Kernels for Problems Parameterized Above Tight Lower Bounds

Gregory Gutin

Department of Computer Science
Royal Holloway, University of London

WORKER 2009, 12th Sept 2009
Various Parameterizations
Strictly Above/Below Expectation Method
Linear Ordering Problem PALB
Lin-2 PALB
Betweenness PALB
Exact $r$-SAT PALB

Outline

1. Various Parameterizations
2. Strictly Above/Below Expectation Method
3. Linear Ordering Problem PALB
4. Lin-2 PALB
5. Betweenness PALB
6. Exact $r$-SAT PALB
Outline

1. Various Parameterizations

2. Strictly Above/Below Expectation Method

3. Linear Ordering Problem PALB

4. Lin-2 PALB

5. Betweenness PALB

6. Exact $r$-SAT PALB
Various Parameterizations

- **Standard** parameterizations: the parameter is the size of a set to optimize.
- Parameterizations using structural parameters such as treewidth, cliquewidth, the number of vertices to delete to make $G$ bipartite, etc.
- Parameterizations above and below tight bounds; initiated in [Mahajan and Raman, 1999].
Acyclic Subgraphs of Digraphs: Standard Parameterization

- Given a digraph $D = (V, A)$, find an acyclic subgraph $H = (V, B)$ of $D$ with the maximum number of arcs.
- Standard parameterization: $k = |B|$. Namely, does $D$ have an acyclic subgraph with at least $k$ arcs?
- But $|B| \geq |A|/2$. So if $k \leq |A|/2$ the answer is YES otherwise $|A| < 2k$, i.e., a linear kernel.
- $k$ is supposed to be small (for $2^{2k} k^{O(1)}$ to be tractable), but $k = |A|/2 - 1$ is not small.
Parameterization Above Tight Lower Bound: Does $D = (V, A)$ have an acyclic subgraph with at least $|A|/2 + k$ arcs? [ASPALB]

The bound is tight: For symmetric digraphs, $k = 0$: a digraph $D$ is symmetric if $xy \in A$ implies $yx \in A$.

Mahajan, Raman and Sikdar (2009): Is ASPALB fixed-parameter tractable?
**Outline**

1. Various Parameterizations
2. **Strictly Above/Below Expectation Method**
3. Linear Ordering Problem PALB
4. Lin-2 PALB
5. Betweenness PALB
6. Exact $r$-SAT PALB
Strictly Above/Below Expectation Method (SABEM)

- SABEM was recently introduced by Gutin, Kim, Szeider and Yeo.
- Apply some reduction rules to reduce the problem to its special case.
- Introduce a random variable $X$ such that the answer to the problem parameterized ALB is YES iff $\text{Prob}(X \geq k) > 0$.
- Use some probabilistic inequities to the reduced problem to obtain a problem kernel from $\text{Prob}(X \geq k) > 0$. 

Strictly Above/Below Expectation Method: Symmetric Case

- $X$ is symmetric, i.e., $X$ and $-X$ have the same distribution.
- If $X$ is discrete, then $X$ is symmetric iff
  \[ \text{Prob}(X = a) = \text{Prob}(X = -a) \text{ for each real } a. \]
- If $X$ is symmetric, then \( \text{Prob}(X \geq \sqrt{\mathbb{E}(X^2)}) > 0. \)
- If $k \leq \sqrt{\mathbb{E}(X^2)}$ then YES. Otherwise, \( \sqrt{\mathbb{E}(X^2)} < k \) and we may get a kernel.
Various Parameterizations
Strictly Above/Below Expectation Method
Linear Ordering Problem PALB
Lin-2 PALB
Betweenness PALB
Exact $r$-SAT PALB

Strictly Above/Below Expectation Method: Asymmetric Case

Lemma (Alon, Gutin, Krivelevich, 2004)

Let $X$ be a real random variable and suppose that its first, second and forth moments satisfy $\mathbb{E}(X) = 0$, $\mathbb{E}(X^2) = \sigma^2 > 0$ and $\mathbb{E}(X^4) \leq b\sigma^4$, respectively. Then $\text{Prob}(X > \frac{\sigma}{4\sqrt{b}}) \geq \frac{1}{4^{4/3}b}$.

Lemma (Bourgain, 1980)

Let $f = f(x_1, \ldots, x_n)$ be a polynomial of degree $r$ in $n$ variables $x_1, \ldots, x_n$ with domain $\{-1, 1\}$. Define a random variable $X$ by choosing a vector $(\varepsilon_1, \ldots, \varepsilon_n) \in \{-1, 1\}^n$ uniformly at random and setting $X = f(\varepsilon_1, \ldots, \varepsilon_n)$. Then $\mathbb{E}(X^4) \leq 2^{6r}(\mathbb{E}(X^2))^2$. 
Outline

1. Various Parameterizations
2. Strictly Above/Below Expectation Method
3. Linear Ordering Problem PALB
4. Lin-2 PALB
5. Betweenness PALB
6. Exact $r$-SAT PALB
**Reduction Rule for Linear Ordering Problem PALB**

- **Linear Ordering PALB**: each arc $ij$ has positive integral weight $w_{ij}$, does $D = (V, A)$ have an acyclic subgraph of weight at least $W/2 + k$, where $W = \sum_{ij \in A} w_{ij}$?

- Reduction rule: Assume $D$ has a directed 2-cycle $iji$;
  - if $w_{ij} = w_{ji}$ delete the cycle,
  - if $w_{ij} > w_{ji}$ delete the arc $ji$ and replace $w_{ij}$ by $w_{ij} - w_{ji}$,
  - if $w_{ji} > w_{ij}$ delete the arc $ij$ and replace $w_{ji}$ by $w_{ji} - w_{ij}$.

- Thus, we’ve reduced **Linear Ordering PALB** to the one on oriented graphs.
Let $D = (V, A)$ be an oriented graph, let $n = |V|$. Consider a random bijection: $\alpha : V \rightarrow \{1, \ldots, n\}$ and a random variable $X(\alpha) = \frac{1}{2} \sum_{ij \in A} \varepsilon_{ij}(\alpha)$, where $\varepsilon_{ij}(\alpha) = w_{ij}$ if $\alpha(i) < \alpha(j)$ and $\varepsilon_{ij}(\alpha) = -w_{ij}$, otherwise.

It is easy to see that 

$$X(\alpha) = \sum \{w_{ij} : ij \in A, \alpha(i) < \alpha(j)\} - W/2.$$ 

Thus, the answer is YES iff there is an $\alpha : V \rightarrow \{1, \ldots, n\}$ such that $X(\alpha) \geq k$. 
Lemma

\[ \mathbb{E}(X^2) \geq \frac{W^{(2)}}{12}, \text{ where } W^{(2)} = \sum_{ij \in A} w_{ij}^2. \]

Since \( X \) is symmetric, we have \( \text{Prob}(X \geq \sqrt{W^{(2)}/12}) > 0. \)
Hence, if \( \sqrt{W^{(2)}/12} \geq k \), there is an \( \alpha : V \to \{1, \ldots, n\} \) such that \( X(\alpha) \geq k \) and, thus, the answer is \text{YES}. Otherwise, \( |A| \leq W^{(2)} < 12 \cdot k^2 \). Thus, we have:

**Theorem (Gutin, Kim, Szeider, Yeo)**

**Linear Ordering PALB** has a quadratic kernel.
Outline

1. Various Parameterizations
2. Strictly Above/Below Expectation Method
3. Linear Ordering Problem PALB
4. Lin-2 PALB
5. Betweenness PALB
6. Exact $r$-SAT PALB
Lin-2 PALB

A system of \( m \) linear equations \( e_1, \ldots, e_m \) in \( n \) variables \( z_1, \ldots, z_n \) over \( \text{GF}(2) \), and each equation \( e_j \) has a positive integral weight \( w_j \). The problem \( \text{MAX LIN}-2 \) asks for an assignment of values to the variables that maximizes the total weight of the satisfied equations.
**Lin-2 PALB**

- A system of $m$ linear equations $e_1, \ldots, e_m$ in $n$ variables $z_1, \ldots, z_n$ over GF(2), and each equation $e_j$ has a positive integral weight $w_j$. The problem MAX LIN-2 asks for an assignment of values to the variables that maximizes the total weight of the satisfied equations.

- Let $W = w_1 + \cdots + w_m$. A greedy-type algorithm guarantees a solution of weight $\geq W/2$.

- **LIN-2 PALB**: Does the system have a solution of weight $\geq W/2 + k$?

- Mahajan, Raman and Sikdar (2009): What is the parameterized complexity of LIN-2 PALB?
Reduction Rules for Lin-2 PALB

The Same LHS Rule

- If two equations $e_j, e_p$ have the same LHS and RHS, replace them by one with the weight $w_j + w_p$.
- If two equations $e_j, e_p$ have the same LHS, but different RHS, replace them by one (or none) with the weight $|w_j - w_p|$.

Rank Rule

Let $A$ be the matrix of the coefficients of the variables in $S$, let $t = \text{rank} A$ and let columns $a^{i_1}, \ldots, a^{i_t}$ of $A$ be linearly independent. Then delete all variables not in $\{z_{i_1}, \ldots, z_{i_t}\}$ from the equations of $S$. 

Gregory Gutin

Problems Parameterized Above TLB
Let $l_j \subseteq \{1, 2, \ldots, n\}$ be the set of indices of the variables in $e_j$, and let $b_j \in \{0, 1\}$ be the RHS of $e_j$.

Define a random variable $X = \sum_{j=1}^{m} X_j$, where $X_j = (-1)^{b_j} w_j \prod_{i \in l_j} \varepsilon_i$ and all $\varepsilon_i$ are independent uniform random variables on $\{-1, 1\}$. 

**SABEM for Lin-2 PALB**

- Various Parameterizations
- Strictly Above/Below Expectation Method
- Linear Ordering Problem PALB
- Lin-2 PALB
- Betweenness PALB
- Exact $r$-SAT PALB
Let \( l_j \subseteq \{1, 2, \ldots, n\} \) be the set of indices of the variables in \( e_j \), and let \( b_j \in \{0, 1\} \) be the RHS of \( e_j \).

Define a random variable \( X = \sum_{j=1}^{m} X_j \), where \( X_j = (-1)^{b_j} w_j \prod_{i \in l_j} \varepsilon_i \) and all \( \varepsilon_i \) are independent uniform random variables on \( \{-1, 1\} \).

We set \( x_i = 0 \) if \( \varepsilon_i = 1 \) and \( x_i = 1 \), otherwise, for each \( i \). Observe that \( X_j = w_j \) if \( e_j \) is satisfied and \( X_j = -w_j \), otherwise.

The weight of the satisfied equations is at least \( W/2 + k \) if and only if \( X \geq 2k \).
Let $S$ be reduced under the Same LHS Rule.

We have $\mathbb{E}(X) = 0$ and $\mathbb{E}(X^2) = \sum_{j=1}^{m} w_j^2 \geq m$.

Gutin, Kim, Szeider, Yeo found ‘quadratic’ kernels in three cases.

In general, the parameterized complexity of Lin-2 PALB remains unknown.
Case 1: There exists a set $U$ of variables such that each equation of $S$ contains an odd number of variables from $U$.

- $X$ is symmetric.
- The same approach as above: YES or the number of equations $m = O(k^2)$.
- Use the Rank Rule and get $n \leq m = O(k^2)$. 
Case 2: The number of variables in each equation is bounded by $r = O(1)$.

- $X$ is not symmetric.
- By the inequality of Alon, Gutin, Krivelevich and Bourgain’s inequality: $YES$ or the number of equations $m = O(k^2)$.
- Use the Rank Rule and get $n \leq m = O(k^2)$. 
Case 3: No variable appears in more than $\rho = O(1)$ equations.

- $X$ is not symmetric.
- By the inequality of Alon, Gutin, Krivelevich and direct bound $\mathbb{E}(X^4) \leq 2\rho^2(\mathbb{E}(X^2))^2$: YES or the number of equations $m = O(k^2)$.
- Use the Rank Rule and get $n \leq m = O(k^2)$. 
Outline

1. Various Parameterizations
2. Strictly Above/Below Expectation Method
3. Linear Ordering Problem PALB
4. Lin-2 PALB
5. Betweenness PALB
6. Exact $r$-SAT PALB
Let $V = \{v_1, \ldots, v_n\}$ be a set of variables and let $C$ be a set of $m$ betweenness constraints of the form $(v_i, \{v_j, v_k\})$.

Given a bijection $\alpha : V \to \{1, \ldots, n\}$, we say that a constraint $(v_i, \{v_j, v_k\})$ is satisfied if either $\alpha(v_j) < \alpha(v_i) < \alpha(v_k)$ or $\alpha(v_k) < \alpha(v_i) < \alpha(v_j)$.

**Betweenness**: find a bijection $\alpha$ satisfying the max number of constraints in $C$.

**Tight Lower Bound**: $m/3$, the expectation number of satisfied constraints is $m/3$.

**Betweenness PALB**: Is there $\alpha$ that satisfies $\geq m/3 + \kappa$ constraints? ($\kappa$ is the parameter)
Various Parameterizations
Strictly Above/Below Expectation Method
Linear Ordering Problem PALB
Lin-2 PALB
Betweenness PALB
Exact $r$-SAT PALB

**Difficulties**

- Benny Chor’s question in Niedermeier’s book (2006): What is the parameterized complexity of Betweenness PALB?
- Difficult to estimate $\mathbb{E}(X^2)$, practically impossible to do $\mathbb{E}(X^4)$, but we cannot use Bourgain’s inequality as $X$ is not a polynomial of constant-bounded degree.
- What to do?
- Gutin, Kim, Mnich and Yeo: Betweenness PALB has a quadratic kernel.
We call a triple $A$, $B$, $C$ of distinct betweenness constraints **complete** if $\text{vars}(A) = \text{vars}(B) = \text{vars}(C)$.

Rule: if $C$ contains a complete triple of constraints, delete these constraints from $C$ and delete from $V$ any variable that appears only in the triple.

**Lemma**

Let $(V, C)$ be an instance of \textsc{Betweenness PALB} and let $(V', C')$ be obtained from $(V, C)$ by applying the reduction rule as long as possible. Then $(V, C)$ is a \textsc{Yes}-instance of \textsc{Betweenness PALB} if and only if so is $(V', C')$. 
Way Around Difficulties-1

- An instance \((V,C)\), where \(V\) is the set of variables and \(C = \{C_1, \ldots, C_m\}\) is the set of betweenness constraints.
- A random function \(\phi : V \rightarrow \{0, 1, 2, 3\}\).
- \(\phi\)-compatible bijections \(\alpha\): if \(\phi(v_i) < \phi(v_j)\) then \(\alpha(v_i) < \alpha(v_j)\).
Way Around Difficulties-2

- Let $\alpha$ be a random $\phi$-compatible bijection and $\nu_p(\alpha) = 1$ if $C_p$ is satisfied and 0, otherwise.
- Let the weights $w(C_p, \phi) = \mathbb{E}(\nu_p(\alpha)) - 1/3$ and $w(C, \phi) = \sum_{p=1}^m w(C_p, \phi)$.

Lemma

If $w(C, \phi) \geq \kappa$ then $(V, C)$ is a Yes-instance of Betweenness PALB.

Thus, to solve Betweenness PALB, it suffices to find $\phi$ for which $w(C, \phi) \geq \kappa$.
- We may forget about bijections $\alpha$!
Way Around Difficulties-3

- Let $X_p = w(C_p, \phi)$, and $X = \sum_{p=1}^{m} X_p$.
- If $\phi$ is a random function from $V$ to $\{0, 1, 2, 3\}$ then $X, X_1, \ldots, X_m$ are random variables.

| $|\{\phi(v_i), \phi(v_j), \phi(v_k)\}|$ | Relation | Value of $X_p$ | Prob. |
|---------------------------------|----------|---------------|-------|
| 1                              | $\phi(v_i) = \phi(v_j) = \phi(v_k)$ | 0              | 1/16  |
| 2                              | $\phi(v_i) \neq \phi(v_j) = \phi(v_k)$ | $-1/3$          | 3/16  |
| 2                              | $\phi(v_i) \in \{\phi(v_j), \phi(v_k)\}$ | $1/6$           | 6/16  |
| 3                              | $\phi(v_i)$ is between $\phi(v_j)$ and $\phi(v_k)$ | $2/3$           | 2/16  |
| 3                              | $\phi(v_i)$ is not between $\phi(v_j)$ and $\phi(v_k)$ | $-1/3$          | 4/16  |
Way Around Difficulties-4

**Lemma**

We have $E[X] = 0$.

**Lemma**

$X$ can be expressed as a polynomial of degree 6 in independent uniformly distributed random variables on $\{-1, 1\}$.

**Lemma**

For an irreducible instance $(V, C)$ we have $E[X^2] \geq \frac{11}{768} m$.

Use of PC.
Various Parameterizations
Strictly Above/Below Expectation Method
Linear Ordering Problem PALB
Lin-2 PALB
Betweenness PALB
Exact $r$-SAT PALB

Outline

1. Various Parameterizations
2. Strictly Above/Below Expectation Method
3. Linear Ordering Problem PALB
4. Lin-2 PALB
5. Betweenness PALB
6. Exact $r$-SAT PALB
**Exact $r$-SAT**

- **Exact $r$-SAT**: A CNF formula $F$ which contains $m$ clauses each with $r$ literals. Is there a truth assignment satisfying all $m$ clauses of $F$?

- **Max Exact $r$-SAT**: Find a truth assignment satisfying the max number of clauses.

- **Tight Lower Bound**: $(2^r - 1)m/2^r$. 

**Exact \( r \)-SAT PALB**

- **Exact \( r \)-SAT PALB**: Is there a truth assignment satisfying
  \[ \geq ((2^r - 1)m + k)/2^r \] clauses?

- Mahajan, Raman and Sikdar (2009): The parameterized complexity of **Exact \( r \)-SAT PALB** for each fixed \( r \)?

- Gutin, Kim, Szider and Yeo (SODA 2010): **Exact 2-SAT PALB** has a kernel with \( O(k^2) \) variables.
A pair of distinct clauses $Y$ and $Z$ has a conflict if there is a literal $p \in Y$ such that $\overline{p} \in Z$.

An $r$-CNF formula $F$ is semicomplete if the number of clauses is $m = 2^r$ and every pair of distinct clauses of $F$ has a conflict.

Lemma: Every truth assignment to a semicomplete $r$-CNF formula satisfies exactly $2^r - 1$ clauses.

Reduction Rule: Delete all semicomplete formulas. This will not change the answer to Exact $r$-SAT PALB.
**Exact 2-SAT PALB**

- **Exact 2-SAT PALB**: Is there a truth assignment satisfying \( \geq \frac{3}{4} (m + k) \) clauses?

- A variable \( x \) in \( F \) is **insignificant** if for each literal \( y \) we have \( xy \in F \) iff \( \bar{x}y \in F \). We may set \( x = 1 \) for each insignificant variable.

- A variable \( x \) in \( F \) is **significant** if it is not insignificant.

**Theorem (Significant Variables Theorem)**

Let \( F \) be a 2-CNF formula without semicomplete formulas. If \( F \) has more than \( k^2 \) significant variables, then the answer to Exact 2-SAT PALB is **Yes**.

This implies a kernel with \( O(k^2) \) variables.
Key Lemma

- $c(\ell)$ is the number of clauses containing literal $\ell$
- $\epsilon(xy) = 1$ if $xy \in F$ and $\epsilon(xy) = 0$, otherwise.

Lemma

For each subset $R = \{x_1, \ldots, x_q\} \subseteq \text{vars}(F)$ the maximum number of satisfiable clauses $\text{sat}(F) \geq (3m + k_R)/4$, where

$$k_R = \sum_{1 \leq i \leq q} (c(x_i) - c(\overline{x_i})) + \sum_{1 \leq i < j \leq q} \epsilon(x_i \overline{x_j}) + \epsilon(\overline{x_i} x_j) - \epsilon(x_i x_j) - \epsilon(\overline{x_i} \overline{x_j}).$$

Proof: Set $x_i = 1$ for all $x_i \in R$ and $\text{Prob}(x_i = 1) = 1/2$ for all $x_i \not\in R$. Show that $\mathbb{E}[\text{sat}(F)] = (3m + k_R)/4$. 

Gregory Gutin  Problems Parameterized Above TLB
Auxiliary Graph

- **Auxiliary graph** \( G = (V, E) \), where \( V = \text{vars}(F) \) and \( xy \in E \) iff there exists a clause \( C \in F \) with \( \text{vars}(C) = \{x, y\} \).
- \( w(x) = c(x) - c(\bar{x}) \)
- \( w(xy) = \epsilon(x_i\bar{x}_j) + \epsilon(\bar{x}_i x_j) - \epsilon(x_i x_j) - \epsilon(\bar{x}_i \bar{x}_j) \).
Proof of Significant Variables Theorem

- $X \subseteq \text{vars}(F)$; $F_X$ is obtained from $F$ by replacing each $x \in X$ by $\overline{x}$.
- We have $\text{sat}(F) = \text{sat}(F_X)$.
- $G_X$ is obtained from $G$ by $X$-switching: reversing the signs of $w(x)$, $w(xy)$ for each $x \in X$.
- If there exist $X \subseteq V$ and subgraph $Q$ such that its total weight in $G_X$ is $\geq k$, then $\text{sat}(F) \geq 3(m + k)/4$.
- If $F$ has more than $k^2$ significant variables, then there exist such $X$ and $Q$.
- We use graph matching theory: the Tutte-Berge formula for maximum matching.
Thank you!

- Questions?
- Comments?