



# Strong Nash equilibria and mixed strategies

Eleonora Braggion<sup>1</sup> · Nicola Gatti<sup>1</sup> · Roberto Lucchetti<sup>1</sup> · Tuomas Sandholm<sup>2</sup> · Bernhard von Stengel<sup>3</sup>

Accepted: 17 June 2020 / Published online: 6 July 2020  
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

## Abstract

We study *strong Nash equilibria* in mixed strategies in finite games. A Nash equilibrium is strong if no coalition of players can jointly deviate so that all players in the coalition get strictly better payoffs. Our main result concerns games with two players and states that if a game admits a strong Nash equilibrium, then the payoff pairs in the support of the equilibrium lie on a straight line in the players' utility space. As a consequence, the set of games that have a strong Nash equilibrium in which at least one player plays a mixed strategy has measure zero. We show that the same property holds for games with more than two players, already when no coalition of two players can profitably deviate. Furthermore, we show that, in contrast to games with two players, in a strong Nash equilibrium an outcome that is strictly Pareto dominated may occur with positive probability.

**Keywords** Noncooperative games · Strong Nash equilibrium · Mixed strategies · Pareto efficiency

---

✉ Nicola Gatti  
nicola.gatti@polimi.it

Eleonora Braggion  
eleonora.braggion@lastminute.com

Roberto Lucchetti  
roberto.lucchetti@polimi.it

Tuomas Sandholm  
sandholm@cs.cmu.edu

Bernhard von Stengel  
b.von-stengel@lse.ac.uk

<sup>1</sup> Politecnico di Milano, Milan, Italy

<sup>2</sup> Carnegie Mellon University, Pittsburgh, USA

<sup>3</sup> London School of Economics and Political Science, London, UK

## 1 Introduction

It is well known that in a non-cooperative game, selfish behavior can cause players to be worse off than they could be by collaborating. The most famous example is the Prisoners' Dilemma (Luce and Raiffa 1957), where strictly dominating strategies for the players lead to a bad outcome for both. The *strong Nash equilibrium* by Aumann (1959) gets around this paradox, as a solution concept that is resilient against coalitional deviations. A strategy profile is a strong Nash equilibrium if no coalition of players can jointly deviate so that all players in the coalition get strictly better payoffs (because this applies to single-player coalitions, it is a Nash equilibrium). Strong Nash equilibrium outcomes are also called *weakly Pareto efficient* for each coalition (Miettinen 1999, Definition 2.5.1). A further refinement is *super strong Nash equilibrium* (Rozenfeld 2007), which requires Pareto efficiency for every coalition (that is, no coalition can improve a player's payoff without hurting at least one other member of the coalition). There are classes of games that have a strong Nash equilibrium but no super strong Nash equilibrium (Gourvès and Monnot 2009), so the distinction between the two solution concepts is meaningful.

The strong Nash equilibrium concept is commonly criticized as too demanding because it allows for unlimited private communication among the players. Moreover, in many games, a strong Nash equilibrium does not exist. For these reasons, among others, relaxations have been proposed. In this paper, in addition to the strong Nash equilibrium, we study a relaxation called *k-strong Nash equilibrium*. In an  $n$ -player game, a  $k$ -strong Nash equilibrium is a Nash equilibrium where no coalition of  $k$  or fewer players can deviate so that all its members strictly benefit (Andelman et al. 2009). The rationale is that in many practical situations only small coalitions can be formed, for example in a network that connects the players. Another relaxation is *coalition-proof Nash equilibrium*, which is a Nash equilibrium that is resilient against those coalitional deviations that are self-enforcing; coalition-proof Nash equilibria can be Pareto inefficient (Bernheim et al. 1987, Table 4). Another generalization of strong Nash equilibrium is *strong correlated equilibrium*, in which the presence of a correlating device allows players to use correlated strategies (Einy and Peleg 1995). Similarly, in Milgrom and Roberts (1996), coalitions that can jointly deviate are also allowed to correlate their mixed strategies in the first place. However, none of these relaxations guarantee existence in general finite games.

Since strong Nash equilibrium is an appealing solution concept and it may not exist, a crucial question is whether or not a given game admits such an equilibrium (Aumann 1959; Conitzer and Sandholm 2008; Nessah and Tian 2014) give existence conditions for continuous convex games. In this paper, we show that, in the space of games with a given finite number of players and strategies for each player, the set of games with a strong Nash equilibrium where at least one player randomizes over at least two pure strategies has measure zero. Dubey (1986) considers generic games with smooth payoff functions, and shows that interior Nash equilibria are generically not Pareto efficient and therefore not "super strong" in our terminology, called "strong" by Dubey. He notes that, as a consequence, super strong Nash equilibria of finite games exist generically only in pure strategies. Our results are similar, but

apply to strong Nash equilibria, and are proved with direct geometric arguments about mixed strategies.

Most of our study concerns two-player games. We give a precise description of games that admit strong and super strong Nash equilibria. For two players, the payoff vectors are points in two-dimensional space. We show that in a strong Nash equilibrium in mixed strategies, the payoff vectors in the support of the equilibrium lie on a straight line which has negative slope or is vertical or horizontal. If that line has negative slope, then the game restricted to the support of the equilibrium is *strictly competitive* (and therefore, via a positive-affine change of payoffs, equivalent to a zero-sum game, see Adler et al. 2009). If the line is vertical or horizontal, then in the equilibrium support one player is indifferent among all actions.

For games with three or more players, we strengthen the result by Dubey (1986) by showing that the set of games that have a 2-strong Nash equilibrium such that at least one player randomizes over at least two pure strategies has measure zero. Moreover, we show that these games can have strong and super strong Nash equilibria in which the game restricted to the support of the equilibrium may contain outcomes that are strictly Pareto dominated; this is significantly different to the two-player case.

## 2 Games with two players

We always consider finite games with a fixed number of players and strategies per player. In this section, we consider two-player games, where player 1 has  $m$  pure strategies, always denoted by  $i = 1, \dots, m$ , and player 2 has  $n$  pure strategies  $j = 1, \dots, n$ . This defines an  $m \times n$  bimatrix game  $(A, B)$  with payoff pairs  $(a_{ij}, b_{ij})$  for each row  $i$  and column  $j$ . Mixed strategies of player 1 and 2 are denoted by letters  $x$  and  $y$  with mixed strategy probabilities  $x_i$  and  $y_j$ , respectively.

All vectors are column vectors. The all-zero vector is denoted by  $\mathbf{0}_m$  in  $\mathbb{R}^m$  and by  $\mathbf{0}_n$  in  $\mathbb{R}^n$ . Similarly, the all-one vector is denoted by  $\mathbf{1}_m$  or  $\mathbf{1}_n$ . The  $i$ th *unit vector* (of any dimension) that has component 1 in row  $i$  and zeros otherwise is denoted by  $e_i$ . Inequalities between vectors like  $x \geq \mathbf{0}_m$  or  $x > \mathbf{0}_m$  hold between all components. The transpose of a column vector  $x$  is the row vector  $x^T$ . Scalars are treated like  $1 \times 1$  matrices in matrix multiplication, and are therefore multiplied to the right to column vectors and to the left to row vectors. We sometimes write pairs in  $\mathbb{R}^2$  as row vectors  $(a, b)$ .

For definiteness, we recall the notions of strong, super strong, and  $k$ -strong Nash equilibrium.

**Definition 1** A Nash equilibrium is called *strong* if no coalition of players can deviate from their mixed strategies so that all members of the coalition strictly benefit. It is called *super strong* if no coalition of players can deviate so that all members weakly benefit, and at least one member strictly benefits. If either property holds for all coalitions with up to  $k$  members, then the equilibrium is called  $k$ -strong or  $k$ -super-strong, respectively.

Thus, 1-strong equilibria are just Nash equilibria. Strong equilibria were introduced by Aumann (1959). The terminology “super strong” is due to Rozenfeld (2007), called “strong” by Dubey (1986). The property of super strong equilibrium means that for any coalition of players, the payoffs to the coalition members are Pareto optimal.

We first consider fully mixed equilibria. These are equilibria with full support, that is, every pure strategy is played with positive probability.

**Lemma 1** *Consider a fully mixed Nash equilibrium  $(x, y)$  of  $(A, B)$  where the payoffs in  $A$  and  $B$  are shifted (by adding a constant to all payoffs in  $A$ , and another constant to all payoffs in  $B$ ) so that both equilibrium payoffs are zero. Then*

$$Ay = \mathbf{0}_m, \quad x^\top B = \mathbf{0}_n^\top. \tag{1}$$

*Suppose  $(x, y)$  is super strong (that is, the equilibrium payoffs are Pareto-optimal). Then*

$$By = \mathbf{0}_m, \quad x^\top A = \mathbf{0}_n^\top. \tag{2}$$

*If  $(x, y)$  is strong then at least one of the conditions (2) holds.*

**Proof** We repeatedly use the simple observation that  $x > \mathbf{0}_m$  and  $x^\top v = 0$  for some non-zero vector  $v$  in  $\mathbb{R}^m$  imply that  $v$  has at least one positive (and at least one negative) component.

Condition (1) holds by the best response condition for Nash equilibria because the equilibrium has full support. In particular,  $Ay = \mathbf{0}_m$  implies that any strategy of player 1 is a best response to  $y$ .

We now consider the players’ own expected payoffs. Suppose  $By \neq \mathbf{0}_m$ . Then  $(By)_i > 0$  for some  $i$  because  $x^\top (By) = 0$ . Therefore, player 1 can help player 2 by playing the pure strategy  $i$ , given by the  $i$ th unit vector  $e_i$  in  $\mathbb{R}^m$ . Then  $e_i^\top Ay = 0$  and  $e_i^\top By = (By)_i > 0$ , so player 1 benefits weakly and player 2 benefits strictly by the change from  $(x, y)$  to  $(e_i, y)$ . Hence  $(x, y)$  is not super strong. The same applies if  $x^\top A \neq \mathbf{0}_n^\top$  and therefore  $(x^\top A)_j > 0$  for some  $j$  where player 2 helps player 1 by playing  $j$ . This shows that (2) holds if  $(x, y)$  is super strong.

Suppose  $(x, y)$  is strong and neither condition in (2) holds. We show that the players can change to a mixed strategy pair  $(\hat{x}, \hat{y})$  so that both players strictly benefit, a contradiction. As before, consider pure strategies (unit vectors)  $e_i$  in  $\mathbb{R}^m$  and  $e_j$  in  $\mathbb{R}^n$  so that  $(By)_i > 0$  and  $(x^\top A)_j > 0$ . For  $\delta$  and  $\varepsilon$  in  $\mathbb{R}$  let

$$\hat{x} = x(1 - \delta) + e_i \delta, \quad \hat{y} = y(1 - \varepsilon) + e_j \varepsilon \tag{3}$$

which for sufficiently small  $|\delta|$  and  $|\varepsilon|$  are mixed strategies because, for example if  $\delta < 0$ , we have  $x_i > 0$  and  $x_k < 1$  for  $k \neq i$ . Then

$$\begin{aligned}
 \hat{x}^T A \hat{y} &= \hat{x}^T A y (1 - \epsilon) + \hat{x}^T A e_j \epsilon \\
 &= 0 + ((1 - \delta)x^T A e_j + \delta e_i^T A e_j) \epsilon \\
 &= ((1 - \delta)(x^T A)_j + \delta a_{ij}) \epsilon \\
 &= ((x^T A)_j + \delta(a_{ij} - (x^T A)_j)) \epsilon
 \end{aligned}
 \tag{4}$$

which for sufficiently small  $|\delta|$  is positive for any  $\epsilon > 0$ . Similarly,

$$\begin{aligned}
 \hat{x}^T B \hat{y} &= (1 - \delta)x^T B \hat{y} + \delta e_i B \hat{y} \\
 &= \delta e_i B (y (1 - \epsilon) + e_j \epsilon) \\
 &= \delta((By)_i + (b_{ij} - (By)_i) \epsilon)
 \end{aligned}
 \tag{5}$$

which for sufficiently small  $|\epsilon|$  is positive for any  $\delta > 0$ . So both players strictly benefit for sufficiently small positive  $\delta$  and  $\epsilon$ , as claimed.

As will be used later, note that if  $(By)_i < 0$ , then in the same manner player 1 can improve player 2’s payoff by *reducing* the mixed strategy probability  $x_i$  by choosing  $\delta < 0$  in (3) and (5). □

Lemma 1 raises the question if  $(x, y)$  can be a strong equilibrium if only one of the conditions (2) hold. According to the following lemma, this is only possible when one player has a constant payoff matrix. Then any Nash equilibrium is trivially strong because the coalition of both players cannot deviate so that both players strictly benefit.

**Lemma 2** Consider a fully mixed Nash equilibrium  $(x, y)$  of  $(A, B)$  with equilibrium payoffs  $(0, 0)$  so that (1) holds, let

$$x^T A = \mathbf{0}_n^T, \quad By \neq \mathbf{0}_m,
 \tag{6}$$

and let  $(x, y)$  be a strong equilibrium. Then  $A$  is the all-zero matrix.

**Proof** Let  $i$  be any pure strategy of player 1 and  $(By)_i \neq 0$ . We claim the  $i$ th row  $e_i^T A$  of  $A$  is the zero row  $\mathbf{0}_n$ . Suppose otherwise that  $a_{ij} \neq 0$  for some  $j$ . Consider  $\hat{x}$  and  $\hat{y}$  as in (3). Then (4) and  $x^T A = \mathbf{0}_n^T$  imply

$$\hat{x}^T A \hat{y} = \delta a_{ij} \epsilon.
 \tag{7}$$

We choose  $|\delta|$  and  $|\epsilon|$  small enough with  $\delta$  of the same sign as  $(By)_i$  so that  $\hat{x}^T B \hat{y} > 0$  in (5), and  $\epsilon$  of the same sign as  $\delta a_{ij}$  so that  $\hat{x}^T A \hat{y} > 0$  in (7). Then both players strictly benefit by playing  $(\hat{x}, \hat{y})$ , that is,  $(x, y)$  is not strong, which proves the claim.

Hence, if  $(By)_i \neq 0$  for all  $i$ , then  $A$  is the all-zero matrix. It remains to consider the case where  $By \neq \mathbf{0}_m$  and  $By$  has some zero component. Let, say,  $(By)_1 > 0$  and  $(By)_i = 0$  for some  $i$ , say  $i = 2$ . Suppose that  $e_2^T A \neq \mathbf{0}_n$ , with  $a_{2j} > 0$  for some column  $j$ , which exists because  $(e_2^T A)y = e_2^T (Ay) = 0$ . We again show that then  $(x, y)$  is not strong, by letting player 1 help player 2 by playing the first two rows, and player 2 increasing  $y_j$ . Let

$$\hat{x} = e_1 \frac{1}{2} + e_2 \frac{1}{2}, \quad \hat{y} = y(1 - \epsilon) + e_j \epsilon. \tag{8}$$

Then, using  $Ay = \mathbf{0}_m$  and  $e_1^\top A = \mathbf{0}_n$  (because  $(By)_1 > 0$ ),

$$\begin{aligned} \hat{x}^\top A \hat{y} &= \hat{x}^\top Ay(1 - \epsilon) + \hat{x}^\top A e_j \epsilon \\ &= \left( \frac{1}{2} e_1^\top A e_j + \frac{1}{2} e_2^\top A e_j \right) \epsilon \\ &= \frac{1}{2} a_{2j} \epsilon \end{aligned} \tag{9}$$

and, using  $e_1^\top By = (By)_1 > 0$  and  $e_2^\top By = (By)_2 = 0$ ,

$$\begin{aligned} \hat{x}^\top B \hat{y} &= \left( \frac{1}{2} e_1^\top B + \frac{1}{2} e_2^\top B \right) (y(1 - \epsilon) + e_j \epsilon) \\ &= \frac{1}{2} (e_1^\top By(1 - \epsilon) + e_1^\top B e_j \epsilon + e_2^\top B e_j \epsilon) \\ &= \frac{1}{2} ((By)_1 + [b_{1j} + b_{2j} - (By)_1] \epsilon) \end{aligned} \tag{10}$$

which for small enough positive  $\epsilon$  is positive, as well as  $\hat{x}^\top A \hat{y}$  in (9), so that both players benefit strictly. This shows the desired contradiction when  $A$  has any non-zero row. So  $A$  is the all-zero matrix as claimed.  $\square$

Hence, for a fully mixed strong equilibrium  $(x, y)$  of a bimatrix game  $(A, B)$  with non-constant payoff matrices, (2) holds.

In preparation for Theorem 1 below we need the following lemma, which is about points in the plane that have the origin in the interior of their convex hull.

**Lemma 3** *Consider points  $u^1, \dots, u^s$  in  $\mathbb{R}^2$ , not all on a line through the origin  $(0, 0)$ , and positive  $\alpha_1, \dots, \alpha_s \in \mathbb{R}$  with  $\alpha_1 + \dots + \alpha_s = 1$  so that  $\alpha_1 u^1 + \dots + \alpha_s u^s = (0, 0)$ . Then there is some  $\epsilon > 0$  so that  $(\epsilon, \epsilon)$  is the convex combination of only one or two of these points.*

**Proof** Let  $C$  be the convex hull of the points  $u^1, \dots, u^s$ . Because  $\alpha_j > 0$  for all  $j$ , the origin  $(0, 0)$  belongs to the relative interior of  $C$ , which is the interior of  $C$  because by assumption  $C$  is not a line segment. Let  $\epsilon = \max\{t \mid (t, t) \in C\}$ , where  $\epsilon > 0$ . Clearly  $(\epsilon, \epsilon)$  belongs to the boundary of  $C$ , which is a polygon, and is therefore either a vertex or on an edge of  $C$ , which proves the claim.  $\square$

**Theorem 1** *Let  $(x, y)$  be a fully mixed strong Nash equilibrium of  $(A, B)$  with equilibrium payoffs  $(0, 0)$ , where neither  $A$  nor  $B$  is a constant matrix. Then for all  $i, j$ , the pairs  $(a_{ij}, b_{ij})$  in  $\mathbb{R}^2$  lie on a line with negative slope  $-\lambda$  through the origin, for some  $\lambda > 0$ . That is,  $b_{ij} = -\lambda a_{ij}$  for all  $i, j$ , and the game  $(\lambda A, B)$  is zero-sum.*

**Proof** By Lemmas 1 and 2, conditions (1) and (2) hold. Consider the pairs  $u_{ij} = (a_{ij}, b_{ij})$  in  $\mathbb{R}^2$ , arranged in an  $m \times n$  matrix in the same way as  $A$  and  $B$ . We first show that any two points  $u_{ij}$  and  $u_{il}$  in the same row  $i$  are on a line through the origin,

that is, are linearly dependent. If not, then  $e_i^\top Ay = 0$  and  $e_i^\top By = 0$  by (1) and (2) with  $y > \mathbf{0}_n$ , so that Lemma 3 applied to the points  $u_{i1}, \dots, u_{im}$  shows that  $(\epsilon, \epsilon)$  for some  $\epsilon > 0$  is a convex combination of  $u_{ij}$  and  $u_{il}$ , say, which defines a mixed strategy  $\hat{y}$  that mixes columns  $j$  and  $l$ . Together with the pure strategy  $i$  of player 1, it gives the positive payoffs  $(e_i^\top A\hat{y}, e_i^\top B\hat{y}) = (\epsilon, \epsilon)$  for both players which shows that  $(x, y)$  is not a strong Nash equilibrium. Similarly, any two points  $u_{ij}$  and  $u_{kj}$  in the same column  $j$  are linearly dependent.

Next, suppose that not all the points  $u_{ij}$  are on a single line through the origin. Because  $(x^\top Ay, x^\top By) = (0, 0)$  and  $x > \mathbf{0}_m$  and  $y > \mathbf{0}_n$ , the origin is a convex combination of these points  $u_{ij}$  with the positive coefficients  $x_i y_j$ , and we again apply Lemma 3. Then

$$(\epsilon, \epsilon) = (1 - \beta)u_{ij} + \beta u_{kl} \tag{11}$$

with  $\epsilon > 0$  for some  $i, j, k, l$  and  $\beta \in [0, 1]$ . If  $u_{ij} > (0, 0)$  or  $u_{kl} > (0, 0)$ , then both players obtain positive payoffs by playing  $(i, j)$  or  $(k, l)$ , so this is not the case, and  $0 < \beta < 1$ . Hence, the two points  $u_{ij}$  and  $u_{kl}$  are in different quadrants of  $\mathbb{R}^2$  and not in the strictly positive quadrant, and the line through them does not contain  $(0, 0)$  because it contains  $(\epsilon, \epsilon)$ . (Observe that  $a_{ij}b_{ij} \leq 0$  and  $a_{kl}b_{kl} \leq 0$ , and that this sign pattern would be preserved if  $(a_{ij}, b_{ij})$  and  $(a_{kl}, b_{kl})$  were scalar multiples of each other and would therefore apply to  $(1 - \beta)(a_{ij}, b_{ij}) + \beta(a_{kl}, b_{kl})$  which contradicts (11).) That is,  $u_{ij}$  and  $u_{kl}$  are linearly independent, where necessarily  $i \neq k$  and  $j \neq l$ . This implies that  $u_{il} = u_{kj} = (0, 0)$  because otherwise both  $u_{ij}$  and  $u_{kl}$  would be scalar multiples of  $u_{il}$  or  $u_{kj}$ . Then the mixed strategies  $\hat{x} = e_i(1 - \beta) + e_k\beta$  and  $\hat{y} = e_j\frac{1}{2} + e_l\frac{1}{2}$  give expected payoffs  $(\hat{x}^\top A\hat{y}, \hat{x}^\top B\hat{y}) = (\frac{\epsilon}{2}, \frac{\epsilon}{2})$ , which again shows that  $(x, y)$  is not a strong Nash equilibrium.

So all pairs  $u_{ij} = (a_{ij}, b_{ij})$  in  $\mathbb{R}^2$  lie on a line through the origin. This line is neither vertical nor horizontal since neither  $A$  nor  $B$  is a constant matrix, and it cannot have positive slope because otherwise there would be at least one point  $u_{ij} > (0, 0)$  and therefore  $(x, y)$  would not be a strong Nash equilibrium. So the slope of the line is negative, as claimed. □

We explicitly observe that as a consequence of Theorem 1, in a two-player game where no player is indifferent over all outcomes, a strong Nash equilibrium in fully mixed strategies is automatically a super strong Nash equilibrium.

By restricting Theorem 1 to the support of a mixed equilibrium, we obtain the following theorem as a corollary. It deals with the case of strong Nash equilibria where at least one player plays a mixed, not pure, strategy.

**Theorem 2** *Let  $(A, B)$  be a bimatrix game. Consider a strong Nash equilibrium such that at least one player plays a mixed, non-pure, strategy. Then either the equilibrium is super strong and all payoff pairs of  $(A, B)$  restricted to the equilibrium support lie on a line with negative slope, or one player is indifferent among all outcomes in the equilibrium support.*

### 3 Games with more than two players

In this section we consider games with more than two players. Consider an equilibrium where at least one of the players, say player 1, plays a mixed strategy that is not pure. If there is a further player who also mixes, let this be player 2; if apart from player 1 all players play a pure strategy, let any of them be player 2. Fix the strategies of the remaining players (at  $z$ , say) as in the considered equilibrium. Then this defines a two-player game between players 1 and 2. If the equilibrium is 2-strong, then the two players cannot both strictly benefit with a joint deviation. By Theorem 1 (or Lemma 2), the payoffs to the two players induced by  $z$  have to be zero-sum after suitable scaling (or all-zero for one player). The following theorem asserts that this is not a generic property in the following sense: for a sufficiently small full-dimensional neighborhood (ball) around the game, considered as a point in the payoff space, and a continuous probability distribution on that ball for choosing a nearby game, the probability that the nearby Nash equilibrium of the perturbed game is 2-strong is zero.

**Theorem 3** *In the payoff space of games with a fixed finite number of players with given finite numbers of pure strategies, the property that a Nash equilibrium where at least one player plays a mixed strategy is 2-strong is not generic.*

**Proof** Consider a 2-strong Nash equilibrium of the game where at least one player plays a mixed, non-pure strategy. If only one such player exists, then his payoffs in his support (for the pure strategies of the other players) have to be equal, which is not a generic property. (This covers also games with only one player.)

Hence, consider two players, say player 1 and 2, who use non-pure mixed strategies. Discard all pure strategies of player 1 and 2 not in the support of these mixed strategies, which we call  $x$  and  $y$ , now with full support, and let  $z$  be the profile of (in general mixed) strategies of the remaining players in this equilibrium  $(x, y, z)$ . This defines payoff matrices  $A(z)$  and  $B(z)$  to player 1 and player 2, which are multilinear expressions in the mixed strategy probabilities in  $z$  and the payoffs to player 1 and 2 for the pure strategy profiles in the support of  $(x, y, z)$ .

Consider independent small Gaussian perturbations around each payoff of the game, truncated and re-scaled to an  $\varepsilon$ -ball around the given game (as a point in the payoff space); Gaussian noise is useful because the sum of independent Gaussian distributions is again Gaussian, which we will use shortly. If the game is generic, then it has only finitely Nash equilibria, and the Nash equilibrium is an isolated point in the mixed strategy space which does not disappear when the payoffs are slightly perturbed (see Govindan and Wilson (2001) and references therein). Hence, for sufficiently small positive  $\varepsilon$  there is a unique equilibrium  $(\hat{x}, \hat{y}, \hat{z})$  nearby (with the same support as  $x, y, z$ ) for every game in the  $\varepsilon$ -ball. Moreover, the payoffs and their perturbations for pure strategy profiles outside the support of  $(x, y, z)$  do not matter.

By Theorem 1 (or Lemma 2), the payoffs in  $(A(z), B(z))$ , after suitable scaling, have to be zero-sum (or all-zero for one player). This seems clearly like a nongeneric property (and is obviously so if 1 and 2 are the only players), but the Nash

equilibrium  $(\hat{x}, \hat{y}, \hat{z})$  varies with the perturbation (via the equations (1) in the best-response condition), which may somehow preserve also the property of being 2-strong.

The following argument shows that it does not. Let the supports of  $x$  and  $y$  be  $I = \{1, \dots, m\}$  and  $J = \{1, \dots, n\}$ , respectively. Let  $K$  be the set of profiles of pure strategies of the remaining players in the support of  $z$ , with corresponding payoffs  $a_{ijk}$  and  $b_{ijk}$  to player 1 and 2 in the original game, for any  $(i, j, k) \in I \times J \times K$ . We now add *additional* small independent Gaussian noise  $\delta_j$  to  $a_{ijk}$  and  $\epsilon_i$  to  $b_{ijk}$ . The resulting noise is still Gaussian (assuming we do the truncation to the  $\epsilon$ -ball later). This changes the entries  $(a_{ij}(z), b_{ij}(z))$  of  $(A(z), B(z))$  to  $(a_{ij}(z) + \delta_j, b_{ij}(z) + \epsilon_i)$ , defining a perturbed bimatrix game  $(A', B')$ . Adding any constants  $\delta_j$  to the columns of the row player and  $\epsilon_i$  to the rows of the column player leaves the Nash equilibrium  $(\hat{x}, \hat{y})$  of  $(A', B')$  unchanged. However, it is no longer 2-strong, because the payoff pairs no longer lie on a line (with probability one). This is most easily seen because the necessary conditions (2) change to  $B'\hat{y} = (\epsilon_1, \epsilon_2, \dots, \epsilon_m)^\top$  and  $\hat{x}^\top A' = (\delta_1, \delta_2, \dots, \delta_n)$ , which (with probability one) for  $m \geq 2$  and  $n \geq 2$  are no longer constant vectors.

Hence, being a 2-strong Nash equilibrium is a not a generic property, as claimed. □

Correspondingly, the stronger conditions for a Nash equilibrium to be  $k$ -strong for  $k \geq 3$ , or  $k$ -super-strong for  $k \geq 2$ , are nongeneric as well.

We end this section with two examples of simple three-player games with two pure strategies per player. In the first game, a strong Nash equilibrium exists in fully mixed strategies. As a consequence, every outcome of the game is played, at the equilibrium, with positive probability. Among these outcomes, one is Pareto dominated (by the strong equilibrium). This shows a very different situation compared to the two-player case, where for a strong Nash equilibrium in fully mixed strategies, *all* outcomes are Pareto efficient.

**Lemma 4** *Consider the following game with three players, where players 1 and 2 choose a row and column, respectively, of  $M_1$  and  $M_2$ , and player 3 chooses between  $M_1$  and  $M_2$ :*

$$M_1 = \begin{pmatrix} (2, 0, 0) & (0, 2, 0) \\ (0, 0, 2) & (0, 0, 0) \end{pmatrix}, \quad M_2 = \begin{pmatrix} (0, 0, 0) & (0, 0, 2) \\ (0, 2, 0) & (2, 0, 0) \end{pmatrix}.$$

*Then the fully mixed strategy profile where each player plays  $(\frac{1}{2}, \frac{1}{2})$  is a super strong Nash equilibrium with payoffs  $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ . These payoffs Pareto dominate the outcome  $(0, 0, 0)$  which is in the support of the equilibrium.*

**Proof** Consider an arbitrary mixed strategy profile with probabilities  $p, q, r$  for the second strategy of players 1, 2, 3, respectively. We claim that the (fully) mixed NE where  $p = q = r = \frac{1}{2}$  is super strong. To prove this, observe that for the grand coalition, this means there is no  $p, q, r$  so that each player benefits at least weakly: player 1,

$$2(1-p)(1-q)(1-r) + 2pqr \geq \frac{1}{2}, \quad (12)$$

player 2,

$$2(1-p)q(1-r) + 2p(1-q)r \geq \frac{1}{2}, \quad (13)$$

and player 3,

$$2p(1-q)(1-r) + 2(1-p)qr \geq \frac{1}{2}, \quad (14)$$

and at least one of them benefits strictly, that is, one of these inequalities is strict. The sum of (12) and (13) states

$$2(1-p)(1-r) + 2pr \geq 1,$$

that is,  $2 - 2p - 2r + 4pr \geq 1$  or equivalently

$$1 - 2r \geq 2p(1 - 2r). \quad (15)$$

Similarly, the sum of (12) and (14) states  $2(1-q)(1-r) + 2qr \geq 1$  or equivalently

$$1 - 2r \geq 2q(1 - 2r). \quad (16)$$

The sum of (13) and (14) states  $2(1-p)q + 2p(1-q) \geq 1$ , that is,  $2q + 2p - 4pq \geq 1$  which is equivalent to both

$$2q(1 - 2p) \geq 1 - 2p \quad (17)$$

and

$$2p(1 - 2q) \geq 1 - 2q. \quad (18)$$

Consider first a possible deviation of players 1 and 2 with player 3 staying fixed at  $r = \frac{1}{2}$ . For this to be profitable for player 1 or player 2, one of (12) or (13) and therefore (15) would have to be strict, which is false for  $r = \frac{1}{2}$ . Similarly, the coalition of players 1 and 3 when  $q = \frac{1}{2}$  cannot strictly benefit player 1 or 3 because then one of (12) or (14) and therefore (16) would have to be strict, which is false for  $q = \frac{1}{2}$ . Finally, the coalition of players 2 and 3 when  $p = \frac{1}{2}$  cannot strictly benefit player 2 or 3 because then one of (13) or (14) and therefore (17) would have to be strict, which is false for  $p = \frac{1}{2}$ .

So no two-player coalition can profitably deviate, even when strictly benefiting only one player.

For the grand coalition, there is no profitable deviation that strictly benefits at least one player because, in essence, (15) and (16) imply that  $\frac{1}{2} - r$ ,  $\frac{1}{2} - p$ , and  $\frac{1}{2} - q$  have the same sign, whereas (17) and (18) imply that  $\frac{1}{2} - p$  and  $\frac{1}{2} - q$  have opposite sign. In detail, consider  $(p, q, r)$  so that the inequalities (12), (13), (14) hold, and at least one of them strictly as a benefit to the respective player. If this is player 1, then (12) and therefore (15) and (16) hold strictly, which requires  $r \neq \frac{1}{2}$ . Divide both

strict inequalities by  $2(1 - 2r)$ , which gives  $\frac{1}{2} > p$  and  $\frac{1}{2} > q$  if  $\frac{1}{2} > r$  and  $\frac{1}{2} < p$  and  $\frac{1}{2} < q$  if  $\frac{1}{2} < r$ . In the first case, (17) divided by the positive term  $2(1 - 2p)$  implies  $q \geq \frac{1}{2}$ , in the second case,  $q \leq \frac{1}{2}$ , in both cases a contradiction.

Similarly, if player 2 strictly benefits and (13) holds strictly, then so does (15) (hence,  $r \neq \frac{1}{2}$ ) and (17). Suppose  $\frac{1}{2} > r$ . As before, (15) holding strictly implies  $\frac{1}{2} > p$  and (17) holding strictly implies  $q > \frac{1}{2}$ , but (16) implies  $\frac{1}{2} \geq q$ , a contradiction. The reverse contradictory inequalities hold if  $\frac{1}{2} < r$ .

In the same manner, if player 3 strictly benefits and (14) holds strictly, then so does (16) (so  $r \neq \frac{1}{2}$ ) and (18). Suppose  $\frac{1}{2} > r$ . As before, (16) holding strictly implies  $\frac{1}{2} > q$  and (18) holding strictly implies  $p > \frac{1}{2}$ , but (15) implies  $\frac{1}{2} \geq p$ , a contradiction. The reverse contradictory inequalities hold if  $\frac{1}{2} < r$ .  $\square$

In the second example instead we show that, even in simple cases, the notion of  $k$ -strong equilibrium can be weaker than the notion of  $(k + 1)$ -strong equilibrium.

**Lemma 5** *Assume players 1 and 2 choose a row and column, respectively, of  $M_1$  and  $M_2$ , and player 3 chooses between  $M_1$  and  $M_2$ :*

$$M_1 = \begin{pmatrix} (1, -1, 0) & (-1, 1, 0) \\ (-1, 1, 0) & (1, -1, 0) \end{pmatrix}, \quad M_2 = \begin{pmatrix} (-2, -2, -2) & (-2, -2, -2) \\ (-2, -2, -2) & (1, 1, 1) \end{pmatrix}.$$

*Then the strategy profile  $\left(\left(\frac{1}{2}, \frac{1}{2}\right), \left(\frac{1}{2}, \frac{1}{2}\right), (1, 0)\right)$  is a 2-super-strong Nash equilibrium that is not a strong Nash equilibrium.*

**Proof** Suppose player 3 plays  $(1, 0)$ . Then the payoffs to player 1 and 2 in  $M_1$  define a zero-sum “matching pennies” game with unique equilibrium strategies  $\left(\frac{1}{2}, \frac{1}{2}\right)$  for players 1 and 2, and payoff 0 to all players. Hence, these three strategies define a Nash equilibrium because no player can profitably deviate. This strategy profile is also a 2-super-strong Nash equilibrium. If player 3’s strategy is fixed, then players 1 and 2 cannot jointly improve their payoffs because the game between them is zero-sum (and their payoffs pairs lie on a line with negative slope). If player 1’s strategy is fixed at  $\left(\frac{1}{2}, \frac{1}{2}\right)$ , then irrespective of player 2’s strategy player 3 would receive negative expected payoff by playing  $M_2$  with positive probability, and therefore player 3 continues to play  $(1, 0)$ , in which case both players always receive payoff zero and nobody can improve. This holds symmetrically when player 2’s strategy is fixed at  $\left(\frac{1}{2}, \frac{1}{2}\right)$ . This shows that the strategy profile is 2-super-strong.

However, the coalition of all three players can profitably change to all players playing their second strategy, with payoffs  $(1, 1, 1)$ .  $\square$

For games with three or more players, a remaining open question is a suitable characterization of games that admit a strong Nash equilibrium.

## References

- Adler I, Daskalakis C, Papadimitriou CH (2009) A note on strictly competitive games. In: Leonardi S (ed) International workshop on internet and network economics (WINE), Lecture Notes in Computer Science, vol 5929. Springer, Berlin, pp 471–474
- Andelman N, Feldman M, Mansour Y (2009) Strong price of anarchy. *Games Econ Behav* 65(2):289–317
- Aumann RJ (1959) Acceptable points in general cooperative  $n$ -person games. In: Luce RD, Tucker AW (eds) Contribution to the theory of games IV, *Annals of Mathematics Studies*, vol 40. Princeton University Press, Princeton, pp 287–324
- Bernheim B, Peleg B, Whinston MD (1987) Coalition-proof Nash equilibria I. Concepts. *J Econ Theory* 42(1):1–12
- Conitzer V, Sandholm T (2008) New complexity results about Nash equilibria. *Games Econ Behav* 63(2):621–641
- Dubey P (1986) Inefficiency of Nash equilibria. *Math Oper Res* 11(1):1–8
- Einy E, Peleg B (1995) Coalition-proof communication equilibria. In: Burnett WA, Moulin H, Sales M, Schofield NJ (eds) Social choice, welfare, and ethics. Cambridge University Press, Cambridge, pp 266–290
- Gourvès L, Monnot J (2009) On strong equilibria in the max cut game. In: Leonardi S (ed) International workshop on internet and network economics (WINE), Lecture Notes in Computer Science, vol 5929. Springer, Berlin, pp 608–615
- Govindan S, Wilson R (2001) Direct proofs of generic finiteness of Nash equilibrium outcomes. *Econometrica* 69(3):765–769
- Luce RD, Raiffa H (1957) Games and decisions. Wiley, New York
- Miettinen K (1999) Nonlinear multiobjective optimization. Kluwer Academic Publishers, Boston
- Milgrom P, Roberts J (1996) Coalition-proofness and correlation with arbitrary communication possibilities. *Games Econ Behav* 17(1):113–128
- Nessah R, Tian G (2014) On the existence of strong Nash equilibria. *J Math Anal Appl* 414(2):871–885
- Rozenfeld O (2007) Strong equilibrium in congestion games. Master's thesis, Computer Science Department, Technion-Israel Institute of Technology, Haifa

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.