{(n)} \right) \right] - M \lambda \frac{1}{1 - \beta} \tag{9}$$

is a Lagrangian relaxation of (8), whose optimal value $V^L(s; \lambda)$ gives a lower bound on $V^R(s)$. The Lagrangian dual problem is to find an optimal value $\lambda^*(s)$ of $\lambda$ giving the best such lower bound, which we denote by $V^D(s)$:

$$\max_{\lambda \geq 0} V^L(s; \lambda). \tag{10}$$

Note that (10) is a convex optimization problem, since $V^L(s; \lambda)$ is concave in $\lambda$. Although weak duality ($V^R(s) \geq V^D(s)$) is ensured, satisfaction of strong duality ($V^R(s) = V^D(s)$) calls for further investigation.

Problem (9) decomposes into the $N$ subproblems

$$\min_{\pi^{(n)} \in \Pi^{(n)}} E^{(n)}_{s_t^{(n)}} \left[ \sum_{t=0}^{\infty} \beta^t \left( C^{(n)}(s_t^{(n)}, a_t^{(n)}) + \lambda a_t^{(n)} \right) \right], \tag{11}$$

where $\Pi^{(n)}$ is the class of nonanticipative tracking policies for target $n$ in isolation. Note that in target $n$’s subproblem (11) multiplier $\lambda$ represents an additional tracking cost, to be added to the target’s regular measurement cost $h^{(n)}$.

B. Indexability: Whittle’s Marginal Productivity Index Policy

Consider now target $n$’s subproblem (11) treating measurement charge $\lambda$ as a scalar parameter, which is now allowed to take negative values. The following defines a key structural property of such a parametric restless bandit subproblem, termed indexability, which simplifies its solution and hence that of Lagrangian dual (10).

Definition 1: We say that subproblem (11) is indexable if there exists an index $\lambda^*(s^{(n)})$ which is a scalar function of the target’s STEV state $s \in S$ such that, for any value of
multiplier $\lambda \in \mathbb{R}$, the active action $a_t^{(n)} = 1$ (measuring the target) is optimal in state $s_t^{(n)} = x$ iff $\lambda^{x(n)}(s) \geq \ldots$ as follows. Letting $\phi(s, z) \equiv 1_{B(z)}(s)\phi_1(s) + 1_{B^c(z)}(s)\phi_0(s)$, where $1_B(s)$ is the indicator of set $B$

We shall refer to (12) as the target’s function we denote by $\pi$ giving the corresponding ETD cost incurred.

C. Sufficient Indexability Conditions and Index Evaluation

Yet, indexability needs to be established for the model at hand. For such a purpose, the first author introduced in work reviewed in [14] sufficient indexability conditions for discrete-state restless bandits based on satisfaction on partial conservation laws (PCLs), along with an index algorithm.

The first author has extended the scope of such conditions to real-state restless bandits in results announced in [15], as reviewed next. The ensuing discussion focuses on a single-project restless bandit modeling the optimal tracking of a single target, whose label $n$ is henceforth dropped from the notation. We thus write, e.g., the target’s state and action processes as $s_t \in S \triangleq [0, \infty)$ and $a_t \in \{0, 1\}$, respectively.

We evaluate the performance of an admissible tracking policy $\pi \in \Pi$ along two dimensions: the work measure

$$g(s, \pi) \triangleq E^s_\pi \left[ \sum_{t=0}^{\infty} \beta^t a_t \right],$$

giving the ETD number of times the target is measured under policy $\pi$ starting at $s_0 = s$; and the cost measure

$$f(s, \pi) \triangleq E^s_\pi \left[ \sum_{t=0}^{\infty} \beta^t C(s_t, a_t) \right],$$

giving the corresponding ETD cost incurred.

The target’s optimal tracking problem (11) is

$$\min_{\pi \in \Pi} f(s, \pi) + \lambda g(s, \pi). \quad (12)$$

We shall refer to (12) as the target’s $\lambda$-charge subproblem. Problem (12) is a real-state MDP, whose optimal value function we denote by $V^*(s; \lambda)$.

In order to solve (12) it suffices to consider deterministic stationary policies, which are naturally represented by their active (state) sets, i.e., the set of STEV states where they prescribe the active action (measure the target). For an active set $B \subseteq S$, we shall refer to the $B$-active policy.

We shall focus attention on the family of threshold policies. For a given threshold level $z \in \mathbb{R^c} \triangleq \mathbb{R} \cup \{-\infty, \infty\}$, the $z$-threshold policy measures the target in STEV state $s$ iff $s > z$, so its active set is $B(z) \triangleq \{s \in S : x > z\}$. Note that $B(z) = \{z, \infty\}$ for $s \geq 0$, $B(z) = S = [0, \infty)$ for $z < 0$, and $B(z) = \emptyset$ for $z = \infty$. We denote by $g(s, z)$ and $f(s, z)$ the corresponding work and reward measures.

For fixed $z$, work measure $g(s, z)$ is characterized as the unique solution to the functional equation

$$g(s, z) = \begin{cases} 1 + \beta g(\phi^1(s), z), & s > z \\ \beta g(\phi^0(s), z), & s \leq z, \end{cases} \quad (13)$$

whereas cost measure $f(s, z)$ is characterized by

$$f(s, z) = \begin{cases} C(s, 1) + \beta f(\phi^1(s), z), & s > z \\ C(s, 0) + \beta f(\phi^0(s), z), & s \leq z. \end{cases} \quad (14)$$

We shall use the marginal counterparts of such measures. For threshold $z$ and action $a$, denote by $\langle a, z \rangle$ the policy that takes action $a$ in the initial slot and adopts the $z$-threshold policy thereafter. Define the marginal work measure

$$w(s, z) \triangleq g(s, 1(z)) - g(s, 0(z)), \quad (15)$$

and the marginal cost measure

$$c(s, z) \triangleq f(s, 0(z)) - f(s, 1(z)). \quad (16)$$

If $w(s, z) \neq 0$, define further the MP measure

$$\lambda(s, z) \triangleq \frac{c(s, z)}{w(s, z)}. \quad (17)$$

The following definition extends to the real-state setting a corresponding definition introduced by the first author in [11] for discrete-state restless bandits.

**Definition 2:** We say that subproblem (12) is PCL-indexable (with respect to threshold policies) if:

(i) positive marginal work: $w(s, z) > 0, s \in S, z \in \mathbb{R}$;

(ii) nondecreasing index: the index defined by

$$\lambda^*(s) \triangleq \lambda(s, s), \quad s \in S. \quad (18)$$

is monotonene decreasing in $s$.

The next result, which extends the scope of a corresponding result in [11] for discrete-state restless bandits to the real-state setting, states the validity of the PCL-based sufficient indexability conditions deployed in this paper. It further shows how to evaluate the Whittle’s MP index.

**Theorem 3:** If subproblem (12) is PCL-indexable, then it is indexable and the $\lambda^*(s)$ in (18) is its Whittle’s MP index.

IV. INDEXABILITY ANALYSIS AND INDEX EVALUATION

In the sequel we focus, without loss of generality, on the case where the target’s tracking cost is $h = 0$.

The indexability analysis of the present model is based on evaluation and analysis of work and cost measures $g(s, z)$ and $f(s, z)$, from which their marginal counterparts $w(s, z)$ and $c(s, z)$ are immediately obtained.

Consider the iterates $a_t(s, z)$ and $\phi_t(s, z)$, which are the action and STEV state processes $a_t$ and $s_t$ generated under the $z$-threshold policy starting at $s_0 = s$. They are recursively computed as follows. Letting $\phi_t(s, z) \equiv 1_{B(z)}(s)\phi^1(s) + 1_{B^c(z)}(s)\phi^0(s)$, where $1_B(s)$ is the indicator of set $B$
and \( B^c(z) \equiv \mathbb{R} \setminus B(z) \), \( \phi_0(s, z) \equiv s \) and \( \phi_t(s, z) \equiv \phi(\phi_{t-1}(s, z), z) \) for \( t \geq 1 \). Further, \( a_0(s, z) \equiv 1_{B(z)}(s) \), and \( a_t(s, z) \equiv 1_{B^c(z)}(\phi_t(s, z)) \) for \( t \geq 1 \).

We can thus evaluate work measure \( g(s, z) \) and cost measure \( f(s, z) \) by computing the infinite series

\[
g(s, z) = \sum_{t=0}^{\infty} \beta^t a_t(s, z)
\]

\[
f(s, z) = dr \sum_{t=0}^{\infty} \beta^t \phi_{t+1}(s, z),
\]

truncating them to a finite number \( T \) of terms.

From these, we can readily compute the marginal work and cost measures \( w(s, z) \) and \( c(s, z) \) via (15)–(16). In turn, we can use the latter to obtain the index \( \lambda^* (s) \) via (18).

A. The Myopic Index

The simplest case to consider is the myopic case \( \beta = 0 \), under which \( g(s, z) = a_0(s, z) \), \( f(s, z) = \phi_1(s, z) \), \( w(s, z) = 1 \), \( c(s, z) = dr [\phi_0(s) - \phi_1(s)] \), and hence \( \lambda(s, z) = c(s, z) \) and \( \lambda^*(s) = dr [\phi_0(s) - \phi_1(s)] = dr(\theta + s)^2/(1 + \theta + s) \). Since \( (d/ds)\lambda^*(s) = dr(\theta + s)(2 + \theta + s)/(1 + \theta + s)^2 > 0 \), the myopic index \( \lambda^*(s) \) is increasing. Therefore, both conditions in Definition 2 hold and, by Theorem 3, the target’s optimal tracking problem is indexable and \( \lambda^*(s) \) is its Whittle’s MP index.

B. Verification of PCL-indexability

Based on the above results we obtain the following result.

Proposition 4: The single-target tracking problem is PCL-indexable. Therefore, it is indexable, and the index \( \lambda^*(s) \) calculated above is its Whittle’s MP index.

We can also extend the result of Proposition 4 to the average criterion. Thus, denoting by \( \lambda^*_\beta (s) \) the MP index for discount factor \( \beta \), it holds that \( \lambda^*_\beta (s) \) increases monotonically to a finite limiting index \( \lambda^*(s) \) as \( \beta \nearrow 1 \).

V. COMPUTATIONAL EXPERIMENTS

A. MP Index Evaluation

We have implemented a MATLAB script for index evaluation using the above results. The MP was then computed for a target instance with parameters \( d = 1 \), \( r = 1 \), and \( q = 5 \), so \( \theta = 5 \). The series in (19) were approximately evaluated by truncating them to \( T = 10^3 \) terms for \( \beta = 0.1, 0.2, \ldots, 0.9 \), and to \( T = 10^5 \) terms for \( \beta = 0.9999 \). For each \( \beta \), the index \( \lambda^*(s) \) was evaluated on a grid of \( s \) values of width \( 10^{-3} \).

Fig. 1 shows the results. As required by the PCL-indexability conditions, in each case the index \( \lambda^*(s) \) was monotone nondecreasing (in fact, strictly increasing) in \( s \). Note that the index \( \lambda^*(s) \) is continuous in \( s \), being also piecewise differentiable. Further, for fixed \( s \) the index \( \lambda^*(s) \) is increasing in \( \beta \), converging as \( \beta \nearrow 1 \) to a limiting index that can be used for average-criterion problem (6), which we have approximated by taking \( \beta = 0.9999 \). For each \( s \), the time to compute \( \lambda^*(s) \) was negligible.
### TABLE I
BENCHMARKING RESULTS.

<table>
<thead>
<tr>
<th>q&lt;sup&gt;(1)&lt;/sup&gt;</th>
<th>myopic</th>
<th>TEV</th>
<th>MP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/2</td>
<td>47.586</td>
<td>44.494</td>
<td>40.677</td>
</tr>
<tr>
<td>1</td>
<td>49.196</td>
<td>46.454</td>
<td>43.709</td>
</tr>
<tr>
<td>2</td>
<td>51.395</td>
<td>49.477</td>
<td>47.818</td>
</tr>
<tr>
<td>3</td>
<td>53.096</td>
<td>51.586</td>
<td>50.857</td>
</tr>
<tr>
<td>4</td>
<td>54.806</td>
<td>54.183</td>
<td>53.309</td>
</tr>
<tr>
<td>6</td>
<td>56.908</td>
<td>56.142</td>
<td>55.396</td>
</tr>
<tr>
<td>7</td>
<td>59.406</td>
<td>59.212</td>
<td>58.927</td>
</tr>
<tr>
<td>8</td>
<td>60.433</td>
<td>60.743</td>
<td>60.433</td>
</tr>
<tr>
<td>9</td>
<td>61.939</td>
<td>62.273</td>
<td>61.939</td>
</tr>
<tr>
<td>10</td>
<td>63.444</td>
<td>63.802</td>
<td>63.444</td>
</tr>
</tbody>
</table>


