ELSEVIER

Theoretical Computer Science 148 (1995) 133–140

Theoretical Computer Science

Note

The fourth moment in Luby's distribution ☆

Devdatt P. Dubhashi [a], Grammati E. Pantziou [b], Paul G. Spirakis [c], Christos D. Zaroliagis [d,*]

[a] *BRICS [1], Department of Computer Science, University of Aarhus, Ny Munkegade, DK-8000 Aarhus C, Denmark*

[b] *Department of Computer Science, University of Central Florida, Orlando FL 32816, USA*

[c] *Computer Technology Institute, P.O. Box 1122, 26110 Patras, Greece; and Department of Computer Science and Engineering, Patras University, 26500 Patras, Greece*

[d] *Max-Planck-Institut für Informatik, Im Stadtwald, 66123 Saarbrücken, Germany*

Received October 1993; revised February 1995
Communicated by J. Díaz

Abstract

Luby (1988) proposed a way to derandomize randomized computations which is based on the construction of a small probability space whose elements are 3-wise independent. In this paper we prove some new properties of Luby's space. More precisely, we analyze the fourth moment and prove an interesting technical property which helps to understand better Luby's distribution. As an application, we study the behavior of random edge cuts in a weighted graph.

Keywords: Fourth moment; Full independence; k-wise independence; Derandomization

1. Introduction

During the last years there is a growing interest in techniques for removing randomness from parallel (and sequential) algorithms. These techniques were originated by [7, 8] and generalized in [1, 2, 4, 6, 9–11]. The approach usually followed can be summarized as follows: The random variables which are considered are defined over a *smaller* probability space, specially designed, containing only a polynomial number of sample points. In that space, the random variables are only k-wise independent

☆ This work was partially supported by the EU ESPRIT Basic Research Action No. 7141 (ALCOM II), by the Greek Ministry of Education and by NSF grant CDA-9211155. Part of this work has been done while the first author was with the Max-Planck-Institut für Informatik, Saarbrücken.

* Corresponding author. Email: zaro@mpi-sb.mpg.de.

[1] Basic Research in Computer Science, Centre of the Danish National Research Foundation.

0304-3975/95/$09.50 © 1995 — Elsevier Science B.V. All rights reserved
SSDI 0304-3975(95)00056-9

(for constant k) but this is usually enough to replace the analysis of the randomized algorithm with fully independent random variables.

In most cases, only 3-wise (or 2-wise) independence is enough. However, in some instances, this is not sufficient [3]. The algorithms described in that paper can be derandomized successfully only because of the 4-wise independence property. In particular, an explicit example is given where Luby's distribution [10], which is 3-wise but not 4-wise independent, cannot be used for the derandomization. But perhaps, it can be hoped that by relating the fourth moments under Luby's distribution and the fully independent distribution, one can use Luby's distribution in some other cases. This, so called fourth moment issue [2, 3], is very interesting technically because it might indicate the dividing barrier between the two probability spaces, namely k-wise and complete independence.

In this paper, we prove some new properties of Luby's probability space, as defined in [10]. More precisely, we examine the fourth moment of this space. We compute the joint probability that four random variables take particular values, and compare it to the corresponding joint probability under the fully independent distribution. The proof of the result is interesting in its own right and may lead to a general methodology of proving such results. We also relate precisely, the fourth moments under the two distributions. As an application, we study the behavior of random edge cuts in a weighted graph. Based on Luby's probability space, it is easy to construct a linear sized space of edge cuts. We then prove that this smaller space has *bigger* variance compared with the variance in the fully independent space of all possible edge cuts taken equiprobably (which is exponential in size). Thus, the smaller space can be a good predictor of extreme values of random variables defined on the larger space, possibly leading to *NC* algorithms for better approximations to the maximum edge cut problem.

The paper is organized as follows. In Section 2 we present the new properties of Luby's distribution and the fourth moment bound. In Section 3 we discuss the applications of these properties in the computation of edge cuts in weighted graphs.

2. Properties of Luby's sample space

Luby in [10], considers random variables $X_1,\ldots,X_n$, for a positive integer n, defined on the sample space $(\Omega, \Pr)$, where $\Omega = GF(2)^{k+1} = \{0,1\}^{k+1}, k = \lceil \log n \rceil$ and Pr the equiprobable measure, i.e., for each point $\omega \in \Omega$ we have $\Pr(\omega) = 2^{-(k+1)}$. Let $\mathbf{i} \in \{0,1\}^k$ denote the binary representation $\langle i_1,\ldots,i_k\rangle$ of the integer i for $1 \leqslant i \leqslant n$. At a point $\omega = \langle \omega_1,\ldots,\omega_{k+1}\rangle$, the random variable X_i, for $1 \leqslant i \leqslant n$, takes the values given by the formula:

$$X_i(\omega) = \mathbf{i} \cdot \omega + \omega_{k+1}, \tag{1}$$

where the notation $\mathbf{i}\cdot\omega$ denotes $i_1\omega_1 + \cdots + i_k\omega_k$. (Note that in this section, all operations are under $GF(2)$. Also, the reader is assumed to be familiar with basic linear algebra terminology and results; see e.g. [12]).

An alternate but equivalent description is as follows: Let $\mathbf{L}$ be an $n \times (k+1)$ matrix over $GF(2)$, whose ith row is $[\mathbf{i}, 1] = [i_1, \ldots, i_k, 1]$, for $1 \leqslant i \leqslant n$. Then at the point $\omega \in \Omega$ (where now $\omega = [\omega_1, \ldots, \omega_{k+1}]^T$), the random variables take the values given by the vector $\mathbf{L} \cdot \omega$. We call the values taken by the random variables at a point ω, their *labels* at ω.

We call a set of integers *dependent* if their binary representations are dependent as vectors over $GF(2)$, and *independent* otherwise. The matrix $\mathbf{L}$ has some interesting properties which we give in the following (easy) proposition.

Proposition 1. (i) *Any three rows of* $\mathbf{L}$ *are linearly independent.* (ii) *Any four rows of* $\mathbf{L}$ *are linearly independent unless they correspond to dependent integers, that is, to integers such that the binary representation of any one of them is the sum of the binary representations of the other three.*

Proof. First note that no row is $\mathbf{0}$ on account of the last column. Hence, the only way for two rows to be linearly dependent is if their sum is $\mathbf{0}$. However, this is impossible as the binary representation of two distinct integers have a position where they differ. Thus any two rows are linearly independent. This in turn implies that the only way for any three rows to be dependent is if their sum is $\mathbf{0}$, which is impossible since the last column in such a sum is necessarily non-zero. Hence any three rows are linearly independent. From the independence of any three rows, it follows that the only way that any four rows can be dependent is if their sum is $\mathbf{0}$. This happens iff they correspond to integers with the stated property. □

These properties of $\mathbf{L}$ imply the following properties of the distribution of the random variables X_i, $i = 1, 2, \ldots, n$, defined above. (*Remark*: The first three are well-known; we add the last, interesting property.)

Lemma 2. *Let* i, j, l, m *be distinct integers between* 1 *and* n (*so necessarily* $n \geqslant 4$ *below*) *and* b_i, b_j, b_l, b_m *be an arbitrary bit pattern. Then, the following hold in Luby's distribution*:

1. $\Pr[X_i = b_i] = 1/2$.
2. $\Pr[X_i = b_i \wedge X_j = b_j] = 1/4$.
3. $\Pr[X_i = b_i \wedge X_j = b_j \wedge X_l = b_l] = 1/8$.
4. $\Pr[X_i = b_i \wedge X_j = b_j \wedge X_l = b_l \wedge X_m = b_m]$

$$= \begin{cases} 1/16 & \text{if } \mathbf{i} + \mathbf{j} + \mathbf{l} \neq \mathbf{m}, \\ 1/8 & \text{if } \mathbf{i} + \mathbf{j} + \mathbf{l} = \mathbf{m} \text{ and } b_i + b_j + b_l = b_m, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Since the proofs of (1)–(3) are similar, we shall prove the strongest one (3). Take the subsystem of $\mathbf{L} \cdot \omega = \mathbf{b}$ corresponding to the rows i, j, l, to get $\mathbf{L}' \cdot \omega = [b_i, b_j, b_l]^T$. Take further a full rank square submatrix of $\mathbf{L}'$ to form the square system $\mathbf{L}'' \cdot [\omega_{i'}, \omega_{j'}, \omega_{l'}]^T = [b_i, b_j, b_l]^T$. Since the coefficient matrix is non-singular, this has

a unique solution. Fixing these three co-ordinates of ω as per the unique solution, and the rest to zeroes gives one point $\omega^* \in \Omega$ giving X_i, X_j, X_l, the respective labels b_i, b_j, b_l.

Let $\mathbf{C}$ be a $3 \times n$ matrix (over $GF(2)$) with rows corresponding to i', j', l' such that each row has all zeroes except for the position given by the corresponding integer where it is 1. Note that $\mathbf{C}$ has full rank. Now, ω gives the same labels as ω^* to X_i, X_j, X_l iff $\mathbf{CL}(\omega - \omega^*) = \mathbf{0}$, i.e. iff $\omega - \omega^* \in \mathrm{Ker}\,\mathbf{CL}$. Since $\dim(\Omega) = \dim(\mathrm{Ker}\,\mathbf{CL}) + \dim(\mathbf{CL})$ (see e.g. [12, Theorem 6.8]) and $\dim(\mathbf{CL}) = \mathrm{rank}(\mathbf{CL}) = 3$, it follows that $\dim(\mathrm{Ker}\,\mathbf{CL}) = (k+1) - 3 = k - 2$. Hence $|\mathrm{Ker}\,\mathbf{CL}| = 2^{k-2}$ and consequently the probability in question is $2^{k-2}/2^{k+1} = 1/8$.

Turning now to the final property (4), we have that if $\mathbf{i} + \mathbf{j} + \mathbf{l} \neq \mathbf{m}$, then the rows corresponding to these integers are independent, and the proof as above gives the stated probability. Otherwise, if $\mathbf{i} + \mathbf{j} + \mathbf{l} = \mathbf{m}$, then the label of X_m is determined by those of X_i, X_j and X_l via

$$X_m(\omega) = \mathbf{m} \cdot \omega + \omega_{k+1} = (\mathbf{i} + \mathbf{j} + \mathbf{l}) \cdot \omega + \omega_{k+1}$$

$$= \mathbf{i} \cdot \omega + \omega_{k+1} + \mathbf{j} \cdot \omega + \omega_{k+1} + \mathbf{l} \cdot \omega + \omega_{k+1} = X_i(\omega) + X_j(\omega) + X_l(\omega).$$

Hence if $b_i + b_j + b_l \neq b_m$, there are no points of Ω corresponding to these labels and the probability is zero. Otherwise, the probability is the same as that of the event that the three random variables take on a fixed label pattern which is computed in (3). □

One can compare Luby's distribution to the fully independent distribution where each X_i is equiprobably 0 or 1 independently. For this, we shall make use of the following notation.

Notation 1. We shall denote the statistics of Luby's distribution with operators subscripted by L, for example, E_L, σ_L^2 and those of the fully independent distribution by the subscript I, for example, E_I, σ_I^2. If no subscripts appear, then the result holds for both distributions.

The following lemma relates the moments between the two distributions.

Lemma 3. *Let $X = X_1 + \cdots + X_n$. We have that $E_I[X^a] = E_L[X^a]$ for $1 \leqslant a \leqslant 3$ and*

$$E_L[X^4] = E_I[X^4] + \frac{1}{64}\binom{n}{3}.$$

Proof. The statements of the third and lower moments follow from the 3-wise independence of Luby's distribution (Lemma 2). For the fourth moment, we observe by expanding that the only difference will come from terms of the form $E_L[X_i X_j X_l X_m]$ for distinct i, j, l, m. In turn, this equals the probability computed in Lemma 2 above, applied to the bit pattern consisting of all ones. For integers i, j, l, m which are independent, this is the same as that for the fully independent distribution. For integers i, j, l, m which are dependent, this exceeds the probability of the fully independent distribution

by 1/16. The result now follows from the fact that there are a total of $\binom{n}{3}/4$ such dependent tuples of integers. □

3. Computing edge cuts in a weighted graph

Let $G = (V, E)$ be a graph with weights $W_e > 0$ for each $e \in E$. Let also (V_1, V_2) be a partition of V into two disjoint sets V_1 and V_2. Then a *cut* $\mathscr{C}$ in G is the set of edges with one endpoint in V_1 and the other in V_2. The *weight* of the cut $\mathscr{C}$, is the sum of the weights of all edges in $\mathscr{C}$. The problem of asking whether there is cut in a graph G with weight at least K ($K > 0$) is known as the *max-cut* problem and is also known that it is an NP-complete problem [5].

Consider the application of Luby's distribution to compute a random cut $\mathscr{C}$ in a graph G defined as above. Each vertex $v \in V$ picks a label $X_v \in \{0, 1\}$ and an edge is in $\mathscr{C}$ iff its endpoints have different labels. For any given edge, the probability that it is in $\mathscr{C}$ is $\frac{1}{2}$ if the labels are picked either uniformly and independently from $\{0, 1\}$, or using Luby's scheme. The latter part of the above statement holds because of the 2-wise independence of Luby's distribution.

For two distinct edges, the probability that they are both in the cut under the fully independent distribution, is $\frac{1}{4}$. The next proposition computes this probability under Luby's distribution.

Proposition 4. *Let e, e' be fixed edges of G. Then, the following hold:* (i) *If e and e' share a vertex, then* $\Pr_L[e \in \mathscr{C} \wedge e' \in \mathscr{C}] = \frac{1}{4}$. (ii) *If e and e' are disjoint, but their endpoints correspond to independent integers, then* $\Pr_L[e \in \mathscr{C} \wedge e' \in \mathscr{C}] = \frac{1}{4}$. (iii) *If e and e' are disjoint, but their endpoints are dependent integers, then* $\Pr_L[e \in \mathscr{C} \wedge e' \in \mathscr{C}] = \frac{1}{2}$.

Proof. Follows easily from the probabilities computed in Lemma 2. □

Let C be the random variable denoting the weight of the cut $\mathscr{C}$. Then, we have (using the *Iversonian* APL notation $[P]$ which denotes 1 if the boolean property P is true and 0 otherwise),

$$C = \sum_{e=(u,v)} [X_u \neq X_v] W_e = \sum_{e=(u,v)} (X_u + X_v - 2X_u X_v) W_e = \sum_e Y_e W_e, \tag{2}$$

since X_u, X_v are 0-1 valued. Here we denote, for $e = (u, v)$,

$$Y_e := X_u + X_v - 2X_u X_v .$$

Hence,

$$E[C] = E\left[\sum_e Y_e W_e\right]$$
$$= \sum_{e=(u,v)} E[(X_u + X_v - 2X_u X_v) W_e]$$

$$= \sum_{e=(u,v)} (E[X_u] + E[X_v] - 2E[X_u X_v])\, W_e$$
$$= \sum_{e=(u,v)} (E[X_u] + E[X_v] - 2E[X_u]E[X_v])\, W_e$$
$$= 1/2 \sum_{e=(u,v)} W_e. \tag{3}$$

(In the third line, for Luby's distribution, we use the 2-wise independence property.)

Next we compute the second moment. From (2), we have,

$$C^2 = \sum_e Y_e W_e \cdot \sum_{e'} Y_{e'} W_{e'} = \sum_{e,e'} Y_e Y_{e'} W_e W_{e'}$$
$$= \sum_{e \cap e' = \emptyset} Y_e Y_{e'} W_e W_{e'} + \sum_{e \cap e' \neq \emptyset} Y_e Y_{e'} W_e W_{e'}$$
$$= \sum_{\substack{e=(u,v) \\ e'=(w,z)}} (X_u + X_v - 2X_u X_v)(X_w + X_z - 2X_w X_z) W_e W_{e'}$$
$$+ \sum_{\substack{e=(u,v) \\ e'=(u,w)}} (X_u + X_v - 2X_u X_v)(X_u + X_w - 2X_u X_w) W_e W_{e'}$$
$$= \sum_{\substack{e=(u,v) \\ e'=(w,z)}} W_e W_{e'} \left(\sum_{a=u,v} \sum_{b=w,z} X_a X_b - 2 \sum_{a \neq b \neq c \neq a} X_a X_b X_c + 4 X_u X_v X_w X_z \right)$$
$$+ \sum_{\substack{e=(u,v) \\ e'=(u,w)}} W_e W_{e'} (X_u + X_v X_w - X_u X_w - X_u X_v).$$

Hence it follows that

$$E[C^2] = \sum_{\substack{e=(u,v) \\ e'=(w,z)}} 4 W_e W_{e'} E[X_u X_v X_w X_z] + 1/4 \sum_{e \cap e' \neq \emptyset} W_e W_{e'}. \tag{4}$$

Here we use for Luby's distribution, the 3-wise independence property and the probabilities computed in Lemma 2.

For the fully independent distribution, we immediately have that for distinct u, v, w, z, $E_I[X_u X_v X_w X_z] = \frac{1}{16}$. Thus we conclude from (4) that

$$E_I[C^2] = 1/4 \sum_{e \cap e' = \emptyset} W_e W_{e'} + 1/4 \sum_{e \cap e' \neq \emptyset} W_e W_{e'}. \tag{5}$$

For Luby's distribution, we use the probabilities computed in Lemma 2 to get:

$$E_L[X_u X_v X_w X_z] = \begin{cases} 1/16 & \text{if } u, v, w, z \text{ are all distinct and independent;} \\ 1/8 & \text{otherwise.} \end{cases}$$

Hence, using (4),

$$E_L[C^2] = 1/4 \sum_{\substack{e \cap e' = \emptyset \\ \neg D(e,e')}} W_e W_{e'} + 1/2 \sum_{D(e,e')} W_e W_{e'} + 1/4 \sum_{e \cap e' \neq \emptyset} W_e W_{e'}, \tag{6}$$

where we use $D(e,e')$ to denote that the endpoints of e,e' are disjoint dependent integers. Comparing (5) and (6), we get:

$$E_L[C^2] = E_I[C^2] + 1/4 \sum_{D(e,e')} W_e W_{e'} .$$

Since by (3) $E_I[C] = E_L[C]$, we conclude:

Theorem 5. *The variance under Luby's distribution and the variance under the fully independent distribution are related as follows*:

$$\sigma_L^2[C] = \sigma_I^2[C] + 1/4 \sum_{D(e,e')} W_e W_{e'},$$

where $D(e,e')$ denotes that e,e' have disjoint endpoints which are dependent integers.

Thus the variance under Luby's distribution is at least as big as the variance under the fully independent distribution. Potentially, this can be used to get a better predictor of extreme values as implied by the following observation:

Observation 6. Given a weighted graph G, we can compute in NC a cut with weight either at most $1/2 \sum_e W_e - \alpha$ or at least $1/2 \sum_e W_e + \alpha$, where

$$\alpha^2 := \sigma_I^2(C) + 1/4 \sum_{D(e,e')} W_e W_{e'} .$$

To see that such a cut exists, use the variance under Luby's distribution. Further since Luby's sample space has only linear size, we can exhaustively search it for the "good" point in NC.

Remark. Under Luby's distribution, the random variable $X = X_1 + X_2 + \cdots + X_n$ is symmetrically distributed around its mean $E[X] = n/2$. To see this, considerfor each point $\omega = \langle \omega_1, \ldots, \omega_k, \omega_{k+1} \rangle$ the point $\omega' := \langle \omega_1, \ldots, \omega_k, 1 - \omega_{k+1} \rangle$ and compute:

$$X(\omega') = \sum_{i=1}^{n} X_i(\omega') = \sum_{i=1}^{n} (1 - X_i(\omega)) = n - X(\omega)$$

Thus for each point ω such that $X(\omega) = n/2 - \alpha$, there corresponds the unique point ω' such that $X(\omega') = n/2 + \alpha$. Hence for each α, $\Pr[X = n/2 - \alpha] = \Pr[X = n/2 + \alpha]$. Unfortunately, this property no longer holds for the variable C we are interested in. We suspect (but cannot prove) that nevertheless, the distribution of C is "shifted upwards" in the sense that if $\Pr[C = E[C] - \alpha] > 0$, then also $\Pr[C = E[C] + \alpha] > 0$ for any $\alpha > 0$. This would give us a predictor of an extreme value for max–cut.

4. Conclusion

We presented here some new properties of Luby's probability space [10]. In particular, we analyzed the fourth moment and gave an application to the behavior of random

edge cuts in a weighted graph. It would be very interesting if the new properties of Luby's distribution presented in this paper can find other applications too.

References

[1] N. Alon, L. Babai and A. Itai, A fast and simple randomized parallel algorithm for the maximal independent set problem, *J. Algorithms* **7** (1986) 567–583.

[2] B. Berger, Using randomness to design efficient deterministic algorithms, Ph.D. Thesis, Dept. of Electrical Engineering and Computer Science, MIT, 1990.

[3] B. Berger, The fourth moment method, *Proc. 2nd ACM-SIAM Symp. on Discrete Algorithms* (1991) 373–383.

[4] B. Berger and J. Rompel, Simulating ($\log^c n$)-wise independence in *NC*, *Proc. 30th IEEE Symp. on FOCS* (1989) 2–7.

[5] M. Garey and D. Johnson, *Computers and Intractability – A Guide to the Theory of* NP-*Completeness* (Freeman, San Francisco, 1979).

[6] M. Goldberg and T. Spencer, A new parallel algorithm for the maximal independent set problem, *SIAM J. Comput.* **18**(2) (1989) 419–427.

[7] A. Joffe, On a set of almost deterministic k-independent random variables, *Ann. Probab.* **2** (1974) 161–162.

[8] R. Karp and A. Wigderson, A fast parallel algorithm for the maximal independent set problem, *J. ACM* **32** (1985) 762–773.

[9] M. Luby, A simple parallel algorithm for the maximal independent set problem, *SIAM J. Comput.* **15** (1986) 1036–1053.

[10] M. Luby, Removing randomness in parallel computation without a processor penalty, *Proc. 29th IEEE Symp. on FOCS* (1988) 162–174.

[11] R. Motwani, J. Naor and M. Naor, The probabilistic method yields deterministic parallel algorithms, *Proc. 30th IEEE Symp. on FOCS* (1989) 8–13.

[12] B. Noble and J.W. Daniel, *Applied Linear Algebra* (Prentice-Hall, Englewood Cliffs, NJ, 3rd ed., 1988).