

# New algorithms for approximate Nash equilibria in bimatrix games

**Citation for published version (APA):**

Bosse, H., Byrka, J., & Markakis, E. (2007). New algorithms for approximate Nash equilibria in bimatrix games. In X. Deng, & F. C. Graham (Eds.), *Proceedings of the 3rd International Workshop on Internet and Network Economics (WINE 2007) 12-14 December 2007, San Diego, California, USA* (pp. 17-29). (Lecture Notes in Computer Science; Vol. 4858). Springer. [https://doi.org/10.1007/978-3-540-77105-0\\_6](https://doi.org/10.1007/978-3-540-77105-0_6)

**DOI:**

[10.1007/978-3-540-77105-0\\_6](https://doi.org/10.1007/978-3-540-77105-0_6)

**Document status and date:**

Published: 01/01/2007

**Document Version:**

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

# New Algorithms for Approximate Nash Equilibria in Bimatrix Games\*

Hartwig Bosse, Jaroslaw Byrka, and Evangelos Markakis

CWI (Center for Math and Computer Science)

Amsterdam, The Netherlands

{bosse, j.byrka, vangelis}@cwi.nl

**Abstract.** We consider the problem of computing additively approximate Nash equilibria in non-cooperative two-player games. We provide a new polynomial time algorithm that achieves an approximation guarantee of 0.36392. Our work improves the previously best known  $(0.38197 + \epsilon)$ -approximation algorithm of Daskalakis, Mehta and Papadimitriou [6].

First, we provide a simpler algorithm, which also achieves 0.38197. This algorithm is then tuned, improving the approximation error to 0.36392. Our method is relatively fast, as it requires solving only one linear program and it is based on using the solution of an auxiliary zero-sum game as a starting point.

## 1 Introduction

A Nash equilibrium of a bimatrix game is a pair of strategies, such that no player has an incentive to deviate (unilaterally). In a series of works [8,4,2], it was established that computing a Nash equilibrium is PPAD-complete even for two-player games. The focus has since then been on algorithms for approximate equilibria.

In this work we use the notion of *additive approximation* and consider the problem of computing approximate Nash equilibria in bimatrix games. Under the usual assumption that the payoff matrices are normalized to be in  $[0, 1]^{n \times n}$ , we say that a pair of strategies is an  $\epsilon$ -Nash equilibrium if no player can gain more than  $\epsilon$  by unilaterally deviating to another strategy. In [3] it was proved that it is PPAD-complete to find an  $\epsilon$ -Nash equilibrium when  $\epsilon$  is of the order  $\frac{1}{\text{poly}(n)}$ . For constant  $\epsilon$  however, the problem is still open. In [11], it was shown that for any constant  $\epsilon > 0$ , an  $\epsilon$ -Nash equilibrium can be computed in subexponential time ( $n^{O(\log n / \epsilon^2)}$ ). As for polynomial time algorithms, it is fairly simple to obtain a 3/4-approximation (see [9] for a slightly better result) and even better a 1/2-approximation [5]. Recently, an improved approximation for

---

\* The first author was supported by NWO. The second and third author were supported by the EU Marie Curie Research Training Network, contract numbers MRTN-CT-2003-504438-ADONET and MRTN-CT-2004-504438-ADONET respectively.

$\epsilon = \frac{3-\sqrt{5}}{2} + \zeta \approx 0.38197 + \zeta$  for any  $\zeta > 0$  was obtained by Daskalakis, Mehta and Papadimitriou [6].

We provide two new algorithms for approximate Nash equilibria. The first one achieves exactly the same factor as [6] but with a simpler and faster technique. The second one, which is an extension of the first and has a more involved analysis, achieves an improved approximation of 0.36392. Both algorithms are based on solving a single linear program in contrast to [6], which may require to solve up to  $n^{O(\frac{1}{\zeta^2})}$  linear programs for a  $(0.38197 + \zeta)$ -approximation.

The main idea of our algorithms is as follows: we first find an equilibrium (say  $x^*, y^*$ ) in the zero-sum game  $R - C$ , where  $R$  and  $C$  are the payoff matrices of the two players. If  $x^*, y^*$  is not a good solution for the original game, then the players take turns and switch to appropriate strategies. Roughly speaking, the probabilities of switching are chosen such that the incentives to deviate become the same for both players. As a result, these probabilities are particular functions in the parameters of the underlying problem. The final part of the analysis then is to choose among these functions so as to minimize the approximation error. The intuition behind using the auxiliary zero-sum game  $R - C$  is that a unilateral switch from  $x^*, y^*$  that improves the payoff of one player cannot hurt the other, as explained in the proof of Theorem 1. We should note that the use of certain zero-sum games has also been considered in [10] for obtaining *well-supported* approximate equilibria, which is a stronger notion of approximation.

In an independent work, Spirakis and Tsaknakis [13] have obtained another algorithm achieving an improved approximation of 0.3393. Their technique is also based on linear programming but seems unrelated to ours, and requires solving a polynomial number of linear programs.

## 2 Notation and Definitions

Consider a two person game  $G$ , where for simplicity the number of available (pure) strategies for each player is  $n$ . Our results still hold when the players do not have the same number of available strategies. We will refer to the two players as the row and the column player and we will denote their  $n \times n$  payoff matrices by  $R, C$  respectively. Hence, if the row player chooses strategy  $i$  and the column player chooses strategy  $j$ , the payoffs are  $R_{ij}$  and  $C_{ij}$  respectively.

A *mixed strategy* for a player is a probability distribution over the set of his pure strategies and will be represented by a vector  $x = (x_1, x_2, \dots, x_n)^T$ , where  $x_i \geq 0$  and  $\sum x_i = 1$ . Here  $x_i$  is the probability that the player will choose his  $i$ th pure strategy. The  $i$ th pure strategy will be represented by the unit vector  $e_i$ , that has 1 in the  $i$ th coordinate and 0 elsewhere. For a mixed strategy pair  $x, y$ , the payoff to the row player is the expected value of a random variable which is equal to  $R_{ij}$  with probability  $x_i y_j$ . Therefore the payoff to the row player is  $x^T R y$ . Similarly the payoff to the column player is  $x^T C y$ .

A Nash equilibrium [12] is a pair of strategies  $x^*, y^*$  such that no player has an incentive to deviate unilaterally. Since mixed strategies are convex combinations of pure strategies, it suffices to consider only deviations to pure strategies:

**Definition 1.** A pair of strategies  $x^*, y^*$  is a Nash equilibrium if:

- (i) For every pure strategy  $e_i$  of the row player,  $e_i^T R y^* \leq (x^*)^T R y^*$ , and
- (ii) For every pure strategy  $e_i$  of the column player,  $(x^*)^T C e_i \leq (x^*)^T C y^*$ .

Assuming that we normalize the entries of the payoff matrices so that they all lie in  $[0, 1]$ , we can define the notion of an additive  $\epsilon$ -approximate Nash equilibrium (or simply  $\epsilon$ -Nash equilibrium) as follows:

**Definition 2.** For any  $\epsilon > 0$ , a pair of strategies  $x^*, y^*$  is an  $\epsilon$ -Nash equilibrium iff:

- (i) For every pure strategy  $e_i$  of the row player,  $e_i^T R y^* \leq (x^*)^T R y^* + \epsilon$ , and
- (ii) For every pure strategy  $e_i$  of the column player,  $(x^*)^T C e_i \leq (x^*)^T C y^* + \epsilon$ .

In other words, no player will gain more than  $\epsilon$  by unilaterally deviating to another strategy. A stronger notion of approximation was introduced in [4], namely  $\epsilon$ -well-supported equilibria. We do not consider this approximation concept here. See [10] for new results on well-supported equilibria.

### 3 A $(\frac{3-\sqrt{5}}{2})$ -Approximation

In this section, we provide an algorithm that achieves exactly the same factor as in [6], which is  $(3 - \sqrt{5})/2$ , but by using a different and simpler method. In the next section we show how to modify our algorithm in order to improve the approximation.

Given a game  $G = (R, C)$ , where the entries of  $R$  and  $C$  are in  $[0, 1]$ , let  $A = R - C$ . Our algorithm is based on solving the zero-sum game  $(A, -A)$  and then modifying appropriately the solution, if it does not provide a good approximation. It is well known that zero-sum games can be solved efficiently using linear programming. The decision on when to modify the zero-sum solution depends on a parameter of the algorithm  $\alpha \in [0, 1]$ . We first describe the algorithm parametrically and then show how to obtain the desired approximation.

#### Algorithm 1

Let  $\alpha \in [0, 1]$  be a parameter of the algorithm.

1. Compute an equilibrium  $(x^*, y^*)$  for the zero-sum game defined by the matrix  $A = R - C$ .
2. Let  $g_1, g_2$  be the incentives to deviate for the row and column player respectively if they play  $(x^*, y^*)$  in the game  $(R, C)$ , i.e.,  $g_1 = \max_{i=1, \dots, n} e_i^T R y^* - (x^*)^T R y^*$  and  $g_2 = \max_{i=1, \dots, n} (x^*)^T C e_i - (x^*)^T C y^*$ . WLOG, assume, that  $g_1 \geq g_2$  (the statement of the algorithm would be completely symmetrical if  $g_1 < g_2$ ).
3. Let  $r_1 \in \operatorname{argmax}_{e_i} e_i^T R y^*$  be an optimal response of the row player to the strategy  $y^*$ . Let  $b_2 \in \operatorname{argmax}_{e_i} r_1^T C e_i$  be an optimal response of the column player to the strategy  $r_1$ .

4. Output the following pair of strategies,  $(\hat{x}, \hat{y})$ , depending on the value of  $g_1$  with respect to the value of  $\alpha$ :

$$(\hat{x}, \hat{y}) = \begin{cases} (x^*, y^*), & \text{if } g_1 \leq \alpha \\ (r_1, (1 - \delta_2) \cdot y^* + \delta_2 \cdot b_2), & \text{otherwise} \end{cases}$$

where  $\delta_2 = \frac{1-g_1}{2-g_1}$ .

**Theorem 1.** *Algorithm 1 outputs a  $\max\{\alpha, \frac{1-\alpha}{2-\alpha}\}$ -approximate Nash equilibrium.*

*Proof.* If  $g_1 \leq \alpha$  (recall that we assumed  $g_1 \geq g_2$ ), then clearly  $(x^*, y^*)$  is an  $\alpha$ -approximate Nash equilibrium.

Suppose  $g_1 > \alpha$ . We will estimate the satisfaction of each player separately. Suppose  $b_1$  is an optimal response for the row player to  $\hat{y}$ , i.e.,  $b_1 \in \operatorname{argmax}_{e_i} e_i^T R \hat{y}$ . The row player plays  $r_1$ , which is a best response to  $y^*$ . Hence  $b_1$  can be better than  $r_1$  only when the column player plays  $b_2$ , which happens with probability  $\delta_2$ . Formally, the amount that the row player can earn by switching is at most:

$$\begin{aligned} b_1^T R \hat{y} - r_1^T R \hat{y} &= (1 - \delta_2)(b_1^T R y^* - r_1^T R y^*) + \delta_2(b_1^T R b_2 - r_1^T R b_2) \\ &\leq \delta_2 \cdot b_1^T R b_2 \leq \delta_2 = \frac{1-g_1}{2-g_1} \end{aligned}$$

The first inequality above comes from the fact that  $r_1$  is a best response to  $y^*$  and the second comes from our assumption that the entries of  $R$  and  $C$  are in  $[0, 1]$ .

Consider the column player. The critical observation is that the column player also benefits (when he plays  $y^*$ ) from the switch of the row player from  $x^*$  to  $r_1$ . In particular, since  $(x^*, y^*)$  is an equilibrium for the zero-sum game  $(R - C, C - R)$ , the following inequalities hold:

$$(x^*)^T R e_j - (x^*)^T C e_j \geq (x^*)^T R y^* - (x^*)^T C y^* \geq e_i^T R y^* - e_i^T C y^*, \quad \forall i, j = 1, \dots, n \tag{1}$$

If  $e_i = r_1$ , we get from (1) that  $r_1^T C y^* \geq r_1^T R y^* - (x^*)^T R y^* + (x^*)^T C y^*$ . But we know that  $r_1^T R y^* - (x^*)^T R y^* = g_1$ , which implies:

$$r_1^T C y^* \geq g_1 + (x^*)^T C y^* \geq g_1 \tag{2}$$

Inequality (2) shows that any deviation of the row player from  $x^*, y^*$ , that improves his payoff, guarantees at least the same gain to the column player as well. Now we can estimate the incentive of the column player to change his strategy. He plays  $\hat{y}$  while he would prefer to play an optimal response to  $\hat{x}$  which is  $b_2$ . Since  $b_2$  is played with probability  $\delta_2$ , by switching he could earn:

$$\begin{aligned} \hat{x}^T C b_2 - \hat{x}^T C \hat{y} &= r_1^T C b_2 - r_1^T C \hat{y} \\ &= r_1^T C b_2 - ((1 - \delta_2)r_1^T C y^* - \delta_2 \cdot r_1^T C b_2) \\ &= (1 - \delta_2)(r_1^T C b_2 - r_1^T C y^*) \\ &\leq (1 - \delta_2)(1 - g_1) = \delta_2 = \frac{1-g_1}{2-g_1} \end{aligned}$$

The last inequality above follows from (2). The probability  $\delta_2$  was chosen so as to equalize the incentives of the two players to deviate in the case that  $g_1 > \alpha$ .

It is now easy to check that the function  $(1 - g_1)/(2 - g_1)$  is decreasing, hence the incentive for both players to deviate is at most  $(1 - \alpha)/(2 - \alpha)$ . Combined with the case when  $g_1 \leq \alpha$ , we get a  $\max\{\alpha, \frac{1-\alpha}{2-\alpha}\}$ -approximate equilibrium.

In order to optimize the approximation factor of Algorithm 1, we only need to equate the two terms,  $\alpha$  and  $\frac{1-\alpha}{2-\alpha}$ , which then gives:

$$\alpha^2 - 3\alpha + 1 = 0 \tag{3}$$

The solution to (3) in the interval  $[0, 1]$  is  $\alpha = \frac{3-\sqrt{5}}{2} \approx 0.38197$ . Note that  $\alpha = 1 - 1/\phi$ , where  $\phi$  is the golden ratio. Since  $\alpha$  is an irrational number, we need to ensure that we can still do the comparison  $g_1 \leq \alpha$  to be able to run Algorithm 1 (note that this is the only point where the algorithm uses the value of  $\alpha$ ). But to test  $g_1 \leq 3 - \sqrt{5}/2$ , it suffices to test if  $(3 - 2g_1)^2 \geq 5$  and clearly  $g_1$  is a polynomially sized rational number. Concerning complexity, zero-sum games can be solved in polynomial time by linear programming. All the other steps of the algorithm require only polynomial time. Therefore, Theorem 1 implies:

**Corollary 1.** *We can compute in polynomial time a  $\frac{3-\sqrt{5}}{2}$ -approximate Nash equilibrium for bimatrix games.*

## 4 An Improved Approximation

In this section we obtain a better approximation of  $1/2 - 1/(3\sqrt{6}) \approx 0.36392$  by essentially proposing a different solution in the cases where Algorithm 1 approaches its worst case guarantee. We first give some motivation for the new algorithm. From the analysis of Algorithm 1, one can easily check that as long as  $g_1$  belongs to  $[0, 1/3] \cup [1/2, 1]$ , we can have a  $1/3$ -approximation if we run the algorithm with any  $\alpha \in [1/3, 1/2)$ . Therefore, the bottleneck for getting a better guarantee is when the maximum incentive to deviate is in  $[1/3, 1/2]$ . In this case, we will change the algorithm so that the row player will play a mix of  $r_1$  and  $x^*$ . Note that in Algorithm 1, the probability of playing  $r_1$  is either 0 or 1 depending on the value of  $g_1$ . This probability will now be a more complicated function of  $g_1$ , derived from a certain optimization problem. As for the column player, we again compute  $b_2$  which is now the best response to the mixture of  $r_1$  and  $x^*$  - not only to  $r_1$ . Then we compute an appropriate mixture of  $b_2$  and  $y^*$ . Again, the probability of playing  $b_2$  is chosen so as to equate the incentives of the two players to defect. Finally we should note that our modification will be not on  $[1/3, 1/2]$  but instead on a subinterval of the form  $[1/3, \beta]$ , where  $\beta$  is derived from the optimization that we perform in our analysis.

### Algorithm 2

1. Compute an equilibrium  $(x^*, y^*)$  for the zero-sum game defined by the matrix  $A = R - C$ .

2. As in Algorithm 1, let  $g_1, g_2$  be the incentives to deviate for the row and column player respectively if they play  $(x^*, y^*)$  in the original game, i.e.,  $g_1 = \max_{i=1, \dots, n} e_i^T R y^* - (x^*)^T R y^*$  and  $g_2 = \max_{i=1, \dots, n} (x^*)^T C e_i - (x^*)^T C y^*$ . WLOG, assume, that  $g_1 \geq g_2$ .
3. Let  $r_1 \in \operatorname{argmax}_{e_i} e_i^T R y^*$  be an optimal response of the row player to the strategy  $y^*$ .
4. The row player will play a mixture of  $r_1$  and  $x^*$ , where the probability of playing  $r_1$  is given by:

$$\delta_1 = \delta_1(g_1) = \begin{cases} 0, & \text{if } g_1 \in [0, 1/3] \\ \Delta_1(g_1), & \text{if } g_1 \in (1/3, \beta] \\ 1, & \text{otherwise} \end{cases}$$

where  $\Delta_1(g_1) = (1 - g_1) \left( -1 + \sqrt{1 + \frac{1}{1-2g_1} - \frac{1}{g_1}} \right)$ .

5. Let  $b_2$  be an optimal response of the column player to  $((1 - \delta_1)x^* + \delta_1 r_1)$ , i.e.,  $b_2 \in \operatorname{argmax}_{e_i} ((1 - \delta_1)x^* + \delta_1 r_1)^T C e_i$ . Let also  $h_2 = (x^*)^T C b_2 - (x^*)^T C y^*$ , i.e., the gain from switching to  $b_2$  if the row player plays  $x^*$ .
6. The column player will play a mixture of  $b_2$  and  $y^*$ , where the probability of playing  $b_2$  is given by:

$$\delta_2 = \delta_2(\delta_1, g_1, h_2) = \begin{cases} 0, & \text{if } g_1 \in [0, 1/3] \\ \max\{0, \Delta_2(\delta_1, g_1, h_2)\}, & \text{if } g_1 \in (1/3, \beta] \\ \frac{1-g_1}{2-g_1}, & \text{otherwise} \end{cases}$$

where  $\Delta_2(\delta_1, g_1, h_2) = \frac{\delta_1 - g_1 + (1 - \delta_1)h_2}{1 + \delta_1 - g_1}$ .

7. Output  $(\hat{x}, \hat{y}) = ((1 - \delta_1)x^* + \delta_1 r_1, (1 - \delta_2)y^* + \delta_2 b_2)$ .

In our analysis, we will take  $\beta$  to be the solution to  $\Delta_1(g_1) = 1$  in  $[1/3, 1/2]$ , which coincides with the root of the polynomial  $x^3 - x^2 - 2x + 1$  in that interval and it is:

$$\beta = \frac{1}{3} + \frac{\sqrt{7}}{3} \cos\left(\frac{1}{3} \tan^{-1}\left(3\sqrt{3}\right)\right) - \frac{\sqrt{21}}{3} \sin\left(\frac{1}{3} \tan^{-1}\left(3\sqrt{3}\right)\right) \quad (4)$$

Calculations show  $0.445041 \leq \beta \leq 0.445042$ . The emergence of  $\beta$  in our analysis arises in the proof of Lemma 1.

*Remark 1.* The actual probabilities  $\delta_1$  and  $\delta_2$  as well as the number  $\beta$  can be irrational numbers. However, for any constant  $\epsilon > 0$ , we can take approximations of high enough accuracy of all the square roots that are involved in the calculations so that the final loss in the approximation ratio will be at most  $\epsilon$ . From now on, for ease of exposition, we will carry out the analysis of Algorithm 2, as if we can compute exactly all the expressions involved.

Note that for  $g_1 \in [\frac{1}{3}, \frac{1}{2}]$  and  $\delta_1 \in [0, 1]$  the denominators that appear in the functions  $\Delta_1, \Delta_2$  do not vanish. The following lemma ensures that  $\hat{x}$  is a valid strategy.

**Lemma 1.** For  $g_1 \in (1/3, \beta]$  we have  $\Delta_1(g_1) \in [0, 1]$ .

The proof of Lemma 1 is based on showing that the function  $\Delta_1$  is increasing in  $[1/3, \beta]$  and that it maps  $[1/3, \beta]$  to  $[0, 1]$ . Due to lack of space we omit it in this version.

Now we bound the incentives of players to deviate. Let  $F$  be the following function:

$$F(\delta_1, g_1, h_2) := \frac{(\delta_1(1 - g_1 - h_2) + h_2)(1 - (1 - \delta_1)h_2)}{1 + \delta_1 - g_1} \quad (5)$$

**Lemma 2.** The pair of strategies  $(\hat{x}, \hat{y})$  is a  $\lambda$ -Nash equilibrium for game  $(R, C)$  with

$$\lambda \leq \begin{cases} g_1 & \text{if } g_1 \leq 1/3 \\ \max_{h_2 \in [0, g_1]} \begin{cases} F(\delta_1, g_1, h_2) & \text{if } \Delta_2(\delta_1, g_1, h_2) \geq 0 \\ (1 - \delta_1)g_1 & \text{if } \Delta_2(\delta_1, g_1, h_2) < 0 \end{cases} & \text{if } g_1 \in (1/3, \beta] \\ \frac{1 - g_1}{2 - g_1} & \text{if } g_1 > \beta \end{cases} \quad (6)$$

*Proof.* In the case that  $g_1 \in [0, 1/3] \cup [\beta, 1]$ , the answer follows from the proof of Theorem 1. The interesting case is when  $g_1 \in [1/3, \beta]$ .

*Case 1:*  $g_1 \leq 1/3$

$(\hat{x}, \hat{y}) = (x^*, y^*)$  which is by definition a  $g_1$ -approximate Nash equilibrium.

*Case 2a:*  $g_1 \in (1/3, \beta]$  and  $\Delta_2(\delta_1, g_1, h_2) \geq 0$

Recall that Lemma 1 implies  $\hat{x}$  is a valid strategy in Case 2. Observe, that  $\delta_2(g_1, \delta_1, h_2) = \Delta_2(g_1, \delta_1, h_2) = \frac{\delta_1 - g_1 + (1 - \delta_1)h_2}{1 + \delta_1 - g_1} \leq 1$  is a valid probability, and therefore  $\hat{y}$  is a valid mixed strategy too.

We estimate the incentive for the row player to deviate from  $\hat{x}$ . If  $b_1$  is an optimal response to  $\hat{y}$ , then the gain from switching is at most:

$$\begin{aligned} b_1^T R \hat{y} - \hat{x}^T R \hat{y} &= (b_1 - \hat{x})^T R \hat{y} = \\ &= \delta_2(b_1 - \hat{x})^T R b_2 + (1 - \delta_2)(b_1 - \hat{x})^T R y^* \\ &\leq \delta_2(1 - \hat{x}^T R b_2) + (1 - \delta_2)(b_1 - \hat{x})^T R y^* \\ &= \delta_2(1 - \delta_1 r_1^T R b_2 - (1 - \delta_1)(x^*)^T R b_2) + (1 - \delta_2)(b_1 - \delta_1 r_1 - (1 - \delta_1)x^*)^T R y^* \end{aligned}$$

By (1) we have  $(x^*)^T R b_2 \geq (x^*)^T C b_2 - (x^*)^T C y^* + (x^*)^T R y^* \geq h_2$ . Also  $r_1$  is a best response to  $y^*$ , hence  $(b_1 - r_1)^T R y^* \leq 0$  and  $(b_1 - x^*)^T R y^* \leq g_1$ . Therefore, the gain from deviating is at most:

$$b_1^T R \hat{y} - \hat{x}^T R \hat{y} \leq \delta_2(1 - (1 - \delta_1)h_2) + (1 - \delta_2)(1 - \delta_1)g_1 = \text{EST}_1.$$

We now estimate the incentive of the column player to switch. The best response to  $\hat{x}$  for the column player is  $b_2$ , which is played with probability  $\delta_2$ . Thus the incentive to deviate from  $\hat{y}$  is:

$$\begin{aligned} \hat{x}^T C b_2 - \hat{x}^T C \hat{y} &= (1 - \delta_2)(\hat{x}^T C b_2 - \hat{x}^T C y^*) \\ &= (1 - \delta_2)((1 - \delta_1)(x^*)^T C b_2 - x^*)^T C y^* + \delta_1(r_1^T C b_2 - r_1^T C y^*) \\ &\leq (1 - \delta_2)((1 - \delta_1)h_2 + \delta_1(1 - g_1)) = \text{EST}_2 \end{aligned}$$



The last inequality follows from the definitions of  $g_1$  and  $h_2$ . It remains to observe that our choice of  $\delta_2(\delta_1, g_1, h_2) = \frac{\delta_1 - g_1 + (1 - \delta_1)h_2}{1 + \delta_1 - g_1}$  makes these estimates both equal to  $F(\delta_1, g_1, h_2)$ :

$$\text{EST}_1 = \text{EST}_2 = \frac{(\delta_1(1 - g_1 - h_2) + h_2)(1 - (1 - \delta_1)h_2)}{\delta_1 + 1 - g_1} = F(\delta_1, g_1, h_2).$$

*Case 2b:*  $g_1 \in (1/3, \beta]$  and  $\Delta_2(\delta_1, g_1, h_2) < 0$

Then  $\hat{y} = y^*$  and the best response of the row player is  $r_1$ . Hence he can improve his payoff by at most

$$r_1^T R y^* - \hat{x}^T R y^* = r_1^T R y^* - (\delta_1 \cdot r_1^T R y^* + (1 - \delta_1)((x^*)^T R y^*)) = (1 - \delta_1)g_1$$

while the column player can improve by at most

$$\hat{x}^T C b_2 - \hat{x}^T C y^* = \delta_1(r_1^T C b_2 - r_1^T C y^*) + (1 - \delta_1)((x^*)^T C b_2 - (x^*)^T C y^*)$$

By (1) we can see that  $r_1^T C y^* \geq g_1$ . Hence

$$\hat{x}^T C b_2 - \hat{x}^T C y^* \leq \delta_1(1 - g_1) + (1 - \delta_1)h_2$$

It is easy to check that  $\Delta_2(g_1, \delta_1, h_2) < 0$  implies  $\delta_1(1 - g_1) + (1 - \delta_1)h_2 < (1 - \delta_1)g_1$ . Therefore the maximum incentive to deviate in this case is at most  $(1 - \delta_1)g_1$ . Combining Case 2a and Case 2b, and taking the worst possible case over the range of  $h_2$  (recall that  $h_2 \leq g_2 \leq g_1$ ), we get precisely the expression in the statement of Lemma 2.

*Case 3:*  $g_1 > \beta$

Notice that in this case, the players are playing the same strategies as in Algorithm 1, when  $g_1 \geq \alpha$ . By the analysis in the proof of Theorem 1, we see that the maximum incentive is  $(1 - g_1)/(2 - g_1)$ . This completes the proof.

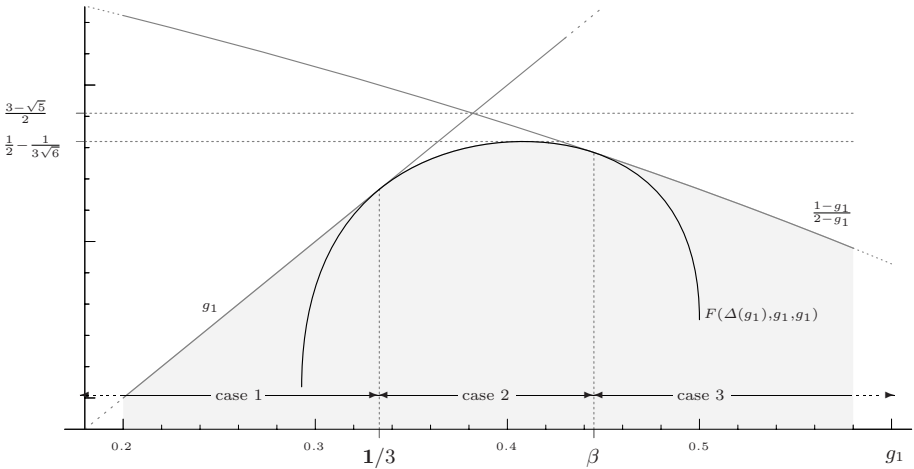
We will now argue that our choice of  $\Delta_1(g_1)$  is optimal for any  $g_1 \in (\frac{1}{3}, \beta]$  and that the expression (6) from Lemma 2 achieves an improvement over Algorithm 1. For this, we need to find the worst possible approximation in Case 2 of Lemma 2. In particular, we need to look at the maxima of the following function:

$$P(g_1) := \min_{\delta_1 \in [0,1]} \max_{h_2 \in [0, g_1]} \begin{cases} F(\delta_1, g_1, h_2) & \text{if } \Delta_2(\delta_1, g_1, h_2) \geq 0 \\ (1 - \delta_1)g_1 & \text{if } \Delta_2(\delta_1, g_1, h_2) < 0 \end{cases} \quad (7)$$

**Lemma 3.** *The tuple  $(\delta_1, h_2) = (\Delta_1(g_1), g_1)$  is an optimal solution for the expression  $P(g_1)$ . Furthermore, the maximum of  $P(g_1)$  over  $g_1$  is  $\frac{1}{2} - \frac{1}{3\sqrt{6}}$ , i.e., the following holds*

$$P(g_1) = F(\Delta_1(g_1), g_1, g_1) \quad \forall g_1 \in [\frac{1}{3}, \frac{1}{2}] \quad (8)$$

$$\max_{g_1 \in [\frac{1}{3}, \beta]} P(g_1) = \frac{1}{2} - \frac{1}{3\sqrt{6}} \leq 0.36392. \quad (9)$$



**Fig. 1.** How the approximation factor depends on  $g_1$

The lemma will be proved in Section 5. Given Remark 1, we are now ready to conclude with the following:

**Theorem 2.** *For any  $\epsilon > 0$ , Algorithm 2 computes a  $(0.36392 + \epsilon)$ -approximate Nash equilibrium.*

*Proof.* By Lemma 2 the output of Algorithm 2,  $(\hat{x}, \hat{y})$  is a pair of mixed strategies for players, such that the incentive of players to deviate is bounded by (6). By Lemma 3 we have that for  $g_1 \in (1/3, \beta]$  the expression (6) is bounded by  $\frac{1}{2} - \frac{1}{3\sqrt{6}} \leq 0.36392$ . It is easy to observe, that for other values of  $g_1$  the expression (6) takes only smaller values. In particular, it is at most  $1/3$  when  $g_1 \in [0, 1/3]$  and at most  $\frac{1-\beta}{2-\beta} \approx 0.3569$  when  $g_1 > \beta$ . The dependence of the approximation on the variable  $g_1$  is presented in Figure 1.

### 5 Proof of Lemma 3

**Fact 3.** *The square function is monotone increasing on the positive domain, i.e.,*

$$a - b \geq 0 \Leftrightarrow a^2 - b^2 \geq 0 \quad \text{holds for all } a, b \in \mathbb{R}, \quad a, b \geq 0 \quad (10)$$

We solved the optimization problem of Lemma 3 in the classic manner, eventually leading to the minimizer  $\Delta_1(g)$ . This procedure is lengthy, so here we give an uninspiring but short proof.

**Proof of Lemma 3 :** Combining (11) and (12) from Lemma 4 below, we obtain:

$$F(\Delta_1(g_1), g_1, g_1) = \min_{\delta_1 \in [0,1]} \max_{h_2 \in [0, g_1]} \begin{cases} F(\delta_1, g_1, h_2) & \text{if } \Delta_2(\delta_1, g_1, h_2) \geq 0 \\ (1 - \delta_1)g_1 & \text{if } \Delta_2(\delta_1, g_1, h_2) < 0 \end{cases}$$

For ease of exposition, we drop the subscripts of the variables from now on. Hence we need to prove  $\max_{g \in [\frac{1}{3}, \beta]} F(\Delta_1(g), g, g) = \frac{1}{2} - \frac{1}{3\sqrt{6}} \leq 0.36392$  where

$$F(\Delta_1(g), g, g) = \frac{1}{4} - \frac{1}{4}(1-2g)(3-2g)(4g-1) + 2(1-g)\sqrt{g(1-2g)(-1+4g-2g^2)}$$

The fact that the radicand of the function  $\Delta_1$  is nonnegative implies that the radicand  $g(1-2g)(-1+4g-2g^2)$  is nonnegative for all  $g \in [1/3, \beta]$ . We now prove that the maximum of  $F(\Delta(g), g, g)$  on  $[\frac{1}{3}, \beta]$  is assumed in  $1/\sqrt{6}$ . Straightforward calculation leads to

$$\mathcal{F}^* := F\left(\Delta(1/\sqrt{6}), 1/\sqrt{6}, 1/\sqrt{6}\right) = \frac{1}{2} - \frac{1}{3\sqrt{6}}$$

Fixing  $g \in [1/3, \beta]$  (arbitrarily), one finds:

$$\begin{aligned} \mathcal{F}^* - F(\Delta_1(g), g, g) &= \\ &= \underbrace{\frac{1}{4} - \frac{1}{3\sqrt{6}} + \frac{1}{4}(1-2g)(3-2g)(4g-1)}_{\geq 0 \text{ (*)}} - \underbrace{2(1-g)\sqrt{g(1-2g)(-1+4g-2g^2)}}_{\geq 0 \text{ (**)}} \end{aligned}$$

Here (\*) and (\*\*) are implied by the choice of  $g$ , i.e.,  $(3-2g) \geq 2(1-g) \geq (1-2g) \geq 0$ , and  $4g-1 \geq 1/3 > 0$  hold. Finally since  $\sqrt{6} > 2$  we have  $\frac{1}{4} - \frac{1}{3\sqrt{6}} > \frac{1}{12} > 0$ .

The inequalities in (\*) and (\*\*) together with (10) lead (after calculations which we omit due to lack of space) to the equivalence:

$$\mathcal{F}^* - F(\Delta_1(g), g, g) \geq 0 \Leftrightarrow \left(\frac{11}{18} + \frac{2}{3\sqrt{6}}(3-g) + (1-g)^2\right) \left(g - \frac{1}{\sqrt{6}}\right)^2 \geq 0$$

Here the second inequality trivially holds since  $(3-g) > 0$  for  $g \in [1/3, \beta]$ .

Thus we showed  $\mathcal{F}^* = F(\Delta_1(1/\sqrt{6}), 1/\sqrt{6}, 1/\sqrt{6}) \geq F(\Delta_1(g), g, g)$ , proving the lemma, since  $g \in [1/3, \beta]$  was chosen arbitrarily and  $1/\sqrt{6} \in [1/3, \beta]$  is implied by  $0.40 \leq 1/\sqrt{6} \leq 0.41 < \beta$ .  $\square$

It now remains to prove the following Lemma:

**Lemma 4.** *For every pair  $(g, \delta) \in [1/3, \beta] \times [0, 1]$  we find*

$$F(\delta, g, g) = \max_{h \in [0, g]} \begin{cases} F(\delta, g, h) & \text{if } \Delta_2(\delta, g, h) \geq 0 \\ (1-\delta)g & \text{if } \Delta_2(\delta, g, h) < 0 \end{cases} \quad (11)$$

$$F(\Delta_1(g), g, g) = \min_{d \in [0, 1]} F(d, g, g) \quad (12)$$

*Proof.* Fix some pair  $(g, \delta) \in [1/3, \beta] \times [0, 1]$ . We rewrite (11) as

$$F(\delta, g, g) \leq \left( \max_{h \in [0, g]} \begin{cases} F(\delta, g, h) & \text{if } \Delta_2(\delta, g, h) \geq 0 \\ (1-\delta)g & \text{if } \Delta_2(\delta, g, h) < 0 \end{cases} \right) \leq \max_{h \in [0, g]} F(\delta, g, g) \quad (13)$$

and prove it as follows: Brief calculation together with  $(1 - g) > 0$  lead to  $\Delta_2(\delta, g, g) = (1 - g)\delta/(1 - g + \delta) \geq 0$ . So there is a  $h^* \in [0, g]$ , namely  $h^* := g$ , such that  $\Delta_2(\delta, g, h^*) \geq 0$ . This implies the first inequality in (13).

Observe that to prove the second inequality in (13), it suffices to show that

$$F(\delta, g, g) \geq (1 - \delta)g \quad \text{and} \quad F(\delta, g, g) \geq F(\delta, g, h) \quad \text{for all } h \in [0, g] \quad (14)$$

both hold – independently of the value of  $\Delta_2$ . Quick calculation proves the first inequality of (14): Recall that the choice on  $(g, \delta)$  implies  $(1 - g) \geq 0$ ,  $2\delta g \geq 0$ , and  $(1 - 2g) \geq 0$ , yielding

$$F(\delta, g, g) - (1 - \delta)g = \frac{(1 - g)\delta}{(1 - g) + \delta} (2\delta g + (1 - 2g)) \geq 0$$

To obtain the second inequality of (14), we show that for the chosen  $\delta, g$ , the function  $F(\delta, g, h)$  is monotone non-decreasing on  $h \in [0, g]$ : Recalling  $h \leq g \leq 1/2$  we find  $(1 - 2h) \geq 0$ , implying

$$\frac{dF(\delta, g, h)}{dh} = \frac{(1 - 2h)(1 - \delta)^2 + g\delta(1 - \delta)}{(1 - g) + \delta} \geq 0$$

This finally proves (14), and thus the second inequality in (13), concluding the proof of (11). To prove (12) fix some  $d \in [0, 1]$  arbitrarily and define  $\mathfrak{p}(g) := g(1 - 2g)(-1 + 4g - 2g^2)$ , which is the radicand appearing in  $F(\Delta_1(g), g, g)$ . Brief calculation leads to

$$\begin{aligned} & (F(d, g, g) - F(\Delta_1(g), g, g)) (1 - g + d) = \\ & \underbrace{(4g - 1)(1 - g)^3 + 2g(1 - 2g)(1 - g)d + g(1 - 2g)d^2}_{\geq 0 \quad (\star)} - \underbrace{2(1 - g + d)(1 - g)\sqrt{\mathfrak{p}(g)}}_{\geq 0 \quad (\star\star)} \end{aligned}$$

To obtain  $(\star)$ , recall  $1/3 < \beta < 1/2$  and observe that the restrictions on  $g, d$  imply  $g, d \geq 0$  as well as  $(4g - 1) \geq 0$ ,  $(1 - g) \geq 0$ , and  $(1 - 2g) \geq 0$ . Moreover we have  $(1 - g + d) > (1 - g) \geq 0$ , showing  $(\star\star)$ . It can also be easily verified that  $\mathfrak{p}(g) \geq 0$  for the chosen  $g$ . Hence exploiting  $(1 - g + d) > 0$  and Fact 3 we obtain that  $F(d, g, g) - F(\Delta_1(g), g, g)$  is nonnegative if and only if the following quantity is nonnegative:

$$((4g - 1)(1 - g)^3 + 2g(1 - 2g)(1 - g)d + g(1 - 2g)d^2)^2 - 4(1 - g + d)^2(1 - g)^2\mathfrak{p}(g)$$

This turns out to be equivalent to:

$$((1 - 3g)(1 - g)^2 + 2g(1 - 2g)(1 - g)d + g(1 - 2g)d^2)^2 \geq 0$$

The last inequality is trivially true, which finally proves (12) since  $(g, d) \in [1/3, \beta] \times [0, 1]$  were chosen arbitrarily.

## 6 Discussion

It is worth noticing that the analysis of both presented algorithms is tight. Tracing all inequalities used, we constructed the following worst-case example, on which the second algorithm yields a 0.36392-approximation of the equilibrium:

$$R = \begin{pmatrix} 0 & \alpha & \alpha \\ \alpha & 0 & 1 \\ \alpha & 1 & 0 \end{pmatrix} \quad C = \begin{pmatrix} 0 & \alpha & \alpha \\ \alpha & 1 & 1/2 \\ \alpha & 1/2 & 1 \end{pmatrix} \quad \text{where } \alpha = 1/\sqrt{6}.$$

In general, our algorithms produce solutions with large support. This is to no surprise, as implied by negative results on the existence of approximate equilibrium strategies with small support [1,7].

The major open question remains whether a polynomial time algorithm for any constant  $\epsilon > 0$  is possible. It would also be interesting to investigate if our methods can be modified to yield better approximations.

## References

1. Althöfer, I.: On sparse approximations to randomized strategies and convex combinations. *Linear Algebra and Applications* 199, 339–355 (1994)
2. Chen, X., Deng, X.: Settling the complexity of 2-player Nash equilibrium. In: *Annual IEEE Symposium on Foundations of Computer Science*, pp. 261–272. IEEE Computer Society Press, Los Alamitos (2006)
3. Chen, X., Deng, X., Teng, S.: Computing Nash equilibria: Approximation and smoothed complexity. In: *Annual IEEE Symposium on Foundations of Computer Science*, pp. 603–612. IEEE Computer Society Press, Los Alamitos (2006)
4. Daskalakis, C., Goldberg, P., Papadimitriou, C.: The complexity of computing a Nash equilibrium. In: *Annual ACM Symposium on Theory of Computing*, pp. 71–78. ACM Press, New York (2006)
5. Daskalakis, C., Mehta, A., Papadimitriou, C.: A note on approximate Nash equilibria. In: *Workshop on Internet and Network Economics*, pp. 297–306 (2006)
6. Daskalakis, C., Mehta, A., Papadimitriou, C.: Progress on approximate Nash equilibria. In: *ACM Conference on Electronic Commerce*, ACM Press, New York (2007)
7. Feder, T., Nazerzadeh, H., Saberi, A.: Approximating Nash equilibria with small-support strategies. In: *ACM Conference on Electronic Commerce*, ACM Press, New York (2007)
8. Goldberg, P., Papadimitriou, C.: Reducibility among equilibrium problems. In: *Annual ACM Symposium on Theory of Computing*, pp. 61–70. ACM Press, New York (2006)
9. Kontogiannis, S., Panagopoulou, P., Spirakis, P.: Polynomial algorithms for approximating Nash equilibria of bimatrix games. In: *Workshop on Internet and Network Economics* (2006)
10. Kontogiannis, S., Spirakis, P.: Efficient algorithms for constant well supported approximate equilibria in bimatrix games. In: *34th International Colloquium on Automata, Languages and Programming* (2007)

11. Lipton, R., Markakis, E., Mehta, A.: Playing large games using simple strategies. In: 4th ACM Conference on Electronic Commerce, pp. 36–41. ACM Press, New York (2003)
12. Nash, J.F.: Non-cooperative games. *Annals of Mathematics* 54, 286–295 (1951)
13. Spirakis, P., Tsaknakis, H.: An optimization approach for approximate Nash equilibria. In: Workshop on Internet and Network Economics (2007)