Electronic Colloquium on Computational Complexity, Report No. 183 (2025)

Symmetric Distributions from Shallow Circuits

Daniel M. Kane* Anthony Ostuni† Kewen Wu‡

Abstract

We characterize the symmetric distributions that can be (approximately) generated by shallow Boolean circuits. More precisely, let $f\colon \{0,1\}^m \to \{0,1\}^n$ be a Boolean function where each output bit depends on at most d input bits. Suppose the output distribution of f evaluated on uniformly random input bits is close in total variation distance to a symmetric distribution $\mathcal{D}$ over $\{0,1\}^n$. Then $\mathcal{D}$ must be close to a mixture of the uniform distribution over n-bit strings of even Hamming weight, the uniform distribution over n-bit strings of odd Hamming weight, and γ-biased product distributions for γ an integer multiple of 2^{-d}. Moreover, the mixing weights are determined by low-degree, sparse $\mathbb{F}_2$-polynomials. This extends the previous classification for generating symmetric distributions that are also uniform over their support.

*University of California, San Diego. Email: dakane@ucsd.edu. Supported by NSF Medium Award CCF-2107547 and NSF CAREER Award CCF-1553288.

†University of California, San Diego. Email: aostuni@ucsd.edu.

‡Institute for Advanced Study. Email: shlw_kevin@hotmail.com. Supported by the National Science Foundation under Grant No. DMS-2424441, and by the IAS School of Mathematics.

ISSN 1433-8092

Contents

1 Introduction

One of the most celebrated results in complexity theory is Håstad's proof [Hås86] that AC^0 circuits[1] require an exponential number of gates to compute the parity function. Surprisingly, the weaker circuit class NC^0 suffices to perform a very similar task. In particular, mapping uniformly random bits $(x_1, x_2, \ldots, x_n)$ to $(x_1 \oplus x_2, x_2 \oplus x_3, \ldots, x_{n-1} \oplus x_n, x_n \oplus x_1)$ produces the uniform distribution over n-bit strings of even parity, or equivalently, over input-output pairs $(x, \mathrm{PARITY}(x))$ [Bab87, BL87]. This observation begs the question: what computational resources are required to (approximately) *generate* specific distributions, as opposed to the traditional task of *computing* specific functions?

A fundamental question in its own right, the complexity of sampling from distributions also has numerous applications. For example, results on the hardness of sampling can be translated to data structure lower bounds [Vio12b, LV11, BIL12, Vio20, CGZ22, Vio23, YZ24, KOW24, AGM+25], provide input-independent quantum-classical separations [WP23, Vio23, KOW24, GKM+25], and are key components to the construction of explicit codes [SS24]. Moreover, techniques and intuition developed in this setting have successfully been applied to pseudorandom generators [Vio12b, LV11, BIL12] and extractors [Vio12c, DW12, Vio14, CZ16, CS16].

While the general problem was considered in early work (see, e.g., [JVV86]), a focus on generating distributions via shallow circuits was advocated for more recently by Viola [Vio12b]. Since then, the field has seen a number of exciting developments (see, e.g., the recent works [FLRS23, Vio23, KOW24, SS24, KOW25, AGM+25] and references therein). One notable takeaway from prior works is that NC^0 circuits can sample very few uniform symmetric distributions (i.e., uniform distributions over a symmetric support), even allowing for small errors. In particular, the line of work [Vio12b, FLRS23, Vio23, KOW24, KOW25] recently culminated in the following classification result, which confirmed a conjecture of Filmus, Leigh, Riazanov, and Sokolov [FLRS23]. For a function $f\colon \{0,1\}^m \to \{0,1\}^n$, let $f(\mathcal{U}^m)$ be the distribution resulting from applying f to $x \sim \mathcal{U}^m$, the uniform distribution over $\{0,1\}^m$.

Theorem 1.1 ([KOW25])**.** *Let $\varepsilon \in [0,1]$ be arbitrary. Assume $f\colon \{0,1\}^m \to \{0,1\}^n$ is computable by an NC^0 circuit of constant depth and $f(\mathcal{U}^m)$ is ε-close in total variation distance to a uniform symmetric distribution where n is sufficiently large. Then $f(\mathcal{U}^m)$ is $O(\varepsilon)$-close to one of the following six special uniform symmetric distributions:*

- *Point distribution on 0^n,*
- *Point distribution on 1^n,*
- *Uniform distribution over $\{0^n, 1^n\}$,*
- *Uniform distribution over n-bit strings with even Hamming weights,*
- *Uniform distribution over n-bit strings with odd Hamming weights,*
- *Uniform distribution over all n-bit strings.*

We emphasize that Theorem 1.1 works with *uniform* symmetric distributions, which does not capture, for example, the $(1/4)$-biased product distribution that is symmetric and easily sampleable by NC^0 circuits.

[1] Recall that these are (families of) Boolean circuits of constant depth and unbounded fan-in gates. They may be contrasted with NC^0 circuits, which are also of constant depth, but have bounded fan-in gates.

1.1 Our Result

In this paper, we extend Theorem 1.1 to handle *arbitrary* symmetric distributions. Let the *locality* of a Boolean function be the largest number of input bits that any output bit depends on. Note that NC^0 is precisely the class of sequences of functions with bounded locality.

Theorem 1.2 (Informal version of Theorem 4.1)**.** *Let $d \geq 0$ be an integer. For any $\varepsilon \in (0, 1]$, there exists some δ such that $\delta \to 0$ as $\varepsilon \to 0$ and the following holds.*

Suppose $f\colon \{0,1\}^m \to \{0,1\}^n$ is a d-local function and $f(\mathcal{U}^m)$ is ε-close in total variation distance to a symmetric distribution where n is sufficiently large in terms of d and ε. Then $f(\mathcal{U}^m)$ is δ-close to some mixture of

1. *The uniform distribution over n-bit strings of even Hamming weight,*
2. *The uniform distribution over n-bit strings of odd Hamming weight, and*
3. *γ-biased product distributions on n bits for γ an integer multiple of 2^{-d}.*

Moreover, the mixing weights are determined by degree-$O_d(1)$ $\mathbb{F}_2$-polynomials with $O_d(n)$ monomials.

The "moreover" conclusion implies this mixture can be exactly produced by $O_d(1)$-local functions (see Remark 4.2).

Learning Structured Distributions. The reconstruction of an unknown probability density function based on observed data is a fundamental problem in both statistics and computer science. The typical setting is the PAC-learning model [Val84, BEHW89, KMR+94]: given access to independent samples of an unknown distribution $\mathcal{D}$, the goal is to output a hypothesis distribution $\mathcal{D}'$ close to $\mathcal{D}$. Much research has been carried out for Gaussian mixtures [DK14], log-concave distributions [DR09], monotone distributions [Bir87], sums of independent integer random variables [DDO+13], junta distributions [ABR16], mixtures of structured distributions [Lin95], and more. The interested reader may wish to consult the survey [Dia16] by Diakonikolas for additional background and references.

A black-box use of Theorem 1.2 in this setting is to learn symmetric distributions that are locally sampleable. Let $\mathcal{D}$ be a symmetric distribution over $\{0,1\}^n$. Assume $\mathcal{D}$ is produced by a d-local function, i.e., $\mathcal{D} = f(\mathcal{U}^m)$ for some d-local function $f\colon \{0,1\}^m \to \{0,1\}^n$. Then for any $\varepsilon > 0$ with n sufficiently large in terms of d and ε, Theorem 1.2 and the standard cover method [Yat85] (see also [Dia16, Theorem 1.5.1]) imply that we can efficiently learn $\mathcal{D}$ up to ε-error in total variation distance with $O_d(1/\varepsilon^2)$ samples with high probability. We believe this should hold for all $\varepsilon > 0$. Indeed, in Subsection 1.2 we propose Conjecture 1.4 for an exact classification of locally sampleable symmetric distributions, which, if true, would imply the desired learning theoretic result.

1.2 Open Problems

Recall that Theorem 1.2 does not work for the case of $\varepsilon = 0$, i.e., *exact* sampling. Indeed, we cannot conclude mixtures of the form given by Theorem 1.2 are the only distributions that can be exactly sampled by NC^0 circuits. We identify the following example, which illustrates that this is not simply a weakness in our analysis, but rather there are other symmetric distributions that can be sampled exactly.

Example 1.3. Let $\mathcal{U}_{1/4}^n$ be the $(1/4)$-biased product distribution over $\{0,1\}^n$. Additionally, let $\texttt{evens}$ and $\texttt{odds}$ denote the uniform distribution over n-bit strings of even Hamming weight and odd Hamming weight, respectively. The distribution[2]

$$\mathcal{P} = \mathcal{U}_{1/4}^n + 2^{-n-1}\texttt{evens} - 2^{-n-1}\texttt{odds}$$

is not of the form given by Theorem 1.2, yet it can be sampled exactly by a bitwise AND of $\texttt{evens}$ and the uniform distribution over $\{0,1\}^n$, which is 3-local. The full details can be found in Appendix A.

Additionally, we suspect that, similarly to the uniform symmetric case [KOW25], one can take the upper bound in Theorem 1.2 to be linear in $\|f(\mathcal{U}^m) - \mathcal{D}\|_{\mathsf{TV}}$ without any dependency on d. Combining with the previous discussion, we conjecture the following strengthening of Theorem 1.2 holds. Observe that Conjecture 1.4 captures Example 1.3.

Conjecture 1.4. *For every $d \in \mathbb{N}$, $\varepsilon \in (0,1)$, and n large enough in terms of d, if $f\colon \{0,1\}^m \to \{0,1\}^n$ is a d-local function and $f(\mathcal{U}^m)$ is ε-close in total variation distance to a symmetric distribution, then $f(\mathcal{U}^m)$ is $O(\varepsilon)$-close to some mixture*

$$\mathcal{M} = \sum_i \alpha_i \cdot g_i(\mathcal{D}_1^{(i)}, \dots, \mathcal{D}_d^{(i)}),$$

where each g_i is a bitwise function and each $\mathcal{D}_j^{(i)}$ is either the uniform distribution over n-bit strings of even Hamming weight or over n-bit strings of odd Hamming weight. Moreover, the mixing weights are determined by degree-$O_d(1)$ $\mathbb{F}_2$-polynomials with $O_d(n)$ monomials, as in Theorem 4.1.

Note the bitwise condition on the above g_i's guarantees the output distribution is symmetric.

Paper Organization. We provide an overview of the proof of Theorem 1.2 in Section 2, as well as a brief comparison of our techniques to those of prior literature. Background material and useful results are collected in Section 3. The bulk of our work is in Section 4, where we state and prove Theorem 4.1, the full version of Theorem 1.2. The appendices contain a number of deferred proofs.

2 Proof Overview

In this section, we sketch the proof of Theorem 1.2 before discussing how the details compare to prior work.

2.1 Proof Overview of Theorem 1.2

We begin with a useful observation from [KOW25]: the total variation distance between the distribution $f(\mathcal{U}^m)$ and any symmetric distribution $\mathcal{P}$ over $\{0,1\}^n$ is, up to constant factors, equal to the distance between the corresponding Hamming weight distributions of $f(\mathcal{U}^m)$ and $\mathcal{P}$ plus the distance between $f(\mathcal{U}^m)$ and its symmetrization (i.e., the distribution resulting from randomly permuting the coordinates of a string $x \sim f(\mathcal{U}^m)$). Expressed symbolically, we have

$$\|f(\mathcal{U}^m) - \mathcal{P}\|_{\mathsf{TV}} = \Theta(\||f(\mathcal{U}^m)| - |\mathcal{P}|\|_{\mathsf{TV}} + \|f(\mathcal{U}^m) - f(\mathcal{U}^m)_{\mathrm{sym}}\|_{\mathsf{TV}}). \tag{1}$$

[2] More formally, $\mathcal{P}$ is the distribution that assigns x probability $\mathcal{U}_{1/4}^n(x) + 2^{-n-1}\texttt{evens}(x) - 2^{-n-1}\texttt{odds}(x)$.

The proof is straightforward, and it essentially follows from several applications of the triangle inequality which show that the two ways $f(\mathcal{U}^m)$ can be far from $\mathcal{P}$ are a weight mismatch or $f(\mathcal{U}^m)$ being far from symmetric, itself (see Lemma 3.4).

By assumption, we know $f(\mathcal{U}^m)$ is ε-close to a symmetric distribution $\mathcal{D}$. Thus, (1) implies $\|f(\mathcal{U}^m) - f(\mathcal{U}^m)_{\mathrm{sym}}\|_{\mathsf{TV}} = O(\varepsilon)$. Hence, it suffices to prove a weaker version of Theorem 1.2 that only compares the weight distributions (Lemma 4.26). More precisely, we want to show that for some δ tending to 0 with ε, the Hamming weight distribution of $f(\mathcal{U}^m)$, denoted $|f(\mathcal{U}^m)|$, is δ-close to a mixture $\mathcal{M}$ of

1. The binomial distribution $\mathsf{Bin}(n, \gamma)$ with n trials and success probability γ for γ an integer multiple of 2^{-d},
2. The binomial distribution $\mathsf{Bin}(n, 1/2)$ conditioned on the outcome being even, and
3. The binomial distribution $\mathsf{Bin}(n, 1/2)$ conditioned on the outcome being odd.

Moreover, we want the mixing weights to be determined by low-degree $\mathbb{F}_2$-polynomials with few monomials. (It will later become clear what "determined" means in this context.) We first prove a weaker version of this result that does not include mixing weight information (Lemma 4.24). Afterwards, we will discuss how to obtain the desired control over the weights. Before proceeding to the details, we first give a brief, high-level overview of the four main steps of the proof and how they fit together.

In the first step, we show that any fixing of the values of input bits which affect many output bits results in the bias of the output weight distribution of $f(\mathcal{U}^m)$ concentrating around a fixed dyadic rational[3] γ multiple of n. This allows us to represent the weight distribution $|f(\mathcal{U}^m)|$ as a mixture of distributions produced by the restricted functions. We then argue in step two that after grouping the parts of the mixture which concentrate around the same γ, each group assigns roughly the same amount of mass to any contiguous interval as the binomial distribution with success probability γ. The third step is to prove a continuity result for each grouped part of the mixture; namely, that the mass assigned to some weight w and to $w + \Delta$ for a small integer Δ are roughly equal. (There is a slight subtlety here in the case of $\gamma = 1/2$, but we will defer its discussion to later in the proof overview.) This allows us to conclude in the fourth and final step that most output weights are assigned comparable mass by the part of the mixture of $|f(\mathcal{U}^m)|$ concentrated around γ and the corresponding binomial distribution. In particular, our weight distribution is close in total variation distance to a mixture of binomial distributions, which is what we wanted to show.

Step 1: Removing Large Influences. In an extremely ideal setting, we might wish that all of f's output bits are roughly independent with bias around $\gamma = a/2^d$ for some integer $0 \le a \le 2^d$; in this case the output weight would resemble the binomial distribution $\mathsf{Bin}(n, \gamma)$. Of course, this is far too much to assume. In actuality, it may be that one specific input bit affects the value of *every* output bit, so none of them are independent.

To progress toward this dream scenario, we follow in the footsteps of many prior works (e.g., [Vio12b, LV11, BIL12, Vio20, Vio23, FLRS23, KOW24, KOW25, GKM+25]) by strategically conditioning on certain input bits to express the distribution as a mixture of more structured sub-distributions. In particular, we will condition on all "high degree" input bits that affect more than n/A output bits for some A to be chosen later. The goal is now to argue that the function resulting from each conditioning produces a distribution which is roughly of the form we originally sought. Note that this will not depend on the actual values that the input bits are set to in the conditioning.

[3]Recall a *dyadic rational* is a number that can be expressed as a fraction whose denominator is a power of two.

We start with a result from [KOW24]: let $\mathcal{D}_q$ be the uniform distribution over n-bit strings of Hamming weight q. If q/n is far from every integer multiple of 2^{-d}, then any d-local function must produce a distribution far (in total variation distance) from $\mathcal{D}_q$. This essentially follows from the fact that the input bits determining any fixed output bit can only be set in 2^d equally likely ways. Observe that symmetric distributions are simply mixtures of $\mathcal{D}_q$ for different q's. Thus we can show (Lemma 4.8) that if $f(\mathcal{U}^m)$ is close to a symmetric distribution $\mathcal{D}$, then for a typical $x \sim \mathcal{U}^m$, the normalized output weight $|f(x)|/n$ has distance at most $n^{-1/(800d)}$ from an integer multiple of 2^{-d}. Note that the closest multiple may be different for different inputs; however, after each conditioning $\rho \in \{0,1\}^S$ on the set of high degree input bits $S \subseteq [m] \coloneqq \{1, 2, \ldots, m\}$, we can strengthen this result (Lemma 4.9) to say that for a typical $x \sim \mathcal{U}^{[m]\setminus S}$, the normalized output weight $|f(x,\rho)|/n$ is close to the *same* integer multiple of 2^{-d}. Here, we will only need "high degree" to correspond to affecting more than $n/O_d(1)$ many output bits (i.e., can take $A = O_d(1)$), although later we will obtain other constraints on how we must set A. Henceforth, we will shorthand $f(x,\rho)$ by $f_\rho(x)$ for clarity.

At a high level, the proof uses the second-moment method. By conditioning on the high degree input bits, we ensure that the variance of the resulting distribution is small, and thus we have concentration around a particular weight. Moreover, this weight must be close to an integer multiple of $n/2^d$, or else our previous insights imply $f(\mathcal{U}^m)$ cannot be close to $\mathcal{D}$. To be slightly more precise, a direct second-moment argument by itself is insufficient to rule out the case that some non-negligible fraction of the time the output weight is close to some other multiple of $n/2^d$. However, one can show (Claim 4.14) that if this occurs, $f_\rho(\mathcal{U}^{[m]\setminus S})$ would also have to assign decent probability mass to weights between these integer multiples, which would again imply $f(\mathcal{U}^m)$ cannot be close to $\mathcal{D}$.

Finally, we combine these deductions to prove that for each conditioning $\rho \in \{0,1\}^S$ of the bits in S, there exists a set $T \coloneqq T_\rho \subseteq [n]$ of size $|T| \le O_{d,k}(1)$ such that every k-tuple of output bits in $[n] \setminus T$ has marginal distribution $\mathcal{U}^k_{\gamma_\rho}$, the γ_ρ-biased product distribution over $\{0,1\}^k$, for γ_ρ an integer multiple of 2^{-d} (see Proposition 4.5). Here, k is some parameter at most $O_d(\log(1/\varepsilon))$. The proof operates by constructing a degree-k multilinear polynomial $P_i \colon \{0,1\}^n \to \{0,1\}$ for each k-tuple with the "wrong" distribution, where the expectation of P_i over $f_\rho(\mathcal{U}^{[m]\setminus S})$ is larger than over the γ-biased product distribution $\mathcal{U}^n_\gamma$ by at least an additive 2^{-kd} term. Note this follows from using our locality bound to view $P_i(f_\rho(\mathcal{U}^{[m]\setminus S}))$ as a polynomial of degree kd in the input bits. Summing the polynomials together into $P = \sum_i P_i$ magnifies the difference in expectations, and if the number of terms (i.e., number of bad k-tuples) is too large, we end up contradicting the assumption that $f(\mathcal{U}^m)$ is close to a symmetric distribution.

Step 2: Kolmogorov Distance. We now group the restricted functions according to their biases, defining $F_\gamma(\mathcal{U}^{[m]\setminus S}) = \mathbb{E}_{\rho:\gamma_\rho=\gamma}[f_\rho(\mathcal{U}^{[m]\setminus S})]$ for each γ. In this second step, we aim to show that every contiguous interval of output weights is assigned roughly the same amount of mass by $|F_\gamma(\mathcal{U}^{[m]\setminus S})|$ and $\mathsf{Bin}(n,\gamma)$. Since this difference in probability mass is convex over mixtures, it suffices to consider the individual distributions $|f_\rho(\mathcal{U}^{[m]\setminus S})|$. Moreover, it is enough to provide a bound on the *Kolmogorov distance*

$$\max_t \left| \mathbf{Pr}\left[|f_\rho(\mathcal{U}^{[m]\setminus S})| \ge t\right] - \mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) \ge t\right] \right|.$$

In a sentence, the bound follows from combining the k-wise independence of most output coordinates with the fact that T is too small to have much of an effect. In the case of $\gamma = 1/2$ and $|T| = 0$, Diakonikolas, Gopalan, Jaiswal, Servedio, and Viola [DGJ+10] gave an upper bound of

roughly $1/\sqrt{k}$ using techniques from approximation theory. (In fact, their result holds for arbitrary threshold functions.) This was generalized by Gopalan, O'Donnell, Wu, and Zuckerman [GOWZ10] to include the case of arbitrary $\gamma \in (0,1)$. In our case, $|T|$ will likely not be 0, but it is small enough that a similar result (Proposition 4.16) still holds. In the edge cases of $\gamma \in \{0,1\}$, we can no longer use [DGJ+10, GOWZ10], nor would the size of T be negligible even if we could, but these special cases can be addressed later via simple arguments (see, e.g., the proof of Lemma 4.23).

For reasons that will become apparent in the subsequent step, we will also need a comparable bound in the case of $\gamma = 1/2$, even accounting for the parity of the output weight. That is, we wish to show

$$\mathbf{Pr}\left[|f_\rho(\mathcal{U}^{[m]\setminus S})| \geq t \text{ and } |f_\rho(\mathcal{U}^{[m]\setminus S})| \text{ is even}\right] \approx \mathbf{Pr}\left[\mathsf{Bin}(n,1/2) \geq t\right] \cdot \mathbf{Pr}\left[|f_\rho(\mathcal{U}^{[m]\setminus S})| \text{ is even}\right].$$

Our analysis here is similar to the previous case, but we now crucially rely on [CHH+20, Theorem 3.1]. In our context, it implies that there is a small set of input bits $R \subseteq [m] \setminus S$ such that re-randomizing over R typically re-randomizes the parity of f_ρ's output weight. Moreover, since R is small and we have already conditioned on input bits of large degree, the vast majority of output bits are unaffected by R. Thus, we can compare $\mathbf{Pr}\left[|f_\rho(\mathcal{U}^{[m]\setminus S})| \geq t\right]$ to $\mathbf{Pr}\left[\mathsf{Bin}(n,1/2) \geq t\right]$ using these unaffected output bits as a proxy for the entirety of the output bits. We briefly note that for the error bounds on the "parity Kolmogorov distance" to be meaningful given our choice of k, we need to take our degree threshold A to be $O_{d,\varepsilon}(1)$ – larger than is necessary for Step 1.

Step 3: Approximate Continuity. Our third step is to argue that for each bias γ, the distribution $F_\gamma(\mathcal{U}^m)$ can be expressed as a mixture $\lambda \cdot E_\gamma + (1-\lambda) \cdot W_\gamma$, where λ is small and W_γ assigns similar probability mass to similar weights. Once again, it suffices to analyze the distributions produced by the individual restricted functions $f_\rho(\mathcal{U}^{[m]\setminus S})$. Here, we show that $f_\rho(\mathcal{U}^{[m]\setminus S})$ can be written as a mixture of distributions E and W, where E is extremely far from every symmetric distribution supported on strings of Hamming weight $\gamma n \pm n^{2/3}$, and the weight distribution of W satisfies a certain continuity property (see Proposition 4.21). More specifically, for positive integers w and Δ, the probability that W's output weight is w differs by no more than $O_d(\Delta/n)$ from the probability that W's output weight is $w+\Delta$. Since $f(\mathcal{U}^m)$ is close to $\mathcal{D}$, one can show that $F_\gamma(\mathcal{U}^m)$ is close to $\mathcal{D}$ conditioned on the output weight being $\gamma n \pm n^{2/3}$ (see Claim 4.25). Thus, proving the above result for $f_\rho(\mathcal{U}^{[m]\setminus S})$ implies minimal mass will typically be assigned to the E part of the mixtures, and so λ (in the mixture defining $F_\gamma(\mathcal{U}^m)$) will be small, as desired.

In our analysis, we will require a structural result about hypergraphs from [KOW24]. Observe that we can associate any function $g\colon \{0,1\}^m \to \{0,1\}^n$ to a hypergraph on the vertex set $[n]$ with an edge for each input bit b containing all of the output bits that depend on b. In the case that g is d-local, this hypergraph has maximum degree at most d. We follow standard hypergraph terminology in defining the *neighborhood* of a vertex v to be the set of vertices sharing an edge with v. We additionally call two neighborhoods N_1, N_2 *connected* if there exist two adjacent (i.e., contained in the same edge) vertices $v_1 \in N_1, v_2 \in N_2$.

By [KOW24, Corollary 4.11], we can find a collection of $r = \Omega_d(n)$ non-connected neighborhoods in $\mathcal{H}$ of size $O_d(1)$ by only removing $O_d(n)$ (with a small implicit constant) edges. Translating the result to f_ρ, we find that there exists a not-too-large set $B \subseteq [m] \setminus S$ of input bits such that any conditioning $\sigma \in \{0,1\}^B$ yields a sub-function $f_{\rho,\sigma}\colon \{0,1\}^{[m]\setminus(S\cup B)} \to \{0,1\}^n$ with many small, pairwise independent collections of output bits. We will define E and W according to the behavior of these neighborhoods, similarly to the arguments in [KOW24, KOW25, GKM+25].

Suppose for at least $r/2$ of the r non-connected neighborhoods $N_1, \ldots, N_r$, we have that $f_{\rho,\sigma}(\mathcal{U}^{[m]\setminus(S\cup A)})$ restricted to the neighborhood N_i has a marginal distribution which differs from

the γ-biased product distribution over N_i, denoted $\mathcal{U}_\gamma^{N_i}$. Since $\mathcal{U}_\gamma^{N_i}$ is pointwise close to the uniform distribution over strings of Hamming weight around γn, the marginal distributions of these neighborhoods (in $f_{\rho,\sigma}$) must also differ from the marginal distributions of the symmetric distribution $\mathcal{D}$. We can then accumulate these errors via concentration bounds to show that $f_{\rho,\sigma}(\mathcal{U}^{[m]\setminus(S\cup B)})$ is far from $\mathcal{D}$ (see Lemma 3.2).

We now set E to be the mixture over all conditions of the bits in B where most resulting neighborhoods differ from the corresponding γ-biased product distribution, and set W to be the mixture over the remaining conditionings. Since each sub-distribution in the mixture E is far from $\mathcal{D}$ by the previous paragraph, a union bound argument implies E, itself, must also be far from $\mathcal{D}$ (see Lemma 3.3). It remains to show the continuity property for W.

Suppose the r non-connected neighborhoods are generated by $v_1, v_2, \ldots, v_r \in [n]$ in the sense that all bits in the i-th neighborhood N_i are either v_i or in an edge that also contains v_i. We further condition on all input bits that do not affect any of $v_1, \ldots, v_r$, so that the value of every output bit outside of $N_1 \cup \cdots \cup N_r$ is fixed. In this way, the output weight of $f_{\rho,\sigma}(\mathcal{U}^{[m]\setminus(S\cup B)})$ becomes a fixed integer (corresponding to the fixed bits outside of the neighborhoods) plus the sum of the neighborhoods' output weights. Since we are in the case where most neighborhoods are extremely structured, we can show that for each $2 \le \ell \le t$, the output weight distribution modulo ℓ of N_i is not constant for a constant fraction of the N_i's with high probability (see Claim 4.22). Finally, we can obtain our desired continuity result through known density comparison theorems for sums of independent, non-constant integer random variables, such as [KOW25, Theorem A.1], which relies on classical anticoncentration tools.

There is one important exception to the above analysis: the case of $\ell = 2$ and $\gamma = 1/2$. Here, we are not guaranteed to typically get many non-connected neighborhoods with variable output weight modulo 2 (even if most of them have the "correct" marginal distribution). To better understand where our reasoning breaks down, consider one of the most surprising examples in this area of work. Define $h\colon \{0,1\}^n \to \{0,1\}^n$ to be

$$h(x_1, \ldots, x_n) = (x_1 \oplus x_2, x_2 \oplus x_3, \ldots, x_{n-1} \oplus x_n, x_n \oplus x_1),$$

so that $h(\mathcal{U}^n)$ is the uniform distribution over n-bit strings of even Hamming weight. Observe that h is extremely simple, only requiring two bits of locality. Additionally, the marginal distribution onto any $k \le n-1$ coordinates is exactly the product distribution $\mathcal{U}_{1/2}^k$.

Let us focus on a particular output bit $y_i = x_i \oplus x_{i+1}$. (For simplicity, assume i is not too close to 1 or n.) Its neighborhood is $y_{i-1} = x_{i-1} \oplus x_i$, y_i, and $y_{i+1} = x_{i+1} \oplus x_{i+2}$, so its Hamming weight modulo 2 is

$$(x_{i-1} \oplus x_i) \oplus (x_i \oplus x_{i+1}) \oplus (x_{i+1} \oplus x_{i+2}) = x_{i-1} \oplus x_{i+2}.$$

If we follow our earlier analysis and fix the value of the input bits that do not affect y_i (i.e., x_{i-1} and x_{i+2}), the neighborhood's Hamming weight modulo two is fixed, regardless of the values of x_i and x_{i+1}.

At a high level, the reason for this exceptional case is that it is the only setting of ℓ and γ for which the binomial distribution $\mathsf{Bin}(t,\gamma)$ mod ℓ can equal $\mathsf{Bin}(t-1,\gamma)$ mod ℓ. In other words, the weight distribution modulo ℓ of $\mathcal{U}_\gamma^k$ conditioned on the first bit being 0 is different than that distribution conditioned on the first bit being 1, unless $\ell = 2$ and $\gamma = 1/2$. This difference means that in the case of $\ell = 2$ and $\gamma = 1/2$, we can only derive a continuity result for weights that are an even distance apart. Hence, we must be cognizant of the output weight's parity throughout much of our analysis, which explains the required parity version of our Kolmogorov distance result mentioned earlier in the proof overview.

Step 4: Putting It Together. At this point, all that remains is to combine the pieces. Recall we have already conditioned on high degree (i.e., larger than $n/O_{d,\varepsilon}(1)$) input bits to find sub-functions $\{f_\rho\}_\rho$, each producing a distribution with some approximate bias γ_ρ. Additionally, the mixtures $F_\gamma(\mathcal{U}^{[m]\setminus S})$ obtained by grouping the sub-functions around their biases each satisfy a Kolmogorov distance bound and approximate continuity.

We partition $\{0, 1, \ldots, n\}$ into consecutive intervals of length $c\sqrt{n}$ for some small $c > 0$, and restrict our attention to the $O(\log(1/\alpha))$ of them that contain all but $O(\alpha)$ of the mass. By our Kolmogorov distance result (Lemma 4.15), we have that $|F_\gamma(\mathcal{U}^{[m]\setminus S})|$ and the binomial distribution $\mathsf{Bin}(n, \gamma)$ assign similar mass to each of these intervals. Moreover, our continuity result (Proposition 4.21) implies the mass in a fixed interval is almost uniformly distributed. Thus for most weights w, $|F_\gamma(\mathcal{U}^{[m]\setminus S})|$ assigns roughly the same mass to w as $\mathsf{Bin}(n, \gamma)$ does. Summing over all weights provides an upper bound on the total variation distance between the two weight distributions. We remark that there is a slight complication in the case of $\gamma = 1/2$, as there Proposition 4.21 only lets us compare weights that are an even distance apart. However, by further subdividing each interval we consider into its even and odd parts, we are able to carry out a similar argument as before.

Finally, we recall that $f(\mathcal{U}^m)$ is a mixture of the distributions $F_\gamma(\mathcal{U}^{[m]\setminus S})$. Since the weight distribution of each $F_\gamma(\mathcal{U}^{[m]\setminus S})$ is close to the weight distribution corresponding to one from Theorem 1.2, their mixture $|f(\mathcal{U}^m)|$ naturally is close to a mixture of those weight distributions (see Lemma 4.24). We conclude by recalling that (1) implies it was sufficient to classify the output weight distribution.

Mixing Weights. The above argument shows that $f(\mathcal{U}^m)$ is close to a mixture $\mathcal{M}$ of the form specified in Theorem 1.2, but without the additional control on the mixing weights. The reason for this shortcoming is that the threshold A at which we consider an input bit "high degree" depends not just on d, but also on ε, and so the mixing weights in the obtained mixture also depend on ε. If we instead chose a threshold that only depended on d, we would not have been able to obtain an effective error bound in the Kolmogorov distance step (i.e., Proposition 4.16) when $\gamma = 1/2$.

The key observation is that the arguments of Step 1 only require conditioning on input bits of degree larger than $n/O_d(1)$. In other words, if we perform some different conditioning κ on the set of input bits $S' \subseteq [m]$ above the weaker threshold $n/O_d(1)$, the distributions generated by the restricted functions will still have output weights which strongly concentrate around some integer multiple of $n/2^d$. Moreover, the mixing weights in this setting *are* integer multiples of $2^{-O_d(1)}$.

It remains to argue that the mixing weights on $\mathcal{M}$, the mixture we proved $f(\mathcal{U}^m)$ is close to by the stronger conditioning ρ, correspond to the mixing weights derived from the weaker conditioning κ (see Lemma 4.26). Since both conditionings produce mixtures of the same distribution, we have

$$\sum_\gamma \mathop{\mathbf{Pr}}_\rho [\gamma_\rho = \gamma] \cdot \mathop{\mathbb{E}}_{\rho:\gamma_\rho=\gamma} f_\rho(\mathcal{U}^{[m]\setminus S}) = f(\mathcal{U}^m) = \sum_\gamma \mathop{\mathbf{Pr}}_\kappa [\gamma_\kappa = \gamma] \cdot \mathop{\mathbb{E}}_{\kappa:\gamma_\kappa=\gamma} f_\kappa(\mathcal{U}^{[m]\setminus S'}).$$

By standard concentration bounds, the probability that $x \sim f(\mathcal{U}^m)$ has Hamming weight close to γn is almost entirely determined by the mass assigned to the distributions whose output weight concentrates around γ. In other words, we know $\mathbf{Pr}_\rho [\gamma_\rho = \gamma] \approx \mathbf{Pr}_\kappa [\gamma_\kappa = \gamma]$ for each γ. By our earlier analysis and the assumption that $f(\mathcal{U}^m)$ is close to $\mathcal{D}$, it must be that the mixing weight on $\mathsf{Bin}(n, \gamma)$ in $\mathcal{M}$ is roughly $\mathbf{Pr}_\rho [\gamma_\rho = \gamma]$, so it must also be close to $\mathbf{Pr}_\kappa [\gamma_\kappa = \gamma]$, which has the form we sought.

The analysis in the case of $\gamma = 1/2$ is similar, but as always, we need to keep track of the output

weight's parity. Here, the analogous equivalence is

$$\mathop{\mathbf{Pr}}_{\rho}\left[\gamma_\rho = 1/2 \wedge |f(\mathcal{U}^m)| \text{ is even}\right] \approx \frac{1}{2^{|S'|}} \sum_{\kappa:\gamma_\kappa = 1/2} \mathop{\mathbf{Pr}}_{x\sim\mathcal{U}^{[m]\setminus S'}}\left[|f_\kappa(x)| \text{ is even}\right]. \tag{2}$$

Since f is d-local with n output bits, the parity of $|f_\kappa(x)|$ can be represented as a degree-d $\mathbb{F}_2$-polynomial with $O_d(n)$ monomials. Thus, the mixing weight for the uniform distribution over n-bit strings of even Hamming weight in $\mathcal{M}$ must be close to the right-hand side of (2), which is the precise sense in which the mixing weights are "determined by low-degree $\mathbb{F}_2$-polynomials". The odd parity case is essentially identical. This concludes the proof overview of Theorem 1.2.

2.2 Technical Comparison to Prior Works

We now briefly survey the approaches of related prior works. To avoid an overly verbose digression, we attempt to focus only on the most relevant papers and do not hope to be exhaustive. Similarly, we restrict our attention within the mentioned works to the results and techniques most pertinent to our own; most, if not all, of the papers discussed contain a number of other interesting results.

The study of the complexity of sampling goes back to at least the 1980s in the influential work of Jerrum, Valiant, and Vazirani [JVV86]. In the context of shallow circuits, several clever sampling constructions (including the one in the introduction) were devised in [Bab87, BL87, IN96], while the first serious treatment of the complexity of sampling (with shallow circuits) appeared in [Vio12b]. There, Viola proved an assortment of sampling-related results, but we will confine our attention to two on the hardness of approximately generating $\mathcal{D}_{n/2}$, the uniform distribution over n-bit strings of Hamming weight $n/2$. (This specific choice is purely for simplicity, and both results also apply to the setting of $\mathcal{D}_{\alpha n}$ for $\alpha \in (0,1)$.)

The first result we will mention [Vio12b, Theorem 1.6] is an unconditional lower bound of $2^{-O(d)} - O(1/n)$ on the distance between $f(\mathcal{U}^m)$, where $f\colon \{0,1\}^m \to \{0,1\}^n$ is a d-local[4] function, and $\mathcal{D}_{n/2}$. The proof, much like our own (see Step 2), relies on a Kolmogorov distance bound obtained from a k-wise independence assumption [DGJ+10, GOWZ10]. If the output distribution $f(\mathcal{U}^m)$ is k-wise independent for some large integer k, then [GOWZ10, Theorem I.5] implies it has the wrong Hamming weight with constant probability, and we are done. Otherwise, there exists a k-tuple $T \subseteq [n]$ of output bits that are not uniformly distributed over $\{0,1\}^k$. In this case, one observes that the probability assigned to any element of $\{0,1\}^k$ is an integer multiple of 2^{-kd}, so the marginal distribution over T must be at least 2^{-kd}-far from uniform. The proof concludes by noting that the marginal distribution of $\mathcal{D}_{n/2}$ onto T is very close to uniform. Note that we use similar granularity ideas in the proof of Proposition 4.5 (in Step 1).

The second result [Vio12b, Theorem 1.3] provides a much stronger lower bound of $1 - 1/\mathsf{poly}(n)$ for the distance between $f(\mathcal{U}^m)$ and $\mathcal{D}_{n/2}$ for any $O(\log n)$-local function f, but is *conditional* on the input size m not substantially exceeding the information-theoretic minimum required to generate the target distribution. The proof begins by partitioning the input bits u of f as $u = (x, y)$ and expressing

$$f(u) = f(x,y) = h(y) \circ g_1(x_1, y) \circ \cdots \circ g_s(x_s, y),$$

where each g_i may depend arbitrarily on y but only on the single bit x_i of x. Additionally, each $g_i(x_i, y) \in \{0,1\}^{O(d)}$. By a greedy approach, this can be accomplished with $s \geq \Omega(n/d^2) = n/\mathsf{polylog}(n)$. The argument proceeds by conditioning on certain input bits (in this case y) and splitting into two scenarios depending on the amount of concentration; this strategy has been employed by several later works, including this one.

[4] In fact, the result is proven for an adaptive version of locality.

In the "concentrated" case where at least $\sqrt{n}$ many g_i's become fixed, the output distribution of $f(x, y)$ has small support. Even taking the union over all choices of such y, one finds that many elements in the support of $\mathcal{D}_{n/2}$ cannot be obtained. Note that this is where the input length restriction originates. Alternatively, at least $s - \sqrt{n} = n/\mathsf{polylog}(n)$ of the g_i's take multiple values. One would like to apply standard anticoncentration results to argue the output weight is too spread out, but as noted earlier in the proof overview (see Step 3), it may be that the Hamming weight of a g_i is fixed, even if the output is not. Viola circumvents this issue by adding an additional "test" to determine whether at most $2\sqrt{n}$ many g_i's can output the all zeros string; in this case, taking two or more values implies the weight cannot be fixed, and thus anticoncentration inequalities can be applied.

We now turn to the recent work of Filmus, Leigh, Riazanov, and Sokolov [FLRS23], who addressed the case of $\mathcal{D}_k$ for $k = o(n)$. They proved [FLRS23, Theorem 1.2] that $\widetilde{\Omega}(\log(n/k))$ bits of locality (even adaptively chosen) are required to generate $\mathcal{D}_k$ to constant error. Moreover, the same result holds for the uniform distribution over strings whose Hamming weight lies in a set S, where $\max_{s \in S} s = k$. Their proof first reduces to the case of considering $\mathcal{D}_1$, and then proceeds via a hitting set vs independent set dichotomy. This can again be viewed as fitting into the concentration vs anticoncentration paradigm. If there are $O(d2^d)$ input bits that together affect every output bit, then one can afford to condition on them and induct on the remaining $(d-1)$-local sub-distributions. Bounds from these distributions can eventually be recombined to deduce bounds on the full distribution using some version of a union bound, such as Lemma 3.3. Otherwise, there are $\Omega(2^d)$ independent output bits, and one can argue it is very likely that two of these bits evaluate to 1, since each (nonzero) output bit is 1 with probability at least 2^{-d}. Interestingly, the authors are able to use the *robust sunflower lemma* [Ros14, ALWZ21] from extremal combinatorics to improve their quantitative bounds; we refer the reader to their original paper for details.

Also recently, strong unconditional locality lower bounds for sampling $\mathcal{D}_{\alpha n}$ where α is a non-dyadic rational were proven independently in [Vio23, KOW24]. At a very high level, both proofs follow from the observation (recorded in [Vio12a]) that the marginal distribution on any output bit of a d-local function f must have probabilities that are integer multiples of 2^{-d}, which cannot accurately approximate (say) $1/3$, corresponding to the marginal distribution for $\mathcal{D}_{n/3}$. We begin by describing the approach of Viola [Vio23, Theorem 25], continuing to focus on the case of $\alpha = 1/3$ for simplicity. First assume by contradiction $\left\| f(\mathcal{U}^m) - \mathcal{D}_{n/3} \right\|_{\mathsf{TV}} \leq 1 - \varepsilon$ for some ε to be determined. The proof proceeds by arguing that there must exist a large subset $R \subseteq \{0,1\}^m$ of the inputs, such that $f(\mathcal{U}(R))$, f evaluated on inputs drawn uniformly at random from R, lands entirely in the support of $\mathcal{D}_{n/3}$ and has almost full min-entropy.[5] The so-called *fixed-set lemma* of Grinberg, Shaltiel, and Viola [GSV18] then implies that R can be further restricted to a large subset $R' \subseteq R$ whose uniform distribution is nearly indistinguishable by d-local functions from a product distribution where each bit is either fixed or uniform. Then the large min-entropy of $f(\mathcal{U}(R))$ (and thus $f(\mathcal{U}(R'))$) combined with the fact that the distribution is contained in the support of $\mathcal{D}_{n/3}$ implies there must be at least one output bit whose marginal distribution is extremely close to $1/3$. This, however, contradicts the fact that $1/3$ cannot accurately be approximated by an integer multiple of 2^{-d}.

The proof in [KOW24] instead operates very similarly to the details of Step 3. Through a graph-theoretic lemma [KOW24, Corollary 4.8], they show that there must be a set of at most $r/2^d$ many input bits upon whose conditioning results in r many independent output bits for $r \approx n/2^{d^2}$. As previously noted, each of these output bits inevitably incurs some error from mismatched marginal distributions, and these errors can be aggregated via standard concentration inequalities (as in

[5] Recall the *min-entropy* of a random variable X is $\log(1/\max_{x \in \mathsf{supp}(X)} \mathbf{Pr}[X = x])$.

Lemma 3.2). Bounds from the individual sub-distributions can once again be recombined via a union bound result like Lemma 3.3.

The work [KOW24] also contains locality bounds for $\mathcal{D}_{\alpha n}$ when α is a dyadic rational. The proof has strong similarities to the non-dyadic case, except that now the marginal distributions on output bits of f do not necessarily disagree with those of the target distribution. The authors overcome this issue by turning to the familiar concentration vs. anticoncentration paradigm. The details have almost entirely been spelled out already in Step 3, as our analysis is nearly identical at that part. We briefly note that the quantitative dependencies here are much worse than in the non-dyadic case, essentially for the reason that one needs to obtain a much richer structure than simply independent output bits, and so the analogous graph-theoretic lemma in this case has poor bounds.[6]

Finally, let us mention the predecessor of this work [KOW25], which classifies what uniform symmetric distributions can be sampled by functions of bounded locality, confirming a conjecture of [FLRS23]. In broad strokes, it follows similarly to the proof overview, so we will be brief. Through a strategic conditioning of input bits, one may reduce to the case where no input bits have large degree and almost all the output bits are k-wise independent. Then one can show the resulting Hamming weight distribution is close in Kolmogorov distance to the binomial distribution, and it satisfies a continuity-type property. Combining these steps together, one classifies the Hamming weight distribution, which is enough to classify the actual distribution, since we assume $f(\mathcal{U}^m)$ is close to a uniform symmetric distribution. Much of the present work is an extension of the ideas in [KOW25] to the general symmetric case, although there are several parts (e.g., much of the analysis in Step 1 and the finer control on the mixing weights) that require substantially new ideas.

We conclude by noting that several works [LV11, BIL12, Vio14, Vio20, CGZ22] prove hardness results against the more powerful classes of AC^0 circuits or read-once branching programs. There is a reasonable overlap in the techniques used with those discussed above, but the proofs of all these results rely on the special properties of certain pseudorandom objects like good codes [LV11, BIL12, CGZ22] or extractors [Vio12c, Vio20].

3 Preliminaries

For a positive integer n, we use $[n]$ to denote the set $\{1, 2, \ldots, n\}$. We use $\mathbb{R}$ to denote the set of real numbers, use $\mathbb{N} = \{0, 1, 2, \ldots\}$ to denote the set of natural numbers, and use $\mathbb{Z}$ to denote the set of integers. For a binary string x, we use $|x|$ to denote its Hamming weight. We use $\log(x)$ and $\ln(x)$ to denote the logarithm with base 2 and e respectively. For $a, b \in \mathbb{R}^{\geq 0}$, we use $a \pm b$ to shorthand a number in the interval $[a - b, a + b]$.

Asymptotics. We use the standard $O(\cdot), \Omega(\cdot), \Theta(\cdot)$ notation, and emphasize that in this paper they only hide universal positive constants that do not depend on any parameter. Occasionally we will use subscripts to suppress a dependence on particular variable (e.g., $O_d(1)$). The notation $\mathsf{poly}(\cdot)$ is also sometimes used to denote a quantity that is polynomial with an unspecified exponent. That is, $\mathsf{poly}(n) = \Theta(n^c)$ for some $c > 0$.

Locality and Hypergraphs. Let $f\colon \{0,1\}^m \to \{0,1\}^n$. We say f is a d-local function if each output bit $i \in [n]$ depends on at most d input bits. Unless otherwise stated, n, m, d are positive integers.

[6] Moreover, the bounds in this graph-theoretic lemma are essentially best possible; see [KOW24, Appendix A].

We sometimes take an alternative view, using hypergraphs to model the dependency relations in f. Let $G = (V, E)$ be an (undirected) hypergraph. For each $i \in V$, we use $I_G(i) \subseteq E$ to denote the set of edges that are incident to i. We say G has maximum degree d if $|I_G(i)| \leq d$ holds for all $i \in V$. Define $N_G(i) = \{i' \in V : I_G(i) \cap I_G(i') \neq \emptyset\}$ to be the neighborhood of i. We visualize the input-output dependencies of a function $f\colon \{0,1\}^m \to \{0,1\}^n$ as a hypergraph on the output bits $[n]$ with an edge for each input bit containing all of the output bits that depend on it. Note that a d-local function corresponds to a hypergraph with maximum degree d.

Probability. Let $\mathcal{P}$ be a (discrete) distribution. We use $x \sim \mathcal{P}$ to denote a random sample x drawn from the distribution $\mathcal{P}$. If $\mathcal{P}$ is a distribution over a product space, then we say $\mathcal{P}$ is a product distribution if its coordinates are independent. In addition, let S be a non-empty set. If S indexes $\mathcal{P}$, we use $\mathcal{P}[S]$ to denote the marginal distribution of $\mathcal{P}$ on coordinates in S. We reserve $\mathcal{U}$ to denote the uniform distribution over $\{0,1\}$.

For a deterministic function f, we use $f(\mathcal{P})$ to denote the output distribution of $f(x)$ given a random $x \sim \mathcal{P}$. For every event $\mathcal{E}$, we define $\mathcal{P}(\mathcal{E})$ to denote the probability that $\mathcal{E}$ occurs under distribution $\mathcal{P}$. In addition, we use $\mathcal{P}(x)$ to denote the probability mass of x under $\mathcal{P}$, and use $\mathsf{supp}\,(\mathcal{P}) = \{x : \mathcal{P}(x) > 0\}$ to denote the support of $\mathcal{P}$.

Let $\mathcal{Q}$ be a distribution. We use $\|\mathcal{P} - \mathcal{Q}\|_{\mathsf{TV}} = \frac{1}{2}\sum_x |\mathcal{P}(x) - \mathcal{Q}(x)|$ to denote their total variation distance.[7] We say $\mathcal{P}$ is ε-close to $\mathcal{Q}$ if $\|\mathcal{P}(x) - \mathcal{Q}(x)\|_{\mathsf{TV}} \leq \varepsilon$, and ε-far otherwise.

Fact 3.1. *Total variation distance has the following equivalent characterizations:*

$$\|\mathcal{P} - \mathcal{Q}\|_{\mathsf{TV}} = \max_{\text{event } \mathcal{E}} \mathcal{P}(\mathcal{E}) - \mathcal{Q}(\mathcal{E}) = \min_{\substack{\text{random variable } (X,Y) \\ X \text{ has marginal } \mathcal{P} \text{ and } Y \text{ has marginal } \mathcal{Q}}} \mathbf{Pr}\left[X \neq Y\right].$$

Let $\mathcal{P}_1, \ldots, \mathcal{P}_t$ be distributions. Then $\mathcal{P}_1 \times \cdots \times \mathcal{P}_t$ is a distribution denoting the product of $\mathcal{P}_1, \ldots, \mathcal{P}_t$. We also use $\mathcal{P}^t$ to denote $\mathcal{P}_1 \times \cdots \times \mathcal{P}_t$ if each $\mathcal{P}_i$ is the same as $\mathcal{P}$. For a finite set $S \subseteq [t]$, we use $\mathcal{P}^S$ to denote the distribution $\mathcal{P}^t$ restricted to the coordinates of S. We say distribution $\mathcal{P}$ is a mixture (or convex combination) of $\mathcal{P}_1, \ldots, \mathcal{P}_t$ if there exist $\alpha_1, \ldots, \alpha_t \in [0,1]$ such that $\sum_{i\in[t]} \alpha_i = 1$ and $\mathcal{P}(x) = \sum_{i\in[t]} \alpha_i \cdot \mathcal{P}_i(x)$ for all x in the sample space. When it is clear from context, we will occasionally write mixtures more simply as $\mathcal{P} = \sum_{i\in[t]} \alpha_i \cdot \mathcal{P}_i$. In the case where the $\mathcal{P}_i$ are all of the form $f_i(\mathcal{U}^m)$ for deterministic functions $f_1, \ldots, f_t\colon \{0,1\}^m \to \{0,1\}^n$, we will occasionally abuse notation by writing $F(\mathcal{U}^m)$ for the mixture $F = \sum_{i\in[t]} \alpha_i \cdot f_i(\mathcal{U}^m)$.

We collect two useful probabilistic results from prior work. The first says that two distributions must be far apart if many of their marginals do not match.

Lemma 3.2 ([KOW24, Lemma 4.2])**.** *Let $\mathcal{P}$, $\mathcal{Q}$, and $\mathcal{W}$ be distributions over an n-dimensional product space, and let $S \subseteq [n]$ be a non-empty set of size s. Assume*

- *$\mathcal{P}[S]$ and $\mathcal{W}[S]$ are two product distributions,*
- *$\|\mathcal{P}[\{i\}] - \mathcal{W}[\{i\}]\|_{\mathsf{TV}} \geq \varepsilon$ holds for all $i \in S$, and*
- *$\mathcal{W}(x) \geq \nu \cdot \mathcal{Q}(x)$ holds for some $\nu > 0$ and all x.*

Then

$$\|\mathcal{P} - \mathcal{Q}\|_{\mathsf{TV}} \geq 1 - 2 \cdot e^{-\varepsilon^2 s/2}/\nu.$$

The second allows us to reason about the distance between a fixed distribution and a mixture by reasoning about the individual distributions composing the mixture.

[7] To evaluate total variation distance, we need two distributions to have the same sample space. This will be clear throughout the paper and thus we omit it for simplicity.

Lemma 3.3 ([Vio20, Corollary 4.2]). *Let $\mathcal{P}_1, \ldots, \mathcal{P}_t$ and $\mathcal{Q}$ be distributions. Assume there exists a value ε such that $\|\mathcal{P}_i - \mathcal{Q}\|_{\mathsf{TV}} \geq 1 - \varepsilon$ for all $i \in [t]$. Then for the balanced mixture $\mathcal{P} = \sum_i \frac{1}{t} \cdot \mathcal{P}_i$, we have*

$$\|\mathcal{P} - \mathcal{Q}\|_{\mathsf{TV}} \geq 1 - t \cdot \varepsilon.$$

Weight Distributions and Symmetrization. If $\mathcal{P}$ is a distribution over $\{0,1\}^n$, we use $|\mathcal{P}|$ to denote the distribution over weights. That is, $|\mathcal{P}|(w) = \sum_{x:|x|=w} \mathcal{P}(x)$. We additionally define the symmetrized distribution $\mathcal{P}_{\text{sym}}$ to be the distribution resulting from randomly permuting the coordinates of a string $x \sim \mathcal{P}$.

We will require the following lemma, which lets us control the distance between two distributions via the distance between their weight distributions, assuming one distribution is symmetric and the other is close to being symmetric.

Lemma 3.4 ([KOW25, Lemma 4.8]). *Let $\mathcal{P}$ and $\mathcal{Q}$ be two distributions on $\{0,1\}^n$ with $\mathcal{Q}$ symmetric. Then*

$$\|\mathcal{P} - \mathcal{Q}\|_{\mathsf{TV}} = \Theta(\||\mathcal{P}| - |\mathcal{Q}|\|_{\mathsf{TV}} + \|\mathcal{P} - \mathcal{P}_{\text{sym}}\|_{\mathsf{TV}}).$$

Binomials and Entropy. Let $\mathcal{H}(x) = x \cdot \log\left(\frac{1}{x}\right) + (1-x) \cdot \log\left(\frac{1}{1-x}\right)$ be the binary entropy function. We will use the following estimate regarding binomial coefficients and the entropy function.

Fact 3.5 (See e.g., [CT06, Lemma 17.5.1]). *For $1 \leq k \leq n-1$, we have*

$$\binom{n}{k} \geq \frac{2^{n \cdot \mathcal{H}(k/n)}}{\sqrt{8k(1-k/n)}}.$$

For positive integer n and parameter $\gamma \in [0,1]$, define $\mathsf{Bin}(n,\gamma)$ to be the binomial distribution of n bits and bias γ, i.e., $x \sim \mathsf{Bin}(n,\gamma)$ is a random integer between 0 and n with probability density function $\binom{n}{x}\gamma^x(1-\gamma)^{n-x}$. We need the following standard estimates, the proofs of which can be found in Appendix B.

Fact 3.6. *Let $\gamma \in (0,1)$, $a < b \in \mathbb{R}$, and $n \in \mathbb{N}^{\geq 1}$. Then the binomial distribution $\mathsf{Bin}(n,\gamma)$ satisfies*

$$\mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) \in [a,b]\right] \leq O\left(\frac{\lceil b-a \rceil}{\sqrt{\gamma(1-\gamma)n}}\right).$$

Fact 3.7. *Let $\gamma \in (0,1)$, $a, b \in \mathbb{N}$, and $n \in \mathbb{N}^{\geq 1}$. Then the binomial distribution $\mathsf{Bin}(n,\gamma)$ satisfies*

$$\left|\mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) = a\right] - \mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) = b\right]\right| \leq O\left(\frac{|b-a|}{\gamma(1-\gamma)n}\right).$$

Concentration and Anti-concentration. We need the following standard (anti-)concentration bounds.

Fact 3.8 (Hoeffding's Inequality). *Assume $X_1, \ldots, X_n$ are independent random variables such that $a \leq X_i \leq b$ holds for all $i \in [n]$. Then for all $\delta \geq 0$, we have*

$$\max\left\{\mathbf{Pr}\left[\frac{1}{n}\sum_{i\in[n]}(X_i - \mathbb{E}[X_i]) \geq \delta\right], \mathbf{Pr}\left[\frac{1}{n}\sum_{i\in[n]}(X_i - \mathbb{E}[X_i]) \leq -\delta\right]\right\} \leq \exp\left\{-\frac{2n\delta^2}{(b-a)^2}\right\}.$$

Fact 3.9 (Chernoff's Inequality)**.** *Assume $X_1, \ldots, X_n$ are independent random variables such that $X_i \in [0,1]$ holds for all $i \in [n]$. Let $\mu = \sum_{i\in[n]} \mathbb{E}[X_i]$. Then for all $\delta \in [0,1]$, we have*

$$\mathbf{Pr}\left[\sum_{i\in[n]} X_i \leq (1-\delta)\mu\right] \leq \exp\left\{-\frac{\delta^2\mu}{2}\right\}.$$

Fact 3.10 (See e.g., [DFKO06, Lemma 2.9])**.** *Assume $f\colon \{0,1\}^n \to \mathbb{R}$ is a degree k polynomial. Let $\mu = \mathbb{E}_{x\sim\{0,1\}^n}[f(x)]$. Then*

$$\mathop{\mathbf{Pr}}_{x\sim\{0,1\}^n}[f(x) \geq \mu] \geq 2^{-O(k)}.$$

Recall that random variables $X_1, \ldots, X_n$ over some domain D are called *k-wise independent* if for any values $d_1, \ldots, d_k \in D$ and indices $i_1, \ldots, i_k \in [n]$, we have

$$\mathbf{Pr}\left[X_{i_1} = d_1, \ldots, X_{i_k} = d_k\right] = \prod_{j=1}^{k} \mathbf{Pr}\left[X_{i_j} = d_j\right].$$

Fact 3.11 (See e.g., [BR94, Lemma 2.2])**.** *Let $k \geq 4$ be an even integer. Suppose $X_1, \ldots, X_n$ are k-wise independent random variables taking values in $[0,1]$. Let $X = X_1 + \cdots + X_n$ and $t > 0$. Then,*

$$\mathbf{Pr}\left[|X - \mathbb{E}[X]| \geq t\right] \leq C_k \cdot \left(\frac{nk}{t^2}\right)^{k/2},$$

where $C_k = 2\sqrt{\pi k} \cdot e^{1/(6k)} \cdot e^{-k/2} \leq 1.0004$.

4 The Characterization

In this section, we prove our main result. Recall `evens` and `odds` denote the uniform distribution over n-bit strings of even Hamming weight and odd Hamming weight, respectively.

Theorem 4.1. *Let $f\colon \{0,1\}^m \to \{0,1\}^n$ be a d-local function. Assume $f(\mathcal{U}^m)$ is ε-close to a symmetric distribution $\mathcal{D}$ over $\{0,1\}^n$. Then if n is sufficiently large in terms of d and ε, $f(\mathcal{U}^m)$ is $O_d\left(\frac{1}{\log(1/\varepsilon)}\right)^{1/5}$-close to a distribution of the form*

$$\sum_{\substack{a\in[0,2^d]\cap\mathbb{Z}\\ a\neq 2^{d-1}}} c_a \cdot \mathcal{U}^n_{a/2^d} + c_e \cdot \texttt{evens} + c_o \cdot \texttt{odds},$$

where each $c_a = c'_a/2^C$ for some integer $0 \leq c'_a \leq 2^C$ and a fixed integer $C = O_d(1)$. Moreover, there exist at most 2^C many degree-d $\mathbb{F}_2$-polynomials $\{p_i\colon \mathbb{F}_2^m \to \mathbb{F}_2\}$, each with $O_d(n)$ monomials, such that

$$c_e = \frac{1}{2^C} \cdot \sum_i \mathop{\mathbf{Pr}}_{x\sim\mathcal{U}^m}[p_i(x) = 0] \quad \text{and} \quad c_o = \frac{1}{2^C} \cdot \sum_i \mathop{\mathbf{Pr}}_{x\sim\mathcal{U}^m}[p_i(x) = 1].$$

Before proceeding to the proof, we make several remarks.

Remark 4.2. Any distribution of this form can be exactly produced by an NC^0 function (with additional input bits and locality):

- Use C bits of locality to select either a product distribution $\mathcal{U}^n_{a/2^d}$ or one of the $\mathbb{F}_2$-polynomials $p_i \colon \mathbb{F}_2^m \to \mathbb{F}_2$.
- If some $\mathcal{U}^n_{a/2^d}$ is selected, we can sample from it with an additional d bits of locality.
- Otherwise an $\mathbb{F}_2$-polynomial $p_i \colon \mathbb{F}_2^m \to \mathbb{F}_2$ is selected. In this case, we wish to produce the distributions `evens` and `odds` with probability $\mathbf{Pr}_{x \sim \mathcal{U}^m}[p_i(x) = 0]$ and $\mathbf{Pr}_{x \sim \mathcal{U}^m}[p_i(x) = 1]$, respectively. Arbitrarily partition the $O_d(n)$ monomials of p_i into n bins of size $O_d(1)$, and set $y \in \{0,1\}^n$ to have the j-th coordinate equal to the sum of the monomials in the j-th bin. Since p_i has degree d, this can be done with $O_d(1)$ bits of locality. Note that y has the right weight distribution, but it may not be symmetric. To remedy this, we sample $z \sim$ `evens` (with fresh randomness) and output $y \oplus z$.

Remark 4.3. An alternative formulation of the last conclusion of Theorem 4.1 is that there exists a degree-$O_d(1)$ $\mathbb{F}_2$-polynomial $P(x, y) = \sum_i \mathbb{1}(x = i) \cdot p_i(y)$[8] with $O_d(n)$ monomials such that $c_e = 2^{-C} \cdot \mathbf{Pr}_z[P(z) = 0]$ and $c_o = 2^{-C} \cdot \mathbf{Pr}_z[P(z) = 1]$. In this formulation, we can still exactly produce distributions of this form via a similar algorithm to the one in Remark 4.2, only now each output bit requires larger locality to compensate for P's larger degree.

Remark 4.4. We have chosen to focus on the most commonly studied setting where the random bits fed to f are unbiased. For readers interested in more general input biases, we note that a similar result to Theorem 4.1 (with the biases of the product distributions and the mixing weights appropriately adjusted) should be provable using the techniques presented here. It is important, however, that the input bits are identically distributed, as the first of our four main steps (see Subsection 4.1) requires the possible output biases to lie in a discrete set.

Our proof will proceed via the four steps outlined in Section 2, each corresponding to its own subsection.

4.1 Removing Large Influences

Our first step is to argue that after conditioning on the high degree input bits of f, almost any small collection of output bits looks identical to those of a γ-biased distribution $\mathcal{U}^n_\gamma$, where γ is an integer multiple of 2^{-d}.

Proposition 4.5. *Let $f \colon \{0,1\}^m \to \{0,1\}^n$ be a d-local function, and let $A \geq 2^{100d}$ be a parameter. Assume $f(\mathcal{U}^m)$ is ε-close to a symmetric distribution over $\{0,1\}^n$ for some $\varepsilon < 2^{-cdA}$, where $c > 0$ is a sufficiently large absolute constant. Further assume that n is sufficiently large in terms of d. Define $S \subseteq [m]$ to be the set of input bits with degree at least n/A.*

Let $k \leq \log(1/\varepsilon)/C_d$ be an arbitrary integer, where $C_d \geq 1$ is a sufficiently large constant depending only on d. Then for each conditioning $\rho \in \{0,1\}^S$ on the bits in S, there exists a subset $T_\rho \subseteq [n]$ of size $|T_\rho| \leq O_{d,k,A}(1)$ such that every k-tuple of output bits in $[n] \setminus T_\rho$ has distribution $\mathcal{U}^k_{\gamma_\rho}$, where $\gamma_\rho = a_\rho/2^d$ and $0 \leq a_\rho \leq 2^d$ is an integer.

In our analysis, we will often need to consider the distance between some bias and its closest integer multiple of 2^{-d}, so we introduce the following notation.

Definition 4.6 (Binary Representation Error)**.** For each $d \in \mathbb{N}$, we use $\mathsf{err}(\gamma, d)$ to denote the minimum distance of γ to an integer multiple of 2^{-d}. In particular, given a binary representation

[8] In an abuse of notation, we identify an integer i with its binary representation.

of γ as $\gamma = \sum_{i \in \mathbb{Z}} a_i \cdot 2^i$ where each $a_i \in \{0,1\}$, we have

$$\mathsf{err}(\gamma, d) = \min \left\{ \sum_{i<-d} a_i \cdot 2^i, \sum_{i<-d} (1-a_i) \cdot 2^i \right\}.$$

The proof of Proposition 4.5 involves a number of similar looking estimates, so we provide a brief overview of the remainder of the section before continuing. It is known from previous work [KOW24] that local functions cannot accurately sample $\mathcal{D}_k$, the uniform distribution over n-bit binary strings of Hamming weight k, so long as k/n has large binary representation error.

Lemma 4.7 ([KOW24, Theorem 5.7])**.** *Let $f\colon \{0,1\}^m \to \{0,1\}^n$ be a d-local function, and let $1 \le k \le n-1$ be an integer. If $\mathsf{err}(k/n, d) \ge \delta$ for some $\delta > 0$, then*

$$\|f(\mathcal{U}^m) - \mathcal{D}_k\|_{\mathsf{TV}} \ge 1 - 4\sqrt{2n} \cdot \exp\left\{-n \cdot \delta^{40d}\right\}.$$

This implies that with high probability, the output weight $|f(x)|$ is close to *some* dyadic rational multiple of n, at least to the degree that f is symmetric (Lemma 4.8). In order to ensure $|f(x)|$ is concentrated around a *fixed* dyadic rational, we condition on input bits of degree at least n/A. This bounds the variance of the weight of f to provide good concentration (Lemma 4.9). However, there is still a chance that some non-negligible fraction of the time, the output weight is close to a different dyadic rational multiple of n. In this case, we can show (Claim 4.14) that the weight distribution must also assign decent probability mass to the weights between these dyadic multiples, which by Lemma 4.7 would contradict our original assumption on the distance between $f(\mathcal{U}^m)$ and $\mathcal{D}$.

Now we proceed toward proving Proposition 4.5. Recall that any symmetric distribution $\mathcal{D}$ is simply a mixture of $\mathcal{D}_k$ for different values of k. Thus, if $f(\mathcal{U}^m)$ is close to a symmetric distribution, Lemma 4.7 implies most of the output weight must have bias close to some multiple of 2^{-d}.

Lemma 4.8. *Let $f\colon \{0,1\}^m \to \{0,1\}^n$ be a d-local function. Assume $f(\mathcal{U}^m)$ is ε-close to a symmetric distribution $\mathcal{D}$ over $\{0,1\}^n$. Then*

$$\Pr_{x \sim \mathcal{U}^m}\left[\mathsf{err}\left(\frac{|f(x)|}{n}, d\right) \ge \frac{1}{n^{1/(800d)}}\right] \le O\left(\varepsilon + e^{-n^{0.9}}\right), \tag{3}$$

where we recall that $\mathsf{err}(\gamma, d)$ is the minimum distance between γ and integer multiples of 2^{-d}.

Proof. Without loss of generality, assume n is sufficiently large. If

$$\Pr_{z \sim \mathcal{D}}\left[\mathsf{err}\left(\frac{z}{n}, d\right) \ge \frac{1}{n^{1/(800d)}}\right] \le 100 \cdot \left(\varepsilon + e^{-n^{0.9}}\right), \tag{4}$$

then (3) holds due to the assumption on $f(\mathcal{U}^m)$. Now we assume (4) does not hold.

For each $0 \le k \le n$, recall that $\mathcal{D}_k$ is the uniform distribution over Hamming weight k strings. Then $\mathcal{D}$ is a mixture of $\{\mathcal{D}_k\}$, i.e., $\mathcal{D} = \sum_k \alpha_k \cdot \mathcal{D}_k$. We say k is *bad* if $\mathsf{err}(k/n, d) \ge n^{-1/(800d)}$. Then the violation of (4) is equivalent to

$$\sum_{\text{bad } k} \alpha_k > 100 \cdot \left(\varepsilon + e^{-n^{0.9}}\right). \tag{5}$$

By Lemma 4.7, for each bad k, there is an event $\mathcal{E}_k$ such that

- under $f(\mathcal{U}^m)$, it happens with probability at most $4\sqrt{2n} \cdot e^{-n^{0.95}}$;
- under $\mathcal{D}_k$, it happens with probability at least $1 - 4\sqrt{2n} \cdot e^{-n^{0.95}} \geq 1/2$.

Hence considering $\mathcal{E} = \bigvee_{\text{bad } k} \mathcal{E}_k$, we have

- under $f(\mathcal{U}^m)$, it happens with probability at most $4n\sqrt{2n} \cdot e^{-n^{0.95}}$ which is at most $e^{-n^{0.9}}$, since we assumed n is sufficiently large;
- under $\mathcal{D}$, it happens with probability at least

$$\sum_{\text{bad } k} \alpha_k \cdot \left(1 - 4\sqrt{2n} \cdot e^{-n^{0.95}}\right) \geq 50 \cdot \left(\varepsilon + e^{-n^{0.9}}\right).$$

This means $f(\mathcal{U}^m)$ is not ε-close to $\mathcal{D}$, a contradiction. □

By conditioning on the high degree input bits, we can reduce the variance of the output weight distribution to obtain a version of Lemma 4.8 where the output weight is concentrated around a *fixed* dyadic rational multiple of n.

Lemma 4.9. *Let $f\colon \{0,1\}^m \to \{0,1\}^n$ be a d-local function with $d \geq 1$, and let $A \geq 2^{100d}$ be a parameter. Assume $f(\mathcal{U}^m)$ is ε-close to a symmetric distribution over $\{0,1\}^n$ for some $\varepsilon < 2^{-cdA}$, where $c > 0$ is a sufficiently large absolute constant. Further assume that n is sufficiently large in terms of d. Define $S \subseteq [m]$ to be the set of input bits with degree at least n/A. Then for each conditioning $\rho \in \{0,1\}^S$ on bits in S, there exists an integer $0 \leq a_\rho \leq 2^d$ such that*

$$\mathop{\mathbf{Pr}}_{x \sim \mathcal{U}^{[m]\setminus S}}\left[\left|\frac{|f(x,\rho)|}{n} - \frac{a_\rho}{2^d}\right| \geq \frac{1}{n^{1/(800d)}}\right] \leq \mathsf{poly}(\varepsilon).$$

For clarity, we prove Lemma 4.9 through a series of claims, the most routine of which have their proofs deferred to Appendix C. The high-level idea is to use the second moment method to show that for any restriction ρ on the high degree input bits, the Hamming weight of $f(x,\rho)$ is typically close to its expectation, which by the previous lemma must be close to a multiple of 2^{-d}. We then turn to a more involved argument to boost the quantitative behavior of these bounds.

Proof. First note that $|S| \leq dA$. Fix an arbitrary $\rho \in \{0,1\}^S$ and define $g\colon \{0,1\}^{m-|S|} \to \{0,1,\ldots,n\}$ by $g(x) = |f(x,\rho)|$. Then $g(x) = \sum_{i=1}^n g_i(x)$, where each $g_i\colon \{0,1\}^{m-|S|} \to \{0,1\}$ is a d-junta (i.e., depends on at most d of its input bits) and every input bit appears in at most n/A different g_i's. Therefore

$$\begin{aligned}
\mathrm{Var}[g] &= \sum_{i,j\in[n]} \mathrm{Cov}\left(g_i, g_j\right) \\
&\leq \sum_{i\in[n]} \#\left\{j \in [n] \,\middle|\, g_j \text{ correlates with } g_i\right\} && \text{(since each } g_i \text{ is Boolean)} \\
&\leq n \cdot dn/A \leq dn^2/2^{100d}. && \text{(since } A \geq 2^{100d})
\end{aligned}$$

Define $p = \mathbb{E}[g]/n$. Then by Chebyshev's inequality, we have

$$\mathop{\mathbf{Pr}}_{x\sim\{0,1\}^{[m]\setminus S}}\left[\left|\frac{|f(x,\rho)|}{n} - p\right| > 2^{-30d}\right] = \mathbf{Pr}\left[\left|\frac{g(x)}{n} - p\right| > 2^{-30d}\right] \leq \frac{d}{2^{40d}} \leq \frac{1}{4}. \tag{6}$$

Combining the above fact that $|f(x,\rho)|/n$ is typically close to p with the fact that it must also typically be close to an integer multiple of 2^{-d} (by Lemma 4.8), we obtain the following claim.

Claim 4.10 (Proved in Appendix C)**.** $\mathsf{err}(p,d) \le 2 \cdot 2^{-30d}$.

By Claim 4.10, there exists an integer $0 \le a \le 2^d$ such that $|p - a/2^d| \le 2 \cdot 2^{-30d}$. Now it suffices to show

$$\mathop{\mathbf{Pr}}_{x \sim \{0,1\}^{[m]\setminus S}} \left[\left| \frac{|f(x,\rho)|}{n} - \frac{a}{2^d} \right| > \frac{1}{n^{1/(800d)}} \right] \le \mathsf{poly}(\varepsilon). \tag{7}$$

We start with two primitive bounds. The first follows from combining Lemma 4.8 with the observation that any event is assigned at most $2^{|S|}$ times more mass by $f(\mathcal{U}^{[m]\setminus S}, \rho)$ than by $f(\mathcal{U}^m)$.

Claim 4.11 (Proved in Appendix C)**.**

$$\mathop{\mathbf{Pr}}_{x \sim \{0,1\}^{[m]\setminus S}} \left[\mathsf{err}\left(\frac{|f(x,\rho)|}{n}, d \right) > \frac{1}{n^{1/(800d)}} \right] \le \mathsf{poly}(\varepsilon).$$

Claim 4.12.

$$\mathop{\mathbf{Pr}}_{x \sim \{0,1\}^{[m]\setminus S}} \left[\left| \frac{|f(x,\rho)|}{n} - \frac{a}{2^d} \right| > \frac{3}{2^{30d}} \right] \le \frac{1}{4}.$$

Proof of Claim 4.12. This follows directly from (6) and Claim 4.10. □

Define

$$\delta \coloneqq \mathop{\mathbf{Pr}}_{x \sim \{0,1\}^{[m]\setminus S}} \left[\left| \frac{|f(x,\rho)|}{n} - \frac{a}{2^d} \right| > 4^{-d} \right].$$

Then by Claim 4.11, we can relate the LHS of (7) to δ, because it is very unlikely that $|f(x,\rho)|/n$ is between $n^{-1/(800d)}$- and 4^{-d}-close to $a/2^d$.

Claim 4.13 (Proved in Appendix C)**.**

$$\mathop{\mathbf{Pr}}_{x \sim \{0,1\}^{[m]\setminus S}} \left[\left| \frac{|f(x,\rho)|}{n} - \frac{a}{2^d} \right| > \frac{1}{n^{1/(800d)}} \right] \le \delta + \mathsf{poly}(\varepsilon).$$

Claim 4.12 implies a constant $1/4$ upper bound on δ. To further improve it, we prove the following claim.

Claim 4.14.

$$\mathop{\mathbf{Pr}}_{x \sim \{0,1\}^{[m]\setminus S}} \left[8^{-d} \le \left| \frac{|f(x,\rho)|}{n} - \frac{a}{2^d} \right| \le 4^{-d} \right] \ge \delta / 2^{20d}.$$

Such a claim is true, because if $|f(x,\rho)|/n$ is typically close to $a/2^d$ but has a δ probability of being far from it, then there must be a decent probability that $|f(x,\rho)|/n$ is close, but not too close, to $a/2^d$, since $|f(x,\rho)|/n$ is roughly continuous in x. Formally, the proof of Claim 4.14 will operate via a coupling argument. We consider two independent inputs x and z where $|f(x,\rho)|/n$ is close to $a/2^d$ but $|f(z,\rho)|/n$ is not. By slowly changing x into z by flipping bits of x on which they disagree in a random order, we can find inputs y where $|f(y,\rho)|/n$ is in the range in question.

Proof of Claim 4.14. Let π be a uniformly random permutation on $[m] \setminus S$. Pick t uniformly at random among $0, 1, \ldots, m - |S|$ and sample $r \sim \mathrm{Bin}(m - |S|, 1/2)$ (i.e., $\mathbf{Pr}[r = s] = 2^{|S|-m}\binom{m-|S|}{s}$ for all $s = 0, 1, \ldots, m$). Define $y \in \{0,1\}^{[m]\setminus S}$ as x with the first t bits in π flipped; and define $z \in \{0,1\}^{[m]\setminus S}$ as x with the first r bits in π flipped. Since y has the same distribution as x, it suffices to show

$$\mathop{\mathbf{Pr}}_{x,z,\pi,t} \left[8^{-d} \le \left| \frac{|f(y,\rho)|}{n} - \frac{a}{2^d} \right| \le 4^{-d} \right] \ge \delta / 2^{20d}. \tag{8}$$

Observe that z is uniform over $\{0,1\}^{[m]\setminus S}$ and is independent of x. Define $\mathcal{E}_x$ to be the event that $\left|\frac{|f(x,\rho)|}{n} - \frac{a}{2^d}\right| > 4^{-d}$ and $\mathcal{E}_z$ to be the event that $\left|\frac{|f(z,\rho)|}{n} - \frac{a}{2^d}\right| \le 3/2^{30d}$. Then by independence and Claim 4.12, we have

$$\mathop{\mathbf{Pr}}_{x,z}\left[\mathcal{E}_z \,|\, \mathcal{E}_x\right] \ge \frac{3}{4}. \tag{9}$$

Recall we wish to show that if $\mathcal{E}_x$ and $\mathcal{E}_z$ both hold, there is a good probability (over t) that $|f(y,\rho)|/n$ is between 8^{-d}- and 4^{-d}-close to $a/2^d$. Note that changing t by 1 only changes $|f(y,\rho)|$ by at most the degree of the relevant input. Since $|f(y,\rho)|$ must pass through the "bad" region, (8) holds as long as it does not pass through too quickly. This is captured by the following event $\mathcal{E}_\pi$.

Denote $m' = m - |S|$ and for each $j \in [m']$ we use $\deg_\pi(j)$ to denote the degree of the j-th input bit under π. Define $L = \lfloor m'/2^{10d} \rfloor$. Let $\mathcal{E}_\pi$ be the event that no L consecutive (under π) input bits starting at a multiple of L have degree sum larger than $n/2^{5d}$; or more formally that $\sum_{j=1}^{L} \deg_\pi(L\cdot i + j) \le n/2^{5d}$ holds for each $i = 0, 1, \ldots, \lfloor m'/L \rfloor$. Whenever $\mathcal{E}_\pi$ holds, we know that any L consecutive (under π) bit flips of x will only change the output weight of f by at most $2\cdot n/2^{5d}$, since any length L interval can only intersect two length L intervals that start at a multiple of L (as in the definition of $\mathcal{E}_\pi$). For $i = 0$, we have

$$\mathop{\mathbb{E}}_{\pi}\left[\sum_{j=1}^{L} \deg_\pi(j)\right] = L \cdot \mathop{\mathbb{E}}_{\pi}\left[\deg_\pi(1)\right] \le \frac{dnL}{m'} \le \frac{dn}{2^{10d}}$$

as the total degree is at most dn. Define the indicator variable I_i to denote that the i-th input bit (in the standard order) is in $\pi(1), \ldots, \pi(L)$. Let $\deg(i)$ be the degree of the i-th input bit (in the standard order). Then we also have

$$\begin{aligned}
\mathop{\mathrm{Var}}_{\pi}\left[\sum_{j=1}^{L} \deg_\pi(j)\right] = \mathop{\mathrm{Var}}_{\pi}\left[\sum_{i=1}^{m'} \deg(i)\cdot I_i\right] &= \sum_i \deg(i)^2 \mathop{\mathrm{Var}}_{\pi}\left[I_i\right] + \sum_{i\neq i'} \deg(i)\deg(i')\,\mathrm{Cov}\left[I_i, I_{i'}\right] \\
&\le \sum_i \deg(i)^2 \cdot \mathop{\mathrm{Var}}_{\pi}\left[I_i\right] && \left(\text{since } \mathrm{Cov}\left[I_i, I_{i'}\right] = \tfrac{L(L-1)}{m'(m'-1)} - \left(\tfrac{L}{m'}\right)^2 \le 0\right) \\
&\le \frac{L}{m'}\sum_i \deg(i)^2 && \left(\text{since } \mathrm{Var}_\pi\left[I_i\right] \le \mathbf{Pr}_\pi\left[I_i = 1\right]\right) \\
&\le \frac{L}{m'} \cdot \max_i \deg(i) \cdot \sum_i \deg(i) \le \frac{L}{m'} \cdot \frac{n}{2^{100d}} \cdot dn,
\end{aligned}$$

since $\deg(i) \le n/A \le n/2^{100d}$ and $\sum_i \deg(i) \le dn$. Therefore

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{\pi}\left[\sum_{j=1}^{L} \deg_\pi(j) \ge n/2^{5d}\right] &\le \mathop{\mathbf{Pr}}_{\pi}\left[\sum_{j=1}^{L} \deg_\pi(j) - \mathbb{E}\left[\sum_{j=1}^{L} \deg_\pi(j)\right] \ge n/2^{5d+1}\right] \\
&\qquad \left(\text{since } 2^{-5d} - d\cdot 2^{-10d} \ge 2^{-5d}/2 \text{ for } d \ge 1\right) \\
&\le \frac{\mathrm{Var}_\pi\left[\sum_{j=1}^{L} \deg_\pi(j)\right]}{4n^2 \cdot 2^{-10d}} && \text{(by Chebyshev inequality)} \\
&\le \frac{d\cdot L}{4\cdot 2^{90d}\cdot m'}.
\end{aligned}$$

Note that the same estimate works for all $i = 1, \ldots, \lfloor m'/L \rfloor$. Hence by independence and a union bound

$$\mathop{\mathbf{Pr}}_{x,\pi}[\mathcal{E}_\pi \mid \mathcal{E}_x] = \mathop{\mathbf{Pr}}_{\pi}[\mathcal{E}_\pi] \geq 1 - \left(1 + \left\lfloor \frac{m'}{L} \right\rfloor\right) \cdot \frac{dL}{4 \cdot 2^{90d} \cdot m'} \geq 1 - \frac{d}{2^{20d}} \geq \frac{7}{8}. \tag{10}$$

Combining (9) and (10), we have $\mathbf{Pr}_{x,z,\pi}[\mathcal{E}_\pi \wedge \mathcal{E}_z \mid \mathcal{E}_x] \geq 5/8$ by a union bound. Since $\mathbf{Pr}_x[\mathcal{E}_x] = \delta$, we have

$$\mathop{\mathbf{Pr}}_{x,z,\pi}[\mathcal{E}_x \wedge \mathcal{E}_z \wedge \mathcal{E}_\pi] \geq \frac{5\delta}{8}. \tag{11}$$

Recall that z is x with the first r bits (in the order of π) flipped. Assuming $\mathcal{E}_x$ and $\mathcal{E}_z$, we know that $\left|\frac{|f(x,\rho)|}{n} - \frac{a}{2^d}\right| > 4^{-d}$ and $\left|\frac{|f(z,\rho)|}{n} - \frac{a}{2^d}\right| \leq 3/2^{30d}$. For each $i = 0, 1, \ldots, m'$, we use $z^{(i)}$ to denote the string x with the first i bits (in the order of π) flipped. Then $z^{(0)} = x$ and $z^{(r)} = z$. Note that the Hamming difference between $f(z^{(i)}, \rho)$ and $f(z^{(i-1)}, \rho)$ is upper bounded by the degree of the i-th input bit (under π), which is further upper bounded by $n/2^{100d}$ by our construction. Since $4^{-d} > 6^{-d} > 3/2^{30d}$, there exists some $i^* \in \{0, 1, \ldots, r\}$ such that

$$6^{-d} \leq \left|\frac{\left|f(z^{(i^*)}, \rho)\right|}{n} - \frac{a}{2^d}\right| \leq 6^{-d} + 2^{-100d}.$$

Assuming $\mathcal{E}_\pi$, for each $j = i^*, i^* - 1, \ldots, i^* - L$ the Hamming difference between $f(z^{(j)}, \rho)$ and $f(z^{(i^*)}, \rho)$ is upper bounded by the total degree of input bits from i^* to $i^* - L$, which is at most $2 \cdot n/2^{5d}$. Hence for each such j, we have

$$8^{-d} \leq 6^{-d} - 2^{-5d+1} \leq \left|\frac{\left|f(z^{(j)}, \rho)\right|}{n} - \frac{a}{2^d}\right| \leq 6^{-d} + 2^{-100d} + 2^{-5d+1} \leq 4^{-d}.$$

In particular, this implies $i^* - L > 0$ as $\left|\frac{|f(z^{(0)},\rho)|}{n} - \frac{a}{2^d}\right| = \left|\frac{|f(x,\rho)|}{n} - \frac{a}{2^d}\right| > 4^{-d}$. At this point, recall that $y = z^{(t)}$ where $t \sim \{0, 1, \ldots, m'\}$. Hence

$$\begin{aligned}
&\mathop{\mathbf{Pr}}_{x,z,\pi,t}\left[8^{-d} \leq \left|\frac{|f(y,\rho)|}{n} - \frac{a}{2^d}\right| \leq 4^{-d} \,\middle|\, \mathcal{E}_x \wedge \mathcal{E}_z \wedge \mathcal{E}_\pi\right] \\
&\geq \mathop{\mathbf{Pr}}_{x,z,\pi,t}[i^* - L \leq t \leq i^* \mid \mathcal{E}_x \wedge \mathcal{E}_z \wedge \mathcal{E}_\pi] \\
&= \frac{L+1}{m'} \geq 2^{-10d}. && \text{(by independence)}
\end{aligned}$$

This, combined with (11), proves (8) and completes the proof of Claim 4.14. □

Given Claim 4.14, we can put a much tighter upper bound on δ, as demanded in Claim 4.13. Observe that $8^{-d} \leq \left|\frac{|f(x,\rho)|}{n} - \frac{a}{2^d}\right| \leq 4^{-d}$ implies that $\left|\frac{|f(x,\rho)|}{n} - \frac{a'}{2^d}\right| \geq 2^{-d} - 4^{-d} \geq 8^{-d}$ holds for any integer $a' \neq a$. Hence when the event in Claim 4.14 happens, we have

$$\mathsf{err}\left(\frac{|f(x,\rho)|}{n}, d\right) \geq 8^{-d}.$$

Claim 4.11 implies this should happen with probability at most $\mathsf{poly}(\varepsilon)$. Hence by Claim 4.14, this means $\delta \leq \mathsf{poly}(\varepsilon)$. Therefore (7) follows directly by Claim 4.13. This completes the proof of Lemma 4.9. □

We now have the requisite tools to prove Proposition 4.5, the main result of Subsection 4.1. We restate it below for convenience.

Proposition 4.5. *Let $f\colon \{0,1\}^m \to \{0,1\}^n$ be a d-local function, and let $A \geq 2^{100d}$ be a parameter. Assume $f(\mathcal{U}^m)$ is ε-close to a symmetric distribution over $\{0,1\}^n$ for some $\varepsilon < 2^{-cdA}$, where $c > 0$ is a sufficiently large absolute constant. Further assume that n is sufficiently large in terms of d. Define $S \subseteq [m]$ to be the set of input bits with degree at least n/A.*

Let $k \leq \log(1/\varepsilon)/C_d$ be an arbitrary integer, where $C_d \geq 1$ is a sufficiently large constant depending only on d. Then for each conditioning $\rho \in \{0,1\}^S$ on the bits in S, there exists a subset $T_\rho \subseteq [n]$ of size $|T_\rho| \leq O_{d,k,A}(1)$ such that every k-tuple of output bits in $[n] \setminus T_\rho$ has distribution $\mathcal{U}^k_{\gamma_\rho}$, where $\gamma_\rho = a_\rho/2^d$ and $0 \leq a_\rho \leq 2^d$ is an integer.

Proof. Fix an arbitrary ρ and shorthand $T = T_\rho, \gamma = \gamma_\rho$. Assume there is a maximal set of R disjoint k-tuples of output bits with the wrong distribution. Then it suffices to show $R = O_{d,k,A}(1)$, since we can set T to be the union of these output bits, and every k-tuple that does not have distribution $\mathcal{U}^k_\gamma$ must intersect T by maximality. Moreover, $|T| \leq k \cdot R = O_{d,k,A}(1)$.

Now for the i-th k-tuple with the wrong distribution, we have a degree-k multilinear polynomial $P_i\colon \{0,1\}^n \to \{0,1\}$ such that

$$\mathop{\mathbb{E}}_{x\sim\{0,1\}^{[m]\setminus S}}[P_i(f(x,\rho))] \geq \mathop{\mathbb{E}}_{z\sim\mathcal{U}^n_\gamma}[P_i(z)] + 2^{-kd}. \tag{12}$$

To see this, let $W \subseteq \{0,1\}^k$ witness the total variation distance δ between $\mathcal{U}^k_\gamma$ and the k-tuple output in $f(\mathcal{U}^{[m]\setminus S}, \rho)$, i.e.,

$$\mathop{\mathbf{Pr}}_{x\sim\{0,1\}^{[m]\setminus S}}[\text{the } k\text{-tuple output of } f(x,\rho) \in W] \geq \mathop{\mathbf{Pr}}_{z\sim\mathcal{U}^k_\gamma}[z \in W] + \delta. \tag{13}$$

Then we define the polynomial $P_i\colon \{0,1\}^n \to \{0,1\}$ as the indicator function that the k-tuple output lies in W. This is a k-junta and is naturally of degree k. Since γ is an integer multiple of 2^{-d}, we know that the probability density function of $\mathcal{U}^k_\gamma$ has granularity 2^{-kd}. In addition, since f is d-local, the probability density function of the k-tuple output of $f(\mathcal{U}^{[m]\setminus S}, \rho)$ also has granularity 2^{-kd}. As the two distributions are different, their total variation distance is $\delta \geq 2^{-kd}$. Putting this into (13) gives (12).

Given the constructions of P_i's, we define $P = \sum_{i\in[R]} P_i$. Then by (12), we have

$$\mathop{\mathbb{E}}_{x\sim\{0,1\}^{[m]\setminus S}}[P(f(x,\rho))] \geq \mathop{\mathbb{E}}_{z\sim\mathcal{U}^n_\gamma}[P(z)] + 2^{-kd} \cdot R.$$

Since each P_i is a degree k polynomial of the output bits of f and f is d-local, P is a degree k polynomial of the output bits of f, and $P(f(x,\rho))$ is degree $k \cdot d$ polynomial of x. By Fact 3.10, with probability at least $2^{-O(kd)}$ over $x \sim \{0,1\}^{[m]\setminus S}$, we have

$$P(f(x,\rho)) \geq \mathop{\mathbb{E}}_{x}[P(f(x,\rho))] \geq \mathop{\mathbb{E}}_{z\sim\mathcal{U}^n_\gamma}[P(z)] + 2^{-kd} \cdot R.$$

Thus by Lemma 4.9 and a union bound, with probability at least $2^{-O(kd)} - \mathsf{poly}(\varepsilon)$, we have

$$P(f(x,\rho)) \geq \mathop{\mathbb{E}}_{z\sim\mathcal{U}^n_\gamma}[P(z)] + 2^{-kd} \cdot R \quad \text{and} \quad \left|\frac{|f(x,\rho)|}{n} - \gamma\right| \leq \frac{1}{n^{1/(800d)}}.$$

Recall that $|S| \le dA$. Hence by randomizing also the coordinates in S, with probability at least $2^{-dA}\left(2^{-O(kd)} - \mathsf{poly}(\varepsilon)\right)$ over $y \sim \{0,1\}^m$, we have

$$P(f(y)) \ge \mathop{\mathbb{E}}_{z\sim\mathcal{U}_\gamma^n}[P(z)] + 2^{-kd}\cdot R \quad \text{and} \quad \left|\frac{|f(y)|}{n} - \gamma\right| \le \frac{1}{n^{1/(800d)}}.$$

Since $f(\mathcal{U}^m)$ is ε-close to a symmetric distribution $\mathcal{D}$ over $\{0,1\}^n$, we also have that with probability at least $2^{-dA}\left(2^{-O(kd)} - \mathsf{poly}(\varepsilon)\right) - \varepsilon$ over $w \sim \mathcal{D}$,

$$P(w) \ge \mathop{\mathbb{E}}_{z\sim\mathcal{U}_\gamma^n}[P(z)] + 2^{-kd}\cdot R \quad \text{and} \quad \left|\frac{w}{n} - \gamma\right| \le \frac{1}{n^{1/(800d)}}. \tag{14}$$

Call an input $w \in \{0,1\}^n$ *large* if

$$P(w) \ge \mathop{\mathbb{E}}_{z\sim\mathcal{U}_\gamma^n}[P(z)] + 2^{-kd}\cdot R.$$

We will show that it is very unlikely for $w \sim \mathcal{D}$ to be large while having weight close to γn. Recall that T is the union of the output bits in each of the R k-tuples with the wrong distribution, and define the distribution $\mathcal{D}^*$ to be the marginal distribution of $\mathcal{D}$ on T conditioned on $w \sim \mathcal{D}$ satisfying

$$\left|\frac{|w|}{n} - \gamma\right| \le \frac{1}{n^{1/(800d)}}. \tag{15}$$

Then by sequentially coupling the $t := |T| = Rk$ bits based on their marginals, we have

$$\begin{aligned}
\left\|\mathcal{U}_\gamma^T - \mathcal{D}^*\right\|_{\mathsf{TV}} &\le \sum_{i=1}^{t} \max_{0\le j\le i-1}\left|\gamma - \frac{|w|-j}{n-i+1}\right| && \text{(by union bound)}\\
&\le \sum_{i=1}^{t}\left(\left|\gamma - \frac{|w|}{n}\right| + \left|\frac{|w|}{n} - \frac{|w|-i+1}{n-i+1}\right|\right)\\
&\le \sum_{i=1}^{t}\left(\left|\gamma - \frac{|w|}{n}\right| + \frac{t}{n-t+1}\right)\\
&\le \frac{t}{n^{1/(800d)}} + \frac{t^2}{n-t+1}. && \text{(by (15))}
\end{aligned}$$

Recall that $P = \sum_{i\in[R]} P_i$ where the $P_i\colon \{0,1\}^n \to \{0,1\}$ are supported on disjoint sets of k-tuples. Thus Fact 3.8 implies

$$\mathop{\mathbf{Pr}}_{x\sim\mathcal{U}_\gamma^n}[x \text{ is large}] = \mathop{\mathbf{Pr}}_{x\sim\mathcal{U}_\gamma^n}\left[\sum_{i\in[R]} P_i(x) - \mathop{\mathbb{E}}_{z\sim\mathcal{U}_\gamma^n}[P_i(z)] \ge 2^{-kd}\cdot R\right] \le \exp\left\{-\frac{R}{2^{2kd}}\right\}.$$

Observing that $\sum_i P_i$ only depends on the bits in T, we can upper bound the event in (14) by

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{w\sim\mathcal{D}}\left[w \text{ is large} \wedge \left|\frac{w}{n} - \gamma\right| \le \frac{1}{n^{1/(800d)}}\right] &\le \mathop{\mathbf{Pr}}_{w\sim\mathcal{D}}\left[w \text{ is large} \,\middle|\, \left|\frac{w}{n} - \gamma\right| \le \frac{1}{n^{1/(800d)}}\right]\\
&\le \mathop{\mathbf{Pr}}_{x\sim\mathcal{U}_\gamma^n}[x \text{ is large}] + \left\|\mathcal{U}_\gamma^T - \mathcal{D}^*\right\|_{\mathsf{TV}}\\
&\le \exp\left\{-\frac{R}{2^{2kd}}\right\} + \frac{t}{n^{1/(800d)}} + \frac{t^2}{n-t+1}.
\end{aligned}$$

Combining with the lower bound from (14), we have

$$\exp\left\{-\frac{R}{2^{2kd}}\right\} + \frac{Rk}{n^{1/(800d)}} + \frac{(Rk)^2}{n - Rk + 1} \geq 2^{-dA}\left(2^{-O(kd)} - \mathsf{poly}(\varepsilon)\right) - \varepsilon.$$

By our assumptions on the size of n, k, and ε, we have $R \leq O_{d,k,A}(1)$ as desired. This completes the proof of Proposition 4.5. □

4.2 Kolmogorov Distance

Our second step toward proving Theorem 4.1 is to show that the output weight distribution of $f(\mathcal{U}^m)$ is close in Kolmogorov distance[9] to some binomial distribution $\mathsf{Bin}(n, \gamma)$, where γ is an integer multiple of 2^{-d}. Moreover in the case of $\gamma = 1/2$, we will show that these distributions are close even accounting for parity.

We will require the fact that biased k-wise independence fools a simple type of threshold function. In particular, we will use the following special case of [GOWZ10, Theorem I.5].

Lemma 4.15. *Let $\gamma \in (0,1), k \in \mathbb{N}$, and $t \in \mathbb{R}^{\geq 0}$. If $\mathcal{D}$ is a distribution on $\{0,1\}^n$ such that every k-tuple of bits in $[n]$ has distribution $\mathcal{U}_\gamma^k$, then*

$$\left|\mathop{\mathbf{Pr}}_{x\sim\mathcal{D}}[x_1 + \cdots + x_n \geq t] - \mathop{\mathbf{Pr}}_{x\sim\mathcal{U}_\gamma^n}[x_1 + \cdots + x_n \geq t]\right| \leq O\left(\frac{\mathsf{polylog}(k)}{\sqrt{k}\cdot\mathsf{poly}(\min\{\gamma, 1-\gamma\})}\right).$$

We now leverage Lemma 4.15 to show the output weight of $f(\mathcal{U}^m)$ is close in Kolmogorov distance to a certain binomial distribution. Note the following can be viewed as a generalization of Lemma 5.2 in (the arXiv version of) [KOW25] to arbitrary biases and mixtures. The proof there holds with minor modifications, but we reproduce it for completeness.

Proposition 4.16. *Let $k \geq 2$ and $\ell \geq 1$ be integers, and let $f_1, \ldots, f_\ell \colon \{0,1\}^m \to \{0,1\}^n$ be d-local functions. For each f_i, suppose no input bit affects more than n/A output bits, and suppose there exists a subset $T_i \subseteq [n]$ of size $|T_i| \leq O_{d,k,A}(1)$ such that every k-tuple of output bits in $[n] \setminus T_i$ has distribution $\mathcal{U}_\gamma^k$, where $\gamma = a/2^d$ for some fixed integer $0 < a < 2^d$. If n is sufficiently large in terms of d, k, and A, then any mixture F of the f_i's and any $t \in \mathbb{R}$ satisfy*

$$\left|\mathbf{Pr}\left[|F(\mathcal{U}^m)| \geq t\right] - \mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) \geq t\right]\right| \leq O\left(\frac{\mathsf{polylog}(k)}{\sqrt{k}\cdot\mathsf{poly}(\min\{\gamma, 1-\gamma\})}\right).$$

Moreover if $a = 2^{d-1}$, then there exists an $\eta \in [0,1]$ such that for any $\delta \in (0, 1/2)$ and $t \in \mathbb{R}$, we have

$$\left|\mathbf{Pr}\left[|F(\mathcal{U}^m)| \geq t \text{ and } |F(\mathcal{U}^m)| \text{ is even}\right] - \eta\,\mathbf{Pr}\left[\mathsf{Bin}(n,1/2) \geq t\right]\right|$$

and

$$\left|\mathbf{Pr}\left[|F(\mathcal{U}^m)| \geq t \text{ and } |F(\mathcal{U}^m)| \text{ is odd}\right] - (1-\eta)\,\mathbf{Pr}\left[\mathsf{Bin}(n,1/2) \geq t\right]\right|$$

are at most

$$O\left(\frac{\log(1/\delta)^{O(d)^d}}{\sqrt{A\delta}} + \frac{\mathsf{polylog}(k)}{\sqrt{k}} + \delta\right).$$

[9] Recall the *Kolmogorov distance* between two distributions $\mathcal{P}, \mathcal{Q}$ is given by $\sup_{t\in\mathbb{R}} |\mathbf{Pr}_{x\in\mathcal{P}}[x \geq t] - \mathbf{Pr}_{y\in\mathcal{Q}}[y \geq t]|$.

Proof. It suffices to prove both results for arbitrary f_i, henceforth denoted f. We also use T to denote T_i in this simpler setting.

We will first handle the general case, assuming $\gamma \le 1/2$ for notational simplicity, and afterwards address the $\gamma = 1/2$ case. Our main tool will be Lemma 4.15. By assumption, each k-tuple of output bits in $[n] \setminus T$ has distribution $\mathcal{U}_\gamma^k$ under $f(\mathcal{U}^m)$, so Lemma 4.15 implies

$$\begin{aligned}
\mathbf{Pr}\left[|f(\mathcal{U}^m)| \ge t\right] &\ge \mathbf{Pr}\left[|f(\mathcal{U}^m)[[n] \setminus T]| \ge t\right] \\
&\ge \mathbf{Pr}\left[\mathsf{Bin}(n - |T|, \gamma) \ge t\right] - O\left(\frac{\mathsf{polylog}(k)}{\sqrt{k} \cdot \mathsf{poly}(\gamma)}\right) && \text{(by Lemma 4.15)} \\
&\ge \mathbf{Pr}\left[\mathsf{Bin}(n, \gamma) \ge t + |T|\right] - O\left(\frac{\mathsf{polylog}(k)}{\sqrt{k} \cdot \mathsf{poly}(\gamma)}\right) \\
&= \mathbf{Pr}\left[\mathsf{Bin}(n, \gamma) \ge t\right] - \mathbf{Pr}\left[t \le \mathsf{Bin}(n, \gamma) < t + |T|\right] - O\left(\frac{\mathsf{polylog}(k)}{\sqrt{k} \cdot \mathsf{poly}(\gamma)}\right) \\
&\ge \mathbf{Pr}\left[\mathsf{Bin}(n, \gamma) \ge t\right] - O\left(\frac{\gamma|T|}{\sqrt{\gamma(1-\gamma)n}} + \frac{\mathsf{polylog}(k)}{\sqrt{k} \cdot \mathsf{poly}(\gamma)}\right) && \text{(by Fact 3.6)} \\
&\ge \mathbf{Pr}\left[\mathsf{Bin}(n, \gamma) \ge t\right] - O\left(\frac{\mathsf{polylog}(k)}{\sqrt{k} \cdot \mathsf{poly}(\gamma)}\right),
\end{aligned}$$

where the final inequality follows from our assumption that n is sufficiently large. A similar upper bound follows from comparing the complement distribution $1^n - f(\mathcal{U}^m)$ to the binomial distribution with bias $1 - \gamma$. This concludes the proof of the general case.

We now turn to the case of $\gamma = 1/2$. The proof will proceed similarly, but with some additional technical work. Here, we require a structural result about low degree $\mathbb{F}_2$-polynomials. The following is essentially [CHH+20, Theorem 3.1].

Theorem 4.17. *Let p be a degree-d polynomial over $\mathbb{F}_2^n$ and $\delta \in (0, 1/2)$. There exists a subset $R \subseteq [n]$ with $|R| \le \log(1/\delta)^{O(d)^d}$ such that if we write $p(x) = p(x_{R^c}, x_R)$ where x_R and x_{R^c} are the coordinates in R and not in R respectively, then with probability at least $1 - \delta$ over the choice of a random value of x_{R^c} we have that*

$$\left|\mathop{\mathbf{Pr}}_{x_R}\left[p(x_{R^c}, x_R) = 1\right] - \mathop{\mathbf{Pr}}_{x}\left[p(x) = 1\right]\right| < \delta. \tag{16}$$

In words, Theorem 4.17 says that we can (with high probability) re-randomize the output of the polynomial p by simply re-randomizing the input coordinates of a small set R.

Define $p\colon \{0,1\}^m \to \{0,1\}$ by $p(x) = |f(x)| \bmod 2$ (i.e., the parity of f's output). Applying Theorem 4.17 yields a set R of at most $\log(1/\delta)^{O(d)^d}$ input bits. Let $S \subseteq [n]$ be the set of output bits in T or affected by input bits in R. Note that

$$|S| \le \frac{n|R|}{A} + |T| \le \frac{n \log(1/\delta)^{O(d)^d}}{A} + O_{d,k,A}(1) = \frac{n \log(1/\delta)^{O(d)^d}}{A},$$

where we used the fact that n is sufficiently large in terms of d, k and A for the last equality.

We will set $\eta = \mathbf{Pr}_{x \sim \mathcal{U}^m}[p(x) = 0] = \mathbf{Pr}[|f(\mathcal{U}^m)| \text{ is even}]$ and show

$$\begin{aligned}
&\mathbf{Pr}\left[|f(\mathcal{U}^m)| \ge t \text{ and } |f(\mathcal{U}^m)| \text{ is even}\right] - \eta\, \mathbf{Pr}\left[\mathsf{Bin}(n, 1/2) \ge t\right] \\
&\ge O\left(\frac{\log(1/\delta)^{O(d)^d}}{\sqrt{A\delta}} + \frac{\mathsf{polylog}(k)}{\sqrt{k}} + \delta\right).
\end{aligned}$$

The case of odd $|f(\mathcal{U}^m)|$ is almost identical, and as above, a similar upper bound follows from analyzing the complement distribution $1^n - f(\mathcal{U}^m)$.

Before proceeding further, we define several variables and events for the sake of future clarity. Let

$$C := \frac{|S|}{2} - \sqrt{\frac{|S|}{\delta}} - |T| \geq \frac{|S|}{2} - \frac{\sqrt{n} \cdot \log(1/\delta)^{O(d)^d}}{\sqrt{A\delta}}, \tag{17}$$

where we again used the fact that n is sufficiently large in terms of d, k, and A. Let EVEN be the event that $|f(\mathcal{U}^m)|$ is even (and similarly for ODD). Additionally, let BIG be the event that $|f(\mathcal{U}^m)[[n] \setminus S]| \geq t - C$. Finally, let GOOD be the event that x_{R^c} satisfies (16) (and BAD be the complement event). We have

$$\begin{aligned}
\mathbf{Pr}\left[|f(\mathcal{U}^m)| \geq t \text{ and } \mathsf{EVEN}\right] &\geq \mathbf{Pr}\left[\mathsf{BIG} \text{ and } |f(\mathcal{U}^m)[S]| \geq C \text{ and } \mathsf{EVEN}\right] \\
&= \mathbf{Pr}\left[\mathsf{BIG} \text{ and } \mathsf{EVEN}\right] - \mathbf{Pr}\left[\mathsf{BIG} \text{ and } |f(\mathcal{U}^m)[S]| < C \text{ and } \mathsf{EVEN}\right] \\
&\geq \mathbf{Pr}\left[\mathsf{BIG}\right] \cdot \mathbf{Pr}\left[\mathsf{EVEN} \;\middle|\; \mathsf{BIG}\right] - \mathbf{Pr}\left[|f(\mathcal{U}^m)[S]| < C\right]. \qquad (18)
\end{aligned}$$

By assumption, each k-tuple of output bits in $[n]\setminus T$ has distribution $\mathcal{U}^k_{1/2}$. Since $T \subseteq S$, Lemma 4.15 implies

$$\mathbf{Pr}\left[\mathsf{BIG}\right] \geq \mathbf{Pr}\left[\mathsf{Bin}(n - |S|, 1/2) \geq t - C\right] - O\left(\frac{\mathsf{polylog}(k)}{\sqrt{k}}\right). \tag{19}$$

Now let $X \sim \mathsf{Bin}(n - |S|, 1/2)$ and $Y \sim \mathsf{Bin}(|S|, 1/2)$. Then $X + Y \sim \mathsf{Bin}(n, 1/2)$ and

$$\begin{aligned}
&\mathbf{Pr}\left[\mathsf{Bin}(n - |S|, 1/2) \geq t - C\right] \\
&= \mathbf{Pr}\left[X \geq t - C\right] = \mathbf{Pr}\left[X + Y \geq t - C + Y\right] \\
&= \mathbf{Pr}[X + Y \geq t] - \mathbf{Pr}[t \leq X + Y < t - C + Y] \\
&= \mathbf{Pr}[\mathsf{Bin}(n, 1/2) \geq t] - \mathop{\mathbb{E}}_{Y}\left[\mathop{\mathbf{Pr}}_{X}[t \leq X + Y < t - C + Y]\right] \\
&\geq \mathbf{Pr}[\mathsf{Bin}(n, 1/2) \geq t] - O\left(\mathop{\mathbb{E}}_{Y}\left[\frac{(Y - C) \cdot 1_{Y \geq C}}{\sqrt{n}}\right]\right) \qquad \text{(by Fact 3.6 and independence of } X, Y) \\
&= \mathbf{Pr}[\mathsf{Bin}(n, 1/2) \geq t] - O\left(\mathop{\mathbb{E}}_{Y}\left[\frac{Y - C}{\sqrt{n}}\right] + \mathop{\mathbb{E}}_{Y}\left[\frac{(C - Y) \cdot 1_{Y < C}}{\sqrt{n}}\right]\right) \\
&\geq \mathbf{Pr}[\mathsf{Bin}(n, 1/2) \geq t] - O\left(\mathop{\mathbb{E}}_{Y}\left[\frac{Y - C}{\sqrt{n}}\right] + \mathop{\mathbb{E}}_{Y}\left[\frac{|Y - |S|/2|}{\sqrt{n}}\right]\right) \qquad \text{(since } C \leq |S|/2) \\
&\geq \mathbf{Pr}[\mathsf{Bin}(n, 1/2) \geq t] - O\left(\frac{(|S|/2 - C) + \sqrt{|S|}}{\sqrt{n}}\right) \qquad \text{(since } Y \sim \mathsf{Bin}(|S|, 1/2)) \\
&= \mathbf{Pr}[\mathsf{Bin}(n, 1/2) \geq t] - O\left(\frac{|S|/2 - C}{\sqrt{n}}\right) \qquad \text{(since } C \leq |S|/2 - \sqrt{|S|/\delta}) \\
&\geq \mathbf{Pr}[\mathsf{Bin}(n, 1/2) \geq t] - O\left(\frac{\log(1/\delta)^{O(d)^d}}{\sqrt{A\delta}}\right). \qquad \text{(by (17))}
\end{aligned}$$

Combining with (19), we have

$$\mathbf{Pr}\left[\mathsf{BIG}\right] \geq \mathbf{Pr}\left[\mathsf{Bin}(n, 1/2) \geq t\right] - O\left(\frac{\log(1/\delta)^{O(d)^d}}{\sqrt{A\delta}} + \frac{\mathsf{polylog}(k)}{\sqrt{k}}\right). \tag{20}$$

To lower bound the conditional probability $\mathbf{Pr}\left[\mathsf{EVEN} \;\middle|\; \mathsf{BIG}\right]$, it will be slightly more convenient to upper bound $\mathbf{Pr}\left[\mathsf{ODD} \;\middle|\; \mathsf{BIG}\right]$. For this, observe that the events BIG and GOOD depend only

on input bits in R^c. Hence by Theorem 4.17, conditioned on events BIG and GOOD, event ODD (which rerandomizes input bits in R) happens with probability at most $1-\eta+\delta$. Moreover, event BIG itself happens with probability at most δ. Therefore

$$\begin{aligned}\mathbf{Pr}\left[\mathsf{ODD} \;\middle|\; \mathsf{BIG}\right] &= \mathbf{Pr}\left[\mathsf{GOOD} \;\middle|\; \mathsf{BIG}\right] \cdot \mathbf{Pr}\left[\mathsf{ODD} \;\middle|\; \mathsf{GOOD \text{ and } BIG}\right] \\ &\quad + \mathbf{Pr}\left[\mathsf{BAD} \;\middle|\; \mathsf{BIG}\right] \cdot \mathbf{Pr}\left[\mathsf{ODD} \;\middle|\; \mathsf{BAD \text{ and } BIG}\right] \\ &\leq \mathbf{Pr}\left[\mathsf{ODD} \;\middle|\; \mathsf{GOOD \text{ and } BIG}\right] + \mathbf{Pr}\left[\mathsf{BAD} \;\middle|\; \mathsf{BIG}\right] \\ &\leq \mathbf{Pr}\left[\mathsf{ODD} \;\middle|\; \mathsf{GOOD \text{ and } BIG}\right] + \frac{\mathbf{Pr}\left[\mathsf{BAD}\right]}{\mathbf{Pr}\left[\mathsf{BIG}\right]} \\ &\leq 1-\eta+\delta+\frac{\delta}{\mathbf{Pr}\left[\mathsf{BIG}\right]} \leq 1-\eta+\frac{2\delta}{\mathbf{Pr}\left[\mathsf{BIG}\right]}.\end{aligned}$$

Combining with (20), we have that

$$\mathbf{Pr}\left[\mathsf{BIG}\right] \cdot \mathbf{Pr}\left[\mathsf{EVEN} \;\middle|\; \mathsf{BIG}\right] \geq \eta\, \mathbf{Pr}\left[\mathsf{Bin}(n,1/2) \geq t\right] - O\left(\frac{\log(1/\delta)^{O(d)^d}}{\sqrt{A\delta}} + \frac{\mathsf{polylog}(k)}{\sqrt{k}} + \delta\right).$$

In light of (18) and to complete our proof, it remains to bound $\mathbf{Pr}\left[|f(\mathcal{U}^m)[S]| < C\right]$. We have

$$\begin{aligned}\mathbf{Pr}\left[|f(\mathcal{U}^m)[S]| < C\right] &\leq \mathbf{Pr}\left[|f(\mathcal{U}^m)[S \setminus T]| < \frac{|S|}{2} - \sqrt{\frac{|S|}{\delta}} - |T|\right] \\ &\leq \mathbf{Pr}\left[|f(\mathcal{U}^m)[S \setminus T]| < \frac{|S \setminus T|}{2} - \sqrt{\frac{|S \setminus T|}{\delta}}\right] < \delta,\end{aligned}$$

where the final inequality follows from Chebyshev's inequality and the observation that the outputs in $S \setminus T$ are 2-wise independent. □

4.3 Approximate Continuity

Our third step in proving Theorem 4.1 is a type of continuity result for the output weight of $f(\mathcal{U}^m)$. The argument will require three ingredients from prior works. The first allows us to find many independent collections of output bits. Recall we may view the input-output dependencies of a d-local function $f\colon \{0,1\}^m \to \{0,1\}^n$ as a hypergraph on the vertex set $[n]$ with one edge for each of the m input bits containing all the output bits it affects. By the locality assumption, no vertex has degree more than d.

Lemma 4.18 ([KOW24, Corollary 4.11])**.** *Let G be a hypergraph on n vertices with maximum degree at most d. For any increasing function $F\colon \mathbb{N} \to \mathbb{N}$, there exists a set S of edges in G whose removal yields at least[10] $r = n/O_{d,F}(1)$ vertices in G whose neighborhoods have size at most $t = O_{d,F}(1)$ and are pairwise non-adjacent, and satisfies $|S| \leq r/F(t)$.*

The neighborhoods from Lemma 4.18 appear in our context as collections of independent groups of output bits. When we analyze the behavior of their Hamming weight, they become a sum of independent integer random variables. This sum has a "continuous" output distribution unless almost all of the integer random variables are nearly constant modulo some integer $s \geq 2$. In fact, a similar property holds even if the integer random variables are only noticeably non-constant modulo s for $s \geq 3$, except now the continuity only holds for weights that are an even distance apart. More formally, we will use following density comparison result, which is a special case of [KOW25, Theorem A.1].

[10] Recall $O_{d,F}(1)$ denotes a quantity whose value is constant once d and F are fixed.

Lemma 4.19 ([KOW25, Theorem A.1])**.** *Let $t \geq 1$ be an integer, and let $X_1, \dots, X_n$ be independent random variables in $\{0, 1, \dots, t\}$. For each $i \in [n]$ and integer $s \geq 1$, define $p_{s,i} = \max_{x \in \mathbb{Z}} \mathbf{Pr}[X_i \equiv x \pmod{s}]$. Suppose for some $L > 0$,*

$$\sum_{i \in [n]} (1 - p_{s,i}) \geq L \cdot n \quad \text{holds for all } s \in \{2, 3, \dots, t\}.$$

Then for any $x \in \mathbb{Z}$ and $\Delta \in \mathbb{Z}$, we have

$$\mathbf{Pr}\left[\sum_{i \in [n]} X_i = x\right] - \mathbf{Pr}\left[\sum_{i \in [n]} X_i = x + \Delta\right] \leq O_{L,t}\left(\frac{|\Delta|}{n}\right).$$

If instead $\sum_{i \in [n]} (1 - p_{s,i}) \geq L \cdot n$ only holds for $s \in \{3, 4, \dots, t\}$, the same conclusion still holds for any even $\Delta \in \mathbb{Z}$.

In order to satisfy the variability assumption of Lemma 4.19, we will need to exploit the fact that two coupled, γ-biased random vectors have different Hamming weight distributions, as long as part of their entries are independent. The following is essentially [KOW24, Lemma 4.4].

Lemma 4.20. *Let (X, Y, Z, W) be a random variable where $X, Z \in \{0,1\}$ and $Y, W \in \{0,1\}^{t-1}$. Let $q \geq \min\{3, t+1\}$ be an integer.*[11] *Assume*

- *X is independent from (Z, W) and Z is independent from (X, Y), and*
- *(X, Y) and (Z, W) have distribution $\mathcal{U}_\gamma^t$ for some $\gamma \in (0, 1)$.*

Then we have

$$\mathbf{Pr}[X + |Y| \equiv Z + |W| \pmod{q}] < 1.$$

Moreover, the same conclusion holds for $q \geq 2$ when $\gamma \neq 1/2$.

We remark that the "moreover" part of the conclusion is not explicitly stated in [KOW24], but the proof can be easily modified to show this. For completeness, we include the full details in Appendix C.

We can now state and prove the main result of this subsection.

Proposition 4.21. *Let $f : \{0,1\}^m \to \{0,1\}^n$ be a d-local function with n sufficiently large in terms of d, and let $\gamma \in (0,1)$ be an integer multiple of 2^{-d}. Then the distribution $f(\mathcal{U}^m)$ can be written as a mixture of distributions E and W, where*

1. *$\|E - \mathcal{D}\|_{\mathsf{TV}} \geq 1 - \exp\{-\Omega_d(n)\}$ for any symmetric distribution $\mathcal{D}$ over $\{0,1\}^n$ with weights $\gamma n \pm n^{2/3}$, and*
2. *For all $w \in \{0, 1, \dots, n\}$ and $\Delta \in \mathbb{Z}$ if $\gamma \neq 1/2$ (or even $\Delta \in \mathbb{Z}$ if $\gamma = 1/2$), we have*

$$\left|\mathbf{Pr}[|W| = w] - \mathbf{Pr}[|W| = w + \Delta]\right| \leq O_d\left(\frac{|\Delta|}{n}\right).$$

The above proposition is very similar to [KOW25, Proposition 4.9], and our proof largely follows the one present there. At a high level, we proceed by using Lemma 4.18 to obtain many independent output neighborhoods. We then classify these neighborhoods according to whether their marginal distributions differ from the γ-biased product distribution or not. The former situation corresponds to the first conclusion of Proposition 4.21, while the latter corresponds to the second.

[11] If $q \geq t + 1$, then one may instead apply Lemma 4.20 with modulus $t + 1$, since $X + |Y| \equiv Z + |W| \pmod{q}$ is equivalent to $X + |Y| = Z + |W|$ for $q \geq t + 1$.

Proof. Let $S \subseteq [m]$ be the set of input coordinates promised by Lemma 4.18, taking $F(t)$ to be a sufficiently large multiple of 2^{2dt}. For each conditioning $\rho \in \{0,1\}^S$, consider the restricted function $f_\rho \colon \{0,1\}^{[m]\setminus S} \to \{0,1\}^n$ defined by $f_\rho(x) = f(x,\rho)$. We call a ρ *good* if at least half of the r non-adjacent neighborhoods $\{N_i\}$ (of size at most $t \le C_d$) promised by Lemma 4.18 satisfy $f_\rho(\mathcal{U}^m)[N_i] = \mathcal{U}_\gamma^{N_i}$. Set

$$E \coloneqq \mathop{\mathbb{E}}_{\rho\,:\,\rho \text{ is not good}} f_\rho(\mathcal{U}^{[m]\setminus S}) \qquad \text{and} \qquad W \coloneqq \mathop{\mathbb{E}}_{\rho\,:\,\rho \text{ is good}} f_\rho(\mathcal{U}^{[m]\setminus S}),$$

so that $f(\mathcal{U}^m) = \nu E + (1-\nu)W$ where $\nu \in [0,1]$ is the fraction of good ρ.

We first prove conclusion 1. Suppose $\rho \in \{0,1\}^S$ is not good. Let N be one of the at least $r/2$ non-adjacent neighborhoods with $f_\rho(\mathcal{U}^m)[N] \neq \mathcal{U}_\gamma^N$. Since N depends on at most dt many input bits and γ is an integer multiple of 2^{-d}, $f_\rho(\mathcal{U}^m)[N]$ is at least 2^{-dt}-far from $\mathcal{U}_\gamma^N$. Moreover, define

$$\nu \coloneqq \min_{x \in \mathsf{supp}(\mathcal{D})} \frac{\mathcal{U}_\gamma^n(x)}{\mathcal{D}(x)}.$$

Let $k = \gamma n \pm n^{2/3}$ be such that $\mathcal{D}$ assigns at least $1/\binom{n}{k}$ probability mass on strings of Hamming weight exactly k. Note that such k exists since we assumed that $\mathcal{D}$ is symmetric and supports on strings of weight $\gamma n \pm n^{2/3}$. Then by choosing an arbitrary x with weight k, we have

$$\begin{aligned}
\nu &\ge \frac{\gamma^k(1-\gamma)^{n-k}}{1/\binom{n}{k}} \ge \frac{2^{n\mathcal{H}(\frac{k}{n})}}{\sqrt{8k(1-\frac{k}{n})}} \cdot \gamma^k(1-\gamma)^{n-k} && \text{(by Fact 3.5)} \\
&= \frac{1}{\sqrt{8|x|(1-\frac{k}{n})}} \cdot \left(\frac{\gamma}{\frac{k}{n}}\right)^k \left(\frac{1-\gamma}{1-\frac{k}{n}}\right)^{n-k} \\
&= \frac{1}{\sqrt{8|x|(1-\frac{k}{n})}} \cdot \left(\frac{\gamma}{\gamma \pm n^{-1/3}}\right)^k \left(\frac{1-\gamma}{1-\gamma \mp n^{-1/3}}\right)^{n-k} \\
&\ge \Theta_d(n^{-1/2}) \cdot \left(\frac{1}{1+O_d(n^{-1/3})}\right)^n \ge \exp\left\{-O_d(n^{2/3})\right\}, && (21)
\end{aligned}$$

where we used the fact that $k = \gamma n \pm n^{2/3}$ and n sufficiently large for the last line. Noting that the non-adjacency of the neighborhoods implies the distributions $f_\rho(\mathcal{U}^m)[N_i]$ and $f_\rho(\mathcal{U}^m)[N_j]$ are independent for $i \neq j$, we can apply Lemma 3.2 to conclude

$$\|f_\rho(\mathcal{U}^m) - \mathcal{D}\|_{\mathsf{TV}} \ge 1 - 2\exp\left\{-r2^{-2dt-1}\right\}/\nu.$$

Then Lemma 3.3 and our choice of $F(t) = \Omega(2^{2dt})$ sufficiently large imply

$$\begin{aligned}
\|E - \mathcal{D}\|_{\mathsf{TV}} &\ge 1 - 2^{|S|} \cdot \left(2\exp\left\{-r2^{-2dt-1}\right\}\right)/\nu \\
&\ge 1 - 2\left(\exp\left\{\frac{r}{F(t)} - \frac{r}{2^{2dt+1}}\right\}\right)/\nu \\
&\ge 1 - \exp\left\{-\Omega_d(n)\right\} \cdot \exp\left\{O_d(n^{2/3})\right\} \ge 1 - \exp\left\{-\Omega_d(n)\right\}. && \text{(by (21))}
\end{aligned}$$

It remains to verify conclusion 2. Fix an arbitrary good ρ for the remainder of the argument. We assume without loss of generality that for some $r' \ge r/2$, the neighborhoods $N(1), \ldots, N(r') \subseteq [n]$ satisfy $f_\rho(\mathcal{U}^{[m]\setminus S})[N_i] = \mathcal{U}_\gamma^{N_i}$.

Let $B \subseteq [m] \setminus S$ be the set of input bits that do not affect any central elements (i.e., the r' output bits that generate $N(1), \ldots, N(r')$), which we henceforth refer to as *extraneous inputs*. For each conditioning $\sigma \in \{0,1\}^B$, we define the restricted functions $f_{\rho,\sigma} \colon \{0,1\}^{[m]\setminus(S\cup B)} \to \{0,1\}^n$ by $f_{\rho,\sigma}(x) = f(x,\rho,\sigma)$, so that

$$f_\rho = \mathop{\mathbb{E}}_{\sigma\in\{0,1\}^B}[f_{\rho,\sigma}].$$

Observe that the value of every output bit in $[n] \setminus (N(1) \cup \cdots \cup N(r'))$ is fixed for all $f_{\rho,\sigma}$, and the weight of the output bits of each neighborhood becomes a random variable

$$X_{\sigma,i} \coloneqq \sum_{j\in N(i)} f_{\rho,\sigma}(\mathcal{U}^{[m]\setminus(S\cup B)})[\{j\}].$$

In particular, the total weight of the output of $f_{\rho,\sigma}(\mathcal{U}^{[m]\setminus(S\cup B)})$ is some constant plus the sum of the $X_{\sigma,i}$'s, which are independent. We would like to claim that Lemma 4.19 can be applied to this situation with high probability.

From here, we proceed similarly to [KOW24, Claims 5.16 & 5.23]. For each integer $s \geq 2$, define

$$p_{\sigma,s,i} = \max_{x\in\mathbb{Z}} \mathbf{Pr}\left[X_{\sigma,i} \equiv x \pmod{s} \,\middle|\, \rho,\sigma\right],$$

where recall we previously fixed a good ρ.

Claim 4.22. For any $i \in [r']$ and $s \geq 3$, there exists some $\sigma \in \{0,1\}^B$ satisfying $p_{\sigma,s,i} < 1$ (i.e., $X_{\sigma,i}$ is not constant modulo s). Moreover if $\gamma \neq 1/2$, the same is true for $s = 2$.

Proof. Consider the neighborhood $N \coloneqq N(i)$ of size $t_i \leq t$, and let $\mathcal{I}$ be the set of input bits that the output bit i depends on. Note $|\mathcal{I}| \leq d$. We randomly sample $\sigma \in \{0,1\}^B$ and two independent $\lambda, \lambda' \in \{0,1\}^{\mathcal{I}}$. Since $f_\rho[N]$ only depends on the bits in $B \cup \mathcal{I}$, we can define $(V, Z) = f_\rho((\sigma,\lambda))$ and $(V', Z') = f_\rho((\sigma,\lambda'))$, where $V, V' \in \{0,1\}$ are the values of N's center and $Z, Z' \in \{0,1\}^{t_i-1}$ are the values of the remaining bits in N. Observe that V is independent from (V', Z') and likewise V' is independent from (V, Z), so we may apply Lemma 4.20 to conclude

$$\mathbf{Pr}\left[V + |Z| \equiv V' + |Z'| \pmod{s}\right] < 1.$$

In particular, there must exist some σ where $X_{\sigma,i}$ is not constant modulo s. □

For our fixed good ρ and some fixed $2 \leq s \leq t$ (or $3 \leq s \leq t$ if $\gamma = 1/2$), let $\mathcal{E}_i$ be the event that $X_{\sigma,i}$ is at least 2^{-d}-far from constant modulo s. Since $X_{\sigma,i}$ depends on at most d input bits (namely, the bits i depends on), if it is not constant modulo s, it must be at least 2^{-d}-far from constant. Furthermore, the bits in the neighborhood $N(i)$ depend on at most dt input bits, so

$$\mathop{\mathbf{Pr}}_{\sigma\in\{0,1\}^B}[\mathcal{E}_i] \geq 2^{-dt}.$$

Recall that the neighborhoods are non-adjacent (after removing the edges in S), so the extraneous bits used to determine $X_{\sigma,i}$ are disjoint from those used to determine $X_{\sigma,j}$ for $i \neq j$. Thus, whether or not X_i is constant modulo s is independent of whether X_j is. Applying Chernoff's inequality (Fact 3.9) with $\delta = 1/2$, we have

$$\mathop{\mathbf{Pr}}_{\sigma\in\{0,1\}^B}\left[\sum_{i\in[r']} \mathcal{E}_i \leq 2^{-dt-1} r'\right] \leq \exp\left\{-\frac{r'}{8}\right\} \leq \exp\left\{-\frac{r}{16}\right\} = \exp\left\{-\Omega_d(n)\right\}.$$

Now call a $\sigma \in \{0,1\}^B$ *fluid* if $\sum_i \mathcal{E}_i \geq 2^{-dt-1} r'$ for *every* $2 \leq s \leq t$ (or $3 \leq s \leq t$ if $\gamma = 1/2$). By a union bound and the fact that $t = O_d(1)$, we find

$$\mathop{\mathbf{Pr}}_{\sigma \in \{0,1\}^B} [\sigma \text{ is fluid}] = 1 - \exp\{-\Omega_d(n)\}. \tag{22}$$

Note any fluid σ satisfies

$$\sum_{i \in [r']} (1 - p_{\sigma,s,i}) \geq 2^{-dt-1} \cdot r' \cdot 2^{-d} \geq 2^{-d(t+1)-1} \cdot r \geq \Omega_d(n),$$

so we can apply Lemma 4.19 to find for any $w \in \mathbb{Z}$ and $\Delta \in \mathbb{Z}$ (or even $\Delta \in \mathbb{Z}$ if $\gamma = 1/2$) that

$$\left| \mathop{\mathbf{Pr}}_{x \sim \mathcal{U}^{[m] \setminus (S \cup B)}} [|f_{\rho,\sigma}(x)| = w] - \mathop{\mathbf{Pr}}_{x \sim \mathcal{U}^{[m] \setminus (S \cup B)}} [|f_{\rho,\sigma}(x)| = w + \Delta] \right| \leq O_d\left(\frac{|\Delta|}{n}\right). \tag{23}$$

Taking the mixture over all such subfunctions, we have

$$\begin{aligned}
&\left| \mathbf{Pr}[|W| = w] - \mathbf{Pr}[|W| = w + \Delta] \right| \\
&= \left| \mathop{\mathbb{E}}_{\text{good } \rho} \left[\mathop{\mathbf{Pr}}_{x \sim \mathcal{U}^{[m] \setminus S}} [|f_\rho(x)| = w] \right] - \mathop{\mathbb{E}}_{\text{good } \rho} \left[\mathop{\mathbf{Pr}}_{x \sim \mathcal{U}^{[m] \setminus S}} [|f_\rho(x)| = w + \Delta] \right] \right| \\
&\leq \mathop{\mathbb{E}}_{\substack{\text{good } \rho \\ \sigma}} \left| \mathop{\mathbf{Pr}}_{x \sim \mathcal{U}^{[m] \setminus (S \cup B)}} [|f_{\rho,\sigma}(x)| = w] - \mathop{\mathbf{Pr}}_{x \sim \mathcal{U}^{[m] \setminus (S \cup B)}} [|f_{\rho,\sigma}(x)| = w + \Delta] \right| \quad \text{(by triangle inequality)} \\
&\leq e^{-\Omega_d(n)} + \mathop{\mathbb{E}}_{\substack{\text{good } \rho \\ \text{fluid } \sigma}} \left| \mathop{\mathbf{Pr}}_{x \sim \mathcal{U}^{[m] \setminus (S \cup B)}} [|f_{\rho,\sigma}(x)| = w] - \mathop{\mathbf{Pr}}_{x \sim \mathcal{U}^{[m] \setminus (S \cup B)}} [|f_{\rho,\sigma}(x)| = w + \Delta] \right| \quad \text{(by (22))} \\
&\leq e^{-\Omega_d(n)} + O_d\left(\frac{|\Delta|}{n}\right). \quad \text{(by (23))}
\end{aligned}$$

If $\Delta = 0$, then $\left| \mathbf{Pr}[|W| = w] - \mathbf{Pr}[|W| = w + \Delta] \right|$ is trivially at most $O_d\left(\frac{|\Delta|}{n}\right)$. Hence we may assume $|\Delta| \geq 1$, in which case $e^{-\Omega_d(n)} + O_d\left(\frac{|\Delta|}{n}\right) \leq O_d\left(\frac{|\Delta|}{n}\right)$. This concludes the proof. □

4.4 Putting It Together

In this final subsection, we combine our earlier results to prove Theorem 4.1. Recall Lemma 3.4 gives that the distance between $f(\mathcal{U}^m)$ and any symmetric distribution $\mathcal{P}$ is

$$\|f(\mathcal{U}^m) - \mathcal{P}\|_{\mathsf{TV}} = \Theta\left(\| |f(\mathcal{U}^m)| - |\mathcal{P}| \|_{\mathsf{TV}} + \|f(\mathcal{U}^m) - f(\mathcal{U}^m)_{\text{sym}}\|_{\mathsf{TV}} \right), \tag{24}$$

where the symmetrization $f(\mathcal{U}^m)_{\text{sym}}$ is the distribution resulting from randomly permuting the coordinates of a string $x \sim f(\mathcal{U}^m)$. Under the assumption $f(\mathcal{U}^m)$ is ε-close to a symmetric distribution $\mathcal{D}$, (24) implies $\|f(\mathcal{U}^m) - f(\mathcal{U}^m)_{\text{sym}}\|_{\mathsf{TV}} \leq O(\varepsilon)$. Thus it remains to show that the weight distribution of $f(\mathcal{U}^m)$ is close to the weight distribution of a mixture of the form specified in Theorem 4.1. We first prove the simpler case where no input bit of f affects many output bits, the output distribution of f restricted to (almost) any small set of output bits looks like a γ-biased product distribution, and the symmetric distribution $\mathcal{D}$ is supported on weights around γn.

Lemma 4.23. *Let $k \geq 2$ and $\ell \geq 1$ be integers, and let $f_1, \dots, f_\ell \colon \{0,1\}^m \to \{0,1\}^n$ be d-local functions. For each f_i, suppose no input bit affects more than n/A output bits, and suppose there exists a subset $T_i \subseteq [n]$ of size $|T_i| \leq O_{d,k,A}(1)$ such that every k-tuple of output bits in $[n] \setminus T_i$ has distribution $\mathcal{U}_\gamma^k$, where $\gamma = a/2^d$ for some fixed integer $0 \leq a \leq 2^d$.*

Furthermore, assume there is some mixture F of the f_i's such that $F(\mathcal{U}^m)$ is ε-close to a symmetric distribution $\mathcal{D}$ over $\{0,1\}^n$ which is supported on strings of weight $\gamma n \pm n^{2/3}$. Then if $\gamma \neq 1/2$ and n is sufficiently large in terms of d, k, A, and ε, we have

$$\| |F(\mathcal{U}^m)| - \mathsf{Bin}(n,\gamma) \|_{\mathsf{TV}} \leq O_d(\varepsilon + k^{-1/5}).$$

Moreover if $\gamma = 1/2$, then there exists a mixture $\mathcal{Q} = \eta|\texttt{evens}| + (1-\eta)|\texttt{odds}|$ satisfying

$$\| |F(\mathcal{U}^m)| - \mathcal{Q} \|_{\mathsf{TV}} \leq O_d\left(\varepsilon + k^{-1/5} + \log(k) \sqrt{\frac{\log(A)^{O(d)^d}}{A^{1/3}} + \frac{\mathsf{polylog}(k)}{\sqrt{k}}} \right).$$

Much of the following proof overlaps with [KOW25, Section 5.3]. The overall idea is to use Proposition 4.21 to argue that the output weight of (most of) $f(\mathcal{U}^m)$ satisfies a continuity property: weights at distance Δ apart are assigned mass that differs by at most $O_d(\Delta/n)$. Since Proposition 4.16 shows any contiguous weight interval is given roughly the same probability as a binomial distribution, our continuity property implies that the output weight of $f(\mathcal{U}^m)$ behaves similarly to a binomial distribution pointwise. In the case of $\gamma = 1/2$, we can combine the precise control Proposition 4.16 provides on the probability of being in a fixed interval and even/odd with the parity continuity of Proposition 4.21 to carry out a similar argument.

Proof. We first address the case of $\gamma \in \{0,1\}$, where several of the tools we have developed (e.g., Proposition 4.16 and Proposition 4.21) do not apply. We will show the argument for $\gamma = 0$; the case of $\gamma = 1$ is essentially identical. Let $T = \bigcup_i T_i$, and observe that by assumption, all output bits in $[n] \setminus T$ are identically zero. Thus,

$$\| |F(\mathcal{U}^m)| - \mathsf{Bin}(n,0) \|_{\mathsf{TV}} = \mathbf{Pr}\left[F(\mathcal{U}^m)[T] \neq 0^T \right]. \tag{25}$$

Let us briefly consider the symmetrized distribution $F(\mathcal{U}^m)_{\mathrm{sym}}$. For clarity, let $t = |T|$. Since any string $x \sim F(\mathcal{U}^m)$ has Hamming weight $|x| \leq t \leq \ell \cdot O_{d,k,A}(1)$, we have

$$\begin{aligned} \mathbf{Pr}\left[F(\mathcal{U}^m)_{\mathrm{sym}}[T] \neq 0^T \right] &\leq 1 - \frac{\binom{n-t}{t}}{\binom{n}{t}} = 1 - \prod_{i=0}^{t-1} \frac{n-t-i}{n-i} \\ &\leq 1 - \left(1 - \frac{t}{n-t}\right)^t \leq 1 - \exp\left\{ -\frac{2t^2}{n-t} \right\} \leq \varepsilon, \end{aligned} \tag{26}$$

since n is sufficiently large in terms of d, k, ℓ, A, and ε. By Lemma 3.4 and our initial assumption, we know that

$$\mathbf{Pr}\left[F(\mathcal{U}^m)[T] \neq 0^T \right] - \mathbf{Pr}\left[F(\mathcal{U}^m)_{\mathrm{sym}}[T] \neq 0^T \right] \leq \| F(\mathcal{U}^m) - F(\mathcal{U}^m)_{\mathrm{sym}} \|_{\mathsf{TV}} = O(\varepsilon).$$

Combining with (25) and (26) yields the desired upper bound of

$$\| |F(\mathcal{U}^m)| - \mathsf{Bin}(n,0) \|_{\mathsf{TV}} \leq \mathbf{Pr}\left[F(\mathcal{U}^m)_{\mathrm{sym}}[T] \neq 0^T \right] + O(\varepsilon) = O(\varepsilon).$$

We now turn to the case of $\gamma \notin \{0, 1/2, 1\}$, assuming for notational convenience that $\gamma \leq 1/2$. Afterwards, we will describe the necessary modifications to handle the remaining $\gamma = 1/2$

case. Define $\kappa := \sqrt{\mathsf{polylog}(k)/(\sqrt{k}\cdot\mathsf{poly}(\gamma))}$ with the exponents on the $\mathsf{polylog}(k)$ and $\mathsf{poly}(\gamma)$ corresponding to those in Proposition 4.16, and partition $\{0,1,\dots,n\}$ into (consecutive) intervals of length $\Theta(\kappa\sqrt{n})$. Observe that for any such interval $\mathcal{I}$, Proposition 4.16 implies

$$\left|\mathbf{Pr}\left[|F(\mathcal{U}^m)|\in\mathcal{I}\right]-\mathbf{Pr}\left[\mathsf{Bin}(n,\gamma)\in\mathcal{I}\right]\right|=O(\kappa^2). \tag{27}$$

Next, we apply Proposition 4.21 to each f_i to write $f_i(\mathcal{U}^m)=\lambda_i E_i+(1-\lambda_i)W_i$, where

1. $\|E_i-\mathcal{D}\|_{\mathsf{TV}}\geq 1-\exp\{-\Omega_d(n)\}$, and
2. For all $w\in\{0,1,\dots,n\}$ and $\Delta\in\mathbb{Z}$, we have

$$\left|\mathbf{Pr}\left[|W_i|=w\right]-\mathbf{Pr}\left[|W_i|=w+\Delta\right]\right|\leq O_d\left(\frac{|\Delta|}{n}\right). \tag{28}$$

Since F is a mixture of the f_i's, there exist $c_1,\dots,c_\ell$ such that

$$F(\mathcal{U}^m)=\sum_i c_i\lambda_i E+\left(1-\sum_i c_i\lambda_i\right)W,$$

where

$$E=\frac{1}{\sum_i c_i\lambda_i}\cdot\sum_i c_i\lambda_i E_i\quad\text{and}\quad W=\frac{1}{1-\sum_i c_i\lambda_i}\cdot\sum_i c_i(1-\lambda_i)W_i.$$

For each $i\in[\ell]$, our distance bound between E_i and $\mathcal{D}$ (Item 1) guarantees an event $\mathcal{E}_i$ with mass at least $1-\exp\{-\Omega_d(n)\}$ in E_i, but mass at most $\exp\{-\Omega_d(n)\}$ in D. Thus if we define $\mathcal{E}=\cup\mathcal{E}_i$, then

$$\varepsilon\geq\|F(\mathcal{U}^m)-\mathcal{D}\|_{\mathsf{TV}}\geq\sum_i c_i\lambda_i(1-\exp\{-\Omega_d(n)\})-\ell\cdot\exp\{-\Omega_d(n)\}.$$

In particular, $\sum_i c_i\lambda_i\leq O(\varepsilon)$, and

$$\|W-F(\mathcal{U}^m)\|_{\mathsf{TV}}\leq O(\varepsilon). \tag{29}$$

By expanding the definition of total variation distance, we find

$$\begin{aligned}
\|\,|W|-\mathsf{Bin}(n,\gamma)\|_{\mathsf{TV}}&=\frac{1}{2}\sum_{w=0}^{n}\left|\mathbf{Pr}\left[|W|=w\right]-\mathbf{Pr}\left[\mathsf{Bin}(n,\gamma)=w\right]\right|\\
&=\frac{1}{2}\sum_{\mathcal{I}}\sum_{w\in\mathcal{I}}\left|\mathbf{Pr}\left[|W|=w\right]-\mathbf{Pr}\left[\mathsf{Bin}(n,\gamma)=w\right]\right|\\
&\leq\sum_{\substack{\mathcal{I}\\ w\in\mathcal{I}}}\left|\mathbf{Pr}\left[|W|=w\right]-\frac{\mathbf{Pr}\left[|W|\in\mathcal{I}\right]}{|\mathcal{I}|}\right|+\left|\frac{\mathbf{Pr}\left[|W|\in\mathcal{I}\right]-\mathbf{Pr}\left[\mathsf{Bin}(n,\gamma)\in\mathcal{I}\right]}{|\mathcal{I}|}\right|\\
&\qquad+\left|\mathbf{Pr}\left[\mathsf{Bin}(n,\gamma)=w\right]-\frac{\mathbf{Pr}\left[\mathsf{Bin}(n,\gamma)\in\mathcal{I}\right]}{|\mathcal{I}|}\right|.
\end{aligned}$$

For any $\delta\in(0,1)$, Fact 3.8 implies that all but $O(\delta)$ of the mass of $\mathsf{Bin}(n,\gamma)$ is supported on $O(\log(1/\delta)/\kappa)$ intervals, which we call *big* (and the remaining intervals *small*). Moreover, $|W|$ also only assigns $C\cdot(\varepsilon+\kappa\log(1/\delta)+\delta)$ mass to small intervals for some sufficiently large constant $C>0$, as otherwise we obtain the contradiction

$$C\cdot(\varepsilon+\kappa\log(1/\delta))\leq C\cdot(\varepsilon+\kappa\log(1/\delta)+\delta)-O(\delta)$$

$$
\begin{aligned}
&\le \sum_{\text{small } \mathcal{I}} \sum_{w \in \mathcal{I}} \mathbf{Pr}\left[|W| = w\right] - \sum_{\text{small } \mathcal{I}} \sum_{w \in \mathcal{I}} \mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) = w\right] \\
&= \sum_{\text{big } \mathcal{I}} \sum_{w \in \mathcal{I}} \mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) = w\right] - \sum_{\text{big } \mathcal{I}} \sum_{w \in \mathcal{I}} \mathbf{Pr}\left[|W| = w\right] \\
&\le \sum_{\text{big } \mathcal{I}} \sum_{w \in \mathcal{I}} \left|\mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) = w\right] - \mathbf{Pr}\left[|W| = w\right]\right| \\
&\le \sum_{\text{big } \mathcal{I}} \sum_{w \in \mathcal{I}} \Big(\left|\mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) = w\right] - \mathbf{Pr}\left[|F(\mathcal{U}^m)| = w\right]\right| \\
&\qquad + \left|\mathbf{Pr}\left[|F(\mathcal{U}^m)| = w\right] - \mathbf{Pr}\left[|W| = w\right]\right| \Big) \\
&\le O(\kappa \log(1/\delta)) + 2 \left\| |F(\mathcal{U}^m)| - |W| \right\|_{\mathsf{TV}} \le O(\varepsilon + \kappa \log(1/\delta)),
\end{aligned}
$$

where the final two inequalities follow from (27) and (29), respectively. Hence,

$$
\begin{aligned}
&\left\| |W| - \mathsf{Bin}(n,\gamma) \right\|_{\mathsf{TV}} \\
&\le O(\varepsilon + \kappa \log(1/\delta) + \delta) + \sum_{\text{big } \mathcal{I}} \sum_{w \in \mathcal{I}} \left| \mathbf{Pr}\left[|W| = w\right] - \frac{\mathbf{Pr}\left[|W| \in \mathcal{I}\right]}{|\mathcal{I}|} \right| \\
&\quad + \left| \frac{\mathbf{Pr}\left[|W| \in \mathcal{I}\right] - \mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) \in \mathcal{I}\right]}{|\mathcal{I}|} \right| + \left| \mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) = w\right] - \frac{\mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) \in \mathcal{I}\right]}{|\mathcal{I}|} \right|. \qquad (30)
\end{aligned}
$$

Clearly,

$$
\begin{aligned}
&\sum_{w \in \mathcal{I}} \left| \frac{\mathbf{Pr}\left[|W| \in \mathcal{I}\right] - \mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) \in \mathcal{I}\right]}{|\mathcal{I}|} \right| \\
&= \left|\mathbf{Pr}\left[|W| \in \mathcal{I}\right] - \mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) \in \mathcal{I}\right]\right| \\
&\le \left|\mathbf{Pr}\left[|W| \in \mathcal{I}\right] - \mathbf{Pr}\left[|F(\mathcal{U}^m)| \in \mathcal{I}\right]\right| + \left|\mathbf{Pr}\left[|F(\mathcal{U}^m)| \in \mathcal{I}\right] - \mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) \in \mathcal{I}\right]\right| \\
&\le \left|\mathbf{Pr}\left[|W| \in \mathcal{I}\right] - \mathbf{Pr}\left[|F(\mathcal{U}^m)| \in \mathcal{I}\right]\right| + O(\kappa^2). \qquad \text{(by (27))}
\end{aligned}
$$

Summing over all big intervals, the first term is at most $2 \left\| W - F(\mathcal{U}^m) \right\|_{\mathsf{TV}} \le O(\varepsilon)$ by (29), and the second term is at most $O\left(\kappa \log(1/\delta)\right)$.

Additionally, note that (28) and the triangle inequality implies

$$
\left| \mathbf{Pr}\left[|W| = w\right] - \frac{\mathbf{Pr}\left[|W| \in \mathcal{I}\right]}{|\mathcal{I}|} \right| \le \max_{y \in \mathcal{I}} \left|\mathbf{Pr}\left[|W| = w\right] - \mathbf{Pr}\left[|W| = y\right]\right| \le O_d\left(\frac{\kappa}{\sqrt{n}}\right).
$$

Summing over all $w \in \mathcal{I}$ gives an upper bound of $O_d(\kappa^2)$, and further summing over big intervals gives $O_d(\kappa \log(1/\delta))$. The sum of the

$$
\left| \mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) = w\right] - \frac{\mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) \in \mathcal{I}\right]}{|\mathcal{I}|} \right|
$$

terms can be bounded similarly, since

$$
\max_{y \in \mathcal{I}} \left|\mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) = w\right] - \mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) = y\right]\right| \le O\left(\frac{\kappa\sqrt{n}}{\gamma(1-\gamma)n}\right) = O_d\left(\frac{\kappa}{\sqrt{n}}\right). \qquad \text{(see Fact 3.7)}
$$

Combining these inequalities together with (30), we find

$$
\left\| |W| - \mathsf{Bin}(n,\gamma) \right\|_{\mathsf{TV}} \le O_d(\varepsilon + \delta + \kappa \log(1/\delta)).
$$

Set $\delta = k^{-1/5}$ and recall $\kappa = \sqrt{\mathsf{polylog}(k)/(\sqrt{k} \cdot \mathsf{poly}(\gamma))}$. Again applying (29), we conclude

$$\||F(\mathcal{U}^m)| - \mathsf{Bin}(n,\gamma)\|_{\mathsf{TV}} \le \||F(\mathcal{U}^m)| - |W|\|_{\mathsf{TV}} + \||W| - \mathsf{Bin}(n,\gamma)\|_{\mathsf{TV}} = O_d(\varepsilon + k^{-1/5}).$$

We now consider the case of $\gamma = 1/2$. The proof is almost identical to the previous case, so we only highlight the relevant differences. In this setting, Proposition 4.21 only provides a continuity guarantee on weights differing by an even integer. Hence, we refine the previously considered intervals into their even and odd parts. Note this only changes the number of intervals in our analysis by a factor of two. Moreover, Proposition 4.16 (applied with $\delta = A^{-1/3}$) now provides a bound on the Kolmogorov distance between $|f(\mathcal{U}^m)|$ and a mixture $\mathcal{M} = \eta|\mathtt{evens}| + (1-\eta)|\mathtt{odds}|$ of

$$O\left(\frac{\log(A)^{O(d)^d}}{A^{1/3}} + \frac{\mathsf{polylog}(k)}{\sqrt{k}}\right) =: \kappa^2.$$

Carrying out the remaining steps with δ set to $k^{-1/5}$ as before, we conclude

$$\begin{aligned}\||F(\mathcal{U}^m)| - \mathcal{M}\|_{\mathsf{TV}} &\le \||F(\mathcal{U}^m)| - |W|\|_{\mathsf{TV}} + \||W| - \mathcal{M}\|_{\mathsf{TV}} \\ &= O_d(\varepsilon + \delta + \kappa\log(1/\delta)) \\ &= O_d\left(\varepsilon + k^{-1/5} + \log(k)\sqrt{\frac{\log(A)^{O(d)^d}}{A^{1/3}} + \frac{\mathsf{polylog}(k)}{\sqrt{k}}}\right).\end{aligned}$$

□

We now use Lemma 4.23 to prove the general case. To obtain that lemma's required assumptions, we first condition on all input bits of large degree. This will certainly result in the setting where no input bit affects many output bits, but using Proposition 4.5, it also lets us conclude the restricted functions generate distributions which resemble γ-biased product distributions. For the final condition of Lemma 4.26, we need the symmetric distribution $\mathcal{D}$ to be supported on weights around γn. Through a somewhat laborious but straightforward calculation, we show in Claim 4.25 that the restrictions of $f(\mathcal{U}^m)$ with bias roughly γ are relatively close to $\mathcal{D}$ conditioned on its output weight being $\gamma n \pm n^{2/3}$, as desired.

Lemma 4.24. *Let $f\colon \{0,1\}^m \to \{0,1\}^n$ be a d-local function. Assume $f(\mathcal{U}^m)$ is ε-close to a symmetric distribution $\mathcal{D}$ over $\{0,1\}^n$ where n is sufficiently large in terms of d and ε, and ε is sufficiently small in terms of d. Then the distribution over the Hamming weight of $f(\mathcal{U}^m)$ is $O_d\left(\frac{1}{\log(1/\varepsilon)}\right)^{1/5}$-close to a convex combination of the form*

$$\sum_{\substack{a\in[0,2^d]\cap\mathbb{Z}\\ a\neq 2^{d-1}}} c_a \cdot \mathsf{Bin}(n, a/2^d) + c_e \cdot |\mathtt{evens}| + c_o \cdot |\mathtt{odds}|.$$

Proof. Let k be the largest even integer less than $\log(1/\varepsilon)/C_d$, where $C_d \ge 1$ is a sufficiently large constant depending only on d. By choosing ε small enough, we may assume $k \ge 4$ (in order to later apply Fact 3.11). Define $S \subseteq [m]$ to be the set of input bits with degree at least n/k. Note that by the locality assumption, $|S| \le dk$. For each conditioning $\rho \in \{0,1\}^S$ on the bits in S, Proposition 4.5 guarantees a subset $T_\rho \subseteq [n]$ of size $|T_\rho| \le O_{d,k}(1)$ and a bias $\gamma_\rho = a_\rho/2^d$, where $0 \le a_\rho \le 2^d$ is an integer, such that every k-tuple of output bits in $[n] \setminus T_\rho$ has distribution $\mathcal{U}^k_{\gamma_\rho}$.

We proceed by grouping the restricted functions according to their biases. More formally, we write $f(\mathcal{U}^m)$ as the mixture

$$f(\mathcal{U}^m) = \sum_\gamma \Pr_\rho[\gamma_\rho = \gamma] \cdot f_\gamma(\mathcal{U}^{[m]\setminus S}) \quad \text{where} \quad f_\gamma(\mathcal{U}^{[m]\setminus S}) := \mathop{\mathbb{E}}_\rho\left[f(\mathcal{U}^{[m]\setminus S}, \rho) \,\middle|\, \gamma_\rho = \gamma\right].$$

Similarly, we let $\mathcal{D}_\gamma$ denote the distribution $\mathcal{D}$ conditioned on the Hamming weight being $\gamma n \pm n^{2/3}$. Since $f(\mathcal{U}^m)$ is close to $\mathcal{D}$ by assumption, $f_\gamma(\mathcal{U}^{[m]\setminus S})$ should be close to $\mathcal{D}_\gamma$. We formalize this intuition in the following claim.

Claim 4.25. If $\mathbf{Pr}_\rho[\gamma_\rho = \gamma] > 0$, then $\left\| f_\gamma(\mathcal{U}^{[m]\setminus S}) - \mathcal{D}_\gamma \right\|_{\mathsf{TV}} \le O_d(\varepsilon)$.

The proof of Claim 4.25 is routine but rather tedious, so we defer the details to Appendix C.

For each γ, we combine Claim 4.25 and Lemma 4.23 to deduce

$$\begin{aligned}
\left\| |f_\gamma(\mathcal{U}^{[m]\setminus S})| - \mathcal{P}_\gamma \right\|_{\mathsf{TV}} &\le O_d\left(\varepsilon + k^{-1/5} + \log(k)\sqrt{\frac{(\log(k))^{O(d)^d}}{k^{1/3}} + \frac{\mathsf{polylog}(k)}{\sqrt{k}}}\right) \\
&\le O_d\left(\varepsilon + \left(\frac{1}{\log(1/\varepsilon)}\right)^{1/5}\right) && \text{(since } k = \Theta_d(\log(1/\varepsilon))\text{)} \\
&\le O_d\left(\frac{1}{\log(1/\varepsilon)}\right)^{1/5}, && (31)
\end{aligned}$$

where

$$\mathcal{P}_\gamma = \begin{cases} \mathsf{Bin}(n,\gamma) & \text{if } \gamma \neq 1/2 \\ \eta|\mathtt{evens}| + (1-\eta)|\mathtt{odds}| & \text{if } \gamma = 1/2 \end{cases}$$

for some $\eta \in [0,1]$. Define the mixture $\mathcal{P} \coloneqq \sum_\gamma \mathbf{Pr}_\rho[\gamma_\rho = \gamma] \cdot \mathcal{P}_\gamma$. Writing the output weight of f as a convex combination of the conditionings, we find

$$\begin{aligned}
\| |f(\mathcal{U}^m)| - \mathcal{P} \|_{\mathsf{TV}} &= \left\| \sum_\gamma \mathop{\mathbf{Pr}}_\rho[\gamma_\rho = \gamma] \cdot |f_\gamma(\mathcal{U}^{[m]\setminus(S\cup R)})| - \sum_\gamma \mathop{\mathbf{Pr}}_\rho[\gamma_\rho = \gamma] \cdot \mathcal{P}_\gamma \right\|_{\mathsf{TV}} \\
&\le \sum_\gamma \mathop{\mathbf{Pr}}_\rho[\gamma_\rho = \gamma] \left\| |f_\gamma(\mathcal{U}^{[m]\setminus(S\cup R)})| - \mathcal{P}_\gamma \right\|_{\mathsf{TV}} && \text{(by triangle inequality)} \\
&\le O_d\left(\frac{1}{\log(1/\varepsilon)}\right)^{1/5}. && \text{(by (31))}
\end{aligned}$$

This completes the proof of Lemma 4.24. □

Lemma 4.24 states that the weight distribution of $f(\mathcal{U}^m)$ must be close to a mixture of specific distributions. The following lemma provides additional information about this mixture by describing the structure of the mixing weights.

Lemma 4.26. *Let $f\colon \{0,1\}^m \to \{0,1\}^n$ be a d-local function. Assume $f(\mathcal{U}^m)$ is ε-close to a symmetric distribution $\mathcal{D}$ over $\{0,1\}^n$ where n is sufficiently large in terms of d and ε, and ε is sufficiently small in terms of d. Then the distribution over the Hamming weight of $f(\mathcal{U}^m)$ is $O_d\left(\frac{1}{\log(1/\varepsilon)}\right)^{1/5}$-close to a convex combination of the form*

$$\sum_{\substack{a \in [0,2^d] \cap \mathbb{Z} \\ a \neq 2^{d-1}}} c_a \cdot \mathsf{Bin}(n, a/2^d) + c_e \cdot |\mathtt{evens}| + c_o \cdot |\mathtt{odds}|,$$

where each $c_a = c'_a/2^C$ for some integer $0 \le c'_a \le 2^C$ and a fixed integer $C = O_d(1)$. Moreover, there exist at most 2^C many degree-d $\mathbb{F}_2$-polynomials $\{p_i\colon \mathbb{F}_2^m \to \mathbb{F}_2\}$, each with $O_d(n)$ monomials, such that

$$c_e = \frac{1}{2^C} \cdot \sum_i \mathop{\mathbf{Pr}}_{x \sim \mathcal{U}^m}[p_i(x) = 0] \quad \text{and} \quad c_o = \frac{1}{2^C} \cdot \sum_i \mathop{\mathbf{Pr}}_{x \sim \mathcal{U}^m}[p_i(x) = 1].$$

Proof. We know from Lemma 4.24 that $f(\mathcal{U}^m)$'s weight distribution is $O_d\left(\frac{1}{\log(1/\varepsilon)}\right)^{1/5}$-close to a convex combination of the form

$$\mathcal{P} = \sum_{\substack{a\in[0,2^d]\cap\mathbb{Z}\\ a\neq 2^{d-1}}} c_a \cdot \mathsf{Bin}(n, a/2^d) + c_e \cdot |\mathtt{evens}| + c_o \cdot |\mathtt{odds}|,$$

so it remains to reason about the mixing weights. To this end, let $S \subseteq [m]$ be the set of input bits with degree at least $n/2^{100d}$. Observe that this is a smaller set than the S used in the proof of Lemma 4.24; this will ultimately provide stronger control over the mixing weights.

For each conditioning $\sigma \in \{0,1\}^S$ on the bits in S, Lemma 4.9 guarantees an integer $0 \le a_\sigma \le 2^d$ such that

$$\mathop{\mathbf{Pr}}_{x\sim\mathcal{U}^{[m]\setminus S}}\left[\left|\frac{|f(x,\sigma)|}{n} - \frac{a_\sigma}{2^d}\right| \ge \frac{1}{n^{1/(800d)}}\right] \le \mathsf{poly}(\varepsilon). \tag{32}$$

This also implies for any integer $b \neq a_\sigma$, we have

$$\begin{aligned}\mathop{\mathbf{Pr}}_{x\sim\mathcal{U}^{[m]\setminus S}}\left[\left|\frac{|f(x,\sigma)|}{n} - \frac{b}{2^d}\right| \le \frac{1}{n^{1/(800d)}}\right] &\le \mathop{\mathbf{Pr}}_{x\sim\mathcal{U}^{[m]\setminus S}}\left[\left|\frac{|f(x,\sigma)|}{n} - \frac{a_\sigma}{2^d}\right| \ge \frac{1}{2^d} - \frac{1}{n^{1/(800d)}}\right]\\ &\le \mathop{\mathbf{Pr}}_{x\sim\mathcal{U}^{[m]\setminus S}}\left[\left|\frac{|f(x,\sigma)|}{n} - \frac{a_\sigma}{2^d}\right| \ge \frac{1}{n^{1/(800d)}}\right] \le \mathsf{poly}(\varepsilon).\end{aligned} \tag{33}$$

Define $p_\sigma\colon \{0,1\}^{[m]\setminus S} \to \{0,1\}$ to be $|f(x,\sigma)| \bmod 2$. Note that p_σ is a sum (modulo 2) of n output bits, each of which depends on at most d input bits. Hence, p_σ can be expressed as a degree-d $\mathbb{F}_2$-polynomial with $O_d(n)$ monomials. We will show that the conclusion of Lemma 4.26 is satisfied by

$$\mathcal{Q} = \sum_{\substack{a\in[0,2^d]\cap\mathbb{Z}\\ a\neq 2^{d-1}}} \frac{c'_a}{2^{|S|}} \cdot \mathsf{Bin}(n, a/2^d) + \frac{c'_e}{2^{|S|}} \cdot |\mathtt{evens}| + \frac{c'_o}{2^{|S|}} \cdot |\mathtt{odds}|, \tag{34}$$

where

$$c'_a = \#\left\{\sigma : a_\sigma = a\right\}, \quad c'_e = \sum_{\sigma:a_\sigma=2^{d-1}} \mathop{\mathbf{Pr}}_{x\sim\mathcal{U}^{[m]\setminus S}}\left[p_\sigma(x)=0\right], \text{ and } c'_o = \sum_{\sigma:a_\sigma=2^{d-1}} \mathop{\mathbf{Pr}}_{x\sim\mathcal{U}^{[m]\setminus S}}\left[p_\sigma(x)=1\right].$$

The proof will proceed in two steps. First, we will show that c_a is essentially the probability that $|f(\mathcal{U}^m)|/n$ is close to $a/2^d$. Second, we will show that this probability is closely approximated by the fraction of conditionings which concentrate around a. (As usual, there are some additional considerations in the $a = 2^{d-1}$ case.)

For clarity, define $c_a = c_e + c_o$ and

- $\delta_a = \mathbf{Pr}_{x\sim\mathcal{U}^m}\left[\left|\frac{|f(x)|}{n} - \frac{a}{2^d}\right| \le \frac{1}{n^{1/(800d)}}\right]$,
- $\delta_e = \mathbf{Pr}_{x\sim\mathcal{U}^m}\left[\left|\frac{|f(x)|}{n} - \frac{1}{2}\right| \le \frac{1}{n^{1/(800d)}} \wedge |f(x)| \text{ is even}\right]$,
- $\delta_o = \mathbf{Pr}_{x\sim\mathcal{U}^m}\left[\left|\frac{|f(x)|}{n} - \frac{1}{2}\right| \le \frac{1}{n^{1/(800d)}} \wedge |f(x)| \text{ is odd}\right]$.

Claim 4.27. For any $0 \le a \le 2^d$, we have $|c_a - \delta_a| \le O_d\left(\frac{1}{\log(1/\varepsilon)}\right)^{1/5}$. Moreover, the same upper bound holds on $|c_e - \delta_e|$ and $|c_o - \delta_o|$.

Proof of Claim 4.27. We first record a number of consequences of Hoeffding's inequality (Fact 3.8) for sufficiently large n. We have

$$\mathop{\mathbf{Pr}}_{w\sim\mathsf{Bin}(n,a/2^d)}\left[\left|\frac{w}{n}-\frac{a}{2^d}\right|\le\frac{1}{n^{1/(800d)}}\right]\ge 1-\exp\left\{-2n^{1-(1/(400d))}\right\}\ge 1-e^{-\sqrt{n}}$$

and

$$\begin{aligned}\mathop{\mathbf{Pr}}_{w\sim\mathsf{Bin}(n,1/2)}\left[\left|\frac{w}{n}-\frac{1}{2}\right|\le\frac{1}{n^{1/(800d)}}\,\middle|\,w\text{ is even}\right]&=1-\mathop{\mathbf{Pr}}_{w\sim\mathsf{Bin}(n,1/2)}\left[\left|\frac{w}{n}-\frac{1}{2}\right|>\frac{1}{n^{1/(800d)}}\,\middle|\,w\text{ is even}\right]\\&\ge 1-\frac{\mathbf{Pr}_{w\sim\mathsf{Bin}(n,1/2)}\left[\left|\frac{w}{n}-\frac{1}{2}\right|>\frac{1}{n^{1/(800d)}}\right]}{\mathbf{Pr}_{w\sim\mathsf{Bin}(n,1/2)}\left[w\text{ is even}\right]}\\&\ge 1-2\exp\left\{-2n^{1-(1/(400d))}\right\}\ge 1-e^{-\sqrt{n}}.\end{aligned}$$

Furthermore for $b\neq a$,

$$\begin{aligned}\mathop{\mathbf{Pr}}_{w\sim\mathsf{Bin}(n,b/2^d)}\left[\left|\frac{w}{n}-\frac{a}{2^d}\right|\le\frac{1}{n^{1/(800d)}}\right]&\le\mathop{\mathbf{Pr}}_{w\sim\mathsf{Bin}(n,b/2^d)}\left[\left|\frac{w}{n}-\frac{b}{2^d}\right|\le\frac{1}{2^d}-\frac{1}{n^{1/(800d)}}\right]\\&\le\exp\left\{-2n\left(\frac{1}{2^d}-\frac{1}{n^{1/(800d)}}\right)^2\right\}\le e^{-\sqrt{n}}.\end{aligned}$$

Thus for any $0\le a\le 2^d$, we have

$$\begin{aligned}\delta'_a\coloneqq\mathop{\mathbf{Pr}}_{w\sim\mathcal{P}}\left[\left|\frac{w}{n}-\frac{a}{2^d}\right|\le\frac{1}{n^{1/(800d)}}\right]&=c_a\cdot\mathop{\mathbf{Pr}}_{w\sim\mathsf{Bin}(n,a/2^d)}\left[\left|\frac{w}{n}-\frac{a}{2^d}\right|\le\frac{1}{n^{1/(800d)}}\right]\\&\quad+\sum_{b\neq a}c_b\cdot\mathop{\mathbf{Pr}}_{w\sim\mathsf{Bin}(n,b/2^d)}\left[\left|\frac{w}{n}-\frac{a}{2^d}\right|\le\frac{1}{n^{1/(800d)}}\right]\\&\le c_a+e^{-\sqrt{n}}.\end{aligned}$$

Similarly, $\delta'_a\ge c_a-e^{-\sqrt{n}}$. Combining, we find that

$$|c_a-\delta_a|\le|c_a-\delta'_a|+|\delta'_a-\delta_a|\le e^{-\sqrt{n}}+\|\mathcal{P}-|f(\mathcal{U}^m)|\|_{\mathsf{TV}}\le O_d\left(\frac{1}{\log(1/\varepsilon)}\right)^{1/5}.$$

For the case of δ_e, we have

$$\begin{aligned}\delta'_e&\coloneqq\mathop{\mathbf{Pr}}_{w\sim\mathsf{Bin}(n,1/2)}\left[\left|\frac{w}{n}-\frac{1}{2}\right|\le\frac{1}{n^{1/(800d)}}\wedge w\text{ is even}\right]\\&=c_e\cdot\mathop{\mathbf{Pr}}_{w\sim\mathsf{Bin}(n,1/2)}\left[\left|\frac{w}{n}-\frac{1}{2}\right|\le\frac{1}{n^{1/(800d)}}\,\middle|\,w\text{ is even}\right]\\&\quad+\sum_{b\neq 2^{d-1}}c_b\cdot\mathop{\mathbf{Pr}}_{w\sim\mathsf{Bin}(n,b/2^d)}\left[\left|\frac{w}{n}-\frac{1}{2}\right|\le\frac{1}{n^{1/(800d)}}\wedge w\text{ is even}\right]\\&\le c_e+\sum_{b\neq 2^{d-1}}c_b\cdot\mathop{\mathbf{Pr}}_{w\sim\mathsf{Bin}(n,b/2^d)}\left[\left|\frac{w}{n}-\frac{1}{2}\right|\le\frac{1}{n^{1/(800d)}}\right]\le c_e+e^{-\sqrt{n}}\end{aligned}$$

and

$$\delta'_e \geq c_e \cdot \mathop{\mathbf{Pr}}_{w\sim\mathsf{Bin}(n,1/2)}\left[\left|\frac{w}{n}-\frac{1}{2}\right| \leq \frac{1}{n^{1/(800d)}} \,\middle|\, w \text{ is even}\right] \geq c_e - e^{-\sqrt{n}}.$$

Thus,

$$|c_e - \delta_e| \leq \left|c_e - \delta'_e\right| + \left|\delta'_e - \delta_e\right| \leq e^{-\sqrt{n}} + \left\|\mathcal{P} - |f(\mathcal{U}^m)|\right\|_{\mathsf{TV}} \leq O_d\left(\frac{1}{\log(1/\varepsilon)}\right)^{1/5}.$$

The analysis of δ_o is essentially identical. □

Recall $p_\sigma(x) = |f(x,\sigma)| \bmod 2$.

Claim 4.28. For any $0 \leq a \leq 2^d$, we have $\left|\delta_a - \frac{\#\{\sigma : a_\sigma = a\}}{2^{|S|}}\right| \leq \mathsf{poly}(\varepsilon)$. Moreover, the same upper bound holds on

$$\left|\delta_e - \frac{1}{2^{|S|}}\sum_{\sigma:a_\sigma=2^{d-1}} \mathop{\mathbf{Pr}}_{x\sim\mathcal{U}^{[m]\setminus S}}[p_\sigma(x)=0]\right| \quad \text{and} \quad \left|\delta_o - \frac{1}{2^{|S|}}\sum_{\sigma:a_\sigma=2^{d-1}} \mathop{\mathbf{Pr}}_{x\sim\mathcal{U}^{[m]\setminus S}}[p_\sigma(x)=1]\right|.$$

Proof of Claim 4.28. We can express δ_a as

$$\frac{1}{2^{|S|}}\left(\sum_{\sigma:a_\sigma=a} \mathop{\mathbf{Pr}}_{x\sim\mathcal{U}^{[m]\setminus S}}\left[\left|\frac{|f(x,\sigma)|}{n}-\frac{a}{2^d}\right| \leq n^{-\frac{1}{800d}}\right] + \sum_{\sigma:a_\sigma\neq a} \mathop{\mathbf{Pr}}_{x\sim\mathcal{U}^{[m]\setminus S}}\left[\left|\frac{|f(x,\sigma)|}{n}-\frac{a}{2^d}\right| \leq n^{-\frac{1}{800d}}\right]\right).$$

If $a_\sigma = a$, then

$$1 - \mathsf{poly}(\varepsilon) \leq \mathop{\mathbf{Pr}}_{x\sim\mathcal{U}^{[m]\setminus S}}\left[\left|\frac{|f(x,\sigma)|}{n}-\frac{a}{2^d}\right| \leq \frac{1}{n^{1/(800d)}}\right] \leq 1$$

by (32). Additionally, if $a_\sigma \neq a$, then

$$0 \leq \mathop{\mathbf{Pr}}_{x\sim\mathcal{U}^{[m]\setminus S}}\left[\left|\frac{|f(x,\sigma)|}{n}-\frac{a}{2^d}\right| \leq \frac{1}{n^{1/(800d)}}\right] \leq \mathsf{poly}(\varepsilon)$$

by (33). Thus,

$$\delta_a \leq \frac{1}{2^{|S|}}\left(\#\{\sigma : a_\sigma = a\} + \#\{\sigma : a_\sigma \neq a\}\cdot\mathsf{poly}(\varepsilon)\right) \leq \frac{\#\{\sigma : a_\sigma = a\}}{2^{|S|}} + \mathsf{poly}(\varepsilon).$$

The lower bound on δ_a follows similarly.

For the case of δ_e, we have

$$\begin{aligned}\delta_e = \frac{1}{2^{|S|}}\Bigg(&\sum_{\sigma:a_\sigma=2^{d-1}} \mathop{\mathbf{Pr}}_{x\sim\mathcal{U}^{[m]\setminus S}}\left[\left|\frac{|f(x,\sigma)|}{n}-\frac{1}{2}\right| \leq \frac{1}{n^{1/(800d)}} \wedge |f(x,\sigma)| \text{ is even}\right] \\ &+ \sum_{\sigma:a_\sigma\neq 2^{d-1}} \mathop{\mathbf{Pr}}_{x\sim\mathcal{U}^{[m]\setminus S}}\left[\left|\frac{|f(x,\sigma)|}{n}-\frac{1}{2}\right| \leq \frac{1}{n^{1/(800d)}} \wedge |f(x,\sigma)| \text{ is even}\right]\Bigg) \\ \leq \frac{1}{2^{|S|}}\Bigg(&\sum_{\sigma:a_\sigma=2^{d-1}} \mathop{\mathbf{Pr}}_{x\sim\mathcal{U}^{[m]\setminus S}}[|f(x,\sigma)| \text{ is even}] + \sum_{\sigma:a_\sigma\neq 2^{d-1}} \mathop{\mathbf{Pr}}_{x\sim\mathcal{U}^{[m]\setminus S}}\left[\left|\frac{|f(x,\sigma)|}{n}-\frac{1}{2}\right| \leq \frac{1}{n^{1/(800d)}}\right]\Bigg)\end{aligned}$$

$$\leq \frac{1}{2^{|S|}} \sum_{\sigma : a_\sigma = 2^{d-1}} \mathop{\mathbf{Pr}}_{x \sim \mathcal{U}^{[m]\setminus S}} [p_\sigma(x) = 0] + \mathsf{poly}(\varepsilon)$$

and

$$\begin{aligned}
\delta_e &\geq \frac{1}{2^{|S|}} \sum_{\sigma : a_\sigma = 2^{d-1}} \mathop{\mathbf{Pr}}_{x \sim \mathcal{U}^{[m]\setminus S}} \left[\left| \frac{|f(x,\sigma)|}{n} - \frac{1}{2} \right| \leq \frac{1}{n^{1/(800d)}} \wedge |f(x,\sigma)| \text{ is even} \right] \\
&\geq \frac{1}{2^{|S|}} \sum_{\sigma : a_\sigma = 2^{d-1}} \left(\mathop{\mathbf{Pr}}_{x \sim \mathcal{U}^{[m]\setminus S}} [|f(x,\sigma)| \text{ is even}] - \mathop{\mathbf{Pr}}_{x \sim \mathcal{U}^{[m]\setminus S}} \left[\left| \frac{|f(x,\sigma)|}{n} - \frac{1}{2} \right| > \frac{1}{n^{1/(800d)}} \right] \right) \\
&\geq \frac{1}{2^{|S|}} \sum_{\sigma : a_\sigma = 2^{d-1}} \mathop{\mathbf{Pr}}_{x \sim \mathcal{U}^{[m]\setminus S}} [p_\sigma(x) = 0] - \mathsf{poly}(\varepsilon).
\end{aligned}$$

The case of δ_o is essentially identical. □

Combining Claim 4.27 and Claim 4.28 gives

$$\left| c_a - \frac{\#\{\sigma : a_\sigma = a\}}{2^{|S|}} \right| \leq O_d\left(\frac{1}{\log(1/\varepsilon)} \right)^{1/5}$$

for all $0 \leq a \leq d$, and

$$\left| c_e - \frac{1}{2^{|S|}} \sum_{\sigma : a_\sigma = 2^{d-1}} \mathop{\mathbf{Pr}}_{x \sim \mathcal{U}^{[m]\setminus S}} [p_\sigma(x) = 0] \right| \leq O_d\left(\frac{1}{\log(1/\varepsilon)} \right)^{1/5}$$

(and similarly for c_o). Hence $|f(\mathcal{U}^m)|$ is $O_d\left(\frac{1}{\log(1/\varepsilon)}\right)^{1/5}$-close to $\mathcal{Q}$ (defined in (34)), as desired. □

Now that we have the appropriate result for weight distributions, our main result Theorem 4.1 quickly follows from Lemma 3.4. We restate Theorem 4.1 below for convenience.

Theorem 4.1. *Let $f\colon \{0,1\}^m \to \{0,1\}^n$ be a d-local function. Assume $f(\mathcal{U}^m)$ is ε-close to a symmetric distribution $\mathcal{D}$ over $\{0,1\}^n$. Then if n is sufficiently large in terms of d and ε, $f(\mathcal{U}^m)$ is $O_d\left(\frac{1}{\log(1/\varepsilon)}\right)^{1/5}$-close to a distribution of the form*

$$\sum_{\substack{a \in [0,2^d] \cap \mathbb{Z} \\ a \neq 2^{d-1}}} c_a \cdot \mathcal{U}^n_{a/2^d} + c_e \cdot \texttt{evens} + c_o \cdot \texttt{odds},$$

where each $c_a = c'_a/2^C$ for some integer $0 \leq c'_a \leq 2^C$ and a fixed integer $C = O_d(1)$. Moreover, there exist at most 2^C many degree-d $\mathbb{F}_2$-polynomials $\{p_i\colon \mathbb{F}_2^m \to \mathbb{F}_2\}$, each with $O_d(n)$ monomials, such that

$$c_e = \frac{1}{2^C} \cdot \sum_i \mathop{\mathbf{Pr}}_{x \sim \mathcal{U}^m} [p_i(x) = 0] \quad \text{and} \quad c_o = \frac{1}{2^C} \cdot \sum_i \mathop{\mathbf{Pr}}_{x \sim \mathcal{U}^m} [p_i(x) = 1].$$

Proof. We will prove

$$\|f(\mathcal{U}^m) - \mathcal{Q}\|_{\mathsf{TV}} \leq C_d \cdot \left(\frac{1}{\log(1/\varepsilon)} \right)^{1/5}, \tag{35}$$

where $\mathcal{Q}$ is a distribution of the form in the theorem statement, and $C_d \geq 1$ is a sufficiently large constant depending only on d. (Importantly, we will want $1/C_d$ to be at most the required upper

bound on ε in the premise of Lemma 4.26.) We assume $\varepsilon \le 1/C_d$, as otherwise the bound in (35) exceeds 1 and trivially holds. We also assume $d \ge 1$, as otherwise f is a constant function, so we can set $\mathcal{Q}$ to be either the 0 or 1-biased product distribution.

Combining our original distance assumption with Lemma 3.4, we have

$$\varepsilon \ge \|f(\mathcal{U}^m) - \mathcal{D}\|_{\mathsf{TV}} = \Theta(\||f(\mathcal{U}^m)| - |\mathcal{D}|\|_{\mathsf{TV}} + \|f(\mathcal{U}^m) - f(\mathcal{U}^m)_{\mathrm{sym}}\|_{\mathsf{TV}}).$$

In particular, $\|f(\mathcal{U}^m) - f(\mathcal{U}^m)_{\mathrm{sym}}\|_{\mathsf{TV}} = O(\varepsilon)$. By Lemma 4.26, we have that the distribution over the Hamming weight of $f(\mathcal{U}^m)$ is $O_d\left(\frac{1}{\log(1/\varepsilon)}\right)^{1/5}$-close to a distribution of the form

$$\sum_{\substack{a \in [0,2^d] \cap \mathbb{Z} \\ a \neq 2^{d-1}}} c_a \cdot \mathsf{Bin}(n, a/2^d) + c_e \cdot |\mathtt{evens}| + c_o \cdot |\mathtt{odds}|,$$

where the mixing weights have the desired properties. Let $\mathcal{M}$ be the symmetric distribution over $\{0,1\}^n$ with weight distribution $|\mathcal{M}| = \mathcal{Q}$. Again applying Lemma 3.4, we have

$$\|f(\mathcal{U}^m) - \mathcal{M}\|_{\mathsf{TV}} = O\left(\||f(\mathcal{U}^m)| - \mathcal{Q}\|_{\mathsf{TV}} + \|f(\mathcal{U}^m) - f(\mathcal{U}^m)_{\mathrm{sym}}\|_{\mathsf{TV}}\right) = O_d\left(\frac{1}{\log(1/\varepsilon)}\right)^{1/5}. \quad \square$$

Acknowledgments

We thank Farzan Byramji for helpful comments on an earlier draft of this paper.

References

[ABR16] Maryam Aliakbarpour, Eric Blais, and Ronitt Rubinfeld. Learning and testing junta distributions. In *Conference on Learning Theory*, pages 19–46. PMLR, 2016. 4

[AGM+25] Yaroslav Alekseev, Mika Göös, Konstantin Myasnikov, Artur Riazanov, and Dmitry Sokolov. Sampling permutations with cell probes is hard. *ECCC preprint - TR25-177*, 2025. 3

[ALWZ21] Ryan Alweiss, Shachar Lovett, Kewen Wu, and Jiapeng Zhang. Improved bounds for the sunflower lemma. *Annals of Mathematics*, 194(3):795–815, 2021. 12

[Bab87] László Babai. Random oracles separate PSPACE from the polynomial-time hierarchy. *Information Processing Letters*, 26(1):51–53, 1987. 3, 11

[BEHW89] Anselm Blumer, Andrzej Ehrenfeucht, David Haussler, and Manfred K Warmuth. Learnability and the vapnik-chervonenkis dimension. *Journal of the ACM (JACM)*, 36(4):929–965, 1989. 4

[BIL12] Chris Beck, Russell Impagliazzo, and Shachar Lovett. Large deviation bounds for decision trees and sampling lower bounds for AC0-circuits. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pages 101–110. IEEE, 2012. 3, 6, 13

[Bir87] Lucien Birgé. Estimating a density under order restrictions: Nonasymptotic minimax risk. *The Annals of Statistics*, pages 995–1012, 1987. 4

[BL87] Ravi B Boppana and Jeffrey C Lagarias. One-way functions and circuit complexity. *Information and Computation*, 74(3):226–240, 1987. 3, 11

[BR94] Mihir Bellare and John Rompel. Randomness-efficient oblivious sampling. In *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pages 276–287. IEEE, 1994. 16

[CGZ22] Eshan Chattopadhyay, Jesse Goodman, and David Zuckerman. The space complexity of sampling. In *13th Innovations in Theoretical Computer Science Conference,(ITCS 2022)*, 2022. 3, 13

[CHH+20] Eshan Chattopadhyay, Pooya Hatami, Kaave Hosseini, Shachar Lovett, and David Zuckerman. XOR lemmas for resilient functions against polynomials. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 234–246, 2020. 8, 26

[CS16] Gil Cohen and Leonard J Schulman. Extractors for near logarithmic min-entropy. In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 178–187. IEEE, 2016. 3

[CT06] Thomas M Cover and Joy A Thomas. Elements of information theory, 2006. 15

[CZ16] Eshan Chattopadhyay and David Zuckerman. Explicit two-source extractors and resilient functions. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 670–683, 2016. 3

[DDO+13] Constantinos Daskalakis, Ilias Diakonikolas, Ryan O'Donnell, Rocco A Servedio, and Li-Yang Tan. Learning sums of independent integer random variables. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 217–226. IEEE, 2013. 4

[DFKO06] Irit Dinur, Ehud Friedgut, Guy Kindler, and Ryan O'Donnell. On the Fourier tails of bounded functions over the discrete cube. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 437–446, 2006. 16

[DGJ+10] Ilias Diakonikolas, Parikshit Gopalan, Ragesh Jaiswal, Rocco A Servedio, and Emanuele Viola. Bounded independence fools halfspaces. *SIAM Journal on Computing*, 39(8):3441–3462, 2010. 7, 8, 11

[Dia16] Ilias Diakonikolas. Learning structured distributions. *Handbook of Big Data*, 267:10–1201, 2016. 4

[DK14] Constantinos Daskalakis and Gautam Kamath. Faster and sample near-optimal algorithms for proper learning mixtures of Gaussians. In *Conference on Learning Theory*, pages 1183–1213. PMLR, 2014. 4

[DR09] Lutz Dümbgen and Kaspar Rufibach. Maximum likelihood estimation of a log-concave density and its distribution function: Basic properties and uniform consistency. *Bernoulli*, pages 40–68, 2009. 4

[DW12] Anindya De and Thomas Watson. Extractors and lower bounds for locally samplable sources. *ACM Transactions on Computation Theory (TOCT)*, 4(1):1–21, 2012. 3

[FLRS23] Yuval Filmus, Itai Leigh, Artur Riazanov, and Dmitry Sokolov. Sampling and certifying symmetric functions. In *Approximation, Randomization, and Combinatorial Optimization. (APPROX/RANDOM)*, 2023. 3, 6, 12, 13

[GKM+25] Daniel Grier, Daniel M. Kane, Jackson Morris, Anthony Ostuni, and Kewen Wu. Quantum advantage from sampling shallow circuits: Beyond hardness of marginals. *arXiv preprint arXiv:2510.07808*, 2025. 3, 6, 8

[GOWZ10] Parikshit Gopalan, Ryan O'Donnell, Yi Wu, and David Zuckerman. Fooling functions of halfspaces under product distributions. In *2010 IEEE 25th Annual Conference on Computational Complexity*, pages 223–234. IEEE, 2010. 8, 11, 25

[GSV18] Aryeh Grinberg, Ronen Shaltiel, and Emanuele Viola. Indistinguishability by adaptive procedures with advice, and lower bounds on hardness amplification proofs. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 956–966. IEEE, 2018. 12

[Hås86] Johan Håstad. *Computational limitations for small depth circuits*. PhD thesis, Massachusetts Institute of Technology, 1986. 3

[IN96] Russell Impagliazzo and Moni Naor. Efficient cryptographic schemes provably as secure as subset sum. *Journal of cryptology*, 9(4):199–216, 1996. 11

[JVV86] Mark R Jerrum, Leslie G Valiant, and Vijay V Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoretical computer science*, 43:169–188, 1986. 3, 11

[KMR+94] Michael Kearns, Yishay Mansour, Dana Ron, Ronitt Rubinfeld, Robert E Schapire, and Linda Sellie. On the learnability of discrete distributions. In *Proceedings of the twenty-sixth annual ACM symposium on Theory of computing*, pages 273–282, 1994. 4

[KOW24] Daniel M Kane, Anthony Ostuni, and Kewen Wu. Locality bounds for sampling Hamming slices. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 1279–1286, 2024. Available at https://arxiv.org/abs/2402.14278. 3, 6, 7, 8, 12, 13, 14, 18, 28, 29, 31, 49

[KOW25] Daniel M Kane, Anthony Ostuni, and Kewen Wu. Locally sampleable uniform symmetric distributions. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, pages 1807–1816, 2025. 3, 5, 6, 8, 9, 13, 15, 25, 28, 29, 33, 48

[Lin95] Bruce G Lindsay. Mixture models: Theory, geometry and applications. In *NSF-CBMS Regional Conference Series in Probability and Statistics*, pages i–163. JSTOR, 1995. 4

[LV11] Shachar Lovett and Emanuele Viola. Bounded-depth circuits cannot sample good codes. In *2011 IEEE 26th Annual Conference on Computational Complexity*, pages 243–251. IEEE, 2011. 3, 6, 13

[Ros14] Benjamin Rossman. The monotone complexity of k-clique on random graphs. *SIAM Journal on Computing*, 43(1):256–279, 2014. 12

[SS24] Ronen Shaltiel and Jad Silbak. Explicit codes for poly-size circuits and functions that are hard to sample on low entropy distributions. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 2028–2038, 2024. 3

[Val84] Leslie G Valiant. A theory of the learnable. *Communications of the ACM*, 27(11):1134–1142, 1984. 4

[Vio12a] Emanuele Viola. Bit-probe lower bounds for succinct data structures. *SIAM Journal on Computing*, 41(6):1593, 2012. 12

[Vio12b] Emanuele Viola. The complexity of distributions. *SIAM Journal on Computing*, 41(1):191–218, 2012. 3, 6, 11

[Vio12c] Emanuele Viola. Extractors for Turing-machine sources. In *International Workshop on Approximation Algorithms for Combinatorial Optimization*, pages 663–671. Springer, 2012. 3, 13

[Vio14] Emanuele Viola. Extractors for circuit sources. *SIAM Journal on Computing*, 43(2):655–672, 2014. 3, 13

[Vio20] Emanuele Viola. Sampling lower bounds: boolean average-case and permutations. *SIAM Journal on Computing*, 49(1):119–137, 2020. 3, 6, 13, 15

[Vio23] Emanuele Viola. New sampling lower bounds via the separator. In *38th Computational Complexity Conference (CCC 2023)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik. Available at https://eccc.weizmann.ac.il/report/2021/073/, 2023. 3, 6, 12

[Wik25] Wikipedia. Vandermonde matrix — Wikipedia, the free encyclopedia. https://en.wikipedia.org/wiki/Vandermonde_matrix, 2025. [Online; accessed 3-November-2025]. 46

[WP23] Adam Bene Watts and Natalie Parham. Unconditional quantum advantage for sampling with shallow circuits. *arXiv preprint arXiv:2301.00995*, 2023. 3

[Yat85] Yannis G Yatracos. Rates of convergence of minimum distance estimators and kolmogorov's entropy. *The Annals of Statistics*, 13(2):768–774, 1985. 4

[YZ24] Huacheng Yu and Wei Zhan. Sampling, flowers and communication. In *15th Innovations in Theoretical Computer Science Conference (ITCS 2024)*, pages 100–1. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2024. 3

A An Example Towards the Exact Characterization

In this appendix, we provide the full details behind the example described in Subsection 1.2. Consider the distribution $\mathcal{P} = \mathcal{U}^n_{1/4} + 2^{-n-1}\texttt{evens} - 2^{-n-1}\texttt{odds}$.

Claim A.1. $\mathcal{P}$ is not a distribution of the form given by Theorem 4.1.

Proof. Suppose by contradiction there exists a mixture of the form specified by Theorem 4.1:

$$\mathcal{Q} = \sum_{\substack{a\in[0,2^d]\cap\mathbb{Z}\\ a\neq 2^{d-1}}} c_a \cdot \mathcal{U}^n_{a/2^d} + c_e \cdot \texttt{evens} + c_o \cdot \texttt{odds} \tag{36}$$

where $\mathcal{P} = \mathcal{Q}$ and $\{c_a\}, c_e, c_o$ are nonnegative values summing to 1. For a subset $S \subseteq [n]$, consider the parity function $\chi_S(x) = (-1)^{\sum_{i \in S} x_i}$. Observe that for any set S and bias γ, we have

$$\mathop{\mathbb{E}}_{x \sim \mathcal{U}_\gamma^n}[\chi_S(x)] = \prod_{i \in S} \mathop{\mathbb{E}}_{x \sim \mathcal{U}_\gamma^n}[(-1)^{x_i}] = (1-2\gamma)^{|S|}.$$

Additionally, if $|S| < n$, we have $\mathbb{E}_{x \sim \texttt{evens}}[\chi_S(x)] = \mathbb{E}_{x \sim \texttt{odds}}[\chi_S(x)] = 0$, and otherwise $|S| = n$, where $\mathbb{E}_{x \sim \texttt{evens}}[\chi_S(x)] = 1$ and $\mathbb{E}_{x \sim \texttt{odds}}[\chi_S(x)] = -1$. Thus,

$$\mathop{\mathbb{E}}_{x \sim \mathcal{P}}[\chi_S(x)] = \mathop{\mathbb{E}}_{x \sim \mathcal{U}_{1/4}^n}[\chi_S(x)] = \frac{1}{2^{|S|}}$$

for every S of size smaller than n. In order for $\mathcal{P} = \mathcal{Q}$, it must be that for every such S, we have

$$\sum_{\substack{a \in [0, 2^d] \cap \mathbb{Z} \\ a \neq 2^{d-1}}} c_a \cdot \left(1 - \frac{a}{2^{d-1}}\right)^{|S|} = \mathop{\mathbb{E}}_{x \sim \mathcal{Q}}[\chi_S(x)] = \mathop{\mathbb{E}}_{x \sim \mathcal{P}}[\chi_S(x)] = \frac{1}{2^{|S|}}.$$

Let $b_a = 1 - \frac{a}{2^{d-1}}$. If we consider the left-hand side of the above equation for $|S| = 0, 1, \ldots, 2^d - 1$, this corresponds to the $2^d \times 2^d$ Vandermonde matrix

$$V = \begin{pmatrix} 1 & b_0 & b_0^2 & \cdots & b_0^{2^d-1} \\ 1 & b_1 & b_1^2 & \cdots & b_1^{2^d-1} \\ 1 & b_2 & b_2^2 & \cdots & b_2^{2^d-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & b_{2^d} & b_{2^d}^2 & \cdots & b_{2^d}^{2^d-1}. \end{pmatrix}.$$

Note that V does not have a row corresponding to $b_{2^{d-1}}$ as $a = 2^{d-1}$ will be handled by `evens` and `odds`. Since the b_a's are all distinct, V is invertible (see, e.g., [Wik25]), so the c_a's are uniquely determined to be $c_a = 1$ for $a = 2^{d-2}$ and 0 otherwise. Since we assume (36) is a mixture, we know that $\{c_a\}$'s and c_e, c_o are nonnegative values that sum to 1; this enforces $c_e = c_o = 0$. That is, $\mathcal{Q} = \mathcal{U}_{1/4}^n$, creating the contradiction

$$\frac{1}{2^n} = \mathop{\mathbb{E}}_{x \sim \mathcal{U}_{1/4}^n}[\chi_{[n]}(x)] = \mathop{\mathbb{E}}_{x \sim \mathcal{Q}}[\chi_{[n]}(x)] = \mathop{\mathbb{E}}_{x \sim \mathcal{P}}[\chi_{[n]}(x)] = \frac{1}{2^n} + \frac{1}{2^{n+1}} - \frac{-1}{2^{n+1}} = \frac{1}{2^{n-1}}. \qquad \square$$

Claim A.2. $\mathcal{P}$ can be sampled exactly by a 3-local function.

Proof. Let `all` be the uniform distribution over $\{0,1\}^n$. Consider the distribution $\mathcal{Q}$ defined by sampling $x \sim$ `evens` and $y \sim$ `all`, and returning the bitwise AND $z = x \wedge y$. Clearly, $\mathcal{Q}$ can be sampled with a 3-local function (where $x \sim$ `evens` is sampled using the telescoping construction discussed in the introduction). We will show $\mathcal{P} = \mathcal{Q}$.

For a nonnegative integer k, observe that

$$\begin{aligned} \mathop{\mathbf{Pr}}_{z \sim \mathcal{Q}}\left[z = 1^k 0^{n-k}\right] &= \sum_{S \subseteq [n-k]} \mathop{\mathbf{Pr}}_{x \sim \texttt{evens}}\left[x = 1^k S\right] \mathop{\mathbf{Pr}}_{y \sim \texttt{all}}\left[y = 1^k \cdots, y|_S = 0^{|S|}\right] \\ &= \sum_{S \subseteq [n-k]} \frac{1 + (-1)^{k+|S|}}{2^n} \cdot \frac{1}{2^{k+|S|}} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{S\subseteq[n-k]} \left(\frac{1}{2^{n+k}} \cdot \frac{1}{2^{|S|}}\right) + (-1)^k \sum_{S\subseteq[n-k]} \left(\frac{1}{2^{n+k}} \left(-\frac{1}{2}\right)^{|S|}\right) \\
&= \frac{1}{2^{n+k}} \cdot \left(\frac{3}{2}\right)^{n-k} + \frac{(-1)^k}{2^{n+k}} \cdot \left(\frac{1}{2}\right)^{n-k} \\
&= \frac{3^{n-k} + (-1)^k}{4^n} = \mathop{\mathbf{Pr}}_{z\sim\mathcal{P}}\left[z = 1^k 0^{n-k}\right].
\end{aligned}
$$

Since both $\mathcal{P}$ and $\mathcal{Q}$ are symmetric distributions, the above calculation implies they must be equal. □

B Missing Proofs in Section 3

B.1 Proof of Fact 3.6

Proof of Fact 3.6. It suffices to assume $\gamma \in (0, 1/2]$ and show for any integer k, we have

$$
\mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) = k\right] \le \frac{O(1)}{\sqrt{\gamma n}}. \tag{37}
$$

To this end, we draw samples from $\mathsf{Bin}(n,\gamma)$ in the following way.

- For each $i \in [n]$, sample an unbiased random coin $B_i \in \{0,1\}$ and sample an independent $W_i \in \{0,1\}$ with probability $\mathbf{Pr}[W_i = 1] = 2\gamma$ and $\mathbf{Pr}[W_i = 0] = 1 - 2\gamma$.
- Define $X_i = W_i \cdot B_i$ for each $i \in [n]$. Then output $\sum_{i\in[n]} X_i$.

Now define $\mathcal{E}$ to be the event that $\sum_{i\in[n]} W_i \le \gamma n$. Then by Fact 3.9 with $\delta = 1/2$ and $\mu = 2\gamma n$, we have

$$
\mathbf{Pr}[\mathcal{E}] \le e^{-\gamma n/4}. \tag{38}
$$

For fixed $W = (W_1, \ldots, W_n)$ under which $\mathcal{E}$ does not happen, let $S = \{i \in [n] : W_i = 1\}$. Then,

$$
\mathbf{Pr}\left[\sum_{i\in[n]} X_i = k \,\middle|\, W\right] = \mathbf{Pr}\left[\sum_{i\in S} B_i = k\right] = \mathbf{Pr}\left[\mathsf{Bin}(|S|, 1/2) = k\right] \le \frac{O(1)}{\sqrt{|S|}} \le \frac{O(1)}{\sqrt{\gamma n}}, \tag{39}
$$

where we use $|S| \ge \gamma n$ for the last inequality.

Now we prove (37):

$$
\begin{aligned}
\text{LHS of (37)} = \mathbf{Pr}\left[\sum_{i\in[n]} X_i = k\right] &\le \mathbf{Pr}[\mathcal{E}] + \mathbf{Pr}\left[\sum_{i\in[n]} X_i = k \,\middle|\, \neg\mathcal{E}\right] \\
&\le e^{-\gamma n/4} + \frac{O(1)}{\sqrt{\gamma n}} && \text{(by (38) and (39))} \\
&\le \frac{O(1)}{\sqrt{\gamma n}} = \text{RHS of (37)}.
\end{aligned}
$$

□

B.2 Proof of Fact 3.7

Proof of Fact 3.7. By potentially swapping a and b or by replacing (γ, a, b) with $(1-\gamma, n-a, n-b)$, we may assume $b \geq a$ and $\mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) = a\right] \geq \mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) = b\right]$. From here, we proceed similarly to the proof of Fact 3.6: define $X_i = W_i^{(X)} \cdot B_i^{(X)}$ and $Y_i = W_i^{(Y)} \cdot B_i^{(Y)}$, where each $B_i^{(\cdot)} \in \{0,1\}$ is an independent unbiased random coin, and each $W_i^{(\cdot)} \in \{0,1\}$ is an independent random variable satisfying $\mathbf{Pr}[W_i^{(\cdot)} = 1] = 2\gamma$ and $\mathbf{Pr}[W_i^{(\cdot)} = 0] = 1 - 2\gamma$. For each $w \in \{0,1\}^n$, we additionally define $S_w = \{i \in [n] : w_i = 1\}$. Then,

$$
\begin{aligned}
&\mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) = a\right] - \mathbf{Pr}\left[\mathsf{Bin}(n,\gamma) = b\right] \\
&= \sum_w \mathbf{Pr}\left[W^{(X)} = w\right] \mathbf{Pr}\left[\sum_{i\in[n]} X_i = a \,\middle|\, W^{(X)} = w\right] \\
&\qquad - \sum_w \mathbf{Pr}\left[W^{(Y)} = w\right] \mathbf{Pr}\left[\sum_{i\in[n]} Y_i = b \,\middle|\, W^{(Y)} = w\right] \\
&= \sum_w \mathbf{Pr}\left[W^{(X)} = w\right] \left(\mathbf{Pr}\left[\sum_{i\in S_w} B_i^{(X)} = a\right] - \mathbf{Pr}\left[\sum_{i\in S_w} B_i^{(Y)} = b\right]\right) \\
&= \sum_w \mathbf{Pr}\left[W^{(X)} = w\right] \left(\mathbf{Pr}\left[\mathsf{Bin}(|S_w|, 1/2) = a\right] - \mathbf{Pr}\left[\mathsf{Bin}(|S_w|, 1/2) = b\right]\right) \\
&\leq \mathbf{Pr}\left[W^{(X)} = 0^n\right] + \sum_{w \neq 0^n} \mathbf{Pr}\left[W^{(X)} = w\right] \cdot O\left(\frac{b-a}{|S_w|}\right),
\end{aligned}
$$

where the final inequality is somewhat standard (see, e.g., [KOW25, Fact A.3]).

For the remainder of the argument we will assume $\gamma \in (0, 1/2]$; the remaining case is similar. Define $\mathcal{E}$ to be the event that $\sum_{i\in[n]} W_i^{(X)} \leq \gamma n$. Then by Fact 3.9 with $\delta = 1/2$ and $\mu = 2\gamma n$, we have

$$\mathbf{Pr}[\mathcal{E}] \leq e^{-\gamma n/4},$$

so we may continue our calculation as follows:

$$
\begin{aligned}
&\mathbf{Pr}\left[W^{(X)} = 0^n\right] + \sum_{w \neq 0^n} \mathbf{Pr}\left[W^{(X)} = w\right] \cdot O\left(\frac{b-a}{|S_w|}\right) \\
&\leq \mathbf{Pr}[\mathcal{E}] + \sum_{w : |S_w| \geq \gamma n} \mathbf{Pr}\left[W^{(X)} = w\right] \cdot O\left(\frac{b-a}{|S_w|}\right) \\
&\leq e^{-\gamma n/4} + O\left(\frac{b-a}{\gamma n}\right) \\
&\leq O\left(\frac{b-a}{\gamma(1-\gamma)n}\right).
\end{aligned}
$$

□

C Missing Proofs in Section 4

C.1 Missing Proofs in Subsection 4.1

C.1.1 Proof of Claim 4.10

Proof of Claim 4.10. Assume $\mathsf{err}(p,d) > 2\cdot 2^{-30d}$. Then for any $x \in \{0,1\}^{m-|S|}$ with $\left|\frac{|f(x,\rho)|}{n} - p\right| \le 2^{-30d}$, we have $\mathsf{err}\left(\frac{|f(x,\rho)|}{n}, d\right) \ge 2^{-30d}$. Combining with (6), we then have

$$\mathop{\mathbf{Pr}}_{y\sim\{0,1\}^m}\left[\mathsf{err}\left(\frac{|f(y)|}{n}, d\right) \ge 2^{-30d}\right] \ge \mathbf{Pr}\left[y_S = \rho\right] \cdot \mathop{\mathbf{Pr}}_{x\sim\{0,1\}^{[m]\setminus S}}\left[\left|\frac{|f(x,\rho)|}{n} - p\right| \le 2^{-30d}\right] \ge 2^{-|S|-1}.$$

Recall $|S| \le dA$ and $\varepsilon < 2^{-cdA}$ for some large constant $c > 0$. Thus for sufficiently large n, this contradicts Lemma 4.8. □

C.1.2 Proof of Claim 4.11

Proof of Claim 4.11. Observe that

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{x\sim\{0,1\}^{[m]\setminus S}}\left[\mathsf{err}\left(\frac{|f(x,\rho)|}{n}, d\right) > \frac{1}{n^{1/(800d)}}\right] &= \mathop{\mathbf{Pr}}_{y\sim\{0,1\}^m}\left[\mathsf{err}\left(\frac{|f(y)|}{n}, d\right) > \frac{1}{n^{1/(800d)}} \,\middle|\, y_S = \rho\right] \\
&\le 2^{|S|} \cdot \mathop{\mathbf{Pr}}_{y\sim\{0,1\}^m}\left[\mathsf{err}\left(\frac{|f(y)|}{n}, d\right) > \frac{1}{n^{1/(800d)}}\right] \\
&\le 2^{dA} \cdot O\left(\varepsilon + e^{-n^{0.9}}\right) && \text{(by Lemma 4.8)} \\
&\le \mathsf{poly}(\varepsilon) && \text{(since } \varepsilon < 2^{-cdA}\text{)}
\end{aligned}$$

for sufficiently large n. □

C.1.3 Proof of Claim 4.13

Proof of Claim 4.13. Observe that the LHS event has the following two cases:

- $\mathsf{err}\left(\frac{|f(x,\rho)|}{n}, d\right) > n^{-1/(800d)}$. By Claim 4.11, this happens with probability $\mathsf{poly}(\varepsilon)$.
- $\mathsf{err}\left(\frac{|f(x,\rho)|}{n}, d\right) \le n^{-1/(800d)}$ but $\left|\frac{|f(x,\rho)|}{n} - \frac{a}{2^d}\right| > \frac{1}{n^{1/(800d)}}$. Then it means $\left|\frac{|f(x,\rho)|}{n} - \frac{a'}{2^d}\right| \le n^{-1/(800d)}$ for some $a' \ne a$. Then

$$\left|\frac{|f(x,\rho)|}{n} - \frac{a}{2^d}\right| \ge 2^{-d} - n^{-1/(800d)} > 4^{-d}$$

as n is sufficiently large in terms of d. This happens with probability at most δ. □

C.2 Missing Proofs in Subsection 4.3

For the convenience of the reader, we include a full proof of Lemma 4.20. We emphasize that aside from minor modifications to handle the case of $\gamma = 1/2$ and relax some quantitative dependencies, the proof is essentially copied verbatim from [KOW24].

Proof of Lemma 4.20. We assume without loss of generality $\gamma \in (0, 1/2]$ by flipping the zeros and ones of (X, Y, Z, W) if necessary. Observe that this preserves the congruence. If $t = 1$ then we have that $\mathbf{Pr}[X + |Y| = Z + |W|] = \mathbf{Pr}[X = Z]$ as $q \geq 2$. Since X and Z are independent copies of $\mathcal{U}_\gamma^1$, we find

$$\mathbf{Pr}[X = Z] = \mathbf{Pr}[X = 1]^2 + (1 - \mathbf{Pr}[X = 1])^2 = \gamma^2 + (1 - \gamma)^2 < 1, \tag{40}$$

where we use the fact that $\gamma \in (0, 1/2]$.

Now we assume $t, q \geq 2$. Expand $\mathbf{Pr}[X + |Y| \equiv Z + |W| \pmod q]$ as

$$\sum_{x,z \in \{0,1\}} \mathbf{Pr}[X = x, Z = z] \, \mathbf{Pr}[x + |Y| \equiv z + |W| \pmod q \mid X = x, Z = z]. \tag{41}$$

For fixed x and z, consider the distribution of $x + |Y| \bmod q$ conditioned on $X = x, Z = z$. Since Z is independent from (X, Y), it is the same as the distribution, denoted by $\mathcal{P}_x$, of $x + |Y| \bmod q$ conditioned on $X = x$. Similarly define $\mathcal{Q}_z$ as the distribution of $z + |W| \bmod q$ conditioned on $Z = z$ (or equivalently, conditioned on $Z = z, X = x$).

Since (X, Y) has distribution $\mathcal{U}_\gamma^t$, $\mathcal{P}_0$ has distribution $\mathcal{D}_0$, the distribution of $|V| \bmod q$ for $V \sim \mathcal{U}_\gamma^{t-1}$. Similarly, $\mathcal{Q}_1$ has distribution $\mathcal{D}_1$, the distribution of $1 + |V| \bmod q$ for $V \sim \mathcal{U}_\gamma^{t-1}$. Hence by Fact 3.1,

$$\mathbf{Pr}[|Y| \equiv 1 + |W| \pmod q \mid X = 0, Z = 1] \leq 1 - \|\mathcal{P}_0 - \mathcal{Q}_1\|_{\mathsf{TV}} = 1 - \|\mathcal{D}_0 - \mathcal{D}_1\|_{\mathsf{TV}}.$$

The same bound holds for $\mathbf{Pr}[1 + |Y| \equiv |W| \pmod q \mid X = 1, Z = 0]$. Plugging back into (41), we have

$$\begin{aligned}
\mathbf{Pr}[X + |Y| \equiv Z + |W| \pmod q] &\leq \mathbf{Pr}[X = Z] + \mathbf{Pr}[X \neq Z] \cdot (1 - \|\mathcal{D}_0 - \mathcal{D}_1\|_{\mathsf{TV}}) \\
&= \mathbf{Pr}[X = Z] + (1 - \mathbf{Pr}[X = Z]) \, (1 - \|\mathcal{D}_0 - \mathcal{D}_1\|_{\mathsf{TV}}) \\
&= 1 - \|\mathcal{D}_0 - \mathcal{D}_1\|_{\mathsf{TV}} \, (1 - \mathbf{Pr}[X = Z]).
\end{aligned}$$

We know $\mathbf{Pr}[X = Z] < 1$ by (40), so the desired result follows from showing $\|\mathcal{D}_0 - \mathcal{D}_1\|_{\mathsf{TV}} > 0$ for any choice of $q \geq 3$, as well as for $q = 2$ if $\gamma \neq 1/2$.

For this we use Fourier analysis. Let $\omega_q = e^{2\pi i/q}$ be the primitive q-th root of unit. We consider the following quantity

$$Q = \left| \mathop{\mathbb{E}}_{X \sim \mathcal{D}_0} \left[\omega_q^X\right] - \mathop{\mathbb{E}}_{X \sim \mathcal{D}_1} \left[\omega_q^X\right] \right|.$$

On the one hand, we have

$$Q \leq \sum_{c \in \mathbb{Z}/q\mathbb{Z}} \left|\omega_q^c \cdot (\mathcal{D}_0(c) - \mathcal{D}_1(c))\right| = \sum_{c \in \mathbb{Z}/q\mathbb{Z}} |\mathcal{D}_0(c) - \mathcal{D}_1(c)| = 2 \cdot \|\mathcal{D}_0 - \mathcal{D}_1\|_{\mathsf{TV}}. \tag{42}$$

On the other hand, we have

$$\begin{aligned}
Q &= \left|(1 - \gamma + \gamma \cdot \omega_q)^{t-1} - \omega_q \cdot (1 - \gamma + \gamma \cdot \omega_q)^{t-1}\right| && \text{(by the definition of } \mathcal{D}_0 \text{ and } \mathcal{D}_1) \\
&= |1 - \omega_q| \cdot |1 - \gamma + \gamma \cdot \omega_q|^{t-1}. && (43)
\end{aligned}$$

Let $r = \sin^2\left(\frac{\pi}{q}\right)$. Then

$$|1 - \omega_q| = \sqrt{\left(1 - \cos\left(\frac{2\pi}{q}\right)\right)^2 + \sin^2\left(\frac{2\pi}{q}\right)} = 2 \cdot \left|\sin\left(\frac{\pi}{q}\right)\right| = 2\sqrt{r}$$

and

$$|1-\gamma+\gamma\cdot\omega_q| = \sqrt{\left(1-\gamma+\gamma\cdot\cos\left(\frac{2\pi}{q}\right)\right)^2+\gamma^2\cdot\sin^2\left(\frac{2\pi}{q}\right)}$$
$$= \sqrt{1-4\gamma(1-\gamma)\cdot\sin^2\left(\frac{\pi}{q}\right)} = \sqrt{1-4\gamma(1-\gamma)r}.$$

Combining these with (42) and (43), we have

$$\|\mathcal{D}_0-\mathcal{D}_1\|_{\mathsf{TV}} \geq \sqrt{r\cdot(1-4\gamma(1-\gamma)r)^{t-1}},$$

which is strictly larger than 0 unless $\gamma = 1/2$ and $q = 2$. This completes the proof. □

C.3 Missing Proofs in Subsection 4.4

Proof of Claim 4.25. For clarity, we define/recall the following notation:

- $\mathcal{E}^*(x)$ is the event that $|x| = \gamma n \pm n^{2/3}$,
- $p_\gamma = \mathbf{Pr}_\rho[\gamma_\rho = \gamma]$,
- $\mathcal{F}_\gamma = \mathbb{E}_\rho\left[f(\mathcal{U}^{[m]\setminus S},\rho)\,\middle|\,\gamma_\rho=\gamma\right]$,
- $\mathcal{G}_\gamma$ is $f(\mathcal{U}^m)$ conditioned on $\mathcal{E}^*$,
- $\mathcal{D}_\gamma$ is $\mathcal{D}$ conditioned on $\mathcal{E}^*$.

We will individually bound $\|\mathcal{F}_\gamma-\mathcal{G}_\gamma\|_{\mathsf{TV}}$ and $\|\mathcal{G}_\gamma-\mathcal{D}_\gamma\|_{\mathsf{TV}}$, and obtain our claim via the triangle inequality. Throughout the following, let $\mathcal{E}$ be an arbitrary event.

We first compare $\mathcal{F}_\gamma$ and $\mathcal{G}_\gamma$. By definition,

$$\mathop{\mathbf{Pr}}_{x\sim\mathcal{G}_\gamma}[\mathcal{E}(x)] = \frac{\mathbf{Pr}_{x\sim f(\mathcal{U}^m)}[\mathcal{E}(x)\wedge\mathcal{E}^*(x)]}{\mathbf{Pr}_{x\sim f(\mathcal{U}^m)}[\mathcal{E}^*(x)]} = \frac{\sum_\alpha p_\alpha\,\mathbf{Pr}_{x\sim\mathcal{F}_\alpha}[\mathcal{E}(x)\wedge\mathcal{E}^*(x)]}{\sum_\alpha p_\alpha\,\mathbf{Pr}_{x\sim\mathcal{F}_\alpha}[\mathcal{E}^*(x)]}$$
$$= \frac{p_\gamma\,\mathbf{Pr}_{x\sim\mathcal{F}_\gamma}[\mathcal{E}(x)\wedge\mathcal{E}^*(x)] + \sum_{\alpha\neq\gamma}p_\alpha\,\mathbf{Pr}_{x\sim\mathcal{F}_\alpha}[\mathcal{E}(x)\wedge\mathcal{E}^*(x)]}{p_\gamma\,\mathbf{Pr}_{x\sim\mathcal{F}_\gamma}[\mathcal{E}^*(x)] + \sum_{\alpha\neq\gamma}p_\alpha\,\mathbf{Pr}_{x\sim\mathcal{F}_\alpha}[\mathcal{E}^*(x)]}. \tag{44}$$

We will separately bound each of the four terms appearing in (44). Note that if $\alpha\neq\gamma$, then $|\alpha-\gamma|\geq 2^{-d}$. Thus, if $|x|$ is close to γn, it must be reasonably far from αn. That is,

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\alpha}[\mathcal{E}^*(x)] &= \mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\alpha}\left[\big||x|-\gamma n\big|\leq n^{2/3}\right] \\
&\leq \mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\alpha}\left[\big||x|-\alpha n\big|\geq\big|\alpha n-\gamma n\big|-n^{2/3}\right] \\
&\leq \mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\alpha}\left[\big||x|-\alpha n\big|\geq 2^{-d}n-n^{2/3}\right] && \text{(since } |\alpha-\gamma|\geq 2^{-d}\text{)} \\
&\leq \mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\alpha}\left[\big||x|-\alpha n\big|\geq 2^{-d-1}n\right]. && \text{(since } n\gg d\text{)}
\end{aligned}$$

Recall for each restricted function $f_\rho(\mathcal{U}^{[m]\setminus S})$ there exists some subset $T_\rho\subseteq[n]$ of size $|T_\rho|\leq O_{d,k}(1)$ such that every k-tuple of bits in $[n]\setminus T_\rho$ sampled from $f_\rho(\mathcal{U}^{[m]\setminus S})$ has distribution $\mathcal{U}^k_{\gamma_\rho}$ (i.e., is k-wise independent for some even k). Define T to be the union over all T_ρ in the mixture $\mathcal{F}_\alpha$, and note

that $|T| \le \ell \cdot O_{d,k}(1)$. Let $\overline{x}$ denote the restriction of x to the bits in $[n] \setminus T$. Then we may continue the previous chain of inequalities by

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\alpha}[\mathcal{E}^*(x)] &\le \mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\alpha}\left[\left||\overline{x}| - \alpha n\right| \ge 2^{-d-1}n - |T|\right] \\
&\le \mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\alpha}\left[\left||\overline{x}| - \alpha(n-|T|)\right| \ge 2^{-d-1}n - (1+\alpha)|T|\right] \\
&\le \mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\alpha}\left[\left||\overline{x}| - \mathbb{E}[\overline{x}]\right| \ge 2^{-d-2}n\right] && \text{(since } n \gg d,k,\ell\text{)} \\
&\le 2\left(\frac{nk}{(2^{-d-2}n)^2}\right)^{k/2} && \text{(by Fact 3.11)} \\
&\le \left(\frac{2^{2d+5}\cdot k}{n}\right)^{k/2}. && (45)
\end{aligned}$$

Now consider $p_\gamma \mathbf{Pr}_{x\sim\mathcal{F}_\gamma}[\mathcal{E}^*(x)]$. We know by assumption that $p_\gamma > 0$, so there must exist some setting ρ of the bits in S such that $\gamma_\rho = \gamma$. Hence, we in fact have the stronger lower bound

$$p_\gamma \ge 2^{-|S|}. \tag{46}$$

For the remaining factor, we find

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\gamma}[\mathcal{E}^*(x)] &= \mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\gamma}\left[\left||x| - \gamma n\right| \le n^{2/3}\right] \\
&\ge 1 - \mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\gamma}\left[\left||\overline{x}| - \gamma(n-|T|)\right| > n^{2/3} - (1+\gamma)|T|\right] \\
&\ge 1 - \mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\gamma}\left[\left||\overline{x}| - \mathbb{E}[\overline{x}]\right| > \frac{n^{2/3}}{2}\right] && \text{(since } n \gg d,k\text{)} \\
&\ge 1 - 2\left(\frac{nk}{(n^{2/3}/2)^2}\right)^{k/2} && \text{(by Fact 3.11)} \\
&\ge 1 - \left(\frac{8k}{n^{1/3}}\right)^{k/2}. && (47)
\end{aligned}$$

Finally, we consider $p_\gamma \mathbf{Pr}_{x\sim\mathcal{F}_\gamma}[\mathcal{E}(x) \wedge \mathcal{E}^*(x)]$. We can again use (46) to lower bound p_γ. Additionally,

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\gamma}[\mathcal{E}(x) \wedge \mathcal{E}^*(x)] &= \mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\gamma}[\mathcal{E}(x)] - \mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\gamma}[\mathcal{E}(x) \wedge \neg\mathcal{E}^*(x)] \\
&\ge \mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\gamma}[\mathcal{E}(x)] - \left(1 - \mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\gamma}[\mathcal{E}^*(x)]\right) \\
&\ge \mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\gamma}[\mathcal{E}(x)] - \left(\frac{8k}{n^{1/3}}\right)^{k/2}, && (48)
\end{aligned}$$

where the final inequality uses (47). Substituting (45), (46), (47), and (48) into (44), we find that

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{x\sim\mathcal{G}_\gamma}[\mathcal{E}(x)] &= \frac{p_\gamma \mathbf{Pr}_{x\sim\mathcal{F}_\gamma}[\mathcal{E}(x)\wedge\mathcal{E}^*(x)] + \sum_{\alpha\neq\gamma} p_\alpha \mathbf{Pr}_{x\sim\mathcal{F}_\alpha}[\mathcal{E}(x)\wedge\mathcal{E}^*(x)]}{p_\gamma \mathbf{Pr}_{x\sim\mathcal{F}_\gamma}[\mathcal{E}^*(x)] + \sum_{\alpha\neq\gamma} p_\alpha \mathbf{Pr}_{x\sim\mathcal{F}_\alpha}[\mathcal{E}^*(x)]} \\
&\le \frac{\mathbf{Pr}_{x\sim\mathcal{F}_\gamma}[\mathcal{E}(x)] + \frac{1}{p_\gamma}\max_{\alpha\neq\gamma}\mathbf{Pr}_{x\sim\mathcal{F}_\alpha}[\mathcal{E}^*(x)]}{\mathbf{Pr}_{x\sim\mathcal{F}_\gamma}[\mathcal{E}^*(x)]}
\end{aligned}$$

$$\leq \frac{\mathbf{Pr}_{x\sim\mathcal{F}_\gamma}\left[\mathcal{E}(x)\right] + 2^{|S|}\left(\frac{2^{2d+5}\cdot k}{n}\right)^{k/2}}{1-\left(\frac{8k}{n^{1/3}}\right)^{k/2}}.$$

Rearranging gives

$$\mathop{\mathbf{Pr}}_{x\sim\mathcal{G}_\gamma}\left[\mathcal{E}(x)\right] - \mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\gamma}\left[\mathcal{E}(x)\right] \leq \left(\frac{8k}{n^{1/3}}\right)^{k/2} + 2^{|S|}\left(\frac{2^{2d+5}\cdot k}{n}\right)^{k/2} \leq n^{-k/10}, \tag{49}$$

since n is sufficiently large in terms of d, k, and ε. Similarly, we find that

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{x\sim\mathcal{G}_\gamma}\left[\mathcal{E}(x)\right] &= \frac{p_\gamma \mathbf{Pr}_{x\sim\mathcal{F}_\gamma}\left[\mathcal{E}(x)\wedge\mathcal{E}^*(x)\right] + \sum_{\alpha\neq\gamma} p_\alpha \mathbf{Pr}_{x\sim\mathcal{F}_\alpha}\left[\mathcal{E}(x)\wedge\mathcal{E}^*(x)\right]}{p_\gamma \mathbf{Pr}_{x\sim\mathcal{F}_\gamma}\left[\mathcal{E}^*(x)\right] + \sum_{\alpha\neq\gamma} p_\alpha \mathbf{Pr}_{x\sim\mathcal{F}_\alpha}\left[\mathcal{E}^*(x)\right]} \\
&\geq \frac{\mathbf{Pr}_{x\sim\mathcal{F}_\gamma}\left[\mathcal{E}(x)\wedge\mathcal{E}^*(x)\right]}{1+\frac{1}{p_\gamma}\max_{\alpha\neq\gamma}\mathbf{Pr}_{x\sim\mathcal{F}_\alpha}\left[\mathcal{E}^*(x)\right]} \\
&\geq \frac{\mathbf{Pr}_{x\sim\mathcal{F}_\gamma}\left[\mathcal{E}(x)\right] - \left(\frac{8k}{n^{1/3}}\right)^{k/2}}{1+2^{|S|}\left(\frac{2^{2d+5}\cdot k}{n}\right)^{k/2}},
\end{aligned}$$

or equivalently

$$\mathop{\mathbf{Pr}}_{x\sim\mathcal{F}_\gamma}\left[\mathcal{E}(x)\right] - \mathop{\mathbf{Pr}}_{x\sim\mathcal{G}_\gamma}\left[\mathcal{E}(x)\right] \leq \left(\frac{8k}{n^{1/3}}\right)^{k/2} + 2^{|S|}\left(\frac{2^{2d+5}\cdot k}{n}\right)^{k/2} \leq n^{-k/10}. \tag{50}$$

Combining (49) and (50) yields

$$\|\mathcal{F}_\gamma - \mathcal{G}_\gamma\|_{\mathsf{TV}} \leq n^{-k/10}. \tag{51}$$

We now compare $\mathcal{G}_\gamma$ and $\mathcal{D}_\gamma$. For clarity, define

$$\frac{A}{B} \coloneqq \frac{\mathbf{Pr}_{x\sim f(\mathcal{U}^m)}\left[\mathcal{E}(x)\wedge\mathcal{E}^*(x)\right]}{\mathbf{Pr}_{x\sim f(\mathcal{U}^m)}\left[\mathcal{E}^*(x)\right]} = \mathop{\mathbf{Pr}}_{x\sim\mathcal{G}_\gamma}\left[\mathcal{E}(x)\right]$$

and

$$\frac{A'}{B'} \coloneqq \frac{\mathbf{Pr}_{y\sim\mathcal{D}}\left[\mathcal{E}(y)\wedge\mathcal{E}^*(y)\right]}{\mathbf{Pr}_{y\sim\mathcal{D}}\left[\mathcal{E}^*(y)\right]} = \mathop{\mathbf{Pr}}_{x\sim\mathcal{D}_\gamma}\left[\mathcal{E}(x)\right].$$

By our initial assumption, we know

$$\max\left\{|A-A'|, |B-B'|\right\} \leq \|f(\mathcal{U}^m) - \mathcal{D}\|_{\mathsf{TV}} \leq \varepsilon. \tag{52}$$

Thus,

$$\begin{aligned}
\left|\mathop{\mathbf{Pr}}_{x\sim\mathcal{G}_\gamma}\left[\mathcal{E}(x)\right] - \mathop{\mathbf{Pr}}_{x\sim\mathcal{D}_\gamma}\left[\mathcal{E}(x)\right]\right| &= \frac{|AB'-A'B|}{BB'} \leq \frac{A+B}{B(B-\varepsilon)}\cdot\varepsilon && \text{(by (52))} \\
&\leq \frac{2B}{B(B-\varepsilon)}\cdot\varepsilon \\
&\leq 2\varepsilon\cdot\left(\frac{1}{p_\gamma\mathbf{Pr}_{x\sim\mathcal{F}_\gamma}\left[\mathcal{E}^*(x)\right]-\varepsilon}\right)
\end{aligned}$$

$$\leq 2\varepsilon \cdot \left(\frac{1}{2^{-|S|}\left(1 - \left(\frac{8k}{n^{1/3}}\right)^{k/2}\right) - \varepsilon} \right) \qquad \text{(by (46) \& (47))}$$

$$\leq 2\varepsilon \cdot \left(\frac{1}{2^{-2|S|} - \varepsilon} \right). \qquad \text{(since } n \text{ large in terms of } k, \varepsilon\text{)}$$

Recall that $|S| \leq dk$, where $k \leq \log(1/\varepsilon)/C_d$ for some sufficiently large constant $C_d > 0$ depending only on d. Hence,

$$\left| \mathop{\mathbf{Pr}}_{x \sim \mathcal{G}_\gamma} [\mathcal{E}(x)] - \mathop{\mathbf{Pr}}_{x \sim \mathcal{D}_\gamma} [\mathcal{E}(x)] \right| \leq O_d(\varepsilon). \tag{53}$$

Combining (51) and (53), we conclude

$$\|\mathcal{F}_\gamma - \mathcal{D}_\gamma\|_{\mathsf{TV}} \leq \|\mathcal{F}_\gamma - \mathcal{G}_\gamma\|_{\mathsf{TV}} + \|\mathcal{G}_\gamma - \mathcal{D}_\gamma\|_{\mathsf{TV}} \leq n^{-k/10} + O_d(\varepsilon) \leq O_d(\varepsilon)$$

for large enough n. □

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il