# Generalization of the Khrapchenko method for *k*-ary bases

Sergeev I. S.
MVK seminar, 2021
SCCS seminar, 2023

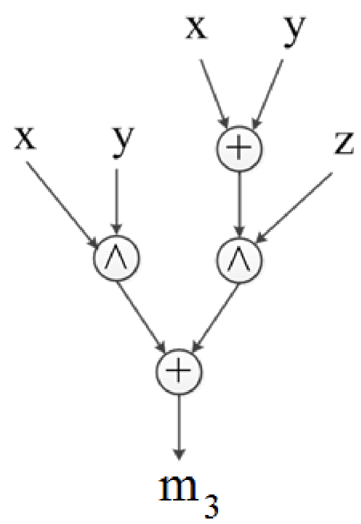
Valeriy Mikhailovich KHRAPCHENKO (1936-2019)

Department of Theoretical Cybernetics,
Keldysh Institute of Applied Math., Acad. Sci. USSR, since 1966

~20 reports at MVK seminar in Moscow Univ., 1963 – 2000

The method of formula complexity lower bounds – 1971

## *FORMULAE*




$\Phi_{\mathcal{B}}(f)$ — formula complexity of the function over a basis $\mathcal{B}$

*Sensitivity* of a boolean function $f$:

$$s(f) = \max_{N \subset f^{-1}(0),\, P \subset f^{-1}(1)} \frac{|R(N,P)|^2}{|N| \cdot |P|},$$

where $R(N,P)$ is the set of pairs of neighboring vectors from $N$ and $P$.

$\boxed{\Phi_{\mathcal{B}_0}(f) \geq s(f)}$ (V. M. Khrapchenko, 1971)

Corollary: $\Phi_{\mathcal{B}_0}(l_n) \geq n^2$.

In arbitrary basis: $\Phi_{\mathcal{B}}(l_n) \succeq n^{\Gamma_{\mathcal{B}}}$, where $\Gamma_{\mathcal{B}}$ is the shrinkage exponent of a basis $\mathcal{B}$.

$\mathcal{U}_k$ — maximal basis of $k$-ary functions, where the complexity of the linear function $l_n$ is linear.

$\mathcal{U}_k \sim \{\text{all monotone } k\text{-ary functions}, \overline{x}, 1\}$.

$\mathcal{U}_2 \sim \mathcal{B}_0$, $\mathcal{U}_3 \sim \{m_3(x,y,z), \overline{x}, 1\}$.

$\Gamma_{\mathcal{U}_k} \geq 1 + \frac{1}{3k-4}$ (N. A. Peryazev, 1995; D. Yu. Cherukhin, 2000)

$\Gamma_{\mathcal{U}_3} \geq \frac{4}{3}$ (H. Chockler, U. Zwick, 2001)

$$\Phi_{\mathcal{U}_k}(l_n) \succeq n^{1+\frac{1}{3k-4}}, \qquad n^{4/3} \preceq \Phi_{\mathcal{U}_3}(l_n) \preceq n^{1.74}.$$

*Formal complexity measures of functions* (over a basis $\mathcal{B}$)

1) $\mu(0) = \mu(1) = 0$.

2) $\mu(x) = 1$.

3) For any $g \in \mathcal{B}$: $\mu(g(f_1, \ldots, f_s)) \leq \mu(f_1) + \ldots + \mu(f_s)$.

Property: $\mu(f) \leq \Phi_{\mathcal{B}}(f)$.

$s(f)$ — formal complexity measure over $\mathcal{U}_2$ (M. S. Paterson; A. E. Andreev)

*Hypothesis*: $s^{\chi}(f)$ — formal complexity measure over $\mathcal{U}_k$.

*Khrapchenko exponent* $\chi_{\mathcal{B}}$ — maximal $\chi$, such that for any function $f$ (computable in the basis $\mathcal{B}$),

$$\Phi_{\mathcal{B}}(f) \geq s^{\chi}(f).$$

As follows from the definition, $\Phi_{\mathcal{B}}(l_n) \succeq n^{2\chi_{\mathcal{B}}}$.

$\chi_{\mathcal{U}_2} = 1$.

$\chi_{\mathcal{B}} \geq \frac{1}{2}$. Trivially, for a basis $\mathcal{B}$ of $k$-ary functions: if $\mathcal{B} \not\subset \mathcal{U}_k$, then $\chi_{\mathcal{B}} = \frac{1}{2}$.

*Special complexity measures of bipartite graphs*

$$G = (A, B, E), \qquad s(G) = \max_{X \subset A,\, Y \subset B} \frac{|E \cap (X \times Y)|^2}{|X| \cdot |Y|}$$

For a boolean function: $G_f = (N, P, R(N, P))$, $N = f^{-1}(0)$, $P = f^{-1}(1)$.

By definition, $s(f) = s(G_f)$.

*Coverings of graphs*

Set of graphs $\{G_i = (A_i, B_i, E_i)\}$ — *covering* of a graph $G = (A, B, E)$, if:

1) $A_i \subset A, \quad B_i \subset B, \quad E_i = E \cap (A_i \times B_i)$;

2) $E = \bigcup E_i$.

Covering is *monotone*, if additionally:

3) For any set of indices $I$, either $A \setminus \bigcup_{i \in I} A_i = \emptyset$, or $B \setminus \bigcup_{i \notin I} B_i = \emptyset$.

The reasoning for 3): let $f = \varphi(f_1, \ldots, f_k)$, where $\varphi$ is monotone.

$A = f^{-1}(0), \quad B = f^{-1}(1), \qquad A_i = f_i^{-1}(0) \cap A, \quad B_i = f_i^{-1}(1) \cap B.$

Assume that for some $I$, there exist vectors

$$\alpha \in A \setminus \bigcup_{i \in I} A_i, \qquad \beta \in B \setminus \bigcup_{i \notin I} B_i, \qquad \text{i.e.}$$

$$f(\alpha) = 0, \quad f_i(\alpha) = 1 \;\; \forall i \in I, \qquad f(\beta) = 1, \quad f_i(\beta) = 0 \;\; \forall i \notin I.$$

Then $f_i(\alpha) \geq f_i(\beta)$ for any $i$ — a contradiction with $f(\alpha) < f(\beta)$.

For any $k \geq 2$, the *complexity exponent* $\chi_k$ of bipartite graphs (correspondingly, the *monotone complexity exponent* $\chi_k^*$) is defined as the maximal number $\chi$, such that for any bipartite graph $G$, and for any its covering (monotone covering) $G_1, \ldots, G_k$,

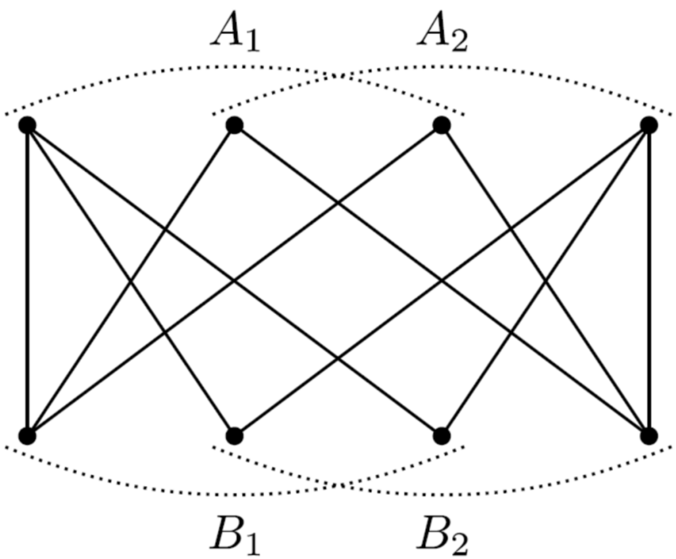
$$s^\chi(G_1) + \ldots + s^\chi(G_k) \geq s^\chi(G).$$

**Claim 1.** $\chi_k \leq \chi_k^* \leq \chi_{U_k}$.

**Claim 2.** $\chi_2^* = 1$.

**Claim 3.** $\chi_2 < 0.95$.

$s(G) = 25/4$
$s(G_1) = s(G_2) = 3$



**Theorem 1 (upper bound).** *For any $k \geq 2$,*

$$\chi_{U_k} \leq \log_{\lceil k/2 \rceil (\lfloor k/2 \rfloor + 1)} k < \frac{1}{2} + \frac{1}{2\log_2(k/2)}.$$

*Proof.* Consider the majority function $m_k$ of $k$ variables.

Choose as $N$ and $P$ the neighboring layers of a boolean cube, where the function changes its value, that is, the sets of vectors of the weight $p-1$, and of the weight $p$, where $p = \lceil k/2 \rceil$.

Obviously, $\Phi_{U_k}(m_k) = k$. On the other hand,

$$s(m_k) \geq \frac{|R(N,P)|}{|P|} \cdot \frac{|R(N,P)|}{|N|} = p(k-p+1) = \lceil k/2 \rceil (\lfloor k/2 \rfloor + 1) = (L_{U_k}(m_k))^{1/\chi},$$

where $\chi = \log_{\lceil k/2 \rceil (\lfloor k/2 \rfloor + 1)} k$. Therefore, for larger $\chi$, the condition $L_{U_k}(m_k) \geq s^{\chi}(m_k)$ does not hold. □

For instance, $\chi_{U_3} \leq \log_4 3 \approx 0.792$.

**Theorem 2 (lower bound).** *For any $k \geq 2$,*

$$\chi_k \geq \frac{1}{2} + \frac{1}{10 \ln k}.$$

*Proof.*

Let $\{G_i = (A_i, B_i, E_i)\}$, $i = 1, \ldots, k$, be a covering of a graph $G = (A, B, E)$.

We have to check that for $\chi = 1/2 + 1/(10 \ln k)$,

$$s^\chi(G_1) + \ldots + s^\chi(G_k) \geq s^\chi(G).$$

W.l.o.g. assume $s(G) = \frac{|E|^2}{|A| \cdot |B|}$. (otherwise consider a subgraph)

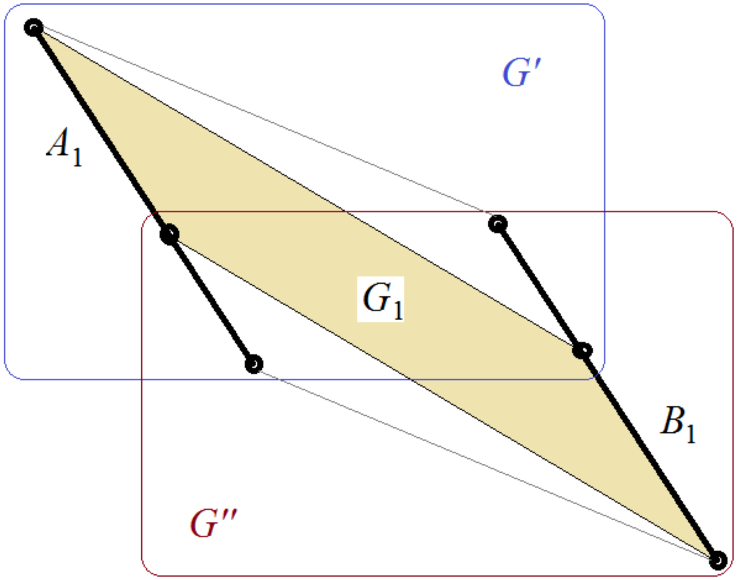
Denote $a_i = |A_i|/|A|$, $b_i = |B_i|/|B|$, $e_i = |E_i|/|E|$.

Induction on $k$. In the degenerate case $k = 1$, set $\chi_1 = 1$ for convenience. Further, we prove a transition from $k-1$ to $k \geq 2$. Consider the two cases:

**I.** One of the covering graphs contains a huge portion of both parts of $G$.

**II.** There is no such graph in the covering.

Case **I.** $\boxed{a_1 + b_1 \geq 1.63}$




$$G' = (A, B \setminus B_1, E') \approx G \setminus B_1,$$

$$G'' = (A \setminus A_1, B, E'') \approx G \setminus A_1,$$

where $E' = E \setminus (A \times B_1)$ and $E'' = E \setminus (A_1 \times B)$.

Graph $G'$ has a covering $G'_2, \ldots, G'_k$, where $G'_i = (A_i, B_i \setminus B_1, E_i \cap E') \approx G_i \setminus B_1$.

Graph $G''$ has a covering $G''_2, \ldots, G''_k$, where $G''_i = (A_i \setminus A_1, B_i, E_i \cap E'') \approx G_i \setminus A_1$.

Informally, $G' \cup G'' = G \setminus E_1$.

Consequently, graphs $G'$ and $G''$ together contain at least $(1 - e_1)|E|$ edges.

$$\max\{s(G'),\, s(G'')\} = \max\{\sqrt{s(G')},\, \sqrt{s(G'')}\}^2 \geq$$

$$\max\left\{\frac{|E'|}{\sqrt{|A|\cdot|B\setminus B_1|}},\, \frac{|E''|}{\sqrt{|A\setminus A_1|\cdot|B|}}\right\}^2 \geq \frac{(1-e_1)^2|E|^2}{(\sqrt{1-a_1}+\sqrt{1-b_1}\,)^2|A|\cdot|B|}. \quad (1)$$

$$\text{since} \qquad \max\left\{\frac{x_1}{y_1},\, \frac{x_2}{y_2}\right\} \geq \frac{x_1+x_2}{y_1+y_2} \qquad \text{for} \quad y_1, y_2 > 0.$$

For $\chi = 1/2 + 1/(10\ln(k-1))$ (for $\chi = 1$ in the case $k = 2$),

$$s^\chi(G_2) + \ldots + s^\chi(G_k) \geq \max\{s^\chi(G_2') + \ldots + s^\chi(G_k'),\; s^\chi(G_2'') + \ldots + s^\chi(G_k'')\}$$

$$\geq \max\{s^\chi(G'),\, s^\chi(G'')\} \geq \left(\frac{(1-e_1)^2|E|^2}{(\sqrt{1-a_1}+\sqrt{1-b_1}\,)^2|A|\cdot|B|}\right)^\chi$$

by the inductive assumption and by (1).

The obtained inequality

$$s^\chi(G_2) + \ldots + s^\chi(G_k) \ge \left(\frac{(1-e_1)^2|E|^2}{(\sqrt{1-a_1}+\sqrt{1-b_1}\,)^2|A|\cdot|B|}\right)^\chi$$

holds also for any smaller $\chi$ due to the Minkowski inequality

$$\sum x_i^p \ge \left(\sum x_i\right)^p, \quad p \le 1, \quad x_i \ge 0.$$

Hence, for $\chi = 1/2 + 1/(10\ln k)$,

$$\begin{aligned} s^\chi(G_1) + \ldots + s^\chi(G_k) &\ge \left(\left(\frac{e_1^2}{a_1 b_1}\right)^\chi + \left(\frac{(1-e_1)^2}{(\sqrt{1-a_1}+\sqrt{1-b_1})^2}\right)^\chi\right)\left(\frac{|E|^2}{|A|\cdot|B|}\right)^\chi \\ &\ge \left[\left(\frac{e_1}{t}\right)^{2\chi} + \left(\frac{1-e_1}{2\sqrt{1-t}}\right)^{2\chi}\right]\left(\frac{|E|^2}{|A|\cdot|B|}\right)^\chi, \end{aligned}$$

where $t = (a_1 + b_1)/2$.

Further, we apply the Hölder's inequality

$$\sum x_i y_i \leq \left(\sum x_i^p\right)^{1/p} \left(\sum y_i^q\right)^{1/q}, \quad \frac{1}{p} + \frac{1}{q} = 1, \quad x_i, y_i \geq 0,$$

with parameters

$$x_1 = \frac{e_1}{t}, \quad x_2 = \frac{1 - e_1}{2\sqrt{1-t}}, \quad y_1 = t, \quad y_2 = 2\sqrt{1-t}, \quad p = 2\chi, \quad q = \frac{2\chi}{2\chi - 1}.$$

So we obtain

$$1 = e_1 + (1 - e_1) = x_1 y_1 + x_2 y_2 \leq \left(\left(\frac{e_1}{t}\right)^p + \left(\frac{1 - e_1}{2\sqrt{1-t}}\right)^p\right)^{1/p} \left(t^q + \left(2\sqrt{1-t}\right)^q\right)^{1/q}.$$

Therefore,

$$\left[\left(\frac{e_1}{t}\right)^{2\chi} + \left(\frac{1 - e_1}{2\sqrt{1-t}}\right)^{2\chi}\right] \geq \frac{1}{\left(t^{\frac{2\chi}{2\chi-1}} + (2\sqrt{1-t}\,)^{\frac{2\chi}{2\chi-1}}\right)^{2\chi-1}} = \left(t^{\frac{2\chi}{2\chi-1}} + (4(1-t))^{\frac{\chi}{2\chi-1}}\right)^{1-2\chi}.$$

Easy to check, that for $\chi < 1$ and $t \in [0.815, 1]$,

$$g(t) = t^{\frac{2\chi}{2\chi-1}} + (4(1-t))^{\frac{\chi}{2\chi-1}} \leq 1.$$

Since $t = (a_1 + b_1)/2 \geq 0.815$ by the case requirement, the chain of inequalities finishes as

$$s^\chi(G_1) + \ldots + s^\chi(G_k) \geq \left[ \left(\frac{e_1}{t}\right)^{2\chi} + \left(\frac{1-e_1}{2\sqrt{1-t}}\right)^{2\chi} \right] \left(\frac{|E|^2}{|A| \cdot |B|}\right)^\chi$$
$$\geq (g(t))^{1-2\chi} \left(\frac{|E|^2}{|A| \cdot |B|}\right)^\chi \geq \left(\frac{|E|^2}{|A| \cdot |B|}\right)^\chi = s^\chi(G).$$

Case **II.** $\boxed{a_i + b_i \leq 1.63 \text{ for any } i.}$

We apply the Hölder's inequality with parameters $x_i = e_i$, $y_i = 1$, $p = 2\chi$, $q = 2\chi/(2\chi - 1)$:

$$1 \leq e_1 + \ldots + e_k \leq \left(e_1^{2\chi} + \ldots + e_k^{2\chi}\right)^{\frac{1}{2\chi}} k^{\frac{2\chi-1}{2\chi}}.$$

It follows from $a_i b_i \leq 0.815^2$ that for $\chi = 1/2 + 1/(10\ln k)$,

$$s^\chi(G_1) + \ldots + s^\chi(G_k) \geq \left(\left(\frac{e_1}{0.815}\right)^{2\chi} + \ldots + \left(\frac{e_k}{0.815}\right)^{2\chi}\right) s^\chi(G)$$
$$\geq \frac{1}{k^{2\chi-1} \cdot 0.815^{2\chi}} \cdot s^\chi(G) \geq \frac{s^\chi(G)}{e^{1/5} \cdot 0.815} > s^\chi(G). \quad \square$$

**Corollary 1.**

$$\frac{1}{2} + \frac{1}{10\ln k} \leq \chi_k \leq \chi_k^* \leq \chi_{U_k} \leq \frac{1}{2} + \frac{1}{2\log_2(k/2)}.$$

*Complexity of a linear function*

**Theorem 3.** *For $k \geq 2$, it holds that $\Phi_{U_{2k}}(l_n) \leq 2n^{1+1/\log_2 k}$.*

1) Any function $f$ of $k$ variables may be implemented by a formula over the basis $U_{2k}$ in such way that any variable is repeated at most twice.

$$\text{DNF for } f(x_1, \ldots, x_k) \quad \xrightarrow{y_i = \overline{x}_i} \quad \varphi(x_1, \ldots, x_k, y_1, \ldots, y_k) \in U_{2k}.$$

2) Verify by induction that $\Phi_{U_{2k}}(l_n) \leq n \cdot 2^{\lceil \log_k n \rceil}$.

For $n = n_1 + \ldots + n_k$, due to 1) we have

$$\Phi_{U_{2k}}(l_n) \leq 2(\Phi_{U_{2k}}(l_{n_1}) + \ldots + \Phi_{U_{2k}}(l_{n_k})).$$

If $n \in (k^{d-1}, k^d]$, then we can always take $n_i \leq k^{d-1}$.

**Corollary 2.** $\Phi_{U_k}(l_n) \asymp n^{1+\Theta(1/\ln k)}$.

$\boxed{k = 3}$

**Theorem 4.** $\chi_3^* \geq 0.769.$ $(\chi_3^* \leq \log_4 3 \approx 0.792)$

[a lengthy proof is omitted]

**Corollary 3.** $n^{1.53} \preceq \Phi_{U_3}(l_n).$ $(\preceq n^{1.74})$

# References

[1] V. M. Khrapchenko. *Method of determining lower bounds for the complexity of P-schemes,* Math. Notes. 1971. **10**(1), 474—479.

[2] N. A. Peryazev. *Complexity of representations of Boolean functions by formulas in nonmonolinear bases,* in Diskret. Mat. Inform. (Izd. Irkutsk Univ., Irkutsk, 1995), Vol. 2.

[3] D. Yu. Cherukhin. *On the complexity of the realization of a linear function by formulas in finite Boolean bases*, Discrete Math. Appl. 2000. **10**(21), 147–157.

[4] H. Chockler, U. Zwick. *Which bases admit non-trivial shrinkage of formulae?* Comput. Complexity. 2001. **10**, 28–40.

[5] I. S. Sergeev. *Formula complexity of a linear function in a $k$-ary basis.* Math. Notes. 2021. **109**(3), 445–458.