On Distributionally Robust Generalized Nash Games Defined over the Wasserstein Ball

JOTA manuscript No.
(will be inserted by the editor)

On distributionally robust generalized Nash games defined over the Wasserstein ball

Filippo Fabiani, Barbara Franci

Received: date / Accepted: date

Abstract In this paper we propose an exact, deterministic, and fully continuous reformulation of generalized Nash games characterized by the presence of soft coupling constraints in the form of distributionally robust (DR) joint chance-constraints (CCs). We first rewrite the underlying uncertain game introducing mixed-integer variables to cope with DR-CCs, where the integer restriction actually amounts to a binary decision vector only, and then extend it to an equivalent deterministic problem with one additional agent handling all those introduced variables. Successively we show that, by means of a careful choice of tailored penalty functions, the extended deterministic game with additional agent can be equivalently recast in a fully continuous setting.

Keywords Uncertain Nash games · Distributionally robust optimization · Mixed-integer programming · Exact penalty functions

Mathematics Subject Classification (2000) 91A15 · 90C15 · 90C11 · more

1 Introduction

Among available multi-agent modelling paradigms, game theory [16] holds the promise of covering a wide variety of potential behaviours emerging both at an individual and a collective level, since it inherently captures the selfish nature of the competitive, yet interacting, agents involved. This reason, together with

Filippo Fabiani, Corresponding author
IMT School for Advanced Studies Lucca
Piazza San Francesco 19, 55100 Lucca, Italy
filippo.fabiani@imtlucca.it

Barbara Franci
Department of Advanced Computing Sciences, Maastricht University
P.O. Box 616, NL–6200 MD Maastricht, The Netherlands
b.franci@maastrichtuniversity.nl

offering an elegant mathematical framework to model applications ranging from economics and finance [25, 44], to engineering [10, 23, 28], biology and social sciences [7, 17], has recently fostered a renewed interest in game theory.

Facing real-world applications in a game-theoretic context often requires tackling the associated technical issues complicating the analysis of the game at hand, such as for instance the presence of uncertainty affecting the main parameters of the game [20, 32]. This is the central focus of this paper, where we will consider generalized Nash games characterized by the presence of *soft coupling constraints* in the form of distributionally robust (DR) joint chance constraints (CCs). The term "soft" here refers to the fact that DR–CCs allow for some violation (in probability) of uncertain constraints, as opposed to more conservative formulations adopted in the robust optimization literature. In addition, this type of DR–CCs often offer a data-driven solution to handle CCs in an almost agnostic scenario, namely when neither the probability distribution nor the support set of the uncertain variable are available, and require that all the uncertain constraints shall be simultaneously satisfied for all the probability distributions within a prescribed ambiguity set [34, 43]. Crucial is hence the choice of such an ambiguity set. In particular, we will assume that this latter is represented in the form of a Wasserstein ball, which has recently attracted a lot of research interest due to its asymptotic consistency properties enabling for tractable, i.e., convex, reformulations [5, 24].

1.1 Related work

The analysis of games affected by the presence of uncertainty has attracted several research efforts in the last few years, essentially because random variables can appear in different ways both in the agents' cost functions and constraints. According to the richness of information available on the underlying uncertainty, different techniques have also been employed to characterize the robustness of the resulting game equilibrium solutions.

A well-studied framework, for instance, considers an expected value formulation of uncertain cost functions, thus implicitly entailing that the agents are risk-neutral as they look for average performance [14, 15, 35]. On the contrary, worst-case approaches [13] are usually more conservative.

The largest number of contributions, however, focus on the case of random variables affecting the feasible set of the game [2, 11, 12]. In the specific case in which the problem at hand allows for some constraint violation, one may adopt for example a CC–based formulation. This latter establishes that the considered uncertain constraints are satisfied with high probability, typically quantified through a prescribed confidence level to be fixed [31, 37, 38]. Existence of equilibria for CC–games has been recently addressed in [31, 37]. Due to their not-immediate tractability, however, reformulations of CC have also been proposed as a second order cone [40], or mixed-integer [1] optimization problems for instance. Besides, this type of CC–based framework requires one full knowledge on the probability distribution characterizing the uncertain variable. Relaxing this (often unpractical) requirement directly leads to the DR counterpart of CC–games, where the distribution of the uncertain parameters is itself uncertain [38]. For DR–games, the existence of equilibria has been recently investigated in [30, 38], while, in case of finite strategy sets, it was also studied in [27, 29, 31, 38, 39].

A distinct feature of (typically nonconvex) DR–games is represented by the ambiguity set, i.e., the set of possible probability distributions for the uncertainty affecting the game, which plays a fundamental role in DR optimization and related tractability. Different formulations are in fact available, which rely on the so-called Wasserstein ball [6, 24, 29, 45], as well as on assumptions on mean and covariance [38,43]. The choice of the ambiguity set also allows one to make connections with risk-averse optimization [34, 45] and to the possibility of determining convex reformulations of the problem [24, 43]. Most of these works, however, do not consider a game-theoretic setup. For a comprehensive collection of results in distributionally robust optimization and its connection with game theory, we refer to the survey papers [2, 34].

1.2 Summary of contribution

In contrast with the aforementioned works, we propose an exact, deterministic, and fully continuous reformulation of generalized Nash games characterized by the presence of nonconvex coupling constraints in the form of DR–CCs. In particular we show that, by suitably tuning two design parameters, any equilibrium solution to a DR–CCs generalized Nash game coincides with that of a deterministic game with an additional dummy agent. First, we rewrite the original, uncertain game by introducing mixed-integer (MI) variables to cope with DR–CCs, where the integer restrictions are actually binary ones. Then, we extend the resulting game to an equivalent deterministic problem with one additional agent handling those MI variables. In case the dummy agent chooses tailored penalty functions to relax the imposed binary restrictions, then we show the extended deterministic game with additional agent can be equivalently recast (and hence solved) in a fully continuous setting.

Notation. Let $\mathbb{R}$ indicate the set of real numbers and let $\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$. $\mathbb{B}$ indicates the binary set $\{0, 1\}$. $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ denotes the standard inner product and $\|\cdot\|$ represents the associated Euclidean norm with related dual $\| \cdot \|_*$. Id_K indicates the identity operator of dimension K. With $\mathbf{1}_K$ (respectively, $\mathbf{0}_K$) we denote the column vector with K elements all equal to 1 (resp., 0). The operator $\mathrm{col}(\cdot)$ stacks its arguments in column vectors or matrices of compatible dimensions. Given N column vectors $x_1, \ldots, x_N \in \mathbb{R}^n$, we denote with $\boldsymbol{x} \coloneqq \mathrm{col}\,(x_1, \ldots, x_N) = \left[x_1^\top, \ldots, x_N^\top\right]^\top$. For a given matrix $A \in \mathbb{R}^{m \times n}$ (respectively, vector $b \in \mathbb{R}^m$), we denote with $A_{h,:}$ (resp., b_h) the h-th row of matrix A (resp., element of vector b), as opposed to $A_{:,h}$ that instead identifies the h-th column of A. Given a set $\mathcal{I}$, $|\mathcal{I}|$ indicates its cardinality. Furthermore, we denote with

$$d_{\mathrm{W}}(\mathbb{P}_1, \mathbb{P}_2) = \inf_{\mathbb{P} \in \mathcal{P}(\mathbb{P}_1, \mathbb{P}_2)} \mathbb{E}_{\mathbb{P}}[\|\tilde{\boldsymbol{\xi}}_1 - \tilde{\boldsymbol{\xi}}_2\|]$$

the standard Wasserstein distance between two distributions $\mathbb{P}_1$ and $\mathbb{P}_2$ on $\mathbb{R}^K$. It can be interpreted as the minimal transportation cost of moving $\mathbb{P}_1$ to $\mathbb{P}_2$ supposing that the cost of moving a Dirac point mass from $\boldsymbol{\xi}_1$ to $\boldsymbol{\xi}_2$ is $\|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2\|$, where $\tilde{\boldsymbol{\xi}}_1 \sim \mathbb{P}_1, \tilde{\boldsymbol{\xi}}_2 \sim \mathbb{P}_2$, and $\mathcal{P}(\mathbb{P}_1, \mathbb{P}_2)$ represents the set of all distributions on $\mathbb{R}^K \times \mathbb{R}^K$ with marginals $\mathbb{P}_1$ and $\mathbb{P}_2$.

2 Problem statement

In this paper we consider generalized Nash games with ℓ-soft coupling constraints modelled in the form of joint CCs, $\mathcal{H} := \{1, \ldots, \ell\}$, which mathematically read as:

$$\forall i \in \mathcal{N} : \begin{cases} \min\limits_{x_i \in \mathcal{X}_i} & J_i(x_i, \boldsymbol{x}_{-i}) \\ \text{s.t.} & \mathbb{P}\left[\max\limits_{h \in \mathcal{H}} \{A_{h,:}\boldsymbol{x} - b_h(\xi)\} > 0\right] \leq \varepsilon \,, \end{cases} \tag{1}$$

where $x_i \in \mathbb{R}^{n_i}$ denotes the local decision variable of agent $i \in \mathcal{N} := \{1, \ldots, N\}$, belonging to the local constraint set $\mathcal{X}_i \subseteq \mathbb{R}^{n_i}$. We let $\boldsymbol{x}_{-i} := \mathrm{col}((x_j)_{j \in \mathcal{N} \setminus \{i\}}) \in \mathbb{R}^{n-n_i}$ stack the decision variable of all the agents but the i-th one, $n := \sum_{i \in \mathcal{N}} n_i$ and analogously, $\mathcal{X}_{-i} = \prod_{j \in \mathcal{N} \setminus \{i\}} \mathcal{X}_j$. With a slight abuse of notation we define the collective strategy profile as $\boldsymbol{x} := \mathrm{col}((x_j)_{j \in \mathcal{N}}) = (x_i, \boldsymbol{x}_{-i}) \in \mathcal{X} = \prod_{i \in \mathcal{N}} \mathcal{X}_i$. Additionally, $J_i : \mathbb{R}^n \to \mathbb{R}$ represents the local cost function each agent aims at minimizing, while $A \in \mathbb{R}^{\ell \times n}$ and $b : \mathbb{R}^l \to \mathbb{R}^\ell$ describe the collection of constraints and shared resources coupling the decisions of the agents taking part to the game. Each agent aims at securing the largest portion of shared resources to achieve individual performance specifications, according to some metric identified by $J(\cdot)$.

Here, we assume the vector of shared resources b being affected by uncertainty $\xi \in \Xi \subseteq \mathbb{R}^l$, distributed according to a (possibly unknown) probability distribution $\mathbb{P}$. This makes $A\boldsymbol{x} - b(\xi) \leq 0$ a constraint to be satisfied in probability, thus nonconvex. As opposed to traditional robust approaches [4] where this type of constraints are required to be met with probability one, possibly yielding overly conservative solutions, with (joint) CC one usually relaxes this hard requirement by allowing for a certain degree of constraint violations, controlled through a safety parameter $\varepsilon \in [0, 1)$. We refer to (1) as a CC generalized Nash game. Next, we make an assumption that will hold throughout:

Standing Assumption 1 *For all $i \in \mathcal{N}$, $\mathcal{X}_i$ is nonempty and compact.* □

As frequently happens in practice, however, an explicit expression for the probability $\mathbb{P}$ and the support set Ξ might be unknown, and even if it was, verifying a CC in the form (1) is known to be NP-hard under generic distributions [26]. On the other hand, in many situations we could have available a set of K independent and identically distributed (i.i.d.) realizations of the uncertainty, $\{\hat{\xi}^{(k)}\}_{k\in\mathcal{K}}$, $\mathcal{K} \coloneqq \{1,\ldots,K\}$, in the form of historical data. For instance, this could be the case in power and energy systems [3, 8], demand-side management [42], traffic coordination and control [9, 48]. The available uncertainty realizations, i.e., the samples, enable us to consider a DR counterpart of the CC generalized Nash game in (1) to edge against possible shifts in the probability distribution. In particular, we can design an ambiguity set through the Wasserstein ball $\mathcal{B}_\theta$ centred around the empirical distribution $\hat{\mathbb{P}}_K \coloneqq \frac{1}{K}\sum_{k\in\mathcal{K}} \delta_{\hat{\xi}^{(k)}}$ with radius $\theta > 0$, i.e.,

$$\mathcal{B}_\theta(\hat{\mathbb{P}}_K) \coloneqq \left\{ \mathbb{Q} \in \mathscr{M}(\Xi) \;\middle|\; d_{\mathrm{W}}(\hat{\mathbb{P}}_K, \mathbb{Q}) \le \theta \right\}.$$

Here, $\mathscr{M}(\Xi)$ denotes the space of all probability distributions supported on Ξ and $d_{\mathrm{W}} : \mathscr{M}(\Xi) \times \mathscr{M}(\Xi) \to \mathbb{R}_{\ge 0}$ represents the Wasserstein distance between two probability distributions [41]. The radius θ is typically chosen in a data-driven fashion [24] and, as in traditional distributionally robust optimization, regulates the conservativeness of the solution obtained (if any).

The DR generalized Nash game associated to the CC game in (1) hence reads as:

$$\forall i \in \mathcal{N} : \begin{cases} \min\limits_{x_i \in \mathcal{X}_i} & J_i(x_i, \boldsymbol{x}_{-i}) \\ \text{s.t.} & \sup\limits_{\mathbb{Q}\in\mathcal{B}_\theta(\hat{\mathbb{P}}_K)} \mathbb{Q}\left[\max\limits_{h\in\mathcal{H}}\{A_{h,:}\boldsymbol{x} - b_h(\xi)\} > 0\right] \le \varepsilon\,. \end{cases} \tag{2}$$

By introducing $\mathcal{X} \coloneqq \prod_{i\in\mathcal{N}} \mathcal{X}_i$, we say that a collective strategy profile $\bar{\boldsymbol{x}}$ is *feasible* if i) $\bar{\boldsymbol{x}} \in \mathcal{X}$, and ii) it satisfies the DR–CC, i.e., the worst-case probability distribution $\mathbb{Q}$ over $\mathcal{B}_\theta(\hat{\mathbb{P}}_K)$ satisfies $\mathbb{Q}\left[\max_{h\in\mathcal{H}}\{A_{h,:}\bar{\boldsymbol{x}} - b_h(\xi)\} > 0\right]$ with confidence at most ε. The nonconvex feasible set characterizing the DR generalized Nash game in (2) can therefore be defined as:

$$\mathcal{F} \coloneqq \left\{ \boldsymbol{x} \in \mathcal{X} \;\middle|\; \sup\limits_{\mathbb{Q}\in\mathcal{B}_\theta(\hat{\mathbb{P}}_K)} \mathbb{Q}\left[\max\limits_{h\in\mathcal{H}}\{A_{h,:}\boldsymbol{x} - b_h(\xi)\} > 0\right] \le \varepsilon \right\}.$$

In the considered framework we are interested in discussing the (almost) standard notion of generalized Nash equilibrium (GNE) formalized next:

Definition 2.1 A collective strategy profile $\boldsymbol{x}^\star \in \mathcal{F}$ is a DR–GNE of the DR–generalized Nash game in (2) if, for all $i \in \mathcal{N}$, $J_i(x_i^\star, \boldsymbol{x}_{-i}^\star) \le J_i(z_i, \boldsymbol{x}_{-i}^\star)$ for all z_i such that $(z_i, \boldsymbol{x}_{-i}^\star) \in \mathcal{F}$. □

To proceed further with our analysis, we shall necessarily postulate the existence of an equilibrium for the DR–generalized Nash game in (2).

Standing Assumption 2 *The* DR*–generalized Nash game in* (2) *admits at least one* DR–GNE. □

Remark 2.1 In some particular instances, integrating Standing Assumption 1 with the convexity of each set $\mathcal{X}_i$, and assuming further that each cost function $J_i(\cdot)$ is continuous in both arguments and, for all $\boldsymbol{x}_i \in \mathcal{X}_{-i}$, $x_i \mapsto J_i(x_i, \boldsymbol{x}_{-i})$ is convex, guarantees that at least an equilibrium exists for the DR–generalized Nash game in (2) – see, e.g., [30, 31, 37]. In the general framework considered here, however, identifying conditions ensuring the existence of at least an equilibrium in the spirit of Definition 2.1 is still an open problem. □

We note that the feasible set of the DR generalized Nash game in (2) is a conservative subset of the feasible region characterizing the CC game in (1), and therefore any equilibrium solution obtained from the DR counterpart of (1), according to Definition 2.1, directly satisfies the same optimality conditions onto a smaller feasible region, thus amounting to a local CC–GNE for (1). In fact, the collective feasible region of the DR game in (2) is smaller than that of the CC game in (1) as we are requiring to satisfy the linear constraints with respect to all the probability distributions falling in the ambiguity set $\mathcal{B}_\theta(\hat{\mathbb{P}}_K)$. This is in accordance to standard distributionally robust paradigms [24, 43], since we are working under the assumption that the true distribution $\mathbb{P}$ is not available. In particular, the main source of conservativeness here is due to the size of the ambiguity set containing the true, yet unknown, probability distribution $\mathbb{P}$, and hence the size of the radius θ employed to design a Wasserstein ball around the empirical distribution $\hat{\mathbb{P}}_K$.

3 An equivalent reformulation

Albeit inherently nonconvex, it has been shown in [6] that, whenever the metric determining the Wasserstein distance is induced by a linear norm, i.e., $\|\cdot\|_1$ or $\|\cdot\|_\infty$, and the uncertainty ξ affects linearly the vector of shared resources b, e.g.,

$$b(\xi) \coloneqq D\xi + d$$

with $D \in \mathbb{R}^{\ell \times l}$ and $d \in \mathbb{R}^\ell$, the DR-CC in (2) can be equivalently encoded through a collection of MI linear inequalities. In particular, we note that the resulting DR generalized Nash game in (2) will turn out to be a deterministic Nash game with binary variables entering in the constraints only. By letting $D \neq 0$ (otherwise the CC could be absorbed directly in $\mathcal{X}$), we are then entitled to equivalently consider the following collection of MI optimization programs:

$$\forall i \in \mathcal{N} : \begin{cases} \min\limits_{x_i, t, \boldsymbol{s}, \boldsymbol{q}} & J_i(x_i, \boldsymbol{x}_{-i}) \\ \text{s.t.} & \bar{A}_i x_i + \sum\limits_{j \in \mathcal{N} \setminus \{i\}} \bar{A}_j x_j + Ct - E\boldsymbol{s} - mE\boldsymbol{q} \leq \bar{D}\hat{\boldsymbol{\xi}} + \bar{d}\,, \\ & \boldsymbol{1}_K^\top \boldsymbol{s} - \varepsilon K t \leq -\theta K\,, \\ & \boldsymbol{1}_K t - \boldsymbol{s} - m(\boldsymbol{1}_K - \boldsymbol{q}) \leq 0\,, \\ & x_i \in \mathcal{X}_i,\ \boldsymbol{s} \geq 0,\ \boldsymbol{q} \in \mathbb{B}^K\,, \end{cases} \tag{3}$$

with $\bar{A}_i \coloneqq (A_{:,i} \otimes \mathbf{1}_K)$, for all $i \in \mathcal{N}$, $C \coloneqq (L \otimes \mathbf{1}_K)$ with $L \coloneqq \mathrm{col}((\|D_{h,:}\|_*)_{h\in\mathcal{H}})$, $E \coloneqq (L \otimes \mathrm{Id}_K)$, $\bar{D} \coloneqq (D \otimes I_K)$, $\hat{\boldsymbol{\xi}} \coloneqq \mathrm{col}((\hat{\xi}^{(k)})_{k\in\mathcal{K}})$ and $\bar{d} \coloneqq (d \otimes \mathbf{1}_K)$. The parameter $m > 0$ represents a tuning term that will be key to show the equivalence. Note that, with the intermediate reformulation in (3) the auxiliary continuous and binary variables $t \in \mathbb{R}$, $\boldsymbol{s} \coloneqq \mathrm{col}((s_k)_{k\in\mathcal{K}}) \in \mathbb{R}^K_+$ and $\boldsymbol{q} \coloneqq \mathrm{col}((q_k)_{k\in\mathcal{K}}) \in \mathbb{B}^K$ are momentarily (i.e., before proposing the final reformulation) considered as "common" among agents, since their role is to take care of the DR–CC coupling the decisions of the population of agents. In [45], it is shown that the DR–CC reformulation in (3) also admits an elegant interpretation in terms of conditional value-at-risk (CVaR).

Next, we note that the MI generalized Nash game in (3) can be massaged into a modified MI game including an additional, dummy agent $N+1$, whose strategy decision process is complicated by the presence of binary restrictions. Specifically, grouping together the elements in (3) involving t, $\boldsymbol{s}$ and $\boldsymbol{q}$ (i.e., matrices, constraints) and letting $x_{N+1} \coloneqq \mathrm{col}(t, \boldsymbol{s}, \boldsymbol{q})$ yields the following collection of $N+1$ optimization problems:

$$\forall i \in \mathcal{N} : \begin{cases} \min\limits_{x_i \in \mathcal{X}_i} & J_i(x_i, \boldsymbol{x}_{-i}) \\ \text{s.t.} & \bar{A}_i x_i + \sum\limits_{j\in\mathcal{N}\setminus\{i\}} \bar{A}_j x_j + m\bar{A}_{N+1} x_{N+1} \le \bar{D}\hat{\boldsymbol{\xi}} + \bar{d}\,, \end{cases} \tag{4}$$

$$i = N+1 : \begin{cases} \min\limits_{x_{N+1} \in \mathcal{X}_{N+1}} & 0 \\ \text{s.t.} & \sum\limits_{j\in\mathcal{N}} \bar{A}_j x_j + m\bar{A}_{N+1} x_{N+1} \le \bar{D}\hat{\boldsymbol{\xi}} + \bar{d}\,, \end{cases} \tag{5}$$

where $\mathcal{X}_{N+1} \coloneqq \left\{x_{N+1} \in \mathbb{R} \times \mathbb{R}^K_{\ge 0} \times \mathbb{B}^K \mid mPx_{N+1} \le v\right\} \subseteq \mathbb{R}^{n_{N+1}}$, $n_{N+1} \coloneqq 2K+1$, and matrix P and vector v collect the (local) constraints in (3) involving t, $\boldsymbol{s}$ and $\boldsymbol{q}$ only. Specifically, they have the following structures: $\bar{A}_{N+1} = \left[\frac{1}{m}C \;\; -\frac{1}{m}E \;\; E\right] \in \mathbb{R}^{\ell K \times (2K+1)}$,

$$P = \begin{bmatrix} -\frac{1}{m}\varepsilon K & \frac{1}{m}\mathbf{1}_K^\top & \mathbf{0}_K^\top \\ \frac{1}{m}\mathbf{1}_K & -\frac{1}{m}\,\mathrm{Id}_K & \mathrm{Id}_K \end{bmatrix} \mathbb{R}^{(1+K)\times(2K+1)}, \text{ and } v = \begin{bmatrix} -\theta K \\ m\mathbf{1}_K \end{bmatrix} \in \mathbb{R}^{1+K}.$$

It hence turns out that the $(N+1)$-th agent only is affected by integer restrictions, i.e., aims at solving an MI feasibility problem (5), while the original set of agents is facing a standard generalized Nash game (4) and it has *only* to make decisions within the new feasible set, parametrized by $m > 0$:

$$\mathcal{F}_m \coloneqq \left\{(\boldsymbol{x}, x_{N+1}) \in \mathcal{X} \times \mathcal{X}_{N+1} \;\middle|\; \bar{A}\boldsymbol{x} + m\bar{A}_{N+1}x_{N+1} \le \bar{D}\hat{\boldsymbol{\xi}} + \bar{d}\right\}.$$

Tackling optimization problems with MI variables or binary restrictions is known to be challenging. Available approaches in the literature, such as [18, 19, 22, 33, 36, 47, 49], require one to design a tailored (typically concave) penalty function to relax the integrality restrictions, thus allowing one to recover an entirely continuous problem at the price of possibly minimizing a nonconvex cost function. Thus, in its most general formulation the MI optimization problem in (5) characterizing agent $N+1$ could read as

$$i = N+1: \begin{cases} \min\limits_{x_{N+1} \in \hat{\mathcal{X}}_{N+1}} & r_{N+1}(x_{N+1}) + p_{N+1}(x_{N+1}, \epsilon) \\ \text{s.t.} & \sum\limits_{j \in \mathcal{N}} \bar{A}_j x_j + m\bar{A}_{N+1} x_{N+1} \leq \bar{D}\hat{\boldsymbol{\xi}} + \bar{d}\,, \end{cases} \tag{6}$$

where $\hat{\mathcal{X}}_{N+1}$ denotes the relaxed version of $\mathcal{X}_{N+1}$, which is thus defined over $[0,1]^K$ rather than on $\mathbb{B}^K$, $r_{N+1} : \mathbb{R}^{n_{N+1}} \to \mathbb{R}$ is some regularization function, while $p_{N+1} : \mathbb{R}^{n_{N+1}} \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ is a penalty function with tolerance $\epsilon > 0$. We make the following technical assumptions on the cost functions just introduced:

Standing Assumption 3 *The following conditions hold true:*

(i) The function $r_{N+1} : \mathbb{R}^{n_{N+1}} \to \mathbb{R}$ is bounded on $\hat{\mathcal{X}}_{N+1}$, and there exists an open set $\mathcal{A} \supset \mathcal{X}_{N+1}$ and real numbers α, $\beta > 0$, such that, for all x_{N+1}, $y_{N+1} \in \mathcal{A}$, $r_{N+1}(\cdot)$ satisfies the following condition:

$$|r_{N+1}(x_{N+1}) - r_{N+1}(y_{N+1})| \leq \beta \|x_{N+1} - y_{N+1}\|^{\alpha}.$$

(ii) The function $p_{N+1} : \mathbb{R}^{n_{N+1}} \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ satisfies the following conditions:
  (a) For all $x_{N+1}, y_{N+1} \in \mathcal{X}_{N+1}$, and $\epsilon > 0$, $p_{N+1}(x_{N+1}, \epsilon) = p_{N+1}(y_{N+1}, \epsilon)$;
  (b) There exists a value $\hat{\epsilon} > 0$ and, for all $z_{N+1} \in \mathcal{X}_{N+1}$, there exists a neighborhood $\mathcal{S}(z_{N+1})$ such that for all $x_{N+1} \in \mathcal{S}(z_{N+1}) \cap (\hat{\mathcal{X}}_{N+1} \setminus \mathcal{X}_{N+1})$, and $\epsilon \in (0, \hat{\epsilon}]$, $p(x_{N+1}, \epsilon) - p(z_{N+1}, \epsilon) \geq \hat{\beta} \|x_{N+1} - z_{N+1}\|^{\alpha}$, where $\hat{\beta} > \beta$ and α as in Standing Assumption 3.(i). In addition, by introducing $\mathcal{S} = \bigcup_{z_{N+1} \in \mathcal{X}_{N+1}} \mathcal{S}(z_{N+1})$, there exists $\hat{x}_{N+1} \notin \mathcal{S}$ such that:

$$\begin{cases} \lim\limits_{\epsilon \to 0} \left(p(\hat{x}_{N+1}, \epsilon) - p(z_{N+1}, \epsilon)\right) = +\infty, \ \textit{for all } z_{N+1} \in \mathcal{X}_{N+1}, \\ p(x_{N+1}, \epsilon) \geq p(\hat{x}_{N+1}, \epsilon), \ \textit{for all } x_{N+1} \in \cap(\hat{\mathcal{X}}_{N+1} \setminus \mathcal{S}), \epsilon > 0. \end{cases}$$

□

We refer to [21, 22] for a comprehensive discussion on functions satisfying Standing Assumption 3 and to Remark 3.1 for some examples. Then, by letting $J_{N+1}(\boldsymbol{x}, x_{N+1}) \coloneqq r_{N+1}(x_{N+1}) + p_{N+1}(x_{N+1}, \epsilon)$, we are now ready to relate the equilibria of the original DR generalized Nash game in (2) with those of the new Nash game (4) with dummy agent $N+1$ choosing a relaxed strategy (6) as follows:

Theorem 3.1 *There exist finite values $\tilde{m}$, $\tilde{\epsilon} > 0$ such that, for all $m \geq \tilde{m}$ and $\epsilon \in (0, \tilde{\epsilon}]$, if some $(\boldsymbol{x}^\star, x_{N+1}^\star)$ is an equilibrium solution to the generalized Nash game* (4)–(6)*, then $\boldsymbol{x}^\star \in \mathcal{F}$ and amounts to a* DR–GNE *of the* DR *generalized Nash game in* (2)*.* □

Proof To prove the result we essentially show that the following equivalences separately hold true:

$$\text{DR game}\,(2) \overset{\text{(i)}}{\Longleftrightarrow} \text{MI game}\,(3) \overset{\text{(ii)}}{\Longleftrightarrow} \text{augmented Nash game}\,(4)-(6)\,,$$

where, with a slight abuse of notation, the equivalence symbol indicates that the games have the same equilibria. (i) To prove this we essentially need to show that a large enough, though finite, value of $m > 0$ is sufficient to make the reformulation (3) of (2) exact, according to [6, Prop. 2]. On the other hand, this implicitly requires one to show that the quantities involved in the MI optimization problems in (3) are bounded. The fact that some $\boldsymbol{x}^\star$ which solves (3), together with some $\mathrm{col}(t^\star, \boldsymbol{s}^\star, \boldsymbol{q}^\star)$, coincides with a DR–GNE of the DR generalized Nash game in (2) then follows by noting that each cost function in (2) is not affected by the underlying reformulation. Then, by rewriting the MI constraints in (3) for each component we obtain

$$\begin{cases} \dfrac{D_{h,:}\hat{\xi}^{(k)} + d_h - A_{h,:}\boldsymbol{x}}{\|D_{h,:}\|_*} + mq_k \geq t - s_k\,,\ \forall h \in \mathcal{H},\, k \in \mathcal{K}\,, \\ \varepsilon K t - \mathbf{1}_K^\top \boldsymbol{s} \geq \theta K\,, \\ t - s_k - m(1 - q_k) \leq 0\,,\ \forall k \in \mathcal{K}\,. \end{cases} \tag{7}$$

In the spirit of the reformulation proposed in [6, §2.4], at optimality $D_{h,:}\hat{\xi}^{(k)} + d_h - A_{h,:}\boldsymbol{x} < 0$ for all $h \in \mathcal{H}$ implies that $q_k = 1$, otherwise $q_k = 0$. Essentially, the role of each binary element q_k is to flag those samples at which violation of the coupling constraints occurs, thereby activating the less restrictive between the first and the third constraint in (7). Then, we first note that the second constraint in (7) implicitly forces the variable t to be nonnegative, since $\mathbf{1}_K^\top \boldsymbol{s} + \theta K$ is always positive.

In addition, for a given pair $(\boldsymbol{x}, t)$, combining together the first and the third constraint in (7) with $s_k \geq 0$ allows us to obtain a closed-form expression characterizing s_k at optimality when $m > 0$ takes large enough values, i.e.,

$$s_k^\star(\boldsymbol{x}, t) = \mathrm{proj}_{\mathbb{R}_{\geq 0}}[t - \min_{h\in\mathcal{H}}\{(D_{h,:}\hat{\xi}^{(k)} + d_h - A_{h,:}\boldsymbol{x})/\|D_{h,:}\|_*\}]. \tag{8}$$

This latter expression is bounded from above provided that t is so, since $\boldsymbol{x}$ takes values in the compact set $\mathcal{X}$, $D \neq 0$, and hence $(D_{h,:}\hat{\xi}^{(k)} + d_h - A_{h,:}\boldsymbol{x})/\|D_{h,:}\|_*$ is bounded, for all h and k (each uncertainty realization is, indeed, such that $\hat{\xi}^{(k)} < \infty$). For the sake of contradiction, let us thus assume that t can grow indefinitely. Then, for a sufficiently large, yet finite, value of t we have that $\boldsymbol{s}$ is large as well, according to its definition in (8). Hence, from the second relation in (7), we obtain $\varepsilon K t - \mathbf{1}_K^\top \boldsymbol{s} \sim \varepsilon K t - Kt = -(1-\varepsilon)Kt \geq \theta K > 0$, which is violated as $\varepsilon \in [0, 1)$ and $t \geq 0$. It hence follows that also t shall be bounded from above, thus concluding this part.

(ii) Once fixed $m \geq \tilde{m}$ to some finite value, showing the equivalence between the MI generalized Nash game in (3) and the new game with an additional, dummy agent in (4)–(5) is straightforward. We indeed note that, in view of Standing Assumption 1 and the result just proved, for any $\boldsymbol{x} \in \mathcal{X}$ the decision variables of the $(N+1)$-th agent take bounded values on $\mathcal{X}_{N+1}$, and hence also on $\hat{\mathcal{X}}_{N+1}$. Moreover, the collective feasible sets described in (3) and (4)–(5) are the same, hence, if $\boldsymbol{x}^\star$ is an equilibrium of (3), i.e., for all $i \in \mathcal{N}$, $J_i(x_i^\star, \boldsymbol{x}_{-i}^\star) \leq J_i(z_i, \boldsymbol{x}_{-i}^\star)$ for all z_i such that $(z_i, \boldsymbol{x}_{-i}^\star)$ satisfies the constraints in (3), then it is also such for (4)–(5), i.e., for all $i \in \mathcal{N} \cup \{N+1\}$, $J_i(x_i^\star, \boldsymbol{x}_{-i}^\star) \leq J_i(z_i, \boldsymbol{x}_{-i}^\star)$ for all z_i such that $(z_i, \boldsymbol{x}_{-i}^\star) \in \mathcal{F}_m$.

Therefore, with the conditions on the cost functions $r_{N+1}(\cdot)$ and $p_{N+1}(\cdot, \epsilon)$ reported in Standing Assumption 3, we can rely on [21, Th. 2.1] to show that there exists some finite $\tilde{\epsilon} > 0$ such that, for all $0 < \epsilon \leq \tilde{\epsilon}$, the relaxed version of the MI problem in (5), i.e., (6), has the same minima of the original program. This consideration concludes the chain of equivalences, and thus the proof. ■

Remark 3.1 Some examples of penalty functions that can be used to relax the MI generalized Nash game in (4)–(5) and turn it into the continuous game (4)–(6) can be found in [22,36]. The most used penalty function in the binary case is

$$p(x_{N+1}, \varepsilon) = \frac{1}{\varepsilon} \sum_{i \in \mathcal{I}_q} x_{N+1,i}(1 - x_{N+1,i}), \tag{9}$$

which dates back to [18, 19, 33]. Since the only binary variable of agent $N+1$ is $\boldsymbol{q}$, we denote with $\mathcal{I}_q$ the set of indices corresponding to $\boldsymbol{q}$ in the decision variable $x_{N+1} = \mathrm{col}(t, \boldsymbol{s}, \boldsymbol{q})$. For this specific choice, bounds on the penalty parameter ε can be obtained [19] that depend, among other quantities, on the maximum and minimum value attained by the cost function of agent $N+1$. Although these bounds can be obtained for a single optimization problem (at least in the linear case [18, 19]), computing it for the game in (4)–(6) might not be immediate since the value of ε is also indirectly related to the constant m appearing in the constraints, and hence it is subject of current research.

Other penalty functions can be found in [46], which however mostly consider binary variables in $\{-1, 1\}^K$. For the $\{0, 1\}^K$-case, the possible penalty functions rely on the fact that $\boldsymbol{q} \in \mathbb{B}^K$ if and only if $\boldsymbol{q} \in \left\{\boldsymbol{q} \in \mathbb{R}^K \mid 0 \leq \boldsymbol{q} \leq 1, 0 \leq \boldsymbol{w} \leq 1, \langle 2\boldsymbol{q} - 1, 2\boldsymbol{w} - 1\rangle = n_q, \forall \boldsymbol{w}\right\}$ which equivalently happens if and only if $\boldsymbol{q} \in \left\{\boldsymbol{q} \in \mathbb{R}^K \;\middle|\; 0 \leq \boldsymbol{q} \leq 1, \|2\boldsymbol{w} - 1\|_2^2 \leq n_q, \langle 2\boldsymbol{q} - 1, 2\boldsymbol{w} - 1\rangle = n_q, \forall \boldsymbol{w}\right\}$, $n_q = |\mathcal{I}_q|$ [46, Section 4]. It follows that

$$p(x_{N+1}, \varepsilon) = \frac{1}{\varepsilon}\left(n_q - \sum_{i \in \mathcal{I}_q} (2x_{N+1,i} - 1)(2w_i - 1)\right).$$

Note that, in this case, the variable $\boldsymbol{w}$ becomes part of the decision variable of agent $N+1$, i.e., $x_{N+1} = \mathrm{col}(t, \boldsymbol{s}, \boldsymbol{q}, \boldsymbol{w})$, and its constraints ($0 \leq \boldsymbol{w} \leq 1$ or $\|2\boldsymbol{w} - 1\|_2^2 \leq n_q$) should be included in the local constraint set $\hat{\mathcal{X}}_{N+1}$ in (6). □

4 Conclusion

In this paper we have proposed an exact, deterministic, and fully continuous reformulation of a generalized Nash game characterized by distributionally robust, joint chance-constraints. Our results leverage the fact that the original problem can be rewritten first as a mixed-integer generalized Nash game with one additional agent in case of a Wasserstein ambiguity set, and then as a continuous one after a careful choice of a suitable penalty function.

As a final remark, we stress that the proposed results can be further improved by finding appropriate bounds for the design parameter needed to tune the penalty term, which is non-trivial in a game-theoretic context. However, this analysis lies outside the preliminary study carried out in this paper, and therefore we leave it for future work.

References

1. L. Adam and M. Branda. Nonlinear chance-constrained problems: Optimality conditions, regularization and solvers. *Journal of Optimization Theory and Applications*, 170, 2016.
2. M. Aghassi and D. Bertsimas. Robust game theory. *Mathematical Programming*, 107(1):231–273, 2006.
3. A. Arrigo, C. Ordoudis, J. Kazempour, Z. De Grève, J.-F. Toubeau, and F. Vallée. Wasserstein distributionally robust chance-constrained optimization for energy and reserve dispatch: An exact and physically-bounded formulation. *European Journal of Operational Research*, 296(1):304–322, 2022.
4. A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust Optimization*, volume 28. Princeton University Press, 2009.
5. D. Bertsimas, V. Gupta, and N. Kallus. Data-driven robust optimization. *Mathematical Programming*, 167(2):235–292, 2018.
6. Z. Chen, D. Kuhn, and W. Wiesemann. Data-driven chance constrained programs over wasserstein balls. *Operations Research*, 2022.
7. A. M. Colman. *Game Theory and its Applications: In the Social and Biological Sciences*. Psychology Press, 2013.
8. C. Duan, L. Jiang, W. Fang, J. Liu, and S. Liu. Data-driven distributionally robust energy-reserve-storage dispatch. *IEEE Transactions on Industrial Informatics*, 14(7):2826–2836, 2017.
9. F. Fabiani. Pursuing robust decisions in uncertain traffic equilibrium problems. In *2021 60th IEEE Conference on Decision and Control (CDC)*, pages 5116–5121. IEEE, 2021.
10. F. Fabiani and S. Grammatico. Multi-vehicle automated driving as a generalized mixed-integer potential game. *IEEE Transactions on Intelligent Transportation Systems*, 21(3):1064–1073, 2019.
11. F. Fabiani, K. Margellos, and P. J. Goulart. On the robustness of equilibria in generalized aggregative games. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pages 3725–3730. IEEE, 2020.
12. F. Fabiani, K. Margellos, and P. J. Goulart. Probabilistic feasibility guarantees for solution sets to uncertain variational inequalities. *Automatica*, 137:110120, 2022.
13. F. Fele and K. Margellos. Probably approximately correct Nash equilibrium learning. *IEEE Transactions on Automatic Control*, 66(9):4238–4245, 2020.
14. B. Franci and S. Grammatico. A distributed forward–backward algorithm for stochastic generalized Nash equilibrium seeking. *IEEE Transactions on Automatic Control*, 66(11):5467–5473, 2020.
15. B. Franci and S. Grammatico. Stochastic generalized Nash equilibrium seeking in merely monotone games. *IEEE Transactions on Automatic Control*, 2021.

16. D. Fudenberg and J. Tirole. *Game Theory*. MIT press, 1991.
17. S. Hummert, K. Bohl, D. Basanta, A. Deutsch, S. Werner, G. Theißen, A. Schroeter, and S. Schuster. Evolutionary game theory: cells as players. *Molecular BioSystems*, 10(12):3044–3065, 2014.
18. B. Kalantari and J. B. Rosen. Penalty for zero–one integer equivalent problem. *Mathematical Programming*, 24(1):229–232, 1982.
19. B. Kalantari and J. B. Rosen. Penalty formulation for zero-one nonlinear programming. *Discrete Applied Mathematics*, 16(2):179–182, 1987.
20. M. J. Kochenderfer. *Decision Making Under Uncertainty: Theory and Application*. MIT press, 2015.
21. S. Lucidi and F. Rinaldi. Exact penalty functions for nonlinear integer programming problems. *Journal of Optimization Theory and Applications*, 145(3):479–488, 2010.
22. S. Lucidi and F. Rinaldi. An exact penalty global optimization approach for mixed-integer programming problems. *Optimization Letters*, 7(2):297–307, 2013.
23. Z. Ma, D. S. Callaway, and I. A. Hiskens. Decentralized charging control of large populations of plug-in electric vehicles. *IEEE Transactions on Control Systems Technology*, 21(1):67–78, 2011.
24. P. Mohajerin Esfahani and D. Kuhn. Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming*, 171(1):115–166, 2018.
25. R. B. Myerson. Nash equilibrium and the history of economic theory. *Journal of Economic Literature*, 37(3):1067–1082, 1999.
26. A. Nemirovski and A. Shapiro. Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17(4):969–996, 2007.
27. H. N. Nguyen, A. Lisser, and V. V. Singh. Random games under elliptically distributed dependent joint chance constraints. *Journal of Optimization Theory and Applications*, 195:249–264, October 2022.
28. A. Orda, R. Rom, and N. Shimkin. Competitive routing in multiuser communication networks. *IEEE/ACM Transactions on Networking*, 1(5):510–521, 1993.
29. G. Peng, T. Zhang, and Q. Zhu. A data-driven distributionally robust game using Wasserstein distance. In *International Conference on Decision and Game Theory for Security*, pages 405–421. Springer, 2020.
30. S. Peng, A. Lisser, V. V. Singh, N. Gupta, and E. Balachandar. Games with distributionally robust joint chance constraints. *Optimization Letters*, 15(6):1931–1953, 2021.
31. S. Peng, V. V. Singh, and A. Lisser. General sum games with joint chance constraints. *Operations Research Letters*, 46(5):482–486, 2018.
32. D. Prelec and G. Loewenstein. Decision making over time and under uncertainty: A common approach. *Management Science*, 37(7):770–786, 1991.
33. M. Raghavachari. On connections between zero-one integer programming and concave programming under linear constraints. *Operations Research*, 17(4):680–684, 1969.
34. H. Rahimian and S. Mehrotra. Distributionally robust optimization: A review. *arXiv preprint arXiv:1908.05659*, 2019.
35. U. Ravat and U. V. Shanbhag. On the characterization of solution sets of smooth and nonsmooth convex stochastic Nash games. *SIAM Journal on Optimization*, 21(3):1168–1199, 2011.
36. F. Rinaldi. New results on the equivalence between zero-one programming and continuous concave programming. *Optimization Letters*, 3(3):377–386, 2009.
37. V. V. Singh, O. Jouini, and A. Lisser. Existence of Nash equilibrium for chance-constrained games. *Operations Research Letters*, 44(5):640–644, 2016.
38. V. V. Singh, O. Jouini, and A. Lisser. Distributionally robust chance-constrained games: Existence and characterization of Nash equilibrium. *Optimization Letters*, 11(7):1385–1405, 2017.
39. V. V. Singh and A. Lisser. A characterization of Nash equilibrium for the games with random payoffs. *Journal of Optimization Theory and Applications*, 178(3):998–1013, September 2018.
40. V. V. Singh and A. Lisser. A second-order cone programming formulation for two player zero-sum games with chance constraints. *European Journal of Operational Research*, 275(3):839–845, 2019.

41. C. Villani. *Topics in Optimal Transportation*, volume 58. American Mathematical Society, 2021.
42. L. Wang, C. Jiang, K. Gong, R. Si, H. Shao, and W. Liu. Data-driven distributionally robust economic dispatch for distribution network with multiple microgrids. *IET Generation, Transmission & Distribution*, 14(24):5712–5719, 2020.
43. W. Wiesemann, D. Kuhn, and M. Sim. Distributionally robust convex optimization. *Operations Research*, 62(6):1358–1376, 2014.
44. D. E. Wildasin. Nash equilibria in models of fiscal competition. *Journal of Public Economics*, 35(2):229–240, 1988.
45. W. Xie. On distributionally robust chance constrained programs with Wasserstein distance. *Mathematical Programming*, 186(1):115–155, 2021.
46. G. Yuan and B. Ghanem. Binary optimization via mathematical programming with equilibrium constraints. *arXiv preprint arXiv:1608.04425*, 2016.
47. G. Yuan and B. Ghanem. An exact penalty method for binary optimization based on MPEC formulation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, 2017.
48. Y. Zhang, Z. Zhang, A. Lim, and M. Sim. Robust data-driven vehicle routing with time windows. *Operations Research*, 69(2):469–485, 2021.
49. W. X. Zhu. Penalty parameter for linearly constrained 0–1 quadratic programming. *Journal of Optimization Theory and Applications*, 116(1):229–239, 2003.