The Parity of Set Systems under Random Restrictions with Applications to Exponential Time Problems

Andreas Björklund[1] Holger Dell[2] Thore Husfeldt[3]

September 16, 2015

Abstract

We reduce the problem of detecting the existence of an object to the problem of computing the parity of the number of objects in question. In particular, when given any non-empty set system, we prove that randomly restricting elements of its ground set makes the size of the restricted set system an odd number with significant probability. When compared to previously known reductions of this type, ours excel in their simplicity: For graph problems, restricting elements of the ground set usually corresponds to simple deletion and contraction operations, which can be encoded efficiently in most problems. We find three applications of our reductions:

1. An exponential-time algorithm: We show how to decide Hamiltonicity in directed n-vertex graphs with running time 1.9999^n provided that the graph has at most 1.0385^n Hamiltonian cycles. We do so by reducing to the algorithm of Björklund and Husfeldt (FOCS 2013) that computes the parity of the number of Hamiltonian cycles in time 1.619^n.
2. A new result in the framework of Cygan et al. (CCC 2012) for analyzing the complexity of NP-hard problems under the Strong Exponential Time Hypothesis: If the parity of the number of Set Covers can be determined in time 1.9999^n, then Set Cover can be decided in the same time.
3. A structural result in parameterized complexity: We define the parameterized complexity class $\oplus$W[1] and prove that it is at least as hard as W[1] under randomized fpt-reductions with bounded one-sided error; this is analogous to the classical result NP $\subseteq$ RP$^{\oplus \mathrm{P}}$ by Toda (SICOMP 1991).

1. Lund University, Sweden
2. Saarland University and Cluster of Excellence (MMCI), Germany; Simons Institute and UC Berkeley
3. Lund University, Sweden and IT University of Copenhagen, Denmark

1 Introduction

A set family $\mathscr{F}$ with an odd number of elements is of course nonempty. In the present paper we study randomized reductions where the opposite holds with significant probability: We reduce the *decision problem* of determining if $|\mathscr{F}|$ is non-zero to the *parity problem* of determining if $|\mathscr{F}|$ is odd. Originally such decision-to-parity reductions were obtained as corollaries to various "isolation lemmas," such as the one of Valiant and Vazirani (1986), where the reduction is to the *unambiguous problem* of distinguishing between $|\mathscr{F}| = 0$ and $|\mathscr{F}| = 1$. Our decision-to-parity reduction is not a reduction to the unambiguous problem, and it has a *much* simpler structure than existing isolation lemmas in that it computes random restrictions of the universe. In our applications, such restrictions simply correspond to random deletions or contractions of the edges.

Organization. In §1.1, we state the main lemma of this paper and discuss its relationship with and consequences for probabilistic polynomial identity tests as well as various isolation lemmas. We prove the Main Lemma in §2, and we discuss its applications for Hamiltonicity in §3, for Set Cover in §4, and for W[1] in §5. We complete this paper in §6 by proving that our decision-to-parity reductions are optimal in a certain black-box model of restriction-based reductions.

1.1 Set Systems under Random Reductions

Let $\mathscr{F}$ denote a family of sets. We present our reductions in a general combinatorial setting, but for the sake of concreteness we invite the reader to think of $\mathscr{F}$ as the family of all vertex subsets forming a k-clique, or the family of all edge subsets forming a Hamiltonian cycle. For instance, in the house graph in Fig. 1, the family $\{\{1,2,3,4,7\},\{1,3,4,6,8\}\}$ corresponds to the Hamiltonian cycles.

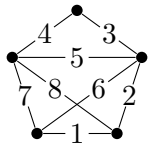

Fig. 1

Let U be the ground set of $\mathscr{F}$, that is, $\mathscr{F} \subseteq 2^U$. A *restriction* is a function

$$\rho\colon U \to \{0,1,*\}\,.$$

The *restricted family* $\mathscr{F}\restriction_\rho$ consists of all sets $F \in \mathscr{F}$ that satisfy $i \in F$ for all i with $\rho(i) = 1$ and $i \notin F$ for all i with $\rho(i) = 0$. A *random restriction* is a distribution over restrictions ρ where $\rho(i)$ is randomly sampled for each i independently subject to $\Pr_\rho(\rho(i) = 0) = p_0$ and $\Pr_\rho(\rho(i) = 1) = p_1$. We define $p_* = 1 - (p_0 + p_1)$. We are interested in the event that the number of sets in the restricted family $\mathscr{F}\restriction_\rho$ is odd, and we write this event as $\oplus\mathscr{F}\restriction_\rho$.

Lemma 1 (Main Lemma). *Let $\mathscr{F}$ be a nonempty family of sets over a universe of size at most n, and let each set have size at most k. Let ρ denote a random restriction with the parameters p_0, p_1, and p_*.*

(i) If $p_0 \geq p_*$, then

$$\Pr_\rho\Big(\oplus\mathscr{F}\restriction_\rho\Big) \geq (1-p_1)^{n-2k} \cdot p_*^k\,. \tag{1}$$

(ii) If $p_0 < p_*$, then

$$\Pr_\rho\Big(\oplus\mathscr{F}\restriction_\rho\Big) \geq (1-p_1)^{n-2k} \cdot \left(\frac{p_0}{p_*}\right)^{\min(k,\log|\mathscr{F}|)} \cdot p_*^k\,. \tag{2}$$

All of our applications are based on random restrictions with $p_1 = 0$; in this case, the success probabilities do not depend on the size of the underlying ground set.

Examples. Consider the graph of Fig. 1, where $|U| = 8$, $k = 5$, and $|\mathscr{F}| = 2$, and assume $p_1 = 0$. The restriction ρ results in an odd number of Hamiltonian cycles exactly if $\rho(1) = \rho(3) = \rho(4) = *$ and either $\rho(2) = \rho(7) = *$ or $\rho(6) = \rho(8) = *$ (but not both). For $p_0 = \frac{1}{2}$ this happens with probability $\frac{12}{256} = \frac{3}{64}$, slightly better than the bound $\frac{1}{32}$ promised by (1). If we set $p_0 = \frac{1}{5}$ then (2) promises the better bound $\Pr_\rho(\oplus\mathscr{F}\restriction_\rho) = (\frac{4}{5})^5 \cdot \frac{1}{4} = \frac{256}{3125} \geq 0.081$. For completeness, direct calculation shows that $\Pr_\rho(\oplus\mathscr{F}\restriction_\rho) = 4 \cdot (\frac{4}{5})^6 \cdot \frac{1}{5} + 2 \cdot (\frac{4}{5})^5 \cdot (\frac{1}{5})^2 = \frac{18432}{78125} \geq 0.235$, so the bound is far from tight in this example.

A simple example that attains (1) with equality is the singleton family $\mathscr{F}$ consisting only of the set $\{1,\ldots,k\}$. Then one easily computes $\Pr_\rho(\oplus\mathscr{F}\restriction_\rho) = \Pr_\rho(\rho(1) = \cdots = \rho(k) = *) = (1-p_0)^k$. For an example attaining (2) with equality, consider the family $\mathscr{F}$ of sets F satisfying $\{1,\ldots,(1-\epsilon)k\} \subseteq F \subseteq \{1,\ldots,k\}$, where $0 < \epsilon \leq \frac{1}{2}$ holds and ϵk is an integer. Then $|\mathscr{F}| = 2^{\epsilon k}$. There is but one restriction ρ for which the event $\oplus\mathscr{F}\restriction_\rho$ happens, namely when $\rho(i) \neq 0$ for all $i \leq (1-\epsilon)k$ and $\rho(i) = 0$ for all $i > (1-\epsilon)k$. Thus, with $p_0 = \epsilon$ we have

$$\Pr_\rho(\oplus\mathscr{F}\restriction_\rho) = (1-p_0)^{(1-\epsilon)k} p_0^{\epsilon k} = (1-p_0)^k \left(\frac{p_0}{1-p_0}\right)^{\log|\mathscr{F}|}.$$

Connection with Probabilistic Polynomial Identity Tests. The Main Lemma with $p_1 = 0$ can be expressed in terms of polynomials over finite fields instead of restricted set systems by considering the nonempty set system $\mathscr{F}$ as the nonzero polynomial

$$p(X_1,\ldots,X_n) = \sum_{F\in\mathscr{F}} \prod_{i\in F} X_i$$

in the polynomial ring $\mathrm{GF}(2)[X_1,\ldots,X_n]$. Let $x_1,\ldots,x_n \in \mathrm{GF}(2)$ be chosen independently and uniformly at random. The Main Lemma implies

$$\Pr_{x_1,\ldots,x_n}\Big(p(x_1,\ldots,x_n) = 0\Big) \leq 1 - 2^{-k}\,,$$

where k is the total degree of p; since p is multilinear, k corresponds to the maximum number of variables occurring in a monomial.

Thus, our Main Lemma can be understood as a variant of the well-known probabilistic polynomial identity test of DeMillo and Lipton (1978), Schwartz (1980), and Zippel (1979) (cf. Arora and Barak 2009, Lemma 7.5). In its standard form, that lemma bounds the probability by $k/2$, where the 2 stems from the size of the finite field GF(2); we usually have $k/2 \geq 1$ and so the bound is vacuous. Nevertheless, variants of the lemma for small finite fields have been studied. In particular, the basic form of the Main Lemma where $p_0 = \frac{1}{2}$ and $p_1 = 0$ appears in Cohen and Tal (2013, Lemma 2.2) and Vassilevska Williams et al. (2015, Lemma 2.2.). For smaller values of p_0, the Main Lemma yields as a corollary the following probabilistic polynomial identity test, which may be new and of independent interest; it applies to *sparse* polynomials over GF(2).

Corollary 2. *Let p be a non-zero polynomial over* GF(2) *in the variables $X_1, \ldots, X_n$, with total degree* $\deg(f)$ *and at most $2^{\epsilon \deg(f)}$ monomials. Let $x_1, \ldots, x_n \in \{0,1\}$ be sampled from the distribution where* $\Pr(x_i = 0) = \epsilon \leq \frac{1}{2}$ *holds for each i independently. Then*

$$\Pr_{x_1,\ldots,x_n}\Big(p(x_1,\ldots,x_n) = 0\Big) \leq 1 - 2^{-H(\epsilon)\deg(f)}\,.$$

Comparison to isolation lemmas based on linear equations. In their seminal paper, Valiant and Vazirani (1986) prove an isolation lemma that can be described for non-empty set systems $\mathscr{F}$ over a ground set U of size n as follows: Suppose we know $s = \log|\mathscr{F}|$ for some s. Then we sample a function $h : \{0,1\}^U \to \{0,1\}^s$ at random from a family of pairwise uniform hash functions. We interpret h as mapping subsets of U to vectors in $\{0,1\}^s$. We define the restricted family $\mathscr{F}_{h=0}$ as

$$\mathscr{F}_{h=0} = \Big\{\, F \in \mathscr{F} \;:\; h(F) = (0,\ldots,0) \,\Big\}\,.$$

Valiant and Vazirani (1986) prove that $\mathscr{F}_{h=0}$ has exactly one element with probability at least $\frac{1}{4}$. Since the cardinality of $\mathscr{F}$ is not known, the value of s must be guessed at random from $\{1,\ldots,n\}$, and the success probability for the whole construction becomes $\Omega(1/n)$. In particular,

$$\Pr_h(\oplus\mathscr{F}_{h=0}) \geq \Omega(\tfrac{1}{n})\,.$$

The procedure we just described is useful for problems that are sufficiently rich to express the condition $h(F) = 0$. In particular, the set of all affine linear functions $h : \mathrm{GF}(2)^n \to \mathrm{GF}(2)^s$ is often used as the family of hash functions; these functions have the form $h(x) = Ax + b$ for some suitable matrix A and vector b over GF(2). Thus the condition $h(F) = 0$ becomes a set of linear equations over GF(2), which can be expressed as a polynomial-size Boolean formula – in fact most natural NP-complete problems are able to express linear constraints with only a polynomial overhead in instance size.

In the exponential time setting, we cannot afford such polynomial blow-up and many problems, including the satisfiability of k-CNF formulas, are not known to be able to

efficiently express arbitrary linear constraints. Nevertheless, Calabro et al. (2003) are able to design an isolation lemma for k-CNF satisfiability, essentially by considering *sparse* linear equation systems, that is, systems where each equation depends only on k variables. Things seem to get even worse for problems such as Set Cover, where we are unable to efficiently express sparse linear equations. This is where our random restrictions come into play since they are much simpler than linear equations; in terms of CNF formulas, they correspond to adding singleton clauses like (x_i) or $(\neg x_i)$.

Neglecting, for a moment, the fact that we may be unable to express the necessary constraints, let us compare the guarantees of Valiant and Vazirani (1986) and the Main Lemma: we only achieve oddness instead of isolation, but we do so with probability 2^{-k} instead of $\Omega(\frac{1}{n})$ — our probability is better if $n \geq 2^k$.

Comparison to isolation lemmas based on minimizing weight. Another isolation lemma for k-CNF satisfiability suitable for the exponential-time setting is due to Traxler (2008) and is based on the isolation lemma of Mulmuley, Vazirani, and Vazirani (1987). Their construction associates random weights $w(x) \in \{1, \ldots, 2|U|\}$ with each element in the ground set. One then considers for each $r \in \{0, \ldots, 2k|U|\}$ the subfamily of sets of weight exactly r, formally defined as

$$\mathscr{F}_{w,r} = \left\{ F \in \mathscr{F} \;:\; \textstyle\sum_{x \in F} w(x) = r \right\}.$$

The isolation lemma of Mulmuley, Vazirani, and Vazirani (1987) says that there is a unique set $F \in \mathscr{F}$ of minimum weight r with probability at least $\frac{1}{2}$. In particular, for this $r = r(\mathscr{F}, w)$ we have $\Pr_w\big(\oplus \mathscr{F}_{w,r} \;\big|\; r = r(\mathscr{F}, w) \big) \geq \frac{1}{2}$. Since r is not known, we sample it uniformly at random, which yields the overall success probability

$$\Pr_{w,r}(\oplus \mathscr{F}_{w,r}) \geq \Omega(\tfrac{1}{kn}).$$

The difficulty with this approach is that, when the weighted instance of, say, Set Cover is translated back to an unweighted instance, the parameters are not preserved because the weights are taken from a set of nonconstant size. On the other hand, the weights 0 and 1 can be expressed in many problems as simple deletions or contractions.

We can view the Main Lemma in the weight-minimization framework as follows: sample random weights $w(x) \in \{0, 1\}$ independently for each x such that $w(x) = 0$ holds with probability p_0, and define the weight of $F \in \mathscr{F}$ as $\prod_{x \in F} w(x)$; by taking the logarithm, we note that minimizing the product is identical to minimizing the sum. The Main Lemma yields a lower bound on the probability that the number of sets with nonzero weight is odd. For comparison with Traxler (2008), note that we only achieve oddness instead of isolation, but we do so with probability 2^{-k} instead of $\Omega(\frac{1}{kn})$, which is much better when k is small.

Other parity lemmas and optimality. Not all decision-to-parity reductions are based on an isolation procedure: Naik, Regan, and Sivakumar (1995) use a small-bias sample space to design a randomized polynomial-time procedure that maps any Boolean formula F,

whose set of satisfying assignments corresponds to a set family $\mathscr{F}$, to a formula F', whose family of satisfying assignments $\mathscr{F}'$ is a subfamily of $\mathscr{F}$; the guarantee is that, if $\mathscr{F}$ is not empty, then $\Pr(\oplus\mathscr{F}') \geq \frac{1}{2} - \epsilon$; they achieve such an algorithm for any constant $\epsilon > 0$.

The constraints in the construction of Naik, Regan, and Sivakumar (1995) are linear equations, which we do not know how to encode into less expressive problems such as Set Cover. On the other hand, restrictions of families often correspond to contractions or deletions, which are typically easy to express. Nevertheless, the success probability of Naik, Regan, and Sivakumar (1995) is much better than the one guaranteed by the Main Lemma, and one may wonder whether this is an artifact of our proof. Alas, we prove in §6 that this is not the case: no decision-to-parity reduction that is based on random restrictions can have a better success probability than what is achieved by the Main Lemma.

2 Proof of the Main Lemma

We bootstrap the main lemma from the following fine-grained variant of the DeMillo–Lipton–Schwartz–Zippel lemma for multilinear polynomials over GF(2).

Lemma 3 (Fine-grained DeMillo–Lipton–Schwartz–Zippel).
Let f be a non-zero multilinear polynomial in the variables $X_1, \ldots, X_n$ over GF(2). *Let* $\deg(f)$ *be the maximum degree of the polynomial and $|f|$ be the number of monomials. Let $q_0, q_1 \in [0,1]$ so that $q_0 + q_1 = 1$. In the following, we sample $x = x_1 \ldots x_n$ from $\{0,1\}^n$ by setting $x_i = 1$ with probability q_1 for each i independently.*

(i) If $q_0 \leq q_1$, then

$$\Pr_x\Big(f(x) = 1\Big) \geq \left(\frac{q_0}{q_1}\right)^{\min(\deg(f), \log|f|)} \cdot q_1^{\deg(f)}\,.$$

(ii) If $q_0 \geq q_1$, then

$$\Pr_x\Big(f(x) = 1\Big) \geq q_1^{\deg(f)}\,.$$

Proof. The proof is by induction on the number n of variables. If $n = 0$ then $f = 1$, and so the probability is equal to one.

For the induction step, suppose that $n > 0$. Let $A, B \subseteq [n]$ be maximal disjoint sets such that

$$f = g \cdot \prod_{i \in A} X_i \cdot \prod_{i \in B} (1 - X_i)\,, \tag{3}$$

where g is a polynomial in the variables X_i for $i \notin A \cup B$. Since f is a non-zero polynomial, so is g.

In the special case that $g = 1$, we have $\deg(f) = |A| + |B|$ and $|f| = 2^{|B|}$, and so $|B| = \log|f|$ and $|A| = \deg(f) - \log|f|$. Therefore, $f(x) = 1$ holds with probability

exactly $q_1^{|A|} \cdot q_0^{|B|} = q_1^{\deg(f)-\log|f|} q_0^{\log|f|}$. Since $\log|f| \leq \deg(f)$, this is exactly what we claimed for $q_0 \leq q_1$. Moreover, if $q_0 \geq q_1$, we observe that $q_1^{\deg(f)}(q_0/q_1)^{\log|f|} \geq q_1^{\deg(f)}$.

Thus it remains to consider the case where g is not identically one. Without loss of generality, let X_1 be a variable that appears in g. Then g can be decomposed uniquely as $g = X_1 g_1 + (1 - X_1)g_0$ where g_0 and g_1 are polynomials in the variables X_i for $i \notin A \cup B \cup \{1\}$. If g_0 or g_1 were identically zero, then either A or B could have been extended by one; therefore, neither g_0 nor g_1 are identically zero. The decomposition of g transfers to a decomposition $f = X_1 f_1 + (1 - X_1)f_0$ where f_0 and f_1 are non-zero polynomials in the variables $X_2, \ldots, X_n$. We prepare to apply the induction hypothesis by conditioning on the value of $x_1 \in \{0, 1\}$ as follows:

$$\Pr_x\Big(f(x) = 1\Big) = q_1 \cdot \Pr_{x_2,\ldots,x_n}\Big(f_1(x_2, \ldots, x_n) = 1\Big) + q_0 \cdot \Pr_{x_2,\ldots,x_n}\Big(f_0(x_2, \ldots, x_n) = 1\Big)$$

Since f_0 and f_1 have fewer than n variables, the induction hypothesis applies. Note that $\deg(f_0)$ and $\deg(f_1)$ are both at most $\deg(f)$, and $|f_0|$ and $|f_1|$ are both at most $|f|$.

Let us first consider the easier case $q_0 \geq q_1$, where a simple application of the induction hypothesis yields the claim:

$$\Pr_x\Big(f(x) = 1\Big) \geq q_1 \cdot q_1^{\deg(f_1)} + q_0 \cdot q_1^{\deg(f_0)} \geq (q_1 + q_0) q_1^{\deg(f)} = q_1^{\deg(f)}\,.$$

Now assume that $q_0 \leq q_1$. Note that $\min(a', b') \leq \min(a, b)$ holds whenever $a' \leq a$ and $b' \leq b$, and that $(q_0/q_1)^{c'} \geq (q_0/q_1)^c$ holds whenever $c' \leq c$. Therefore,

$$\begin{aligned}
\Pr_x\Big(f(x) = 1\Big) &\geq q_1 \cdot \left(\frac{q_0}{q_1}\right)^{\min(\deg(f_1),\log|f_1|)} \cdot q_1^{\deg(f_1)} \\
&\quad + q_0 \cdot \left(\frac{q_0}{q_1}\right)^{\min(\deg(f_0),\log|f_0|)} \cdot q_1^{\deg(f_0)} \\
&\geq \left(\frac{q_0}{q_1}\right)^{\min(\deg(f),\log|f|)} \cdot q_1^{\deg(f)}\,.
\end{aligned}$$

This finishes the proof of the lemma. ∎

Let $[n] = \{1, \ldots, n\}$. We define the distribution $\mathcal{D}(p_0, p_1, n)$ over the set of all restrictions $\rho : [n] \to \{0, 1, *\}$ as follows: For each $i \in [n]$ independently, we sample $\rho(i)$ at random so that $\rho(i) = b$ holds with probability exactly p_b for $b \in \{0, 1, *\}$ where p_* is defined as $1 - (p_0 + p_1)$.

Lemma 1 (Main Lemma, restated). *Let $\mathscr{F}$ be a non-empty family of subsets of $[n]$, and let k be the size of the largest set in $\mathscr{F}$.*

(i) If $p_0 \geq p_*$, then

$$\Pr_{\rho\sim\mathcal{D}(p_0,p_1,n)}\Big(\oplus\mathscr{F}\restriction_\rho\Big) \geq (1 - p_1)^{n-2k} \cdot p_*^k\,.$$

(ii) If $p_0 < p_*$, then

$$\Pr_{\rho\sim\mathcal{D}(p_0,p_1,n)}\Big(\oplus\mathscr{F}\restriction_\rho\Big) \geq (1-p_1)^{n-2k}\cdot\left(\frac{p_0}{p_*}\right)^{\min(k,\log|\mathscr{F}|)}\cdot p_*^k\,.$$

Proof. We define

$$f^{\mathscr{F}} = \sum_{F\in\mathscr{F}}\prod_{i\in F} X_i\,.$$

Moreover, for a "core" set $C\subseteq\{1,\ldots,n\}$, we further define

$$f^{\mathscr{F},C} = \sum_{\substack{F\in\mathscr{F}\\ F\supseteq C}}\prod_{i\in F\setminus C} X_i\,.$$

Clearly $f^{\mathscr{F},\emptyset} = f^{\mathscr{F}}$, and $f^{\mathscr{F},C}$ is a multilinear polynomial in the variables x_i for $i\in[n]\setminus C$. For any restriction ρ, we set $C=\rho^{-1}(1)$. We now have

$$\Big|\mathscr{F}\restriction_\rho\Big| \bmod 2 = f^{\mathscr{F},C}(x)\,. \tag{4}$$

where for all i with $\rho(i)\neq 1$, we let

$$x_i = \begin{cases} 0 & \text{if } \rho(i)=0, \text{ and}\\ 1 & \text{if } \rho(i)=*\,.\end{cases}$$

To see (4), first note that $f^{\mathscr{F},C}(x)$ is equal to the parity of the set M of monomials $\prod_{i\in I}X_i$ of $f^{\mathscr{F},C}$ such that $x_i=1$ holds for all $i\in I$. Now we establish a bijective mapping $\pi:\mathscr{F}\restriction_\rho\to M$. Let $F\in\mathscr{F}\restriction_\rho$. Then $F\in\mathscr{F}$ and $F\supseteq C$ holds, and so $\pi(F):=\prod_{i\in F\setminus C}X_i$ is a monomial of $f^{\mathscr{F},C}$. Moreover, all $i\in F\setminus C$ satisfy $\rho(i)=*$, and so $x_i=1$. The mapping π is clearly injective, for if $F,F'\supseteq C$ and $F\neq F'$, then $\pi(F)\neq\pi(F')$. It is also surjective: Let $\prod_{i\in I}X_i$ be a monomial in M and let $F=I\cup C$. Since $\rho(i)=*$ holds for all $i\in I$ and $\rho(i)=1$ for all $i\in C$, we have $F\in\mathscr{F}\restriction_\rho$ and $\pi(F)=\prod_{i\in I}X_i$.

Using (4), we reduce to the previous lemma. For this, we first sample $C\subseteq[n]$ by including each $i\in[n]$ in the set C independently with probability p_*. Let $q_0=\frac{p_0}{p_0+p_*}$ and $q_1=\frac{p_*}{p_0+p_*}$; for each $i\in[n]\setminus C$ independently, we next sample $x_i\in\{0,1\}$ so that $x_i=0$ holds with probability exactly q_0. Note that such (C,x) stand in one-to-one correspondence with restrictions ρ, and the resulting distribution is exactly $\mathcal{D}(p_0,p_1,n)$. Thus we have

$$\Pr_\rho\Big(\oplus\mathscr{F}\restriction_\rho\Big) = \mathop{\mathbf{E}}_C\Pr_x\big(f^{\mathscr{F},C}(x)=1\big) = \mathop{\mathbf{E}}_C\Big(\Pr_x\big(f^{\mathscr{F},C}(x)=1\big)\;\Big|\; f^{\mathscr{F},C}\neq 0\Big)\cdot\Pr_C\Big(f^{\mathscr{F},C}\neq 0\Big)\,.$$

Note that $f^{\mathscr{F},C}$ has $n-|C|$ variables, at most $|\mathscr{F}|$ monomials, and degree at most k. Thus the previous lemma implies a lower bound on $\Pr_x\Big(f^{\mathscr{F},C}(x)=1\Big)$ whenever $f^{\mathscr{F},C}$ is not identically zero. We condition on this event, and so it remains to provide a lower bound on the probability that $f^{\mathscr{F},C}$ is not identically zero for random C. Since $\mathscr{F}$ contains a

set F of size k, the probability that $\rho(i) = 1$ holds for all $i \notin F$ is at least $(1-p_1)^{n-k}$. This event implies that the polynomial is not zero. Thus, we obtain the following in the case that $p_0 \leq p_*$:

$$\begin{aligned}\Pr_\rho\Big(\oplus\mathscr{F}\restriction_\rho\Big) &\geq (1-p_1)^{n-k} \cdot \left(\frac{q_0}{q_1}\right)^{\min(k,\log|\mathscr{F}|)} \cdot q_1^k \\ &= (1-p_1)^{n-k} \cdot \left(\frac{p_0}{p_*}\right)^{\min(k,\log|\mathscr{F}|)} \cdot \left(\frac{p_*}{p_0+p_*}\right)^k \\ &= (1-p_1)^{n-2k} \cdot \left(\frac{p_0}{p_*}\right)^{\min(k,\log|\mathscr{F}|)} \cdot p_*^k\,.\end{aligned}$$

Similarly, in the case that $p_0 \geq p_*$, we get:

$$\begin{aligned}\Pr_\rho\Big(\oplus\mathscr{F}\restriction_\rho\Big) &\geq (1-p_1)^{n-k} \cdot q_1^k = (1-p_1)^{n-k} \cdot \left(\frac{p_*}{p_0+p_*}\right)^k \\ &= (1-p_1)^{n-2k} \cdot p_*^k\,.\end{aligned}$$

■

3 Directed Hamiltonicity

The most straightforward algorithmic application of our reductions is to translate a decision problem to its corresponding parity problem. This is useful in case a faster variant is known for the parity version. In the regime of exponential time problems, we currently know a single candidate for this approach: Björklund and Husfeldt (2013) recently found an algorithm that computes the parity of the number of Hamiltonian cycles in a directed n-vertex graph in $O(1.619^n)$ time, but we do not know how to decide Hamiltonicity in directed graphs in time $(2-\Omega(1))^n$. We devise such an algorithm in the special case that the number of Hamiltonian cycles is guaranteed to be small. Let $H\colon [0,1] \to \mathbf{R}$ denote the binary entropy function given by $H(\epsilon) = -(1-\epsilon)\log_2(1-\epsilon) - \epsilon\log_2\epsilon$.

Theorem 4. *For all $\epsilon > 0$, there is a randomized $O(2^{(0.6942+H(\epsilon))n})$ time algorithm to detect a Hamiltonian cycle in a given directed n-vertex graph G with at most $2^{\epsilon n}$ Hamiltonian cycles.*

In particular, if the number of Hamiltonian cycles is known to be bounded by 1.0385^n, we decide Hamiltonicity in time $O(1.9999^n)$.

Discussion and related work. The best time bound currently known for directed Hamiltonicity is $2^n/\exp(\Omega(\sqrt{n/\log n}))$ due to Björklund (2012). In particular, no 1.9999^n algorithm is known. There are no insightful hardness arguments to account for this situation; for instance, there is no lower bound under the Strong Exponential Time Hypothesis. We do know an $O(1.657^n)$ time algorithm for Hamiltonicity detection in undirected graphs (Björklund 2014) and an $O(1.888^n)$ time algorithm for bipartite directed graphs (Cygan,

Kratsch, and Nederlof 2013). The existence of a $(2-\Omega(1))^n$ algorithm for the general case is currently an open question.

Is Theorem 4 further evidence for a $(2-\Omega(1))^n$ time algorithm for directed Hamiltonicity? We are undecided about this. For a counterargument, consider another problem where a restriction of the solution set leads to a $(2-\Omega(1))^n$ time algorithm, without making the general case seem easier: Counting the number of perfect matchings in a bipartite $2n$-vertex graph. It is not known how to solve the general problem faster than $2^n/\exp(\Omega(\sqrt{n/\log n}))$, but when there are not too many matchings, they can be counted in time $(2-\Omega(1))^n$ (Björklund, Husfeldt, and Lyckberg 2015).

We remark that when the input graph is bipartite, we could reduce to the faster parity algorithm of Björklund and Husfeldt (2013), which runs in time $1.5^n\,\mathrm{poly}(n)$. For this class of graphs, our constructions imply that there is a randomized algorithm to detect a Hamiltonian cycle in time $O(2^{(0.5848+H(\epsilon))n})$ if the input graph has at most $2^{\epsilon n}$ Hamiltonian cycles. In particular, if the number of Hamiltonian cycles is at most $O(1.0431^n)$, the resulting bound is better than the bound $O(1.888^n)$ of Cygan, Kratsch, and Nederlof (2013). Similarly, for the undirected (non-bipartite) case, we can beat the $O(1.657^n)$ bound of Björklund (2014) for the undirected case for instances with at most $O(1.0024^n)$ cycles.

In summary, detecting a Hamiltonian cycle seems to become easier when we know that there are few of them. Currently, this result appears to be the most interesting application of the Main Lemma. However, it is unclear if future work on Hamiltonicity will prove it to be a central linchpin in our final understanding, or render it completely useless—it could still turn out that the decision problem in the general case is *easier* than the parity problem.

Proof. We now prove Theorem 4. We begin with a simple corollary expressing the second part of our Main lemma in terms of the binary entropy function.

Corollary 5. *Let $\mathscr{F}$ be a nonempty family of sets, each of size at most k. Assume $|\mathscr{F}| \le 2^{\epsilon k}$ holds for some ϵ with $0 < \epsilon \le \frac{1}{2}$. Let ρ denote a random restriction with $p_0 = \epsilon$ and $p_1 = 0$. Then*

$$\Pr_\rho(\oplus\mathscr{F}\restriction_\rho) \ge 2^{-H(\epsilon)k}\,. \tag{5}$$

Proof. By (2) of the Main Lemma, we have

$$\Pr_\rho(\oplus\mathscr{F}\restriction_\rho) \ge (1-\epsilon)^k\left(\frac{\epsilon}{1-\epsilon}\right)^{\epsilon k} = 2^{-H(\epsilon)k}\,. \qquad \blacksquare$$

Our algorithm to reduce from Hamiltonicity to its parity version is very simple:

Algorithm H (*Hamiltonicity*) *Given a directed graph on n vertices with arc set A, this algorithm decides whether the graph contains a Hamiltonian cycle.*

H1 (Remove arcs at random.) Construct the arc subset A' by removing each $a \in A$ independently at random with probability $p_0 = \log|\mathscr{F}|/k$.

H2 (Compute parity.) Return "yes" if the algorithm of Björklund and Husfeldt (2013) determines that the parity of the number of Hamiltonian cycles of the graph given by A' is odd. Otherwise return "no."

Proof (of Theorem 4). The running time of H is dominated by the call to the algorithm of Björklund and Husfeldt (2013), which runs in time within a polynomial factor of $1.619^n \leq 2^{0.6942n}$. To compute the success probability of H, let $\mathscr{F}$ be the family of all subsets of A that form a directed Hamiltonian cycle and assume that $|\mathscr{F}| \leq 2^{\epsilon n}$ for some ϵ with $0 \leq \epsilon \leq \frac{1}{2}$. By identifying arc subsets A' with restrictions ρ in the canonical way, we see that $\mathscr{F}\restriction_\rho$ is the family of directed Hamiltonian cycles in the subgraph given by A'. The algorithm is successful if and only if $\oplus\mathscr{F}\restriction_\rho$ holds. Thus, by Corollary 5 with $k = n$, the success probability is at least $2^{-H(\epsilon)n}$. Finally, we amplify this to a constant by repeating the algorithm $2^{H(\epsilon)n}$ times. ∎

4 Decision-to-Parity Reductions for Set Cover and Hitting Set

For Set Cover and Hitting Set, we establish a strong connection between the parity and decision versions, namely that computing the parity of the number of solutions cannot be much easier than finding one.

Consider as input a family $\mathscr{F}$ of m subsets of some universe U with n elements. A subfamily $\mathscr{C} \subseteq \mathscr{F}$ is *covering* if the union of all $C \in \mathscr{C}$ equals U. The Set Cover problem is given a set family $\mathscr{F}$ and a positive integer t to decide if there is a covering subfamily with at most t sets. The problem's parity analogue $\oplus$ Set Covers is to determine the parity of the number covering subfamilies with at most t sets.

Dually, a set $H \subseteq U$ is a *hitting set* if H intersects F for every $F \in \mathscr{F}$. The Hitting Set problem is given a set family $\mathscr{F}$ and a positive integer t to decide if there exists a hitting set of size at most t. The parity analogue $\oplus$ Hitting Sets is to determine the parity of the number of hitting sets of size at most t. We prove the following theorem.

Theorem 6. *Let $c \geq 1$.*

(i) If $\oplus$ Set Covers *can be solved in time $d^n \cdot \mathrm{poly}(n+m)$ for all $d > c$, then the same is true for* Set Cover.

(ii) If $\oplus$ Hittings Sets *can be solved in time $d^m \cdot \mathrm{poly}(n+m)$ for all $d > c$, then the same is true for* Hitting Set.

Discussion and related work. Theorem 6 should be understood in the framework of Cygan et al. (2012), where it establishes a new reduction in their network of reductions. Our results are complementary to the alternative parameterization, with n and m exchanged in Theorem 6, which is already known: The isolation lemma of Calabro et al. (2003) in combination with Cygan et al. (2012) implies that if $\oplus$ Hitting Sets can be solved in time $d^n \cdot \mathrm{poly}(n+m)$ for all $d > c$, then the same is true for Hitting Set.

Proof. Let $\mathscr{F}$ be a family of subsets some universe U of size n. The following preposterous algorithm is the core of the reduction from Set Cover to $\oplus$ Set Covers:

S1 (Remove sets at random.) Remove each $F \in \mathscr{F}$ with probability $\frac{1}{2}$.

Lemma 7. *Let t and n be positive integers with $t \leq n$, and let $\mathscr{F}$ be a family of subsets of U of size n. If $\mathscr{F}$ contains a set cover of size at most t, then with probability at least 2^{-t}, the number of set covers of size at most t in the output of* S1 *is odd.*

Proof. Let $\mathscr{C}$ be the family of set covers of size at most t, that is, the family of all subsets $C \subseteq \mathscr{F}$ with $|C| \leq t$ and $\bigcup_{F \in C} F = U$. Then the output of S1 can be viewed as a restricted family $\mathscr{F}\restriction_\rho$ for a random restriction ρ with $p_0 = \frac{1}{2}$ and $p_1 = 0$. Then $\mathscr{F}\restriction_\rho$'s family of set covers of size at most t is $\mathscr{C}\restriction_\rho$. The Main Lemma applied to $\mathscr{F}$ yields the claim that $\oplus\mathscr{C}\restriction_\rho$ holds with probability at least 2^{-t}. ∎

Proof (of Theorem 6). Let $(\mathscr{F}, t)$ be an instance of Set Cover. S1 transforms it into a new instance $(\mathscr{F}', t')$ with $t' = t$. Clearly if the input to S1 is a no-instance, so is its output. Conversely, if $(\mathscr{F}, t)$ is a yes-instance, Lemma 7 guarantees that S1 outputs a yes-instance $(\mathscr{F}', t')$ with probability at least 2^{-t}. In a second step S2, the output of S1 is fed to a hypothetical algorithm for $\oplus$ Set Covers. Overall this algorithm has a running time of $c^n \operatorname{poly}(n+m)$, but its success probability 2^{-t} is too small.

To counteract the exponential dependency on t, we use an idea of Cygan et al. (2012, Theorem 4.8) to preprocess the instance $(\mathscr{F}, t)$: Let q be any positive integer and let $(\mathscr{F}, t)$ be the overall input. In a new first step S0, we add dummy sets to $\mathscr{F}$ that each contain fresh elements and we increment t for each dummy set added; we stop once t/q is an integer – note that t has increased at most by $q-1$, which is a constant. Next we apply the "powering" step S1a which constructs the family $\mathscr{F}^q$ of unions $F_1 \cup \cdots \cup F_q$ for each choice of q sets $F_1, \ldots, F_q \in \mathscr{F}$. Set $t' = t/q \leq n/q$. Then $(\mathscr{F}^q, t')$ is a yes-instance if and only if $(\mathscr{F}, t)$ is. S1a takes time $m^q = \operatorname{poly}(m)$. Next we run S1 on $(\mathscr{F}^q, t')$ as before, which yields an instance $(\mathscr{F}', t')$, which we send to the assumed Set Cover algorithm in S2. Overall, our procedure takes time $c^n \operatorname{poly}(n+m)$ and has success probability at least $2^{-t'} \geq 2^{n/q}$. To amplify this to a constant, we repeat the procedure $2^{n/q}$ times, which leads to a running time of $c^n 2^{n/q} \operatorname{poly}(n+m)$. In particular, $2^{1/q} \to 1$ as $q \to \infty$, so in the framework of Cygan et al. (2012) the growth rate of Set Cover in terms of n is indeed at most c. Formally, we also need to observe the following: if each set in the family is bounded in size by k, then each set family sent as a query to the oracle has sets of size at most $qk = O(k)$. ∎

Hitting Set. We can apply the parity reduction to Hitting Set as well. This is "dual" to the previous section, and also follows from the fact that Set Cover has an algorithm that runs in time $c^n \operatorname{poly}$ if and only if Hitting Set has an algorithm that runs in time $c^m \operatorname{poly}$. The core of the reduction now looks like this:

P1 (Delete points at random.) For each $i \in U$ with independent probability $\frac{1}{2}$, remove $i \in U$ from U and replace every set $F \in \mathscr{F}$ by $F \setminus \{i\}$.

Theorem 8. *If* $\oplus$ Hitting Sets *can be solved in time* $c^m \cdot \mathrm{poly}(n+m)$*, then* Hitting Set *can be solved in time* $c^m \cdot \mathrm{poly}(n+m)$*.*

5 Consequences for Parameterized Complexity

We define the parameterized complexity class $\oplus$W[1] in terms of its complete problem $\oplus$ Multicolored Cliques: This problem is given a graph G and a coloring $c\colon V(G) \to [k]$ to decide if there is an odd number of *multicolored cliques*, that is, cliques of size exactly k where each color is used exactly once. Formally we treat $\oplus$ Multicolored Cliques as an ordinary decision problem. We let $\oplus$W[1] be the class of all parameterized problems that have an *fpt-reduction* to $\oplus$ Multicolored Clique. We recall from Flum and Grohe (2006, Def. 2.1) that fpt-reductions are deterministic many-to-one reductions that run in fixed-parameter tractable time and that map an instance with parameter k to an instance with parameter at most $f(k)$. We prove the following connection between W[1] and $\oplus$W[1] as a consequence of the Main Lemma.

Theorem 9. *There is a randomized fpt-reduction from* Multicolored Clique *to* $\oplus$ Multicolored Cliques *with one-sided error at most* $\frac{1}{2}$*; errors may only occur on yes-instances.*

Discussion and related work. Our motivation for Theorem 9 stems from structural complexity: Toda's theorem (Toda 1991) states that $\mathrm{PH} \subseteq \mathrm{P}^{\#\mathrm{P}}$, that is, every problem in the polynomial-time hierarchy reduces to counting satisfying assignments of Boolean formulas. Theorem 9 aspires to be a step towards an interesting analogue of Toda's theorem in parameterized complexity. In particular, the first step of Toda's proof is

$$\mathrm{NP} \subseteq \mathrm{RP}^{\oplus\mathrm{P}}\,, \tag{6}$$

or in words: there is a randomized polynomial-time oracle reduction from Sat to $\oplus$ Sat with bounded error and which can only err on positive instances; the existence of such a reduction follows from the isolation lemma. Using a trick that we also rely on in the proof of Theorem 9, Toda (1991) is able to turn this reduction into a many-to-one reduction. In terms of structural complexity, the existence of such a many-to-one reduction from Sat to $\oplus$ Sat then implies

$$\mathrm{NP} \subseteq \mathrm{RP}^{\oplus\mathrm{P}[1]}\,, \tag{7}$$

where the notation [1] indicates that the number of queries to the $\oplus$P-oracle is at most one. Theorem 9 is a natural and direct parameterized complexity analogue of (7), but for obvious reasons we decided not to state it as $\mathrm{W}[1] \subseteq \mathrm{RFPT}^{\oplus\mathrm{W}[1][1]}$.

Montoya and Müller (2013, Theorem 8.6) prove a parameterized complexity analogue of the isolation lemma. Implicit in their work is a W[1]-analogue of (6); more precisely, they obtain a reduction with similar specifications as the one in Theorem 9, but with two main differences: While their reduction guarantees uniqueness rather than just oddness, it is only a many-to-many and not a many-to-one reduction. Nevertheless, their reduction can be turned into a many-to-one reduction, even if many queries are made, as follows:

Suppose their reduction outputs a sequence $G_1, \ldots, G_t$ of queries to k-Clique such that at least one of them has exactly one clique of size k. Then we can take a disjoint union of a random subset of them: Pick a random set $S \subseteq \{1, \ldots, t\}$ and compute $G' = \dot{\bigcup}_{i \in S} G_i$. By a standard argument, we observe that the probability that G' has an odd number of k-cliques is exactly $\frac{1}{2}$. Hence the contribution of our work in the $\oplus$W[1]-setting lies more in the fact that our reduction is very simple, whereas the one of Montoya and Müller (2013, Theorem 8.7) is more complex.

We remark that Theorem 9 reveals a body of algorithmic open problems, the most intriguing of which, perhaps, is the question whether $\oplus$ k-Paths is fixed-parameter tractable or $\oplus$W[1]-hard. Note that $\oplus$ k-Matchings is polynomial-time solvable by a reduction to the determinant, which is established using a standard interpolation argument in the matching polynomial.

Proof. We now prove Theorem 9. In the proof, we apply the Main Lemma with $p_0 = \frac{1}{2}$ to the family of all multicolored cliques. The success probability of this application is $\geq 2^{-k}$, which we amplify to a constant by repeating the reduction $t = O(2^k)$ times independently. To combine the various independent trials back into a single instance, we use a trick also used by Toda (1991), which we restate here for completeness.

Lemma 10 (OR-composition for $\oplus$ Multicolored Cliques).
There is a polynomial-time algorithm A with the following specification: for all graphs $G_1, \ldots, G_t$ with k vertex colors each, the algorithm produces a graph $G' = A(G_1, \ldots, G_t; k)$ with $k' = tk$ colors such that G' has an odd number of multicolored cliques if and only if at least one G_i has an odd number of multicolored cliques.

Proof. Let $G_1, \ldots, G_t$ and k be given as input. Let $k' = tk$. First we add a fresh disjoint multicolored clique of size k to each graph to obtain the graphs $G_1^{+1}, \ldots, G_t^{+1}$. We assume that all $k' = tk$ colors are distinct. Now we compute the "clique sum" of the t graphs, that is, we compute the graph H constructed as follows: Starting from the disjoint union of the graphs, we add all edges between vertices from distinct graphs. Finally, the reduction produces the output $G' = H^{+1}$, that is, H where we added a fresh disjoint multicolored clique of size k'. Let N_i be the number of multicolored cliques in G_i. It is easy to see that the number of multicolored cliques in G' is $1 + \prod_{i=1}^{k}(N_i + 1)$. In particular, this number is odd if and only if at least one N_i is odd, which proves the correctness of the reduction. ∎

We are ready to prove the theorem.

Proof (of Theorem 9). Let (G, k, c) be an instance of Multicolored Clique. Let $\mathscr{F} = \{\, S \subseteq V(G) : S \text{ is a multicolored clique} \,\}$. Then $\Pr(\oplus \mathscr{F}\restriction_\rho) \geq 2^{-k}$ holds by the Main Lemma. This fact motivates the following reduction: For each vertex independently, we flip a coin and remove it with probability $\frac{1}{2}$. If the input does not contain a multicolored clique, the output does not contain one either. If the input does contain a multicolored clique, then, with probability at least 2^{-k}, the output contains an odd number of multicolored cliques. Repeating this reduction $t = O(2^k)$ times independently produces

graphs $G_1, \ldots, G_t$, which we combine into a single instance G' using Lemma 10. Overall, the reduction takes time $2^k \operatorname{poly}(n)$ and the parameter of the output is $t \cdot k = f(k)$, so this is an fpt-reduction. It remains to prove that the error probability of the overall reduction is bounded by a constant: If G does not have a multicolored clique, then, with probability one over the random choices of the reduction, the output G' has an even number of multicolored cliques. On the other hand, if G has a multicolored clique, then, with probability at most $\left(1 - 2^{-k}\right)^t \leq \exp(-2^{-k} \cdot t) \leq \frac{1}{2}$, the output G' has an even number of multicolored cliques. ∎

Inspired by Theorem 9, we stumbled on a body of algorithmic open problems, the most intriguing of which, perhaps, is the question whether $\oplus$ k-Paths is fixed-parameter tractable or $\oplus$W[1]-hard.

6 Black-box Optimality of the Main Lemma

We consider a framework similar to Dell et al. (2013), who provide evidence that Valiant and Vazirani (1986) achieves the best possible success probability – they prove that every isolation procedure that acts in a certain black-box model can have success probability at most $O(\frac{1}{n})$. Here, we prove that the probability guarantees of Lemma 1 cannot be significantly improved by exchanging the distribution $\mathcal{D}(p_0, p_1, n)$ with $p_0 = \min\{\frac{1}{2}, \frac{c}{k}\}$ and $p_1 = 0$ for any other distribution $\mathcal{D}$ over the set of restrictions $\rho : [n] \to \{0, *\}$.

Lemma 11. *For all positive integers c, k and n with $c \leq k \leq n$ and all distributions $\mathcal{D}$ over restrictions $\rho : [n] \to \{0, *\}$, there is a non-empty family $\mathscr{F}$ of cost at most c and with sets of size at most k such that:*

$$\Pr_{\rho \sim \mathcal{D}}\left(\oplus\mathscr{F}\restriction_\rho\right) \leq q \doteq \begin{cases} 2^{-k} & \text{if } \frac{c}{k} = 1\,, \\ 2^{-k} \cdot 2 & \text{if } \frac{1}{2} \leq \frac{c}{k} < 1\,, \\ 2^{-H(c/k)k} \cdot \sqrt{8k} & \text{if } 0 \leq \frac{c}{k} \leq \frac{1}{2}\,. \end{cases}$$

This lemma establishes the optimality of Lemma 1 in a "black-box" model of restriction-based decision-to-parity reductions, in which the restriction computed by the reduction may only depend on the given parameters c, k, and n, and not on any other aspect of the set family $\mathscr{F}$.

Proof. The claim is that the following quantity is at most q:

$$\sup_{\mathcal{D}} \inf_{\mathscr{F}} \Pr_{\rho \sim \mathcal{D}}\left(\oplus\mathscr{F}\restriction_\rho\right) = \inf_{\mathcal{D}'} \sup_{\rho} \Pr_{\mathscr{F} \sim \mathcal{D}'}\left(\oplus\mathscr{F}\restriction_\rho\right).$$

The equality follows from the minimax theorem. We define a suitable distribution $\mathcal{D}'$ on non-empty families $\mathscr{F} \subseteq \binom{[n]}{\leq k}$ to bound the right-hand side. The distribution $\mathcal{D}'$ chooses uniformly at random a set $S \subseteq [k]$ of size at least $s \doteq k - c$. The distribution produces the family $\mathscr{F}$ that consists of all sets S' that satisfy $S \subseteq S' \subseteq [k]$. Note that

all families $\mathscr{F}$ obtained in this fashion have cost at most c because they are extremal families whose set of irrelevant vertices is $\overline{S}$, which is of size c.

We claim that every function $\rho : [n] \to \{0, *\}$ satisfies

$$\Pr_{\mathscr{F} \sim \mathcal{D}'}\left(\oplus\mathscr{F}\restriction_\rho\right) \leq q\,.$$

Since $\mathcal{D}'$ only outputs families over $[k]$, we can assume without loss of generality that $n = k$. Furthermore, $\mathcal{D}'$ only outputs families $\mathscr{F}$ that are upwards closed over $[k]$ and that have a unique minimal element S. For such families, $\oplus\mathscr{F}\restriction_\rho$ holds if and only if $\rho(i) = 0$ holds for all $i \in \overline{S}$ and $\rho(i) = *$ for all $i \in S$. Let r be the number of elements $i \in [k]$ for which $\rho(i) = *$. If $r < s$, then the probability is zero and $q \geq 0$ is an upper bound. Otherwise there is exactly one choice of S so that the event happens, and we have

$$\Pr_{\mathscr{F} \sim \mathcal{D}'}\left(\oplus\mathscr{F}\restriction_\rho\right) = \left(\sum_{i=s}^{k} \binom{k}{i}\right)^{-1}.$$

If $s = 0$, this probability equals 2^{-k}. If $s \leq \frac{k}{2}$, the probability is at most $\left(2^k/2\right)^{-1} = 2{\cdot}2^{-k}$. Finally if $s \geq \frac{k}{2}$, we use the fact that

$$\binom{k}{\geq s} = \binom{k}{\leq c} \geq \frac{1}{\sqrt{8c(1-c/k)}} \cdot 2^{H(c/k)\cdot k} \geq \frac{1}{\sqrt{8k}} \cdot 2^{H(c/k)k}\,.$$

The first inequality is a standard lower bound on the size of the Hamming ball of radius c, see Ash (1965, p. 121). Thus we can bound the probability from above by $\left(2^{H(c/k)k}/\sqrt{8k}\right)^{-1} = 2^{-H(c/k)k} \cdot \sqrt{8k}$. This finishes the proof of the lemma. ∎

Acknowledgments. We would like to thank Johan Håstad for pointing out a simpler proof of the main lemma. Moreover, we are grateful to the following people for references and clarifying discussions: Radu Curticapean, Moritz Müller, Srikanth Srinivasan, Ryan Williams.

AB and TH are supported by the Swedish Research Council, grant VR 2012-4730: Exact Exponential-time Algorithms.

References

Arora, Sanjeev, and Boaz Barak. 2009. *Computational Complexity: A Modern Approach.* Cambridge University Press. ISBN: 978-0-521-42426-4.

Ash, Robert B. 1965. *Information Theory.* Dover Publications, Inc. ISBN: 0-486-66521-6.

Björklund, Andreas. 2012. "Below all subsets for permutational counting problems." arXiv: `1211.0391 [cs:DS]`.

Björklund, Andreas. 2014. "Determinant sums for undirected Hamiltonicity." *SIAM Journal on Computing* 43 (1): 280–299.

Björklund, Andreas, and Thore Husfeldt. 2013. "The parity of directed Hamiltonian cycles." In *Proceedings of the 54th Annual Symposium on Foundations of Computer Science (FOCS),* 727–735. doi:10.1109/FOCS.2013.83.

Björklund, Andreas, Thore Husfeldt, and Isak Lyckberg. 2015. "Computing the Permanent Modulo a Prime Power." In preparation.

Calabro, Chris, Russell Impagliazzo, Valentine Kabanets, and Ramamohan Paturi. 2003. "The complexity of unique k-SAT: An isolation lemma for k-CNFs." In *Proceedings of the 18th Annual Conference on Computational Complexity (CCC).* doi:10.1109/CCC.2003.1214416.

Cohen, Gil, and Avishay Tal. 2013. *Two structural results for low degree polynomials and applications.* Tech report TR13-145. Electronic Colloquium on Computational Complexity (ECCC). http://eccc.hpi-web.de/report/2013/145.

Cygan, Marek, Holger Dell, Daniel Lokshtanov, Dániel Marx, Jesper Nederlof, Yoshio Okamoto, Ramamohan Paturi, Saket Saurabh, and Magnus Wahlström. 2012. "On problems as hard as CNF-SAT." In *Proceedings of the 27th Annual Conference on Computational Complexity (CCC),* 74–84. doi:10.1109/CCC.2012.36.

Cygan, Marek, Stefan Kratsch, and Jesper Nederlof. 2013. "Fast Hamiltonicity checking via bases of perfect matchings." In *Proceedings of the 45th Annual Symposium on Theory of Computing (STOC),* 301–310. doi:10.1145/2488608.2488646.

Dell, Holger, Valentine Kabanets, Dieter van Melkebeek, and Osamu Watanabe. 2013. "Is Valiant–Vazirani's isolation probability improvable?" *Computational Complexity* 22 (2): 345–383. doi:10.1007/s00037-013-0059-7.

Flum, Jörg, and Martin Grohe. 2006. *Parameterized Complexity Theory.* Springer.

Montoya, Juan Andrés, and Moritz Müller. 2013. "Parameterized random complexity." *Theory of Computing Systems* 52 (2): 221–270. doi:10.1007/s00224-011-9381-0.

Mulmuley, Ketan, Umesh V. Vazirani, and Vijay V. Vazirani. 1987. "Matching is as easy as matrix inversion." *Combinatorica* 7 (1): 105–113. doi:10.1007/BF02579206.

Naik, Ashish V., Kenneth W. Regan, and D. Sivakumar. 1995. "On quasilinear time complexity theory." *Theoretical Computer Science* 148 (2): 325–349. doi:10.1016/0304-3975(95)00031-Q.

Toda, Seinosuke. 1991. "PP is as hard as the polynomial-time hierarchy." *SIAM Journal on Computing* 20 (5): 865–877. doi:10.1137/0220053.

Traxler, Patrick. 2008. "The time complexity of constraint satisfaction." In *Proceedings of the 3rd International Workshop on Parameterized and Exact Computation (IWPEC),* 190–201. doi:10.1007/978-3-540-79723-4_18.

Valiant, Leslie G., and Vijay V. Vazirani. 1986. "NP is as easy as detecting unique solutions." *Theoretical Computer Science* 47:85–93. doi:10.1016/0304-3975(86)90135-0.

Vassilevska Williams, Virginia, Josh Wang, Ryan Williams, and Huacheng Yu. 2015. "Finding four-node subgraphs in triangle time." In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4–6, 2015,* 1671–1680. doi:10.1137/1.9781611973730.111.