On Stability of Multistage Stochastic Decision Problems

Alexander Mänz[1] and Silvia Vogel[2]

[1] ASPECTA Lebensversicherung AG, Germany. AMaenz@aspecta.com
[2] Technische Universitäet Ilmenau, Germany. Silvia.Vogel@tu-ilmenau.de

Summary. The paper considers a general multistage stochastic decision problem which contains Markovian decision processes and multistage stochastic programming problems as special cases. The objective functions, the constraint sets and the probability measures are approximated. Making use of the Bellman Principle, (semi) convergence statements for the optimal value functions and the optimal decisions at each stage are derived. The considerations rely on stability assertions for parametric programming problems which are extended and adapted to the multistage case. Furthermore, new sufficient conditions for the convergence of objective functions which are integrals with respect to decision-dependent probability measures are presented. The paper generalizes results by Langen(1981) with respect to the convergence notions, the integrability conditions and the continuity assumptions.

1 Introduction

Many decision processes go in several steps: The decision maker, who wants to minimize a certain cost functional, chooses an action, obtains further information, reacts to this new aspects, again obtains new information, and so on, up to a finite horizon m. Costs arise at each step or at the end of the decision process only and they can depend on all states and actions observed so far. Often it is assumed that the information which becomes available between the actions can be modelled as random variable, which will be called state and whose distribution is known in advance. Normally, these random variables are not independent and, moreover, their distributions are influenced by foregoing actions.

Markovian decision processes and multistage stochastic programming problems are well-investigated models for decision processes of that kind. Despite several differences they have a similar structure (see e.g. [3]). One important common feature is that the decision maker tries to optimize the expected cost functional.

In the following we will assume that the random total costs, given a sequence of decisions $(x_1, x_2, \ldots, x_m)$ and a sequence of random states

$(S_1, \dots, S_{m+1})$, have the following form:

$$F(S_1, \dots, S_{m+1}, x_1, \dots, x_m) := \sum_{k=1}^{m} c_k(S_1, \dots, S_{k+1}, x_1, \dots, x_k),$$

i.e. we have a sum of stage costs. The terminal costs can be included in $c_m(S_1, \dots, S_{m+1}, x_1, \dots, x_m)$. The aim then consists in finding a sequence $(\vartheta_1, \dots, \vartheta_m)$ of non-anticipative deterministic decision functions that yields minimal expected total costs where the expectation is taken with respect to the common distribution of $S_1, \dots, S_{m+1}$. Usually there are also constraints for the decisions.

Under some natural assumptions on the set of admissible decisions, which will be specified in the next section, the Bellman Principle is applicable, which enables the decision maker to determine the optimal sequence of decision functions (at least theoretically) in a 'backward procedure'. According to the Bellman Principle, at stage k, one has to solve an optimization problem of the following form

$$\min_{x_k \in D_k(\overline{s}_k, \overline{x}_{k-1})} \int_{\mathbb{S}} c_k(\overline{s}_k, s_{k+1}, \overline{x}_{k-1}, x_k) + \Phi_{k+1}(\overline{s}_k, s_{k+1}, \overline{x}_{k-1}, x_k) dP_{k+1|\overline{s}_k, \overline{x}_k}(s_{k+1})$$

where s_k means a realization of the random variable S_k. $\overline{s}_k = (s_1, \dots, s_k)$ describes the so-called state-history and $\overline{x}_{k-1} = (x_1, \dots, x_{k-1})$ the decision history, which are known when the decision x_k has to be chosen. $\Phi_{k+1}(\overline{s}_k, s_{k+1}, \overline{x}_{k-1}, x_k)$ denotes the minimal expected future costs. $P_{k+1|\overline{s}_k, \overline{x}_k}$ is the probability distribution of S_{k+1}, given $\overline{s}_k$ and $\overline{x}_k$. Here $D_k(\overline{s}_k, \overline{x}_{k-1})$ describes the set of admissible decisions, which will also be called 'constraint set'.

Unfortunately, there is often a lack of information about the true probability measures and they have to be approximated. Furthermore, the optimal 'future costs' are usually determined with a certain approximation error only. Hence there is a need for stability statements that clarify under what conditions the optimal values and optimal decision functions of the approximate problems come close to the corresponding quantities for the true problem.

Stability for multistage problems has been dealt with in an L^p-setting for stochastic programming problems with linear or linear-quadratic objective functions ([4], [19]) or via a stage-wise approach, which mainly relies on backward recursion (cf. [10] for Markovian decision processes and [6] for multistage stochastic programming problems).

We shall also use the stage-wise approach and derive qualitative stability results. A general model will be considered, which includes Markovian decision processes and multistage stochastic programming as well. In contrast to most stochastic programming models, we will allow for probability measures that depend on foregoing decisions. Approximations of the state space as considered by Langen [10], however, will not be considered, because approximations of this kind could be widely covered by an appropriate choice of the probability measures. Apart from this exception, we shall give weak sufficient stability conditions which generalize the results by Langen [10] with respect to the convergence notions, the integrability conditions and continuity assumptions. Allowing for discontinuous integrands opens e.g. the possibility to deal with probabilistic objective functions and/or constraints.

The considerations rely on qualitative stability results for one-stage stochastic programs where the probability measure does not depend on the decision ([8], [9]) and extend them to the multistage case and decision-dependent probability measures.

The paper is organized as follows. In Section 2 we provide the mathematical model. Section 3 deals with general parameter-dependent one-stage optimization problems. In Section 4 the special form of the objective functions as parameter-dependent integrals is taken into account. Section 5 combines the results to the multistage setting.

2 Mathematical Model

We base the considerations on the following model: A stage consists of an observation of a state and an action which follows that observation. This agreement is in accordance with the point of view in Markovian decision models. In multistage stochastic programming problems a stage usually starts with an action, thus our model has to be specialized to apply to this case. In order to investigate a model which is as general as possible, we decided to consider stage costs c_k which may depend also on s_{k+1}, c.f. [15].

In what follows, m denotes the number of stages, the so-called horizon, and by N_m we mean the set $\{1, \ldots, m\}$.

We base our considerations on the investigations in [5] and [15]. The states or observations s_k at stage k are assumed to be elements of a standard Borel space $\mathbb{S}$, i.e. a non-empty Borel subset of a complete, separable, metric space, provided with the system of Borel subsets (to simplify presentation, here and in the following, we omit an additional symbol for the system of Borel subsets). The actions or decisions x_k are taken from a standard Borel space $\mathbb{A}$. The sets of possible actions can be constrained by certain conditions which can depend on the history so far. These conditions are described by means of multifunctions D_k. In order to explain these multifunctions we will use the following abbreviations: Let $\mathbb{H}_{1,k} := \mathbb{S}^k$, $k \in N_{m+1}$, and $\mathbb{H}_{2,k} := \mathbb{A}^k$, $k \in N_m$.

Now $D_k : \mathbb{H}_{1,k} \times \mathbb{H}_{2,k-1} \to 2^{\mathbb{A}}$, $k \in N_m \setminus \{1\}$, and $D_1 : \mathbb{H}_{1,1} \to 2^{\mathbb{A}}$ are multifunctions which determine for histories $(\overline{s}_k, \overline{x}_{k-1})$ and s_1 the constraint sets or sets of admissible actions $D_k(\overline{s}_k, \overline{x}_{k-1})$ and $D_1(s_1)$, respectively. We assume that all multifunctions D_k are closed-valued.

The probability measures $P_{k+1|\cdot,\cdot} : \mathbb{H}_{1,k} \times \mathbb{H}_{2,k} \to \mathcal{P}(\mathbb{S})$, $k \in N_m \setminus \{1\}$, describe how the state history $\overline{s}_k \in \mathbb{H}_{1,k}$ and the decision history $\overline{x}_k \in \mathbb{H}_{2,k}$ influence the probability distribution of the observation in stage $k+1$. $\mathcal{P}(\mathbb{S})$

means the set of all probability measures on the σ-field of Borel sets $\mathcal{B}(\mathbb{S})$ of $\mathbb{S}$. $P_1 \in \mathcal{P}(\mathbb{S})$ is the distribution of the first state.

The aim now consists in finding an optimal strategy (or policy, plan), i.e. a sequence of decision rules which tells the decision maker at each stage how to decide, given the foregoing observations and actions. Thus we define a strategy as a sequence $\vartheta = (\delta_k)_{k=1,\dots,m}$ of decision functions $\delta_1 : \mathbb{H}_{1,1} \to \mathbb{A}$ and $\delta_k : \mathbb{H}_{1,k} \times \mathbb{H}_{2,k-1} \to \mathbb{A}$. A sequence $(\overline{s}_k)_{k \in N_{m+1}}$ of observation histories and a strategy ϑ then define recursively a sequence of actions $(x_k(\overline{s}_k, \vartheta))_{k \in N_m}$ and decision histories $(\overline{x}_k(\overline{s}_k, \vartheta))_{k \in N_m}$ via

$$\overline{x}_1(s_1, \vartheta) = x_1(\overline{s}_1, \vartheta) := \delta_1(s_1),$$

$$x_k(\overline{s}_k, \vartheta) := \delta_k(\overline{s}_k, \overline{x}_{k-1}(\overline{s}_{k-1}, \vartheta)),\ \overline{x}_k(\overline{s}_k, \vartheta) := (\overline{x}_{k-1}(\overline{s}_{k-1}, \vartheta), x_k(\overline{s}_k, \vartheta)).$$

Thus probability measures $P_{k+1|\overline{s}_k,\vartheta}$ on $\mathcal{B}(\mathbb{S})$ can be defined by

$$P_{k+1|\overline{s}_k,\vartheta}(B) := P_{k+1|\overline{s}_k,\overline{x}_k(\overline{s}_k,\vartheta)}(B).$$

We assume that $\delta_k,\ k = 1, \dots, m$, are Borel-measurable functions of their arguments. In order to guarantee this property for an optimal strategy we suppose that the cost functions $c_k : \mathbb{H}_{1,k+1} \times \mathbb{H}_{2,k} \to \mathbb{R} \cup \{+\infty\}$ are measurable with respect to the product sigma field of all arguments and that the graphs of the constraint multifunctions are measurable. Furthermore, we suppose that for each $B \in \mathcal{B}(\mathbb{S})$ the functions $(\overline{s}_k, \overline{x}_k) \to P_{k+1|\overline{s}_k,\overline{x}_k}(B)$ are Borel-measurable.

Then we can base our considerations on the measurable space $[\Omega_T, \Sigma]$ with

$$\Omega_T = \mathbb{S}^{m+1},\ \Sigma = \bigotimes_{i=1}^{m+1} \mathcal{B}(\mathbb{S}) \quad \text{and} \quad S_k(\omega) = s_k \text{ for } \omega = \overline{s}_{m+1} \in \Omega.$$

Using the abbreviation $\overline{S}_i := (S_1, \dots, S_i)$, a probability measure P_ϑ on $[\Omega_T, \Sigma]$ is defined by $P_\vartheta(S_1 \in B) := P_1(B)$, and

$$P_\vartheta(S_k \in B | \overline{S}_{k-1} = \overline{s}_{k-1}) := P_{k|\overline{s}_{k-1},\vartheta}(B),\ \ B \in \mathcal{B}(\mathbb{S}),\ k \geq 2.$$

We will call a strategy $\vartheta = (\delta_k)_{k=1,\dots,m}$ *admissible*, if $\delta_1(s_1) \in D_1(s_1)$ for all $s_1 \in \text{supp}P_1$, and, for $k \in N_m \setminus \{1\}$,

$$\delta_k(\overline{s}_k, \overline{x}_{k-1}(\overline{s}_{k-1}, \vartheta)) \in D_k(\overline{s}_k, \overline{x}_{k-1}(\overline{s}_{k-1}, \vartheta)) \text{ for all } \overline{s}_k \in \text{supp}P_{k|\overline{s}_{k-1},\vartheta}$$

where supp denotes the support of a probability measure. The set of admissible strategies will be denoted by Θ.

We exclude induced constraints, i.e. we assume that $D_k(\overline{s}_k, \overline{x}_{k-1}(\overline{s}_{k-1}, \vartheta))$ is nonempty for all admissible ϑ, all $k \in N_m \setminus \{1\}$, and all $\overline{s}_k \in \text{supp}P_{k|\overline{s}_{k-1},\vartheta}$.

Now, for a given strategy ϑ, the random total costs can be written in the form

$$F_\vartheta(\omega) := \sum_{k=1}^{m} c_k(\overline{S}_{k+1}(\omega), \overline{x}_k(\overline{S}_k(\omega), \vartheta)).$$

The task for the decision maker consists in finding a strategy $\vartheta^* \in \Theta$ such that

$$\min_{\vartheta \in \Theta} \mathbb{E}_\vartheta F_\vartheta = \mathbb{E}_{\vartheta^*} F_{\vartheta^*}$$

where $\mathbb{E}_\vartheta$ denotes the expectation with respect to P_ϑ. We assume that there is at least one strategy ϑ such that $\mathbb{E}_\vartheta F_\vartheta < \infty$.

Given a history $(\overline{s}_m, \overline{x}_{m-1}) \in \mathbb{H}_{1,m} \times \mathbb{H}_{2,m-1}$, an optimal decision x_m^* can be obtained by

$$\inf_{x \in D_m(\overline{s}_m, \overline{x}_{m-1})} \int_{\mathbb{S}} c_m(\overline{s}_m, s, \overline{x}_{m-1}, x) dP_{m+1|\overline{s}_m, \overline{x}_{m-1}, x}(s)$$
$$= \int_{\mathbb{S}} c_m(\overline{s}_m, s, \overline{x}_{m-1}, x_m^*) dP_{m+1|\overline{s}_m, \overline{x}_{m-1}, x_m^*}(s) =: \Phi_m(\overline{s}_m, \overline{x}_{m-1}).$$

Furthermore, for $k = m-1, \dots, 1$, x_k^* is obtained by

$$\inf_{x \in D_k(\overline{s}_k, \overline{x}_{k-1})} \int_{\mathbb{S}} c_k(\overline{s}_k, s, \overline{x}_{k-1}, x) + \Phi_{k+1}(\overline{s}_k, s, \overline{x}_{k-1}, x) dP_{k+1|\overline{s}_k, \overline{x}_{k-1}, x}(s)$$
$$= \int_{\mathbb{S}} c_k(\overline{s}_k, s, \overline{x}_{k-1}, x_k^*) + \Phi_{k+1}(\overline{s}_k, s, \overline{x}_{k-1}, x_k^*) dP_{k+1|\overline{s}_k, \overline{x}_{k-1}, x_k^*}(s)$$
$$=: \Phi_k(\overline{s}_k, \overline{x}_{k-1}).$$

In order to avoid permanent distinction between the cases $k = 1$ and $k > 1$, here and in the following we assume that dependence on a parameter x_{k-1} for $k = 1$ is ignored.

The above equations open the possibility to carry over results from one-stage optimization problems to the multi-stage case. Note that with the agreement $\Phi_{m+1}(\overline{s}_{m+1}, \overline{x}_m) := 0 \;\; \forall (\overline{s}_{m+1}, \overline{x}_m) \in \mathbb{H}_{1,m+1} \times \mathbb{H}_{2,m}$ there is a uniform structure for $k = m, \dots, 1$.

It should be mentioned that Markovian decision processes as investigated by Langen [10] fit into this framework with the following agreements: $c_k(\overline{s}_{k+1}, \overline{x}_k) = \beta(s_1, x_1, s_2) \cdot \ldots \cdot \beta(s_{k-1}, x_{k-1}, s_k) \cdot r(s_k, x_k)$ where β denotes the bounded discount factor and r the reward function. Furthermore, $D_k(\overline{s}_k, \overline{x}_{k-1}) = \tilde{D}(s_k)$ and $P_{k+1|\overline{s}_k, \overline{x}_k}(B) = q(s_k, x_k, B)$ for a transition function q.

The well-investigated two-stage stochastic programming problems are obtained via $m = 2$, $c_1(s_1, s_2, x_1,) = \tilde{c}_1(x_1)$, $c_2(s_1, s_2, s_3, x_1, x_2) = \tilde{c}_2(x_1, s_2, x_2)$, $P_{2|s_1, x_1} = \tilde{P}_2$, $P_{3|\overline{s}_2, \overline{x}_2}$ arbitrary, $D_1(s_1) = \tilde{D}_1$, $D_2(s_1, x_1, s_2) = \tilde{D}_2(x_1, s_2)$.

Now we assume that each of the determining components $(D_k)_{k \in N_m}$, $(P_{k+1|\cdot,\cdot})_{k \in N_{m+1}}$, and $(c_k)_{k \in N_m}$ of our original model is approximated by a sequence in a suitable sense. Consequently, we have to investigate approximate models

$$(DM^{(n)}) \qquad (D_k^{(n)})_{k \in N_m}, \; (P_{k|\cdot,\cdot}^{(n)})_{k \in N_{m+1}}, \; (c_k^{(n)})_{k \in N_m}.$$

In the following, the original model will be indicated by the superscript '$^{(0)}$':

$$(DM^{(0)}) \qquad (D_k^{(0)})_{k \in N_m}, \; (P_{k|\cdot,\cdot}^{(0)})_{k \in N_{m+1}}, \; (c_k^{(0)})_{k \in N_m}.$$

For all problems $(DM^{(n)})$, $n \in N_0 := \{0, 1, \ldots, \}$, we impose the same assumptions as for the original problem. Hence we can proceed as indicated above and solve an optimization problem at each stage.

We will use the following abbreviations for $n \in N_0$:

$$\begin{aligned} \Phi_{m+1}^{(n)}(\overline{s}_{m+1}, \overline{x}_m) &:= 0, \text{ and, for } k \in N_m, \\ \varphi_k^{(n)}(\overline{s}_{k+1}, \overline{x}_k) &:= c_k^{(n)}(\overline{s}_{k+1}, \overline{x}_k) + \Phi_{k+1}^{(n)}(\overline{s}_{k+1}, \overline{x}_k), \\ f_k^{(n)}(\overline{s}_k, \overline{x}_{k-1}, x_k) &:= \int_{\mathbb{S}} \varphi_k^{(n)}(\overline{s}_k, s, \overline{x}_{k-1}, x_k) dP_{k+1|\overline{s}_k, \overline{x}_{k-1}, x_k}^{(n)}(s), \\ \Phi_k^{(n)}(\overline{s}_k, \overline{x}_{k-1}) &:= \inf_{x_k \in D_k^{(n)}(\overline{s}_k, \overline{x}_{k-1})} f_k^{(n)}(\overline{s}_k, \overline{x}_{k-1}, x_k). \end{aligned}$$

Furthermore, we introduce - for the original and the approximate problems - the so-called solution sets for each stage $k \in N_m$, which contain the optimal decisions:

$$W_k^{(n)}(\overline{s}_k, \overline{x}_{k-1}) := \left\{ x_k \in D_k^{(n)}(\overline{s}_k, \overline{x}_{k-1}) \ : \ f_k^{(n)}(\overline{s}_k, \overline{x}_{k-1}, x_k) = \Phi_k^{(n)}(\overline{s}_k, \overline{x}_{k-1}) \right\}.$$

Our aim consists in deriving conditions which ensure that the approximate problems yield strategies $\vartheta^{(n)}$ such that $\lim\limits_{n \to \infty} \mathbb{E}_{\vartheta^{(n)}} F_{\vartheta^{(n)}} = \mathbb{E}_{\vartheta^*} F_{\vartheta^*}$.

Sufficient for this equality are, for instance, the following two conditions (where $H_{1,k}$ and D_k are suitable sets which will be specified in Section 5):

(a) for each $k \in \{2, \ldots, m\}$, all $\overline{s}_k \in H_{1,k}$, all $\overline{x}_{k-1}^{(0)} \in D_{k-1}$, all sequences $(\overline{s}_k^{(n)}, \overline{x}_{k-1}^{(n)})_{n \in N} \to (\overline{s}_k^{(0)}, \overline{x}_{k-1}^{(0)})$, one has $\lim\limits_{n \to \infty} \Phi_k^{(n)}(\overline{s}_k^{(n)}, \overline{x}_{k-1}^{(n)}) = \Phi_k^{(0)}(\overline{s}_k^{(0)}, \overline{x}_{k-1}^{(0)})$,

(b) for P_1-almost all s, $\lim\limits_{n \to \infty} \Phi_1^{(n)}(s) = \Phi_1^{(0)}(s)$.

Furthermore, it will be shown that the conditions which guarantee these equations also yield K-$\limsup\limits_{n \to \infty} W_k^{(n)}(\overline{s}_k^{(0)}, \overline{x}_{k-1}^{(0)}) \subset W_k^{(0)}(\overline{s}_k^{(0)}, \overline{x}_{k-1}^{(0)})$ for all $\overline{s}_k \in H_{1,k}$ and all $\overline{x}_{k-1}^{(0)} \in D_{k-1}$, and K–$\limsup\limits_{n \to \infty} W_1^{(n)}(s) \subset W_1^{(0)}(s)$ for P_1-almost all s.

Moreover, we will provide conditions under which even 'pointwise' convergence with respect to the state histories (together with continuous convergence with respect to the decisions) yields the desired statement.

The so-called outer Kuratowski-Painlevé-Limes K–$\limsup\limits_{n \to \infty}$ is defined in the following section. Aiming at approximating the whole solution set of the original problem would require rather strong conditions and is in fact more than one really needs.

3 Stability of Parametric One-Stage Problems

We consider the optimization problem which occurs at a fixed stage k. In comparison to one-stage problems, in the multistage-stage setting there is mainly one new aspect that has to be coped with, namely dependance on the parameter 'history', which can occur in the constraint sets, the integrands and the probability measures. Because for stability investigations the states and the actions have to be handled in a different way, we will distinguish the state and the decision history in the functions and multifunctions under consideration.

We shall investigate optimization problems of the following form

$$(P^{(n)}(\overline{s},\overline{x})) \qquad \min_{x\in D^{(n)}(\overline{s},\overline{x})} f^{(n)}(\overline{s},\overline{x},x),$$

with $n \in N_0$. Here $\overline{s}$ denotes an element of a standard Borel space $\mathbb{H}_1$ and $\overline{x}$ denotes an element of a standard Borel space $\mathbb{H}_2$. The optimal value functions will be denoted by $\Phi^{(n)}$ and the solution set multifunctions by $W^{(n)}$.

The optimization problems which occur at stage $k \in N_m$ in the 'backward procedure' have form $(P^{(n)}(\overline{s},\overline{x}))$. $\overline{s}$ may be regarded as 'state history' and $\overline{x}$ as 'decision history', which is available when the new decision has to be chosen. $\Phi^{(n)}(\overline{s},\overline{x})$ means the optimal 'rest costs' and $W^{(n)}(\overline{s},\overline{x})$ denotes the set of optimal decisions given the history $(\overline{s},\overline{x})$. As the optimal value function at stage k is a main part of the objective function for stage $k-1$, we aim at deriving stability assertions for the optimal value functions which are known to be desirable for the objective functions. Continuous convergence of the objective functions of a sequence of optimizations problems has proved to be an appropriate convergence notion for stability considerations. If the constraint sets remain fixed, the continuous convergence condition can be weakened. In one-stage problems often epi-convergence is imposed. However, the sum of two epi-convergent sequences is in general not epi-convergent. Hence we adapt a condition which was considered by Langen (for maximization problems) and in [10] called upper-semi-continuous convergence. We shall call this kind of convergence lower semicontinuous pointwise convergence. For the constraint sets and the solution sets we need the concept of Kuratowski-Painlevé convergence.

In [1] and [13] nowadays classical stability results for parametric optimization problems are compiled. Here we have to deal with three kinds of parameters. Firstly, the upper (approximation) index $^{(n)}$ can be interpreted as parameter. Therefore semicontinuous behavior in [1], [13] appears here in the form of semi-approximations. Furthermore, there are the parameters $\overline{s}$ and $\overline{x}$. They play a different role in stability considerations, see below. Hence we have to modify the classical results to apply to our special parametric sequences.

We start recalling the definition of Kuratowski-Painlevé-convergence:

Definition 1. *Let $(M_n)_{n\in\mathbb{N}}$ be a sequence of nonempty sets in $\mathbb{A}$. Then the Limes superior (in the Kuratowski-Painlevé sense) or 'outer limit' K–$\limsup_{n\to\infty} M_n$ and the Limes inferior (in the Kuratowski-Painlevé sense) or 'inner limit' K–$\liminf_{n\to\infty} M_n$ are defined by*

$$K\text{–}\limsup_{n\to\infty} M_n := \left\{ x \in \mathbb{A} \mid \begin{array}{l} \exists (x_n)_{n\in\mathbb{N}} \to x \text{ such that} \\ \forall n \in \mathbb{N}\ \exists m > n:\ x_m \in M_m \end{array} \right\},$$
$$K\text{–}\liminf_{n\to\infty} M_n := \left\{ x \in \mathbb{A} \mid \begin{array}{l} \exists (x_n)_{n\in\mathbb{N}} \to x \text{ such that} \\ \exists n_0 \in \mathbb{N}\ \forall n > n_0:\ x_n \in M_n \end{array} \right\}.$$

If both limits coincide, the Kuratowski-Painlevé -Limes $K\text{–}\lim_{n\to\infty}$ exists:

$$K\text{–}\lim_{n\to\infty} M_n := K\text{–}\limsup_{n\to\infty} M_n = K\text{–}\liminf_{n\to\infty} M_n.$$

We have to extend these notions to multifunctions $\{C^{(n)},\ n \in N_0\}$ which map into the Borel sets of $\mathbb{A}$ and are defined on the cross product of standard Borel spaces $\tilde{\mathbb{H}}_1 \times \tilde{\mathbb{H}}_2$. $\tilde{\mathbb{H}}_1$ and $\tilde{\mathbb{H}}_2$ may be different from $\mathbb{H}_1$ and $\mathbb{H}_2$, respectively. We introduce the new notions, because we do not need semicontinuous behavior with respect to all history parameters. Semicontinuous behavior with respect to the actions is assumed in the stability statements for the optimization problems and cannot be dispensed with. Semicontinuity assumptions with respect to the states are convenient for the derivation of sufficient conditions in Section 4. They can, however, often be replaced with pointwise convergence, compare, e.g. Theorem 4. Semicontinuous behavior with respect to s_1 is never needed. Thus an element of $\tilde{\mathbb{H}}_1$ can be understood as a whole state history for the stage under consideration or the first state only. Then the elements of $\tilde{\mathbb{H}}_2$ are in the first case the action histories and in the second case the state histories except s_1 and the complete action histories. Thus we always have $\tilde{\mathbb{H}}_1 \times \tilde{\mathbb{H}}_2 = \mathbb{H}_1 \times \mathbb{H}_2$.

In the following G_i denotes a Borel subset of $\tilde{\mathbb{H}}_i$, $i = 1, 2$.

Definition 2. *Let $\{C^{(n)},\ n \in N_0\}$ be a family of multifunctions $C^{(n)} : \tilde{\mathbb{H}}_1 \times \tilde{\mathbb{H}}_2 \to 2^{\mathbb{A}}$. The sequence $(C^{(n)})_{n\in N}$ is said to be*

(i) an inner semi-approximation to $C^{(0)}$ on G_2 given G_1 (abbreviated $C^{(n)} \xrightarrow[G_2|G_1]{K-i} C^{(0)}$) if

$$\forall s \in G_1\ \forall y \in G_2\ \forall (y_n)_{n\in\mathbb{N}} \to y : K\text{–}\limsup_{n\to\infty} C^{(n)}(s, y_n) \subset C^{(0)}(s, y),$$

(ii) an outer semi-approximation to $C^{(0)}$ on G_2 given G_1 (abbreviated $C^{(n)} \xrightarrow[G_2|G_1]{K-o} C^{(0)}$) if

$$\forall s \in G_1\ \forall y \in G_2\ \forall (y_n)_{n\in\mathbb{N}} \to y : K\text{–}\liminf_{n\to\infty} C^{(n)}(s, y_n) \supset C^{(0)}(s, y).$$

(iii) convergent in the Kuratowski-Painlevé sense to $C^{(0)}$ on G_2 given G_1 (abbreviated $C^{(n)} \xrightarrow[G_2|G_1]{K} C^{(0)}$) if

$$C^{(n)} \xrightarrow[G_2|G_1]{K-i} C^{(0)} \quad \text{and} \quad C^{(n)} \xrightarrow[G_2|G_1]{K-o} C^{(0)}.$$

Now we introduce the convergence notions for sequences of functions we shall deal with. $\tilde{\mathbb{A}}$ denotes a standard Borel space. In the following $\tilde{\mathbb{A}}$ will usually be interpreted as $\tilde{\mathbb{H}}_2 \times \mathbb{A}$.

Definition 3. *Let $\{f^{(n)}, n \in N_0\}$ be a family of functions $f^{(n)} : \tilde{\mathbb{H}}_1 \times \tilde{\mathbb{A}} \to \overline{\mathbb{R}}$ and C a Borel subset of $\tilde{\mathbb{A}}$. The sequence $(f^{(n)})_{n\in N}$ is said to be a*

(i) lower semicontinuous approximation to $f^{(0)}$ on C given G_1 (abbreviated $f^{(n)} \xrightarrow[C|G_1]{l} f^{(0)}$) if

$$\forall s \in G_1 \; \forall y \in C \; \forall (y_n)_{n\in\mathbb{N}} \to y : \liminf_{n\to\infty} f^{(n)}(s, y_n) \geq f^{(0)}(s, y),$$

(ii) upper semicontinuous approximation to $f^{(0)}$ on C given G_1 (abbreviated $f^{(n)} \xrightarrow[C|G_1]{u} f^{(0)}$), if

$$-f^{(n)} \xrightarrow[C|G_1]{l} -f^{(0)},$$

(iii) continuously convergent to $f^{(0)}$ on C given G_1 (abbreviated $f^{(n)} \xrightarrow[C|G_1]{c} f^{(0)}$) if

$$f^{(n)} \xrightarrow[C|G_1]{l} f^{(0)} \text{ and } f^{(n)} \xrightarrow[C|G_1]{u} f^{(0)},$$

(iv) lower semicontinuously pointwise convergent to $f^{(0)}$ on C given G_1 (abbreviated $f^{(n)} \xrightarrow[C|G_1]{l-p} f^{(0)}$) if

$$f^{(n)} \xrightarrow[C|G_1]{l} f^{(0)} \quad \text{and} \quad \forall s \in G_1 \; \forall y \in C : \lim_{n\to\infty} f^{(n)}(s, y) = f^{(0)}(s, y).$$

In order to employ results from parametric programming in our setting, the following lemmas are helpful. $\mathbb{A}_1$ denotes an auxiliary metric space.

Recall that a multifunction $\hat{C} : \tilde{\mathbb{H}}_2 \times \mathbb{A}_1 \to 2^{\mathbb{A}}$ is *closed* at a point (y_0, λ_0), if for all pairs of sequences $(y_n, \lambda_n)_{n\in N}$ and $(x_n)_{n\in N}$ with the properties $(y_n, \lambda_n) \to (y_0, \lambda_0)$, $x_n \in \hat{C}(y_n, \lambda_n)$ and $x_n \to x_0$ the property $x_0 \in \hat{C}(y_0, \lambda_0)$ follows. A multifunction $\hat{C} : \tilde{\mathbb{H}}_2 \times \mathbb{A}_1 \to 2^{\mathbb{A}}$ is *lower semicontinuous (l.s.c.) in the sense of Berge* at a point (y_0, λ_0), if for each open set Q satisfying $Q \cap \hat{C}(y_0, \lambda_0) \neq \emptyset$ there exists a neighborhood $U(y_0, \lambda_0)$ of (y_0, λ_0) such that for all $(y, \lambda) \in U(y_0, \lambda_0)$ the set $\hat{C}(y, \lambda) \cap Q$ is non-empty.

Lemma 1. *Let a family $\Lambda := \{\lambda_n, n \in N_0\}$ of elements of $\mathbb{A}_1$ with $\lambda_n \to \lambda_0$ and a multifunction $\hat{C} : \tilde{\mathbb{H}}_1 \times \tilde{\mathbb{H}}_2 \times \Lambda \to 2^{\mathbb{A}}$ be given. Suppose that $C^{(n)}(s, y) := \hat{C}(s, y, \lambda_n)$, $n \in N_0$, $\lambda_n \in \Lambda$. Furthermore, assume that for all $s \in G_1$ the multifunction $\hat{C}(s, \cdot, \lambda_0)$ is closed-valued. Then*

(i) $C^{(n)} \xrightarrow[G_2|G_1]{K-i} C^{(0)} \iff \forall s \in G_1,\ \hat{C}(s, \cdot, \cdot)$ is closed on $G_2 \times \{\lambda_0\}$,

(ii) $C^{(n)} \xrightarrow[G_2|G_1]{K-o} C^{(0)} \iff \forall s \in G_1,\ \hat{C}(s, \cdot, \cdot)$ is l.s.c. in the sense of Berge on $G_2 \times \{\lambda_0\}$.

Lemma 2. *Let a family $\Lambda := \{\lambda_n,\ n \in N_0\}$ of elements of $\mathbb{A}_1$ with $\lambda_n \to \lambda_0$, a function $\hat{f} : \tilde{\mathbb{H}}_1 \times \tilde{\mathbb{A}} \times \Lambda \to \overline{\mathbb{R}}$ and a Borel subset $C \subset \tilde{\mathbb{A}}$ be given. Suppose that $f^{(n)}(s,y) := \hat{f}(s,y,\lambda_n)$, $n \in N_0$, $\lambda_n \in \Lambda$. Furthermore, assume that for all $s \in G_1$ the function $\hat{f}(s,\cdot,\lambda_0)$ is l.s.c. Then, $f^{(n)} \xrightarrow[C|G_1]{l} f^{(0)} \iff \forall s \in G_1$, $\hat{f}(s,\cdot,\cdot)$ is l.s.c. on $C \times \{\lambda_0\}$.*

Combining these assertions, corresponding statements can be derived for continuous convergence and lower semicontinuous pointwise convergence. The proofs of the lemmas are straightforward and will be omitted. Note that the closed-valuedness and lower semicontinuity, respectively, are needed for the '$\Rightarrow$'-direction of the proofs only.

In order to formulate the stability results for our setting, we use the following assumptions. Let $C(G_1,G_2) := \{(\overline{y},x) :\ x \in D^{(0)}(\overline{s},\overline{y}),\ \overline{s} \in G_1,\ \overline{y} \in G_2\}$.

(A1) For all $\overline{s} \in G_1$, the function $f^{(0)}(\overline{s},\cdot,\cdot)$ is u.s.c. on $C(G_1,G_2)$ and $f^{(n)} \xrightarrow[C(G_1,G_2)|G_1]{u} f^{(0)}$.

(A2) For all $\overline{s} \in G_1$ and $\overline{y} \in G_2$, there exists $x^{(0)}(\overline{s},\overline{y}) \in W^{(0)}(\overline{s},\overline{y})$ such that $f^{(0)}(\overline{s},\cdot,\cdot)$ is u.s.c. at $(\overline{y},x^{(0)})$ and $f^{(n)} \xrightarrow[\{(\overline{y},x^{(0)}(\overline{s},\overline{y}))\}|\{\overline{s}\}]{u} f^{(0)}$.

Now, for instance, the following statements can be proved making use of results in [1, Chapter 4] and [13].

Theorem 1. *(i) Let (A1) or (A2) hold and assume that $D^{(n)} \xrightarrow[G_2|G_1]{K-o} D^{(0)}$. Then $\Phi^{(0)}(\overline{s},\cdot)$ is u.s.c. on G_2 and $\Phi^{(n)} \xrightarrow[G_2|G_1]{u} \Phi^{(0)}$.*

(ii) Let $f^{(0)}(\overline{s},\cdot,\cdot)$ be l.s.c. on $C(G_1,G_2)$ for all $\overline{s} \in G_1$ and assume that $f^{(n)} \xrightarrow[C(G_1,G_2)|G_1]{l} f^{(0)}$, $D^{(n)} \xrightarrow[G_2|G_1]{K-i} D^{(0)}$. Furthermore, suppose that for all $\overline{s} \in G_1$ and all $\overline{y} \in G_2$ there is a compact set K such that for all sequences $(y_n)_{n\in N} \to \overline{y}$ there is an n_0 with $D^{(n)}(\overline{s},y_n) \subset K\ \forall n \geq n_0$. Then, $\Phi^{(0)}(\overline{s},\cdot)$ is l.s.c. on G_2 and $\Phi^{(n)} \xrightarrow[G_2|G_1]{l} \Phi^{(0)}$.

(iii) Let $\Phi^{(n)} \xrightarrow[G_2|G_1]{u} \Phi^{(0)}$. Furthermore, assume that, for all $\overline{s} \in G_1$, $f^{(0)}(\overline{s},\cdot,\cdot)$ is l.s.c. on $C(G_1,G_2)$, $f^{(n)} \xrightarrow[C(G_1,G_2)|G_1]{l} f^{(0)}$, and $D^{(n)} \xrightarrow[G_2|G_1]{K-i} D^{(0)}$. Then, $W^{(n)} \xrightarrow[G_2|G_1]{K-i} W^{(0)}$.

If the constraint set does not vary with n, the continuity and continuous convergence conditions can be weakened.

Theorem 2. *Suppose that for all $\overline{s} \in G_1$ and all $\overline{y} \in G_2$ there is a nonempty compact set $D(\overline{s},\overline{y})$ with $D^{(n)}(\overline{s},\overline{y}) = D(\overline{s},\overline{y})\ \forall n \in N_0$. Furthermore, assume that for all $\overline{s} \in G_1$ the function $f^{(0)}(\overline{s},\cdot,\cdot)$ is l.s.c. on $C(G_1,G_2)$, and $f^{(n)} \xrightarrow[C(G_1,G_2)|G_1]{l-p} f^{(0)}$. Then $\Phi^{(0)}(\overline{s},\cdot)$ is l.s.c. on G_2, $\Phi^{(n)} \xrightarrow[G_2|G_1]{l-p} \Phi^{(0)}$ and, for all $\overline{s} \in G_1$ and all $\overline{y} \in G_2$, the inclusion $K\text{–}\limsup_{n\to\infty} W^{(n)}(\overline{s},\overline{y}) \subset W^{(0)}(\overline{s},\overline{y})$ holds.*

Proof. Taking Theorem 1(ii) into account, we still have to show that for all $\overline{s} \in G_1$ and all $\overline{y} \in G_2$ $\limsup_{n\to\infty} \Phi^{(n)}(\overline{s},\overline{y}) \leq \Phi^{(0)}(\overline{s},\overline{y})$ and K–$\limsup_{n\to\infty} W^{(n)}(\overline{s},\overline{y}) \subset W^{(0)}(\overline{s},\overline{y})$ hold. Let $\overline{s} \in G_1$ and $\overline{y} \in G_2$ be fixed. $f^{(0)}(\overline{s},\overline{y},\cdot)$ being l.s.c. and $D(\overline{s},\overline{y})$ being compact, there is an $x \in D(\overline{s},\overline{x})$ such that $\Phi^{(0)}(\overline{s},\overline{y}) = f^{(0)}(\overline{s},\overline{y},x)$. Consequently,

$$\limsup_{n\to\infty} \Phi^{(n)}(\overline{s},\overline{y}) \leq \limsup_{n\to\infty} f^{(n)}(\overline{s},\overline{y},x) \leq f^{(0)}(\overline{s},\overline{y},x) = \Phi^{(0)}(\overline{s},\overline{y}).$$

Now, assume that there is a sequence $(x_{n_k})_{k\in N}$ with $x_{n_k} \in W^{(n_k)}(\overline{s},\overline{y})$ and $x_{n_k} \to x_0 \notin W^{(0)}(\overline{s},\overline{y})$. $x_{n_k} \in W^{(n_k)}(\overline{s},\overline{y})$ implies $x_0 \in D(\overline{s},\overline{y})$. Otherwise there is an $x \in W^{(0)}(\overline{s},\overline{y})$, consequently, $f^{(0)}(\overline{s},\overline{y},x) < f^{(0)}(\overline{s},\overline{y},x_0)$. Thus, because of $\lim_{n\to\infty} \Phi^{(n)}(\overline{s},\overline{y}) = \Phi^{(0)}(\overline{s},\overline{y})$, we have

$$\lim_{k\to\infty} f^{(n_k)}(\overline{s},\overline{y},x_{n_k}) = f^{(0)}(\overline{s},\overline{y},x) < f^{(0)}(\overline{s},\overline{y},x_0)$$

in contradiction to $\liminf_{n\to\infty} f^{(n)}(\overline{s},\overline{y},x_n) \geq f^{(0)}(\overline{s},\overline{y},x_0)$. □

Combining and specializing the above results, e.g. Theorem 2.8 in [10] can be derived (if approximations of the state space are not taken into account).

For multifunctions $D^{(n)}$, which are described by inequality constraints, sufficient conditions are available (see e.g. [1, 17]). Semicontinuous convergence of the constraint functions plays a central role in these statements too.

4 Sufficient Conditions for Continuous Convergence and Epi-Convergence

In this section we investigate lower semicontinuous convergence for functions which are integrals. The results can then be employed to obtain sufficient conditions for either continuous convergence or (together with assertions on pointwise convergence) for lower semicontinuous pointwise convergence and hence also for epi-convergence. Corollary 3.4 in [10] and further results that rely on Corollary 3.4 give sufficient conditions assuming weak convergence of the probability measures, continuous convergence or upper-semi-continuous convergence of the integrands and uniform (with respect to the decision and the history) boundedness of the integrands. We will - among other generalizations - particularly weaken the uniform boundedness condition and the convergence condition with respect to the states.

For $i_n = n$ and $i_n = 0$, we shall investigate functions $f^{(i_n)} : \mathbb{H}_1 \times \mathbb{H}_2 \times \mathbb{A} \to \overline{\mathbb{R}}$ of the following form:

$$f^{(i_n)}(\overline{s},\overline{x},x) = \int_{\mathbb{S}} \varphi^{(i_n)}(\overline{s},s,\overline{x},x)dP^{(i_n)}_{\overline{s},\overline{x},x}(s)$$

where $P^{(i_n)}_{\overline{s},\overline{x},x}$, $n \in N_0$, are probability measures on $\mathcal{B}(\mathbb{S})$ and $\varphi^{(i_n)} : \mathbb{H}_1 \times \mathbb{S} \times \mathbb{H}_2 \times \mathbb{A} \to \overline{\mathbb{R}}$, $n \in N_0$, are integrands which are supposed to be measurable with respect to the product-σ-algebra of the arguments and integrable with respect to the probability measures under consideration. The 'parameter' i_n has been introduced in order to reduce the effort for the notation and the proof of the results, because usually the same considerations lead either (for $i_n = n$) to semi-approximation properties of $(f^{(n)})_{n\in N}$ or (for $i_n = 0$) to semicontinuity of $f^{(0)}$.

Sufficient conditions for semicontinuous convergence of sequences of functions which are integrals with respect to a probability measure that does not depend on the decision are given in [9]. We will extend the results of [9] to the parameter-dependent case. Two approaches are suggested: The first one (so-called direct approach, which was suggested by P. Lachout), assumes weak convergence of the probability measures, a lower semi-approximation property for the integrands and lower equi-integrability defined below. It can be employed to generalize Theorem 3.3 in [10]. The second approach (so called pointwise approach [18] or scalarization [7]) reduces convergence considerations for sequences of functions to convergence of sequences of real values. It is especially favorable in a random setting, but works in our case as well. It may be regarded as a bridge to results of asymptotic statistics and limit theorems of probability theory. Furthermore, it does not assume that the integrands 'behave semicontinuously' with respect to the state history.

The direct approach uses the following definition [9]:

Definition 4. *Let a sequence $(\hat{\varphi}^{(n)})_{n\in\mathbb{N}}$ of Borel-measurable functions $\hat{\varphi}^{(n)} : \mathbb{S} \to \overline{\mathbb{R}}$ and a sequence $(P^{(n)})_{n\in\mathbb{N}}$ of probability measures on $\mathcal{B}(\mathbb{S})$ be given. The family $\{(\hat{\varphi}^{(n)}, P^{(n)}), n \in \mathbb{N}\}$ is called lower equi-integrable, if there exists a $k \in \mathbb{N}$ such that*

$$\lim_{\Delta\to\infty} \inf_{n\geq k} \int_{\mathbb{S}} \hat{\varphi}^{(n)}(s)\chi_{\{\hat{\varphi}^{(n)}(s)<-\Delta\}}dP^{(n)}(s) = 0. \tag{1}$$

Let $G_1 \subset \mathbb{H}_1$ and $G_2 \subset \mathbb{H}_2$ be given.

Theorem 3. *Assume that for all $\overline{s} \in G_1$, $\overline{x}^{(0)} \in G_2$, $x^{(0)} \in D^{(0)}(\overline{s},\overline{x}^{(0)})$ and all sequences $(\overline{x}^{(n)}, x^{(n)})_{n\in N} \to (\overline{x}^{(0)}, x^{(0)})$ the following assumptions are satisfied for $i_n = n$ and $i_n = 0$:*

(i) $P^{(i_n)}_{\overline{s},\overline{x}^{(n)},x^{(n)}} \xrightarrow{w} P^{(0)}_{\overline{s},\overline{x}^{(0)},x^{(0)}}$,

(ii) $\liminf_{n\in\mathbb{N}} \varphi^{(i_n)}(\overline{s}, s^{(n)}, \overline{x}^{(n)}, x^{(n)}) \geq \varphi^{(0)}(\overline{s}, s, \overline{x}^{(0)}, x^{(0)})$ *for* $P^{(0)}_{\overline{s},\overline{x}^{(0)},x^{(0)}}$*-almost all s and all sequences $(s^{(n)})_{n\in N} \to s$,*

(iii) the functions $\varphi^{(n)}(\overline{s}, \cdot, \overline{x}^{(n)}, x^{(n)})$ are $P^{(n)}_{\overline{s},\overline{x}^{(n)},x^{(n)}}$-integrable for all $n \in \mathbb{N}_0$ and the family $\{(\varphi^{(i_n)}(\overline{s}, \cdot, \overline{x}^{(n)}, x^{(n)}),\ P^{(i_n)}_{\overline{s},\overline{x}^{(n)},x^{(n)}}), n \in \mathbb{N}\}$ is lower equi-integrable.

Then, for all $\overline{s} \in G_1$*, the function* $f^{(0)}(\overline{s},\cdot,\cdot)$ *is l.s.c. on* $C(G_1,G_2)$ *and* $f^{(n)} \xrightarrow[C(G_1,G_2)|G_1]{l} f^{(0)}$.

Proof. We follow the proof of Theorem 3.1 in [9]. Although Theorem 3.1 is formulated for functions which are defined on $\mathbb{R}^p \times \mathbb{R}^m$ only, it holds for functions which are defined on cross-products of metric spaces. Let $P_n := P^{(i_n)}_{\overline{s},\overline{x}^{(n)},x^{(n)}}$; $P_0 := P^{(0)}_{\overline{s},\overline{x}^{(0)},x^{(0)}}$; $\xi_n := 0$ and $\varphi_n(\overline{x},x,s) := \varphi^{(i_n)}(\overline{s},s,\overline{x},x)$. Then application of Theorem 3.1 to the case $i_n = 0$ yields the lower semicontinuity result for $f^{(0)}$, and application to $i_n = n$ the assertion concerning $f^{(n)}$. □

The pointwise approach can also be applied to parameter-dependent probability measures. The following result is in the spirit of Theorem 3.2 (i) in [9]. We need the following auxiliary quantities: for $\overline{s} \in G_1 \subset \mathbb{H}_1$, a family $\{(\overline{x}^{(n)},x^{(n)}), n \in N_0\}$, $\varepsilon > 0$ and $i_n = 0$ and $i_n = n$, respectively, we define

$$Z^{(i_n)}_{\varepsilon}(\overline{s},\overline{x}^{(0)},x^{(0)},\overline{x}^{(n)},x^{(n)}) := \int_{\mathbb{S}} \inf_{(\overline{y},y)\in U_\varepsilon(\overline{x}^{(0)},x^{(0)})} \varphi^{(i_n)}(\overline{s},s,\overline{y},y)\,dP^{(i_n)}_{\overline{s},\overline{x}^{(n)},x^{(n)}}(s)$$

where $U_\varepsilon(\overline{x}^{(0)},x^{(0)})$ denotes a closed ball of radius ε and center $(\overline{x}^{(0)},x^{(0)})$.

Theorem 4. *Let, for given sets* $G_1 \subset \mathbb{H}_1$ *and* $G_2 \subset \mathbb{H}_2$*, the following assumptions be satisfied for each* $\overline{s} \in G_1$*, each* $(\overline{x}^{(0)},x^{(0)}) \in C(G_1,G_2)$*, all sequences* $(\overline{x}^{(n)},x^{(n)})_{n\in N} \to (\overline{x}^{(0)},x^{(0)})$*,* $i_n = n$ *and* $i_n = 0$*:*

(i) $\varphi^{(0)}(\overline{s},s,\cdot,\cdot)$ *is l.s.c. at* $(\overline{x}^{(0)},x^{(0)})$ *for* $P^{(0)}_{\overline{s},\overline{x}^{(0)},x^{(0)}}$*-almost all* s*.*

(ii) There is an $\bar{\varepsilon} > 0$ *such that* $Z^{(0)}_{\bar{\varepsilon}}(\overline{s},\overline{x}^{(0)},x^{(0)},\overline{x}^{(0)},x^{(0)}) > -\infty$ *and* $Z^{(i_n)}_{\varepsilon}(\overline{s},\overline{x}^{(0)},x^{(0)},\overline{x}^{(n)},x^{(n)})$*,* $Z^{(0)}_{\varepsilon}(\overline{s},\overline{x}^{(0)},x^{(0)},\overline{x}^{(0)},x^{(0)})$ *exist for each* $0 < \varepsilon < \bar{\varepsilon}$ *and each* $n \in N$*.*

(iii) $\forall \varepsilon \in (0,\bar{\varepsilon})$*,*

$$\liminf_{n\to\infty} Z^{(i_n)}_{\varepsilon}(\overline{s},\overline{x}^{(0)},x^{(0)},\overline{x}^{(n)},x^{(n)}) \geq Z^{(0)}_{\varepsilon}(\overline{s},\overline{x}^{(0)},x^{(0)},\overline{x}^{(0)},x^{(0)}).$$

Then, for all $\overline{s} \in G_1$*, the function* $f^{(0)}(\overline{s},\cdot,\cdot)$ *is l.s.c. on* $C(G_1,G_2)$ *and* $f^{(n)} \xrightarrow[C(G_1,G_2)|G_1]{l} f^{(0)}$.

Proof. Let $\overline{s} \in G_1$ and $(\overline{x}^{(0)},x^{(0)}) \in C(G_1,G_2)$ be fixed. According to the monotone convergence lemma we have

$$\sup_{\varepsilon>0} Z^{(0)}_{\varepsilon}(\overline{s},\overline{x}^{(0)},x^{(0)},\overline{x}^{(0)},x^{(0)}) = f^{(0)}(\overline{s},\overline{x}^{(0)},x^{(0)}).$$

Furthermore, for each $0 < \varepsilon < \bar{\varepsilon}$ and each $(\overline{x}^{(n)},x^{(n)})_{n\in N} \to (\overline{x}^{(0)},x^{(0)})$, the relation

$$\liminf_{n\to\infty} f^{(i_n)}(\overline{s},\overline{x}^{(n)},x^{(n)})$$

$$\geq \liminf_{n\to\infty} \int \inf_{(\bar{y},y)\in U_\varepsilon(\overline{x}^{(0)},x^{(0)})} \varphi^{(i_n)}(\overline{s},s,\bar{y},y)dP^{(i_n)}_{\overline{s},\overline{x}^{(n)},x^{(n)}}(s)$$
$$\geq Z^{(0)}_\varepsilon(\overline{s},\overline{x}^{(0)},x^{(0)},\overline{x}^{(0)},x^{(0)})$$

holds. □

Assumption (iii) can be supplemented by several sufficient conditions. If the probability measure does not depend on the decision, considerations in [9] can be employed. In the general case, e.g. one of the following approaches may be used: if there exists a dominating measure for all probability measures, (iii) is implied by suitable semicontinuity assumptions (with respect to the actions) of the Radon-Nikodym-derivatives. Furthermore, laws of large numbers for triangular arrays are often helpful. Eventually, there is the possibility to proceed via weak convergence of probability measures. Then, however, semicontinuity with respect to the states is needed.

The above theorems generalize Langen's Theorems 3.3 and 3.5. The boundedness condition is weakened considerably. Moreover, the semicontinuity assumption with respect to the states can either be omitted or at least restricted to almost all state histories. Thus probabilities among the objective and/or constraint functions can be taken into account. For the treatment of probabilities see e.g. [9].

5 Stability of Multistage Problems

We come back to the m-stage problem. We can combine Theorem 1 or Theorem 2 with Theorem 3 or Theorem 4 in order to derive stability statements for the multistage case. We will, for example, demonstrate how Theorem 4 together with Theorem 1 can be employed (making use of Theorem 2 the continuity and approximation assumptions with respect to foregoing actions could be further weakened).

In order to make clear in what points semicontinuous behavior with respect to the states is really needed, we introduce the following sets:

Let Θ^* be the set of all optimal strategies $\vartheta^* = (\delta^*_k)_{k\in N_m}$ for the original problem with $\delta^*_k(\overline{s}_k,\overline{x}_{k-1}) = x^*_k(\overline{s}_k,\overline{x}_{k-1})$. Each ϑ^* induces a probability measure P_{ϑ^*} on Σ.

Consider a standard Borel space $H_1(\Theta^*) \subset \mathbb{S}^{m+1}$ with $P_{\vartheta^*}(H_1(\Theta^*)) = 1$ $\forall\, \vartheta^* \in \Theta^*$, and define

$$\begin{aligned} H_{1,k} &:= \{\overline{s}_k : (\overline{s}_k, s_{k+1},\ldots,s_{m+1}) \in H_1(\Theta^*)\},\ k=1,\ldots,m,\\ D_1 &:= \{x_1 \in D^{(0)}_1(\overline{s}_1) : \overline{s}_1 \in H_{1,1}\},\\ D_k &:= \{(\overline{x}_{k-1},x_k) : \overline{x}_{k-1} \in D_{k-1}, x_k \in D^{(0)}_k(\overline{s}_k,\overline{x}_{k-1}), \overline{s}_k \in H_{1,k}\},\\ &\quad\ k=2,\ldots,m. \end{aligned}$$

Eventually, let, for $k \in N_m$ and $i_n = 0$ and $i_n = n$, respectively,

$$Z_{k,\varepsilon}^{(i_n)}(\overline{s}_k, \overline{x}_{k-1}^{(0)}, x_k^{(0)}, \overline{x}_{k-1}^{(n)}, x_k^{(n)}) := \int_{\mathbb{S}} \sup_{(\overline{y},y)\in U_\varepsilon(\overline{x}_{k-1}^{(0)}, x_k^{(0)})} |\varphi_k^{(i_n)}(\overline{s}_k, s, \overline{y}, y)| dP^{(i_n)}_{k|\overline{s}_k, \overline{x}_{k-1}^{(n)}, x_k^{(n)}}(s).$$

Theorem 5. *Let the following assumptions be satisfied for each* $k \in N_m$, *all* $\overline{s}_k \in H_{1,k}$, *all* $\overline{x}_k^{(0)} \in D_k$, *all sequences* $(\overline{x}_k^{(n)})_{n\in N} \to \overline{x}_k^{(0)}$, $i_n = n$ *and* $i_n = 0$:

(i) $\lim_{n\to\infty} |c_k^{(i_n)}(\overline{s}_k, s^{(0)}, \overline{x}_k^{(n)}) - c_k^{(0)}(\overline{s}_k, s^{(0)}, \overline{x}_k^{(0)})| = 0$ *for* $P_{k+1|\overline{s}_k, \overline{x}_k^{(0)}}$*-almost all* $s^{(0)}$.

(ii) $D_k^{(n)}(\overline{s}_k, \overline{x}_{k-1}^{(n)}) \xrightarrow{K} D_k^{(0)}(\overline{s}_k, \overline{x}_{k-1}^{(0)})$.

(iii) $\exists$ *compact* K $\exists n_0$ $\forall n \geq n_0$ *such that* $D_k^{(n)}(\overline{s}_k, \overline{x}_{k-1}^{(n)}) \subset K$.

(iv) There is an $\tilde{\varepsilon} > 0$ *such that* $Z_{k,\tilde{\varepsilon}}^{(0)}(\overline{s}_k, \overline{x}_{k-1}^{(0)}, x_k^{(0)}, \overline{x}_{k-1}^{(0)}, x_k^{(0)}) < \infty$ *and* $Z_{k,\varepsilon}^{(n)}(\overline{s}_k, \overline{x}_{k-1}^{(0)}, x_k^{(0)}, \overline{x}_{k-1}^{(n)}, x_k^{(n)})$ *exist for each* $\varepsilon \in (0, \tilde{\varepsilon})$ *and each* $n \in N$ *and* $\forall\, 0 < \varepsilon < \tilde{\varepsilon}$,

$$\lim_{n\to\infty} Z_{k,\varepsilon}^{(i_n)}(\overline{s}_k, \overline{x}_{k-1}^{(0)}, x_k^{(0)}, \overline{x}_{k-1}^{(n)}, x_k^{(n)}) = Z_{k,\varepsilon}^{(0)}(\overline{s}_k, \overline{x}_{k-1}^{(0)}, x_k^{(0)}, \overline{x}_{k-1}^{(0)}, x_k^{(0)}).$$

Then, for $k = 2, \ldots, m$, *the function* $\Phi_k^{(0)}(\overline{s}_k, \cdot)$ *is continuous on* D_{k-1}, $\Phi_k^{(n)} \xrightarrow[D_{k-1}|H_{1,k}]{c} \Phi_k^{(0)}$, $W_k^{(n)} \xrightarrow[D_{k-1}|H_{1,k}]{K-i} W_k^{(0)}$, $\lim_{n\to\infty} \Phi_1^{(n)}(s) = \Phi_1^{(0)}(s)$, *and* $W_1^{(n)}(s) \xrightarrow{K-i} W_1^{(0)}(s)$ *for* P_1*-almost all* s.

Proof. We proceed by backward induction. Because of $\Phi_{m+1}^{(i_n)}(\overline{s}_{m+1}, \overline{x}_m) = 0$ for all $(\overline{s}_{m+1}, \overline{x}_m) \in \mathbb{H}_{1,k+1} \times \mathbb{H}_{2,k}$ and all $n \in N$, we have $\varphi_m^{(i_n)} = c_m^{(i_n)}$. Applying Theorem 4 to $c_m^{(i_n)}$ and $-c_m^{(i_n)}$, $G_1 = H_{1,m}$, $G_2 = D_{m-1}$ and $C(G_1, G_2) = D_m$, we obtain the continuity of $f_m^{(0)}(\overline{s}_m, \cdot)$ on D_m and $f_m^{(i_n)} \xrightarrow[D_m|H_{1,m}]{c} f_m^{(0)}$. This, together with the assumptions (ii) and (iii) gives, by Theorem 1, in case $i_n = 0$ the continuity of $\Phi_m^{(0)}(\overline{s}_m, \cdot)$ on D_{m-1} and, for $i_n = n$, $\Phi_m^{(n)} \xrightarrow[D_{m-1}|H_{1,m}]{c} \Phi_m^{(0)}$, and $W_m^{(n)} \xrightarrow[D_{m-1}|H_{1,m}]{K-i} W_m^{(0)}$. For the stage k we can proceed in the same way. The continuity assumptions for $\varphi_k^{(0)}$ are satisfied because of (i) and the continuity of $\Phi_{k+1}^{(0)}$. Integrability is assumed in (iv). □

Acknowledgements. The authors are grateful to the referees for helpful remarks and suggestions.

References

1. B. Bank, J. Guddat, D. Klatte, B. Kummer, and K. Tammer. Nonlinear Parametric Optimization. Akademie-Verlag Berlin, 1982.
2. D.P. Bertsekas and S.E. Shreve. Stochastic Optimal Control: The Discrete Time Case. Academic Press, New York, 1978.

3. J. Dupačová and K. Sladký. Comparison of Multistage Stochastic Programs with Recourse and Stochastic Dynamic Programs with Discrete Time. Z. Angew. Math. Mech. 82: 11-12: 753-765, 2002.
4. O. Fiedler and W. Römisch. Stability of multistage stochastic programming. Ann. Oper. Res. 56: 79-93, 1995.
5. K. Hinderer. Foundations of Non-Stationary Dynamic Programming with Discrete Time Parameter. Lecture Notes in Oper. Res. and Math. Syst., 33, 1970.
6. V. Kaňková. A Remark on Analysis of Multistage Stochastic Programs: Markov Dependence. Z. Angew. Math. Mech. 82, 11-12: 781-793, 2002.
7. L.A. Korf and R.J.-B. Wets. Random lsc functions: an ergodic theorem. Math. Oper. Res. 26: 421-445, 2001.
8. S. Vogel and P. Lachout. On continuous convergence and epi-convergence of random functions - Theory and relations. Kybernetika 39: 75-98, 2003.
9. S. Vogel and P. Lachout. On continuous convergence and epi-convergence of random functions - Sufficient conditions and applications. Kybernetika 39: 99-118, 2003.
10. H.-J. Langen. Convergence of Dynamic Programming Models. Math. Oper. Res. 6: 493-512, 1981.
11. A. Mänz. Stabilität mehrstufiger stochastischer Optimierungsprobleme. Diploma Thesis, TU Ilmenau, 2003.
12. P.H. Müller and V. Nollau (ed.). Steuerung stochastischer Prozesse. Akademie-Verlag, Berlin, 1984.
13. S.M. Robinson. Local epi-continuity and local optimization. Math. Programming 37: 208 - 222, 1987.
14. R.T. Rockafeller and R. J.-B. Wets. Variational Analysis. Springer, 1998.
15. M. Schäl. Conditions for Optimality in Dynamic Programming and for the Limit of n-Stage Optimal Policies to Be Optimal. Z. Wahrscheinlichkeitstheorie verw. Gebiete 32: 179-196, 1975.
16. S. Vogel. Stochastische Stabilitätskonzepte. Habilitation, TU Ilmenau, 1991.
17. S. Vogel. A stochastic approach to stability in stochastic programming. J. Comput. and Appl. Mathematics, Series Appl. Analysis and Stochastics 56: 65-96, 1994.
18. S. Vogel.On stability in stochastic programming - Sufficient conditions for continuous convergence and epi-convergence. Preprint TU Ilmenau, 1994.
19. J. Wang. Stability of multistage stochastic programming. Ann. Oper. Res. 56: 313-322, 1995.