Math. Program., Ser. B (2019) 174:99–127
https://doi.org/10.1007/s10107-018-1279-z

FULL LENGTH PAPER

Generalized conditioning based approaches to computing confidence intervals for solutions to stochastic variational inequalities

Michael Lamm[1] · **Shu Lu**[2]

Received: 29 September 2016 / Accepted: 13 April 2018 / Published online: 4 May 2018
© Springer-Verlag GmbH Germany, part of Springer Nature and Mathematical Optimization Society 2018

Abstract Stochastic variational inequalities (SVI) provide a unified framework for the study of a general class of nonlinear optimization and Nash-type equilibrium problems with uncertain model data. Often the true solution to an SVI cannot be found directly and must be approximated. This paper considers the use of a sample average approximation (SAA), and proposes a new method to compute confidence intervals for individual components of the true SVI solution based on the asymptotic distribution of SAA solutions. We estimate the asymptotic distribution based on one SAA solution instead of generating multiple SAA solutions, and can handle inequality constraints without requiring the strict complementarity condition in the standard nonlinear programming setting. The method in this paper uses the confidence regions to guide the selection of a single piece of a piecewise linear function that governs the asymptotic distribution of SAA solutions, and does not rely on convergence rates of the SAA solutions in probability. It also provides options to control the computation procedure and investigate effects of certain key estimates on the intervals.

Mathematics Subject Classification 90C33 · 90C15 · 65K10 · 62F25

✉ Michael Lamm
Michael.Lamm@sas.com

Shu Lu
shulu@email.unc.edu

[1] SAS Institute Inc, SAS Campus Drive, Cary, NC 27513, USA

[2] Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, 355 Hanes Hall, CB#3260, Chapel Hill, NC 27599-3260, USA

1 Introduction

Variational inequalities (VI) have been used as a tool to study the mathematical structure and develop algorithms for a general class of nonlinear optimization and Nash-type equilibrium problems. For a comprehensive treatment on this subject see the book [11]. For the various applications of variational inequalities, see [3,13,14,17,18,21,35] and references therein. A solution to the variational inequality defined by a closed convex set $S \subset \mathbb{R}^n$ and a function $f : S \to \mathbb{R}^n$ is a point $x \in S$ that satisfies the inclusion $-f(x) \in \mathbf{N}_S(x)$, where $\mathbf{N}_S(x) \subset \mathbb{R}^n$ is the normal cone to S at x:

$$\mathbf{N}_S(x) = \{v \in \mathbb{R}^n \mid \langle v, s - x\rangle \leq 0 \text{ for each } s \in S\}.$$

For many problems of interest, data defining the problem are subject to uncertainty. Several formulations of stochastic variational inequalities (SVI) have been proposed to incorporate uncertain model data into the VI framework, including the expected value formulation considered in [20,24,25,46,52] , the expected residual minimization formulation studied in [1,4,6,7,12,33,54], and two-stage or multistage SVI formulations [5,42].

In this paper we consider the expected value formulation of an SVI. To give its definition, let $(\Omega, \mathscr{F}, P)$ be a probability space, and ξ be a d-dimensional random vector defined on Ω and supported on a closed subset Ξ of $\mathbb{R}^d$. Let O be an open subset of $\mathbb{R}^n$ and F be a measurable function from $O \times \Xi$ to $\mathbb{R}^n$, such that for each $x \in O$ the expectation $E\|F(x, \xi)\| < \infty$. Let $f_0(x) = E[F(x, \xi)]$ for each $x \in O$. The SVI problem is to find a point $x \in S \cap O$ that satisfies

$$0 \in f_0(x) + \mathbf{N}_S(x). \tag{1}$$

The SVI in the formulation (1) represents the first order necessary conditions of a large class of single stage stochastic optimization and stochastic equilibrium problems. This SVI formulation is closely related to statistical estimation such as M-estimation and has been used to study statistical inference for sparse penalized regression; see [23,25,32,49,51,53] and references therein.

Throughout this paper we assume that the set S defining the SVI (1) is a polyhedral convex set in $\mathbb{R}^n$. A solution to (1) will be referred to as a true solution and will be denoted by x_0. The expectation used to define the function f_0 often cannot be easily evaluated, and the true solution x_0 must be approximated. This paper considers the use of a sample average approximation (SAA) of (1) defined formally in Sect. 2, and the computation of confidence intervals for individual components of the true solution based on the solution to an SAA problem.

1.1 Motivation and related work

For the extensive study on stability and inference for stochastic optimization and stochastic variational inequalities, see [2,8,10,20,25,27,28,36,43,45–47,50,52] and references therein. Our work aims to complement the existing literature with the fol-

lowing features: (a) our interest is in confidence regions and intervals for the true SVI solution; (b) we estimate the asymptotic distribution based on a SAA solution without generating multiple SAA solutions; (c) we can handle inequality constraints and allow the true solution to be degenerate in the sense that the corresponding normal map solution of (1), defined formally in Sect. 2, is allowed to lie on the boundary of a cell in the normal manifold, in other words, the true solution does not need to satisfy strict complementarity in the standard nonlinear programming setting.

A general polyhedral convex set, if not affine, has a nonsmooth boundary that includes all of its proper faces. A consequence of such nonsmoothness is a piecewise linear structure in the asymptotic distribution of solutions to the SAA problems when the true solution is degenerate in the sense described above. Traditional approaches to performing inference are not applicable due to this piecewise structure, and alternative techniques are needed for developing and analyzing new inferential methods.

In [26,29–31] the authors have applied sensitivity analysis of constrained optimization and variational inequalities to exploit properties of the piecewise linear structure to construct asymptotically exact confidence regions and intervals. In [26], methods were proposed for computing generally valid individual confidence intervals for solutions to the SVI and its normal map formulation. These methods addressed the challenges posed by the piecewise structure of the SAA solutions' asymptotic distribution by using a conditioning based approach described in Sect. 2. However, the methods developed in [26] face two limitations. First, they require exponential convergence rates of the SAA solutions to obtain a consistent estimator of a certain piecewise linear function needed to compute the confidence intervals. Second, the computation of the confidence intervals can be sensitive to misestimation of the piecewise structure of the asymptotic distribution. Addressing these limitations serves as the primary motivation for the methods proposed in this paper.

1.2 Contributions

The methods proposed in this paper generalize the conditioning based approaches for computing individual confidence intervals proposed in [26]. Instead of working with an estimation of a global piecewise linear function as in [26], in this paper we directly target the estimation at a single piece of that piecewise linear function. Moreover, the methods that we propose are designed to allow for separate consideration of properties of the estimates for the cone corresponding to the selection function and the matrix representation of the selection function.

By allowing for a separate consideration of properties of these estimates the proposed methods can incorporate a greater level of flexibility. As a result, instead of requiring the estimate of the selection function to be consistent, we allow an incorrect estimation to occur under a certain probability and then adjust for the potential misestimation during the interval computation stage. The new methods in this paper therefore do not rely on exponential convergence rates of the SAA solutions. The increased flexibilities can also be exploited to generate a family of confidence intervals to account for different possibilities for the cone and the potentially piecewise structure and to investigate the sensitivity of the intervals to certain estimates. These

options are especially beneficial at moderate sample sizes when certain asymptotic properties of the estimates may not be sufficiently realized.

The proposed methods are justified with weak convergence results in Theorems 1 and 2. Key to implementing these results is the ability to estimate from the sample data a cone, or family of cones, that satisfies the conditions of Theorems 1 and 2. In Theorem 3 we show how such a cone can be obtained by using confidence regions for the true solution proposed in [30]. Combined, these results allow the methods developed in this paper to overcome the limitations of the methods proposed in [26].

The remainder of this paper is organized as follows. In Sect. 2 we formally introduce the SAA formulation of (1), the normal map formulation of a variational inequality, and review previous work on the computation of confidence sets for the true solution. Section 3 introduces the new methods for computing individual confidence intervals and contains the main theoretical results of this paper. In Sect. 4 we apply the proposed methods to a numerical example. For this example we are able to compute the true solution and evaluate the performance of the proposed methods; we observe that both the interval coverage rates and the interval lengths are in line with our expectations.

2 Background

In this section we formally define the SAA problem, introduce the normal map formulation of variational inequalities, discuss pertinent properties of piecewise linear functions, and review previous work on the construction of confidence sets for the true solution. Throughout, we use $\mathcal{N}(\mu, \Sigma)$ to denote a normal random vector with mean μ and covariance matrix Σ, $Z \sim \mathcal{N}(0, I_n)$ to denote a standard normal random vector, and $Y_n \Rightarrow Y$ to denote the weak convergence of random variables Y_n to Y. The probability of an event A is denoted by $\Pr(A)$ and the $(1-\alpha) * 100\%$ percentile of a χ^2 random variable with l degrees of freedom is denoted by $\chi^2_l(\alpha)$. We denote the jth component of a vector x by $(x)_j$ and the jth row of a matrix M by $(M)_j$. For a set $C \subset \mathbb{R}^n$, we use $\text{cone}(C)$ to denote the cone generated by C: that is, the set $\{\lambda x \mid \lambda \geq 0, x \in C\}$. We use riC to denote the relative interior of C. Finally, $\|\cdot\|$ will denote the norm of an element in a normed space; unless a specific norm is stated it can be any norm, as long as the same norm is used in all related contexts.

The SAA problem considers a variational inequality where the function f_0 in (1) is estimated with a sample average function. Let $\xi^1, \ldots, \xi^N$ be independent and identically distributed (i.i.d) random variables with the same distribution as that of ξ. The sample average function $f_N(x, \omega) : O \times \Omega \to \mathbb{R}^n$ is defined to be

$$f_N(x, \omega) = N^{-1} \sum_{i=1}^{N} F(x, \xi^i(\omega)). \tag{2}$$

The sample average function gives rise to the SAA problem of finding a point $x \in S \cap O$ such that

$$0 \in f_N(x) + \mathbf{N}_S(x). \tag{3}$$

Throughout this paper, we shall refer to a point that satisfies (3) as an SAA solution and will denote it by x_N. Conditions under which the SAA solutions are known to be consistent estimators of x_0 have been well studied, see [10,19,20,25,46]. In addition to consistency, results about the exponential rate of convergence of SAA solutions in probability were obtained in [52]; see also [48].

A tool we shall use extensively for analyzing the asymptotic behavior of SAA solutions and for the computation of confidence intervals is the normal map formulation of a variational inequality. Taking f_0, S and O as above, the *normal map* [11,37,40,44] induced by f_0 and S is a function $f^{\mathrm{nor}}_{0,S} : \Pi_S^{-1}(O) \to \mathbb{R}^n$ defined as

$$f^{\mathrm{nor}}_{0,S}(z) = f_0 \circ \Pi_S(z) + z - \Pi_S(z).$$

Here Π_S denotes the Euclidian projector onto the set S and $f_0 \circ \Pi_S$ is the composite function of f_0 and Π_S. A point $x \in O \cap S$ is a solution to the SVI (1) only if the point $z = x - f_0(x)$ satisfies

$$f^{\mathrm{nor}}_{0,S}(z) = 0. \tag{4}$$

When the above equality is satisfied, one also has $\Pi_S(z) = x$. We refer to (4) as the normal map formulation of (1) and denote its solution by z_0. Similarly, we shall use $f^{\mathrm{nor}}_{N,S}$ to denote the normal map induced by f_N and S and denote the solution to the normal map formulation of (3) by z_N.

Since S is a polyhedral convex set, the Euclidian projector Π_S is a piecewise affine function. For a piecewise affine function ψ we denote its selection functions by ψ_j, $j = 1, \ldots, k$. Every piecewise affine function admits a corresponding polyhedral subdivision $\Gamma = \{P_1, \ldots, P_k\}$ [44, Proposition 2.2.3]. Such a subdivision is a finite collection of n-dimensional polyhedral sets whose union equals $\mathbb{R}^n$ while the intersection of any distinct sets is either empty or a proper face of the sets. When each set P_j is a polyhedral cone the subdivision is referred to as a conical subdivision. A polyhedral subdivision Γ corresponds to the piecewise affine function ψ if the restriction of ψ to each set $P_j \in \Gamma$ agrees with its selection function ψ_j. We will denote the selection function of ψ corresponding to the set $P_j \in \Gamma$ by $\psi|_{P_j}$. When each selection function ψ_j is linear, ψ is piecewise linear and the corresponding subdivision is conical.

For the Euclidian projector onto a polyhedral convex set S the corresponding polyhedral subdivision can be obtained from the n-cells in the *normal manifold* of S. The n-cells in the normal manifold of S are constructed from the nonempty faces of S. Let F be a nonempty face of S. Then on the relative interior of F the normal cone to S is a constant cone, denoted as $\mathbf{N}_S(ri\,F)$, and the corresponding n-cell C_F is defined to be $C_F = F + \mathbf{N}_S(ri\,F)$. Each k-dimensional face of an n-cell is called a k-cell in the normal manifold for $k = 0, 1, \ldots, n$. The relative interiors of all cells in the normal manifold of S form a partition of $\mathbb{R}^n$.

A piecewise affine function ψ will not have a linear Fréchet derivative (F-derivative) at points x such that $\psi(x) = \psi_j(x) = \psi_k(x)$ for two distinct selections functions ψ_j and ψ_k. While not F-differentiable at all points, a piecewise affine function will be

B-differentiable everywhere. A function $h : \mathbb{R}^n \to \mathbb{R}^m$ is B-differentiable at a point x if there exists a positive homogeneous function $H : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$h(x + v) = h(x) + H(v) + o(v).$$

The function H is then called the B-derivative of h at x, and we denote it by $dh(x)$. If $dh(x)$ is also linear, then it is the classical F-derivative. For a piecewise affine function h with selection functions h_i and the polyhedral subdivision Γ, we can give a specific representation of $dh(x)$ by writing $\Gamma(x) = \{P_i \in \Gamma | x \in P_i\}$, $I = \{i \mid P_i \in \Gamma(x)\}$ and $\Gamma'(x) = \{K_i = \text{cone}(P_i - x) | i \in I\}$. That is, $\Gamma(x)$ is the collection of elements in Γ that contain x, and $\Gamma'(x)$ is the "globalization" of $\Gamma(x)$ along with a shift of the origin. With this notation, the B-derivative $dh(x)$ is piecewise linear with the family of selection functions given by $\{dh_i(x) | i \in I\}$ and the corresponding conical subdivision given by $\Gamma'(x)$.

For a polyhedral convex set S the form of the B-derivative $d\Pi_S(x)(\cdot)$ is also closely related to the normal manifold of S. First we define the tangent cone to a polyhedral convex set S at $x \in S$ to be

$$T_S(x) = \{v \in \mathbb{R}^n | \quad \text{there exists } t > 0 \text{ such that } x + tv \in S\},$$

and the critical cone to S at a point $z \in \mathbb{R}^n$ to be

$$K(z) = T_S(\Pi_S(z)) \cap \{z - \Pi_S(z)\}^{\perp}.$$

As shown in [39, Corollary 4.5] and [34, Lemma 5], for any point $z \in \mathbb{R}^n$ and any sufficiently small $h \in \mathbb{R}^n$ the equality

$$\Pi_S(z + h) = \Pi_S(z) + \Pi_{K(z)}(h) \tag{5}$$

holds, which implies

$$d\Pi_S(z) = \Pi_{K(z)}. \tag{6}$$

The connection to the normal manifold of S follows from the fact that for all points z in the relative interior of a k-cell the critical cone $K(z)$ is a constant cone; see [31, Theorem 8]. The B-derivative $d\Pi_S(z)(\cdot)$ is therefore the same function for all z in the relative interior of a k-cell. For points z and z' in the relative interior of different k-cells $d\Pi_S(z)(\cdot)$ and $d\Pi_S(z')(\cdot)$ can be quite different. As a result small changes in the choice of z can result in significant changes in the form of $d\Pi_S(z)(\cdot)$ and the B-derivative is not continuous with respect to the point z at which it is taken.

Next, we introduce the two assumptions used in this paper. Note that the derivatives that appear in the assumptions are the classic F-derivatives.

Assumption 1 (a) $E\|F(x, \xi)\|^2 < \infty$ for all $x \in O$.
(b) The map $x \mapsto F(x, \xi(\omega))$ is continuously differentiable on O for a.e. $\omega \in \Omega$.
(c) There exists a square integrable random variable C such that

$$\|F(x, \xi(\omega)) - F(x', \xi(\omega))\| + \|d_x F(x, \xi(\omega)) - d_x F(x', \xi(\omega))\| \leq C(\omega)\|x - x'\|,$$

for a.e. $\omega \in \Omega$.

Assumption 1 implies that f_0 is continuously differentiable on O with $df_0(x) = E[d_x F(x, \xi)]$, see, e.g., [46, Theorem 7.44]. By the chain rule of B-differentiability, $f_{0,S}^{\mathrm{nor}}$ is B-differentiable with its B-derivative at z_0 given by

$$df_{0,S}^{\mathrm{nor}}(z_0)(h) = df_0(x_0) \circ d\Pi_S(z_0)(h) + h - d\Pi_S(z_0)(h). \tag{7}$$

For any nonempty compact subset X of O, let $C^1(X, \mathbb{R}^n)$ be the Banach space of continuously differentiable mappings $f : X \to \mathbb{R}^n$, equipped with the norm

$$\|f\|_{1,X} = \sup_{x \in X} \|f(x)\| + \sup_{x \in X} \|df(x)\|.$$

In addition to ensuring the differentiability of f_N, Assumption 1 also guarantees the almost sure convergence of the sample average function f_N to f_0 as an element of $C^1(X, \mathbb{R}^n)$ [46, Theorem 7.48], as well as the weak convergence of $\sqrt{N}(f_N - f_0)$ in $C^1(X, \mathbb{R}^n)$ as shown in [31, Theorem 5]. In the rest of this paper, let X be a compact neighborhood of x_0 in O.

Assumption 2 Suppose that x_0 solves the variational inequality (1). Let $z_0 = x_0 - f_0(x_0)$, $L = df_0(x_0)$, $K_0 = T_S(x_0) \cap \{z_0 - x_0\}^{\perp}$, and assume that $L_{K_0}^{\mathrm{nor}}$ is a homeomorphism from $\mathbb{R}^n$ to $\mathbb{R}^n$, where $L_{K_0}^{\mathrm{nor}}$ is the normal map induced by L and K_0.

Since x_0 solves (1) and z_0 is defined as $x_0 - f_0(x_0)$, we have $x_0 = \Pi_S(z_0)$. In view of (5), $\Pi_S(z) - x_0$ coincides with $\Pi_{K_0}(z - z_0)$ for z in a neighborhood of z_0. It follows that $L_{K_0}^{\mathrm{nor}}$ is exactly $df_{0,S}^{\mathrm{nor}}(z_0)$, the B-derivative of $f_{0,S}^{\mathrm{nor}}$ at z_0, and that $L_{K_0}^{\mathrm{nor}}$ is a global homeomorphism if and only if $L_S^{\mathrm{nor}}(\cdot)$ is a local homeomorphism at z_0 (see also [40, Theorem 5.2]). Moreover, based on the relation between variational inequalities and normal maps, it can be shown that $L_S^{\mathrm{nor}}(\cdot)$ is a local homeomorphism at z_0 if and only if there exist neighborhoods X of x_0 in $\mathbb{R}^n$ and V of 0 in $\mathbb{R}^n$ such that the following linear variational inequality

$$v \in f_0(x_0) + L(x - x_0) + \mathbf{N}_S(x)$$

has a unique solution in X for any $v \in V$ that is Lipschitz continous with respect to v. The latter condition is known as strong regularity, and is a sufficient condition for the variational inequality (1) to have a locally unique solution in a neighborhood of x_0 under small perturbations of f_0 [38,41]. This sufficient condition is also necessary, when the perturbation to f_0 includes a parameter on the right-hand-side of (1) [9]. In this sense, Assumption 2 is the weakest condition possible to ensure (1) to have a locally unique solution under small perturbations of f_0.

It was shown in [37,40,44] that $L_{K_0}^{\mathrm{nor}}$ is a global homeomorphism if and only if it is *coherent oriented*, meaning that all matrices representing its linear selection functions are of the same nonzero determinantal sign. In particular, if the restriction of L on the linear span of K is positive definite, then $L_{K_0}^{\mathrm{nor}}$ is coherent oriented.

To relate the homeomorphism condition in Assumption 2 with other conditions used in nonlinear programming, consider a nonlinear program in $\mathbb{R}^n$ with p inequality

and q equality constraints, with its objective and constraints being twice continuously differentiable. The corresponding Karush–Kuhn–Tucker (KKT) system can be written as a variational inequality (1) with $S = \mathbb{R}^n \times \mathbb{R}^p_+ \times \mathbb{R}^q$. Suppose that the KKT system has a solution. If that solution satisfies the linear independence constraint qualification (LICQ) and a strong second-order sufficient condition, then the homeomorphism condition holds [9,38]. In particular, Assumption 2 does not require the KKT solution to satisfy the strict complementarity condition, which requires multipliers corresponding to active constraints to be nonzero. Under strict complementarity, the normal map solution z_0 corresponding to the KKT solution lies in the interior of a cell in the normal manifold of S, K_0 is a subspace, $L^{\text{nor}}_{K_0}$ is a linear map, and $f^{\text{nor}}_{0,S}$ is differentiable at z_0. By relaxing strict complementarity, we allow z_0 to be located on the boundary of a cell in the normal manifold of S, $L^{\text{nor}}_{K_0}$ to be piecewise linear, and $f^{\text{nor}}_{0,S}$ to be nondifferentiable (but B-differentiable) at z_0.

Assumptions 1 and 2 are used in [31, Theorem 7] to show that for a.e. $\omega \in \Omega$ there exits an N_ω such that for $N \geq N_\omega$ the SAA problems have locally unique solutions x_N and z_N that converge to x_0 and z_0 respectively, with

$$\sqrt{N}(z_N - z_0) \Rightarrow \left(L^{\text{nor}}_{K_0}\right)^{-1}(Y_0), \tag{8}$$

and

$$\sqrt{N}(\Pi_S(z_N) - \Pi_S(z_0)) \Rightarrow \Pi_{K_0} \circ \left(L^{\text{nor}}_{K_0}\right)^{-1}(Y_0). \tag{9}$$

Here, Y_0 is a normal random vector in $\mathbb{R}^n$ with zero mean and the same covariance matrix as $F(x_0, \xi)$. For related work on the weak convergence of SAA solutions of SVIs and optimal solutions to constrained stochastic optimization problems, see, [10,25,46] and references therein.

The above weak convergence results have been used to obtain asymptotically exact confidence regions for z_0 in [29–31]. Below we review some results that will be used in the next section. Let Σ_0 be the covariance matrix of Y_0 and Σ_N be the sample covariance matrix of $\{F(x_N, \xi^i)\}_{i=1}^N$; then Σ_N converges almost surely to Σ_0 [30, Lemma 3.6]. Following the notation used earlier, let $\Gamma(z_0)$ denote the collection of n-cells in the normal manifold of S that contain z_0. [30, Proposition 3.5] says that under Assumptions 1 and 2,

$$\lim_{N\to\infty} \Pr\left(z_N \in \cup_{P_j \in \Gamma(z_0)} ri(P_j)\right) = 1, \tag{10}$$

which implies that $d\Pi_S(z_N)$ is a linear function with high probability even if $d\Pi_S(z_0)$ is piecewise linear. Accordingly, the B-derivative $df^{\text{nor}}_{N,S}(z_N)$ of $f^{\text{nor}}_{N,S}$ at z_N, defined as

$$df^{\text{nor}}_{N,S}(z_N)(h) = df_N(x_N) \circ d\Pi_S(z_N)(h) + h - d\Pi_S(z_N)(h),$$

is a linear function with high probability and therefore not a consistent estimator of $df^{\text{nor}}_{0,S}(z_0)$ in general.

If Σ_0 is nonsingular, then Σ_N is nonsingular for large N almost surely, and the following set (which is an ellipsoid with high probability)

$$Q_N = \left\{ z \in \mathbb{R}^n \,\middle|\, N \left[df_{N,S}^{\mathrm{nor}}(z_N)(z - z_N) \right]^T \Sigma_N^{-1} \left[df_{N,S}^{\mathrm{nor}}(z_N)(z - z_N) \right] \le \chi_n^2(\alpha) \right\}$$

satisfies $\lim_{N\to\infty} \Pr\{z_0 \in Q_N\} = 1-\alpha$ as shown in [30]. In general, one can decompose Σ_N as

$$\Sigma_N = U_N^T \Delta_N U_N$$

where U_N is an orthogonal $n \times n$ matrix and Δ_N is a diagonal matrix with monotonically decreasing elements. Let $\rho > 0$ be the minimum of all positive eigenvalues of Σ_0 and D_N be the upper-left submatrix of Δ_N whose diagonal elements are at least $\rho/2$. Let l_N be the number of rows in D_N, $(U_N)_1$ be the submatrix of U_N that consists of its first l_N rows, and $(U_N)_2$ be the submatrix that consists of the remaining rows of U_N. It was shown in [30] that for any $\varepsilon > 0$ and integer N, the set $R_{N,\varepsilon}$ defined as

$$R_{N,\varepsilon} = \left\{ z \in \mathbb{R}^n \,\middle|\, \begin{array}{l} N \left[df_{N,S}^{\mathrm{nor}}(z_N)(z - z_N) \right]^T (U_N)_1^T D_N^{-1} (U_N)_1 \left[df_{N,S}^{\mathrm{nor}}(z_N)(z - z_N) \right] \le \chi_{l_N}^2(\alpha) \\ \\ \| \sqrt{N} (U_N)_2 df_{N,S}^{\mathrm{nor}}(z_N)(z - z_N) \|_\infty \le \varepsilon \end{array} \right\} \tag{11}$$

satisfies $\lim_{N\to\infty} \Pr\left\{z_0 \in R_{N,\varepsilon}\right\} = 1 - \alpha$. Since the nonsingular case can be treated as a specialization of the singular case with $l = n$, in the rest of this paper we refer only to the singular case and consider regions $R_{N,\varepsilon}$.

While $df_{N,S}^{\mathrm{nor}}(z_N)$ can be used to obtain asymptotically exact confidence regions as above, it does not lead to asymptotically exact confidence intervals for individual components of z_0 because it is not a consistent estimator of $df_{0,S}^{\mathrm{nor}}(z_0)$ and therefore does not provide sufficient information on the distribution of individual components of z_N. It was shown in [30] that the following interval

$$(z_N)_j \pm \sqrt{\frac{\chi_1^2(\alpha) \left(df_{N,S}^{\mathrm{nor}}(z_N)^{-1} \Sigma_N df_{N,S}^{\mathrm{nor}}(z_N)^{-T} \right)_{j,j}}{N}} \tag{12}$$

is not guaranteed to be asymptotically exact without more restrictive assumptions on the true solution z_0.

Methods for computing asymptotically exact individual confidence intervals for $(z_0)_j$ and $(x_0)_j$ are proposed in [26] by using a consistent estimator for $df_{0,S}^{\mathrm{nor}}(z_0)$ developed in [29]. To reduce the amount of computation associated with the piecewise structure, the authors propose a conditioning-based approach in [26] to eliminate calculations for irrelevant regions. This is done by defining two functions $\eta_j^\alpha(\cdot,\cdot)$ and $h_j^\alpha(\cdot,\cdot,\cdot)$ to be used to compute confidence intervals for $(z_0)_j$ and $(x_0)_j$ respectively. To define the function $\eta_j^\alpha(\cdot,\cdot)$, let $\psi : \mathbb{R}^n \to \mathbb{R}^n$ be a piecewise linear homeomorphism with a family of selection functions $\{M_1, \ldots, M_l\}$ and the corresponding conical subdivision $\{K_1, \ldots, K_l\}$. For any choice of cone K_i, $i = 1, \ldots, l$, component $j =$

$1, \ldots, n$ and $\alpha \in (0, 1)$ we first define $\eta_j^\alpha(\psi, x)$ for points $x \in \text{int}K_i$ as the unique and strictly positive number satisfying

$$\Pr\left(\left|\left(\psi^{-1}(Z)\right)_j\right| \leq \eta_j^\alpha(\psi, x),\ \psi^{-1}(Z) \in K_i\right) = (1-\alpha)\Pr\left(\psi^{-1}(Z) \in K_i\right) \tag{13}$$

where Z denotes a standard normal random vector and $\left(\psi^{-1}(Z)\right)_j$ stands for the jth component of the random variable $\psi^{-1}(Z)$. For points $x \in \bigcap_{s=1}^k K_{i_s}$ define $\eta_j^\alpha(\psi, x) = \max_{s=1,\ldots,k} \eta_j^\alpha(\psi, x_{i_s})$ where $x_{i_s} \in \text{int}K_{i_s}$. The definition of $h_j^\alpha(\cdot, \cdot, \cdot)$ is analogous. Let ψ and ν be piecewise linear functions from $\mathbb{R}^n$ to $\mathbb{R}^n$ that share a common conical subdivision, $\{K_1, \ldots, K_l\}$, with ν invertible. Then for any choice of cone $K_i, i = 1, \ldots, l$, component $j = 1, \ldots, n$ and $\alpha \in (0, 1)$ define $h_j^\alpha(\psi, \nu, x)$ for points $x \in \text{int}K_i$ to be

$$h_j^\alpha(\psi, \nu, x) = \inf\left\{\ell \geq 0 \;\middle|\; \frac{\Pr\left(\left|\left(\psi\circ\nu^{-1}(Z)\right)_j\right| \leq \ell \text{ and } \nu^{-1}(Z) \in K_i\right)}{\Pr\left(\nu^{-1}(Z) \in K_i\right)} \geq (1-\alpha)\right\}. \tag{14}$$

For points $x \in \bigcap_{s=1}^v K_{i_s}$ define $h_j^\alpha(\psi, \nu, x) = \max_{s=1,\ldots,v} h_j^\alpha(\psi, \nu, x_{i_s})$ where $x_{i_s} \in \text{int}K_{i_s}$.

In [26], the confidence intervals for $(z_0)_j$ are obtained by evaluating $\eta_j^\alpha(\psi, x)$, where ψ is a transformation of a consistent estimator of $df_{0,S}^{\text{nor}}(z_0)$ so that the distribution of $\psi^{-1}(Z)$ approximates that of $\sqrt{N}(z_N - z_0)$, and the location of x is determined by z_N and that estimator of $df_{0,S}^{\text{nor}}(z_0)$ and is used to identify the cone in the conical subdivision associated with $df_{0,S}^{\text{nor}}(z_0)$ that contains $z_N - z_0$. Confidence intervals of $(x_0)_j$ are obtained by evaluations of h_j^α which is analogous to η_j^α but more involved. This approach to computing individual confidence intervals is dependent on the consistency of the estimator for $df_{0,S}^{\text{nor}}(z_0)$ and the accurate estimate for the piecewise structure of the asymptotic distribution of SAA solutions. The consistency of that estimator is ensured by requiring z_N to converge to z_0 at an exponential rate, which is in turn guaranteed by replacing Assumption 1 with a stronger assumption on the moment generating functions of $F(X, \xi) - f_0(x)$ and $dF(X, \xi) - df_0(x)$.

In the next section we will develop new methods that do not use that consistent estimator and do not require the exponential convergence rate of z_N. The conditioning-based approach will be extended with generalizations of η_j^α and h_j^α. The new methods also provide a natural mechanism for investigating the sensitivity of interval widths to different estimates for the piecewise structure in the asymptotic distribution of the SAA solutions.

3 New methods on computation of individual confidence intervals

In this section we develop new methods for computing confidence intervals for $(z_0)_j$ and $(x_0)_j$. The basic idea is to estimate the selection function of $df_{S,0}^{\text{nor}}(z_0)$ that contains

$z_N - z_0$ in its corresponding cone, and then compute confidence intervals based on that estimate. In contrast to methods in [26] which first obtain an estimate of the function $df^{\text{nor}}_{S,0}(z_0)$ and then identify a selection function of it, the new methods directly target at estimation of that selection function and not the global piecewise structure. As a result, the new methods are more flexible than the methods in [26]. In particular, the new methods do not require the exponential convergence rate of z_N. The new methods also allow a misestimation of the selection function to occur within a given probability, and allow for the use of multiple cones to estimate the cone corresponding to the selection function.

Below, Sect. 3.1 discusses how to estimate the selection function of $df^{\text{nor}}_{S,0}(z_0)$, Sect. 3.2 introduces the functions $\tilde{\eta}^{\alpha}_j$ and $\tilde{h}^{\alpha_2}_j$ to be used for interval computation and proves the limiting coverage properties of the intervals, and Sect. 3.3 provides a specific method to estimate the cone corresponding to the selection function.

3.1 Estimation of a selection function of $df^{\text{nor}}_{S,0}(z_0)$

For a given sample we need to estimate the selection function of $df^{\text{nor}}_{S,0}(z_0)$ that contains $z_N - z_0$ in its corresponding cone. Using the notation from the previous section let $\Gamma(z_0) = \{P_1, \ldots, P_l\}$ denote the n-cells in the normal manifold of S that contain z_0. Then the cones $K_i = \text{cone}(P_i - z_0)$, $i = 1, \ldots, l$, make up the elements of the conical subdivision associated with $df^{\text{nor}}_{S,0}(z_0)$. Let $\Gamma'(z_0) = \{K_1, \ldots, K_l\}$ and let $K_{i(z_N)}$ denote the element of $\Gamma'(z_0)$ with

$$z_N - z_0 \in K_{i(z_N)};$$

by (10) the choice of $K_{i(z_N)}$ is unique with high probability. Let $M_{i(z_N)}$ denote the matrix representation for the selection function corresponding to the cone $K_{i(z_N)}$. Then by definition for all $h \in K_{i(z_N)}$

$$M_{i(z_N)}h = df^{\text{nor}}_{0,S}(z_0)(h) = df_0(x_0) \circ d\Pi_S(z_0)(h) + h - d\Pi_S(z_0)(h).$$

The proposed methods will require an estimate $M_N(\omega)$ for $M_{i(z_N)}$ that satisfies

$$\lim_{N\to\infty} \Pr\left(\sup_{h\in\mathbb{R}^n, h\neq 0} \frac{\|M_N h - M_{i(z_N)}h\|}{\|h\|} \leq \varepsilon \right) = 1 \tag{15}$$

for all $\varepsilon > 0$. To obtain such an estimate M_N we make use of the following lemma.

Lemma 1 *Suppose that Assumptions 1 and 2 hold. Then the estimate*

$$M_N := df^{\text{nor}}_{N,S}(z_N) = df_N(x_N) \circ d\Pi_S(z_N) + I_n - d\Pi_S(z_N)$$

satisfies (15).

Proof Let $\Gamma(z_0)$ denote the set of n-cells in the normal manifold of S that contain z_0, and for each N define the event

$$A_N = \left\{\omega \,\middle|\, z_N \in \cup_{P_j \in \Gamma(z_0)} ri(P_j)\right\}.$$

When A_N occurs the SAA solution is contained in the interior of an n-cell $P_N(\omega) \in \Gamma(z_0)$. For any $z \in P_N(\omega)$, the restriction of $d\Pi_S(z)$ to the cone $K_z(\omega) = \text{cone}(P_N(\omega) - z)$ satisfies $d\Pi_S(z)|_{K_z(\omega)} = d\Pi_S(z_N)$.

By [30, Proposition 3.5], under Assumptions 1 and 2 $\lim_{N\to\infty} \Pr(A_N) = 1$. Therefore

$$\lim_{N\to\infty} \Pr\left(d\Pi_S(z_0)\Big|_{K_{i(z_N)}} = d\Pi_S(z_N)\right) = 1 \tag{16}$$

and,

$$\begin{aligned}
&\lim_{N\to\infty} \Pr\left(\sup_{h\in\mathbb{R}^n, h\neq 0} \frac{\|M_N h - M_{i(z_N)} h\|}{\|h\|} \leq \varepsilon\right) \\
&= \lim_{N\to\infty} \Pr\left(\sup_{h\in\mathbb{R}^n, h\neq 0} \frac{\|df_N(x_N)\circ d\Pi_S(z_N)(h) - df_0(x_0)\circ d\Pi_S(z_N)(h)\|}{\|h\|} \leq \varepsilon\right) \\
&= 1
\end{aligned}$$

where the final equality follows from the a.s. convergence of f_N to f_0 in $C^1(X, \mathbb{R}^n)$. □

Estimation of the cone $K_{i(z_N)}$ demands more work. Without knowing the rate at which z_N converges to z_0, we do not have an estimate that equals the cone $K_{i(z_N)}$ with probability converging to one. We therefore relax the conditions required for an estimate of $K_{i(z_N)}$ in two ways. First, we allow $K_{i(z_N)}$ to be estimated with a set of cones instead of a single cone. Second, we only require that set contain $K_{i(z_N)}$ with a limiting probability greater than a specified lower bound. Formally, our methods will make use of a family of polyhedral cones of dimension n, denoted by $\mathscr{K}_N^{\alpha_1}$, that satisfy

$$\liminf_{N\to\infty} \Pr\left(K_{i(z_N)} \in \mathscr{K}_N^{\alpha_1}\right) \geq 1 - \alpha_1. \tag{17}$$

We will provide a computationally efficient method in Sect. 3.3 for choosing $\mathscr{K}_N^{\alpha_1}$ based on an SAA solution z_N, and prove that it satisfies (17).

3.2 Limiting coverage probabilities of confidence intervals

We first give a generalization of the function η_j^α in (13), to be used to compute the width of a confidence interval for $(z_0)_j$. Let M be an invertible $n \times n$ matrix and

$\mathscr{K} = \{\tilde{K}_1, \ldots, \tilde{K}_m\}$ be a collection of polyhedral convex sets of dimension n. We define the function $\tilde{\eta}_j^{\alpha}(\mathscr{K}, M)$ as

$$\tilde{\eta}_j^{\alpha}(\mathscr{K}, M) = \inf\left\{ \ell \geq 0 \,\middle|\, \frac{\Pr\left(\left|(M^{-1}Z)_j\right| \leq \ell, \text{ and } M^{-1}Z \in \tilde{K}_i\right)}{\Pr\left(M^{-1}Z \in \tilde{K}_i\right)} \geq 1-\alpha \text{ for all } \tilde{K}_i \in \mathscr{K} \right\}.$$

To compute an confidence interval's width, the value of M at which we evaluate $\tilde{\eta}_j^{\alpha}$ is chosen so that $M^{-1}Z$ provides an estimate for random variable $(L_{K_0}^{\text{nor}})^{-1}(Y_0)$, as defined in (8), when $(L_{K_0}^{\text{nor}})^{-1}(Y_0) \in K_{i(z_N)}$. The set of cones $\mathscr{K}$ at which we evaluate $\tilde{\eta}_j^{\alpha}$ provide a set of estimates for $K_{i(z_N)}$. The ability to include multiple cones in $\mathscr{K}$ provides a natural means for addressing the potential sensitivity of the interval's width to misestimation of the piecewise structure in the distribution of $(L_{K_0}^{\text{nor}})^{-1}(Y_0)$.

The following convergence result establishes the desired properties of intervals computed using $\tilde{\eta}_j^{\alpha}$. As before, $\Gamma'(z_0) = \{K_1, \ldots, K_l\}$ denotes the conical subdivision associated with $df_{0,S}^{\text{nor}}(z_0)$, with $df_{0,S}^{\text{nor}}(z_0)|_{K_i} = M_i$ and $K_i = \text{cone}(P_i - z_0)$ where $P_1, \ldots, P_l$ are the n-cells in the normal manifold of S that contain z_0.

Theorem 1 *Suppose that Assumptions 1 and 2 hold, that $\mathscr{K}_N^{\alpha_1}$ satisfies (17) for $\alpha_1 \in (0, 1]$, and that the determinant of Σ_0 is strictly positive. Then for every $j = 1, \ldots, n$ and $\alpha_2 \in (0, 1)$,*

$$\liminf_{N\to\infty} \Pr\left(\sqrt{N}\left|(z_N - z_0)_j\right| \leq \tilde{\eta}_j^{\alpha_2}\left(\mathscr{K}_N^{\alpha_1}, \Sigma_N^{-1/2} df_{N,S}^{\text{nor}}(z_N)\right)\right) \geq 1 - (\alpha_1 + \alpha_2).$$

Proof For each N define a function $\Phi_0(z_N) : \mathbb{R}^n \to \mathbb{R}^n$ as

$$\Phi_0(z_N)(h) = df_N(\Pi_S(z_N)) \circ d\Pi_S(z_0)(h) + h - d\Pi_S(z_0)(h) \text{ for each } h \in \mathbb{R}^n, \quad (18)$$

and define the event $A_N = \left\{K_{i(z_N)} \in \mathscr{K}_N^{\alpha_1}\right\}$. Let B be a fixed neighborhood of z_0 such that $B \cap (z_0 + K_i) = B \cap P_i$ for $i = 1, \ldots, l$. Then

$$\begin{aligned}
&\liminf_{N\to\infty} \Pr\left(\sqrt{N}\left|(z_N - z_0)_j\right| \leq \tilde{\eta}_j^{\alpha_2}\left(\mathscr{K}_N^{\alpha_1}, \Sigma_N^{-1/2} df_{N,S}^{\text{nor}}(z_N)\right)\right) \\
&\quad \geq \liminf_{N\to\infty} \sum_{i=1}^{l} \Pr\left(\sqrt{N}\left|(z_N - z_0)_j\right|\right. \\
&\quad \left. \leq \tilde{\eta}_j^{\alpha_2}\left(\left\{K_{i(z_N)}\right\}, \Sigma_N^{-1/2} df_{N,S}^{\text{nor}}(z_N)\right);\ A_N;\ z_N \in B \cap \text{int} P_i\right).
\end{aligned}$$

When A_N holds and $z_N \in B \cap \text{int} P_i$, we have $\Phi_0(z_N)|_{K_{i(z_N)}} = df_{N,S}^{\text{nor}}(z_N)$, and for any $x_i \in \text{int}(K_{i(z_N)})$

$$\tilde{\eta}_j^{\alpha_2}\left(\left\{K_{i(z_N)}\right\}, \Sigma_N^{-1/2} df_{N,S}^{\text{nor}}(z_N)\right) = \eta_j^{\alpha_2}\left(\Sigma_N^{-1/2}\Phi_0(z_N), x_i\right). \quad (19)$$

Under Assumptions 1 and 2, for all $\varepsilon > 0$,

$$\lim_{N\to\infty} \Pr\left(\sup_{h\in\mathbb{R}^n, h\neq 0} \frac{\left\| \Phi_0(z_N)(h) - df^{\mathrm{nor}}_{0,S}(z_0)(h) \right\|}{\|h\|} \leq \varepsilon \right) = 1.$$

By applying [26, Lemma 3] and the same argument used to show [26, Theorem 6], we find

$$\eta_j^{\alpha_2}\left(\Sigma_N^{-1/2} \Phi_0(z_N), x_i \right) \Rightarrow \eta_j^{\alpha_2}\left(\Sigma_0^{-1/2} df^{\mathrm{nor}}_{0,S}(z_0), x_i \right),$$

and

$$\begin{aligned} &\lim_{N\to\infty} \sum_{i=1}^{l} \Pr\left(\sqrt{N} |(z_N - z_0)_j| \leq \eta_j^{\alpha_2}\left((\Sigma_N^{-1/2} \Phi_0(z_N), x_i \right);\ z_N \in B \cap \mathrm{int} P_i \right) \\ &\quad = 1 - \alpha_2. \end{aligned} \tag{20}$$

Next we observe that,

$$\begin{aligned} &\liminf_{N\to\infty} \sum_{i=1}^{l} \Pr\left(\sqrt{N} |(z_N - z_0)_j| \right. \\ &\quad \leq \tilde{\eta}_j^{\alpha_2}\left(\left\{ K_{i(z_N)} \right\}, \Sigma_N^{-1/2} df^{\mathrm{nor}}_{N,S}(z_N) \right);\ A_N;\ z_N \in B \cap \mathrm{int} P_i \Big) \\ &\quad = \liminf_{N\to\infty} \sum_{i=1}^{l} \Big[\Pr\left(\sqrt{N} |(z_N - z_0)_j| \right. \\ &\quad \leq \eta_j^{\alpha_2}\left(\Sigma_N^{-1/2} \Phi_0(z_N), x_i \right);\ A_N;\ z_N \in B \cap \mathrm{int} P_i \Big) \Big] \\ &\quad \geq \lim_{N\to\infty} \sum_{i=1}^{l} \Big[\Pr\left(\sqrt{N} |(z_N - z_0)_j| \right. \\ &\quad \leq \eta_j^{\alpha_2}\left(\Sigma_N^{-1/2} \Phi_0(z_N), x_i \right);\ z_N \in B \cap \mathrm{int} P_i \Big) \Big] - \alpha_1 \end{aligned}$$

where the first equality follows from (19), and the inequality follows from the definition of A_N and (17). Combining this with (20) proves the theorem. □

In the proof of Theorem 1 the nonsingularity assumption about Σ_0 is used to ensure that the limit (20) is well defined. This assumption can be replaced with the condition described after [26, Theorem 6] without affecting the convergence result. That condition requires that for all cones $K_i \in \Gamma'(z_0)$ and matrices M_i satisfying $df^{\mathrm{nor}}_{0,S}(z_0)|_{K_i} = M_i$ the polyhedra K_i are continuity sets with respect to the random vector $(M_i)^{-1} Y_0$ and that $\Pr((M_i^{-1} Y_0)_j = 0$ must equal zero.

Next, we extend the definition of h_j^{α} in (14) to be used to directly compute confidence intervals for $(x_0)_j$. As above, let M be an invertible $n \times n$ matrix and

$\mathscr{K} = \{\tilde{K}_1, \ldots, \tilde{K}_m\}$ be a collection of polyhedral convex sets of dimension n. In addition, let Q be an $n \times n$ matrix. We define the function $\tilde{h}_j^{\alpha_2}(\mathscr{K}, Q, M)$ as

$$\tilde{h}_j^{\alpha_2}(\mathscr{K}, Q, M) = \inf\left\{\ell \geq 0 \,\middle|\, \frac{\Pr\left(|(Q)_j M^{-1} Z| \leq \ell, \text{ and } M^{-1} Z \in \tilde{K}_i\right)}{\Pr\left(M^{-1} Z \in \tilde{K}_i\right)} \geq 1 - \alpha_2 \text{ for all } \tilde{K}_i \in \mathscr{K}\right\}.$$

Theorem 2 *Suppose that Assumptions 1 and 2 hold, that $\mathscr{K}_N^{\alpha_1}$ satisfies (17) for $\alpha_1 \in (0, 1]$, and that the determinant of Σ_0 is strictly positive. Then for every $j = 1, \ldots, n$ and $\alpha_2 \in (0, 1)$,*

$$\liminf_{N\to\infty} \Pr\left(\sqrt{N}|(x_N - x_0)_j| \leq \tilde{h}_j^{\alpha_2}\left(\mathscr{K}_N^{\alpha_1}, d\Pi_S(z_N), \Sigma_N^{-1/2} df_{N,S}^{\text{nor}}(z_N)\right)\right) \geq 1 - (\alpha_1 + \alpha_2).$$

Proof For each N define $\Phi_0(z_N) : \mathbb{R}^n \times \mathbb{R}^n$ as in (18), and define the event $A_N = \left\{K_{i(z_N)} \in \mathscr{K}_N^{\alpha_1}\right\}$. As before let B be a fixed neighborhood of z_0 such that $B \cap (z_0 + K_i) = B \cap P_i$ for $i = 1, \ldots, l$ and let $u_i \in \text{int}K_i$. Then using the same arguments in Theorem 1 and [26, Theorem 7] it follows that,

$$\begin{aligned}
&\liminf_{N\to\infty} \Pr\left(\sqrt{N}|(x_N - x_0)_j| \leq \tilde{h}_j^{\alpha_2}\left(\mathscr{K}_N^{\alpha_1}, d\Pi_S(z_N), \Sigma_N^{-1/2} df_{N,S}^{\text{nor}}(z_N)\right)\right) \\
&\geq \liminf_{N\to\infty} \sum_{i=1}^{l} \Pr\left(\sqrt{N}|(x_N - x_0)_j| \right. \\
&\left. \leq \tilde{h}_j^{\alpha_2}\left(\left\{K_{i(z_N)}\right\}, d\Pi_S(z_N), \Sigma_N^{-1/2} df_{N,S}^{\text{nor}}(z_N)\right);\ A_N;\ z_N \in B \cap \text{int}P_i\right) \\
&= \liminf_{N\to\infty} \sum_{i=1}^{l} \Pr\left(\sqrt{N}|(x_N - x_0)_j| \leq h_j^{\alpha_2}\left(d\Pi_S(z_0), \Sigma_N^{-1/2}\Phi_0(z_N), u_i\right);\ A_N;\ z_N \in B \cap \text{int}P_i\right) \\
&\geq \lim_{N\to\infty} \sum_{i=1}^{l} \left[\Pr\left(\sqrt{N}|(x_N - x_0)_j| \leq h_j^{\alpha_2}\left(d\Pi_S(z_0), \Sigma_N^{-1/2}\Phi_0(z_N), u_i\right);\ z_N \in B \cap \text{int}P_i\right)\right] - \alpha_1 \\
&\geq 1 - (\alpha_1 + \alpha_2).
\end{aligned}$$

□

As in the proof of Theorem 1, the assumption that Σ_0 is invertible is used to ensure that the final limit in the proof of Theorem 2 is well defined, and can similarly be relaxed to the conditions described after [26, Theorem 7]. By its definition, the function $\tilde{h}_j^{\alpha}$ will return a value of zero when the jth component of $d\Pi_S(z_N)$ is zero. Under the conditions assumed by Theorem 2, it follows from the a.s. convergence of z_N to z_0 and Eqs. (5), (6), and (16) that

$$\lim_{N\to\infty} \Pr\left((d\Pi_S(z_N))_j = 0\right) = \lim_{N\to\infty} \Pr\left((d\Pi_S(z_N))_j = 0 \quad \text{and} \quad (x_N)_j = (x_0)_j\right)$$

and the value of zero returned by $\tilde{h}_j^{\alpha}$ is asymptotically correct. However, the value of $\tilde{h}_j^{\alpha}$ may incorrectly equal zero when z_N is contained in an n-cell P_k for which the

jth component of $\Pi_S|_{P_k}$ is zero but $z_0 \notin P_k$, an event which occurs with probability converging to zero by (16). In practice, to guard against such cases, when the jth component of an estimate $d\Pi_S(z_N)$ is equal to zero we replace it by the unit vector e_j whose jth element is one. This replacement does not effect the results of Theorem 2 and would return a value of $\tilde{h}_j^\alpha = \tilde{\eta}_j^\alpha$.

To evaluate $\tilde{\eta}_j^\alpha$ and $\tilde{h}_j^\alpha$ the approach described in [26, Section 4] can be applied. This approach requires a search over values of $\ell > 0$ and evaluating probabilities of the form

$$\Pr\left(\left|(Q)_j M^{-1} Z\right| \leq \ell, \quad \text{and} \quad M^{-1} Z \in K_i\right).$$

Monte Carlo or Quasi-Monte Carlo methods as discussed and implemented in [15, 16] can be used to evaluate those probabilities. In practice we have found that this implementation provides sufficient accuracy for the evaluation $\tilde{\eta}_j^\alpha$ and $\tilde{h}_j^\alpha$ so long as the probability

$$\Pr\left(M^{-1} Z \in K_i\right)$$

is greater than 10^{-6} in magnitude.

3.3 Identification of $\mathscr{K}_N^{\alpha_1}$

For the methods of Theorems 1 and 2 to be computationally tractable we would like to limit the number of cones $K_i \in \mathscr{K}_N^{\alpha_1}$. In Theorem 3 below we show that from the sample data we can identify a single cone to be included in $\mathscr{K}_N^{\alpha_1}$ to satisfy (17). To do so we use the asymptotically exact confidence regions in (11) to identify a subset of cells in the normal manifold of S, and select a cell with the lowest dimension from this subset. The proof of Theorem 3 does not require Σ_N to be invertible and will use the following proposition which provides a closed form expression for the simultaneous confidence regions computed from $R_{N,\varepsilon}$. The simultaneous confidence intervals are computed by finding the minimal axis-aligned bounding box that contains $R_{N,\varepsilon}$. When $df_{N,S}^{\text{nor}}(z_N)$ is a linear function, the resulting simultaneous confidence intervals are given by

$$\left[(z_N)_1 - w_{N,1}^\varepsilon, (z_N)_1 + w_{N,1}^\varepsilon\right] \times \cdots \times \left[(z_N)_n - w_{N,n}^\varepsilon, (z_N)_n + w_{N,n}^\varepsilon\right] \tag{21}$$

where $w_{N,j}^\varepsilon$ is the optimal value of the following problem in which $(U_N)_1$, $(U_N)_2$, D_N are defined in Proposition 1.

$$\begin{aligned} &\text{maximize} && (w)_j \\ &\text{subject to:} && N\left[df_{N,S}^{\text{nor}}(z_N)(w)\right]^T (U_N)_1^T D_N^{-1} (U_N)_1 \left[df_{N,S}^{\text{nor}}(z_N)(w)\right] \leq \chi^2_{l_N}(\alpha) \\ &&& \|\sqrt{N}(U_N)_2 df_{N,S}^{\text{nor}}(z_N)(w)\|_\infty \leq \varepsilon. \end{aligned} \tag{22}$$

Proposition 1 *Suppose that* $df^{\text{nor}}_{N,S}(z_N)$ *is a linear homeomorphism and* Σ_N *has decomposition* $\Sigma_N = U_N^T \Delta_N U_N$, *where* U_N *is an orthogonal matrix with rows* $u_{N,1}, \ldots, u_{N,n}$ *and* Δ_N *is a diagonal matrix with elements* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. *For a choice of* l_N *with* $\lambda_{l_N} > 0$, *let* D_N *be a diagonal matrix with elements* $\lambda_1, \ldots, \lambda_{l_N}$,

$$(U_N)_1 = \begin{bmatrix} u_{N,1} \\ \ldots \\ u_{N,l_N} \end{bmatrix} \quad \text{and} \quad (U_N)_2 = \begin{bmatrix} u_{N,l_N+1} \\ \vdots \\ u_{N,n} \end{bmatrix}.$$

Then for each $j = 1, \ldots, n$, *the optimal value of* (22) *is an affine function of* ε *given by*

$$w^{\varepsilon}_{N,j} = \sqrt{\frac{\chi^2_{l_N}(\alpha) \sum_{i=1}^{l_N} \left(c_{N,j} u^T_{N,i}\right)^2 \lambda_i}{N}} + \frac{\varepsilon}{\sqrt{N}} \sum_{i=l_N+1}^{n} \left| c_{N,j} u^T_{N,i} \right|, \tag{23}$$

where $c_{N,j}$ *is the* (jth) *row of* $(df^{\text{nor}}_{N,S}(z_N))^{-1}$.

Proof Let V_N and T_N be the subspaces spanned by $\{u^T_{N,1}, \ldots, u^T_{N,l_N}\}$ and $\{u^T_{N,l_N+1}, \ldots, u^T_{N,n}\}$, respectively. Then V_N is the orthogonal complement of T_N and any vector $df^{\text{nor}}_{N,S}(z_N)(w)$ can be decomposed as

$$df^{\text{nor}}_{N,S}(z_N)(w) = v + t$$

with $v \in V_N$ and $t \in T_N$. The result (23) can be proved by formulating (22) in terms of v and t and then separating it into two problems. □

In the proof of Theorem 3 we will use the same notation as in Theorem 1. With this notation, the conical subdivision of $d\Pi_S(z_0)$ is comprised of sets $K_i = \text{cone}(P_i - z_0)$ where $P_1, \ldots, P_l$ are all n-cells in the normal manifold of S that contain z_0. The element of the conical subdivision of $d\Pi_S(z_0)$ that contains $z_N - z_0$ for a particular sample shall be denoted by $K_{i(z_N)}$.

Theorem 3 *Suppose that Assumptions* 1 *and* 2 *hold. Let* $\alpha_1 \in (0, 1)$ *and* $\varepsilon > 0$ *be fixed. Let* $R_{N,\varepsilon}$ *be a* $(1 - \alpha_1) * 100\%$ *confidence region for* z_0 *as given in Eq.* (11), P_N *be an* n*-cell in the normal manifold of* S *with* $z_N \in P_N$, *and* C_{i_N} *be a cell that has the smallest dimension of all cells that intersect* $R_{N,\varepsilon}$ *and* P_N. *Then for* $\tilde{z}_{i_N} \in ri\, C_{i_N}$, $\mathscr{K}_N^{\alpha_1} = \left\{cone(P_N - \tilde{z}_{i_N})\right\}$ *satisfies* (17).

Proof Let $C_1, \ldots, C_m$ denote all of the cells in the normal manifold of S, and $\mathscr{I} = \{i \mid z_0 \notin C_i\}$ be the collection of indices for the cells that do not contain z_0. For each cell C_i define the function

$$d_i(z) = d(z, C_i) = \min_{x \in C_i} \|x - z\|.$$

Then $\min_{i \in \mathscr{I}} d_i(z_0) = \delta > 0$ since $d_i(z_0) > 0$ for each $i \in \mathscr{I}$ and $\mathscr{I}$ is a finite set. Let C_{i_0} denote the unique cell that contains z_0 in its relative interior. As shown in [31,

Proposition 5.1] the cell C_{i_0} is the cell of lowest dimension to contain z_0. Therefore for any cell C_i with $C_i \neq C_{i_0}$ and dimension less than or equal to that of C_{i_0}, $z_0 \notin C_i$ and $i \in \mathscr{I}$.

For any $i \in \mathscr{I}$ and $z \in C_i$,

$$\|z_N - z\| \geq \|z_0 - z\| - \|z_N - z_0\| \geq \delta - \|z_N - z_0\|. \tag{24}$$

Let G_N denote the event that $\delta/2 \leq \min_{i \in \mathscr{I}} d_i(z_N)$. By (24) and the almost sure convergence of z_N to z_0, the probability of G_N occurring converges to one.

Next, consider the simultaneous confidence intervals (21) for z_0. As in Proposition 1 let $\Sigma_N = U_N^T \Delta_N U_N$, where U_N is an orthogonal matrix with rows $u_{N,1}, \ldots, u_{N,n}$ and Δ_N is a diagonal matrix with elements $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. For each $j = 1, \ldots, n$, let $c_{N,j}$ be the jth row of $df^{\text{nor}}_{N,S}(z_N)^{-1}$. Then the half-width $w^{\varepsilon}_{N,j}$ of the jth interval in (21) is given by (23).

Let $w_N = (w^{\varepsilon}_{N,1}, \ldots, w^{\varepsilon}_{N,j})$. Define the event A_N to be $\left\{ \|w_N\| < \delta/2 \text{ and } z_0 \in R_{N,\varepsilon} \right\}$, and let B be a fixed neighborhood of z_0 such that $B \cap (z_0 + K_i) = B \cap P_i$ for $i = 1, \ldots, l$. Then,

$$\begin{aligned}
&\liminf_{N\to\infty} \Pr\left(K_{i(z_N)} = \text{cone}(P_N - \tilde{z}_{i_N})\right) \\
&= \liminf_{N\to\infty} \sum_{i=1}^{l} \Pr\left(K_i = \text{cone}(P_i - \tilde{z}_{i_N});\ z_N \in B \cap \text{int} P_i\right) \\
&\geq \liminf_{N\to\infty} \sum_{i=1}^{l} \Pr\left(K_i = \text{cone}(P_i - \tilde{z}_{i_N});\ A_N;\ G_N;\ z_N \in B \cap \text{int} P_i\right) \\
&= \liminf_{N\to\infty} \sum_{i=1}^{l} \Pr\left(A_N;\ G_N;\ z_N \in B \cap \text{int} P_i\right).
\end{aligned}$$

The final equality holds for the following reason. Suppose A_N and G_N both occur, with $z_N \in B \cap \text{int} P_i$. Then $P_N = P_i$, and C_{i_0} intersects with P_N because P_i contains z_0. When the event A_N occurs, for any $z \in R_{N,\varepsilon}$,

$$\|z_N - z\| \leq \|w_N\| < \delta/2.$$

On the other hand, when G_N occurs,

$$\delta/2 \leq \min_{i \in \mathscr{I}} d_i(z_N),$$

and thus no cell with index $i \in \mathscr{I}$ intersects with $R_{N,\varepsilon}$. Therefore each cell that intersects with $R_{N,\varepsilon}$ necessarily contains z_0 and C_{i_0}. It follows that C_{i_0} is the cell of lowest dimension to intersect with $R_{N,\varepsilon}$ and P_N, i.e., $C_{i_0} = C_{i_N}$. Consequently $K_i = \text{cone}(P_i - z_0) = \text{cone}(P_i - \tilde{z}_{i_N})$.

It follows that

$$\begin{aligned}
&\liminf_{N\to\infty} \Pr\left(K_{i(z_N)} = \text{cone}(P_N - \tilde{z}_{i_N})\right) \\
&\geq \liminf_{N\to\infty} \sum_{i=1}^{l} \Pr\left(A_N;\ G_N;\ z_N \in B \cap \text{int}P_i\right) \\
&= \liminf_{N\to\infty} \Pr\left(A_N;\ G_N\right) \\
&\geq \lim_{N\to\infty} \Pr\left(z_0 \in R_{N,\varepsilon}\right) = 1 - \alpha_1
\end{aligned}$$

and $\mathscr{K}_N^{\alpha_1} = \left\{\text{cone}(P_N - \tilde{z}_{i_N})\right\}$ satisfies (17). □

When the k-cell that contains z_0 in its relative interior is a singleton so that $C_{i_0} = \{z_0\}$, we have

$$\lim_{N\to\infty} \Pr\left(K_{i(z_N)} = \text{cone}(P_N - \tilde{z}_{i_N})\right) = 1 - \alpha_1$$

and (17) is satisfied with equality. Alternatively if z_0 is contained in the interior of an n-cell

$$\lim_{N\to\infty} \Pr\left(K_{i(z_N)} = \text{cone}(P_N - \tilde{z}_{i_N})\right) = 1$$

so in general it is not possible to sharpen the results of Theorem 3. The proof of Theorem 3 can also be applied to justify the use of simultaneous confidence intervals computed from $R_{N,\varepsilon}$ to identify the set of k-cells from which C_{i_N} is chosen. When the set S is a box, working with the simultaneous confidence intervals has the computational benefit of allowing us to identify the cell C_{i_N} by making n componentwise comparisons.

In Theorem 3 we use $(1 - \alpha_1) * 100\%$ confidence regions to identify which cell in the normal manifold of S contains z_0 in its relative interior and which cone in the conical subdivision of $df_{0,S}^{\text{nor}}(z_0)$ contains $z_N - z_0$. The limiting probability of making an incorrect choice is bounded above by α_1. This is different from the methods of [26] which estimate $df_{0,S}^{\text{nor}}(z_0)$ based on the exponential rate of convergence of z_N. The new approach allows us to construct intervals to meet or exceed the $(1 - \alpha) * 100\%$ confidence level, with $\alpha = \alpha_1 + \alpha_2$, by balancing between the control on cone estimation error (α_1) and the interval noncoverage rate given the correct cone (α_2).

Theorems 1 and 2 provide a general framework for confidence interval computation. In addition to accommodating the balance between the two types of error α_1 and α_2, they also allow for the option of including multiple cones in the set $\mathscr{K}_N^{\alpha_1}$. This gives an opportunity to balance between increased empirical coverage and the amount of computation required, which may range from the choice of a single cone as in Theorem 3 to choosing all the cones generated from the faces of P_N. In the later case α_1 can be set to zero, but the required computations may be intractable for problems of even modest size due to the large number of cones involved.

To end this section we illustrate the estimation of $K_{i(z_N)}$ using a single cone with the following example for a single SAA solution. Let $S = \mathbb{R}^2_+$ and

$$F(x, \xi) = \begin{bmatrix} \xi_1 & \xi_2 \\ \xi_3 & \xi_4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} \xi_5 \\ \xi_6 \end{bmatrix}$$

where the random vector ξ is uniformly distributed over the box $[0, 2] \times [0, 1] \times [0, 2] \times [0, 4] \times [-1, 1] \times [-1, 1]$. For this example f_0 has a closed form expression and the true SVI is given by

$$0 \in \begin{bmatrix} 1 & 1/2 \\ 1 & 2 \end{bmatrix} x + \mathbf{N}_{\mathbb{R}^2_+}(x)$$

and has true solutions $z_0 = x_0 = 0$. The B-derivative $df^{\text{nor}}_{0,S}(z_0)$ is a piecewise line function represented by the matrices

$$\begin{bmatrix} 1 & 1/2 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1/2 \\ 0 & 2 \end{bmatrix}, \text{ and } \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

in each of the orthants $\mathbb{R}^2_+$, $\mathbb{R}_+ \times \mathbb{R}_-$, $\mathbb{R}_- \times \mathbb{R}_+$, and $\mathbb{R}^2_-$. An SAA problem with $N = 200$ is given by

$$0 \in \begin{bmatrix} 1.0887 & 0.4712 \\ 1.0357 & 1.9836 \end{bmatrix} x + \begin{bmatrix} 0.0307 \\ 0.0654 \end{bmatrix} + \mathbf{N}_{\mathbb{R}^2_+}(x).$$

The SAA solution is $z_N = (-0.0307, -0.0654) \in \mathbb{R}_- \times \mathbb{R}_- = K_{i(z_N)}$, and the matrices $M_{i(z_N)}$ and M_N are equal to the identity matrix I_2. The 97.5% confidence region for z_0 is equal to

$$Q_N = \left\{ z \in \mathbb{R}^2 \middle| (z - z_N)^T \begin{bmatrix} 3.3798 & -0.0549 \\ -0.0549 & 2.8518 \end{bmatrix} (z - z_N) \leq 0.03689 \right\}$$

which is contained in the rectangle $[-0.1352, 0.0738] \times [-0.1792, 0.0483]$. Edges of the rectangle are conservative simultaneous confidence intervals for $(z_0)_1$ and $(z_0)_2$.

As seen in Fig. 1 both Q_N and the rectangle intersect all of the faces of $\mathbb{R}_- \times \mathbb{R}_-$, and using either set leads to the selection of $\mathscr{K}_N^{.025} = K_{i(z_N)} = \mathbb{R}_- \times \mathbb{R}_-$. Evaluating $\tilde{\eta}_1^{0.025}$ and $\tilde{\eta}_2^{0.025}$ at the estimates $\mathscr{K}_N^{.025}$ and M_N, we obtain the 95% individual confidence intervals $[-0.1173, 0.0558]$ for $(z_0)_1$ and $[-0.1597, 0.0288]$ for $(z_0)_2$. Note that $[-0.1173, 0.0558] \times (z_N)_2 \subset Q_N$ and $(z_N)_1 \times [-0.1597, 0.0288] \subset Q_N$. For this example $d\Pi_S(z_N)$ is equal to the zero matrix, so using the conservative approach to evaluating $\tilde{h}_j^{0.025}$ described after Theorem 2, $\tilde{h}_j^{0.025} = \tilde{\eta}_j^{0.025}$, and the 95% individual confidence intervals for $(x_0)_1$ and $(x_0)_2$ are given by $[0, 0.0866]$ and $[0, 0.0942]$ respectively.

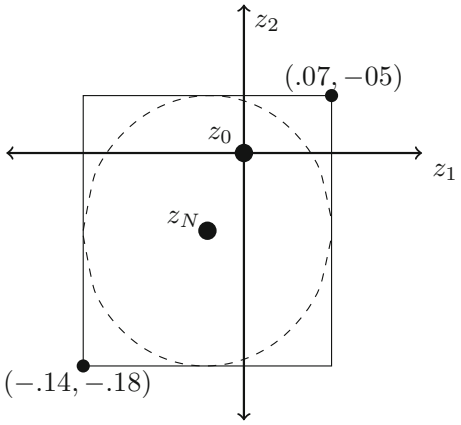

Fig. 1 Confidence region Q_N (dashed) and simultaneous confidence intervals

4 A numerical example

In this section we apply the methods proposed in Sect. 3 to a stochastic Cournot-Nash equilibrium problem. The example is adapted from a model for the European natural gas market previously considered in [20] and [22]. Cournot competition is used to describe the competitive behavior of firms making simultaneous decisions regarding their levels of production and distribution of a homogeneous product across various markets.

To formally state the problem, assume that the decision made by each firm i, $i = 1, \ldots, m$, is represented by a vector x_i of dimension d_i and that each firm's decision is constrained to be in a convex set $S_i \subset \mathbb{R}^{d_i}$. Let $x = (x_1, \ldots, x_m)$ denote the concatenation of all players' decisions and denote the random vector by ξ. With $\phi_{i_0}(x) = E[\Phi_i(x, \xi)]$ denoting the expected profit function for player i, $x^* = (x_1^*, \ldots, x_m^*)$ is a Cournot-Nash equilibrium if

$$x_i^* \in arg\, max_{x_i \in S_i} \phi_{i0}\left(x_1^*, \ldots, x_i, \ldots, x_m^*\right) \quad \text{for each} \quad i = 1, \ldots, m.$$

The first order necessary condition for player i's profit maximization problem is

$$0 \in -\frac{\partial \phi_{i_0}}{\partial x_i}(x) + \mathbf{N}_{S_i}(x_i).$$

Let $S = \prod_{i=1}^m S_i$ and

$$f_0(x) = \left(-\tfrac{\partial \phi_{1_0}}{\partial x_1}(x), \ \ldots \ -\tfrac{\partial \phi_{m_0}}{\partial x_m}(x)\right).$$

A necessary condition for x^* to be a Cournot-Nash equilibrium is

$$0 \in f_0(x) + \mathbf{N}_S(x). \tag{25}$$

For (25) to fit the framework of an SVI we require a function $F(x, \xi)$ such that $f_0(x) = E[F(x, \xi)]$ and $E\|F(x, \xi)\| < \infty$ for all $x \in S$. The natural candidate

$$F(x, \xi) = \left(-\frac{\partial \Phi_i}{\partial x_1}(x, \xi), \ \ldots \ - \frac{\partial \Phi_m}{\partial x_m}(x, \xi) \right)$$

will meet these criteria if the profit functions Φ_i satisfy the conditions of Assumption 1. In this case the SVI (25) gives rise to the SAA problem

$$0 \in f_N(x) + \mathbf{N}_S(x) \tag{26}$$

where

$$f_N = N^{-1} \sum_{k=1}^{N} F(x, \xi^k).$$

In the natural gas market example we consider in this section the competing firms are four gas producers, indexed by i, with six markets indexed by j. Producers decide on the quantity of gas to ship each year during four time periods, indexed by t, to the six markets. There are 24 decision variables for each producer, denoted by $x_{i,j}^t$, corresponding to the amount of natural gas shipped by producer i to market j each year in time period t. The decision vectors x_i are required to be nonnegative so each player's set of feasible decisions is $S_i = \mathbb{R}_+^{24}$ and the concatenation of all players' decisions is given by $x = (x_1, \ldots, x_4) \in S_1 \times \cdots \times S_4 \subset \mathbb{R}^{96}$.

In the model's formulation, the following parameters are used:

- D_j^t: the domestic gas production of market j each year in time period t
- c_i^t: the constant marginal transportation cost for shipping for producer i in time period t
- e_j^t: the price elasticity of demand for natural gas in market j in time period t
- y_t: the number of years in time period t, taken to be 5 years for time periods 1, 2, 3, and 20 years for time period 4.

In time period t, the yearly production cost for producer i is given by

$$G_i(x) = a_i - b_i \ln \left(X_i - \sum_{j=1}^{6} x_{i,j}^t \right),$$

where a_i, b_i and X_i are parameters. The parameter X_i provides an upper bound on the yearly production of producer i. Values for the parameters indexed by player i are given in Table 1 and values for the parameters indexed by market j are given in Table 2.

The uncertainty in the problem is associated with the price of natural gas in the different markets. The price of natural gas in market j for time period t, denoted by P_j^t, is determined by the total amount of natural gas available annually, as well as ξ^t

Table 1 Producer parameter values

Producer	a	b	X	c^1	c^2	c^3	c^4
Russia	1.606	51	80	0.58	0.56	0.55	0.55
Netherlands	1.212	67	80	0.14	0.13	0.13	0.12
Norway	1.507	85	80	0.35	0.34	0.34	0.33
Algeria	2.102	96	80	0.70	0.69	0.64	0.62

Table 2 Values for price elasticity e and domestic gas production D

Market	Period 1		Period 2		Period 3		Period 4	
BelLux	−1.07	0.00	− 1.26	0.00	− 1.34	0.00	− 1.42	0.00
FRGer	− 1.46	13.70	− 1.58	13.80	− 1.68	13.80	−1.79	13.80
France	− 0.81	4.80	− 1.19	2.90	− 1.57	3.00	− 2.01	3.00
Italy	− 1.15	10.40	− 1.36	10.00	− 1.45	10.00	− 1.54	10.40
Netherlands	− 0.94	22.93	− 1.13	20.96	− 1.29	24.11	− 1.45	23.90
UK	− 0.61	33.70	− 0.87	35.00	− 1.10	37.00	− 1.30	38.00

the random price of oil in time period t. It is given by

$$P_j^t(x, \xi^t) = p_j^t(\xi^t) \left(\frac{Q_j^t(x)}{q_j^t(\xi^t)} \right)^{1/e_j^t}.$$

In the above equation, $Q_j^t(x) = D_j^t + \sum_{i=1}^{6} x_{i,j}^t$ is the total amount of natural gas available in market j annually throughout time period t. The functions $p_j^t(\xi^t)$ and $q_j^t(\xi^t)$ provide the base price and the base demand for natural gas as a function of the price of oil, and are defined as

$$p_j^t(\xi^t) = p0_j^t \left(\xi^t / or_t\right) \qquad \text{and} \qquad q_j^t(\xi^t) = q0_j^t \left(\xi^t / or_t\right)^{\eta_t}$$

with the following parameters:

- $p0_j^t$: reference price of natural gas in market j in time period t
- $q0_j^t$: reference demand for natural gas in market j in time period t
- or_t: reference price for oil in time period t
- η_t: the elasticity relating the relative demand for natural gas to the relative price of oil.

We assume that the prices of oil in each time period are independent and uniformly distributed with lower and upper bounds L_t and U_t. The values for the parameters in the base price and demand functions are given in Tables 3 and 4.

Table 3 Reference prices $p0$ and demands $q0$

Market	Period 1		Period 2		Period 3		Period 4	
BelLux	5.12	7.8	2.56	9.4	3.41	9.4	5.12	9.5
FRGer	5.27	40.7	2.64	46.2	3.52	46.5	5.27	44.6
France	5.25	23.6	2.62	28.3	3.50	9.8	5.25	28.5
Italy	5.15	25.3	2.57	34.9	3.43	37.5	5.15	37.2
Netherlands	5.16	28.9	2.58	29.9	3.44	32.2	5.16	29.7
UK	4.54	43.8	2.27	50.3	3.03	56.4	4.54	53.7

Table 4 Time period parameters in base price demand function

t	η_t	or_t	L_t	U_t
1	−0.10	30	16	34
2	−0.12	15	12	18
3	−0.24	30	24	36
4	−0.36	35	28	42

Assuming a fixed annual interest rate of $r = 0.1$, for each time period we use the factor f_t to express the future value of money with

$$f_t = \left(\frac{(1+r)^{y_t} - 1}{r}\right)\left(\frac{1}{(1+r)^{\sum_{s=1}^{t} y_s}}\right).$$

The net present value profit function for producer i is then defined to be

$$\Phi_i(x,\xi) = \sum_{t=1}^{4} f_t \left[\sum_{j=1}^{6}\left(P_j^t(x,\xi^t) - c_i^t\right)x_{i,j}^t - G_i(x)\right]. \tag{27}$$

Taking the expectation of (27) reduces to calculating $E\left[P_j^t(x,\xi^t)\right]$ and provides us with an expression for the expected profit function for player i. Under the assumption that the oil prices are uniformly distributed we have

$$E\left[P_j^t(x,\xi^t)\right] = \frac{p0_j^t\left(Q_j^t(x)q0_j^t\right)^{1/e_j^t} or_t^{\eta_t/e_j^t - 1}\left(U_t^{2-\eta_t/e_j^t} - L_t^{2-\eta_t/e_j^t}\right)}{(U_t - L_t)(2 - \eta_t/e_j^t)}.$$

Using knowledge of the true SVI we can observe that $d(f_0)_S(z_0)$ is a linear function and Σ_0 is degenerate but satisfies the alternative conditions required to prove Theorems 1 and 2. Since z_0 is in the interior of an n-cell for this example,

Table 5 Summary of coverage rates for $(z_0)_j$, $N = 200$ and $N = 2000$

Percentile	$N = 200$			$N = 2000$		
	$\upsilon_j^{.05}$ (%)	$\tilde{\eta}_j^{.04}$ (%)	$\tilde{\eta}_j^{.025}$ (%)	$\upsilon_j^{.05}$ (%)	$\tilde{\eta}_j^{.04}$ (%)	$\tilde{\eta}_j^{.025}$ (%)
MIN	88.05	88.70	89.05	94.60	95.70	97.20
Q1	94.60	95.57	97.07	94.85	96.00	97.43
MEDIAN	94.85	95.78	97.45	95.18	96.15	97.78
Q3	94.95	95.88	97.55	95.36	96.41	98.15
MAX	96.85	97.60	98.30	95.50	96.60	98.40

$$\lim_{N\to\infty} \Pr\left(K_{i(z_N)} = \text{cone}(P_N - \tilde{z}_{i_N})\right) = 1$$

and the intervals computed using $\tilde{\eta}_j^{\alpha_2}$ and $\tilde{h}_j^{\alpha_2}$ will be conservative for large sample sizes.

The performance of the obtained 95% confidence intervals is examined by generating 2000 replications at each sample size of $N = 200$, and 2000. To compute the width of an interval $\tilde{\eta}_j^{\alpha_2}$ or $\tilde{h}_j^{\alpha_2}$ we consider values of $\alpha_2 = 0.04, 0.025$ and $\alpha_1 = 0.05 - \alpha_2$. For each replication $\mathscr{K}_N^{\alpha_1}$ is chosen to contain a single cone determined using the approach described in Theorem 3 but with the set $R_{N,\varepsilon}$ replaced by the simultaneous confidence intervals (21). For all replications Σ_N is degenerate and the expression (23) is evaluated with $\varepsilon = 0$ and l_N determined by the number of eigenvalues of Σ_N greater than $N^{-1/3}$. In addition to the intervals computed using $\tilde{\eta}_j^{\alpha_2}$ and $\tilde{h}_j^{\alpha_2}$ we consider the intervals (12). Since z_0 is contained in the relative interior of an n-cell for this example the intervals (12) are asymptotically exact. These intervals are computed for a value of $\alpha = 0.05$ and we denote their half-widths by υ_j^{α}. Finally, we also consider individual confidence intervals for $(x_0)_j$ obtained from the projection of the intervals (12) onto the set S.

The performance of all the methods were largely in line with expectations. In Tables 5 and 6 we provide the five-number summaries for the coverage rates across all components of $(z_0)_j$ and $(x_0)_j$ for each interval. For both values of $\alpha_1 = 0.01$ and 0.025 the set of $\mathbb{R}^{96}$ was chosen for $\mathscr{K}_N^{\alpha_1}$ across all replications with $N = 2000$. For these replications the interval widths $\tilde{\eta}_j^{\alpha_2}$ reduced to the standard expressions used in (12). All differences in the interval widths for these samples were therefore due to the differences in the values of $\chi_1^2(0.05)$, $\chi_1^2(0.04)$, and $\chi_1^2(0.025)$, and the intervals computed using $\tilde{\eta}_j^{\alpha_2}$ were conservative in their coverage of $(z_0)_j$.

At the smaller sample size of $N = 200$, the estimates for $\mathscr{K}_N^{\alpha_1}$ were more conservative. For 689 replications with $\alpha_2 = 0.025$ and 941 replications with $\alpha_2 = 0.04$ the estimates of $\mathscr{K}_N^{\alpha_1}$ contained a cone generated from a face of P_N with dimension of $n - 2$ or below, and $\tilde{\eta}_j^{\alpha_2}$ lacked a closed form expression. For these replications the differences in $\tilde{\eta}_j^{\alpha_2}$ and υ_j^{α} were still largely driven by the different choices of α_2 and α. The ratio of $\tilde{\eta}_j^{0.04}$ and $\upsilon_j^{0.05}$ differed from $\sqrt{\chi_1^2(0.04)/\chi_1^2(0.05)}$ by no more than ten percent for 99.59% of the intervals for which $\tilde{\eta}_j^{0.04}$ lacked a closed form expression.

Table 6 Summary of coverage rates for $(x_0)_j$, $N = 200$ and $N = 2000$

Percentile	$N = 200$			$N = 2000$		
	$\upsilon_j^{.05}$ (%)	$\tilde{h}_j^{.04}$ (%)	$\tilde{h}_j^{.025}$ (%)	$\upsilon_j^{.05}$ (%)	$\tilde{h}_j^{.04}$ (%)	$\tilde{h}_j^{.025}$ (%)
MIN	88.20	88.70	89.05	94.60	95.70	97.20
Q1	94.75	95.70	97.08	94.85	96.00	97.50
MEDIAN	94.90	95.83	97.45	95.30	96.25	97.95
Q3	95.05	95.95	97.60	95.40	96.5	98.35
MAX	100	100	100	100	100	100

Similarly the ratio of $\tilde{\eta}_j^{0.025}$ and $\upsilon_j^{0.05}$ was within ten percent of $\sqrt{\chi_1^2(0.025)/\chi_1^2(0.05)}$ for 99.84% of the replications for which $\tilde{\eta}_j^{0.025}$ lacked a closed form expression. One can also include other candidate cones in $\mathscr{K}_N^{\alpha_1}$ when computing $\eta_j^{\alpha_2}$. Because the simultaneous confidence intervals for z_0 tend to be conservative in their coverage rates, cones generated from faces of P_N with dimension larger than the face identified using the approach of Theorem 3 are likely to be of the most interest. For this example that would suggest including the whole space $\mathbb{R}^{96}$, which in view of the above observations will not cause significant changes in the interval widths.

Overall at the sample size of $N = 200$ the proposed methods performed well. As was the case for the replications at the larger sample size, while the proposed methods were conservative in their coverage of $(z_0)_j$ this was not due to excessively wide intervals. The proposed intervals are therefore providing informative measures of the variability in each component of $(z_N)_j$, even with the conservative adjustments used to generalize the methods described in [26]. Among the 96 components of z_0, there were only two components, namely $(z_0)_{74}$ and $(z_0)_{83}$, for which the three computed intervals had coverages rates below 90% at the sample size of $N = 200$. There were 184 and 195 replications for which the signs of $(z_N)_{74}$ and $(z_0)_{74}$, and $(z_N)_{83}$ and $(z_0)_{83}$ differed. For these replications z_N and z_0 were not contained in the same n-cell and the intervals for these two components were negatively affected. This is reflected in the "MIN" row in Table 5 under $N = 200$.

In computing confidence intervals for $(x_0)_j$, we used the more conservative implementation described after Theorem 2, which will not return a value of $\tilde{h}_j^{\alpha_2} = 0$, to evaluate $\tilde{h}_j^{\alpha_2}$. Since $S = \mathbb{R}_+^{96}$, for this example $\tilde{h}_j^{\alpha_2} = \tilde{\eta}_j^{\alpha_2}$ for all components. The coverage of $(x_0)_j$ and $(z_0)_j$ therefore only differ for replications where $(z_N)_j < 0 = (x_N)_j$. As can be seen from Table 6, the coverage rates for most components meet or exceed the specified level. A majority of the intervals are conservative due to the nature of the proposed methods and the relaxed conditions. However, the intervals are not overly wide and provide useful information about the uncertainty for each individual component of the SAA solutions.

Acknowledgements Research of Michael Lamm and Shu Lu is supported by National Science Foundation under the Grants DMS-1109099 and DMS-1407241. Research of Michael Lamm took place during his graduate study in the Department of Statistics and Operations at the University of North Carolina at Chapel

Hill. We thank the three anonymous referees for comments and suggestions that have helped to improve the presentation of this paper.

References

1. Agdeppa, R.P., Yamashita, N., Fukushima, M.: Convex expected residual models for stochastic affine variational inequality problems and its application to the traffic equilibrium problem. Pac. J. Optim. **6**(1), 3–19 (2010)
2. Anitescu, M., Petra, C.: Higher-order confidence intervals for stochastic programming using bootstrapping. Technical Report ANL/MCS-P1964-1011, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL (2011)
3. Attouch, H., Cominetti, R., Teboulle, M.: Forward: Special issue on nonlinear convex optimization and variational inequalities. Math. Program. **116**(1–2), 1–3 (2009). https://doi.org/10.1007/s10107-007-0116-6
4. Chen, X., Fukushima, M.: Expected residual minimization method for stochastic linear complementarity problems. Math. Oper. Res. **30**, 1022–1038 (2005)
5. Chen, X., Pong, T.K., Wets, R.J.B.: Two-stage stochastic variational inequalities: an ERM-solution procedure (2015) (preprint)
6. Chen, X., Wets, R.J.B., Zhang, Y.: Stochastic variational inequalities: residual minimization smoothing sample average approximations. SIAM J. Optim. **22**(2), 649–673 (2012)
7. Chen, X., Zhang, C., Fukushima, M.: Robust solution of monotone stochastic linear complementarity problems. Math. Program. **117**, 51–80 (2009)
8. Dentcheva, D., Römisch, W.: Differential stability of two-stage stochastic programs. SIAM J. Optim. **11**(1), 87–112 (2000)
9. Dontchev, A.L., Rockafellar, R.T.: Characterizations of strong regularity for variational inequalities over polyhedral convex sets. SIAM J. Optim. **6**(4), 1087–1105 (1996)
10. Dupacova, J., Wets, R.: Asymptotic behavior of statistical estimators and of optimal solutions of stochastic optimization problems. Ann. Stat. **16**(4), 1517–1549 (1988)
11. Facchinei, F., Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems, vol. I. Springer, New York (2003)
12. Fang, H., Chen, X., Fukushima, M.: Stochastic R_0 matrix linear complementarity problems. SIAM J. Optim. **18**, 482–506 (2007)
13. Ferris, M.C., Pang, J.S.: Complementarity and Variational Problems: State of the Art. SIAM, Philadelphia (1997)
14. Ferris, M.C., Pang, J.S.: Engineering and economic applications of complementarity problems. SIAM Rev. **39**, 669–713 (1997)
15. Genz, A., Bretz, F.: Computation of Multivariate Normal and t Probabilities. Lecture Notes in Statistics. Springer, Heidelberg (2009)
16. Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F., Hothorn, T.: mvtnorm: multivariate normal and t distributions (2013). http://CRAN.R-project.org/package=mvtnorm. R package version 0.9-9996
17. Giannessi, F., Maugeri, A. (eds.): Variational Inequalities and Network Equilibrium Problems. Plenum Press, New York (1995)
18. Giannessi, F., Maugeri, A., Pardalos, P.M. (eds.): Equilibrium Problems: Nonsmooth Optimization and Variational Inequality Models. Nonconvex optimization and its applications, vol. 58. Kluwer Academic Publishers, Dordrecht (2001)
19. Gürkan, G., Pang, J.S.: Approximations of Nash equilibria. Math. Program. **117**(1–2), 223–253 (2009). https://doi.org/10.1007/10107-007-0156-y
20. Gürkan, G., Yonca Özge, A., Robinson, S.M.: Sample-path solution of stochastic variational inequalities. Math. Program. **84**, 313–333 (1999)
21. Harker, P.T., Pang, J.S.: Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms, and applications. Math. Program. **48**, 161–220 (1990)
22. Haurie, A., Zaccour, G., Legrand, J., Smeers, Y.: A stochastic dynamic nash-cournot model for the European gas market. Technical Report G-87-24, École des hautes études commerciales, Montréal, Québec, Canada (1987)
23. Huber, P.: The behavior of maximum likelihood estimates under nonstandard conditions. In: LeCam L., Neyman J. (eds.) Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics, pp. 221–233. University of California Press, Berkeley, CA (1967)

24. Jiang, H., Xu, H.: Stochastic approximation approaches to the stochastic variational inequality problem. IEEE Trans. Autom. Control **53**(6), 1462–1475 (2008)
25. King, A.J., Rockafellar, R.T.: Asymptotic theory for solutions in statistical estimation and stochastic programming. Math. Oper. Res. **18**, 148–162 (1993)
26. Lamm, M., Lu, S., Budhiraja, A.: Individual confidence intervals for true solutions to expected value formulations stochastic variational inequalities. Math. Prog. Ser. B **165**(1), 151–196 (2017)
27. Lan, G., Nemirovski, A., Shapiro, A.: Validation analysis of mirror descent stochastic approximation method. Math. Program. **134**(2), 425–458 (2012)
28. Linderoth, J., Shapiro, A., Wright, S.: The empirical behavior of sampling methods for stochastic programming. Ann. Oper. Res. **142**, 215–241 (2006)
29. Lu, S.: A new method to build confidence regions for solutions of stochastic variational inequalities. Optimization **63**(9), 1431–1443 (2014)
30. Lu, S.: Symmetric confidence regions and confidence intervals for normal map formulations of stochastic variational inequalities. SIAM J. Optim. **24**(3), 1458–1484 (2014)
31. Lu, S., Budhiraja, A.: Confidence regions for stochastic variational inequalities. Math. Oper. Res. **38**, 545–568 (2013)
32. Lu, S., Liu, Y., Yin, L., Zhang, K.: Confidence intervals and retions for the lasso by using stochastic variational inequality techniques in optimization. J. R. Stat. Soc. Ser. B **79**(2), 589–611 (2017)
33. Luo, M., Lin, G.: Expected residual minimization method for stochastic variational inequality problems. J. Optim. Theory Appl. **140**, 103–116 (2009)
34. Pang, J.: Newton's method for B-differentiable equations. Math. Oper. Res. **15**, 311–341 (1990)
35. Pang, J.S., Ralph, D.: Forward: Special issue on nonlinear programming, variational inequalities, and stochastic programming. Math. Program. **117**(1–2), 1–4 (2009). https://doi.org/10.1007/s10107-007-0169-6
36. Phelps, C., Royset, J.O., Gong, Q.: Optimal control of uncertain systems using sample average approximations. SIAM J. Control Optim. **54**(1), 1–29 (2016)
37. Ralph, D.: On branching numbers of normal manifolds. Nonlinear Anal. Theory Methods Appl. **22**, 1041–1050 (1994)
38. Robinson, S.M.: Strongly regular generalized equations. Math. Oper. Res. **5**(1), 43–62 (1980)
39. Robinson, S.M.: An implicit-function theorem for a class of nonsmooth functions. Math. Oper. Res. **16**(2), 292–309 (1991)
40. Robinson, S.M.: Normal maps induced by linear transformations. Math. Oper. Res. **17**(3), 691–714 (1992)
41. Robinson, S.M.: Sensitivity analysis of variational inequalities by normal-map techniques. In: Giannessi, F., Maugeri, A. (eds.) Variational Inequalities and Network Equilibrium Problems, pp. 257–269. Plenum Press, New York (1995)
42. Rockafellar, R.T., Wets, R.J.B.: Stochastic variational inequalities: single-stage to multistage. Math. Program. Ser. B **165**(1), 331–360 (2017)
43. Römisch, W.: Stability of stochastic programming problems. In: Ruszczyński, A., Shapiro, A. (eds.) Handbooks in Operations Research and Management Science, vol. 10, pp. 483–554. Elsevier, Amsterdam (2003)
44. Scholtes, S.: Introduction to Piecewise Differentiable Equations. Springer, New York (2012)
45. Shapiro, A.: Asymptotic behavior of optimal solutions in stochastic programming. Math. Oper. Res. **18**, 829–845 (1993)
46. Shapiro, A., Dentcheva, D., Ruszczyński, A.P.: Lectures on Stochastic Programming: Modeling and Theory. Society for Industrial and Applied Mathematics and Mathematical Programming Society, Philadelphia, PA (2009)
47. Shapiro, A., Homem-de Mello, T.: On the rate of convergence of optimal solutions of monte carlo approximations of stochastic programs. SIAM J. Optim. **11**(1), 70–86 (2000)
48. Shapiro, A., Xu, H.: Stochastic mathematical programs with equilibrium constraints, modeling and sample average approximation. Optimization **57**, 395–418 (2008)
49. Stefanski, L.A., Boos, D.D.: The calculus of M-estimation. Am. Stat. **56**(1), 29–38 (2002)
50. Vogel, S.: Universal confidence sets for solutions of optimization problems. SIAM J. Optim. **19**(3), 1467–1488 (2008)
51. Wald, A.: Note on the consitency of the maximum likelihood estimate. Ann. Math. Stat. **20**, 595–601 (1949)

52. Xu, H.: Sample average approximation methods for a class of stochastic variational inequality problems. Asia Pac. J. Oper. Res. **27**(1), 103–119 (2010)
53. Yin, L., Lu, S., Liu, Y.: Confidence intervals for sparse penalized regression with random designs (2015) (**submitted for publication**)
54. Zhang, C., Chen, X., Sumlee, A.: Robust Wardrop's user equilibrium assignment under stochastic demand and supply. Transp. Res. B **45**(3), 534–552 (2011)