Bare quantum simultaneity versus classical interactivity in communication complexity

Dmitry Gavinsky[*]

January 5, 2021

Abstract

A relational bipartite communication problem is presented that has an efficient quantum simultaneous-messages protocol, but no efficient classical two-way protocol.

1 Introduction

The physical universe that we have the pleasure to experiment with may not be classical. For instance, from the point of view of Popperian methodology [Pop34], the statements _the Lorentz transformation adequately describes the kinematics of the Universe_ and _the Bell inequalities can be violated_ both are plausible, and together they imply that the Universe is _non-local_ – in particular, _non-classical_.[1]

Among various non-classical theories, _quantum mechanics_ is, probably, the most adequate: first of all, it is very accurate in predicting the _probabilities_ of experimental outcomes (and there are reasons to believe that this is the best we can hope for prophesy-wise); on the other hand, quantum mechanics is compatible with our other plausible theories in the _regimes that have been tested experimentally_ (as well as in those ones that we can expect to be able to test any time soon). It is interesting to identify and investigate those scenarios where the predictions of quantum mechanics are _most non-classical_, as studying physical reality can be entertaining and might be of philosophical significance.[2]

This problem makes sense, in particular, in the context of _computational complexity theory_: e.g., we may ask whether computational devices that have at their disposal the flexibility

---

[1] Despite its apparent formality, the above statement overlooks various _hidden assumptions_: e.g., that of _unpredictable input_ in the Bell-violating experiments. Such assumptions are likely unavoidable in any mathematical treatment of the phenomenal reality. These problems, albeit of highest interest and importance, are beyond our current scope.

[2] The significance of understanding quantum mechanics seems to be at least two-fold. On the one hand, the theory is among the frontiers of physical research, and analytical investigation of the possibility of a priori physical knowledge has led to some of the deepest ontological and epistemological doctrines so far. On the other hand, quantum mechanics is intimately related to the problem of causation, which might be not as fundamental as the problem of a priori synthetic knowledge, but is nevertheless very important.

offered by quantum mechanics are qualitatively stronger than their classical counterparts. Obviously, the question would sound most natural in the context of *Turing machines*, but addressing that seems to be beyond our current understanding. Viewing this line of research as an endeavour to understand the "quantum aspect" of physical reality, we would like to minimise the number of assumptions made in the analysis, in particular, to avoid postulating unproven classical hardness of computational problems.

One of the richest computational models that we already know how to analyse is the setting of *communication complexity*, and in this work we will be investigating *the superiority of quantum over classical models in the context of two-player communication*. Here is a brief informal introduction of some concepts of communication complexity.

- In the model of *bipartite* communication there are two <u>players</u>, *Alice* and *Bob*, who receive one portion of input each: Alice gets x and Bob gets y. Their goal is to use the allowed type of communication (as described next) in order to compute an answer that would be correct with respect to the received pair (x, y).

- The three main bipartite <u>layouts</u> are *two-way (interactive) communication*, *one-way communication* and *simultaneous message passing (SMP)*. In the first case the players can exchange messages interactively before answering, in the second case only Alice can send a message to Bob (who then answers), in the third case both Alice and Bob send one message each to a third participant – the *referee* (who then answers).

- Communication problems determine which answers are correct. The three main <u>types of problems</u> are *total functions*, *partial functions* and *relations*: in the first case there is exactly one correct answer for each possible pair of input values, and the set of those pairs equals the direct product of possible inputs of Alice and possible inputs of Bob; the second case is similar, but the set of possible inputs can be arbitrary; in the third case multiple correct answers for the same pair of input values are allowed.

- An *efficient* solution is a communication protocol where the players use at most poly-logarithmic (in the input length) amount of communication and produce a right answer with high confidence.

- Communication models can be strengthened by *shared randomness*, which corresponds to allowing the players to use *mixed strategies* (this can be helpful only in the weakest among the layouts – the *SMP*), or by *shared entanglement*, which allows the players to share any (input-independent) quantum state and use it while running the protocol.[3]

Quantum communication complexity has been an active area of research over the last few decades. Among numerous results in the field, many of which surprised us a lot when they came out, the most relevant to this work are the following:

- In 1998 a *partial function* was demonstrated [BCW98] for which in *zero-error regime* quantum protocols had exponential advantage over the classical ones (both one-way and interactive).

- In 1999 a *partial function* was demonstrated [Raz99] that had an efficient *quantum two-way protocol*, but no efficient *classical two-way protocol*.

- In 2001 a *total function* was demonstrated [BCWdW01] that had an efficient *quantum simultaneous-messages protocol without shared randomness*, but no efficient *classical simultaneous-messages protocol without shared randomness*.

- In 2004 a *relation* was demonstrated [BYJK04] that had an efficient *quantum simultaneous-*

---

[3] Sometimes in this work we call a model *bare* to emphasise that it allows no shared resources.

messages protocol without shared randomness, but no efficient *classical one-way protocol*.

- In 2007 a *partial function* was demonstrated [GKK$^+$08] that had an efficient *quantum one-way protocol*, but no efficient *classical one-way protocol*.
- In 2008 a *relation* was demonstrated [Gav08] with an efficient *quantum one-way protocol*, but no efficient *classical two-way protocol*.
- In 2010 a *partial function* was demonstrated [KR11] with an efficient *quantum one-way protocol*, but no efficient *classical two-way protocol*.
- In 2016 a *partial function* was demonstrated [Gav20] with an efficient *quantum simultaneous-messages protocol with shared entanglement*, but no efficient *classical two-way protocol*.
- In 2017 a *partial function* was demonstrated [Gav19] with an efficient *quantum simultaneous-messages protocol without shared randomness*, but no efficient *classical simultaneous-messages protocol*, even *with shared randomness*.

Looking for the regimes where the predictions of quantum mechanics are as far as possible from those of classical mechanics, we can ask the following questions:

- What is the *weakest* quantum communication model that outperforms its classical counterpart?
- What is the *strongest* classical communication model that is outperformed by the quantum counterpart?

Both of these questions have been answered: the weakest quantum model, *SMP*, can outperform the classical *SMP* [BCWdW01], and the strongest classical model, two-way, can be outperformed by the quantum two-way [Raz99].

Can these two separations be "united"? In other words:

- Can quantum *SMP* (the quantum weakest) outperform classical two-way communication (the classical strongest)?

Although some of the separations mentioned earlier come rather close to giving a positive answer, no such example has been known before. In particular, the strongest classical model known to be outperformed by quantum *SMP* is classical *SMP* with shared randomness [Gav19], and the weakest quantum model known to outperform classical two-way communication is quantum *SMP* with shared entanglement [Gav20].[4]

This work demonstrates a relational problem that has a quantum simultaneous-messages protocol (with no shared resources) of cost $O(\log n)$, and whose classical two-way complexity is $\Omega\big(n^{1/3}\big)$. In other words, *qualitative advantage over the strongest feasible model of classical communication is inherent even in the weakest reasonable quantum model.*

The communication problem will be called *Gap Hamming Relation (GHR)*, and it can be viewed as a relational version of the *Gap Hamming Distance (GHD)* promise function, well known in the context of classical communication complexity.

2 Preliminaries

We will write $[n]$ to denote the set $\{1, \ldots, n\} \subset \mathbb{N}$ and $[n] - 1$ to denote $\{0, \ldots, n - 1\}$. Let (a, b), $[a, b]$, $[a, b)$ and $(a, b]$ denote the corresponding open, closed and half-open intervals in $\mathbb{R}$. We will use "|" to denote divisibility.

---

[4] Intuitively, shared entanglement is a significant additional "quantum resource": the players are supposed to remain *entangled* – that is, to maintain an inherently non-classical condition – for indefinitely long period of time between preparing the "set-up" and actually receiving the input values and executing the protocol.

Let x_i address the i'th bit of x for $x \in \{0,1\}^n$ and $i \in [n]$. Similarly, for $S \subseteq [n]$, let x_S denote the $|S|$-bit string, consisting of (naturally-ordered) bits of x, whose indices are in S. To avoid ambiguity, we will sometimes write $x|_S$ ($x|_i$) instead of x_S (x_i). For a set (or a family) A, we will write $A|_i$ and $A|_S$ to address, respectively, $\{x_i | x \in A\}$ and $\{x_S | x \in A\}$. For $y \in \{0,1\}^n$, let $x \oplus y$ stand for the bit-wise XOR between the two strings.

Let $|x|$ denote the Hamming weight of x. At times we will assume (the trivial) isomorphisms among the n-bit strings, the subsets of $[n]$ and the elements of $[2^n] - 1$. In particular, the notation $\binom{[n]}{k}$ will stand for $\{x \in \{0,1\}^n \big| |x| = k\}$, and $x \cap y$ will address the set $\{i \in [n] \big| x_i = y_i = 1\}$. We will write $\langle x, y \rangle$ for $(|x \oplus y| \mod 2)$, letting it extend naturally to the cases of x, y or both being elements of $[2^n] - 1$.

Let σ_i denote the permutation of $[n]$ which is the i'th cyclic shift (i.e., $\sigma_i(j) = i + j$ if $i + j \leq n$ and $i + j - n$ otherwise). For $x \in \{0,1\}^n$, denote by $\sigma_j(x)$ the j-bit cyclic shift of x – that is, the string of n bits where on $\sigma_j(i)$'th position is x_i for each $i \in [n]$.

We use square brackets $[\ldots]$ to denote events and capital letters from a "mildly calligraphic" font to represent random variables. We will consider probabilities, expectations and so on, conditioned on random variables (sometimes in addition to conditioning on certain events), viewing the corresponding expressions as random variables by themselves. E.g., if $\mathcal{X}$ and $\mathcal{Y}$ are random variables, then $\boldsymbol{Pr}\big[\mathcal{X} = 1 \big| \mathcal{Y} = 1\big] \in \mathbb{R}$, but both $\boldsymbol{Pr}\big[\mathcal{X} = 1 \big| \mathcal{Y}\big]$ and $\boldsymbol{E}\big[\mathcal{X} = 1 \big| \mathcal{Y}\big]$ are random variables whose values are determined by that of $\mathcal{Y}$.

Let $\mathcal{U}_A$ denote the uniform distribution over the elements of A. Sometimes (e.g., in subscripts) we will write "$\Subset A$" instead of "$\sim \mathcal{U}_A$". We will write things like $\{0,1\}^{n+n}$ when the set should be viewed as $\{0,1\}^n \times \{0,1\}^n$: e.g., $(\mathcal{X}, \mathcal{Y}) \Subset \{0,1\}^{n+n}$ refers to the uniform distribution of two independent n-bit strings, and $A \times B \subseteq \{0,1\}^{n+n}$ will be used when $A \times B$ is a combinatorial rectangle in $\{0,1\}^n \times \{0,1\}^n$.

We will denote by $H(\ldots)$ the binary entropy of a discrete random variable, and by $\boldsymbol{I}\big[\ldots : \ldots\big]$ the corresponding mutual information: $H(\mathcal{X}) = \sum_{x \in \mathrm{supp}(\mathcal{X})} \boldsymbol{Pr}[\mathcal{X} = x] \cdot \log\big(\frac{1}{\boldsymbol{Pr}[\mathcal{X}=x]}\big)$ and $\boldsymbol{I}\big[\mathcal{X} : \mathcal{Y}\big] = H(\mathcal{X}) + H(\mathcal{Y}) - H(\mathcal{X}, \mathcal{Y})$ for every $\mathcal{X}$ and $\mathcal{Y}$ – naturally generalising the notation to the event- and variable-conditional cases.

We let $\perp$ and $\top$ denote, respectively, the false and the true values of Boolean predicates. More generally, we will use the Boolean domain $\{\perp, \top\}$ instead of $\{0,1\}$ when there is, subjectively, a *strong logical flavour* to the values.

We will use Dirac's notation (*bra-ket*) for quantum messages and manipulations with them.

3 Communication complexity

For an excellent survey of classical communication complexity, see [KN97]. Quantum communication models differ from their classical counterparts in two aspects: the players are allowed to send *quantum messages* (consequently, the complexity is measured in *qubits*) and to perform arbitrary *quantum operations* locally.[5] An excellent survey of quantum communication complexity is [BCMdW10].

Of central importance to this work is the model of bipartite *Simultaneous Message Passing (SMP)*, where there are three participants: *players* Alice and Bob, and *the referee*. An *SMP-*

---

[5] We say that a communication model allows *shared entanglement* if the players can share any (input-independent) quantum state and use it in the protocol (in the case of simultaneous message passing, entanglement is only allowed between Alice and Bob). Models with shared entanglement will not be used in this work, but they are mentioned in some discussions.

protocol has the following structure: Alice receives x and sends her message to the referee; at the same time, Bob receives y and sends his message to the referee; the referee uses the content of the two received messages to compute the answer.

In the *quantum version* of *SMP*, denoted by $\mathcal{Q}^{\|}$, the players can send quantum messages and the referee can apply any quantum measurement to compute the answer.

Another important communication model is that of *interactive (two-way) communication*. We will be primarily interested in the *classical randomised version*, denoted by $\mathcal{R}$: here two players, Alice and Bob, receive the two parts of input x and y and interact in order to compute their answer.

A problem is said to have an efficient solution (in a given model) if for any constant $\varepsilon > 0$ there exists a protocol that transfers at most *poly-logarithmic* (in the input length) number of (qu)bits and produces a correct answer with probability at least ε. Note that if the considered problem is *functional*, then it is sufficient for a protocol to succeed with any probability $1/2 + \Omega(1)$: a constant number of independent repetitions would allow reducing the error below ε. On the other hand, in the *relational* case error reduction by repetition is not necessarily possible, as there can be multiple correct answers; accordingly, in order to claim that a relation has an efficient solution, one has to make sure that there is an efficient protocol for *every* constant positive error bound.

3.1   Gap Hamming Distance problem and its variations

As usual, we use n to denote the input length of a computational problem. To avoid unnecessary complication of the notation, in this work we will always assume that $n > 1$ and $2 \,|\, \log n$.

The *Gap Hamming Distance (GHD)* is the following bipartite functional problem:

$$GHD_d : \{0,1\}^{n+n} \to \{0,1\};$$

$$GHD_d(x,y) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } |x \oplus y| \geq n/2 + d; \\ 0 & \text{if } |x \oplus y| \leq n/2 - d; \\ \text{undefined} & \text{otherwise.} \end{cases}$$

In the above definition $d \in [n/2]$ is a parameter; the most interesting regime (both in general and for us) is that of $d \in \Theta(\sqrt{n})$, so we will write GHD to denote $GHD_{\sqrt{n}}$.

It has been demonstrated [CR11] that the $\mathcal{R}$-complexity of GHD is $\Omega(n)$; moreover, a nearly-trivial lower bound of $\Omega(\sqrt{n})$ has been known as folklore since the introduction of GHD in [IW03].

In this work we will relax GHD in order to let the model $\mathcal{Q}^{\|}$ solve it efficiently, at the same time preserving – up to a polynomial – the $\mathcal{R}$-hardness of the problem.[6] Our relaxation will turn the original problem into a relation (the possibility of a functional separation is left as an open problem, cf. Section 7).

Informally, GHD can be viewed as making the players decide whether x and y are slightly closer together or slightly further apart than the expected $n/2$; the high-level idea behind turning it into an easier relational problem is to ask them for a *transformation* (of some pre-determined form) that maps x to a string, whose distance from y is somehow biased.

---

[6] As we are interested in *structural* separation between $\mathcal{Q}^{\|}$ and $\mathcal{R}$, we only need a lower bound of $n^{\Omega(1)}$ on the $\mathcal{R}$-complexity of any problem that admits a feasible solution in $\mathcal{Q}^{\|}$ – that is, a protocol of complexity poly-log(n).

Let us have a look at the following candidate problem:
$$\widetilde{GHR} \subseteq \{0,1\}^{n+n} \times \{0,1\}^n;$$
$$\widetilde{GHR} \overset{\text{def}}{=} \left\{(x,y,\tau) \,\middle|\, |\tau \oplus x \oplus y| \leq n/2 - \sqrt{n}\right\}.$$

That is, the valid answers for the input pair (x,y) are all such $\tau \in \{0,1\}^n$ that make $\tau \oplus x$ at most $(n/2 - \sqrt{n})$-far from y. This relaxation is *too radical*, as $\widetilde{GHR}$ is easy even for classical one-way communication with shared randomness (let alone $\mathcal{R}$, which is a strictly stronger model). Indeed, Alice and Bob can use shared randomness to generate a sufficiently long sequence of uniformly-random n-bit strings $\mathcal{Z}_1, \ldots, \mathcal{Z}_t$, then Alice can send to Bob the index of $\mathcal{Z}_{i_0}$ that is closest to x (with high probability the distance will be less than $n/2 - \sqrt{n}$, as t is a large enough constant) and Bob can output $\tau = \mathcal{Z}_{i_0} \oplus y$.

To make our relation harder than $\widetilde{GHR}$ (in order to preserve the $\mathcal{R}$-hardness of the original problem), we reduce the number of allowed transformations (that is, of possible answers) to polynomial. On the other hand, the structure of the allowed transformations makes the new problem more convenient for $\mathcal{Q}^{\|}$-protocols.

Notation.

$$\forall s \in \{0,1\}^{\log n}, \, i \in [n]: \quad \tau_s|_i \overset{\text{def}}{=} \langle s, i \rangle; \quad [7]$$

$$\forall j \in [n], \, s \in \{0,1\}^{\log n}: \quad \Delta_{j,s}(x,y) \overset{\text{def}}{=} \left|\sigma_j(\tau_s \oplus x) \oplus y\right|.$$

$$\aleph(x,y) \overset{\text{def}}{=} \begin{cases} \top & \text{if} \quad \displaystyle\sum_{\substack{j,s: \\ n/2-\sqrt{n}/2 \leq \Delta_{j,s}(x,y) \leq n/2+\sqrt{n}/2}} \left(\Delta_{j,s}(x,y) - \frac{n}{2}\right)^2 \leq \frac{n^3}{9}; \\ \bot & \text{otherwise.} \end{cases} \tag{1}$$

That is, $\{\tau_s\}$ is the standard Fourier basis for $\{0,1\}^{\log n} \to \{\pm 1\}$ in the "$\{0,1\}$-form" (in particular, $|\tau_{s_1} \oplus \tau_{s_2}| \equiv n/2$ for all $s_1 \neq s_2$), and allowed are the n^2 transformations that combine bit-flipping according to some τ_s with a cyclic rotation σ_j. The predicate $\aleph(x,y)$ asserts that both values are, in some sense, *typical*.[8]

Definition 1 (*Gap Hamming Relation (GHR)*)**.** *The bipartite relation GHR is a subset of* $\{0,1\}^n \times \{0,1\}^n \times \left([n] \times \{0,1\}^{\log n}\right)^{\log n}$ *– that is, the input consists of two n-bit strings and the output is a sequence of $\log n$ pairs (j_i, s_i). The sequence $((j_1,s_1), \ldots, (j_{\log n}, s_{\log n}))$ is a valid answer to $GHR(x,y)$ if the following holds:*

$$\begin{cases} \left|\left\{i : \Delta_{j_i,s_i}(x,y) \notin \left[n/2 - \sqrt{n}/2, \, n/2 + \sqrt{n}/2\right]\right\}\right| \geq \log n/2 & \text{if } \aleph(x,y); \\ \top & \text{otherwise.} \end{cases}$$

---

[7] Recall that $\langle x, y \rangle$ stands for $(|x \oplus y| \mod 2)$.

[8] The constants used in the definition of the predicate are such that, on the one hand, $\aleph(\mathcal{X}, \mathcal{Y})$ is satisfied almost always by *uniformly-random $\mathcal{X}$ and $\mathcal{Y}$* (showing this will be important for the analysis), while on the other hand, $\aleph(x,y)$ testifies that our quantum *SMP*-protocol exhibits certain "super-classical" behaviour when its input is (x,y). The structure of our communication problem $GHR(x,y)$ will be such that this behaviour results in a correct answer, while every classical approach to solving GHR is either fruitless or inefficient.

Intuitively, if we ignore conditioning on $\aleph(x, y)$, then a valid answer to $GHR(x, y)$ is a sequence of $\log n$ transformations, at least half of which map the value of x either *somewhat close to* or *somewhat far from* the value of y. We will see that GHR is easy for $\mathcal{Q}^{\|}$ but hard for $\mathcal{R}$.

Intuition behind the GHR problem

The players are required to find a short list of transformations of x that are likely to make its distance from y slightly different from the expected $n/2$ for $(x, y) \in \{0, 1\}^{n+n}$. Intuitively one can view the family of n^2 outcomes of the allowed transformations

$$\left(\sigma_j(\tau_s \oplus \mathcal{X}) \right)_{j \in [n], \, s \in \{0,1\}^{\log n}}$$

as nearly-independent and uniformly distributed when $\mathcal{X} \in \{0, 1\}^n$. [9] Then the challenge faced by the players is *to find among these n^2 random strings some that are biased either away from or towards y*. Note that the required bias is $\sqrt{n}/2$, and we expect most of the n^2 distances to be around the expected $n/2$ within the order of magnitude of the standard deviation of the corresponding binomial distribution – that is, inside $n/2 \pm O(\sqrt{n})$. [10]

The reason why, intuitively, this problem may have an efficient quantum simultaneous-message protocol is the following. If we let Alice and Bob send to the referee, respectively, $1/\sqrt{n} \cdot \sum_{i=1}^n (-1)^{x_i} \cdot |i\rangle$ and $1/\sqrt{n} \cdot \sum_{i=1}^n (-1)^{y_i} \cdot |i\rangle$, then a simple multi-outcome quantum measurement exists, such that every outcome (j, s) combines applying an allowed transformation to the state representing x with estimating the resulting inner product with the state representing y. That is, the measurement can be interpreted as inspecting the values of

$$\left| \frac{1}{n} \cdot \sum_{i=1}^n (-1)^{\left(\sigma_j(\tau_s \oplus x)\right)_i + y_i} \cdot \langle i | i \rangle \right| = \left| \frac{2 \cdot \left| \sigma_j(x \oplus \tau_s) \oplus y \right|}{n} - 1 \right| = \left| \frac{2 \cdot \Delta_{j,s}(x, y)}{n} - 1 \right|$$

for every (j, s), followed by outputting one of the pairs in a manner that reflects the observed distribution of $\Delta_{j,s}(x, y)$. As the probability of the corresponding measurement outcome is proportional to the square of the inner product, the whole procedure favours those outcomes (j, s) that represent the transformations introducing more bias. See Section 4 for details.

The reason why, intuitively, GHR may be hard for classical interactive protocols is that solving the problem requires estimating the Hamming distance between a pair of uniformly-random n-bit string, which is known to be hard: as mentioned earlier, the $\mathcal{R}$-complexity of GHD is $\Omega(n)$. We have already seen how this intuition breaks down when the family of allowed transformations is exponentially-large, but allowing only polynomially-many choices indeed preserves the classical hardness of the original GHD: when the number of choices is very limited, *a protocol must find out more about each possibility*. See Section 6 for details.

Hence, GHR is a plausible candidate for revealing qualitative advantage of bare quantum simultaneity over classical interactivity: On the one hand, an efficient $\mathcal{Q}^{\|}$-protocol can sample among the allowed transformations of x those that make its distance from y biased away from the expected $n/2$; on the other hand, we may expect that the distances from the interval $n/2 \pm \tilde{O}(\sqrt{n})$ between a pair of n-bit strings are well obscured from efficient classical protocols.

---

[9] Of course, the n-bit values themselves are very far from being independent: the claimed intuitive independence only makes sense in the context of their distance to a uniformly-random $\mathcal{Y}$. Formalising and quantifying this intuition will be the technical core of the analysis of GHR's complexity.

[10] And all the n^2 distances are, almost certainly, inside $n/2 \pm O\left(\sqrt{n \cdot \log n}\right)$.

7

4 Solving *GHR* with simultaneous quantum messages

Let

$$\forall j \in [n],\ s \in \{0,1\}^{\log n}\ :\ \left|u_j^s\right\rangle \overset{\text{def}}{=} \frac{1}{\sqrt{n}} \cdot \sum_{i=1}^{n} (-1)^{\tau_s|_i} \cdot |i\rangle \big|\sigma_j(i)\big\rangle,$$

then

$$\left\{ \left|u_j^s\right\rangle \Big| j \in [n],\ s \in \{0,1\}^{\log n} \right\}$$

is an orthonormal basis for $\mathbb{R}^{n+n}$. [11] Indeed, if $j_1 \neq j_2$, then

$$\left\langle u_{j_1}^{s_1}\Big|u_{j_2}^{s_2}\right\rangle = \frac{1}{n} \cdot \sum_{i=1}^{n} (-1)^{\tau_{s_1}(i)+\tau_{s_2}(i)} \cdot \langle i|i\rangle \cdot \left\langle \sigma_{j_1}(i)\big|\sigma_{j_2}(i)\right\rangle = 0;$$

otherwise, if $j_1 = j_2$ and $s_1 \neq s_2$, then

$$\left\langle u_{j_1}^{s_1}\Big|u_{j_2}^{s_2}\right\rangle = \frac{1}{n} \cdot \sum_{i=1}^{n} (-1)^{\tau_{s_1}(i)+\tau_{s_2}(i)} \cdot \langle i|i\rangle \cdot \left\langle \sigma_{j_1}(i)\big|\sigma_{j_2}(i)\right\rangle = \frac{1}{n} \cdot \sum_{i=1}^{n} (-1)^{\tau_{s_1}(i)+\tau_{s_2}(i)} = 0,$$

as $\{\tau_s\}$ is the "$\{0,1\}$-form" of a Fourier basis for $\{0,1\}^{\log n} \to \{\pm 1\}$.

A protocol for *GHR*. Upon receiving the input, Alice and Bob construct, respectively,

$$|\phi_{Al}\rangle \overset{\text{def}}{=} \frac{1}{\sqrt{n}} \cdot \sum_{i=1}^{n} (-1)^{x_i} \cdot |i\rangle \ \text{ and } \ |\phi_{Bo}\rangle \overset{\text{def}}{=} \frac{1}{\sqrt{n}} \cdot \sum_{i=1}^{n} (-1)^{y_i} \cdot |i\rangle,$$

and send $\log n$ copies of both the states to the referee. The referee then independently applies to $\log n$ copies of the state $|\phi_{Al}\rangle|\phi_{Bo}\rangle$ the complete projective measurement

$$\left\{ \left|u_j^s\right\rangle\!\left\langle u_j^s\right| : j \in [n],\ s \in \{0,1\}^{\log n} \right\}$$

and outputs the list of obtained outcomes $((j_1, s_1), \ldots, (j_{\log n}, s_{\log n}))$ as his answer.

The complexity of this $\mathcal{Q}^{\|}$-protocol is $O\big((\log n)^2\big)$.

Correctness. Assume that $\aleph(x, y)$ holds with respect to the received input (x, y) – otherwise any answer would be correct by the definition of *GHR*. For $i \in [\log n]$, let $(\mathcal{J}_i, \mathcal{S}_i)$ be random variables that take values (j_i, s_i) – that is, the outcome of the i'th referee's measurement. We want to analyse the probability of the event

$$e_i \overset{\text{def}}{=} \left[\Delta_{\mathcal{J}_i,\mathcal{S}_i}(x,y) \in \left[\frac{n}{2} - \frac{\sqrt{n}}{2}, \frac{n}{2} + \frac{\sqrt{n}}{2} \right] \right],$$

computed with respect to the action of the protocol (x and y are fixed, subject to $\aleph(x, y)$).

---

[11] All quantum states considered in this work can be represented as unit vectors in the real space of the corresponding dimension, so we do not lose generality by ignoring the case of complex vectors.

Fix $j \in [n]$ and $s \in \{0,1\}^{\log n}$, then for any $i \in [\log n]$:

$$\boldsymbol{Pr}\big[(\mathcal{J}_i, \mathcal{S}_i) = (j,s)\big] = \left|\left(\langle\phi_{Al}|\langle\phi_{Bo}|\right)\big|u_j^s\rangle\right|^2$$

$$= \frac{1}{n^3} \cdot \left|\left(\sum_{k,j=1}^{n}(-1)^{x_k+y_j} \cdot \langle k|\langle j|\right)\left(\sum_{k=1}^{n}(-1)^{\tau_s|_k} \cdot |k\rangle\big|\sigma_j(k)\rangle\right)\right|^2 \qquad (2)$$

$$= \frac{1}{n^3} \cdot \left(\sum_{k=1}^{n}(-1)^{x_k+\tau_s|_k+y_{\sigma_j(k)}}\right)^2$$

$$= \frac{4}{n^3} \cdot \left(\big|\sigma_j(x\oplus\tau_s)\oplus y\big| - \frac{n}{2}\right)^2 \;=\; \frac{4}{n^3}\cdot\left(\Delta_{j,s}(x,y) - \frac{n}{2}\right)^2$$

and

$$\boldsymbol{Pr}[e_i] = \sum_{\substack{j,s: \\ n/2-\sqrt{n}/2 \le \Delta_{j,s}(x,y) \le n/2+\sqrt{n}/2}} \boldsymbol{Pr}\big[(\mathcal{J}_i, \mathcal{S}_i) = (j,s)\big]$$

$$= \frac{4}{n^3} \cdot \sum_{\substack{j,s: \\ n/2-\sqrt{n}/2 \le \Delta_{j,s}(x,y) \le n/2+\sqrt{n}/2}} \left(\Delta_{j,s}(x,y) - \frac{n}{2}\right)^2 \le \frac{4}{9},$$

where the inequality is the assumption that $\aleph(x,y)$ holds (cf. (1)). [12]

As the outcomes of the referee's measurements are independent,

$$\boldsymbol{Pr}[\text{the protocol errs}]$$

$$\le \boldsymbol{Pr}\left[\left|\left\{i : \Delta_{j_i,s_i}(\mathcal{X},\mathcal{Y}) \in \left[\frac{n}{2}-\frac{\sqrt{n}}{2}, \frac{n}{2}+\frac{\sqrt{n}}{2}\right]\right\}\right| > \frac{\log n}{2}\middle|\aleph(\mathcal{X},\mathcal{Y})\right]$$

$$= \boldsymbol{Pr}\left[\text{more than half of the } e_i\text{-s take place}\middle|\aleph(\mathcal{X},\mathcal{Y})\right] \in n^{-\Omega(1)},$$

as follows from Hoeffding's inequality (Fact 2).

From the above analysis it follows readily that an $\mathcal{Q}^{\|}$-protocol of cost $O(\log n)$ would suffice for guaranteeing accuracy at least $1 - \varepsilon$ for arbitrarily small constant ε: As the definition of GHR allows repeated elements in the output sequence, the players can send sufficiently large *constant* number t of the states $|\phi_{Al}\rangle$ and $|\phi_{Bo}\rangle$ to the referee, who can measure them to generate $((j_1,s_1),\ldots,(j_t,s_t))$, and then output this sequence repeated $\lceil \log n/t \rceil$ times, trimming the last copy if necessary.

Corollary 1. $\mathcal{Q}^{\|}(GHR) \in O(\log n)$.

---

[12] Note that the value (of the right-hand side) of (2) is always at most $1/n$ – that is, the measurement outcome is not determined even if the input is *perfectly aligned* with one of the allowed transformations. This is crucial: while the states $\big|u_j^s\rangle$ form an orthonormal basis, the concept of input's perfect alignment is *not necessarily* monogamous with respect to the transformations (think about $x = y = \bar{0}$), and this is reflected by the uncertainty in the outcome. This is also why we need the power of relations for our separation.

5 Concentration bounds

Let us start by listing some known bounds that this work uses.

Fact 1 (*Markov's and Chebyshev's inequalities*)**.** *Let $\mathcal{X}$ be a random variable, then*

$$\forall t \geq 0 : \boldsymbol{Pr}\big[|\mathcal{X}| \geq t\big] \leq \frac{\boldsymbol{E}\big[|\mathcal{X}|\big]}{t} \quad \text{and} \quad \boldsymbol{Pr}\Big[\big|\mathcal{X} - \boldsymbol{E}[\mathcal{X}]\big| \geq t\Big] \leq \frac{\boldsymbol{Var}[\mathcal{X}]}{t^2}.$$

One of the most convenient formulations of Chernoff-like statements – that is, a tail bound for the sum of mutually-independent random variables – is the well-known Hoeffding's inequality.

Fact 2 (*Hoeffding's inequality*)**.** *Let $\mathcal{X}_1, \ldots, \mathcal{X}_m$ be mutually-independent random variables, such that $\forall i \in [m] : a_i \leq \mathcal{X}_i \leq b_i$. Then*

$$\forall t \geq 0 : \boldsymbol{Pr}\left[\left|\sum_{i=1}^m \mathcal{X}_i - \boldsymbol{E}\left[\sum_{i=1}^m \mathcal{X}_i\right]\right| \geq t\right] \leq 2 \cdot e^{\frac{-2t^2}{\sum_{j=1}^m (b_j - a_j)^2}}.$$

In some situations it will be useful to have a tail-of-sum bound with more relaxed conditions and somewhat tighter guarantees, which we prove next. Intuitively, the following formulation combines some advantages of the one by Angluin and Valiant [AV79] (which beats qualitatively the classical statement in the regime of low expectations) with the universality of Azuma's inequality applied to the Doob martingale (which gets rid of the independence requirement).[13]

Theorem 1 (*A relaxed Chernoff bound*)**.** *Let $\mathcal{X}_1, \ldots, \mathcal{X}_m$ be random variables with countable supports[14], such that $\forall i \in [m] : 0 \leq \mathcal{X}_i \leq 1$. Then*

$$
\begin{aligned}
\forall a, t \geq 0 : \boldsymbol{Pr}\left[\sum_{i=1}^m \mathcal{X}_i \leq a - t \,\middle|\, \sum_{i=1}^m \boldsymbol{E}\Big[\mathcal{X}_i \Big| \mathcal{X}_1, \ldots, \mathcal{X}_{i-1}\Big] \geq a\right] &\leq e^{\frac{-t^2}{2a}}, \\
\boldsymbol{Pr}\left[\sum_{i=1}^m \mathcal{X}_i \geq a + t \,\middle|\, \sum_{i=1}^m \boldsymbol{E}\Big[\mathcal{X}_i \Big| \mathcal{X}_1, \ldots, \mathcal{X}_{i-1}\Big] \leq a\right] &\leq e^{\frac{-t^2}{2a+t}}.
\end{aligned}
\tag{3}
$$

In particular,

$$
\begin{aligned}
\boldsymbol{Pr}\left[\sum_{i=1}^m \mathcal{X}_i \leq a - t\right] &\leq e^{\frac{-t^2}{2a}} + \boldsymbol{Pr}\left[\sum_{i=1}^m \boldsymbol{E}\Big[\mathcal{X}_i \Big| \mathcal{X}_1, \ldots, \mathcal{X}_{i-1}\Big] < a\right], \\
\boldsymbol{Pr}\left[\sum_{i=1}^m \mathcal{X}_i \geq a + t\right] &\leq e^{\frac{-t^2}{2a+t}} + \boldsymbol{Pr}\left[\sum_{i=1}^m \boldsymbol{E}\Big[\mathcal{X}_i \Big| \mathcal{X}_1, \ldots, \mathcal{X}_{i-1}\Big] > a\right],
\end{aligned}
$$

---

[13] See [Sin18] for an excellent overview of Chernoff-like concentration bounds.

[14] The continuous version can be obtained along similar lines under the integrability assumption.

and

$$\forall \mu, t \geq 0 : \sum_{i=1}^{m} \boldsymbol{E}\left[\mathcal{X}_i \middle| \mathcal{X}_1, \ldots, \mathcal{X}_{i-1}\right] \geq \mu \implies \boldsymbol{Pr}\left[\sum_{i=1}^{m} \mathcal{X}_i \leq \mu - t\right] \leq e^{\frac{-t^2}{2\mu}},$$

$$\sum_{i=1}^{m} \boldsymbol{E}\left[\mathcal{X}_i \middle| \mathcal{X}_1, \ldots, \mathcal{X}_{i-1}\right] \leq \mu \implies \boldsymbol{Pr}\left[\sum_{i=1}^{m} \mathcal{X}_i \geq \mu + t\right] \leq e^{\frac{-t^2}{2\mu+t}}. \tag{4}$$

The three formulations are, essentially, equivalent and given only for convenience. Note that we use a *cut-off limit a* (μ in (4)), which can lead to loss of tightness in some situations:[15] a more thorough alternative would be to take into account the actual distribution of $\sum \boldsymbol{E}\left[\mathcal{X}_i \middle| \mathcal{X}_1, \ldots, \mathcal{X}_{i-1}\right]$, which would likely result in a less convenient formulation.

We prove it via a standard approach to this type of inequalities, due to Bernstein – namely, by analysing the expectation of the target value's exponent. The following lemma will be the technical core.

Lemma 1. *Let $\mathcal{X}_1, \ldots, \mathcal{X}_m$ be random variables with countable supports, such that $\forall i \in [m] : 0 \leq \mathcal{X}_i \leq 1$ and*

$$\sum_{i=1}^{m} \boldsymbol{E}\left[\mathcal{X}_i \middle| \mathcal{X}_1, \ldots, \mathcal{X}_{i-1}\right] \leq \mu.$$

Then

$$\forall u \geq 0 : \boldsymbol{E}\left[e^{u \cdot \sum_{i=1}^{m} \mathcal{X}_i}\right] \leq \left((e^u - 1) \cdot \frac{\mu}{m} + 1\right)^m.$$

Proof of Lemma 1. We argue by induction on m.

Base case: $m = 1$.
$$\boldsymbol{E}\left[e^{u \cdot \mathcal{X}_1}\right] = \sum_{x} \boldsymbol{Pr}[\mathcal{X}_1 = x] \cdot e^{ux} \leq \sum_{x} \boldsymbol{Pr}[\mathcal{X}_1 = x] \cdot \left((e^u - 1) \cdot x + 1\right) \leq (e^u - 1) \cdot \mu + 1,$$

where the summing is over $x \in \mathrm{supp}(\mathcal{X}_1)$ and the first inequality follows from the convexity of $e^{u \cdot x}$ on $x \in [0,1]$.

Inductive step: $m > 1$. Let $p_1 \overset{\text{def}}{=} \boldsymbol{E}[\mathcal{X}_1]$, then
$$\boldsymbol{E}\left[e^{u \cdot \sum_{i=1}^{m} \mathcal{X}_i}\right] = \sum_{x} \boldsymbol{Pr}[\mathcal{X}_1 = x] \cdot e^{u \cdot x} \cdot \boldsymbol{E}\left[e^{u \cdot \sum_{i=2}^{m} \mathcal{X}_i} \middle| \mathcal{X}_1 = x\right]$$

$$\leq \sum_{x} \boldsymbol{Pr}[\mathcal{X}_1 = x] \cdot e^{u \cdot x} \cdot \left((e^u - 1) \cdot \frac{\mu - p_1}{m - 1} + 1\right)^{m-1}$$

$$= \boldsymbol{E}\left[e^{u \cdot \mathcal{X}_1}\right] \cdot \left((e^u - 1) \cdot \frac{\mu - p_1}{m - 1} + 1\right)^{m-1},$$

---

[15] E.g., it is likely that the bound guaranteed by Lemma 5 loses a multiplicative log-factor due to this.

as follows from the inductive assumption and the fact that $\sum_{i=2}^{m} \boldsymbol{E}\big[\mathcal{X}_i\big|\mathcal{X}_1 = x, \mathcal{X}_2, \ldots, \mathcal{X}_{i-1}\big] \le \mu - p_1$ for any $x \in \text{supp}(\mathcal{X}_1)$, and the required

$$
\ldots \; \le \; \big((e^u - 1) \cdot p_1 + 1\big) \cdot \left((e^u - 1) \cdot \frac{\mu - p_1}{m - 1} + 1\right)^{m-1}
$$

$$
= \left(\big((e^u - 1) \cdot p_1 + 1\big)^{\frac{1}{m}} \cdot \left((e^u - 1) \cdot \frac{\mu - p_1}{m - 1} + 1\right)^{\frac{m-1}{m}}\right)^{m}
$$

$$
\le \left(\frac{(e^u - 1) \cdot p_1 + 1}{m} + \frac{(e^u - 1) \cdot (\mu - p_1)}{m} + \frac{m - 1}{m}\right)^{m} = \left((e^u - 1) \cdot \frac{\mu}{m} + 1\right)^{m}
$$

follows from the base case and the weighted arithmetic/geometric means inequality. $\blacksquare_{Lemma\ 1}$

We are ready to prove the relaxed Chernoff bound.

Proof of Theorem 1. First suppose

$$
\sum_{i=1}^{m} \boldsymbol{E}\big[\mathcal{X}_i\big|\mathcal{X}_1, \ldots, \mathcal{X}_{i-1}\big] \le \mu,
$$

then

$$
\forall A \ge \mu, \, u \ge 0 : \boldsymbol{Pr}\left[\sum_{i=1}^{m} \mathcal{X}_i \ge A\right] \; \le \; \boldsymbol{Pr}\left[e^{u \cdot \sum_{i=1}^{m} \mathcal{X}_i} \ge e^{u \cdot A}\right]
$$

$$
\le \; e^{-u \cdot A} \cdot \boldsymbol{E}\left[e^{u \cdot \sum_{i=1}^{m} \mathcal{X}_i}\right] \; \le \; e^{-u \cdot A} \cdot \left((e^u - 1) \cdot \frac{\mu}{m} + 1\right)^{m},
$$

where the last two inequalities are, respectively, Markov's (Fact 1) and Lemma 1. Plugging in

$$
e^u \stackrel{\text{def}}{=} \frac{A \cdot (m - \mu)}{\mu \cdot (m - A)},
$$

where the right-hand side is at least 1 (as $A \ge \mu$), we obtain

$$
\boldsymbol{Pr}\left[\sum_{i=1}^{m} \mathcal{X}_i \ge A\right] \le \left(\frac{\mu \cdot (m - A)}{A \cdot (m - \mu)}\right)^{A} \cdot \left(\frac{A \cdot (m - \mu) - \mu \cdot (m - A)}{\mu \cdot (m - A)} \cdot \frac{\mu}{m} + 1\right)^{m}
$$
$$
= \left(\frac{\mu}{A}\right)^{A} \cdot \left(\frac{m - \mu}{m - A}\right)^{m-A},
$$
(5)

which holds whenever $\sum_{i=1}^{m} \boldsymbol{E}\big[\mathcal{X}_i\big|\mathcal{X}_1, \ldots, \mathcal{X}_{i-1}\big] \le \mu$.

Letting $\forall i \in [m] : \overline{\mathcal{X}_i} \equiv 1 - \mathcal{X}_i$, we get that if $\sum_{i=1}^{m} \boldsymbol{E}\big[\mathcal{X}_i\big|\mathcal{X}_1, \ldots, \mathcal{X}_{i-1}\big] \ge \mu$, then

$$
\forall A \le \mu : \boldsymbol{Pr}\left[\sum_{i=1}^{m} \mathcal{X}_i \le A\right] = \boldsymbol{Pr}\left[\sum_{i=1}^{m} \overline{\mathcal{X}_i} \ge m - A\right] \le \left(\frac{m - \mu}{m - A}\right)^{m-A} \cdot \left(\frac{\mu}{A}\right)^{A}, \quad (6)
$$

where the inequality is (5) (note that $\sum_{i=1}^{m} \boldsymbol{E}\big[\overline{\mathcal{X}_i}\big|\overline{\mathcal{X}_1}, \ldots, \overline{\mathcal{X}_{i-1}}\big] \le m - \mu$). [16]

---

[16] The identically-looking right-hand sides of (5) and (6) are semantically different, as the former corresponds to $[A \ge \mu]$ and the latter to $[A \le \mu]$.

Next we are going to use the facts that

$$\forall x \geq 0 : \frac{2x}{2+x} \leq \ln(1+x) \leq x \quad \text{and} \quad \forall x \in [0,1] : \frac{x^2 - 2x}{2 - 2x} \leq \ln(1-x) \leq -x.$$

From (5) it follows that

$$\forall \mu, t \geq 0 : \sum_{i=1}^{m} \boldsymbol{E}\Big[\mathcal{X}_i \big| \mathcal{X}_1, \ldots, \mathcal{X}_{i-1}\Big] \leq \mu \implies$$

$$\ln\left(\boldsymbol{Pr}\left[\sum_{i=1}^{m} \mathcal{X}_i \geq \mu + t\right]\right) \leq (\mu + t) \cdot \ln\left(\frac{\mu}{\mu + t}\right) + (m - \mu - t) \cdot \ln\left(1 + \frac{t}{m - \mu - t}\right)$$

$$\leq t - (\mu + t) \cdot \ln\left(1 + \frac{t}{\mu}\right) \ \leq \ t - (\mu + t) \cdot \frac{2t}{2\mu + t} = -\frac{t^2}{2\mu + t},$$

and from (6) it follows that

$$\forall \mu \geq 0, \ t \in [0, \mu] : \sum_{i=1}^{m} \boldsymbol{E}\Big[\mathcal{X}_i \big| \mathcal{X}_1, \ldots, \mathcal{X}_{i-1}\Big] \geq \mu \implies$$

$$\ln\left(\boldsymbol{Pr}\left[\sum_{i=1}^{m} \mathcal{X}_i \leq \mu - t\right]\right) \leq (\mu - t) \cdot \ln\left(\frac{\mu}{\mu - t}\right) + (m - \mu + t) \cdot \ln\left(1 - \frac{t}{m - \mu + t}\right)$$

$$\leq -t - (\mu - t) \cdot \ln\left(1 - \frac{t}{\mu}\right) \ \leq \ -t - \left(-t + \frac{t^2}{2\mu}\right) = -\frac{t^2}{2\mu}.$$

Thus, we have established (4).

Let η denote the distribution of $\mathcal{X}_1, \ldots, \mathcal{X}_m$, and let η' be their distribution, conditioned on $\left[\sum_{i=1}^{m} \boldsymbol{E}_\eta[\mathcal{X}_i | \mathcal{X}_1, \ldots, \mathcal{X}_{i-1}] \geq a\right]$ (assuming that the event's probability is positive). Then it is easy to see that

$$\sum_{i=1}^{m} \boldsymbol{E}_{\eta'}\Big[\mathcal{X}_i \big| \mathcal{X}_1, \ldots, \mathcal{X}_{i-1}\Big] \geq a,$$

which, together with (4), implies the first inequality in (3). The second one follows by a similar argument. ∎ *Theorem 1*

Lastly, there are cases where typical Chernoff-like bounds are insufficient: aiming for universality, they usually compromise tightness. The following ad hoc statement deals with concentration near the expectation under symmetric binomial distributions.

Lemma 2. *Let $2|m$ and $\mathcal{X}_1, \ldots, \mathcal{X}_m$ be mutually-independent random variables, such that $\forall i \in [m] : \boldsymbol{Pr}[\mathcal{X}_i = 0] = \boldsymbol{Pr}[\mathcal{X}_i = 1] = 1/2$. Then for every $a < b$ in $[m] \cup \{0\}$:*

$$\boldsymbol{Pr}\left[a \leq \sum_{i=1}^{m} \mathcal{X}_i \leq b\right]$$

$$\geq \sqrt{\frac{2}{\pi m}} \cdot (b - a) - \sqrt{\frac{8}{9\pi m^3}} \cdot \left(\left(b - \frac{m}{2}\right)^3 - \left(a - \frac{m}{2}\right)^3\right) - O\left(\frac{1}{m}\right).$$ [17]

13

To prove it, we will use Stirling's approximation in the following form.

Fact 3 (*Stirling's approximation*)**.**

$$n! \in \sqrt{2\pi n} \cdot \left(\frac{n}{e}\right)^n \cdot \left(1 + O\left(\frac{1}{n}\right)\right).$$

Proof of Lemma 2.

$$\mathbf{Pr}\left[\sum_{i=1}^m \mathcal{X}_i = \frac{m}{2}\right] = \binom{m}{m/2} \Big/ 2^m = 2^{-m} \cdot \frac{m!}{\left(\left(\frac{m}{2}\right)!\right)^2} \in \sqrt{\frac{2}{\pi m}} \cdot \left(1 \pm O\left(\frac{1}{m}\right)\right),$$

where the containment is Fact 3. For any $d \in [m/2] \cup \{0\}$,

$$\frac{\mathbf{Pr}\left[\sum \mathcal{X}_i = \frac{m}{2} + d\right]}{\mathbf{Pr}\left[\sum \mathcal{X}_i = \frac{m}{2}\right]} = \frac{\mathbf{Pr}\left[\sum \mathcal{X}_i = \frac{m}{2} - d\right]}{\mathbf{Pr}\left[\sum \mathcal{X}_i = \frac{m}{2}\right]} = \frac{\left(\left(\frac{m}{2}\right)!\right)^2}{\left(\frac{m}{2} - d\right)! \cdot \left(\frac{m}{2} + d\right)!}$$

$$= \frac{\left(\frac{m}{2} - d + 1\right) \cdots \left(\frac{m}{2}\right)}{\left(\frac{m}{2} + 1\right) \cdots \left(\frac{m}{2} + d\right)} > \left(\frac{\frac{m}{2} - d}{\frac{m}{2}}\right)^d = \left(1 - \frac{2d}{m}\right)^d \geq 1 - \frac{2d^2}{m}.$$

Then

$$\mathbf{Pr}\left[a \leq \sum_{i=1}^m \mathcal{X}_i \leq b\right] \geq \mathbf{Pr}\left[\sum \mathcal{X}_i = \frac{m}{2}\right] \cdot \int_{a - \frac{m}{2}}^{b - \frac{m}{2}} \left(1 - \frac{2x^2}{m}\right) dx$$

$$\geq \left(\sqrt{\frac{2}{\pi m}} \cdot \left(1 - O\left(\frac{1}{m}\right)\right)\right) \cdot \left(b - a - \frac{2}{3m} \cdot x^3 \Big|_{a - \frac{m}{2}}^{b - \frac{m}{2}}\right)$$

$$\geq \sqrt{\frac{2}{\pi m}} \cdot (b - a) - \sqrt{\frac{8}{9\pi m^3}} \cdot \left(\left(b - \frac{m}{2}\right)^3 - \left(a - \frac{m}{2}\right)^3\right) - O\left(\frac{1}{m}\right),$$

as desired. ∎ *Lemma 2*

6 The $\mathcal{R}$-complexity of *GHR*: a lower bound

The definition of *GHR* relies on the predicate $\aleph(\cdot, \cdot)$, and it will be useful to know how likely $\aleph(\mathcal{X}, \mathcal{Y})$ is to hold with respect to our main input distribution for *GHR* – the uniform.

Theorem 2.

$$\Pr_{(\mathcal{X}, \mathcal{Y}) \in \{0,1\}^{n+n}} [\aleph(\mathcal{X}, \mathcal{Y})] \in 1 - O\left(\frac{(\ln n)^2}{\sqrt{n}}\right).$$

To prove it, we will use several auxiliary lemmas. The first one is a concentration statement for the bit-wise XOR of a uniformly-random n-bit string with its own cyclic shift.

---

[17] Here again, we do not aim for tightness beyond the needs, preferring simplicity and self-containment of the argument: in particular, we will use the convenient $[(1 - 2d/m)^d \geq 1 - 2d^2/m]$ instead of the less wasteful $[(1 - 2d/m)^d \geq \exp\left(-2d^2/(m - 2d)\right)]$ or directly analysing the left-hand side. The assumption that $2|m$ is made in order to simplify the notation: dropping it would not affect the claimed bound.

Lemma 3. *Let $i \in [n-1]$, $s \in \{0,1\}^n$ and $\mathcal{A} \sim \mathcal{U}_{\{0,1\}^n}$. Then*

$$\forall t \geq 0 : \boldsymbol{Pr}\left[\left|\left|\mathcal{A} \oplus \sigma_i(\mathcal{A}) \oplus s\right| - \frac{n}{2}\right| \geq t\right] \leq 4 \cdot e^{\frac{-t^2}{2n}}.$$

Proof of Lemma 3. Let $k|n$ be the orbit length of σ_i in $[n]$ and $\left\{o_i \big| i \in [n/k]\right\}$ be the corresponding partition of $[n]$ into orbits. For each $i \in [n/k]$, let c_i be the lexicographically-first element of o_i and $o_i' \overset{\text{def}}{=} o_i \setminus \{c_i\}$ (note that the sets $\cup_i o_i'$ and $\left\{c_1, \ldots, c_{n/k}\right\}$ partition $[n]$).

Then

$$\left(\mathcal{A} \oplus \sigma_i(\mathcal{A})\right)|_{\cup o_i'} \sim \mathcal{U}_{\{0,1\}^{\frac{n \cdot (k-1)}{k}}},$$

and therefore,

$$\boldsymbol{Pr}\left[\left|\left|\left(\mathcal{A} \oplus \sigma_i(\mathcal{A}) \oplus s\right)|_{\cup o_i'}\right| - \frac{n \cdot (k-1)}{2k}\right| \geq \frac{t}{2}\right] \leq 2 \cdot e^{\frac{-t^2}{2n}}$$

by Hoeffding's inequality (Fact 2). Similarly, $\left(\mathcal{A} \oplus \sigma_i(\mathcal{A})\right)|_{\{c_i\}} \subseteq \{0,1\}^{n/k}$ and

$$\boldsymbol{Pr}\left[\left|\left|\left(\mathcal{A} \oplus \sigma_i(\mathcal{A}) \oplus s\right)|_{\{c_i\}}\right| - \frac{n}{2k}\right| \geq \frac{t}{2}\right] \leq 2 \cdot e^{\frac{-t^2}{2n}}.$$

By the union bound,

$$\boldsymbol{Pr}\left[\left|\left|\mathcal{A} \oplus \sigma_i(\mathcal{A}) \oplus s\right| - \frac{n}{2}\right| \geq t\right] \leq \boldsymbol{Pr}\left[\left|\left|\left(\mathcal{A} \oplus \sigma_i(\mathcal{A}) \oplus s\right)|_{\cup o_i'}\right| - \frac{n \cdot (k-1)}{2k}\right| \geq \frac{t}{2}\right]$$

$$+ \boldsymbol{Pr}\left[\left|\left|\left(\mathcal{A} \oplus \sigma_i(\mathcal{A}) \oplus s\right)|_{\{c_i\}}\right| - \frac{n}{2k}\right| \geq \frac{t}{2}\right]$$

$$\leq 4 \cdot e^{\frac{-t^2}{2n}},$$

as desired. $\blacksquare_{Lemma\ 3}$

Second, we need a simple bound on the expectation with respect to "anti-correlating" probability distribution.

Lemma 4. *Let μ be a probability distribution with finite support[18] A and $f : A \to \mathbb{R}$, such that $\forall a_1, a_2 \in A : f(a_1) \leq f(a_2) \iff \mu(a_1) \geq \mu(a_2)$. Then*

$$\mathop{\boldsymbol{E}}_{\mathcal{X} \sim \mu}\left[f(\mathcal{X})\right] \leq \mathop{\boldsymbol{E}}_{\mathcal{X} \subseteq A}\left[f(\mathcal{X})\right].$$

---

[18] The continuous version can be obtained along similar lines under the integrability assumption.

Proof of Lemma 4.

$$\underset{\mathcal{X} \sim \mu}{\boldsymbol{E}}\big[f(\mathcal{X})\big] = \sum_{x \in A} f(x) \cdot \mu(x)$$

$$= \frac{1}{|A|} \cdot \left(\left(\sum_{x \in A} f(x) \right) \cdot \left(\sum_{x \in A} \mu(x) \right) + \frac{1}{2} \cdot \sum_{x_1 \neq x_2} \big(f(x_1) - f(x_2)\big) \cdot \big(\mu(x_1) - \mu(x_2)\big) \right)$$

$$= \underset{\mathcal{X} \Subset A}{\boldsymbol{E}}\big[f(\mathcal{X})\big] + \frac{1}{2 \cdot |A|} \cdot \sum_{x_1 \neq x_2} \big(f(x_1) - f(x_2)\big) \cdot \big(\mu(x_1) - \mu(x_2)\big)$$

$$\leq \underset{\mathcal{X} \Subset A}{\boldsymbol{E}}\big[f(\mathcal{X})\big],$$

as

$$\forall x_1, x_2 \in A : \ \big(f(x_1) - f(x_2)\big) \cdot \big(\mu(x_1) - \mu(x_2)\big) \leq 0$$

by assumption. ∎ *Lemma 4*

The third auxiliary lemma bounds the dependence between the Hamming weight of a uniformly-random n-bit string $\mathcal{A}$ and that of $\mathcal{A} \oplus s$ for some fixed $s \in \{0,1\}^n$.

Lemma 5. *Let $s \in \{0,1\}^n$, $\big||s| - n/2\big| \leq d$ and $\mathcal{A} \sim \mathcal{U}_{\{0,1\}^n}$. Then there exists a random variable $\widetilde{\mathcal{A}}$, such that*

$$\boldsymbol{I}\Big[\big|\mathcal{A} \oplus \widetilde{\mathcal{A}} \oplus s\big| : |\mathcal{A}|\Big] = 0$$

and

$$\forall t \geq 0 : \ \boldsymbol{Pr}\left[\big|\widetilde{\mathcal{A}}\big| \geq \frac{16d^2}{n} + t\right] \leq 4 \cdot e^{\ln n - \frac{t}{4 \ln n}} + 2 \cdot e^{\frac{-t^2 n}{224 \cdot t \cdot n \cdot \ln n + 2048 \cdot d^2 \cdot \ln n}}.$$

More specifically,

$$\forall t, l \geq 0 :$$
$$\boldsymbol{Pr}\left[\big|\widetilde{\mathcal{A}}\big| \geq \frac{(2d+2l)^2}{n} + t \ \bigg| \ \big||\mathcal{A}| - n/2\big| \leq l\right] \leq 2 \cdot \left(e^{\frac{-t^2 n}{32 \cdot (2d+2l)^2 \cdot \ln n + 80 t n}} + e^{\ln n - \frac{t}{2 \ln n}} \right).$$

As we will be dealing with the uniform distribution over n-bit strings, typical for us will be the regime of $l, d \in O(\sqrt{n})$, in which case the above statement reads as follows: *the value of $|\mathcal{A} \oplus s|$ is poly-logarithmically-close*[19] *in absolute distance to being independent from $|\mathcal{A}|$ –* clearly, this is a very strong bound on the dependence between a pair of random values whose standard deviations are in $\Theta(\sqrt{n})$.

Proof of Lemma 5. Assume $|s| \geq n/2$ and let $d \overset{\text{def}}{=} |s| - n/2$ (the case of $|s| < n/2$ can be treated similarly). Let $a_1, \ldots, a_{n/2+d}$ be the coordinates where $s(a_i) = 1$ and $b_1, \ldots, b_{\frac{n}{2}-d}$ be the coordinates where $s(b_i) = 0$.

---

[19] The above lemma gives meaningful upper bound for $t \in \Omega\big((\ln n)^2\big)$, which is sufficient for our needs. We would expect the actual $\big|\widetilde{\mathcal{A}}\big|$ – as constructed in the following proof – to be almost always in $O(\ln n)$, or even in $o(\ln n)$: cf. Footnote 15.

We will construct the distribution of $\widetilde{\mathcal{A}}$ bit by bit, then argue that $\left|\mathcal{A} \oplus \widetilde{\mathcal{A}} \oplus s\right|$ doesn't depend on $|\mathcal{A}|$. The high-level intuition behind the construction will be to handle one by one the pairs of coordinates (a_i, b_j) of $\widetilde{\mathcal{A}}$: in this case the fact that $s(a_i) = 1$ and $s(b_j) = 0$ allows for a very efficient symmetric argument, which will be described in Stage 2 below. But due to the fact that the numbers of 0- and 1-coordinates in the string s can be different, we have to treat the "excess" separately, and that will be the purpose of Stage 1 (if we were not assuming for simplicity that $|s| \geq n/2$, we would end up with two very symmetric versions of Stage 1, while Stage 2 would remain unchanged).

Stage 1. Here we describe the distribution of $\widetilde{\mathcal{A}}_{a_1, \ldots, a_{2d}}$ (skipping to Stage 2 if $d = 0$).

Let k_0 be the value of $|\mathcal{A}|$. By the bit-symmetry of both the distribution of $\mathcal{A}$ and the assumption $[|\mathcal{A}| = k_0]$, it holds that

$$\forall i \in [n] : \boldsymbol{Pr}\left[\mathcal{A}_i = 1 \middle| |\mathcal{A}| = k_0\right] = \frac{k_0}{n};$$

in particular, $\boldsymbol{Pr}\left[\mathcal{A}_{a_1} = 1 \middle| |\mathcal{A}| = k_0\right] = k_0/n$. Let:

$$\boldsymbol{Pr}\left[\widetilde{\mathcal{A}}_{a_1} = 1\right] \stackrel{\text{def}}{=} \begin{cases} 1 - \frac{n}{2k_0} & \text{if } 2k_0 > n \text{ and } \mathcal{A}_{a_1} = 1; \\ 1 - \frac{n}{2n - 2k_0} & \text{if } 2k_0 < n \text{ and } \mathcal{A}_{a_1} = 0; \\ 0 & \text{otherwise.} \end{cases}$$

Note that *conditioned on* $[|\mathcal{A}| = k_0]$, it is the case that

$$\boldsymbol{Pr}\left[\mathcal{A}_{a_1} \oplus \widetilde{\mathcal{A}}_{a_1} = 1\right]$$
$$= \boldsymbol{Pr}\left[\mathcal{A}_{a_1} = 1\right] \cdot \boldsymbol{Pr}\left[\widetilde{\mathcal{A}}_{a_1} = 0 \middle| \mathcal{A}_{a_1} = 1\right] + \boldsymbol{Pr}\left[\mathcal{A}_{a_1} = 0\right] \cdot \boldsymbol{Pr}\left[\widetilde{\mathcal{A}}_{a_1} = 1 \middle| \mathcal{A}_{a_1} = 0\right]$$
$$= \begin{cases} \frac{k_0}{n} \cdot \frac{n}{2k_0} & \text{if } 2k_0 > n, \\ \frac{k_0}{n} + \left(1 - \frac{k_0}{n}\right) \cdot \left(1 - \frac{n}{2n - 2k_0}\right) & \text{if } 2k_0 < n, \\ \frac{k_0}{n} & \text{if } 2k_0 = n \end{cases} = \frac{1}{2}. \tag{7}$$

The distribution of $\widetilde{\mathcal{A}}_{a_i}$ for each $i \in [2d]$ will depend on $|\mathcal{A}| = k_0$ and on $\left|\mathcal{A}_{a_1, \ldots, a_{i-1}}\right|$: namely, on $k_{i-1} \stackrel{\text{def}}{=} k_0 - \sum_{j=1}^{i-1} \mathcal{A}_{a_j}$ $\left(= \left|\mathcal{A}_{[n] \setminus \{a_1, \ldots, a_{i-1}\}}\right|\right)$. Additionally, it will depend on the value of $\mathcal{A}_{a_i}$.

Let $n_{i-1} \stackrel{\text{def}}{=} n - i + 1$; note that for every $t \in [n] \setminus \{a_1, \ldots, a_{i-1}\}$ the probability of $[\mathcal{A}_t = 1]$, conditioned on the content of $\mathcal{A}_{a_1, \ldots, a_{i-1}}$ and on $[|\mathcal{A}| = k_0]$, equals $\frac{k_{i-1}}{n_{i-1}}$. Generalising the case of $\widetilde{\mathcal{A}}_{a_1}$, define:

$$\forall i \in [2d] : \boldsymbol{Pr}\left[\widetilde{\mathcal{A}}_{a_i} = 1\right] \stackrel{\text{def}}{=} \begin{cases} 1 - \frac{n_{i-1}}{2k_{i-1}} & \text{if } 2k_{i-1} > n_{i-1} \text{ and } \mathcal{A}_{a_i} = 1; \\ 1 - \frac{n_{i-1}}{2n_{i-1} - 2k_{i-1}} & \text{if } 2k_{i-1} < n_{i-1} \text{ and } \mathcal{A}_{a_i} = 0; \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

Let $\overline{\mathcal{A}} \stackrel{\text{def}}{=} \mathcal{A} \oplus \widetilde{\mathcal{A}}$. Then, similarly to (7), it hods that

$$\boldsymbol{Pr}\left[\overline{\mathcal{A}}_{a_i} = 1 \middle| \mathcal{A}_{a_1, \ldots, a_{i-1}}, |\mathcal{A}|\right] \equiv \frac{1}{2}. \tag{9}$$

17

Let $i \in [2d]$. As $\overline{\mathcal{A}}_{a_1,\dots,a_i} = \mathcal{A}_{a_1,\dots,a_i} \oplus \widetilde{\mathcal{A}}_{a_1,\dots,a_i}$ and the distribution of $\widetilde{\mathcal{A}}_{a_1,\dots,a_i}$ only depends on $|\mathcal{A}|$ and $\mathcal{A}_{a_1,\dots,a_i}$,

$$\boldsymbol{I}\!\left[\overline{\mathcal{A}}_{a_1,\dots,a_i} : \mathcal{A}\,\middle|\,\mathcal{A}_{a_1,\dots,a_i},\, |\mathcal{A}|\right] = 0,$$

and therefore,

$$H\!\left(\overline{\mathcal{A}}_{a_1,\dots,a_i}\,\middle|\,\mathcal{A}_{a_1,\dots,a_i},\, |\mathcal{A}|\right) = H\!\left(\overline{\mathcal{A}}_{a_1,\dots,a_i}\,\middle|\,\mathcal{A}_{a_1,\dots,a_i},\, |\mathcal{A}|,\, \mathcal{A}\right) = H\!\left(\overline{\mathcal{A}}_{a_1,\dots,a_i}\,\middle|\,\mathcal{A}\right). \tag{10}$$

As the distribution of $\widetilde{\mathcal{A}}$ only depends on $\mathcal{A}$ and there are no other "inter-bit" dependencies,

$$\boldsymbol{I}\!\left[\widetilde{\mathcal{A}}_{a_i} : \widetilde{\mathcal{A}}_{a_1,\dots,a_{i-1}}\,\middle|\,\mathcal{A}\right] = \boldsymbol{I}\!\left[\overline{\mathcal{A}}_{a_i} : \overline{\mathcal{A}}_{a_1,\dots,a_{i-1}}\,\middle|\,\mathcal{A}\right] = 0,$$

and therefore,

$$H\!\left(\overline{\mathcal{A}}_{a_1,\dots,a_{i-1}}\,\middle|\,\overline{\mathcal{A}}_{a_i},\, \mathcal{A}\right) = H\!\left(\overline{\mathcal{A}}_{a_1,\dots,a_{i-1}}\,\middle|\,\mathcal{A}\right) = H\!\left(\overline{\mathcal{A}}_{a_1,\dots,a_{i-1}}\,\middle|\,\mathcal{A}_{a_1,\dots,a_{i-1}},\, |\mathcal{A}|\right),$$

where the latter equality is (10). So,

$$
\begin{aligned}
0 &= H\!\left(\overline{\mathcal{A}}_{a_1,\dots,a_{i-1}}\,\middle|\,\mathcal{A}_{a_1,\dots,a_{i-1}},\, |\mathcal{A}|\right) - H\!\left(\overline{\mathcal{A}}_{a_1,\dots,a_{i-1}}\,\middle|\,\overline{\mathcal{A}}_{a_i},\, \mathcal{A},\, \mathcal{A}_{a_1,\dots,a_{i-1}},\, |\mathcal{A}|\right) \\
&= \boldsymbol{I}\!\left[\overline{\mathcal{A}}_{a_1,\dots,a_{i-1}} : \overline{\mathcal{A}}_{a_i},\, \mathcal{A}\,\middle|\,\mathcal{A}_{a_1,\dots,a_{i-1}},\, |\mathcal{A}|\right] \geq \boldsymbol{I}\!\left[\overline{\mathcal{A}}_{a_1,\dots,a_{i-1}} : \overline{\mathcal{A}}_{a_i}\,\middle|\,\mathcal{A}_{a_1,\dots,a_{i-1}},\, |\mathcal{A}|\right].
\end{aligned}
\tag{11}
$$

Therefore,

$$
\begin{aligned}
H\!\left(\overline{\mathcal{A}}_{a_i}\,\middle|\,\mathcal{A}_{a_1,\dots,a_{i-1}},\, |\mathcal{A}|\right) &= H\!\left(\overline{\mathcal{A}}_{a_i}\,\middle|\,\mathcal{A}_{a_1,\dots,a_{i-1}},\, |\mathcal{A}|,\, \overline{\mathcal{A}}_{a_1,\dots,a_{i-1}}\right) \\
&\leq H\!\left(\overline{\mathcal{A}}_{a_i}\,\middle|\,\overline{\mathcal{A}}_{a_1,\dots,a_{i-1}},\, |\mathcal{A}|\right)
\end{aligned}
$$

and

$$
\begin{aligned}
H\!\left(\overline{\mathcal{A}}_{a_1,\dots,a_{2d}}\,\middle\|\,|\mathcal{A}|\right) &= \sum_{i=1}^{2d} H\!\left(\overline{\mathcal{A}}_{a_i}\,\middle|\,\overline{\mathcal{A}}_{a_1,\dots,a_{i-1}},\, |\mathcal{A}|\right) \\
&\geq \sum_{i=1}^{2d} H\!\left(\overline{\mathcal{A}}_{a_i}\,\middle|\,\mathcal{A}_{a_1,\dots,a_{i-1}},\, |\mathcal{A}|\right) = 2d,
\end{aligned}
\tag{12}
$$

where the concluding equality is (9).

Let us estimate $\left|\widetilde{\mathcal{A}}_{a_1,\dots,a_{2d}}\right|$. If we denote $l_i \stackrel{\text{def}}{=} |k_i - {}^{n_i}\!/{}_2|$, then from (8) it follows that

$$\forall i \in [2d]: \ \boldsymbol{Pr}\!\left[\widetilde{\mathcal{A}}_i = 1\,\middle|\,\widetilde{\mathcal{A}}_{[i-1]}\right] \in \left\{\frac{2l_{i-1}}{n_{i-1} + 2l_{i-1}},\, 0\right\}$$

always. As $0 \leq l_i \leq \left|\,|\mathcal{A}| - {}^{n}\!/{}_2\right| + {}^{i}\!/{}_2$ always,

$$\boldsymbol{Pr}\!\left[\widetilde{\mathcal{A}}_i = 1\,\middle|\,\widetilde{\mathcal{A}}_{[i-1]}\right] \leq \frac{2l_0 + 2d}{n},$$

where $l_0 = \left|\,|\mathcal{A}| - {}^{n}\!/{}_2\right|$, and therefore,

$$\sum_{i=1}^{2d} \boldsymbol{E}\!\left[\widetilde{\mathcal{A}}_i\,\middle|\,\widetilde{\mathcal{A}}_{[i-1]}\right] \leq \frac{4dl_0 + 4d^2}{n}.$$

It follows from the relaxed Chernoff bound (Theorem 1) that

$$\forall t, l \geq 0 : \boldsymbol{Pr}\left[\left|\widetilde{\mathcal{A}}_{a_1,\ldots,a_{2d}}\right| \geq \frac{4dl + 4d^2}{n} + t \,\middle|\, \left|\left|\mathcal{A}\right| - n/2\right| = l\right] \leq e^{\frac{-t^2 \cdot n}{8dl + 8d^2 + tn}}. \qquad (13)$$

Note that in the typical regime of $t, l \in O(\sqrt{n})$ the right-hand side of (13) is $e^{-\Omega(t)}$.

Stage 2. Here we describe the distribution of $\widetilde{\mathcal{A}}_{a_{2d+1},\ldots,a_{\frac{n}{2}+d},b_1,\ldots,b_{\frac{n}{2}-d}}$.

Let k_{2d} denote the Hamming weight of the remaining bits of $\widetilde{\mathcal{A}}$ after the completion of Stage 1, namely

$$k_{2d} \stackrel{\text{def}}{=} \left|\mathcal{A}_{a_{2d+1},\ldots,a_{\frac{n}{2}+d},b_1,\ldots,b_{\frac{n}{2}-d}}\right|$$

(note that if $d = 0$ and Stage 1 has been skipped, then $k_{2d} = k_0 = |\mathcal{A}|$). As before, $\forall i \in [n] - 1 : n_i = n - i$.

Our tactics will be to handle one by one the $\frac{n}{2} - d$ pairs (a_{2d+i}, b_i) and toss $\widetilde{\mathcal{A}}_{a_{2d+i},b_i}$ in a way that will guarantee – conditional on all the earlier choices – that the distribution of every $\left|\mathcal{A}_{a_{2d+i},b_i} \oplus \widetilde{\mathcal{A}}_{a_{2d+i},b_i} \oplus s_{a_{2d+i},b_i}\right|$ is the same as the distribution of $\mathcal{B}_1 + \mathcal{B}_2$ for $(\mathcal{B}_1, \mathcal{B}_2) \leftarrow \{0,1\}^2$.

Conditional on $\left[\left|\mathcal{A}_{a_{2d+1},\ldots,a_{\frac{n}{2}+d},b_1,\ldots,b_{\frac{n}{2}-d}}\right| = k_{2d}\right]$, the distribution of $(\mathcal{A}_{a_{2d+1}}, \mathcal{A}_{b_1})$ (as well as of any other pair of coordinates) is perfectly symmetric – although most likely not independent. Accordingly, conditional on $\left[\mathcal{A}_{a_{2d+1}} \neq \mathcal{A}_{b_1}\right]$, the events $\left[\left|\mathcal{A}_{a_{2d+1},b_1} \oplus s_{a_{2d+1},b_1}\right| = 0\right]$ and $\left[\left|\mathcal{A}_{a_{2d+1},b_1} \oplus s_{a_{2d+1},b_1}\right| = 2\right]$ are equiprobable – just like the corresponding events with respect to $|\mathcal{B}_1 + \mathcal{B}_2|$ – so, we only need to use $\widetilde{\mathcal{A}}_{a_{2d+1},b_1}$ in order to balance the probabilities of $\left[\mathcal{A}_{a_{2d+1}} \oplus \widetilde{\mathcal{A}}_{a_{2d+1}} = \mathcal{A}_{b_1} \oplus \widetilde{\mathcal{A}}_{b_1}\right]$ and $\left[\mathcal{A}_{a_{2d+1}} \oplus \widetilde{\mathcal{A}}_{a_{2d+1}} \neq \mathcal{A}_{b_1} \oplus \widetilde{\mathcal{A}}_{b_1}\right]$ (trying to keep the Hamming weight of $\widetilde{\mathcal{A}}$ as small as possible).

It holds that

$$\boldsymbol{Pr}\left[\mathcal{A}_{a_{2d+1}} \neq \mathcal{A}_{b_1}\right] = \frac{2k_{2d} \cdot (n_{2d} - k_{2d})}{n_{2d} \cdot (n_{2d} - 1)}.$$

First let us define a random variable $\mathcal{Z}_1$, such that $\boldsymbol{Pr}\left[\mathcal{A}_{a_{2d+1}} \oplus \mathcal{A}_{b_1} \oplus \mathcal{Z}_1 = 1\right] = 1/2$, afterwards we will use it to construct the distribution of $\widetilde{\mathcal{A}}_{a_{2d+1},b_1}$. Let:

$$\boldsymbol{Pr}[\mathcal{Z}_1 = 1]$$
$$\stackrel{\text{def}}{=} \begin{cases} 1 - \frac{n_{2d} \cdot (n_{2d}-1)}{4k_{2d} \cdot (n_{2d}-k_{2d})} & \text{if } \left|\frac{n_{2d}}{2} - k_{2d}\right| < \frac{\sqrt{n_{2d}}}{2} \text{ and } \mathcal{A}_{a_{2d+1}} \neq \mathcal{A}_{b_1}; \\ 1 - \frac{n_{2d} \cdot (n_{2d}-1)}{2n_{2d} \cdot (n_{2d}-1) - 4k_{2d} \cdot (n_{2d}-k_{2d})} & \text{if } \left|\frac{n_{2d}}{2} - k_{2d}\right| > \frac{\sqrt{n_{2d}}}{2} \text{ and } \mathcal{A}_{a_{2d+1}} = \mathcal{A}_{b_1}; \\ 0 & \text{otherwise.} \end{cases} \qquad (14)$$

Note that

$$n_{2d} \cdot (n_{2d} - 1) = 4k_{2d} \cdot (n_{2d} - k_{2d}) \iff \left|\frac{n_{2d}}{2} - k_{2d}\right| = \frac{\sqrt{n_{2d}}}{2},$$

and therefore the probability of $[\mathcal{Z}_1 = 1]$ is well-defined.

Then
$$\boldsymbol{Pr}\Big[\mathcal{A}_{a_{2d+1}} \oplus \mathcal{A}_{b_1} \oplus \mathcal{Z}_1 = 1\Big] = \boldsymbol{Pr}\Big[\mathcal{A}_{a_{2d+1}} \neq \mathcal{A}_{b_1}\Big] \cdot \boldsymbol{Pr}\Big[\mathcal{Z}_1 = 0\Big|\mathcal{A}_{a_{2d+1}} \neq \mathcal{A}_{b_1}\Big]$$
$$+ \boldsymbol{Pr}\Big[\mathcal{A}_{a_{2d+1}} = \mathcal{A}_{b_1}\Big] \cdot \boldsymbol{Pr}\Big[\mathcal{Z}_1 = 1\Big|\mathcal{A}_{a_{2d+1}} = \mathcal{A}_{b_1}\Big]$$
$$= \begin{cases} \frac{2k_{2d}\cdot(n_{2d}-k_{2d})}{n_{2d}\cdot(n_{2d}-1)} \cdot \frac{n_{2d}\cdot(n_{2d}-1)}{4k_{2d}\cdot(n_{2d}-k_{2d})} & \text{if } \left|\frac{n_{2d}}{2} - k_{2d}\right| < \frac{\sqrt{n_{2d}}}{2}, \\ \text{\textcolor{red}{\ast}} & \text{if } \left|\frac{n_{2d}}{2} - k_{2d}\right| > \frac{\sqrt{n_{2d}}}{2}, \qquad (15) \\ \frac{2\left(\frac{n_{2d}}{2} - \frac{\sqrt{n_{2d}}}{2}\right)\left(\frac{n_{2d}}{2} + \frac{\sqrt{n_{2d}}}{2}\right)}{n_{2d}\cdot(n_{2d}-1)} & \text{if } \left|\frac{n_{2d}}{2} - k_{2d}\right| = \frac{\sqrt{n_{2d}}}{2} \end{cases}$$
$$= \frac{1}{2},$$

where $\textcolor{red}{\ast}$ stands for
$$\frac{2k_{2d}\cdot(n_{2d}-k_{2d})}{n_{2d}\cdot(n_{2d}-1)}$$
$$+ \left(1 - \frac{2k_{2d}\cdot(n_{2d}-k_{2d})}{n_{2d}\cdot(n_{2d}-1)}\right) \cdot \left(1 - \frac{n_{2d}\cdot(n_{2d}-1)}{2n_{2d}\cdot(n_{2d}-1) - 4k_{2d}\cdot(n_{2d}-k_{2d})}\right)$$
$$= 1 - \frac{n_{2d}\cdot(n_{2d}-1)}{2n_{2d}\cdot(n_{2d}-1) - 4k_{2d}\cdot(n_{2d}-k_{2d})} + \frac{2k_{2d}\cdot(n_{2d}-k_{2d})}{2n_{2d}\cdot(n_{2d}-1) - 4k_{2d}\cdot(n_{2d}-k_{2d})} = \frac{1}{2}.$$

Let us convert the content of $\mathcal{Z}_1$ into a distribution of $\widetilde{\mathcal{A}}_{a_{2d+1}}$ and $\widetilde{\mathcal{A}}_{b_1}$:
$$[\mathcal{Z}_1 = 0] \implies [\widetilde{\mathcal{A}}_{a_{2d+1},b_1} \equiv (0,0)];$$
$$\boldsymbol{Pr}\Big[\widetilde{\mathcal{A}}_{a_{2d+1},b_1} = (0,1)\Big|\mathcal{Z}_1 = 1\Big] = \boldsymbol{Pr}\Big[\widetilde{\mathcal{A}}_{a_{2d+1},b_1} = (1,0)\Big|\mathcal{Z}_1 = 1\Big] = \frac{1}{2}.$$

Now it is easy to see that the distribution of $\left|\mathcal{A}_{a_{2d+1},b_1} \oplus \widetilde{\mathcal{A}}_{a_{2d+1},b_1} \oplus s_{a_{2d+1},b_1}\right|$, conditioned on $\left[\left|\mathcal{A}_{a_{2d+1},\ldots,a_{\frac{n}{2}+d},b_1,\ldots,b_{\frac{n}{2}-d}}\right| = k_{2d}\right]$, is the same as the distribution of $\mathcal{B}_1 + \mathcal{B}_2$ for $(\mathcal{B}_1, \mathcal{B}_2) \in \{0,1\}^2$. [20] By the symmetry of all the involved distributions with respect to the bit-positions a_{2d+1} and b_1, we only need to check that
$$\boldsymbol{Pr}\Big[\mathcal{A}_{a_{2d+1}} \oplus \widetilde{\mathcal{A}}_{a_{2d+1}} = \mathcal{A}_{b_1} \oplus \widetilde{\mathcal{A}}_{b_1}\Big] = \boldsymbol{Pr}\Big[\mathcal{A}_{a_{2d+1}} \oplus \widetilde{\mathcal{A}}_{a_{2d+1}} \neq \mathcal{A}_{b_1} \oplus \widetilde{\mathcal{A}}_{b_1}\Big],$$
and this follows trivially from (15) and the symmetry of $\widetilde{\mathcal{A}}$'th distribution with respect to those two bit-positions. In other words,
$$\left|\mathcal{A}_{a_{2d+1},b_1} \oplus \widetilde{\mathcal{A}}_{a_{2d+1},b_1} \oplus s_{a_{2d+1},b_1}\right| = \begin{cases} 0 & \text{with probability } 1/4; \\ 1 & \text{with probability } 1/2; \qquad (16) \\ 2 & \text{with probability } 1/4. \end{cases}$$

---

[20] Note that the distribution of $\left(\mathcal{A}_{a_{2d+1},b_1} \oplus \widetilde{\mathcal{A}}_{a_{2d+1},b_1}\right)$ can be non-uniform, and this can be intuitively viewed as the *resource* that allows savings in terms of $\left|\widetilde{\mathcal{A}}\right|$ in Stage 2 in comparison to Stage 1. Now we do not have to worry if the probabilities of $\left[\mathcal{A}_{a_{2d+1},b_1} \oplus \widetilde{\mathcal{A}}_{a_{2d+1},b_1} = (0,0)\right]$ and $\left[\mathcal{A}_{a_{2d+1},b_1} \oplus \widetilde{\mathcal{A}}_{a_{2d+1},b_1} = (1,1)\right]$ are different from each other, which indeed happens in the above construction if $\left|\frac{n_{2d}}{2} - k_{2d}\right|$ is larger than $\frac{\sqrt{n_{2d}}}{2}$.

The distribution of $\mathcal{Z}_i$ for each $i \in [n/2-d]$ will depend, naturally, on $\left|\mathcal{A}_{a_{2d+1},\ldots,a_{\frac{n}{2}+d},b_1,\ldots,b_{\frac{n}{2}-d}}\right| = k_{2d}$, as well as on $\left|\mathcal{A}_{a_{2d+1},\ldots,a_{2d+i-1},b_1,\ldots,b_{i-1}}\right|$: namely, on

$$k_{2d+2i-2} \overset{\text{def}}{=} k_{2d} - \sum_{j=1}^{i-1}\left(\mathcal{A}_{a_{2d+j}} + \mathcal{A}_{b_j}\right) \quad \left(= \left|\mathcal{A}_{a_{2d+i},\ldots,a_{\frac{n}{2}+d},b_i,\ldots,b_{\frac{n}{2}-d}}\right|\right).$$

Additionally, it will depend on $\mathcal{A}_{a_{2d+i}} \oplus \mathcal{A}_{b_i}$.

Note that

$$\forall i \in [n/2 - d],\, j \in \left\{a_{2d+i},\ldots,a_{\frac{n}{2}+d}, b_i,\ldots,b_{\frac{n}{2}-d}\right\}:$$

$$\boldsymbol{Pr}\left[\mathcal{A}_j = 1 \middle| \left|\mathcal{A}_{a_{2d+i},\ldots,a_{\frac{n}{2}+d},b_i,\ldots,b_{\frac{n}{2}-d}}\right| = k_{2d+2i-2}\right] = \frac{k_{2d+2i-2}}{n_{2d+2i-2}}$$

and – as $\mathcal{A}_{a_{2d+i}}$ and $\mathcal{A}_{b_i}$ are not, in general, independent –

$$\boldsymbol{Pr}\left[\mathcal{A}_{a_{2d+i}} \neq \mathcal{A}_{b_i} \middle| \left|\mathcal{A}_{a_{2d+i},\ldots,a_{\frac{n}{2}+d},b_i,\ldots,b_{\frac{n}{2}-d}}\right| = k_{2d+2i-2}\right] = \frac{2k_{2d+2i-2} \cdot (n_{2d+2i-2} - k_{2d+2i-2})}{n_{2d+2i-2} \cdot (n_{2d+2i-2} - 1)}.$$

Generalising the case of $\mathcal{Z}_1$ (14), define for $\ast \overset{\text{def}}{=} 2d + 2i - 2$:

$$\forall i \in [n/2 - d]:$$

$$\boldsymbol{Pr}[\mathcal{Z}_i = 1] \overset{\text{def}}{=} \begin{cases} 1 - \frac{n_\ast \cdot (n_\ast - 1)}{4k_\ast \cdot (n_\ast - k_\ast)} & \text{if } \left|\frac{n_\ast}{2} - k_\ast\right| < \frac{\sqrt{n_\ast}}{2} \text{ and } \mathcal{A}_{a_{2d+i}} \neq \mathcal{A}_{b_i}; \\ 1 - \frac{n_\ast \cdot (n_\ast - 1)}{2n_\ast \cdot (n_\ast - 1) - 4k_\ast \cdot (n_\ast - k_\ast)} & \text{if } \left|\frac{n_\ast}{2} - k_\ast\right| > \frac{\sqrt{n_\ast}}{2} \text{ and } \mathcal{A}_{a_{2d+i}} = \mathcal{A}_{b_i}; \\ 0 & \text{otherwise.} \end{cases}$$

$$(17)$$

Then, similarly to (15), it holds that

$$\boldsymbol{Pr}\left[\mathcal{A}_{a_{2d+i}} \oplus \mathcal{A}_{b_i} \oplus \mathcal{Z}_i = 1 \middle| \mathcal{A}_{a_1,\ldots,a_{2d+i-1},b_1,\ldots,b_{i-1}}, |\mathcal{A}|\right] \equiv \frac{1}{2}. \qquad (18)$$

Define the distribution of $\widetilde{\mathcal{A}}_{a_{2d+i}}$ and $\widetilde{\mathcal{A}}_{b_i}$:

$$[\mathcal{Z}_i = 0] \implies [\widetilde{\mathcal{A}}_{a_{2d+i},b_i} \equiv (0,0)];$$

$$\boldsymbol{Pr}\left[\widetilde{\mathcal{A}}_{a_{2d+i},b_i} = (0,1) \middle| \mathcal{Z}_i = 1\right] = \boldsymbol{Pr}\left[\widetilde{\mathcal{A}}_{a_{2d+i},b_i} = (1,0) \middle| \mathcal{Z}_i = 1\right] = \frac{1}{2}.$$

Similarly to (16), conditional on $\mathcal{A}_{a_1,\ldots,a_{2d+i-1},b_1,\ldots,b_{i-1}}$ and $|\mathcal{A}|$ it holds that

$$\left|\left(\mathcal{A} \oplus \widetilde{\mathcal{A}} \oplus s\right)_{a_{2d+i},b_i}\right| = \begin{cases} 0 & \text{with probability } 1/4; \\ 1 & \text{with probability } 1/2; \\ 2 & \text{with probability } 1/4. \end{cases} \qquad (19)$$

From now and until the end of Stage 2, all the distributions are assumed to be conditioned on

$$\left[\left|\mathcal{A}_{a_{2d+1},\ldots,a_{\frac{n}{2}+d},b_1,\ldots,b_{\frac{n}{2}-d}}\right| = k_{2d}\right] \qquad (20)$$

for some $k_{2d} \in [n - 2d] \cup \{0\}$, unless stated otherwise.

Let us argue that the distribution of

$$\left|\left(\mathcal{A}\oplus\widetilde{\mathcal{A}}\oplus s\right)_{a_{2d+1},\dots,a_{\frac{n}{2}+d},b_1,\dots,b_{\frac{n}{2}-d}}\right|$$

equals the distribution of $|\mathcal{B}|$ for $\mathcal{B}\sim\mathcal{U}_{\{0,1\}^{n-2d}}$. We prove it by induction.

Let $i\in[n/2-d]$, we claim that the distribution of $\left|\left(\mathcal{A}\oplus\widetilde{\mathcal{A}}\oplus s\right)_{a_{2d+1},\dots,a_{2d+i},b_1,\dots,b_i}\right|$ equals that of $\left|\mathcal{B}_{[2i]}\right|$. For $i=1$ the statement follows readily from (19), so assume that $i>1$ and the distribution of $\left|\left(\mathcal{A}\oplus\widetilde{\mathcal{A}}\oplus s\right)_{a_{2d+1},\dots,a_{2d+i-1},b_1,\dots,b_{i-1}}\right|$ equals that of $\left|\mathcal{B}_{[2i-2]}\right|$.

By the argument, fully analogous to that behind (11), it holds that

$$\mathbf{I}\left[\left(\mathcal{A}\oplus\widetilde{\mathcal{A}}\oplus s\right)_{a_{2d+1},\dots,a_{2d+i-1},b_1,\dots,b_{i-1}}:\left(\mathcal{A}\oplus\widetilde{\mathcal{A}}\oplus s\right)_{a_{2d+i},b_i}\middle|\mathcal{A}_{a_{2d+1},\dots,a_{2d+i-1},b_1,\dots,b_{i-1}}\right]=0.$$

Then (19) implies that for any $w\in\{0,1\}^n$, the distribution of $\left(\mathcal{A}\oplus\widetilde{\mathcal{A}}\oplus s\right)_{a_{2d+i},b_i}$ conditioned on $\left|\left(\mathcal{A}\oplus\widetilde{\mathcal{A}}\oplus s\right)_{a_{2d+1},\dots,a_{2d+i-1},b_1,\dots,b_{i-1}}\right|$ equals that of $\left|\mathcal{B}_{[2]}\right|$, and therefore the distribution of

$$\left|\left(\mathcal{A}\oplus\widetilde{\mathcal{A}}\oplus s\right)_{a_{2d+1},\dots,a_{2d+i},b_1,\dots,b_i}\right|$$

$$=\left|\left(\mathcal{A}\oplus\widetilde{\mathcal{A}}\oplus s\right)_{a_{2d+1},\dots,a_{2d+i-1},b_1,\dots,b_{i-1}}\right|+\left|\left(\mathcal{A}\oplus\widetilde{\mathcal{A}}\oplus s\right)_{a_{2d+i},b_i}\right|$$

equals that of $\left|\mathcal{B}_{[2i]}\right|$, as required for the induction step.

Completion. Let us no longer implicitly assume (20). We have just seen that the distribution of

$$\left|\left(\mathcal{A}\oplus\widetilde{\mathcal{A}}\oplus s\right)_{a_{2d+1},\dots,a_{\frac{n}{2}+d},b_1,\dots,b_{\frac{n}{2}-d}}\right|,$$

conditioned on

$$\left|\mathcal{A}_{a_{2d+1},\dots,a_{\frac{n}{2}+d},b_1,\dots,b_{\frac{n}{2}-d}}\right|=|\mathcal{A}|-\left|\mathcal{A}_{a_1,\dots,a_{2d}}\right|,$$

equals that of $|\mathcal{B}|$ for $\mathcal{B}\sim\mathcal{U}_{\{0,1\}^{n-2d}}$. It is clear from the construction and the symmetry of the considered distributions of $\mathcal{A}$, that additional conditioning on the content of $\mathcal{A}_{a_1,\dots,a_{2d}}$ doesn't affect the above statement.

From (12):

$$H\left(\left(\mathcal{A}\oplus\widetilde{\mathcal{A}}\right)_{a_1,\dots,a_{2d}}\middle||\mathcal{A}|\right)=2d\implies H\left(\left(\mathcal{A}\oplus\widetilde{\mathcal{A}}\oplus s\right)_{a_1,\dots,a_{2d}}\middle||\mathcal{A}|\right)=2d,$$

and therefore the distribution of

$$\left|\left(\mathcal{A}\oplus\widetilde{\mathcal{A}}\oplus s\right)_{a_1,\dots,a_{2d}}\right|,$$

conditioned on $|\mathcal{A}|$, equals that of $|\mathcal{B}|$ for $\mathcal{B}\sim\mathcal{U}_{\{0,1\}^{2d}}$.

By the argument, fully analogous to that behind (11), it holds that

$$\boldsymbol{I}\left[\left(\mathcal{A}\oplus\widetilde{\mathcal{A}}\oplus s\right)_{a_1,\dots,a_{2d}}:\left(\mathcal{A}\oplus\widetilde{\mathcal{A}}\oplus s\right)_{a_{2d+1},\dots,a_{\frac{n}{2}+d},b_1,\dots,b_{\frac{n}{2}-d}}\middle|\,\mathcal{A}_{a_1,\dots,a_{2d}},\,|\mathcal{A}|\right]=0.$$

Therefore, conditioned on $[|\mathcal{A}|=k_0]$,

$$\left|\mathcal{A}\oplus\widetilde{\mathcal{A}}\oplus s\right|=\left|\mathcal{A}_{a_{2d+1},\dots,a_{\frac{n}{2}+d},b_1,\dots,b_{\frac{n}{2}-d}}\right|+\left|\left(\mathcal{A}\oplus\widetilde{\mathcal{A}}\oplus s\right)_{a_1,\dots,a_{2d}}\right|$$

is distributed like $|\mathcal{B}|$ for $\mathcal{B}\sim\mathcal{U}_{\{0,1\}^n}$ – in particular, the fact that the above distribution doesn't depend on k_0 implies that

$$\boldsymbol{I}\left[\left|\mathcal{A}\oplus\widetilde{\mathcal{A}}\oplus s\right|:|\mathcal{A}|\right]=0,$$

as required.

It remains to argue that $\left|\widetilde{\mathcal{A}}\right|$ is likely to be small. Towards estimating $\left|\widetilde{\mathcal{A}}_{a_1,\dots,a_{2d}}\right|$ we will use (13), now let us analyse $\left|\widetilde{\mathcal{A}}_{a_{2d+1},\dots,a_{\frac{n}{2}+d},b_1,\dots,b_{\frac{n}{2}-d}}\right|$, as constructed during Stage 2.

Let $\forall j\in[n]-1:\ l_j\stackrel{\text{def}}{=}|k_j-{n_j}/{2}|$ and $\forall i\in[{n}/{2}-d]:\ *\stackrel{\text{def}}{=}2d+2i-2$, then it follows from (17) that

$$\text{if }l_*=\frac{\sqrt{n_*}}{2},\text{ then }\boldsymbol{Pr}[\mathcal{Z}_i=1]=0;$$

$$\text{if }l_*<\frac{\sqrt{n_*}}{2},\text{ then }\boldsymbol{Pr}[\mathcal{Z}_i=1]\le 1-\frac{n_*\cdot(n_*-1)}{4k_*\cdot(n_*-k_*)}=1-\frac{n_*^2-n_*}{4\left(\frac{n_*}{2}-l_*\right)\left(\frac{n_*}{2}+l_*\right)}$$

$$=\frac{n_*-4l_*^2}{n_*^2-4l_*^2}\le\frac{1}{n_*};$$

$$\text{if }l_*>\frac{\sqrt{n_*}}{2},\text{ then }\boldsymbol{Pr}[\mathcal{Z}_i=1]\le 1-\frac{n_*^2-n_*}{n_*^2-2n_*+4l_*^2}$$

$$=\frac{4l_*^2-n_*}{n_*^2-2n_*+4l_*^2}\le\frac{4l_*^2}{n_*^2-n_*+4l_*^2}<\frac{4l_*^2}{n_*^2}.$$

Therefore, $\forall i\in[{n}/{2}-d]:$

$$\boldsymbol{Pr}[\mathcal{Z}_i=1]\le\frac{1}{n_{2d+2i-2}}+\frac{4l_{2d+2i-2}^2}{n_{2d+2i-2}^2}=\frac{1}{n_{2d+2i-2}}+\left(\frac{2k_{2d+2i-2}}{n_{2d+2i-2}}-1\right)^2.$$

By construction, subject to the assignment in (17), every $\mathcal{Z}_i$ is independent from all other variables – in particular,

$$\forall i\in\left[\frac{n}{2}-d\right]:\ \boldsymbol{Pr}\left[\mathcal{Z}_i=1\middle|\mathcal{Z}_{[i-1]}\right]\le\frac{1}{n_{2d+2i-2}}+\left(\frac{2k_{2d+2i-2}}{n_{2d+2i-2}}-1\right)^2. \qquad (21)$$

For all $i\in[n]-1$, let K_i be a random variable that takes the value of k_i (in particular, $K_0=|\mathcal{A}|$). For $u\ge 0$, denote by $\boldsymbol{e}_u$ the event

$$\left[\forall i\in[n]-1:\ \left|K_i-|\mathcal{A}|\cdot\frac{n_i}{n}\right|\le u\cdot\sqrt{n_i}\right].$$

Then

$$\boldsymbol{Pr}[\neg\boldsymbol{e}_u] \leq \sum_{i=0}^{n-1} \boldsymbol{Pr}\left[\left|K_i - |\mathcal{A}| \cdot \frac{n_i}{n}\right| > u \cdot \sqrt{n_i}\right] \leq \sum_{i=0}^{n-1} 2 \cdot e^{\frac{-2u^2 n_i}{n_i}} = 2n \cdot e^{-2u^2}, \qquad (22)$$

where the second inequality is Hoeffding's (Fact 2).

Assuming $\boldsymbol{e}_u$ and letting $l \stackrel{\text{def}}{=} \left||\mathcal{A}| - {}^n\!/_2\right|$,

$$\forall i \in [n] - 1 : \quad \left|\frac{2k_i}{n_i} - 1\right| \leq \left|\frac{2 \cdot |\mathcal{A}|}{n} - 1\right| + \frac{u}{\sqrt{n_i}} = \frac{2l}{n} + \frac{u}{\sqrt{n_i}},$$

and therefore (21) gives

$$\forall i \in \left[\frac{n}{2} - d\right] : \quad \boldsymbol{Pr}\left[\mathcal{Z}_i = 1 \middle| \mathcal{Z}_{[i-1]}, \boldsymbol{e}_u\right] \leq \frac{1 + 2u^2}{n_{2d+2i-2}} + \frac{8l^2}{n^2}$$

and

$$\sum_{i \in \left[\frac{n}{2} - d\right]} \boldsymbol{Pr}\left[\mathcal{Z}_i = 1 \middle| \mathcal{Z}_{[i-1]}, \boldsymbol{e}_u\right] \leq \frac{4l^2}{n} + \left(1 + 2u^2\right) \cdot \sum_{i \in \left[\frac{n}{2} - d\right]} \frac{1}{n_{2d+2i-2}}$$

$$\leq \frac{4l^2}{n} + \left(1 + 2u^2\right) \cdot \ln n.$$

Combined with (22), this gives

$$\boldsymbol{Pr}\left[\sum_{i \in \left[\frac{n}{2} - d\right]} \boldsymbol{Pr}\left[\mathcal{Z}_i = 1 \middle| \mathcal{Z}_{[i-1]}\right] > \frac{4l^2}{n} + \left(1 + 2u^2\right) \cdot \ln n\right] \leq \boldsymbol{Pr}[\neg\boldsymbol{e}_u] \leq 2n \cdot e^{-2u^2}.$$

Applying the relaxed Chernoff bound (Theorem 1) with $a \stackrel{\text{def}}{=} \frac{4l^2}{n} + \left(1 + 2u^2\right) \cdot \ln n$ gives

$$\forall u, t, l \geq 0 :$$

$$\boldsymbol{Pr}\left[|\mathcal{Z}| \geq \frac{4l^2}{n} + \left(1 + 2u^2\right) \cdot \ln n + t \middle| \left||\mathcal{A}| - {}^n\!/_2\right| = l\right]$$

$$\leq e^{\frac{-t^2}{8l^2/n + (2+4u^2) \cdot \ln n + t}} + 2 \cdot e^{\ln n - 2u^2}.$$

As $\left|\widetilde{\mathcal{A}}_{a_{2d+1},\dots,a_{\frac{n}{2}+d},b_1,\dots,b_{\frac{n}{2}-d}}\right| \equiv |\mathcal{Z}|$ by construction, combined with (13), this gives

$$\forall u, t, l \geq 0 :$$

$$\boldsymbol{Pr}\left[\left|\widetilde{\mathcal{A}}\right| \geq \frac{4dl + 4d^2 + 4l^2}{n} + \left(1 + 2u^2\right) \cdot \ln n + 2t \middle| \left||\mathcal{A}| - {}^n\!/_2\right| = l\right]$$

$$\leq e^{\frac{-t^2 \cdot n}{8dl + 8d^2 + tn}} + e^{\frac{-t^2 \cdot n}{8l^2 + (2+4u^2) \cdot n \cdot \ln n + tn}} + 2 \cdot e^{\ln n - 2u^2}.$$

Let us denote $T \stackrel{\text{def}}{=} 4t\sqrt{\ln n}$ and choose $u^2 \stackrel{\text{def}}{=} \frac{t}{\sqrt{\ln n}} = \frac{T}{4 \ln n}$. Assume for now that $T \geq 4 \ln n$, then

$$\frac{4dl + 4d^2 + 4l^2}{n} + \left(1 + 2u^2\right) \cdot \ln n + 2t \leq \frac{(2d + 2l)^2}{n} + T,$$

24

and so,

$$\forall l \geq 0,\ T \geq 4\ln n:\ \boldsymbol{Pr}\left[\left|\widetilde{\mathcal{A}}\right| \geq \frac{(2d+2l)^2}{n} + T\ \middle|\ \left|\left|\mathcal{A}\right| - {}^n\!/_2\right| = l\right] \leq \textcolor{red}{\ast},$$

where

$$\textcolor{red}{\ast} = e^{\frac{-t^2 \cdot n}{8dl + 8d^2 + tn}} + e^{\frac{-t^2 \cdot n}{8l^2 + (2+4u^2)\cdot n \cdot \ln n + tn}} + 2 \cdot e^{\ln n - 2u^2}$$

$$\leq 2 \cdot e^{\frac{-T^2 n}{32 \cdot (2d+2l)^2 \cdot \ln n + 80Tn}} + 2 \cdot e^{\ln n - \frac{T}{2\ln n}}.$$

As we always assume that $n \geq 4$, the condition $[T \geq 4\ln n]$ can be removed, as otherwise $2 \cdot e^{\ln n - \frac{T}{2\ln n}} > 1$ and the inequality holds trivially. So,

$$\forall l,\ T \geq 0:$$

$$\boldsymbol{Pr}\left[\left|\widetilde{\mathcal{A}}\right| \geq \frac{(2d+2l)^2}{n} + T\ \middle|\ \left|\left|\mathcal{A}\right| - {}^n\!/_2\right| = l\right] \leq 2 \cdot \left(e^{\frac{-T^2 n}{32 \cdot (2d+2l)^2 \cdot \ln n + 80Tn}} + e^{\ln n - \frac{T}{2\ln n}}\right).$$

As a function of s and l, the above bound on $\boldsymbol{Pr}\left[\left|\widetilde{\mathcal{A}}\right| > s \middle| \left|\left|\mathcal{A}\right| - {}^n\!/_2\right| = l\right]$ is monotonically increasing in l, and therefore,

$$\boldsymbol{Pr}\left[\left|\widetilde{\mathcal{A}}\right| \geq \frac{(2d+2l)^2}{n} + T\ \middle|\ \left|\left|\mathcal{A}\right| - {}^n\!/_2\right| \leq l\right] \leq 2 \cdot \left(e^{\frac{-T^2 n}{32 \cdot (2d+2l)^2 \cdot \ln n + 80Tn}} + e^{\ln n - \frac{T}{2\ln n}}\right),$$

as required.

To complete the proof, observe that Hoeffding's inequality (Fact 2) implies that

$$\forall l \geq 0:\ \boldsymbol{Pr}\left[\left|\left|\mathcal{A}\right| - \frac{n}{2}\right| \geq l\right] \leq 2 \cdot e^{\frac{-2l^2}{n}}.$$

Letting $l \stackrel{\text{def}}{=} \max\left\{d, \frac{\sqrt{Tn}}{4}\right\} \implies (2d+2l)^2 \leq 16d^2 + Tn$, we obtain

$$\forall T \geq 0:\boldsymbol{Pr}\left[\left|\widetilde{\mathcal{A}}\right| \geq \frac{16d^2}{n} + 2T\right]$$

$$\leq \boldsymbol{Pr}\left[\left|\widetilde{\mathcal{A}}\right| \geq \frac{(2d+2l)^2}{n} + T\ \middle|\ \left|\left|\mathcal{A}\right| - \frac{n}{2}\right| \leq l\right] + \boldsymbol{Pr}\left[\left|\left|\mathcal{A}\right| - \frac{n}{2}\right| \geq l\right]$$

$$\leq 2 \cdot e^{\frac{-T}{2}} + 2 \cdot e^{\ln n - \frac{T}{2\ln n}} + 2 \cdot e^{\frac{-T^2 n}{32 \cdot (16d^2 + Tn) \cdot \ln n + 80Tn}}$$

$$\leq 4 \cdot e^{\ln n - \frac{T}{2\ln n}} + 2 \cdot e^{\frac{-T^2 n}{112 \cdot \ln n \cdot T \cdot n + 512 \cdot d^2 \cdot \ln n}},$$

and the result follows by substituting $t \stackrel{\text{def}}{=} 2T$ and noticing that increasing d leads to shrinkage of the event whose probability is bounded and, at the same time, growing of the claimed upper bound (therefore the assumption $\left[\left|\left|s\right| - {}^n\!/_2\right| = d\right]$ may be replaced by $\left[\left|\left|s\right| - {}^n\!/_2\right| \leq d\right]$).

$\blacksquare$ *Lemma 5*

We are ready to investigate the likelihood of $\aleph(\mathcal{X}, \mathcal{Y})$.

Proof of Theorem 2. Recall the definition of $\Delta_{j,s}$, as given in (1), and let us define "weighted indicators":

$$w : [n] \cup \{0\} \to [0, 1/4]; \ w(x) \overset{\text{def}}{=} \begin{cases} \frac{1}{n} \cdot \left(x - \frac{n}{2}\right)^2 & \text{if } \frac{n}{2} - \frac{\sqrt{n}}{2} \leq x \leq \frac{n}{2} + \frac{\sqrt{n}}{2}; \\ 0 & \text{otherwise.} \end{cases}$$

$$\forall j \in [n], \ s \in \{0,1\}^{\log n} : \ \mathcal{W}_{j,s} \overset{\text{def}}{=} w\left(\Delta_{j,s}(\mathcal{X}, \mathcal{Y})\right).$$

Then we want to show that

$$\boldsymbol{Pr}\left[\sum_{j \in [n], \, s \in \{0,1\}^{\log n}} \mathcal{W}_{j,s} > \frac{n^2}{9}\right] \in O\left(\frac{(\ln n)^2}{\sqrt{n}}\right). \tag{23}$$

Towards understanding the behaviour of the sum of $\mathcal{W}_{j,s}$-s we analyse the dependencies among them. Let $(j_1, s_1) \neq (j_2, s_2)$ and define random variables:

$$\text{for } i \in \{1, 2\} : \ \mathcal{D}_i \overset{\text{def}}{=} \Delta_{j_i, s_i}(\mathcal{X}, \mathcal{Y});$$

$$\mathcal{Z} \overset{\text{def}}{=} \sigma_{j_1}\left(\tau_{s_1} \oplus \mathcal{X}\right) \oplus \sigma_{j_2}\left(\tau_{s_2} \oplus \mathcal{X}\right);$$

$$\mathcal{V} \overset{\text{def}}{=} \sigma_{j_1}\left(\tau_{s_1} \oplus \mathcal{X}\right) \oplus \mathcal{Y}.$$

Note that

$$\mathcal{D}_1 \equiv |\mathcal{V}|, \ \mathcal{D}_2 \equiv |\mathcal{Z} \oplus \mathcal{V}|, \ \mathcal{W}_{j_1, s_1} \equiv w(\mathcal{D}_1) \ \text{and} \ \mathcal{W}_{j_2, s_2} \equiv w(\mathcal{D}_2).$$

As $(\mathcal{X}, \mathcal{Y}) \sim \mathcal{U}_{\{0,1\}^{n+n}}$,

$$\mathcal{V} \sim \mathcal{U}_{\{0,1\}^n} \ \text{and} \ \boldsymbol{I}\left[\mathcal{V} : \mathcal{Z}\right] = 0.$$

Apply Lemma 5 to define a random variable $\widetilde{\mathcal{V}}$, such that

$$\boldsymbol{I}\left[|\mathcal{V}| : \left|\mathcal{Z} \oplus \mathcal{V} \oplus \widetilde{\mathcal{V}}\right| \, \middle|\, \mathcal{Z}\right] = 0$$

and

$$\forall c, d \geq 0 : \boldsymbol{Pr}\left[\left|\widetilde{\mathcal{V}}\right| \geq \frac{16d^2}{n} + c \cdot (\ln n)^2 \, \middle|\, \left|\, |\mathcal{Z}| - \frac{n}{2}\right| \leq d\right] \tag{24}$$

$$\leq 4 \cdot e^{(1 - c/4) \cdot \ln n} + 2 \cdot e^{\frac{-c^2 \cdot n \cdot (\ln n)^3}{224 \cdot c \cdot n \cdot (\ln n)^2 + 2048 \cdot d^2}}.$$

Then

$$\boldsymbol{I}\left[|\mathcal{V}| : \left|\mathcal{Z} \oplus \mathcal{V} \oplus \widetilde{\mathcal{V}}\right|\right] \leq \boldsymbol{I}\left[|\mathcal{V}| : \mathcal{Z}, \left|\mathcal{Z} \oplus \mathcal{V} \oplus \widetilde{\mathcal{V}}\right|\right]$$

$$= \boldsymbol{I}\left[|\mathcal{V}| : \mathcal{Z}\right] + \boldsymbol{I}\left[|\mathcal{V}| : \left|\mathcal{Z} \oplus \mathcal{V} \oplus \widetilde{\mathcal{V}}\right| \, \middle|\, \mathcal{Z}\right] = 0.$$

As $\mathcal{D}_2 = |\mathcal{Z} \oplus \mathcal{V}|$,

$$\mathcal{D}_2 \in \left[\left|\left|\mathcal{Z} \oplus \mathcal{V} \oplus \widetilde{\mathcal{V}}\right| - \left|\widetilde{\mathcal{V}}\right|\right|, \left|\mathcal{Z} \oplus \mathcal{V} \oplus \widetilde{\mathcal{V}}\right| + \left|\widetilde{\mathcal{V}}\right|\right].$$

Define
$$\widetilde{\mathcal{D}}_2 \overset{\text{def}}{=} \left| \mathcal{Z} \oplus \mathcal{V} \oplus \widetilde{\mathcal{V}} \right| - \mathcal{D}_2,$$

then
$$\boldsymbol{I}\!\left[\mathcal{D}_1 : \mathcal{D}_2 + \widetilde{\mathcal{D}}_2\right] = \boldsymbol{I}\!\left[\left|\mathcal{V}\right| : \left|\mathcal{Z} \oplus \mathcal{V} \oplus \widetilde{\mathcal{V}}\right|\right] = 0$$

and

$$
\begin{aligned}
\boldsymbol{Cov}\!\left[\mathcal{W}_{j_1,s_1}, \mathcal{W}_{j_2,s_2}\right] &= \boldsymbol{Cov}\!\left[w(\mathcal{D}_1),\, w(\mathcal{D}_2 + \widetilde{\mathcal{D}}_2) + \left(w(\mathcal{D}_2) - w(\mathcal{D}_2 + \widetilde{\mathcal{D}}_2)\right)\right] \\
&= \boldsymbol{Cov}\!\left[w(\mathcal{D}_1),\, w(\mathcal{D}_2 + \widetilde{\mathcal{D}}_2)\right] + \boldsymbol{Cov}\!\left[w(\mathcal{D}_1),\, w(\mathcal{D}_2) - w(\mathcal{D}_2 + \widetilde{\mathcal{D}}_2)\right] \\
&= \boldsymbol{Cov}\!\left[w(\mathcal{D}_1),\, w(\mathcal{D}_2) - w(\mathcal{D}_2 + \widetilde{\mathcal{D}}_2)\right] \\
&\leq \frac{1}{4} \cdot \boldsymbol{E}\!\left[\left| w(\mathcal{D}_2) - w(\mathcal{D}_2 + \widetilde{\mathcal{D}}_2) - \boldsymbol{E}\!\left[w(\mathcal{D}_2) - w(\mathcal{D}_2 + \widetilde{\mathcal{D}}_2)\right]\right|\right] \\
&\leq \frac{1}{2} \cdot \boldsymbol{E}\!\left[\left| w(\mathcal{D}_2) - w(\mathcal{D}_2 + \widetilde{\mathcal{D}}_2)\right|\right].
\end{aligned}
\tag{25}
$$

Note that $\left|\widetilde{\mathcal{D}}_2\right| \leq \left|\widetilde{\mathcal{V}}\right|$ by construction (notice the difference in the meanings of $|\ldots|$), therefore $\left|\widetilde{\mathcal{D}}_2\right|$ is also subject to (24). Let us make the bound more comfortable by analysing the behaviour of $\left|\,|\mathcal{Z}| - \frac{n}{2}\right|$. If $j_1 = j_2$, then $s_1 \neq s_2$ and we have
$$\mathcal{Z} = \sigma_{j_1}\!\left(\tau_{s_1} \oplus \mathcal{X}\right) \oplus \sigma_{j_1}\!\left(\tau_{s_2} \oplus \mathcal{X}\right) = \sigma_{j_1}\!\left(\tau_{s_1} \oplus \tau_{s_2}\right),$$

which implies $|\mathcal{Z}| \equiv \frac{n}{2}$ by the construction of τ_s-s. If, on the other hand, $j_1 \neq j_2$, then
$$\mathcal{Z} = \sigma_{j_1}\!\left(\tau_{s_1}\right) \oplus \sigma_{j_2}\!\left(\tau_{s_2}\right) \oplus \sigma_{j_1}(\mathcal{X}) \oplus \sigma_{j_2}(\mathcal{X}) = \sigma_{j_1}\!\left(\tau_{s_1}\right) \oplus \sigma_{j_2}\!\left(\tau_{s_2}\right) \oplus \mathcal{X}' \oplus \sigma_{j_2 - j_1}(\mathcal{X}')$$

for $\mathcal{X}' \overset{\text{def}}{=} \sigma_{j_1}(\mathcal{X}) \sim \mathcal{U}_{\{0,1\}^n}$, and therefore,
$$\boldsymbol{Pr}\!\left[\left|\,|\mathcal{Z}| - \frac{n}{2}\right| \geq \sqrt{n} \cdot \ln n\right] \leq 4 \cdot e^{\frac{-(\ln n)^2}{2}}$$

by Lemma 3. Combining this with (24), we obtain that $\forall c \geq 0$:

$$\boldsymbol{Pr}\!\left[\left|\widetilde{\mathcal{D}}_2\right| \geq (16 + c) \cdot (\ln n)^2\right]$$
$$\leq \boldsymbol{Pr}\!\left[\left|\widetilde{\mathcal{V}}\right| \geq (16 + c) \cdot (\ln n)^2 \,\middle|\, \left|\,|\mathcal{Z}| - \frac{n}{2}\right| \leq \sqrt{n} \cdot \ln n\right] + \boldsymbol{Pr}\!\left[\left|\,|\mathcal{Z}| - \frac{n}{2}\right| > \sqrt{n} \cdot \ln n\right]$$
$$\leq 4 \cdot e^{(1 - c/4) \cdot \ln n} + 2 \cdot e^{\frac{-c^2 \cdot \ln n}{224 \cdot c + 2048}} + 4 \cdot e^{\frac{-(\ln n)^2}{2}},$$

and choosing $c \overset{\text{def}}{=} 121$ gives
$$\boldsymbol{Pr}\!\left[\left|\widetilde{\mathcal{D}}_2\right| \geq 137 \cdot (\ln n)^2\right] < \frac{10}{\sqrt{n}}.\tag{26}$$

We are ready to analyse $\boldsymbol{E}\left[\left|w(\mathcal{D}_2) - w(\mathcal{D}_2 + \widetilde{\mathcal{D}}_2)\right|\right]$. Let $\xi \overset{\text{def}}{=} 137 \cdot (\ln n)^2$ and

$$I_\xi \overset{\text{def}}{=} \left[\frac{n - \sqrt{n}}{2} - \xi, \; \frac{n - \sqrt{n}}{2} + \xi\right] \cup \left[\frac{n + \sqrt{n}}{2} - \xi, \; \frac{n + \sqrt{n}}{2} + \xi\right],$$

$$J_\xi \overset{\text{def}}{=} \left[\frac{n - \sqrt{n}}{2} + \xi, \; \frac{n + \sqrt{n}}{2} - \xi\right].$$

Then by the construction of $w(\dots)$:

$$\boldsymbol{E}\left[\left|w(\mathcal{D}_2) - w(\mathcal{D}_2 + \widetilde{\mathcal{D}}_2)\right|\right]$$

$$\leq \left(\boldsymbol{Pr}\left[\left|\widetilde{\mathcal{D}}_2\right| > \xi\right] + \boldsymbol{Pr}\left[\mathcal{D}_2 \in I_\xi\right]\right) \cdot \frac{1}{4}$$

$$+ \boldsymbol{E}\left[w(\mathcal{D}_2) - w(\mathcal{D}_2 + \widetilde{\mathcal{D}}_2)\middle|\left|\widetilde{\mathcal{D}}_2\right| \leq \xi, \; \mathcal{D}_2 \in J_\xi\right]$$

$$< \frac{3}{\sqrt{n}} + \frac{\boldsymbol{Pr}\left[\mathcal{D}_2 \in I_\xi\right]}{4} + \boldsymbol{E}\left[w(\mathcal{D}_2) - w(\mathcal{D}_2 + \widetilde{\mathcal{D}}_2)\middle|\left|\widetilde{\mathcal{D}}_2\right| \leq \xi, \; \mathcal{D}_2 \in J_\xi\right],$$

where the latter inequality is (26).

$$\boldsymbol{Pr}\left[\mathcal{D}_2 \in I_\xi\right] \leq \boldsymbol{Pr}\left[\mathcal{D}_2 \in I_\xi\middle|\mathcal{D}_2 \in \left[\frac{n - \sqrt{n}}{2} - \xi, \; \frac{n + \sqrt{n}}{2} + \xi\right]\right]$$

$$\leq \frac{|I_\xi|}{\left|\left[\frac{n - \sqrt{n}}{2} - \xi, \; \frac{n + \sqrt{n}}{2} + \xi\right]\right|} \quad < \quad \frac{4\xi}{\sqrt{n}},$$

where the second inequality is due to the fact that $\mathcal{D}_2$ comes from the symmetric binomial distribution centred (and therefore maximised) at $n/2$ and I_ξ is a union of two symmetric edges of the considered interval $\left[\frac{n - \sqrt{n}}{2} - \xi, \; \frac{n + \sqrt{n}}{2} + \xi\right]$.

$$\boldsymbol{E}\left[w(\mathcal{D}_2) - w(\mathcal{D}_2 + \widetilde{\mathcal{D}}_2)\middle|\left|\widetilde{\mathcal{D}}_2\right| \leq \xi, \; \mathcal{D}_2 \in J_\xi\right]$$

$$\leq \sup_{x_0 \in \left[\frac{n - \sqrt{n}}{2}, \; \frac{n + \sqrt{n}}{2}\right]} \left\{\left|\frac{dw(x)}{dx}[x_0]\right|\right\} \cdot \xi \quad = \quad \frac{\xi}{\sqrt{n}}.$$

Putting this together gives

$$\boldsymbol{E}\left[\left|w(\mathcal{D}_2) - w(\mathcal{D}_2 + \widetilde{\mathcal{D}}_2)\right|\right] < \frac{3\xi}{\sqrt{n}} = \frac{411 \cdot (\ln n)^2}{\sqrt{n}}$$

and then from (25):

$$\forall (j_1, s_1) \neq (j_2, s_2) : \boldsymbol{Cov}\left[\mathcal{W}_{j_1, s_1}, \mathcal{W}_{j_2, s_2}\right] \leq \frac{206 \cdot (\ln n)^2}{\sqrt{n}}. \tag{27}$$

Now – still towards understanding the behaviour of the sum of $\mathcal{W}_{j,s}$-s – let us estimate their expectations.

$$\forall (j,s): \ \mathbf{E}\big[\mathcal{W}_{j,s}\big] = \mathbf{Pr}\left[\Delta_{j,s}(\mathcal{X},\mathcal{Y}) \in \left[\frac{n}{2} - \frac{\sqrt{n}}{2}, \frac{n}{2} + \frac{\sqrt{n}}{2}\right]\right]$$

$$\cdot \mathbf{E}\left[\frac{1}{n} \cdot \left(\Delta_{j,s}(\mathcal{X},\mathcal{Y}) - \frac{n}{2}\right)^2 \Bigg| \Delta_{j,s}(\mathcal{X},\mathcal{Y}) \in \left[\frac{n}{2} - \frac{\sqrt{n}}{2}, \frac{n}{2} + \frac{\sqrt{n}}{2}\right]\right]$$

$$\leq \frac{1}{n} \cdot \mathbf{E}\left[\left(\Delta_{j,s}(\mathcal{X},\mathcal{Y}) - \frac{n}{2}\right)^2 \Bigg| \Delta_{j,s}(\mathcal{X},\mathcal{Y}) \in \left[\frac{n}{2} - \frac{\sqrt{n}}{2}, \frac{n}{2} + \frac{\sqrt{n}}{2}\right]\right],$$

where $\Delta_{j,s}(\mathcal{X},\mathcal{Y}) = |\sigma_j(\tau_s \oplus \mathcal{X}) \oplus \mathcal{Y}|$ comes from the symmetric binomial distribution $B(n,1/2)$. Let $S \equiv \Delta_{j,s}(\mathcal{X},\mathcal{Y})$, then

$$n \cdot \mathbf{E}\big[\mathcal{W}_{j,s}\big] \leq \mathop{\mathbf{E}}_{S \sim B(n,1/2)}\left[\left(S - \frac{n}{2}\right)^2 \Bigg| S \in \left[\frac{n}{2} - \frac{\sqrt{n}}{2}, \frac{n}{2} + \frac{\sqrt{n}}{2}\right]\right].$$

As for any $a_1, a_2 \in \{n/2 - \sqrt{n}/2, \ldots, n/2 + \sqrt{n}/2\}$ it, obviously, holds that

$$\mathbf{Pr}[S = a_1] \leq \mathbf{Pr}[S = a_2] \iff (a_1 - n/2)^2 \geq (a_2 - n/2)^2,$$

Lemma 4 implies that

$$\mathop{\mathbf{E}}_{S \sim B(n,1/2)}\left[\left(S - \frac{n}{2}\right)^2 \Bigg| S \in \left[\frac{n}{2} - \frac{\sqrt{n}}{2}, \frac{n}{2} + \frac{\sqrt{n}}{2}\right]\right] \leq \mathop{\mathbf{E}}_{S \in \left\{\frac{n}{2} - \frac{\sqrt{n}}{2}, \ldots, \frac{n}{2} + \frac{\sqrt{n}}{2}\right\}}\left[\left(S - \frac{n}{2}\right)^2\right].$$

So,

$$\mathbf{E}\big[\mathcal{W}_{j,s}\big] \leq \frac{1}{n} \cdot \mathop{\mathbf{E}}_{S \in \left\{\frac{n}{2} - \frac{\sqrt{n}}{2}, \ldots, \frac{n}{2} + \frac{\sqrt{n}}{2}\right\}}\left[\left(S - \frac{n}{2}\right)^2\right]$$

$$= \frac{1}{n} \cdot \mathop{\mathbf{E}}_{S \in \left\{-\frac{\sqrt{n}}{2}, \ldots, \frac{\sqrt{n}}{2}\right\}}\left[S^2\right] < \frac{n + 3\sqrt{n} + 2}{12n} \leq \frac{1}{11},$$

as n is large enough, and

$$\mathbf{E}\left[\sum_{j \in [n], \, s \in \{0,1\}^{\log n}} \mathcal{W}_{j,s}\right] < \frac{n^2}{11}.$$

Combining this with (27) and applying Chebyshev's inequality (Fact 1), we get

$$\mathbf{Pr}\left[\sum_{j \in [n], \, s \in \{0,1\}^{\log n}} \mathcal{W}_{j,s} > \frac{n^2}{9}\right] \leq \mathbf{Pr}\left[\left|\sum_{j,s} \mathcal{W}_{j,s} - \mathbf{E}\left[\sum_{j,s} \mathcal{W}_{j,s}\right]\right| > \frac{n^2}{50}\right]$$

$$\leq 2500 \cdot \frac{\mathbf{Var}\left[\sum_{j,s} \mathcal{W}_{j,s}\right]}{n^4}$$

$$= 2500 \cdot \frac{\sum_{j,s} \boldsymbol{Var}\big[\mathcal{W}_{j,s}\big] + \sum_{(j_1,s_1)\neq(j_2,s_2)} \boldsymbol{Cov}\big[\mathcal{W}_{j_1,s_1}, \mathcal{W}_{j_2,s_2}\big]}{n^4}$$

$$\leq 2500 \cdot \frac{n^2 \cdot \frac{1}{16} + n^4 \cdot \frac{206 \cdot (\ln n)^2}{\sqrt{n}}}{n^4} \quad \leq \quad \frac{517500 \cdot (\ln n)^2}{\sqrt{n}},$$

which implies the required (23).   ∎ *Theorem 2*

Now that we have satisfied our curiosity concerning the likelihood of $\aleph(\mathcal{X}, \mathcal{Y})$ with respect to the uniform distribution, let us analyse the $\mathcal{R}$-complexity of *GHR*.

6.1 Spectral stability of large rectangles

Lower-bounding the $\mathcal{R}$-complexity of a communication problem is often based on showing that it admits no large nearly-monochromatic rectangles with respect to certain input distribution.[21] We will use this approach in our analysis of *GHR*, and the input distribution that we consider will be the uniform.

A core structural property of large rectangles that will be used in the analysis is their stability with respect to the spectrum of Hamming distances between the two input values. Namely, we will claim that if $(\mathcal{X}, \mathcal{Y})$ is sampled at uniform from a large rectangle, then $|\mathcal{X} \oplus \mathcal{Y}|$ is distributed, in certain sense, similarly to how it is distributed when $(\mathcal{X}, \mathcal{Y}) \Subset \{0,1\}^{n+n}$.

Note that large rectangles can fully control the *parity* of $|\mathcal{X} \oplus \mathcal{Y}|$: say, if both A and B are the set of strings of even Hemming weight, then $2|\,|x \oplus y|$ for every $(x,y) \in A \times B$. Also it is relatively easy to control the *marginal* values of $|\mathcal{X} \oplus \mathcal{Y}|$: say, an efficient protocol can use binary search to find a poly-logarithmic number of coordinates where the pair of inputs disagree and then test the equality over the rest of coordinates, thus checking whether the distance between the two input values is at most poly-logarithmic. On the other hand, it is intuitively obvious that, other than knowing the parity of $|\mathcal{X} \oplus \mathcal{Y}|$, a large rectangle cannot have a good command of this value's distribution near $n/2$. The following statement formalises this intuition. [22]

Theorem 3. *Let* $A \times B \subseteq \{0,1\}^{n+n}$ *be a rectangle,* $k \in [n]-1$ *and* $\mu_{A \times B} \overset{def}{=} \log\left(\frac{n}{\mathcal{U}_{\{0,1\}^{n+n}}(A \times B)}\right)$. *If* $\sqrt{n \cdot \mu_{A \times B}} \leq n/14$ *and* $|n/2 - k| \leq n/14$, *then*

$$1 - O\left(\frac{(\mu_{A \times B})^{\frac{3}{2}}}{\sqrt{n}} + \left|\frac{1}{2} - \frac{k}{n}\right| \cdot \mu_{A \times B}\right)$$

$$\leq \frac{\boldsymbol{Pr}_{(\mathcal{X},\mathcal{Y}) \Subset A \times B}\big[|\mathcal{X} \oplus \mathcal{Y}| \in \{k, k+1\}\big]}{\boldsymbol{Pr}_{(\mathcal{X},\mathcal{Y}) \Subset \{0,1\}^{n+n}}\big[|\mathcal{X} \oplus \mathcal{Y}| \in \{k, k+1\}\big]}$$

$$\leq 1 + O\left(\left(\sqrt{\frac{\log n}{n}} + \left|\frac{1}{2} - \frac{k}{n}\right|\right) \cdot \mu_{A \times B}\right),$$

---

[21] A set of inputs to a relational problem is called (nearly-)monochromatic if some answer is correct for (almost) all of the set elements.

[22] We've tried to give Theorem 3 a modular formulation, but did not struggle to optimise it beyond our requirements, so some of the parameters can possibly be improved.

where the right-hand side holds only as long as it is $1 + o(1)$. [23]

In other words, large rectangles are not too biased with respect to the likely values of $|\mathcal{X} \oplus \mathcal{Y}|$, as long as we consider at least pairs of consecutive points in the spectrum (this is necessary since singletons are affected by the parity control, as discussed above).

We will use the following result due to Razborov [Raz92].

Fact 4. *Let μ_j be the uniform distribution over $\left\{ (x, y) \in \binom{[4l-1]}{l} \times \binom{[4l-1]}{l} \,\middle|\, |x \cap y| = j \right\}$ for $j \in \{0, 1\}$. Then for every $A, B \subseteq \{0, 1\}^{4l-1}$:*

$$\mu_1(A \times B) \geq \frac{1}{45} \cdot \mu_0(A \times B) - 2^{-\Omega(l)}.$$

Intuitively, it says that only very small rectangles – those of size $2^{-\Omega(l)}$ – can contain much more disjoint pairs than pairs intersecting over exactly one element (this was used in [Raz92] to show that the $\mathcal{R}$-complexity of the bipartite disjointness function was $\Omega(n)$).

Proof of Theorem 3. For any $s \subseteq [n] \cup \{0\}$, define the *relative weight* of s (with respect to $A \times B$) as

$$\mathrm{rw}(s) = \frac{\boldsymbol{Pr}_{(\mathcal{X}, \mathcal{Y}) \in A \times B}\big[|\mathcal{X} \oplus \mathcal{Y}| \in s \big]}{\boldsymbol{Pr}_{(\mathcal{X}, \mathcal{Y}) \in \{0,1\}^{n+n}}\big[|\mathcal{X} \oplus \mathcal{Y}| \in s \big]},$$

and for every $i \in [n] \cup \{0\}$ let $\mathrm{rw}(i) \overset{\mathrm{def}}{=} \mathrm{rw}(\{i\})$.

Assuming that $A \times B$ exhibits spectral irregularity – that is, a deviation of $\mathrm{rw}(\cdot)$ from 1 – we will find a *representative pair of coordinates* that, on the one hand, are both not too far from $n/2$, and on the other hand, whose relative weights are somewhat apart from each other: in combination with Fact 4 this will lead to an upper bound on the size of $A \times B$.

Our way of finding a representative pair will depend on $\mathrm{rw}(\{k, k+1\})$.

The case of $\mathrm{rw}(\{k, k+1\}) > 1$. For

$$\lambda_+ \overset{\mathrm{def}}{=} \min\big\{ \mathrm{rw}(\{k, k+1\}) - 1, \, 1 \big\},$$

let t_+ be the minimal distance from $n/2$ of a pair of coordinates with relative weight at most $1 + \lambda_+/3$ and same parity pattern as in $\{k, k+1\}$:

$$t_+ \overset{\mathrm{def}}{=} \min_i \left\{ \left| \frac{n}{2} - i \right| \,\middle|\, \mathrm{rw}(\{i, i+1\}) \leq 1 + \frac{\lambda_+}{3}, \, 2|(i+k) \right\}.$$

The integer interval

$$T_+ \overset{\mathrm{def}}{=} \left\{ \frac{n}{2} - t_+ + 2, \ldots, \frac{n}{2} + t_+ - 1 \right\}$$

partitions into pairs $(n/2 - t_+ + 2, \, n/2 - t_+ + 3), \ldots, (n/2 + t_+ - 2, \, n/2 + t_+ - 1)$, each satisfying $[\mathrm{rw}(\{j, j+1\}) > 1 + \lambda_+/3]$. Accordingly,

$$\mathrm{rw}(T_+) > 1 + \frac{\lambda_+}{3}$$

---

[23] More precisely, the upper bound holds if the expression inside $O(\cdot)$ is less than some universal constant, otherwise not necessarily.

and

$$1 \geq \Pr_{(\mathcal{X},\mathcal{Y}) \in A \times B}\left[(\mathcal{X},\mathcal{Y}) \in T_+\right]$$

$$\geq \left(1 + \frac{\lambda_+}{3}\right) \cdot \Pr_{(\mathcal{X},\mathcal{Y}) \in \{0,1\}^{n+n}}\left[(\mathcal{X},\mathcal{Y}) \in T_+\right]$$

$$\geq \left(1 + \frac{\lambda_+}{3}\right) \cdot \Pr_{(\mathcal{X},\mathcal{Y}) \in \{0,1\}^{n+n}}\left[\left||\mathcal{X} \oplus \mathcal{Y}| - \frac{n}{2}\right| < t_+ - 1\right],$$

as $\left[\left||\mathcal{X} \oplus \mathcal{Y}| - {n}/{2}\right| < t_+ - 1\right]$ implies $\left[(\mathcal{X},\mathcal{Y}) \in T_+\right]$. By Hoeffding's inequality (Fact 2),

$$\Pr_{(\mathcal{X},\mathcal{Y}) \in \{0,1\}^{n+n}}\left[\left||\mathcal{X} \oplus \mathcal{Y}| - \frac{n}{2}\right| \geq t_+ - 1\right] < 2 \cdot e^{\frac{4t_+ - 2t_+^2}{n}},$$

and therefore

$$\left(1 + \frac{\lambda_+}{3}\right) \cdot \left(1 - 2 \cdot e^{\frac{4t_+ - 2t_+^2}{n}}\right) < 1,$$

which, due to $\lambda_+ \in (0,1]$, implies that for large enough n,

$$1 - 2 \cdot e^{\frac{4t_+ - 2t_+^2}{n}} < \frac{1}{1 + \frac{\lambda_+}{3}} \leq 1 - \frac{\lambda_+}{4} \implies e^{\frac{4t_+ - 2t_+^2}{n}} > \frac{\lambda_+}{8} \implies t_+ < \sqrt{n \cdot \ln\left(\frac{1}{\lambda_+}\right)}. \quad (28)$$

Let $a \in \{{n}/{2} - t_+,\ {n}/{2} + t_+\}$ be such that $\mathrm{rw}(\{a,\ a + 1\}) \leq 1 + \lambda_+/3$. As in the case that we've just considered $\mathrm{rw}(\{k,\ k + 1\}) \geq 1 + \lambda_+$ and $0 < \lambda_+ \leq 1$, it holds that $(1 + \Lambda) \cdot \mathrm{rw}(\{a,\ a + 1\}) \leq \mathrm{rw}(\{k,\ k + 1\})$ for $\Lambda \overset{\mathrm{def}}{=} \lambda_+/3 > 0$. Note that

$$\left|{n}/{2} - a\right| = t_+, \quad \Lambda \in \Theta(\lambda_+) \quad (29)$$

and $2|(a + k)$ hold by construction.

The case of $\mathrm{rw}(\{k,\ k + 1\}) < 1$. For

$$\lambda_- \overset{\mathrm{def}}{=} 1 - \mathrm{rw}(\{k,\ k + 1\}),$$

let t_- be the minimal distance from ${n}/{2}$ of a pair of coordinates with relative weight at least $1 - \lambda_-/2$ and same parity pattern as in $\{k,\ k + 1\}$:

$$t_- \overset{\mathrm{def}}{=} \min_i\left\{\left|\frac{n}{2} - i\right| \,\middle|\, \mathrm{rw}(\{i,\ i + 1\}) \geq 1 - \frac{\lambda_-}{2},\ 2|(i + k)\right\}.$$

The integer interval

$$T_- \overset{\mathrm{def}}{=} \left\{\frac{n}{2} - t_- + 2, \ldots, \frac{n}{2} + t_- - 1\right\}$$

partitions into pairs $({n}/{2} - t_- + 2,\ {n}/{2} - t_- + 3), \ldots, ({n}/{2} + t_- - 2,\ {n}/{2} + t_- - 1)$, each satisfying $[\mathrm{rw}(\{j,\ j + 1\}) < 1 - \lambda_-/2]$. Accordingly,

$$\mathrm{rw}(T_-) < 1 - \frac{\lambda_-}{2}.$$

To bound t_- from above, let us use the fact that the subset of $A \times B$ where $\mathrm{rw}(|x+y|) \geq 1 - \lambda_-/2$ must *compensate* the missing $\lambda_-/2$-portion in the rest of the rectangle:

$$\Pr_{(\mathcal{X},\mathcal{Y}) \Subset A \times B}\left[\left|\,|\mathcal{X} \oplus \mathcal{Y}| - \frac{n}{2}\right| \notin T_-\right] = \sum_{i \notin T_-} \Pr_{(\mathcal{X},\mathcal{Y}) \Subset A \times B}\left[\left|\,|\mathcal{X} \oplus \mathcal{Y}| - \frac{n}{2}\right| = i\right]$$

$$\geq \sum_{i \notin T_-} \left(\Pr_{(\mathcal{X},\mathcal{Y}) \Subset A \times B}\left[\left|\,|\mathcal{X} \oplus \mathcal{Y}| - \frac{n}{2}\right| = i\right] - \Pr_{(\mathcal{X},\mathcal{Y}) \Subset \{0,1\}^{n+n}}\left[\left|\,|\mathcal{X} \oplus \mathcal{Y}| - \frac{n}{2}\right| = i\right]\right)$$

$$= \sum_{i \in T_-} \left(\Pr_{(\mathcal{X},\mathcal{Y}) \Subset \{0,1\}^{n+n}}\left[\left|\,|\mathcal{X} \oplus \mathcal{Y}| - \frac{n}{2}\right| = i\right] - \Pr_{(\mathcal{X},\mathcal{Y}) \Subset A \times B}\left[\left|\,|\mathcal{X} \oplus \mathcal{Y}| - \frac{n}{2}\right| = i\right]\right)$$

$$\geq \frac{\lambda_-}{2} \cdot \Pr_{(\mathcal{X},\mathcal{Y}) \Subset \{0,1\}^{n+n}}\left[\left|\,|\mathcal{X} \oplus \mathcal{Y}| - \frac{n}{2}\right| \in T_-\right]$$

$$\geq \frac{\lambda_-}{2} \cdot \Pr_{(\mathcal{X},\mathcal{Y}) \Subset A \times B}\left[\left|\,|\mathcal{X} \oplus \mathcal{Y}| - \frac{n}{2}\right| \in T_-\right],$$

from where by adding $\frac{\lambda_-}{2} \cdot \boldsymbol{Pr}_{A \times B}\left[\left|\,|\mathcal{X} \oplus \mathcal{Y}| - \frac{n}{2}\right| \notin T_-\right]$ to both the sides we get

$$\frac{\lambda_-}{2} \leq \frac{3}{2} \cdot \Pr_{(\mathcal{X},\mathcal{Y}) \Subset A \times B}\left[\left|\,|\mathcal{X} \oplus \mathcal{Y}| - \frac{n}{2}\right| \notin T_-\right]$$

$$= \frac{3}{2} \cdot \Pr_{(\mathcal{X},\mathcal{Y}) \Subset \{0,1\}^{n+n}}\left[\left|\,|\mathcal{X} \oplus \mathcal{Y}| - \frac{n}{2}\right| \notin T_- \,\Big|\, (\mathcal{X},\mathcal{Y}) \in A \times B\right]$$

$$\leq \frac{3}{2} \cdot \frac{\boldsymbol{Pr}_{(\mathcal{X},\mathcal{Y}) \Subset \{0,1\}^{n+n}}\left[\left|\,|\mathcal{X} \oplus \mathcal{Y}| - \frac{n}{2}\right| \notin T_-\right]}{\mathcal{U}_{\{0,1\}^{n+n}}(A \times B)}$$

$$\leq \frac{3}{2} \cdot \frac{\boldsymbol{Pr}_{(\mathcal{X},\mathcal{Y}) \Subset \{0,1\}^{n+n}}\left[\left|\,|\mathcal{X} \oplus \mathcal{Y}| - \frac{n}{2}\right| \geq t_- - 1\right]}{\mathcal{U}_{\{0,1\}^{n+n}}(A \times B)},$$

as $\left[\left|\,|\mathcal{X} \oplus \mathcal{Y}| - n/2\right| \notin T_-\right]$ implies $\left[\left|\,|\mathcal{X} \oplus \mathcal{Y}| - n/2\right| \geq t_- - 1\right]$. By Hoeffding's inequality (Fact 2),

$$\Pr_{(\mathcal{X},\mathcal{Y}) \Subset \{0,1\}^{n+n}}\left[\left|\,|\mathcal{X} \oplus \mathcal{Y}| - \frac{n}{2}\right| \geq t_- - 1\right] < 2 \cdot e^{\frac{4t_- - 2t_-^2}{n}},$$

and so, for large enough n,

$$e^{\frac{4t_- - 2t_-^2}{n}} > \frac{\lambda_-}{6} \cdot \mathcal{U}_{\{0,1\}^{n+n}}(A \times B) \implies t_- < \sqrt{n \cdot \ln\left(\frac{1}{\lambda_- \cdot \mathcal{U}_{\{0,1\}^{n+n}}(A \times B)}\right)}. \tag{30}$$

Let $a \in \{n/2 - t_-, \, n/2 + t_-\}$ be such that $\mathrm{rw}(\{a, a+1\}) \geq 1 - \lambda_-/2$. As in the case that we've just considered $\mathrm{rw}(\{k, k+1\}) = 1 - \lambda_-$ for $\lambda_- > 0$, it holds that $(1+\Lambda) \cdot \mathrm{rw}(\{k, k+1\}) \leq \mathrm{rw}(\{a, a+1\})$ for $\Lambda \stackrel{\text{def}}{=} \lambda_-/2 > 0$. Note that

$$\left|n/2 - a\right| = t_-, \quad \Lambda \in \Theta(\lambda_-) \tag{31}$$

and $2|(a+k)$ hold by construction.

Completion. We need a pair of coordinates b_1, $b_2 \in [n] - 1$ with the following properties:

$$\begin{cases} b_1 < b_2 \text{ and } 2|(b_2 - b_1); \\ (1 + \Lambda) \cdot \mathrm{rw}(\{b_1, \, b_1 + 1\}) \leq \mathrm{rw}(\{b_2, \, b_2 + 1\}) \text{ for some } \Lambda \in (0, 1]; \\ \mathrm{rw}(\{b_2, \, b_2 + 1\}) \geq 1/2. \end{cases} \quad (32)$$

The pair of coordinates that we have identified in the previous stage (k and a) almost satisfy these requirements: it may happen that *the ordering is wrong* (say, $k > a$ but $\mathrm{rw}(\{k, \, k+1\}) < \mathrm{rw}(\{a, \, a+1\})$). However, this situation can be addressed by a simple transformation of the rectangle under consideration: if we let

$$B' \stackrel{\text{def}}{=} \Big\{ \neg b \Big| b \in B \Big\},$$

where "$\neg$" stands for the bit-wise negation of a binary string, then, obviously,

$$\forall i \in [n] \cup \{0\} : \quad \underset{(\mathcal{X}, \mathcal{Y}) \in A \times B}{\boldsymbol{Pr}} \big[|\mathcal{X} \oplus \mathcal{Y}| = i \big] = \underset{(\mathcal{X}, \mathcal{Y}) \in A \times B'}{\boldsymbol{Pr}} \big[|\mathcal{X} \oplus \mathcal{Y}| = n - i \big].$$

That is, the distributions of $|\mathcal{X} \oplus \mathcal{Y}|$ corresponding to the rectangles $A \times B$ and $A \times B'$ are symmetrically opposite, and we will get all the three conditions of (32) satisfied either for the rectangle $A \times B$ with the coordinates k and a, or for the rectangle $A \times B'$ with the coordinates $n - k$ and $n - a$.

For notational convenience (and without loss of generality), let us assume the former case: That is, we are assuming that a pair of points b_1, $b_2 \in [n] - 1$ satisfy the conditions listed at (32) with parameters provided by the appropriate version of our pair-picking algorithm (note that the conditions are implicitly referring to the rectangle $A \times B$, as the relative weight $\mathrm{rw}(\cdot)$ is defined with respect to it).

For the following argument we will use a pair of coordinates: either $(b_1, \, b_2)$ or $(b_1+1, \, b_2+1)$, let us make the choice. By the nature of the binomial distribution,

$$\forall i \in [n] - 1 : \quad \underset{(\mathcal{X}, \mathcal{Y}) \in \{0,1\}^{n+n}}{\boldsymbol{Pr}} \big[|\mathcal{X} \oplus \mathcal{Y}| = i + 1 \big] = \frac{i+1}{n-i} \cdot \underset{(\mathcal{X}, \mathcal{Y}) \in \{0,1\}^{n+n}}{\boldsymbol{Pr}} \big[|\mathcal{X} \oplus \mathcal{Y}| = i \big],$$

and therefore,

$$\begin{aligned} \mathrm{rw}(\{i, \, i+1\}) &= \frac{\boldsymbol{Pr}_{A \times B} \big[|\mathcal{X} \oplus \mathcal{Y}| = i \big] + \boldsymbol{Pr}_{A \times B} \big[|\mathcal{X} \oplus \mathcal{Y}| = i + 1 \big]}{\boldsymbol{Pr}_{\{0,1\}^{n+n}} \big[|\mathcal{X} \oplus \mathcal{Y}| = i \big] + \boldsymbol{Pr}_{\{0,1\}^{n+n}} \big[|\mathcal{X} \oplus \mathcal{Y}| = i + 1 \big]} \\ &= \frac{\boldsymbol{Pr}_{A \times B} \big[|\mathcal{X} \oplus \mathcal{Y}| = i \big]}{\boldsymbol{Pr}_{\{0,1\}^{n+n}} \big[|\mathcal{X} \oplus \mathcal{Y}| = i \big] \cdot \left(1 + \frac{i+1}{n-i} \right)} \\ &\quad + \frac{\boldsymbol{Pr}_{A \times B} \big[|\mathcal{X} \oplus \mathcal{Y}| = i + 1 \big]}{\boldsymbol{Pr}_{\{0,1\}^{n+n}} \big[|\mathcal{X} \oplus \mathcal{Y}| = i + 1 \big] \cdot \left(1 + \frac{n-i}{i+1} \right)} \\ &= \mathrm{rw}(i) \cdot \frac{n-i}{n+1} + \mathrm{rw}(i+1) \cdot \frac{i+1}{n+1}. \end{aligned}$$

Accordingly, our assumptions regarding b_1 and b_2 imply that $1/2 < \mathrm{rw}(b_2) + \mathrm{rw}(b_2 + 1)$ and

$$(1 + \Lambda) \cdot \mathrm{rw}(b_1) \cdot (n - b_1) + (1 + \Lambda) \cdot \mathrm{rw}(b_1 + 1) \cdot (b_1 + 1)$$
$$\leq \mathrm{rw}(b_2) \cdot (n - b_2) + \mathrm{rw}(b_2 + 1) \cdot (b_2 + 1).$$

If we let

$$\delta_{max} \overset{\text{def}}{=} \frac{\max\{|n - 2 \cdot b_1|, |n - 2 \cdot b_2|\} + 2}{n} \quad \left(= \frac{\max\{|n - 2 \cdot a|, |n - 2 \cdot k|\} + 2}{n} \right),$$

then

$$(1 + \Lambda) \cdot \mathrm{rw}(b_1) \cdot (1 - \delta_{max}) + (1 + \Lambda) \cdot \mathrm{rw}(b_1 + 1) \cdot (1 - \delta_{max})$$
$$\leq \mathrm{rw}(b_2) \cdot (1 + \delta_{max}) + \mathrm{rw}(b_2 + 1) \cdot (1 + \delta_{max})$$

and

$$(1 + \Lambda) \cdot (1 - 2\delta_{max}) \cdot \Big(\mathrm{rw}(b_1) + \mathrm{rw}(b_1 + 1)\Big) \leq \mathrm{rw}(b_2) + \mathrm{rw}(b_2 + 1). \tag{33}$$

Recall that $\mathrm{rw}(b_2) + \mathrm{rw}(b_2 + 1) > 1/2$ and $\Lambda \in (0, 1]$, consider the following cases. If $\mathrm{rw}(b_2 + 1) < \Lambda/7$, then $\mathrm{rw}(b_2) > 5/14$ and from (33):

$$(1 + \Lambda) \cdot (1 - 2\delta_{max}) \cdot \Big(\mathrm{rw}(b_1) + \mathrm{rw}(b_1 + 1)\Big) < \mathrm{rw}(b_2) + \frac{\Lambda}{7} < \mathrm{rw}(b_2) \cdot \left(1 + \frac{2\Lambda}{5}\right),$$

implying

$$\mathrm{rw}(b_2) > \frac{1 + \Lambda}{1 + \frac{2\Lambda}{5}} \cdot (1 - 2\delta_{max}) \cdot \mathrm{rw}(b_1) > \left(1 + \frac{\Lambda}{5}\right) \cdot (1 - 2\delta_{max}) \cdot \mathrm{rw}(b_1).$$

Similarly, if $\mathrm{rw}(b_2) < \Lambda/7$, then $\mathrm{rw}(b_2 + 1) > 5/14$ and

$$\mathrm{rw}(b_2 + 1) > \left(1 + \frac{\Lambda}{5}\right) \cdot (1 - 2\delta_{max}) \cdot \mathrm{rw}(b_1 + 1).$$

Lastly, if $\mathrm{rw}(b_2), \mathrm{rw}(b_2 + 1) \geq \Lambda/7$, then it follows from (33) that at least one of

$$\mathrm{rw}(b_2) \geq (1 + \Lambda) \cdot (1 - 2\delta_{max}) \cdot \mathrm{rw}(b_1), \ \ \mathrm{rw}(b_2 + 1) \geq (1 + \Lambda) \cdot (1 - 2\delta_{max}) \cdot \mathrm{rw}(b_1 + 1)$$

must hold.

In every case, we have identified a pair of coordinates $c_1, c_2 \in [n] \cup \{0\}$, such that the following holds:

$$\begin{cases} c_2 > c_1 \text{ and } 2|(c_2 - c_1); \\ |n/2 - c_1|, |n/2 - c_2| \leq \max\Big\{|n/2 - a|, |n/2 - k|\Big\} + 1; \\ \mathrm{rw}(c_2) \geq (1 + \Lambda/5 - 4\delta_{max}) \cdot \mathrm{rw}(c_1) \text{ for } \delta_{max} = \Big(\max\{|n - 2a|, |n - 2k|\} + 2\Big)/n; \\ \mathrm{rw}(c_2) \geq \Lambda/7. \end{cases} \tag{34}$$

Now let

$$l \overset{\text{def}}{=} \left\lfloor \frac{c_2}{4(c_2 - c_1)} \right\rfloor$$

and assume that

$$3 \cdot c_2 \leq 4 \cdot c_1 \tag{35}$$

– that is, $l \geq 1$ (in the end we will make sure that this holds). Suppose that we are given some $(x, y) \in \binom{[4l-1]}{l} \times \binom{[4l-1]}{l}$ and the goal is to distinguish the case of $[x \cap y = \emptyset]$ from that of $[|x \cap y| = 1]$, using the assumed irregularity of $A \times B$ (as witnessed by the coordinates c_1 and

35

c_2) – that will let us derive the desired conclusion from the hardness of distinguishing the two cases (as expressed by Fact 4).

Consider the following randomised algorithm that we will denote by Ξ.

1. Map x to $\mathcal{X}_1 \in \{0,1\}^{12l-3}$ by converting every x_i, $i \in [4l-1]$ into $\mathcal{X}_1|_{3i-2,\dots,3i} \in \{0,1\}^3$ as follows:

$$\mathcal{X}_1|_{3i-2,\dots,3i} = \begin{cases} 100 & \text{if } x_i = 0; \\ 010 & \text{if } x_i = 1. \end{cases}$$

2. Map y to $\mathcal{Y}_1 \in \{0,1\}^{12l-3}$ by mapping every y_i, $i \in [4l-1]$ to $\mathcal{Y}_1|_{3i-2,\dots,3i} \in \{0,1\}^3$ as follows:

$$\mathcal{Y}_1|_{3i-2,\dots,3i} = \begin{cases} 001 & \text{if } x_i = 0; \\ 010 & \text{if } x_i = 1. \end{cases}$$

3. Let $\mathcal{X}_2, \mathcal{Y}_2 \in \{0,1\}^{\frac{(12l-3)\cdot(c_2-c_1)}{2}}$ be the concatenations of $\frac{c_2-c_1}{2}$ copies of, respectively, $\mathcal{X}_1$ and $\mathcal{Y}_1$.

4. Let $\mathcal{X}_3 \in \{0,1\}^n$ be the concatenations of $\mathcal{X}_2$ with $0^{n-\frac{(12l-3)\cdot(c_2-c_1)}{2}}$.

5. Let $\mathcal{Y}_3 \in \{0,1\}^n$ be the concatenations of $\mathcal{Y}_2$ with $1^{c_2-(4l-1)(c_2-c_1)}$ and with $0^{n-c_2-\frac{(4l-1)(c_2-c_1)}{2}}$.

6. Let $\mathcal{S}$ be a uniformly-random permutation of $[n]$ and $\mathcal{X}_4, \mathcal{Y}_4 \in \{0,1\}^n$ contain the outcomes of, respectively, $S(\mathcal{X}_3)$ and $S(\mathcal{Y}_3)$.

7. Let $\mathcal{T} \sim \mathcal{U}_{\{0,1\}^n}$ and $\mathcal{X}_5, \mathcal{Y}_5 \in \{0,1\}^n$ contain the outcomes of, respectively, $\mathcal{T} \oplus \mathcal{X}_4$ and $\mathcal{T} \oplus \mathcal{Y}_4$.

8. *Accept* if $(\mathcal{X}_5, \mathcal{Y}_5) \in A \times B$ and *reject* otherwise.

In the above description of the procedure Ξ, let us view $\mathcal{S}$, $\mathcal{T}$, $\mathcal{X}_i$-s and $\mathcal{Y}_i$-s as random variables: some of them hold values that are fixed by those of x and y, others possess entropy of their own. Note however that once the values of $\mathcal{S}$ and $\mathcal{T}$ are fixed, Ξ becomes a deterministic algorithm with input (x,y).

We claim that if $x \cap y = \emptyset$, then $(\mathcal{X}_5, \mathcal{Y}_5) \sim \mathcal{U}_{\{0,1\}^{n+n}}$ subject to $|\mathcal{X}_5 \oplus \mathcal{Y}_5| = c_2$, and if $|x \cap y| = 1$, then $(\mathcal{X}_5, \mathcal{Y}_5) \sim \mathcal{U}_{\{0,1\}^{n+n}}$ subject to $|\mathcal{X}_5 \oplus \mathcal{Y}_5| = c_1$. Indeed, suppose that $x \cap y = \emptyset$, then the distance between every pair of corresponding bit-triples of $\mathcal{X}_1$ and $\mathcal{Y}_1$ is exactly 2 and therefore $|\mathcal{X}_1 \oplus \mathcal{Y}_1| = 8l - 2$. Consequently, $|\mathcal{X}_2 \oplus \mathcal{Y}_2| = (c_2 - c_1)(4l - 1)$ and

$$|\mathcal{X}_3 \oplus \mathcal{Y}_3| = (c_2 - c_1)(4l - 1) + c_2 - (4l - 1)(c_2 - c_1) = c_2.$$

If, on the other hand, $|x \cap y| = 1$, then the distance between all but one pair of corresponding bit-triples of $\mathcal{X}_1$ and $\mathcal{Y}_1$ is 2 and the remaining pair of triples both equal 010, and therefore $|\mathcal{X}_1 \oplus \mathcal{Y}_1| = 8l - 4$, $|\mathcal{X}_2 \oplus \mathcal{Y}_2| = (c_2 - c_1)(4l - 2)$ and

$$|\mathcal{X}_3 \oplus \mathcal{Y}_3| = (c_2 - c_1)(4l - 2) + c_2 - (4l - 1)(c_2 - c_1) = c_2 - (c_2 - c_1) = c_1.$$

The above statement about the conditional distribution of $(\mathcal{X}_5, \mathcal{Y}_5)$ follows by noticing that steps (6) and (7) map every $(\mathcal{X}_3, \mathcal{Y}_3)$ to uniformly-random $(\mathcal{X}_5, \mathcal{Y}_5)$, subject to $[|\mathcal{X}_5 \oplus \mathcal{Y}_5| = |\mathcal{X}_3 \oplus \mathcal{Y}_3|]$.

As

$$\forall i \in [n] \cup \{0\} : \underset{\mathcal{U}_{\{0,1\}^{n+n}}}{\boldsymbol{Pr}}\Big[(\mathcal{X}, \mathcal{Y}) \in A \times B\,\big|\,|\mathcal{X} \oplus \mathcal{Y}| = i\Big]$$

$$= \underset{\mathcal{U}_{\{0,1\}^{n+n}}}{\boldsymbol{Pr}}\Big[|\mathcal{X} \oplus \mathcal{Y}| = i\,\big|\,(\mathcal{X}, \mathcal{Y}) \in A \times B\Big] \cdot \frac{\boldsymbol{Pr}_{\mathcal{U}_{\{0,1\}^{n+n}}}\big[(\mathcal{X}, \mathcal{Y}) \in A \times B\big]}{\boldsymbol{Pr}_{\mathcal{U}_{\{0,1\}^{n+n}}}\big[|\mathcal{X} \oplus \mathcal{Y}| = i\big]}$$

$$= \mathrm{rw}(i) \cdot \mathcal{U}_{\{0,1\}^{n+n}}(A \times B),$$

it follows by (34) that

$$\frac{\boldsymbol{Pr}\big[\Xi \text{ accepts}\big|x \cap y = \emptyset\big]}{\boldsymbol{Pr}\big[\Xi \text{ accepts}\big|\,|x \cap y| = 1\big]} = \frac{\boldsymbol{Pr}_{\mathcal{U}_{\{0,1\}^{n+n}}}\big[(\mathcal{X}, \mathcal{Y}) \in A \times B\big|\,|\mathcal{X} \oplus \mathcal{Y}| = c_2\big]}{\boldsymbol{Pr}_{\mathcal{U}_{\{0,1\}^{n+n}}}\big[(\mathcal{X}, \mathcal{Y}) \in A \times B\big|\,|\mathcal{X} \oplus \mathcal{Y}| = c_1\big]}$$

$$\geq 1 + \Lambda/5 - 4\delta_{max}$$

and

$$\boldsymbol{Pr}\Big[\Xi \text{ accepts}\Big|x \cap y = \emptyset\Big] \geq \frac{\Lambda \cdot \mathcal{U}_{\{0,1\}^{n+n}}(A \times B)}{7}.$$

Note that conditioned on the values of $\mathcal{S}$ and $\mathcal{T}$, the set of pairs (x, y) accepted by Ξ forms a combinatorial rectangle in $\binom{[4l-1]}{l} \times \binom{[4l-1]}{l}$. Moreover, if we run k independent instances of Ξ for the same input (x, y) and *accept if and only if all k executions have accepted*, even then the set of accepted pairs still forms a combinatorial rectangle (conditions on the values of both $\mathcal{S}$ and $\mathcal{T}$ in each instance), as the intersection of many rectangles is a rectangle.

For the version with k repetitions the following holds, due to the mutual independence of the instances:

$$\forall (x, y) \in \binom{[4l-1]}{l} \times \binom{[4l-1]}{l} :$$

$$\frac{\boldsymbol{Pr}\big[Accept\big|x \cap y = \emptyset\big]}{\boldsymbol{Pr}\big[Accept\big|\,|x \cap y| = 1\big]} \geq \left(1 + \frac{\Lambda}{5} - 4\delta_{max}\right)^k$$

and

$$\boldsymbol{Pr}\Big[Accept\Big|x \cap y = \emptyset\Big] \geq \left(\frac{\Lambda \cdot \mathcal{U}_{\{0,1\}^{n+n}}(A \times B)}{7}\right)^k.$$

Let $\mathcal{X}$ and $\mathcal{Y}$ be random variables that respectively take the values x and y, and let $\mathcal{U}$ denote their uniform distribution (i.e., $\mathcal{U}_{\binom{[4l-1]}{l} \times \binom{[4l-1]}{l}}$). Let $\Xi^k_{s_1,\dots,s_k,\,t_1,\dots,t_k}$ denote the k-repetition version of Ξ, where the i'th instance runs with $\mathcal{S} = s_i$, $\mathcal{T} = t_i$. Then the standard counting argument implies that there exist such values of $s_1, \dots, s_k, t_1, \dots, t_k$ that the (deterministic) procedure $\Xi^k_{s_1,\dots,s_k,\,t_1,\dots,t_k}$ satisfies both

$$\frac{\boldsymbol{Pr}_{(\mathcal{X},\mathcal{Y})\sim\mathcal{U}}\Big[\Xi^k_{s_1,\dots,s_k,t_1,\dots,t_k}(\mathcal{X},\mathcal{Y}) \text{ accepts}\Big|\mathcal{X} \cap \mathcal{Y} = \emptyset\Big]}{\boldsymbol{Pr}_{(\mathcal{X},\mathcal{Y})\sim\mathcal{U}}\Big[\Xi^k_{s_1,\dots,s_k,t_1,\dots,t_k}(\mathcal{X},\mathcal{Y}) \text{ accepts}\Big|\,|\mathcal{X} \cap \mathcal{Y}| = 1\Big]} \geq \frac{1}{2} \cdot \left(1 + \frac{\Lambda}{5} - 4\delta_{max}\right)^k$$

and

$$\underset{(\mathcal{X},\mathcal{Y})\sim\mathcal{U}}{\boldsymbol{Pr}}\Big[\Xi^k_{s_1,\dots,s_k,t_1,\dots,t_k}(\mathcal{X},\mathcal{Y}) \text{ accepts}\Big|\mathcal{X} \cap \mathcal{Y} = \emptyset\Big] \geq \frac{1}{2} \cdot \left(\frac{\Lambda \cdot \mathcal{U}_{\{0,1\}^{n+n}}(A \times B)}{7}\right)^k.$$

As the acceptance region of $\Xi^k_{s_1,\dots,s_k,\,t_1,\dots,t_k}$ forms a rectangle, Fact 4 implies that

$$\Pr_{(\mathcal{X},\mathcal{Y})\sim\mathcal{U}}\Big[\Xi^k_{s_1,\dots,s_k,\,t_1,\dots,t_k}(\mathcal{X},\mathcal{Y})\text{ accepts}\Big|\,|\mathcal{X}\cap\mathcal{Y}|=1\Big]$$

$$\geq\frac{1}{45}\cdot\Pr_{(\mathcal{X},\mathcal{Y})\sim\mathcal{U}}\Big[\Xi^k_{s_1,\dots,s_k,\,t_1,\dots,t_k}(\mathcal{X},\mathcal{Y})\text{ accepts}\Big|\mathcal{X}\cap\mathcal{Y}=\emptyset\Big]-2^{-\Omega(l)},$$

and so,

$$\frac{90\cdot\Pr_{(\mathcal{X},\mathcal{Y})\sim\mathcal{U}}\Big[\Xi^k_{s_1,\dots,s_k,\,t_1,\dots,t_k}(\mathcal{X},\mathcal{Y})\text{ accepts}\Big|\mathcal{X}\cap\mathcal{Y}=\emptyset\Big]}{\big(1+\Lambda/5-4\delta_{max}\big)^k}$$

$$\geq\Pr_{(\mathcal{X},\mathcal{Y})\sim\mathcal{U}}\Big[\Xi^k_{s_1,\dots,s_k,\,t_1,\dots,t_k}(\mathcal{X},\mathcal{Y})\text{ accepts}\Big|\mathcal{X}\cap\mathcal{Y}=\emptyset\Big]-2^{-\Omega(l)}.$$

In other words,

$$\Pr_{(\mathcal{X},\mathcal{Y})\sim\mathcal{U}}\Big[\Xi^k_{s_1,\dots,s_k,\,t_1,\dots,t_k}(\mathcal{X},\mathcal{Y})\text{ accepts}\Big|\mathcal{X}\cap\mathcal{Y}=\emptyset\Big]\cdot\left(1-\frac{90}{\big(1+\Lambda/5-4\delta_{max}\big)^k}\right)\leq 2^{-\Omega(l)},$$

and so,

$$\frac{1}{2}\cdot\left(\frac{\Lambda\cdot\mathcal{U}_{\{0,1\}^{n+n}}(A\times B)}{7}\right)^k\cdot\left(1-\frac{90}{\big(1+\Lambda/5-4\delta_{max}\big)^k}\right)\leq 2^{-\Omega(l)}.$$

Assuming $\Lambda>20\cdot\delta_{max}$, we can pick sufficiently large

$$k\in O\left(\frac{1}{\log(1+\Lambda/5-4\delta_{max})}\right)=O\left(\frac{1}{\Lambda-20\cdot\delta_{max}}\right),$$

such that

$$\log\left(\frac{1}{\Lambda}\right)+\log\left(\frac{1}{\mathcal{U}_{\{0,1\}^{n+n}}(A\times B)}\right)\in\Omega\left(\frac{l}{k}\right)=\Omega\Big(l\cdot(\Lambda-20\cdot\delta_{max})\Big).$$

Substituting the definition of l, we get

$$\log\left(\frac{1}{\Lambda}\right)+\log\left(\frac{1}{\mathcal{U}_{\{0,1\}^{n+n}}(A\times B)}\right)\in\Omega\left(\frac{c_2}{c_2-c_1}\cdot(\Lambda-20\cdot\delta_{max})\right).$$

Let

$$d_{max}\stackrel{\text{def}}{=}\max\left\{\left|\frac{n}{2}-a\right|,\ \left|\frac{n}{2}-k\right|\right\},\tag{36}$$

then

$$\Lambda-20\cdot\delta_{max}=\Lambda-20\cdot\frac{\max\{|n-2\cdot a|,\,|n-2\cdot k|\}+2}{n}=\Lambda-40\cdot\frac{d_{max}+1}{n}$$

and the assumption

$$d_{max}\leq\frac{\Lambda\cdot n}{41}\tag{37}$$

38

guarantees that $\Lambda - 20 \cdot \delta_{max} \in \Omega(\Lambda)$ and $1/\Lambda < n$ (note that $d_{max} \geq 1$ is implied by its definition, as $a \neq k$), and therefore,

$$\log\left(\frac{n}{\mathcal{U}_{\{0,1\}^{n+n}}(A \times B)}\right) \in \Omega\left(\frac{c_2 \cdot \Lambda}{c_2 - c_1}\right) \subseteq \Omega\left(\frac{n \cdot \Lambda}{d_{max}}\right). \tag{38}$$

As Λ must either violate (37) or satisfy (38) (or both), we may conclude

$$\Lambda \in O\left(\frac{d_{max}}{n} \cdot \log\left(\frac{n}{\mathcal{U}_{\{0,1\}^{n+n}}(A \times B)}\right)\right). \tag{39}$$

If $\mathrm{rw}(\{k, k+1\}) > 1$, then (28), (29), (36) and (39) imply

$$\lambda_+ \in O\left(\frac{t_+ + \left|\frac{n}{2} - k\right|}{n} \cdot \log\left(\frac{n}{\mathcal{U}_{\{0,1\}^{n+n}}(A \times B)}\right)\right)$$

$$\subseteq O\left(\left(\sqrt{\frac{\log n}{n}} + \left|\frac{1}{2} - \frac{k}{n}\right|\right) \cdot \log\left(\frac{n}{\mathcal{U}_{\{0,1\}^{n+n}}(A \times B)}\right)\right),$$

and if $\mathrm{rw}(\{k, k+1\}) < 1$, then (30), (31), (36) and (39) imply

$$\lambda_- \in O\left(\frac{t_- + \left|\frac{n}{2} - k\right|}{n} \cdot \log\left(\frac{n}{\mathcal{U}_{\{0,1\}^{n+n}}(A \times B)}\right)\right)$$

$$\subseteq O\left(\frac{\sqrt{n \cdot \ln\left(\frac{1}{\lambda_- \cdot \mathcal{U}_{\{0,1\}^{n+n}}(A \times B)}\right)} + \left|\frac{n}{2} - k\right|}{n} \cdot \log\left(\frac{n}{\mathcal{U}_{\{0,1\}^{n+n}}(A \times B)}\right)\right)$$

$$\subseteq O\left(\frac{\left(\log \frac{n}{\mathcal{U}_{\{0,1\}^{n+n}}(A \times B)}\right)^{\frac{3}{2}}}{\sqrt{n}} + \left|\frac{1}{2} - \frac{k}{n}\right| \cdot \log\left(\frac{n}{\mathcal{U}_{\{0,1\}^{n+n}}(A \times B)}\right)\right),$$

as required. Lastly, for (35) to hold it is enough for $d_{max} \leq n/14$ to hold, which is guaranteed by the assumed $\left[\sqrt{n \cdot \mu_{A \times B}}, |n/2 - k| \leq n/14\right]$. ∎ *Theorem 3*

It remains to see how spectral stability of large rectangles implies hardness of *GHR* for classical interactive protocols.

Theorem 4. $\mathcal{R}(GHR) \in \Omega(n^{1/3})$.

Proof of Theorem 4. If there is a randomised protocol of cost c that solves *GHR* with error at most $1/8$, then for any fixed distribution some choice of the random seed guarantees at most the same error. So, a deterministic protocol Υ of cost c solves *GHR* with error at most $1/8$

when $(\mathcal{X}, \mathcal{Y}) \sim \mathcal{U}_{\{0,1\}^{n+n}}$. As there are at most $n^{2\log n}$ different possible answers, Υ partitions $\{0,1\}^{n+n}$ into at most $2^{c+2(\log n)^2}$ rectangles, each marked with some answer value.[24]

Let $V_1 \subseteq \{0,1\}^{n+n}$ be the union of the rectangles of Υ whose size is at least $2^{-c-2(\log n)^2-4}$ (with respect to $\mathcal{U}_{\{0,1\}^{n+n}}$), then $\mathcal{U}_{\{0,1\}^{n+n}}(V_1) \geq 15/16$. Let $V_2 \subseteq \{0,1\}^{n+n}$ be the union of such rectangles $C \times D$ of Υ that satisfy

$$\Pr_{(\mathcal{X},\mathcal{Y}) \subseteq C \times D}\left[\neg \aleph(\mathcal{X}, \mathcal{Y})\right] \leq 16 \cdot \Pr_{(\mathcal{X},\mathcal{Y}) \subseteq \{0,1\}^{n+n}}\left[\neg \aleph(\mathcal{X}, \mathcal{Y})\right],$$

then $\mathcal{U}_{\{0,1\}^{n+n}}(V_2) \geq 15/16$ and $\mathcal{U}_{\{0,1\}^{n+n}}(V_1 \cap V_2) \geq 7/8$. Accordingly,

$$\Pr_{(\mathcal{X},\mathcal{Y}) \sim \mathcal{U}_{\{0,1\}^{n+n}}}\left[\Upsilon(\mathcal{X}, \mathcal{Y}) \text{ is right} \middle| (\mathcal{X}, \mathcal{Y}) \in V_1 \cap V_2\right] \geq 1 - \frac{1}{8}\bigg/\frac{7}{8} = \frac{6}{7}.$$

This implies existence of a rectangle $A \times B \subseteq V_1 \cap V_2$, where, due to Theorem 2,

$$\Pr_{(\mathcal{X},\mathcal{Y}) \subseteq A \times B}\left[\neg \aleph(\mathcal{X}, \mathcal{Y})\right] \in O\left(\frac{(\ln n)^2}{\sqrt{n}}\right).$$

Let us denote by "$((j_1, s_1), \ldots, (j_{\log n}, s_{\log n}))$" the answer of Υ on $A \times B$: we know that it is correct with probability at least $6/7$; by the definition of GHR, this means

$$\frac{6}{7} \leq \Pr_{(\mathcal{X},\mathcal{Y}) \subseteq A \times B}\left[\left|\left\{i \in [\log n] : \Delta_{j_i, s_i}(\mathcal{X}, \mathcal{Y}) \notin \left[\frac{n}{2} - \frac{\sqrt{n}}{2}, \frac{n}{2} + \frac{\sqrt{n}}{2}\right]\right\}\right| \geq \frac{\log n}{2}\right]$$
$$+ \Pr_{(\mathcal{X},\mathcal{Y}) \subseteq A \times B}\left[\neg \aleph(\mathcal{X}, \mathcal{Y})\right].$$

For large enough n,

$$\frac{3}{4} \leq \Pr_{(\mathcal{X},\mathcal{Y}) \subseteq A \times B}\left[\left|\left\{i \in [\log n] : \Delta_{j_i, s_i}(\mathcal{X}, \mathcal{Y}) \notin \left[\frac{n}{2} - \frac{\sqrt{n}}{2}, \frac{n}{2} + \frac{\sqrt{n}}{2}\right]\right\}\right| \geq \frac{\log n}{2}\right]$$
$$\leq \frac{2}{\log n} \cdot \mathbf{E}_{(\mathcal{X},\mathcal{Y}) \subseteq A \times B}\left[\left|\left\{i \in [\log n] : \Delta_{j_i, s_i}(\mathcal{X}, \mathcal{Y}) \notin \left[\frac{n}{2} - \frac{\sqrt{n}}{2}, \frac{n}{2} + \frac{\sqrt{n}}{2}\right]\right\}\right|\right]$$
$$= \frac{2}{\log n} \cdot \sum_{i=1}^{\log n} \Pr_{(\mathcal{X},\mathcal{Y}) \subseteq A \times B}\left[\Delta_{j_i, s_i}(\mathcal{X}, \mathcal{Y}) \notin \left[\frac{n}{2} - \frac{\sqrt{n}}{2}, \frac{n}{2} + \frac{\sqrt{n}}{2}\right]\right],$$

and therefore for some $(j_0, s_0) \in [n] \times \{0,1\}^{\log n}$:

$$\Pr_{(\mathcal{X},\mathcal{Y}) \subseteq A \times B}\left[\left|\sigma_{j_0}\left(\tau_{s_0} \oplus \mathcal{X}\right) \oplus \mathcal{Y}\right| \notin \left[\frac{n}{2} - \frac{\sqrt{n}}{2}, \frac{n}{2} + \frac{\sqrt{n}}{2}\right]\right] \geq \frac{3}{8}.$$

---

[24] As the answer length is poly-logarithmic in the case of GHR, we could assume without loss of generality that the protocol's last message always contains the answer (then a deterministic protocol of cost c would partition the input space into at most 2^c rectangles). We avoid that assumption in order to admit efficient protocols for communication problems with long answers.

Let $A' \stackrel{\text{def}}{=} \{\sigma_{j_0}(\tau_{s_0} \oplus a) | a \in A\}$, then

$$\underset{(\mathcal{X},\mathcal{Y}) \Subset A' \times B}{\boldsymbol{Pr}} \left[|\mathcal{X} \oplus \mathcal{Y}| \in \left[\frac{n}{2} - \frac{\sqrt{n}}{2}, \frac{n}{2} + \frac{\sqrt{n}}{2} \right] \right] \leq \frac{5}{8} \tag{40}$$

and, since $|A'| = |A|$ and $A \times B \subseteq V_1$,

$$\mathcal{U}_{\{0,1\}^{n+n}}(A' \times B) \geq 2^{-c-2(\log n)^2 - 4}. \tag{41}$$

Aiming to benefit from Theorem 3, now we need a lower bound on

$$\underset{(\mathcal{X},\mathcal{Y}) \Subset \{0,1\}^{n+n}}{\boldsymbol{Pr}} \left[|\mathcal{X} \oplus \mathcal{Y}| \in \left[\frac{n}{2} - \frac{\sqrt{n}}{2}, \frac{n}{2} + \frac{\sqrt{n}}{2} - 1 \right] \right]$$

$$= \underset{\mathcal{X} \Subset \{0,1\}^n}{\boldsymbol{Pr}} \left[|\mathcal{X}| \in \left[\frac{n}{2} - \frac{\sqrt{n}}{2}, \frac{n}{2} + \frac{\sqrt{n}}{2} - 1 \right] \right].$$

The bound has to be above $5/8$, and a typical Chernoff-like estimation would lack the required tightness, so we use Lemma 2:

$$\underset{(\mathcal{X},\mathcal{Y}) \Subset \{0,1\}^{n+n}}{\boldsymbol{Pr}} \left[|\mathcal{X} \oplus \mathcal{Y}| \in \left[\frac{n}{2} - \frac{\sqrt{n}}{2}, \frac{n}{2} + \frac{\sqrt{n}}{2} - 1 \right] \right] \geq \sqrt{\frac{2}{\pi}} - \sqrt{\frac{1}{18\pi}} - o(1) > \frac{21}{32}$$

for large enough n. Combining this with (40) and noting that $[n/2 - \sqrt{n}/2, n/2 + \sqrt{n}/2 - 1]$ can be partitioned into pairs $\{n/2 - \sqrt{n}/2, n/2 - \sqrt{n}/2 + 1\}, \dots, \{n/2 + \sqrt{n}/2 - 2, n/2 + \sqrt{n}/2 - 1\}$, we conclude that for some $k \in \{n/2 - \sqrt{n}/2, \dots, n/2 + \sqrt{n}/2 - 2\}$:

$$\frac{\boldsymbol{Pr}_{(\mathcal{X},\mathcal{Y}) \Subset A' \times B}\left[|\mathcal{X} \oplus \mathcal{Y}| \in \{k, k+1\} \right]}{\boldsymbol{Pr}_{(\mathcal{X},\mathcal{Y}) \Subset \{0,1\}^{n+n}}\left[|\mathcal{X} \oplus \mathcal{Y}| \in \{k, k+1\} \right]} < \frac{5}{8} \bigg/ \frac{21}{32} = \frac{20}{21}.$$

In the terminology of the spectral stability statement (Theorem 3) this implies

$$\frac{(\mu_{A' \times B})^{3/2}}{\sqrt{n}} + \left| \frac{1}{2} - \frac{k}{n} \right| \cdot \mu_{A' \times B} \in \Omega(1),$$

that is,

$$\log\left(\frac{n}{\mathcal{U}_{\{0,1\}^{n+n}}(A' \times B)} \right) = \mu_{A' \times B} \in \Omega\left(n^{1/3}\right),$$

and the required $\left[c \in \Omega\left(n^{1/3}\right) \right]$ follows from (41). $\blacksquare$ *Theorem 4*

7 Conclusions

From Corollary 1 and Theorem 4:

Corollary 2.

$$\mathcal{Q}^{\|}(GHR) \in O(\log n) \quad and \quad \mathcal{R}(GHR) \in \Omega\left(n^{1/3}\right).$$

This demonstrates that quantum *SMP*, which is the weakest reasonable quantum model, can outperform classical two-way communication, which is the strongest model of feasible classical communication.

What interesting questions are still worth asking? One possible way to complicate the challenge would be to ask for examples where a quantum model outperforms a substantially reinforced classical one (say, *AM*, where protocols can use both randomness and *non-determinism*). Two possible objections come to mind: First, it seems that most of the known "supermodels" either can be handled by, essentially, the same techniques as $\mathcal{R}$ (and therefore are of limited theoretical interest), or are still beyond the analytical reach (like *AM*) – in the latter case making an excellent subject of investigation besides the context of quantum mechanics. Second, comparing a model that cannot be viewed as classically-feasible to a quantum model would lack the philosophical significance and the appeal of reflecting the *super-classical* nature of quantum mechanics.[25]

We think that the most interesting remaining problem in this area is to analyse the possibility of outperforming classical communication by quantum *SMP* on *more elementary levels* than relational problems. Recall the hierarchy of *problem types* that we mentioned earlier: total functions, partial functions and relations. Obviously, in this context total functions are more elementary than partial functions, while relations are the most sophisticated. Viewed as *tools for demonstrating the advantage of one communication model over another*, total functions are the weakest and relations are the strongest – accordingly, the arguments based on total functions are the hardest to obtain (maybe even impossible in some cases), while the arguments based on relations (like the one presented in this work) are the easiest.

In terms of the *known* super-classical merits of $\mathcal{Q}^{\parallel}$, the three types indeed form a hierarchy:

- A *total function* is known that has an efficient $\mathcal{Q}^{\parallel}$-protocol, but no efficient classical simultaneous-messages protocol without shared randomness [BCWdW01].[26]
- A *partial function* is known that has an efficient $\mathcal{Q}^{\parallel}$-protocol, but no efficient classical simultaneous-messages protocol, even with shared randomness [Gav19].
- Our *relation GHR* has an efficient $\mathcal{Q}^{\parallel}$-protocol, but no efficient classical two-way protocol.

Can the first two positions be improved? Namely:

- Can a quantum simultaneous-messages protocol outperform classical simultaneous-messages protocols with shared randomness over a total function?
- Can a quantum simultaneous-messages protocol outperform classical one-way protocols over a partial function?

Acknowledgements

---

[25] These arguments are rather subjective, and intentionally so. For instance, one could view possible advantage of quantum two-way communication over *AM* as a very natural indirect comparison between the quantum and the classical worlds (as *AM* is known to be stronger than $\mathcal{R}$); on the other hand, comparing $\mathcal{Q}^{\parallel}$ to $\mathcal{R}$ can be viewed as too indirect and too unnatural, as it mixes very different communication layouts.

[26] This is the only known example of qualitative advantage of quantum communication for a total function.

References

[AV79]   D. Angluin and L. Valiant. Fast Probabilistic Algorithms for Hamiltonian Paths and Matchings. *Journal of Computer and System Sciences 18*, pages 155–193, 1979.

[BCMdW10]   H. Buhrman, R. Cleve, S. Massar, and R. de Wolf. Nonlocality and Communication Complexity. *Reviews of Modern Physics 82(1)*, pages 665–698, 2010.

[BCW98]   H. Buhrman, R. Cleve, and A. Wigderson. Quantum vs. Classical Communication and Computation. *Proceedings of the 30th Symposium on Theory of Computing*, pages 63–68, 1998.

[BCWdW01]   H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf. Quantum Fingerprinting. *Physical Review Letters 87(16)*, article 167902, 2001.

[BYJK04]   Z. Bar-Yossef, T. S. Jayram, and I. Kerenidis. Exponential Separation of Quantum and Classical One-Way Communication Complexity. *Proceedings of 36th Symposium on Theory of Computing*, pages 128–137, 2004.

[CR11]   A. Chakrabarti and O. Regev. An Optimal Lower Bound on the Communication Complexity of Gap-Hamming-Distance. *Proceedings of the 43rd Symposium on Theory of Computing*, pages 51–60, 2011.

[Gav08]   D. Gavinsky. Classical Interaction Cannot Replace a Quantum Message. *Proceedings of the 40th Symposium on Theory of Computing*, pages 95–102, 2008.

[Gav19]   D. Gavinsky. Quantum Versus Classical Simultaneity in Communication Complexity. *IEEE Transactions on Information Theory 65(10)*, pages 6466–6483, 2019.

[Gav20]   D. Gavinsky. Entangled Simultaneity Versus Classical Interactivity in Communication Complexity. *IEEE Transactions on Information Theory 66(7)*, pages 4641–4651, 2020.

[GKK+08]   D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf. Exponential Separations for One-Way Quantum Communication Complexity, with Applications to Cryptography. *SIAM Journal on Computing 38(5)*, pages 1695–1708, 2008.

[IW03]   P. Indyk and D. Woodruff. Tight Lower Bounds for the Distinct Elements Problem. *Proceedings of the 44th Annual Symposium on Foundations of Computer Science*, pages 283–289, 2003.

[KN97]   E. Kushilevitz and N. Nisan. Communication Complexity. *Cambridge University Press*, 1997.

[KR11]   B. Klartag and O. Regev. Quantum One-Way Communication Can Be Exponentially Stronger than Classical Communication. *Proceedings of the 43st Symposium on Theory of Computing*, pages 31–40, 2011.

[Pop34]   K. Popper. Logik der Forschung: Zur Erkenntnistheorie der modernen Natur-wissenschaft. *Tübingen: Siebeck*, 1934.

[Raz92]   A. Razborov. On the Distributional Complexity of Disjointness. *Theoretical Computer Science 106(2)*, pages 385–390, 1992.

[Raz99]   R. Raz. Exponential Separation of Quantum and Classical Communication Complexity. *Proceedings of the 31st Symposium on Theory of Computing*, pages 358–367, 1999.

[Sin18]   A. Sinclair. Randomness and Computation. *UC Berkeley*, lecture notes, 2018.