Electronic Colloquium on Computational Complexity, Report No. 98 (2020)

Impossibility of Derandomizing the Isolation Lemma for all Families

Manindra Agrawal[1], Rohit Gurjar[2], and Thomas Thierauf*[3]

[1]IIT Kanpur
[2]IIT Bombay
[3]Aalen University

July 4, 2020

Abstract

The *Isolation Lemma* states that when random weights are assigned to the elements of a finite set E, then in any given family of subsets of E, exactly one set has the minimum weight, with high probability. In this note, we present two proofs for the fact that it is impossible to efficiently derandomize the Isolation Lemma for arbitrary families.

The first proof is from Chari, Rohatgi and Srinivasan and uses the potential method. An alternate proof is due to the first author of this note. It uses the polynomial method. However, it is not written anywhere. The main purpose of this note is to present that proof. Additionally we show that the above lower bounds are almost tight with respect to various parameters.

1 Introduction

The *Isolation Lemma* by Mulmuley, Vazirani, and Vazirani [MVV87] is a powerful lemma that states that for any family of subsets of a set, one can *isolate* a subset in the family by assigning small random weights to the elements. Formally, let E be a finite set. For any weight function $w\colon E \to \mathbb{Z}$, consider its natural extension to all subsets of E as $w(M) = \sum_{e\in M} w(e)$. A weight function w is said to be *isolating for a family* $\mathcal{F} \subseteq 2^E$ of subsets of E, if there is a unique minimum weight set in $\mathcal{F}$.

Lemma 1.1 (Isolation Lemma [MVV87])**.** *Let E be a finite set, $|E| = m$, and $\mathcal{F} \subseteq 2^E$ be a family of subsets of E. Let $w\colon E \to [W]$ be a random weight function, where for each $e \in E$, the weight $w(e)$ is chosen uniformly and independently at random. Then*

$$\Pr[w \text{ is isolating for } \mathcal{F}] \;\geq\; 1 - \frac{m}{W}.$$

*Supported in part by DFG grant TH 472/5-1

ISSN 1433-8092

A proof of the lemma can be found in the original paper, or nowadays, in many text books on computational complexity (e.g., [MR95]). One can also see Joel Spencer's rewording of this proof on Wikipedia. An improved probability bound was given by Ta-Shma [Ta-15].

To appreciate the lemma, note that the size of the family $\mathcal{F}$ could be exponentially larger than the size of the weight range W. Hence, there can be exponentially many collisions among the weights of the sets in $\mathcal{F}$. But still, with a good probability, the minimum weight set in $\mathcal{F}$ will be unique. Moreover, the lemma does not assume any property about the family $\mathcal{F}$.

It is also interesting to note that the Isolation Lemma is in some sense a *black-box* method for isolation: to obtain an isolating weight function w, the family $\mathcal{F}$ is ignored, w is simply chosen at random, independently of $\mathcal{F}$. But implicitly, $\mathcal{F}$ plays a role. A weight function w that is isolating for $\mathcal{F}$ might *not* be isolating for another family $\mathcal{F}'$, or vice versa. The Isolation Lemma just states that for both families, a large fraction of all possible weight functions is isolating. However, the two fractions might greatly differ.

The original motivation of Mulmuley, Vazirani, and Vazirani to come up with the Isolation Lemma was to design a randomized parallel algorithm for the matching problem. The ground set E is the set of edges of a given graph G, and family $\mathcal{F}$ is the set of all perfect matchings in G. The Isolation Lemma yields a unique perfect matching of minimum weight. This can be found efficiently by a parallel algorithm. Since then, the lemma has found numerous applications in design of randomized algorithms, for example,

- parallel algorithms for the minimum cost flow problem [OS93, LP12] and linear matroid intersection [NSV94],
- algebraic algorithms for problems on disjoint paths in a graph [BH14, HN18],
- belief propagation [GSW12] and interior point methods [DS08] for flow problems,
- randomness efficient polynomial identity testing [KS01],
- an exponential time algorithm for lattice isomorphism [HR14].

Also see [BV04, Tra08, EW10, KBB$^+$11, LP12]) for more algorithmic applications. Likewise, many results in complexity theory have also used this lemma, for example,

- detecting unique solutions is as hard as NP [MVV87],
- results on unambiguous and parity complexity classes involving nondeterministic small space computation or small depth circuits [Wig94, GW96, ARZ99, RA00],
- results in dynamic complexity [DHK14, DKM$^+$18].

The basic intuition behind many of these applications is that the Isolation Lemma brings down the number of solutions to one, and it so happens in many settings that finding the solution is easier when there is only one solution. In many of the algorithmic applications, the only use of randomization is the Isolation Lemma, and thus, these algorithms can be made deterministic if the Isolation Lemma can be derandomized.

The *derandomization* of the Isolation Lemma is a challenging open problem. It means to *deterministically* construct an isolating weight function for a given family. In the *white-box* version of the problem, the isolating weight function may depend on the family $\mathcal{F}$. However, as the Isolation Lemma has a black-box character as described above, the standard meaning of derandomization asks for a black-box solution. That is, we are asking for a deterministic construction of an isolating weight function w without actually knowing the family $\mathcal{F}$, but with the promise that $\mathcal{F}$ belongs to a certain class $\mathcal{C}$ of interest. As we will see in Section 2, coming up with a single weight function for every family in class $\mathcal{C}$ might not be possible. Thus, one is allowed to construct a collection $\mathcal{W}$ of weight functions such that for every family in $\mathcal{C}$, some weight function in $\mathcal{W}$ is isolating. The only input parameter to compute $\mathcal{W}$ is m, the size of the ground set E. The construction is called *efficient* if it runs in time $\text{poly}(m)$ and also W, the size of the range of the weight functions in $\mathcal{W}$, is bounded by $\text{poly}(m)$.

The Isolation Lemma has been completely or partially derandomized for many specific classes of families, for example, perfect matchings in special classes of graphs [GK87, DK98, AHT07, DKR10, AGGT16] and (s,t)-paths in various classes of graphs [BTV09, KT16, vMP19]. Black-box (partial) derandomizations are also known for perfect matchings in general graphs [FGT16, ST17] and vertices of box-totally dual integral polytopes [GTV18]. On the other hand, Chari, Rohatgi and Srinivasan [CRS95] have ruled out a derandomization for arbitrary families:

An efficient *derandomization of the Isolation Lemma*
for the class of all families is impossible.

Their proof uses a potential function argument. The first author [Agr07], at that time unaware of the earlier result, made the same statement but without a proof. The main purpose of this note is to present an alternate proof of this statement using the polynomial method.

The result is similar in spirit to the fact that it is not possible to fool all possible Boolean functions on $\{0,1\}^n$ with a pseudorandom generator of seed length less than n.

2 Defining the problem

In this section, we formally define what we mean by an *efficient* derandomization of the Isolation Lemma.

Derandomization via large weights. First, observe that it is trivial to derandomize the Isolation Lemma with exponential weights. For $E = \{e_1, e_2, \ldots, e_m\}$, let $W \geq 2^{m-1}$ and for $i = 1, 2, \ldots, m$, define

$$w(e_i) = 2^{i-1}.$$

Then *any* subset $A \subseteq E$ has a unique weight. So in particular, w is isolating for any family $\mathcal{F} \subseteq 2^E$.

In the applications, one usually wants the weights to polynomially bounded however. As we will see next, this is not possible in general.

Derandomization via large weights is optimal. A simple counting argument rules out the existence of a substantially smaller isolating weight function for all possible families. Observe that for any weight function $w\colon E \to [W]$, the weight of any set $A \subseteq E$ is bounded by $w(A) \leq w(E) \leq mW$. Since E has 2^m subsets, when $mW < 2^m$, there will be two sets A_1, A_2 with the same weight. Therefore, w is not isolating for the family $\mathcal{F} = \{A_1, A_2\}$. Hence, for $W < 2^m/m$, there is no isolating function for all families $\mathcal{F}$.

Collections of weight functions. Since a single weight function cannot work, a more reasonable goal would be to construct a small collection $\mathcal{W}$ of weight functions with a small range such that for each family $\mathcal{F} \subseteq 2^E$, one of the weight functions in $\mathcal{W}$ is isolating. This can be seen as a *seeded* isolating weight function. That is, to construct a random weight function using a small number of random bits such that for each family, it is isolating with a nonzero probability.

Definition 2.1 (Isolating collection of weight functions)**.** *Let $\mathcal{W}$ be a collection of weight functions on a set E.*

1. *$\mathcal{W}$ is* isolating for a family $\mathcal{F} \subseteq 2^E$*, if there exists a function $w \in \mathcal{W}$ that is isolating for $\mathcal{F}$.*
2. *$\mathcal{W}$ is* isolating *if it is isolating for every family $\mathcal{F} \subseteq 2^E$.*
3. *$\mathcal{W}$ is* isolating for a class $\mathcal{C} \subseteq 2^{2^E}$ *of families of subsets of E, if it is isolating for every family $\mathcal{F} \in \mathcal{C}$.*

Derandomization via large collections. It is again trivial to derandomize the Isolation Lemma with exponentially many weight functions. For any subset $A \subseteq E$, define the weight function w_A as the characteristic function of A. For $e \in E$,

$$w_A(e) = \begin{cases} 1, & \text{if } e \in A, \\ 0, & \text{otherwise.} \end{cases}$$

Then the collection $\mathcal{W} = \{ w_A \mid A \subseteq E \}$ of size 2^m is isolating. Let $\mathcal{F} \subseteq 2^E$ be any family of sets. Let $A \in \mathcal{F}$ be a set of minimum cardinality in $\mathcal{F}$. Then the complementary weight function $w_{\overline{A}}$ isolates A: We have $w_{\overline{A}}(A) = 0$ and $w_{\overline{A}}(B) > 0$, for all other sets $B \in \mathcal{F}$, because any such B contains an $e \notin A$.

Efficient derandomization task. In an *efficient* derandomization, all paramaters involved should be polynomially bounded in m, the size of the ground set E.

> An *efficient derandomization of the Isolation Lemma* would mean to come up with a deterministic polynomial-time algorithm that computes a polynomial size collection of weight functions with polynomial range that isolates every family.

This is a considerably stronger requirement compared to the Isolation lemma. Clearly, the randomized procedure is replaced by a deterministic one, this is our main intention. But also

note that logically, there is a swap of quantifiers because we are asking for a black-box solution. First one has to compute the set of weight functions that then should work for all families. In the Isolation Lemma, it is the other way round. First the family is fixed, and then one chooses the weight function. This way gives more freedom, even if the choice of the weight function does not really involve the given family as explained above.

It turns out that we are actually asking for too much. Such a general derandomization is impossible.

3 A lower bound on the size of a collection and its range

Let $\mathcal{W}$ be an isolating collection weight functions. The following theorem gives a lower bound on the number of functions in $\mathcal{W}$ and their ranges in terms of the size of the ground set.

Theorem 3.1 ([CRS95, Agr07]). *Let* E *be a finite set,* $|E| = m$*, and* $\mathcal{W}$ *be an isolating collection of* N *weight functions* $E \to [W]$*. Then*

$$N(mW + 1) \geq 2^m .$$

It follows that when the number of weight functions N is polynomially bounded, then the range W has to be exponential, or vice versa.

Remark. Like we ruled out the existence of a single isolating weight function, there seems to be no straightforward counting argument to rule out a small collection of isolating weight functions. Just the fact that the number of possible families is doubly-exponential is not enough to argue the impossibility of efficient isolation. Indeed, if w is the weight function assigning all elements weight 1, then w is isolating for any family that contains the empty set. That means a single weight assignment w is isolating for half of all possible families.

To show the impossibility of efficient isolation, Chari et al. [CRS95] used a potential function argument, while we present a proof via the polynomial method. In essence our argument is that, if we have a small isolating collection of weight functions, then we will have a small set of points such that any nonzero polynomial has a nonzero evaluation on at least one of the points. One can get a contradiction by constructing a multilinear polynomial that vanishes on a given small set of points.

We will actually show a slightly more general bound from which Theorem 3.1 follows. For a collection of subsets $\mathcal{S} \subseteq 2^E$ of E, we consider the class $\mathcal{C} = 2^{\mathcal{S}}$ of all families that take sets only from $\mathcal{S}$. Now we give the bound on the parameters N and W for weight functions that are isolating for $\mathcal{C}$, in terms of the size of $\mathcal{S}$. Theorem 3.1 is the special case when $\mathcal{S} = 2^E$.

Theorem 3.2 ([CRS95, Agr07]). *Let* E *be a finite set, where* $|E| = m$*. Let* $\mathcal{S} \subseteq 2^E$ *and* $\mathcal{C} = 2^{\mathcal{S}}$*. Let* $\mathcal{W}$ *be a collection of* N *weight functions* $E \to [W]$ *that is isolating for* $\mathcal{C}$*. Then*

$$N(mW + 1) \geq |\mathcal{S}| . \tag{1}$$

First Proof, via polynomial method. For simplicity, let $E = [m]$ and let $\mathcal{S} \subseteq 2^E$. Let $p(z)$ be a multilinear polynomial in m variables $z = (z_1, z_2, \ldots, z_m)$. That is, p is a sum of multilinear

monomials, where each monomial is a product $\alpha_M \prod_{i \in M} z_i$ of some variables, each of degree one, for some $M \subseteq [m]$, and nonzero coefficient α_M. We associate M with the monomial. Let $\mathcal{F}_p \subseteq 2^E$ be the family of monomials of p with nonzero coefficients. Then we can write

$$p(z) = \sum_{M \in \mathcal{F}_p} \alpha_M \prod_{i \in M} z_i \,.$$

We consider only multilinear polynomials where $\mathcal{F}_p \subseteq \mathcal{S}$, i.e., polynomials with *support* in $\mathcal{S}$.

By our assumption, there is a weight function $w \in \mathcal{W}$ that is isolating for $\mathcal{F}_p$. Let $M^* \in \mathcal{F}_p$ be the unique minimum weight set in $\mathcal{F}_p$ w.r.t. w. Consider the following substitution of variables: for a new variable t,

$$z_j \mapsto t^{w(j)}.$$

Note that the substitution replaces a monomial M by

$$\prod_{j \in M} z_j \mapsto t^{w(M)}.$$

Let $q(t)$ be the polynomial obtained from $p(z)$ after this substitution. We claim that $q(t)$ is a nonzero polynomial. This is because the term $t^{w(M^*)}$ in $q(t)$ is coming from a unique monomial of $p(z)$ and cannot be canceled with any other term.

Note that the degree of $q(t)$ is at most mW. Thus, $q(t)$ is nonzero for at least one of the points in $[mW + 1]$. We conclude that any multilinear polynomial $p(z)$ supported on a subset of $\mathcal{S}$ gives a nonzero evaluation on at least one of the points in

$$T = \left\{ (t^{w(1)}, t^{w(2)}, \ldots, t^{w(m)}) \mid w \in \mathcal{W}, \quad t \in [mW + 1] \right\}.$$

Now, we argue that the size of T must be large.

Claim 3.3. *Let $H \subseteq \mathbb{R}^m$ be a set of points with $|H| < |\mathcal{S}|$. Then there exists a nonzero multilinear polynomial $p(z)$ supported on some subset of $\mathcal{S}$ which is zero at all points in H.*

Proof. We find $p(z)$ by setting up a system of linear equations. The $|\mathcal{S}|$ coefficients of $p(z)$, say $\{\alpha_M\}_{M \in \mathcal{S}}$, are the unknowns. Each point $h \in H$ gives us a homogeneous linear equation,

$$\sum_{M \in \mathcal{S}} \alpha_M \prod_{j \in M} h_j = 0.$$

When the number of constraints $|H|$ is smaller than the number of unknowns $|\mathcal{S}|$, there exists a non-trivial solution. This proves the claim. □

Claim 3.3 implies that $|T| = N(mW + 1) \geq |\mathcal{S}|$. □

We also present the proof of Chari, Rohatgi and Srinivasan [CRS95] that uses a very different technique. It actually gives a slighty stronger bound than (1), namely

$$NmW + 1 \geq |\mathcal{S}| \,. \tag{2}$$

Second Proof, via potential method [CRS95]. Consider the following algorithm. It successively takes out a set A from $\mathcal{S}$ that is isolated by some $w \in \mathcal{W}$. This defines an order on the sets in $\mathcal{S}$. We will argue via this order.

- Initialize $\mathcal{F} \leftarrow \mathcal{S}$
- while $|\mathcal{F}| > 1$ do
  - find some $w \in \mathcal{W}$ that is isolating for $\mathcal{F}$. Let $A \in \mathcal{F}$ be the isolated set w.r.t. w
  - Update $\mathcal{F} \leftarrow \mathcal{F} - A$

The *while* loop does exactly $|\mathcal{S}| - 1$ iterations. At any iteration, we define the *potential function* $\Phi(\mathcal{F})$ with respect to the current family $\mathcal{F}$, that sums up the weight $w(A)$ of a minimum weight set $A \in \mathcal{F}$, for every $w \in \mathcal{W}$,

$$\Phi(\mathcal{F}) = \sum_{w \in \mathcal{W}} \min_{A \in \mathcal{F}} w(A).$$

The potential Φ increases in every iteration by at least 1, because with respect to some $w \in \mathcal{W}$, we remove the *unique* minimizing set from $\mathcal{F}$. Hence, the potential values at the start of the loop and at the end of the loop must differ by at least $|\mathcal{S}| - 1$.

The potential at the start of the loop is $\Phi_0 = \Phi(\mathcal{S}) \geq 0$. Let A_0 be the remaining set in $\mathcal{F}$ when the algorithm halts. Then the potential at the end of the loop is $\Phi_1 = \sum_{w \in \mathcal{W}} w(A_0) \leq NmW$. Thus, we have

$$|\mathcal{S}| - 1 \leq \Phi_1 - \Phi_0 \leq \Phi_1 \leq NmW.$$

This shows (2). □

Minimizing the number of random bits

Any derandomization question can also be interpreted as that of minimizing the number of random bits used. Chari et al. [CRS95] had presented their lower bound in these terms. Suppose we want to construct a weight function $w\colon E \to [W]$ using a randomized procedure such that for each family $\mathcal{F}$ in a given class $\mathcal{C}$, w is isolating for $\mathcal{F}$ with a nonzero probability. Theorem 3.2 also implies a lower bound on the number of random bits needed in any such randomized construction. Observe that if we use r random bits in the procedure, then essentially, w will be sampled from a collection $\mathcal{W}$ of weight functions with $N = |\mathcal{W}| \leq 2^r$. The requirement of nonzero probability just means that the collection $\mathcal{W}$ should be isolating for $\mathcal{C}$. From Theorem 3.2 we get that $2^r(nW + 1) \geq |S|$, and therefore $r \geq \log|\mathcal{S}| - \log(mW + 1)$.

Corollary 3.4. *Let E be a finite set, where $|E| = m$. Let $\mathcal{S} \subseteq 2^E$ and $\mathcal{C} = 2^{\mathcal{S}}$. Let w be a weight function $E \to [W]$ constructed via a randomized procedure such that for every family $\mathcal{F}$ in $\mathcal{C}$, w is isolating for $\mathcal{F}$ with nonzero probability. Then the procedure needs to use at least $\log|\mathcal{S}| - \log(mW + 1)$ many random bits.*

Generalization to multisets

Klivans and Spielman [KS01] gave a generalized version of the Isolation Lemma that works on a family of multisets, where the weight of a multiset is defined analogously, i.e., by taking the multiplicity of the elements into account.

Lemma 3.5 (Generalized Isolation Lemma [KS01]). *Let* E *be a finite set,* $|E| = m$*, and* $\mathcal{F}$ *be a family of multisets over* E*, where the multiplicity of every element is at most* d*. Let* $w\colon E \to [W]$ *be a random weight function, where for each* $e \in E$*, the weight* $w(e)$ *is chosen uniformly and independently at random. Then*

$$\Pr[w \text{ is isolating for } \mathcal{F}] \;\geq\; 1 - \frac{dm}{W}.$$

We can generalize our lower bound in Theorem 3.1 to the setting of multisets, with an appropriate dependence on parameter d. Both the above proofs can be extended to work in this setting. In the polynomial method proof, we need to replace *multilinear* monomials with monomials of *individual degree* d.

Corollary 3.6. *Let* E *be a finite set,* $|E| = m$*, and* $\mathcal{W}$ *be a collection of* N *weight functions* $E \to [W]$ *that is isolating for all families of multisets over* E *with multiplicities bounded by* d*. Then*

$$N(mdW + 1) \geq (d+1)^m.$$

Haviv and Regev [HR14] generalized Lemma 3.5 even further to a setting where they want multiple sets/multisets to be isolated, instead of just one. One of the versions of their lemma requires that for some parameter ℓ, the ℓ smallest weight multisets should be unique, i.e., have pairwise different weights. For randomly chosen weights, they show that this is achieved with probability at least

$$1 - \frac{d(2d+1)\ell^2 m}{W}.$$

Clearly, the lower bound given in Corollary 3.6 applies here too (by discarding the parameter ℓ).

Lower bounds for circuit families

The classes we consider in Theorem 3.2 can capture other interesting classes that have been studied in the context of derandomizing the Isolation Lemma. Arvind and Mukhopadhyay [AM08] showed that the derandomization of the Isolation Lemma implies certain circuit lower bounds. They considered classes of families based on circuits and showed a connection to black-box derandomization of *polynomial identity testing* (PIT).

Let C be a Boolean circuit with m inputs and $A \subseteq [m]$ be a set. Let $\chi_A \in \{0,1\}^m$ be the characteristic string of A. With circuit C, we associate the following family $\mathcal{F}_C \subseteq 2^{[m]}$ of sets,

$$\mathcal{F}_C = \{\, A \mid C(\chi_A) = 1 \,\}.$$

For a parameter s, the class $\mathcal{C}_{m,s}$ consists of all families that correspond to circuits of size at most s,

$$\mathcal{C}_{m,s} = \{\, \mathcal{F}_C \mid \text{circuit } C \text{ has } m \text{ inputs and size } \leq s \,\}.$$

Note that $\mathcal{C}_{m,s}$ contains exponentially many families in m and s. Theorem 3.2 also gives a lower bound for $\mathcal{C}_{m,s}$.

Corollary 3.7. *Let $m, s \geq 1$ and $\mathcal{W}$ be a collection of N weight functions $[m] \to [W]$ that is isolating for $\mathcal{C}_{m,s}$. Then*

$$mN(mW + 1) \geq s\,.$$

Proof. First observe that for any family $\mathcal{F} \subseteq 2^{[m]}$, one can construct a trivial circuit of size $m|\mathcal{F}|$: put one OR-gate on top of $|\mathcal{F}|$ AND-gates, one for every set $A \in \mathcal{F}$, each of which is connected to the m inputs, namely, to input x_i if $i \in A$, and to $\overline{x}_i$ otherwise.

Let $\mathcal{S} \subseteq 2^{[m]}$ be an arbitrary collection of subsets with $|\mathcal{S}| = s/m$. By the observation just made, we have $\mathcal{S} \subseteq \mathcal{C}_{m,s}$. Thus, $\mathcal{W}$ is also isolating for $2^{\mathcal{S}}$. Hence, the claim follows from Theorem 3.2. □

Corollary 3.7 rules out the possibility to efficiently derandomize the Isolation Lemma for exponential-size circuits. But it leaves open that an efficient derandomization is possible for polynomial-size circuits.

4 An upper bound on the size of a collection and its range

Theorem 3.1 gives a lower bound on the size of a collection $\mathcal{W}$ to be a isolating for all possible families $\mathcal{F} \subseteq 2^E$. We now show that the lower bound is tight up to a multiplicative factor of $2m$.

Theorem 4.1. *Let $E = [m]$ be a finite set. Given any numbers W, N such that*

$$NW \geq 2^{m+1}\,,$$

then there exists a collection $\mathcal{W}$ of at most N weight functions $E \to [W]$ that is isolating.

Proof. The construction of $\mathcal{W}$ will be a combination of the two strategies we discussed – the one with large weights and the other with large collections. Define k as maximum such that $2^k \leq W$. We partition E into $E = E_0 \cup E_1$, where $E_0 = [k]$ and $E_1 = [m] - [k]$. On E_0 we define the weight function w_0. For $e \in E_0$, let

$$w_0(e) = 2^{e-1}\,.$$

We combine w_0 with a collection of weight functions on E_1. For any subset $A_1 \subseteq E_1$, define the weight function w_{A_1} as follows. For $e \in E$, let

$$w_{A_1}(e) = \begin{cases} w_0(e), & \text{if } e \in E_0, \\ 2^k, & \text{if } e \in A_1, \\ 0, & \text{otherwise.} \end{cases}$$

Define the collection $\mathcal{W} = \{\, w_{A_1} \mid A_1 \subseteq E_1 \,\}$. Note that $W < 2^{k+1}$ and so, by the assumption in the lemma, we have $N \geq 2^{m-k} = |\mathcal{W}|$. It remains to show that $\mathcal{W}$ is isolating.

Let $\mathcal{F} \subseteq 2^E$ be any family of sets. We exhibit a specific set $A^* \in \mathcal{F}$ that is isolated by $\mathcal{W}$. For any set $A \in \mathcal{F}$, consider its size within E_1, i.e. $|A \cap E_1|$. Fix one set in $\mathcal{F}$, say $\widehat{A}$, that has minimum size within E_1. There might be several sets $A \in \mathcal{F}$ that agree with $\widehat{A}$ on E_1. Let

$$\mathcal{F}_0 = \left\{\, A \in \mathcal{F} \mid A \cap E_1 = \widehat{A} \cap E_1 \,\right\}.$$

From $\mathcal{F}_0$, select the unique set, say A^*, that has minimum weight on E_0, i.e. $w(A^* \cap E_0)$ is minimum.

Let $A_1 = E_1 - A^*$. We claim that the weight function w_{A_1} isolates A^*: We have $w_{A_1}(A^* \cap E_1) = 0$ and $w_{A_1}(A^* \cap E_0) < 2^k$. Thus, $w_{A_1}(A^*) < 2^k$. On the other hand, for any set $A \in \mathcal{F} - \mathcal{F}_0$, we have $w_{A_1}(A) \geq 2^k$, because any such A contains an element $e \in A_1$ which has weight 2^k. Moreover, by construction, all sets in $\mathcal{F}_0$ have distinct weights and A^* has the unique minimum weight among them. □

Next, we generalize Theorem 4.1 to the setting of Theorem 3.2, i.e., for class $\mathcal{C} = 2^{\mathcal{S}}$, for any $\mathcal{S} \subseteq 2^E$. However, we do not come close to the lower bound of Theorem 3.2. Here, we do not have an explicit construction, but use the *probabilistic method*.

Lemma 4.2. *Let E be a finite set, where $|E| = m$. Let $\mathcal{S} \subseteq 2^E$ and $\mathcal{C} = 2^{\mathcal{S}}$. Given numbers W, N such that*

$$N \log \frac{W}{m} > |\mathcal{S}|\,, \tag{3}$$

there exists a collection $\mathcal{W}$ of N weight functions $E \to [W]$ that is isolating for $\mathcal{C}$.

Proof. Suppose we construct $\mathcal{W}$ by choosing N weight functions from $[W]^m$ uniformly and independently at random. By Lemma 1.1, for a given family $\mathcal{F} \in 2^{\mathcal{S}}$ and a randomly chosen w, we have

$$\Pr[w \text{ is not isolating for } \mathcal{F}] \;\leq\; \frac{m}{W}\,.$$

Since the weight functions in $\mathcal{W}$ are chosen independently, we have

$$\Pr[\mathcal{W} \text{ is not isolating for } \mathcal{F}] \;\leq\; \left(\frac{m}{W}\right)^N\,.$$

By the union bound, we get

$$\Pr[\,\exists \mathcal{F} \in 2^{\mathcal{S}} \;\; \mathcal{W} \text{ is not isolating for } \mathcal{F}] \leq\; 2^{|\mathcal{S}|} \left(\frac{m}{W}\right)^N \;<\; 1\,,$$

where the last inequality uses (3). Thus, $\mathcal{W}$ is isolating for $\mathcal{C}$ with a nonzero probability. □

Open Question

We have shown in Theorem 4.1 that the lower bound of Theorem 3.1 with respect to all families is almost tight. However, for the class $\mathcal{C} = 2^{\mathcal{S}}$, we have a gap between the lower bound and the upper bound in Theorem 3.2 and Lemma 4.2, respectively. Can we improve Lemma 4.2 to show an isolating collection of weight functions for any parameters W, N with $NW \geq |S|\,\mathrm{poly}(m)$?

References

[AGGT16] Rahul Arora, Ashu Gupta, Rohit Gurjar, and Raghunath Tewari. Derandomizing isolation lemma for $K_{3,3}$-free and K_5-free bipartite graphs. In *Proceedings of the 33rd Symposium on Theoretical Aspects of Computer Science (STACS)*, volume 47 of *LIPIcs*, pages 10:1 – 10:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016.

[Agr07] Manindra Agrawal. Rings and integer lattices in computer science. Barbados Workshop on Computational Complexity, 2007. Lecture no. 9. https://www.cs.mcgill.ca/~denis/barbados07/barbados2007.ps.

[AHT07] Manindra Agrawal, Thanh Minh Hoang, and Thomas Thierauf. The polynomially bounded perfect matching problem is in NC^2. In *Proceedings of the 24th International Symposium on Theoretical Aspects of Computer Science (STACS)*, volume 4393 of *Lecture Notes in Computer Science*, pages 489 – 499. Springer, 2007.

[AM08] Vikraman Arvind and Partha Mukhopadhyay. Derandomizing the isolation lemma and lower bounds for circuit size. In *Proceedings of Approximation, Randomization and Combinatorial Optimization. Algorithms and Techniques, 11th International Workshop, APPROX, and 12th International Workshop, RANDOM*, pages 276 – 289, 2008.

[ARZ99] Eric Allender, Klaus Reinhardt, and Shiyu Zhou. Isolation, matching, and counting uniform and nonuniform upper bounds. *Journal of Computer and System Sciences*, 59(2):164 – 181, 1999.

[BH14] Andreas Björklund and Thore Husfeldt. Shortest two disjoint paths in polynomial time. In *Proceedings of the 41st International Colloquium on Automata, Languages, and Programming (ICALP)*, volume 8572 of *Lecture Notes in Computer Science*, pages 211 – 222. Springer, 2014.

[BTV09] Chris Bourke, Ragunath Tewari, and N. V. Vinodchandran. Directed planar reachability is in unambiguous log-space. *ACM Transactions on Computation Theory*, 1:4:1 – 4:17, 2009.

[BV04] Rene Beier and Berthold Vöcking. Typical properties of winners and losers in discrete optimization. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing (STOC)*, pages 343 – 352, 2004.

[CRS95] Suresh Chari, Pankaj Rohatgi, and Aravind Srinivasan. Randomness-optimal unique element isolation withapplications to perfect matching and related problems. *SIAM Journal on Computing*, 24(5):1036 – 1050, 1995.

[DHK14] Samir Datta, William Hesse, and Raghav Kulkarni. Dynamic complexity of directed reachability and other problems. In *Proceedings of the 41st International Colloquium on Automata, Languages, and Programming (ICALP)*, volume 8572 of *Lecture Notes in Computer Science*, pages 356 – 367. Springer, 2014.

[DK98] Elias Dahlhaus and Marek Karpinski. Matching and multidimensional matching in chordal and strongly chordal graphs. *Discrete Applied Mathematics*, 84(13):79 – 91, 1998.

[DKM+18] Samir Datta, Raghav Kulkarni, Anish Mukherjee, Thomas Schwentick, and Thomas Zeume. Reachability is in DynFO. *Journal of the ACM*, 65(5), August 2018.

[DKR10] Samir Datta, Raghav Kulkarni, and Sambuddha Roy. Deterministically isolating a perfect matching in bipartite planar graphs. *Theory of Computing Systems*, 47:737–757, 2010.

[DS08] Samuel I. Daitch and Daniel A. Spielman. Faster approximate lossy generalized flow via interior point algorithms. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing (STOC)*, pages 451 – 460, 2008.

[EW10] Jeff Erickson and Pratik Worah. Computing the shortest essential cycle. *Discrete & Computational Geometry*, 44:912 – 930, 2010.

[FGT16] Stephen A. Fenner, Rohit Gurjar, and Thomas Thierauf. Bipartite perfect matching is in quasi-NC. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 754 – 763, 2016.

[GK87] Dima Grigoriev and Marek Karpinski. The matching problem for bipartite graphs with polynomially bounded permanents is in NC. In *Proceedings of the 28th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 166 – 172, 1987.

[GSW12] David Gamarnik, Devavrat Shah, and Yehua Wei. Belief propagation for min-cost network flow: Convergence and correctness. *Operations Research*, 60(2):410 – 428, 2012.

[GTV18] Rohit Gurjar, Thomas Thierauf, and Nisheeth K. Vishnoi. Isolating a vertex via lattices: Polytopes with totally unimodular faces. In *Proceedings of the 45th International Colloquium on Automata, Languages, and Programming (ICALP)*, volume 107 of *LIPIcs*, pages 74:1 – 74:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018.

[GW96] Anna Gál and Avi Wigderson. Boolean complexity classes vs. their arithmetic analogs. In *Proceedings of the 7th International Conference on Random Structures and Algorithms*, pages 99 – 111, 1996.

[HN18] Hiroshi Hirai and Hiroyuki Namba. Shortest (a+b)-path packing via hafnian. *Algorithmica*, 80(8):2478–2491, 2018.

[HR14] Ishay Haviv and Oded Regev. On the lattice isomorphism problem. In *Proceedings of the 25th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 391 – 404. SIAM, 2014.

[KBB+11] Yashodhan Kanoria, Mohsen Bayati, Christian Borgs, Jennifer T. Chayes, and Andrea Montanari. Fast convergence of natural bargaining dynamics in exchange networks. In *Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1518 – 1537. SIAM, 2011.

[KS01] Adam Klivans and Daniel A. Spielman. Randomness efficient identity testing of multivariate polynomials. In *Proceedings of the 33rd Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 216 – 223, 2001.

[KT16] Vivek Anand T. Kallampally and Raghunath Tewari. Trading determinism for time in space bounded computations. In *Proceedings of the 41st International Symposium on Mathematical Foundations of Computer Science (MFCS)*, pages 10:1 – 10:13, 2016.

[LP12] Andrzej Lingas and Mia Persson. A fast parallel algorithm for minimum-cost small integral flows. In *Proceedings of the 18th International Conference on Parallel Processing (Euro-Par)*, pages 688 – 699. Springer Berlin Heidelberg, 2012.

[MR95] Rajeev Motwani and Prabhakar Raghavan. *Randomized Algorithms*. Cambridge University Press, USA, 1995.

[MVV87] Ketan Mulmuley, Umesh V. Vazirani, and Vijay V. Vazirani. Matching is as easy as matrix inversion. *Combinatorica*, 7:105–113, 1987.

[NSV94] H. Narayanan, Huzur Saran, and Vijay V. Vazirani. Randomized parallel algorithms for matroid union and intersection, with applications to arboresences and edge-disjoint spanning trees. *SIAM Journal on Computing*, 23(2):387 – 397, 1994.

[OS93] James B. Orlin and Clifford Stein. Parallel algorithms for the assignment and minimum-cost flow problems. *Operations Research Letters*, 14(4):181–186, 1993.

[RA00] Klaus Reinhardt and Eric Allender. Making nondeterminism unambiguous. *SIAM Journal on Computing*, 29(4):1118–1131, 2000.

[ST17] Ola Svensson and Jakub Tarnawski. The matching problem in general graphs is in quasi-NC. In *Proceedings of the 58th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 696 – 707, 2017.

[Ta-15] Noam Ta-Shma. A simple proof of the isolation lemma. Technical Report TR15-080, Electronic Colloquium on Computational Complexity (ECCC), 2015.

[Tra08] Patrick Traxler. The time complexity of constraint satisfaction. In *Proceedings of the 3rd International Workshop on Parameterized and Exact Computation*, pages 190 – 201. Springer Berlin Heidelberg, 2008.

[vMP19] Dieter van Melkebeek and Gautam Prakriya. Derandomizing isolation in space-bounded settings. *SIAM Journal on Computing*, 48(3):979 – 1021, 2019.

[Wig94] Avi Wigderson. NL/poly $\subseteq \oplus$ L/poly. In *Proceedings of the 9th Annual Structure in Complexity Theory Conference*, pages 59 – 62, 1994.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il