Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 117 (2007) 1357–1371
stochastic processes and their applications
www.elsevier.com/locate/spa

On Gittins' index theorem in continuous time

Peter Bank[a,*], Christian Küchler[b]

[a] Department of Mathematics, Columbia University, 2990 Broadway, Mail Code 4433, New York, NY 10027, USA
[b] Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany

Received 12 June 2006; received in revised form 10 January 2007; accepted 22 January 2007
Available online 25 January 2007

Abstract

We give a new and comparably short proof of Gittins' index theorem for dynamic allocation problems of the multi-armed bandit type in continuous time under minimal assumptions. This proof gives a complete characterization of optimal allocation strategies as those policies which follow the current leader among the Gittins indices while ensuring that a Gittins index is at an all-time low whenever the associated project is not worked on exclusively. The main tool is a representation property of Gittins index processes which allows us to show that these processes can be chosen to be pathwise lower semi-continuous from the right and quasi-lower semi-continuous from the left. Both regularity properties turn out to be crucial for our characterization and the construction of optimal allocation policies.
© 2007 Elsevier B.V. All rights reserved.

MSC: 60G07; 60G40; 60H25; 91B16; 91B28

Keywords: Gittins index; Multi-armed bandits; Representation theorem

0. Introduction

Multi-armed bandit problems arise when a limited amount of resources, usually thought of as working time, has to be allocated to a number of different projects. These offer dynamically varying, random rewards when worked on, and so one has to trade-off those projects which currently offer the highest rewards against those projects which possibly will yield even better payoffs after they have been worked on for some time.

* Corresponding author. Tel.: +1 2128545545; fax: +1 2128548962.
E-mail addresses: pbank@math.columbia.edu (P. Bank), ckuechler@math.hu-berlin.de (C. Küchler).
URLs: http://www.math.columbia.edu/~pbank (P. Bank), http://www.math.hu-berlin.de/~ckuechler (C. Küchler).

0304-4149/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.spa.2007.01.006

Typically, these problems are rather difficult (if not even impossible) to solve due to their high dimensionality and the complex interdependencies between the projects. For independent projects, however, it was shown first by Gittins and Jones [6] that there exists a project-specific dynamic performance measure, later called the Gittins index of a project, such that optimal allocations are obtained from an index policy which (essentially) amounts to focussing at each point only on those projects which exhibit a maximal Gittins index. This celebrated result was subsequently extended from Gittins' and Jones' original discrete-time, Markovian framework to a completely general continuous-time setting; see, e.g., Whittle [15], Varaiya, Walrand and Buyukkoc [13], Mandelbaum [10], Weber [14], El Karoui and Karatzas [3,4], Kaspi and Mandelbaum [8,9].

These accounts are based on two equivalent definitions of the Gittins index. First, there is the definition as an indifference threshold with respect to early retirement of a project like, e.g., in Whittle [15]. Also El Karoui and Karatzas [3,4] use this for their martingale approach in continuous time. Second, there is Gittins' forward characterization of his index as the properly discounted maximal expected future rewards per expected time used for obtaining these. In continuous time, this characterization is the starting point for the excursion theoretic approach pursued by Kaspi and Mandelbaum [8,9].

By contrast, our approach to Gittins' index theorem is based on a characteristic representation property which relates the Gittins index to the accumulated future expected rewards from a given project; see Corollary 2.1. This property, in conjunction with a novel partial-integration argument, not only allows for a comparably short proof of Gittins' index theorem under minimal assumptions, see Theorem 2, it actually also allows us to give necessary and sufficient conditions for optimality, a result already conjectured by Kaspi and Mandelbaum [9]. Another important point in our analysis is the lower semi-continuity of Gittins index processes. This new regularity result is crucial not only for the continuous-time construction of Gittins indices as proper optional processes, but it actually turns out to be indispensable even for the existence of optimal allocation strategies. Our Theorem 1 shows how this regularity follows from the representation property of Gittins indices. This result also complements the representation theorem of Bank and El Karoui [1].

1. Problem formulation

We shall follow Mandelbaum [10] in order to formalize our allocation problem as a multi-parameter control problem. To this end, let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space on which we have N real-valued processes $h^p = (h^p_s,\ s \geq 0)$ satisfying

$$\mathbb{E}\int_0^\infty \mathrm{e}^{-\alpha s}|h^p_s|\mathrm{d}s < \infty \quad (p = 1, \dots, N).$$

We will interpret h^p_s as the reward rate from project p after we have spent already $s \geq 0$ units of time on this particular project. When working on a project we typically will learn more about its future prospects. This gain of information is modeled by a project-specific filtration $\mathbb{F}^p = (\mathscr{F}^p_s)_{s\geq 0}$ which satisfies the usual conditions of right-continuity and completeness. We assume each h^p is progressively measurable with respect to the filtration $\mathbb{F}^p$ associated with the corresponding project.

An allocation strategy is given by a collection $\mathbf{S}$ of increasing processes $(S^p_t,\ t \geq 0)$ $(p = 1, \dots, N)$ specifying how much time should be spent on each project up to calendar time t.

Feasibility is ensured by the constraint

$$\sum_{p=1}^{N} S_t^p = t \quad \text{for all } t \geq 0.$$

In addition, we require the strategy to be nonanticipative in the sense that for each time $t \geq 0$ the vector $\mathbf{S}_t = (S_t^p,\ p = 1, \dots, N)$ can be viewed as a stopping time for the multi-parameter filtration generated by $\mathbb{F}^p\ (p = 1, \dots, N)$:

$$\{\mathbf{S}_t \leq s\} \triangleq \bigcap_{p=1}^{N} \{S_t^p \leq s^p\} \in \mathscr{F}_{\mathbf{s}} \triangleq \bigvee_{p=1}^{N} \mathscr{F}_{s^p}^p \quad \text{for any } \mathbf{s} \triangleq (s^1, \dots, s^p) \in \mathbb{R}_+^N.$$

The class of all such allocation strategies $\mathbf{S}$ will be denoted by $\mathscr{S}$.

With this notation we can now introduce the rewards generated by a dynamic allocation strategy $\mathbf{S} \in \mathscr{S}$ as

$$\mathscr{R}(\mathbf{S}) \triangleq \sum_{p=1}^{N} \int_0^\infty \mathrm{e}^{-\alpha t} h_{S_t^p}^p \,\mathrm{d}S_t^p,$$

where $\alpha > 0$ is a constant discount factor. The optimization problem we want to study amounts to maximizing the expected rewards:

$$\text{Maximize } \mathbb{E}\mathscr{R}(\mathbf{S}) \text{ over } \mathbf{S} \in \mathscr{S}.$$

A naive approach to the above allocation problem would be to always work on the project which currently exhibits the highest reward rate. It is easy to see that, while such a myopic approach may typically produce reasonable results, it need not be optimal in general. Indeed, despite the discounting of future rewards, it may be worthwhile to engage in a project with presently low rewards if progress in this project is likely to lead to significantly higher rewards in the (not too far) future which will then compensate for the foregoing of *initially* higher rewards obtainable from other projects. However, if the rewards are decreasing for all projects, it is well known that the myopic strategy is indeed optimal as shown by the following lemma whose proof is given in the Appendix for the sake of completeness:

Lemma 1.1. *In the deteriorating case where all reward processes $h^p\ (p = 1, \dots, N)$ are right-continuous and decreasing, an allocation strategy $\mathbf{S}$ maximizes $\mathbb{E}\mathscr{R}(\mathbf{S})$ over $\mathbf{S} \in \mathscr{S}$ if and only if it is myopically following the leading reward:*

$$h_{S_t^p}^p = \max_{q=1,\dots,N} h_{S_t^q}^q \quad \textit{whenever } \mathrm{d}S_t^p > 0\ (p = 1, \dots, N).$$

Here and in the sequel it is convenient to use the following notation.

Notation. *Let $(A_t,\ t \geq 0)$ be left-continuous and increasing. Following Mandelbaum [10], we call $t \geq 0$ a point of right increase for S and write*

$$\mathrm{d}A_t > 0 \quad \textit{if } A_t < A_u \textit{ for all } u > t.$$

Similarly, we use expressions like $\mathrm{d}A_t < \mathrm{d}t$ to indicate that $t \geq 0$ is a point of right increase for $t - A_t$.

Points of right increase fit neatly together with semi-continuity from the right as pointed out in the following lemma which we will use for our characterization of optimal allocation strategies and whose proof we defer to the Appendix.

Lemma 1.2. *Let* $(A_t, t \geq 0)$ *be left-continuous and increasing. If* $(k_t, t \geq 0)$ *is nonnegative with lower semi-right-continuous paths, i.e.,* $\liminf_{s \searrow t} k_s \geq k_t$ *for all* $t \geq 0$*, we have*

$$k_t = 0 \text{ whenever } \mathrm{d}A_t > 0 \quad \text{if and only if} \quad \int_0^\infty k_t \mathrm{d}A_t = 0.$$

2. Gittins index processes and their representation properties

As described by Lemma 1.1, focussing myopically on current reward rates will be optimal in the case of deteriorating rewards. It is well known, however, that in general one should focus instead on the projects' Gittins indices. These are dynamic performance measures which at each point in time indicate the future prospects of each project, taking into account all the information available gathered about the project so far.

While the mathematical construction of these Gittins index processes is rather straight forward in discrete time, it becomes surprisingly subtle in continuous time for technical as well as conceptual reasons. Indeed, as is well known, optimal policies crucially depend on the running infimum of these processes and so these processes sample paths must be sufficiently regular for such infima to become right-continuous and adapted processes. This technical difficulty will be overcome – without any further assumptions on the filtrations – by our Theorem 1 which we shall use to construct Gittins indices as the unique optional solution with lower semi-right-continuous paths of some stochastic representation problem. This pathwise lower semi-continuity from the right will be needed also for our *necessary and sufficient* conditions for optimality of an allocation strategy; cf. the proof of Theorem 2 in the following section.

Even more important on a conceptual level is the fact that one also needs these Gittins index processes to be quasi-lower semi-continuity from the left—this is in order to ensure the mere existence of (optimal) index strategies as will become clear in Section 4 below; see in particular Remark 4.2.

2.1. A stochastic representation problem

In order to address the aforementioned regularity issues of Gittins index processes, we shall investigate more deeply their extraordinarily useful connection with certain stochastic representation problems which was already observed but not further investigated in Bank and El Karoui [1]. The following theorem supplements their main results accordingly; to clarify the structural assumptions, it is formulated in a slightly more general form than necessary for our later applications.

Theorem 1. *Let* X *be an optional process of class* (D) *and continuous in expectation with* $X_\infty = 0$*, defined on a filtered probability space* $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P})$ *satisfying the usual hypotheses of right-continuity and completeness. Consider an optional random measure* μ *on* $[0, \infty)$ *with finite expected total mass* $\mathbb{E}\mu([0, \infty)) < \infty$*, full support* $\operatorname{supp} \mu = [0, \infty)$ *and no atoms almost surely.*

Then there exists a unique optional process ξ *with lower semi-right-continuous paths, i.e.,*

$$\liminf_{u \searrow s} \xi_u \geq \xi_s \quad \textit{for all } s \geq 0 \quad \mathbb{P}\text{-}a.s.,$$

and the property that for every stopping time S *we have* $1_{[S,\infty)}(s) \inf_{u \in [S,s]} \xi_u \in L^1(\mathbb{P} \otimes \mu(\mathrm{d}s))$ *with*

$$X_S = \mathbb{E}\left[\int_S^\infty \inf_{u \in [S,s]} \xi_u \mu(\mathrm{d}s) \middle| \mathscr{F}_S \right]. \tag{1}$$

In addition, ξ *is also quasi-lower semi-continuous from the left:*

$$\liminf_{u \nearrow S} \xi_u \geq \xi_S \quad \textit{for every finite predictable stopping time } S,$$

and we have the alternative equivalent descriptions of ξ *as*

$$\xi_S = \inf \left\{ m \in \mathbb{R} \,\middle|\, \begin{array}{l} m \geq \mathbb{E}[X_S + m\mu((S,T])| \mathscr{F}_S]\, \mathbb{P}\text{-}a.s. \\ \\ \textit{for all stopping times } T \geq S \end{array} \right\} \tag{2}$$

and

$$\xi_S = \operatorname*{ess\,sup}_{T > S \textit{ stopping time}} \frac{\mathbb{E}[X_T - X_S| \mathscr{F}_S]}{\mathbb{E}[\mu((S,T])| \mathscr{F}_S]} \tag{3}$$

at every finite stopping time S.

Proof. Existence of an optional ξ with (1) follows with $f(t,l) \triangleq -l$ and $\xi \triangleq -L$ from Theorem 3 and Remark 2.1 in Bank and El Karoui [1]. From their Eq. (23) we also infer that ξ can actually be chosen via the following threshold principle (2) or, equivalently, via

$$\xi_S = \inf\{m \in \mathbb{R} \mid Y_S^m = X_S\} \quad \text{for every finite stopping time } S \geq 0$$

where

$$Y_S^m = \operatorname*{ess\,sup}_{T \geq S \text{ stopping time}} \mathbb{E}[X_T + m\mu([S,T))| \mathscr{F}_S] \quad (m \in \mathbb{R}).$$

Our assumptions on X and μ ensure that these Snell envelopes have right-continuous sample paths. Moreover, they can be chosen in such a way that $m \mapsto Y_s^m(\omega)$ is continuous and increasing for any $\omega \in \Omega$, $s \geq 0$. Granted the claimed lower semi-continuity from the right (proven below), Theorem 1 in Bank and El Karoui [1] shows that ξ must satisfy (3) at every finite stopping time S and so ξ is uniquely determined as an optional process. So property (2) implies (1) which in turn implies (3). It is also readily checked that conversely (3) implies (2).

By Proposition 2 in Dellacherie and Lenglart [2], to show that ξ is indeed pathwise lower semi-continuity from the right, it suffices to prove $\lim_n \xi_{S_n} \geq \xi_S$, for any sequence of stopping times $S^n \downarrow S$ such that $\zeta \triangleq \lim_n \xi_{S_n}$ exists almost surely. Now, for such a sequence of stopping times, we indeed have that, for $\varepsilon > 0$,

$$Y_S^{\zeta+\varepsilon} = \lim_n Y_{S_n}^{\zeta+\varepsilon} = \lim_n X_{S^n} = X_S.$$

As $\varepsilon > 0$ was arbitrary, this suffices to conclude $\zeta \geq \xi_S$ from the above threshold description of ξ_S.

In order to prove ξ's quasi-lower semi-continuity from the left, let $S^n \uparrow S$ be an announcing sequence of stopping times for the finite predictable stopping time S. Note that since $S^n < S$ on $\{S > 0\}$ we have

$$\eta^n(s) \triangleq 1_{[S^n,\infty)}(s) \inf_{u\in[S^n,s]} \xi_u \to 1_{[S,\infty)}(s) \inf_{u\in[S,s]} \xi_u \wedge \liminf_{u\nearrow S} \xi_u \quad \mathbb{P}\otimes\mu\text{-a.e.} \quad \text{for } n\uparrow\infty.$$

So, granted $\mathbb{P}\otimes\mu$-uniform integrability of the family $(\eta^n, n = 1, 2, \ldots)$, we can invoke the representation property of ξ and use Lebesgue's theorem to conclude

$$\begin{aligned}\mathbb{E}X_S &= \lim_n \mathbb{E}X_{S^n} = \lim_n \mathbb{E}\int_0^\infty \eta^n(s)\mu(\mathrm{d}s) = \mathbb{E}\int_S^\infty \inf_{u\in[S,s]} \xi_u \wedge \liminf_{u\nearrow S} \xi_u \mu(\mathrm{d}s)\\ &\le \mathbb{E}\int_S^\infty \inf_{u\in[S,s]} \xi_u \mu(\mathrm{d}s) = \mathbb{E}X_S.\end{aligned}$$

It follows that equality must in fact hold true everywhere in the above calculation and so $\liminf_{u\uparrow S} \xi_u \ge \xi_S$ due to the lower semi-right-continuity of ξ.

So to conclude our claim it remains to prove that the family $(\eta^n,\ n = 1, 2, \ldots)$ is $\mathbb{P}\otimes\mu$-uniformly integrable. In fact, since $\eta^n(s) \ge 0 \wedge \inf_{u\in[0,s]} \xi_u \in L^1(\mathbb{P}\otimes\mu(\mathrm{d}s))$, it suffices to prove this for $(\eta^n \vee 0,\ n = 1, 2, \ldots)$ or, equivalently,

$$\limsup_{c\uparrow\infty} \sup_n \mathbb{E}\int_0^\infty \eta^n 1_{\{\eta^n>c\}} \mathrm{d}\mu = 0.$$

To this end, let us introduce $T_c^n \triangleq \inf\{t \ge S^n \mid \xi_t \le c\}$, $T_c \triangleq \inf\{t \ge S \mid \xi_t \le c\}$ and note that $T_c^n \searrow S^n$ as $c \uparrow \infty$ and $T_c^n \le T_c$ almost surely, with equality holding true on $\{T_c^n \ge S\}$. Moreover, we have

$$\mathbb{E}\int_0^\infty \eta^n 1_{\{\eta^n>c\}} \mathrm{d}\mu = \mathbb{E}\int_{S^n}^{T_c^n} \inf_{u\in[S^n,s]} \xi_u \mu(\mathrm{d}s) = \mathbb{E}X_{S^n} - \mathbb{E}X_{T_c^n},$$

where for the last identity we used again the representation property of ξ as well as the observation that $\inf_{u\in[S^n,T_c^n]} \xi_u = \xi_{T_c^n}$ by definition of T_c^n and lower semi-continuity from the right. It follows that for $m = 1, 2, \ldots$, we obtain the estimate

$$\sup_{n\ge m} \mathbb{E}\int_0^\infty \eta^n 1_{\{\eta^n>c\}} \mathrm{d}\mu \le \sup_{n\ge m} \mathbb{E}X_{S^n} - \inf_{\substack{S^m\le T\le S\\ T \text{ a stopping time}}} \mathbb{E}X_T + \mathbb{E}\int_S^{T_c} \inf_{u\in[S,s]} \xi_u \mu(\mathrm{d}s). \tag{4}$$

Indeed, due to the previous identity we can write

$$\mathbb{E}\int_0^\infty \eta^n 1_{\{\eta^n>c\}} \mathrm{d}\mu = \mathbb{E}X_{S^n} - \mathbb{E}X_{T_c^n\wedge S} + \mathbb{E}\left[X_S - X_{T_c^n\vee S}\right]$$

where, for $n \ge m$, $T \triangleq T_c^n \wedge S$ is a stopping time between S^m and S as in the above infimum. Moreover, since $T_c^n = T_c$ on $\{T_c^n > S\}$ are level passage times for ξ after time S, employing the representation property in the same way as above allows us to rewrite the last expectation as

$$\mathbb{E}\left[X_S - X_{T_c^n\vee S}\right] = \mathbb{E}\int_S^{T_c^n} \inf_{u\in[S,s]} \xi_u \mu(\mathrm{d}s) 1_{\{T_c^n>S\}} \le \mathbb{E}\int_S^{T_c} \inf_{u\in[S,s]} \xi_u \mu(\mathrm{d}s),$$

where the final estimate holds true since the infimum over ξ is nonnegative over $[S, T_c]$ for $c \ge 0$ by definition of T_c.

Now, since X is continuous in expectation and $S^m \uparrow S$, the supremum and the infimum on the right side of (4) both converge to $\mathbb{E}X_S$ as $m \uparrow \infty$ and so their difference will be smaller than any given $\varepsilon > 0$ for m large enough. As $T_c \searrow S$ for $c \uparrow \infty$, we can choose c so large that the last term in (4) becomes less than ε along with the terms $\mathbb{E}\int_0^\infty \eta^n 1_{\{\eta^n > c\}} \mathrm{d}\mu$ for n less than our previously chosen m. This accomplishes our proof. $\square$

2.2. *Gittins index processes*

We now can use the above representation theorem in order to construct the desired Gittins index processes for every project $p = 1, \dots, N$. To this end, we choose X in Theorem 1 to be the $\mathbb{F}^p$-optional projection of $(\int_t^\infty \mathrm{e}^{-\alpha s} h_s^p \mathrm{d}s,\ t \geq 0)$ and we set $\mu(\mathrm{d}s) \triangleq \mathrm{e}^{-\alpha s}\mathrm{d}s$ and obtain the following immediate corollary:

Corollary 2.1. *For every project $p = 1, \dots, N$, there exists a unique $\mathbb{F}^p$-optional process g^p which is lower semi-continuous from the right and quasi-lower semi-continuous from the left that satisfies the following three equivalent properties:*

(i) *For every $\mathbb{F}^p$-stopping time $S \geq 0$,*

$$g_S^p = \inf\left\{ m \in \mathbb{R} \left| \begin{array}{l} m \geq \mathbb{E}\left[\displaystyle\int_S^T \mathrm{e}^{-\alpha(s-S)} h_s^p \mathrm{d}s + \mathrm{e}^{-\alpha(T-S)} m \,\middle|\, \mathscr{F}_S^p \right] \\ \textit{for all } \mathbb{F}^p\textit{-stopping times } T \geq S \end{array} \right.\right\}.$$

(ii) *For every $\mathbb{F}^p$-stopping time $S \geq 0$,*

$$g_S^p = \operatorname*{ess\,sup}_{T > S\mathbb{F}^p\text{-stopping time}} \frac{\mathbb{E}\left[\int_S^T \mathrm{e}^{-\alpha s} h_s^p \mathrm{d}s \,\middle|\, \mathscr{F}_S^p \right]}{\mathbb{E}\left[\int_S^T \mathrm{e}^{-\alpha s} \mathrm{d}s \,\middle|\, \mathscr{F}_S^p \right]}.$$

(iii) *For every $\mathbb{F}^p$-stopping time $S \geq 0$, the process $1_{[S,\infty)}(s)\mathrm{e}^{-\alpha s} \inf_{u \in [S,s]} g_u^p$ is $\mathbb{P} \otimes \mathrm{d}s$-integrable and*

$$\mathbb{E}\left[\int_S^\infty \mathrm{e}^{-\alpha s} h_s^p \mathrm{d}s \,\middle|\, \mathscr{F}_S^p \right] = \mathbb{E}\left[\int_S^\infty \mathrm{e}^{-\alpha s} \inf_{u \in [S,s]} g_u^p \mathrm{d}s \,\middle|\, \mathscr{F}_S^p \right]. \quad \square$$

The above process g^p is called the *Gittins index process* of project p. Its quasi-lower semi-continuity from the left indicates that Gittins indices weigh the gathering of information about future rewards against the size of current rewards: if new information about a project can be learned at a predictable time, the Gittins index shortly before will be based on the best possible outcome of this information in order to fully reveal the benefit of the additional information which can be gathered with just a little extra effort; once this information has been obtained, the index can only drop from (and will never rise above) its inevitably overoptimistic previous level. This is of course not to say that Gittins indices cannot jump upwards at all—it only means that upward jumps come as a complete surprise, i.e., at a totally inaccessible stopping time.

The pathwise lower semi-continuity from the right is important for the construction of Gittins indices as proper *optional* processes rather than as a collection of random variables each of which is only determined up to a $\mathbb{P}$-null set. While the latter construction is suggested by the classical descriptions (i) and (ii) of Gittins indices described in the preceding corollary, it has the drawback that one cannot really consider the running infima of Gittins index processes needed in the discussion of index policies.

As we shall see, only the above property (iii) is needed for our full characterization of the solution to our allocation problem; this representation property was noted before, but actually not used to prove Gittins' index theorem in, e.g., Mandelbaum [10] and El Karoui and Karatzas [3]. Property (i) in the above corollary is the classical characterization of Gittins indices as the minimal 'retirement reward' which will make one stop a project immediately; see Gittins and Jones [6], Whittle [15], Weber [14]. Property (ii) is Gittins' [5] forward characterization of his index. It is at the heart of the excursion theoretic proofs in Kaspi and Mandelbaum [8,9]; see also Varaiya, Walrand and Buyukkoc [13] and Presman and Sonin [11].

3. Gittins' index theorem

Let us now focus on this paper's main result, the complete characterization of optimal continuous-time allocation strategies as the class of so-called Gittins index strategies. This result was conjectured by Kaspi and Mandelbaum [9], whose excursion theoretic approach already showed that index strategies are indeed optimal. As we shall see, our novel approach via the representation property of Gittins indices allows for a rather short proof of this result which reveals in addition that the index strategies are in fact the *only* optimal allocation policies.

A crucial feature of a project's Gittins index process is that it can be calculated without considering any of the other projects. As a well-known consequence, this allows one to reduce the generally multi-dimensional allocation problem to the problem of separately calculating the N Gittins index processes g^p $(p = 1, \dots, N)$. This, of course, can only hold true provided that our projects are not inextricably intertwined with each other. We thus make the following standard assumption:

Assumption 3.1. *The σ-algebras $\mathscr{F}^p_\infty$ $(p = 1, \dots, N)$ are independent.*

We can now give the following ramification of Gittins' celebrated index theorem:

Theorem 2. *Under Assumption 3.1, an allocation strategy $\mathbf{S} \in \mathscr{S}$ is optimal if and only if it is an index strategy, i.e., if and only if almost surely*

(i) *the strategy follows the leading Gittins index in the sense that*

$$g^p_{S^p_t} = \max_{q=1,\dots,N} g^q_{S^q_t} \quad \textit{whenever } \mathrm{d}S^p_t > 0 \ (p = 1, \dots, N),$$

and

(ii) *whenever a project is not engaged full-time, its Gittins index is at an all-time low:*

$$g^p_{S^p_t} = \inf_{u \in [0, S^p_t]} g^p_u \quad \textit{whenever } \mathrm{d}S^p_t < \mathrm{d}t \ (p = 1, \dots, N).$$

Proof. Consider an arbitrary allocation strategy $\mathbf{S}$. We first use partial integration to rewrite the rewards from project $p = 1, \dots, N$ as follows:

$$\begin{aligned}\int_0^\infty \mathrm{e}^{-\alpha t} h^p_{S^p_t} \mathrm{d}S^p_t &= -\int_0^\infty \mathrm{e}^{-\alpha(t - S^p_t)} \mathrm{d} \int_{S^p_t}^\infty \mathrm{e}^{-\alpha s} h^p_s \mathrm{d}s \\ &= \int_0^\infty \mathrm{e}^{-\alpha s} h^p_s \mathrm{d}s + \int_0^\infty \left\{ \int_{S^p_t}^\infty \mathrm{e}^{-\alpha s} h^p_s \mathrm{d}s \right\} \mathrm{d}\mathrm{e}^{-\alpha(t - S^p_t)}.\end{aligned}$$

Hence the expected rewards from project p are

$$\mathbb{E} \int_0^\infty \mathrm{e}^{-\alpha t} h^p_{S^p_t} \mathrm{d}S^p_t = \mathbb{E} \int_0^\infty \mathrm{e}^{-\alpha s} h^p_s \mathrm{d}s + \mathbb{E} \int_0^\infty \mathbb{E}\left[\int_{S^p_t}^\infty \mathrm{e}^{-\alpha s} h^p_s \mathrm{d}s \,\middle|\, \mathscr{F}_{\mathbf{S}_t} \right] \mathrm{d}\mathrm{e}^{-\alpha(t - S^p_t)},$$

where the passage from the $\{\ldots\}$-term to its optional projection with respect to $\mathbb{F}^{\mathbf{S}} = (\mathscr{F}_{\mathbf{S}_t})_{t\geq 0}$ is justified since $\mathrm{e}^{-\alpha(t-S_t^p)}$ $(t \geq 0)$ is a nonincreasing process adapted to this filtration; see, e.g., Théorème (1.33) in Jacod [7] or Theorem (VI.20.6)(i) in Rogers and Williams [12].

Due to our independence Assumption 3.1, the optional projection of $(\int_{S_t^p}^\infty \mathrm{e}^{-\alpha s} h_s^p \mathrm{d}s,\ t \geq 0)$ w.r.t. $\mathbb{F}^{\mathbf{S}}$ is indistinguishable from the optional projection of $(\int_S^\infty \mathrm{e}^{-\alpha s} h_s^p \mathrm{d}s,\ S \geq 0)$ w.r.t. $\mathbb{F}^p$ when the latter is evaluated along the time change described by $(S_t^p,\ t \geq 0)$; see Lemma 3.2 below. By Corollary 2.1(iii), this latter optional projection is given by

$$\mathbb{E}\left[\int_S^\infty \mathrm{e}^{-\alpha s} \inf_{u\in[S,s]} g_u^p \mathrm{d}s \,\middle|\, \mathscr{F}_S^p\right]$$

at every $\mathbb{F}^p$-stopping time S. Again by Lemma 3.2, the evaluation of this last optional projection along the time change described by $(S_t^p,\ t \geq 0)$ is indistinguishable from the $\mathbb{F}^{\mathbf{S}}$-optional projection of $(\int_{S_t^p}^\infty \mathrm{e}^{-\alpha s} \inf_{u\in[S_t^p,s]} g_u^p \mathrm{d}s,\ t \geq 0)$. Hence, the expected rewards from project p can be written as

$$\begin{aligned}\mathbb{E}\int_0^\infty \mathrm{e}^{-\alpha t} h_{S_t^p}^p \mathrm{d}S_t^p &= \mathbb{E}\int_0^\infty \mathrm{e}^{-\alpha s} \inf_{u\in[0,s]} g_u^p \mathrm{d}s \\ &\quad + \mathbb{E}\int_0^\infty \left\{\int_{S_t^p}^\infty \mathrm{e}^{-\alpha s} \inf_{u\in[S_t^p,s]} g_u^p \mathrm{d}s\right\} \mathrm{d}\mathrm{e}^{-\alpha(t-S_t^p)}.\end{aligned}$$

Now consider the auxiliary reward process $\tilde{h}_s^p \triangleq \inf_{u\in[0,s]} g_u^p$ and note that

$$\tilde{h}_s^p \leq \inf_{u\in[S_t^p,s]} g_u^p \quad \text{for } s \geq S_t^p.$$

Since $\mathrm{e}^{-\alpha(t-S_t^p)}$ is nonincreasing, it follows that

$$\begin{aligned}\mathbb{E}\int_0^\infty \mathrm{e}^{-\alpha t} h_{S_t^p}^p \mathrm{d}S_t^p &\leq \mathbb{E}\int_0^\infty \mathrm{e}^{-\alpha s} \tilde{h}_s^p \mathrm{d}s + \mathbb{E}\int_0^\infty \left\{\int_{S_t^p}^\infty \mathrm{e}^{-\alpha s} \tilde{h}_s^p \mathrm{d}s\right\} \mathrm{d}\mathrm{e}^{-\alpha(t-S_t^p)} \\ &= \mathbb{E}\int_0^\infty \mathrm{e}^{-\alpha t} \tilde{h}_{S_t^p}^p \mathrm{d}S_t^p. \end{aligned} \tag{5}$$

Here the last identity follows by repeating our initial calculation for the reward stream $\tilde{h}^p$ instead of h^p which, being right-continuous, decreasing and adapted, coincides with its own Gittins index: $\tilde{g}^p = \tilde{h}^p$.

Finally summing over p in (5) and denoting by $\tilde{\mathscr{R}}(\mathbf{S})$ the rewards generated by $\mathbf{S}$ when viewed as an allocation strategy for the auxiliary projects $\tilde{h}^p$ $(p = 1, \ldots, N)$ shows that

$$\mathbb{E}\mathscr{R}(\mathbf{S}) \leq \mathbb{E}\tilde{\mathscr{R}}(\mathbf{S}) \leq \sup_{\mathbf{S}'\in\mathscr{S}} \mathbb{E}\tilde{\mathscr{R}}(\mathbf{S}'). \tag{6}$$

Let us verify that in fact equality holds true everywhere in (6) if $\mathbf{S}$ is an index strategy, i.e., a strategy with properties (i) and (ii) as stated in our theorem; existence of such a strategy is proven in Theorem 3 below. Indeed, Inequality (5) results from estimating the integrand $I_t^p \triangleq \int_{S_t^p}^\infty \mathrm{e}^{-\alpha s} \inf_{u\in[S_t^p,s]} g_u^p \mathrm{d}s$ against $\tilde{I}_t^p \triangleq \int_{S_t^p}^\infty \mathrm{e}^{-\alpha s} \inf_{u\in[0,s]} g_u^p \mathrm{d}s$. Clearly I^p has lower semi-right-continuous paths and $\tilde{I}^p$ is continuous. Hence, $k^p \triangleq I^p - \tilde{I}^p \geq 0$ is lower semi-right-continuous and, moreover, it vanishes precisely at those times t when $g_{S_t^p}^p = \inf_{u\in[0,S_t^p]} g_u^p$, since g^p is lower semi-right-continuous. By property (ii) of an index strategy this holds true whenever $\mathrm{d}e^{-\alpha(t-S_t^p)} < 0$. Thus, by Lemma 1.2, equality holds true in (5) for any $p = 1, \dots, N$ and thus also in the first estimate in (6). In addition such a strategy will also satisfy equality in the second estimate, since following the leader among the g^p $(p = 1, \dots, N)$ entails in particular that the strategy also follows the leader among our auxiliary deteriorating reward processes $\tilde{h}^p$ $(p = 1, \dots, N)$. The latter condition, however, was already identified as sufficient for equality in the second estimate; see Lemma 1.1.

As a consequence, any index strategy $\mathbf{S}$ yields a maximal expected reward

$$\mathbb{E}\mathscr{R}(\mathbf{S}) = \sup_{\mathbf{S}' \in \mathscr{S}} \mathbb{E}\tilde{\mathscr{R}}(\mathbf{S}') = \sup_{\mathbf{S}' \in \mathscr{S}} \mathbb{E}\mathscr{R}(\mathbf{S}')$$

and is therefore optimal.

Conversely, it follows that an optimal strategy $\mathbf{S} \in \mathscr{S}$ must satisfy equality in both estimates of (6). Equality in the first estimate requires equality in (5) for every $p = 1, \dots, N$, a condition which using Lemma 1.2 as in the previous paragraph is seen to amount to property (ii) of an index strategy. By the same lemma, the second estimate will hold true as an equality only if

$$\inf_{u \in [0, S_t^p]} g_u^p = \max_{q=1,\dots,N} \inf_{u \in [0, S_t^q]} g_u^q \quad \text{whenever } \mathrm{d}S_t^p > 0.$$

Now, when $\mathrm{d}S_t^p > 0$, all other projects must have $\mathrm{d}S_t^q < \mathrm{d}t$ and so, due to the previously established property (ii), $\inf_{u \in [0, S_t^q]} g_u^q = g_{S_t^q}^q$ for $q \neq p$. It follows that

$$g_{S_t^p}^p \geq \inf_{u \in [0, S_t^p]} g_u^p = \max_{q=1,\dots,N} \inf_{u \in [0, S_t^q]} g_u^q \geq \max_{q \neq p} g_{S_t^q}^q \quad \text{whenever } \mathrm{d}S_t^p > 0.$$

So $\mathbf{S}$ also has to satisfy property (i). $\square$

The following lemma is needed in our proof of Gittins' Index Theorem 2. We record it separately as it highlights the importance of the independence Assumption 3.1.

Lemma 3.2. *For any strategy $\mathbf{S} \in \mathscr{S}$, let the σ-algebra*

$$\mathscr{F}_{\mathbf{S}_t} \triangleq \left\{ A \in \bigvee_{p=1}^{N} \mathscr{F}_\infty^p \,\middle|\, A \cap \{\mathbf{S}_t \leq \mathbf{s}\} \in \mathscr{F}_{\mathbf{s}} \quad \text{for all } \mathbf{s} \in \mathbb{R}_+^N \right\}$$

describe the information gathered up to time $t \geq 0$ and use $\mathbb{F}^{\mathbf{S}} = (\mathscr{F}_{\mathbf{S}_t})_{t \geq 0}$ to denote the corresponding filtration. Fix $p \in \{1, \dots, N\}$ and let $H^p = (H_s^p, s \geq 0)$ be an $\mathscr{F}_\infty^p \otimes \mathscr{B}[0, \infty)$-measurable bounded stochastic process.

Under Assumption 3.1, the optional projection of $(H_{S_t^p}^p,\ t \geq 0)$ with respect to $\mathbb{F}^{\mathbf{S}}$ coincides with the $\mathbb{F}^p$-optional projection G^p of $(H_s^p,\ s \geq 0)$ evaluated along $(S_t^p,\ t \geq 0)$:

$$\mathbb{E}\left[H_{S_T^p}^p \,\middle|\, \mathscr{F}_{\mathbf{S}_T} \right] = G_{S_T^p}^p \quad \text{for any } \mathbb{F}^{\mathbf{S}}\text{-stopping time } T \tag{7}$$

where G^p satisfies

$$\mathbb{E}\left[H_S^p \,\middle|\, \mathscr{F}_S^p \right] = G_S^p \quad \text{for any } \mathbb{F}^p\text{-stopping time } S.$$

Proof. It is easy to see that the class $\mathscr{H}^p$ of those processes H^p for which the assertion of our lemma holds true forms a monotone class. By the monotone class theorem it therefore suffices to show that $\mathscr{H}^p$ contains all $H^p : \Omega \times [0, \infty) \to \mathbb{R}$ of the form

$$H_s^p(\omega) = B^p(\omega) 1_{[s_0, \infty)}(s)$$

for some bounded, $\mathscr{F}^p_\infty$-measurable B^p and some $s_0 \geq 0$. These H^p are right-continuous and have left limits and so it suffices to verify (7) for deterministic times $T \equiv t \in \mathbb{R}_+$. Moreover, we obtain that G^p is the right-continuous version of $(\mathbb{E}\left[B^p \mid \mathscr{F}^p_s\right] 1_{[s_0,\infty)}(s),\ s \geq 0)$. Now take $A \in \mathscr{F}_{\mathbf{S}_t}$ and note that with $s^n_i \triangleq s_0 + i/n$ $(i = 0, 1, \ldots, n = 1, 2, \ldots)$ we can compute

$$\begin{aligned}
\mathbb{E}\left[H^p_{S^p_t} 1_A\right] &= \lim_n \sum_{i=1}^{\infty} \mathbb{E}\left[B^p 1_{A\cap\{s^n_i \leq S^p_t < s^n_{i+1}\}}\right] \\
&= \lim_n \sum_{i=0}^{\infty} \mathbb{E}\left[\mathbb{E}\left[B^p \,\middle|\, \mathscr{F}^p_{s^n_{i+1}} \vee \bigvee_{q\neq p} \mathscr{F}^q_\infty\right] 1_{A\cap\{s^n_i \leq S^p_t < s^n_{i+1}\}}\right] \\
&= \lim_n \sum_{i=0}^{\infty} \mathbb{E}\left[G^p_{s^n_{i+1}} 1_{A\cap\{s^n_i \leq S^p_t < s^n_{i+1}\}}\right] \\
&= \mathbb{E}\left[G^p_{S^p_t} 1_A\right].
\end{aligned}$$

Indeed, the first identity follows by right-continuity of H^p, the second is due to $A \cap \{s^n_i \leq S^p_t < s^n_{i+1}\} \in \mathscr{F}^p_{s^n_{i+1}} \vee \bigvee_{q\neq p} \mathscr{F}^q_\infty$ by definition of $\mathscr{F}_{\mathbf{S}_t}$, the third identity follows from our assumption of independence 3.1, and the last holds because of right-continuity of G^p. $\square$

4. Construction of index strategies

In this section we shall complete the proof of Theorem 2 by explicitly constructing an index strategy. This will be achieved by combining techniques from El Karoui and Karatzas [4] and Kaspi and Mandelbaum [8] with the quasi-lower left-continuity of the Gittins index processes. As we shall see, this latter property allows one to avoid any further assumption on the underlying filtrations $\mathbb{F}^p$ $(p = 1, \ldots, N)$ such as quasi-left-continuity (cf., e.g., El Karoui and Karatzas [4]). As already pointed out by Kaspi and Mandelbaum [9], the main issue here is that, for index strategies to exist, upward jumps of different Gittins index processes may never occur away from the same all-time low: if such a joint jump was conceivable, the required properties of an index strategy would contradict each other; see Remark 4.2 below. Complementing the rather involved excursion theoretic Proposition 10 of Kaspi and Mandelbaum [9], we show how the quasi-lower semi-continuity of Gittins index processes from the left allows one to easily rule out the existence of such unpleasant common jumps:

Lemma 4.1. *Assumption 3.1 entails that no two Gittins index processes jump upward from the same new all-time low. More precisely, the random project-specific level sets*

$$J^p \triangleq \left\{ m \in \mathbb{R} \,\middle|\, m = \inf_{u\in[0,s]} g^p_u \text{ is not attained on } [0,s] \text{ for some } s \in [0,\infty) \right\}$$

are disjoint almost surely:

$$\mathbb{P}[J^p \cap J^q = \varnothing \text{ for all } p \neq q] = 1.$$

Proof. Note first that the random sets J^p $(p = 1, \ldots, N)$ are independent by Assumption 3.1 and also at most countable because each g^p has lower semi-right-continuous paths. It thus suffices to prove that $\mathbb{P}[m \in J^p] = 0$ for fixed $m \in \mathbb{R}$ and $p = 1, \ldots, N$. To this end, consider the sequence

$$S^{n,p} \triangleq \inf\{s \geq 0 \mid m < g^p_s \leq m + 1/n\}$$

which announces when $\inf_{u\in[0,.]} g^p_u$ creeps down to the level m strictly from above, i.e., $(S^{n,p},\ n = 1, 2, \ldots)$ is an announcing sequence for

$$S^p \triangleq \inf\left\{s \geq 0 \mid \inf_{u\in[0,s)} g^p_u = m < g^p_v \text{ for all } v \in [0, s)\right\}.$$

It follows that S^p is an $\mathbb{F}^p$-predictable stopping time and, as g^p is quasi-lower semi-continuous from the left, we have

$$g^p_{S^p} \leq \liminf_{u\uparrow S^p} g^p_u = \inf_{u\in[0,S^p)} g^p_u = m \quad \text{a.s. on } \{S^p < \infty\}$$

by definition of S^p. It follows that almost surely $m \notin J^p$ whenever this level is at all attained by $\inf_{u\in[0,s]} g^p_u$ for some $s \in [0, \infty)$. □

Now we can complete the construction of El Karoui and Karatzas [4] and arrive at the following result.

Theorem 3. *Under Assumption 3.1, there exists an index strategy.*

Proof. For every project $p = 1, \ldots, N$, we consider the $\mathbb{F}^p$-stopping times

$$\sigma^p_m \triangleq \inf\{s \geq 0 \mid g^p_s \leq m\} \quad (m \in \mathbb{R}),$$

the time for which project p has to be worked on before its Gittins index is less than or equal to m for the first time. Let us in addition denote by

$$\tau_m \triangleq \sum_{p=1}^{N} \sigma^p_m \quad (m \in \mathbb{R})$$

the time needed by *any* strategy which always follows the leading Gittins index until each of these indices is less than or equal to m. Observe that, in particular, the index strategy to be constructed will and must have these τ_m $(m \in \mathbb{R})$ as level passage times in this sense.

The mappings $m \mapsto \sigma^p_m$ $(p = 1, \ldots, N)$ and $m \mapsto \tau_m$ are obviously decreasing. In particular, the latter mapping provides a random partition of the time axis $[0, \infty)$ into 'jump intervals' of the form $[\tau_{m+}, \tau_{m-})$ and their complement onto which $m \mapsto \tau_m$ is mapping continuously.

If $t \geq 0$ lies in this complement, there is a unique m such that $\tau_m = t$ and, to ensure that $\mathbf{S}$ follows the leader among the Gittins indices and in accordance with our observation on τ_m, we put $S^p_t \triangleq \sigma^p_m$ so that each project $p = 1, \ldots, N$ is either stalled (when $g^p_{S^p_t} < m$) or worked on with Gittins index $g^p_{S^p_t} = m$.

Let us now construct the index strategy on a jump interval like $[\tau_{m+}, \tau_{m-})$. We first focus on the subinterval $[\tau_{m+}, \tau_m)$. This is nonempty if there is at least one project p_m such that $\sigma^{p_m}_{m+} < \sigma^{p_m}_m$, i.e., if there is a project whose Gittins index approaches the level $m = \inf_{u\in[0,\sigma^{p_m}_{m+})} g^{p_m}_u$ strictly before actually passing it for the first time. Hence, level m is contained in the set J^{p_m} associated to this project by Lemma 4.1. It follows from this lemma that p_m is actually the only project for which m is such a level. Consequently, $\sigma^p_{m+} = \sigma^p_m$ for all $p \neq p_m$ and so our observation on τ_m suggests to put

$$S^p_t \triangleq \begin{cases} \sigma^{p_m}_{m+} + (t - \tau_{m+}) & \text{if } p = p_m, \\ \sigma^p_{m+} = \sigma^p_m & \text{otherwise} \end{cases} \quad \text{for } t \in [\tau_{m+}, \tau_m)$$

in order to ensure the excursion property of an index strategy as well as that we follow the leading Gittins index.

Finally, for the complementary subinterval $[\tau_m, \tau_{m-})$, we follow one of Kaspi and Mandelbaum [9]'s priority schemes and first activate project $p = 1$ until its Gittins index falls below m, before we abandon this project and engage in project $p = 2$ in the same manner etc. This procedure ensures the excursion property. El Karoui and Karatzas [4] formalize this procedure by considering the partition $[\tau_m, \tau_{m-}) = \cup_{p=1}^{N} [\tau_m^{p-1}, \tau_m^p)$ where

$$\tau_m^0 \triangleq \tau_m, \qquad \tau_m^p \triangleq \tau_m^{p-1} + \left(\sigma_{m-}^p - \sigma_m^p\right) \quad (p = 1, \dots, N)$$

and defining

$$S_t^p \triangleq \begin{cases} \sigma_{m-}^p & p = 1, \dots, q-1 \\ \sigma_m^p + t - \tau^{p-1} & \text{if } p = q \\ \sigma_m^p & p = q+1, \dots, N. \end{cases} \quad \text{for } t \in [\tau_m^{q-1}, \tau_m^q) \quad (q = 1, \dots, N).$$

Now S_t^p is defined for every $t \geq 0$ and $p = 1, \dots, N$ and we obtain an allocation strategy $\mathbf{S} \triangleq (S_t^p, t \geq 0) \in \mathscr{S}$. It is readily checked that $\mathbf{S}$ has the desired properties (i) and (ii) and is therefore an index strategy. □

Remark 4.2. The above construction shows that the quasi-lower semi-continuity from the left of Gittins index processes is essential for the existence of an index strategy. More precisely, as already observed in, e.g., Kaspi and Mandelbaum [9], the excursion property of an index strategy could otherwise contradict the requirement that an index strategy follows the leading index. Indeed, it would be conceivable that the index processes of two different projects could approach a common new all-time low without actually attaining it. This would forbid us to retire either project because of the excursion property. At the same time, however, the 'follow the leader' property of an index strategy would require us to retire the project which afterwards exhibits the smaller Gittins index—a dilemma which we show cannot actually occur with positive probability.

Acknowledgments

We are grateful to Mihai Sirbu for pointing out an error in an earlier version of our proof for Lemma 4.1. We also would like to thank the anonymous referee whose careful comments helped to clarify the presentation. The first author was supported by Deutsche Forschungsgemeinschaft through grant BA 2276/1-1, and NSF through grant DMS-0505021, which is gratefully acknowledged. The second author was supported by BMBF through project 03SF0312E, which is gratefully acknowledged.

Appendix

Proof of Lemma 1.1 (*Optimality of Myopic Strategies in the Deteriorating Case*). The following proof is essentially the time change argument due to El Karoui and Karatzas [4]; we only extend it to allow for possibly negative reward processes. We shall use the same notation as in our proof of Theorem 3 above.

Due to our assumption that the reward processes h^p $(p = 1, \dots, N)$ are right-continuous and decreasing the solution of the allocation problem turns out to coincide with the completely anticipating, pathwise solution, and is thus independent from the considered probability measure and filtrations. These are therefore omitted in the sequel for the sake of simplicity.

To see this we use Fubini's theorem and integration by parts to rewrite the (pathwise) reward from a strategy $\mathbf{S}$ as

$$\begin{aligned}
\mathscr{R}(\mathbf{S}) &= \sum_{p=1}^{N}\int_0^\infty \mathrm{e}^{-\alpha t}\left\{\int_0^\infty 1_{\left\{m<h^p_{S^p_t}\right\}}\mathrm{d}m - \int_{-\infty}^0 1_{\left\{h^p_{S^p_t}<m\right\}}\mathrm{d}m\right\}\mathrm{d}S^p_t \\
&= \sum_{p=1}^{N}\int_0^\infty \left\{\int_0^\infty \mathrm{e}^{-\alpha t}\mathrm{d}\left(S^p_t\wedge\sigma^p_m\right)\right\}\mathrm{d}m - \sum_{p=1}^{N}\int_{-\infty}^0 \left\{\int_0^\infty \mathrm{e}^{-\alpha t}\mathrm{d}\left(S^p_t\vee\sigma^p_m\right)\right\}\mathrm{d}m \\
&= \int_0^\infty\left\{\int_0^\infty\left(\sum_{p=1}^{N} S^p_t\wedge\sigma^p_m\right)\alpha\mathrm{e}^{-\alpha t}\mathrm{d}t\right\}\mathrm{d}m \\
&\quad - \int_{-\infty}^0\left\{\int_0^\infty\sum_{p=1}^{N}\left(S^p_t-\sigma^p_m\right)^+\alpha\mathrm{e}^{-\alpha t}\mathrm{d}t\right\}\mathrm{d}m \\
&\le \int_0^\infty\left\{\int_0^\infty (t\wedge\tau_m)\alpha\mathrm{e}^{-\alpha t}\mathrm{d}t\right\}\mathrm{d}m - \int_{-\infty}^0\left\{\int_0^\infty (t-\tau_m)^+\alpha\mathrm{e}^{-\alpha t}\mathrm{d}t\right\}\mathrm{d}m \\
&= \frac{1}{\alpha}\int_0^\infty\left\{1-\mathrm{e}^{-\alpha\tau_m}\right\}\mathrm{d}m - \frac{1}{\alpha}\int_{-\infty}^0 \mathrm{e}^{-\alpha\tau_m}\mathrm{d}m.
\end{aligned}$$

This yields an upper bound for the reward from an arbitrary allocation strategy $\mathbf{S}$. In fact, this estimate is sharp. Indeed, equality will hold true everywhere in the above calculation iff for any $t \ge 0$ the conditions

$$\sum_{p=1}^{N} S^p_t\wedge\sigma^p_m = t\wedge\tau_m \quad \text{for all } m\ge 0$$

and

$$\sum_{p=1}^{N}(S^p_t-\sigma^p_m)^+ = (t-\tau_m)^+ \quad \text{for all } m<0$$

are satisfied, i.e., iff

$$S^p_{\tau_m} = \sigma^p_m \quad \text{for all } m\in\mathbb{R},\ p=1,\dots,N.$$

The last condition is known as *synchronization identity* and was used for the first time in Dynamic Allocation Theory by El Karoui and Karatzas [4]. For the deteriorating reward processes considered here, it is clearly satisfied if and only if $\mathbf{S}$ follows the leading Gittins index. Such strategies which follow the leader among the Gittins indices do in fact exist as shown by Theorem 3. As a consequence, the above upper bound is indeed attained and so we can conclude the asserted characterization of optimal allocation policies. □

Proof of Lemma 1.2 (*Semi-Continuity and Points of Increase*). Let us first prove the 'only if' part of our assertion and take a lower semi-right-continuous $k \ge 0$ for which there is a $t \ge 0$ with $k_t > 0$ but $\mathrm{d}A_t > 0$. Then we actually have $k_u \ge \varepsilon$ for $u \in [t, t+\varepsilon)$ for some sufficiently small $\varepsilon > 0$ and we also have $A_t < A_{t+\varepsilon}$. So

$$\int_0^\infty k_u\mathrm{d}A_u \ge \int_{[t,t+\varepsilon)} k_u\mathrm{d}A_u \ge \varepsilon(A_{t+\varepsilon}-A_t) > 0.$$

For the 'if' part we note that any lower semi-right-continuous $k \geq 0$ is the limit of an increasing sequence of right-continuous, piecewise constant $k^n \geq 0$ $(n = 1, 2, \ldots)$. If, in addition, k vanishes at times t when $\mathrm{d}A_t > 0$, these $0 \leq k^n \leq k$ will inherit this property. But then k^n vanishes on each of its intervals of constancy which contains a point of increase of A. So

$$0 = \int_0^\infty k_t^n \mathrm{d}A_t \nearrow \int_0^\infty k_t \mathrm{d}A_t \quad \text{as } n \uparrow \infty$$

by monotone convergence. □

References

[1] P. Bank, N. El Karoui, A stochastic representation theorem with applications to optimization and obstacle problems, Ann. Probab. 32 (1B) (2004) 1030–1067.

[2] C. Dellacherie, E. Lenglart, Sur des problèmes de régularisation, de récollement et d'interpolation en théorie des processus, in: J. Azema, M. Yor (Eds.), Séminaire de Probabilités XVI, in: Lecture Notes in Mathematics, vol. 920, Springer, 1982, pp. 298–313.

[3] N. El Karoui, I. Karatzas, Dynamic allocation problems in continuous time, Ann. Appl. Probab. 4 (1994) 255–286.

[4] N. El Karoui, I. Karatzas, Synchronization and optimality for multi-armed bandit problems in continuous time, Mat. Apl. Comput. 16 (2) (1997) 117–151.

[5] J.C. Gittins, Bandit processes and dynamic allocation indices (with discussion), J. Roy. Statist. Soc. Ser. B 41 (1979) 148–164.

[6] J.C. Gittins, D.M. Jones, A dynamic allocation index for the sequential design of experiments, in: Progress in Statistics (European Meeting Statisticians, Budapest, 1972), in: Colloq. Math. Soc. János Bolyai, Vol. 9, North-Holland, Amsterdam, 1974, pp. 241–266.

[7] J. Jacod, Calcul Stochastique et Problèmes de Martingales, in: Lecture Notes in Mathematics, vol. 714, Springer, 1979.

[8] H. Kaspi, A. Mandelbaum, Lévy bandits: Multi-armed bandits driven by Lévy processes, Ann. Appl. Probab. 5 (2) (1995) 541–565.

[9] H. Kaspi, A. Mandelbaum, Multi-armed bandits in discrete and continuous time, Ann. Appl. Probab. 8 (4) (1998) 1270–1290.

[10] A. Mandelbaum, Continuous multi-armed bandits and multiparameter processes, Ann. Probab. 15 (4) (1987) 1527–1556.

[11] È.L. Presman, I.N. Sonin, Sequential control with incomplete information, in: Economic Theory, Econometrics, and Mathematical Economics, Academic Press Inc., San Diego, CA, 1990. The Bayesian approach to multi-armed bandit problems, Translated from the 1982 Russian original, edited and with a preface by E.A. Medova-Dempster and M.A.H. Dempster.

[12] L.C.G. Rogers, D. Williams, Foundations, in: Diffusions, Markov Processes, and Martingales, vol. 1–2, Cambridge University Press, Cambridge, 2000. Cambridge Mathematical Library, Reprint of the second (1994) edition.

[13] P.P. Varaiya, J.C. Walrand, C. Buyukkoc, Extensions of the multiarmed bandit problem: The discounted case, IEEE Trans. Automat. Control 30 (5) (1985) 426–439.

[14] R. Weber, On the Gittins index for multiarmed bandits, Ann. Appl. Probab. 2 (4) (1992) 1024–1033.

[15] P. Whittle, Multi-armed bandits and the Gittins index, J. Roy. Statist. Soc. Ser. B 42 (2) (1980) 143–149.