Deterministic Approximate Counting of Depth-2 Circuits

Michael Luby* Boban Veličković† Avi Widgerson ‡

Abstract

We describe deterministic algorithms which for a given depth-2 circuit F approximate the probability that on a random input F outputs a specific value α. Our approach gives an algorithm which for a given GF[2] multivariate polynomial p and given $\epsilon > 0$ approximates the number of zeros of p within a multiplicative factor $1+\epsilon$. The algorithm runs in time $\exp(\exp(O(\sqrt{\log(n/\epsilon)})))$, where n is the size of the circuit. We also obtain an algorithm which given a DNF formula F and $\epsilon > 0$ approximates the number of satisfying assignments of F within a factor of $1+\epsilon$ and runs in time $\exp(O((\log(n/\epsilon))^4))$.

1 Introduction

This paper deals with the problem of approximating the accepting probability of general depth-2 boolean circuits. Examples of boolean functions which can be computed by such circuits are DNF formulas, polynomials over $GF[2]$, threshold functions, etc. There are easy probabilistic algorithms which for a given circuit F and a real parameter $\epsilon > 0$ approximate the probability that on a random input F outputs 0. The algorithm simply chooses $N = O(\ln(1/\delta)/\epsilon^2)$ assignments uniformly at random and outputs the ratio Y of the assignments which satisfy F. The probability that Y is within ϵ of $\Pr[F = 0]$ is at least $1 - \delta$. For the case of DNF formulas and polynomials over $GF[2]$ it has been shown that these algorithms can be transformed to efficient probabilistic algorithms which approximate the accepting probability of F in the relative sense, i.e. that produce a real Y such that $|\Pr[F = 0] - Y| \leq \epsilon \Pr[F = 0]$ (see [KL1] and [KL2]).

The problem of finding such approximation algorithms which are deterministic is significantly harder and has generated quite a bit of attention recently. For DNF formulas there have been produced deterministic approximation algorithms which run in time

*International Computer Science Institute and UC Berkeley, research supported in part by NSF Grant CCR-9016468 and grant No. 89-00312 from the United States-Israel Binational Science Foundation (BSF), Jerusalem, Israel

†Department of Mathematics, York University, Toronto, research partially supported by NSF, research partially done while at UC Berkeley

‡Hebrew University and Princeton University, research partially done while visiting the International Computer Science Institute

$\exp\left(\log(n/\epsilon)\right)^{O(1)})$, where n is the size of the input (see [NW] and [LV]). We treat the case of general symmetric depth-2 boolean circuits and present a deterministic approximation algorithm which given such a ciruit F and $\epsilon > 0$ produces an ϵ-approximation of $\Pr[F = 0]$ and runs in time $\exp(\exp(O(\sqrt{\log(n/\epsilon)})))$, where n is the size of the circuit. The most interesting application of this algorithm is for approximating the number of zeros of multivariate polynomials over $GF[2]$. Although the running time is probably very far from the optimal, the running time is $\exp((n/\epsilon)^{o(1)})$, which is significantly better than the trivial bound of $\exp(n)$. Also, in the case of DNF formulas we modify the algorithm presented in [NW] to improve its running time to $\exp(\log(n/\epsilon)^4)$.

The paper is organized as follows. In §2 we present the outline of the main construction. §3 deals with the notion of a d-nearly disjoint family for a given (m, n, k)-system and contains a construction of such families. In §4 we deal with multiparty communication protocols and present the approximation algorithm for general symmetric depth-2 circuits. This algorithm relies crucially on the lower bound result for the multiparty complexity of the generalized inner product function due to Babai, Nisan, and Szegedy [BNS]. In §5 we show that the algorithm from [NW] in the case of DNF formulas can be modified to significantly improve its running time. Finally, in §6 we make some remarks on converting the algorithms presented in this paper to relative approximation algorithms.

2 Construction outline

Let us consider a general depth-2 circuit which computes a boolean function F on n variables $x_1, \ldots, x_n$. Suppose that the first level of the circuit has m nodes which compute functions $T_1, \ldots, T_m$. In all our applications each T_i will be the conjunction of some of the variables or their negations. Let C_i be the set of indices of the variables on which T_i depends. We are interested in approximating the probability μ that F takes value 0 when all the variables are randomly, uniformly, and independently chosen. Let $\epsilon > 0$ be the error allowed. We follow the standard approach to such approximation problems which is to replace the uniform distribution $\mathcal{U}$ on $\{0,1\}^n$ by another efficiently constructible distribution $\mathcal{V}$ which has much smaller sample space such that the probability with respect to $\mathcal{V}$ that F evaluates to 0 is within ϵ of μ. The algorithm then queries exhaustively all sample points of $\mathcal{V}$ and outputs the ratio of points for which F equals 0.

We now describe the construction of $\mathcal{V}$. We start with an integer l which is much smaller than n, define n random variables $X_1, \ldots, X_n$ on $\{0,1\}^l$ and let $\mathcal{V}$ be the probability distribution on $\{0,1\}^n$ generated in this way. We arrange things so that the probability ν that F evaluates to 0 with these random variables as inputs is within ϵ of μ. The idea of Nisan and Wigderson [NW] is to choose n subsets of $[l]$, say $S_1, \ldots, S_n$, which are nearly disjoint in a certain technical sense and have size r and a simple function g on r variables which is hard to approximate in an appropriate model of computation. Let $y_1, \ldots, y_l$ be uniformly distributed and independent $\{0,1\}$-random variables and then let X_i be equal to $g(y_{S_i})$. Here y_{S_i} denotes the string consisting of the coordinates of y whose index is in S_i in the natural ordering. We show that if $|\mu - \nu| \geq \epsilon$ then there is an algorithm A with inputs from $\{0,1\}^r$ which computes the correct value of $g(z)$ on at least $1/2 + \epsilon/n$ fraction of the inputs. This algorithm has special form and the choice of g is made such that there is no algorithm of such form which approximates it. It follows that we must have $|\mu - \nu| \leq \epsilon$. In the case of DNF formulas g is, as in [NW], XOR and we use the lower bounds of Håstad and Bopanna (see [H]). In the case of modular or threshold circuits we use the generalized inner product function and a lower bound on the multiparty communication complexity of this function due to Babai, Nisan and Szegedy [BNS].

3 Nearly disjoint sets

Let I be a set of size l and $\mathcal{S}$ a family of n subsets of I each of size at least r. For a positive integer d we say that $\mathcal{S}$ is *d-nearly disjoint* provided $|S \cap T| < d$, for any two distinct $S, T \in \mathcal{S}$. We also call such a family an *(n,l,d,r)-design.* The existence of (n,l,d,r)-designs for appropriate values of n, l, d and r was used by Nisan and Wigderson [NW] in their approximation algorithm. We now refine this notion and use it to derive an improved algorithm.

Suppose we are given in addition two integer parameters m and k and a family $\mathcal{C}$ of m subsets of $[n]$ each of size at most k. We call such a family $\mathcal{C}$ an *(m,n,k)-system.* Let us say that a family $\mathcal{S} = \{S_1, \ldots, S_n\}$ of subsets of I is *d-nearly disjoint* for $\mathcal{C}$ if for every $i \in [n]$ and $j \in [m]$ $|S_i \cap \bigcup_{i' \in C_j \setminus \{i\}} S_{i'}| < d$. Note that a d-nearly disjoint family is simply a d-nearly disjoint family for the family of all singletons from $[n]$. If, as before, the size of each member of $\mathcal{S}$ is at least r we say that $\mathcal{S}$ a *(n,l,d,r)-design for $\mathcal{C}$*. We shall need to construct efficiently (n,l,d,r)-designs for a given (m,n,k)-system. We first prove by a probabilistic argument the existence of such designs with appropriate parameters and then show how this can be converted to an efficient construction.

Let us fix nl biased $2d$-independent $\{0,1\}$-valued random variables x_{ij}, for $i \in [n]$ and $j \in [l]$, such that x_{ij} is equal 1 with probability $2r/l$ and 0 otherwise. For a fixed setting $x = \langle x_{i,j} : i \in [n], j \in [l] \rangle$ to the sequence of random variables and for each $i \in [n]$, let $S_i(x) = \{j : x_{ij} = 1\}$, and as before let $\mathcal{S}(x) = \{S_1(x), \ldots, S_n(x)\}$.

Lemma 1 *Let $\mathcal{C}$ be a given (m,n,k)-system. The probability that $\mathcal{S}(x)$ is not a (n,l,d,r)-design for $\mathcal{C}$ is at most*

$$\frac{n}{r^d} + mn \left(\frac{12r^2k}{dl} \right)^d .$$

PROOF : Note first that for each i the expected value of $|S_i|$ is $2r$. By the $2d$-moment extension of Chebyshev's inequality $\Pr[|S_i| < r] < r^{-d}$. Note that this bounds uses only

$2d$-wise independence between the random variables. Summing up over all $i \in [n]$ we deduce that the probability that there exists i such that $|S_i| < r$ is at most n/r^d.

Fix $i \in [n]$ and $j \in [m]$. We give an upper bound of the probability that $|S_i \cap \bigcup_{i' \in C_j \setminus \{i\}} S_{i'}| \geq d$. For a fixed $A \subseteq [l]$ of size d the probability that A is contained in both S_i and $\bigcup_{i' \in C_j \setminus \{i\}} S_{i'}$ is at most $\left(\frac{4r^2k}{l^2}\right)^d$. This again uses only $2d$-wise independence. Summing this up over all possible subsets of $[l]$ of size d we obtain the upper bound

$$\frac{l!}{d!(l-d)!}\left(\frac{4r^2k}{l^2}\right)^d.$$

Summing up over all possible pairs $(i,j) \in [n] \times [m]$ and combining the two upper bounds we obtain the desired result. ■

We shall apply this lemma in two different situation with different sets of parameters. We state this in the following corollary.

Corollary 2 *For a given (m,n,k)-system $\mathcal{C}$ the probability that $\mathcal{S}(x)$ is not an (n,l,d,r)-design for $\mathcal{C}$ is at most ϵ for the following choices of parameters:*

(a) $d = \sqrt{\log(nm/\epsilon)}$, $r = 2^{4d}$, *and* $l = k2^{9d}$.

(b) $d = \log(mn/\epsilon)$, $r = d^2$, *and* $l = 24kd^3$.

Note that in Lemma 1 we use only $2d$-wise independence of the random variables. It will be crucial in our applications that there exist efficiently constructible $2d$-wise independent probability distributions with small sample spaces (see [L], [ABI]). This allows us to convert these existence arguments into constructions by searching exhaustively all sample points. We state the result we shall need in the following proposition.

Proposition 3 *There is an explicit construction of a sample space of size $(nl)^{O(d)}$ for nl $\{0,1\}$-valued $2d$-wise independent random variables $X_{i,j}$, for $(i,j) \in [n] \times [m]$, such that $X_{i,j}$ is equal to 1 with probability $2r/l$.*

The size of the sample space can further be reduced if instead of truly $2d$-wise independent random variables we consider $2d$-wise δ-dependent variables for some small δ. In this case the size of the sample space can be made polynomial in $\log(nl)$, d, and $1/\delta$, (see [NN], [AGHP], [EGLNV]).

4 General symmetric gates

4.1 Multiparty communication protocols

Let us recall the definition of d-party communication protocols [CFL]. Suppose $[r] = R_1 \dot\cup R_2 \dot\cup \cdots \dot\cup R_d$ is a partition of $[r]$ into d disjoint parts. Imagine d players, $P_1, P_2, \ldots, P_d$, who have access to an input string $z \in \{0,1\}^r$ such that P_i knows all bits of z except those in positions beloging to R_i. We consider protocols in which the players exchange bits according to the part of the input they know and previous messages. For a function f on r boolean variables let the d-party complexity of f $c_d(f)$ be defined as the smallest number of bits which the players need to exchange in order to compute f on any given input. Clearly $c_d(f)$ depends on the particular partition of $[r]$ used. We shall need the following observation of Håstad and Goldmann [HG]. We sketch the proof for completeness.

Lemma 4 *Suppose a boolean function f can be computed by a depth-2 circuit in which the top gate is any symmetric function and the bottom gates are arbitrary but have fan-in less than d. Then for any given partition of $[r]$ into d disjoint parts $c_d(f) \leq d\log(m)$.*

PROOF : Fix a partition $[r] = R_1 \dot\cup \cdots \dot\cup R_d$. Assign each bottom gate to the first player P_i such that the indices of the variables in this gate do not belong to R_i. This can be done since each bottom gate uses less than d variables and there are d players. Note that P_i can evaluate this gate. Now given an input z each player simply broadcasts the number of gates assigned to him which evaluate to 1. As the top gate is symmetric, this information suffices to compute f. ■

We now define the generalized inner-product function $g_{s,d}$ considered in [BNS]. Suppose $r = sd$ and let $z_{i,j}$ be doubly indexed boolean variables where $i \in [s]$ and $j \in [d]$. Let $g_{s,d} : \{0,1\}^r \to \{0,1\}$ be defined as follows:

$$g_{s,d}(z) = \Sigma_{i=1}^{s} \Pi_{j=1}^{d} z_{i,j}$$

where the sum is calculated modulo 2. We will be interested in communication protocols among d players P_i, for $i \in [d]$, where P_i knows the value of $z_{i',j}$ for all $(i',j) \in [s] \times [d]$ except for $i' = i$. Let $c_d(f)$ denote the d-party communication complexity of a function f relative to this partition. Recall that for two boolean functions f and g on r variables the bias of f and g is defined as:

$$\mathrm{Bias}(f,g) = |\Pr[f(z) = g(z)] - \Pr[f(z) \neq g(z)]|.$$

where $z \in \{0,1\}^r$ is chosen uniformly at random. We shall use the following result of Babai, Nisan, and Szegedy [BNS].

Theorem 1 *Let $r = sd$ and let f be a boolean function on r variables such that $c_d(f) \leq t$. Then $Bias(f, g_{s,d}) \leq 2^t(1 - 4^{1-d})^s$.*

4.2 The pseudo-random generator

We now present a construction of a pseudo-random generator which allows us to approximate the accepting probability of a depth-2 circuit whose top gate is any symmetric boolean function σ.

Suppose we are given such a boolean circuit which computes a function F on n variables together with some $\epsilon > 0$. We are interested in approximating within ϵ the probability μ that on a random input F outputs 1. We think of the first level of the circuit as computing monomials $T_1, \ldots, T_m$ over (possibly negated) boolean variables $x_1, \ldots, x_n$ and of the top level as computing $F = \sigma(T_1, \ldots, T_m)$. Let C_i denote the set of $j \in [n]$ such that x_j or its negation appears in T_i.

Note that for every $i \in [m]$ if C_i is larger than $\log(2m/\epsilon)$ and we replace T_i by the function which is constantly equal to 0 we change the output of the circuit on at most $\frac{\epsilon}{2m}$ of the inputs. Since the number of such i is at most m by replacing T_i in this way for each of them we obtain a new function which differs from the old one on at most $\epsilon/2$ fraction of the inputs and is computed by a similar circuit as the old one. Any $\epsilon/2$-approximation of the accepting probability of the new circuit would then be an ϵ-approximation of the accepting probability of the old one. Thus, by changing ϵ to $\epsilon/2$ if necessary we may assume without loss of generality that the size of C_i is at most $k = \log(2m/\epsilon)$, for each i.

Let $\mathcal{C} = \{C_1, \ldots, C_m\}$. We fix positive integers d, l and r such that there exists an (n, l, d, r)-design $\mathcal{S} = \{S_1, \ldots, S_n\}$ for $\mathcal{C}$. The key ingredient to our generator is a function $g : \{0,1\}^r \rightarrow \{0,1\}$ which will typically be hard to compute in a certain model of computation. We define the function $G_{\mathcal{S}} : \{0,1\}^l \rightarrow \{0,1\}^n$ by

$$G_{\mathcal{S}}(y) = \langle g(y_{S_1}), \ldots, g(y_{S_n}))\rangle,$$

where for each $i \in [n]$ y_{S_i} is the string consisting of the first r coordinates of y whose index is in S_i ordered in the natural ordering. $G_{\mathcal{S}}$ is the generator. The quality of $G_{\mathcal{S}}$ with respect to F is measured by

$$\delta(F, G_{\mathcal{S}}) = |\Pr[F(x) = 1] - \Pr[F(G_{\mathcal{S}}(y)) = 1]|$$

where the inputs x to F and y to $G_{\mathcal{S}}$ are chosen uniformly at random. As described in the outline we relate this value to the complexity of approximating the function g. The first part of the proof of the following lemma is based on a similar argument from [NW].

Lemma 5 *If $\delta(F, G_{\mathcal{S}}) \geq \epsilon$ then there exists a boolean function h over r variables which can be computed by a d-party communication protocol for any partition of $[r]$ into d disjoint sets in $2d\log(m)$ rounds and such that $Bias(h, g) \geq \epsilon/n$.*

PROOF : Let $\nu = \Pr[F(G_{\mathcal{S}}(y)) = 1]$. The first part of the analysis is an interpolation technique which is borrowed from cryptography. We describe a sequence $\mathcal{U}_1, \ldots, \mathcal{U}_{n+1}$

of probability distributions on $\{0,1\}^n$. Let y_i, for $i = 1, \ldots, l$ and x_j, for $j = 1, \ldots, n$ be independent uniformly distributed $\{0,1\}$-random variables and define $\mathcal{U}_i$ to be the distribution given by the sequences of random variables $g(y_{S_1}), \ldots, g(y_{S_{i-1}}), x_i, \ldots, x_n$. Let μ_i be the probability that F evaluates to 0 with respect to $\mathcal{U}_i$. Then $\mu_1 = \mu$ and $\mu_{n+1} = \nu$. Assume that $|\mu - \nu| > \epsilon$ and without loss of generality that $\mu > \nu$. Then there is some i such that $\mu_i - \mu_{i+1} > \epsilon/n$. Fix such an i and define a function h for predicting the output of g as follows.

For fixed values $y_j = b_j$, for $j \in [l] \setminus S_i$ and $x_j = a_j$, for $j > i$ the value of $F(g(y_{S_1}), \ldots, g(y_{S_{i-1}}), x_i, \ldots, x_n)$ depends only on y_{S_i} and x_i. Let us denote this function by $H_{a,b}$. By an averaging argument we show that there are values $\bar{a}$ and $\bar{b}$ such that

$$\Pr[H_{\bar{a},\bar{b}}(y_{S_i}, x_i) = 0] - \Pr[H_{\bar{a},\bar{b}}(y_{S_i}, g(y_{S_i})) = 0] > \epsilon/n.$$

Fix such values $\bar{a}$ and $\bar{b}$. Given y_{S_i} compute $H_{\bar{a},\bar{b}}(y_{S_i}, 0)$ and $H_{\bar{a},\bar{b}}(y_{S_i}, 1)$. If they are equal let $h(y_{S_i})$ be some fixed bit $\tau \in \{0,1\}$. We choose τ which gives the correct value of $g(y_{S_i})$ for at least $1/2$ of such y_{S_i}. If $H_{\bar{a},\bar{b}}(y_{S_i}, 0)$ and $H_{\bar{a},\bar{b}}(y_{S_i}, 1)$ are not equal let $h(y_{S_i})$ be $\rho \in \{0,1\}$ such that $H_{\bar{a},\bar{b}}(y_{S_i}, \rho) = 1$. One can then easily show that such h predicts the value of g with advantage at least ϵ/n.

Let us now verify that h can be computed by a d-communication protocol in $2d\log(m)$ rounds for any partition of $[r]$ into d disjoint sets. This is where we use the fact that $\mathcal{S}$ is an (n,l,d,r)-design for $\mathcal{C}$. Note that $H_{\bar{a},\bar{b}}(y_{S_i}, 0)$ can be written as $\sigma(T_1^0(y_{S_i}), \ldots, T_m^0(y_{S_i}))$ where $T_j^0(y_{S_i})$ is obtained from T_j by making the appropriate substitutions and fixing of variables. Now $T_j^0(y_{S_i})$ depends only on the variables whose index is in $S_i \cap \cup_{i' \in C_j \setminus \{i\}} S_{i'}$ and since $\mathcal{S}$ is an (n,l,d,r)-design for $\mathcal{C}$ this set has size less than d. Thus we are in a position to apply Lemma 4 to conclude that $H_{\bar{a},\bar{b}}(y_{S_i}, 0)$ can be computed by a d-party communication protocol in at most $d\log(m)$ rounds. Similarly one can compute $H_{\bar{a},\bar{b}}(y_{S_i}, 1)$ in another $d\log(m)$ rounds. Computing $h(y_{S_i})$ from these two values does not require any more communication. Thus the total number of communication bits needed to compute $h(y_{S_i})$ is at most $2d\log(m)$. ■

let $\mathcal{F}_{n,m}$ denote the collection of all depth-2 boolean circuits on n variables whose first level consists of at most m conjunctions and whose top level is a fixed but arbitrary symmetric function of these conjunctions. The following theorem is the main result of this paper.

Theorem 2 *Let n and m be positive integers and let $\epsilon > 0$. Then there is an integer $t = \exp(O(\sqrt{\log(nm/\epsilon)}))$ and an explicit generator $G : \{0,1\}^t \to \{0,1\}^n$ such that, for any $F \in \mathcal{F}_{n,m}$, $\delta(F,G) < \epsilon$.*

PROOF : We assemble the pieces of the construction. Let $k = \log(3nm/\epsilon)$, $d = \sqrt{\log(3nm/\epsilon)}$, $r = 2^{4d}$, and $l = k2^{9d}$. Now using Proposition 3 fix $t' = O(d\log(nl))$ such that over $\{0,1\}^{t'}$ there exist explicitly defined nl biased $\{0,1\}$-valued random variables $X_{i,j}$ for $(i,j) \in [n] \times [l]$, which are $2d$-wise independent and such that $X_{i,j}$ equals 1 with probability $2r/l$.

For $w \in \{0,1\}^{t'}$ let $\mathcal{S}(w) = \{S_1(w), \ldots, S_n(w)\}$ where $S_i(w) = \{j : X_{i,j}(w) = 1\}$. Let $s = r/d$ and let $g_{s,d}$ denote the generalized inner product function as defined in Section 4.1. Let $t = t' + l$ and define a generator $G : \{0,1\}^t \to \{0,1\}^n$ as follows. Think of an input to G as a pair (w, y) where $w \in \{0,1\}^{t'}$ and $y \in \{0,1\}^l$. Then let:

$$G(w,y) = (g_{s,d}(y_{S_1(w)}), \ldots, g_{s,d}(y_{S_n(w)})).$$

Here, as before, y_S denotes the projection of y to the first r coordinates of S if S has size at least r and otherwise equals 0^r.

Given a depth-2 circuit $F \in \mathcal{F}_{n,m}$ we show that $\delta(F, G) < \epsilon$. Suppose the first level of F consists of conjunctions $T_1, \ldots, T_m$ and let C_i be the set of variables on which T_i depends. Let F' be the circuit obtained from F by replacing T_i by 0 for every i such that C_i has size more than k. Then clearly

$$|\Pr[F') = 1] - \Pr[F = 1]| \leq \epsilon/3.$$

Let $\mathcal{C} = \{C_i : i \in [m] \text{ and } |C_i| \leq k\}$. Then by Corollary 2, part (b), $\mathcal{S}(w)$ is an (n,l,d,r)-design for $\mathcal{C}$ with probability at least $1 - \epsilon/3$.

Fix $w = w_0$ such that $\mathcal{S}(w_0)$ is an (n,l,d,r)-design for $\mathcal{C}$ and let $G_{\mathcal{S}(w)} : \{0,1\}^l \to \{0,1\}^n$ be the induced generator. We claim that $\delta(F', G_{\mathcal{S}(w_0)}) < \epsilon/3$. Otherwise by Lemma 5 there is a boolean function $h : \{0,1\}^r \to \{0,1\}$ which can be computed by a d-party communication protocol for any partition of $[r]$ into d parts in $2d\log(m)$ rounds such that $\text{Bias}(h, g_{s,d}) \geq \frac{\epsilon}{3n}$. On the other hand by Theorem 1 it follows that for such h $\text{Bias}(h, g_{s,d}) \leq m^{2d}(1 - 4^{1-d})^s < \frac{\epsilon}{3n}$, a contradiction.

Putting all of these pieces together we conclude that $\delta(F, G) < \epsilon$, as desired. ■

Corollary 6 *There is a deterministic algorithm A which given integers n, m, and $\epsilon > 0$, and $F \in \mathcal{F}_{n,m}$ outputs Y such that $|\Pr[F = 0] - Y| \leq \epsilon$. The running time of the algorithm is* $\exp(\exp(O(\sqrt{\log(nm/\epsilon)})))$.

5 Approximating DNF formulas

In this section we present a deterministic algorithm A which on input a DNF formula F and ϵ outputs a real number Y such that $|\Pr[F = 0] - Y| \leq \epsilon$. The running time of the algorithm is $\exp(O(\log(n/\epsilon)^4))$ where n is the size of the formula. The algorithm is a slight modification of the algorithm presented in [NW] but uses our refined notion of (n,l,d,r)-design for a given (m,n,k)-system $\mathcal{C}$. The running time is better than the running time of the algorithm presented in [LV] for formulas which have unrestricted clause length. Since the arguments in this case are very similar to the ones in [NW] or in §4 we shall be sketchy in our description. As in [NW] we use the following lower bound result of Boppana and Håstad [H, Chapter 8].

Theorem 3 *Let F be any DNF formula on r boolean variables with clauses of size at most b. Let Σ_r denote the XOR function on r variables. Then $Bias(F, \Sigma_r) \leq 2^{-r/b}$.*

For integers n and m let $\mathcal{D}_{n,m}$ denote the collection of all DNF formulas on n boolean variables with m clauses.

Theorem 4 *For every integers n, m and every $\epsilon > 0$ there is $t = O(\log(nm/\epsilon)^4)$ and an explicit generator $G : \{0,1\}^t \rightarrow \{0,1\}^n$ such that, for every $F \in \mathcal{D}_{n,m}$, $\delta(F, G) \leq \epsilon$.*

PROOF: Let $d = k = \log(3nm/\epsilon)$, $r = d^2$ and $l = 24kd^3$. Use Proposition 3 to find $t' = O(\log(3nm/\epsilon)^4)$ and explicitly constructed $2d$-wise independent $\{0,1\}$-valued random variables $X_{i,j}$, for $(i,j) \in [n] \times [m]$ such that $X_{i,j}$ is equal to 1 with probability $2r/l$. For $w \in \{0,1\}^{t'}$ let $\mathcal{S}(w) = \{S_1(w), \ldots, S_n(w)\}$, where $S_i(w) = \{j : X_{i,j}(w) = 1\}$. Then by Corollary 2 (b) for any given (m,n,k)-system $\mathcal{C}$ $\mathcal{S}(w)$ is an (n,l,d,r)-design for $\mathcal{C}$ for all but at most $\epsilon/3$ fraction of w's. Let Σ_r denote the XOR function on r boolean variables. Let $t = t' + l$ and define a genarator $G : \{0,1\}^t \rightarrow \{0,1\}^n$ as follows. Represent an input to G as a pair (w, y) where $w \in \{0,1\}^{t'}$ and $y \in \{0,1\}^l$. Then let:

$$G(w,y) = (\Sigma_r(y_{S_1(w)}), \ldots, \Sigma_r(y_{S_n(w)})).$$

Again y_S represents the projection of y to the first r coordinates of S if S has size at least r and otherwise equals 0^r.

The analysis now proceeds as in [NW] and uses Theorem 3 above. ■

Corollary 7 *There is a deterministic algorithm A which given integers n, m, and a real $\epsilon > 0$, and $F \in \mathcal{D}_{n,m}$ outputs Y such that $|\Pr[F = 0] - Y| \leq \epsilon$. The running time of the algorithm is $\exp(O(\log(nm/\epsilon)^4))$.*

6 Absolute versus relative error

All the approximation algorithms we have presented in this paper deal with absolute error approximation. They can easily be converted to relative error approximation algorithms, i.e. algorithms that output Y such that $|\Pr[F = 0] - Y| \leq \epsilon \Pr[F = 0]$, using the reduction techniques presented in [KL1] for DNF formulas and in [KL2] for polynomials over $GF[2]$. We leave the details to the interested reader.

References

[ABI] N. Alon, L. Babai, A. Itai, "A Fast and Simple Randomized Parallel Algorithm for the Maximal Independent Set Problem", *Journal of Algorithms*, 7, pp. 567–583, 1986.

[AGHP] N. Alon, O. Goldreich, J. Håstad, R. Peralta, "Simple constructions of almost k-wise independent random variables", *FOCS 1990*.

[BNS] Babai, L, Nisan, N., Szegedy, M., "Multiparty protocols and logspace-hard pseudo-random sequences", preliminary version in *STOC*, 1989, pp. 1-11, journal version in *JCSS*, Vol. 45, No. 2, October 1992, pp. 204-232.

[CFL] Chandra, Furst, and Lipton, "Multiparty protocols, *FOCS 1983*.

[EGLNV] G. Evan, O. Goldreich, M. Luby, N. Nisan, B. Veličković, "Approximation of general independent distributions", *STOC 1992*, pp. 10-16.

[H] J. Håstad, "Computational limitations for small depth circuits", Ph.D. thesis, *MIT press*, 1986.

[HG] J. Håstad and M. Goldmann, "On the power of small depth threshold circuits", *FOCS 1990*, pp. 610-618.

[KL1] R. Karp and M. Luby, "Monte-Carlo algorithms for enumeration and reliability problems", 24^{th} *STOC*, 1983, pp. 54-63.

[KL2] M. Karpinski and M. Luby, "Approximating the Number of Solutions to a *GF[2]* Formula," a preliminary version appears in the proceedings of *Second Annual Symposium on Discrete Algorithms*, 1991, pp. 300-303, final version accepted to *Journal of Algorithms*.

[NN] M. Naor and S. Naor, "Small Bias Probability Spaces: Efficient Constructions and Applications", *STOC 90*.

[NW] N. Nisan and A. Widgerson, "Hardness vs. Randomness", *FOCS 88*, pp. 2-11.

[LV] M. Luby and B. Veličković, "On Deterministic Approximation of DNF", *STOC 1991*, pp. 430-438.

[R] A. A. Raborov, "Lower bounds on the size of dounded depth circuits in basis $\{\&, \wedge, \oplus\}$, (in Russian), *Usp. Mat. Nauk,* vol. *41*:4, 1986, pp. 219-220.

[S] R. Smolensky, "Algebraic mathods in the theory of lower bounds for boolean circuit complexity", *STOC 1987*, pp. 77-82.