

EÖTVÖS LORÁND UNIVERSITY

FACULTY OF SCIENCE

LÁSZLÓ VARGA

MSC IN MATHEMATICS, 1ST YEAR

PROBLEMS IN PPA
AND A GENERALIZATION OF
OLSON'S THEOREM

TUDOMÁNYOS DIÁKKÖRI DOLGOZAT

Advisor:

László Végh

London School of Economics and Political Science

Egerváry Research Group on Combinatorial Optimization

(Department of Operations Research, ELTE)



30th December 2012

Abstract

In complexity theory, search problems request a certificate instead of providing only the yes/no answer to a decision question. In [1], Papadimitriou began to study such search problems, where the existence of a witness is guaranteed by a mathematical theorem. The class PPA (Polynomial Parity Argument) is one of the complexity classes introduced by Papadimitriou, it contains such search problems, where the witness is established via the simple parity argument.

In the first part of this work, we prove that the search problem corresponding to the Combinatorial Nullstellensatz ([2]) over $\mathbb{F}_2$ is contained in PPA. This result answers an open question by D. West ([4]). Then, we give a new proof of Olson's theorem ([3]) via Combinatorial Nullstellensatz, which shows that the corresponding search problem belongs to PPA. As a consequence, the problem of finding 2^d -divisible subgraphs in a graph with more than $n \cdot (2^d - 1) - 2^{d-1}$ edges is also contained to PPA.

In the second part of this work, we investigate a generalization of Olson's theorem. Roughly speaking, we wish to determine how many vectors could be given such that none of the sum of subsequences is in a certain subset modulo prime powers. In most cases, our result gives better estimations than previously known results. We also present our conjectures and the scope of the future work. Most of these results are new, some of them has already contained in my thesis [6].

1 Introduction

Complexity theory usually addresses decision questions that can be answered by yes or no, for example, determine whether a given graph contains a Hamiltonian path. Decision problems are classified into complexity classes such as P and NP. The above problem is NP-complete, meaning intuitively that it is hard to solve: we cannot expect a polynomial time algorithm to decide this unless $P=NP$.

Still, there are certain settings when the decision question is not the right way to address the problem. An eminent example is finding a mixed Nash equilibrium in a finite two-player game. The decision problem would ask the question whether a mixed Nash equilibrium exists. The answer is trivially yes, due to Nash theorem, but this does not provide an actual mixed Nash equilibrium – finding one is a highly important and well-studied question.

Another significant example is finding the prime factorization of a positive integer. The factorization obviously exists but no polynomial time algorithm is known to find it.

A further interesting example related to Hamiltonian paths is the following: given a Hamiltonian path, find another Hamiltonian path in the graph or in the complement graph. The existence of this other Hamiltonian path is guaranteed by Lovász's theorem. (See e.g. [1] or [6].)

Such examples motivate the study of search problems: instead of providing only the yes/no answer to a decision question, we request a certificate verifying the answer.

In [1], Papadimitriou initiated a study of such search problems, where the existence of a witness is guaranteed by a mathematical theorem. No polynomial time algorithm is known for the above three examples, however, they do not seem to be as hard as solving NP-complete problems either. Papadimitriou defined several interesting classes of search problems that lie between the classes FP and FNP – these complexity classes of the search problems correspond to P and NP, respectively ('F' stands for 'functional'). The existence of witnesses in these classes are guaranteed by simple graph theoretic arguments. However, these arguments do not provide polynomial time algorithms for finding the witness.

We present one of these classes and its basic notations in the following part of this section. This class, the class PPA (Polynomial Parity Argument) is based on a simple parity fact in graphs.

In Section 2, we discuss the complexity of a general algebraic theorem, the Combinatorial Nullstellensatz by Alon ([2]), which guarantees roots of special polynomials. We can solve Combinatorial Nullstellensatz over $\mathbb{F}_2$ in polynomial time, if the polynomial is given explicitly and we prove that it is in PPA, if the polynomial is given implicitly. Any other results about the complexity of Combinatorial Nullstellensatz were not previously known.

In Section 4, we study further problems in Graph Theory, and we prove that they are in PPA. These results are the consequences of a combinatorial-algebraic theorem, Olson theorem which we discuss in Section 3 .

We give a self-interesting extension of Olson's theorem which is concerned with the sum of subsequences modulo prime power. Olson's theorem states that if there are given sufficiently many vectors of integers, there exists a subset of the vectors which fullfils simultaneous divisibility. For instance, it is a generalization of the well-known Erdős-Ginzburg-Ziv-theorem.

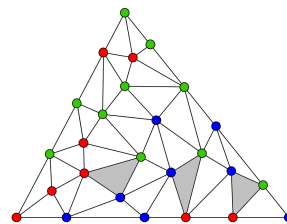
We extend this theorem to being in certain subsets modulo prime powers instead of the divisibility. We proved this generalization via Combinatorial Nullstellensatz.

1.1 The Parity Argument, the class PPA

The following observation is elementary, but it has numerous application in various theorems.

Every finite graph has an even number of odd-degree nodes.

For instance, Sperner's theorem, which states that a triangulation of a right-colored triangle always has a triangle with three colors, can be proved via this Parity Argument. Then, Brouwer's fixpoint theorem can be reduced to Sperner's, and the Nash theorem can be proved using Brouwer's theorem.



In [1], Papadimitriou defines the class PPA (Polynomial Parity Argument), which contains such search problems that are based on the Parity Argument.

We define the class PPA in a slightly different way, but following the same lines. First, we present the basic notations of it.

A problem is in PPA if and only if it is polynomially reducible to the following problem.

SHORT END OF THE LINE. Given an integer n and an undirected finite graph $G = (V, E)$ in which each node has at most two incident edges. The graph is described in the following way: it can be assumed that each node has a unique code from Σ^n , that is $V \subseteq \Sigma^n$. Furthermore, there is a polynomial algorithm (in n) which computes for a node its neighbours: for a $v \in V$, it lists the neighbours of v . At last, there is a node ε which has exactly one incident edge. Find another node which has exactly one incident edge.

The name of the problem refers to its trivial solution: from the given node ε , we can start and in each step we can move to the another neighbour of the actual node. At the end of the line, we find a node with one incident edge. Unfortunately, this trivial algorithm terminates only after exponential time in the worst case.

However, in general cases, it is more comfortable to use the following problem, the **END OF THE LINE**. The **SHORT END OF THE LINE** and the **END OF THE LINE** are polynomially reducible to each other, so their complexity is polynomially equivalent. (For details, please see [1] or [6].)

For defining this problem, we need the following notion of the *pairing function*.

Definition 1.1 *Given an undirected finite graph $G = (V, E)$. ϕ is called a pairing function, if $\phi : V \times V \rightarrow V$ such that for each node $v \in V$ it pairs up neighbours of v . That is, if v is an even-degree node, then for every node w with $\{v, w\} \in E$ there is another node $w' \neq v$ such that $\{v, w'\} \in E$, $\phi(v, w) = w'$ and $\phi(v, w') = w$. If v has odd degree, all neighbours but one similarly can be paired up. For v , there exists exactly one node w such that $\{v, w\} \in E$ and $\phi(v, w) = w$.*

END OF THE LINE. We are given an integer n and an undirected finite graph $G = (V, E)$. The graph is described in the following way: it can be assumed that each node has a unique code from Σ^n , that is $V \subseteq \Sigma^n$. Furthermore, we are given a polynomial time edge recognition algorithm and a polynomial time algorithm which computes $\phi(v, w)$ for a node v and one of its neighbours w . At last, we are given a node ε which has odd number of edges and a node δ which shows it: $\phi(\varepsilon, \delta) = \delta$. Find another node which has odd number of edges.

The name of the problem also refers to its trivial solution: from the given node ε , we can start and in each step we can move to $\phi(v, w)$ from v where w is the previous visited node. At last, at the end of the line, we find another node which has odd edges.

Hence, in a reduction to **END OF THE LINE**, we have to present the nodes of the graph and its edges which can be checked in polynomial time, the node ε called standard leaf and the pairing function, that is, for a fixed node v , we should pair up its neighbours. We say that $\phi(v, w)$ is the mate of w at v .

In [1], Sperner's, Brouwer's and Nash theorem are proved to be contained in PPA as we expected. Papadimitriou proved that some other problems are also in PPA.

For instance, he showed that the following problem CHÉVALLEY MOD 2 belongs to PPA.

Theorem 1.2 (Chévalley) *Let $\mathbb{F}$ be a finite field with characteristic p . Let $p_1, p_2, \dots, p_m$ be polynomials in n variables over $\mathbb{F}$. Suppose that $\sum_{i=1}^m \deg(p_i) < n$. Then, the number of common solutions of the polynomial equation system $p_i(x_1, \dots, x_n) = 0$ ($i = 1 \dots m$) is divisible by p . In particular, if there is a solution, there exists another.*

CHÉVALLEY MOD 2. Given polynomials $p_1, p_2, \dots, p_m$ in $\mathbb{F}_2[x_1, \dots, x_n]$ such that $\sum_{i=1}^m \deg(p_i) < n$. Also, given a root $(c_1, c_2, \dots, c_n) \in \mathbb{F}_2^n$ of the equation system $p_i(\mathbf{x}) = 0$ ($i = 1, \dots, m$). Chévalley's theorem states that the number of roots is even. Find another root.

The complexity of COMBINATORIAL NULLSTELLENSATZ MOD 2 and relationship with PPA and CHÉVALLEY MOD 2 is an open question by D. West ([4]). In Section 2, we investigate these problems.

In Proposition 2.10, another proof will be presented for showing that CHÉVALLEY MOD 2 is in PPA, reducing to COMBINATORIAL NULLSTELLENSATZ MOD 2 and in Theorem 2.9, we prove that COMBINATORIAL NULLSTELLENSATZ MOD 2 is in PPA. The proof of Theorem 2.9 is similar to Papadimitriou's proof about CHÉVALLEY MOD 2, but it requires more complex techniques.

The class PPA is defined by a problem in Graph Theory, so it is not so wondrous that there are also non-algebraic problems in PPA, for instance, ANOTHER HAMILTONIAN PATH belongs to PPA. We have already mentioned this problem in the previous section. In this problem, given a graph G and an Hamiltonian path in G , find another Hamiltonian path in G or in its complement $\overline{G}$. Lovász proved that the number of Hamiltonian path in a graph and its complement is even, if the number of nodes at least 4, hence, such Hamilton path exists.

In Section 4, we will study new graph problems and prove that they are in PPA. For instance, we proved that finding 2^d -divisible subgraphs in a graph with more than $n \cdot (2^d - 1) - 2^{d-1}$ edges is contained in PPA and it is not known to be solvable in polynomial time. The existence of such a subgraph is guaranteed by a theorem proved by Alon et al. in [5].

For more details about the class PPA, please see [1] or [6].

2 Complexity of finding roots over $\mathbb{F}_2$

How difficult is finding roots of a polynomial? Over $\mathbb{R}$ it usually seems an impossible task to find the exact roots, moreover, even determining the number of roots seems a difficult task. What about polynomials over the simplest algebraic field, $\mathbb{F}_2$? Obviously, it is not too complex: we should substitute only the element 0 and the element 1. We are done because the number of possible roots is tiny. However, what about a multivariable polynomial? Does it have any roots? How difficult is finding them? We study these questions over $\mathbb{F}_2$.

It is easy to verify that determining whether a general multivariable polynomial has any roots or not is an NP-complete problem. This can be proved via reducing the well-known NP-complete problem, 3-SAT. Also, finding a root if there exists one is an NP-hard problem.

Furthermore, there is an interesting theorem that gives a sufficient condition to existence of a root of a multivariable polynomial over an arbitrary field F . The Combinatorial Nullstellensatz, presented by Alon in [2], is a general algebraic theorem, but it has numerous applications in Combinatorial Number Theory, in Graph Theory and in Combinatorics, too.

The complexity of finding a root of polynomial whose existence is guaranteed by the Combinatorial Nullstellensatz is not a well-studied area, in general, it is not known to be solvable in polynomial time. An open question in [4] is about the complexity of Combinatorial Nullstellensatz and its relationship to the class PPA and Chévalley's theorem.

In this chapter, we present our answer to these questions: Combinatorial Nullstellensatz can be solved in polynomial time if the polynomial is given explicitly and it is in PPA, if it is given implicitly and Chévalley's theorem is polynomially reducible to it.

2.1 Complexity in general: NP-completeness and NP-hardness

Now, we mention two problems about polynomials over $\mathbb{F}_2$ in general. These results show that finding a root of a polynomial given implicitly, in general form is hard. So, investigating polynomials in special forms and complexity of finding roots of these polynomial such as in Combinatorial Nullstellensatz can be an interesting question.

If a polynomial is given implicitly, the number of terms of the polynomial might be exponentially large, even though, we can describe it in polynomial space.

ROOT OF A POLYNOMIAL MOD 2. Let p_{ij} ($i = 1 \dots k$, $j = 1 \dots m(i)$) be polynomials in $\mathbb{F}_2[x_1, \dots, x_n]$ and assume $k, m(i)$ and the size of every p_{ij} is polynomially bounded in n . Let $f = \sum_{i=1}^k \left(\prod_{j=1}^{m(i)} p_{ij} \right)$. Determine if this polynomial does have any root.

Proposition 2.1 ([11]) *3-SAT is polynomially reducible to ROOT OF A POLYNOMIAL MOD 2. Consequently, ROOT OF A POLYNOMIAL MOD 2 is NP-complete.*

It is easy to check that the following problem is NP-hard. Note that, similarly, finding a non-root if exists, is also an NP-hard problem.

EXISTING ROOT OF A POLYNOMIAL MOD 2. Let us given a polynomial $f = \sum_{i=1}^k \left(\prod_{j=1}^{m(i)} p_{ij} \right)$ as above. Suppose that there exists an $\mathbf{s} \in \mathbb{F}_2^n$ vector such that $f(\mathbf{s}) = 0$. Find such a vector.

Proposition 2.2 ROOT OF A POLYNOMIAL MOD 2 is polynomially reducible to EXISTING ROOT OF A POLYNOMIAL MOD 2. Consequently, it is NP-hard.

2.2 Complexity of Combinatorial Nullstellensatz

Theorem 2.3 (Combinatorial Nullstellensatz, Alon) Let F be an arbitrary field, and let $f = f(x_1, x_2, \dots, x_n)$ be a polynomial in $F[x_1, \dots, x_n]$. Suppose the degree of f is $\sum_{i=1}^n t_i$, where each t_i is a nonnegative integer, and suppose the coefficient of $\prod_{i=1}^n x_i^{t_i}$ is nonzero. Then, if $S_1, S_2, \dots, S_n$ are subsets of F with $|S_i| > t_i$, there exists an $(s_1, s_2, \dots, s_n) \in S_1 \times S_2 \times \dots \times S_n$ such that $f(s_1, s_2, \dots, s_n) \neq 0$.

As we have already mentioned, we study the complexity of Combinatorial Nullstellensatz over $\mathbb{F}_2$. Thereby we answer an open question in [4].

A vector $(s_1, s_2, \dots, s_n) \in S_1 \times S_2 \times \dots \times S_n$ is called appropriate if it satisfies the condition of Combinatorial Nullstellensatz: $f(s_1, s_2, \dots, s_n) \neq 0$.

Note that if $t_i = 0$ for some indices, then we could choose $|S_i| = 1$, so in an appropriate $(s_1, s_2, \dots, s_n)$ vector, we have to choose the single element of S_i to s_i and therefore, we could replace x_i by the only element of S_i in f . Hence, we may assume that $S_i = \mathbb{F}_2$ for every index i and the problem is finding an $(s_1, s_2, \dots, s_n) \in \mathbb{F}_2^n$ such that $f(s_1, s_2, \dots, s_n) \neq 0$.

Also, note that if the size of f (the number of terms) is polynomially bounded, there is a polynomial algorithm for finding an appropriate vector.

COMBINATORIAL NULLSTELLENSATZ MOD 2 WITH POLYNOMIALLY BOUNDED TERMS. Given a polynomial f in n variables with polynomially bounded number of terms. Suppose that the degree of f is n and the coefficient of $x_1 x_2 \dots x_n$ is nonzero. Find an appropriate vector $(s_1, s_2, \dots, s_n) \in \mathbb{F}_2^n$ such that $f(s_1, s_2, \dots, s_n) \neq 0$.

Proposition 2.4 COMBINATORIAL NULLSTELLENSATZ MOD 2 WITH POLYNOMIALLY BOUNDED TERMS can be solved in polynomial time.

Proof. First, we can replace the term $x_{i_1}^{t_{i_1}} x_{i_2}^{t_{i_2}} \dots x_{i_k}^{t_{i_k}}$ to $x_{i_1} x_{i_2} \dots x_{i_k}$, because these are equal due to the fact $0^t = 0, 1^t = 1$ in $\mathbb{F}_2$.

Substitute 0 and 1 to x_1 : let $g(x_2, \dots, x_n) = f(0, x_2, \dots, x_n)$ and $h(x_2, \dots, x_n) = f(1, x_2, \dots, x_n)$. If in f the coefficient of $x_2 x_3 \dots x_n$ is nonzero, then in g the coefficient of $x_2 x_3 \dots x_n$ will be also nonzero. If it is zero, in h the coefficient of $x_2 x_3 \dots x_n$ will be nonzero. Then, substitute 0 and 1 to x_2 and in one of them the coefficient of $x_3 x_4 \dots x_n$ will be nonzero.

Iterating this method, we obtain a constant nonzero polynomial, and this means that the substitution $(s_1, s_2, \dots, s_n) \in \mathbb{F}_2^n$ vector is appropriate.

Due to the polynomially bounded number of terms, these substitutions can be done in polynomial time, so we can find an appropriate vector in polynomial time. $\square$

Hence, we will deal with the case where the size of f might be exponentially large, however, f is given implicitly, e.g. as the sum of products of polynomial, with a polynomial size description. We must guarantee that the coefficient of $x_1 x_2 \dots x_n$ is nonzero, so we require that the maximum possible degree is at most n , and term $x_1 x_2 \dots x_n$ can be obtained only one way.

In [6], we investigated the complexity of Combinatorial Nullstellensatz in the following form, similarly to ROOT OF A POLYNOMIAL MOD 2 and EXISTING ROOT OF A POLYNOMIAL MOD 2. f might be exponentially large, but we can describe it in polynomial space.

SHORT COMBINATORIAL NULLSTELLENSATZ MOD 2. Given a polynomial f in n variables in a form $f = \sum_{i=1}^k \left(\prod_{j=1}^{m(i)} p_{ij} \right)$, where p_{ij} is a polynomial in $\mathbb{F}_2[x_1, \dots, x_n]$, $k, m(i)$ and the size of p_{ij} (the number of monomials) is polynomially bounded. Suppose that the degree of f is n and the coefficient of $x_1 x_2 \dots x_n$ is nonzero and they can be checked in polynomial time. Find an appropriate vector $(s_1, s_2, \dots, s_n) \in \mathbb{F}_2^n$ such that $f(s_1, s_2, \dots, s_n) \neq 0$.

In [6], we showed that SHORT COMBINATORIAL NULLSTELLENSATZ MOD 2 is in PPA and it does not seem to be trivially solvable in polynomial time. Our proof is similar to Papadimitriou's proof which shows that CHÉVALLEY MOD 2 is in PPA, but it requires additional ideas.

Now, let us consider polynomials in more complicated form.

Definition 2.5 Let $s \leq \log_2 n$. We say that the polynomial f in n variables is in general form, if f can be presented as $f = \prod_{j_1=1}^{m_1} \left(\sum_{i_1=1}^{k_1} \left(\dots \left(\prod_{j_s=1}^{k_s} \left(\sum_{i_s=1}^{m_s} p_{\underline{i}, \underline{j}} \right) \right) \dots \right) \right)$ where $p_{\underline{i}, \underline{j}}$ is a monomial where $\underline{i} = (i_1, \dots, i_s)$, $\underline{j} = (j_1, \dots, j_s)$ and $k_1, \dots, k_s, m_1, \dots, m_s$ are polynomially bounded.

Due to $s \leq \log_2 n$, the description of f is polynomially bounded in n . We will see that it can be assumed in investigation of Combinatorial Nullstellensatz.

Note that, usually, the integer k -s and m -s depend on the previous integers k -s and m -s, however, we can assume that these are constants because it can be written some nullpolynomial instead of missing parts in sums, and constant one polynomial in products.

Remark 2.6 A term of decomposed form of f can be represented with a tree T . Each node in level l has m_l child nodes and each node has a label number between 1 and k_l . It reflects that which subterm is chosen from the sum. Each leaf of the tree has an unique path $\underline{j}$ from the root, and let the labels of this path be $\underline{i}$. This leaf corresponds to $p_{\underline{i}, \underline{j}}$ and the term which is represented by this tree is the product of $p_{\underline{i}, \underline{j}}$ -s which are corresponded by the leaves. This tree has also polynomial size description.

Let us consider the following notations. Let $0 < r < s$. For $\underline{i}' = (i_1, \dots, i_r)$, $\underline{j}' = (j_1, \dots, j_r)$, let $f_{\underline{i}', \underline{j}'} = \prod_{j_{r+1}=1}^{m_{r+1}} \left(\sum_{i_{r+1}=1}^{k_{r+1}} \left(\dots \left(\prod_{j_s=1}^{k_s} \left(\sum_{i_s=1}^{m_s} p_{\underline{i}, \underline{j}} \right) \right) \dots \right) \right)$ be the inner part of the f .

Let $pre((i_1, \dots, i_s), r) = (i_1, \dots, i_r)$ and $pre((j_1, \dots, j_s), r) = (j_1, \dots, j_r)$.

Furthermore, for the term of decomposed form of f which is represented with tree T , let $Pre(T, r) = \{(pre(\underline{i}, r), pre(\underline{j}, r)) : \underline{i}, \underline{j} \text{ is a leaf in } T\}$, that is, the leaves of T restricted for the first r levels.

Definition 2.7 Let us define the maximum possible degree $\widehat{\deg}(f)$ in the following recursive manner. For a polynomial f with polynomially bounded number of terms, let the maximum possible degree of f be equal to its degree: $\widehat{\deg}(f) = \deg(f)$. The maximum possible degree of $f = \prod_{j=1}^m f_j$ is $\widehat{\deg}(f) = \sum_{j=1}^m \widehat{\deg}(f_j)$ and the maximum possible degree of $f = \sum_{i=1}^k f_i$ is $\widehat{\deg}(f) = \max_{i=1}^k \widehat{\deg}(f_i)$.

Note that $\widehat{\deg}(f) \geq \deg(f)$ for any polynomial f and the maximum possible degree of a polynomial can be computed from the general form in polynomial time in the description of f . (Note that, in s , it can be exponential time.)

Now let us define the *capped general form* and the *top term*. The degree of the top term will be strictly greater than the others and in this way, we can guarantee that $x_1 x_2 \dots x_n$ can be obtained only one way in a capped general form.

Definition 2.8 We say that f is in capped general form if it contains only a single monomial or it is in general form and satisfies the following conditions:

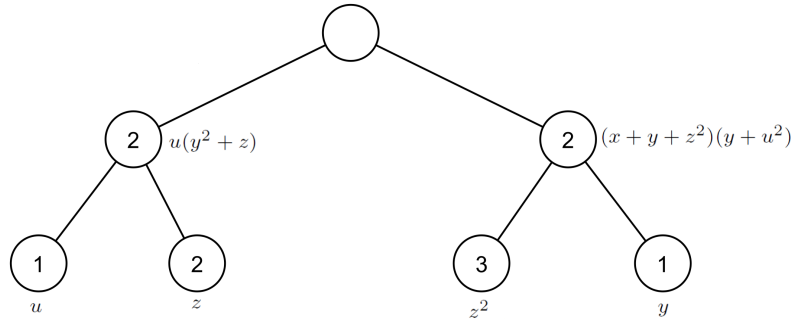
If $f = \sum_{i=1}^k f_i$, there exists an index $1 \leq i^* \leq k$ such that f_{i^*} is in capped general form and $\widehat{\deg}(f_{i^*}) > \widehat{\deg}(f_i)$ for any $i \neq i^*$. In this case the top term of f is the top term of f_{i^*} .

If $f = \prod_{j=1}^m f_j$, for any $1 \leq j \leq m$, the polynomial f_j is in capped general form. In this case the top term of f is the product of top terms of f_j .

Also note that if there is a polynomial f in general form with $\widehat{\deg}(f) \leq n$ and with $m_i > 1$, then $s \leq \log_2 n$. So the size of representation of f and size of trees representing terms are polynomially bounded in n .

COMBINATORIAL NULLSTELLENSATZ MOD 2. Given a polynomial f in n variables which is in capped general form, $\widehat{\deg}(f) = n$ and the top term of f is $x_1 x_2 \dots x_n$. Find an appropriate vector $(s_1, s_2, \dots, s_n) \in \mathbb{F}_2^n$ such that $f(s_1, s_2, \dots, s_n) \neq 0$.

Such a vector exists due to Combinatorial Nullstellensatz. Now, let us prove the main result of this section. We construct a graph and nodes correspond to the vectors and the terms. The nodes



A tree T corresponding to the term $u \cdot z \cdot z^2 \cdot y$ in polynomial $f(x, y, z, u) = ((x + y)(x^2 + z) + u(y^2 + z))(xy + (x + y + z^2)(y + u^2))$. The labels in the circles correspond to i_l 's and the polynomial next to circle is the corresponding inner part $f_{\underline{i}, \underline{j}}$ of f .

with odd degree correspond to the appropriate vectors and the term $x_1x_2 \dots x_n$. As we mentioned in the Introduction, we have to present a pairing function. It is easy at the terms, but it is more complicated at the vectors. The main idea here is the following: we pair up a tree corresponding to a term at last level as Papadimitriou presented in proof about CHÉVALLEY MOD 2, but if a term is lonely, we could pair up at above level. This concept of lonely terms is the basically new technique in our proof.

Theorem 2.9 *COMBINATORIAL NULLSTELLENSATZ MOD 2 is polynomially reducible to END OF THE LINE. Consequently, COMBINATORIAL NULLSTELLENSATZ MOD 2 is in PPA.*

Proof. We shall construct a graph Γ whose odd-degree nodes precisely correspond to appropriate vectors $\underline{s}$ such that $f(\underline{s}) \neq 0$ and furthermore to the standard leaf.

The graph is bipartite. The nodes of the left side are all the vectors in $\mathbb{F}_2^n$. The nodes of the other side are the terms of the polynomial f . Each term is represented in the above way, with a tree. There is an edge between vector $\underline{x}$ and term t if and only if $t(\underline{x}) = 1$.

It is easy to see that any vector $\underline{x}$ is appropriate if and only if its degree is odd.

All nodes in the side of terms except the top term of f have even degree in Γ . f is in topped general form, so the only term $x_1x_2 \dots x_n$ is the top of f , any other term has degree at most $n - 1$, hence it does not contain all variables. Let t be any such term, and let x_l be a variable not appearing in t . Then, if t is connected to $(s_1, s_2, \dots, 0, \dots, s_n)$, it is also connected to $(s_1, s_2, \dots, 1, \dots, s_n)$, so the degree of these nodes are even.

Therefore, odd-degree nodes are precisely $\underline{s}$ vectors such that $f(\underline{s}) \neq 0$ and the node corresponding to the term $x_1x_2 \dots x_n$. Let the node corresponding to the term $x_1x_2 \dots x_n$ be the standard leaf. (It is connected only to $(1, 1, \dots, 1)$.)

The nodes of this graph have exponentially large degrees, so we must exhibit a pairing function between the edges out of a node.

For a node corresponding to the term $t(\underline{x}) \neq x_1x_2 \dots x_n$, we pair up the vector $\underline{x}$ for which $t(\underline{x}) = 1$ to $(x_1, x_2, \dots, 1 - x_l, \dots, x_n)$ where x_l is such a variable which does not appear in t . (We choose the smallest such index l .)

For a vector $\underline{x}$ which is not an appropriate vector, we should pair up the terms such that $t(\underline{x}) = 1$. We will show by induction on s that they can be paired up. Suppose that the term t which we should pair up is represented by the tree T .

So, from now on, let us consider $\underline{x}$ and tree T fixed, we are going to present the mate of T .

In case $s = 1$, pick the smallest j_1 such that $\sum_{i_1=1}^{m_1} p_{i_1, j_1}$ is zero at $\underline{x}$. (There exists such j_1 because the product of such sums is zero.) So there is an even number of monomials in the sum such that $p_{i_1, j_1}(\underline{x}) = 1$, so one can pair these monomials by a pairing function. If the mate of i_1 is i'_1 , the mate of tree T is defined as T with replacing i_1 by i'_1 .

Now, let us consider the case $s > 1$. If $f_{\underline{i}, \underline{j}}$ is zero at $\underline{x}$ for some $(\underline{i}, \underline{j}) \in \text{Pre}(T, s - 1)$, there exists an index j_s such that $\sum_{i_s=1}^{m_s} p_{(\underline{i}, i_s), (\underline{j}, j_s)}(\underline{x})$ is zero. Pick the lexicographically first

such $(\underline{i}', \underline{j}') \in \text{Pre}(T, s-1)$ and j_s . There is an even number of monomials in the sum such that $p_{(\underline{i}', i_s), (\underline{j}', j_s)}(\underline{x}) = 1$, so one can pair these monomials by a pairing function. If the mate of i_s is i'_s , the mate of tree T is defined as T with replacing i_s by i'_s .

Otherwise, $f_{\underline{i}', \underline{j}'}$ is nonzero at $\underline{x}$ for every $(\underline{i}', \underline{j}') \in \text{Pre}(T, s-1)$, so $\sum_{i_s=1}^{m_s} p_{(\underline{i}', i_s), (\underline{j}', j_s)}(\underline{x}) = 1$ for every $(\underline{i}', \underline{j}') \in \text{Pre}(T, s-1)$ and every j_s . We can pair up the monomials such that $p_{(\underline{i}', i_s), (\underline{j}', j_s)}(\underline{x}) = 1$, one of them does not have a mate, denote its index by $\omega_{\underline{i}', (\underline{j}', j_s)}$. We call it a lonely index.

Look at the tree T . If there exists a label at level s such that it is not the lonely $\omega_{(\underline{i}', (\underline{j}', j_s))}$, pick the lexicographically first and if the mate of i_s is i'_s , the mate of tree T is defined as T with replacing i_s by i'_s .

If every label at level s is lonely, let us consider the subtree T' of T with only the first $s-1$ levels, considering $f_{\underline{i}', \underline{j}'}$ as a monomial for $(\underline{i}', \underline{j}') \in \text{Pre}(T, s-1)$. By induction, it can be paired up. If i_{s-1} is replaced by i'_{s-1} , we can replace each $\omega_{(i_1, \dots, i_{s-1}), (\underline{j}', j_s)}$ by $\omega_{(i_1, \dots, i'_{s-1}), (\underline{j}', j_s)}$.

Note that, if $\underline{x}$ is an appropriate vector, the pairing of the terms such that $t(\underline{x}) = 1$ can be obtained similarly, but there exists a totally lonely term t . Then, the mate of t is defined as t .

Observe that this gives a bijection and a correct pairing function. We presented a polynomial algorithm that computes the mate of an edge out of a node, so the proof is complete. $\square$

2.3 A consequence of Combinatorial Nullstellensatz

Let us recall the following search problem, which is known to be in PPA ([1]). Using that Chévalley's theorem can be proved via a reduction to the Combinatorial Nullstellensatz, we are give an alternative proof of the PPA membership.

CHÉVALLEY MOD 2. We are given polynomials $p_1, p_2, \dots, p_m$ in $\mathbb{F}_2[x_1, \dots, x_n]$ such that $\sum_{i=1}^m \deg(p_i) < n$. Also, we are given a root $(c_1, c_2, \dots, c_n) \in \mathbb{F}_2^n$ of the equation system $p_i(\underline{x}) = 0$ ($i = 1, \dots, m$). Chévalley's theorem states that the number of roots is even. Find another root.

Proposition 2.10 *CHÉVALLEY MOD 2 is polynomially reducible to COMBINATORIAL NULLSTELLENSATZ MOD 2.*

Proof. Given polynomials $p_1, p_2, \dots, p_m$ and the root $\underline{c} = (c_1, c_2, \dots, c_n) \in \mathbb{F}_2^n$, we shall construct the polynomial f as follows:

$$f = \prod_{i=1}^m (1 + p_i) + \prod_{i=1}^n (x_i - c_i + 1)$$

Obviously, f is in capped general form and the top term is $x_1 x_2 \dots x_n$, because $\sum_{i=1}^m \deg(p_i) < n$. So the conditions of COMBINATORIAL NULLSTELLENSATZ MOD 2 hold. However, the vectors $\underline{s} = (s_1, s_2, \dots, s_n) \in \mathbb{F}_2^n$ such that $f(\underline{s}) \neq 0$ are precisely the roots of the equation system different from $\underline{c}$. $\square$

3 The generalization of Olson's theorem

In Combinatorial Number Theory, there are numerous theorems which state that in a sufficiently large set of integers there is a subset which has fine properties.

For instance, the following are widely known.

Theorem 3.1 *Any multiset of m integers has a subset such that the sum of whose elements is divisible by m .*

Theorem 3.2 (Erdős-Ginzburg-Ziv, 1961) *Any multiset of $2k-1$ integers has a subset of size k such that the sum of whose elements is divisible by k .*

Olson's theorem is a similar theorem, it states about simultaneous divisibility. It easy to check that the above theorems for prime power follow from Olson's. For Theorem 3.1, let $n = 1$ and a_{1j} ($1 \leq j \leq m$) be the given numbers and for the Erdős-Ginzburg-Ziv Theorem, let $n = 2$ and a_{1j} ($1 \leq j \leq m = 2k - 1$) be the given numbers and a_{2j} ($1 \leq j \leq m = 2k - 1$) be equal to 1.

Theorem 3.3 (Olson,[3]) *Let n, m be positive integers and let p be an arbitrary prime. Let $d_1 \geq d_2 \geq \dots \geq d_n$ nonnegative integers and for $1 \leq i \leq n, 1 \leq j \leq m$ let a_{ij} be an integer. If $m > \sum_{i=1}^n (p^{d_i} - 1)$, then there is a subset $\emptyset \neq J \subseteq \{1, 2, \dots, m\}$ such that $\sum_{j \in J} a_{ij}$ is divisible by p^{d_i} for every $1 \leq i \leq n$.*

Olson's theorem is tight in the following sense: for arbitrary numbers d_i , if $m = \sum_{i=1}^n (p^{d_i} - 1)$ there exists integers a_{ij} such that the proper nontrivial subset does not exist.

There are two natural question related to Olson's theorem.

1. Do similar inequalities hold for arbitrary integer k instead of p ?
2. What about the inequalities if we do not require p^{d_i} -divisibility, only being in certain subsets modulo p^{d_i} .

The first question is wide open. Generally, for an arbitrary Abelian group G , let us define the Davenport constant of G : $D(G)$ is the smallest integer D so that every sequence of G with length at least D has a nontrivial subsequence which sums to zero. ([10]). Roughly speaking, Olson's theorem is the only significant theorem about determining $D(G)$. It states that $D(\mathbb{Z}_{p^{d_1}} \times \mathbb{Z}_{p^{d_2}} \times \dots \times \mathbb{Z}_{p^{d_n}}) = 1 + \sum_{i=1}^n (p^{d_i} - 1)$. The standing conjecture that $D(\mathbb{Z}_k^n) = 1 + n(k - 1)$ for arbitrary integer k , but it is still open.

There are some results in connection with second question.

Alon, Friedland and Kalai proved the following generalization of Olson's theorem in [5]. Note that, they presented an application in Graph Theory, which we will show in Section 4.

Theorem 3.4 (Alon, Friedland, Kalai, [5]) Let $\text{card}_p(Q)$ denote the number of distinct elements in Q modulo p . Let n, m be positive integers and let p be a prime. Let $d_1 \geq d_2 \geq \dots \geq d_n$ nonnegative integers and for $1 \leq i \leq n, 1 \leq j \leq m$ let a_{ij} be an integer. Then for arbitrary subset Q_i of $\{0, 1, \dots, p^{d_i} - 1\}$ ($1 \leq i \leq n$) containing zero, if $m > \sum_{i=1}^n (p^{d_i} - \text{card}_p(Q_i))$, then there is a subset $\emptyset \neq J \subseteq \{1, 2, \dots, m\}$ such that

$$\sum_{j \in J} a_{ij} \equiv q_i \pmod{p^{d_i}} \quad \text{for some } q_i \in Q_i \text{ for every } 1 \leq i \leq n$$

However, this theorem is usually not tight and it is an interesting question to determine the tight values. Formally, we investigate the following general problem.

Definition 3.5 Let us be given a prime p and $d_1 \geq d_2 \geq \dots \geq d_n$ nonnegative integers and $Q_1, Q_2, \dots, Q_n$ sets such that each of them contains zero and $Q_i \subseteq \mathbb{Z}_{p^{d_i}}$ for every $1 \leq i \leq n$. Denote $(d_1, d_2, \dots, d_n)$ by $\mathbf{d}$ and $(Q_1, Q_2, \dots, Q_n)$ by $\mathbf{Q}$.

Determine the minimum value $F(\mathbf{d}, \mathbf{Q})$ such that for every $m > F(\mathbf{d}, \mathbf{Q})$ and for arbitrary integers a_{ij} ($1 \leq i \leq n, 1 \leq j \leq m$) there exists a nontrivial subset $J \subseteq \{1, 2, \dots, m\}$ that fulfills the following condition:

$$\sum_{j \in J} a_{ij} \equiv q_i \pmod{p^{d_i}} \quad \text{for some } q_i \in Q_i \text{ for every } 1 \leq i \leq n. \quad (\clubsuit)$$

Conversely, we are looking for the maximum value $m = F(\mathbf{d}, \mathbf{Q})$ such that there exists such integers a_{ij} ($1 \leq i \leq n, 1 \leq j \leq m$) that any nontrivial subset $J \subseteq \{1, 2, \dots, m\}$ does not fulfill the condition $(\clubsuit)$.

Note that Q_i must contains zero. Otherwise, $F(\mathbf{d}, \mathbf{Q}) = \infty$, because if all a_{ij} 's is zero, there never exists nontrivial subsequence whose sum is in Q_i .

Let $\mathbf{0} = (\{0\}, \{0\}, \dots, \{0\})$. Then, Olson's theorem states that $F(\mathbf{d}, \mathbf{0}) = \sum_{i=1}^n (p^{d_i} - 1)$.

Similarly, we can formalize the generalization by Alon et al. Their theorem states that $F(\mathbf{d}, \mathbf{Q}) \leq \sum_{i=1}^n (p^{d_i} - \text{card}_p(Q_i))$.

In [6], we presented another generalisation of Olson's theorem applying Chévalley's theorem.

We determine the exact values of $F(\mathbf{d}, \mathbf{Q})$ assuming that the set Q_i have a specific structure: they are so-called R -zero sets (see Definition 3.6). This property asserts that written in base p , they must have digit 0 in a fixed set R of positions.

Definition 3.6 Let R be a subset of $\{0, 1, \dots, d-1\}$. Let us define the set $\Omega \subseteq \{0, 1, 2, \dots, p^d - 1\}$ by the following property: $c \in \Omega$ if and only if $c^{(r)} = 0$ for every $r \in R$ in the form of c in base p . We call Ω the R -zero set modulo p^d . Let $\kappa(R) = (p-1) \sum_{r \in R} p^r$.

Theorem 3.7 (R-zero Olson) Let n, m be positive integers and let p be a prime. Let $d_1 \geq d_2 \geq \dots \geq d_n$ nonnegative integers and for $1 \leq i \leq n, 1 \leq j \leq m$ let a_{ij} be an integer. Let R_i be an arbitrary subset of $\{0, 1, \dots, d_i-1\}$ and denote the R_i -zero set modulo p^{d_i} by Ω_i . If $m > \sum_{i=1}^n \kappa(R_i)$, then there is a subset $\emptyset \neq J \subseteq \{1, 2, \dots, m\}$ such that

$$\sum_{j \in J} a_{ij} \equiv \omega_i \pmod{p^{d_i}} \quad \text{for some } \omega_i \in \Omega_i \text{ for every } 1 \leq i \leq n.$$

That is, $F(\mathbf{d}, \mathbf{\Omega}) \leq \sum_{i=1}^n \kappa(R_i)$.

Moreover, these bounds are tight: $F(\mathbf{d}, \mathbf{\Omega}) = \sum_{i=1}^n \kappa(R_i)$, because if $m = \sum_{i=1}^n \kappa(R_i)$, there exists integers a_{ij} such that the proper nontrivial subset does not exist, e. g. -1 is chosen $\kappa(R_i)$ times at each row. It is easy to verify that in this case, proper subset does not exist.

We do not include the proof of this theorem, as it will follow as a special case of the more general Theorem 3.18. There we give a nontrivial upper bound on $F(\mathbf{d}, \mathbf{Q})$ for arbitrary sets Q_i in terms of structural properties of the sets Q_i .

Now, we mention some inequalities about $F(\mathbf{d}, \mathbf{Q})$. They are easy to verify.

Proposition 3.8 *If $\mathbf{d} = (d_1, d_2, \dots, d_n)$ and $\mathbf{Q} = (Q_1, Q_2, \dots, Q_n)$, then*

$$F(d_1, Q_1) + F(d_2, Q_2) + \dots + F(d_n, Q_n) \leq F(\mathbf{d}, \mathbf{Q})$$

Proposition 3.9 *For arbitrary $Q_i \subseteq \{0, 1, \dots, p^{d_i} - 1\}$, let $\overline{Q_i}$ be the largest $\overline{Q_i} = \Omega_i \subseteq Q_i$ such that Ω_i is the R_i -zero subset modulo p^{d_i} for some $R_i \subseteq \{0, 1, \dots, d_i - 1\}$. Then, $F(\mathbf{d}, \mathbf{Q}) \leq \sum_{i=1}^n \kappa(R_i)$*

Proposition 3.10 *Let $c_i Q_i = \{c_i q_i \pmod{p^{d_i}} : q_i \in Q_i\}$ with $(c_i, p) = 1$, so c_i is not divisible by p . Let $\mathbf{c} \cdot \mathbf{Q}$ be $(c_1 Q_1, c_2 Q_2, \dots, c_n Q_n)$. Then $F(\mathbf{d}, \mathbf{Q}) = F(\mathbf{d}, \mathbf{c} \cdot \mathbf{Q})$*

Furthermore, it is easy to see the following proposition.

Proposition 3.11 *Let Q be an arbitrary subset of $\mathbb{Z}_{p^d}$ containing zero. Then, at most $p^d - |Q|$ numbers can be chosen such that none of the sum of subsequences is in Q modulo p^d . Hence, for $n = 1$, $F(d, Q) \leq p^d - |Q|$.*

Proof. Suppose that none of the sum of subsequences of numbers $a_1, a_2, \dots, a_k$ is in Q modulo p^d . Let $s_j = \sum_{i=1}^j a_i$ be some subsequences. These integers s_j are pairwise incongruent modulo p^d , otherwise their subtraction is zero modulo p^d , but it would be in Q . So s_j 's are distinct integers which are outside of Q , so $k \leq p^d - |Q|$ and the proposition follows. $\square$

Our main goal is to improve and integrate these results to a full theory. In the next section we present the basic tools and after that, the new results about determining the value $F(\mathbf{d}, \mathbf{Q})$.

3.1 The basic tool of our results

At first, we investigate the residue of binomial coefficients modulo p . Lucas' theorem gives a nice relationship between the residue of $\binom{c}{b}$ and the form of c and b in base p .

Theorem 3.12 (Lucas, [8]) *Let p be a prime, let $c^{(d-1)} \dots c^{(1)} c^{(0)}$ be the form of c and $b^{(d-1)} \dots b^{(1)} b^{(0)}$ be the form of b in base p . Then*

$$\binom{c}{b} \equiv \binom{c_0}{b_0} \cdot \binom{c_1}{b_1} \cdots \binom{c_{d-1}}{b_{d-1}} \pmod{p}$$

Proof. The proof is based on the following equation.

$$\begin{aligned} \sum_{b=0}^c \binom{c}{b} x^b &= (1+x)^c = \prod_{r=0}^{d-1} \left((1+x)^{p^r} \right)^{c_r} \equiv \prod_{r=0}^{d-1} \left(1+x^{p^r} \right)^{c_r} \pmod{p} = \\ &= \prod_{r=0}^{d-1} \left(\sum_{s_r=0}^{c_r} \binom{c_r}{s_r} x^{s_r \cdot p^r} \right) = \sum_{b=0}^c \left(\sum_{(s_0, s_1, \dots, s_{d-1}): \sum s_r p^r = b} \prod_{r=0}^{d-1} \binom{c_r}{s_r} \right) x^b \end{aligned}$$

That is,

$$\binom{c}{b} \equiv \sum_{(s_0, s_1, \dots, s_{d-1}): \sum s_r p^r = b} \prod_{r=0}^{d-1} \binom{c_r}{s_r} \pmod{p}$$

However, $0 \leq s_r \leq c_r < p$, so the inner sum contains at most one term: $s_r = b_r$ for every r if $b_r \leq c_r$ for every r , and it is an empty sum if $b_r > c_r$ for any r . In the first case, the theorem immediately follows. In the latter case the sum over the empty set is zero, and for any $b_r > c_r$, $\binom{c_r}{b_r}$ is also zero. $\square$

We use the following corollary of Lucas' theorem.

Corollary 3.13 *Let $c^{(d-1)} \dots c^{(1)} c^{(0)}$ be the form of c in base p . Then,*

$$\binom{c}{p^r} \equiv c^{(r)} \pmod{p}$$

Proof. It is easy to see that if $b = p^r$ then $b^{(r)} = 1$ and $b^{(i)} = 0$, so

$$\binom{c}{p^r} \equiv \binom{c^{(0)}}{0} \cdot \binom{c^{(1)}}{0} \cdot \dots \cdot \binom{c^{(r)}}{1} \cdot \dots \cdot \binom{c^{(d-1)}}{0} \equiv c^{(r)} \pmod{p} \quad \square$$

The following definition and lemma show the consequence of Corollary 3.13 and Lucas's theorem: how can we encode integers by using polynomials over $\mathbb{F}_p$.

Definition 3.14 *Let $f = \sum_{i=1}^k p_i$ be a sum of monomials with coefficient one. So p_i is a monomial over $\mathbb{F}_p$ in the form $\prod_{j=1}^n x_j^{t_{ij}}$, with coefficient 1, where each t_{ij} is a nonnegative integer. (For instance, we write $x_i^2 x_j + x_i^2 x_j + x_i^2 x_j$ instead of $3x_i^2 x_j$.)*

Let $\Psi_p^{(r)}(f) \in \mathbb{F}_p[x_1, \dots, x_n]$ be the following polynomial:

$$\sum_{1 \leq i_1 < i_2 < \dots < i_{p^r} \leq k} p_{i_1} p_{i_2} \dots p_{i_{p^r}}$$

Lemma 3.15 *Let $f = \sum_{i=1}^k p_i$ be a sum of monomials with coefficient one. So p_i is a monomial over $\mathbb{F}_p$ in the form $\prod_{j=1}^n x_j^{t_{ij}}$, with coefficient 1, where each t_{ij} is a nonnegative integer. Let $\underline{s} = (s_1, s_2, \dots, s_n) \in \{0, 1\}^n$. Further, let $c = f(\underline{s}) = \sum_{i=1}^k \prod_{j=1}^n s_j^{t_{ij}}$ and let $c^{(d-1)} \dots c^{(1)} c^{(0)}$ be the form of c in base p . Then,*

$$c^{(r)} = \gamma \in \{0, 1, \dots, p-1\} \quad \text{if and only if} \quad \Psi_p^{(r)}(f)(\underline{s}) = \gamma \in \mathbb{F}_p \cong \{0, 1, \dots, p-1\}$$

Proof. It is easy to see that the number of terms in $\Psi_p^{(r)}(f)$ that are 1 at $\underline{s}$ is precisely $\binom{c}{p^r}$. Due to Corollary 3.13, $\binom{c}{p^r} \equiv c^{(r)} \pmod{p}$. Hence, $\Psi_p^{(r)}(f)(\underline{s})$ equals $c^{(r)}$, because both are congruent to $\binom{c}{p^r}$ modulo p and both are in $\mathbb{F}_p \cong \{0, 1, \dots, p-1\}$. This shows the lemma. $\square$

Note that $\deg(\Psi_p^{(r)}(f)) \leq p^r \cdot \deg(f)$.

3.2 The main generalization and the price of a forbidden set

We are ready to formulate our main result, the extension of Theorem 3.7 from R-zero sets to arbitrary set containing zero. Our approach is to describe a subset Q of $\mathbb{Z}_{p^d}$ by a set of polynomial equalities such that each element in $\mathbb{Z}_{p^d} \setminus Q$ satisfies at least one of them. We associate certain prices to these constraints and the bound on $F(\mathbf{d}, \mathbf{Q})$ will be given in term of such prices of the Q_i 's.

These polynomial equalities refer to the digits in base p . For example, if Q is the R-zero set modulo p^d , then we can describe it by forbidding integer $\gamma \in \{1, 2, \dots, p-1\}$ at the position $r \in R$.

Now, let us define the fine conditions and the price of a forbidden set.

Definition 3.16 Let α, β be arbitrary numbers in $\mathbb{Z}_{p^d}$ and let γ be arbitrary number in $\mathbb{F}_p$. Then, $\Psi_p^{(r)}(\alpha \cdot T + \beta) = \gamma$ is called a fine condition C for T , if $T = 0$ does not satisfy it. The price of this fine condition is $pr(C) = p^r$. The price of a set of fine conditions is equal to their sum: $pr(\{C_1, C_2, \dots, C_k\}) = \sum_{i=1}^k pr(C_i)$.

Definition 3.17 Suppose that B is a subset of $\mathbb{Z}_{p^d}$ which does not contain zero. We say that B is covered by $\mathcal{C}$ if every $b \in B$ satisfies at least one fine condition in $\mathcal{C} = \{C_1, C_2, \dots, C_k\}$. The price of the set B is $pr(B) = \min\{pr(\mathcal{C}) : B \text{ is covered by } \mathcal{C}\}$.

Theorem 3.18 Let n, m be positive integers and let p be a prime. Let $d_1 \geq d_2 \geq \dots \geq d_n$ nonnegative integers and for $1 \leq i \leq n, 1 \leq j \leq m$ let a_{ij} be an integer. Let Q_i be an arbitrary subset of $\mathbb{Z}_{p^{d_i}}$ which contains zero. If $m > \sum_{i=1}^n pr(\mathbb{Z}_{p^{d_i}} \setminus Q_i)$, then there is a subset $\emptyset \neq J \subseteq \{1, 2, \dots, m\}$ such that

$$\sum_{j \in J} a_{ij} \equiv q_i \pmod{p^{d_i}} \quad \text{for some } q_i \in Q_i \text{ for every } 1 \leq i \leq n.$$

That is, $F(\mathbf{d}, \mathbf{Q}) \leq \sum_{i=1}^n pr(\mathbb{Z}_{p^{d_i}} \setminus Q_i)$.

Proof. Let B_i be the complementary set of Q_i , that is, $\mathbb{Z}_{p^{d_i}} \setminus Q_i$. Due to the definitions, there exists a set of fine conditions $\mathcal{C}_i$ which covers B_i and its price $pr(B_i)$.

Then, let us consider the following polynomial in $\mathbb{F}_p[x_1, \dots, x_m]$:

$$f(\mathbf{x}) = \prod_{i=1}^n \prod_{(\Psi_p^{(r)}(\alpha \cdot s + \beta) = \gamma) \in \mathcal{C}_i} \left(\Psi_p^{(r)} \left(\alpha \cdot \sum_{j=1}^m a_{ij} x_j + \beta \right) - \gamma \right) - c \cdot \prod_{j=1}^m (1 - x_j)$$

where c is a nonzero constant which will be defined later.

The degree of the first part of the polynomial is $\sum_{i=1}^n \sum_{C \in \mathcal{C}_i} pr(C) = \sum_{i=1}^n pr(B_i) < m$, so the degree of the polynomial f is m , and the coefficient of $x_1 x_2 \dots x_m$ is $-c \cdot (-1)^m \neq 0$.

If $\mathbf{x} = \mathbf{0}$, then the first part is nonzero, because zero does not satisfy any fine linear condition. Let c be the value of the first part at $\mathbf{0}$. So, $f(\mathbf{0}) = c - c = 0$.

Let $t_j = 1, S_j = \{0, 1\}$. Then, the conditions of Combinatorial Nullstellensatz hold, so there exists an $\mathbf{s} \in \{0, 1\}^m$ such that $f(\mathbf{s}) \neq 0$.

For this $\underline{s} \in \{0, 1\}^m$, let T_i be $\sum_{j=1}^m a_{ij}s_j$. The vector $\underline{s}$ contains at least one 1 due to $f(\underline{0}) = 0$, so the second part of the polynomial f is zero, and hence, the first part must be nonzero at vector $\underline{s}$.

This means that T_i does not satisfy any condition in $\mathcal{C}_i$, so it is not B_i , hence it must be in Q_i .

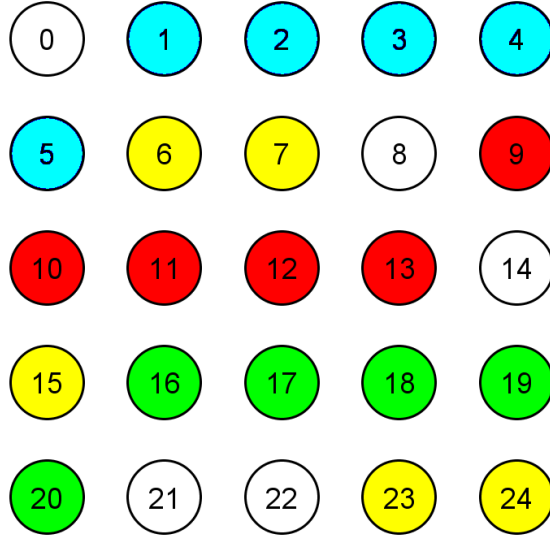
So, $J = \{j \in \{1, 2, \dots, m\} | s_j = 1\}$ fulfills the condition

$$\sum_{j \in J} a_{ij} \equiv q_i \pmod{p^{d_i}} \quad \text{for some } q_i \in Q_i \text{ for every } 1 \leq i \leq n. \quad \square$$

Remark 3.19 Formally, let $C_{r,\gamma}$ be the fine linear condition $\Psi_p^{(r)}(T) = \gamma$. Let $\mathcal{C}_i = \{C_{r,\gamma} : r \in R_i, \gamma \in \{1, 2, \dots, p-1\}\}$. Then $pr(\mathcal{C}_i) = \kappa(R_i)$, so $pr(\Omega_i) \leq \kappa(R_i)$ and it implies the R -zero Olson theorem.

Naturally, one can ask about the relationship between Theorem 3.18 and Theorem 3.4 of Alon et al. We investigate the question with given examples, and our result always gives stronger estimations than Alon et al., so it seems and we conjecture that for arbitrary sets Q_i our estimation is better, although, we have not proved it yet for arbitrary prime p , only in the case $p = 2$.

For instance, let $p^d = 25$, $Q = \{0, 8, 14, 21, 22\}$. Find fine conditions with price 20 such that they covers $\mathbb{Z}_{p^d} \setminus Q$. It can be solved, the following picture shows a solution, but it does not seem too trivial.



A solution for covering $\mathbb{Z}_{p^d} \setminus Q$ by fine conditions with price 20

4 An application of Olson's theorem: p^d -divisible subgraphs

In this section we discuss an application of Olson's theorem which was published in [5]. We present their results and then we apply our proof of Olson's theorem via Combinatorial Nullstellensatz to show that some natural problem is in PPA.

We also mention a more general question about subgraphs modulo prime power.

4.1 The existence of p^d -divisible subgraphs

A graph is called p^d -divisible, if every degree is divisible by p^d . In [5], Alon, Friedland and Kalai discussed the following question. Given is a prime p and a prime power p^d and an integer n . The problem is to determine the smallest value m such that for every graph on n vertices there is a nonempty p^d -divisible subgraph, if the graph has m edges. The nonempty p^d -divisible subgraph is a nonempty subset of edges such that for every vertex of the graph, the number of edges which are incident to it is divisible by p^d for a p prime. Conversely, there how many edges can a graph have such that it contains no nontrivial p^d -divisible subgraph. Denote the maximum number of these edges by $f(n, p^d)$.

The authors proved the following corollary of the original Olson's theorem and using it in the case $p = 2$, they derived the result on p^d -divisible subgraphs (Theorem 4.2).

Corollary 4.1 ([5]) *Let n, m be positive integers and let p be a prime. Let $d_1 \geq d_2 \geq \dots \geq d_n \geq 1$ positive integers and for $1 \leq i \leq n, 1 \leq j \leq m$ let a_{ij} be an integer such that $\sum_{i=1}^n a_{ij}$ is divisible by p for every j index. If $m > p^{d_n-1} - 1 + \sum_{i=1}^{n-1} (p^{d_i} - 1)$, then there is a subset $\emptyset \neq J \subseteq \{1, 2, \dots, m\}$ such that $\sum_{j \in J} a_{ij}$ is divisible by p^{d_i} for every $1 \leq i \leq n$.*

Theorem 4.2 (Alon, Friedland, Kalai, [5]) *For the maximum number of edges of a graph G on n vertices that contains no nontrivial p^d -divisible subgraph,*

$$f(n, p^d) \leq \begin{cases} (p^d - 1) \cdot n & \text{if } p \text{ is an odd prime.} \\ (2^d - 1) \cdot n - 2^{d-1} & \text{if } p = 2 \end{cases}$$

Proof. Denote the number of edges of G by m . Let $a_{ij} = 1$ if and only if the j^{th} edge is incident to the i^{th} vertex. (That is, $((a_{ij}))$ is the incidence matrix of G .)

For an odd prime p , suppose $m > (p^d - 1) \cdot n$. According to Olson's theorem with $d_1 = d_2 = \dots = d_n = d$, there is a nonempty J subset of edges such that $\sum_{j \in J} a_{ij}$ is divisible by p^d for every i , so there is nontrivial p^d -divisible subgraph.

For $p = 2$, suppose $m > (2^d - 1) \cdot n - 2^{d-1}$. According to the Corollary 4.1 with $d_1 = d_2 = \dots = d_n = d$, there is a nonempty J subset of edges such that $\sum_{j \in J} a_{ij}$ is divisible by p^d for every i , so there is nontrivial p^d -divisible subgraph. The conditions of corollary are true, because $\sum_{i=1}^n a_{ij} = 2$ for every index j . $\square$

Proposition 4.3 ([5]) *If $n \geq 3$, there exists graphs that has $(p^d - 1) \cdot n$ edges (if $p > 2$) or $(2^d - 1) \cdot n - 2^{d-1}$ edges (if $p = 2$), but does not contain nontrivial p^d -divisible subgraph. Consequently,*

$$f(n, p^d) = \begin{cases} (p^d - 1) \cdot n & \text{if } p \text{ is an odd prime.} \\ (2^d - 1) \cdot n - 2^{d-1} & \text{if } p = 2 \end{cases}$$

Proof. For $p > 2$ let G_0 be the graph which is obtained from a triangle by replacing each edges by $p^d - 1$ parallel edges. For $p = 2$, let G_0 be the similar graph, but replace one of edges by $2^{d-1} - 1$ edges instead of $2^d - 1$ edges. Then add to G_0 $n - 3$ new vertices and join each of them by $p^d - 1$ edges to vertices of G_0 .

It is easy to see that this graph does not contains p^d -divisible subgraph, it has n vertices and $(p^d - 1) \cdot n$ edges (if $p > 2$) or $(2^d - 1) \cdot n - 2^{d-1}$ edges (if $p = 2$). $\square$

4.2 Complexity aspects of Olson's theorem and 2^d -divisible subgraphs

OLSON MOD 2^d . Let us be given the integers n and m such that $m > n \cdot (2^d - 1)$ and given integers a_{ij} ($1 \leq i \leq n$, $1 \leq j \leq m$). Find a $\emptyset \neq J \subseteq \{1, 2, \dots, m\}$ such that $\sum_{j \in J} a_{ij}$ is divisible by 2^d for every i .

EVEN-SUM OLSON MOD 2^d . Given the integers n and m such that $m > n \cdot (2^d - 1) - 2^{d-1}$ and given integers a_{ij} ($1 \leq i \leq n$, $1 \leq j \leq m$) such that $\sum_{i=1}^n a_{ij}$ is even for every j index. Find an $\emptyset \neq J \subseteq \{1, 2, \dots, m\}$ such that $\sum_{j \in J} a_{ij}$ is divisible by 2^d for every i .

Note that, d is not the part of the input, it is fixed for the problem. Such subset exists according to Olson's theorem, which can be proved via Combinatorial Nullstellensatz. It implies the following proposition.

Proposition 4.4 *OLSON MOD 2^d and EVEN-SUM OLSON MOD 2^d are reducible to COMBINATORIAL NULLSTELLENSATZ MOD 2. Consequently, they are in PPA.*

According to the application of Olson's theorem, we get the following natural problem.

2^d -DIVISIBLE SUBGRAPH. Given a $G = (V, E)$ graph, where $|V| = n$ and $|E| = m$. Suppose that $m > n \cdot (2^d - 1) - 2^{d-1}$. Find a 2^d -divisible subgraph, that is, an $\emptyset \neq F \subseteq E$ such that for every $v \in V$, the number of incident edges of F is divisible by 2^d .

Note that, such a subgraph exists according to Theorem 4.2.

For $d = 1$, this is the well-known fact that in a graph, if the number of edges is at least as the number of vertices, there is an even subgraph, namely, there is a cycle in the graph. To find a cycle in such a graph, there is polynomial algorithm but for $d > 1$, we do not know whether it is in P.

Proposition 4.5 *2^d -DIVISIBLE SUBGRAPH is reducible to EVEN-SUM OLSON MOD 2^d . Consequently, it is in PPA.*

Proof. The proof of Theorem 4.2 shows the reduction. For the graph G , we shall construct the incidence matrix of G , where $m > n \cdot (2^d - 1) - 2^{d-1}$ and $\sum_{i=1}^n a_{ij} = 2$ is even for every j index. So the conditions of EVEN-SUM OLSON MOD 2^d are true and the right $\emptyset \neq J \subseteq \{1, 2, \dots, m\}$ subsets precisely correspond to the 2^d -divisible subgraphs. $\square$

Similarly, 2^d -DIVISIBLE SUBGRAPH IN HYPERGRAPHS can be defined and proved that they are in PPA.

4.3 A special case: even subgraph of a hypergraph

There is an interesting special case of finding 2^d -divisible subgraphs in hypergraphs. Let $G = (V, E)$ be a hypergraph on $|V| = n$ with $|E| = m > n$. Then, there is an even subgraph: there exists an $\emptyset \neq F \subseteq E$ such that for every $v \in V$, the number of incident hyperedges of F is even. More precisely, $|\{e \in F : v \in e\}|$ is even for every $v \in V$.

As we have already mentioned, it is well-known for graph and it holds with $m \leq n$ yet. It is worth checking that for hypergraphs the condition $m > n$ is tight: there exists hypergraph with $m = n$ such that does not contains nontrivial even subgraph.

For example, let $V = \{a, b, c, d\}$ and $E = \{\{a, b, c\}, \{a, b, d\}, \{a, c, d\}, \{b, c, d\}\}$. It is easy to check that this hypergraph does not contains nontrivial even subgraph.

Furthermore, because of the above observations, it is PPA, and in what follows, we show that it can be solved in polynomial time.

EVEN SUBGRAPH OF A HYPERGRAPH. Given a $G = (V, E)$ hypergraph, where $|V| = n$ and $|E| = m$. Suppose that $m > n$. Find an even subgraph: an $\emptyset \neq F \subseteq E$ such that for every $v \in V$, the number of incident hyperedges of F is even.

Proposition 4.6 EVEN SUBGRAPH OF A HYPERGRAPH is in P.

Proof. We present a polynomial time algorithm for this search problem.

It is easy to see that if $\underline{x} \in \mathbb{F}_2^m$, $\underline{0} = (0, 0, \dots, 0) \in \mathbb{F}_2^n$ and A is the incidence matrix of G , $F = \{e \in E : x_e = 1\}$ corresponds to an even subgraph if and only if, $A\underline{x} = \underline{0}$ over $\mathbb{F}_2$.

Due to the condition $m > n$, according to linear algebra, there exists a nontrivial solution $\underline{x}$ and it can be found in polynomial time using e.g. Gaussian elimination. $\square$

Note that Gaussian elimination runs in $O(n^3)$, though, in graphs an even subgraph can be found in $O(n)$. It is an interesting question whether even subgraph of a hypergraph can be found more efficiently.

5 Open questions and future work

Complexity of Combinatorial Nullstellensatz

It is an interesting question which problems are PPA-complete. At present, no natural PPA-complete problem is known, only a variant of Sperner lemma on non-orientable manifolds is known to be PPA-complete ([9]).

COMBINATORIAL NULLSTELLENSATZ MOD 2 is suspected to be PPA-complete, because it seems fairly general algebraic theorem and for example, CHÉVALLEY MOD 2 is reducible to it.

The generalization of Olson's theorem

We make the following conjecture about Olson's theorem. Note that, it is true for $n = 1$ due to Proposition 3.11.

Conjecture 5.1 $F(\mathbf{d}, \mathbf{Q}) \leq \sum_{i=1}^n |\mathbb{Z}_{p^{d_i}} \setminus Q_i| = \sum_{i=1}^n p^{d_i} - |Q_i|$.

Remark 5.2 *Let B be an arbitrary subset of $\mathbb{Z}_{p^d}$ which does not contain zero. In order to prove Conjecture 5.1, one may conjecture that the price of B is at most the size of B : $pr(B) \leq |B|$*

Unfortunately, it is not true. For instance, in case $p^d = 2^2 = 4$, $B = \{2\}$ can not be covered by any fine linear condition with price 1.

Also, we conjecture that Theorem 3.4 can be proved via Theorem 3.18, that is, $pr(\mathbb{Z}_{p^d} \setminus Q) \leq p^d - \text{card}_p(Q)$. Note that, it is true for $p = 2$.

Subgraphs modulo prime power

More general, the following quantity for graphs (hypergraphs) can be studied.

Definition 5.3 *Let Q is a subset of $\{0, 1, \dots, p^d - 1\}$ containing the 0. A subgraph is called Q -subgraph modulo p^d if every degree in it is in Q modulo p^d . Let $f(n, p^d, Q)$ be the maximum number of edges (hyperedges) if a graph (hypergraph) has n vertices and it has not a nontrivial Q -subgraph.*

Of course, $f(n, p^d, Q) \leq F(\mathbf{d}, \mathbf{Q})$ where $\mathbf{d} = (d, d, \dots, d)$ and $\mathbf{Q} = (Q, Q, \dots, Q)$.

Question 5.4 *When does $f(n, p^d, Q) = F(\mathbf{d}, \mathbf{Q})$ hold? For instance, is $f(n, p^d, \Omega) = n \cdot \kappa(R)$ true for every n, d and $R \subseteq \{0, 1, \dots, d - 1\}$?*

Complexity of p^d -divisible subgraphs

It is an interesting question whether there are more special case where a 2^d -divisible subgraph can be found in polynomial time. Note that, finding a p -divisible subgraph is also not known to be in P, for prime $p > 2$. The following question is still open and it is in our scope.

Question 5.5 *Can 4-DIVISIBLE SUBGRAPH be solvable in polynomial time?*

Acknowledgements

I am really grateful to my advisor, László Végh for introduction to this area of combinatorics and his help to answer my questions. His comments and valuable suggestions were extremely helpful for me. I would like to also express my thanks to Dömötör Pálvölgyi, with who discussed the complexity extensions of Combinatorial Nullstellensatz.

References

- [1] C. H. Papadimitriou: *On the Complexity of the Parity Argument and Other Inefficient Proofs of Existence*. Journal of Computer and System Sciences Vol. 48, Issue 3, June 1994, pp. 498-532
- [2] N. Alon: *Combinatorial Nullstellensatz*. Combinatorics, Probability and Computing, 8, 1999, pp. 7-29
- [3] J. E. Olson: *A Combinatorial Problem on Finite Abelian Groups, I*. Journal of Number Theory, Vol. 1, Issue 1, January 1969, pp. 8-10
- [4] D. B. West: *Research Experiences for Graduate Students in Combinatorics (Open Problem Collection)* <http://www.math.uiuc.edu/~west/regs/combnull.html>
- [5] N. Alon, S. Friedland, G. Kalai: *Regular subgraphs of almost regular graphs*. Journal of Combinatorial Theory, Series B
- [6] L. Varga: *Complexity classes based on parity argument*. BSc Thesis, June 2012, http://www.cs.elte.hu/blobs/diplomamunkak/bsc_alkmat/2012/varga_laszlo.pdf
- [7] L. Végh: *Játékelmélet (Game Theory)*. (Available only in Hungarian), 2010 <http://www.cs.elte.hu/~veghal/orak/jatekelm-2010dec10.pdf>
- [8] N. J. Fine: *Binomial coefficients modulo a prime*. The American Mathematical Monthly, Vol. 54, No. 10, Part 1, Dec., 1947, pp. 589-592
- [9] M. Grigni: *A Sperner Lemma Complete for PPA*. Information Processing Letters, 77(5-6) pp. 255-259, March 2001
- [10] *Open Problem Garden, Davenport's constant*. http://garden.irmacs.sfu.ca/?q=op/davenports_constant
- [11] M. R. Garey and D. S. Johnson: *Computers and Intractability: A guide to the theory of NP-completeness*, W. H. Freeman and co., New York, 1979.