

## ON ws-CONVERGENCE OF PRODUCT MEASURES

ERIK J. BALDER

A number of fundamental results, centered around extensions of Prohorov's theorem, is proven for the ws-topology for measures on a product space. These results contribute to the foundations of stochastic decision theory. They also subsume the principal results of Young measure theory, which only considers product measures with a fixed, common marginal. Specializations yield the criterion for relative ws-compactness of Schäl (1975), the refined characterizations of ws-convergence of Galdéano and Truffert (1997, 1998), and a new version of Fatou's lemma in several dimensions. In a separate, nonsequential development, a generalization is given of the relative ws-compactness criterion of Jacod and Mémin (1981). New applications are given to the existence of optimal equilibrium distributions over player-action pairs in game theory and the existence of most optimistic scenarios in stochastic decision theory.

**1. Introduction.** Several models in operations research, statistics, economics, and optimal control share a common primitive structure. It is characterized by the following aspects:

(1) *Decision rules in their most natural form are functions from some underlying outcome space  $\Omega$  into a set  $S$  of decisions.* If the decision maker adopts a certain function  $f$ , then to each outcome  $\omega$  in  $\Omega$  she associates the decision  $f(\omega)$  in  $S$ .

(2) *The cost of each decision rule is evaluated by aggregation over outcome-decision pairs.* Letting  $g(\omega, s)$  denote the immediate cost of taking decision  $s$  under the outcome  $\omega$ , the aggregate cost of using decision rule  $f$ , which leads to the outcome-decision pairs  $(\omega, f(\omega))$ ,  $\omega \in \Omega$ , could be given by  $\int_{\Omega} g(\omega, f(\omega))\mu(d\omega)$ , where  $\mu$  is some distribution over the outcome space.

We give some examples: (i) Based on an observation  $\omega$  the statistician-decision maker must decide about accepting or rejecting a certain hypothesis. Her decision rules are called test functions and their expected cost is called risk. She is interested in minimizing this risk over a certain subclass of test functions. (ii) Based on the condition  $\omega$  of a machine, the OR-decision maker must decide to perform preventive maintenance or to renew the machine (or not). He does so by means of a (one-period) maintenance policy, and is interested in their expected cost. (iii) Based on a (nonstochastic) time variable  $\omega$ , the controller-decision maker must take an instantaneous control decision. She calls her decision rules control functions and is interested in minimizing some total cost integral over time, whose integrand may also involve instantaneous values of an associated trajectory, over a subset of feasible control functions. (iv) To each player  $\omega$  in a game, the economist-decision maker wishes to assign a certain action. He calls his decision rules action profiles. Although he has no immediate integral criterion by which to judge the effectiveness of an action profile, a related criterion of this type can often be constructed by artificial means (see Example 3.1 below).

It seems rather obvious that the mathematician-decision maker should be interested in the analytical properties of aggregate cost functionals of the type  $J_g : f \mapsto \int_{\Omega} g(\omega, f(\omega))\mu(d\omega)$ ,

Received January 3, 2000; revised December 18, 2000.

*MSC 2000 subject classification.* Primary: 60B05, 62C99.

*OR/MS subject classification.* Primary: Decision analysis; secondary: probability.

*Key words.* Weak-strong topology for product measures, tightness conditions, weak-strong compactness, Prohorov's theorem, Komlós' theorem, Young measures, existence of optimal decision rules, Fatou's lemma in several dimensions.

such as their (semi)continuity, inf-compactness, etc. At its bare minimum, such a study would seem to require that  $\Omega$  be equipped with a  $\sigma$ -algebra  $\mathcal{A}$  and one or more measures  $\mu$ . On the other hand, as is already demonstrated by the case where there is only a single outcome, the action space  $S$  would need a topology for such an analysis to proceed; for additional measurability purposes, one would then use the Borel  $\sigma$ -algebra  $\mathcal{B}(S)$  on  $S$ . Further, given the need for aggregation over  $\Omega$ , the decision rules in this model would have to be measurable as functions from  $(\Omega, \mathcal{A})$  into  $(S, \mathcal{B}(S))$ .

Unless  $\Omega$  is discrete, it is mathematically naive to expect that these barest of requirements would lead to an analytically satisfactory structure for the decision sciences mentioned above. For instance, if  $\Omega$  is the usual Lebesgue unit interval  $[0, 1]$  and if  $S$  is the two-element set  $\{0, 1\}$ , then it is well known that the sequence  $(r_n)$  of Rademacher functions, as defined in Example 1.1, does not have a limit in any classical sense (and neither does any of its subsequences). By Tychonov's theorem, it has convergent *subnets* in the product topology on  $\{-1, 1\}^{[0,1]}$ , but since we wish to integrate over  $[0, 1]$  this observation is to no avail. It is true that by adding  $\frac{1}{2}$  to  $S$  (i.e., by assuming that  $\frac{1}{2}$  allows an interpretation as a decision and by taking  $S := \{0, \frac{1}{2}, 1\}$ ), one can save the day for classical limits in this example:  $(r_n)$  converges weakly in  $L^1([0, 1])$  to the constant  $r_\infty \equiv \frac{1}{2}$ . However, only for a very limited class of cost integrands  $g$  does such weak convergence of  $(r_n)$  to  $r_\infty$  entail convergence of the corresponding risks  $(J_g(r_n))$  to  $J_g(r_\infty)$ . This deficiency can be detected quickly by considering the quadratic cost  $g(\omega, s) := s^2$ . Then  $J_g(r_n) = \int_0^1 r_n^2 = \frac{1}{2}$  for all  $n \in \mathbb{N}$ , but  $J_g(r_\infty) = \frac{1}{4}$ . It can be shown that, for convergence of  $(J_g(r_n))$  to  $J_g(r_\infty)$  to go through in this example,  $g(\omega, s)$  must be linear in  $s$  among other things—note that this also forces us to extend the decision space  $S$  further to  $\mathbb{R}$ .

Given the difficulties of determining the precise cost integrand in many decision problems and the concomitant need for flexibility in their theoretical description, this does not seem to be the right way to proceed. Rather, the analysis should be able to treat a more substantial class of cost integrands  $g$ . The solution is to extend the decision space much more radically, by allowing for *mixed* decisions, i.e., probability measures on the original decision space  $S$  (admittedly, in some decision contexts, mixed decisions could occur quite naturally and already be a part of the standard model). This idea is due to L.C. Young; in turn, his seminal work on generalized curves in the calculus of variations and relaxed control functions in optimal control theory was inspired by Hilbert's twentieth problem. For instance, in the simple example above, we can associate to  $(r_n)$  the mixed decision rule  $\delta_\infty$  that assigns to each  $\omega$  the probability measure that places probability  $\frac{1}{2}$  on decision 0 and  $\frac{1}{2}$  on decision 1. One then gets

$$(1.1) \quad \lim_n J_g(r_n) = \tilde{J}_g(\delta_\infty),$$

not just for  $g(\omega, s) := s^2$  considered above, but for all bounded,  $\mathcal{A} \otimes \mathcal{B}(S)$ -measurable cost integrands  $g : \Omega \times S \rightarrow \mathbb{R}$  for which  $g(\omega, s)$  is continuous in the decision variable  $s$ . Here we denote by

$$\tilde{J}_g(\delta) := \int_\Omega \left[ \int_S g(\omega, s) \delta(\omega)(ds) \right] \mu(d\omega),$$

the aggregate cost of any mixed decision rule  $\delta$ . Mathematically speaking, such mixed decision rules are precisely transition probabilities with respect to  $(\Omega, \mathcal{A})$  and  $(S, \mathcal{B}(S))$  (see below). Going one step further, we can also denote (1.1) as

$$(1.2) \quad \lim_n \int_{\Omega \times S} g d\pi_n = \int_{\Omega \times S} g d\pi_\infty,$$

where  $\pi_n(A \times B) := \mu(A \cap r_n^{-1}(B))$  and  $\pi_\infty(A \times B) := \int_A \delta_\infty(\omega)(B) \mu(d\omega)$  define finite measures on  $(\Omega \times S, \mathcal{A} \otimes \mathcal{B}(S))$ . This is the preferred form in which this paper studies both ordinary decision rules, such as  $r_n$ , and generalized decision rules, such as  $\delta_\infty$ . Also, in this

form they are equipped with the *ws*-topology, which is our main subject. For instance, in (1.2) the sequence  $(\pi_n)$  is stated to converge to  $\pi_\infty$  in the *ws*-topology (cf., Theorem 3.1 below). To summarize, the trade-off discussed here is that “good” convergence properties for the decision rules, with respect to a large class of cost integrands, are bought at the price of enlarging the space of decision rules by allowing for mixed decision rules and/or their associated measures on outcome-decision pairs.

After giving this motivation for the paper, we describe its mathematical framework in greater detail. Let  $(\Omega, \mathcal{A})$  be an abstract measurable space and let  $S$  be a topological space that is completely regular and Suslin; we equip  $S$  with its Borel  $\sigma$ -algebra  $\mathcal{B}(S)$ . As is explained below (see (2.3)ff.), no essential loss of generality is suffered by supposing  $S$  to be a metric Suslin space instead of a completely regular Suslin space, so the reader is at liberty to think of  $S$  as a metric Suslin space. Recall here that a Polish space is a separable space that can be equipped with a metric for which it is complete. Recall also from Definitions III.67 and III.79 in Dellacherie and Meyer (1975) that a Suslin [respectively Lusin] space is a Hausdorff topological space which is the image of a Polish space under a continuous [respectively continuous injective] mapping. Hence, both Polish and Lusin spaces are special instances of Suslin spaces. Another useful reference is Schwartz (1975). Different names and definitions are sometimes attached to these notions as well. For instance, a Borel space in the sense of Definition 7.7 in Bertsekas and Shreve (1978) (that is, a topological space that is homeomorphic to a Borel subset of a Polish space) is precisely a metrizable Lusin space. Indeed, by the second part in Theorem III.17 of Dellacherie and Meyer (1975) such a space is metrizable Lusin and, conversely, any metrizable Lusin space is a Borel space by Definition III.16 of Dellacherie and Meyer (1975).

Let  $\mathcal{M}(\Omega \times S)$  be the set of all finite nonnegative measures on  $(\Omega \times S, \mathcal{A} \otimes \mathcal{B}(S))$ . On this set the following *weak-strong* topology (*ws*-topology for short) was introduced by Schäl (1975) (here, as usual,  $\mathcal{C}_b(S)$  stands for the space of all bounded continuous functions on  $S$ ).

**DEFINITION 1.1.** The *ws-topology* on  $\mathcal{M}(\Omega \times S)$  is the coarsest topology for which all functionals  $\pi \mapsto \int_{A \times S} c(s)\pi(d(\omega, s))$ ,  $A \in \mathcal{A}$ ,  $c \in \mathcal{C}_b(S)$ , are continuous.

This is one of several equivalent definitions discussed in Theorem 3.7 of Schäl (1975). The *ws*-topology is called the “measurable-continuous topology” by Jacod and Mémin (1981) and the “narrow topology” by Galdéano and Truffert (1997, 1998). If a sequence  $(\pi_n)$  in  $\mathcal{M}(\Omega \times S)$  converges in the *ws*-topology to a limit  $\pi_\infty \in \mathcal{M}(\Omega \times S)$ , then this will be indicated by  $\pi_n \rightarrow^{ws} \pi_\infty$ . A standard argument shows the *ws*-topology to be Hausdorff. It “straddles” two classical topologies on  $\mathcal{M}(\Omega)$  and  $\mathcal{M}(S)$ ; here  $\mathcal{M}(\Omega)$  is the set of all nonnegative finite measures on  $(\Omega, \mathcal{A})$  and  $\mathcal{M}(S)$  the set of all such measures on  $(S, \mathcal{B}(S))$ .

**DEFINITION 1.2.**

(i) The *s-topology* on  $\mathcal{M}(\Omega)$  is the coarsest topology for which all functionals  $\lambda \mapsto \lambda(A)$ ,  $A \in \mathcal{A}$ , are continuous.

(ii) The *w-topology* on  $\mathcal{M}(S)$  is the coarsest topology for which all functionals  $\nu \mapsto \int_S c(s)\nu(ds)$ ,  $c \in \mathcal{C}_b(S)$ , are continuous.

Clearly, the *s*-topology is the finest topology on  $\mathcal{M}(\Omega)$  for which  $\pi \mapsto \pi^\Omega := \pi(\cdot \times S)$ , the marginal projection from  $\mathcal{M}(\Omega \times S)$  onto  $\mathcal{M}(\Omega)$ , is continuous. Similarly, the *w*-topology is the finest one on  $\mathcal{M}(S)$  for which  $\pi \mapsto \pi^S := \pi(\Omega \times \cdot)$ , the marginal projection from  $\mathcal{M}(\Omega \times S)$  onto  $\mathcal{M}(S)$ , is continuous. Conversely, it is not possible to describe the *ws*-topology solely in terms of these marginal topologies, because different measures in  $\mathcal{M}(\Omega \times S)$  may have the same marginal projections. Important compactness results for the *s*-topology can be found in Gänsler (1971). The *w*-topology is well known under the name *weak* (or *narrow*) topology. It has been studied extensively in probability and measure theory; e.g., cf., Ash (1972), Billingsley (1968), Dellacherie and Meyer (1975), and Schwartz (1975). Schäl (1975) gives some fundamental results for the *ws*-topology (for  $S$  separable and metric). These include Theorem 3.7 of Schäl (1975), which extends the classical portmanteau theorem. In Theorem 3.10 of Schäl (1975), he also gives a criterion

for relative  $ws$ -compactness, but only in terms of  $w$ -compactness in  $\mathcal{M}(\Omega \times S)$  (see Corollary 2.2). For this, he additionally supposes  $\Omega$  to be topological. As also shown by him, the  $ws$ -topology leads naturally to a topology for policy-induced measures, the  $ws^\infty$ -topology, that is useful for existence in stochastic dynamic programming; see Nowak (1988) and Balder (1989b, 1992) for related subsequent work. Independently, Jacod and Mémmin (1981) also study the  $ws$ -topology. Their choice for a Polish space  $S$  opens up a richer variety of results (the present paper's more frugal choice for a completely regular Suslin space  $S$  does the same). While their portmanteau-type Proposition 2.4 is still covered by Theorem 3.7 of Schäl (1975), their Theorem 2.16 goes considerably further. In their Theorem 2.8, Jacod and Mémmin (1981) also give a relative  $ws$ -compactness result that goes further than the corresponding result of Schäl (1975) in that it addresses the situation where the measurable space  $(\Omega, \mathcal{A})$  is abstract (but with  $S$  Polish, as already mentioned before). The portmanteau-type results of Jacod and Mémmin (1981), cited above, were recently refined by Galdéano (1997) in her doctoral thesis and by Galdéano and Truffert (1998), notably in connection with variational convergence. Like Jacod and Mémmin (1981), they use abstract  $(\Omega, \mathcal{A})$  and Polish  $S$ .

The foundations for the  $ws$ -topology in Schäl (1975), which lie in statistical decision theory, have much in common with the foundations of what is now often called Young measure theory. See Warga (1972) and, more recently, Balder (1984b, 1995, 2000a, b), Valadier (1990), and Pedregal (1997). The principal object of study there is the so-called narrow topology (alias Young measure topology) for transition probabilities, which we now recall. A *transition measure* (alias *kernel*) with respect to  $(\Omega, \mathcal{A})$  and  $(S, \mathcal{B}(S))$  is a mapping  $\tilde{\delta} : \Omega \mapsto \mathcal{M}(S)$ , such that  $\tilde{\delta}(\cdot)(B) : \omega \mapsto \tilde{\delta}(\omega)(B)$  is  $\mathcal{A}$ -measurable for every  $B \in \mathcal{B}(S)$ . We denote the set of all such transition measures by  $\mathcal{T}(\Omega; S)$ . A *transition probability* (alias *Markov kernel*) is a transition measure  $\tilde{\delta} \in \mathcal{T}(\Omega; S)$  which takes only values in the set  $\mathcal{P}(S)$  of all probability measures on  $S$ . Let  $\mathcal{R}(\Omega; S)$  be the set of all such transition probabilities; thus, for  $\tilde{\delta} \in \mathcal{T}(\Omega; S)$ , we have  $\tilde{\delta} \in \mathcal{R}(\Omega; S)$  if and only if  $\tilde{\delta}(\cdot)(S) \equiv 1$ . See §2.6 in Ash (1972), Definition IX.1 of Dellacherie and Meyer (1975), and §III.2 of Neveu (1965) for technical backup. In Young measure theory, the measurable space  $(\Omega, \mathcal{A})$  is endowed with a *fixed* measure  $\mu \in \mathcal{M}(\Omega)$ . Now there corresponds to every  $\tilde{\delta} \in \mathcal{T}(\Omega; S)$ —and in particular to every  $\tilde{\delta} \in \mathcal{R}(\Omega; S)$ —a canonical product measure (possibly infinite) on  $(\Omega \times S, \mathcal{A} \otimes \mathcal{B}(S))$ ; it is given by

$$(\mu \otimes \tilde{\delta})(E) := \int_{\Omega} \tilde{\delta}(\omega)(E_{\omega}) \mu(d\omega), \quad E \in \mathcal{A} \otimes \mathcal{B}(S).$$

E.g., see §2.6 in Ash (1972) or §III.2 in Neveu (1965). Here,  $E_{\omega}$  denotes the section of  $E$  at  $\omega$ . Observe that  $\mu \otimes \tilde{\delta} \in \mathcal{M}(\Omega \times S)$  whenever  $\tilde{\delta}(\cdot)(S)$  is  $\mu$ -integrable. The following definition was given in Balder (1984b); as we mentioned above, it is stated for a fixed measure  $\mu$ :

**DEFINITION 1.3.** The *narrow topology* on  $\mathcal{R}(\Omega; S)$  is the coarsest topology for which all functionals  $\delta \mapsto \int_{A \times S} c(s)(\mu \otimes \delta)(d(\omega, s))$ ,  $A \in \mathcal{A}$ ,  $c \in \mathcal{C}_b(S)$ , are continuous.

Note that this is one of several equivalent definitions; cf., Theorem 2.2 of Balder (1988). Clearly, the mapping  $\delta \mapsto \mu \otimes \delta$  is continuous from  $\mathcal{R}(\Omega; S)$  into  $\mathcal{M}(\Omega \times S)$ . Moreover, if we identify transition probabilities that only differ on a  $\mu$ -null set, then Definitions 1.1 and 1.3 entail that the mapping  $\delta \mapsto \mu \otimes \delta$  is a homeomorphism between  $\mathcal{R}(\Omega; S)$ , endowed with the (quotient) narrow topology, and the subset  $\Pi_{\mu} := \{\pi \in \mathcal{M}(\Omega \times S) : \pi^{\Omega} = \mu\}$  of  $\mathcal{M}(\Omega \times S)$ , endowed with the relative  $ws$ -topology. Indeed, by a well-known disintegration result (see (2.1) below),  $\Pi_{\mu}$  is precisely the set  $\{\mu \otimes \delta : \delta \in \mathcal{R}(\Omega; S)\}$ . Hence, the  $ws$ -topology generalizes the narrow topology for transition probabilities. As shown by the

following example, the connections are less direct in the opposite direction:

EXAMPLE 1.1. Let  $\Omega$  be the unit interval  $[0, 1]$ , equipped with the Lebesgue  $\sigma$ -algebra  $\mathcal{A}$ , and the Lebesgue measure  $\lambda$ . Let  $r_1(\omega) := 1$  if  $\omega \in [0, \frac{1}{2}]$  and  $r_1(\omega) := 0$  if  $\omega \in (\frac{1}{2}, 1]$ , and extend  $r_1$  to  $\mathbb{R}$  by periodicity with period 1. Let  $r_n(\omega) := r_1(2^{n-1}\omega)$ . Consider the sequence  $(\mu_n)$  in  $\mathcal{M}(\Omega)$ , given by  $\mu_n(A) := \int_A r_n d\lambda$ . Then it follows by standard arguments that  $\mu_n(A) \rightarrow \mu_\infty(A) := \lambda(A)/2$  for every  $A \in \mathcal{A}$ . Consider also the sequence  $(\delta_n)$  in  $\mathcal{R}(\Omega; S)$ , defined by  $\delta_n(\omega) := \epsilon_{r_n(\omega)}$  (we use  $S := \{0, 1\}$ ). Here,  $\epsilon_a$  is the usual notation for the Dirac point measure at  $a \in \{0, 1\}$ . By the same sort of argument (see Example 2.6 in Balder 1988), it follows that  $(\delta_n)$  converges narrowly to the constant transition probability  $\delta_\infty \in \mathcal{R}(\Omega; S)$ , defined by  $\delta_\infty(\omega) := (\epsilon_0 + \epsilon_1)/2$ . This holds both when  $\Omega$  is equipped with  $\lambda$  or with  $\mu_\infty = \lambda/2$ . Now  $\mu_n \otimes \delta_n \xrightarrow{ws} \pi_\infty$ , with  $\pi_\infty := \mu_\infty \otimes \epsilon_1$ . To see this, observe that for every  $A \in \mathcal{A}$  and  $c \in \mathcal{C}_b(S)$ , one has  $\int_{A \times S} c d(\mu_n \otimes \delta_n) = \int_A r_n(\omega) c(r_n(\omega)) \lambda(d\omega) = \mu_n(A) c(1) \rightarrow \mu_\infty(A) c(1)$ . Consequently, we do *not* have  $\mu_n \otimes \delta_n \xrightarrow{ws} \mu_\infty \otimes \delta_\infty$ .

While this example shows that the reverse direction is not without some intricacy, this paper will show that, nevertheless, the reverse route is still a viable one, which leads to many new results for the ws-topology. Our principal tool on this route is a canonical re-decomposition of the product measures. Namely, relative compactness and related questions for the ws-topology, including all questions involving the ws-convergence of *sequences*, can essentially be resolved by a rather refined apparatus developed for the study of narrow convergence of transition probabilities, that is to say, by modern Young measure theory. Given the results already obtained within this theory (see Balder 1984b, 1988, 1995, 2000a, b), we shall describe the route in detail, but not all the details to which it leads, for this would be unnecessarily repetitive. Instead, we present some major results that have currently no counterpart whatsoever in the cited literature on the ws-topology. These include the following: (i) Theorem 2.2, a simultaneous generalization of Prohorov’s Theorem 2.1 and Komlós Theorem 2.3, (ii) Theorem 2.6, a complete, useful characterization of sequential ws-convergence in terms of Komlós-convergence (i.e., in terms of *pointwise w*-convergence of averages), (iii) Theorem 3.2, an upper semicontinuity result for the pointwise support sets of a ws-convergent sequence, and (iv) Theorems 2.4 and 5.1; these form two further extensions of Prohorov’s Theorem 2.1 and generalize the above-mentioned compactness criteria of Schäl (1975) and Jacod and Mémin (1981) (see Corollary 2.2 and Theorem 5.2). The usefulness of these results is illustrated by some applications, including a new version of Fatou’s lemma in several dimensions (Theorem 4.1). An application to the upper semicontinuity of Cournot-Nash equilibrium correspondences is given in Balder and Yannelis (2000).

**2. Three Prohorov-type theorems.** Recall from Theorem 1 of Valadier (1973) that a measure  $\pi \in \mathcal{M}(\Omega \times S)$  can be decomposed (or disintegrated) as follows: There exists a transition probability  $\delta_\pi$  in  $\mathcal{R}(\Omega; S)$  (the latter set was introduced in §1), such that

$$(2.1) \quad \pi(E) = \int_\Omega \delta_\pi(\omega)(E_\omega) \pi^\Omega(d\omega), \quad E \in \mathcal{A} \otimes \mathcal{B}(S).$$

Notice that this decomposition yields a transition *probability*  $\delta_\pi$  (determined uniquely by  $\pi$  modulo a  $\mu$ -null set), even though  $\pi$  itself need not be a probability measure. In terms of §1, (2.1) states that  $\pi$  can be decomposed into the product measure  $\pi^\Omega \otimes \delta_\pi$ . Observe also that the condition in Valadier (1973) that the marginal  $\pi^S$  of  $\pi$  be Radon follows by Theorem III.69 of Dellacherie and Meyer (1975), in view of the fact that  $S$  is Suslin. Now suppose that  $\Pi \subset \mathcal{M}(\Omega \times S)$  is such that the collection  $\Pi^\Omega$  of its  $\Omega$ -marginals, defined by  $\Pi^\Omega := \{\pi^\Omega : \pi \in \Pi\}$ , is dominated by some  $\mu \in \mathcal{M}(\Omega)$  (from now on, this will be called *marginal domination* of  $\Pi$  by  $\mu$ ). Correspondingly, for any  $\pi \in \Pi$ , we indicate by  $\tilde{\phi}_\pi \in \mathcal{L}^1_{\mathbb{R}}(\Omega, \mu)$  an arbitrary but fixed version of the Radon-Nikodym density

of  $\pi^\Omega$  with respect to  $\mu$ . Then (2.1) can be restated as follows (from now on we call this *redcomposition*):

$$(2.2) \quad \pi(E) = \int_{\Omega} \tilde{\delta}_\pi(\omega)(E_\omega)\mu(d\omega), \quad E \in \mathcal{A} \otimes \mathcal{B}(S).$$

That is to say, every  $\pi \in \Pi$  can also be decomposed as  $\mu \otimes \tilde{\delta}_\pi$ , where  $\tilde{\delta}_\pi \in \mathcal{T}(\Omega; S)$  is now a transition *measure*; it is given by  $\tilde{\delta}_\pi(\omega) := \check{\phi}_\pi(\omega)\delta_\pi(\omega)$ . Observe that this implies  $\check{\phi}_\pi = \tilde{\delta}_\pi(\cdot)(S)$ . The notation for  $\delta_\pi$ ,  $\tilde{\delta}_\pi$  and  $\check{\phi}_\pi$  introduced here will also be used in the sequel. Particular examples of marginally dominated sets  $\Pi$  are:

- (i) Any sequence  $(\pi_n)$  in  $\mathcal{M}(\Omega \times S)$ .
- (ii) Any subset  $\Pi$  of  $\mathcal{M}(\Omega \times S)$  for which  $\Pi^\Omega$  is relatively  $s$ -compact.

Here the first case is evident (e.g.,  $\mu := \sum_n 2^{-n} \pi_n^\Omega / (1 + \pi_n(\Omega \times S))$  marginally dominates  $(\pi_n)$ ), and the second case follows by Proposition 2.2 below. To some extent, the fact that sequences are always marginally dominated, regardless of relative  $s$ -compactness of the marginals, explains the finer results that we shall obtain for sequences. The following definition is classical; see Billingsley (1968), Dellacherie and Meyer (1975), or Schwartz (1975):

DEFINITION 2.1. A set  $M \subset \mathcal{M}(S)$  is *tight* if for every  $\epsilon > 0$  there is a compact  $K_\epsilon \subset S$  such that  $\sup_{\nu \in M} \nu(S \setminus K_\epsilon) < \epsilon$ .

We recall Prohorov’s famous theorem (Theorem 1.12, page 170 of Prohorov 1956). It asserts that tightness in the classical sense of Definition 2.1, together with boundedness, constitutes a sufficient condition for both relative sequential  $w$ -compactness and relative (topological)  $w$ -compactness in  $\mathcal{M}(S)$ :

THEOREM 2.1 (PROHOROV). *If  $M \subset \mathcal{M}(S)$  is tight and bounded, then*

- (i)  *$M$  is relatively sequentially  $w$ -compact;*
- (ii)  *$M$  is relatively  $w$ -compact.*

Recall that  $M$  is said to be *bounded* if  $\sup_{\nu \in M} \nu(S)$  is finite. Part (i) of this theorem can be found in Theorem 6.1 of Billingsley (1968) and part (ii) in Theorem III.59 of Dellacherie and Meyer (1975). A fine point in part (i) is that Billingsley (1968) requires  $S$  to be metrizable. However, our completely regular Suslin space  $S$  has a *weak metric*, i.e., a metric  $d$  whose topology is not finer than the original topology on  $S$  (e.g., see Theorem III.66 of Dellacherie and Meyer 1975). Indeed, observe that by complete regularity the functions in  $\mathcal{C}_b(S)$  separate the points of  $S$ . Hence, by the Suslin property and Lemma III.31 of Castaing and Valadier (1977), a countable subcollection  $(c_i)$  in  $\mathcal{C}_b(S)$  already separates the points. So

$$(2.3) \quad d(s, z) := \sum_{i=1}^{\infty} 2^{-i} (1 + \|c_i\|_\infty)^{-1} |c_i(s) - c_i(z)|$$

forms a weak metric on  $S$  (here  $\|c_i\|_\infty := \sup_S |c_i|$ ). It follows that  $(S, d)$  is also Suslin, and on compact sets the two topologies are actually equivalent. Moreover, the corresponding Borel  $\sigma$ -algebras coincide by Corollary 2, page 101, of Schwartz (1975). From these facts it is not hard to deduce that the above part (i) of the theorem still holds in our setting (cf. Theorems 2.4, 2.5 in Balder 2000a). We now extend tightness as in Definition 2.1 in two versions. The first of these comes from Young measure theory (see Berliocchi and Lasry 1973, and Balder 1979, 1984b), where it is simply called *tightness*. We shall use it to extend Theorem 2.1(i), i.e., the sequential part of Prohorov’s theorem, in two different forms (see Theorem 2.2 and Corollary 2.1 below). The second version of tightness, which we call *ws-tightness*, is more demanding. It serves for extensions to the  $ws$ -topology of both the sequential part (i) of Prohorov’s Theorem 2.1 and the nonsequential part (ii). This is done in Theorems 2.4 and 5.1 respectively.

DEFINITION 2.2.

(i) A set  $\Pi \subset \mathcal{M}(\Omega \times S)$  is *tight* if there exists a  $\mathcal{A} \otimes \mathcal{B}(S)$ -measurable function  $h : \Omega \times S \rightarrow [0, +\infty]$ , such that the set  $\{s \in S : h(\omega, s) \leq \beta\}$  is compact for every  $\omega \in \Omega$  and  $\beta \in \mathbb{R}_+$  and such that  $\sup_{\pi \in \Pi} \int_{\Omega \times S} h d\pi < +\infty$ .

(ii) A set  $\Pi \subset \mathcal{M}(\Omega \times S)$  is *ws-tight* if  $\Pi$  is tight and  $\Pi^\Omega$  is relatively  $s$ -compact.

Observe that ws-tightness of  $\Pi \subset \mathcal{M}(\Omega \times S)$  implies that  $\sup_{\pi \in \Pi} \pi(\Omega \times S) < +\infty$ , i.e.,  $\Pi$  is *bounded* (just note that  $\lambda \mapsto \lambda(\Omega)$  is  $s$ -continuous on the compact  $s$ -closure of  $\Pi^\Omega$ ). To compare the new definition of tightness with the classical one in Definition 2.1, we give an equivalent version of part (i) of Definition 2.2. We do so by means of the following proposition (cf. Jawhar 1984, and Exercise 10 on page 109 of Bourbaki 1974, Chapter 5).

PROPOSITION 2.1. *For every  $\Pi \subset \mathcal{M}(\Omega \times S)$  the following are equivalent:*

- (a)  $\Pi$  is tight in the sense of Definition 2.2(i).
- (b) For every  $\epsilon > 0$ , there exists a compact-valued multifunction  $\Gamma_\epsilon : \Omega \rightarrow 2^S$ , with  $\mathcal{A} \otimes \mathcal{B}(S)$ -measurable graph  $\text{gph } \Gamma_\epsilon$ , such that  $\sup_{\pi \in \Pi} \pi((\Omega \times S) \setminus \text{gph } \Gamma_\epsilon) < \epsilon$ .

Here  $\text{gph } \Gamma_\epsilon := \{(\omega, s) \in \Omega \times S : s \in \Gamma_\epsilon(\omega)\}$ .

PROOF. (a)  $\Rightarrow$  (b): Let  $h$  be as in Definition 2.2(i). Take  $\Gamma_\epsilon(\omega) := \{s \in S : h(\omega, s) \leq \sigma/\epsilon\}$ , with  $\sigma := \sup_{\pi \in \Pi} \int h d\pi$ . Then, clearly,  $\Gamma_\epsilon$  has a measurable graph and compact values. To see that the inequality holds also, we simply observe that  $\sigma \geq \int_{(\Omega \times S) \setminus \text{gph } \Gamma_\epsilon} \sigma/\epsilon d\pi$  holds for all  $\pi \in \Pi$ .

(b)  $\Rightarrow$  (a): Take  $\epsilon := 3^{-n}$ ; rather than taking finite unions of multifunctions, we can suppose without loss of generality that the multifunctions  $\Gamma_{1/3^n}$  are pointwise nondecreasing (in  $n$ ). Now, set  $h(\omega, s) := 2^n$  if  $s \in \Gamma_{1/3^{n+1}}(\omega) \setminus \Gamma_{1/3^n}(\omega)$  and  $h(\omega, s) := 0$  if  $s \in \Gamma_{1/3}(\omega)$ . Then it is easy to see that  $h$  has the properties required in Definition 2.2(i).  $\square$

It is clear from this proposition that classical tightness as in Definition 2.1 is generalized by tightness in the sense of Definition 2.2 (simply trivialize the space  $(\Omega, \mathcal{A})$  by taking  $\Omega$  equal to a singleton or by setting  $\mathcal{A} := \{\emptyset, \Omega\}$ ).

PROPOSITION 2.2. *For every  $\Pi \subset \mathcal{M}(\Omega \times S)$ , the following are equivalent:*

- (a)  $\Pi^\Omega$  is relatively  $s$ -compact.
- (b)  $\Pi^\Omega$  is relatively sequentially  $s$ -compact.
- (c)  $\Pi^\Omega$  is dominated by a measure  $\mu \in \mathcal{M}(\Omega)$  and the corresponding collection  $\{\tilde{\phi}_\pi : \pi \in \Pi\}$  of densities is uniformly  $\mu$ -integrable.

PROOF. Each of (a), (b), and (c) implies boundedness of  $\Pi^\Omega$  (i.e.,  $\sup_{\pi \in \Pi} \pi^\Omega(S) < +\infty$ ). So the equivalences hold by Theorem 2.6 of Gänssler (1971). Observe that 2.6(iii) of Gänssler (1971) states only uniform absolute continuity, but, in combination with  $\sup_{\pi \in \Pi} \int_\Omega \tilde{\phi}_\pi d\mu = \sup_{\pi \in \Pi} \pi^\Omega(S) < +\infty$ , this yields uniform  $\mu$ -integrability as stated in (c) (apply Proposition II.5.2 in Neveu 1965).  $\square$

This shows that the relative  $s$ -compactness condition in Definition 2.2(ii) can be stated in several alternative ways. The next result applies in particular when  $\Pi \subset \mathcal{M}(\Omega \times S)$  is ws-tight; its version for the narrow topology for transition probabilities is well known. The proof does not make any use of the Suslin property of  $S$  (it only uses the separability and metrizability of  $S$ ); thus, this proposition extends Remark 3.11 of Schäl (1975).

PROPOSITION 2.3. *Suppose that  $\mathcal{A}$  is countably generated and  $S$  is metrizable (i.e.,  $S$  is metrizable Suslin). Then every  $\Pi \subset \mathcal{M}(\Omega \times S)$ , such that  $\Pi^\Omega$  is relatively  $s$ -compact, is metrizable for the ws-topology.*

PROOF. By Proposition 2.2, there exists a dominating measure  $\mu \in \mathcal{M}(\Omega)$  for  $\Pi^\Omega$ . By hypothesis, there exists a countable (at most) algebra  $\mathcal{A}_0 \subset \mathcal{A}$ , which generates the  $\sigma$ -algebra  $\mathcal{A}$ . Let us write  $\mathcal{A}_0 := \{A_j : j \in \mathbb{N}\}$ . By Proposition 7.19 of Bertsekas and Shreve (1978), there exists a countable subset  $(c'_i)$  of  $\mathcal{C}_b(S)$ , such that for any net  $(\nu_\gamma)$  in  $\mathcal{M}(S)$

and any  $\bar{\nu} \in \mathcal{M}(S)$ , the following is true:  $\lim_{\gamma} \int_S c'_i d\nu_{\gamma} = \int_S c'_i d\bar{\nu}$  for every  $i \in \mathbb{N}$  implies  $\bar{\nu} = w - \lim_{\gamma} \nu_{\gamma}$ . Now set

$$\rho(\pi, \pi') := \sum_{i,j} 2^{-i-j} (1 + \|c'_i\|_{\infty})^{-1} \left| \int_{A_j \times S} c'_i d\pi - \int_{A_j \times S} c'_i d\pi' \right|.$$

First, observe that this defines a metric on  $\mathcal{M}(\Omega \times S)$ , which is not finer than the ws-topology. It remains to prove that  $\bar{\pi} = ws - \lim_{\gamma} \pi_{\gamma}$  for any net  $(\pi_{\gamma})$  in  $\Pi$ , and any  $\bar{\pi} \in \Pi$ , such that  $\lim_{\gamma} \rho(\pi_{\gamma}, \bar{\pi}) = 0$ . To this end, let  $A \in \mathcal{A}$  and  $c \in \mathcal{C}_b(S)$  be arbitrary. Define  $\bar{\pi}^A := \bar{\pi}(A \times \cdot)$  and  $\pi_{\gamma}^A := \pi_{\gamma}(A \times \cdot)$  in  $\mathcal{M}(S)$ . By the above property of  $(c'_i)$ , the hypothesis  $\lim_{\gamma} \rho(\pi_{\gamma}, \bar{\pi}) = 0$  implies  $\bar{\pi}^A = w - \lim_{\gamma} \pi_{\gamma}^A$  for every  $A$ . In particular, this gives  $\lim_{\gamma} \int_{A_j \times S} c d\pi_{\gamma} = \int_{A_j \times S} c d\bar{\pi}$  for every  $A_j$ . By Theorem 1.3.11 in Ash (1972), there exists for every  $\epsilon > 0$  a set  $A_j \in \mathcal{A}_0$ , such that  $\int_{\Omega} |1_A - 1_{A_j}| d\mu < \epsilon$ . Using boundedness of  $\Pi$ , it follows that on  $\Pi$ , the functional  $\pi \mapsto \int_{A \times S} c d\pi$  is the uniform limit of a certain sequence of functionals  $\pi \mapsto \int_{A_j \times S} c d\pi$ . Therefore, we conclude that  $\lim_{\gamma} \int_{A \times S} c d\pi_{\gamma} = \int_{A \times S} c d\bar{\pi}$ .  $\square$

REMARK 2.1. In their Proposition 2.10, Jacod and Mémmin (1981) claim that  $\mathcal{M}(\Omega \times S)$ , as a whole, is metrizable for the ws-topology if  $\mathcal{A}$  is countably generated, regardless of any  $s$ -compactness of marginals. The present author does not know a counterexample to this claim, but wishes to point out that the proof of Proposition 2.10 on page 535 of Jacod and Mémmin (1981) is unconvincing. Namely, for  $\Omega := [0, 1]$  and trivial  $S$  it already breaks down for the sequence  $(\epsilon_{1/2^n})$  and  $\epsilon_0$  in  $\mathcal{M}(\Omega)$ . In that situation,  $\mathcal{A}_0$ , the algebra of finite disjoint unions of right-open and left-closed intervals with rational endpoints, generates  $\mathcal{A} := \mathcal{B}([0, 1])$ . But while  $\epsilon_{1/2^n}(A) \rightarrow \epsilon_0(A)$  for every  $A \in \mathcal{A}_0$ , which is in complete accordance with the hypotheses on page 535 of Jacod and Mémmin (1981), we have  $\epsilon_{1/2^n}(B) \not\rightarrow \epsilon_0(B)$  for  $B := \{1/2^j : j \in \mathbb{N}\}$ .

The remainder of this section is devoted to three different extensions of the sequential part (i) of Theorem 2.1, and to an associated characterization of sequential ws-convergence. Given this sequential orientation, it should not come as a surprise that it only makes use of the *sequential* compactness of the subsets of  $S$  used in Definition 2.2(i) (by using the weak metric of (2.3), it is clear that such sequential compactness is implied by compactness—note that the converse need not be true). In other words, for the sole purpose of extending part (i) of Prohorov’s Theorem 2.1, one could phrase Definition 2.2(i) in terms of sequential compactness; this was done in Balder (1989c, 1990, 1995, 2000a, b). In Balder (1989c, 1990), the following intermediate, nontopological mode of convergence was introduced and studied in a more abstract context. For sequences  $(\pi_n)$  in  $\mathcal{M}(\Omega \times S)$ , we shall use it to characterize ws-convergence completely in terms of the associated sequence  $(\tilde{\delta}_{\pi_n})$  in  $\mathcal{T}(\Omega; S)$ .

DEFINITION 2.3. Given  $\mu \in \mathcal{M}(\Omega)$ , a sequence  $(\tilde{\delta}_n)$  of transition measures in  $\mathcal{T}(\Omega; S)$  *K-converges under  $\mu$*  to  $\tilde{\delta}_{\infty} \in \mathcal{T}(\Omega; S)$  (notation:  $\tilde{\delta}_n \xrightarrow{\mu, K} \tilde{\delta}_{\infty}$ ) if for every subsequence  $(\tilde{\delta}_{n_j})$  of  $(\tilde{\delta}_n)$  there is a  $\mu$ -null set  $N$  in  $\mathcal{A}$ —possibly depending on that subsequence—such that

$$(2.4) \quad \frac{1}{m} \sum_{j=1}^m \tilde{\delta}_{n_j}(\omega) \xrightarrow{w} \tilde{\delta}_{\infty}(\omega) \quad \text{in } \mathcal{M}(S) \text{ for every } \omega \in \Omega \setminus N.$$

EXAMPLE 2.1.

(i) Independent and identically distributed sequences in  $\mathcal{R}(\Omega; S)$  provide concrete and interesting examples of *K*-convergence. For instance, let  $\Omega := [0, 1]$  be equipped with the Lebesgue measure  $\mu$  and let  $(r_n)$  be the sequence of Rademacher functions  $r_n(\omega) := \text{sgn}(\sin(2^n \pi \omega))$  (set  $\text{sgn } 0 := 1$  by default). For  $S := \{1, -1\}$ , we can define  $\delta_n(\omega) := \epsilon_{r_n(\omega)}$ . Then the random measures  $\delta_n : [0, 1] \rightarrow \mathcal{P}(\{1, -1\})$  are independent and identically distributed. By Kolmogorov’s strong law of large numbers, which can be applied to every

subsequence of  $(\delta_n)$  (observe that  $\mathcal{P}(\{1, -1\})$  has dimension 1), we obtain  $\delta_n \rightarrow^{\mu, K} \delta_\infty$  with  $\delta_\infty \equiv (\epsilon_1 + \epsilon_{-1})/2$ .

(ii) A less interesting illustration of  $K$ -convergence is as follows. Let  $(\tilde{\delta}_n)$  and  $\tilde{\delta}_\infty$  be given in  $\mathcal{T}(\Omega; S)$  with  $\tilde{\delta}_n(\omega) \rightarrow^w \tilde{\delta}_\infty(\omega)$  in  $\mathcal{M}(S)$  for  $\mu$ -a.e.  $\omega$  in  $\Omega$ . Concretely, for  $\Omega := [0, 1]$ , equipped with the Lebesgue  $\sigma$ -algebra  $\mathcal{A}$  and the Lebesgue measure  $\mu$ , and for  $S := \{0\}$  we could take the following sequence  $(\tilde{\delta}_n)$ . For  $\omega \in [0, 1/n]$  let  $\tilde{\delta}_n(\omega)(\{0\}) := n$  and for  $\omega \in (1/n, 1]$  let  $\tilde{\delta}_n(\omega)(\{0\}) := 0$ . Also, let  $\tilde{\delta}_\infty(\omega)(\{0\}) = 0$  for all  $\omega \in \Omega$ . This example also shows that, unlike  $ws$ -convergence in  $\mathcal{M}(\Omega \times S)$  and  $K$ -convergence in  $\mathcal{R}(\Omega; S)$ ,  $K$ -convergence in the space of transition measures  $\mathcal{T}(\Omega; S)$  need not preserve aggregate measure in the limit. Notably, in the above situation,  $(\mu \otimes \tilde{\delta}_n)(\Omega \times S)$  equals 1 for all  $n \in \mathbb{N}$ , but it equals 0 for  $n = \infty$ .

In Example 2.1(i) Kolmogorov’s theorem is actually applied uncountably many times (viz. once for each subsequence). Each such application yields an exceptional  $\mu$ -null set  $N$  (i.e., the null set that figures in Kolmogorov’s limit statement). While Definition 2.3 allows for this, it does not mean perforce that the total number of exceptional null sets  $N$  involved in Definition 2.3 is uncountably infinite as well. For instance, in Example 2.1(ii), one and the same null set can serve for all subsequences. The following fact, however, is elementary: for any  $(\alpha_n)$  and  $\alpha_\infty$  in  $\mathbb{R}$ :

$$(2.5) \quad \lim_m \frac{1}{m} \sum_{j=1}^m \alpha_{n_j} = \alpha_\infty \quad \text{for every subsequence } (\alpha_{n_j}) \text{ of } (\alpha_n) \text{ implies } \alpha_n \rightarrow \alpha_\infty.$$

This means that in Example 2.1(i), the uncountable number of applications of Kolmogorov’s theorem is indeed matched by an uncountable number of exceptional null sets. This finding underlines the importance of the null sets in Definition 2.3: Their plurality distinguishes stronger from weaker modes of convergence in  $\mathcal{T}(\Omega; S)$ .

Next, we state a useful lower semicontinuity property of  $K$ -convergence. This combines a Fatou- and a Fatou-Vitali-type result. Recall here that a *normal integrand* on  $\Omega \times S$  is an  $\mathcal{A} \otimes \mathcal{B}(S)$ -measurable function  $g : \Omega \times S \rightarrow (-\infty, +\infty]$ , such that  $g(\omega, \cdot)$  is lower semicontinuous on  $S$  for every  $\omega \in \Omega$ .

**PROPOSITION 2.4.** *If  $\tilde{\delta}_n \rightarrow^{\mu, K} \tilde{\delta}_\infty$  for  $(\tilde{\delta}_n)$  and  $\tilde{\delta}_\infty$  in  $\mathcal{T}(\Omega; S)$  and  $\mu \in \mathcal{M}(\Omega)$ , then the following hold:*

- (i)  $\liminf_n \int_{\Omega \times S} g d(\mu \otimes \tilde{\delta}_n) \geq \int_{\Omega \times S} g d(\mu \otimes \tilde{\delta}_\infty)$  for every nonnegative normal integrand  $g$  on  $\Omega \times S$ .
- (ii)  $\liminf_n \int_{\Omega \times S} g d(\mu \otimes \tilde{\delta}_n) \geq \int_{\Omega \times S} g d(\mu \otimes \tilde{\delta}_\infty)$  for every normal integrand  $g$  on  $\Omega \times S$  that is bounded below, provided that  $(\tilde{\delta}_n(\cdot)(S))$  is  $\mu$ -uniformly integrable.

**PROOF.** (i) Let  $\beta := \liminf_n \int_{\Omega \times S} g d(\mu \otimes \tilde{\delta}_n)$ . For elementary reasons, there is a subsequence  $(\mu \otimes \tilde{\delta}_{n_j})$  of  $(\mu \otimes \tilde{\delta}_n)$ , such that  $\beta = \lim_j \int_{\Omega \times S} g d(\mu \otimes \tilde{\delta}_{n_j})$ . Set  $\psi_n(\omega) := \int_S g(\omega, s) \tilde{\delta}_n(\omega)(ds)$  for  $n \in \mathbb{N} \cup \{\infty\}$ . Then  $\beta = \lim_j \int_\Omega \psi_{n_j} d\mu$ . Of course, this implies also  $\beta = \lim_m \int_\Omega (1/m) \sum_{j=1}^m \psi_{n_j} d\mu$ . Now (2.4) gives  $\liminf_m (1/m) \sum_{j=1}^m \psi_{n_j}(\omega) \geq \psi_\infty(\omega)$  for  $\mu$ -a.e.  $\omega$ , because the function  $g(\omega, \cdot)$  is lower semicontinuous and nonnegative on  $S$  (apply Theorem III.55 of Dellacherie and Meyer 1975). Hence, an application of Fatou’s lemma gives  $\beta \geq \int_\Omega \psi_\infty d\mu$ . Since  $\int_\Omega \psi_\infty d\mu = \int_{\Omega \times S} g d(\mu \otimes \tilde{\delta}_\infty)$ , the proof of (i) is finished.

(ii) Again we set  $\beta := \liminf_n \int_{\Omega \times S} g d(\mu \otimes \tilde{\delta}_n)$ . As before, there exists a subsequence  $(\tilde{\delta}_{n'})$  of  $(\tilde{\delta}_n)$  for which  $\beta = \lim_{n'} \int_{\Omega \times S} g d(\mu \otimes \tilde{\delta}_{n'})$ . By uniform integrability of  $(\tilde{\delta}_{n'}(\cdot)(S))$  and the Dunford-Pettis theorem, there exist a further subsequence  $(\tilde{\delta}_{n'_j})$  of  $(\tilde{\delta}_{n'})$  and a function  $\psi_* \in \mathcal{L}^1_{\mathbb{R}}(\Omega, \mu)$ , such that

$$(2.6) \quad (\tilde{\delta}_{n'_j}(\cdot)(S)) \text{ converges to } \psi_* \text{ in the weak topology } \sigma(\mathcal{L}^1_{\mathbb{R}}(\Omega, \mu), \mathcal{L}^\infty(\Omega)).$$

By (2.4) we have  $(1/m) \sum_{j=1}^m \tilde{\delta}_{n'_j}(\omega) \rightarrow^w \tilde{\delta}_\infty(\omega)$  for  $\mu$ -a.e.  $\omega$ , so in particular  $(1/m) \times \sum_{j=1}^m \tilde{\delta}_{n'_j}(\omega)(S) \rightarrow \tilde{\delta}_\infty(\omega)(S)$ . Because of (2.6), the same averages  $(1/m) \sum_{j=1}^m \tilde{\delta}_{n'_j}(\cdot)(S)$  also converge weakly to  $\psi_*$  in  $\sigma(\mathcal{L}_\mathbb{R}^1, \mathcal{L}_\mathbb{R}^\infty)$ . As is well known, these two facts together imply  $\tilde{\delta}_\infty(\omega)(S) = \psi_*(\omega)$  for  $\mu$ -a.e.  $\omega$  (use the Lebesgue-Vitali theorem). By hypothesis, there is a constant  $\alpha \in \mathbb{R}$ , such that  $g \geq -\alpha$ . So  $g + \alpha$  is a nonnegative normal integrand on  $\Omega \times S$ . By part (i),

$$\beta + \alpha \liminf_j (\mu \otimes \tilde{\delta}_{n'_j})(\Omega \times S) \geq \int_{\Omega \times S} g d(\mu \otimes \tilde{\delta}_\infty) + \alpha(\mu \otimes \tilde{\delta}_\infty)(\Omega \times S).$$

Here  $\liminf_j (\mu \otimes \tilde{\delta}_{n'_j})(\Omega \times S) = \lim_j \int_\Omega \tilde{\delta}_{n'_j}(\cdot)(S) d\mu = \int_\Omega \psi_* d\mu$ , as follows by (2.6). So, in view of  $\tilde{\delta}_\infty(\cdot)(S) = \psi_*$   $\mu$ -a.e., the inequality simplifies to  $\beta \geq \int_{\Omega \times S} g d(\mu \otimes \tilde{\delta}_\infty)$ .  $\square$

For  $(\tilde{\delta}_n)$  in  $\mathcal{R}(\Omega; S) \subset \mathcal{T}(\Omega; S)$  uniform integrability as in part (ii) of the above proposition holds trivially by  $\tilde{\delta}_n(\cdot)(S) \equiv 1$  for all  $n$ . Hence, the distinction between parts (i) and (ii) in the above proposition is not encountered in Young measure theory.

Our first and central extension of Theorem 2.1(i) can now be stated. It states that tightness is a sufficient condition for “relative compactness” for  $K$ -convergence in  $\mathcal{M}(\Omega \times S)$  (as Komlós-convergence is nontopological, parentheses are called for). Recall from §2 that a sequence  $(\pi_n)$  is *always* marginally dominated by some measure  $\mu \in \mathcal{M}(\Omega)$ , causing every  $\pi_n$ ,  $n \in \mathbb{N}$ , to be decomposable as  $\mu \otimes \tilde{\delta}_{\pi_n}$ , by virtue of (2.2).

**THEOREM 2.2.** *If  $(\pi_n)$  in  $\mathcal{M}(\Omega \times S)$  is tight, bounded, and marginally dominated by  $\mu \in \mathcal{M}(\Omega)$ , then there exist a subsequence  $(\tilde{\delta}_{\pi_{n'}})$  of  $(\tilde{\delta}_{\pi_n})$ , and a transition measure  $\tilde{\delta}_* \in \mathcal{T}(\Omega; S)$ , such that  $\tilde{\delta}_*(\cdot)(S)$  is  $\mu$ -integrable and  $\tilde{\delta}_{\pi_{n'}} \rightarrow^{\mu, K} \tilde{\delta}_*$ .*

If one trivializes  $(\Omega, \mathcal{A})$ , then it is easy to see that Theorem 2.2 reduces to part (i) of Prohorov’s Theorem 2.1 (use (2.5)). On the other hand, if one trivializes  $S$ , then Theorem 2.2 reduces to Komlós’ theorem, which is as follows (see Komlós 1967 or Chatterji 1973):

**THEOREM 2.3 (KOMLÓS).** *Let  $(\psi_n)$  be a sequence in  $\mathcal{L}_\mathbb{R}^1(\Omega, \mu)$ , such that  $\sup_n \int_\Omega |\psi_n| d\mu < +\infty$ . Then there exist a subsequence  $(\psi_{n'})$  of  $(\psi_n)$  and a function  $\psi_* \in \mathcal{L}_\mathbb{R}^1(\Omega, \mu)$ , such that for every further subsequence  $(\psi_{n'_j})$  of  $(\psi_{n'})$  there is a  $\mu$ -null set  $N$ —possibly depending on that subsequence—such that*

$$\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{j=1}^m \psi_{n'_j}(\omega) = \psi_*(\omega) \quad \text{for every } \omega \text{ in } \Omega \setminus N.$$

Beautiful connections exist between Theorem 2.3 and Kolmogorov’s strong law of large numbers; e.g., see Aldous (1977) and Chatterji (1985). Very directly—e.g., see the exercise on page 217 of Stout (1974) or see Valadier (1991)—Theorem 2.3 implies the SLLN, and hence extends it to sequences of nonindependent random variables. Therefore, the SLLN is also extended by Theorem 2.2, which generalizes Theorem 2.3, as we saw above. Theorem 2.2 also extends the very similar Prohorov-Komlós theorem for transition probabilities in Theorem 5.1 of Balder (1990); however, that result does not reduce to Komlós’ theorem if  $S$  is trivial. Our second extension of Theorem 2.1(i) is as follows:

**COROLLARY 2.1.** *If  $(\pi_n)$  in  $\mathcal{M}(\Omega \times S)$  is tight, bounded, and marginally dominated by  $\mu \in \mathcal{M}(\Omega)$ , then there exist a subsequence  $(\pi_{n'})$  of  $(\pi_n)$ , a measure  $\pi_* \in \mathcal{M}(\Omega \times S)$ , marginally dominated by  $\mu$ , and a nonincreasing sequence  $(A_p)$  of sets in  $\mathcal{A}$ , such that  $\lim_p \mu(A_p) = 0$  and*

$$\lim_{n'} \int_{A \times S} c(s) \pi_{n'}(d(\omega, s)) = \int_{A \times S} c(s) \pi_*(d(\omega, s))$$

for every  $p \in \mathbb{N}$ ,  $A \in \mathcal{A}$ ,  $A \subset \Omega \setminus A_p$  and  $c \in \mathcal{C}_b(S)$ .

PROOF. By Theorem 2.2, there exists a subsequence  $(\pi_{n'})$  of  $(\pi_n)$  and  $\tilde{\delta}_* \in \mathcal{F}(\Omega; S)$ , such that  $\tilde{\delta}_*(\cdot)(S)$  is  $\mu$ -integrable and  $\tilde{\delta}_{\pi_{n'}} \rightarrow^{\mu, K} \tilde{\delta}_*$ . The former means that  $\pi_* := \mu \otimes \tilde{\delta}_*$  is well defined in  $\mathcal{M}(\Omega \times S)$  (see §1). Also, we have  $\sup_n \int_{\Omega} \tilde{\phi}_{\pi_n} d\mu = \sup_n \pi_n(\Omega \times S) < +\infty$ . Hence, by the biting lemma (see Gaposhkin 1972, or Brooks and Chacon 1977, p. 17), there exists a sequence  $(A_p)$  that decreases to a null set, such that  $(\tilde{\phi}_{\pi_n})$  is uniformly  $\mu$ -integrable over  $\Omega \setminus A_p$  for every fixed  $p \in \mathbb{N}$ . For  $A \subset \Omega \setminus A_p$ ,  $p \in \mathbb{N}$ , it remains to invoke Proposition 2.4(ii) twice: Set  $\Omega := A$  and set first  $g(\omega, s) := c(s)$ , and then  $g(\omega, s) := -c(s)$ . This gives  $\lim_n \int_{A \times S} c d(\mu \otimes \tilde{\delta}_{\pi_n}) = \int_{A \times S} c d(\mu \otimes \tilde{\delta}_*)$ . In view of (2.2) and the definition of  $\pi_*$ , this finishes the proof.  $\square$

We shall now give a quick proof of Theorem 2.2 by means of the abstract generalization of Komlós' Theorem 2.3, given in Theorem 2.1 of Balder (1990) (see Balder and Hess 1996 for further developments in this direction). This proof requires only a slight extension of the demonstration of Theorem 5.1 of Balder (1990), as given in §5 of that reference. A more elaborate proof of Theorem 2.2, starting directly from Theorem 2.3, could be given along the lines of the proof of Theorem 3.8 in Balder (2000a). (To do this, one first obtains a preliminary subsequence by applying Theorem 2.3 to  $(\tilde{\phi}_{\pi_n})$  and then proceeds similarly to pages 32–33 in Balder 2000a.)

PROOF OF THEOREM 2.2. To apply Theorem 2.1 of Balder (1990), we slightly modify the substitutions made in §5 of Balder (1990). We now take its  $E$  to be  $\mathcal{M}(S)$ , equipped with the  $w$ -topology; this takes the place of  $E = \mathcal{P}(S)$  as in Balder (1990). Consequently, the last line on page 12 of that reference must be adapted as follows:  $h(\omega, x) := \int_S h'(\omega, s)x(ds) + x(S)$ ,  $x \in E$  (here  $h'$  plays the same role as  $h$  in Definition 2.2). This causes  $h(\omega, \cdot)$  to be sequentially  $w$ -inf-compact on  $\mathcal{M}(S)$  for every  $\omega \in \Omega$  by Prohorov's Theorem 2.1. Also, the definition of  $a^g$  on page 13 of Balder (1990) must be slightly adapted: We still define  $a^g : \Omega \times \mathcal{M}(S) \rightarrow \mathbb{R}$  by  $a^g(\omega, x) := \int_{\Omega \times S} g(\omega, s)x(ds)$ , but this time we use the bounded Carathéodory functions, i.e., bounded  $\mathcal{A} \otimes \mathcal{B}(S)$ -measurable  $g : \Omega \times S \rightarrow \mathbb{R}$ , such that  $g(\omega, \cdot)$  is continuous on  $S$  for every  $\omega \in \Omega$ . Let  $\|g\|_{\infty} := \sup_{\Omega \times S} |g|$ ; then the inequality  $|a^g(\omega, x)| \leq \|g\|_{\infty}x(S) \leq \|g\|_{\infty}h(\omega, x)$  shows that Condition (B) on page 3 of Balder (1990) continues to hold. The result now follows from Theorem 2.1 of that same reference, as shown in its §5.  $\square$

Observe that Theorem 2.2 and Corollary 2.1 require tightness, but not  $ws$ -tightness. This allows for situations where all marginal projections  $\pi_n^{\Omega}$ ,  $n \in \mathbb{N}$ , are absolutely continuous with respect to some given  $\mu \in \mathcal{M}(\Omega)$ , but where  $\pi_*^{\Omega}$  is not absolutely continuous with respect to  $\mu$ :

EXAMPLE 2.2. Let  $\Omega := [0, 1]$  be equipped with the Lebesgue  $\sigma$ -algebra  $\mathcal{A}$  and the Lebesgue measure  $\mu$ . Let  $S := \{0\}$  and define  $\pi_n \in \mathcal{M}(\Omega \times S)$  by  $\pi_n(A \times S) := n\mu(A \cap [0, 1/n])$ . Here all  $\pi_n^{\Omega}$ ,  $n \in \mathbb{N}$ , are absolutely continuous with respect to  $\mu$ . Now  $(\pi_n)$  is tight (take  $\Gamma_{\epsilon} \equiv S = \{0\}$  in Proposition 2.1), but not  $ws$ -tight (notice that  $\pi_n^{\Omega} \rightarrow^w \epsilon_0$ , but not  $\pi_n^{\Omega} \rightarrow^s \epsilon_0$ ). Yet Corollary 2.1 applies, and from the preceding analysis, one sees immediately that any nonincreasing sequence  $(A_p)$  will do for which  $\bigcap_p A_p = \{0\}$ . For  $(\pi_{n'})$ , one can simply take  $(\pi_n)$  itself, and for  $\pi_*$  the null measure in  $\mathcal{M}(\Omega \times S)$ .

Our third generalization of Prohorov's Theorem 2.1(i) is a full-fledged generalization to  $ws$ -convergence. It requires the full force of  $ws$ -tightness (to see that it generalizes, one just trivializes  $(\Omega, \mathcal{A})$  again). This third generalization also includes the sequential versions of Prohorov's theorem for narrow convergence of transition probabilities in Balder (1989c, 1990, 2000a). As we know from §1, these have for  $\Pi^{\Omega}$  a singleton  $\{\mu\}$ . A nonsequential companion result is Theorem 5.1, given below. It extends the remaining part (ii) of Prohorov's Theorem 2.1.

THEOREM 2.4. *If  $\Pi \subset \mathcal{M}(\Omega \times S)$  is  $ws$ -tight, then  $\Pi$  is relatively sequentially  $ws$ -compact.*

PROOF. Let  $(\pi_n)$  be any sequence in  $\Pi$ , and let  $\mu \in \mathcal{M}(\Omega)$  be as in Proposition 2.2. By Theorem 2.2, there exists a subsequence  $(\pi_{n'})$  of  $(\pi_n)$  and  $\delta_* \in \mathcal{T}(\Omega; S)$ , such that  $\delta_*(\cdot)(S)$  is  $\mu$ -integrable and  $\delta_{\pi_{n'}} \rightarrow^{\mu, K} \delta_*$ . Proposition 2.2,  $(\phi_{\pi_n})$  is uniformly  $\mu$ -integrable. One now proceeds as in the proof of Corollary 2.1 to prove  $\pi_{n'} \xrightarrow{ws} \pi_* := \mu \otimes \delta_*$  by means of Proposition 2.4.  $\square$

Theorem 2.4 can be augmented to deal with situations where  $S$  is metrizable or Polish:

THEOREM 2.5. For  $\Pi \subset \mathcal{M}(\Omega \times S)$  consider the following statements:

- (a)  $\Pi^\Omega \subset \mathcal{M}(\Omega)$  is relatively  $s$ -compact and  $\Pi^S := \{\pi^S : \pi \in \Pi\} \subset \mathcal{M}(S)$  is tight.
- (b)  $\Pi$  is  $ws$ -tight.
- (c) Every sequence in  $\Pi$  is  $ws$ -tight.
- (d)  $\Pi$  is relatively sequentially  $ws$ -compact.
- (e)  $\Pi^\Omega \subset \mathcal{M}(\Omega)$  is relatively  $s$ -compact and  $\Pi^S \subset \mathcal{M}(S)$  is relatively sequentially  $w$ -compact.

The following hold:

- (i) In general, (a)  $\Rightarrow$  (b)  $\Rightarrow$  (c)  $\Rightarrow$  (d)  $\Rightarrow$  (e).
- (ii) If  $S$  is metrizable, then (c)  $\Leftrightarrow$  (d)  $\Leftrightarrow$  (e).
- (iii) If  $S$  is Polish, then (a)  $\Leftrightarrow$  (b)  $\Leftrightarrow$  (c)  $\Leftrightarrow$  (d)  $\Leftrightarrow$  (e).

Observe that in parts (ii)–(iii) the  $w$ -topology on  $\mathcal{M}(S)$  is metrizable (apply Theorem III.60 of Dellacherie and Meyer 1975); hence, in (e)  $\Pi^S \subset \mathcal{M}(S)$  is also relatively  $w$ -compact.

We show now that the criterion for relative  $ws$ -compactness in Theorem 3.10 of Schäl (1975) follows directly from part (ii) of the above theorem, in combination with the metrization Proposition 2.3. This result of Schäl (1975) has a metrizable Lusin space  $\Omega$ , with  $\mathcal{A} = \mathcal{B}(\Omega)$ . This allows him to consider the  $w$ -topology on  $\mathcal{M}(\Omega \times S)$ , but it is considerably more than we require here. On the other hand, Schäl (1975) uses a separable metric  $S$ , whereas we use a metrizable Suslin space  $S$ , so his result does not follow in its entirety from our result. Note also that Schäl’s result contains a third equivalent property which we do not wish to consider here.

COROLLARY 2.2. Suppose that  $\Omega$  is a topological space, with  $\mathcal{A} := \mathcal{B}(\Omega)$  countably generated, and suppose that  $S$  is metrizable. For every  $\Pi \subset \mathcal{M}(\Omega \times S)$  the following are equivalent:

- (a)  $\Pi$  is relatively  $ws$ -compact.
- (b)  $\Pi$  is relatively  $w$ -compact and  $\Pi^\Omega$  is relatively  $s$ -compact.

PROOF. (a)  $\Rightarrow$  (b): Elementary; the  $w$ -topology on  $\mathcal{M}(\Omega \times S)$  is coarser than the  $ws$ -topology, and  $\pi \mapsto \pi^\Omega$  is continuous.

(b)  $\Rightarrow$  (a): Continuity of  $\pi \mapsto \pi^\Omega$  causes the marginal projection onto  $\Omega$  of the  $ws$ -closure  $\overline{\Pi}$  of  $\Pi$  to be contained in the  $s$ -closure of  $\Pi^\Omega$ . Hence, that projection is relatively  $s$ -compact. It follows by Proposition 2.3 that  $\overline{\Pi}$  is metrizable. Therefore, it is enough to prove relative sequential  $ws$ -compactness of  $\Pi$ . But (b) implies elementarily that  $\Pi^S$  is relatively  $w$ -compact; hence, it is also relatively sequentially  $w$ -compact (recall that  $\mathcal{M}(S)$  is metrizable—see the comments just after Theorem 2.5). So the desired relative sequential compactness of  $\Pi$  follows by Theorem 2.5(ii).  $\square$

LEMMA 2.1. If  $\Pi \subset \mathcal{M}(\Omega \times S)$  is such that  $\Pi^S \subset \mathcal{M}(S)$  is tight in the sense of Definition 2.1, then  $\Pi$  is tight in the sense of Definition 2.2(i).

PROOF. Since  $\Pi^S$  is tight, there exist compact sets  $K_\epsilon$ ,  $\epsilon > 0$ , as in Definition 2.1. Then part (b) of Proposition 2.1 applies, with  $\Gamma_\epsilon$  the constant multifunction equal to  $K_\epsilon$ . By Proposition 2.1, this shows that  $\Pi$  is tight in the sense of Definition 2.2(i).  $\square$

LEMMA 2.2. *Suppose that  $S$  is metrizable.*

- (i) *Every  $w$ -convergent sequence in  $\mathcal{M}(S)$  is tight in the classical sense of Definition 2.1.*
- (ii) *Every sequence  $(\pi_n)$  in  $\mathcal{M}(\Omega \times S)$ , such that  $(\pi_n^S)$  is  $w$ -convergent (and in particular every  $ws$ -convergent sequence  $(\pi_n)$ ) is tight in the sense of Definition 2.2(i).*

PROOF. (i) Let  $(\nu_n)$  be  $w$ -convergent in  $\mathcal{M}(S)$ . Since  $S$  is Suslin, every single measure  $\nu_n$ ,  $n \in \mathbb{N}$ , is tight (alias Radon) by Theorem III.69 in Dellacherie and Meyer (1975). So, given the metrizability of  $S$ , it follows by Theorem 8 on page 241 of Billingsley (1968) (see also LeCam 1957) that the entire sequence of  $(\nu_n)$  is tight.

(ii) For  $(\pi_n)$  in  $\mathcal{M}(\Omega \times S)$  and  $\nu_\infty \in \mathcal{M}(S)$ , let  $\pi_n^S \rightarrow^w \nu_\infty$ . By part (i),  $(\pi_n^S)$  is tight in the classical sense. So by Lemma 2.1,  $(\pi_n)$  is tight in the sense of Definition 2.2(i).  $\square$

PROOF OF THEOREM 2.5. (a)  $\Rightarrow$  (b): This follows directly from Lemma 2.1.

(b)  $\Rightarrow$  (c): A fortiori.

(c)  $\Rightarrow$  (d): Apply Theorem 2.4.

(d)  $\Rightarrow$  (e): Recall that the marginal projections  $\pi \mapsto \pi^\Omega$  and  $\pi \mapsto \pi^S$  are continuous.

(e)  $\Rightarrow$  (c) (if  $S$  is metrizable): Let  $(\pi_n)$  be an arbitrary sequence in  $\mathcal{M}(\Omega \times S)$ . Then  $(\pi_n^S)$  is relatively sequentially  $w$ -compact, so Lemma 2.2(ii) applies:  $(\pi_n)$  is tight. Since  $\Pi^\Omega$  is relatively  $s$ -compact,  $(\pi_n)$  is also  $ws$ -tight.

(e)  $\Rightarrow$  (a) (if  $S$  is Polish): Since  $S$  is Polish, the converse Prohorov Theorem 6.2 in Billingsley (1968) applies. Hence, the relative (sequential)  $w$ -compactness of  $\Pi^S$  implies tightness of  $\Pi^S$  in the sense of Definition 2.1.  $\square$

By following ideas of Balder (1995, 2000a), we can completely characterize the  $ws$ -convergence of sequences in  $\mathcal{M}(\Omega \times S)$ . This is done by means of Theorem 2.2, provided that the Suslin space  $S$  is metrizable for its original topology. A similar characterization can also be given for nonmetrizable  $S$ , but it would only hold for tight sequences; cf. Balder (1995, 2000a). As applications in the next section will show, this characterization forms a powerful tool to study  $ws$ -convergence and  $ws$ -closure. It extends Corollary 3.16 of Balder (1995) and Theorem 4.8 of Balder (2000a).

THEOREM 2.6. *Suppose that  $S$  is metrizable. For every  $(\pi_n)$  in  $\mathcal{M}(\Omega \times S)$ , marginally dominated by  $\mu \in \mathcal{M}(\Omega)$ , and every  $\pi_\infty$  in  $\mathcal{M}(\Omega \times S)$  the following are equivalent:*

- (a)  $\pi_n \rightarrow^{ws} \pi_\infty$  in  $\mathcal{M}(\Omega \times S)$ ,
- (b)  $(\tilde{\phi}_{\pi_n})$  is uniformly  $\mu$ -integrable,  $\pi_\infty^\Omega$  is absolutely continuous with respect to  $\mu$  and every subsequence  $(\pi_{n'})$  of  $(\pi_n)$  has a further subsequence  $(\pi_{n''})$ , such that  $\tilde{\delta}_{\pi_{n''}} \rightarrow^{\mu, K} \tilde{\delta}_{\pi_\infty}$ .

PROOF. (a)  $\Rightarrow$  (b): Uniform integrability of  $(\tilde{\phi}_{\pi_n})$  holds by Proposition 2.2. Also, continuity of the marginal projection on  $S$  gives  $\pi_n^S \rightarrow^w \pi_\infty^S$  in  $\mathcal{M}(S)$ . By Lemma 2.2(ii), it follows that  $(\pi_n)$  is tight. Since  $(\pi_n)$  is also evidently bounded, we may invoke Theorem 2.2. This gives that to every subsequence  $(\pi_{n'})$  of  $(\pi_n)$  there corresponds a further subsequence  $(\pi_{n''})$  and a  $\tilde{\delta}_* \in \mathcal{F}(\Omega; S)$ , such that  $\tilde{\delta}_{\pi_{n''}} \rightarrow^{\mu, K} \tilde{\delta}_*$ . It remains to show that  $\tilde{\delta}_* = \tilde{\delta}_{\pi_\infty}$   $\mu$ -a.e. (observe that  $\pi_\infty^\Omega$  is absolutely continuous with respect to  $\mu$  by Definition 1.2). We already saw that  $(\tilde{\phi}_{\pi_n})$  is uniformly  $\mu$ -integrable, so it follows by Proposition 2.4 that  $\pi_n \rightarrow^{ws} \mu \otimes \tilde{\delta}_*$  (see the proof of Corollary 2.1). Since the  $ws$ -topology is Hausdorff, (a) gives  $\mu \otimes \tilde{\delta}_* = \pi_\infty = \mu \otimes \tilde{\delta}_{\pi_\infty}$ , whence  $\tilde{\delta}_*(\omega) = \tilde{\delta}_{\pi_\infty}(\omega)$  for  $\mu$ -a.e.  $\omega$ .

(b)  $\Rightarrow$  (a): Similar to the proof of Corollary 2.1, Proposition 2.4 implies that every subsequence  $(\pi_{n'})$  of  $(\pi_n)$  has a further subsequence  $(\pi_{n''})$ , such that  $\pi_{n''} \rightarrow^{ws} \pi_\infty$ . By contraposition, this fact immediately implies (a).  $\square$

**3. Developments and applications.** We begin this section by giving some applications of Theorem 2.6, the characterization result for  $ws$ -convergence of sequences in  $\mathcal{M}(\Omega \times S)$ . The following characterization of  $ws$ -convergence could be made part of a broader portmanteau-type theorem, quite similar to what was done in Balder (1995, 2000a).

**THEOREM 3.1.** *Suppose that  $S$  is metrizable. For every  $(\pi_n)$  and  $\pi_\infty$  in  $\mathcal{M}(\Omega \times S)$ , the following are equivalent:*

- (a)  $\pi_n \xrightarrow{ws} \pi_\infty$  in  $\mathcal{M}(\Omega \times S)$ ,
- (b)  $\lim_n \int_{\Omega \times S} g d\pi_n = \int_{\Omega \times S} g d\pi_\infty$  for every bounded  $\mathcal{A} \otimes \mathcal{B}(S)$ -measurable function  $g : \Omega \times S \rightarrow \mathbb{R}$ , such that  $g(\omega, \cdot)$  is continuous on  $S$  for every  $\omega \in \Omega$ .
- (c)  $\liminf_n \int_{\Omega \times S} g d\pi_n \geq \int_{\Omega \times S} g d\pi_\infty$  for every normal integrand  $g$  on  $\Omega \times S$ , such that

$$\limsup_{\alpha \rightarrow \infty} \int_n \max(-g, 0) d\pi_n = 0.$$

In (c) the following *integration convention* is made: We set  $\int_{\Omega \times S} g d\pi_\infty := \int_{\Omega \times S} \max(g, 0) d\pi_\infty - \int_{\Omega \times S} \max(-g, 0) d\pi_\infty$ , it being understood that  $(+\infty) - (+\infty)$  is by definition equal to  $+\infty$ .

**PROOF.** (a)  $\Rightarrow$  (c): Fix any  $\alpha \geq 0$  and let  $\beta_\alpha := \liminf_n \int_{\Omega \times S} \max(g, -\alpha) d\pi_n$ . There is a subsequence  $(\pi_{n'})$  of  $(\pi_n)$  for which  $\beta_\alpha = \lim_{n'} \int_{\Omega \times S} \max(g, -\alpha) d\pi_{n'}$ . By Theorem 2.6,  $(\pi_{n'})$  has a further subsequence  $(\pi_{n''})$  with  $\tilde{\delta}_{\pi_{n''}} \xrightarrow{\mu, K} \tilde{\delta}_{\pi_\infty}$  and  $(\tilde{\phi}_{\pi_{n''}})$  is  $\mu$ -uniformly integrable. Then Proposition 2.4(ii) implies  $\beta_\alpha \geq \int_{\Omega \times S} \max(g, -\alpha) d\pi_\infty$ . In turn, this gives  $\beta_\alpha \geq \int_{\Omega \times S} g d\pi_\infty$ . Letting  $\alpha$  go to infinity gives the desired inequality, because

$$\int_{\Omega \times S} g d\pi_n \geq \int_{\Omega \times S} \max(g, -\alpha) d\pi_n - \int_{\{g \leq -\alpha\}} \max(-g, 0) d\pi_n,$$

for all  $n \in \mathbb{N}$ .

(c)  $\Rightarrow$  (b)  $\Rightarrow$  (a): Elementary: For the first implication, apply (c) to both  $g$  and  $-g$ , and for the second one apply (b) to  $g(\omega, s) := 1_A(\omega)c(s)$ .  $\square$

Our next application of Theorem 2.6 is an upper semicontinuity result for the pointwise support sets of a ws-convergent sequence. Similar results were obtained for narrow convergence of transition probabilities in Balder (1995, 2000a, 2000b). Recall that the *support* of a measure  $\nu$  in  $\mathcal{M}(S)$  is defined as follows:

$$\text{supp } \nu := \bigcap_{F \subset S} \{F : F \text{ closed, } \nu(S \setminus F) = 0\}.$$

Recall also from Dal Maso (1993) that the Kuratowski upper limit set (alias *limes superior*)  $Ls_n B_n$  of a sequence  $(B_n)$  of subsets of  $S$  is defined as the set of all  $s \in S$ , such that  $(s_{n_j})$  converges to  $s$  for some subsequence  $(s_{n_j})$ ,  $s_{n_j} \in B_{n_j}$ . If  $S$  is metrizable, it is easy to see that the following identity holds:

$$(3.1) \quad Ls_n B_n = \bigcap_{p=1}^{\infty} \text{cl} \left( \bigcup_{n=p}^{\infty} B_n \right).$$

**THEOREM 3.2.** *Suppose that  $S$  is metrizable. If  $\pi_n \xrightarrow{ws} \pi_\infty$  for  $(\pi_n)$  and  $\pi_\infty$  in  $\mathcal{M}(\Omega \times S)$ , then*

$$\text{supp } \tilde{\delta}_{\pi_\infty}(\omega) \subset Ls_n \text{supp } \tilde{\delta}_{\pi_n}(\omega) \quad \text{for } \mu\text{-a.e. } \omega \text{ in } \Omega,$$

for every marginally dominating measure  $\mu \in \mathcal{M}(\Omega)$ . Moreover,

$$\text{supp } \delta_{\pi_\infty}(\omega) \subset Ls_n \text{supp } \delta_{\pi_n}(\omega) \quad \text{for } \pi_\infty^\Omega\text{-a.e. } \omega \text{ in } \Omega,$$

whence

$$\pi_\infty(\{(\omega, s) \in \Omega \times S : s \notin Ls_n \text{supp } \delta_{\pi_n}(\omega)\}) = 0.$$

LEMMA 3.1. *Suppose that  $S$  is metrizable. If  $\tilde{\delta}_n \rightarrow^{\mu, K} \tilde{\delta}_\infty$  for  $(\tilde{\delta}_n)$  and  $\tilde{\delta}_\infty$  in  $\mathcal{T}(\Omega; S)$  and for  $\mu \in \mathcal{M}(\Omega)$ , then*

$$\text{supp } \tilde{\delta}_\infty(\omega) \subset Ls_n \text{supp } \tilde{\delta}_n(\omega) \text{ for } \mu\text{-a.e. } \omega \text{ in } \Omega.$$

PROOF. Because of (2.4) there is a  $\mu$ -null set  $N$  with  $\nu_{m, \omega} := (1/m) \sum_{n=1}^m \tilde{\delta}_n(\omega) \rightarrow^w \tilde{\delta}_\infty(\omega)$  in  $\mathcal{M}(S)$  for every  $\omega \notin N$ . Fix an arbitrary  $\omega \notin N$ . For every  $p \in \mathbb{N}$ , the portmanteau Theorem 2.1 of Billingsley (1968) gives  $\tilde{\delta}_\infty(\omega)(G_p) \leq \liminf_m \nu_{m, \omega}(G_p) \leq \liminf_m (1/m) \sum_{n=1}^{p-1} \tilde{\delta}_n(\omega)(S) = 0$ , where  $G_p$  denotes the open set  $\Omega \setminus \text{cl } \bigcup_{n \geq p} \text{supp } \tilde{\delta}_n(\omega)$ . It follows that  $\tilde{\delta}_\infty(\omega)(\bigcup_p G_p) = 0$ . By (3.1), this proves the result.  $\square$

For nonmetrizable  $S$ , this lemma continues to hold, but in a more complicated form. This can be gleaned from analogous results for narrow convergence in  $\mathcal{R}(\Omega; S)$  given in Balder (1995, 2000a).

PROOF OF THEOREM 3.2. By Theorem 2.6 there is a subsequence  $(\pi_{n'})$  of  $(\pi_n)$ , such that  $\tilde{\delta}_{\pi_{n'}} \rightarrow^{\mu, K} \tilde{\delta}_{\pi_\infty}$ . So the first result follows by Lemma 3.1, for the inclusion  $Ls_{n'} \text{supp } \tilde{\delta}_{\pi_{n'}}(\omega) \subset Ls_n \text{supp } \tilde{\delta}_{\pi_n}(\omega)$  is evident.

The second result is an obvious consequence of the first one: For every  $n \in \mathbb{N}$  and  $\omega$ , one has trivially  $\text{supp } \tilde{\delta}_{\pi_n}(\omega) \subset \text{supp } \delta_{\pi_n}(\omega)$  by (2.2), with equality of these two sets whenever  $\tilde{\phi}_{\pi_\infty}(\omega) > 0$ . Observe here that (2.2) continues to hold for  $\pi_\infty$  because of Definition 1.2. The third result also follows from (2.2).  $\square$

REMARK 3.1. As follows from Lemma 3.1,  $\tilde{\delta}_*$  in Theorem 2.2 has the following property:

$$\text{supp } \tilde{\delta}_*(\omega) \subset Ls_n \text{supp } \tilde{\delta}_{\pi_n}(\omega) \text{ for } \mu\text{-a.e. } \omega \text{ in } \Omega.$$

As a first application where Theorem 3.2 comes in handy, we generalize the main compactness result of Yushkevich (1997):

PROPOSITION 3.1. *Suppose that  $S$  is metrizable. Let  $\Gamma : \Omega \rightarrow 2^S$  be a multifunction such that*

$$\Gamma(\omega) \text{ is compact for every } \omega \in \Omega,$$

$$\text{gph } \Gamma := \{(\omega, s) \in \Omega \times S : s \in \Gamma(\omega)\} \text{ is } \mathcal{A} \otimes \mathcal{B}(S)\text{-measurable.}$$

Also, let  $M \subset \mathcal{M}(\Omega)$  be an  $s$ -compact set. Then  $\Pi_\Gamma := \{\pi \in \mathcal{M}(\Omega \times S) : \pi^\Omega \in M, \pi((\Omega \times S) \setminus \text{gph } \Gamma) = 0\}$  is sequentially  $ws$ -compact.

This extends Theorem 1 of Yushkevich (1997), where  $M$  is a singleton and  $(\Omega, \mathcal{A})$  is a measurable Lusin space (note that compactness is understood to be sequential compactness in that reference—see page 459 of Yushkevich 1997). Because Yushkevich (1997) works with a singleton  $M$ , his version of the above proposition could also be proven by means of standard Young measure theory (this fact was also observed in Yushkevich 1997).

PROOF. Clearly,  $\Pi_\Gamma$  is  $ws$ -tight by Proposition 2.1. So by Theorem 2.4,  $\Pi_\Gamma$  is relatively sequentially  $ws$ -compact. Therefore, any sequence  $(\pi_n)$  in  $\Pi_\Gamma$  has a subsequence  $(\pi_{n'})$  that  $ws$ -converges to some  $\pi_* \in \mathcal{M}(\Omega \times S)$ . Observe already that this implies  $\pi_*^\Omega \in M$  by  $(ws, s)$ -continuity of  $\pi \mapsto \pi^\Omega$ . By (2.2) and by definition of  $\Pi_\Gamma$ , we have  $\text{supp } \tilde{\delta}_{\pi_{n'}}(\omega) \subset \Gamma(\omega)$  for all  $n'$  for  $\mu$ -a.e.  $\omega$ . Since  $\Gamma(\omega)$  is certainly closed for every  $\omega$ , it follows by Theorem 3.2 that  $\text{supp } \tilde{\delta}_\infty(\omega)$  is also contained in  $\Gamma(\omega)$  for  $\mu$ -a.e.  $\omega$ . Hence,  $\pi_* \in \Pi_\Gamma$ .  $\square$

Next, following Balder (1984a, b, 1995, 2000a, b), we enrich  $S$  by considering  $S \times \widehat{\mathbb{N}}$ . Here  $\widehat{\mathbb{N}} := \mathbb{N} \cup \{\infty\}$  is the Alexandrov compactification of  $\mathbb{N}$  (which is metrizable and compact), and  $S \times \widehat{\mathbb{N}}$  is equipped with the product topology. For  $n \in \widehat{\mathbb{N}}$  let  $\epsilon_n \in \mathcal{P}(\widehat{\mathbb{N}})$  be the Dirac probability measure concentrated at the point  $n$ . It turns out that such enrichment can

be obtained entirely *for free*:

LEMMA 3.2. *Suppose that  $S$  is metrizable. For every  $(\nu_n)$  and  $\nu_\infty$  in  $\mathcal{M}(S)$  the following are equivalent:*

- (a)  $\nu_n \xrightarrow{w} \nu_\infty$  in  $\mathcal{M}(S)$ ,
- (b)  $\nu_n \times \epsilon_n \xrightarrow{w} \nu_\infty \times \epsilon_\infty$  in  $\mathcal{M}(S \times \widehat{\mathbb{N}})$ .

The nontrivial implication (a)  $\Rightarrow$  (b) follows directly from Corollary 2.6 in Balder (2000a). The following result, which generalizes Corollary 4.9 in Balder (2000a), is now immediate by Theorem 2.6.

THEOREM 3.3. *Suppose that  $S$  is metrizable. For every  $(\pi_n)$  and  $\pi_\infty$  in  $\mathcal{M}(\Omega \times S)$ , the following are equivalent:*

- (a)  $(\pi_n)$  converges in the ws-topology to  $\pi_\infty \in \mathcal{M}(\Omega \times S)$ ,
- (b)  $(\pi_n \times \epsilon_n)$  converges in the ws-topology to  $\pi_\infty \times \epsilon_\infty \in \mathcal{M}(\Omega \times (S \times \widehat{\mathbb{N}}))$ .
- (c)  $\liminf_n \int_{\Omega \times S} g(\omega, s, n) \pi_n(d(\omega, s)) \geq \int_{\Omega \times S} g(\omega, s, \infty) \pi_\infty(d(\omega, s))$  for every normal integrand  $g$  on  $\Omega \times (S \times \widehat{\mathbb{N}})$  which is bounded from below.

The refined portmanteau-type theorems for ws-convergence, obtained by Galdéano (1997) and Galdéano and Truffert (1998), follow easily from Theorem 3.3 and the preceding results. This is quite similar to applications of Young measure theory to lower closure type results in Balder (1995, 2000a, b). For instance, Theorem 2.1 of Galdéano and Truffert now follows by invoking Theorem 3.1 and the “free enrichment principle” explained above. As another example, we shall now essentially derive Theorem 1.2 of Galdéano and Truffert (who use a Polish space  $S$ ):

PROPOSITION 3.2. *Suppose that  $S$  is metrizable. For every  $(\pi_n)$  and  $\pi_\infty$  in  $\mathcal{M}(\Omega \times S)$ , the following are equivalent:*

- (a)  $\pi_n \xrightarrow{ws} \pi_\infty$  in  $\mathcal{M}(\Omega \times S)$ ,
- (b)  $\pi_n(\Omega \times S) \rightarrow \pi_\infty(\Omega \times S)$  and  $\limsup_n \pi_n(\text{gph } \Gamma_n) \leq \pi_\infty(\text{gph } \Gamma_\infty)$  for every collection  $\{\Gamma_n : n \in \mathbb{N} \cup \{\infty\}\}$  of multifunctions  $\Gamma_n : \Omega \rightarrow 2^S$ , such that

$$\begin{aligned} \text{gph } \Gamma_n & \text{ is } \mathcal{A} \otimes \mathcal{B}(S)\text{-measurable for every } n \in \mathbb{N} \cup \{\infty\}, \\ \Gamma_n(\omega) & \text{ is closed for every } \omega \in \Omega \text{ and } n \in \mathbb{N} \cup \{\infty\}, \\ Ls_n \Gamma_n(\omega) & \subset \Gamma_\infty(\omega) \text{ for every } \omega \in \Omega. \end{aligned}$$

PROOF. (a)  $\Rightarrow$  (b): The first statement in (b) is obvious. To prove the second one, we define  $g : \Omega \times S \times \widehat{\mathbb{N}} \rightarrow \{-1, 0\}$  by  $g(\omega, s, n) := -1_{\text{gph } \Gamma_n}(\omega, s)$ . Then it follows easily from the given properties of  $(\Gamma_n)$  that  $g(\omega, \cdot, \cdot)$  is lower semicontinuous on  $S \times \widehat{\mathbb{N}}$  for every  $\omega \in \Omega$ . In view of (a), we can apply Theorem 3.3(c) to  $g$ , which easily yields the upper semicontinuity statement in (b).

(b)  $\Rightarrow$  (a): By Definition 1.1 it is clear that (a) holds if and only if  $\pi_n^A \xrightarrow{w} \pi_\infty^A$  for an arbitrary  $A \in \mathcal{A}$ , where  $\pi_n^A := \pi_n(A \times \cdot)$ . Hence, by the portmanteau Theorem 2.1 of Billingsley (1968), here considered in  $\mathcal{M}(S)$  instead of  $\mathcal{P}(S)$ , it is enough to prove that  $\limsup_n \pi_n(A \times F) \leq \pi_\infty(A \times F)$  for every closed  $F \subset S$ . Define  $F_n$  as the set of all  $s \in S$  with  $\text{dist}(s, F) \leq 1/n$ . Then the  $F_n$  are closed and  $Ls_n F_n = F_\infty := F$ . So we may apply (b) to  $\Gamma_n(\omega) := F_n$  for  $\omega \in A$  and  $\Gamma_n(\omega) := \emptyset$  otherwise. This gives precisely  $\limsup_n \pi_n(A \times F) \leq \pi_\infty(A \times F)$ .  $\square$

Next, we present an application, which touches upon much of the material gathered thus far, to the subject of equilibrium distributions in continuum game theory.

EXAMPLE 3.1. Let  $\Omega$  be an abstract set of players, equipped with a countably generated  $\sigma$ -algebra  $\mathcal{A}$ . Such models, which have a continuum of players, follow Aumann (1964) and Schmeidler (1973). Let  $S$  be a compact metric space of actions, and let  $\mathcal{P}(\Omega \times S)$  denote

the set of all probability measures in  $\mathcal{M}(\Omega \times S)$ . Let  $U : \Omega \times S \times \mathcal{P}(\Omega \times S) \rightarrow [-\infty, +\infty]$  be a given function, whose interpretation is as follows: When faced with the probability distribution  $\pi$  over player-action pairs, player  $\omega$  derives utility  $U(\omega, s, \pi)$  from taking the action  $s$ . Following Mas-Colell (1984), but considerably generalizing his notions at the same time, we say that a probability measure  $\pi_* \in \mathcal{P}(\Omega \times S)$  is a *Cournot-Nash equilibrium distribution* over player-action pairs if

$$\pi_* (\{(\omega, s) \in \Omega \times S : s \in \operatorname{argmax}_{s' \in S} U(\omega, s', \pi_*)\}) = 1.$$

Thus, such a probability measure  $\pi_*$  assigns mass 1 to those player-action pairs  $(\omega, s) \in \Omega \times S$  whose player  $\omega$  is entirely satisfied with the action  $s$  in that  $U(\omega, s', \pi_*)$  attains its maximum for  $s' = s$ . Let  $\Pi_{\text{CNE}} \subset \mathcal{P}(\Omega \times S)$  denote the set of all such Cournot-Nash equilibrium distributions. Let  $Q \subset \mathcal{P}(\Omega)$  be a given  $s$ -compact set of probability measures, representing the set of feasible distributions for the population of players. The set of all feasible distributions over player-action pairs is then defined by

$$\Pi_Q := \{\pi \in \mathcal{P}(\Omega \times S) : \pi^\Omega \in Q\}.$$

Also, let  $\ell : \Omega \times S \rightarrow [0, +\infty]$  be a given normal integrand and consider the minimization problem

$$(\text{MP}) : \inf_{\pi \in \Pi_{\text{CNE}} \cap \Pi_Q} \int_{\Omega \times S} \ell d\pi.$$

We present conditions under which (MP) possesses an optimal solution. For instance, if  $\int_{\Omega \times S} \ell d\pi$  measures the degree of dispersion of the distribution  $\pi$  (e.g., its variance), then such an optimal solution would be a feasible Cournot-Nash equilibrium distribution with minimum dispersion. Our conditions are as follows: (i)  $U(\omega, \cdot, \cdot)$  is upper semicontinuous on  $S \times \mathcal{P}(\Omega \times S)$  for every  $\omega \in \Omega$ , (ii)  $V(\omega, \cdot)$  is lower semicontinuous on  $\mathcal{P}(\Omega \times S)$  for every  $\omega \in \Omega$ , where  $V(\omega, \pi) := \sup_{s \in S} U(\omega, s, \pi)$ . We shall argue that (a)  $\Pi_Q$  is ws-compact, (b)  $\Pi_{\text{CNE}} \cap \Pi_Q \neq \emptyset$  and (c)  $\Pi_{\text{CNE}}$  is ws-closed.

Observe that  $\Pi_Q$  is ws-tight (by compactness of  $S$ , one can set  $h \equiv 0$  in Definition 2.2). So  $\Pi_Q$  is relatively ws-compact by Theorem 2.4 (notice that  $\Pi_Q$  is ws-metrizable by Proposition 2.3). Also, it is easy to see that  $\Pi_Q$  is ws-closed, so  $\Pi_Q$  is ws-compact. This proves (a). Next, we prove (b) by a fixed point argument on  $\Pi_Q$ , which is subdivided in several steps. In passing, we shall also prove (c). We start by observing that for every  $\pi \in \Pi_Q$  one has  $\pi \in \Pi_{\text{CNE}}$  if and only if  $\pi \in \Phi(\pi)$ , where

$$(3.2) \quad \Phi(\pi) := \{\pi' \in \Pi_Q : \pi'(\{(\omega, s) \in \Omega \times S : s \in \operatorname{argmax}_{s' \in S} U(\omega, s', \pi)\}) = 1\}.$$

In terms of  $V$ , this can be rewritten as

$$\Phi(\pi) := \{\pi' \in \Pi_Q : \pi'(\{(\omega, s) \in \Omega \times S : U(\omega, s, \pi) = V(\omega, \pi)\}) = 1\}.$$

*Step 1.* Trivially, for every  $\pi \in \Pi_Q$  the set  $\Phi(\pi)$  is convex.

*Step 2.* Let  $\pi \in \Pi_Q$  be arbitrary. We prove that  $\Phi(\pi)$  is nonempty by means of measurable selection arguments. By Lemma III.14 of Castaing and Valadier (1977) the function  $V(\cdot, \pi)$  is universally measurable on  $\Omega$ , so it is *a fortiori* measurable with respect to the  $\mu$ -completion  $\mathcal{A}_\mu$  of  $\mathcal{A}$ . Here,  $\mu \in \mathcal{M}(\Omega)$  marginally dominates the  $s$ -compact set  $Q$  (apply Proposition 2.2). Therefore, the set  $F := \{(\omega, s) \in \Omega \times S : U(\omega, s, \pi) = V(\omega, \pi)\}$  is  $\mathcal{A}_\mu \otimes \mathcal{B}(S)$ -measurable. So by the von Neumann-Aumann measurable selection theorem (Theorem III.22 in Castaing and Valadier 1977) there exists a  $\mathcal{A}_\mu$ -measurable function  $f : \Omega \rightarrow S$ , such that  $(\omega, f(\omega)) \in F$  for all  $\omega \in \Omega$ . Hence,  $\pi' \in \mathcal{M}(\Omega \times S)$ , defined by

$$\pi'(E) := \int_{\Omega} 1_{E_\omega}(f(\omega)) \pi^\Omega(d\omega),$$

satisfies  $\pi'(F) = 1$  and  $(\pi')^\Omega = \pi^\Omega \in Q$  (observe that the above integral makes sense since  $\Pi^\Omega$  is dominated by  $\mu$ ). This is to say that  $\pi'$  belongs to  $\Phi(\pi)$ , which is therefore nonempty.

*Step 3.* We prove that the multifunction  $\Phi : \Pi_Q \rightarrow 2^{\Pi_Q}$  is upper semicontinuous. By ws-compactness of  $\Pi_Q$ , already established, it is enough to prove that the graph of  $\Phi$  is closed: if  $\pi'_n \in \Phi(\pi_n)$  for all  $n \in \mathbb{N}$  and  $\pi_n \xrightarrow{ws} \pi_\infty$  and  $\pi'_n \xrightarrow{ws} \pi'_\infty$  in  $\mathcal{M}(\Omega \times S)$  then, evidently, for all  $n \in \mathbb{N}$ ,

$$\int_{\Omega \times S} [\arctan V(\omega, \pi_n) - \arctan U(\omega, s, \pi_n)] \pi'_n(d(\omega, s)) = 0$$

(the arctangent transformation serves to make the integrand bounded). Define  $g : \Omega \times (S \times \widehat{\mathbb{N}}) \rightarrow [0, +\infty]$  as follows:

$$g(\omega, s, n) := \arctan V(\omega, \pi_n) - \arctan U(\omega, s, \pi_n).$$

Then the semicontinuity conditions for  $U$  and  $V$  imply that  $g(\omega, \cdot, \cdot)$  is lower semicontinuous on  $S \times \widehat{\mathbb{N}}$  for every  $\omega \in \Omega$ . Moreover, by the proof of Step 2 it follows that  $g$  is  $\mathcal{A}_\mu \otimes \mathcal{B}(S \times \widehat{\mathbb{N}})$ -measurable. So  $g$  is a normal integrand on  $\Omega \times (S \times \widehat{\mathbb{N}})$  and we can apply Theorem 3.3(c). This gives

$$\int_{\Omega \times S} [\arctan V(\omega, \pi_\infty) - \arctan U(\omega, s, \pi_\infty)] \pi'_\infty(d(\omega, s)) \leq 0.$$

Taking into consideration that  $g$  is nonnegative and the strict monotonicity of the arctangent transformation, this implies

$$\pi'_\infty(\{(\omega, s) \in \Omega \times S : U(\omega, s, \pi_\infty) = V(\omega, \pi_\infty)\}) = 1,$$

i.e.,  $\pi'_\infty \in \Phi(\pi_\infty)$ . This proves the desired graph-closedness of  $\Phi$ .

*Step 4.* Since  $\Phi : \Pi_Q \rightarrow 2^{\Pi_Q}$  has been shown to be upper semicontinuous with closed convex values in the compact convex set  $\Pi_Q$ , it follows by an application of the Kakutani fixed point theorem that there is  $\pi_* \in \Pi_Q$ , such that  $\pi_* \in \Phi(\pi_*)$ . Hence, the set  $\Pi_{\text{CNE}} \cap \Pi_Q$  of feasible equilibrium distributions is nonempty. Observe that the graph-closedness of  $\Phi$ , proven in Step 3, implies that  $\Pi_{\text{CNE}}$  is closed. This follows immediately from (3.2). Our proof of (a), (b), and (c) above is therefore complete. This guarantees that  $\Pi_{\text{CNE}} \cap \Pi_Q$  is nonempty and compact. Finally, the objective function  $\pi \mapsto \int_{\Omega \times S} \ell d\pi$  is ws-lower semicontinuous on  $\mathcal{M}(\Omega \times S)$  by Theorem 3.1(c). Thus, the desired existence of an optimal solution of (MP) follows immediately by invoking the Weierstrass existence theorem.

The approach followed in the above example can be extended even further (for instance to  $\omega$ -dependent action spaces). In the standard literature on the subject the set  $Q$  is merely a singleton and  $\ell \equiv 0$  (i.e., the result in the example comes down to an equilibrium distribution existence result).

**4. A new multidimensional Fatou lemma.** A well-known area of application of the Young measure apparatus is formed by lower closure results “without convexity;” see Balder (1984a, b, c, 1985, 1995, 2000a, b). We illustrate the usefulness of the results developed thus far by giving a new type of Fatou’s lemma in several dimensions:

**THEOREM 4.1.** *Given  $\mu \in \mathcal{M}(\Omega)$  and  $d \in \mathbb{N}$ , let  $(\tilde{\phi}_n)$  and  $\tilde{\phi}_\infty$  be nonnegative functions in  $\mathcal{L}^1_{\mathbb{R}}(\Omega, \mu)$ , such that  $(\tilde{\phi}_n)$  converges to  $\tilde{\phi}_\infty \in \mathcal{L}^1_{\mathbb{R}}(\Omega, \mu)$  in the weak topology  $\sigma(\mathcal{L}^1_{\mathbb{R}}(\Omega, \mu), \mathcal{L}^\infty_{\mathbb{R}}(\Omega))$ . Let  $(f_n)$  be a sequence of  $\mathcal{A}$ -measurable functions from  $\Omega$  into  $\mathbb{R}^d$  such that  $\phi_n f_n^i$  is  $\mu$ -integrable for every  $n \in \mathbb{N}$ , and such that*

$$a^i := \lim_n \int_\Omega \tilde{\phi}_n(\omega) f_n^i(\omega) \mu(d\omega) \quad \text{exists for } i = 1, \dots, d,$$

$$(\max(-f_n^i, 0) \tilde{\phi}_n) \quad \text{is uniformly } \mu\text{-integrable for } i = 1, \dots, d.$$

Then there exists an  $\mathcal{A}$ -measurable function  $f_*$  from  $\Omega$  into  $\mathbb{R}^d$ , such that  $\int_{\Omega} \tilde{\phi}_{\infty} f_*^i d\mu \leq a^i$  for  $i = 1, \dots, d$  and

$$(4.1) \quad f_*(\omega) \in Ls_n\{f_n(\omega)\} \quad \text{for } \mu\text{-a.e. } \omega \text{ in } \Omega \text{ with } \tilde{\phi}_{\infty}(\omega) > 0.$$

PROOF. Take  $S := \mathbb{R}^d$  and define  $\pi_n \in \mathcal{M}(\Omega \times S)$  by

$$\pi_n(E) := \int_{\Omega} \tilde{\phi}_n(\omega) 1_{E_{\omega}}(f_n(\omega)) \mu(d\omega), \quad E \in \mathcal{A} \otimes \mathcal{B}(S).$$

Of course,  $\sup_n \int_{\Omega} \max(-f_n^i, 0) \tilde{\phi}_n d\mu < +\infty$  holds for every  $i$ , by the uniform integrability hypothesis. Together with the existence of the limit  $a^i$ , this means that

$$(4.2) \quad \sup_n \int_{\Omega} \tilde{\phi}_n |f_n^i| d\mu = \sup_n \int_{\Omega \times S} |s^i| \pi_n(d(\omega, s)) < +\infty, \quad i = 1, \dots, d.$$

Hence, for  $h(\omega, s) := \sum_{i=1}^d |s^i|$  we obtain  $\sup_n \int_{\Omega \times S} h d\pi_n < +\infty$ . Also, it is obvious that the set  $\{s \in \mathbb{R}_+^d : \sum_{i=1}^d |s^i| \leq \beta\}$  is compact for every  $\beta \in \mathbb{R}_+$ . Therefore, part (i) of Definition 2.2 is fulfilled. Part (ii) of that definition is also fulfilled, because

$$\pi_n^{\Omega}(A) = \int_A \tilde{\phi}_n d\mu \rightarrow \int_A \tilde{\phi}_{\infty} d\mu =: \lambda(A) \quad \text{for every } A \in \mathcal{A}.$$

Hence,  $(\pi_n)$  is ws-tight. By Theorem 2.4, there exist a subsequence  $(\pi_{n'})$  of  $(\pi_n)$  and a measure  $\pi_*$  in  $\mathcal{M}(\Omega \times S)$ , such that  $\pi_{n'} \rightarrow^{ws} \pi_*$ . Then it follows by Theorem 3.1 that

$$(4.3) \quad a^i = \lim_{n'} \int_{\Omega \times S} s^i \pi_{n'}(d(\omega, s)) \geq \int_{\Omega \times S} s^i \pi_*(d(\omega, s)) \quad \text{for } i = 1, \dots, d.$$

The preceding gives  $\lambda = \pi_*^{\Omega}$ . Hence, by (2.1),  $\pi_*$  disintegrates as  $\pi_* = \lambda \otimes \delta_*$  for some transition probability  $\delta_*$  with respect to  $(\Omega, \mathcal{A})$  and  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ . So the above can be rewritten as  $a^i \geq \int_{\Omega} [\int_S s^i \delta_*(\omega)(ds)] \lambda(d\omega)$  for  $i = 1, \dots, d$ . Besides, Theorem 3.2 gives  $\text{supp } \delta_*(\omega) \subset Ls_n\{f_n(\omega)\}$  for  $\lambda$ -a.e.  $\omega$ .

The space  $\Omega$  can now be partitioned into a nonatomic part  $\Omega^{\text{na}}$  and a purely atomic part  $\Omega^{\text{pa}}$ . First, we deal with  $\Omega^{\text{pa}}$  which is, by its definition, the union of at most countably many  $\mu$ -atoms  $A_j$ , with  $\mu(A_j) > 0$ . On each  $A_j$  the functions  $\tilde{\phi}_n$  and  $f_n$  are a.e. constant, say with values  $\beta_{n,j} \in \mathbb{R}$  and  $s_{n,j} \in \mathbb{R}^d$ . We now split  $\Omega^{\text{pa}}$  further into  $\tilde{A}$ , the union of all  $A_j$  for which  $\lambda(A_j) > 0$  and its complement  $\Omega^{\text{pa}} \setminus \tilde{A}$ . Then it is evident that  $\Omega^{\text{pa}} \setminus \tilde{A}$  has  $\lambda$ -measure zero. On all  $A_j$  weak convergence of  $(\tilde{\phi}_n)$  to  $\tilde{\phi}_{\infty}$  comes down to  $\lim_n \beta_{n,j} = \beta_{\infty,j}$ . Also, (4.2) implies that  $\sup_n \sum_j \beta_{n,j} |s_{n,j}^i| < +\infty$  for  $i = 1, \dots, d$ . Hence, it follows that  $\sup_n |s_{n,j}^i| < +\infty$  for every  $j$  with  $\beta_{\infty,j} > 0$  (that is, with  $\lambda(A_j) > 0$ ). Hence, by a preliminary diagonal subsequence selection argument, we can suppose without loss of generality that on  $\tilde{A}$  the sequence  $(f_n)$  converges pointwise  $\lambda$ -a.e. some limit function  $f_*$ . Since  $\text{supp } \delta_*(\omega) \subset Ls_n\{f_n(\omega)\} = \{f_*(\omega)\}$  for  $\lambda$ -a.e.  $\omega$  in  $\tilde{A}$ , we conclude that  $\delta_*(\omega)$  is the point measure  $\epsilon_{f_*(\omega)}$  for such  $\omega$ . Clearly, this meets (4.1) on  $\tilde{A}$ .

Next, on  $\Omega^{\text{na}}$  the measure  $\mu$  is nonatomic, whence also  $\lambda$ , which is  $\mu$ -absolutely continuous. Thus, an application of Lyapunov's theorem for Young measures (Theorem 5.3 in Balder 2000a) gives the existence of a measurable function  $f_*$  from  $\Omega^{\text{na}}$  into  $S$ , such that  $f_*(\omega) \in Ls_n\{f_n(\omega)\}$  for  $\lambda$ -a.e.  $\omega$  in  $\Omega^{\text{na}}$  and  $\int_{\Omega^{\text{na}}} f_*^i d\lambda = \int_{\Omega^{\text{na}}} [\int_S s^i \delta_*(\omega)(ds)] \lambda(d\omega) \leq a^i$  for all  $i$ .

Finally, substituting the effect of these decompositions into (4.3) gives

$$a^i \geq \int_{\Omega^{\text{na}}} \left[ \int_S s^i \delta_*(\omega)(ds) \right] \lambda(d\omega) + \int_{\tilde{A}} f_*^i d\lambda = \int_{\Omega^{\text{na}} \cup \tilde{A}} f_*^i d\lambda \quad \text{for } i = 1, \dots, d.$$

By choosing  $f_* \equiv 0$  on the  $\lambda$ -null set  $\Omega^{pa} \setminus \tilde{A}$ , it is easy to see that  $f_*$  is now as stated in the theorem.  $\square$

Theorem 4.1 generalizes the multidimensional Fatou lemma of Balder (1984a), which subsumes both the original Fatou lemma of Schmeidler (1970) and the one of Artstein (1979). All those results work with  $\tilde{\phi}_n \equiv 1$  for all  $n$ , and the above result does not seem to follow from any of them. The following example shows that the positivity condition  $\tilde{\phi}_\infty(\omega) > 0$  in (4.1) is indispensable.

EXAMPLE 4.1. Let  $\Omega := [0, 1]$  be equipped with the Lebesgue  $\sigma$ -algebra  $\mathcal{A}$  and with the Lebesgue measure  $\mu$ . Let  $d := 1$ ,  $\tilde{\phi}_n \equiv n^{-1}$ ,  $\tilde{\phi}_\infty \equiv 0$  and  $f_n \equiv n$ . Then  $\lim_n \int_\Omega \tilde{\phi}_n f_n d\mu = 1$ , and  $Ls_n f_n(\omega) = \emptyset$  for all  $\omega$ . By  $\tilde{\phi}_\infty \equiv 0$ , this is still in agreement with (4.1).

We conclude with a new application of Theorem 4.1. Applications of the multidimensional Fatou lemma of Balder (1984a), which we just generalized in terms of Theorem 4.1, were already given in Balder (1984c) to existence in optimal control problems with singular components, and by Balder (1984a) and Balder and Pistorius (2000) to existence of optimal consumption plans in economics and finance. The usefulness of multidimensional Fatou lemmas has been known for some time in mathematical economics; see for instance Greenberg et al. (1979) for applications to competitive equilibria, which employ an earlier version of the multidimensional Fatou lemma (generalized by both Balder 1984a and Theorem 4.1).

EXAMPLE 4.2. A decision maker is uncertain about the state of nature in  $\Omega := \mathbb{R}$ , equipped with the Lebesgue  $\sigma$ -algebra  $\mathcal{A}$  and the Lebesgue measure  $\lambda$ , which she believes to be distributed according to a normal distribution with variance 1 and unknown mean  $\theta \in [-\theta_0, \theta_0]$ . Her  $\theta_0$  is a given bound. Denote the corresponding normal densities by  $p_\theta$ . A “most optimistic scenario” for the decision maker is defined to be an optimal solution of the minimization problem

$$(P) : \quad \text{minimize } J^0(\theta, u) := \int_\Omega g^0(\omega, u(\omega)) p_\theta(\omega) \lambda(d\omega),$$

over all possible decision rules  $u$  and all  $\theta$ ,  $|\theta| \leq \theta_0$ , subject to certain constraints,

$$J^i(\theta, u) := \int_A g^i(\omega, u(\omega)) p_\theta(\omega) \lambda(d\omega) \leq \alpha^i, \quad i = 1, \dots, m.$$

Here  $\alpha^1, \dots, \alpha^m$  are given constants in  $\mathbb{R}$ . Also, a *decision rule* is defined to be a measurable function  $u$  from  $\Omega := \mathbb{R}$  into  $Z := \mathbb{R}^p$ , such that  $u(\omega) \in U(\omega)$  for all  $\omega \in \Omega$ , where  $U : \Omega \rightarrow 2^Z$  is a compact nonempty-valued multifunction with  $\mathcal{A} \times \mathcal{B}(Z)$ -measurable graph  $\text{gph } U$ . Further, the functions  $g^i : \text{gph } U \rightarrow (-\infty, +\infty]$ ,  $i = 0, \dots, m$ , are  $\mathcal{A} \otimes \mathcal{B}(Z)$ -measurable, and  $g^i(\omega, \cdot)$  is lower semicontinuous on  $U(\omega)$  for every  $\omega \in \Omega$ . Moreover, we suppose that  $\gamma := \inf_{0 \leq i \leq m} \inf_{(\omega, z) \in \text{gph } U} g^i(\omega, z) > -\infty$ . Hence, the above integrals exist.

We shall now prove the existence of a “most optimistic scenario”  $(\theta_*, u_*)$  for problem (P) by means of Theorem 4.1, supposing that (P) has at least one feasible solution pair  $(u, \theta)$ . Let  $a^0 := \inf(P)$ ; then there exists a minimizing sequence  $(\theta_n, u_n)$  for (P). By compactness of  $[-\theta_0, \theta_0]$  we may suppose, without loss of generality, that  $(\theta_n)$  converges to some  $\theta_* \in [-\theta_0, \theta_0]$ . Also, by compactness of  $[\gamma, \alpha^i]$ , we may suppose without loss of generality that  $(J^i(\theta_n, u_n))$  converges to some  $a^i \in [\gamma, \alpha^i]$  for  $i = 1, \dots, m$ . By continuity of  $\theta \mapsto p_\theta(\omega)$  for each  $\omega \in \mathbb{R}$ , it follows from Scheffé’s Theorem 16.11 in Billingsley (1986) that  $\int_\Omega |p_{\theta_n} - p_{\theta_*}| d\lambda \rightarrow 0$ . Hence,  $\mu_n \rightarrow^s \mu_\infty$  for  $\mu_n(A) := \int_A p_{\theta_n} d\lambda$  and  $\mu_\infty(A) := \int_A p_{\theta_*} d\lambda$ . Now define  $f_n : \Omega \rightarrow \mathbb{R}^{m+1}$  by

$$f_n^i(\omega) := g^i(\omega, u_n(\omega)) p_{\theta_n}(\omega), \quad i = 0, \dots, m;$$

then it is evident that  $f_n^i$  is  $\lambda$ -integrable for every  $n$  and that  $\lim_n \int_{\Omega} f_n^i d\lambda = a^i$ . By Theorem 4.1 there exists an  $\mathcal{A}$ -measurable function  $f_*$  from  $\Omega$  into  $\mathbb{R}^{m+1}$ , such that

$$(4.4) \quad \int_{\Omega} f_*^i p_{\theta_*} d\lambda \leq a^i \quad \text{for } i = 0, \dots, m,$$

and such that for  $\lambda$ -a.e.  $\omega$  in  $\Omega$  there exists a subsequence  $(f_{n_{\omega}})$ —possibly depending upon  $\omega$ —with

$$(4.5) \quad f_{n_{\omega}}(\omega) \rightarrow f_*(\omega).$$

For all coordinates  $i = 0, \dots, m$  we have here  $f_{n_{\omega}}^i(\omega) := g^i(\omega, u_{n_{\omega}}(\omega))$ , with  $u_{n_{\omega}}(\omega)$  in the compact subset  $U(\omega)$ . By taking a convergent subsequence in (4.5) and by subsequently using the lower semicontinuity of  $g^i(\omega, \cdot)$ , it follows that for  $\lambda$ -a.e.  $\omega$  there exists at least one point  $z_{\omega} \in U(\omega)$  for which  $f_*^i(\omega) = \lim_{n_{\omega}} f_{n_{\omega}}^i(\omega) \geq g^i(\omega, z_{\omega})$ ,  $i = 0, \dots, m$ . By the implicit measurable selection Theorem III.38 in Castaing and Valadier (1977) it thus follows that there exists a measurable selection  $u_*$  of  $U$  with the same inequalities, i.e.,  $f_*^i(\omega) \geq g^i(\omega, u_*(\omega))$  for  $i = 0, \dots, m$ . If we substitute this in (4.4), we find

$$J^i(\theta_*, u_*) \leq a^i \leq \alpha^i, \quad i = 0, \dots, m.$$

So  $(\theta_*, u_*)$  meets the constraints of  $(P)$  and  $J^0(\theta_*, u_*) \leq a^0 := \inf(P)$ . Hence,  $(\theta_*, u_*)$  is an optimal solution of  $(P)$ .

**5. A nonsequential Prohorov-type theorem.** Here we extend the nonsequential (i.e., topological) part (ii) of Prohorov’s Theorem 2.1 to the ws-topology. We show it to generalize the corresponding criterion for relative ws-compactness in Jacod and Mémín (1981). Our proof uses truncation of transition measures and reduces the situation to one where results from Young measure theory can be applied. We just mention that several other purely topological results from Young measure theory can be extended so as to yield counterparts for the ws-topology.

**THEOREM 5.1.** *Suppose that  $\Pi \subset \mathcal{M}(\Omega \times S)$  is ws-tight. Then  $\Pi$  is relatively ws-compact.*

Together with Theorem 2.4, this completely extends Prohorov’s Theorem 2.1 to the ws-topology. Theorem 5.1 can be stated differently when  $S$  is a Polish space. We present the following counterpart to Theorem 2.5:

**THEOREM 5.2.** *For  $\Pi \subset \mathcal{M}(\Omega \times S)$ , consider the following statements:*

- (a)  $\Pi^{\Omega} \subset \mathcal{M}(\Omega)$  is relatively  $s$ -compact and  $\Pi^S \subset \mathcal{M}(S)$  is tight.
- (b)  $\Pi$  is ws-tight.
- (c)  $\Pi$  is relatively ws-compact.
- (d)  $\Pi^{\Omega} \subset \mathcal{M}(\Omega)$  is relatively  $s$ -compact and  $\Pi^S \subset \mathcal{M}(S)$  is relatively  $w$ -compact.

*The following hold:*

- (i) In general (a)  $\Rightarrow$  (b)  $\Rightarrow$  (c)  $\Rightarrow$  (d).
- (ii) If  $S$  is Polish, then (a)  $\Leftrightarrow$  (b)  $\Leftrightarrow$  (c)  $\Leftrightarrow$  (d).

The proof is almost completely contained in that of Theorem 2.5 and will be omitted. Theorem 2.8 of Jacod and Mémín (1981), who use a Polish space  $S$  throughout, comes down to the equivalence (c)  $\Leftrightarrow$  (d) in the above theorem.

In the remainder of this section, we prove Theorem 5.1 by means of an extension of Prohorov’s theorem for the narrow topology for transition probabilities. This result was given in Theorem 2.3 of Balder (1988) for a metrizable Lusin space  $S$ . Subsequently, in Theorem 2.2 of Balder (1989a), it was extended to the situation where  $S$  is completely regular and Suslin, as used in this paper. For the reader’s convenience, we include its proof as given in Balder (1989a). Recall that the narrow topology was defined in Definition 1.3.

**THEOREM 5.3 (THEOREM 2.2 OF BALDER 1989a).** *If for  $\mu \in \mathcal{M}(\Omega)$  and  $\Delta \subset \mathcal{R}(\Omega; S)$  the set  $\{\mu \otimes \delta : \delta \in \Delta\}$  in  $\mathcal{M}(\Omega \times S)$  is tight, then  $\Delta$  is relatively narrowly compact.*

**PROOF.** The proof is divided into two parts.

*Preliminary case.* First, we suppose in addition that  $S$  is metrizable. To prove relative compactness of  $\Delta$  for the narrow topology, it is enough to demonstrate that Theorem 2.3 in Balder (1988) remains valid for a metrizable Suslin space  $S$  instead of the metrizable Lusin space used there. Observe first that everything said on pages 266–270 of that same reference continues to hold for a metrizable Suslin space  $S$ , except for the line that immediately follows the definition of the function  $\hat{h}$ . Recall this definition from page 270 of Balder (1988):  $\hat{h} := h$  on  $\Omega \times S$  and  $\hat{h} := +\infty$  on  $\Omega \times (\widehat{S} \setminus S)$ . Here  $h$  is as in Definition 2.2 and  $\widehat{S}$  stands for the Hilbert cube compactification of  $S$ . To prove that  $\hat{h}$  is  $\mathcal{A} \otimes \mathcal{B}(\widehat{S})$ -measurable, the metrizable Lusin hypothesis of Balder (1988) is of immediate use, since it implies that  $S$  belongs to  $\mathcal{B}(\widehat{S})$  by Definition III.16 of Dellacherie and Meyer (1975). However, in case  $S$  is merely metrizable Suslin we can still prove that  $\hat{h}$  is  $\mathcal{A}_\mu \otimes \mathcal{B}(\widehat{S})$ -measurable and end up with a standard  $\mathcal{A} \otimes \mathcal{B}(\widehat{S})$ -measurable modification of  $\hat{h}$ . Here  $\mathcal{A}_\mu$  stands for the  $\mu$ -completion of the  $\sigma$ -algebra  $\mathcal{A}$ . This goes as follows. Let  $\hat{d}$  be a metric on the Hilbert cube and let  $\beta \in \mathbb{R}$  be arbitrary. Observe that the set  $\hat{h}^{-1}[0, \beta]$  in  $\Omega \times S$  equals  $C := h^{-1}[0, \beta]$ . Define  $u(\omega, \hat{s}) := \inf_{s \in C_\omega} \hat{d}(\hat{s}, s)$ ; if  $C_\omega = \emptyset$  we set  $u(\omega, \cdot)$  equal to  $+\infty$ . By the measurable projection Theorem III.23 in Castaing and Valadier (1977),  $u(\cdot, \hat{s})$  is  $\mathcal{A}_\mu$ -measurable for every  $\hat{s} \in \widehat{S}$ . Also,  $u(\omega, \cdot)$  is clearly continuous on  $\widehat{S}$  for every  $\omega \in \Omega$ . By Lemma III.14 of Castaing and Valadier (1977), it follows that  $u$  is  $\mathcal{A}_\mu \otimes \mathcal{B}(\widehat{S})$ -measurable. Now, by Definition 2.2,  $C_\omega$  is compact in  $S$ , whence in  $\widehat{S}$ ; of course, this also means that  $C_\omega$  is closed in  $\widehat{S}$ . Hence,  $C$  coincides with  $u^{-1}(\{0\})$ . We conclude therefore that  $\hat{h}$  is measurable with respect to  $\mathcal{A}_\mu \otimes \mathcal{B}(\widehat{S})$ . At this point, the approximation argument involving the  $\mu$ -completion of  $\mathcal{A}$  on page 269 of Balder (1988) can be imitated (or, more directly, Lemma A.1 in Balder 1984b can be applied). This gives a  $\mathcal{A} \otimes \mathcal{B}(\widehat{S})$ -measurable modification  $\hat{h} : \Omega \times \widehat{S} \rightarrow [0, +\infty]$  of  $\hat{h}$ , for which  $\hat{h}(\omega, \cdot) = \hat{h}(\omega, \cdot)$  for  $\mu$ -a.e.  $\omega$ . After this, the proof on pages 270–271 of Balder (1988) can be resumed to conclude that  $\Delta$  is relatively compact for the narrow topology.

*General case.* By (2.3) we already demonstrated that  $S$  can be given a weak metric  $d$ , whose topology is not finer than the given topology on  $S$ . Moreover, we recorded there that the resulting metric space  $(S, d)$  is also Suslin and that its Borel  $\sigma$ -algebra coincides with the original  $\sigma$ -algebra  $\mathcal{B}(S)$  on  $S$ . Now observe that  $h$  in Definition 2.2 is *a fortiori*, such that for every  $\omega \in \Omega$ , the function  $h(\omega, \cdot)$  is inf-compact for the  $d$ -topology on  $S$ . By the preliminary case above it follows that  $\Delta$  is certainly relatively “new-narrowly” compact, where “new-narrowly” indicates that we have switched from the original topology to the  $d$ -topology on  $S$ . We now finish by demonstrating that, as a consequence of the given tightness, the new-narrow topology coincides on  $\Sigma := \{\delta \in \mathcal{R}(\Omega; S) : I_h(\delta) \leq \sigma\}$  with the original narrow topology. Here,  $I_h(\delta) := \int_{\Omega \times S} h d(\mu \otimes \delta)$  and  $\sigma := \sup_{\delta \in \Delta} I_h(\delta)$ . Evidently, on all of  $\mathcal{R}(\Omega; S)$  the new-narrow topology is certainly not finer than the original narrow topology. So it remains to prove the converse inclusion, relative to  $\Sigma$ . For this it is enough, by Theorem 2.2 in Balder (1988), to prove that  $\delta \mapsto I_g(\delta)$  is new-narrowly continuous for any  $\mathcal{A} \times \mathcal{B}(S)$ -measurable  $g : \Omega \times S \rightarrow [0, +\infty]$ , such that  $g(\omega, \cdot)$  is lower semicontinuous for every  $\omega \in \Omega$ . Let  $g_\epsilon := g + \epsilon h$ ,  $\epsilon > 0$ . Then every  $g_\epsilon$  is  $\mathcal{A} \times \mathcal{B}(S)$ -measurable, and such that, for every  $\omega \in \Omega$ , the function  $g_\epsilon(\omega, \cdot)$  is inf-compact; *a fortiori*, the latter makes  $g_\epsilon(\omega, \cdot)$  also  $d$ -inf-compact, whence  $d$ -lower semicontinuous. So, again by Theorem 2.2 of Balder (1988), the functional  $I_{g_\epsilon}$  is new-narrowly lower semicontinuous on all of  $\mathcal{R}(\Omega; S)$ . The identity  $I_g(\delta) = \sup_{\epsilon > 0} (I_{g_\epsilon}(\delta) - \epsilon \sigma)$ , which holds for every  $\delta \in \Sigma$ , then implies that  $I_g$  is new-narrowly lower semicontinuous on  $\Sigma$ .  $\square$

**PROOF OF THEOREM 5.1.** By Proposition 2.2, there exists a dominating measure  $\mu$  for  $\Pi^\Omega$  such that the corresponding set of densities  $\{\tilde{\phi}_\pi : \pi \in \Pi\}$  is uniformly integrable with

respect to  $\mu$ . For  $p \in \mathbb{N}$  and  $\tilde{\delta} \in \mathcal{T}(\Omega; S)$ , we define  $\tilde{\delta}^p \in \mathcal{T}(\Omega; S)$  by truncation:

$$\tilde{\delta}^p(\omega) := \begin{cases} \tilde{\delta}(\omega) & \text{if } \tilde{\delta}(\omega)(S) \leq p, \\ \text{null measure on } S & \text{otherwise.} \end{cases}$$

Fix  $p$ . Since  $\tilde{\delta}_\pi^p \leq \tilde{\delta}_\pi$  for every  $\pi \in \Pi$ , tightness of  $\Pi$  as in Definition 2.2 implies

$$\sup_{\pi \in \Pi} \int_{\Omega} \left[ \int_S h(\omega, s) \tilde{\delta}_\pi^p(\omega)(ds) \right] \mu(d\omega) \leq \sup_{\pi \in \Pi} \int_{\Omega \times S} h d\pi < +\infty,$$

in view of (2.2). This shows that for  $\Delta^p := \{(1/p)\tilde{\delta}_\pi^p : \pi \in \Pi\}$  the tightness condition of Theorem 5.3 is met. The fact that  $\Delta^p$  does not lie in  $\mathcal{R}(\Omega; S)$ , but in the set of all transition *subprobabilities* with respect to  $(\Omega, \mathcal{A})$  and  $(S, \mathcal{B}(S))$  does not impede application of Theorem 5.3, since it is well known that this theorem extends immediately to transition subprobabilities (as do most other results on Young measures). Theorem 5.3 now implies that  $\Delta^p$  is relatively compact for the narrow topology. Hence,  $\Pi^p := \{\mu \otimes \tilde{\delta}_\pi^p : \pi \in \Pi\}$  is relatively ws-compact. We define  $T^p : \Pi \rightarrow \Pi^p$  by  $T^p(\pi) := \mu \otimes \tilde{\delta}_\pi^p$ . Let  $\mathcal{U}$  be an arbitrary ultrafilter on  $\Pi$ . To prove relative ws-compactness of  $\Pi$ , it is enough to prove that  $\mathcal{U}$  ws-converges in  $\mathcal{M}(\Omega \times S)$ . By Proposition 4.12 of Choquet (1969) the collection  $T^p(\mathcal{U})$  is an ultrafilter on  $\Pi^p$ . By relative ws-compactness of  $\Pi^p$ , demonstrated above, it follows that  $T^p(\mathcal{U})$  ws-converges to some limit in the ws-closure of  $\Pi^p$  (apply Proposition 4.15 of Choquet 1969). Clearly, this limit must be of the form  $\mu \otimes \tilde{\eta}_p$ , with  $\tilde{\eta}_p \in \mathcal{T}(\Omega; S)$  such that  $\tilde{\eta}_p(\omega)(S) \leq p$  for  $\mu$ -a.e.  $\omega$  (use Definition 1.1). Uniformly in  $p$ , the following bound obviously holds:

$$(5.1) \quad (\mu \otimes \tilde{\eta}_p)(\Omega \times S) \leq \sup_{\pi \in \Pi} \pi(\Omega \times S) < +\infty.$$

Further, by Definition 1.1 the inequality  $\tilde{\delta}_\pi^{p+1}(\omega)(B) \geq \tilde{\delta}_\pi^p(\omega)(B)$  for all  $\omega \in \Omega$  and all  $B \in \mathcal{B}(S)$  leads to  $(\mu \otimes \tilde{\eta}_{p+1})(A \times B) \geq (\mu \otimes \tilde{\eta}_p)(A \times B)$  for every  $A \in \mathcal{A}$  and  $B \in \mathcal{B}(S)$  (first, take  $B$  to be open for the weak metric  $d$  and use Theorem A6.6 of Ash (1972); thereafter, approximate as in Corollary 4.3.7 of that same reference). This implies  $\mu \otimes \tilde{\eta}_{p+1} \geq \mu \otimes \tilde{\eta}_p$  on  $\mathcal{A} \times \mathcal{B}(S)$ . Because of this monotonicity, the limit  $\pi_* := \lim_p \mu \otimes \tilde{\eta}_p$  forms a measure on  $\mathcal{A} \otimes \mathcal{B}(S)$ , which is bounded by (5.1), so  $\pi_*$  belongs to  $\mathcal{M}(\Omega \times S)$ . We claim that the ultrafilter  $\mathcal{U}$  ws-converges to  $\pi_*$ . To this end, fix any  $A \in \mathcal{A}$  and  $c \in \mathcal{C}_b(S)$ . Then the above definition of truncation gives for every  $p \in \mathbb{N}$  and  $\pi \in \Pi$ ,

$$\alpha_\pi^p := \left| \int_{A \times S} c d\pi - \int_A \left[ \int_S c(s) \tilde{\delta}_\pi^p(\omega)(ds) \right] \mu(d\omega) \right| \leq \|c\|_\infty \int_{\{\omega \in \Omega : \tilde{\delta}_\pi(\omega) > p\}} \tilde{\phi}_\pi d\mu,$$

where we use (2.2) and the associated identity  $\tilde{\delta}_\pi(\cdot)(S) = \tilde{\phi}_\pi$ . By uniform  $\mu$ -integrability of  $\{\tilde{\phi}_\pi : \pi \in \Pi\}$ , this implies  $\lim_{p \rightarrow \infty} \sup_{\pi \in \Pi} \alpha_\pi^p = 0$ . Now for any  $p$ ,

$$\left| \int_{A \times S} c d\pi - \int_{A \times S} c d\pi_* \right| \leq \alpha_\pi^p + \beta_\pi^p + \gamma^p,$$

with  $\beta_\pi^p := \left| \int_{A \times S} c d(\mu \otimes \tilde{\delta}_\pi^p) - \int_{A \times S} c d(\mu \otimes \tilde{\eta}_p) \right|$  and  $\gamma^p := \left| \int_{A \times S} c d(\mu \otimes \tilde{\eta}_p) - \int_{A \times S} c d\pi_* \right|$ . For any fixed  $p$ , the above ws-convergence of  $T^p(\mathcal{U})$  to  $\mu \otimes \tilde{\eta}_p$  implies that  $\beta_\pi^p$  converges to 0 along  $\mathcal{U}$ . Finally,  $\lim_p \gamma^p = 0$  follows by an obvious application of the monotone convergence theorem for the positive and negative parts of the bounded function  $c$ , using (5.1). Together, this proves that  $\int_{A \times S} c d\pi$  converges to  $\int_{A \times S} c d\pi_*$  along  $\mathcal{U}$ . Since  $c$  and  $A$  were arbitrary, this proves that  $\mathcal{U}$  ws-converges to  $\pi_*$  in  $\mathcal{M}(\Omega \times S)$ .  $\square$

**Acknowledgments.** I am indebted to an editor and a referee for useful suggestions about presentation.

## References

- Aldous, D. J. 1977. Limit theorems for subsequences of arbitrarily-dependent sequences of random variables. *Z. Wahrscheinlichkeitsth. Verw. Geb.* **40** 59–82.
- Artstein, Z. 1979. A note on Fatou's lemma in several dimensions. *J. Math. Econom.* **6** 277–282.
- Ash, R. B. 1972. *Real Analysis and Probability*, Academic Press, New York.
- Balder, E. J. 1979. On a useful compactification for optimal control problems. *J. Math. Anal. Appl.* **72** 391–398.
- . 1984a. A unifying note on Fatou's lemma in several dimensions. *Math. Oper. Res.* **9** 267–275.
- . 1984b. A general approach to lower semicontinuity and lower closure in optimal control theory. *SIAM J. Control Optim.* **22** 570–598.
- . 1984c. Existence results without convexity conditions for general problems of optimal control with singular components. *J. Math. Anal. Appl.* **101** 527–539.
- . 1985. An extension of Prohorov's theorem for transition probabilities with applications to infinite-dimensional lower closure problems. *Rend. Circ. Mat. Palermo (II)* **34** 427–447.
- . 1988. Generalized equilibrium results for games with incomplete information. *Math. Oper. Res.* **13** 265–276.
- . 1989a. On Prohorov's theorem for transition probabilities, *Sém. Anal. Convexe* **19** 9.1–9.11.
- . 1989b. On compactness of the space of policies in stochastic dynamic programming. *Stochastic Processes Appl.* **32** 141–150.
- . 1989c. Unusual applications of a.e. convergence. G. A. Edgar, L. Sucheston, eds. *Almost Everywhere Convergence* Academic Press, New York, 31–53.
- . 1990. New sequential compactness results for spaces of scalarly integrable functions. *J. Math. Anal. Appl.* **151** 1–16.
- . 1991. On equivalence of strong and weak convergence in  $L_1$ -spaces under extreme point conditions. *Israel J. Math.* **75** 21–47.
- . 1992. Existence without explicit compactness in stochastic dynamic programming. *Math. Oper. Res.* **17** 572–580.
- . 1995. Lectures on Young Measures, *Cahiers du Centre de Recherche de Mathématiques de la Décision (CEREMADE)* **9517** Université Paris-Dauphine, Paris, France.
- . 2000a. New fundamentals of Young measure theory. A. Ioffe, S. Reich, I. Shafir, eds. *Calculus of Variations and Optimal Control*. Chapman and Hall/CRC Research Notes in Mathematics **411** CRC Press, Boca Raton, FL, 24–48.
- . 2000b. Lectures on Young measure theory and its applications in economics (Proceedings Grado School on Measure Theory and Real Analysis). *Rend. Sem. Matem. Trieste* **31** 1–69.
- , C. Hess. 1996. Two generalizations of Komlós' theorem with lower closure-type applications. *J. Convex Anal.* **3** 25–44.
- , M. Pistorius. 2001. On an optimal consumption problem for  $p$ -integrable consumption plans. *Econ. Theory* **17** 721–737.
- , N. C. Yannellis. 2000. Upper semicontinuity of the Cournot-Nash equilibrium correspondence: A general approach. Forthcoming.
- Berliocchi, H., J. M. Lasry. 1973. Intégrales normales et mesures paramétrées en calcul des variations. *Bull. Soc. Math. France* **101** 129–184.
- Bertsekas, D. P., S. E. Shreve. 1978. *Stochastic Optimal Control: the Discrete Time Case* Academic Press, New York.
- Billingsley, P. 1968. *Convergence of Probability Measures* Wiley, New York.
- . 1986. *Probability and Measure* Wiley, New York.
- Bourbaki, N. 1974. *Topologie Générale* Hermann, Paris, France.
- Brooks, J., R. V. Chacon. 1980. Continuity and compactness of measures. *Adv. Math.* **37** 16–26.
- Castaing, C., M. Valadier. 1977. *Convex Analysis and Measurable Multifunctions* Lecture Notes in Mathematics **580** Springer-Verlag, Berlin, Germany.
- Chatterji, S. D. 1973. Les martingales et leurs applications analytiques. J. L. Bretagnolle et al., eds. *Ecole d'Été de Probabilités: Processus Stochastiques*. Lecture Notes in Mathematics **307** Springer-Verlag, Berlin, Germany, 27–164.
- . 1985. A subsequence principle theory. *Jber. d. Dt. Math.-Verein.* **87** 91–107.
- Choquet, G. 1969. *Lectures on Analysis* Benjamin, Reading, MA.
- Dal Maso, G. 1993. *An Introduction to  $\Gamma$ -Convergence* Birkhäuser, Boston, MA.
- Dellacherie, C., P.-A. Meyer. 1975. *Probabilités et Potentiel* Hermann, Paris, France (English translation: North-Holland, Amsterdam, The Netherlands).
- Galdéano, F. 1997. Convergence étroite de mesures définies sur un espace produit Thèse de doctorat, Université de Perpignan, Académie de Montpellier, Perpignan, France.
- , A. Truffert. 1998. On the narrow convergence of measures defined on a product space. Preprint.

- Gänssler, P. 1971. Compactness and sequential compactness inspaces of measures. *Z. Wahrscheinlichkeitsth. Verw. Geb.* **17** 124–146.
- Gaposhkin, V. F. 1972. Convergence and limit theorems for sequences of random variables. *Theory Probab. Appl.* **17**(3) 379–400.
- Greenberg, J., B. Shitovitz, and A. Wieczorek. 1979. Existence of equilibria in atomless production economies with price dependent preferences. *J. Math. Econom.* **6** 31–41.
- Jacod, J., J. Mémin. 1981. Sur un type de convergence intermédiaire entre la convergence en loi et la convergence en probabilité. J. Azéma, M. Yor, eds. *Séminaire de Probabilités XV*, Lecture Notes in Mathematics **850** Springer-Verlag, Berlin, Germany, 529–546.
- , ———. 1983. Rectification à: Sur un type de convergence intermédiaire entre la convergence en loi et la convergence en probabilité. J. Azéma, M. Yor, eds. *Séminaire de Probabilités XVII*, Lecture Notes in Mathematics **986** Springer-Verlag, Berlin, Germany, 509–511.
- Jawhar, A. 1984. Mesures de transition et applications. *Sém. Analyse Convexe Montpellier* **14** 13.1–13.62.
- Komlós, J. 1967. A generalisation of a problem of Steinhaus. *Acta Math. Acad. Sci. Hungar.* **18** 217–229.
- LeCam, L. M. 1957. Convergence in distribution of stochastic processes. *Univ. California Publ. Statist.* **2**(11) 207–236.
- Mas-Colell, A. 1984. On a theorem of Schmeidler. *J. Math. Econom.* **13** 201–206.
- Neveu, J. 1965. *Mathematical Foundations of the Calculus of Probability* Holden-Day, San Francisco, CA.
- Nowak, A. 1988. On the weak topology on a space of probability measures induced by policies. *Bull. Polish Acad. Sci. Math.* **36** 181–186.
- Pedregal, P. 1997. *Parametrized Measures and Variational Principles* Birkhäuser, Basel, Switzerland.
- Prohorov, Yu. V. 1956. Convergence of random processes and limit theorems in probability theory. *Theory Probab. Appl.* **1** 157–214.
- Schäl, M. 1975. On dynamic programming: compactness of the space of policies. *Stochastic Processes Appl.* **3** 345–364.
- Schmeidler, D. 1970. Fatou’s lemma in several dimensions. *Proc. Amer. Math. Soc.* **24** 300–306.
- Schmeidler, D. 1973. Equilibrium points of non-atomic games. *Proc. Amer. Math. Soc.* **24** 300–306.
- Schwartz, L. 1975. *Radon Measures* Oxford University Press, Oxford, U.K.
- Stout, W. F. 1974. *Almost Sure Convergence* Academic Press, New York.
- Valadier, M. 1973. Désintégration d’une mesure sur un produit. *C. R. Acad. Sci. Paris* **276** 33–35.
- . 1990. Young measures. A. Cellina, ed. *Methods of Nonconvex Analysis* Lecture Notes in Math. **1446** Springer-Verlag, Berlin, Germany, 152–188.
- . 1991. La loi forte des grands nombres comme conséquence d’autres théorèmes. *Sém. Analyse Convexe Montpellier* **21** 15.1–15.6.
- Warga, J. 1972. *Optimal Control of Differential and Functional Equations*. Academic Press, New York.
- Yushkevich, A. 1997. The compactness of a policy space in dynamic programming via an extension theorem for Carathéodory functions. *Math. Oper. Res.* **22** 458–467.