NATIONAL RESEARCH UNIVERSITY
HIGHER SCHOOL OF ECONOMICS

M. Gavrilovich, V. Kreps

ON A CLASS OF OPTIMIZATION PROBLEMS WITH NO “EFFECTIVELY COMPUTABLE” SOLUTION

BASIC RESEARCH PROGRAM

WORKING PAPERS

SERIES: ECONOMICS
WP BRP 112/EC/2015

M. Gavrilovich[1], *V. Kreps*[2]

ON A CLASS OF OPTIMIZATION PROBLEMS WITH NO "EFFECTIVELY COMPUTABLE" SOLUTION

It is well-known that large random structures may have non-random macroscopic properties. We give an example of non-random properties for a class of large optimization problems related to the computational problem $MAXFLS^{=}$ of calculating the maximal number of consistent equations in a given overdetermined system of linear equations. A problem of this kind is faced by a decision maker (an Agent) choosing the means to protect a house from natural disasters. For this class we establish the following. There is no "efficiently computable" optimal strategy for the Agent. When the size of a random instance of the optimization problem goes to infinity the probability that the uniform mixed strategy of the Agent is ε optimal goes to one. Moreover, there is no "efficiently computable" strategy for the Agent which is substantially better for each instance of the optimization problem.

Keywords: optimization, concentration of measure, probabilistically checkable proofs

JEL Classification: C61

[1] Misha Gavrilovich, Higher School of Economics at St.Petersburg, mishap@sdf.org
[2] Viktoria Kreps, Higher School of Economics at St.Petersburg, vita_kreps@mail.ru

1. Introduction

It is well-known that macroscopic properties of large random objects are often not random; sometimes this is called concentration of measure. The simplest example is the law of large numbers: the arithmetic mean of many independent identically distributed (i.d.d) variables is non-random. An example from physics is delivered by the properties of a gas: pressure and temperature depend on the average speed of molecules and not their individual velocities. A simple example in game theory is a matrix game with a random square payoff matrix of fixed size. The matrix elements are i.d.d. Gaussian random variables. Both players know the matrix realized in the game. Consider the same game repeated n times; each time the matrix is chosen independently. If the number n of repetitions goes to infinity, then the value of this n-stage game divided by n approaches the expectation of the value of the 1-stage game. The latter is not a random variable.

In game theory a similar concentration of measure phenomena were studied in [1–5]. We show the concentration of measure for a special class of random optimization problems of large size which have no optimal effectively computable solutions.

First we show that the uniform mixed strategy using all the pure strategies is equiprobably close to the optimal solution with an overwhelming probability when the size of the problem is large enough. Note that this “almost” optimal strategy does not depend on the parameters of the game and thus may be used when no information is available about the problem (except its size).

Then we show that there is no effectively computable strategy which is substantially better for all the problems of the class considered.

2. Basic notions and the formulation of the problem

First we introduce the preliminary notions necessary to define the class of optimization problems. Let $C_{N,K}$, $K >> N$ be a overdetermined unsolvable system of K linear equations with N variables, for example with real coefficients.

$MAXFLS^{=}$ is the following computational problem: *given such a system C, find the maximal possible number $M = M(C_{N,K}$ of consistent (i.e. simultaneously satisfiable) equations, and the values of the variables satisfying that many equations.* It is well-known that $MAXFLS^{=}$ is NP-hard (cf. [6]), i.e. for each computational problem in the class NP there is a polynomial time algorithm reducing it to $MAXFLS^{=}$.

Consider systems of linear equations of the following form (*). Let x_r, $r = 1, \dots, N$, take values 0 and 1; each equation has exactly 3 variables and is of form $x_r + x_s + x_t = a_{r,s,t}$ modulo 2, where $a_{r,s,t}$ is also equal to either 0 or 1, $1 \le r < s < t \le N$. We assume that the equations are distinct and each pair of equations is consistent.

It is easy to see that the maximal possible number K of different equations in such a system is bounded by

$$K \le 2 \cdot \frac{N(N-1)(N-2)}{6}.$$

Consider a random assignment of values to the variables in which each variable takes both possible values 0 and 1 independently and equiprobably with probability $1/2$.

For each equation the expectation that it is satisfied by such an assignment is $1/2$, and therefore the expectation of the maximal possible number $M(C_{N,K})$ of consistent equations is at least $K/2$, i.e. for the maximal possible number $M = M(C_{N,K})$ of simultaneously satisfied equations of the system $C = C_{N,K}$ the following inequality holds:

$$M(C_{N,K})/K \geq 1/2. \tag{1}$$

Hastad [7] considers systems of form (*) and proves that it is NP-hard to approximately compute $M(C_{N,K})$ beyond the random threshold, namely for each $\varepsilon > 1/2$, that it is NP-hard to determine whether

$$M/K > 1/2 + \varepsilon. \tag{2}$$

We call a system $C_{N,K}$ of equations *degenerate* iff each equation in the system has its opposite, i.e. the equation with the same left-hand and different right-hand side.

Given a non-degenerate system $C_{N,K}$ of form (*), consider the following optimization problem.

An Agent wants to build a house protected from as many natural disasters as possible. The Agent controls N parameters x_r, $r = 1, \ldots, N$ taking values 0 and 1. $\boldsymbol{x} = (x_1, \ldots, x_N)$—a pure strategy of the Agent. There are 2^N pure strategies. The Agent is able to randomize his strategies.

There are K possible natural disasters, $K >> N$. We consider the case when the natural disaster happens once. The disaster is chosen by Nature at random according to a probability distribution $\boldsymbol{p} = (p_1, \ldots, p_K)$ on the set $\{1, \ldots, K\}$ known to the Agent; here p_j are positive rational number.

The house is protected from the disaster j iff the j-th equation of the system $C_{N,K}$ is satisfied by the values of the parameters assigned by the Agent. The system is known to the Agent.

If the house is protected, we set the Agent's payoff to be $+1$; if it is not protected, we set the Agent's payoff to be 0.

The agent tries to maximize his expected payoff. The (mixed) strategy of the Agent which maximizes his expected payoff, is called *optimal*.

For this class of optimization problems denoted by $\Gamma(C_{N,K}, \boldsymbol{p})$ we show that:

1. the expectation of the Agent's payoff is at least $1/2$ when the Agent plays optimally;
2. there is no "effectively computable" optimal strategies of the Agent, i.e. there is no algorithm working in polynomial time and realizing the optimal strategy for each optimization problem of the class described above;
3. the Agent may ensure payoff $1/2$ by uniformly mixing all the pure strategies whatever probability distribution the Nature uses;
4. for some natural measure on the set of all optimization problems $\Gamma(C_{N,K}, \boldsymbol{p})$ this uniform mixed strategy is ε-optimal for $N \to \infty$ and $K/N \to \infty$, except on a set of small measure;
5. the Agent has no "efficiently computable" strategy which is significantly better for *all* optimization problems than his uniform mixed strategy if Nature plays her uniform

strategy choosing all the disasters with the same probability $1/K$.

Let us clarify the notion of the effectively computable mixed strategy for a class C of optimization problems. Consider an arbitrary probabilistic algorithm which, given (a description of) an optimization problem of class C, outputs a pure Agent strategy with a certain probability. We think that such an algorithms “computes” a mixed strategy for an arbitrary optimization problem of class C. If the algorithm always finishes the computation in a time bounded by a polynomial in the length of the description of the optimization problem, we say that the algorithm realizes an efficiently computable strategy.

3. Preliminary results

Lemma 1. $M(C_{N,K})/K = 1/2$ *iff the system* $C_{N,K}$ *is degenerate.*

Proof. It is easy to check that for degenerate systems the equality $M/K = 1/2$ holds.

Assume the equality $M(C_{N,K})/K = 1/2$ holds.

This means the maximum number of simultaneously satisfied equations is exactly equal to the average number (by the uniform “counting” measure) and therefore each assignment satisfies exactly $K/2$ equations. Let us prove the system is degenerate.

For this we consider how the maximum number of the simultaneously satisfied equations changes when a single variable changes its value, then when two variables change their values, and, finally, when three variables change their values.

Fix an arbitrary assignment $x_1 = c_1, \dots, x_N = c_N$. Consider sets $X_i(c_1, \dots, c_N)$, $1 \le i \le N$ of equations satisfied by $x_1 = c_1, \dots, x_N = c_N$ and containing occurrences of variable x_i. Similarly, consider the set $X'_i(c_1, \dots, c_N)$ of unsatisfied equations containing occurrences of the same variable x_i.

Consider the assignments where a single variable $x_i = c_i + 1$ changes its value and notice $|X_i| = |X'_i|$.

$$X_i(c_1, \dots, c_i, \dots, c_N) = X'_i(c_1, \dots, c_i + 1, \dots, c_N)$$

and

$$X'_i(c_1, \dots, c_i, \dots, c_N) = X_i(c_1, \dots, c_i + 1, \dots, c_N).$$

Analogously, consider assignments where two variables $x_r = c_r + 1$, $x_s = c_s + 1$ change their values.

Under this change of the assignment the following equations become unsatisfied which were satisfied before:

$$X_r(c_1, \dots, c_i, \dots, c_N) \cup X_s(c_1, \dots, c_i, \dots, c_N) \setminus (X_r(c_1, \dots, c_i, \dots, c_N) \cap X_s(c_1, \dots, c_i, \dots, c_N)).$$

Analogously, under this change of the assignment the following equations become satisfied which were unsatisfied before:

$$X'_r(c_1, \dots, c_i, \dots, c_N) \cup X'_s(c_1, \dots, c_i, \dots, c_N) \setminus (X'_r(c_1, \dots, c_i, \dots, c_N) \cap X'_s(c_1, \dots, c_i, \dots, c_N)).$$

Hence, the cardinality of these sets are equal; let us now calculate their size.

By assumption the latter two sets have the same number of elements. The number is $|X_r| + |X_s| - |X_r \cap X_t|$. To simplify notation, omit $(c_1, \dots, c_i, \dots, c_N)$.

By the inclusion-exclusion principle their size is equal to $|X_r| + |X_s| - |X_r \cap X_t|$ and $|X'_r| + |X'_s| - |X'_r \cap X'_t|$, respectively.

Now, $|X_r| = |X'_r|, |X_s| = |X'_s|$ implies $|X_r \cap X_s| = |X'_r \cap X'_s|$.

Analogously, $|X_r \cap X_t| = |X'_r \cap X'_t|$, $|X_s \cap X_t| = |X'_s \cap X'_t|$.

Now consider assignments where three variables $x_r = c_r + 1$, $x_s = c_s + 1$ and $x_t = c_t + 1$ change their values. Using the inclusion-exclusion principle in a similar way, we get

$$|X_r \cap X_s \cap X_t| = |X_r| + |X_s| + |X_t| - |X_s \cap X_t| - |X_r \cap X_t| - |X_s \cap X_t|$$

and analogously

$$|X'_r \cap X'_s \cap X'_t| = |X'_r| + |X'_s| + |X'_t| - |X'_s \cap X'_t| - |X'_r \cap X'_t| - |X'_s \cap X'_t|.$$

Hence $|X_r \cap X_s \cap X_t| = |X'_r \cap X'_s \cap X'_t|$ and these sets are either both empty or have a single element.

Finally, notice that $|X_r \cap X_s \cap X_t| = 1$ iff the system contains the equation $x_r + x_s + x_t = c_r + c_s + c_t$, and that $|X'_r \cap X'_s \cap X'_t| = 1$ iff the system contains the "opposite" equation $x_r + x_s + x_t = c_r + c_s + c_t + 1$ which is inconsistent with the former equation.

The same argument goes through if these two sets are both empty. This finishes the proof that the system is degenerate. □

Claim 1. *(a) The Agent has a pure optimal strategy* $x^* = x^*(C_{N,K}, \boldsymbol{p})$.

(b) If the Agent uses the optimal strategy, the payoff expectation $Val(\Gamma(C_{N,K}, \boldsymbol{p}))$ *is at least* $1/2$. *The equality holds iff the system contains a pair of contradictory equations (a degenerate system).*

(c) There is no "effectively computable" optimal strategy for the Agent, i.e. there is no algorithm which takes at most polynomial time (in K*) and realizes the Agent's optimal strategy for each problem* $\Gamma(C_{N,K}, \boldsymbol{p})$ *of the form described above.*

Proof. Let us show that for an arbitrary mixed strategy $\boldsymbol{p}$ of Nature with positive rational components $p_j > 0$ there exists the best answer of the Agent as a pure strategy.

First consider the case when $\boldsymbol{p}$ is the uniform distribution: $\boldsymbol{p}$: $p_j = 1/K$.

Let $\boldsymbol{p}$, $p_j = 1/K$, $j = 1, \dots, K$, represent the uniform strategy of Nature. It is easy to see that there is a unique optimal strategy for the Agent corresponding to this strategy of Nature, namely this strategy is pure and assigns to variables the values which satisfy the maximal number

$M(C_{N,K})$ of equations. We know it is NP-hard to compute this strategy of the Agent, as it is NP-hard to compute the number $M(C_{N,K})$. In other words, the (mixed) optimal strategy is NP-hard to compute.

Using Lemma 1, inequality 1 implies item (b), and item (c) is implied by the fact that the problem of finding $M(C_{N,K})$ is NP-hard.

Consider the general case: Nature plays the strategy represented by $\boldsymbol{p} = (p_1, \dots, p_K)$ where $p_j > 0$ are rational, $j = 1, \dots, K$. Let $D = D(\boldsymbol{p})$ be the least common multiple of the denominators of p_j, $j = 1, \dots, K$. Thus $p_j = d_j/D$ where $\sum_{j=1}^{K} d_j = D$ and d_j are integer. Note that $D > K$ unless the distribution $\boldsymbol{p}$ is uniform and $D = K$.

Consider the system $\tilde{C}_{N,D}$ of D linear equations of N variables which, for $j = 1, \dots, K$, contains d_j copies of j-th equation of the system C. Let $\boldsymbol{q}$ be the uniform distribution on the equations of the system $\tilde{C}_{N,D}$. Now consider the game instances $\Gamma(C_{N,K}, \boldsymbol{p})$ where Nature plays $\boldsymbol{p}$ and $\Gamma(\tilde{C}_{N,D}, \boldsymbol{q})$ where Nature plays $\boldsymbol{q}$. They are equivalent and the corresponding optimal strategies for the Agent are the same, to assign to variables the values which satisfy the maximal number $M(\tilde{C}_{N,D})$ of equations of system $\tilde{C}_{N,D}$. □

Claim 2. *For any distribution* $\boldsymbol{p}$ *the Agent's mixed strategy where he equiprobably (with probability* $1/2^N$*) picks a pure strategy to use, gives payoff* $1/2$.

Proof. Recall that we assume that all the components of the distribution $\boldsymbol{p}$ are positive.

An argument analogous to the proof of inequality (1) shows that the payoff is $1/2$ when the Agent uses the uniformly mixing strategy: the expectation that each equation is satisfied by the assignment is equal to $1/2$. Hence, the expectation of the payoff is equal to $1/2$ as well.

This strategy is optimal iff the system is degenerate, i.e. the system $C_{N,K}$ has a pair of contradictory equations.

4. ε-optimality of the Agent's uniform strategy

In this section we consider probability measures on a set of systems (optimization problems) and show that the uniform strategy is ε-optimal with a probability approaching 1 when $N \to \infty$ and $K/N \to \infty$.

In Theorem 1 below we assume that the distribution $p_{\boldsymbol{uni}}$ is uniform and we denote $\Gamma(C_{N,K}, p_{\boldsymbol{uni}})$ as the corresponding optimization problem. We also assume that the system $C_{N,K}$ contains no pairs of inconsistent equations.

Consider the uniform probability measure on the family of the systems of the form (*) containing exactly K different equations. Let $\mu_{N,K}$ denote this measure. We show that for a random system of the form (*), chosen according to measure $\mu_{N,K}$, without pairs of inconsistent equations, the Agent's uniform strategy which randomly chooses from each of the pure strategies with equal probability, is asymptotically ε-optimal except for a set of small $\mu_{N,K}$-measures.

Let $a(C_{N,K})$ denote the difference $M(C_{N,K}/K - 1/2$. Fix $\varepsilon > 0$. Consider the set of systems

$UNI_{N,K}(\varepsilon)$,

$$UNI_{N,K}(\varepsilon) = \{C_{N,K}: a(C_{N,K}) < \varepsilon\}.$$

If $C_{N,K} \in UNI_{N,K}(\varepsilon)$, then Claim 2 implies that the Agent's mixed strategy which uses equiprobably (with probability $1/2^N$) each pure strategy, is the ε-optimal strategy in the optimization problem $\Gamma(C_{N,K}, p_{\boldsymbol{uni}})$.

Let $\lceil(1/2-\varepsilon)K\rceil$ denote the least integer not less than $(1/2-\varepsilon)K$, and let $\binom{K}{\lceil(\varepsilon+1/2)K\rceil}$ denote the the number of $\lceil(1/2-\varepsilon)K\rceil$-element subsets of a set of size K.

Theorem 1. *Fix $\varepsilon > 0$. The probability $\mu_{N,K}$ of systems $C_{N,K}$ such that the Agent's uniform strategy is ε-optimal for $\Gamma(C_{N,K}, p_{\boldsymbol{uni}})$, goes to 1 when $N \to \infty$ and $K/N \to \infty$.*

Namely, this probability satisfies the inequalities

$$(I)\ 1 - \ {}_2F_1(1, \lceil(1/2-\varepsilon)K\rceil, \lceil(1/2+\varepsilon)K\rceil + 1, -1) \cdot 2^{N-K} \cdot \binom{K}{\lceil(\varepsilon+1/2)K\rceil}$$

$$\leq \mu_{N,K}(UNI_{N,K}(\varepsilon))$$
$$\leq 1 - \ {}_2F_1(1, \lceil(1/2-\varepsilon)K\rceil, \lceil(1/2+\varepsilon)K\rceil + 1, -1) \cdot 2^{-K} \cdot \binom{K}{\lceil(\varepsilon+1/2)K\rceil},$$

where ${}_2F_1(a,b,c,d)$ is the hypergeometric function (cf. [8]). For some constant $c > 0$ it holds

$$(II)\ 1 - 2^N \int_K^\infty e^{-t^2}\, dt \cdot exp(2\varepsilon^2 K + c)) \leq \mu(UNI_{N,K}(\varepsilon))$$

$$\leq 1 - \int_K^\infty e^{-t^2}\, dt \cdot exp(2\varepsilon^2 K - c)).$$

Note that the estimate (II) is more explicit but less precise.

Proof. As the measure $\mu_{N,K}$ is uniform, it is enough to find an upper bound for the fraction (among all systems) of the systems such that $M/K > (\frac{1}{2}+\varepsilon)$, i.e. those systems for which some assignment satisfies at least $(\frac{1}{2}+\varepsilon)$-fraction of the equations. It is enough to estimate the number of equations with fixed left-hand sides which satisfy at least $(\frac{1}{2}+\varepsilon)$-fraction of the equations on a given assignment.

So let us fix the left hand sides of all the K equations of the system. Further, fix an arbitrary assignment of values to the variables; there are 2^N of them. A system with given left hand sides of the equations is fully determined by the set of its equations which are satisfied by a given assignment of values to the variables.

Let $\binom{K}{>L}$ denote the number of subsets of the set of K elements which contain more than L

elements. Note that $\binom{K}{>(\frac{1}{2}+\varepsilon)K}$ is the number of systems with given left hand sides such that a given assignment satisfies more then $(\frac{1}{2}+\varepsilon)$-fraction of the number of the equations. In terms of probability theory, $\binom{K}{>(\frac{1}{2}+\varepsilon)K} = 2^K P_\varepsilon$ where P_ε is the probability of that more than $(\frac{1}{2}+\varepsilon)K$ heads occur in the Bernoulli trial of K tosses of a uniform coin. Here we use that there are no two equations have the same left hand sides.

This number may be written explicitly as the hyper-geometric function

$$\binom{K}{> (\frac{1}{2}+\varepsilon)K} = 2^{-K}\binom{K}{\lceil(\frac{1}{2}+\varepsilon)K\rceil} \quad {}_2\mathrm{F}_1\left(1, \lceil\left(-\frac{1}{2}+\varepsilon\right)K\rceil;\ \lceil\left(\frac{1}{2}+\varepsilon\right)K\rceil + 1;\ -1\right).$$

We can also use estimates of Littlewood [9,10]:

$$P_\varepsilon = \int_K^\infty e^{-t^2}\, dt \cdot exp(2\varepsilon^2 K + c_0 K^{-1/2} + O(K^{-1})).$$

Constant c may be given explicitly but we only remark that $c < 0.6$.

Estimates (II) follow from (I) and the fact that $c_0 K^{-1/2} + O(K^{-1})$ are bounded both from above and below by a constant.

Finally, note that in this way we may have counted each system several times, but at most 2^N, and this implies the estimates. □

We remark that the measure $\mu_{N,K}$ is not natural in the sense that it is not generated by a natural stochastic process. This measure may be generated by the process which adds a new equation to the set of equations already added. However, this process needs to check whether the equation does not belong to the set already and this is what it makes unnatural from the point of view of the theory of stochastic processes.

Now let us show that the proof of Theorem 1 goes through for a larger class of optimization problems and measures satisfying certain symmetry assumptions. Namely, Theorem 1 generalizes to problems $\Gamma(C_{N,K}, \boldsymbol{p})$ where the distribution $\boldsymbol{p}$ and the number K of equations are arbitrary. The class of probability measures is defined with help of the following parameters.

Fix three constants $\delta(N),\ 0 < \delta(N) < 1,\ N < K_{min} < K_{max}$.

Define the following conditions on a measure ν_N of the countable set $\Gamma(C_{N,K}, \boldsymbol{p})$ of systems.

1. the symmetry measure of a system (i.e., the measure of a singleton set) does not depend on the right hand sides of the equations occurring in the system;
2. the error measure of the set of systems with K equations where $K_{min} < K < K_{max}$, is more than $1 - \delta(N)$.

Let us give an example of a measure which is natural from the point of view of probability theory and satisfies condition (1) of uniformity of the measure's restriction to a set of systems with fixed left hand sides.

Define the following random process which builds a system of not more than K different equations. Independently and equiprobably pick an equation K times. This defines a probability measure on the set of systems of linear equations with multiplicities and the corresponding optimization problem $\Gamma(C_{N,K'},\boldsymbol{p})$ with a random number of equations $K' < K$. The probability p_j, $1 \le j \le K'$, is proportional to the number of times the corresponding equation was chosen by the random process. At each step, and for each equation, the probability of picking the equation is equal to the probability of picking the equation with the same left hand side and the opposite right hand side; this implies the measure satisfies condition (1) of uniformity of the restriction.

Let us now show that this measure also satisfies condition (2). Take $K_{min} = (1-\varepsilon)K$ and $K_{max} = (1+\varepsilon)K$. To estimate error $\delta(N)$, we use Bernoulli tails estimates, for example Theorem 1.5 [11] gives a bound $\delta(N) < \frac{K^{-1/2}}{\varepsilon} e^{-\varepsilon^2 K/6}$ for $K < N^3/2$ i $3/K < \varepsilon < 1/12$.

Let us introduce the following notation to state Theorem 2. Let $a(\Gamma(C_{N,K},\boldsymbol{p})) = Val(\Gamma(C_{N,K},\boldsymbol{p}) - 1/2$. Fix $\varepsilon > 0$. Consider the set of systems $UNI'_{N,K}(\varepsilon)$ defined as

$$UNI'_{N,K}(\varepsilon) = \{\Gamma(C_{N,K},\boldsymbol{p}) : a(\Gamma(C_{N,K},\boldsymbol{p})) < \varepsilon\}.$$

Theorem 2. Fix $\varepsilon > 0$. Let ν_N denote a measure satisfying conditions (1) and (2).

The probability $\nu_N(\mathrm{UNI'}_{N,K}(\varepsilon))$ that in a random (according to measure ν_N) instance of the optimization problem $\Gamma(C_{N,K},\mathbf{p})$ the uniform strategy of the Agent is $\boldsymbol{\varepsilon}$-optimal goes to 1 when $N \to \infty$ and $K/N \to \infty$.

Namely, this probability ν_N of the set of the optimization problems such that the uniform strategy is $\boldsymbol{\varepsilon}$-optimal, satisfies the following inequalities:

$$\text{(I)}\ \ 1 - 2^{N-K_{min}} \binom{K_{min}}{\lceil(\varepsilon + 1/2)K_{min}\rceil}\ {}_2F_1(1, \lceil(1/2-\varepsilon)K_{min}\rceil, \lceil(1/2+\varepsilon)K_{min}\rceil + 1, -1) - \delta(N) \le \nu_N(\mathrm{UNI'}_{N,K}(\varepsilon))$$

$$\nu_N(\mathrm{UNI'}_{N,K}(\varepsilon)) \le 1 - 2^{-K_{max}} \binom{K_{max}}{\lceil(\varepsilon + 1/2)K_{max}\rceil}\ {}_2F_1(1, \lceil(1/2-\varepsilon)K_{max}\rceil, \lceil(1/2+\varepsilon)K_{max}\rceil + 1, -1) + \delta(N).$$

For some constant $c > 0$

$$\text{(II)}\ \ 1 - 2^N \int_{K_{min}}^{\infty} e^{-t^2}\, dt \cdot \exp(2\varepsilon^2 K_{min} + c) - \delta(N) \le \nu_N(\mathrm{UNI'}_{N,K}(\varepsilon))$$

$$\le 1 - \int_{K_{max}}^{\infty} e^{-t^2}\, dt \cdot \exp(2\varepsilon^2 K_{max} - c)) + \delta(N).$$

Proof. Pick a number K' and pick K' many left hand sides of equations. Consider the measure induced on the systems with these left hand sides. Their number is $2^{K'}$ and by condition (1) the measure induced on these systems is uniform. Hence we may use the estimates (I), (II) for the uniform measure in the proof of Theorem 1. Note that the dependence on K' in these estimates is monotone; also note that by (2) we have the inequality $K_{\min} < K < K_{\max}$ holds on a set of measure at least $1 - \delta(N)$. Hence, the lower bound in Theorem 2 is obtained from the lower bound in Theorem 1 by replacing K by $K_{\min}$ and subtracting the error term $\delta(N)$. Similarly, the upper bound in Theorem 2 is obtained from the upper bound in Theorem 1 by replacing K by $K_{\max}$ and adding the error term $\delta(N)$. □

5. The unimprovability of the uniform strategy of the Agent in the class of effectively computable strategies

As Theorems 1 and 2 imply, the uniform strategy of the Agent is close to being optimal for all problems in $\Gamma(C_{N,K}, \mathbf{p})$ except a set of small measure which approach 0 when the size of the system goes to infinity. For these exceptional problems it holds that

$$1/2 + \varepsilon < Val(\Gamma(C_{N,K}, \mathbf{p}) \leq 1.$$

In this section we investigate a "worst" case for the Agent to use the uniform strategy, namely when his payoff expectation $\mathrm{Val}(\Gamma(C_{N,K}, \mathbf{p}))$ is close to 1,

$$\mathrm{Val}(\Gamma(C_{N,K}, \mathbf{p}) \geq 1 - \varepsilon.$$

Assuming that Nature chooses equations equiprobably, $\mathbf{p}$: $p_j = 1/K$, $j = 1, \dots, K$, we show that there is no “effectively computable” mixed Agent’s strategy which is substantially better for problems $\Gamma(C_{N,K}, \mathbf{p})$ under consideration. From now on we omit the parameter $\mathbf{p}$ as it is fixed.

To demonstrate this we use the result of Hastad [6] which establishes that for a fixed $\varepsilon > 0$ it is NP-hard to check whether there exists an assignment of values of the variables which satisfy at least $(1/2 + \varepsilon)K$ many equations in an arbitrary system of form (*).

Let us clarify the motivation and methods used by Hastad. Using Fourier transforms, Hastad [6] constructed a probabilistically checkable proof system (PCP, cf. (Arora & Barak [12]) with minimal parameters.

Usually, an algorithm used to check a proof takes the whole of proof as input and reads the whole proof. A probabilistically checkable proof is a proof rewritten in such a way that that a modified algorithm checks it reading 3 bits at random and finds an error with probability at least $1/2$.

Dinur [13] gives a new construction of probabilistically checkable proofs which uses expander graphs.

We note that although the size of the probabilistically checkable proofs built is bounded by $n \times (\log n)^{O(1)}$ where n is the length of a classical proof, the constant appearing $O(1)$ may be quite large.

We use an equivalent reformulation of the result of Hastad which says that for the considered class of optimization problems constructing an optimal solution is asymptotically as hard as constructing an approximate solution with a given rate of approximation.

Consider the subfamily $\Gamma_{1-\varepsilon} = \Gamma_{1-\varepsilon}(C_{N,K})$ of the systems $\Gamma(C_{N,K})$ of form (*) in which the Agent's optimal payoff is at least $1 - \varepsilon$.

Reformulation of the Hastad's result. Fix $\varepsilon > 0$. It is NP-hard to find a mixed strategy for the Agent which gives him payoff at least $1/2 + \varepsilon$ in any problem $\Gamma_{1-\varepsilon}$.

Corollary. Assume $P \neq NP$. Fix $\varepsilon > 0$. There is no "effectively computable" mixed strategy which gives the Agent a payoff of at least $1/2 + \varepsilon$ in an arbitrary optimization problem in $\Gamma_{1-\varepsilon}$, i.e. there is no algorithm working in polynomial time and computing a possibly mixed strategy with payoff at least $1/2 + \varepsilon$ in an arbitrary problem $\Gamma_{1-\varepsilon}$.

6. Conclusions

We consider a class of optimization problems interpreted as problems arising when an Agent wants to build a house and choose ways to protect it from natural disasters. We showed that in this class there are no effectively computable optimal strategies. However, the uniformly mixing strategy—to choose equiprobably from all available pure strategies—is close to being optimal for an overwhelming majority of optimization problems when their size is large enough. Therefore, it is advisable to use this strategy even when little is known about the parameters of the optimization problem.

Under the assumption that natural disasters happen with the same probability, we showed that there is no effective strategy which is always essentially better than the uniform strategy for all the problems, in particular those problems where the uniform strategy is not close to optimal.

The authors would like to thank Fedor Sandomirsky for valuable remarks and references.

References

[1] D. G. Kabe (1983) Laws of large numbers for random payoff games J. Ind. Math. Soc., Vol.33 (Part 1), 73-87.

[2] J. Jonasson (2004) On the optimal strategy in a random game Elect. Comm. in Probab. Vol.9 , 132-139.

[3] A. McLennan, J. Berg (2005) Asymptotic expected number of Nash equilibria of two-player normal form games Games and Economic Behavior, Vol.51, I.2, 264вЂ“295.

[4] A. Tanikawa (2011) On stability of the values of random zero-sum games International Journal of Innovative Computing, Information and Control Vol. 7, I. 1, 133-140.

[5] Sandomirskiy (2015) On typical properties of large zero-sum games The Int. Conference SING-GTM2015. Abstracts. Eds. by L.Petrosjan and N.Zenkevich, St.Petersburg.

[6] E.Amaldi, V.Kann (1995) The complexity and approximability of finding maximum feasible subsystems of linear relations. Theoretical Computer Science Vol.147, 181-210.

[7] J.Hastad (1997) Some optimal inapproximability results. Proceedings of 29th Annual Symposium on Theory of Computation, El Paso, 1-10.

[8] Y.Luke (1975) Mathematical functions and their approximations. Academic Press.

[9] B. D. McKay (1989) On Littlewood's estimate for the binomial distribution. Adv. Appl. Prob., http://cs.anu.edu.au/ Brendan.McKay/papers/littlewood2.pdf

[10] J. E. Littlewood (1969) On the probability in the tail of a binomial distribution. Adv. Appl. Prob., , 43вЂ“72.

[11] B. Bollobas (2001) Random Graphs. Cambridge studies in advanced mathematics, 2nd ed.

[12] S.Arora, B. Barak (2009) Computational Complexity: A Modern Approach Cambridge University Press.

[13] I.Dinur (2005) The PCP theorem by gap amplification. Technical Report TR05-046, ECCC, 2005. Revision 1, available from http://eccc.uni-trier.de/eccc-reports/2005/TR05-046/.

CORRESPONDING AUTHOR

Misha Gavrilovich
Higher School of Economics at St.Petersburg, mishap@sdf.org