Random Variables with Measurability Constraints with Application to Opportunistic Scheduling

Michael J. Neely
University of Southern California
https://viterbi-web.usc.edu/~mjneely/

Abstract

This paper proves a representation theorem regarding sequences of random elements that take values in a Borel space and are measurable with respect to the sigma algebra generated by an arbitrary union of sigma algebras. This, together with a related representation theorem of Kallenberg, is used to characterize the set of multidimensional decision vectors in a discrete time stochastic control problem with measurability and causality constraints, including opportunistic scheduling problems for time-varying communication networks. A network capacity theorem for these systems is refined, without requiring an implicit and arbitrarily complex extension of the state space, by introducing two measurability assumptions and using a theory of constructible sets. An example that makes use of well known pathologies in descriptive set theory is given to show a nonmeasurable scheduling scheme can outperform all measurable scheduling schemes.

I. Introduction

Let $(\Omega, \mathcal{F})$ and $(\Gamma, \mathcal{G})$ be two measurable spaces. Let $X : \Omega \to \mathbb{R}$ be a Borel measurable function and let $Y : \Omega \to \Gamma$ be a measurable function. The Doob-Dynkin lemma states that X is $\sigma(Y)$-measurable if and only if $X = h(Y)$ for some Borel measurable function $h : \Gamma \to \mathbb{R}$ [1][2][3]. Suppose we know only that X is $\sigma(\mathcal{H}_1 \cup \mathcal{H}_2)$-measurable, where $\mathcal{H}_1 \subseteq \mathcal{F}$ and $\mathcal{H}_2 \subseteq \mathcal{F}$ are two subsigma algebras on Ω. Is it necessarily true that $X = h(Y_1, Y_2)$ for some Borel measurable function $h : [0,1]^2 \to \mathbb{R}$ and some Borel measurable functions $Y_i : \Omega \to [0,1]$ such that Y_i is $\mathcal{H}_i$-measurable for each $i \in \{1,2\}$?

This question motivates the more general question of characterizing all sequences of Borel measurable functions that satisfy certain *measurability constraints*. Fix K as a nonempty set that is finite or countably infinite. For each $k \in K$ let $X_k : \Omega \to \mathbb{R}$ be a function. Let J be a nonempty set with arbitrarily large cardinality. For each $j \in J$, let $\mathcal{H}_j \subseteq \mathcal{F}$ be a given subsigma algebra on Ω. We characterize all $(X_k)_{k \in K}$ that satisfy

$$X_k \text{ is } \sigma(\cup_{j \in J_k} \mathcal{H}_j)\text{-measurable} \quad \forall k \in K \tag{1}$$

where J_k are given nonempty sets that satisfy $J_k \subseteq J$ for all $k \in K$. The first result is that $(X_k)_{k \in K}$ satisfies (1) if and only if

$$X_k = h_k((Y_j)_{j \in \tilde{J}_k}) \quad \forall k \in K \tag{2}$$

for some Borel measurable functions $Y_j : \Omega \to [0,1]$ that are $\mathcal{H}_j$-measurable for each $j \in J$, some countable subsets $\tilde{J}_k \subseteq J_k$, and some Borel measurable functions $h_k : [0,1]^{\tilde{J}_k} \to \mathbb{R}$ for each $k \in K$. Measurability of each function h_k is with respect to the product sigma algebra on $[0,1]^{\tilde{J}_k}$. Observe that each X_k in (2) draws from the the *same* collection of functions $(Y_j)_{j \in J}$ (rather than defining variables $Y_{j,k}$ separately for each k). In particular, a single function Y_j can be used to represent the influence of the sigma algebra $\mathcal{H}_j$ whenever that influence is required. A special case of this result gives an affirmative answer to the question posed in the first paragraph. The result (1)-(2) immediately generalizes to allow X_k to be a random element of any Borel space, such as the space $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$ for some positive integer m.

A. Applications to stochastic control

The measurability constraints (1) have applications to stochastic control. For example, consider a discrete time system that operates over time slots $k \in \{1, 2, 3, \ldots\}$ according to some probability triplet $(\Omega, \mathcal{F}, P)$. Let $S_k : \Omega \to \Omega_S$ be the *system state* that can be observed at time k, which is a random element associated with some measurable space $(\Omega_S, \mathcal{F}_S)$ with arbitrary structure. Every step k the system controller observes S_k and chooses a *decision* that affects a vector of attributes $X_k \in \mathbb{R}^m$, where m is some fixed positive integer. The vector X_k is required to satisfy the following *system constraints*

$$X_k \in C(S_k) \quad \forall k \in \{1, 2, 3, \ldots\} \tag{3}$$

where $C : \Gamma \to Pow(\mathbb{R}^m)$ is a given set-valued function that maps the observed state S_k to a subset of $\mathbb{R}^m$ that consists of all decision options for X_k (where $Pow(\mathbb{R}^m)$ denotes the power set of $\mathbb{R}^m$). The next state S_{k+1} can be influenced by the prior states and decisions according to some model supported by the probability space, such as a Markov chain model. The m components of X_k can represent rewards, prices, power expenditures, and so on, associated with time slot k, and can also include values that affect the next state. This work is motivated by the application of *opportunistic scheduling*, where

$(S_k)_{k=1}^{\infty}$ are independent and identically distributed (i.i.d.) random channel states that are sequentially observed in a wireless communication system at the start of each slot k, X_k is a vector of transmission rates over m different channels, and $C(S_k)$ is the set of all possible transmission rate vectors that can be supported on slot k when the observed channel state is S_k.

Consider *causal and measurable* decision policies that are constrained to make decisions that yield valid random variables and are based only on observations of the past. Assume the decisions can be *stochastic*, so they can be informed by an external source of randomness that is represented by some sigma algebra $\mathcal{G} \subseteq \mathcal{F}$ on Ω. For example, $\mathcal{G}$ might be the sigma algebra generated by an infinite sequence of i.i.d. random elements in some arbitrary measurable space and whose values are selected by an independent computing device at time 0 (before any control decisions are made). Then we require:

$$X_k \text{ is } \sigma(\sigma(S_1) \cup \cdots \cup \sigma(S_k) \cup \mathcal{G})\text{-measurable} \quad \forall k \in \{1, 2, 3, \ldots\} \tag{4}$$

where $\sigma(S_i)$ is the sigma algebra generated by the random element S_i. Under any such causal and measurable decision policy the result (2) implies

$$X_k = h_k(Y_1, \ldots, Y_k, R) \quad \forall k \in \{1, 2, 3, \ldots\} \tag{5}$$

for some Borel measurable functions h_k, some $\mathcal{G}$-measurable random variable R that takes values in $[0, 1]$, and some $\sigma(S_i)$-measurable random variables Y_i that take values in $[0, 1]$. It is interesting that the *same* random variables $R, \{Y_i\}_{i=1}^{\infty}$ can be used to construct X_k for all time steps k. In particular:

- While the observed random elements S_i are associated with an arbitrarily complex measurable space $(\Omega_S, \mathcal{F}_S)$ where Ω_S has arbitrary cardinality, it suffices to boil these random elements down to real-valued random variables $Y_i : \Omega \to [0, 1]$ where each Y_i is a measurable function of S_i.
- While the external source of randomness is from an arbitrarily complex sigma algebra $\mathcal{G}$ on Ω, it suffices to boil it down to a single draw of a random variable $R : \Omega \to [0, 1]$ that is $\mathcal{G}$-measurable.

The constraint $X_k \in C(S_k)$ seems to require knowledge of the full value of S_k, while the form (5) says this constraint must be sustained only by observing the "boiled" variables $Y_1, \ldots, Y_k, R$ (all of which take values in $[0, 1]$). In particular, all policies that satisfy (3)-(4) are characterized according to the following choices:

1) Choose a single $\mathcal{G}$-measurable random variable $R : \Omega \to [0, 1]$.
2) Choose Borel measurable functions $\theta_k : \Omega_S \to [0, 1]$ from which $Y_k = \theta_k(S_k)$ are defined for all $k \in \{1, 2, 3, \ldots\}$.
3) For each $k \in \{1, 2, 3, \ldots\}$, define a Borel measurable function $h_k : [0, 1]^{k+1} \to \mathbb{R}^m$ such that

$$h_k(\theta_1(S_1(\omega)), \theta_2(S_2(\omega)), \ldots, \theta_k(S_k(\omega)), R(\omega)) \in C(S_k(\omega)) \quad \forall \omega \in \Omega \tag{6}$$

If the constraint (6) is impossible to meet, then no causal decision policy that meets the required measurability constraints exists. Sufficient conditions for (6) are given in Section IV using two measurability assumptions that include the existence of a *measurable choice function*. Measurable choice is a classical problem in descriptive set theory and conditions for existence in certain cases are found in the *selection theorems* of [4][5][6][7][8][9][10]. In particular, the works [6][7][8] use measurable choice to establish cost minimizing policies for economics and dynamic programming applications. Our work gives a simple application to the multidimensional *capacity region* in the opportunistic scheduling problem. A theory of *constructible sets* from [11], together with measurable choice, is used to refine the capacity results of [12][13][14]. We also apply classical pathological cases from descriptive set theory to show an example where a nonmeasurable policy produces significantly larger time averages in comparison to any measurable policy.

B. Related work

The Doob-Dynkin lemma is proven on page 603 in [1] (see also Lemma 1.13 in [2], and [3]). Recent discussion of this lemma is in [15]. The Doob-Dynkin lemma can be used to directly characterize all $\sigma(\mathcal{H}_1 \cup \mathcal{H}_2)$-measurable functions $X : \Omega \to \mathbb{R}$ in the special case when $\mathcal{H}_i = \sigma(Y_i)$ for some random variables Y_i for $i \in \{1, 2\}$. In that special case the Doob-Dynkin lemma implies $X = h(Y_1, Y_2)$. The difficulty is that the sigma algebras $\mathcal{H}_1$ and $\mathcal{H}_2$ can be arbitrarily complex, including sigma algebras that cannot be generated by any real-valued random variable. An early version of this question was addressed by the author on StackExchange in [16] using Dynkin's multiplicative class theorem (see Theorem 18.51 in [17]) together with several techniques that are refined and generalized in the current paper.[1] Rather than using a multiplicative class argument, the current paper establishes a related sigma algebra fact that is of interest in its own right.

For the probability space $(\Omega, \mathcal{F}, P)$ used in the stochastic control problem, consider the special case when we are given some measurable space $(\Omega_Q, \mathcal{F}_Q)$ and we are told $\mathcal{G} = \sigma(Q)$ for some random element $Q : \Omega \to \Omega_Q$ that is measurable with respect to $(\Omega, \mathcal{F})$ and $(\Omega_Q, \mathcal{F}_Q)$. The causal and measurable constraint (4) is thus equivalent to

$$X_k \text{ is } \sigma(\sigma(S_1) \cup \cdots \cup \sigma(S_k) \cup \sigma(Q))\text{-measurable} \quad \forall k \in \{1, 2, 3, \ldots\}$$

[1]The question of X being $\sigma(\mathcal{H}_1 \cup \mathcal{H}_2)$-measurable was posed by the author as a StackExchange question in [16]. Users initially conjectured the representation $X = h(Y_1, Y_2)$ was generally impossible but suggested proving a weaker representation by a monotone class argument; the strong result was eventually proven by the author using Dynkin's multiplicative class theorem [17].

from which the Doob-Dynkin lemma immediately implies

$$X_k = h_k(S_1, S_2, \ldots, S_k, Q) \quad \forall k \in \{1, 2, 3, \ldots\} \tag{7}$$

for some measurable function $h_k : \Omega_S^k \times \Omega \rightarrow \mathbb{R}$. However, the reason (5) is stronger (and nontrivial) is that the $\mathcal{G}$-measurable random variable $R : \Omega \rightarrow [0,1]$ takes values only in $[0,1]$ regardless of the complexity of the random element Q that generates $\mathcal{G}$; Similarly each Y_j is $\mathcal{H}_j$-measurable and takes values on $[0,1]$.

An important representation theorem related to (5) is given by Kallenberg in Proposition 5.13 of [2]: There it is shown that if X is a random element of a Borel space and S is a random element of an arbitrary measurable space, and if the probability space is *extended* (using standard product space concepts) to include a random variable U that is uniformly distributed over $[0,1]$ and that is independent of everything else, then $X = g(S, W)$ almost surely, where g is some measurable function and W is some random variable that is uniformly distributed over $[0,1]$ and independent of S. It is not difficult to strengthen this result to *surely* rather than *almost surely* (this is done in Section IV-A for completeness). When applied to the stochastic control problem, if we assume $(\Omega, \mathcal{F}, P)$ is the already-extended space and $\mathcal{G} = \sigma(U)$, the result immediately implies

$$X_k = g_k(S_k, W_k) \quad \forall k \in \{1, 2, 3, \ldots\}$$

where for each $k \in \{1, 2, 3, \ldots\}$, g_k is a Borel measurable function and W_k is a random variable that is uniformly distributed over $[0,1]$ and independent of S_k. However, the g_k functions cannot be viewed as defining a control policy because the value W_k and its structure within the g_k function can depend on the realizations of $S_1, \ldots, S_{k-1}$.

Selection theorems for measurable choice are developed by Blackwell and Ryll-Nardzewski [4], Kuratowski and Ryll-Nardzewski [5], and Von Neumann [9] (see also [10][18]). Measurable choice for economics and dynamic programming are considered by Maitra [7], Aumann [8], and Dubins and Savage [6]. For example, [7] considers a set S for *current states* and a set A for *action choices*, where S is a Borel subset of a Polish space and A is a compact metric space, and shows (see also [6]) that if u is a bounded upper semi-continuous function on $S \times A$ then there is a measurable choice function $\psi : S \rightarrow A$ such that

$$u(s, \psi(s)) = \max_{a \in A} u(s, a) \quad \forall s \in S$$

Continuous time control with measurable choice is in [19].

Fundamental optimality properties for dynamic programming with general state and action sets are in [20][21][7][22]. For example, Blackwell in [20] considers one step of a finite stage dynamic program with Borel spaces A, S, H where A is the set of possible actions, S the set of current states, and H the set of historical states from the past (see also [6]). The one-step goal is to observe $s \in S$ and $h \in H$ and choose an action $a \in A$ to maximize a utility $u(s, a)$ (so the utility depends only on the current state and action). Mild conditions imply that for any policy that chooses a as a measurable function of both s and h, and for any $\epsilon > 0$, there is a measurable *memoryless strategy* that chooses $a \in A$ based only on the current state s that achieves utility at most ϵ worse (for almost all a, h defined in a probabilistic sense). However, [20] also gives a counter-example to show this is impossible without the mild conditions. This counter-example is similar in spirit to the example in Section V-A of the current paper. However, the structure of our example is different: It treats the infinite horizon opportunistic scheduling problem; It uses a different pathological set from descriptive set theory than the one used in [20]; It compares a nonmeasurable policy to all possible measurable policies, rather than comparing a measurable policy of two variables to all possible measurable policies in one variable. Optimality of stationary policies in multi-step dynamic programs over Borel spaces is considered in [21][7] and related nonstationary problems are in [22]. Nonmeasurable gambling strategies are treated in [6].

Tassiulas and Ephremides establish the *capacity region* for a class of time-varying networks in [14] and prove that a max-weight rule stabilizes the network whenever possible. Capacity regions for more general systems that choose $X_k \in C(S_k)$ are treated in [13][12][23], see also related problems of network utility maximization [24][25][26][27] and energy minimization [28]. The general result in [12] makes implicit assumptions regarding measurability and probability space extension. The current paper refines a capacity theorem from [12] without extending the space by introducing two measurability assumptions, including a measurable choice assumption, together with a property of *constructible sets* from [11].

The field of descriptive set theory was initiated in the classic works of Souslin [29] and Lusin [30]. Souslin showed existence of a two dimensional Borel set that has a non-Borel projection onto the first dimension. Examples of multidimensional Borel sets that do not contain a measurable choice function are developed by Blackwell [31], Novikoff [32], Sierpiński [33], and Addison [34] (see also Example 5.1.7 in [10]). In [35] Sierpiński constructs a subset of $[0,1]$ that has inner measure 0 and outer measure 1 (see also [36][37]). These classic pathological examples are used in Section V to show examples where nonmeasurable decisions can be used in the opportunistic scheduling problem to enable time averages that are superior to those achieved by any measurable policy.

II. PRELIMINARIES

A. Terminology

Let $\mathbb{N} = \{1, 2, 3, \ldots\}$ denote the natural numbers, $\mathbb{R}$ the real numbers, and $\mathcal{B}(\mathbb{R})$ the standard Borel sigma algebra on $\mathbb{R}$. For $A \in \mathcal{B}(\mathbb{R})$ define $\mathcal{B}(A) = \{B \in \mathcal{B}(\mathbb{R}) : B \subseteq A\}$. A *measurable space* is a pair $(\Omega, \mathcal{F})$ where Ω is a nonempty set and $\mathcal{F}$

is a sigma algebra on Ω. Suppose $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ are two measurable spaces. Let $\mathcal{H} \subseteq \Omega_1$ be another sigma algebra on Ω_1. With respect to the measurable space $(\Omega_2, \mathcal{F}_2)$, a function $g : \Omega_1 \to \Omega_2$ is said to be *$\mathcal{H}$-measurable* if

$$g^{-1}(A) \in \mathcal{H} \quad \forall A \in \mathcal{F}_2$$

where $g^{-1}(A) = \{\omega \in \Omega_1 : g(\omega) \in A\}$. With respect to the two measurable spaces $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$, a function $g : \Omega_1 \to \Omega_2$ is said to be *measurable* if it is $\mathcal{F}_1$-measurable. Two measurable spaces $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ are *isomorphic* if there is a bijective function $b : \Omega_1 \to \Omega_2$ that is measurable and has a measurable inverse; such a function is called an *isomorphism*. A measurable space $(\Omega, \mathcal{F})$ is called a *Borel space* if it is isomorphic to $(A, \mathcal{B}(A))$ for some $A \in \mathcal{B}([0, 1])$. If $(\Omega_2, \mathcal{F}_2)$ is a Borel space then a measurable function $g : \Omega_1 \to \Omega_2$ is sometimes referred to as a *Borel measurable function* as a reminder that the target space is a Borel space.

Fix J as a nonempty set (possibly uncountably infinite). Let $(\Omega_j, \mathcal{F}_j)$ be measurable spaces for each $j \in J$. Define

$$\times_{j\in J}\Omega_j = \{(x_j)_{j\in J} : x_j \in \Omega_j \quad \forall j \in J\}$$

Define $\mathcal{C}$ as the collection of subsets of $\times_{j\in J}\Omega_j$ of the form $\times_{j\in J}A_j$ for some sets A_j that satisfy: (i) $A_j \in \mathcal{F}_j$ for all $j \in J$; (ii) $A_j = \Omega_j$ for all but at most one index $j \in J$. Define the product sigma algebra on $\times_{j\in J}\Omega_j$, also called the cylindrical sigma algebra, as

$$\otimes_{j\in J}\mathcal{F}_j = \sigma(\mathcal{C})$$

where $\sigma(\mathcal{C})$ denotes the sigma algebra generated by the collection of sets $\mathcal{C}$. For a given measurable space $(\Omega, \mathcal{F})$ define $\Omega^J = \times_{j\in J}\Omega$ and define its product sigma algebra as $\otimes_{j\in J}\mathcal{F}$. A special case of interest is $[0, 1]^J$ with product sigma algebra $\otimes_{j\in J}\mathcal{B}([0, 1])$ (this measurable space is a Borel space whenever J is a finite or countably infinite set).

A *probability space* is a triplet $(\Omega, \mathcal{F}, P)$ where $(\Omega, \mathcal{F})$ is a measurable space and $P : \mathcal{F} \to [0, 1]$ is a probability measure. A *random variable* is a measurable function $X : \Omega \to \mathbb{R}$. A *random element* is a measurable function $S : \Omega \to \Omega_S$ where $(\Omega_S, \mathcal{F}_S)$ is some given measurable space. By $U \sim \mathcal{U}[0, 1]$ we mean that $U : \Omega \to [0, 1]$ is a random variable that is uniformly distributed over $[0, 1]$.

B. Standard results

Lemma 1: There is an isomorphism $\phi : [0, 1] \to [0, 1]^{\mathbb{N}}$. [See Theorem A.47 in [38] and Chapter 13 of [39].]

Lemma 2: If D is an uncountably infinite Borel measurable subset of a Borel space then there is an isomorphism $b : D \to [0, 1]$. [This is a result of Kuratowski in [40], see also statement and proof in Theorem 3.3.13 of [10].]

Lemma 3: Let J be a nonempty set (possibly uncountably infinite). Let $(\Omega, \mathcal{F})$ and $(\Omega_j, \mathcal{F}_j)$ for $j \in J$ be measurable spaces. Then [see similar Lemmas 1.7, 1.8 in [2]]:

- Composition: If $f : \Omega_1 \to \Omega_2$ and $g : \Omega_2 \to \Omega_3$ are measurable functions, the composition $g \circ f$ is measurable.
- Multidimensional expansion: Let $Y_j : \Omega \to \Omega_j$ be measurable functions for each $j \in J$. The function $Y : \Omega \to \times_{j\in J}\Omega_j$ given by $Y = (Y_j)_{j\in J}$ is measurable with respect to $(\Omega, \mathcal{F})$ and $(\times_{j\in J}\Omega_j, \otimes_{j\in J}\mathcal{F}_j)$. In particular, if $\mathcal{H}_j$ is another sigma algebra on Ω for each $j \in J$, and if Y_j is $\mathcal{H}_j$-measurable, then Y is $\sigma(\cup_{j\in J}\mathcal{H}_j)$-measurable.

III. Representation of Borel measurable functions

Throughout this section assume: $(\Omega, \mathcal{F})$ is a measurable space; J is a nonempty set (possibly uncountably infinite); $\mathcal{H}_j \subseteq \mathcal{F}$ is a subsigma algebra on Ω for each $j \in J$.

Proposition 1: Define $\mathcal{C}$ as the set of functions $X : \Omega \to [0, 1]$ of the form $X = h(\vec{Y})$ where $h : [0, 1]^J \to [0, 1]$ is measurable, $\vec{Y} = (Y_j)_{j\in J}$, and $Y_j : \Omega \to [0, 1]$ is $\mathcal{H}_j$-measurable for each $j \in J$. Define $\mathcal{Z}$ as the following collection of subsets of Ω:

$$\mathcal{Z} = \{X^{-1}(B) \subseteq \Omega : B \in \mathcal{B}([0, 1]), X \in \mathcal{C}\}$$

Then

a) $\mathcal{Z}$ is a sigma algebra on Ω.

b) $\sigma(\cup_{j\in J}\mathcal{H}_j) = \mathcal{Z}$.

c) $X : \Omega \to [0, 1]$ is $\sigma(\cup_{j\in J}\mathcal{H}_j)$-measurable if and only if $X \in \mathcal{C}$.

Proof: (Part (a) of Proposition 1) We show $\mathcal{Z}$ satisfies the three properties of a sigma algebra on Ω:

1) To show $\Omega \in \mathcal{Z}$, define the measurable functions $h = 0$, $Y_j = 0$ for all $j \in J$, and $X = h(\vec{Y}) = 0 \in \mathcal{C}$. Define $B = [0, 1] \in \mathcal{B}([0, 1])$. Then $\Omega = X^{-1}(B) \in \mathcal{Z}$.
2) Fix $A \in \mathcal{Z}$. We want to show $A^c \in \mathcal{Z}$. Since $A \in \mathcal{Z}$ there exists $X \in \mathcal{C}$ and $B \in \mathcal{B}([0, 1])$ such that $A = X^{-1}(B)$. Then $A^c = X^{-1}(B^c) \in \mathcal{Z}$.

3) Let $\{A_n\}_{n=1}^{\infty}$ be an infinite sequence of sets in $\mathcal{Z}$. We want to show $\cup_{n=1}^{\infty} A_n \in \mathcal{Z}$. It suffices to show $\cap_{n=1}^{\infty} A_n^c \in \mathcal{Z}$. For each positive integer n there exists $X_n \in \mathcal{C}$ and $B_n \in \mathcal{B}([0,1])$ such that $A_n = X_n^{-1}(B_n)$ and so $A_n^c = X_n^{-1}(B_n^c)$. Let $\phi : [0,1] \to [0,1]^{\mathbb{N}}$ be an isomorphism (recall Lemma 1). Define

$$\begin{aligned} X &= \phi^{-1}\left((X_n)_{n=1}^{\infty}\right) \\ B &= \phi^{-1}\left(\times_{n=1}^{\infty} B_n^c\right) \end{aligned} \tag{8}$$

Since ϕ^{-1} maps measurable sets to measurable sets we have $B \in \mathcal{B}([0,1])$. Then:

$$\begin{aligned} \cap_{n=1}^{\infty} A_n^c &= \{\omega \in \Omega : X_n(\omega) \in B_n^c \quad \forall n \in \mathbb{N}\} \\ &= \left\{\omega \in \Omega : \phi^{-1}\left((X_n(\omega))_{n=1}^{\infty}\right) \in \phi^{-1}\left(\times_{n=1}^{\infty} B_n^c\right)\right\} \\ &= X^{-1}(B) \end{aligned}$$

Considering the structure of set $\mathcal{Z}$, it remains to show that $X \in \mathcal{C}$. Fix $n \in \mathbb{N}$. Since $X_n \in \mathcal{C}$ we have

$$X_n = h^{(n)}(\vec{Y}^{(n)}) \tag{9}$$

for some measurable function $h^{(n)} : [0,1]^J \to [0,1]$ and some $\vec{Y}^{(n)} = (Y_j^{(n)})_{j \in J}$ such that $Y_j^{(n)} : \Omega \to [0,1]$ is $\mathcal{H}_j$-measurable for all $j \in J$. For each $j \in J$ define $W_j : \Omega \to [0,1]$ by

$$W_j = \phi^{-1}(Y_j^{(1)}, Y_j^{(2)}, Y_j^{(3)}, \ldots) \tag{10}$$

Note that W_j is a composition of the measurable function $\phi^{-1} : [0,1]^{\mathbb{N}} \to [0,1]$ with the $\mathcal{H}_j$-measurable function $Z : \Omega \to [0,1]^{\mathbb{N}}$ given by $Z(\omega) = (Y_j^{(1)}(\omega), Y_j^{(2)}(\omega), Y_j^{(3)}(\omega), \ldots)$ and hence W_j is itself $\mathcal{H}_j$-measurable (recall Lemma 3). Write function ϕ according to its components $\phi = (\phi_1, \phi_2, \phi_3, \ldots)$ and note that each component function $\phi_n : [0,1] \to [0,1]$ is measurable. For each $j \in J$ we have from (10)

$$\begin{aligned} (Y_j^{(1)}, Y_j^{(2)}, Y_j^{(3)}, \ldots) &= \phi(W_j) \\ &= (\phi_1(W_j), \phi_2(W_j), \phi_3(W_j), \ldots) \end{aligned}$$

and so $Y_j^{(n)} = \phi_n(W_j)$ for all $j \in J, n \in \mathbb{N}$, that is,

$$\vec{Y}^{(n)} = (\phi_n(W_j))_{j \in J}$$

Substituting the above equality into (9) yields

$$X_n = h^{(n)}((\phi_n(W_j))_{j \in J}) \tag{11}$$

Define the function $\alpha^{(n)} : [0,1]^J \to [0,1]$ for each $x = (x_j)_{j \in J}$ by

$$\alpha^{(n)}(x) = h^{(n)}((\phi_n(x_j))_{j \in J})$$

Define $\vec{W} = (W_j)_{j \in J}$. Using this and the definition of $\alpha^{(n)}$ in (11) gives:

$$X_n = \alpha^{(n)}(\vec{W}) \tag{12}$$

Define the function $h : [0,1]^J \to [0,1]$ by

$$h(x) = \phi^{-1}(\alpha^{(1)}(x), \alpha^{(2)}(x), \alpha^{(3)}(x), \ldots) \quad \forall x \in [0,1]^J$$

The functions $\alpha^{(n)}$ and h are formed by compositions and multidimensional expansions of measurable functions and so they are themselves measurable (recall Lemma 3). By definition of h it holds that

$$\begin{aligned} h(\vec{W}) &= \phi^{-1}(\alpha^{(1)}(\vec{W}), \alpha^{(2)}(\vec{W}), \alpha^{(3)}(\vec{W}), \ldots) \\ &\overset{(a)}{=} \phi^{-1}(X_1, X_2, X_3, \ldots) \\ &\overset{(b)}{=} X \end{aligned}$$

where (a) holds by substituting (12); (b) holds by definition of X in (8). Thus, $X \in \mathcal{C}$.

□

Proof: (Part (b) of Proposition 1) To show that $\mathcal{Z} \subseteq \sigma(\cup_{j \in J} \mathcal{H}_j)$, fix $A \in \mathcal{Z}$. By definition of $\mathcal{Z}$, there exists $B \in \mathcal{B}([0,1])$ and $X \in \mathcal{C}$ such that $A = X^{-1}(B)$, where $X = h(\vec{Y})$ for some measurable function $h : [0,1]^J \to [0,1]$ and some vector-valued function $\vec{Y} = (Y_j)_{j \in J}$ composed of $\mathcal{H}_j$-measurable functions $Y_j : \Omega \to [0,1]$ for each $j \in J$. Thus

$$\begin{aligned} A &= X^{-1}(B) \\ &= \{\omega \in \Omega : h(\vec{Y}) \in B\} \end{aligned} \tag{13}$$

Lemma 3 ensures that $h(\vec{Y})$ is $\sigma(\cup_{j\in J}\mathcal{H}_j)$-measurable, and so the right-hand-side of (13) is a set in $\sigma(\cup_{j\in J}\mathcal{H}_j)$, which implies the desired conclusion $A \in \sigma(\cup_{j\in J}\mathcal{H}_j)$.

We now show $\sigma(\cup_{j\in J}\mathcal{H}_j) \subseteq \mathcal{Z}$. Fix $m \in J$. Let A_m be a subset of Ω such that $A_m \in \mathcal{H}_m$. Define $\vec{Y} = (Y_j)_{j\in J}$ by $Y_j = 0$ if $j \neq m$ and

$$Y_m(\omega) = \begin{cases} 1 & \text{if } \omega \in A_m \\ 0 & \text{else} \end{cases}$$

It is clear that Y_j is $\mathcal{H}_j$-measurable for all $j \in J$. Define the measurable function $h : [0,1]^J \rightarrow [0,1]$ by $h((x_j)_{j\in J}) = x_m$. Define $B = \{1\} \in \mathcal{B}([0,1])$. Then

$$A_m = \{\omega \in \Omega : h(\vec{Y}(\omega)) \in B\}$$

so by definition of $\mathcal{Z}$ we have $A_m \in \mathcal{Z}$. This holds for all $m \in J$ and $A_m \in \mathcal{H}_m$ so

$$\cup_{j\in J}\mathcal{H}_j \subseteq \mathcal{Z}$$

Taking the sigma algebra of both sides gives

$$\sigma(\cup_{j\in J}\mathcal{H}_j) \subseteq \sigma(\mathcal{Z})$$

Part (a) implies that $\sigma(\mathcal{Z}) = \mathcal{Z}$, which completes the proof. □

Proof: (Part (c) of Proposition 1) Suppose $X \in \mathcal{C}$. Then $X = h(\vec{Y})$ for some measurable h and for $\vec{Y} = (Y_j)_{j\in J}$ with $Y_j : \Omega \rightarrow [0,1]$ being $\mathcal{H}_j$-measurable for all $j \in J$. Lemma 3 implies that X is $\sigma(\cup_{j\in J}\mathcal{H}_j)$-measurable.

Now suppose $X : \Omega \rightarrow [0,1]$ is $\sigma(\cup_{j\in J}\mathcal{H}_j)$-measurable. It is well known that X is the pointwise limit of *simple* functions X_m, so that

$$X(\omega) = \lim_{m\rightarrow\infty} X_m(\omega) \quad \forall \omega \in \Omega \tag{14}$$

where for each positive integer m the function $X_m : \Omega \rightarrow [0,1]$ has the form

$$X_m = \sum_{i=1}^{k_m} v_{i,m} 1_{\{X \in I_{i,m}\}} \tag{15}$$

where k_m is some positive integer; $I_{1,m}, I_{2,m}, \ldots, I_{k_m,m}$ are some disjoint sets in $\mathcal{B}([0,1])$ whose union is $[0,1]$; 1_A is an indicator function that is 1 if event A is true and 0 else; $v_{i,m}$ are some real numbers in $[0,1]$ for each $i \in \{1,\ldots,k_m\}$.

Since X is $\sigma(\cup_{j\in J}\mathcal{H}_j)$-measurable, we have for each positive integer m and each $i \in \{1,\ldots,k_m\}$:

$$\{X \in I_{i,m}\} \in \sigma(\cup_{j\in J}\mathcal{H}_j) = \mathcal{Z}$$

where the final equality holds by part (b). It follows by definition of $\mathcal{Z}$ that

$$\{X \in I_{i,m}\} = \{\omega \in \Omega : X_{i,m}(\omega) \in B_{i,m}\} \tag{16}$$

for some $B_{i,m} \in \mathcal{B}([0,1])$ and some $X_{i,m} \in \mathcal{C}$. Substituting (16) into (15) and using (14) we obtain

$$X = \limsup_{m\rightarrow\infty} \sum_{i=1}^{k_m} v_{i,m} 1_{\{X_{i,m} \in B_{i,m}\}} \tag{17}$$

where we have used the fact that the limit exists and so must be equal to the $\limsup$.

By definition of $\mathcal{C}$, each function $X_{i,m} \in \mathcal{C}$ has the form

$$X_{i,m} = h^{(i,m)}((Y_j^{(i,m)})_{j\in J}) \tag{18}$$

for measurable functions $h^{(i,m)} : [0,1]^J \rightarrow [0,1]$ and some $\mathcal{H}_j$-measurable functions $Y_j^{(i,m)} : \Omega \rightarrow [0,1]$ for $j \in J$. Let L be the (countably infinite) set of all indices (i,m) such that $m \in \mathbb{N}$ and $i \in \{1,\ldots,k_m\}$. Let $\phi : [0,1] \rightarrow [0,1]^L$ be an isomorphism. For each $j \in J$ define $W_j = \phi^{-1}((Y_j^{(i,m)})_{(i,m)\in L})$. Since W_j is the composition of the measurable function ϕ^{-1} with the multidimensional expansion of $\mathcal{H}_j$-measurable functions, it is itself $\mathcal{H}_j$-measurable (recall Lemma 3). Define $\phi_{i,m}$ as the (i,m) component of the ϕ function for each $(i,m) \in L$. Then from (18)

$$X_{i,m} = h^{(i,m)}\left((\phi_{i,m}(W_j))_{j\in J}\right)$$

Define $\vec{W} = (W_j)_{j\in J}$. Then

$$(X_{i,m})_{(i,m)\in L} = \alpha(\vec{W}) \tag{19}$$

where $\alpha : [0,1]^J \rightarrow [0,1]^L$ is the measurable function defined for $x = (x_j)_{j\in J}$ by component functions $\alpha_{i,m}(x)$ for each $(i,m) \in L$ by

$$\alpha_{i,m}(x) = h^{(i,m)}\left((\phi_{i,m}(x_j))_{j\in J}\right)$$

Define the measurable function $g : [0,1]^L \to [0,1]$ for each $x = (x_{i,m})_{(i,m)\in L}$ by

$$g(x) = \limsup_{m\to\infty} \sum_{i=1}^{k_m} v_{i,m} 1_{\{x_{i,m}\in B_{i,m}\}}$$

where we observe the limsup is in the set $[0,1]$ because for each m, all $v_{i,m}$ values are in $[0,1]$ and at most one term in the sum is nonzero. It follows that

$$\begin{aligned} X &\overset{(a)}{=} g((X_{i,m})_{(i,m)\in L}) \\ &\overset{(b)}{=} g(\alpha(\vec{W})) \end{aligned}$$

where (a) holds by (17); (b) holds by (19). We can now define the measurable function $h : [0,1]^J \to [0,1]$ by $h(x) = g(\alpha(x))$ and we see that $X = h(\vec{W})$, where $\vec{W} = (W_j)_{j\in J}$ for $W_j : \Omega \to [0,1]$ being $\mathcal{H}_j$-measurable for all $j \in J$. It follows that X has the required form for inclusion in the set $\mathcal{C}$. □

Now fix K as a finite or countably infinite set. For each $k \in K$ let $(V_k, \mathcal{F}_k)$ be a Borel space. We consider measurable functions $X_k : \Omega \to V_k$.

Proposition 2: Fix J as a nonempty set (possibly uncountably infinite). For each $j \in J$, let $\mathcal{H}_j \subseteq \mathcal{F}$ be a sigma algebra on Ω. Fix functions $X_k : \Omega \to V_k$ for $k \in K$, where $(V_k, \mathcal{F}_k)$ are given Borel spaces. For each $k \in K$, fix $J_k \subseteq J$. Then

$$X_k \text{ is } \sigma(\cup_{j\in J_k}\mathcal{H}_j)\text{-measurable} \quad \forall k \in K \tag{20}$$

if and only if for each $k \in K$ we have

$$X_k = h_k((Y_j)_{j\in\tilde{J}_k}) \tag{21}$$

where $h_k : [0,1] \to V_k$ is some measurable function, $Y_j : \Omega \to [0,1]$ are some $\mathcal{H}_j$-measurable functions for each $j \in J$, and $\tilde{J}_k$ is a finite or countably infinite subset of J_k for each $k \in K$.

Proof: For the reverse direction, it is clear from Lemma 3 that if $(X_k)_{k\in K}$ has the given form $X_k = h_k((Y_j)_{j\in\tilde{J}_k})$ then (20) holds. To prove the forward direction, suppose that (20) holds. Fix $k \in K$. Since $(V_k, \mathcal{F}_k)$ is a Borel space, there is a set $D_k \in \mathcal{B}([0,1])$ and an isomorphism $b_k : V_k \to D_k$. Define $Z_k : \Omega \to [0,1]$ by $Z_k = b_k(X_k)$. Lemma 3 implies that Z_k is $\sigma(\cup_{j\in J_k}\mathcal{H}_j)$-measurable. By Proposition 1 we have $Z_k = g^{(k)}\left((Y_j^{(k)})_{j\in J_k}\right)$ with $Y_j^{(k)} : \Omega \to [0,1]$ being $\mathcal{H}_j$-measurable for all $j \in J$, and $g^{(k)} : [0,1]^{J_k} \to [0,1]$ is measurable. For every such real-valued measurable function $g^{(k)}$, there is a countable subset $\tilde{J}_k \subseteq J_k$ for which the function only depends on the variables y_j for $j \in \tilde{J}_k$ [see, for example, related Exercise 1.1.22 in [41] and Section 3.13d in [3]]. Thus, we modify the $g^{(k)}$ functions to $f^{(k)} : [0,1]^{\tilde{J}_k} \to [0,1]$ measurable for which

$$Z_k = f^{(k)}\left((Y_j^{(k)})_{j\in\tilde{J}_k}\right) \tag{22}$$

Let $\phi : [0,1] \to [0,1]^K$ be an isomorphism. For each $j \in J$ define

$$Y_j = \phi^{-1}((Y_j^{(k)})_{k\in K})$$

Since each function $Y_j^{(k)}$ is $\mathcal{H}_j$-measurable, Y_j is also $\mathcal{H}_j$-measurable (recall Lemma 3). For each $k \in K$ let ϕ_k denote the kth component of ϕ. Then

$$\phi_k(Y_j) = Y_j^{(k)}$$

which gives by substitution into (22):

$$\begin{aligned} Z_k &= f^{(k)}\left((\phi_k(Y_j))_{j\in\tilde{J}_k}\right) \\ &= \alpha^{(k)}\left((Y_j)_{j\in\tilde{J}_k}\right) \end{aligned} \tag{23}$$

where $\alpha^{(k)} : [0,1]^{\tilde{J}_k} \to [0,1]$ is defined as the measurable function for each $x = (x_j)_{j\in\tilde{J}_k}$ by

$$\alpha^{(k)}(x) = f^{(k)}\left((\phi_k(x_j))_{j\in\tilde{J}_k}\right)$$

Substituting the definition $Z_k = b_k(X_k)$ into the left-hand-side of (23) gives

$$b_k(X_k) = \alpha^{(k)}\left((Y_j)_{j\in\tilde{J}_k}\right)$$

Taking $b_k^{-1}(\cdot)$ of both sides gives

$$X_k = b_k^{-1}\left(\alpha^{(k)}\left((Y_j)_{j\in\tilde{J}_k}\right)\right)$$

This holds for all $k \in K$ and has the desired form $X_k = h_k\left((Y_j)_{j\in\tilde{J}_k}\right)$ when the measurable function $h_k : [0,1]^{\tilde{J}_k} \to V_k$ is defined by $h_k(x) = b_k^{-1}(\alpha^{(k)}(x))$ for all $x \in [0,1]^{\tilde{J}_k}$. $\square$

Corollary 1: Let $(V, \mathcal{F}_V)$ be a Borel space. Let J be a nonempty set and let $(\Omega_j, \mathcal{F}_j)$ be measurable spaces for each $j \in J$. If $f : \times_{j\in J}\Omega_j \to V$ is a measurable function with respect to $(\times_{j\in J}\Omega_j, \otimes_{j\in J}\mathcal{F}_j)$ and $(V, \mathcal{F}_V)$ and $\omega = (\omega_j)_{j\in J}$ then

$$f(\omega) = h((\theta_j(\omega_j))_{j\in\tilde{J}}) \quad \forall \omega \in \times_{j\in J}\Omega_j$$

where $\tilde{J} \subseteq J$ is a finite or countably infinite set, $\theta_j : \Omega_j \to [0,1]$ is a measurable function for each $j \in \tilde{J}$, and $h : \times_{j\in\tilde{J}}\Omega_j \to V$ is some measurable function with respect to $(\times_{j\in\tilde{J}}\Omega_j, \otimes_{j\in J}\mathcal{F}_j)$ and $(V, \mathcal{F}_V)$.

Proof: Define $S_j : \Omega \to \Omega_j$ by $S_j(\omega) = \omega_j$ for $j \in J$. Define $\mathcal{H}_j = \sigma(S_j)$. Then f is $\sigma(\cup_{j\in J}\mathcal{H}_j)$-measurable and Proposition 2 implies $f = h((Y_j)_{j\in\tilde{J}})$ for a countable subset $\tilde{J} \subseteq J$, a measurable function h, and for Y_j being $\sigma(S_j)$ measurable for each $j \in \tilde{J}$. The Doob-Dynkin lemma implies $Y_j = \theta_j(S_j) = \theta_j(\omega_j)$ for $j \in \tilde{J}$. $\square$

IV. Stochastic control

Throughout this section we fix a probability triplet $(\Omega, \mathcal{F}, P)$. Let $(\Omega_S, \mathcal{F}_S)$ be a measurable space and let $(\Omega_X, \mathcal{F}_X)$ be a Borel space. Consider a discrete time system that evolves over time slots $k \in \{1, 2, 3, \ldots\}$. Let $(S_k)_{k=1}^{\infty}$ be a sequence of random elements of the form $S_k : \Omega \to \Omega_S$. The value S_k represents a system characteristic or state at time k. Let $\mathcal{G} \subseteq \mathcal{F}$ be a sigma algebra on Ω that is used as a source of randomness to facilitate stochastic decisions. Let $(X_k)_{k=1}^{\infty}$ be a sequence of random elements of the form $X_k : \Omega \to \Omega_X$. Each X_k represents a decision that is made at time k based on observing $S_1, \ldots, S_k$. Assume decisions for each step k are made to ensure

$$X_k \text{ is } \sigma(\sigma(S_1) \cup \cdots \cup \sigma(S_k) \cup \mathcal{G})\text{-measurable} \tag{24}$$

$$X_k \in C(S_k) \tag{25}$$

where $C : \Omega_S \to Pow(\Omega_X)$ is a set-valued map and $Pow(\Omega_X)$ is the set of all subsets of Ω_X. Constraint (24) is the *causal and measurable* constraint. Constraint (25) is a system constraint that restricts the X_k value to a set that depends on S_k. Values of S_{k+1} are determined by some probability rule on the system and are possibly dependent on $S_1, \ldots, S_k$ and $X_1, \ldots, X_k$. A special case is when S_k represents the state of a discrete time Markov chain and there is some transition probability kernel that specifies the conditional distribution of S_{k+1} given S_k and X_k.

Decisions X_k can be vector valued with components that represent power expenditures, costs, or rewards incurred or earned by different parts of the system at time k. We want to characterize all decision elements $(X_k)_{k=1}^{\infty}$ that satisfy (24)-(25). Proposition 2 ensures that if (24)-(25) hold then

$$X_k = h_k(Y_1, \ldots, Y_k, R) \in C(S_k) \quad \forall k \in \mathbb{N}$$

for some Borel measurable functions $h_k : \Omega_S^k \times [0,1] \to \Omega_X$, some $\mathcal{G}$-measurable random variable $R : \Omega \to [0,1]$, and some random variables $Y_k = \theta_k(S_k)$ for some measurable functions $\theta_k : \Omega_S \to [0,1]$. It immediately follows that

$$X_k = g_k(S_1, \ldots, S_k, R) \in C(S_k) \quad \forall k \in \mathbb{N} \tag{26}$$

where $g_k : \Omega_S^k \times [0,1] \to \Omega_X$ is defined

$$g_k(s_1, \ldots, s_k, r) = h_k(\theta_1(s_1), \ldots, \theta_k(s_k), r)$$

Consider the following additional assumptions:

Assumption 1: There is a deterministic measurable choice function $\psi : \Omega_S \to \Omega_X$ such that $\psi(s) \in C(s)$ for all $s \in \Omega_S$.

Assumption 2: $\{(s, x) \in \Omega_S \times \Omega_X : x \in C(s)\} \in \mathcal{F}_S \otimes \mathcal{F}_X$

Both assumptions hold if Ω_S is a finite or countably infinite set, $\mathcal{F}_S = Pow(\Omega_S)$, and $C(s)$ is a nonempty subset of $\mathcal{F}_X$ for each $s \in \Omega_S$. Assumptions 1-2 also hold in the case when a vector of resources $P_k \in \mathbb{R}^a$ (such as power allocations) is chosen on each slot $k \in \mathbb{N}$ and affects a vector of rewards $R_k \in \mathbb{R}^b$ (such as transmission rates over links of a communication system) via $R_k = f(S_k, P_k)$, where a, b are given positive integers, Ω_P is a given Borel measurable subset of $\mathbb{R}^a$, $f : \Omega_S \times \Omega_P \to \mathbb{R}^b$ is a given measurable function, and

$$C(s) = \{(p, f(s, p)) \in \mathbb{R}^{a+b} : p \in \Omega_P\} \quad \forall s \in \Omega_S \tag{27}$$

Indeed, Assumption 1 holds for (27) because $\psi(s) = (0, f(s, 0))$ is a deterministic measurable choice function; Assumption 2 can be seen to hold for (27) by defining the measurable function $g : \Omega_S \times \Omega_P \times \mathbb{R}^b \to \mathbb{R}^b$ by $g(s, p, r) = r - f(s, p)$ and observing that $g^{-1}(\{0\})$ is measurable. More general sufficient conditions for existence of a deterministic measurable choice function are given in the *selection theorems* of Blackwell and Ryll-Nardzewski [4], Kuratowski and Ryll-Nardzewski [5], Dubins and Savage [6], Maitra [7], Aumann [8], Schäl [42], Von Neumann [9], Srivastava [10], and Cascales, Kadets, Rodríguez [18].

Lemma 4: Suppose Assumptions 1 and 2 hold. The sequence $(X_k)_{k=1}^{\infty}$ of Borel measurable random elements of the form $X_k : \Omega \rightarrow \Omega_X$ satisfies (24)-(25) if and only if there are measurable functions $v_k : \Omega_S^k \times [0,1] \rightarrow \Omega_X$ for each $k \in \mathbb{N}$ such that

$$v_k(s_1, \ldots, s_k, r) \in C(s_k) \quad \forall (s_1, \ldots, s_k, r) \in \Omega_S^k \times [0,1] \tag{28}$$

and a $\mathcal{G}$-measurable random variable $R : \Omega \rightarrow [0,1]$ such that

$$X_k = v_k(S_1, \ldots, S_k, R) \quad \forall k \in \mathbb{N} \tag{29}$$

Proof: Suppose $(X_k)_{k=1}^{\infty}$ satisfy (24)-(25). Then (26) holds for some measurable functions $g_k : \Omega_S^k \times [0,1] \rightarrow \Omega_X$ and some $\mathcal{G}$-measurable random variable $R : \Omega \rightarrow [0,1]$. Define $v_k : \Omega_S^k \times [0,1] \rightarrow \Omega_X$ by

$$v_k(s_1, \ldots, s_k, r) = \begin{cases} g_k(s_1, \ldots, s_k, r) & \text{if } g_k(s_1, \ldots, s_k, r) \in C(s_k) \\ \psi(s_k) & \text{else} \end{cases}$$

Assumptions 1, 2, and measurability of g_k imply that v_k is measurable. Since $\psi(s) \in C(s)$ for all $s \in \Omega_S$, function v_k satisfies (28). By (26) and definition of v_k we obtain (29).

Conversely, suppose there are v_k functions and a random variable R that satisfy (28)-(29). Properties (28)-(29) imply $X_k \in C(S_k)$ for all k, while measurability of v_k and the structure $X_k = v_k(S_1, \ldots, S_k, R)$ ensure (by the Doob-Dynkin lemma) that X_k is $\sigma(S_1, \ldots, S_k, R)$-measurable. Since $\sigma(R) \subseteq \mathcal{G}$ it holds that X_k is $\sigma(\sigma(S_1) \cup \cdots \cup \sigma(S_k) \cup \mathcal{G})$-measurable, so that (24)-(25) hold. □

The v_k functions and the random variable $R : \Omega \rightarrow [0,1]$ in the above result completely specify a causal and measurable control policy: At time 0, generate a $\mathcal{G}$-measurable random variable $R : \Omega \rightarrow [0,1]$. At each step $k \in \{1, 2, 3, \ldots\}$, observe $(S_1, \ldots, S_k)$ and make the decision $X_k = v_k(S_1, \ldots, S_k, R)$. The above lemma ensures that, if Assumptions 1-2 hold, all policies that satisfy (24)-(25) can be specified in this way.

A. Another representation

The following representation theorem from Kallenberg [2] bears some resemblance to (26) and uses the concept of a *randomization variable* U.

Theorem 1: (Proposition 5.13 in [2]) Fix $(\Omega, \mathcal{F}, P)$ as a probability triplet and let $X : \Omega \rightarrow \Omega_X$ and $S : \Omega \rightarrow \Omega_S$ be random elements where $(\Omega_X, \mathcal{F}_X)$ is a Borel space and $(\Omega_S, \mathcal{F}_S)$ is a measurable space. Suppose there is a random variable $U \sim \mathcal{U}[0,1]$ that is independent of (S, X) (U is called a *randomization variable*). Then

$$X = f(S, R) \quad \text{almost surely}$$

for some measurable function $f : \Omega_S \times [0,1] \rightarrow \Omega_X$ and some random variable $R \sim \mathcal{U}[0,1]$ that is independent of S. Further, R is $\sigma(S, X, U)$-measurable.

The next simple corollary changes "almost surely" to "surely."

Corollary 2: Under the same assumptions as Theorem 1 we can ensure $X = g(S, W)$ surely for some measurable function $g : \Omega_S \times [0,1] \rightarrow \Omega_X$ and some random variable $W \sim \mathcal{U}[0,1]$ that is independent of S and that is $\sigma(S, X, U)$-measurable.

Proof: First consider the case when Ω_X is an uncountably infinite set. Theorem 1 implies $X = f(S, R)$ almost surely for some measurable f and some random variable $R \sim \mathcal{U}[0,1]$ that is independent of S. Let C be an uncountable Borel measurable subset of $[0,1]$ that has measure 0, such as the Cantor set. Let $b : C \rightarrow \Omega_X$ be an isomorphism (recall Lemma 2). Define the random variable $W : \Omega \rightarrow [0,1]$ by

$$W = \begin{cases} R & \text{if } X = f(S, R) \text{ and } R \notin C \\ b^{-1}(X) & \text{else} \end{cases}$$

Since $P[X = f(S, R)] = 1$ and $P[R \notin C] = 1$ we have that $P[W = R] = 1$ and so W is also uniformly distributed over $[0,1]$ and independent of S. By definition of W we have

$$W \notin C \implies (W = R \text{ and } X = f(S, R)) \tag{30}$$
$$W \in C \implies W = b^{-1}(X) \tag{31}$$

Define the measurable function $g : \Omega_S \times [0,1] \rightarrow \Omega_X$ by

$$g(s, w) = \begin{cases} f(s, w) & \text{if } w \notin C \\ b(w) & \text{else} \end{cases}$$

It remains to show $X = g(S, W)$. If $W \notin C$ then by definition of g we have

$$\begin{aligned} g(S, W) &= f(S, W) \\ &\overset{(a)}{=} X \end{aligned}$$

where (a) holds by (30). If $W \in C$ then by definition of g we have

$$g(S, W) = b(W) = b(b^{-1}(X)) = X$$

where we have used (31). The case when Ω_X is finite or countably infinite is similar and proceeds by defining C as a subset of $[0, 1]$ with the same cardinality as Ω_X. □

Corollary 3: If random elements $(S_k)_{k=1}^{\infty}$ and $(X_k)_{k=1}^{\infty}$ satisfy $X_k \in C(S_k)$ for all $k \in \mathbb{N}$ (where each $X_k : \Omega \to \Omega_X$ is measurable with respect to the Borel space $(\Omega_X, \mathcal{F}_X)$; each $S_k : \Omega \to \Omega_S$ is measurable with respect to the general measurable space $(\Omega_S, \mathcal{F}_S)$), and if there is a random variable $U \sim \mathcal{U}[0, 1]$ that is independent of $(S_k, X_k)_{k=1}^{\infty}$, then

a) For all $k \in \mathbb{N}$ we (surely) have $X_k = g_k(S_k, W_k) \in C(S_k)$ for some measurable function $g_k : \Omega_S \times [0, 1] \to \Omega_X$ and some random variable $W_k \sim \mathcal{U}[0, 1]$ that is independent of S_k.

b) If Assumptions 1-2 hold then for all $k \in \mathbb{N}$ there is a measurable function $v_k : \Omega_S \times [0, 1] \to \Omega_X$ that satisfies $v_k(s, r) \in C(s)$ for all $(s, r) \in \Omega_S \times [0, 1]$ such that

$$X_k = v_k(S_k, W_k) \tag{32}$$

where the random variables W_k are the same as in part (a).

Proof: Part (a) follows immediately from Corollary 2. To prove (b), fix $k \in \{1, 2, 3, \dots\}$ and define

$$v_k(s, r) = \begin{cases} g_k(s, r) & \text{if } g_k(s, r) \in C(s) \\ \psi(s) & \text{else} \end{cases}$$

where g_k is from part (a). Assumptions 1 and 2 and measurability of g_k ensure measurability of v_k. Since $\psi(s) \in C(s)$ for all $s \in \Omega_S$, it is clear that $v_k(s, r) \in C(s)$ for all $(s, r) \in \Omega_S \times [0, 1]$. By part (a) it holds that $X_k = v_k(S_k, W_k)$. □

The equality (32) has a simpler structure than (29). However, the v_k functions in (29) completely specify a causal and measurable control policy. In contrast, the v_k functions in (32) do not specify a control policy because each W_k may have some unknown dependence on $S_1, \dots, S_{k-1}$ as well as on additional sources of (potentially noncausal) randomness.

B. Opportunistic scheduling

The following special case is of interest in the area of wireless networks. Fix $m \in \mathbb{N}$ and let $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$ be the measurable space for the decision elements X_k. There are m different wireless links that can change over time according to states $(S_k)_{k=1}^{\infty}$, where S_k describes the state of all channels on slot k. At the start of each slot $k \in \mathbb{N}$ we observe S_k and then choose a transmission rate vector $X_k \in C(S_k)$, where $C(S_k) \subseteq \mathbb{R}^m$ is the set of all transmission rate options available when the channel state is S_k (different rate options arise, for example, from different modulation and coding choices). This is called an *opportunistic scheduling system* because the state S_k is known before X_k is decided. Control strategies for such systems consider network stability [14][23], utility maximization [13][12][24][25][26][27], and energy minimization [28]. Assume $(S_k)_{k=1}^{\infty}$ are identically distributed random elements associated with a measurable space $(\Omega_S, \mathcal{F}_S)$ and a distribution $\lambda : \mathcal{F}_S \to [0, 1]$:

$$\lambda(A) = P[S_k \in A] \quad \forall A \in \mathcal{F}_S$$

The full sequence $(S_k)_{k=1}^{\infty}$ is "chosen by nature" at time 0. In a causal decision scenario, on step k the controller only knows the values of $S_1, \dots, S_k$ before choosing $X_k \in C(S_k)$. In a noncausal scenario the full $(S_k)_{k=1}^{\infty}$ sequence is known. Assume $\mathcal{G} \subseteq \mathcal{F}$ is a subsigma algebra independent of $\sigma((S_k)_{k=1}^{\infty})$ that is used as a source of randomness to facilitate stochastic decisions. Assume there is a random variable $U \sim \mathcal{U}[0, 1]$ that is $\mathcal{G}$-measurable.

The work [12] defines the *network rate region* Γ as the set of all expectations of X_1 that can be achieved on the first slot, shows this set is the same for all slots, and determines the fundamental *capacity region* (see also [14][23]) when such a transmission system is used for single and multi-hop queueing networks.[2] The argument in [12] implicitly allows expanding the probability space to ensure the sigma algebra $\mathcal{G}$ is complex enough to emulate an independent virtual system with identical stochastics over any number of virtual slots before the slot 1 decision on the actual system is made. The next results do not require expanding the probability space and allow $\mathcal{G}$ to be as simple as $\mathcal{G} = \sigma(U)$.

Assumption 3: For the function $C : \Omega_S \to Pow(\mathbb{R}^m)$, there is a bounded subset $D \subseteq \mathbb{R}^m$ such that $C(s)$ is nonempty and $C(s) \subseteq D$ for all $s \in \Omega_S$.[3]

Definition 1: Given a distribution $\lambda : \mathcal{F}_S \to [0, 1]$ and a function $C : \Omega_S \to Pow(\mathbb{R}^m)$ that satisfies Assumption 3, define the *rate region* $\Gamma \subseteq \mathbb{R}^m$ as the set of all expectation vectors $\mathbb{E}[v(S, U)]$ that can be achieved by some measurable function $v : \Omega_S \times \mathbb{R} \to \mathbb{R}^m$ that satisfies $v(s, w) \in C(s)$ for all $s \in \Omega_S$ and $w \in \mathbb{R}$, and on a probability space with independent random elements S and U such that S has distribution λ and $U \sim \mathcal{U}[0, 1]$.

Define $\overline{\Gamma}$ as the closure of the set Γ. Using Corollary 3b, it is straightforward to show that Assumptions 1, 2, 3 imply that Γ is nonempty, bounded, and convex, while $\overline{\Gamma}$ is nonempty, compact, and convex. It can be shown the definition of Γ is unchanged

[2]For 1-hop networks the *capacity region* is the set of all vectors that are dominated by a vector in the closure of Γ. For multi-hop networks the capacity region depends on all possible multi-hop flow allocations available on graphs associated with points in the closure of Γ [23][13][43].

[3]Assumption 3 is mainly for convenience and can be replaced by the weaker assumption that expectations of random vectors $X_k \in C(S_k)$ are finite.

if one allows U to be a random variable of *any* distribution, provided that U and S are independent. The next lemma shows that $\overline{\Gamma}$ captures all time average expectations of X_k that can be achieved at any time k by a measurable decision policy for choosing $X_k \in C(S_k)$, regardless of whether or not the policy is causal. Sample path time averages are also considered in the lemma using a theory of *constructible sets* [11]. Counter-examples in Section V show that time averages can be far outside the set $\overline{\Gamma}$ if the controller can make nonmeasurable decisions.

Proposition 3: Suppose Assumptions 1, 2, 3 hold for the opportunistic scheduling problem with identically distributed random elements $(S_k)_{k=1}^{\infty}$ with some distribution λ. Let $(X_k)_{k=1}^{\infty}$ be a sequence of (Borel measurable) random vectors that satisfy $X_k \in C(S_k)$ surely for each $k \in \mathbb{N}$. Then

a) For all $k \in \mathbb{N}$ we have $\mathbb{E}[X_k] \in \Gamma$ and $\frac{1}{k}\sum_{i=1}^{k}\mathbb{E}[X_i] \in \Gamma$.

b) If $(S_k)_{k=1}^{\infty}$ is i.i.d. and S_k is independent of $(X_1, \ldots, X_{k-1})$ for all $k \in \{2,3,4,\ldots\}$ then for all $k \in \mathbb{N}$

$$\mathbb{E}[X_k|\mathcal{H}_k] \in \overline{\Gamma} \quad \text{almost surely}$$

where $\mathcal{H}_k = \sigma(X_1, \ldots, X_{k-1})$ for $k \geq 2$ and $\mathcal{H}_1 = \{\phi, \Omega\}$.

c) If $(S_k)_{k=1}^{\infty}$ is i.i.d. and S_k is independent of $(X_1, \ldots, X_{k-1})$ for all $k \in \{2,3,4,\ldots\}$ then

$$\lim_{k\to\infty} \text{dist}\left(\frac{1}{k}\sum_{i=1}^{k} X_i, \overline{\Gamma}\right) = 0 \quad \text{almost surely}$$

where $\text{dist}(x, \overline{\Gamma})$ is the Euclidean distance between a point $x \in \mathbb{R}^m$ and the compact and convex set $\overline{\Gamma} \subseteq \mathbb{R}^m$.

Proof: Without loss of generality, for parts (a)-(b) we can assume existence of a random variable $U \sim \mathcal{U}[0,1]$ of the form $U : \Omega \to [0,1]$ that is independent of $(S_k, X_k)_{k=1}^{\infty}$. Indeed, if this does not hold then we can extend the probability space to a new space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ such that

$$\tilde{\Omega} = \Omega \times [0,1], \quad \tilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{B}([0,1]), \quad \tilde{P} = P \otimes \mu$$

where μ is the standard Borel measure on Borel subsets of $[0,1]$. Each outcome of the new sample space has the form $\tilde{\omega} = (\omega, t)$ where $\omega \in \Omega$ and $t \in [0,1]$. Then define $\tilde{S}_k : \tilde{\Omega} \to \Omega_S$ and $\tilde{X}_k : \tilde{\Omega} \to \mathbb{R}^m$ by

$$\tilde{S}_k(\omega, t) = S_k(\omega) \quad , \quad \tilde{X}_k(\omega, t) = X_k(\omega)$$

Also define $U : \tilde{\Omega} \to [0,1]$ by $U(\omega, t) = t$ and observe that $U \sim \mathcal{U}[0,1]$ and U is independent of $(\tilde{X}_k, \tilde{S}_k)_{k=1}^{\infty}$. Then $(\tilde{X}_k, \tilde{S}_k)_{k=1}^{\infty}$ on the extended probability space has the same distribution as $(X_k, S_k)_{k=1}^{\infty}$ on the original space. Thus, X_k and $\tilde{X}_k$ have the same expectation (useful for part (a)); $f(X_1, \ldots, X_{k-1})$ and $f(\tilde{X}_1, \ldots, \tilde{X}_{k-1})$ have the same distribution for any measurable function f (useful for part (b)).

To prove (a), suppose there is a sequence of random vectors $(X_k)_{k=1}^{\infty}$ that satisfy $X_k \in C(S_k)$ surely for each $k \in \{1,2,3,\ldots\}$. Assuming existence of $U \sim \mathcal{U}[0,1]$ that is independent of $(S_k, X_k)_{k=1}^{\infty}$, apply Corollary 3b to obtain

$$X_k = v_k(S_k, W_k) \tag{33}$$

for some measurable function $v_k : \Omega_S \times [0,1] \to \mathbb{R}^m$ that satisfies $v_k(s,v) \in C(s)$ for all $s \in \Omega_S, t \in [0,1]$ and some random variable $W_k \sim \mathcal{U}[0,1]$ that is independent of S_k. By definition of Γ it holds that $\mathbb{E}[X_k] \in \Gamma$. This holds for all $k \in \mathbb{N}$. Convexity of Γ ensures that $\frac{1}{k}\sum_{i=1}^{k}\mathbb{E}[X_k] \in \Gamma$ for all $k \in \mathbb{N}$.

To prove (b), assume $(S_k)_{k=1}^{\infty}$ are i.i.d. and fix $k \in \{2,3,4,\ldots\}$. Let Z be a version of $\mathbb{E}[X_k|\mathcal{H}_k]$. Since $\overline{\Gamma}$ is a closed subset of $\mathbb{R}^m$ it is Borel measurable and $\{Z \notin \overline{\Gamma}\}$ is an event. Suppose $P[Z \notin \overline{\Gamma}] > 0$ (we reach a contradiction). Since $\overline{\Gamma}$ is compact and convex it is *constructible*, meaning it is the countable intersection of closed half-spaces (see Proposition 7.5.6 in [11]):

$$\overline{\Gamma} = \cap_{j=1}^{\infty}\{x \in \mathbb{R}^m : a_j^{\top} x \leq b_j\} \tag{34}$$

for some $a_j \in \mathbb{R}^m$ and $b_j \in \mathbb{R}$ for $j \in \mathbb{N}$. Then

$$P\left[Z \notin \overline{\Gamma}\right] = P[\cup_{j=1}^{\infty}\{a_j^{\top} Z > b_j\}]$$

Since the above probability is positive, there must be an index $i \in \mathbb{N}$ for which $P[a_i^{\top} Z > b_i] > 0$. Define the event $A = \{a_i^{\top} Z > b_i\}$. Then $P[A] > 0$. Since Z is a version of $\mathbb{E}[X_k|\mathcal{H}_k]$ it is $\mathcal{H}_k$-measurable and $A \in \mathcal{H}_k$. By the defining property of a conditional expectation it holds that

$$\mathbb{E}[Z 1_A] = \mathbb{E}[X_k 1_A]$$

and so

$$\frac{\mathbb{E}\left[a_i^{\top} Z 1_A\right]}{P[A]} = \frac{a_i^{\top}\mathbb{E}[X_k 1_A]}{P[A]}$$

Since $U \sim \mathcal{U}[0,1]$ is independent of the random element $(S_k, 1_A) \in \Omega_S \times \{0,1\}$ we have by Corollary 3b

$$X_k = v(S_k, 1_A, W)$$

for some random variable $W \sim \mathcal{U}[0,1]$ that is independent of $(S_k, 1_A)$, and for some measurable function $v : \Omega_S \times \{0,1\} \times [0,1] \to \mathbb{R}^m$ such that

$$v(s,b,w) \in C(s) \quad \forall s \in \Omega_S, b \in \{0,1\}, w \in [0,1]$$

Then

$$\frac{\mathbb{E}\left[a_i^\top Z 1_A\right]}{P[A]} = \frac{a_i^\top \mathbb{E}\left[v(S_k, 1_A, W)1_A\right]}{P[A]} \tag{35}$$

$$= a_i^\top \mathbb{E}\left[v(S_k, 1, W)|A\right] \tag{36}$$

$$= a_i^\top \mathbb{E}\left[v(S_k, 1, W)\right] \tag{37}$$

where we have used the fact that (S_k, W) is independent of A (recall that W is independent of $(S_k, 1_A)$ and S_k is independent of the $\mathcal{H}_k$-measurable random variable 1_A, so that $S_k, W, 1_A$ are mutually independent). Defining $y = \mathbb{E}\left[v(S_k, 1, W)\right]$, it follows by definition of Γ that $y \in \Gamma \subseteq \overline{\Gamma}$ and so $a_i^\top y \leq b_i$ (recall (34)). Thus

$$\frac{\mathbb{E}\left[a_i^\top Z 1_A\right]}{P[A]} \leq b_i$$

and so

$$\mathbb{E}\left[(a_i^\top Z - b_i)1_A\right] \leq 0 \tag{38}$$

By definition of event $A = \{a_i^\top Z > b_i\}$, the random variable $(a_i^\top Z - b_i)1_A$ is nonnegative, so (38) implies

$$(a_i^\top Z - b_i)1_A = 0 \quad \text{almost surely}$$

However, $(a_i^\top Z - b_i)1_A > 0$ if and only if $1_A > 0$, which contradicts the fact that $P[A] > 0$.

To prove (c), define $M_1 = X_1 - \mathbb{E}\left[X_1\right]$ and define $M_k = X_k - \mathbb{E}\left[X_k|\mathcal{H}_k\right]$ for $k \in \{2,3,4,\ldots\}$. Then $\{M_k\}_{k=1}^\infty$ is a zero mean sequence of bounded random vectors in $\mathbb{R}^m$ where each component forms a martingale difference, and so the law of large numbers for martingale differences implies that $\frac{1}{k}\sum_{i=1}^k M_i$ converges to 0 almost surely [44]. Thus, $\frac{1}{k}\sum_{i=1}^k X_i - \frac{1}{k}\sum_{i=1}^k \mathbb{E}\left[X_k|\mathcal{H}_k\right]$ converges to 0 almost surely. However, part (b) implies $\mathbb{E}\left[X_k|\mathcal{H}_k\right] \in \overline{\Gamma}$ almost surely, and convexity of $\overline{\Gamma}$ implies $\frac{1}{k}\sum_{i=1}^k \mathbb{E}\left[X_k|\mathcal{H}_k\right] \in \overline{\Gamma}$ almost surely. □

Conversely, if $(S_k)_{k=1}^\infty$ is identically distributed with distribution λ and $U \sim \mathcal{U}[0,1]$ is independent of $(S_k)_{k=1}^\infty$, then for any $x \in \overline{\Gamma}$ it is straightforward to see there is a causal and measurable decision policy that chooses $X_k \in C(S_k)$ to ensure:

$$\lim_{k\to\infty} \tfrac{1}{k}\textstyle\sum_{i=1}^k \mathbb{E}\left[X_i\right] = x$$

and if $(S_k)_{k=1}^\infty$ are i.i.d. then we can further ensure

$$\lim_{k\to\infty} \tfrac{1}{k}\textstyle\sum_{i=1}^k X_i = x \quad \text{almost surely}$$

This is done by first mapping the random variable $U \sim \mathcal{U}[0,1]$ to a sequence of i.i.d. $\mathcal{U}[0,1]$ random variables $(U_1, U_2, U_3, \ldots)$ by using the bits corresponding to the binary expansion of U. Then fix any sequence of points $x_k \in \Gamma$ that satisfy $x_k \to x$ and define $X_k = v_k(S_k, U_k)$ for $k \in \mathbb{N}$ where v_k is the corresponding measurable function in the definition of Γ that ensures $X_k \in C(S_k)$ and $\mathbb{E}\left[X_k\right] = x_k$ for all $k \in \mathbb{N}$.

V. Counter-examples

This section considers pathological cases for the opportunistic scheduling problem with i.i.d. channel states $(S_k)_{k=1}^\infty$ and only $m = 1$ channel. We use $(\Omega_S, \mathcal{F}_S) = ([0,1], \mathcal{B}([0,1]))$ throughout.

A. *Nonmeasurable policies*

This example gives an opportunistic scheduling system for which a nonmeasurable policy produces larger time averages in comparison to any measurable policy. It is similar in spirit to the example given by Blackwell in [20] which shows that a single decision at one step of a finite stage dynamic program can bring arbitrarily more utility if it makes a measurable decision based on both the current state and memory, rather than only on the current state (even if the utility function depends only on the current state and the current decision). However, the structure of that example is different from ours and compares two measurable policies rather than a nonmeasurable policy in comparison to any measurable decision.

Define

$$\Omega = [0,1]^{\mathbb{N}}, \quad \mathcal{F} = \otimes_{k\in\mathbb{N}}\mathcal{B}([0,1]), \quad P = \otimes_{k\in\mathbb{N}}\mu$$

where μ is the standard Borel measure on $[0,1]$. Each outcome has the structure $\omega = (\omega_0, \omega_1, \omega_2, \ldots)$. Define $U : \Omega \to [0,1]$ and $S_k : \Omega \to [0,1]$ by $U(\omega) = \omega_0$ and $S_k(\omega) = \omega_k$ for $k \in \{1,2,3,\ldots\}$. Then $(S_k)_{k=1}^\infty$ are i.i.d. $\mathcal{U}[0,1]$ variables; $U \sim \mathcal{U}[0,1]$ is independent of $(S_k)_{k=1}^\infty$.

Fix $A \subseteq [0,1]$ as a set with inner measure 0 and outer measure 1 (such sets exist under the Axiom of Choice [35][36][37]). In particular, neither A nor its complement $A^c = [0,1]\backslash A$ contains a Borel subset of positive measure. Define $C : [0,1] \rightarrow Pow(\mathbb{R})$ by

$$C(s) = \begin{cases} \{0,1\} & \text{if } s \in A \\ \{0,2\} & \text{if } s \in A^c \end{cases}$$

A measurable choice function for this system is $\psi(s) = 0$ for all $s \in [0,1]$ and so Assumption 1 holds. However, it can be shown that Assumption 2 fails.

Consider any decision policy that is measurable, meaning that it produces valid random variables $(X_k)_{k=1}^{\infty}$ that satisfy $X_k \in C(S_k)$ surely for all $k \in \mathbb{N}$. For each fixed $k \in \mathbb{N}$ define the set

$$D_k = \{\omega \in \Omega : \omega_k \in A\}$$

It turns out that every $\mathcal{F}$-measurable subset of D_k has measure 0, and every $\mathcal{F}$-measurable subset of D_k^c also has measure 0 (proof postponed to the next paragraph). Since X_k is a valid random variable we know $\{X_k = 1\}$ and $\{X_k = 2\}$ are valid events (i.e., in $\mathcal{F}$). Since $X_k \in C(S_k)$ we have $\{X_k = 1\} \subseteq D_k$ and so $P[X_k = 1] = 0$. Likewise, $\{X_k = 2\} \subseteq D_k^c$ and so $P[X_k = 2] = 0$. It follows that $X_k = 0$ almost surely for all $k \in \mathbb{N}$ and so $\lim_{k\rightarrow\infty} \frac{1}{k}\sum_{i=1}^{k} X_i = 0$ almost surely. On the other hand, the (nonmeasurable) policy that chooses $X_k = 1$ if $S_k \in A$ and $X_k = 2$ otherwise surely yields $X_k \in C(S_k)$ and $X_k \geq 1$ for all $k \in \mathbb{N}$, so $\liminf_{k\rightarrow\infty} \frac{1}{k}\sum_{i=1}^{k} X_i \geq 1$ surely.

It remains to show that every $\mathcal{F}$-measurable subset of D_k has probability 0 (the corresponding proof for D_k^c is similar). Let Z be a $\mathcal{F}$-measurable subset of D_k. Let $\pi_k(Z)$ be the projection of Z onto the dimension k. Then $\pi_k(Z) \subseteq A$. Since $Z \in \mathcal{F}$, Z can be viewed as a Borel measurable subset of the Polish space $[0,1]^{\mathbb{N}}$, and since the projection onto one dimension is a continuous function, the set $\pi_k(Z)$ is an analytic subset of $[0,1]$ and hence a Lebesgue measurable subset of $[0,1]$ [10]. It follows that $\pi_k(Z) = B \cup R$ where B is a Borel set and R is a subset of a Borel set $\tilde{R}$ with $\mu(\tilde{R}) = 0$. Since $B \subseteq \pi_k(Z) \subseteq A$ we have $\mu(B) = 0$ (recall that all Borel measurable subsets of A have measure 0). Therefore

$$\begin{aligned} P[Z] &\leq P\left[\omega \in \Omega : \omega_k \in B \cup \tilde{R}\right] \\ &= \mu(B \cup \tilde{R}) \\ &\leq \mu(B) + \mu(\tilde{R}) = 0 \end{aligned}$$

B. *Randomization without deterministic measurable choice*

This example shows a probability space with random elements $X_k \in C(S_k)$ for all $k \in \mathbb{N}$, so that a form of randomized choice exists, without the measurable choice assumption (Assumption 1). Famous examples in the field of descriptive set theory by Blackwell [31], Novikoff [32], Sierpiński [33], and Addison [34] prove existence of a Borel measurable set $A \subseteq [0,1]^2$ with a projection $\pi_1(A)$ onto the first component that satisfies $\pi_1(A) = [0,1]$ and such that there is no measurable function $\psi : [0,1] \rightarrow [0,1]$ that satisfies $(x, \psi(x)) \in A$ for all $x \in [0,1]$ (see also Example 5.1.7 in [10]). Define

$$C(s) = \{y : (s,y) \in A\} \quad \forall s \in [0,1]$$

Assumption 2 holds for this system because $\{(s,x) \in [0,1]^2 : x \in C(s)\} = A$ is a Borel set, while Assumption 1 fails because there is no measurable choice function $\psi(s)$. Nevertheless a probability space can have i.i.d. random vectors $(S_k, X_k)_{k=1}^{\infty}$ that satisfy $X_k \in C(S_k)$ surely for all $k \in \mathbb{N}$ as follows: Let $(U_k)_{k=0}^{\infty}$ be i.i.d. $\mathcal{U}[0,1]$ variables. Since A is an uncountably infinite Borel measurable subset of $\mathbb{R}^2$, there is an isomorphism $b : [0,1] \rightarrow A$. Define $(S_k, X_k) = b(U_k)$ for all $k \in \mathbb{N}$. Then $(S_k, X_k)_{k=1}^{\infty}$ are indeed i.i.d. random vectors and $(S_k, X_k) \in A$ and so $X_k \in C(S_k)$ surely for all $k \in \mathbb{N}$. Fix $k \in \mathbb{N}$. Using randomization variable U_0 with Corollary 2 implies $X_k = h_k(S_k, W_k)$ for some Borel measurable function h_k and some random variable $W_k \sim \mathcal{U}[0,1]$ that is independent of S_k. Since Assumption 2 holds, any probability space that contains a random element (S, W) with the same distribution as (S_k, W_k) yields

$$P[h_k(S, W) \in C(S)] = 1$$

VI. Conclusion

This paper considers sequences of Borel measurable functions where each function X_k is constrained to be measurable with respect to the sigma algebra generated by the union of an arbitrary number of sigma algebras associated with index k. Specifically, each X_k is $\sigma(\cup_{j\in J_k}\mathcal{H}_j)$-measurable, where J_k is an arbitrary index set and $\mathcal{H}_j$ is a sigma algebra for each $j \in J_k$. It is shown that each X_k can be expressed as the composition of a Borel function h_k with some real-valued measurable functions Y_j for $j \in J_k$, each Y_j being measurable with respect to the sigma algebra $\mathcal{H}_j$. The same Y_j functions can be used to represent the influence of $\mathcal{H}_j$ for all indices k in which $j \in J_k$. For applications to stochastic control, this enables functional representations of all possible decision vectors that satisfy causality and measurability constraints. This leads to a refined theorem on network capacity for opportunistic scheduling in time-varying wireless networks. The theorem uses two measurability assumptions, including an assumption on existence of a measurable choice function. By utilizing classical pathological counter-examples in the field of descriptive set theory, an example opportunistic scheduling system is developed for which a nonmeasurable policy yields significantly better time averages in comparison to any measurable policy.

ACKNOWLEDGEMENTS

This work was supported in part by one or more of: NSF CCF-1718477, NSF SpecEES 1824418.

REFERENCES

[1] J. L. Doob. *Stochastic Processes*. Wiley, 1953.
[2] O. Kallenberg. *Foundations of Modern Probability, 2nd ed., Probability and its Applications*. Springer-Verlag, 2002.
[3] D. Williams. *Probability with Martingales*. Cambridge Mathematical Textbooks, Cambridge University Press, 1991.
[4] D. Blackwell and C. Ryll-Nardzewski. Non-existence of everywhere proper conditional distributions. *Ann. Math. Statist.*, 34(1):223–225, 1963.
[5] K. Kuratowski and C. Ryll-Nardzewski. A general theorem on selectors. *Bull. Polish Acad. des Sciences*, 13:397–403, 1965.
[6] L. E. Dubins and L. J. Savage. *How to Gamble if you Must: Inequalities for Stochastic Processes*. Courier Corp., 2014.
[7] A. Maitra. Discounted dynamic programming on compact metric spaces. *Indian Journal of Statistics, Series A*, 30(2):211–216, 1968.
[8] R. J. Aumann. Measurable utility and the measurable choice theorem. In G. T. Guilbaud, editor, *La Decision*, chapter 15, pages 15–26. Editions du Centre National de la Recherche Scientifique, Paris, 1969.
[9] J. Von Neumann. On rings of operators: Reduction theory. *Annals of Mathematics*, 50:401–485, 1949.
[10] S. M. Srivastava. *A Course on Borel Sets*. Springer New York, NY, 1998.
[11] J. M. Borwein and J. D. Vanderwerff. *Convex Functions: Constructions, Characterizations and Counterexamples*. Cambridge University Press, 2010.
[12] M. J. Neely. *Stochastic Network Optimization with Application to Communication and Queueing Systems*. Morgan & Claypool, 2010.
[13] L. Georgiadis, M. J. Neely, and L. Tassiulas. Resource allocation and cross-layer control in wireless networks. *Foundations and Trends in Networking*, vol. 1, no. 1, pp. 1-149, 2006.
[14] L. Tassiulas and A. Ephremides. Dynamic server allocation to parallel queues with randomly varying connectivity. *IEEE Transactions on Information Theory*, vol. 39, no. 2, pp. 466-478, March 1993.
[15] G. Taraldsen. Optimal learning from the Doob-Dynkin lemma. *arXiv:1801.00974*, 2018.
[16] *https://math.stackexchange.com/questions/3752941/if-x-is-sigma-mathcalg-cup-mathcalh-measurable-can-we-write-x-fg*.
[17] B. K. Driver. *Analysis Tools with Examples*. Springer, New York, 2004.
[18] B. Cascales, V. Kadets, and J. Rodríguez. Measurability and selections of multi-functions in banach spaces. *Journal of Convex Analysis*, 17:229–240, 02 2010.
[19] C. Olech. Existence theorems for optimal problems with vector-valued cost function. *Trans. of American Mathematical Society*, pages 159–180, Feb. 1969.
[20] D. Blackwell. Memoryless strategies in finite-stage dynamic programming. *Annals of Mathematical Statistics*, 1963.
[21] D. Blackwell. Discounted dynamic programming. *Annals of Mathematical Statistics*, 1964.
[22] M. Schäl. On dynamic programming: Compactness of the space of policies. *Stochastic Processes and their Applications*, 3:345–364, 1975.
[23] M. J. Neely, E. Modiano, and C. E Rohrs. Dynamic power allocation and routing for time varying wireless networks. *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 1, pp. 89-103, January 2005.
[24] X. Liu, E. K. P. Chong, and N. B. Shroff. A framework for opportunistic scheduling in wireless networks. *Computer Networks*, vol. 41, no. 4, pp. 451-474, March 2003.
[25] A. Eryilmaz and R. Srikant. Joint congestion control, routing, and MAC for stability and fairness in wireless networks. *IEEE Journal on Selected Areas in Communications, Special Issue on Nonlinear Optimization of Communication Systems*, vol. 14, pp. 1514-1524, Aug. 2006.
[26] A. Stolyar. Maximizing queueing network utility subject to stability: Greedy primal-dual algorithm. *Queueing Systems*, vol. 50, no. 4, pp. 401-457, 2005.
[27] M. J. Neely, E. Modiano, and C. Li. Fairness and optimal stochastic control for heterogeneous networks. *IEEE/ACM Transactions on Networking*, vol. 16, no. 2, pp. 396-409, April 2008.
[28] M. J. Neely. Energy optimal control for time varying wireless networks. *IEEE Transactions on Information Theory*, vol. 52, no. 7, pp. 2915-2934, July 2006.
[29] M. Souslin. Sur une definition des ensembles B sans nombres tranfinis. *C. R. Acad. Sciences, Paris*, 164:88–91, 1917.
[30] N. Lusin. Sur la classification de M. Baire. *C. R. Acad. des Sci, Paris*, 164:91–94, 1917.
[31] D. Blackwell. A Borel set not containing a graph. *The Annals of Mathematical Statistics*, 39(4):1345–1347, 1968.
[32] D. Novikoff. Sur les functions implicites mesurables. *B. Fund. Math.*, 17:8–25, 1931.
[33] W. Sierpiński. Sur deux complémentaires analytiques non séparables. *B. Fund. Math.*, 17:296–297, 1931.
[34] J. W. Addison. Separation principles in the hierarchies of classical and effective descriptive set theory. *Fund. Math*, 46:123–135, 1958.
[35] W. Sierpiński and N. Lusin. Sur une décompostion d'un intervalle en une infinité non dénumberable d'ensembles non measurables. *Comptes Rendus de l'Academie des Sciences*, pages 422–426, 1917.
[36] A. Abian. Partition of nondenumerable closed sets of reals. *Czechoslovak Math. J.*, 26:207–210, 1976.
[37] D. Rizzolo. Strange uniform variables. *ArXiv:1301.7148v1*, 2013.
[38] L. Breiman. *Probability*. SIAM, Philadelphia, 2nd edition, 1968-92.
[39] R. M. Dudley. *Real Analysis and Probability*. Wadsworth Pub. Co., Pacific Grove, CA, 1989.
[40] K. Kuratowski. Sur une généralisation de la notion d'homéomorphie. *Fund. Math*, 22:206–220, 1934.
[41] A. Dembo. Probability theory: Stat 310 / math 230. *Stanford University; unpublished notes*, 2013.
[42] M. Schäl. A selection theorem for optimization problems. *Arch. Math.*, 25:219–224, 1974.
[43] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Transactions on Automatic Control*, vol. 37, no. 12, pp. 1936-1948, Dec. 1992.
[44] Y. S. Chow. On a strong law of large numbers for martingales. *Ann. Math Statist*, vol. 38, no. 2, 1967.