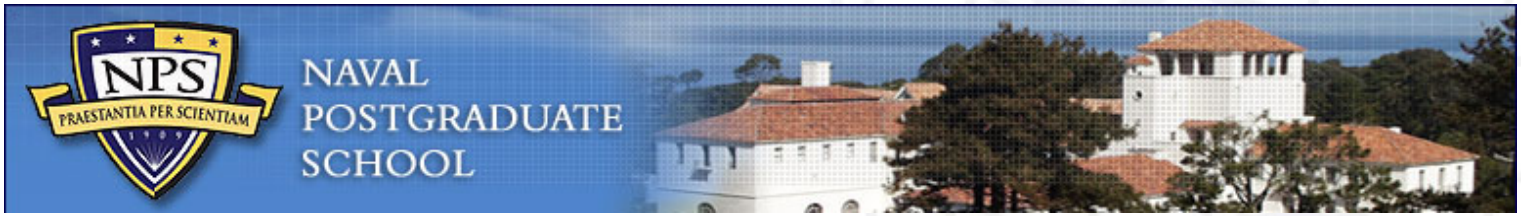
West Coast Number Theory, Pacific Grove, CA, December 2016

Bisecting Binomial Coefficients (*Keeping things fair in a sequence*)

Pante Stanica

Department of Applied Mathematics
Naval Postgraduate School
Monterey, CA 93943, USA; pstanica@nps.edu

*Also associated to IMAR (Institute of Mathematics of the Romanian Academy)



The history of a problem I

- Motivated by an attempt to construct symmetric Boolean functions with various cryptographic properties (resilience, avalanche features), Mitchell (1990), Jefferies (1991), Gopalakrishnan et al. (1993), von zur Gathen and Roche (1997), Cusick & Li (2005), Castro-Medina (2011–2016), among others, study a seemingly "innocent" problem, namely the binomial coefficients bisection (BCB), which we shall describe below.
- The connection between symmetric Boolean functions and binomial coefficients is rather immediate.
- A Boolean function $f : \mathbb{V}_n \to \mathbb{F}_2$ is symmetric if its output value $f(\mathbf{x})$ only depends upon the (Hamming) weight of its input, $\mathrm{wt}(\mathbf{x})$ (number of nonzero bits of $\mathbf{x}$).

The history of a problem II

- There are $\binom{n}{w}$ vector $\mathbf{x}$ of weight $\mathrm{wt}(\mathbf{x}) = w$, and f is constant on each such set of vectors. Thus, f can be "compressed" into an $n+1$ vector of values corresponding to each partition class of cardinality $\binom{n}{w}$, $0 \leq w \leq n$.
- Now, if one further imposes balancedness on f (in addition to symmetry), that is its weight is $\mathrm{wt}(f) = 2^{n-1}$, so we have a two set partition I, J, of the binomial coefficients $\binom{n}{w}$ so that the function f has value $b \in \{0, 1\}$ on the vectors of weight in I and value $\bar{b}$ on vectors in J.
- Thus, we are prompted in studying these splitting (bisections) of binomial coefficients.

The history of a problem III

- If $\sum_{i=0}^{n} \delta_i \binom{n}{i} = 0$, $\delta_i \in \{-1, 1\}$, then we call $(\delta_0, \ldots, \delta_n)$ a solution of the equation $\sum_{i=0}^{n} x_i \binom{n}{i} = 0$, $x_i \in \{-1, 1\}$.
  Certainly, for such a solution, denoting by $I = \{i \mid \delta_i = 1\}$ and $J = \{i \mid \delta_i = -1\}$, we obtain a *bisection*
  $$\sum_{i \in I} \binom{n}{i} = \sum_{i \in J} \binom{n}{i} = 2^{n-1}.$$
- Note (see also [Cusick & Li, 2005]): if n is even, then $\pm(1, -1, 1, -1, \ldots,)$ and if n is odd then $(\delta_0, \ldots, \delta_{(n-1)/2}, -\delta_{(n-1)/2}, \ldots, -\delta_0)$ are $2^{\frac{n+1}{2}}$ solutions. These are called *trivial* solutions.

The history of a problem IV

- There are sporadic cases when non-trivial solutions do appear. In general, when $n \equiv 2 \pmod 6$, because of the identity $\binom{n}{(n+1)/3} = \binom{n}{(n+1)/3-1} + \binom{n}{n-((n+1)/3-1)}$, nontrivial solutions always appear. Besides these results, all is known about the bisection of binomial coefficients is mostly computational.
- E.g.: All known values of n for which non-trivial bisections exist, $n \leq 128$ (von zur Gathen and Roche, '97); all non-trivial bisections for $n \leq 28$ (Jefferies, '91); rediscovered by Cusick et al. (2005).

Our approach to the problem I

- The binomial coefficients bisection can be thought of as a subset sum problem. The view we take is the following: a binomial coefficients bisection $\sum_{i \in I} \binom{n}{i} = \sum_{i \in \bar{I}} \binom{n}{i}$ will generate a solution to the Boolean equation

$$\sum_{i=0}^{n} x_i \binom{n}{i} = 2^{n-1}, x_i \in \{0, 1\}$$

by taking $x_i = 1$ for $i \in I$ and $x_i = 0$, for $i \in \bar{I}$. Certainly, the reciprocal is true, as well, and so, we have an equivalence between these two problems.

Our approach to the problem II

- Further, given a set of positive integers $A = \{a_1, \ldots, a_N\}$ and $b \leq \frac{1}{2} \sum_i a_i$, $b \in \mathbb{N}$, one investigates the Boolean equation

$$\sum_{i=1}^{N} x_i a_i = b, \ x_i \in \{0, 1\}.$$

- The advantage of this approach is that these equations were studied before by analytical number theory methods and much (well, some) is known.
- In general, these problems are well known to be NP-complete [Garey–Johnson, 1979] and have many applications in cryptography, such as the Merkle-Hellman cryptosystem (1978).

Our approach to the problem III

- The density of a set $\mathcal{S} = \{a_1, \ldots, a_N\}$ is defined as

$$d(\mathcal{S}) = \frac{N}{\log_2 \left(\max\limits_{1 \leq i \leq N} a_i \right)}$$

  – in terms of knapsack cryptosystems,

$$d(\mathcal{S}) = \frac{\text{bit size of the plaintext}}{\text{average bit size of the cyphertext}}$$

- For binomial coefficients $\mathbf{P}_n = \left\{ \binom{n}{0}, \binom{n}{1}, \ldots, \binom{n}{n} \right\}$, using the well-known inequalities

$$\frac{4^{\lfloor n/2 \rfloor}}{2\lfloor n/2 \rfloor + 1} \leq \binom{n}{\lfloor n/2 \rfloor} \leq 4^{\lfloor n/2 \rfloor}$$

Our approach to the problem IV

the density becomes

$$\frac{n+1}{2\lfloor n/2\rfloor - \log_2(2\lfloor n/2\rfloor + 1)} \leq d(\mathbf{P}) = \frac{n+1}{\log_2(\max_i \binom{n}{i})} = \frac{n+1}{\log_2 \binom{n}{\lfloor n/2\rfloor}} \leq \frac{n+1}{2\lfloor n/2\rfloor},$$

and so,

$$d(\mathbf{P}) \to 1, \text{ as } n \to \infty.$$

- Lagarias and Odlyzko (1985) showed that almost all the subset sum problem with density $d < 0.6463\ldots$ can be solved in polynomial time with a single call to an oracle that can find (in polynomial time with high probability) the shortest vector in a special lattice.
- Coster et al. (1992) improved the bound to $d < 0.9408\ldots$. Since for the case of binomial coefficients, the density is $d \to 1$ (as $n \to \infty$), none of these methods are applicable.

The underlying method I

- We recall here the following important result of Freiman (1980) (see also [Buzytsky (1982), Chaimovich, Freiman, Galil (1989)]).

Theorem (Freiman '80)

Let $A = \{a_1, a_2, \ldots, a_N\}$ *and* $b \leq \frac{1}{2}\sum_{i=1}^{N} a_i$. *The number of Boolean solutions for the equation*

$$\sum_{i=1}^{N} a_i x_i = b, \ x_i \in \{0, 1\}$$

is precisely $\int_0^1 e^{-2\pi i x b} \prod_{j=1}^{N} \left(1 + e^{2\pi i x a_j}\right) d\,x.$

The underlying method II

- Applying Freiman's paradigm to the bisection of binomial coefficients we immediately infer the next result.

Theorem (S., 2016)

The number of binomial coefficients bisections for fixed n is exactly

$$J_n = \int_0^1 e^{-2^n \pi i x} \prod_{j=0}^{n} \left(1 + e^{2\pi i x \binom{n}{j}}\right) dx = 2^{n+1} \int_0^1 \prod_{j=0}^{n} \cos\left(\pi x \binom{n}{j}\right) dx.$$

- We easily recovered the J_n data of [Jefferies (1991), Cusick & Li (2005)] for $2 \leq n \leq 29$: $2, 4, 2, 8, 2, 16, 6, 32, 2, 64, 2, 144, 14, 256, 2, 512, 2, 1024, 6, 2048, 2, 4096, 50, 8192, 6, 16384, 2, 34816$.

First bound for the number of bisections I

- Our next result gives the first nontrivial upper bound for the number of bisections for odd n, in the literature.

Theorem (S., 2016)

The number J_n of binomial coefficients bisections for odd n is upper bounded by

$$J_n \leq \binom{n+1}{\frac{n+1}{2}} \sim \frac{2^{n+1}}{\sqrt{\pi(n+1)/2}}, \text{ as } n \to \infty.$$

Proof.

- Write $J_n = 2^{n+1} \int_0^1 \prod_{j=0}^{(n-1)/2} \cos^2\left(\pi x \binom{n}{j}\right) dx$, n odd.

First bound for the number of bisections II

- Set $B := (n-1)/2$. By Hölder inequality,

$$\int_0^1 \prod_{j=0}^{B} \cos^2\left(\pi x \binom{n}{j}\right) dx \le \left(\prod_{j=0}^{B} \int_0^1 \cos^{2(B+1)}\left(\pi x \binom{n}{j}\right) dx\right)^{1/(B+1)}. \quad (1)$$

- Using $\int \cos^m(ax)dx = \frac{1}{ma}\cos^{m-1}(ax)\sin(ax) + \frac{m-1}{m}\int \cos^{m-2}(ax)dx$, then, with $m = 2(B+1)$, $a = \pi\binom{n}{j}$, we compute

$$\begin{aligned}
\int_0^1 \cos^{2(B+1)}\left(\pi x \binom{n}{j}\right) dx &= \frac{1}{2(B+1)\pi\binom{n}{j}} \cos^{2B+1}\left(\pi x \binom{n}{j}\right) \sin\left(\pi x \binom{n}{j}\right)\Big|_0^1 \\
&\quad + \frac{2B+1}{2B+2}\int_0^1 \cos^{2B}\left(\pi x \binom{n}{j}\right) dx \\
&= \cdots\cdots \\
&= \frac{(2B+1)(2B-1)\cdots 1}{(2B+2)2B\cdots 2} \\
&= \frac{(2B+2)!}{2^{2(B+1)}((B+1)!)^2} = \frac{1}{2^{2(B+1)}}\binom{2(B+1)}{B+1}.
\end{aligned}$$

First bound for the number of bisections III

Replacing this into the upper bound of (1) we get

$$J_n \le 2^{n+1}\left(\prod_{k=0}^{B}\frac{1}{2^{2(B+1)}}\binom{2(B+1)}{B+1}\right)^{1/(B+1)}$$

$$= \binom{2(B+1)}{B+1} = \binom{n+1}{\frac{n+1}{2}} \sim \frac{2^{n+1}}{\sqrt{\pi(n+1)/2}}, \text{ as } n \to \infty,$$

using Stirling's $n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$. □

Better bound on J_n

- Quite a bit more complicated to get a better bound.

Theorem (Ionascu-Martinsen-S. 2016)

Let $\alpha_s = \frac{n}{\lfloor \frac{n}{2} \rfloor - s}$ *and* $b_i = \binom{n}{i}$, $n \geq 5$. *Then*

$$2^{-(n+2)} J_n \leq \frac{\operatorname{erf}\left(\frac{\pi \sqrt{\binom{2n}{n}}}{2\binom{n}{\lfloor n/2 \rfloor}} \right)}{2\sqrt{\pi \binom{2n}{n}}} + \sum_{s=1}^{\lfloor n/2 \rfloor - 1} \exp\left(-\frac{\pi^2 2^{2n(H(\alpha_s)+o(1))}}{4(\lfloor n/2 \rfloor + 1) b^2_{\lfloor n/2 \rfloor - s + 1}} \right) \frac{s}{(\lfloor n/2 \rfloor - s + 1) b_{\lfloor n/2 \rfloor - s + 1}},$$

where $H(\alpha) = -\alpha \log_2(\alpha) - (1-\alpha)\log_2(1-\alpha)$ *is the binary entropy function and* $\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$ *is the error function.*

- One can find that the expression above is $O\left(\frac{2^n}{n}\right)$ (in fact, $J_n \leq \frac{2^{n+2}}{n}$).

Some exact counts I

- f is *SAC* [Webster – Tavares (1985)]: complementing any of the n input bits the output changes with probability $1/2$.
- f is *SAC* of order k ($SAC(k)$– [Forré (1988)], $0 \leq k \leq n-2$, if whenever k input bits are fixed, the resulting function of $n-k$ variables is *SAC*.

Theorem (I.M.S. 2016)

If p is a prime number, then $J_{p-1} = 2$.

- This implies conjecture $Q2, Q4$ of Cusick and Li (2005): thus, there are only four symmetric $SAC(k)$ functions for infinitely many n.

Some exact counts II

- Based upon our computational data, we conjecture: $J_{2^{2k}} = 2$, $J_{2^{2k+1}} = 6$, $k \geq 1$.

Theorem (Ionascu-Martinsen-S., 2016)

We have:

1. *If* $n = k^2 - 2$, $k \geq 4$ *even, then* $J_n \geq 10$, $J_{n-1} \geq 2^{\frac{n+1}{2}} + 2^{\frac{n+1}{2}-3}$ (*tight*).
2. *If* $k \equiv 0, 1 \pmod 3$ *and* $n = \frac{F_{4k+1}+2F_{4k}-6}{5}$, *then* $J_n \geq 2^{\frac{n+1}{2}} + 2^{\frac{n-3}{2}}$.
3. *Let* $n = 4k^2 + 16k + 13, k \geq 0$. *Then, there are at least* $2^{(n+1)/2-3}$ *nontrivial bisections for the binomial coefficients* $\left\{\binom{n}{j}\right\}_{0 \leq j \leq n}$, *and so,* $J_n \geq 2^{\frac{n+1}{2}} + 2^{\frac{n-1}{2}}$.

Table: Number of Binomial Coefficients Bisections

n	J_n	n	J_n	n	J_n
1	2	18	2	35	$2^{18}+24$
2	2	19	2^{10}	36	2
3	2^2	20	6	37	2^{19}
4	2	21	2^{11}	38	38
5	2^3	22	2	39	2^{20}
6	2	23	2^{12}	40	2
7	2^4	24	50	41	$2^{21}+15\cdot 2^{11}$
8	6	25	2^{13}	42	2
9	2^5	26	6	43	2^{22}
10	2	27	2^{14}	44	134
11	2^6	28	2	45	2^{23}
12	2	29	$2^{15}+2^{11}$	46	2
13	2^7+2^4	30	2	47	$2^{24}+2^{20}$
14	14	31	$2^{16}+5\cdot 2^7$	48	4098
15	2^8	32	6	49	2^{25}
16	2	33	$2^{17}+2^{14}$	50	6
17	2^9	34	130	51	2^{26}

Are there 2^k-sections? I

- It is a natural question to ask whether a splitting of binomial coefficients of size other than two do exist.
- As for the bisection, we say that we have a 2^k-section of a set of integers A if there is a partition of the set A of cardinality 2^k such that the sum on each partition set is $\frac{1}{2^k}\sum_{x\in A_j} x$, $1 \leq j \leq 2^k$.

Theorem (S., 2016 ☺)

Let $n \geq 1$. For $k \geq 2$, there are no 2^k-sections of binomial coefficients $\left\{\binom{n}{j}\right\}_{0\leq j\leq n}$.

▸ Skip2End!

Are there 2^k-sections? II

Proof.

- The result is easy to show for $1 \leq n \leq 10$, so we assume that $n \geq 10$.
- Freiman (1996) considered the system of equations

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m = b_1$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m = b_2$$

where $(0,0) \neq (a_{1j}, a_{2,j}) \in \mathbb{Z}^2$, $(b_1, b_2) \in \mathbb{Z}^2$, and he showed that the number of solutions $x_j \in \{0,1\}$ of the above system is exactly

$$J_{b_1,b_2} = 2^m \int_G \int e^{-2\pi i(xb_1+yb_2)} \prod_{j=1}^{m} \frac{1}{2}\left(1 + e^{2\pi i(xa_{1j}+ya_{2j})}\right) dx\, dy,$$

where $G = \{(x,y) \,|\, x, y \in \mathbb{R}, |x| \leq \frac{1}{2}, |y| \leq \frac{1}{2}\}$.

Are there 2^k-sections? III

- Let $n \geq 10$ be fixed, and assume $\exists$ 2^k-section, $k \geq 2$ (let k largest with this property). We consider such a 2^k-section and partition the binomial coefficients $\binom{n}{j}$ in 2^k (disjoint) sets A_i s.t. $\sum_{j \in A_i} \binom{n}{j} = 2^{n-k}$, $1 \leq i \leq 2^k$.
- Since k is largest with this property (certainly, $k < n$), one of the sets, w.l.o.g., say A_1, cannot be bisected further. We next consider the system

$$\sum_{j \in \cup_{i=2}^{2^k} A_i} x_j \binom{n}{j} + \sum_{j \in A_1} x_j \cdot 0 = (2^k - 1)2^{n-k}$$

$$\sum_{j \in \cup_{i=2}^{2^k} A_i} x_j \cdot 0 + \sum_{j \in A_1} x_j \binom{n}{j} = 2^{n-k},$$

which must have a solution.

Are there 2^k-sections? IV

- By Freiman's system paradigm the # of solutions is exactly

$$
\begin{aligned}
J_{(2^k-1)\,2^{n-k},2^{n-k}} &= 2^{n+1}\int_{-1/2}^{1/2}\int_{-1/2}^{1/2} e^{-2\pi 2^{n-k}((2^k-1)x+y)} \\
&\quad\cdot \prod_{j\in\cup_{i=2}^{2^k}A_i} \frac{1}{2}\left(1+e^{2\pi i x\binom{n}{j}}\right)\prod_{j\in A_1}\frac{1}{2}\left(1+e^{2\pi i y\binom{n}{j}}\right) d\,x\,d\,y \\
&= 2^{n+1}\int_{-1/2}^{1/2} e^{-(2^k-1)\pi 2^{n-k+1}x}\prod_{j\in\cup_{i=2}^{2^k}A_i}\frac{1}{2}\left(1+e^{2\pi i x\binom{n}{j}}\right) \\
&\quad\cdot\int_{-1/2}^{1/2} e^{-\pi 2^{n-k+1}y}\prod_{j\in A_1}\frac{1}{2}\left(1+e^{2\pi i y\binom{n}{j}}\right) \\
&= 2^{n+1}\int_{-1/2}^{1/2}\prod_{j\in\cup_{i=2}^{2^k}A_i}\cos\left(\pi x\binom{n}{j}\right)\int_{-1/2}^{1/2}\prod_{j\in A_1}\cos\left(\pi x\binom{n}{j}\right).
\end{aligned}
$$

Are there 2^k-sections? V

- We let $\langle , \rangle$ be the regular Euclidean scalar product, and observe that

$$\prod_{j \in A_1} \cos\left(\pi i x \binom{n}{j}\right) = \frac{1}{2^{|A_1|-1}} \sum_{\theta \in \{-1,1\}^{|A_1|-1}} \cos\left(\pi i x \left\langle (1,\theta), (\binom{n}{j})_{j \in A_1} \right\rangle\right).$$

- Note that $\left\langle (1,\theta), (\binom{n}{j})_{j \in A_1} \right\rangle \equiv \sum_{j \in A_1} \binom{n}{j} = 2^{n-k} \equiv 0$ (mod 2), for all $\theta \in \{-1,1\}^{|A_1|-1}$.
- Moreover, the scalar product $\left\langle (1,\theta), (\binom{n}{j})_{j \in A_1} \right\rangle \neq 0$, since we assumed that A_1 cannot be bisected further.

Are there 2^k-sections? VI

- Therefore, the integral

$$\int_{-1/2}^{1/2} \prod_{j\in A_1} \cos\left(\pi x \binom{n}{j}\right)$$

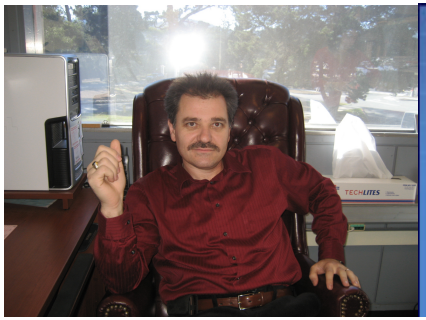
$$= \frac{1}{2^{|A_1|-1}} \int_{-1/2}^{1/2} \sum_{\theta\in\{-1,1\}^{|A_1|-1}} \cos\left(\pi x \left\langle (1,\theta), (\binom{n}{j})_{j\in A_1} \right\rangle\right)$$

$$= \frac{1}{2^{|A_1|-1}} \sum_{\theta\in\{-1,1\}^{|A_1|-1}} \int_{-1/2}^{1/2} \cos\left(\pi x \left\langle (1,\theta), (\binom{n}{j})_{j\in A_1} \right\rangle\right)$$

$$= \frac{1}{2^{|A_1|-1}\pi \left\langle (1,\theta), (\binom{n}{j})_{j\in A_1} \right\rangle} \sum_{\theta\in\{-1,1\}^{|A_1|-1}} \sin\left(\pi x \left\langle (1,\theta), (\binom{n}{j})_{j\in A_1} \right\rangle\right) \Big|_{-1/2}^{1/2}$$

$$= 0,$$

since $\left\langle (1,\theta), (\binom{n}{j})_{j\in A_1} \right\rangle \equiv 0 \pmod 2$, which shows that our assumption that, for $k \geq 2$, there are 2^k-sections of binomial coefficients is false. The proof is done. □



Theorem (Pante Stanica)

Thank you for your attention!

Proof.

None required! (Also, lunch time is almost upon us ☺) □

References I


P.L. Buzytsky, *An effective formula for the number of solutions of linear Boolean equations*, SIAM J. Alg. Disc. Meth. 3:2 (1982), 182–186.


M. Chaimovich, G. Freiman, Z. Galil, *Solving dense subset-sum problems by using analytical number theory*, Journal of Complexity 5 (1989), 271–282.


M.J. Coster, A. Joux, B.A. LaMacchia, A.M. Odlyzko, C.-P. Schnorr, J. Stern, *Improved low-density subset sum algorithms*, Comput. Complexity 2 (1992), 111C–128.


T.W. Cusick, Y. Li, *k-th order symmetric SAC boolean functions and bisecting binomial coefficients*, Discrete Appl. Math. 149 (2005), 73–86.


R. Forré, *The strict avalanche criterion: spectral properties of Boolean functions and an extended definition*, Adv. in Cryptology – Crypto. '88, pp. 450–468.


G.A. Freiman, *An analytical method of analysis of linear Boolean equations*, Ann. N.Y. Acad. Sci. 337 (1980), 97–102.


G.A. Freiman, On Solvability of a System of Two Boolean Linear Equations, *Number Theory: New York Seminar 1991–1995*, 135–150.


J. Fuller, W. Millan, Linear redundancy in *S*-boxes, *Fast Software Encryption* 2003 (Berlin: Springer LNCS 2887, 2003), 74–86.


M.R. Garey, D.S. Johnson, Computer and Intractability: A Guide to the Theory of NP-Completeness, W.H. Freeman and CO., San Francisco, 1979.


J. von zur Gathen, J. Roche, *Polynomials with two values*, Combinatorica 17 (1997), 345–362.

References II


K. Gopalakrishnan, D.G. Hoffman, D.R. Stinson, *A note on a conjecture concerning symmetric resilient functions*, Inform. Proc. Lett. 47 (1993), 139–143.


N. Jefferies, *Sporadic partitions of binomial coefficients*, Elec. letters 27:15 (1991), 134–136.


J.C. Lagarias, A.M. Odlyzko, *Solving Low-Lensity Subset Sum Problems*, J. Assoc. Comp. Mach. 32:1 (1985), 229–246.


R. Merkle, M. Hellman, *Hiding Information and Signatures in Trapdoor Knapsacks*, IEEE Trans. Inf. Theory 24:5 (1978), 525–530.


C. Mitchell, *Enumerating boolean functions of cryptographic significance*, J. Cryptology 2 (1990), 155–170.


M.B. Nathanson, *Additive Number Theory: Inverse Problems and the Geometry of Sumsets*, Graduate Texts in Mathematics, Vol. 165, Springer-Verlag, New York, 1996.


P. Stănică, *Good Lower and Upper Bounds on Binomial Coefficients*, J. Inequalities in Pure and Applied Math., Vol.2, Issue 3, Art. 30 (2001).


A.F. Webster, S.E. Tavares, *On the design of S-boxes*, Advances in Cryptology – Crypto. 1985, pp. 523–534.