FULL LENGTH PAPER

Approximations of Nash equilibria

Gül Gürkan · Jong-Shi Pang

Abstract   Inspired by previous works on approximations of optimization problems and recent papers on the approximation of Walrasian and Nash equilibria and on stochastic variational inequalities, the present paper investigates the approximation of Nash equilibria and clarifies the conditions required for the convergence of the approximate equilibria via a direct approach, a variational approach, and an optimization approach. Besides directly addressing the issue of convergence of Nash equilibria via approximation, our investigation leads to a deeper understanding of various notions of functional convergence and their interconnections; more importantly, the investigation yields improved conditions for convergence of the approximate Nash equilibria via the variational approach. An illustrative application of our results to the

G. Gürkan (✉)
Department of Econometrics and Operations Research,
Tilburg University, Tilburg, 5037 AB, The Netherlands
e-mail: ggurkan@uvt.nl

J.-S. Pang
Department of Mathematical Sciences,
Rensselaer Polytechnic Institute,
Troy, NY 12180-3590, USA
e-mail: pangj@rpi.edu

approximation of a Nash equilibrium in a competitive capacity expansion model under uncertainty is presented.

Keywords   Variational inequalities · Stochastic · Equilibrium models · Epiconvergence · Competitive capacity expansion under uncertainty

1 Introduction

Since the classic paper by Wets [34] on functional epiconvergence and the elegant review of Kall [17] on related convergence concepts, the subject of approximation of optimization problems has been extensively investigated and fruitfully developed along various directions. Today, the subject is subsumed under the broad context of variational analysis for which a comprehensive study is excellently documented in the treatise by Rockafellar and Wets [29]. In general, the issue of approximation of optimization problems arises in a number of ways, most prominently, (a) in dealing with stochastic optimization problems which typically involve some multi-dimensional integral functions that need to be approximated, (b) in the convergence analysis of such iterative algorithms as barrier and penalty methods for solving constrained optimization problems as unconstrained problems, and (c) in smoothing schemes whereby nonsmooth functions are approximated by sequences of smooth functions. We refer to the handbook edited by Ruszczyński and Shapiro [31], particularly Chapts. 6 and 7, and to [26,28] on the subject of approximation of stochastic programs, to the classic paper by Attouch and Wets [2] and the recent paper by Levy [22] on the convergence analysis of iterative methods for (deterministic) optimization problems using variational analytic tools, and to Sect. 11.8 in [10] for a systematic view of smoothing methods for solving nonsmooth equations with application to finite-dimensional variational inequalities and complementarity problems.

In contrast to optimization problems, the approximation of Nash equilibria is a subject that has so far received only minimal attention in the literature of mathematical programming. In essence, all the reasons, practical and theoretical, for studying the approximation of optimization problems persist in the computation of Nash equilibria. Among these, the challenge to solve realistic Nash equilibrium problems with uncertainty and extended Nash problems with coupled constraints and nonsmooth objective functions provides a prime motivation to investigate the approximation of Nash equilibria. Unlike individual optimization problems, Nash equilibrium problems involve several optimization problems that are linked together by the well-known Nash equilibrium concept [24]. As such, existing approximation theory for optimization problems is not directly applicable to Nash problems in their primitive formulations. A systematic investigation of such a theory for the latter problems is the primary goal of this paper.

Two recent papers, [13] and [16], provide much of the inspiration for our work. The former paper [13] analyzes the convergence of the sample-path approach to solv-

ing stochastic variational inequalities. Yielding a first convergence theory for the approximation of Nash equilibria, this variational approach requires the differentiability of the individual players' objective functions. In contrast, the latter paper [16] employs a minsup formulation of the Nash equilibrium problem together with the concept of "lopsided convergence" of bivariate functions [3,4] to study the convergence of Nash equilibria, without requiring the players' objectives to be differentiable. For an application of the latter theory to the related Walrasian equilibrium problem, see [15]. The two approaches differ not only in the differentiability requirement, but also in the "continuous convergence" of the players' approximate objective functions: the former imposed on the partial gradients of these functions relative to the individual players' variables, and the latter on the objective functions themselves.

The main contributions of our work are as follows. First, we introduce a componentwise epiconvergence notion for multi-component functions, which generalizes that of "epiconvergence" of uni-component functions of one (multi-dimensional) variable, the latter being a fundamental concept in variational analysis [29]; we show that the new notion is sufficient for the convergence of the approximate Nash equilibria. We next show how the multi-epiconvergence property can be replaced by some variational conditions on the directional derivatives and/or the subgradients when the players' objective functions are convex for fixed rival players' strategies. In the process of our analysis, we will see that the continuous convergence properties employed in the references [13,16] can be significantly weakened; furthermore, our derived results will clarify the connections between these asymptotic properties, which are imposed on the functions and their (partial) gradients, respectively. We also clarify the connections to the lopsided convergence approach of Jofré and Wets [16] and present a competitive capacity expansion model under uncertainty to illustrate the need for an approximation theory of Nash equilibria.

2 Convergence of Nash equilibria: the direct approach

Consider a multi-player Nash equilibrium problem where there are N players, each labelled by an integer $\nu = 1, \ldots, N$. Player ν has a cost function $\theta_\nu(x)$ that depends on all players' strategies $x = (x^{\nu'})_{\nu'=1}^N$, where each component $x^{\nu'}$ represents a strategy of player ν', which is required to belong to the closed convex set $X^{\nu'} \subseteq \Re^{n_{\nu'}}$, for some positive integer $n_{\nu'}$. For every fixed but arbitrary tuple of rival players' strategies $x^{-\nu} \in X^{-\nu} \equiv \prod_{\nu' \neq \nu} X^{\nu'}$, player ν solves the following optimization problem in her own variable x^ν:

$$\text{minimize} \quad \theta_\nu(x) \quad \text{subject to} \quad x^\nu \in X^\nu.$$

A tuple of strategies $\widehat{x} \equiv (\widehat{x}^\nu)_{\nu=1}^N \in X \equiv \prod_{\nu=1}^N X^\nu$ is a *Nash equilibrium* if for all $\nu = 1, \ldots, N$,

$$\theta_\nu(\widehat{x}) \leq \theta_\nu(x^\nu, \widehat{x}^{-\nu}) \quad \forall x^\nu \in X^\nu.$$

We are interested in a deeper understanding of approximation of a Nash equilibrium via several of its equivalent formulations. At the most basic level, we distinguish between the cases depending on whether each player' objective function $\theta_\nu(\cdot, x^{-\nu})$ is convex and/or differentiable for given $x^{-\nu}$ and on the asymptotic properties of the approximating functions. It turns out that there are significant differences between the conditions needed in each case and clarifying the connections between these conditions is part of our goal.

We begin with the above definition of a Nash equilibrium; no convexity or differentiability is assumed on the cost functions. In this approach, we approximate the cost functions directly and compute a sequence of Nash equilibria based on such an approximation. The key property required for the convergence of the so-generated sequence is the following concept of functional convergence for multi-component functions, which extends the concept of epiconvergence [29] for uni-component functions. Letting $n \equiv \sum_{\nu=1}^N n_\nu$, we introduce the following basic definition.

Definition A family of multi-component functions

$$\left\{ \theta_{\nu,1}, \theta_{\nu,2}, \dots \right\}_{\nu=1}^N, \tag{1}$$

where each $\theta_{\nu,k} : \Re^n \to \Re$, *multi-epiconverges* to the functions $\{\theta_\nu\}_{\nu=1}^N$ on the set X if the following two conditions hold for every $\nu = 1, \dots, N$ and every $x \in X$:

(Ma) for every sequence $\{x^{-\nu,k}\} \subset X^{-\nu}$ converging to $x^{-\nu}$, a sequence $\{x^{\nu,k}\} \subset X^\nu$ converging to x^ν exists such that

$$\limsup_{k \to \infty} \theta_{\nu,k}(x^k) \leq \theta_\nu(x), \quad \text{where } x^k \equiv (x^{\nu',k})_{\nu'=1}^N,$$

(Mb) for every sequence $\{x^k\} \subset X$ converging to x,

$$\liminf_{k \to \infty} \theta_{\nu,k}(x^k) \geq \theta_\nu(x).$$

(Having introduced the above definition, we should immediately note that in Definition 3.3.1 of [8] and in Definition 4 of [9], Escobar introduced a convergence concept for discontinuous games, i.e., games where the players' objective functions may fail to be continuous. Specialized to the continuous case, Escobar's definition reduces to multi-epiconvergence defined above. Escobar also established a convergence result, Theorem 6 in [9], that coincides (for continuous games) with our Theorem 1 below. Notwithstanding this overlap, we developed our theory independently of that of the references, which primarily deal with discontinuous games. Indeed, we are grateful to Alejandro Jofré who pointed us to Escobar's work after he has seen a draft of our paper.)

The multi-epiconvergence concept differs from the lopsided convergence concept for bivariate functions introduced by Attouch and Wets [3,4]. The two concepts are introduced for different purposes: multi-epiconvergence for the direct approach of approximating Nash equilibria (see below), and lopsided convergence for approximating saddle points which are equilibria to a two-person zero-sum Nash game. Yet,

there are some connections between the two convergence concepts in the context of the Nikaido–Isoda function; see Sect. 4.1 for details. To prepare for our main result, Theorem 1 below, we note that, in general, the family (1) multi-epiconverges to the functions $\{\theta_v\}_{v=1}^{N}$ on the set X if and only if for every $v = 1, \ldots, N$ and every sequence $\{x^{-v,k}\} \subset X^{-v}$ converging to some $x^{-v,\infty} \in X^{-v}$, the sequence of uni-component functions $\{\psi_{v,k}\}$, where

$$\psi_{v,k}(x^v) \equiv \theta_{v,k}(x^v, x^{-v,k}), \quad x^v \in X^v \tag{2}$$

epiconverges to the uni-component function

$$\psi_{v,\infty}(x^v) \equiv \theta_v(x^v, x^{-v,\infty}), \quad x^v \in X^v \tag{3}$$

on the set X^v. This observation connects the multi-epiconvergence concept for multi-component functions to the well-known epiconvergence concept for uni-component functions in several ways. One, due to well-known connections between functional epiconvergence and graphical convergence of subdifferentials (see e.g. Attouch [1] and [29, Theorem 12.35]), it is possible to restate the multi-epiconvergence of the family (1) in terms of the set convergence of the graphs of subdifferentials $\partial \psi_{v,k}$, under the convexity of the functions $\psi_{v,k}$. (For extensions of Attouch's theorem to nonconvex functions, see [23,27].) Two, multi-epiconvergence implies the following *diagonal epiconvergence*: namely, for every $v = 1, \ldots, N$ and every fixed $x^{-v} \in X^{-v}$, the sequence of uni-component functions $\{\theta_{v,k}(\cdot, x^{-v})\}$ epiconverges to the uni-component function $\theta_v(\cdot, x^{-v})$. Third, via the epiconvergence of the sequence $\{\psi_{v,k}\}$, the convergence of Nash equilibria can be established very easily.

Theorem 1 *Let each X^v be a closed subset of $\Re^{n_v}$. Suppose that the family (1) multi-epiconverges to the functions $\{\theta_v\}_{v=1}^{N}$ on the set X. If the sequence $\{x^k \equiv (x^{v,k})_{v=1}^{N}\}$, where each x^k is a Nash equilibrium tuple of the pairs $(\theta_{v,k}, X^v)_{v=1}^{N}$, converges to $x^\infty \equiv (x^{v,\infty})_{v=1}^{N}$, then x^∞ is a Nash equilibrium tuple of the pairs $(\theta_v, X^v)_{v=1}^{N}$.*

Proof Since each X^v is a closed set, we have $x^\infty \in X$. By definition,

$$x^{v,k} \in \arg\min \{\theta_{v,k}(x^v, x^{-v,k}) : x^v \in X^v\} \quad \forall k.$$

Since the sequence of functions $\{\theta_{v,k}(\cdot, x^{-v,k})\}$ epiconverges to the function $\theta_v(\cdot, x^{-v,\infty})$ on X^v, it follows from the theory of epiconvergence that

$$x^{v,\infty} \in \arg\min \{\theta_v(x^v, x^{-v,\infty}) : x^v \in X^v\}.$$

Since this holds for all $v = 1, \ldots, N$, the desired Nash property of the tuple x^∞ follows readily.   □

Theorem 1 assumes no convexity on the players' objective functions. We can obtain a version of the result when the approximating functions $\theta_{v,k}(\cdot, x^{-v,k})$ are convex. This "convex" version is based on a known result in convex analysis which states that if the sequence of convex functions $\{f_k\}$ converges pointwise to the finite function f (i.e.,

$\lim_{k\to\infty} f_k(z) = f(z)$ for every z), then f must itself be convex and $\{f_k\}$ continuously converges to f; i.e., $\lim_{k\to\infty} f_k(z^k) = f(z^\infty)$ for every sequence $\{z^k\}$ converging to z^∞.

Corollary 1 *The conclusion of Theorem 1 holds under the conditions*:

(cc) *for every $\nu = 1, \ldots, N$ and all $k = 1, 2, \ldots$, the functions $\theta_{\nu,k}(\cdot, x^{-\nu})$ are convex for all fixed $x^{-\nu} \in X^{-\nu}$, and*

(Pc) *for every $\nu = 1, \ldots, N$ and every sequence $\{x^{-\nu,k}\} \subset X^{-\nu}$ converging to some $x^{-\nu,\infty} \in X^{-\nu}$, the sequence of uni-components functions $\{\psi_{\nu,k}\}$ defined by (2) converges pointwise to the uni-component function $\psi_{\nu,\infty}$ defined by (3) on X^ν.*

$\hfill\square$

It is worth pointing out that the condition (Pc) is somewhat stronger than the pointwise convergence of the sequence of functions $\{\theta_{\nu,k}\}$ to the function θ_ν on X. The latter condition asserts that $\lim_{k\to\infty} \theta_{\nu,k}(x) = \theta_\nu(x)$ for every $x \in X$. In contrast, the former condition (Pc) stipulates that

$$\lim_{k\to\infty} \theta_{\nu,k}(x^\nu, x^{-\nu,k}) = \theta_\nu(x^\nu, x^{-\nu})$$

for every $x^\nu \in X^\nu$ and every sequence $\{x^{-\nu,k}\} \subset X^{-\nu}$ converging to $x^{-\nu} \in X^{-\nu}$. For later comparison, it is useful also to formally define the "diagonal" version of the conditions (Ma) and (Mb), respectively, which we have informally mentioned above.

Definition The family (1) of multi-component functions *diagonally epiconverges* to the functions $\{\theta_\nu\}_{\nu=1}^N$ on the set X if for every $\nu = 1, \ldots, N$ and every $x^{-\nu} \in X^{-\nu}$, the sequence of functions $\{\theta_{\nu,k}(\cdot, x^{-\nu})\}$ epiconverges to the function $\theta_\nu(\cdot, x^{-\nu})$ on X^ν; i.e., the following two conditions hold for every $\nu = 1, \ldots, N$ and every $x \in X$:

(Da)   a sequence $\{x^{\nu,k}\} \subset X^\nu$ converging to x^ν exists such that

$$\limsup_{k\to\infty} \theta_{\nu,k}(x^{\nu,k}, x^{-\nu}) \leq \theta_\nu(x),$$

(Db)   for every sequence $\{x^{\nu,k}\} \subset X^\nu$ converging to x^ν,

$$\liminf_{k\to\infty} \theta_{\nu,k}(x^{\nu,k}, x^{-\nu}) \geq \theta_\nu(x).$$

Clearly, (Ma) $\Rightarrow$ (Da) and (Mb) $\Rightarrow$ (Db). The reverse implications hold under an "upper equicontinuity" property defined below. Notice that this definition involves only the sequence of approximating functions and does not involve the limit function.

Definition The multi-component functions (1) are *off-diagonally equicontinuous* (odec) on X if for every sequence $\{x^k \equiv (x^{\nu,k})_{\nu=1}^N\} \subset X$ converging to some $x^\infty \in X$, it holds that, for all $\nu = 1, \ldots, N$,

$$\liminf_{k\to\infty} \theta_{\nu,k}(x^{\nu,k}, x^{-\nu,\infty}) \leq \liminf_{k\to\infty} \theta_{\nu,k}(x^k)$$

$$\leq \limsup_{k\to\infty} \theta_{\nu,k}(x^k) \leq \limsup_{k\to\infty} \theta_{\nu,k}(x^{\nu,k}, x^{-\nu,\infty}). \quad (4)$$

Clearly, if for every $\nu = 1, \ldots, N$, the sequence of functions $\{\theta_{\nu,k}\}$ converges continuously (Cc) to θ_ν on X, then for every sequence $\{x^k\} \subset X$ converging to $x^\infty \in X$, equalities hold throughout (4); moreover, all the limits there are equal to $\theta_\nu(x^\infty)$. If the functions $\theta_{\nu,k}$ are differentiable and for all bounded subsets S of X,

$$\limsup_{k \to \infty} \max_{\nu' \neq \nu} \sup_{x \in S} \| \nabla_{\nu'} \theta_{\nu,k}(x) \| < \infty \quad \forall \nu = 1, \ldots, N,$$

then (4) also holds. This can easily be seen from the following identity, which is a consequence of the mean-value theorem for multivariate functions,

$$\theta_{\nu,k}(x^k) = \theta_{\nu,k}(x^{\nu,k}, x^{-\nu,\infty})$$

$$+ \sum_{\nu' \neq \nu} \int_0^1 \nabla_{\nu'} \theta_{\nu,k}(x^{\nu,k}, x^{-\nu,\infty} + \tau(x^{-\nu,k} - x^{-\nu,\infty}))^{\mathrm{T}} (x^{-\nu,k} - x^{-\nu,\infty}) \, d\tau.$$

Without requiring differentiability of the functions involved, the following result summarizes the connections between all the convergence concepts introduced so far.

Theorem 2 *Let each X^ν be a closed subset of $\Re^{n_\nu}$. Let the family of multi-component functions* (1) *and the players' objective functions $\{\theta_\nu\}_{\nu=1}^N$ be given. The following implications hold.*

(A)   (cc) + (Pc) $\Leftrightarrow$ (cc) + (Cc);
(B)   (Cc) $\Rightarrow$ (odec) + (Da) + (Db);
(C)   (odec) + (Da) $\Rightarrow$ (Ma), and (odec) + (Db) $\Rightarrow$ (Mb);
(D)   (Ma) + (Mb) $\Rightarrow$ convergence of Nash equilibria.

Proof It suffices to prove only the implications in (C). Assume (odec) and (Da). Let $x \in X$ and let $\{x^{-\nu,k}\} \subset X^{-\nu}$ be any sequence converging to $x^{-\nu} \in X^{-\nu}$. Let $\{x^{\nu,k}\} \subset X^\nu$ be a sequence converging to x^ν corresponding to $x^{-\nu} \in X^{-\nu}$ as stipulated by (Da). We have,

$$\limsup_{k \to \infty} \theta_{\nu,k}(x^k) \leq \limsup_{k \to \infty} \theta_{\nu,k}(x^{\nu,k}, x^{-\nu}) \leq \theta_\nu(x),$$

where the first inequality follows from (odec) and the second from (Da). Thus (Ma) holds. Similarly, for any $x \in X$ and any sequence $\{x^k\} \subset X$ converging to x, we have,

$$\liminf_{k \to \infty} \theta_{\nu,k}(x^k) \geq \liminf_{k \to \infty} \theta_{\nu,k}(x^{\nu,k}, x^{-\nu}) \geq \theta_\nu(x),$$

where the first inequality follows from (odec) and the second from (Db). Thus (Mb) holds.   □

3 The variational approach

A distinguished feature of the direct approach is that it does not require the differentiability of the functions involved; moreover, the approximation is applied to the

scalar-valued functions $\{\theta_\nu\}_{\nu=1}^N$ themselves. In contrast, the approximation in the variational approach described in this section requires the differentiability of these cost functions and is applied to their (partial) gradients, which are vector functions.

Specifically, assume throughout this subsection that the functions $\{\theta_\nu\}_{\nu=1}^N$ are once differentiable on an open set containing X. It follows that [10] under (cc), $\widehat{x} \equiv (\widehat{x}^\nu)_{\nu=1}^N \in X$ is a Nash equilibrium if and only if the following variational inequality (VI) holds:

$$\sum_{\nu=1}^N (x^\nu - \widehat{x}^\nu)^{\mathrm{T}} \nabla_{x^\nu} \theta_\nu(\widehat{x}) \geq 0 \quad \forall x \equiv (x^\nu)_{\nu=1}^N \in X.$$

In [13], a convergence result of Nash equilibria was established based on an approximation of the vector function $F : \Re^n \to \Re^n$, where $F(x) \equiv (\nabla_{x^\nu} \theta_\nu(x))_{\nu=1}^N$. We call F the *VI function* of the Nash game. Assuming continuous convergence to the latter function, the proof of following result is straightforward (see the proof of Theorem 4).

Theorem 3 *Let θ_ν be differentiable on open set containing the closed convex set $X^\nu \subseteq \Re^{n_\nu}$. Suppose that the sequence of vector functions $\{F^k\}$ converges continuously to the VI function F on X. If $\{x^k\} \subset X$ is a sequence of vectors converging to the vector x^∞ such that for every k,*

$$(x - x^k)^{\mathrm{T}} F^k(x^k) \geq 0 \quad \forall x \in X,$$

then x^∞ is a Nash equilibrium of the pairs $(\theta_\nu, X^\nu)_{\nu=1}^N$, provided that each function $\theta_\nu(\cdot, x^{-\nu,\infty})$ is convex. □

In general, the continuous convergence of $\{F^k\}$ to F tells us very little about the approximation of the functions $\{\theta_\nu\}_{\nu=1}^N$ themselves. A reason is that the VI function F contains only the "principal derivatives" $\nabla_{x^\nu} \theta_\nu$, which gives no information about the dependence of player ν's objective function θ_ν on the rival players' variables. At first thought, this seems to suggest an advantage in favor of the variational approach because it requires no such information. Nevertheless, one has to keep in mind that the convergence of $\{F^k\}$ to F pertains to vector functions; a natural question is therefore how to construct such a convergent sequence of approximate functions $\{F^k\}$ based only on the principal derivatives of $\{\theta_\nu\}$ and not using the full information of the latter scalar-valued functions themselves. The reference [13] did not address this question, neither will we. Instead, we will investigate a related issue that arises from the following consideration. Namely, suppose we approximate the cost functions $\{\theta_\nu\}_{\nu=1}^N$ directly, and not their derivatives, even if the latter are well behaved. The question is: how is the convergence of "functional approximation" related to the convergence of the gradients? Interestingly, we are not aware of a direct treatment of this issue, even for uni-component functions, although a referee has noted several related results (see Remark 2 below). A known result for convex functions is the following. If $\{f_k\}$ is a sequence of differentiable convex functions converging pointwise to the differentiable function f, then the sequence of gradients $\{\nabla f_k\}$ converges continuously to the

gradient ∇f. (The concept of "Jacobian consistency smoothing" in [10, Definition 11.8.3] is related to the issue, but does not yield a satisfactory answer to the question.)

Before discussing the connection between functional approximation and gradient approximation in the next subsection, we present a result that is in the spirit of Theorem 3 but does not require the functions to the differentiable. To motivate this extension, we note that when we take $F^k \equiv (\nabla_{x^\nu}\theta_{\nu,k})_{\nu=1}^N$ for each k in Theorem 3, the continuous convergence of $\{F^k\}$ to F means that for every sequence $\{x^k\} \subset X$ converging to some $x^\infty \in X$, we have

$$\lim_{k\to\infty} \nabla_{x^\nu}\theta_{\nu,k}(x^k) = \nabla_{x^\nu}\theta_\nu(x^\infty) \quad \forall \nu = 1, \ldots, N. \tag{5}$$

In essence, we will replace the gradients $\nabla_{x^\nu}\theta_{\nu,k}(x^k)$ and $\nabla_{x^\nu}\theta_\nu(x^\infty)$ by the respective subgradients when these gradients do not exist. In what follows, the notation $\partial_{x^\nu}\phi(x)$ refers to the subdifferential of the convex function $\phi(\cdot, x^{-\nu})$ evaluated at x^ν.

Theorem 4 *Let each X^ν be a closed convex subset of $\Re^{n_\nu}$. Suppose that for every $\nu = 1, \ldots, N$ and every $x^{-\nu} \in X^{-\nu}$, the functions $\theta_{\nu,k}(\cdot, x^{-\nu})$ and $\theta_\nu(\cdot, x^{-\nu})$ are all convex. Suppose further that for every $\nu = 1, \ldots, N$ and every sequence $\{x^k\} \subset X$ converging to some x^∞, every sequence of partial subgradients $\{a^{\nu,k}\}$, where $a^{\nu,k} \in \partial_{x^\nu}\theta_{\nu,k}(x^k)$ for all k, has an accumulation point $a^{\nu,\infty}$ belonging to $\partial_{x^\nu}\theta_\nu(x^\infty)$. Then the conclusion of Theorem 1 holds.*

Proof Suppose that the sequence $\{x^k \equiv (x^{\nu,k})_{\nu=1}^N\}$, where each x^k is a Nash equilibrium tuple of the pairs $(\theta_{\nu,k}, X^\nu)_{\nu=1}^N$, converges to $x^\infty \equiv (x^{\nu,\infty})_{\nu=1}^N$. It follows that for every $\nu = 1, \ldots, N$ and every k, there exists $a^{\nu,k} \in \partial_{x^\nu}\theta_{\nu,k}(x^{\nu,k})$ such that

$$(x^\nu - x^{\nu,k})^{\mathrm{T}}a^{\nu,k} \geq 0 \quad \forall x^\nu \in X^\nu.$$

By assumption, the sequence $\{a^{\nu,k}\}$ has an accumulation point $a^{\nu,\infty}$ belonging to $\partial_{x^\nu}\theta_\nu(x^\infty)$. Passing to the limit $k \to \infty$ in such a convergent subsequence, we deduce

$$(x^\nu - x^{\nu,\infty})^{\mathrm{T}}a^{\nu,\infty} \geq 0 \quad \forall x^\nu \in X^\nu,$$

which implies that $x^{\nu,\infty} \in \min_{x^\nu \in X^\nu} \theta_\nu(x^\infty)$. Since this holds for all $\nu = 1, \ldots, N$, the desired Nash property of x^∞ follows. $\qquad\square$

3.1 Functional versus gradient convergence

In this subsection, we deviate slightly from the main discussion of Nash equilibria and address the issue of functional approximation versus gradient approximation. We undertake this discussion because we believe that the issue is significant and worthy of a careful treatment. At the end of the subsection, we present alternative conditions for the convergence of approximate Nash equilibria via the variational approach; see Corollaries 2 and 3.

We begin by giving an example involving uni-component functions to show that the convexity assumption cannot be dropped from the implication: continuous convergence of $\{f_k\} \Rightarrow$ continuous convergence of $\{\nabla f_k\}$. Specifically, derived from a smoothing of the absolute-value function, the example below provides a sequence of continuously convergent functions to a nonconvex function such that a gradient sequence is unbounded and the limit gradient is discontinuous.

Example 1 Consider the function of one variable:

$$\psi(t) \equiv \begin{cases} t^2 \sin\left(\dfrac{1}{t}\right) & \text{if } t \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to see that this function is everywhere differentiable with

$$\psi'(t) = \begin{cases} 2t \sin\left(\dfrac{1}{t}\right) - \cos\left(\dfrac{1}{t}\right) & \text{if } t \neq 0 \\ 0 & \text{if } t = 0. \end{cases}$$

Note that the derivative $\psi'(t)$ is not continuous at $t = 0$. Consider the sequence of functions $\{\psi_k\}$, where for each $k > 0$,

$$\psi_k(t) \equiv t^2 \sin\left(\frac{1}{\sqrt{t^2 + 1/k^2}}\right).$$

It is easy to verify that $\{\psi_k\}$ converges continuously to ψ on the real line. Direct differentiation gives

$$\psi_k'(t) = 2t \sin\left(\frac{1}{\sqrt{t^2 + 1/k^2}}\right) - \frac{t^3}{(t^2 + 1/k^2)^{3/2}} \cos\left(\frac{1}{\sqrt{t^2 + 1/k^2}}\right).$$

We have

$$\psi_k'(1/k) = \frac{2}{k} \sin\left(\frac{k}{\sqrt{2}}\right) - \frac{1}{2^{3/2}} \cos\left(\frac{k}{\sqrt{2}}\right),$$

which has no limit as $k \uparrow \infty$, even though the sequence $\{1/k\}$ converges to zero. □

The converse implication: continuous convergence of $\{\nabla f_k\} \Rightarrow$ continuous convergence of $\{f_k\}$, is clearly not true in general. A trivial example is the sequence of constant functions $\{f_k(x) = k\}$ whose gradients are all equal to zero.

Remark 1 A comment by a referee raises the question of whether the non-implication: continuous convergence of $\{\psi_k\} \not\Rightarrow$ continuous convergence of $\{\nabla\psi_k\}$, might be due to the discontinuity of the limit gradient $\nabla\psi$ (in the example, it is at $t = 0$). Notwithstanding this question, the above example illustrates that the continuous convergence

of a sequence of functions is too weak to imply even the continuity of the gradient of the limit function. From this point of view, the sequence $\{\psi_k\}$ in the example is perhaps not needed because the implication in question is obviously not true by taking the sequence of functions to consist of one single differentiable, but not continuously differentiable, function. □

Assuming that all ψ_k and ψ are differentiable (and convex), Proposition 2 below gives a complete characterization for the continuous convergence of $\{\nabla\psi_k\}$ to $\nabla\psi$ in terms of the continuous convergence of $\{\psi_k\}$ to ψ, by identifying a key missing condition. The next result, which pertains to uni-component functions, shows that if the gradients of the approximating functions converge continuously to the gradient of a given function, then the approximating functions converge continuously to the given function if and only if the latter convergence occurs at a single point.

Proposition 1 *Let $\{\psi_k\}$ and ψ be differentiable scalar functions defined on an open set containing the convex set $X \subseteq \Re^n$. Suppose that $\{\nabla\psi_k\}$ converges continuously to $\nabla\psi$ on X. Then $\{\psi_k\}$ converges continuously to ψ on X if and only if a vector $z \in X$ exists such that*

$$\lim_{k\to\infty} \psi_k(z) = \psi(z).$$

Proof We first show that $\{\nabla\psi_k\}$ is uniformly bounded on compact subsets of X, i.e., for every compact $S \subseteq X$,

$$\sup_k \sup_{x\in S} \|\nabla\psi_k(x)\| < \infty. \tag{6}$$

Assume for contradiction that this is not true. Then there exists a compact set $S \subseteq X$ and a sequence $\{x^k\} \subset S$ such that

$$\lim_{k\to\infty} \|\nabla\psi_k(x^k)\| = \infty.$$

Since S is compact, there is a subsequence $\{x^k : k \in \kappa\}$ that converges to an element x^∞ in $S \subseteq X$, for some infinite index set κ. It follows that $\{\nabla\psi_k(x^k) : k \in \kappa\}$ converges to $\nabla\psi(x^\infty)$, which contradicts the above unbounded limit.

Clearly, it suffices to show the "if" assertion of the proposition. Let $\{x^k\} \subset X$ be an arbitrary sequence converging to some $x \in X$. By the integral form of the mean-value theorem, we can write

$$\psi_k(x^k) = \psi_k(z) + \int_0^1 \nabla\psi_k(z + t(x^k - z))^{\mathrm{T}}(x^k - z)\, dt \quad \text{and}$$

$$\psi(x) = \psi(z) + \int_0^1 \nabla\psi(z + t(x - z))^{\mathrm{T}}(x - z)\, dt.$$

By (6) and Lebesque's dominated convergence theorem, we have

$$\lim_{k\to\infty} \int_0^1 \nabla\psi_k(z+t(x^k-z))^{\mathrm{T}}(x^k-z)\,dt \;=\; \int_0^1 \nabla\psi(z+t(x-z))^{\mathrm{T}}(x-z)\,dt.$$

Hence, the convergence of $\{\psi_k(x^k)\}$ to $\psi(x)$ follows readily.   □

Remark 2 A referee has observed that the proof of the "if" part in Proposition 1 requires only the "boundedness of the Lipschitz modulus". Thus the proposition remains valid if the continuous convergence of $\{\nabla\psi_k\}$ to ψ is replaced by the Lipschitz continuity of ψ_k on X together with the boundedness of the Lipschitz modulus. The proof remains more or less the same; it suffices to use a mean-value theorem for a Lipschitz continuous function and the Clarke subdifferential [7] instead of the gradient. This referee also noted that using Corollary 5.45 and Theorem 5.40 in [29], Proposition 1 can be strengthened as follows: $\{\psi_k\}$ converges continuously on X if and only if $\{\psi_k\}$ converges pointwise on X and is "asymptotically equicontinuous" on X, that is, for any $x \in X$ and any $\varepsilon > 0$, a neighborhood V of x exists such that $\|\psi_k(y) - \psi_k(x)\| \leq \varepsilon$ for all k and all $y \in V$, and ψ_k is bounded on V.   □

To present the promised characterization of the pointwise (thus continuous) convergence of a sequence of differentiable convex functions in terms of the continuous convergence of the gradient functions, we let $\widehat{X}$ be the family of closed convex subsets of X and define

$$\Xi \equiv \left\{ S \in \widehat{X} \;:\; \exists\ \text{bounded } \{x^k\} \text{ such that } x^k \in \arg\min_{x\in S} \psi_k(x) \quad \forall\,k \right\}.$$

Clearly, Ξ contains all the singletons $\{x\}$ with $x \in X$. If X is convex and compact, then $\Xi = \widehat{X}$.

Proposition 2 *Let ψ_k and ψ be differentiable functions defined on an open set containing the convex subset X of $\Re^n$. Consider the following statements (a–d).*

(a) *The sequence of gradients $\{\nabla\psi_k\}$ converges continuously to the gradient $\nabla\psi$ on X and for all $S \in \Xi$,*

$$\lim_{k\to\infty}\ \min_{x\in S}\ \psi_k(x) \;=\; \min_{x\in S}\ \psi(x). \tag{7}$$

(b) *The sequence of gradients $\{\nabla\psi_k\}$ converges continuously to the gradient $\nabla\psi$ on X and there exists $\bar{x} \in X$ such that $\lim_{k\to\infty}\ \psi_k(\bar{x}) = \psi(\bar{x})$.*

(c) *The sequence of functions $\{\psi_k\}$ converges continuously to ψ on X.*

(d) *The sequence of functions $\{\psi_k\}$ converges pointwise to ψ on X.*

It holds that (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d). If in addition each ψ_k is convex, then (d) $\Rightarrow$ (a); moreover, any one of the statements (a–d) is further equivalent to

(e) *the sequence of gradients $\{\nabla\psi_k\}$ converges continuously to the gradient $\nabla\psi$ on X and there exists $S \in \Xi$ such that (7) holds.*

Proof (a) $\Rightarrow$ (b). This is obvious by taking $S = \{\bar{x}\}$.

(b) $\Rightarrow$ (c). This follows from Proposition 1.

(c) $\Rightarrow$ (d). This is obvious.

(d) $\Rightarrow$ (a) under convexity of ψ_k. It suffices to prove that (7) holds for all sets $S \in \Xi$. Let S be any such set. Let $\{x^k\} \subset S$ be a bounded sequence such that $x^k \in \arg\min\{\psi_k(x) : x \in S\}$ for all k. Since S is bounded, $\{x^k\}$ has at least one accumulation point; moreover, every such point must be a minimizer of ψ on S. Indeed, let $\{x^k : k \in \kappa\}$ be a subsequence of $\{x^k\}$ converging to a limit x^∞, which must belong to S by its closedness. For every $x \in S$, we have $\psi_k(x^k) \leq \psi_k(x)$. Passing to the limit $k(\in \kappa) \to \infty$ and using the pointwise (and thus, continuous) convergence of $\{\psi_k\}$ to ψ on X, it follows that $\psi(x^\infty) \leq \psi(x)$, as desired.

Finally, it remains to show that if each ψ_k is convex, then (e) implies (d). By way of contradiction, assume that the sequence $\{\psi_k(\bar{x})\}$ does not converge to $\psi(\bar{x})$ for some $\bar{x} \in X$. There exist a scalar $\varepsilon > 0$ and an infinite subset κ_ε of $\{1, 2, \ldots, \}$ such that $|\psi_k(\bar{x}) - \psi(\bar{x})| > \varepsilon$ for all $k \in \kappa_\varepsilon$. Let $\{x^k\}$ be a bounded sequence such that $x^k \in \arg\min\{\psi_k(x) : x \in S\}$ for all k, where S is as prescribed by (e). By working with an appropriate subsequence of $\{x^k : k \in \kappa_\varepsilon\}$ if necessary, we can assume without loss of generality that the latter sequence converges to some limit x^∞, which must be an element of S by the closedness of S. We claim that $x^\infty \in \arg\min\{\psi(x) : x \in S\}$. Indeed, by the definition of x^k, we have, for every k,

$$(y - x^k)^\mathrm{T}\nabla\psi_k(x^k) \geq 0 \quad \forall\, y \in S.$$

Passing to the limit $k(\in \kappa_\varepsilon) \to \infty$ and by the continuous convergence of $\{\nabla\psi_k\}$ to $\nabla\psi$, we deduce

$$(y - x^\infty)^\mathrm{T}\nabla\psi(x^\infty) \geq 0 \quad \forall\, y \in S,$$

which establishes the claim, by the convexity of ψ. It follows from (7) that

$$\lim_{k(\in\kappa_\varepsilon)\to\infty} \psi_k(x^k) = \psi(x^\infty).$$

By the same proof as that for Proposition 1, we deduce that $\{\psi_k : k \in \kappa_\varepsilon\}$ converges continuously, hence pointwise, to ψ on X. But this contradicts the fact that $|\psi_k(\bar{x}) - \psi(\bar{x})| > \varepsilon$ for all $k \in \kappa_\varepsilon$. Hence, part (e) holds. $\square$

Remark 3 Proposition 2 is reminiscent of Attouch's theorem that characterizes the epiconvergence of convex functions in terms of the graphical convergence of their subdifferentials, which also involves the existence of a point $\bar{x}$ at which the sequence $\{\psi_k(\bar{x})\}$ converges to $\psi(\bar{x})$; see [29, Theorem 12.35]. $\square$

Propositions 1 and 2 can be used to weaken the strong limit hypothesis (5) to a "diagonally continuous convergence" of the gradients; see condition (CcD) below.

We first present a result using the former proposition. No convexity is needed in this case.

Corollary 2 *Let each X^ν be a closed subset of $\Re^{n_\nu}$. Let (1) be a family of differentiable multi-component functions satisfying (odec). Suppose furthermore that for every $\nu = 1, \ldots, N$ and every $x^{-\nu} \in X^{-\nu}$, the function $\theta_\nu(\cdot, x^{-\nu})$ is differentiable,*

(Pc1)   *there exists $\bar{x}^\nu \in X^\nu$ such that $\lim_{k \to \infty} \theta_{\nu,k}(\bar{x}^\nu, x^{-\nu}) = \theta_\nu(\bar{x}^\nu, x^{-\nu})$, and*

(CcD)   *the sequence of principal gradients $\{\nabla_{x^\nu} \theta_{\nu,k}(\cdot, x^{-\nu})\}$ converges continuously to $\nabla_{x^\nu} \theta_\nu(\cdot, x^{-\nu})$ on X^ν.*

Then the conclusion of Theorem 1 holds.

Proof By Proposition 1, it follows that for every $\nu = 1, \ldots, N$ and every $x^{-\nu} \in X^{-\nu}$, the sequence of functions $\{\theta_{\nu,k}(\cdot, x^{-\nu})\}$ converges continuously to the function $\theta_\nu(\cdot, x^{-\nu})$. Hence, conditions (Da) and (Db) both hold. Since (odec) is in place, the corollary follows from Theorem 2.   □

If the functions $\theta_{\nu,k}(\cdot, x^{-\nu})$ and $\theta_\nu(\cdot, x^{-\nu})$ are all convex, then the condition (Pc1) can be replaced by the convergence of the sequence of minimum objective values $\left\{ \min_{x^\nu \in S^\nu} \theta_{\nu,k}(x) \right\}$ to $\min_{x^\nu \in S^\nu} \theta_\nu(x)$ for a suitable subset S^ν of X^ν. In particular, taking S^ν to be X^ν, we have the following result, whose proof is the same as that for Corollary 2 and thus omitted.

Corollary 3 *Let each X^ν be a closed convex subset of $\Re^{n_\nu}$. Let (1) be a family of multi-component functions satisfying (odec). Suppose that for every $\nu = 1, \ldots, N$ and every $x^{-\nu} \in X^{-\nu}$, the functions $\theta_\nu(\cdot, x^{-\nu})$ and $\theta_{\nu,k}(\cdot, x^{-\nu})$ for all k are differentiable and convex and that*

(Pc1′)   *X^ν contains a bounded sequence $\{x^{\nu,k}\}$, where for each k, $x^{\nu,k} \in \arg\min_{x^\nu \in X^\nu}$ $\theta_{\nu,k}(x)$ and*

$$\lim_{k \to \infty} \min_{x^\nu \in X^\nu} \theta_{\nu,k}(x) = \min_{x^\nu \in X^\nu} \theta_\nu(x),$$

(CcD)   *the sequence of principal gradients $\{\nabla_{x^\nu} \theta_{\nu,k}(\cdot, x^{-\nu})\}$ converges continuously to $\nabla_{x^\nu} \theta_\nu(\cdot, x^{-\nu})$ on X^ν.*

Then the conclusion of Theorem 1 holds.   □

We conclude the discussion in this section by raising the question of whether there is an extension of Corollary 3 for nondifferentiable functions wherein the gradients in condition (CcD) are replaced by the respective subgradients as done in Theorem 4. We do not have an answer to the question at this time.

4 The optimization approach

In [15,16], Jofré and Wets employ the "Nikaido–Isoda" (NI) function [25] to formulate an equivalent optimization problem to study the approximation of Nash equilibria.

(Incidentally, such an optimization formulation is not new (cf. e.g. [30]) and has been the basis for a class of "relaxation algorithms" for computing Nash equilibria, even with coupled constraints [5,20,21,33].) Jofré and Wets also approximate the players' strategy sets. Consistent with our developments so far, we will not consider the latter approximation. The main tool that Jofré and Wets used is the theory of lopsided convergence for bivariate functions [2–4]. In order for this theory to be applicable, these authors assumed that the sequence of approximate objective functions converges continuously to the players' objective functions. Our goal in this section is to show that under the multi-epiconvergence condition, it is possible to bypass the lopsided convergence theory and establish the convergence of the sequence of NI minima to a Nash equilibrium directly, thereby strengthening the Jofré-Wets analysis with the weakening of the continuous convergence assumption. Furthermore, our analysis allows the approximate Nash equilibria be computed inexactly according to a certain error criterion.

In what follows, we employ a regularization of the NI function. The benefit of such a regularization is well known from the optimization perspective. Phrased in the game context, the regularization yields unique players' responses to rivals' strategies. Specifically, the *regularized Nikaido–Isoda* function is, by definition, for an arbitrary scalar $c > 0$ the bivariate function: for $x = (x^\nu)_{\nu=1}^N$ and $y = (y^\nu)_{\nu=1}^N$,

$$\phi_c(x, y) \equiv \sum_{\nu=1}^N \left[\theta_\nu(x^\nu, x^{-\nu}) - \theta_\nu(y^\nu, x^{-\nu}) - \frac{c}{2}(x^\nu - y^\nu)^{\mathrm{T}}(x^\nu - y^\nu) \right].$$

Consider the following value function:

$$\chi_c(x) \equiv \max_{y \in X} \phi_c(x, y), \quad x \in X$$
$$= \sum_{\nu=1}^N \left\{ \theta_\nu(x) - \min_{y^\nu \in X^\nu} \left[\theta_\nu(y^\nu, x^{-\nu}) + \frac{c}{2}(x^\nu - y^\nu)^{\mathrm{T}}(x^\nu - y^\nu) \right] \right\}.$$

There are close connections between the function $\chi_c(x)$ and the regularized gap function for a VI [10–12]; since such connections are not relevant to our analysis here, we omit their discussion. Instead, we summarize several basic properties of $\chi_c(x)$ in the following result.

Proposition 3 *Let each X^ν be a closed convex subset of $\Re^{n_\nu}$. Assume that for every $\nu = 1, \ldots, N$ and every $x^{-\nu} \in X^{-\nu}$, $\theta_\nu(\cdot, x^{-\nu})$ is convex. The following statements are valid for any scalar $c > 0$.*

(a) *$\chi_c(x)$ is a well-defined nonnegative function on X.*
(b) *$\widehat{x}$ is a Nash equilibrium if and only if $\widehat{x}$ is a global minimizer of $\chi_c(x)$ on X and $\chi_c(\widehat{x}) = 0$.*
(c) *For every $x \in X$, there exists a unique $y(x) \equiv (y^\nu(x))_{\nu=1}^N$ such that for every $\nu = 1, \ldots, N$,*

$$\operatorname*{arg\,min}_{y^\nu \in X^\nu} \left[\theta_\nu(y^\nu, x^{-\nu}) + \frac{c}{2}(x^\nu - y^\nu)^{\mathrm{T}}(x^\nu - y^\nu) \right] = y^\nu(x),$$

    and $y^\nu(x)$ is continuous in x.

Proof Part (a) is clear; so is the "only if" statement of part (b). To prove the "if" statement, suppose that $\widehat{x} \in X$ satisfies $\chi_c(\widehat{x}) = 0$. We then have $\phi_c(\widehat{x}, y) \leq 0$ for all $y \in X$. Fix an arbitrary ν and let $x^\nu \in X^\nu$ be arbitrary. For any $\tau \in [0, 1]$, let $y^{\nu'} = \widehat{x}^{\nu'}$ for all $\nu' \neq \nu$, and $y^\nu = \tau\widehat{x}^\nu + (1 - \tau)x^\nu$. We then have

$$
\begin{aligned}
0 &\geq \phi_c(\widehat{x}, y) \\
&= \theta_\nu(\widehat{x}^\nu, \widehat{x}^{-\nu}) - \theta_\nu(\tau\widehat{x}^\nu + (1 - \tau)x^\nu, \widehat{x}^{-\nu}) + \frac{c}{2}(1 - \tau)^2 (\widehat{x}^\nu - x^\nu)^{\mathrm{T}}(\widehat{x}^\nu - x^\nu) \\
&\geq (1 - \tau)\left[\theta_\nu(\widehat{x}^\nu, \widehat{x}^{-\nu}) - \theta_\nu(x^\nu, \widehat{x}^{-\nu}) + \frac{c}{2}(1 - \tau)(\widehat{x}^\nu - x^\nu)^{\mathrm{T}}(\widehat{x}^\nu - x^\nu)\right].
\end{aligned}
$$

Dividing by $1 - \tau$ and then letting $\tau \uparrow 1$, we deduce that $\theta_\nu(\widehat{x}) \leq \theta_\nu(x^\nu, \widehat{x}^{-\nu})$ for all $x^\nu \in X^\nu$. Since this holds for all $\nu = 1, \ldots, N$, it follows that $\widehat{x}$ is a Nash equilibrium. Part (c) follows well-known results in parametric optimization in view of the fact that for every $\nu = 1, \ldots, N$ and every $x \in X$, the function $y^\nu \mapsto \theta_\nu(y^\nu, x^{-\nu}) + \frac{c}{2}(x^\nu - y^\nu)^{\mathrm{T}}(x^\nu - y^\nu)$ is strongly convex with modulus at least $c > 0$, which is independent of x.   □

    By part (b) of Proposition 3, the Nash equilibrium problem is equivalent to the minmax problem

$$
\min_{x \in X} \chi_c(x) = \min_{x \in X} \max_{y \in X} \phi_c(x, y)
$$

having a solution $\widehat{x}$ that satisfies $\chi_c(\widehat{x}) = 0$. In terms of the function $y(x)$, we have

$$
\chi_c(x) = \sum_{\nu=1}^{N} (\theta_\nu(x) - \zeta_\nu(x)), \quad \text{where}
$$

$$
\begin{aligned}
\zeta_\nu(x) &\equiv \min_{y^\nu \in X^\nu}\left[\theta_\nu(y^\nu, x^{-\nu}) + \frac{c}{2}(x^\nu - y^\nu)^{\mathrm{T}}(x^\nu - y^\nu)\right] \\
&= \theta_\nu(y^\nu(x), x^{-\nu}) + \frac{c}{2}(x^\nu - y^\nu(x))^{\mathrm{T}}(x^\nu - y^\nu(x)).
\end{aligned}
$$

Note that a vector $x \in X$ satisfying $\chi_c(x) \leq \varepsilon$ for some $\varepsilon > 0$ can be considered an inexact Nash equilibrium. In the theorem below, this is used as an inexact criterion for the approximate Nash equilibria corresponding to a family of multi-component functions $\{\theta_{\nu,1}, \theta_{\nu,2}, \ldots\}_{\nu=1}^{N}$ approximating $\{\theta_\nu\}_{\nu=1}^{N}$. For such a family, we write $\chi_{c,k}(x)$ for the value function corresponding to the k-th subfamily $\{\theta_{\nu,k}\}_{\nu=1}^{N}$; we also write $\zeta_{\nu,k}(x)$ accordingly.

Theorem 5 *Let each $X^\nu \subseteq \Re^{n_\nu}$ be closed convex and $\{\theta_{\nu,1}, \theta_{\nu,2}, \ldots\}_{\nu=1}^{N}$ be a family of functions multi-epiconverging to the functions $\{\theta_\nu\}_{\nu=1}^{N}$ on the set X. Suppose that for every $\nu = 1, \ldots, N$ and every $x^{-\nu} \in X^{-\nu}$, $\theta_\nu(\cdot, x^{-\nu})$ and $\theta_{\nu,k}(\cdot, x^{-\nu})$ for all k are convex. Let $\{\varepsilon_k\}$ be an arbitrary sequence of positive scalars converging to zero*

and let $c > 0$ be an arbitrary scalar. If the sequence $\{x^k \equiv (x^{\nu,k})_{\nu=1}^N\}$, where each $x^k \in X$ satisfying $\chi_{c,k}(x^k) \le \varepsilon_k$, converges to $x^\infty \equiv (x^{\nu,\infty})_{\nu=1}^N$, then x^∞ is a Nash equilibrium tuple of the pairs $(\theta_\nu, X^\nu)_{\nu=1}^N$.

Proof We have

$$\chi_{c,k}(x^k) = \sum_{\nu=1}^N \left(\theta_{\nu,k}(x^k) - \zeta_{\nu,k}(x^k) \right), \quad \text{where}$$

$$\zeta_{\nu,k}(x^k) = \min_{y^\nu \in X^\nu} \left[\theta_{\nu,k}(y^\nu, x^{-\nu,k}) + \frac{c}{2}(x^{\nu,k} - y^\nu)^T(x^{\nu,k} - y^\nu) \right].$$

By the multi-epiconvergence of $\{\theta_{\nu,k}\}_{\nu=1}^N$ to $\{\theta_\nu\}_{\nu=1}^N$ on X, it follows that for every $\nu = 1, \ldots, N$, the sequence of functions

$$\{ \theta_{\nu,k}(\cdot, x^{-\nu,k}) + (c/2)(x^{\nu,k} - \cdot)^T(x^{\nu,k} - \cdot) \}$$

epiconverges to the function $\theta_\nu(\cdot, x^{-\nu,\infty}) + (c/2)(x^{\nu,\infty} - \cdot)^T(x^{\nu,\infty} - \cdot)$ on X^ν. Consequently,

$$\lim_{k\to\infty} \zeta_{\nu,k}(x^k) = \zeta_\nu(x^\infty) \quad \forall \nu = 1, \ldots, N.$$

Since

$$\liminf_{k\to\infty} \sum_{\nu=1}^N \theta_{\nu,k}(x^k) \ge \sum_{\nu=1}^N \liminf_{k\to\infty} \theta_{\nu,k}(x^k) \ge \sum_{\nu=1}^N \theta_\nu(x^\infty),$$

where the second inequality is due to condition (Mb), we obtain

$$\liminf_{k\to\infty} \chi_{c,k}(x^k) \ge \sum_{\nu=1}^N \left(\theta_\nu(x^\infty) - \zeta_\nu(x^\infty) \right) = \chi_c(x^\infty).$$

By assumption, the lim inf on the left side is equal to zero. Since χ_c is a nonnegative function, it follows that $\chi_c(x^\infty) = 0$. Hence x^∞ is a desired Nash equilibrium. $\square$

The advantage of Theorem 5 over Theorem 1 is that each x^k is not required to be an exact Nash equilibrium for $(\theta_{\nu,k}, X^\nu)_{\nu=1}^N$. This is presumably a benefit of using the regularized NI function versus the direct approach. The verification of this benefit in practical computations is beyond the scope of this work, whose aim is to understand the supporting theory of these approaches.

4.1 Multi-epi, lopsided, and system-epi convergence

In this subsection, we briefly discuss the relation between multi-epiconvergence of the players' payoff functions and lopsided convergence applied to the original (un-regularized) Nikaido–Isoda bivariate function:

$$\phi(x, y) \equiv \sum_{v=1}^{N} \left[\theta_v(x^v, x^{-v}) - \theta_v(y^v, x^{-v}) \right].$$

For this purpose, we take a family $\{\theta_{v,1}, \theta_{v,2}, \ldots\}_{v=1}^{N}$ of multi-component functions approximating the players' objective functions $\{\theta_v\}_{v=1}^{N}$. This sequence of approximating functions induces the family $\{\phi_k\}$, where

$$\phi_k(x, y) \equiv \sum_{v=1}^{N} \left[\theta_{v,k}(x^v, x^{-v}) - \theta_{v,k}(y^v, x^{-v}) \right], \quad k = 1, 2, \ldots.$$

By the definition of lopsided convergence (see [2–4, 15, 16]), we say that the family $\{\phi_k\}$ *converges lopsidedly* to ϕ if for every pair $(x, y) \in X \times X$,

(Na) for every sequence $\{x^k\} \subset X$ converging to x, a sequence $\{y^k\} \subset X$ converging to y exists such that

$$\liminf_{k \to \infty} \phi_k(x^k, y^k) \geq \phi(x, y);$$

(Nb) a sequence $\{x^k\} \subset X$ converging to x exists such that for every sequence $\{y^k\} \subset X$ converging to y,

$$\limsup_{k \to \infty} \phi_k(x^k, y^k) \leq \phi(x, y).$$

Under lopsided convergence, it follows that if $\{(x^k, y^k)\}$ is a sequence converging to (x, y) such that, for every k, (x^k, y^k) is a saddle point of ϕ_k on $X \times X$, then (x, y) is a saddle point of ϕ on $X \times X$.

In what follows, we examine the connection between the conditions (Na)–(Nb) and the conditions (Ma)–(Mb). It turns out that these conditions are naturally linked to the approximation of the "system objective function", which is by definition the sum of the players' objective functions: namely, $\Psi(x) \equiv \sum_{v=1}^{N} \theta_v(x)$. For convenience of comparison, we recall that the family $\{\Psi_k\}$, where $\Psi_k(x) \equiv \sum_{v=1}^{N} \theta_{v,k}(x)$, epiconverges to Ψ on X if the following two conditions hold for every $x \in X$:

(Sa) a sequence $\{x^k\} \subset X$ converging to x exists such that

$$\limsup_{k \to \infty} \sum_{v=1}^{N} \theta_{v,k}(x^k) \leq \sum_{v=1}^{N} \theta_v(x),$$

(Sb) for every sequence $\{x^k\} \subset X$ converging to x,

$$\liminf_{k \to \infty} \sum_{v=1}^{N} \theta_{v,k}(x^k) \geq \sum_{v=1}^{N} \theta_v(x).$$

The minimization of the function $\Psi(x)$ on the set X is known as the *system optimization problem*; the latter problem is often used in a system planning context as opposed to the Nash equilibrium problem which is a model describing the non-cooperative behavior of the players.

The following result summarizes the relations between the three sets of conditions (Ma)–(Mb), (Na)–(Nb), and (Sa)–(Sb).

Proposition 4 The following statements hold.

(A) (Ma) + (Mb) $\Rightarrow$ (Na);
(B) (Mb) $\Rightarrow$ (Sb);
(C) under (Ma) and (Mb), (Nb) $\Leftrightarrow$ (Sa).

Proof It would be useful to remind the reader the following elementary fact: for any finite number of sequences $\{a_{v,k}\}_{v=1}^{N}$,

$$\sum_{v=1}^{N} \liminf_{k \to \infty} a_{v,k} \leq \liminf_{k \to \infty} \sum_{v=1}^{N} a_{v,k} \leq \limsup_{k \to \infty} \sum_{v=1}^{N} a_{v,k} \leq \sum_{v=1}^{N} \limsup_{k \to \infty} a_{v,k}.$$

Suppose (Ma) and (Mb) hold. Let $(x, y) \in X \times X$ be given and let $\{x^k\} \subset X$ converge to x. For each $v = 1, \ldots, N$, let $\{y^{v,k}\} \subset X^v$ be a sequence converging to y^v, whose existence is guaranteed by (Ma), such that

$$\limsup_{k \to \infty} \theta_{v,k}(y^{v,k}, x^{-v,k}) \leq \theta_v(y^v, x^{-v}).$$

By (Mb), we have

$$\liminf_{k \to \infty} \theta_{v,k}(x^{v,k}, x^{-v,k}) \geq \theta_v(x^v, x^{-v}).$$

Negating the former inequality and adding all the inequalities for $v = 1, \ldots, N$ yield (Na). Statement (B) is obvious in view of the string of inequalities noted at the beginning of the proof.

Suppose (Mb) and (Sa) hold. Let $(x, y) \in X \times X$ be given. Let $\{x^k\} \subset X$ be a sequence converging to x, whose existence is guaranteed by (Sa), such that

$$\limsup_{k \to \infty} \sum_{v=1}^{N} \theta_{v,k}(x^k) \leq \sum_{v=1}^{N} \theta_v(x). \tag{8}$$

For every sequence $\{y^k\} \subset X$ converging to y, by (Mb), we have, for every $\nu = 1, \ldots, N$,
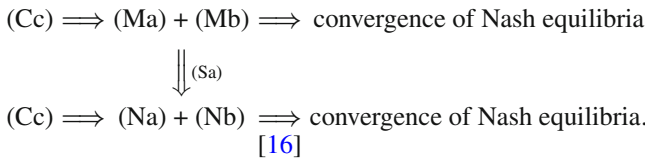
$$\liminf_{k \to \infty} \theta_{\nu,k}(y^{\nu,k}, x^{-\nu,k}) \geq \theta_\nu(y^\nu, x^{-\nu}).$$

Negating this inequality and adding for all ν and to (8) yield (Nb). Conversely, suppose (Ma), (Mb), and (Nb) hold. Let $x \in X$. By (Nb), there exists a sequence $\{x^k\} \subset X$ converging to x such that for any sequence $\{\widetilde{x}^k\} \subset X$ converging to x, we have

$$\limsup_{k \to \infty} \sum_{\nu=1}^{N} \left[\theta_{\nu,k}(x^k) - \theta_{\nu,k}(\widetilde{x}^{\nu,k}, x^{-\nu,k}) \right] \leq 0. \tag{9}$$

By (Ma), for every $\nu = 1, \ldots, N$, we may choose the sequence $\{\widetilde{x}^{\nu,k}\} \subset X^\nu$ such that $\limsup_{k \to \infty} \theta_{\nu,k}(\widetilde{x}^{\nu,k}, x^{-\nu,k}) \leq \theta_\nu(x)$. By (Mb), it follows that $\lim_{k \to \infty} \theta_{\nu,k}(\widetilde{x}^{\nu,k}, x^{-\nu,k}) = \theta_\nu(x)$. Adding these equalities for $\nu = 1, \ldots, N$ and to (9), we deduce (Sa). □

It is natural to ask whether there are further relations between these conditions; regrettably, we don't have an answer for this question. In particular, we don't know whether the three sets of conditions are equivalent if the system objective function is convex. Instead of pursuing such open issues, we summarize the above discussion in the diagram below.

$$(Cc) \Longrightarrow (Ma) + (Mb) \Longrightarrow \text{convergence of Nash equilibria}$$

$$\Big\| (Sa)$$

$$(Cc) \Longrightarrow (Na) + (Nb) \Longrightarrow \text{convergence of Nash equilibria.}$$
[16]

5 Competitive capacity expansion under uncertainty

This section presents an illustrative example of a Nash equilibrium problem where approximation of the players' cost functions is needed for practical solution. The problem is a Nash–Cournot extension of a well-known capacity expansion problem under economic uncertainty. The latter problem is conveniently formulated as a standard two-stage stochastic program with recourse; see e.g. [6,14,18]. The extension from a monopolistic environment to an oligopolistic competition is the topic of discussion herein. While being a natural combination of two well-known problems—a stochastic monopolistic two-stage model and a deterministic oligopolistic Nash–Cournot model, the full model has not been formally presented and analyzed in the literature.

There are finitely many firms, each labelled by an index $\nu = 1, \ldots, N$, who are competing in finitely many spatially separated markets indexed by the elements in the set $\mathcal{N}$. Each firm has the same information on market uncertainty; there is no entry of new firm and there is no firm exit in the model. Firm ν chooses a vector of regional production capacities denoted by $x^\nu \in X^\nu \subseteq \Re^{n_\nu}$, where X^ν is a closed convex set

of admissible capacities. Capacity is installed prior to demand being realized; these supply-side decisions must be made prior to observing market equilibrium. In modeling economic uncertainty, it is reasonable to parameterize price-quantity relationships. Thus we have a regional price $p_j(Q_j, \omega)$ that is a random function defined on the triple $(\Omega, \mathcal{F}, \mathcal{P})$, which denotes the underlying common probability space for all the random variables and random functions in the model with $\omega \in \Omega$, and $Q_j \equiv \sum_{\nu=1}^{N} q_j^{\nu}$ is the total production by all firms in region $j \in \mathcal{N}$. For each realization ω of the model's randomness, and each (deterministic) capacity vector x^{ν}, firm ν determines its production quantities $q^{\nu} \equiv (q_j^{\nu})_{j \in \mathcal{N}}$ by solving a profit maximization problem in anticipation of other firms' productions $q^{-\nu}$:

$$\text{maximize } \pi_{\nu}(q^{\nu}, q^{-\nu}, \omega) \equiv \sum_{j \in \mathcal{N}} \left[p_j(Q_j, \omega)q_j^{\nu} - c_j^{\nu}(q^{\nu}) \right]$$

$$\text{subject to } q^{\nu} \in S^{\nu}(x^{\nu}, \omega) \tag{10}$$

where $S^{\nu}(x^{\nu}, \omega)$ is firm ν's set of admissible productions and $c_j^{\nu}(q^{\nu})$ is firm ν's production cost in region j. In what follows, we assume that

$$S^{\nu}(x^{\nu}, \omega) \equiv \{ q^{\nu} \in \Re^{m_{\nu}} : W^{\nu}q^{\nu} \leq h^{\nu}(\omega) - B^{\nu}(\omega)x^{\nu} \},$$

with W^{ν} being firm ν's recourse matrix, which we assume to be constant, B^{ν} being firm ν's technology matrix, and h^{ν} is a vector, with the latter matrix and vector being random.

When the map $(q^{\nu})_{\nu=1}^{N} \mapsto -\left(\nabla_{q^{\nu}} \pi_{\nu}(q, \omega)\right)_{\nu=1}^{N}$ is strongly monotone, an equilibrium solution $q^*(x, \omega) \equiv \left(q^{\nu,*}(x, \omega)\right)_{\nu=1}^{N}$ of the firms' productions to the Nash subgame defined by the maximization problems (10) exists and is unique. This subgame is the *second-stage* game. Let $\vartheta_{\nu}(x, \omega) \equiv \pi_{\nu}(q^*(x, \omega), \omega)$ be firm ν's maximum profit as a function of all firms' capacity vector x and the random realization ω. The *first-stage* game is to determine a capacity tuple $x^* \equiv (x^{\nu,*})_{\nu=1}^{N}$ such that, for all $\nu = 1, \ldots, N$, $\theta_{\nu}(x^*) \leq \theta_{\nu}(x^{\nu}, x^{-\nu,*})$ for all $x^{\nu} \in X^{\nu}$, where

$$\theta_{\nu}(x) \equiv \eta_{\nu}(x^{\nu}) - \mathbb{E}\vartheta_{\nu}(x, \omega)$$

is firm ν's net cost $\eta_{\nu}(x^{\nu})$ of securing the capacity x^{ν} less the expected maximum profit.

For a given integer $\ell > 0$, let $\{\omega^1, \ldots, \omega^{\ell}\} \subset \Omega$ be ℓ independently and identically distributed samples of the model's randomness. Consider the random functions:

$$\theta_{\nu,\ell}(x) \equiv \eta_{\nu}(x^{\nu}) - \frac{1}{\ell} \sum_{i=1}^{\ell} \vartheta_{\nu}(x, \omega^i) = \eta_{\nu}(x^{\nu}) - \frac{1}{\ell} \sum_{i=1}^{\ell} \pi_{\nu}(q^*(x, \omega^i), \omega^i)$$

and the approximate Nash model wherein player ν's optimization problem is

$$\text{minimize } \theta_{\nu,\ell}(x) \quad \text{subject to } x^{\nu} \in X^{\nu}. \tag{11}$$

Let $\widehat{x}^\ell \equiv (\widehat{x}^{\nu,\ell})_{\nu=1}^N$ denote an equilibrium solution to the latter approximated Nash capacity game, which we assume exists. Notice that this solution is a random vector. (A remark: not implied by the setting so far, the existence of $\widehat{x}^\ell$ is not a trivial matter because the capacity game is actually an *equilibrium problem with equilibrium constraints* whose systematic study is still very much in an infant stage. For one thing, player ν's objective function $\theta_{\nu,\ell}(x^\nu, x^{-\nu})$ is not necessarily a convex function in x^ν for fixed $x^{-\nu}$, due to the non-concavity of $\vartheta_\nu(x^\nu, x^{-\nu}, \omega)$ in x^ν.) In what follows, we wish to establish the almost sure (subsequential) convergence of $(\widehat{x}^\ell)_{\ell=1}^\infty$ to a deterministic equilibrium x^* of the two-stage capacity game. The cornerstone to establish this convergence is to show that, under the assumptions imposed in Proposition 5 below, the sequence of random variables $\left\{ \frac{1}{\ell} \sum_{i=1}^\ell \pi_\nu(q^*(x, \omega^i), \omega^i) \right\}_{\ell=1}^\infty$ continuously converges in x to the expected value $\mathbb{E}_\omega \vartheta_\nu(x, \omega)$ almost surely. Toward this demonstration, we assume that each firm's profit function $\pi_\nu(\cdot, q^{-\nu}, \omega)$ is continuously differentiable and concave for fixed $q^{-\nu}$. It then follows that the vector $q^*(x, \omega)$ is a solution of the VI $(\Xi(x, \omega), \Pi(\cdot, \omega))$ where

$$\Xi(x, \omega) \equiv \prod_{\nu=1}^N S^\nu(x^\nu, \omega) \quad \text{and} \quad \Pi(q, \omega) \equiv -\left(\nabla_{q^\nu} \pi_\nu(q, \omega) \right)_{\nu=1}^N.$$

It is well-known that when the function $\Pi(\cdot, \omega)$ is strongly monotone with a modulus that is valid for all $\omega \in \Omega$, $q^*(x, \omega)$ exists and is the unique solution of the said VI; see [10]. The next result, Proposition 5, will establish an important uniform (in ω) locally Hölderian property of the solution function $q^*(\cdot, \omega)$ under some mild assumptions. Although there are variations of this result in the VI literature, see e.g. [10, Chaps. 5 and 6], Proposition 5 below distinguishes itself both in the setting and the conclusion. For the setting, we note that the constraint set $\Xi(x, \omega)$ of the VI $(\Xi(x, \omega), \Pi(\cdot, \omega))$ depends on the pair (x, ω); for the conclusion, we note that the resulting Hölderian bounds (13) and (14) hold uniformly for all ω. For completeness, we will give a detailed proof of the result below. (See Remarks 4 and 5 for more discussion related to Proposition 5.)

Afterwards, using this rather technical result, we establish in Theorem 6 below the almost sure convergence of the sample average of the players' profits, from which the almost sure convergence of the Nash equilibria of the sampled games to a Nash equilibrium of the stochastic capacity game follows readily. Finally in Sect. 5.1, we present an application of Theorem 6 for isoelastic price-demand functions of the form $p_j(Q_j, \omega) = a_j(\omega) Q_j^{-1/b_j(\omega)}$ with $Q_j > 0$ and appropriate positive random coefficients a_j and b_j. In particular, we demonstrate how to verify some of the conditions of Proposition 5. Clearly, a similar (and simpler) analysis can also be carried out for linear price-demand relationships of the form $p_j(Q_j, \omega) = a_j(\omega) - b_j(\omega) Q_j$ with the assumptions on a_j and b_j modified appropriately.

Proposition 5 *Suppose that conditions* (A) *and* (B) *below hold:*

(A) *for each* $(x, \omega) \in X \times \Omega$, $q^*(x, \omega)$ *exists and is unique;*

(B) $\displaystyle \sup_{\omega \in \Omega} \max \left(\max_{1 \leq \nu \leq N} \| B^\nu(\omega) \|, \ \max_{1 \leq \nu \leq N} \| h^\nu(\omega) \| \right) < \infty.$

Let $\mathcal{X}$ be a bounded subset of X such that

(a) *a constant $\rho > 0$ exists such that for all triples (x, x', ω) in $\mathcal{X} \times \mathcal{X} \times \Omega$,*

$$(q^*(x, \omega) - q^*(x', \omega))^{\mathrm{T}}(\Pi(q^*(x, \omega), \omega) - \Pi(q^*(x', \omega), \omega))$$
$$\geq \rho \| q^*(x, \omega) - q^*(x', \omega) \|^2;$$

(b) *a vector $\bar{x} \in \mathcal{X}$ exists such that $\sup\limits_{\omega \in \Omega} \| q^*(\bar{x}, \omega) \| < \infty$;*

(c) *positive constants $\xi < 2$, $\gamma > 0$, $\beta > 0$, ζ_1, and ζ_2 exist such that for all $(x, x', \omega) \in \mathcal{X} \times \mathcal{X} \times \Omega$,*

$$\| \Pi(q^*(x, \omega), \omega) \| \leq \zeta_1 + \zeta_2 \| q^*(x, \omega) \|^\xi$$

and, for all $\nu = 1, \ldots, N$,

$$\left| \pi_\nu(q^*(x, \omega), \omega) - \pi_\nu(q^*(x', \omega), \omega) \right| \leq \beta \| q^*(x, \omega) - q^*(x', \omega) \|^\gamma. \quad (12)$$

Then positive scalars $L_\mathcal{X}$ and $L'_\mathcal{X}$ exist such that for all triples (x, x', ω) in $\mathcal{X} \times \mathcal{X} \times \Omega$,

$$\| q^*(x, \omega) - q^*(x', \omega) \| \leq L_\mathcal{X} \| x - x' \|^{1/2} \quad (13)$$

and, for all $\nu = 1, \ldots, N$,

$$\left| \pi_\nu(q^*(x, \omega), \omega) - \pi_\nu(q^*(x', \omega), \omega) \right| \leq L'_\mathcal{X} \| x - x' \|^{\gamma/2}. \quad (14)$$

Proof Let $\mathcal{X}$ be as given. For each pair $(x, \omega) \in \mathcal{X} \times \Omega$, there exists $\lambda(x, \omega) \equiv (\lambda^\nu(x, \omega))_{\nu=1}^N$ satisfying the Karush–Kuhn–Tucker conditions of the VI ($\Xi(x, \omega)$, $\Pi(\cdot, \omega)$): for all $\nu = 1, \ldots, N$,

$$0 = \Pi^\nu(q^*(x, \omega), \omega) + (W^\nu)^{\mathrm{T}} \lambda^\nu(x, \omega)$$
$$0 \leq \lambda^\nu(x, \omega) \perp W^\nu q^{\nu,*}(x, \omega) + B^\nu(\omega)x^\nu - h^\nu(\omega) \equiv \mu^\nu(x, \omega) \leq 0. \quad (15)$$

We claim that positive scalars ζ'_1 and ζ'_2 exist such that for each (x, ω), we can choose $\lambda(x, \omega)$ satisfying

$$\| \lambda(x, \omega) \| \leq \zeta'_1 + \zeta'_2 \| q^*(x, \omega) \|^\xi. \quad (16)$$

Indeed, it suffices to choose $\lambda(x, \omega)$ so that, for each ν, the rows of W^ν corresponding to the positive components of $\lambda^\nu(x, \omega)$ are linearly independent. Writing $\mathcal{J}_\nu \equiv \{j : \lambda^\nu(x, \omega)_j > 0\}$, we have, for all $\nu = 1, \ldots, N$,

$$\| \lambda^\nu(x, \omega) \| = \| \lambda^\nu_{\mathcal{J}_\nu}(x, \omega) \| \leq \left\| \left[W^\nu_{\mathcal{J}_\nu \cdot} (W^\nu_{\mathcal{J}_\nu \cdot})^{\mathrm{T}} \right]^{-1} W^\nu_{\mathcal{J}_\nu \cdot} \right\| \| \Pi^\nu(q^*(x, \omega), \omega) \|$$

$$\leq \left\| \left[W^\nu_{\mathcal{J}_\nu \cdot} (W^\nu_{\mathcal{J}_\nu \cdot})^{\mathrm{T}} \right]^{-1} W^\nu_{\mathcal{J}_\nu \cdot} \right\| \left[\zeta_1 + \zeta_2 \| q^*(x, \omega) \|^\xi \right]$$

from which the existence of ζ_1' and ζ_2' follows readily, due to the fact that there are only finitely many index sets. Next we show that

$$\sup_{(x,\omega)\in\mathcal{X}\times\Omega} \| q^*(x,\omega) \| \; < \; \infty. \tag{17}$$

Pre-multiplying the first equation in (15) by $q^{\nu,*}(x,\omega)^{\mathrm{T}}$, summing over ν, and using the strong monotonicity condition (a), we have

$$
\begin{aligned}
0 = \sum_{\nu=1}^{N} & \left\{ q^{\nu,*}(x,\omega)^{\mathrm{T}} \left[\Pi^{\nu}(q^*(x,\omega),\omega) + (W^{\nu})^{\mathrm{T}}\lambda^{\nu}(x,\omega) \right] \right\} \\
\geq\; & \rho \, \| q^*(x,\omega) - q^*(\bar{x},\omega) \|^2 \\
& - \left[\sup_{\omega\in\Omega} \|q^*(\bar{x},\omega)\| \right] \left[2\zeta_1 + \zeta_2 \left(\|q^*(x,\omega)\|^{\xi} + \sup_{\omega\in\Omega} \|q^*(\bar{x},\omega)\|^{\xi} \right) \right] \\
& - \left(\zeta_1' + \zeta_2'\|q^*(x,\omega)\|^{\xi} \right) \left[\max_{1\leq\nu\leq N}\sup_{\omega\in\Omega} \|B^{\nu}(\omega)\| \, \|x\| + \max_{1\leq\nu\leq N}\sup_{\omega\in\Omega} \| h^{\nu}(\omega)\| \right].
\end{aligned}
$$

By (b), the boundedness of $\mathcal{X}$, and the fact that $\xi \in (0, 2)$, (17) follows readily. Clearly, (17) and (16) imply that for the chosen $\lambda(x,\omega)$, we have $\displaystyle\sup_{(x,\omega)\in\mathcal{X}\times\Omega} \|\lambda(x,\omega)\| < \infty$. From (15), we have, for any $(x, x', \omega) \in \mathcal{X} \times \mathcal{X} \times \Omega$,

$$
\begin{aligned}
0 = \sum_{\nu=1}^{N} & \left(q^{\nu,*}(x,\omega) - q^{\nu,*}(x',\omega) \right)^{\mathrm{T}} \left[\Pi^{\nu}(q^*(x,\omega),\omega) - \Pi^{\nu}(q^*(x',\omega),\omega) \right] \\
& + \sum_{\nu=1}^{N} \left[W^{\nu} \left(q^{\nu,*}(x,\omega) - q^{\nu,*}(x',\omega) \right) \right]^{\mathrm{T}} \left(\lambda^{\nu}(x,\omega) - \lambda^{\nu}(x',\omega) \right) \\
\geq\; & \rho \, \| q^*(x,\omega) - q^*(x',\omega) \|^2 \\
& - \max_{1\leq\nu\leq N}\sup_{\omega\in\Omega} \| B^{\nu}(\omega) \| \, \| x - x' \| \, \| \lambda(x,\omega) - \lambda(x',\omega) \|.
\end{aligned}
$$

Since $\lambda(x,\omega)$ and $\lambda(x',\omega)$ are bounded, (13) follows readily. It is clear that (14) follows from (13) and (12).   □

Remark 4 It would be useful to explain a bit about the assumptions in Proposition 5, which will be verified subsequently for the family of power price functions; see Sect. 5.1. Here, we make some general remarks about these assumptions. First of all, it is not uncommon for each set X^{ν} to be bounded in practice as capacity is limited. Nevertheless, the main point of the proposition is the uniformity in ω in the two inequalities (13) and (14). The boundedness of X is not sufficient for such uniformity to hold. As already noted, the existence and uniqueness of $q^*(x,\omega)$ holds if $\Pi(\cdot,\omega)$ is strongly monotone; assumption (a) postulates essentially that such strong monotonicity is uniform in ω. Assumption (B) holds if the functions B^{ν} and h^{ν} are continuous and Ω is compact. (A referee has suggested that this explicit boundedness assumption

could be weakened to a condition of "boundedness by integrable functions"; while we agree that this could be done, we employ (B) in order to partially avoid the technical complexity induced by weakened condition and also since it may easily be justified practically.) The compactness of Ω also provides a sufficient condition for (b) to hold, provided that the solution function $q^*(\bar{x}, \cdot)$ is continuous for some $\bar{x}$. Finally, assumption (c) induces some technical conditions on the profit functions $\pi_v(\cdot, \omega)$. All these assumptions (except (B)) will be verified for the special case of isoelastic price functions; see Lemma 1. $\qquad\square$

Theorem 6 below establishes the almost sure convergence of the sample average of the players' profits, from which the almost sure convergence of the Nash equilibria of the sampled games to a Nash equilibrium of the stochastic capacity game follows readily.

Theorem 6 *Assume that conditions (A) and (B) in Proposition 5 hold, and that each function $\eta_v(x^v)$ is continuous. Let $\{\omega^i\}$ be a sequence of independently and identically distributed samples with compact support Ω. The following two statements are valid.*

(I) *For every sequence $\left\{x^\ell \equiv (x^{v,\ell})_{v=1}^N\right\} \subset \mathcal{X}$ converging to $x^\infty \equiv (x^{v,\infty})_{v=1}^N$ in $\mathcal{X}$, where $\mathcal{X}$ is a bounded subset of X that satisfies conditions (a)–(c) in Proposition 5, it holds that*

$$\lim_{\ell \to \infty} \frac{1}{\ell} \sum_{i=1}^{\ell} \pi_v(q^*(x^\ell, \omega^i), \omega^i) \;=\; \mathbb{E}\pi_v(q^*(x^\infty, \omega), \omega) \quad \text{almost surely.}$$

(II) *For each positive integer ℓ, let $\widehat{x}^\ell \equiv (\widehat{x}^{v,\ell})_{v=1}^N$ be an equilibrium solution to the approximated Nash capacity game where the individual player's optimization problem is given by (11). If $\{\widehat{x}^\ell\}$ is contained in a bounded subset $\mathcal{X}$ of X that satisfies conditions (a)–(c) in Proposition 5, then, almost surely, every accumulation point of $\{\widehat{x}^\ell\}$ is a Nash equilibrium solution of the two-stage capacity game.*

Proof Condition (c) and (17) imply that, for all $v = 1, \ldots, N$,

$$\sup_{(x,\omega) \in \mathcal{X} \times \Omega} \left| \pi_v(q^*(x, \omega), \omega) \right| \;<\; \infty.$$

This implies that $\pi_v(q^*(x, \omega), \omega)$ has a finite variance that is bounded independently of $x \in \mathcal{X}$. Let $\{x^\ell\} \subset \mathcal{X}$ be a sequence converging to $x^\infty \in \mathcal{X}$. Since $\pi_v(q^*(x^\infty, \omega^i), \omega^i)$ are independent and identically distributed samples of $\pi_v(q^*(x^\infty, \omega), \omega)$, it follows from the strong law of large numbers that, for every $v = 1, \ldots, N$,

$$\lim_{\ell \to \infty} \frac{1}{\ell} \sum_{i=1}^{\ell} \left[\pi_v(q^*(x^\infty, \omega^i), \omega^i) - \mathbb{E}\pi_v(q^*(x^\infty, \omega), \omega) \right] = 0 \quad \text{almost surely.}$$

We can write

$$\pi_\nu(q^*(x^\ell, \omega^i), \omega^i) - \mathbb{E}\pi_\nu(q^*(x^\infty, \omega), \omega)$$
$$= \left[\pi_\nu(q^*(x^\ell, \omega^i), \omega^i) - \pi_\nu(q^*(x^\infty, \omega^i), \omega^i) \right]$$
$$+ \left[\pi_\nu(q^*(x^\infty, \omega^i), \omega^i) - \mathbb{E}\pi_\nu(q^*(x^\infty, \omega), \omega) \right].$$

By Proposition 5, positive constants L' and γ' exist such that, for all ω,

$$\left| \pi_\nu(q^*(x^\ell, \omega), \omega) - \pi_\nu(q^*(x^\infty, \omega), \omega) \right| \leq L' \| x^\ell - x^\infty \|^{\gamma'}.$$

Thus the limit in (I) follows. Hence, for every $\nu = 1, \ldots, N$, the sequence of random functions $\{\theta_{\nu,\ell}(x)\}_{\ell=1}^\infty$ converges continuously in x to $\theta_\nu(x)$ almost surely. Statement (II) follows from Theorems 1 and 2(B)–(D).   □

Remark 5 As it is evident from the above proof, a key conclusion that we need from Proposition 5 is the convergence of the sequence $\{\pi_\nu(q^*(x^\ell, \omega), \omega)\}$ to $\pi_\nu(q^*(x^\infty, \omega), \omega)$ uniformly on compact sets almost surely as $\ell \to \infty$. Via the Hölderian inequalities (13) and (14), the setting of Proposition 5 is sufficient for this purpose. More generally, Chap. 6 in [31], particularly Proposition 7 therein, and the paper [32, Sect. 4] address the kind of convergence in part (I) of Theorem 6 under broad conditions that are much weaker than those in Proposition 5. Nevertheless, the latter proposition is useful as an intermediate result for practical application, as illustrated in the next subsection.   □

5.1 Power price functions

Finally, we present an application of Theorem 6 by considering the random isoelastic price function $p_j(Q_j, \omega)$ of the type:

$$p_j(Q_j, \omega) = a_j(\omega) \, Q_j^{-1/b_j(\omega)}, \quad Q_j > 0, \tag{18}$$

for appropriate positive random coefficients a_j and b_j. We first establish a lemma summarizing two properties of such a function.

Lemma 1 *Suppose that each function $c_j^\nu(q^\nu)$ is convex and twice continuously differentiable, and for some positive constants $\gamma_c < 2$, $\beta_{1,c}$, and $\beta_{2,c}$,*

$$\| \nabla c_j^\nu(q^\nu) \| \leq \beta_{1,c} + \beta_{2,c} \| q^\nu \|^{\gamma_c} \quad \forall \nu = 1, \ldots, N \text{ and } j \in \mathcal{N}. \tag{19}$$

If the random variables $a_j(\omega)$ and $b_j(\omega)$ are such that, for all $j \in \mathcal{N}$,

$$0 < \inf_{\omega \in \Omega} a_j(\omega) \leq \sup_{\omega \in \Omega} a_j(\omega) < \infty, \tag{20}$$

and

$$1 < \inf_{\omega \in \Omega} b_j(\omega) \leq \sup_{\omega \in \Omega} b_j(\omega) < \infty, \tag{21}$$

then, for every convex set $\mathcal{Q}$ in which $q \geq 0$,

$$\inf_{q \in \mathcal{Q}} \min_{j \in \mathcal{N}} Q_j > 0,$$

and each $c_k^\nu(q^\nu)$ is Lipschitz continuous, the following two statements hold:

(a′) *a constant $\rho_{\mathcal{Q}} > 0$ exists such that, for all $(q, q', \omega) \in \mathcal{Q} \times \mathcal{Q} \times \Omega$,*

$$(q - q')^{\mathrm{T}}(\Pi(q, \omega) - \Pi(q', \omega)) \geq \rho_{\mathcal{Q}} \|q - q'\|^2,$$

(c′) *positive constants $\beta_{1,q}$, $\beta_{2,q}$, and $L_{\mathcal{Q}}$ exist such that*

$$\sup_{(q,\omega) \in \mathcal{Q} \times \Omega} \|\Pi(q, \omega)\| < \beta_{1,q} + \beta_{2,q} \|q\|^{\gamma_c},$$

and, for all $\nu = 1, \ldots, N$,

$$\left| \pi_\nu(q, \omega) - \pi_\nu(q', \omega) \right| \leq L_{\mathcal{Q}} \|q - q'\| \quad \forall (q, q', \omega) \in \mathcal{Q} \times \mathcal{Q} \times \Omega. \tag{22}$$

Proof For (a′), it suffices to show that

$$\inf_{(q,\omega) \in \mathcal{Q} \times \Omega} \min_{\|z\|=1} z^{\mathrm{T}} J_q \Pi(q, \omega) z > 0$$

since the uniform positive definiteness of partial Jacobian matrix $J_q \Pi(q, \omega)$ guarantees the strong monotonicity of $\Pi(q, \omega)$ in q uniformly for all ω. In general, letting $r_\nu(q, \omega) \equiv \sum_{j \in \mathcal{N}} q_j^\nu p_j(Q_j, \omega)$ be firm ν's revenue function and recalling $Q_j = \sum_{\nu=1}^N q_j^\nu$, we deduce $\frac{\partial r_\nu(q, \omega)}{\partial q_j^\nu} = p_j(Q_j, \omega) + \frac{\partial p_j(Q_j, \omega)}{\partial Q_j} q_j^\nu$, and

$$\frac{\partial}{\partial q_k^{\nu'}} \left(\frac{\partial r_\nu(q, \omega)}{\partial q_j^\nu} \right) = \begin{cases} 0 & \text{if } k \neq j \\[2ex] \dfrac{\partial p_j(Q_j, \omega)}{\partial Q_j} + \dfrac{\partial^2 p_j(Q_j, \omega)}{\partial Q_j^2} q_j^\nu & \text{if } k = j \text{ and } \nu' \neq \nu \\[2ex] 2 \dfrac{\partial p_j(Q_j, \omega)}{\partial Q_j} + \dfrac{\partial^2 p_j(Q_j, \omega)}{\partial Q_j^2} q_j^\nu & \text{if } k = j \text{ and } \nu' = \nu. \end{cases}$$

Therefore, with $z \equiv (z^\nu)_{\nu=1}^N$ and $z^\nu \equiv (z_j^\nu)_{j \in \mathcal{N}}$, it follows that

$$
z^{\mathrm{T}} J_q \Pi(q, \omega) z = \sum_{\nu=1}^N (z^\nu)^{\mathrm{T}} \left[\sum_{j \in \mathcal{N}} \nabla^2 c_j^\nu(q^\nu) \right] z^\nu
$$

$$
- \sum_{j \in \mathcal{N}} \frac{\partial p_j(Q_j, \omega)}{\partial Q_j} \left[\sum_{\nu=1}^N (z_j^\nu)^2 + \left(\sum_{\nu=1}^N z_j^\nu \right)^2 \right]
$$

$$
- \sum_{j \in \mathcal{N}} \frac{\partial^2 p_j(Q_j, \omega)}{\partial Q_j^2} \left(\sum_{\nu=1}^N q_j^\nu z_j^\nu \right) \left(\sum_{\nu=1}^N z_j^\nu \right).
$$

With $p_j(Q_j, \omega)$ given by (18), we have $\frac{\partial p_j(Q_j, \omega)}{\partial Q_j} = -\frac{a_j(\omega)}{b_j(\omega)} Q_j^{-1-1/b_j(\omega)}$, and $\frac{\partial^2 p_j(Q_j, \omega)}{\partial Q_j^2} = \frac{a_j(\omega)}{b_j(\omega)} \left(1 + \frac{1}{b_j(\omega)}\right) Q_j^{-2-1/b_j(\omega)}$, which yield

$$
z^{\mathrm{T}} J_q \Pi(q, \omega) z = \sum_{\nu=1}^N (z^\nu)^{\mathrm{T}} \left[\sum_{j \in \mathcal{N}} \nabla^2 c_j^\nu(q^\nu) \right] z^\nu
$$

$$
+ \sum_{j \in \mathcal{N}} \frac{a_j(\omega)}{b_j(\omega)} Q_j^{-1-1/b_j(\omega)} \left[\sum_{\nu=1}^N (z_j^\nu)^2 + \left(\sum_{\nu=1}^N z_j^\nu \right)^2 \right]
$$

$$
- \sum_{j \in \mathcal{N}} \frac{a_j(\omega)}{b_j(\omega)} Q_j^{-2-1/b_j(\omega)} \left(1 + \frac{1}{b_j(\omega)}\right) \left(\sum_{\nu=1}^N q_j^\nu z_j^\nu \right) \left(\sum_{\nu=1}^N z_j^\nu \right).
$$

Since

$$
\left(\sum_{\nu=1}^N q_j^\nu \right) \left[\sum_{\nu=1}^N (z_j^\nu)^2 + \left(\sum_{\nu=1}^N z_j^\nu \right)^2 \right] - \left(1 + \frac{1}{b_j(\omega)}\right) \left(\sum_{\nu=1}^N q_j^\nu z_j^\nu \right) \left(\sum_{\nu=1}^N z_j^\nu \right)
$$

$$
\geq Q_j \left[1 - \frac{1}{4} \left(1 + \frac{1}{b_j(\omega)}\right)^2 \right] \sum_{\nu=1}^N (z_j^\nu)^2, \quad \text{because } b_j(\omega) > 1,
$$

it follows that,

$$
z^{\mathrm{T}} J_q \Pi(q, \omega) z \geq \sum_{j \in \mathcal{N}} \frac{a_j(\omega)}{b_j(\omega)} Q_j^{-1-1/b_j(\omega)} \left[1 - \frac{1}{4} \left(1 + \frac{1}{b_j(\omega)}\right)^2 \right] \sum_{\nu=1}^N (z_j^\nu)^2
$$

$$
\geq \left\{ \inf_{j \in \mathcal{N}} \frac{a_j(\omega)}{b_j(\omega)} Q_j^{-1-1/b_j(\omega)} \left[1 - \frac{1}{4} \left(1 + \frac{1}{b_j(\omega)}\right)^2 \right] \right\} \| z \|^2.
$$

This establishes (a′). Since

$$
\Pi_j^v(q, \omega) \;=\; a_j(\omega)\, Q_j^{-1/b_j(\omega)}\left[\,1 - \frac{1}{b_j(\omega)}\,\frac{q_j^v}{Q_j}\,\right] + \frac{\partial c_j^v(q^v)}{\partial q_j^v}, \quad \text{and}
$$

$$
\sup_{(q,\omega)\in\mathcal{Q}\times\Omega}\left\{\, a_j(\omega)\, Q_j^{-1/b_j(\omega)}\left[\,1 - \frac{1}{b_j(\omega)}\,\frac{q_j^v}{Q_j}\,\right]\right\} \;<\; \infty,
$$

the existence of the constants $\beta_{1,q}$ and $\beta_{2,q}$ satisfying (c′) follows from the assumption on the functions $c_j^v(q^v)$; cf. (19). We have

$$
\pi_v(q, \omega) - \pi_v(q', \omega) \;=\; \sum_{j\in\mathcal{N}}\Big[\Big(p_j(Q_j,\omega)q_j^v - p_j(Q_j',\omega)q_j^{v,\prime}\Big) \\
-(c_j^v(q^v) - c_j^v(q^{v,\prime}))\Big].
$$

Since Q_j is bounded away from zero on $\mathcal{Q}$, it follows that $p_j(\cdot, \omega)$ is Lipschitz continuous on $\mathcal{Q}$, with a Lipschitz constant independent of $\omega \in \Omega$. Finally, since each $c_j^v(q^v)$ is Lipschitz continuous on $\mathcal{Q}$ and $\sup_{(q,\omega)\in\mathcal{Q}\times\Omega} p_j(Q_j,\omega) < \infty$ for all $j \in \mathcal{N}$, the Lipschitz condition (22) follows readily.   □

Note that (a′) of Lemma 1 guarantees condition (A) of Proposition 5. Therefore, based on the above calculations, we deduce the following convergence result for the two-stage Nash capacity expansion game with price functions given by (18). The proof follows readily from the above analysis and the details are omitted.

Corollary 4 *Let each $p_j(Q, \omega)$ be given by (18) where the coefficients satisfy (20) and (21). Suppose that each function $c_j^v(q^v)$ is convex, twice continuously differentiable, and globally Lipschitz continuous, that condition (B) in Proposition 5 hold, that each function $\eta_v(x^v)$ is continuous, and that $\{q^*(x, \omega) : (x, \omega) \in X \times \Omega\}$ is contained in a set $\mathcal{Q}$ in which $q \geq 0$ and*

$$
\inf_{q\in\mathcal{Q}}\ \min_{j\in\mathcal{N}} Q_j \;>\; 0. \tag{23}
$$

Let $\{\omega^i\}$ be a sequence of independently and identically distributed samples with compact support Ω. For each positive integer ℓ, let $\widehat{x}^\ell \equiv (\widehat{x}^{v,\ell})_{v=1}^N$ be an equilibrium solution to the approximated Nash capacity game where the individual player's optimization problem is given by (11). If $\{\widehat{x}^\ell\}$ is bounded, then, almost surely, every accumulation point of $\{\widehat{x}^\ell\}$ is a Nash equilibrium solution of the two-stage capacity game.   □

We finish the paper by noting that condition (23) is sometimes called "industry boundedness". The condition was used in [19] to show the existence of a Nash–Cournot equilibrium; similarly (20) and (21) are standard assumptions to ensure the good behavior of the regional isoelastic price functions.

References

1. Attouch, H.: Variational Convergence for Functions and Operators. Pitman Press, Boston (1984)
2. Attouch, H., Wets, R.J.B.: Approximation and convergence in nonlinear optimization. In: Mangasarian, O.L., Meyer, R.R., Robinson, S.M. (eds.) Nonlinear Programming 4, pp. 367–394. Academic, New York (1981)
3. Attouch, H., Wets, R.J.B.: Convergence des points min/sup et de points fixes. Computes Rendus de l'Académie Des Sciences de Paris **296**, 657–660 (1983)
4. Attouch, H., Wets, R.J.B.: A convergence theory for saddle functions. Trans. Am. Math. Soc. **280**, 1–41 (1983)
5. Berridge, S., Krawczyk, J.B.: Relaxation algorithms in finding Nash equilibria. Economics Working Paper Archive (1997). [URL: http://econwpa.wustl.edu/eprints/comp/papers/9707/0707002.abs.]
6. Birge, J.R., Louveaux, F.: Introduction to Stochastic Programming. Springer, New York (1997)
7. Clarke, F.H.: Optimization and Nonsmooth Analysis.. Wiley, New York (1983)
8. Escobar Castro, J.F.: Convergencia Variacional en Teoría de Juegos: Existencia, Estabilidad y Aplicaciones a Mercados Eléctricos. Memoria Para Optar a Título de Ingeniero Civil Matemático. Universidad de Chile, Santiago (2003)
9. Escobar J.F.: A continuity result for discontinuous games. Department of Economics. Stanford University, Palo Alto, CA (2005, Preprint)
10. Facchinei, F., Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer, New York (2003)
11. Fukushima, M.: Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems. Math. Program. **53**, 99–110 (1992)
12. Fukushima, M.: Merit functions for variational inequality and complementarity problems. In: Di Pillo, G., Giannessi, F. (eds) Nonlinear Optimization and Applications. Plenum, New York, pp. 155–170 (1996)
13. Gürkan, G., Özge, A.Y., Robinson, S.M.: Sample-path solution of stochastic variational inequalities. Math. Program. **84**, 313–333 (1999)
14. Higle, J.L., Sen, S.: Stochastic Decomposition: A Statistical Method for Large-Scale Stochastic Linear Programming. Kluwer, Dordrecht (1996)
15. Jofré, A., Wets, R.J.B.: Continuity properties of Walras equilibrium points. Ann. Oper. Res. **114**, 229–243 (2002)
16. Jofré, A., Wets, R.J.B.: Variational bivariate functions: convergence and applications. Manuscript, Department of Mathematics. University of California, Davis, CA (2004)
17. Kall, P.: Approximation to optimization problems: an elementary review. Math. Oper. Res. **11**, 9–18 (1986)
18. Kall, P., Wallace, S.W.: Stochastic Programming. Wiley, Chichester (1994)
19. Kolstad, C.D., Mathiesen, L.: Computing Nash–Cournot equilibria. Oper. Res. **39**, 739–748 (1991)
20. Krawczyk, J.B.: Coupled constraint Nash equilibria in environmental games. Resour. Energy Econ. **27**, 157–181 (2005)
21. Krawczyk, J.B., Uryasev, S.: Relaxation algorithms to find Nash equilibria with economic applications. Environ. Model. Assess. **5**, 63–73 (2000)
22. Levy, A.B.: Convergence of successive approximation methods with parameter target sets. Math. Oper. Res. **30**, 765–784 (2005)
23. Levy, A.B., Poliquin, R., Thibault, L.: Partial extensions of Attouch's theorem with applications to proto-derivatives of subgradient mappings. Trans. Am. Math. Soc. **347**, 1260–1294 (1995)
24. Nash, J.F.: Non-cooperative games. Ann. Math. **54**, 286–295 (1951)

25. Nikaido, H., Isoda, K.: Note on noncooperative convex games. Pac. J. Math. **5**(Suppl. 1), 807–815 (1955)
26. Plambeck, E.L., Fu, B.R., Robinson, S.M., Suri, R.: Sample-path optimization of convex stochastic performance functions. Math. Program. **75**, 137–176 (1996)
27. Poliquin, R.: An extension of Attouch's theorem and its applicataion to second-order epi-differentiation of convexly composite functions. Trans. Am. Math. Soc. **332**, 861–874 (1992)
28. Robinson, S.M.: Analysis of sample-path optimization. Math. Oper. Res. **21**, 513–528 (1996)
29. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Springer, Berlin (1998)
30. Rosen, J.B.: Existence and uniqueness for concave n-person games. Econometrica **33**, 520–534 (1965)
31. Ruszczyński, A., Shapiro, A. (eds.) Stochastic Programming. Handbook in Operations Research and Management Science, vol. 10. North-Holland, The Netherlands (2003)
32. Shapiro, A., Xu, H.: Stochastic mathematical programs with equilibrium constraints, modeling and sample average approximation. (URL: http://www.optimization-online.org/DB_HTML/2005/01/1046.html) (revised June 2005)
33. Uryasev, S., Rubinstein, R.T.: On relaxation algorithms in computation of Nash equilibria. IEEE Trans. Automat. Control **39**, 1263–1267 (1994)
34. Wets, R.J.B: Convergence of convex functions, variational inequalities, and convex optimization problems. In: Cottle, R.W., Gianessi, F., Lions, J.L. (eds.) Variational Inequalities and Complementarity Problems., pp. 375–403. Wiley, Chichester (1980)