General Randomness Amplification with Non-signaling Security

Kai-Min Chung* Yaoyun Shi† Xiaodi Wu‡

Abstract

Highly unpredictable events appear to be abundant in life. However, when modeled rigorously, their existence in nature is far from evident. In fact, the world can be deterministic while at the same time the predictions of quantum mechanics are consistent with observations. Assuming that randomness does exist but only in a weak form, could highly random events be possible? This fundamental question was first raised by Colbeck and Renner (*Nature Physics*, 8:450–453, 2012). In this work, we answer this question positively, without the various restrictions assumed in the previous works. More precisely, our protocol uses quantum devices, a single weak randomness source quantified by a general notion of non-signaling min-entropy, tolerates a constant amount of device imperfection, and the security is against an all-powerful non-signaling adversary. Unlike the previous works proving non-signaling security, our result does not rely on any structural restrictions or independence assumptions. Thus it implies a stronger interpretation of the dichotomy statement put forward by Gallego et al. (*Nature Communications*, 4:2654, 2013): "[e]ither our world is fully deterministic or there exist in nature events that are fully random."

Note: This is a new work after our QIP 2014 paper, where the security proved is against a quantum, as opposed to non-signaling, adversary.

*Institute of Information Science, Academia Sinica, Taiwan.

†Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48103, USA.
This research was supported in part by NSF Awards 1216729, 1318070, and 1526928.
‡Department of Computer and Information Science, University of Oregon, Eugene, OR 97403, USA.

1 Introduction

In this work, an event is random if it cannot be perfectly predicted. An event is *perfectly* random, if all the alternatives have an equal chance of occurring. A family of events are said to be *truly* random, if in the limit of the indexing parameters, the events tend to be perfectly random. These notions of randomness are *relative*: an event being random to one party may not necessarily mean that it is random to another. For example, the arrival time of a bus may be completely random to a visitor but much more predictable to a frequent rider. We aim to study randomness in the broadest scope and with the fewest restrictions on the observing party.

To an omnipresent and all-powerful observer (referred to as "Adversary" below), the world is much less random than our experiences suggest. For instance, where a spinning roulette will stop can be perfectly predicted by a party who knows the mechanical parameters of the device. In fact, even the very existence of randomness is not clear. The philosophical debate between determinism and indeterminism has a long history. For example, Baruch Spinoza held the extreme view of strict determinism, that every event, including human behavior, follows from previous events deterministically.

Quantum mechanics may seem to have randomness postulated in its axioms: the Born rule of quantum mechanics assigns a probability for a measurement outcome to occur. However, as pointed out by John Bell [9], one cannot rule out what he called "super-determinism." That is, the world is deterministic yet the predictions of quantum mechanics are consistent with observations. Thus for our investigation to be meaningful, we will assume that randomness, at least in the weakest possible form, does exist, relative to Adversary. A fundamental and the central question studied in this work is:

Question 1.1 (Central Question) *Can true randomness arise from weak randomness?*

As a mathematical question, this has been studied by computer scientists for decades. The whole field of "randomness extractor" precisely aims to generate true randomness from weak randomness by applying a deterministic function. Zuckerman [21] realized that the weakest form of randomness is most appropriately quantified by the *min-entropy*, which characterizes the best chance for Adversary to guess the information correctly. A major finding of the theory is a positive answer to the Central Question: true randomness can indeed be generated deterministically from two or more *independent* weak sources of randomness (e.g., [4, 3]). However, a simple yet far reaching observation is that single-source randomness extraction is impossible: there is no deterministic function that transforms a single source of weak randomness into even a single bit of true randomness. We can never be sure if two random variables are independent by only examining the numbers. Thus it appears that the existence of true randomness will have to rest upon the belief in independence.

Colbeck and Renner [8] were the first to ask the Central Question from the perspective of fundamental physics, by imposing physical laws on the devices used for the randomness production and on the adversary. They showed that the law of physics will indeed allow the production of true randomness, even if just a single weak source is available, as long as certain additional, and justifiable, assumptions hold. In particular, they assume the follow two physical laws. [1]

- Quantum Theory (QT): The prediction of quantum mechanics is correct.
- Non-signaling (NS): Without the exchange of information, what happens in a physical system should not affect the observable behavior of another system.

[1] More precisely, they assume QT for the completeness (i.e., the honest devices are quantum mechanical). For soundness (i.e., the property of the protocol not being fooled by dishonest devices), they have a result assuming only NS and another result (with a stronger parameter describing the weak randomness) also assuming QT behavior of the devices.

QT does not necessarily mean that quantum mechanics is complete (which in turn means the only available operations on a physical system are precisely those described by quantum mechanics). NS is satisfied by all widely accepted physical theories, including quantum mechanics. They made an additional assumption about the weak randomness and proved that true randomness can be produced from the single source of weak randomness. The same conclusion was drawn in many subsequent works under different assumptions on the weak randomness and/or Adversary. While these assumptions were crucial for the derivation of the conclusion, there appears no a priori reason for their validity in Nature. The *central question* with the most general form of weak randomness and non-signal security remained open. Our main result answers this question positively and can be informally stated as follows.

Theorem 1.2 *(Main Theorem; informal) Assuming QT and NS, true randomness can be deterministically and certifiably generated from a single source of weak randomness in the most general form as described by min-entropy.*

As articulated by Gallego *et al.* [11], such results imply a dichotomy statement on the existence of randomness in Nature: "[e]ither our world is fully deterministic or there exist in Nature events that are fully random." Our theorem implies a stronger interpretation of this dichotomy statement: the weak randomness can be a general min-entropy source, without any of the (conditional) independence or structural assumptions required in previous works.

Organization

In Section 2, we briefly review necessary background and main concepts for randomness amplification. We introduce a precise model to formally state our main theorem in Section 3, and present our construction for proving the main theorem and sketch the proof in Section 4. We adopt a slightly informal terminology in the first four sections for simplicity.

Starting from Section 5, we provide a mathematically more formal treatment with further detailed discussions. In Section 5, we introduce a set of rigorous notations and terminologies about NS systems. We also introduce protocols, distances, and min-entropy in this language. The model of general NS DI-RA is discussed in Section 6. We describe our main protocol and prove our main result in Section 7. Three main technical challenges are illustrated and resolved in Section 8, 9, 10, respectively.

2 Technical Background

In the previous section, we have described informally the motivation for the general randomness amplification problem and our main result. In this section, we will review the technical framework — the device-independent randomness amplification (DI-RA) framework introduced by Colbeck and Renner [8] — within which randomness amplification is studied as well as the key components for a desirable protocol. We defer formal definitions to the next section.

Let us first consider what it entails for statements like Theorem 1.2 to be true. We need to design a deterministic procedure that operates on some physical system with a single source of weak randomness with sufficient min-entropy (defined appropriately), and generates certifiable true randomness assuming only QT and NS. Note that since we do not make assumptions on the physical system, we allow the procedure to reject if the system does not follow the prescribed behavior, but there should exist an "honest" physical system (which relies on QT) that makes the procedure accept with high probability (QT completeness). Furthermore, we consider an omnipresent and all-powerful

observer (i.e., Adversary), who knows full initial information about the physical system, with the only restriction being that it cannot communicate with the system after the procedure starts (i.e., satisfying the non-signaling condition). Informally, we require that when the procedure accepts, the output should "look uniform" to this Adversary (NS security). Note that we do not want to assume quantum completeness, so the physical system and Adversary are not restricted to be quantum. Proving Theorem 1.2 amounts to designing such a procedure with QT completeness and NS soundness.

In the DI-RA framework, we consider a physical system that consists of a weak source X with a sufficient min-entropy and a set of devices D, which can be thought of as black-boxes with a classical interface. We assume that the devices do not communicate among themselves, i.e., the NS condition holds. Other than that, we do not make any assumptions on the inner working of the devices.

Protocol Π on the source-device system is a deterministic procedure, each step of which specifies a device and its input, based on the source X and the device outputs in the previous steps. Π outputs a *decision bit* $O \in \{\mathsf{Acc}/\mathsf{Rej}\}$ and one or more *output bits* Z. QT completeness requires the existence of quantum devices that make Π accept with high probability. NS-security requires that when Π accepts, the output Z should "look close to uniform" to the above-mentioned all-powerful but non-signaling Adversary, which is modeled as an extra component in the physical system. An illustration of a source-device system and a protocol is in Fig. (1).

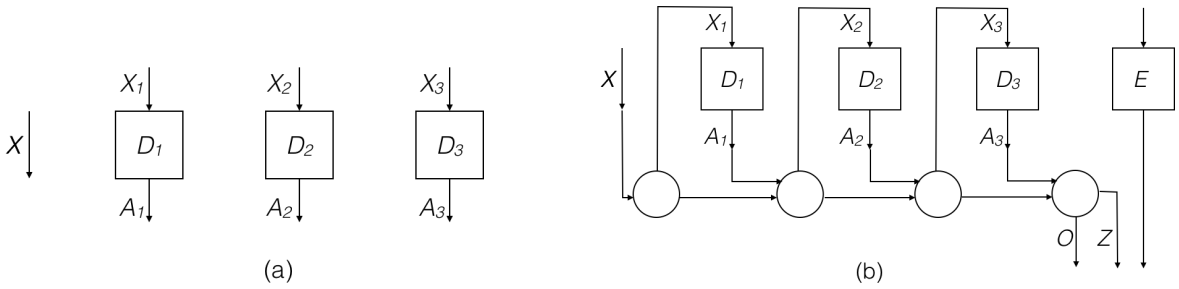

Figure 1: (a) illustrates an example of a weak source X and a set of 3 non-signaling devices $D = (D_1, D_2, D_3)$. Each device D_i can accept an input X_i and output A_i, $i = 1, 2, 3$. That the devices together with X are non-signaling means that all marginals such as $XA_1|X_1$, $A_2A_3|X_2X_3$ are well defined. That X is of min-entropy k to D means that no DI protocol making use of D can output X with $> 2^{-k}$ probability. (b) illustrates a DI protocol making use of XD in (a). The Adversary E is also in non-signaling relation with XD. Each circle is a deterministic function. The protocol has a completeness error ϵ_c if $\mathbf{Pr}[O = \mathsf{Acc}] \geq 1 - \epsilon_c$. It has a soundness error ϵ_s if OZE is ϵ_s-close to $\overline{OZE}$, which is obtained from OZE by replacing Z with uniform output when $O = \mathsf{Acc}$. In general, a DI protocol can be adaptive (by using previous device outputs) in setting device inputs; our protocol is not.

The key idea in DI-RA protocols, first proposed by Colbeck in his Thesis [7], is to certify randomness via Bell violation of non-local games. A non-local game is a one-round game played between a referee R and a set of non-communicating devices, where the referee sends challenges to and receives answers from each device, and decides whether the devices win the game based on the transcript. For example, in the CHSH game [6], R sends uniform challenge bits $x, y \in_R \{0, 1\}$ to two devices, receives two bits $a, b \in \{0, 1\}$ from each device, and accepts (i.e., the devices win) if $a \oplus b = x \wedge y$. The devices are restricted to not communicate, but the device strategy can be classical, quantum, or non-signaling.

We can define the classical and quantum value of a game as the maximum winning probability of optimal classical and quantum devices, respectively. Bell violation refers to non-local games where the quantum value is strictly higher than the classical value. For example, the classical and quantum values of the CHSH game are 75% and $\approx 85\%$, respectively.

A simple yet far-reaching observation, first made by Colbeck [7], is that all deterministic devices are classical, and thus any device strategy with winning probability higher than the classical value must be randomized. This gives an approach to certify randomness in the devices' output by statistically witnessing Bell violation of non-local games. Namely, we can play a game multiple times and compute the empirical winning probability. If it is significantly higher than the classical value (yet achievable by quantum devices), then we can be confident that the devices' outputs should contain certain randomness. Indeed, this is the common paradigm in all existing DI-RA protocols.

There are several issues in turning the above idea into a DI-RA protocol. In particular, if we use the CHSH game, we need independent and uniform input bits $x, y \in_R \{0,1\}$ to certify randomness. The problem is that we have only weak randomness to use. To make Bell violation useful, Colbeck and Renner [8], as well as subsequent authors, model the weak source as a Santha-Vazirani (SV) source, where each bit has entropy even conditioned on the value of previous bits. Namely, X is a δ-SV source if $(1/2 - \delta) \leq \Pr[X_i = x_i | X_1 = x_1, \ldots, X_{i-1} = X_{i-1}] \leq (1/2 + \delta)$ for every $x_1, \ldots, x_i \in \{0,1\}$. They design non-local games where even when the input is sampled from a δ-SV source for sufficiently small δ, all deterministic/classical devices has a bounded winning probability. By the same argument, this bound, when violated, can be used to certify randomness using any SV source. However, for this to work, it is crucial to assume that the SV source is *independent* of the devices. As a consequence, certain independence assumptions are required in most DI-RA protocols.

Another issue is, we have only argued that Bell violation certifies randomness in the devices' outputs, but the output may not be close to uniform. Furthermore, the outputs may even be known to Adversary, e.g., due to the entanglement between the devices and Adversary. Therefore, to obtain a DI-RA protocol, we need to show that (i) Bell violation in fact certifies that the devices' outputs "look random" to Adversary, and (ii) we can extract close to uniform randomness from these output bits. In the quantum setting (i.e., assuming the devices and Adversary are quantum), (i) follows by monogamy of entanglement: in order to achieve high winning probability, the devices need to share strong entanglement, which by monogamy implies that the devices cannot have strong correlation with Adversary and their outputs must "look random" to Adversary. Fortunately, such phenomenon generalizes to the NS setting for some non-local games (e.g., in [15]). For (ii), in the quantum setting, one can use quantum-proof strong randomness extractor when a seed is available (e.g., in [10]). In the NS setting, unfortunately, we show that general "NS-proof" strong randomness extractors do not exist. For the existing DI-RA protocols, different techniques are used to circumvent this impossibility by exploiting the structure of non-local games and/or adding additional independence assumptions.

3 Main Result

The objective of this section is to state our main result precisely. To that end, we will introduce the necessary definitions and concepts, which will also be necessary for the next section. As we shall see, our construction relies on composition of DI-RA sub-protocols. A formal model is important to analyze the composition rigorously. We give a concise presentation here and defer detailed discussion to Sections 5 and 6.

We start by defining non-signaling devices. A collection P of n devices is a family of probability

distributions $\{P_{a_1,...,a_n|x_1,...,x_n}\}_{x_1,...,x_n}$. The pair (x_i, a_i) are said to be the input and output of the device i, respectively, and are drawn from their corresponding alphabets. P is said to be non-signaling if for any subset $\{i_1, i_2, ..., i_k\} \subseteq [n]$ of the devices, the marginal $P_{a_{i_1},...,a_{i_k}|x_{i_1},...,x_{i_k}}$ is well defined (that is, invariant for all input combinations for other devices). While this only models "one-time" devices that take a single input and produce a single output, it is sufficient to describe our protocol.

A classical source can be considered as a device with no input, or equivalently, a fixed input denoted by $\perp$. Thus, a physical system P is simply a collection of non-signaling devices. Once a device D_i in P receives input x_i^* and produces a_i^*, it is turned into a classical source. Namely, the system P is turned from $\{P_{a_1,...,a_n|x_1,...,x_n}\}_{x_1,...,x_n}$ to $\{P_{a_1,...,(x_i^*,a_i^*),...,a_n|x_1,...,\perp,...,x_n}\}_{x_1,...,\perp,...,x_n}$. A physical system P may consists of multiple components, each of which contains multiple devices. We simply denote it as $P = C_1C_2 \dots C_t$, and use $C_{i_1}C_{i_2} \dots C_{i_k}$ to denote the corresponding sub-system (which is well defined by the non-signaling condition).

A non-signaling protocol Π using a physical system P is a deterministic procedure, each step of which specifies a device and its input using the device outputs in the previous steps. At the end, Π outputs a *decision bit* $O \in \{\mathsf{Acc}/\mathsf{Rej}\}$ and may output one or more *output bits* Z. For our purpose, we can assume that Z is only a *single* bit similarly as previous results (e.g., [8]) as it already demonstrates the true randomness. Moreover, our protocol can be readily extended to generate more output bits. Thus, we shall focus on the single output bit case in the rest of the paper for clarity. We can now use the language built so far to define the notions of NS distance, min-entropy, and properties of a DI-RA protocol in an operational way.

We first define NS distance, which generalizes trace distance of quantum states to physical systems. The *NS distance* between two physical systems P_1 and P_2 is $\Delta(P_1, P_2) := \sup_\Pi |\Pi(P_1) - \Pi(P_2)|$, where the supremum is over all non-signaling protocols Π and $\Pi(P_i)$, $i = 1, 2$, is the probability of accepting on device P_i, i.e., $\Pi(P_i) := \Pr[\Pi \text{ accepts when using } P_i]$.

Consider a classical source X and a component D in a physical system P. We say X has at least k bits *min-entropy-to-D*, denoted by $\mathrm{H}_\infty(X|D) \geq k$, if for any non-signaling protocol Π using D (but without access to X), Π outputs X with $\leq 2^{-k}$ probability. Note that the definition generalizes the operational meaning of quantum min-entropy as the guessing probability. We say that X is *uniform-to-D* if X is uniform and independent to D, i.e., $XD \equiv U \otimes P$, where U denotes the uniform distribution (and of the same length as X). Finally, we say X is *ϵ-close-to-uniform-to-D*, or *ϵ-uniform-to-D for short*, if $\Delta(XP, U \otimes P) \leq \epsilon$. As we shall see, the terminology "min-entropy/uniform-to-D" is convenient to reason about composition of protocols.

We next define the syntax and the desired properties of DI-RA protocols. The physical system P for a DI-RA protocol consists of three components $P = XDE$, where X is a classical *Source*, $D = (D_1, ..., D_\ell)$ is a set of *Protocol Devices*, and E is the *Adversary*. We say X is an (n, k) source-to-devices if X has n bits and at least k bits min-entropy-to-D. W.l.o.g., we can model that $E = WD_E$ consists of a classical source W describing information about the Source-Device components, and a device D_E that captures non-signaling correlation between the Devices and Adversary.[2]

A *DI-RA protocol* Π is a non-signaling protocol that applies to the Source-Device sub-system. We define the desired properties of DI-RA protocols below. For the analysis of protocol composition, it is important to identify the conditions under which these desirable properties hold. We do not specify these conditions when defining the properties, but will be explicit about the conditions when Π satisfies the properties.

[2]Technically, we can model E as a single device D_E since W can be absorbed in D_E. Nevertheless, we model W explicitly for its semantic meaning.

- *QT completeness.* There exist "honest" quantum devices such that Π accepts with probability $\geq 1 - \epsilon_c$. The parameter ϵ_c is called the *completeness error*.

- *Robustness.* This is a strengthening of completeness which requires Π to accept with probability $\geq 1 - \epsilon_c$ even when the devices have up to η level of noise from the honest devices (i.e., winning the non-local games with a probability that is at most η below that of the optimal quantum device). The parameter η is the tolerated *noise level*. We assume for simplicity that the noises among different sets of devices are independent, though our analysis can be extended to more general settings.

- *NS soundness (i.e., security).* Informally, we require that when Π accepts, the output Z is close-to-uniform-to-E. This is captured by comparing the output system OZE (i.e., decision bit, output bits, and Adversary) with an "ideal" output system $\overline{OZE}$, which is obtained from OZE by replacing Z with the uniform bits when $O = \mathsf{Acc}$ (this is well defined since both O and Z are classical sources). The soundness requires that $\Delta(OZE, \overline{OZE}) \leq \epsilon_s$, where ϵ_s is called the *soundness error*.

We are now able to state our main theorem precisely. We stress that we do not assume any structural or independence conditions on the source, and the source is weakly random only to Protocol Devices and can be deterministic to Adversary, a strong feature also achieved in our previous quantum security result [5].

Theorem 3.1 *(Main Theorem; formally in Theorem 7.7) For any $\epsilon > 0$, any $n \geq k \geq \mathrm{poly}(1/\epsilon)$, there exists a DI-RA protocol Π that achieves robustness and NS soundness when the Source is an (n, k) source-to-devices, with completeness and soundness errors $\epsilon_c = \epsilon_s = \epsilon$ and a constant noise level $\eta \approx 1.748\%$. Π uses $2^{\mathrm{poly}(1/\epsilon)}$ devices.*

Related works. Colbeck and Renner [8] model X as a *Santha-Vazirani* source, i.e., for some $\delta \in (0, 1/2)$ and all i, $1 \leq i \leq n$, $\Pr[X_i = x_i | X_{1,2...,i-1} = x_{1,2,...,i-1}] \in [1/2 - \delta, 1/2 + \delta]$. They show the first non-trivial DI-RA protocol for such a source when $\delta < 0.058$. Adversary is modeled as holding classical information W on the Source-Device system with no additional NS correlation between Devices and Adversary. All subsequent authors proving NS security (such as [11, 17, 12, 18, 13, 2, 19, 20]) continue restricting X to SV sources while improving on various aspects such as the bias, robustness, the number of devices, and allowing certain NS Device-Adversary correlations. These correlations, however, all require some conditional independence condition(s). More specifically, Gallego *et al.* [11] require the SV source X to be independent of Devices conditioned on W. In [18, 2], X is in addition assumed independent to both Devices and Adversary. The difference is subtle but significant. In particular, the security of [18, 2] does not hold when Adversary knows X. Consequently, we can only use the approach of [11] in our composition construction. Under stronger conditional independence assumptions in [2, 13, 19], the authors reduce the number of devices to a constant from $\mathrm{poly}(1/\epsilon_s)$ in [11]. Wojewódka *et al.* [20] relax the independence assumption, but can only handle SV sources with a constant bias.

Additional assumptions weaken the interpretation of the dichotomy statement. For example, the existing results do not hold when SV source is replaced by the physically more plausible SV *block* source, where entropy is only guaranteed for each block of multiple bits. Conditional independence assumptions are undesirable, as they seem not justified by physical laws, and it seems impossible to perfectly certify independence among general distributions. In contrast, our protocol requires only NS and local min-entropy against Devices for soundness, and QT for completeness.

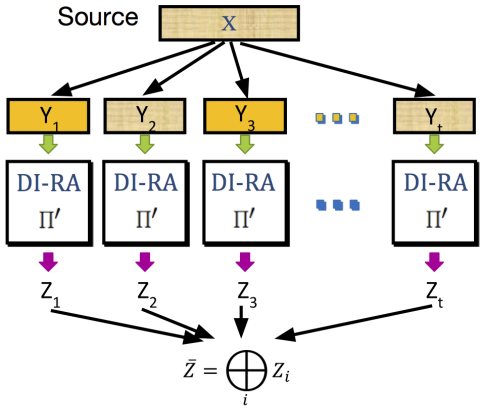

Figure 2: The Source X is used as the input to t instances of a classical seeded randomness extractor $\mathrm{Ext}(X, i)$, each instance corresponds to a value of the seed i, for all possible seed values. One of the outputs Y_{i^*} is ϵ-uniform to its corresponding DI-RA sub-protocol. Such a sub-protocol outputs a globally random output, decoupling the correlations among Y_{i^*}'s. The final output is the XOR of sub-protocol outputs.

4 Our Construction

In this section, we prove our main theorem by constructing a DI-RA protocol for a general weak source. As mentioned above, all existing NS-secure DI-RA protocols directly feed the Source X as challenges to the Devices D for some non-local games, which is the reason that they require X to be a Santha-Vazirani source and independent of D (conditioned on Adversary's classical information W). We circumvent these requirements by the composition approach in our quantum-secure DI-RA protocol [5], where we first apply a classical pre-processing to *improve the quality* of the source. In fact, our construction has the same framework and intuition as [5], but establishing NS security is much more challenging and requires very different techniques. We first review the intuition and then discuss the challenges and our solutions.

As shown in Figure 2, the first step of our protocol Π is a classical pre-processing that turns X into multiple blocks $Y = (Y_1, \ldots, Y_t)$, where some "good" block Y_{i^*} is close to uniform to the devices. We do not know which blocks are "good," thus refer to Y a "somewhere" (approximately) uniform source. Note that we have *no* guarantee that any of the blocks Y_i is *independent* of the devices. Moreover, Y_i may correlated with other blocks and may even be known to Adversary. In the second step, we feed each Y_i to a DI-RA sub-protocol Π' with distinct set of devices, where Π' only needs to "work" on the "good" (i.e., close-to-uniform) block , as opposed to a general weak source. Finally, if we accept (we will discuss the acceptance condition later), we output $Z = \bigoplus_i Z_i$ as final output.

The key intuition here is that Π' is not used to amplify randomness (since the source is already close to uniform), but is used as a *decoupler* to lift "*local uniformity*" to "*global uniformity.*" More precisely, a "good" block Y_{i^*} is locally close-to-uniform to devices but may be correlated with other blocks. The role of Π' is to output Z_{i^*} that is close-to-uniform to both Adversary and the remaining outputs Z_{-i^*}, i.e., decouple the correlation among blocks. This can be inferred by viewing all remaining sub-protocols as part of Adversary for $\Pi'(Y_{i^*})$ and is sufficient to imply that the XOR-ed final output Z is close to uniform to Adversary.

In the following subsections, we discuss these steps in details, and state formally in lemmas what is achieved by each step. The proofs are deferred to later sections. We present the formal construction of our protocol in Section 4.3.

4.1 Obtaining Somewhere Uniform-to-device Source

The goal of the first step is to turn X into a somewhere uniform source $Y = (Y_1, \ldots, Y_t)$ where some "good" block Y_{i^*} is close-to-uniform to the devices. In the quantum setting, this can be done by applying a quantum-proof strong randomness extractor with all possible seeds, i.e., $Y_i = \mathrm{Ext}(X, i)$. It is not hard to show that in this context some Y_{i^*} is close to uniform to all devices D. A natural approach here is to consider the non-signaling analogue. However, as mentioned and presented in Appendix A, we show that "NS-proof" strong randomness extractor does *not* exist.

Nevertheless, we observe that using quantum-proof extractors is an overkill in the sense that it yields a somewhere uniform source Y where the uniform block Y_{i^*} is uniform to *all* devices. For the composition to work, it suffices that Y_{i^*} is only uniform to the corresponding set of devices. For this weaker goal, we show that in fact enumerating seeds for a *classical* strong randomness extractor suffices, even though at the expense of an exponential loss in parameters. More precisely, if $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a classical (k, ϵ)-strong randomness extractor, and $Y_i = \mathrm{Ext}(X, i)$, then Y is somewhere uniform in the sense that some Y_{i^*} is $(2^m \cdot \epsilon)$-close to uniform to the corresponding set of devices.

Note that the exponential loss in error is affordable by setting ϵ sufficiently small. On the downside, this increases the seed length of the extractor, which is the reason that we need $2^{\mathrm{poly}(1/\epsilon_s)}$ devices to achieve soundness error ϵ_s,[3] whereas our quantum-secure protocol [5] only needs $\mathrm{poly}(1/\epsilon_s)$ devices. We do not know if the exponential 2^m loss is necessary, and leave it as an interesting open question.

In the quantum setting, we can let Π accept only when all Π' accept. However, in our case, due to the exponential number of sub-protocols and the implicit dependency between the completeness error and the source length m for Π', we cannot expect that all Π' accept. To handle this, we need to relax the acceptance condition to accept when sufficiently many sub-protocols accept. However, adversarial devices may choose to fail on the "good" blocks, which hurt soundness. To withstand such an attack, we strengthen the requirement on Y to be close to uniform for *most* blocks. This readily follows by adjusting the parameters in the analysis.

We are now ready to state the definition of strengthened somewhere uniform-to-device sources, and the statement achieved by the first step of the protocol. We defer the proof to Section 8 but mention that this is achieved by a *post-selection* argument.

Definition 4.1 *Let P be a physical system with t classical sources $Y = (Y_1, \ldots, Y_t)$ and t device components $D_1, \ldots, D_t$ (each may contain multiple devices). We say that Y is a* $(1-\gamma)$-somewhere uniform-to-device source *if there exist at least $(1-\gamma)$-fraction of $i \in [t]$ such that Y_i is uniform-to-D_i. Similarly, we say that Y is a* $(1-\gamma)$-somewhere ϵ-uniform-to-device source *if there exist at least $(1-\gamma)$-fraction of $i \in [t]$ such that Y_i is ϵ-uniform-to-D_i.*

Theorem 4.2 *Let $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ be a classical (k, ϵ)-strong seeded extractor, and P be a source-device physical system with a classical source X and 2^d device components $D = (D_1, \ldots, D_{2^d})$. For every $i \in [2^d]$, let $Y_i = \mathrm{Ext}(X, i)$. Let $Y = (Y_1, \ldots, Y_{2^d})$. If X has at least k bits of min-entropy-to-D, then Y is a $(1-\gamma)$-somewhere $(2^m \cdot \epsilon/\gamma)$-uniform-to-device source.*

[3] More precisely, the number of blocks is 2^d, where the seed length $d = \Omega(\log 1/\epsilon)$. On the other hand, to achieve soundness error ϵ_s, we need the output length $m = \mathrm{poly}(1/\epsilon_s)$. As a result, we need $2^{\mathrm{poly}(1/\epsilon_s)}$ devices.

4.2 The decoupling sub-protocol Π'

In this step, we need a DI-RA protocol for the sources that are ϵ-*uniform-to-devices*. Note that compared to DI-RA protocols for an SV source, the source being uniform makes the task easier. However, the following two issues in our context make the task more challenging. (i) The source is only uniform-to-devices but not uniform to Adversary (and may be known to Adversary), and (ii) the source is only ϵ-uniform and may not be independent to devices. We discuss the issues in turn.

First issue: local uniformity. We observe that (only) the DI-RA protocol of Gallego *et al.* [11] works without assuming independence between the SV source and Adversary (but requires *independence* between the source and the device conditioned on Adversary's classical information W). Thus, we choose to follow the approach of Gallego *et al.* [11]. Taking the advantage of the source being uniform, we improve the construction to achieve *robustness* by using the bipartite BHK game [1] (see further discussion below). We state the achieved statement as the following lemma with brief further discussion. We defer the proof to Section 9.

Lemma 4.3 *For any $\epsilon > 0$ and $k \geq \text{poly}(1/\epsilon)$, there exists a DI-RA protocol Π that achieves robustness and NS soundness when Source is uniform-to-Device and of length at least k, with completeness and soundness errors $\epsilon_c = \epsilon_s = \epsilon$ and a constant noise level $\eta \approx 1.748\%$. Π uses $\text{poly}(1/\epsilon)$ number of devices.*

Note that Lemma 4.3 is non-trivial since while Source is uniform to Devices, it may be known to Adversary, but the output is required to be uniform-to-Adversary. In fact, to our best knowledge, the only technique that achieves this goal is in [11] (which in turn is based on [15]) at the cost of requiring $\text{poly}(1/\epsilon)$ number of devices (even for uniform-to-Device Sources).[4]

At a high level, the protocol also proceeds by statistically witnessing Bell violation of certain non-local games. However, it is important to use a distinct set of devices for different rounds of the game, and to use disjoint rounds for testing (i.e., witnessing Bell violation) and generating output (referred to as *testing* rounds and *output* rounds, respectively). The reason is subtle: the analysis crucially requires that the protocol decision (which depends only on the testing rounds) and the output round devices are *non-signaling*. This is the reason that we need $\text{poly}(1/\epsilon)$ number of devices.

To handle SV sources, Gallego *et al.* [11] uses a five-party Mermin game, which has quantum value 1 (i.e., there exist quantum devices win the game with probability 1) but for challenges sampled from any SV sources, all classical devices have winning probability bounded away from 1 (but approaches 1 when the bias of SV source decreases). Thus, any SV sources can be used to certify randomness. However, this prevents the protocol to achieve robustness since the protocol can only accept when the devices win all the Mermin games. In contrast, by taking the advantage of the uniform source, we can use the bipartite BHK game originally used in [15] and tolerate a constant level of noise.

Finally, we mention that Gallego *et al.* [11] rely on a non-constructive hash function to extract uniform bits from devices' output. We make the construction explicit by selecting the hash function from t-wise independent hash functions using the uniform source.

[4] For example, [8] assumes that Adversary holds only classical information W, and the works of [18, 2] assume that Source X is independent to both Devices D and Adversary E conditional independence Adversary's classical information W. See Section **??** for further discussion about related works.

Second issue: handling "ϵ-error." The second issue turns out to cause the most challenging technical difficulty in this work. While there is only an "ϵ-error", it breaks the independence between the source and the devices, which is crucial for using non-local games to certify randomness. Also, simple union bound/triangle inequality does not apply since the source is ϵ-uniform-to-device when restricted to the Source-Device sub-system, but not in the global physical system.

In the quantum setting, this issue can be handled generically by a standard "fidelity trick" to bridge the local and global distance: if any two local states A, B are ϵ-close, then for any global state A' of A, one can always find a nearby global state B' of B such that A' and B' are also ϵ'-close when measured in fidelity. Applying the fidelity trick, we can first argue that if locally the source is ϵ-uniform to the devices, then globally the system is $\sqrt{\epsilon}$-close to a system where the source is uniform to the device (the square root loss is due to a conversion from the fidelity to the trace distance). Then by the triangle inequality, if Π' works for uniform-to-device sources, then Π' works for ϵ-uniform sources with a $\sqrt{\epsilon}$ additional error at most.

In the NS setting, the "fidelity trick" does not apply directly due to the lack of quantum structure. We do not know if there is an analogous general statement bridging the local and global distance. Instead, we handle the imperfect source issue by a non-black-box approach, exploiting the structure of our protocol. We state the achieved statement as the following lemma and sketch the main proof idea. We defer the formal proof to Section 10.

Lemma 4.4 *For any $\epsilon > 0$ and $k \geq \mathrm{poly}(1/\epsilon)$, there exists a DI-RA protocol Π that achieves robustness and NS soundness when the Source is $\tilde{\epsilon}$-close to uniform-to-devices and of length at least k for sufficiently small $\tilde{\epsilon} \leq \mathrm{poly}(\epsilon)$. The completeness and soundness errors are $\epsilon_c = \epsilon_s = \epsilon$ and the noise level is a constant $\eta \approx 1.748\%$. Π uses $\mathrm{poly}(1/\epsilon)$ devices.*

To prove Lemma 4.4, we identify a key property in the proof of Lemma 4.3 that is sufficient to complete the analysis. To state the property, we note that the "quality" of the output bit Z is determined by the "quality" of the devices in the output rounds (denoted by D_{out}). In particular, if D_{out} is "bad", then Z is far from uniform. Informally, the property states that the probability that Π accepts but D_{out} is "bad" is small.

Then the most technically involved step is to show that if this property (with slightly weaker parameters) fails, i.e., the probability that Π accepts but D_{out} is "bad" becomes too large, then the source is $\tilde{\epsilon}$-far from uniform-to-devices, i.e., $\Delta(XD, U \otimes D) > \tilde{\epsilon}$. To do so, we need to construct a distinguisher protocol $\tilde{\Pi}$ that distinguish XD from $U \otimes D$ with advantage $\tilde{\epsilon}$. It may seem that the failure of the property directly gives such a distinguisher using the protocol Π. However, whether D_{out} is "bad" or not depends on its device strategy (i.e., as a function of $\{P_{a_1,\ldots,a_m|x_1,\ldots,x_m}\}_{x_1,\ldots,x_m}$ of D_{out}), which cannot be directly observed from the execution of Π (which only generates one sample). In our proof, we construct the reduction distinguisher $\tilde{\Pi}$ by properly modifying Π so that we can "indirectly observe" whether D_{out} is "bad" or not, and show that this is sufficient to distinguish XD from $U \otimes D$.

Remark on completeness. Finally, note that some blocks Y_i can be far from uniform (in fact, even deterministic). We need to make sure that the honest devices for Π' will not be rejected with higher probability when receiving bad inputs, as otherwise we lose completeness. Therefore, we further require Π' to have a small completeness error for any source X. Fortunately, this follows by the property of the BHK game, where the honest device strategy wins with the same (high) probability for every fixed input. We state this fact in the following lemma.

General DI-RA protocol Π

- Let $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ be a classical (k, ϵ_0)-strong seeded randomness extractor.
- Let Π' be a DI-RA protocol that has source length m, completeness and soundness error ϵ', and uses t' devices.
- Π operates on an input classical source X over $\{0,1\}^n$ and $t = 2^d \cdot t'$ devices $D = (D_1, \ldots, D_{2^d})$, where each D_i denotes a set of t' devices, as follows.

Protocol Π

1. For every $i \in [2^d]$, let $Y_i = \mathrm{Ext}(X, i)$. Let $Y = (Y_1, \ldots, Y_{2^d})$.
2. For every $i \in [2^d]$, invoke Π' on the subsystem Y_i, D_i and obtain O_i, Z_i as the output.
3. Let $\eta = 2 \cdot \epsilon'$ (a threshold). If less than η-fraction of Π' rejects (i.e., $O_i = \mathsf{Rej}$), then Π accepts and outputs $O = \mathsf{Acc}$ and $Z = \bigoplus_{i \in [2^d]} Z_i$. Otherwise, Π rejects, i.e., outputs $O = \mathsf{Rej}$ and $Z = \perp$.

Figure 3: Our Main Construction of the General DI-RA Protocol Π.

Lemma 4.5 *The protocol Π in Lemma 4.4 additionally achieves robustness for arbitrary Source, with the same completeness error and noise level.*

4.3 Putting Things Together

We are now ready to present the formal construction of our general DI-RA protocol stated in Theorem 3.1, as shown in Figure 3. The first step of Π uses a classical strong randomness extractor to turn the source X to a $(1-\gamma)$-somewhere $\tilde{\epsilon}$-uniform-to-device source Y for sufficiently small γ and $\tilde{\epsilon}$ by Theorem 4.2. Then Π invokes the decoupler protocol Π' from Lemma 4.4 and 4.5 for each block Y_i using a distinct set of devices D_i, and obtains O_i and Z_i as the output. Π accepts and output $Z = \bigoplus_{i \in [2^d]} Z_i$ if the fraction of rejection is below threshold $\eta = 2\epsilon'$.

We defer the formal analysis of our protocol to Section 7, but provide a proof sketch here. The analysis essentially follows the aforementioned intuition to use the decoupler Π' to lift "locally uniform" good block Y_{i^*} to "globally uniform" Z_{i^*}, which implies that the final output Z is close to uniform-to-Adversary. The main complication is handling errors from the threshold decision, where we need to argue that the allowed at most η-fraction of rejections do not hurt soundness. Intuitively, this follows by the fact that most $(1-\gamma)$-fraction of blocks in Y are good, so some good block Y_{i^*} will be accepted to produce globally uniform Z_{i^*}. We sketch how to formalize this intuition below.

Set $\gamma = \epsilon' = \epsilon/4$, and set the remaining parameters according to Theorem 4.2 and Lemma 4.4 and 4.5. It is not hard to see that Π inherits the robustness of Π' since we set the threshold $\eta = 2\epsilon'$. To prove the soundness, we need to compare the output system OZE with the ideal output system $\overline{OZE}$ which replaces Z with the uniform bits when $O = \mathsf{Acc}$. Note that when $O = \mathsf{Rej}$, OZE and $\overline{OZE}$ are identical so the distinguishing error is 0.

Consider a good block Y_{i^*}. The soundness of $\Pi'(Y_{i^*}, D_{i^*})$ implies that $\Delta(O_{i^*} Z_{i^*} E_{i^*}, \overline{O_{i^*} Z_{i^*} E_{i^*}}) \leq$

ϵ', where E_{i^*} is Adversary for $\Pi'(Y_{i^*}, D_{i^*})$, which includes the remaining components of the system. In particular, it includes O, Z_{-i^*} and E. Thus, $\Delta(O_{i^*}Z_{i^*}OZ_{-i^*}E, \overline{O_{i^*}Z_{i^*}OZ_{-i^*}E}) \leq \epsilon'$, which implies $\Delta(O_{i^*}OZE, \overline{O_{i^*}OZE}) \leq \epsilon'$ since $Z = \bigoplus_{i\in[2^d]} Z_i$ and post-processing can only decrease the distance. Recall that $\overline{O_{i^*}OZE}$ is obtained by replacing Z_{i^*} with a uniform bit when $O_{i^*} = \mathsf{Acc}$, it (informally) means that Z is close to uniform when $\Pi'(Y_{i^*}, D_{i^*})$ accepts.

For intuition, note that if we additionally assume that for some good block Y_{i^*}, $O = \mathsf{Acc}$ implies $O_{i^*} = \mathsf{Acc}$ (i.e., $\Pi'(Y_{i^*}, D_{i^*})$ accepts whenever Π accepts), then the above statement implies Z is close to uniform when Π accepts, i.e., $\Delta(OZE, \overline{OZE}) \leq \epsilon'$, where $\overline{OZE}$ is the ideal output system for Π (instead of $\Pi'(Y_{i^*}, D_{i^*})$). In other words, with the additional assumption, the soundness of $\Pi'(Y_{i^*}, D_{i^*})$ implies the soundness of Π.

Finally, we can remove the assumption at the cost of small additive error using the facts that there are $(1-\gamma)$-fraction of good blocks, and Π accepting implies at least $(1-\eta)$-fraction of Π' accept. By an averaging argument, there exists a good block Y_{i^*} such that $\Pr[O = \mathsf{Acc} \wedge O_{i^*} = \mathsf{Rej}] \leq \gamma + \eta$. Combined with the above argument, we can show that the soundness error of Π is at most $\gamma+\eta+\epsilon' \leq \epsilon$.

5 Preliminaries of Non-signaling (NS) Systems

We will leverage vectors in $\mathbb{R}^n$ to represent NS systems. Here we summarize several operations (and notations) on these vectors. For example, let $v = (v_1, \cdots, v_n) \in \mathbb{R}^n$ and $w = (w_1, \cdots, w_n)$ be two vectors. Let $[n] = \{1, \cdots, n\}$.

- $v \otimes w$: is the tensor product of v, w, which leads to a vector $\mathbb{R}^{nm}$. We can use (i, j) where $i \in [n], j \in [m]$ to label the entree of $v \otimes w$. Hence, we have $(v \otimes w)_{(i,j)} = v_i w_j$ for all (i, j).
- $\times$: for multiplication between scalars, or a scalar's multiplication with a vector. Might be omitted when it is clear from the context, e.g.,

$$\lambda v = (\lambda v_1, \cdots, \lambda v_n) \in \mathbb{R}^n.$$

- $v \cdot w$: is the inner-product of v, w when $n = m$. Namely,

$$v \cdot w = \sum_{i=1}^{n} v_i w_i.$$

- $v \preceq w$: (when $n = m$) means $v_i \leq w_i$ for $i \in [n]$.
- $|v|$: means a new vector whose entrees are the absolute values of the entrees of v. Namely,

$$|v| = (|v_1|, \cdots, |v_n|) \in \mathbb{R}^n.$$

For convenience, we also adopt the following conventions:

- Variables in capital letters are random variables. Variables in small case letters are specific values.
- Bold font means multi-dimension, where the exact dimension will be omitted when clear from the context.

Non-signaling (NS) systems

We propose a set of terminology about the NS systems. It extends the existing notions such as NS strategies and operations on NS boxes, however, in a more systematic way. It is not the only choice of such an extension. However, it provides a set of convenient terminology for our purpose. We hope that this set of terminology also provides a convenient language in general to work with NS systems.

NS Systems & States. For any NS system, one needs to first specify the set of parties that are non-signaling to each other. Let $N_1, \cdots, N_n$ denote n spatially separated parties in the NS system. Each N_i is an *NS box*[5] that takes input X_i with the input alphabet $\mathcal{X}_i$ and outputs A_i with the output alphabet $\mathcal{A}_i$. The notation N_i is a name used to refer to the NS box itself.

Definition 5.1 *An NS system Ξ consists of n NS boxes $N_1, \cdots, N_n$, each with its input alphabet $\mathcal{X}_i$ and output alphabet $\mathcal{A}_i$, for any $i \in [n]$.*

An *NS state* of system Ξ, denoted by Γ, is an NS strategy (in the conventional way) that can be played by NS boxes $N_1, \cdots, N_n$ in the system Ξ subject to the NS constraints among NS boxes. If necessary, we also use $\Xi_{N_1,\cdots,N_n}$ to specifically denote each NS box of the NS system Ξ. An NS strategy refers to a collection of (probabilisitic) correlations, denoted by $\Gamma_{A_1\cdots A_n|X_1\cdots X_n}$, between the inputs and the outputs of NS boxes. Specifically, each entry $\Gamma_{a_1\cdots a_n|x_1\cdots x_n}$ is the probability of outputing $a_1, \cdots, a_n$ conditioned on inputing $x_1, \cdots, x_n$ to each NS box N_i respectively. Again here capital variables (e.g., A, X) denote random variables and small case variables (e.g., a, x) denote specific values of random variables. Mathematically, any NS state Γ in an NS system Ξ is a dimension $\Pi_{i=1}^n |\mathcal{A}_i||\mathcal{X}_i|$ nonnegative vector.

Let S be any subset of $[n] = \{1, \cdots, n\}$ and its complement be $\overline{S}$. Let A_S denote the collection of random variables $A_i, i \in S$ for any subset S. Similarly for X_S and x_S. For any fixed input $x_{\overline{S}}$, we can define the marginal input-output correlations over the system S, as follows,

$$\Gamma_{A_S|X_S x_{\overline{S}}} = \sum_{a_{\overline{S}}} \Gamma_{A_S a_{\overline{S}}|X_S x_{\overline{S}}}. \tag{5.1}$$

The *NS condition* requires the following,

$$\Gamma_{A_S|X_S x_{\overline{S}}} = \Gamma_{A_S|X_S x'_{\overline{S}}}, \forall x_{\overline{S}}, x'_{\overline{S}}, S \subseteq [n]. \tag{5.2}$$

Definition 5.2 *An NS state, denoted Γ, of an NS system Ξ, is an NS strategy that can be played by NS boxes in the system Ξ subject to the NS constraint (Eq. (5.2)). The set of all NS states in the NS system Ξ is denoted by* $\mathrm{NSS}(\Xi)$.

Since the NS condition (Eq. (5.2)) is linear, any convex combination of NS states is still an NS state. Therefore, the set of all NS states $\mathrm{NSS}(\Xi)$ for any NS system Ξ is a *convex* set. We elaborate on a few basic concepts about NS states as follows.

[5] Our formulation implicitly assumes that each NS box only takes an input in once and cannot be reused afterwards. This is not the most general case in NS systems. For example, one can imagine repetitive use of the same NS box for different inputs. However, our formulation is sufficient for the analysis in this paper.

Marginal NS states. The NS condition (Eq. (5.2)) suggests the NS state restricting to the subset S of the system does not depend on the input $x_{\overline{S}}$ to NS boxes in the subset $\overline{S}$ of the system. Thus, we have well-defined *marginal NS states* when we only care about the input-output correlations of a subset of NS boxes. Precisely, a marginal NS state on the subset S, denoted $\Gamma_{A_S|X_S}$, is defined as follows,

$$\Gamma_{A_S|X_S} = \Gamma_{A_S|X_S x_{\overline{S}}}, \tag{5.3}$$

for any specific input $x_{\overline{S}}$ to the subset $\overline{S}$ of NS boxes. For convenience, we will call such NS states as marginal states (or sometimes, local states) when S is clear from the context, and the NS state for the entire NS system as the global state.

Conditional NS states. One can also consider conditioning NS states of the subsystem S on certain inputs and outputs, denoted $a_{\overline{S}}, x_{\overline{S}}$ respectively, of the subsystem $\overline{S}$. The *conditional NS state*, denoted $\Gamma_{A_S|X_S,a_{\overline{S}},x_{\overline{S}}}$, is defined by first fixing $X_{\overline{S}} = x_{\overline{S}}$ and then taking the conditional distribution on $A_{\overline{S}} = a_{\overline{S}}$. It is not hard to show that $\Gamma_{A_S|X_S,a_{\overline{S}},x_{\overline{S}}}$ is still an NS state for the subsystem S.

Furthermore, one could extend the above definition to handle any event ω in the subsystem $\overline{S}$, which is a subset of all possible $\{(a_{\overline{S}}, x_{\overline{S}})\}$. Let $\mathbf{Pr}[a_{\overline{S}}, x_{\overline{S}}|\omega]$ denote the conditional probability of $(a_{\overline{S}}, x_{\overline{S}})$ on the event ω. Then we have,

$$\Gamma_{A_S|X_S,\omega} = \sum_{(a_{\overline{S}},x_{\overline{S}})\in\omega} \mathbf{Pr}[a_{\overline{S}}, x_{\overline{S}}|\omega] \times \Gamma_{A_S|X_S,a_{\overline{S}},x_{\overline{S}}}, \tag{5.4}$$

which is also an NS state for the subsystem S.

The above notion naturally gives rise to the following *sub-normalized* NS states,

$$\Gamma_{\omega,A_S|\perp,X_S} = \sum_{(a_{\overline{S}},x_{\overline{S}})\in\omega} \mathbf{Pr}[a_{\overline{S}}, x_{\overline{S}}] \times \Gamma_{A_S|X_S,a_{\overline{S}},x_{\overline{S}}}, \tag{5.5}$$

for any event ω over the sample space $(a_{\overline{S}}, x_{\overline{S}})$. It is sub-normalized in the sense that it only characterizes the NS state when ω happens, in general with probability smaller than 1.

Special NS states. There are a few special cases of NS states worth special attention. The first case is that any conventional random source (i.e., any random variable with some distribution) can be viewed as a special NS box that ignores the input and directly outputs the corresponding random variable as the output. Precisely, we use $\perp$ to denote the input to such NS boxes. Thus, any random source A is a special NS state $\Gamma_{A|\perp}$, which is denoted by A for convenience. In particular, we will use U_n to denote the uniform distribution over $\{0,1\}^n$, i.e., $\Gamma_{U_n|\perp}$, or U_A to denote the uniform distribution over the range of A. We refer the NS boxes ignoring inputs as the classical- (C-) part of the NS state. Otherwise, it belongs to the NS part. Thus, similar to the classical-quantum states, we can define classical-NS (or C-NS) states.

The second case of NS states are *product* NS states across some cut $S/\overline{S}$ for some subset S. Namely, the NS state $\Gamma_{A_S,A_{\overline{S}}|X_S,X_{\overline{S}}}$ is a product NS state across $S/\overline{S}$ if and only if

$$\Gamma_{A_S,A_{\overline{S}}|X_S,X_{\overline{S}}} = \Gamma_{A_S|X_S} \otimes \Gamma_{A_{\overline{S}}|X_{\overline{S}}}, \tag{5.6}$$

where the $\otimes$ operation is carried out when deeming NS states as vectors. Specifically, the $\otimes$ operation means, for any $x_S, x_{\overline{S}}$, the conditional probabilistic distribution $\Gamma_{A_S,A_{\overline{S}}|x_S,x_{\overline{S}}}$ is a product distribution of its marginal probabilistic distribution $\Gamma_{A_S|x_S}$ and $\Gamma_{A_{\overline{S}}|x_{\overline{S}}}$.

The third case of NS states are C-NS states with the C-part uniform to some part of the NS part. For example, let $\Gamma_{A,\mathbf{B}|\perp,\mathbf{Y}}$ be any C-NS state with classical part A. For any sub-system S, A is called *uniform-to-S*, if

$$\Gamma_{A,\mathbf{B}_S|\perp,\mathbf{Y}_S} = U_A \otimes \Gamma_{\mathbf{B}_S|\mathbf{Y}_S}. \tag{5.7}$$

We also define

$$\mathrm{Ideal}_A(\Gamma_{A,\mathbf{B}|\perp,\mathbf{Y}}) = U_A \otimes \Gamma_{\mathbf{B}|\mathbf{Y}}. \tag{5.8}$$

Protocols over NS States.

We abstract classical interactions with NS boxes in a certain NS system Ξ_1 as *protocols* that map an NS state Γ_1 in NS system Ξ_1 to NS state Γ_2 in NS system Ξ_2. The change of the NS system is potentially due to the use of NS boxes and the introduction of new random variables. Precisely, let $\mathcal{P}(\Xi_1 \to \Xi_2)$ denote the set of protocols mapping from $\mathrm{NSS}(\Xi_1)$ to $\mathrm{NSS}(\Xi_2)$. Let Ξ_1 be any NS system with m NS boxes. Let $N_1, N_2, \cdots N_m$ denote these NS boxes. Label NS boxes with *available* or *unavailable*. If an NS box has no input or it has been used, we label it unavailable. Otherwise, we label it available. Any protocol Π proceeds in the following way.

- Let history be the collection of the indices, inputs and outputs of unavailable boxes. Initially, history only contains the outputs of NS boxes with no inputs.
- Based on the history, the protocol decides to adopt one of the following actions: (1) choose one available box, and generate the input to feed into that box; or (2) stop and output.
- If the choice is (1), then the protocol proceeds to feed that input to the chosen box, and obtain the output of that box. Then the protocol adds the index of the chosen box and its input-output into history, labels this box as unavailable, and goes back to the last step.
- If the choice is (2), then the protocol ends and outputs. Let S be the subset of unavailable boxes when the protocol ends. The output of the protocol is an NS state in a new NS system Ξ_2, in which the set S of boxes are replaced by boxes with no input and (a_i, x_i) as output for each $i \in S$. Namely, the new NS state is $\Gamma_{A_S,X_S,A_{\overline{S}}|X_{\overline{S}}}$. The protocol's output could also contain any random variable that depends on A_S, X_S.

For convenience, Ξ_1 and Ξ_2 might be omitted when they are clear from the context. We use $\Pi(\Gamma_1)$ to denote the outcome of protocol Π on Γ_1, i.e., $\Pi(\Gamma_1) = \Gamma_2$. Moreover,

- Any protocol $\Pi \in \mathcal{P}(\Xi_1 \to \Xi_2)$ is *deterministic* (*randomized*) if all the choices of each step of Π are deterministic (randomized, respectively).
- For any $\Pi_1 \in \mathcal{P}(\Xi_1 \to \Xi_2)$ and $\Pi_2 \in \mathcal{P}(\Xi_2 \to \Xi_3)$, the *composition* of $\Pi_2 \circ \Pi_1$ is a new protocol inside $\mathcal{P}(\Xi_1 \to \Xi_3)$ that executes Π_1 first and then Π_2 second. Namely, $\Pi_2 \circ \Pi_1(\cdot) = \Pi_2(\Pi_1(\cdot))$.

Distinguisher & Distance Measures between NS states.

Now we consider a specific protocol D, called *distinguishers*, which map an NS state to a binary random variable. It is used to distinguish one NS state from another, where the binary variable refers to the index of the NS states. This provides a natural operational definition of the distance between NS states Γ_1 and Γ_2 inside $\mathrm{NSS}(\Xi)$.

Precisely, let $\Delta(Q_1, Q_2)$ denote the *statistical* distance between two distributions Q_1 and Q_2 over the same domain. We define the NS distance as follows.

Definition 5.3 *For any two NS states Γ_1 and Γ_2, and any distinguisher $D \in \mathcal{P}(\Xi \to \{0,1\})$, we define the induced distance by Π between Γ_1 and Γ_2 by*

$$\Delta_D(\Gamma_1, \Gamma_2) = \Delta(D(\Gamma_1), D(\Gamma_2)), \tag{5.9}$$

where $D(\Gamma_1), D(\Gamma_2)$ are two distributions over $\{0,1\}$.

The NS distance *between Γ_1 and Γ_2 is defined by*

$$\Delta(\Gamma_1, \Gamma_2) = max_D\, \Delta_D(\Gamma_1, \Gamma_2), \tag{5.10}$$

where the maximization is over all possible distinguishers D.

The NS distance inherits many properties from the statistical distance, which we summarize as follows.

Triangle Inequality. For any three NS states $\Gamma_1, \Gamma_2, \Gamma_3$, we have

$$\Delta(\Gamma_1, \Gamma_3) \leq \Delta(\Gamma_1, \Gamma_2) + \Delta(\Gamma_2, \Gamma_3). \tag{5.11}$$

Proof: Let $\delta = \Delta(\Gamma_1, \Gamma_3)$ and D_{13} the optimal distinguisher between Γ_1 and Γ_3. Apply D_{13} to $\Gamma_1, \Gamma_2, \Gamma_3$ to obtain three distributions over $\{0,1\}$. By the triangle inequality of the statistical distance,

$$\Delta(D_{13}(\Gamma_1), D_{13}(\Gamma_3)) \leq \Delta(D_{13}(\Gamma_1), D_{13}(\Gamma_2)) + \Delta(D_{13}(\Gamma_2), D_{13}(\Gamma_3)).$$

and by definition,

$$\Delta(D_{13}(\Gamma_1), D_{13}(\Gamma_2)) \leq \Delta(\Gamma_1, \Gamma_2) \text{ and } \Delta(D_{13}(\Gamma_2), D_{13}(\Gamma_3)) \leq \Delta(\Gamma_2, \Gamma_3),$$

which completes the proof.

Data-Processing Inequality. For any two NS states Γ_1, Γ_2, and any protocol $\Pi \in \mathcal{P}(\Xi)$, we have

$$\Delta(\Pi(\Gamma_1), \Pi(\Gamma_2)) \leq \Delta(\Gamma_1, \Gamma_2). \tag{5.12}$$

Proof: Any distinguisher D to distinguish between $\Pi(\Gamma_1)$ and $\Pi(\Gamma_2)$ leads to a distinguisher $D \circ \Pi$ between Γ_1, Γ_2. The proof is completed by definition.

NS distance between C-NS states. Let $\Gamma_{A_1,\mathbf{B}|\perp,\mathbf{Y}}, \Gamma_{A_2,\mathbf{B}|\perp,\mathbf{Y}}$ be C-NS states in some NS system. As any distinguisher can first look at A_1(or A_2) and then decide its strategy for the rest of the system, then we have

$$\Delta(\Gamma_{A_1,\mathbf{B}|\perp,\mathbf{Y}}, \Gamma_{A_2,\mathbf{B}|\perp,\mathbf{Y}}) = \sum_a \Delta(\Gamma_{A_1=a,\mathbf{B}|\perp,\mathbf{Y}}, \Gamma_{A_2=a,\mathbf{B}|\perp,\mathbf{Y}}),$$

where $\Gamma_{A_1=a,\mathbf{B}|\perp,\mathbf{Y}}$ is a sub-normalized state.

Predictor & NS Min-entropy

Another special type of protocols P, called *predictors*, which map NS states to a random string $Z \in \{0,1\}^m$. For any C-NS state, any predictor applied on the NS part (box N_2) is used to predict the C-part (box N_1). Precisely, given a C-NS state $\Gamma_{Z,F|\perp,E}$ ($Z \in \{0,1\}^m$), a predictor P feeds some E into the NS box N_2 and obtains the output F. Namely, a predictor P turns $\Gamma_{Z,F|\perp,E}$ into a probability distribution (Z,F,E). The predictor further bases on F,E to provide a guess Z' of Z. The *guessing probability* of P is thus

$$p_{\text{guess}}(P,\Gamma_{Z,F|\perp,E}) = \mathbf{Pr}[Z = Z']. \tag{5.13}$$

We can thus define the NS min-entropy in the following operational way.

Definition 5.4 *For any C-NS state $\Gamma_{Z,F|\perp,E}$ ($Z \in \{0,1\}^m$), the* NS min-entropy *of Z conditioned on the NS part (box N_2) is defined by*

$$\mathrm{H}_\infty^{\mathsf{ns}}(Z|N_2)_\Gamma = -\log_2(\max_P p_{\text{guess}}(P,\Gamma_{Z,F|\perp,E})),$$

where the maximization is over all possible predictors P. Let $k = \mathrm{H}_\infty^{\mathsf{ns}}(Z|N_2)_\Gamma$. Any such NS state is called an (m,k)-NS-source (or k-NS-source).

It is worth mentioning that the above NS min-entropy definition is consistent with the definitions of quantum and classical (conditional) min-entropies when the underlying system becomes quantum or classical. For convenience, we will use (NS) min-entropy to denote quantum or classical min-entropy when the underlying system is clear from the context.

6 Model of Certifiable Extraction of Randomness

In this section, we formulate a general framework for certifiable extraction of randomness, or the general device-independent randomness amplification (DI-RA), from NS systems. Our motivation is to minimize the assumptions required for such secure extraction.

At the high level, given an NS state including a classical part as the classical source and an NS part of NS boxes as the untrusted (physical) devices, a general DI-RA protocol is to extract certifiable randomness that looks uniform from a certain environment, usually everything in the rest of the world which could include the manufacturer of the untrusted devices. Here we consider the correlation between the untrusted devices and the rest of the world is non-signaling. Such a general DI-RA protocol may accept on certain initial NS states and generate close-to-uniform outputs, or decide to reject when it detects suspicious behaviors of the untrusted devices. By a *certifiable* output, we mean that the chance of accepting an undesirable output (e.g., far from uniform) is small.

There are two types of objects in the model. A classical source is simply a boolean string of finite length. A physical device (also called a box) is a black box with classical interface. Namely, one can feed the device a classical input string and receive a classical output string from it. These devices have prescribed strategies to answer the inputs they receives, and may or may not accept multiple inputs. (We only consider the single input case in this paper.) The devices' answers may correlate with other devices' answers, but we always assume that the devices are spatially separated and satisfied the no-signaling condition. Physical devices are used to model both untrusted devices and the information held by the adversary. Since we do not make any assumptions about the inner-working of physical devices, these devices are *untrusted* by default and only the observed classical input/output can be trusted. Thus, it suffices to model only the input-output correlation of these devices, which is captured by NS boxes.

The NS system in DI-RA. Any DI-RA protocol operates in an NS system Ξ that consists of the following three parts:

- *Classical source* S is a single classical source of randomness.
- *Devices*
$$\mathbf{D} = (D_1, \ldots, D_t)$$
consists of a set of t NS boxes, which represents t untrusted devices.
- *Adversary (or Environment)* E is the rest of the NS system, which could consist of one or several NS boxes.

We will hence use $S, \mathbf{D}, E$ to represent each part of the NS system Ξ. Any NS state $\Gamma_{S,\mathbf{D},E} \in \mathrm{NSS}(\Xi)$ thus represents a particular status of the NS system Ξ. In particular, $\Gamma_{S,\mathbf{D},E}$ specifies the quality and quantity of the classical source S and how it is related to the devices and the environment. It also specifies the strategy of the untrusted devices, which is the strategy of the corresponding NS boxes. Moreover, $\Gamma_{S,\mathbf{D},E}$ specifies the correlation between the devices and the environment, which in turn characterizes the cheating power of a potential adversary who holds the environment.

Source condition. Here we list a few interesting conditions on the classical source. In particular, its relation to the devices and the environment, which refers to different assumptions in the context of certifiable randomness extraction. Let $\Gamma_{S,\mathbf{D},E}$ be any NS state in such an NS system. We call the classical source $S \in \{0,1\}^n$ is

- *local-uniform*, if S is uniform-to-$\mathbf{D}$, i.e., $\Gamma_{S,\mathbf{D}} = U_S \otimes \Gamma_{\mathbf{D}}$; *global-uniform*, if S is uniform-to-$\mathbf{D}, E$, i.e., $\Gamma_{S,\mathbf{D},E} = U_S \otimes \Gamma_{\mathbf{D},E}$.
- ϵ-*local-uniform*, if $\Delta(\Gamma_{S,\mathbf{D}}, U_S \otimes \Gamma_{\mathbf{D}}) \leq \epsilon$; ϵ-*global-uniform*, if $\Delta(\Gamma_{S,\mathbf{D},E}, U_S \otimes \Gamma_{\mathbf{D},E}) \leq \epsilon$.
- *local-*(n,k)*-source*, if S has k NS min-entropy conditioned on $\mathbf{D}$. Namely, $\mathrm{H}_\infty^{\mathsf{ns}}(S|\mathbf{D})_\Gamma = k$; *global-*$(n,k)$*-source*, if S has k NS min-entropy conditioned on $\mathbf{D}, E$. Namely, $\mathrm{H}_\infty^{\mathsf{ns}}(S|\mathbf{D},E)_\Gamma = k$.

Implementation. Strategies of general NS boxes are not known to be implemented by means allowed by quantum mechanics. In cases where such NS boxes can be implemented by quantum means (i.e., through the use of shared quantum states and local measurements), we call these NS states *quantum implementable*.

We also consider the presence of honest noises in the implementation of NS boxes, by either quantum means or super-quantum means. In this way, we can distinguish cheating devices from honest but noisy devices, whose noise is due to some engineering imperfection. There is a lot of freedom to choose the honest noise model. Here we consider a simple and reasonable noise model, parameterized by $0 \leq \eta \leq 1$, which says independently for each NS box and each input, the output will be flipped to a uniform selection over all possible outcomes with probability η. Any protocol that works in the presence of η level of noise is called η-*robust*.

General NS DI-RA. We have defined protocols over NS systems in the previous section. General device-independent randomness amplification protocols are just such *deterministic* protocols over NS systems. We have elaborated on the initial NS system that these protocols operate on. It suffices to describe the output NS system of these protocols.

For any initial NS system $\Xi = (S, \mathbf{D}, E)$, a general NS DI-RA protocol applies to $(S, \mathbf{D})$ subsystem and output a decision bit $O \in \{0, 1\}$, where 0 is for rejecting and 1 for accepting, as well as an output string $Z \in \{0, 1\}^*$, which are stored in the new NS system (O, Z). Let $\Xi' = (O, Z, E)$ denote the final NS system after the protocol terminates. Precisely, we have

Definition 6.1 (General NS DI-RA) *A general device-independent randomness amplification protocol* Π *is a deterministic protocol mapping an input NS state in* $\Xi = (S, \mathbf{D}, E)$ *to an output NS state in* $\Xi = (O, Z, E)$.

For any output NS state $\Gamma_{O,Z,F|\perp,\perp,E}$, where we abuse E to denote the input to the subsystem E with potentially multiple NS boxes and let F denote their output. [6] Intuitively, when $O = 0$ (reject), the quality of output Z is irrelevant as we don't require any guarantee on the output when the protocol rejects. On the other side, when $O = 1$ (accept), we hope that the output Z is uniform to the subsystem E. To capture this intuition, we define an ideal state $\Gamma^{\mathsf{Ideal}}_{O,Z,F|\perp,\perp,E}$ as follows

$$\Gamma^{\mathsf{Ideal}}_{O,Z,F|\perp,\perp,E} = \begin{cases} \Gamma_{O=0,Z,F|\perp,\perp,E}, & O = 0, \\ \mathrm{Ideal}_Z(\Gamma_{O=1,Z,F|\perp,\perp,E}) = (O = 1) \otimes U_Z \otimes \Gamma_{F|E}, & O = 1. \end{cases} \tag{6.1}$$

It then suffices to use the NS distance between $\Gamma_{O,Z,F|\perp,\perp,E}$ and $\Gamma^{\mathsf{Ideal}}_{O,Z,F|\perp,\perp,E}$ to capture the above intuition. Precisely,

Definition 6.2 (Soundness error of the output NS state) *Let* $\Gamma_{O,Z,F|\perp,\perp,E}$ *be the output NS state of a general DI-RA protocol and* $\Gamma^{\mathsf{Ideal}}_{O,Z,F|\perp,\perp,E}$ *the ideal output NS state. We say that* $\Gamma_{O,Z,F|\perp,\perp,E}$ *has a soundness error* ϵ *if*

$$\Delta(\Gamma_{O,Z,F|\perp,\perp,E}, \Gamma^{\mathsf{Ideal}}_{O,Z,F|\perp,\perp,E}) \leq \epsilon. \tag{6.2}$$

We remark that this soundness error definition is the natural NS extension of the definition from our previous work about quantum system [5]. We also note that we define a (technical) strong completeness/robustness property, which is used in the analysis of our composition construction.

Definition 6.3 *Let* $\Xi = (S, \mathbf{D}, E)$ *be the input NS system with* t *devices, i.e.,* $\mathbf{D} = (D_1, \cdots, D_t)$. *Let* $\Xi' = (O, Z, E)$ *be the output NS system where* $Z \in \{0, 1\}^m$. *Any deterministic protocol* $\Pi \in P(\Xi \to \Xi')$ *is called a (general) NS DI-RA protocol with a completeness error* ϵ_c *tolerating an* η *level of noise, and a soundness error* ϵ_s *under source condition* C, *if the following conditions hold.*

- *(Completeness) There exists an NS state* $\Gamma_{S,\mathbf{D},E} \in \mathrm{NSS}(\Xi)$ *such that its output state* $\Gamma_{O,Z,E} = \Pi(\Gamma_{S,\mathbf{D},E}) \in \mathrm{NSS}(\Xi')$ *satisfies*

$$\mathbf{Pr}[O = 1(\mathsf{Accept})] \geq 1 - \epsilon_c.$$

  Moreover, we call this QT completeness *if* $\Gamma_{S,\mathbf{D},E} \in \mathrm{NSS}(\Xi)$ *is quantum implementable.*

[6] We choose to not make the restriction on (E, F) due to multiple NS boxes in the subsystem E explicitly, which is irrelevant to our discussion here. However, one should keep in mind such restriction is already assumed.

- *(Strong completeness) If* Π *has completeness error* ϵ_c *for every fixed source value* $X = x$*.*
- *(NS Soundness under condition C) For any input NS state* $\Gamma_{S,\mathbf{D},E} \in \mathrm{NSS}(\Xi)$ *satisfying condition* C*, its output state* $\Gamma_{O,Z,E}$ *has a soundness error* $\leq \epsilon_s$*. Here condition* C *refers to anyone of local/gloabl-uniform,* ϵ*-local/global-uniform, and local/global-*k*-source defined previously.*
- *((Strong) robustness) If* Π *has completeness error* ϵ_c *even in the presence of an* η*-level of noise. Strong robustness is defined similarly except for every fixed source value* $X = x$*.*

We call any deterministic protocol $\Pi \in P(\Xi \to \Xi')$ an (n, k, t, ϵ) NS DI-RA protocol if the input NS system has t devices and the protocol Π has a completeness error $\epsilon_c = \epsilon$ and a soundness error $\epsilon_s = \epsilon$ when the source is a *local-*(n, k)*-source*. Similarly, we call any such protocol an (n, t, ϵ) NS DI-RA protocol if the input NS system has t devices and the protocol Π has a completeness error $\epsilon_c = \epsilon$ and a soundness error $\epsilon_s = \epsilon$ when the source has n *bits and is local-uniform.*

We don't explicitly write down the output length as it is less significant for our conceptual message. One can think the output length is one bit in this paper. It should also be understood that one can simply compose this one-bit protocol multiple times to output multiple bits if more resources (e.g., devices) are allowed.

7 Main Protocol

We will describe our main protocol in this section. To that end, let us first introduce classical seeded randomness extractors and somewhere uniform sources in the NS setting. Classical seeded extractors are deterministic functions that convert any classical min-entropy source to a marginally uniform output with the help of a short uniform seed. Precisely,

Definition 7.1 (Strong Seeded Extractor) *A function* $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is a classical* (k, ϵ)-strong seeded (randomness) extractor*, if for any min-entropy* $\geq k$ *source* $X \in \{0,1\}^n$*, and for a uniform seed* $Y \in \{0,1\}^d$ *independent of* X*, we have*

$$\Delta(\mathrm{Ext}(X, Y), Y), U_m \otimes Y) \leq \epsilon. \tag{7.1}$$

Note that we use NS notations above which are mathematically sound. As previously defined, all classical random variables are treated as NS boxes with $\perp$ input and our notations when used in this special case are consistent with all classical ones. One of the best known classical extractors is as follows.

Theorem 7.2 ([14]) *For every constant* $\alpha > 0$*, and all positive integers* n, k *and* $\epsilon > 0$*, there is an explicit construction of a strong* (k, ϵ) *extractor* $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *with* $d = O(\log n + \log(1/\epsilon))$ *and* $m \geq (1 - \alpha)k$*.*

We are ready to introduce another important object *somewhere uniform sources* in the NS setting that is critically used in our protocol. We remark that somewhere randomness has been a well-motivated and actively-studied object in the classical literature. We extend and adjust this notion to the NS setting to make it useful for our general NS DI-RA protocols. Also, for technical reasons, we require that most of the blocks are close to uniform-to-device.

General NS DI-RA protocol Π_{amp}

- Let $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ be a classical (k, ϵ_0)-strong seeded randomness extractor.
- Let Π_{dec} be an $(n_{\mathsf{dec}}, t_{\mathsf{dec}}, \epsilon_{\mathsf{dec}})$ NS DI-RA with seed length n_{dec} that uses t_{dec} devices.
- Π_{amp} operates on the input classical source S over $\{0,1\}^n$ and $t_{\mathsf{amp}} = 2^d \cdot t_{\mathsf{dec}}$ devices $\mathbf{D} = (\mathbf{D}_1, \ldots, \mathbf{D}_{2^d})$, where each $\mathbf{D}_i$ denotes a set of t_{dec} devices, as follows.

Protocol Π_{amp}

1. For every $i \in \{0,1\}^d$, let $S_i = \mathrm{Ext}(X, i)$ and invoke Π_{dec} on the subsystem $S_i, \mathbf{D}_i$ and obtain O_i, Z_i as the output.
2. Let $\eta = 2 \cdot \epsilon_{\mathsf{dec}}$ (a threshold). If at most η-fraction of Π_{dec} rejects (i.e., $O_i = \mathsf{Rej}$), then Π_{amp} accepts and outputs $O = \mathsf{Acc}$ and $Z = \bigoplus_{i \in [2^d]} Z_i$. Otherwise, Π_{amp} rejects, i.e., outputs $O = \mathsf{Rej}$ and $Z = \perp$.

Figure 4: The Main Construction of the General NS DI-RA Protocol Π_{amp}.

Definition 7.3 *A source-device NS state $\Gamma_{\mathbf{X},\mathbf{D}}$ with classical source $X = X_1, \ldots, X_t$ and t devices $D_1, \ldots, D_t$ (or t set of devices $\mathbf{D}_1, \ldots, \mathbf{D}_t$) is a* $(1-\gamma)$-somewhere uniform-to-device *NS state if there exist at least $(1-\gamma)$-fraction of $i \in [t]$ such that $\Gamma_{X_i, D_i} = U_{X_i} \otimes \Gamma_{D_i}$.*

$\Gamma_{X,\mathbf{D}}$ is a $(1-\gamma)$-somewhere ϵ-uniform-to-device *NS state if there exists at least $(1-\gamma)$-fraction of $i \in [t]$ such that Γ_{X_i,D_i} is ϵ-close to $U_{X_i} \otimes \Gamma_{D_i}$. Namely, $\Delta(\Gamma_{X_i,D_i}, U_{X_i} \otimes \Gamma_{D_i}) \leq \epsilon$.*

The key change we made in the definition is that each section of the somewhere uniform source is only required to be uniform to a *single* set of devices, instead of being uniform to the collection of all devices. As we discussed before, the latter requirement is not known to be satisfied by the classical construction from seeded extractors because there is no such extractor in the NS setting. However we show it is possible to achieve somewhere ϵ-uniform to a single set of devices by losing some parameters and going through a *post-selection* argument. Precisely, we have, (proof deferred to Section 8)

Theorem 7.4 *Let* $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *be a classical (k, ϵ)-strong seeded extractor. Let $\Gamma_{X,\mathbf{D}}$ be a source-device NS state with a classical source X and 2^d set of devices $\mathbf{D} = (\mathbf{D}_1, \ldots, \mathbf{D}_{2^d})$. For every $i \in [2^d]$, let $Y_i = \mathrm{Ext}(X, i)$. If X has k-bits of NS min-entropy conditioned on $\mathbf{D}$, i.e., $\mathrm{H}^{\mathsf{ns}}_\infty(X|\mathbf{D})_\Gamma \geq k$, then $\Gamma_{\mathbf{Y},\mathbf{D}}$ is a $(1-\gamma)$-somewhere $(2^m \cdot \epsilon/\gamma)$-uniform-to-device source for any $\gamma < 2^{-m}\epsilon^{-1}/100$, where $\mathbf{Y} = (Y_1, \cdots, Y_{2^d})$.*

Our construction of the general NS DI-RA protocol is then given in Fig. 4. As we discussed before, we will need NS DI-RA protocols running on local-uniform sources to serve as the *decouplers*, which is denoted by Π_{dec} in Fig. 4. We also need strong completeness/robustness for composition. We construct such Π_{dec} and leave its construction, proof, and discussions to Section 9. Note that both Π_{dec} and Π_{amp} only output one bit.

Theorem 7.5 *For any $0 < \epsilon_{\mathsf{dec}} < 1$, there exists an η-strong robust $(n_{\mathsf{dec}}, t_{\mathsf{dec}}, \epsilon_{\mathsf{dec}})$ NS DI-RA such that $n = \mathrm{poly}(1/\epsilon_{\mathsf{dec}})$, $t = \mathrm{poly}(1/\epsilon_{\mathsf{dec}})$, and $\eta \approx 1.748\%$.*

We further prove that such Π_{dec} works even when the source is ϵ'-local-uniform for some ϵ'.

Theorem 7.6 *For any $0 < \epsilon_{\mathsf{dec}} < 1$, there exists an η-strong robust $(n_{\mathsf{dec}}, t_{\mathsf{dec}}, \epsilon_{\mathsf{dec}})$ NS DI-RA with $n = \mathrm{poly}(1/\epsilon_{\mathsf{dec}})$, $t = \mathrm{poly}(1/\epsilon_{\mathsf{dec}})$, and $\eta \approx 1.748\%$ such that the soundness error ϵ_{dec} holds when the input source is ϵ'-local-uniform for $\epsilon' \leq \mathrm{poly}(\epsilon_{\mathsf{dec}})$.*

We remark that in the quantum world, such a theorem can be proven in a black-box manner by using the fact that if any two local states A, B are ϵ-close, then for any global state A' of A, one can always find a nearby global state B' of B such that A' and B' are also ϵ'-close when measured in fidelity. Such a statement is not known to hold in the NS setting. Instead, we approach Theorem 7.6 in a very non-black-box way by identifying a critical (but technical) property of Π_{dec} and proving such property still holds when the source is ϵ'-local uniform. This proof is achieved by a complicated NS reduction, which we defer to Section 10.

By putting everything together and setting appropriate parameters, we have our main result

Theorem 7.7 *For any $\epsilon > 0$, any $n \geq k \geq \mathrm{poly}(1/\epsilon)$, there exists a DI-RA protocol Π_{amp} (in Fig. 4) that uses one (n, k) source input, has the completeness and soundness errors $\epsilon_c = \epsilon_s = \epsilon$, tolerates a $\eta \approx 1.748\%$ level of noise, and uses $2^{\mathrm{poly}(1/\epsilon)}$ devices.*

Proof. We set the parameters in Π_{amp} as follows. We first set $\gamma = \epsilon_{\mathsf{dec}} = \epsilon/4$. We then set the parameters $n_{\mathsf{dec}}, t_{\mathsf{dec}}, \epsilon'$ for Π_{dec} based on Theorem 7.6 and ϵ_{dec}, namely, $n_{\mathsf{dec}} = \mathrm{poly}(1/\epsilon_{\mathsf{dec}}), t_{\mathsf{dec}} = \mathrm{poly}(1/\epsilon_{\mathsf{dec}})$ and $\epsilon' = \mathrm{poly}(\epsilon_{\mathsf{dec}})$. Then we set the parameters for Ext as follows: first, set $m = n_{\mathsf{dec}}$ and $\epsilon_0 = 2^{-m}\gamma\epsilon'$. Then set n, k, d large enough based on Theorem 7.2 to satisfy $m = n_{\mathsf{dec}}$ and $\epsilon_0 = 2^{-m}\gamma\epsilon'$. In particular, we have $n = \mathrm{poly}(1/\epsilon)$ and $k = \mathrm{poly}(1/\epsilon)$.

The completeness and robustness properties follow from the strong robustness of Π_{dec}. More precisely, we consider the honest devices for Π_{amp} to be independent copies of honest devices of Π_{dec}. By strong robustness of Π_{dec}, even with $\eta \approx 1.748\%$ level of noise, each Π_{dec} rejects with probability at most ϵ_{dec} independently, for any fixed input source value. Therefore, the fraction of rejected Π_{dec} is $< \eta$ with high probability.

Proving soundness is more involved. Note that the execution of Π_{amp} generates O, Z and $O_i, Z_{ii\in[2^d]}$, and our goal is to show $\Delta(\Gamma_{O,Z,E}, \Gamma^{\mathsf{Ideal}}_{O,Z,E}) \leq \epsilon$. We bound the NS distance in two steps: we will first identify a "good" index $i^* \in [2^d]$, and then use soundness of the i^*-th sub-protocol to bound the NS distance.

In the first step of Π_{amp}, by Theorem 8.1 and the setting of parameters, we know that there are at least $(1-\gamma)$-fraction of $i \in [2^d]$ such that the i-th source-device subsystem $\Gamma_{S_i,\mathbf{D}_i}$ is ϵ'-close to uniform-to-device. Let us denote the set of good indices $G \subseteq [2^d]$. We have the following claim.

Claim 7.8 *There exists $i^* \in G$ such that $\Pr[O = \mathsf{Acc} \wedge O_{i^*} = \mathsf{Rej}] \leq \gamma + \eta$.*

Proof. By definition of Π_{amp}, when $O = \mathsf{Acc}$, there are at least $(1-\eta)$-fraction of $i \in [2^d]$ such that $O_i = \mathsf{Acc}$. Since G contains at least $(1-\gamma)$-fraction of $i \in [2^d]$, we know that when $O = \mathsf{Acc}$, there are at least $(1-\eta-\gamma)$-fraction of $i \in [2^d]$ such that $i \in G$ and $O_i = \mathsf{Acc}$. In other words,

$$\Pr_{i\leftarrow[2^d]}[i \in G \wedge O_i = \mathsf{Acc} | O = \mathsf{Acc}] \geq 1 - \eta - \gamma,$$

where the probability is over Π_{amp} and uniform $i \in [2^d]$. By an averaging argument, there exists some fixed $i^* \in G$ such that $\Pr[O_{i^*} = \mathsf{Acc} | O = \mathsf{Acc}] \geq 1 - \eta - \gamma$, or equivalently, $\Pr[O_{i^*} = \mathsf{Rej} | O = \mathsf{Acc}] \leq$

$\eta + \gamma$. Therefore,

$$\Pr[O = \mathsf{Acc} \wedge O_{i^*} = \mathsf{Rej}] \leq \Pr[O_{i^*} = \mathsf{Rej} | O = \mathsf{Acc}] \leq \eta + \gamma.$$

■

Since $i^* \in G$, the i^*-th source-device subsystem $\Gamma_{S_i, \mathbf{D}_i}$ is ϵ'-close to uniform-to-device. The soundness of i^*-th sub-protocol Π_{dec} implies that

$$\Delta(\Gamma_{O_{i^*}, Z_{i^*}, (O, Z_{-i^*}, E)}, \Gamma^{\mathsf{Ideal}}_{O_{i^*}, Z_{i^*}, (O, Z_{-i^*}, E)}) \leq \epsilon_{\mathsf{dec}},$$

where we view the (O, Z_{-i^*}, E) components as the adversary to the i^*-th sub-protocol Π_{dec}. Since O and O_{i^*} are classical components, we can consider the distance conditioned on the values of O and O_{i^*}. In particular, we have

$$\begin{aligned}
\epsilon_{\mathsf{dec}} &\geq \Delta(\Gamma_{O_{i^*}, Z_{i^*}, (O, Z_{-i^*}, E)}, \Gamma^{\mathsf{Ideal}}_{O_{i^*}, Z_{i^*}, (O, Z_{-i^*}, E)}) \\
&\geq \Pr[O = O_{i^*} = \mathsf{Acc}] \cdot \Delta(\Gamma_{O_{i^*}, Z_{i^*}, (O, Z_{-i^*}, E) | O = O_{i^*} = \mathsf{Acc}}, \Gamma^{\mathsf{Ideal}}_{O_{i^*}, Z_{i^*}, (O, Z_{-i^*}, E) | O = O_{i^*} = \mathsf{Acc}}) \\
&= \Pr[O = O_{i^*} = \mathsf{Acc}] \cdot \Delta(\Gamma_{O_{i^*}, Z_{i^*}, (O, Z_{-i^*}, E) | O = O_{i^*} = \mathsf{Acc}}, \Gamma_{O_{i^*}, U_1, (O, Z_{-i^*}, E) | O = O_{i^*} = \mathsf{Acc}}) \\
&\geq \Pr[O = O_{i^*} = \mathsf{Acc}] \cdot \Delta(\Gamma_{O_{i^*}, Z_{i^*} \oplus Z_{-i^*}, (O, E) | O = O_{i^*} = \mathsf{Acc}}, \Gamma_{O_{i^*}, U_1 \oplus Z_{-i^*}, (O, E) | O = O_{i^*} = \mathsf{Acc}}) \\
&= \Pr[O = O_{i^*} = \mathsf{Acc}] \cdot \Delta(\Gamma_{O_{i^*}, Z, (O, E) | O = O_{i^*} = \mathsf{Acc}}, \Gamma_{O_{i^*}, U_1, (O, E) | O = O_{i^*} = \mathsf{Acc}}), \qquad (7.2)
\end{aligned}$$

where the third line follows by the definition of ideal state, where U_1 is an independent and uniform bit, and the fourth follows by the fact that deterministic operation can only decrease the distance. We can now bound the soundness error as follows.

$$\begin{aligned}
\Delta(\Gamma_{O,Z,E}, \Gamma^{\mathsf{Ideal}}_{O,Z,E}) &\leq \Delta(\Gamma_{O,Z,(O_i,E)}, \Gamma^{\mathsf{Ideal}}_{O,Z,(O_i,E)}) \\
&= \Pr[O = \mathsf{Rej}] \cdot \Delta(\Gamma_{O,Z,(O_{i^*},E) | O = \mathsf{Rej}}, \Gamma^{\mathsf{Ideal}}_{O,Z,(O_{i^*},E) | O = \mathsf{Rej}}) \\
&+ \Pr[O = \mathsf{Acc} \wedge O_{i^*} = \mathsf{Rej}] \cdot \Delta(\Gamma_{O,Z,(O_{i^*},E) | O = \mathsf{Acc}, O_{i^*} = \mathsf{Rej}}, \Gamma^{\mathsf{Ideal}}_{O,Z,(O_{i^*},E) | O = \mathsf{Acc}, O_{i^*} = \mathsf{Rej}}) \\
&+ \Pr[O = O_{i^*} = \mathsf{Acc}] \cdot \Delta(\Gamma_{O,Z,(O_{i^*},E) | O = O_{i^*} = \mathsf{Acc}}, \Gamma^{\mathsf{Ideal}}_{O,Z,(O_{i^*},E) | O = O_{i^*} = \mathsf{Acc}}) \\
&\leq \Pr[O = \mathsf{Rej}] \cdot 0 + \Pr[O = \mathsf{Acc} \wedge O_{i^*} = \mathsf{Rej}] \cdot 1 \\
&+ \Pr[O = O_{i^*} = \mathsf{Acc}] \cdot \Delta(\Gamma_{O,Z,(O_{i^*},E) | O = O_{i^*} = \mathsf{Acc}}, \Gamma_{O,U_1,(O_{i^*},E) | O = O_{i^*} = \mathsf{Acc}}) \\
&\leq (\gamma + \eta) + \epsilon_{\mathsf{dec}} \leq \epsilon,
\end{aligned}$$

where we use the trivial distance bound 1 for the case $(O = \mathsf{Acc} \wedge O_{i^*} = \mathsf{Rej})$, and use Eq 7.2 for the case $O = O_{i^*} = \mathsf{Acc}$.

■

Remark: We use poly$(\cdot)$ to represent polynomial dependence because the specific polynomial does not have a simple form and is less relevant to our conceptual message.

8 Somewhere Uniform-to-device Source

We now show that by using a *classical* (k, ϵ)-strong randomness extractor, we can turn a source X with k-bits of NS min-entropy to devices to a somewhere $2^m \cdot \epsilon$-uniform-to-device source, where m is the output length of the extractor.

Theorem 8.1 *Let* $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *be a classical* (k,ϵ)*-strong seeded extractor. Let* $\Gamma_{X,\mathbf{D}}$ *be a source-device NS state with a classical source* X *and* 2^d *set of devices* $\mathbf{D} = (\mathbf{D}_1, \ldots, \mathbf{D}_{2^d})$*. For every* $i \in [2^d]$*, let* $Y_i = \mathrm{Ext}(X, i)$*. If* X *has* k*-bits of NS min-entropy conditioned on* $\mathbf{D}$*, i.e.,* $\mathrm{H}^{\mathsf{ns}}_\infty(X|\mathbf{D})_\Gamma \geq k$*, then* $\Gamma_{\mathbf{Y},\mathbf{D}}$ *is a* $(1-\gamma)$*-somewhere* $(2^m \cdot \epsilon/\gamma)$*-uniform-to-device source for any* $0 < \gamma < 1$*, where* $\mathbf{Y} = (Y_1, \cdots, Y_{2^d})$*.*

Proof. We prove the theorem by contradiction. Suppose $\Gamma_{\mathbf{Y},\mathbf{D}}$ is not $(1-\gamma)$-somewhere $(2^m \cdot \epsilon/\gamma)$-close-to-uniform-to-device, which means that for at least γ-fraction of $i \in [2^d]$, $\Gamma_{Y_i,\mathbf{D}_i}$ is $2^m \cdot \epsilon/\gamma$-far from $U_{Y_i} \otimes \Gamma_{\mathbf{D}_i}$; let $B \subset [2^d]$ denote the set of such bad blocks. This means that for every $i \in B$, there exist distinguishers A_i that distinguish $\Gamma_{Y_i,\mathbf{D}_i}$ from $U_{Y_i} \otimes \Gamma_{\mathbf{D}_i}$ with advantage $2^m \cdot \epsilon/\gamma$. In general, A_i may choose its input to $\mathbf{D}_i$ depending on the classical source Y_i. To derive a contradiction, we consider a class of *simple* distinguishers: A distinguisher A' for $\Gamma_{Y,\mathbf{D}}$ and $U_Y \otimes \Gamma_{\mathbf{D}}$ is called *simple* if A' chooses it's input to the device *independent* of the classical source Y. We proceed in the following two steps to derive a contradiction:

- On one hand, we use a *post-selection argument* to reduce the distinguishers A_i to *simple* distinguishers A'_i at the cost of losing a factor of 2^m in the distinguishing advantage. Namely, each A'_i is a simple distinguisher that distinguishes $\Gamma_{Y,\mathbf{D}}$ from $U_Y \otimes \Gamma_{\mathbf{D}}$ with advantage $> \epsilon$.
- On the other hand, we observe that since simple distinguishers convert the devices into classical distributions without accessing the classical source, a classical strong seeded extractor can still extract randomness against simple distinguishers. This means that the distinguishing advatange of A'_i should be less than ϵ (on average), and hence, a contradiction.

Post-selection argument. For each $i \in B$, we define A'_i as follows: A'_i makes a uniformly random guess W_i, and supplies $\mathbf{D}_i$ with input $M_i(W_i)$ to obtain output $E_i(W_i)$. Then A'_i check if W_i equals to the classical source Y_i. If so, A'_i outputs whatever A_i outputs, and otherwise, A'_i outputs $\perp$. Clearly, with probability 2^{-m}, the guess is correct and A'_i perfectly simulates A_i, and with probability $1-2^{-m}$, A'_i simply gives up. Thus, the advantage of A'_i is exactly 2^{-m} times the advantage of A_i, which is $> \epsilon/\gamma$.

Extraction against simple distinguishers. Note that for every $i \in B$, A'_i turns the device into a classical distrbution $E'_i = (W_i, M_i(W_i), E_i(W_i))$ without accessing the classical source X. For conveience, let (dummy) $E'_i = \perp$ for every $i \notin B$. Let $E = (E'_1, \ldots, E'_{2^d})$. By the fact that X has k bits of min-entropy to the devices, we have $\mathrm{H}_\infty(X|E) \geq k$. By the property of strong extractor Ext, we have

$$\mathrm{E}_i[\Delta((Y_i, E'_i), (U_{Y_i}, E'_i))] \leq \mathrm{E}_i[\Delta((Y_i, E), (U_{Y_i}, E))] \leq \epsilon.$$

However, by the facts that (i) $|B| \geq \gamma \cdot 2^d$, and (ii) for every $i \in B$, $\Delta((Y_i, E'_i), (U_{Y_i}, E'_i)) > \epsilon/\gamma$, we have

$$\mathrm{E}_i[\Delta((Y_i, E'_i), (U_{Y_i}, E'_i))] > \epsilon,$$

a contradiction. ∎

9 Non-signaling Decouplers

In this section, we modularize and simplify the security proof in [11] by making use of a different nonlocal game (the BHK game [1]) and taking advantage of the fact that the ideal input distribution

to the decoupler is uniform in our case instead of SV. This allows us to reduce the time complexity, efficiently generate random hash functions, and achieve robustness comparing to [11], which might be of independent interest. We also identify a key technical property (Corollary 9.10), which is crucial in handling close-to-uniform sources in Section 10.

We will first introduce the BHK game in Section 9.1 and then describe our protocol and the analysis of its completeness, soundness and robustness in Section 9.2

9.1 BHK games

The BHK_M game (parameterized by M), introduced in [1], proceed as follows. Two spatially separated parties, N_1 and N_2, receive $X, Y \in \{0, \cdots, M-1\}$ and output $A, B \in \{0, 1\}$ respectively. Any input pair (X, Y) is valid if and only if it is from the set $\mathrm{Input}_{\mathrm{BHK}_M} = \{(x, y) : y = x \text{ or } y = x+1 \mod M\}$. Let the indicator function I be defined as $I\{\text{true}\} = 1$, $I\{\text{false}\} = 0$. For any input-output pair (X, Y, A, B), the game value $\mathcal{B}$ is defined as

$$\mathcal{B}[A, B, X, Y] = \frac{1}{2} + M(A \oplus B \oplus I\{X = M-1\}I\{Y = 0\}). \tag{9.1}$$

The expected game value, denoted $\langle\mathcal{B}\rangle$, is over *uniformly* selected valid inputs. One could define the game value vector $|\mathrm{BHK}_M\rangle \in \mathbb{R}^{|A|\times|B|\times|X|\times|Y|}$ such that

$$\langle\mathcal{B}\rangle = \left\langle \mathrm{BHK}_M, \Gamma_{AB|XY} \right\rangle, \tag{9.2}$$

where $\Gamma_{AB|XY}$ is any NS strategy (state) for BHK_M games.Let us consider $\langle\mathcal{B}\rangle$ obtained by *classical* and *quantum* strategies for this game.

Classical Strategies. Without loss of generality, one can restrict to *deterministic* strategies. It is easy to see that any deterministic strategy will incur a penalty of M on at least one valid input, which happens with probability $1/2M$. Thus the expected game value $\langle\mathcal{B}\rangle$ satisfies,

$$\langle\mathcal{B}\rangle \geq \frac{1}{2} + M\frac{1}{2M} = 1.$$

Quantum Strategies. There turns out to be a good quantum strategy for BHK_M games achieving the following expected game value,

$$\langle\mathcal{B}\rangle = \frac{1}{2} + M\sin^2\left(\frac{\pi}{4M}\right) = \frac{1}{2} + O\left(\frac{1}{M}\right). \tag{9.3}$$

The particular quantum strategy is as follows. Let Alice and Bob share an EPR pair. For any $x \in \{0, \cdots, M-1\}$ for Alice, she performs the measurement in the orthogonal basis,

$$\{|0\rangle \mp e^{i\pi\frac{x}{M}}|1\rangle\}_{x\in\{0,\cdots,M-1\}},$$

while for any $y \in \{0, \cdots, M-1\}$ for Bob, he performs the measurement in the orthogonal basis,

$$\{|0\rangle \mp e^{-i\pi\frac{y+1/2}{M}}|1\rangle\}_{y\in\{0,\cdots,M-1\}}.$$

Thus for each valid input (x, y), the strategy loses with probability $\sin^2(\frac{\pi}{4M})$, which leads to Equ. (9.3).

Representation of NS-Strategies. Now we turn to a few properties of the representation of any NS strategy for BHK_M games, which were discovered in [16]. Adopting the notation in [16], we arrange any NS-strategy $\Gamma_{AB|XY}$ vector in $\mathbb{R}^{|A|\times|B|\times|X|\times|Y|}$ as follows,

$$\Gamma_{A,B|X,Y} = \begin{array}{|cc|c|c|} \hline P(0,0|0,0) & P(0,1|0,0) & \cdots & P(0,0|0,M-1) \\ P(1,0|0,0) & P(1,1|0,0) & & \\ \hline \multicolumn{2}{|c|}{\vdots} & \ddots & \vdots \\ \hline \multicolumn{2}{|l|}{P(0,0|M-1,0)} & \cdots & P(0,0|M-1,M-1) \\ \hline \end{array} \tag{9.4}$$

Moreover, we define two vectors $\vec{\mathbf{1}}, \nu$ of the same dimension as

$$\vec{\mathbf{1}} = \begin{array}{|cc|cc|cc|} \hline 1 & 1 & 1 & 1 & & \\ 1 & 1 & 1 & 1 & & \\ \hline & & 1 & 1 & \ddots & \\ & & 1 & 1 & & \\ \hline & & & & \ddots & \begin{matrix} 1 & 1 \\ 1 & 1 \end{matrix} \\ \hline 1 & 1 & & & \begin{matrix} 1 & 1 \\ 1 & 1 \end{matrix} & \\ 1 & 1 & & & & \\ \hline \end{array}, \quad \nu = \frac{1}{2} \begin{array}{|cc|cc|cc|} \hline 0 & 1 & 0 & 1 & & \\ -1 & 0 & -1 & 0 & & \\ \hline & & 0 & -1 & \ddots & \\ & & 1 & 0 & & \\ \hline & & & & \ddots & \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} \\ \hline 1 & 0 & & & \begin{matrix} 0 & -1 \\ 1 & 0 \end{matrix} & \\ 0 & -1 & & & & \\ \hline \end{array}, \tag{9.5}$$

where empty boxes are understood as having zeros

$$\boxed{} = \boxed{\begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix}}, \tag{9.6}$$

and ellipsis between two identical boxes are understood as an arbitrarily large sequence of identical boxes. In the following, we demonstrate a connection between Alice's input-output distribution and the expected game value. Let $\mu = \frac{1}{4M}\vec{\mathbf{1}}$. For any $a \in \{0,1\}$, define

$$\beta_a = \mu + (-1)^a \nu. \tag{9.7}$$

Moreover, we have

$$|\mathrm{BHK}_M\rangle = \mu + |\nu|, \tag{9.8}$$

where $|\nu|$ means entry-wise absolute value. It is easy to see that $\beta_a \preceq |\mathrm{BHK}_M\rangle, \forall a \in \{0,1\}$, where "$\preceq$" means entry-wise "$\leq$".

Now imagine one play k BHK_M games over k pairs of separate boxes in parallel (denote such a game by BHK_M^k) and assign the game value according to the vector $\left|\mathrm{BHK}_M^k\right\rangle = |\mathrm{BHK}_M\rangle^{\otimes k}$. Similarly, for any $\mathbf{a} \in \{0,1\}^k$, one could define

$$\beta_{\mathbf{a}} = \bigotimes_{i=1}^{k} \beta_{a_i}. \tag{9.9}$$

The following lemma provides an alternative characterization of Alice's side input-output distribution $\Gamma_{\mathbf{a}|\mathbf{x}}$ in terms of $\beta_{\mathbf{a}}$ and $\Gamma_{\mathbf{AB}|\mathbf{XY}}$.

Lemma 9.1 ([16], Lemma 6) *Assume* $\mathbf{P}_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}}$ *is an arbitrary* $2k$*-partite non-signaling strategy, then for any* $\mathbf{a} \in \{0,1\}^k$ *and any* $\mathbf{x} \in \{0,\cdots,M-1\}^{\otimes k}$*, we have* [7]

$$\Gamma_{\mathbf{a}|\mathbf{x}} = \beta_{\mathbf{a}} \cdot \Gamma_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}}. \tag{9.10}$$

Observe that, by definition, we have

$$\beta_{\mathbf{a}} = \beta_{a_1} \otimes \beta_{a_2} \otimes \cdots \otimes \beta_{a_k} \preceq |\mathrm{BHK}_M\rangle^{\otimes k} = \left|\mathrm{BHK}_M^k\right\rangle. \tag{9.11}$$

Also note that all entries in $\Gamma_{\mathbf{AB}|\mathbf{XY}}$ are nonnegative, thus we have

Corollary 9.2 *For any* $\mathbf{a} \in \{0,1\}^k$ *and any* $\mathbf{x} \in \{0,\cdots,M-1\}^{\otimes k}$*, we have*

$$\Gamma_{\mathbf{a}|\mathbf{x}} \leq \left\langle \mathrm{BHK}_M^k, \Gamma_{\mathbf{AB}|\mathbf{XY}} \right\rangle. \tag{9.12}$$

Let $\beta_{\text{uniform}} = (\frac{1}{2M}\vec{\mathbf{1}})^{\otimes k}$, which is a normalized "uniform" vector over all valid inputs and satisfies

$$\beta_{\text{uniform}} \cdot \Gamma_{\mathbf{AB}|\mathbf{XY}} = 1, \tag{9.13}$$

for any NS-strategy $\Gamma_{\mathbf{AB}|\mathbf{XY}}$ for BHK_M^k games. Moreover, we have

$$\sum_{\mathbf{a}\in\{0,1\}^k} \beta_{\mathbf{a}} = \left(\frac{1}{2M}\vec{\mathbf{1}}\right)^{\otimes k} = \beta_{\text{uniform}}. \tag{9.14}$$

Strong Monogamy

In this section we modularize and extend the technique in [15] about a strong *monogamy* phenomenon of BHK_M^k games. Roughly, it says that if a marginal NS state (strategy) plays BHK_M^k very well, then there exists a function $h : \{0,1\}^k \to \{0,1\}$ such that the bit $h(\mathbf{a})$, obtained from Alice's output $\mathbf{a} \in \{0,1\}^k$, is almost uniform to the environment (or the rest of the NS system), no matter how Alice, Bob, and the environment are correlated, i.e., no matter what global NS state is shared among Alice, Bob and the environment. We formulate this technique for BHK_M^k games and provide an *efficient* way to find such a function h.

Settings. Assume there is a $2k+1$-partite NS system, in which the first k NS boxes are denoted by $\mathbf{A} = (A_1,\cdots,A_k)$, and the second k NS boxes are denoted by $\mathbf{B} = (B_1,\cdots,B_k)$. The $(2k+1)$th NS box is denoted by E, which stands for the environment. The NS boxes $\mathbf{A}$ and $\mathbf{B}$ are playing the BHK_M^k games. Let $\Gamma_{\mathbf{A},\mathbf{B},F|\mathbf{X},\mathbf{Y},E}$ denote the NS strategy used by $\mathbf{A},\mathbf{B}$ and E, and hence $\Gamma_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}}$, a marginal NS state of $\Gamma_{\mathbf{A},\mathbf{B},F|\mathbf{X},\mathbf{Y},E}$, is the NS strategy used by $\mathbf{A}$ and $\mathbf{B}$ to play BHK_M^k.

We then apply *some* function $h : \{0,1\}^k \to \{0,1\}$ to Alice's output $\mathbf{a} \in \{0,1\}^k$ (given input $\mathbf{x}$) and output $z = h(\mathbf{a})$. Let $\mathcal{A}_z = \{\mathbf{a} \in \{0,1\}^k : h(\mathbf{a}) = z\}$ for any $z \in \{0,1\}$ and $\beta_{\mathcal{A}_z} = \sum_{\mathbf{a}\in\mathcal{A}_z} \beta_{\mathbf{a}}$. It follows from Lemma 9.1 that the distribution of z (i.e., $\Gamma_{z|\mathbf{x}}$) is given by

$$\Gamma_{z|\mathbf{x}} = \sum_{\mathbf{a}\in\mathcal{A}_z} \Gamma_{a|x} = \sum_{\mathbf{a}\in\mathcal{A}_z} \beta_{\mathbf{a}} \cdot \Gamma_{AB|XY} = \beta_{\mathcal{A}_z} \cdot \Gamma_{AB|XY}. \tag{9.15}$$

The particular function h in our consideration should satisfy the following property.

[7] The original statement of Lemma 6 in [16] claims only for input $\mathbf{x} = (0,\cdots,0)$. However, one can invoke the relabeling technique demonstrated in Lemma 7 of the same paper to extend the statement to all $\mathbf{x}$s.

Property 9.3 (Concentration) *Any function $h : \{0,1\}^k \to \{0,1\}$ is* C-concentrated *if and only if for any $z \in \{0,1\}$ such that,*

$$\left|\beta_{\mathcal{A}_z} - \frac{1}{2}\beta_{\text{uniform}}\right| \preceq C \left|\text{BHK}_M^k\right\rangle, \tag{9.16}$$

where β_{uniform} is defined in Eq.(9.13).

Proposition 9.4 *Let $\Gamma_{ABF|XYE}$ be the global NS-states shared among Alice, Bob and the environment, where Alice and Bob each have k NS boxes and play the BHK_M^k game. Let $h : \{0,1\}^k \to \{0,1\}$ a C-concentrated function. Let $\mathbf{a} \in \{0,1\}^k$ be Alice's output given any input $\mathbf{x}$ and $z = h(\mathbf{a})$ be the final output, and hence $\Gamma_{ZF|XE}$ the NS-state after Alice outputs $z = h(\mathbf{a})$. Then we have*

$$\Delta(\Gamma_{ZF|XE}, U_1 \otimes \Gamma_{F|E}) \leq 2C \left\langle \text{BHK}_M^k, \Gamma_{AB|XY} \right\rangle. \tag{9.17}$$

Note that proof of the above proposition is a modularization of the technique in [15] in our NS language. We defer the proof to Appendix B.

Randomness-efficient Generation of Well-concentrated Functions

It then suffices to find C-concentrated functions with reasonable Cs. It was found in [15] that fully random hashing functions from $\{0,1\}^k$ to $\{0,1\}$ are $2^{O(\sqrt{k})}$-concentrated. However, generating such hashing functions requires 2^k uniform bits. In the following we claim that $\Theta(k)$-wise random hashing functions (requiring only $\Theta(k)$ uniform bits to generate) are reasonably concentrated, which leads to an *efficient* protocol to generate desired functions. The proof goes through a careful concentration analysis of $\Theta(k)$-wise random hashing functions which we defer to Appendix B.

Lemma 9.5 *Any t-wise random hash function $h : \{0,1\}^k \to \{0,1\}$ is $2^{\frac{1}{2}(k+O(\log(k)))}$*-concentrated *with probability $1 - 2^{-\Omega(k)}$ when $t = \Theta(k)$.*

9.2 The Protocol

We construct a concrete NS DI-RA protocol with local-uniform soundness (i.e., serving as the decoupler) based on BHK games. We *directly* prove its local-uniform soundness against non-signaling adversaries by making use of a *strong monogamy* phenomenon of BHK games. The protocol (denoted Π_{BHK}) proceeds as in Figure 5. The parameter ϵ_B is the soundness parameter of this protocol, while γ is some parameter related to the robustness and M is the parameter for BHK games.

QT Strong Completeness & Strong Robustness

Our design of protocol Π_{BHK} requires $v_{-r} \leq \alpha n$ to accept. For honest boxes, this roughly requires each play of the BHK_M game (not BHK_M^k) must obtain game value no more than $\frac{1-\gamma}{\sqrt{2}}$. Let $0 \leq \delta_n \leq 1$ denote the noise parameter. Namely, each round's output of playing BHK_M game might be changed with δ_n chance.

Protocol Π_{BHK}

1. Let S be the random bits and parameters ϵ_B, γ, M, both be given as input. Let $k = \Theta((1/\gamma) \cdot \log(1/(\gamma \cdot \epsilon_B)))$, $\alpha = \left((1-\gamma)/\sqrt{2}\right)^k$, $c = (1/2 + M)^k$ and $n = \Theta((c^2/\alpha^2) \cdot \log(1/\epsilon_B))$. The input randomness S of length $\Theta(nk)$ is split into $(\mathbf{X}, \mathbf{Y}, R, H)$.

2. Run BHK_M^k games for n times in parallel on $n \times k$ pairs of spatially separated boxes (nonlocal boxes) with random valid inputs generated from $\mathbf{X}, \mathbf{Y}$.

3. Let $w_1, \ldots, w_n \in [k]$ denote the number of acceptances in each BHK_M^k game and define $v_i = (1/2)^{w_i} \cdot (1/2 + M)^{k - w_i}$ for $i \in [n]$.

4. Use R to select a uniformly random index $r \in [n]$. Let $v_{-r} = \sum_{i \neq r} v_i$.

   4.1 If $v_{-r} \leq \alpha n$, then **Accept**. (Set $O = \mathsf{Acc}$ (i.e., 1).) Moreover, use H to generate a $\Theta(k)$-wise independent hash function $h : \{0,1\}^k \to \{0,1\}$, then output $Z = h(\mathbf{a}_r)$ where $\mathbf{a}_r$ is Alice side's output of the rth BHK_M^k game.

   4.2 Otherwise, **Reject**. (Set $O = \mathsf{Rej}$ (i.e., 0).)

Figure 5: Protocol Π_{BHK} for BHK_M^k games.

QT Strong Completeness. The completeness of Π_{BHK} can be achieved by invoking the quantum strategy demonstrated in Section 9.1, which achieves the rejection probability $\sin^2(\frac{\pi}{4M})$ for each play of BHK_M game, for whatever input (x, y). This implies strong completeness. Taking into account the noise δ_n, for each play of BHK_M game, the honest boxes can achieve game values no more than

$$\frac{1}{2} + M\left(\sin^2\left(\frac{\pi}{4M}\right) + \delta_n\right),$$

which is smaller than $\frac{1-\gamma}{\sqrt{2}}$ by choosing appropriate M and small enough γ and δ_n. When this condition is achieved, by concentration of measures, we have for this strategy, the protocol Π_{BHK} accepts with probability $1 - 2^{-\Omega(nk)}$ on whatever source.

Strong Robustness. To optimize robustness, it suffices to choose appropriate M, γ, δ_n that satisfy the following constraint and maximize δ_n,

$$\frac{1}{2} + M\left(\sin^2\left(\frac{\pi}{4M}\right) + \delta_n\right) \leq \frac{1-\gamma}{\sqrt{2}}.$$

It follows easily that one can choose $M = 6$, $\gamma = 10^{-8}$ to get $\delta_n \approx 1.748\%$.

NS Soundness Analysis of Π_{BHK}

Let $\Gamma_{\mathbf{S},\mathbf{A},\mathbf{B},\mathbf{E}|\perp,\mathbf{X},\mathbf{Y},\mathbf{M}}$ be the NS state of the entire NS system consisting of a seed S that is uniform to the rest of the system, the NS boxes (called the device part) with input $\mathbf{X}, \mathbf{Y}$ and output $\mathbf{A}, \mathbf{B}$ playing n BHK_M^k games, and the environment with input M and output E, which is the rest of the world. Thus, the specific NS strategy to play n BHK_M^k games is the marginal NS state $\Gamma_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}}$.

In the context of protocol Π_{BHK}, the seed S contains $\mathbf{X}, \mathbf{Y}, R, H$, where $\mathbf{X}, \mathbf{Y}$ are fed to the NS boxes as inputs[8], R is the output round, H is the randomness to generate the hash function, all of which are independent of each other. Moreover, we know that

$$\Gamma_{\mathbf{S},\mathbf{A},\mathbf{B}|\perp,\mathbf{X},\mathbf{Y}} = \mathbf{U}_{\mathbf{S}} \otimes \Gamma_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}}.$$

Furthermore, we give n BHK_M^k games an *artificial* order and imagine BHK_M^k games are played sequentially (though on different boxes) in this order.

Probability Space. We are interested in the probability space generated by the seed-device system, i.e., by the marginal NS state $\Gamma_{\mathbf{S},\mathbf{A},\mathbf{B}|\perp,\mathbf{X},\mathbf{Y}}$. It should be understood that the seed-device system connects with the environment (or, the rest of the world) through some NS correlation, given by $\Gamma_{\mathbf{S},\mathbf{A},\mathbf{B},\mathbf{E}|\perp,\mathbf{X},\mathbf{Y},\mathbf{M}}$. Specifically, we consider the probability space consists of $\Omega = (\mathbf{X}, \mathbf{Y}, \mathbf{A}, \mathbf{B}, R, H, O)$, where $O \in \{\mathsf{Acc}, \mathsf{Rej}\}$ is the indicator of accept/reject of the protocol and is a deterministic function of $\mathbf{X}_{-r}, \mathbf{Y}_{-r}, \mathbf{A}_{-r}, \mathbf{B}_{-r}$ when $R = r$.

Let $\mathsf{hist}_{<r}$ be a set of input-output pairs from the first $r-1$ out of n BHK_M^k games. We define the following set of random variables over the probability space Ω. Initialize $\mathsf{hist}_{<1} = \emptyset$. For each $i = 1, \ldots, n$, we define the empirical score v_i and the strategy score μ_i.

- Let the strategy score be $\mu_i = \langle \mathrm{BHK}_M^k | \Gamma_{\mathbf{A}_i,\mathbf{B}_i|\mathbf{X}_i,\mathbf{Y}_i,\mathsf{hist}_{<i}} \rangle$. [9]
- Sample independently uniform bits $\mathbf{x}_i, \mathbf{y}_i$ and sample $\mathbf{a}_i, \mathbf{b}_i$ according to $\Gamma_{\mathbf{A}_i,\mathbf{B}_i|\mathbf{x}_i,\mathbf{y}_i,\mathsf{hist}_{<i}}$. Set $\mathsf{hist}_{<i+1} = \mathsf{hist}_{<i} \circ (\mathbf{x}_i, \mathbf{y}_i, \mathbf{a}_i, \mathbf{b}_i)$.
- Let w_i be the number of accepts in $(\mathbf{x}_i, \mathbf{y}_i, \mathbf{a}_i, \mathbf{b}_i)$ and $v_i = (1/2)^{w_i} \cdot (1/2 + M)^{k-w_i}$ (as in Fig. 5).

Let $\mu = \sum_i \mu_i$ and $v = \sum_i v_i$. Let $\mu_{-r} = \sum_{i \neq r} \mu_i$ and $v_{-r} = \sum_{i \neq r} v_i$. Let Acc denote the event that the protocol accepts (i.e., the event $\{v_{-r} \leq \alpha n\}$). Let Rej denote the event otherwise. Let $\alpha = ((1-\gamma)/\sqrt{2})^k$ and $c = (1/2 + M)^k$.

Analysis. The power of *strong monogamy* of BHK_M^k games allows us to bound the soundness error (distance to uniform) by only looking at the seed-device system. Specifically, we only look at the probability space Ω to analyze the error. One can imagine that Ω is equivalently sampled in the following way.

- We first sample uniform and independent $r \leftarrow R$ and $h \leftarrow H$.
- We then sample uniform and independent $\mathbf{X}_{<r}, \mathbf{Y}_{<r}$ and feed them into the first $r-1$ NS boxes of n BHK_M^k games to produce $\mathsf{hist}_{<r} = (\mathbf{x}_{<r}, \mathbf{y}_{<r}, \mathbf{a}_{<r}, \mathbf{b}_{<r})$.
- We then sample the decision bit $o \leftarrow O|_{r,h,\mathsf{hist}_{<r}}$. Note that once r is fixed, the decision bit o is independent of the input-output pair of the r-th game, because o does not depend on the r-th game's input-output and the NS condition between the r-th game and the rest of games.

[8] We don't distinguish the specific input and all possible inputs to the NS boxes in our notation, as the protocol will use that specific input.

[9] $\Gamma_{\mathbf{A}_i,\mathbf{B}_i|\mathbf{X}_i,\mathbf{Y}_i,\mathsf{hist}_{<i}}$ should be understood as the conditional NS state for the ith game on the event $\mathsf{hist}_{<i}$.

Now, if $o = \mathsf{Rej}$, then we don't get any error. If $o = \mathsf{Acc}$, then we bound the error as follows: If h is bad , then we use the trivial bound 1. If h is good, then the error can be bounded by $2C \cdot \mu_r(\mathsf{hist}_{<r}, \mathsf{Acc})$ by Prop. 9.4. We know that h is good with high probability by Lemma 9.5. It suffices to know that the bad event that $o = \mathsf{Acc}$ yet $\mu_r(\mathsf{hist}_{<r}, \mathsf{Acc})$ is large happens with a very small probability, where

$$\mu_r(\mathsf{hist}_{<r}, \mathsf{Acc}) = \langle \mathrm{BHK}_M^k | \Gamma_{\mathbf{A}_r, \mathbf{B}_r | \mathbf{X}_r, \mathbf{Y}_r, \mathsf{hist}_{<r}, \mathsf{Acc}} \rangle. \tag{9.18}$$

All of the above shall imply that the soundness error is small.

Obstacles & Solutions. There are two major obstacles of doing so. First, we do not know how to reason about $\mu_r(\mathsf{hist}_{<r}, \mathsf{Acc})$ directly, as it is not directly inferred from the Accept/Reject condition of the protocol. Instead, we relate $\mu_r(\mathsf{hist}_{<r}, \mathsf{Acc})$ with $\mu_r(\mathsf{hist}_{<r})$, i.e., the strategy without conditioned on accepting. Specifically, we show that,

Lemma 9.6 *Let r be the random index in Protocol B, for every $r, \mathsf{hist}_{<r}$, then we have*

$$\mu_r(\mathsf{hist}_{<r}, \mathsf{Acc}) \leq \frac{\mu_r(\mathsf{hist}_{<r})}{\Pr[\mathsf{Acc} | r, \mathsf{hist}_{<r}]}.$$

This statement makes sense intuitively: as the score μ_r is a linear function on the strategy, conditioning on the Acc event should increase the score at most by the probability that Acc happens. To make the intuition correct, technically this step crucially relies on the fact that the n rounds BHK_M^k games are played using different devices that are NS to each other. Precisely,

Proof. [of Lemma 9.6] For every $\mathbf{a}_r, \mathbf{b}_r, \mathbf{x}_r, \mathbf{y}_r$, we have

$$\Gamma_{\mathbf{a}_r, \mathbf{b}_r | \mathbf{x}_r, \mathbf{y}_r, \mathsf{hist}_{<r-1}, \mathsf{Acc}} \leq \frac{\Gamma_{\mathbf{a}_r, \mathbf{b}_r | \mathbf{x}_r, \mathbf{y}_r, \mathsf{hist}_{<r-1}}}{\mathbf{Pr}[\mathsf{Acc} | \mathbf{x}_r, \mathbf{y}_r, r, \mathsf{hist}_{<r-1}]},$$

because $\mathbf{Pr}[A|BC] \leq \mathbf{Pr}[A|B]/\mathbf{Pr}[C|B]$ for any event A, B, C. Note that Acc does not depend on the rth game. Moreover, by NS-condition and the fact $\mathbf{x}_r, \mathbf{y}_r$ are independently sampled, different $\mathbf{x}_r, \mathbf{y}_r$ won't disturb the input-output distribution of the rest $n-1$ games. Thus, we have

$$\mathbf{Pr}[\mathsf{Acc} | \mathbf{x}_r, \mathbf{y}_r, r, \mathsf{hist}_{<r-1}] = \mathbf{Pr}[\mathsf{Acc} | r, \mathsf{hist}_{<r-1}]. \tag{9.19}$$

By definition and note that all entries of $\left|\mathrm{BHK}_M^k\right\rangle$ are nonnegative, we have

$$\mu_r(\mathsf{hist}_{<r}, \mathsf{Acc}) \leq \frac{\mu_r(\mathsf{hist}_{<r})}{\Pr[\mathsf{Acc} | r, \mathsf{hist}_{<r}]}.$$

■

Second, we then show the BAD event that $o = \mathsf{Acc}$ yet $\mu_r(\mathsf{hist}_{<r})$ is large happen with small probability. Intuitively, the reason is as follows: by the concentration of martingales, the event that $\mu = \sum \mu_i$ and $v = \sum v_i$ are close with high probability. Moreover, notice that this event is independent of R and thus conditioning on it won't change R's distribution. Now, observe that $o = \mathsf{Acc}$ means v_{-r} is small, which implies that v and μ are also small. Therefore, there cannot be too many $r \in [n]$ that cause $o = \mathsf{Acc}$ yet $\mu_r = \mu_r(\mathsf{hist}_{<r})$ is large. This gives an upper bound on the probability that the BAD event can happen.

The following lemma establishes the fact that μ and v are close with high probability.

Lemma 9.7 *For every $\delta \in (0,1)$, we have* $\Pr[v \le \mu - \delta n] \le e^{-\delta^2 n/2c^2}$.

Proof. We introduce martingales and the relevant concentration property for unfamiliar audience.

Martingale. Define a *filter* $\{\mathcal{F}_i : i = 0, \cdots n\}$ as an increasing sequence of σ-fields such that $\emptyset = \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}_n$ on some probability space. Let $\{X_i\}$ be a sequence of random variables such that X_i is measurable with respect to $\mathcal{F}_i$. We call $\{X_i\}$ a *martingale* with respect to $\{\mathcal{F}\}$ if $\forall i, \mathrm{E}[X_i|\mathcal{F}_{i-1}] = X_{i-1}$.

Lemma 9.8 (Azuma-Hoeffding) *Let X be a martingale associated with a filter $\mathbf{F}$ such that $|x_k - x_{k-1}| \le c_k$ for all k. Then for all integers m and $\lambda \ge 0$,*

$$\Pr[X_m \le \mathrm{E}[X_m] - \lambda] \le e^{-\frac{\lambda^2}{2(\sum_{k=1}^m c_k^2)}}. \tag{9.20}$$

The lemma then follows by observing that the random variables defined above form a martingale and applying Lemma 9.8. Formally, for $i \in \{0, \ldots, n\}$, let $\mathcal{F}_i$ be the σ-field generated by cond_i, which yields a natural filter $\mathbf{F}$: $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_n$. By definition, for every $i \in [n]$ and cond_{i-1}, $\mathrm{E}[v_i|\mathrm{cond}_{i-1}] = \mu_i(\mathrm{cond}_{i-1})$. Thus, define $P_i \stackrel{\text{def}}{=} v_i - \mu_i$, and $Q_i = \sum_{j \le i} P_j$ for $i \in [n]$, and we have that $Q_0 = 0, Q_1, \ldots, Q_n$ forms a martingale corresponding to the filter $\mathbf{F}$ with $|Q_i - Q_{i-1}| \le c$. By Lemma 9.8 with $\lambda = \delta n$ and note that $\mathrm{E}[Q_n] = 0$, we have

$$\Pr[Q_n \le -\delta n] \le e^{-\frac{\lambda^2}{2nc^2}} \le e^{-\delta^2 n/2c^2}.$$

By definition, $Q_n = v - \mu$, thus we have $\Pr[v \le \mu - \delta pn] \le e^{-\delta^2 n/2c^2}$. ■

Now we complete the second step with the following lemma.

Lemma 9.9 *For every $\delta, \kappa \in (0,1)$, let $\eta = \alpha + \delta + c/n$, then we have,*

$$\Pr[o = \mathsf{Acc} \wedge \mu_r(\mathsf{hist}_{<r}) \ge \eta/\kappa] \le e^{-\delta^2 n/2c^2} + \kappa.$$

Proof. Let C denote that event that $v \le \mu - \delta n$, which happens with at most probability $e^{-\delta^2 n/2c^2}$ by Lemma 9.7. Thus,

$$\begin{aligned} \Pr[o = \mathsf{Acc} \wedge \mu_r(\mathsf{hist}_{<r}) \ge \eta/\kappa] \;=\; & \Pr[C \wedge o = \mathsf{Acc} \wedge \mu_r(\mathsf{hist}_{<r}) \ge \eta/\kappa] + \\ & \Pr[\overline{C} \wedge o = \mathsf{Acc} \wedge \mu_r(\mathsf{hist}_{<r}) \ge \eta/\kappa], \end{aligned} \tag{9.21}$$

where the first term on the RHS is upper bounded by $\Pr[C] \le e^{-\delta^2 n/2c^2}$. It suffices to bound the second term then, where $\overline{C}$ holds, namely, $\mu \le v + \delta n$.

Now we apply the observation that the event $\overline{C}$ is independent of random variable R, i.e., $\Pr[R = r|\overline{C}] = \mathbf{Pr}[R = r] = 1/n$. Define a new random variable $\tilde{\mu}_r$ as

$$\tilde{\mu}_r = \mu_r(\mathsf{hist}_{<r}) \text{ if } o = \mathsf{Acc};\ \text{otherwise } = 0.$$

It is easy to see that, for any γ, the event $\tilde{\mu}_r \ge \gamma$ is equivalent to $o = \mathsf{Acc} \wedge \mu_r(\mathsf{hist}_{<r}) \ge \gamma$. Thus,

$$\Pr[\tilde{\mu_r} \ge \gamma|\overline{C}] = \Pr[o = \mathsf{Acc} \wedge \mu_r(\mathsf{hist}_{<r}) \ge \gamma|\overline{C}],$$

where the RHS upper bounds $\Pr[\overline{C} \wedge o = \mathsf{Acc} \wedge \mu_r(\mathsf{hist}_{<r}) \ge \gamma]$.

It suffices to calculate $\mathrm{E}[\tilde{\mu_r}|\overline{C}]$ and apply the Markov inequality. To that end, we have

$$\mathrm{E}[\tilde{\mu_R}|\overline{C}] = \sum_{r=1}^{n} \Pr[R=r|\overline{C}]\,\mathrm{E}[\tilde{\mu_r}|\overline{C}] = \frac{1}{n}\,\mathrm{E}\left[\sum_{r=1}^{n} \tilde{\mu_r}|\overline{C}\right].$$

By the definition of $\tilde{\mu_r}$, it is non-zero only when Acc happens, i.e., $\mathrm{E}[\tilde{\mu_r}|\overline{\mathsf{Acc}},\overline{C}] = 0$. The Acc condition then implies that $v_{-r} \le \alpha n$. Hence $v \le v_{-r} + c$ as c is the largest game value for a single BHK_M^k game. Therefore, with $\eta = \alpha + \delta + c/n$, we have

$$\mathrm{E}\left[\sum_{r=1}^{n} \tilde{\mu_r}|\overline{C}\right] \le \mathrm{E}[\mu|\mathsf{Acc},\overline{C}] \le \mathrm{E}[v+\delta n|\mathsf{Acc},\overline{C}] \le \alpha n + c + \delta n = \eta n.$$

Namely, we have that $\mathrm{E}[\tilde{\mu_R}|\overline{C}] \le \eta$. Thus, by the Markov inequality, we have

$$\Pr[\tilde{\mu_r} \ge \eta/\kappa|\overline{C}] \le \frac{\mathrm{E}[\tilde{\mu_R}|\overline{C}]}{\eta/\kappa} \le \kappa,$$

which bounds the second term in the RHS of Eq. (9.21). Thus, we have

$$\Pr[o = \mathsf{Acc} \wedge \mu_r(\mathsf{hist}_{<r}) \ge \eta/\kappa] \le e^{-\delta^2 n/2c^2} + \kappa,$$

which completes the proof. ■

As a simple corollary of Lemma 9.6 and Lemma 9.9, we have

Corollary 9.10 *For every $\delta, \kappa \in (0,1)$, let $\eta = \alpha + \delta + c/n$, then we have,*

$$\Pr\left[o = \mathsf{Acc} \wedge \left(\mu_r(\mathsf{hist}_{<r}, \mathsf{Acc}) \ge \frac{(\eta/\kappa)}{\Pr[\mathsf{Acc}|r, \mathsf{hist}_{<r}]}\right)\right] \le e^{-\delta^2 n/2c^2} + \kappa.$$

Remark. This is the critical property to derive the soundness of the protocol. As we will see in Section 10, this property still holds when the input is ϵ close to uniform through a rather complicated NS reduction. Once this property is established, one can almost readily derive the soundness of the protocol as follows.

Summary of Parameters. For readability, we summarize our parameter choices (shown in Fig. 5) as follows. The reason behind these choices will become transparent in the proof. The parameters M and γ are used to optimize the robustness of the protocol. Both are constants. E.g., $M = 6$ and $\gamma = 10^{-8}$ for robust noise 1.748%. We set other parameters to achieve soundness error ϵ_B as follows.

- First, we set $k = O(\frac{1}{\gamma}\log\frac{1}{\gamma\epsilon_B}) = O(\log(1/\epsilon_B))$. This determines $\alpha = \left(\frac{(1-\gamma)}{\sqrt{2}}\right)^k = \mathrm{poly}(\epsilon_B)$, $c = (M+1/2)^k = \mathrm{poly}(1/\epsilon_B)$, and $C = 2^{0.5k+O(\log k)} = \mathrm{poly}(1/\epsilon_B)$. Also we set $\delta = \alpha$ so $\eta < 3\alpha$, and $\kappa = \epsilon_B/4$.
- We set $n = O\left(\frac{c^2}{\alpha^2}\log\frac{1}{\epsilon_B}\right) = \mathrm{poly}(1/\epsilon_B)$. The requied seed length is dominated by $O(nk) = \mathrm{poly}(1/\epsilon_B)$.

Theorem 9.11 *Let $\Gamma_{\mathbf{S},\mathbf{A},\mathbf{B},\mathbf{E}|\perp,\mathbf{X},\mathbf{Y},\mathbf{M}}$ be any NS state of the NS system consisting of the seed S, the device $\mathbf{D}=(\mathbf{A},\mathbf{B})$ and the environment E. Assume that the seed S is uniform to the device $\mathbf{D}$, i.e.,*

$$\Gamma_{\mathbf{S},\mathbf{A},\mathbf{B}|\perp,\mathbf{X},\mathbf{Y}} = \mathbf{U}_{\mathbf{S}} \otimes \Gamma_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}}.$$

Then executing the protocol Π_{BHK} (shown in Fig. 5) on $\Gamma_{\mathbf{S},\mathbf{A},\mathbf{B},\mathbf{E}|\perp,\mathbf{X},\mathbf{Y},\mathbf{M}}$ leads to an NS state $\Gamma_{O,Z,E|\perp,\perp,M}$ such that

$$\Delta(\Gamma_{\mathsf{Acc},Z,E|\perp,\perp,M}, U_1 \otimes \Gamma_{\mathsf{Acc},E|\perp,M}) \leq O(\epsilon_B).$$

Or, equivalently, we have that Π_{BHK} has NS soundness error $O(\epsilon_B)$ when the source S is local-uniform.

Proof. Let $C = 2^{0.5k+O(\log(k))}$ and denote the event that the hashing function h is *not* C-concentrated by B_1. By Lemma 9.5, we have $\mathbf{Pr}[\mathrm{B}_1 \wedge \mathsf{Acc}] \leq \mathbf{Pr}[\mathrm{B}_1] \leq 2^{-\Omega(k)}$. Note further that B_1 is independent of the events happening on the device side.

Moreover, let B_2 denote the event that $o = \mathsf{Acc} \wedge \mu_r(\mathsf{hist}_{<r}, \mathsf{Acc}) \geq \frac{(\eta/\kappa)}{\Pr[\mathsf{Acc}|r,\mathsf{hist}_{<r}]}$, where $\eta = \alpha + \delta + c/n$ and $0 < \delta, \kappa < 1$. By Corollary 9.10, we have $\mathbf{Pr}[\mathrm{B}_2] \leq e^{-\delta^2 n/2c^2} + \kappa$.

When either B_1 or B_2 happens, we will use the trivial upper bound 1 to bound the distance. Otherwise, we have $\mu_r(\mathsf{hist}_{<r}, \mathsf{Acc}) \leq \frac{(\eta/\kappa)}{\Pr[\mathsf{Acc}|r,\mathsf{hist}_{<r}]}$ and will use Proposition 9.4 to bound the distance.

Precisely, we follow the way we sample the probability space $\Omega = (r, h, \mathsf{hist}_{<r}, o)$, and apply the above strategy,

$$\begin{aligned}
\Delta(\Gamma_{\mathsf{Acc},Z,E|\perp,\perp,M}, U_1 \otimes \Gamma_{\mathsf{Acc},E|\perp,M}) &\leq \mathbf{Pr}[\mathrm{B}_1 \wedge \mathsf{Acc}] + \mathbf{Pr}[\mathrm{B}_2] \\
&+ \sum_{r,\mathsf{hist}_{<r}} \Pr[r, \mathsf{hist}_{<r}] \cdot \Pr[\mathsf{Acc}|r, \mathsf{hist}_{<r}] \cdot \left(2C \cdot \frac{(\eta/\kappa)}{\Pr[\mathsf{Acc}|r, \mathsf{hist}_{<r}]}\right) \\
&\leq 2^{-\Omega(k)} + (e^{-\delta^2 n/2c^2} + \kappa) + \sum_{r,\mathsf{hist}_{<r}} \Pr[r, \mathsf{hist}_{<r}] \cdot (2C \cdot (\eta/\kappa)) \\
&\leq 2^{-\Omega(k)} + (e^{-\delta^2 n/2c^2} + \kappa) + 2C \cdot (\eta/\kappa).
\end{aligned}$$

To bound the above quantity by ϵ_B, we simply choose $\delta = \alpha$ (which implies $\eta < 3\alpha$) and $\kappa = \epsilon_B/4$ and choose sufficiently large constant in $k = O((1/\gamma)\log(1/\gamma\epsilon_B))$; these are sufficient to bound each term by $\epsilon_B/4$. ■

10 Handling Close-to-uniform Seed

We want to show that the protocol Π_{BHK} works even if the source-device system is only ϵ-local-uniform, i.e.,

$$\Delta(\Gamma_{\mathbf{S},\mathbf{A},\mathbf{B}|\perp,\mathbf{X},\mathbf{Y}}, U_S \otimes \Gamma_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}}) \leq \epsilon.$$

To that end, we will re-investigate the proof of the soundness of Π_{BHK}, identify a key property, and prove its validity (up to some parameter loss) even when the source is only ϵ-local-uniform (see Lemma 10.1). The soundness then follows (see Theorem 10.2) by a similar argument in Theorem 9.11.

The setup. Let $S = (X, Y, R, H)$ denote the source, and U_S denote the uniform seed similar to Section 9. The whole source-device-environment system can be described by $\Gamma_{\mathbf{S},\mathbf{A},\mathbf{B},\mathbf{E}|\perp,\mathbf{X},\mathbf{Y},\mathbf{M}}$, and we know that the source-device part $\Gamma_{\mathbf{S},\mathbf{A},\mathbf{B}|\perp,\mathbf{X},\mathbf{Y}}$ is ϵ-close to $U_S \otimes \Gamma_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}}$. Our goal is to show that the analysis of protocol Π_{BHK} still works for $\Gamma_{\mathbf{S},\mathbf{A},\mathbf{B},\mathbf{E}|\perp,\mathbf{X},\mathbf{Y},\mathbf{M}}$ with small soundness error.

Notation. We say $\Gamma_{S,A,B|\perp,X,Y}$ is the Real NS state, and $\mathbf{U_S} \otimes \Gamma_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}}$ is the Ideal NS state. $\Pi_{\mathrm{BHK}}(\Gamma_{S,A,B|\perp,X,Y})$ denotes the experiment of applying the protocol Π_{BHK} to the NS state $\Gamma_{S,A,B|\perp,X,Y}$, which results in a classical probability distribution.

Strategy. Towards this goal, we need to go into the analysis of Π_{BHK} for uniform seed. Let us first recall the high level structure of the analysis. There, we first bound the probability of two bad events: (i) let BAD_1 denote the event that $h \leftarrow H$ is bad, and (ii) let $\mathsf{BAD}_2(\gamma)$ denote the event that $o = \mathsf{Acc}$ and $\mu_r(\mathsf{hist}_{<r}, \mathsf{Acc}) \geq \gamma$ for some γ. Suppose we can bound $\Pr[\mathsf{BAD}_1] \leq \delta_1$, and $\Pr[\mathsf{BAD}_2(\gamma)] \leq \delta_2$. Then due to Theorem 9.11, the soundness error is bounded by

$$\delta_1 + \delta_2 + 2C \cdot \gamma.$$

We will show that when we apply Π_{BHK} to $\Gamma_{S,A,B|\perp,X,Y}$, it is still the case that both BAD_1 and $\mathsf{BAD}_2(\cdot)$ events happen with small probability (for slightly larger parameters $\delta_1', \delta_2', \gamma'$). This is sufficient to bound the soundness error of Π_{BHK} for $\Gamma_{S,A,B|\perp,X,Y}$ following the same proof of Theorem 9.11. For BAD_1, because it is an event only depending on part of the seed S, it is easy to see that $\Pr[\mathsf{BAD}_1] \leq \delta_1 + \epsilon$, since otherwise, S is not even ϵ-close to U_S marginally.

However, handling BAD_2 is significantly more non-trivial; in particular, we cannot assert that $\Pr[\mathsf{BAD}_2(\gamma)] \leq \delta_2 + \epsilon$. The reason is subtle because that the definition of $\mu_r(\mathsf{hist}_{<r}, \mathsf{Acc})$ *depends* on the underlying probability distribution of the experiment (namely, the strategy of the r-th round boxes conditioned on $\mathsf{hist}_{<r}, \mathsf{Acc}$), and hence $\mathsf{BAD}_2(\gamma)$ refers to *different* events when the underlying probability space is $\Pi_{\mathrm{BHK}}(\Gamma_{S,A,B|\perp,X,Y})$ and $\Pi_{\mathrm{BHK}}(\mathbf{U_S} \otimes \Gamma_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}})$.

To bridge the gap, we develop a novel reduction argument for NS protocols. Specifically, we show that if $\Pr[\mathsf{BAD}_2(2\gamma)] \geq 2\delta_2$ in $\Pi(\mathsf{Real})$, then Real and Ideal are not ϵ-close for sufficiently small ϵ, by constructing a NS distinguisher D that distinguishes Real from Ideal with advantage $> \epsilon$. At a high level, the main idea is to make $\mu_r(\mathsf{hist}_{<r}, \mathsf{Acc})$ "indirectly observed" as a conditional expectation of a random variable in the resulting classical distribution; this is done by applying a modified protocol Π' to the NS state. We can then lower bound the statistical distance between the resulting distribution of $\Pi'(\mathsf{Real})$ and $\Pi'(\mathsf{Ideal})$ based on the gap between expectations $\mu_r(\mathsf{hist}_{<r}, \mathsf{Acc})$.

To that end, we prove the following key lemma.

Lemma 10.1 *Let* $\mathsf{Real} = \Gamma_{S,A,B|\perp,X,Y}$, $\mathsf{Ideal} = \mathbf{U_S} \otimes \Gamma_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}}$. *If in* $\Pi_{\mathrm{BHK}}(\mathsf{Ideal})$,

$$\Pr[o = \mathsf{Acc} \wedge \mu_r(\mathsf{hist}_{<r}, \mathsf{Acc}) \geq \gamma] \leq \delta_2,$$

and Real *and* Ideal *are* ϵ*-close for* $\epsilon < \delta_2\gamma/(4(M+1)^k)$, *then in* $\Pi_{\mathrm{BHK}}(\mathsf{Real})$,

$$\Pr[o = \mathsf{Acc} \wedge \mu_r(\mathsf{hist}_{<r}, \mathsf{Acc}) \geq 2\gamma] \leq 2\delta_2,$$

Proof. Suppose for the sake of contradiction that in $\Pi_{\mathrm{BHK}}(\mathsf{Real})$,

$$\Pr[o = \mathsf{Acc} \wedge \mu_r(\mathsf{hist}_{<r}, \mathsf{Acc}) \geq 2\gamma] > 2\delta_2.$$

We show that Real and Ideal can be distinguished with advantage $> \epsilon$. Define the following slight variant protocol Π'_{BHK} of Π_{BHK}.

> **Protocol** Π'_{BHK}: Π'_{BHK} is identical to Π_{BHK}, except that it does not feed (X_r, Y_r) as input to the r-th round boxes. Instead, it samples independent and uniform input (X'_r, Y'_r) to feed to the r-th round boxes.

Note that using the original protocol Π_{BHK} on ϵ-close-uniform seed might lead to (X_r, Y_r) that is neither independent nor uniform. On the contrary, Π'_{BHK} guarantees to sample independent and uniform input (X'_r, Y'_r) to the r-th round boxes.

We consider a distinguisher D that applies Π'_{BHK} to Real and Ideal and distinguishes the resulting classical distribution. We show that the distinguishing advantage of D, which is the statistical distance of $\Pi'_{\mathrm{BHK}}(\mathsf{Real})$ and $\Pi'_{\mathrm{BHK}}(\mathsf{Ideal})$, is $> \epsilon$. A crucial observation, which is the reason to consider Π'_{BHK}, is that we have

$$\mu_r(\mathsf{hist}_{<r}, \mathsf{Acc}) = \mathrm{E}[v_r | \mathsf{hist}_{<r}, \mathsf{Acc}].$$

We will use this equality to lower bound the statistical distance. For expositional clarity, we introduce the following notation. We focus on the following two "components" of the resulting classical distributions: Let $C_1 = (R, \mathsf{HIST}_{<R}, O)$ and $C_2 = V_R$. Furthermore, let $(C_1^{(\mathsf{Real})}, C_2^{(\mathsf{Real})})$ and $(C_1^{(\mathsf{Ideal})}, C_2^{(\mathsf{Ideal})})$ denote the distributions resulting from $\Pi'_{\mathrm{BHK}}(\mathsf{Real})$ and $\Pi'_{\mathrm{BHK}}(\mathsf{Ideal})$ respectively. With these notations and the above equality, we have

$$\Pr_{c_1 \leftarrow C_1^{(\mathsf{Ideal})}} \left[o = \mathsf{Acc} \wedge \left(\mathrm{E}\left[C_2^{(\mathsf{Ideal})} | C_1^{(\mathsf{Ideal})} = c_1 \right] \geq \gamma \right) \right] \leq \delta_2,$$

but

$$\Pr_{c_1 \leftarrow C_1^{(\mathsf{Real})}} \left[o = \mathsf{Acc} \wedge \left(\mathrm{E}\left[C_2^{(\mathsf{Real})} | C_1^{(\mathsf{Real})} = c_1 \right] \geq 2\gamma \right) \right] > 2\delta_2.$$

We now show that the above implies

$$\Delta(C_1^{(\mathsf{Real})}, C_2^{(\mathsf{Real})}), (C_1^{(\mathsf{Ideal})}, C_2^{(\mathsf{Ideal})}) > \epsilon.$$

Firstly, if $\Delta(C_1^{(\mathsf{Real})}, C_1^{(\mathsf{Ideal})}) > \epsilon$, then we are done. Thus, we can assume $\Delta(C_1^{(\mathsf{Real})}, C_1^{(\mathsf{Ideal})}) \leq \epsilon$. Let's denote event $\mathbf{A} = (o = \mathsf{Acc})$,

$$\mathbf{B} = \left(\mathrm{E}\left[C_2^{(\mathsf{Ideal})} | C_1^{(\mathsf{Ideal})} = c_1 \right] \geq \gamma \right),$$

$$\mathbf{C} = \left(\mathrm{E}\left[C_2^{(\mathsf{Real})} | C_1^{(\mathsf{Real})} = c_1 \right] \geq 2\gamma \right).$$

Note that they are events in the probability space of C_1. In this notation, we consider $(\mathbf{A} \wedge \mathbf{B})$, $(\mathbf{A} \wedge \mathbf{C})$, and $(\mathbf{A} \wedge \mathbf{C} \wedge \bar{\mathbf{B}})$. Suppose $C_1^{(\mathsf{Real})}$ and $C_1^{(\mathsf{Ideal})}$ have identical distribution, then clearly by union bound, $\Pr_{C_1^{(\mathsf{Real})}}[\mathbf{A} \wedge \mathbf{C} \wedge \bar{\mathbf{B}}] \leq 2\delta_2 - \delta_2$. When $\Delta(C_1^{(\mathsf{Real})}, C_1^{(\mathsf{Ideal})}) \geq \epsilon$, we just need another union bound to conclude

$$\Pr_{C_1^{(\mathsf{Real})}}[\mathbf{A} \wedge \mathbf{C} \wedge \bar{\mathbf{B}}] \geq 2\delta_2 - \delta_2 - \epsilon.$$

Namely,

$$\Pr_{c_1 \leftarrow C_1^{(\mathsf{Real})}} \left[o = \mathsf{Acc} \wedge \left(\mathrm{E}\left[C_2^{(\mathsf{Real})} | C_1^{(\mathsf{Real})} = c_1 \right] \geq 2\gamma \right) \wedge \left(\mathrm{E}\left[C_2^{(\mathsf{Ideal})} | C_1^{(\mathsf{Ideal})} = c_1 \right] < \gamma \right) \right] \geq 2\delta_2 - \delta_2 - \epsilon.$$

Because $\epsilon < \delta_2\gamma/(4(M+1)^k)$, we have the above RHS is lower bounded by $\delta_2/2$.

In particular, we have with probability at least δ_2 over $c_1 \leftarrow C_1^{(\mathsf{Real})}$,

$$\mathrm{E}\left[C_2^{(\mathsf{Real})} | C_1^{(\mathsf{Real})} = c_1 \right] - \mathrm{E}\left[C_2^{(\mathsf{Ideal})} | C_1^{(\mathsf{Ideal})} = c_1 \right] > \gamma.$$

By Lemma 10.3 and the fact that C_2 takes value in $[(1/2)^k, (1/2+M)^k]$, it implies with probability at least δ_2 over $c_1 \leftarrow C_1^{(\mathsf{Real})}$,

$$\Delta(C_2^{(\mathsf{Real})}|_{C_1^{(\mathsf{Real})}=c_1}, C_2^{(\mathsf{Ideal})}|_{C_1^{(\mathsf{Ideal})}=c_1}) \geq \gamma/(M+1)^k.$$

By Lemma 10.4, we have

$$\Delta((C_1^{(\mathsf{Real})}, C_2^{(\mathsf{Real})}), (C_1^{(\mathsf{Ideal})}, C_2^{(\mathsf{Ideal})}) > \delta_2/2 \cdot (\gamma/(M+1)^k)/2 > \epsilon.$$

Therefore, the distinguishing advantage is $> \epsilon$, which means Real and Ideal are not ϵ-close, a contradiction. ■

By following a similar argument in Theorem 9.11, we have

Theorem 10.2 *For any $0 < \epsilon_{\mathsf{dec}} < 1$, there exists an η-strong-robust $(n_{\mathsf{dec}}, t_{\mathsf{dec}}, \epsilon_{\mathsf{dec}})$ NS DI-RA with $n = \mathrm{poly}(1/\epsilon_{\mathsf{dec}})$, $t = \mathrm{poly}(1/\epsilon_{\mathsf{dec}})$, and $\eta \approx 1.748\%$ such that the soundness error ϵ_{dec} holds when the input source is ϵ'-local-uniform for $\epsilon' \leq \mathrm{poly}(\epsilon_{\mathsf{dec}})$.*

Proof. (Sketch) We claim that Π_{BHK} is the desired NS DI-RA protocol. Completeness and robustness of Π_{BHK} remains unchanged. It suffices to prove the soundness under the new condition, which almost follows from the argument in proving Theorem 9.11.

Here we highlight the missing step. By Lemma 10.1, in the Ideal distribution, we can set $\gamma = \eta/\kappa = O(\epsilon_{\mathsf{dec}}/C)$ and $\delta_2 = (e^{-\Omega(\delta^2 n/c^2)} + \kappa) = O(\epsilon_{\mathsf{dec}})$. We can conclude that if $\Gamma_{S,A,B|\perp,X,Y}$ is ϵ'-close to $U_S \otimes \Gamma_{\mathbf{A},\mathbf{B}|\mathbf{X},\mathbf{Y}}$ for

$$\epsilon' \leq O(\epsilon_{\mathsf{dec}}^2/C(M+1)^k) = \mathrm{poly}(\epsilon_{\mathsf{dec}}),$$

then

$$\Pr[o = \mathsf{Acc} \wedge \mu_r(\mathsf{hist}_{<r}, \mathsf{Acc}) \geq 2\gamma] \leq 2\delta_2,$$

which is sufficient to imply that the soundness error is at most $O(\epsilon_{\mathsf{dec}})$ by the same analysis as in the proof of Theorem 9.11. ■

Lemmas. We make use of the following lemmas in the proof of the above theorem and defer their proofs to Appendix B.

Lemma 10.3 *Let Y and Y' be two random variables over $[a,b]$ with expectation μ and μ', respectively. Then the statistical distance $\Delta(Y,Y')$ is at least $|\mu - \mu'|/(b-a)$.*

Lemma 10.4 *Let (X,Y) and (X',Y') be distributions. Suppose with probability at least α over $x \leftarrow X$, $\Delta(Y|_{X=x}, Y'|_{X'=x}) \geq \beta$, then*

$$\Delta((X,Y),(X',Y')) \geq \alpha\beta/2.$$

References

[1] J. Barrett, L. Hardy, and A. Kent. No signaling and quantum key distribution. *Phys. Rev. Lett.*, 95:010503, Jun 2005.

[2] F. G. Brandão, R. Ramanathan, K. H. Andrzej Grudka, M. Horodecki, and H. W. Paweł Horodecki, Tomasz Szarek. Realistic noise-tolerant randomness amplification using finite number of devices. *Nature Communications*, 7, 2016. Earlier versions: *QIP 2014* and arXiv:1310.4544.

[3] E. Chattopadhyay and D. Zuckerman. Explicit two-source extractors and resilient functions. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing*, pages 670–683. ACM, 2016.

[4] B. Chor and O. Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM Journal on Computing*, 17(2):230–261, 1988.

[5] K.-M. Chung, X. Wu, and Y. Shi. Physical randomness extractors. *QIP 2014*, arXiv:1402.4797, 2014.

[6] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23:880–884, October 1969.

[7] R. Colbeck. *Quantum And Relativistic Protocols For Secure Multi-Party Computation.* PhD thesis, University of Cambridge, 2006.

[8] R. Colbeck and R. Renner. Free randomness can be amplified. *Nature Physics*, 8:450–453, 2012.

[9] P. C. W. Davies and J. R. Brown. *The Ghost in the Atom: A Discussion of the Mysteries of Quantum Physics.* Cambridge University Press, 1986.

[10] A. De, C. Portmann, T. Vidick, and R. Renner. Trevisan's extractor in the presence of quantum side information. *SIAM Journal on Computing*, 067(258932), 2012.

[11] R. Gallego, L. Masanes, G. de la Torre, C. Dhara, L. Aolita, and A. Acín. Full randomness from arbitrarily deterministic events. *Nature Communications*, 4:2654, 2013.

[12] A. Grudka, K. Horodecki, M. Horodecki, P. Horodecki, M. Pawłowski, and R. Ramanathan. Free randomness amplification using bipartite chain correlations, 2013.

[13] A. Grudka, K. Horodecki, M. Horodecki, P. Horodecki, M. Pawłowski, and R. Ramanathan. Free randomness amplification using bipartite chain correlations. *Phys. Rev. A*, 90:032322, Sep 2014.

[14] V. Guruswami, C. Umans, and S. P. Vadhan. Unbalanced expanders and randomness extractors from Parvaresh–Vardy codes. *J. ACM*, 56(4), 2009.

[15] L. Masanes. Universally composable privacy amplification from causality constraints. *Phys. Rev. Lett.*, 102:140501, Apr 2009.

[16] L. Masanes, R. Renner, M. Christandl, A. Winter, and J. Barrett. Unconditional security of key distribution from causality constraints, 2006.

[17] P. Mironowicz and M. Pawłowski. Amplification of arbitrarily weak randomness, 2013.

[18] R. Ramanathan, F. G. Brandão, A. Grudka, K. Horodecki, M. Horodecki, and P. Horodecki. Robust device independent randomness amplification, 2013. arxiv:1308.4635.

[19] R. Ramanathan, F. G. Brandão, K. Horodecki, M. Horodecki, P. Horodecki, and H. Wojewódka. Randomness amplification against no-signaling adversaries using two devices. arXiv:1504.06313.

[20] H. Wojewodka, F. G. Brandão, A. Grudka, M. Horodecki, K. Horodecki, P. Horodecki, M. Pawlowski, and R. Ramanathan. Amplifying the randomness of weak sources correlated with devices. arXiv:1601.06455.

[21] D. Zuckerman. General weak random sources. In *Foundations of Computer Science, 1990. Proceedings., 31st Annual Symposium on*, pages 534–543 vol.2, 1990.

A Impossibility of NS-proof Strong Randomness Extractor

In this section, we show that NS-proof strong randomness extractor does not exist even for extracting one bit from C-NS sources with $n-1$ bits of min-entropy. We start with the definition.

Definition A.1 (NS-proof Strong Randomness Extractor) *A function* $\mathrm{Ext}:\{0,1\}^n\times\{0,1\}^d\to\{0,1\}$ *is a* NS-proof (k,ϵ)-strong randomness extractor*, if for any C-NS state* $\Gamma_{X,D}$ *with* n *bit source* X *and NS min-entropy* $\mathrm{H}^{\mathsf{ns}}_\infty(X|D)_\Gamma\geq k$*, and for a uniform seed* $Y\in\{0,1\}^d$ *independent of* $\Gamma_{X,D}$*, we have*

$$\Delta(\Gamma_{\mathrm{Ext}(X,Y),Y,D},U_1\otimes Y\otimes\Gamma_D),\leq\epsilon. \tag{A.1}$$

where U_1 *is an independent uniform bit.*

Theorem A.2 *For every* $n,d\in\mathbb{N}$ *and function* $\mathrm{Ext}:\{0,1\}^n\times\{0,1\}^d\to\{0,1\}$*, there exists a C-NS state* $\Gamma_{X,D}$ *with* n *bit source* X *and NS min-entropy* $\mathrm{H}^{\mathsf{ns}}_\infty(X|D)_\Gamma\geq n-1$ *such that*

$$\Delta(\Gamma_{\mathrm{Ext}(X,Y),Y,D},U_1\otimes Y\otimes\Gamma_D)\geq 1/2$$

Proof. We define $\Gamma_{X,D}$ as follows. First, we let the source X to have uniform (marginal) distribution over $\{0,1\}^n$. Then, we define the NS box D with input $E\in\{0,1\}^d$ and single bit output $F\in\{0,1\}$, which correlates with X, as follow. On input $E=e$, the box D outputs $\mathrm{Ext}(X,e)$.

We first verify that $\Gamma_{X,D}=\Gamma_{X,F|\perp,E}$ is indeed a C-NS state with $\mathrm{H}^{\mathsf{ns}}_\infty(X|D)_\Gamma\geq n-1$. Note that the NS condition only requires that X has a well-defined marginal distribution for any input $E=e\in\{0,1\}^d$ to the box, which holds since by construction, X has uniform marginal. Also, $\mathrm{H}^{\mathsf{ns}}_\infty(X|D)_\Gamma\geq n-1$ since D can only learn at most 1 bit information from X.

Now, note that given $Y\otimes\Gamma_D$, a predictor P can easily predict $\mathrm{Ext}(X,Y)$ by querying D with input $E=Y$, who will output $F=\mathrm{Ext}(X,Y)$. Therefore, P can predict $\mathrm{Ext}(X,Y)$ with probability 1, which implies the NS distance

$$\Delta(\Gamma_{\mathrm{Ext}(X,Y),Y,D},U_1\otimes Y\otimes\Gamma_D)\geq 1/2$$

■

B Proofs from the main text

B.1 Proof of Proposition 9.4

Proposition B.1 (Restated) *Let* $\Gamma_{ABF|XYE}$ *be the global NS-states shared among Alice, Bob and the environment, where Alice and Bob each has* k *NS boxes and play the* BHK^k_M *game. Let* h :

$\{0,1\}^k \to \{0,1\}$ *a C-concentrated function. Let* $\mathbf{a} \in \{0,1\}^k$ *be Alice's output given any input* $\mathbf{x}$ *and* $z = h(\mathbf{a})$ *be the final output, and hence* $\Gamma_{ZF|XE}$ *the NS-state after Alice outputs* $z = h(\mathbf{a})$*. Then we have*

$$\Delta(\Gamma_{ZF|XE}, U_1 \otimes \Gamma_{F|E}) \leq 2C \left\langle \mathrm{BHK}_M^k, \Gamma_{AB|XY} \right\rangle. \tag{B.1}$$

Proof. By definition, the above LHS is equivalent to the statistical distance between classical distributions generated by the best distinguisher. The most general distinguisher's strategy for this case is to first look at Z and then decide on the input $E = E(Z)$ based on Z, and then to distinguish between the resultant distributions (Z, F).

Let us fix any such distinguisher D and let $p(Z, F)$ denote the distribution obtained from $\Gamma_{ZF|XE}$ and $q(Z, F)$ the one obtained from $U_1 \times \Gamma_{F|E}$.

It is easy to see that $q(z, f)$ coming from $U_1 \times \Gamma_{F|E}$ satisfies,

$$q(z, f) = \frac{1}{2}\Gamma_{f|e(z)},$$

for any $z \in \{0,1\}$ and any choice of e as a function of z. It is a little tricker to connect $p(z, f)$ to $\Gamma_{ABF|XYE}$, and then $\Gamma_{ZF|XE}$. For any fixed input $\mathbf{x}, z, e, f$, we have

$$\begin{aligned} \Gamma_{z,f|\mathbf{x},e} &= \Gamma_{z|\mathbf{x},e,f} \times \Gamma_{f|\mathbf{x},e} \\ &= (\beta_{\mathcal{A}_z} \cdot \Gamma_{\mathbf{AB}|\mathbf{XY},e,f})\Gamma_{f|e}, \end{aligned}$$

where the second equality is due to Equ.(9.15) and the NS-condition $\Gamma_{f|\mathbf{x}e} = \Gamma_{f|e}$. By definition, $p(z, f)$ is the probability when Alice and Bob obtain z (given $\mathbf{x}$) and the environment's input is $e(z)$, a function of z, namely,

$$p(z, f) = \Gamma_{z,f|\mathbf{x},e(z)} = (\beta_{\mathcal{A}_z} \cdot \Gamma_{\mathbf{AB}|\mathbf{XY},e(z),f})\Gamma_{f|e(z)}.$$

Thus, we can bound the statistical distance between the two distributions as follows.

$$\begin{aligned} \sum_{z,f} |p(z, f) - q(z, f)| &= \sum_{z,f} \Gamma_{f|e(z)} \left| \beta_{\mathcal{A}_z} \cdot \Gamma_{\mathbf{AB}|\mathbf{XY},e(z),f} - \frac{1}{2} \right| \\ &\leq \sum_{z,f} \Gamma_{f|e(z)} \left| \beta_{\mathcal{A}_z} - \frac{1}{2}\beta_{\text{uniform}} \right| \left| \Gamma_{\mathbf{AB}|\mathbf{XY},e(z),f} \right\rangle \\ &\leq C \left\langle \mathrm{BHK}_M^k, \sum_{z,f} \Gamma_{f|e(z)} \times \Gamma_{\mathbf{AB}|\mathbf{XY},e(z),f} \right\rangle \\ &= 2C \left\langle \mathrm{BHK}_M^k, \Gamma_{\mathbf{AB}|\mathbf{XY}} \right\rangle, \end{aligned}$$

where the second inequality is from Equ.(9.13) and the Cauchy-Schwarz inequality. The third inequality is because h is C-concentrated. The last equality is due to the fact that $\sum_f \Gamma_{f|e} \times \Gamma_{\mathbf{AB}|\mathbf{XY},e,f} = \sum_f \Gamma_{\mathbf{AB}f|\mathbf{XY}e} = \Gamma_{\mathbf{AB}|\mathbf{XY}}$ for any e by the NS-condition. ■

B.2 Proof of Lemma 9.5

Here we build a concentration bound about the summation of any t-wise independent random variables in Lemma B.2.

Lemma B.2 *Let $t \geq 4$ be even. Let $X_1, \ldots, X_n$ be t-wise independent random variables over $[-1, 1]$ with $\mathrm{E}[X_i] = 0$ and $\mathrm{Var}[X_i] \leq \sigma$ for every $i \in [n]$. Let $X = \sum_i X_i$. If $\sigma n \geq 2$ then for every $C > 0$, we have*

$$\Pr[|X| \geq C] \leq \left(\frac{\sigma n t^2}{4C^2}\right)^{t/2}.$$

Proof. By a standard moment method, we have

$$\Pr[|X| \geq C] = \Pr[X^t \geq C^t] \leq \frac{\mathrm{E}[X^t]}{C^t}.$$

By definition and linearity of expectation,

$$\mathrm{E}[X^t] = \mathrm{E}\left[\left(\sum_{i=1}^{n} X_i\right)^t\right] = \sum_{i_1,\ldots,i_t} \mathrm{E}[X_{i_1} \cdots X_{i_t}]. \tag{B.2}$$

For each term, if $(i_1, \ldots, i_t)$ consists of d distinct indices, then we can write

$$\mathrm{E}[X_{i_1} \cdots X_{i_t}] = \mathrm{E}[X_{j_1}^{a_1} \cdots X_{j_d}^{a_d}] = \mathrm{E}[X_{j_1}^{a_1}] \cdots \mathrm{E}[X_{j_d}^{a_d}]$$

for some indices $j_1, \ldots, j_d \in [n]$ and exponents $a_1, \ldots, a_d \geq 1$, where the last equality uses t-wise independence. Note that $X_i \in [-1, 1]$, $\mathrm{E}[X_i] = 0$ and $\mathrm{Var}[X_i] \leq \sigma$ implies that $|\mathrm{E}[X_i^a]| \leq \sigma$ for every $a \geq 2$. Thus, we have $\mathrm{E}[X_{j_1}^{a_1}] \cdots \mathrm{E}[X_{j_d}^{a_d}] \leq \sigma^d$. Also note that if $(i_1, \ldots, i_t)$ consists of more than $t/2$ distinct indices, then there must exist some index i^* that appears only once, and since $\mathrm{E}[X_{i^*}] = 0$, the term equals to 0. Finally, note that there are at most $\binom{n}{d} \cdot d^t$ terms that consists of *exactly* d distinct indices. We can upper bound the sum in Eq. (B.2) by

$$\sum_{d=1}^{t/2} \sigma^d \cdot \binom{n}{d} \cdot d^t \leq \sum_{d=1}^{t/2} \frac{\sigma^d n^d}{2^d} \left(\frac{t}{2}\right)^t \leq \left(\frac{\sigma n t^2}{4}\right)^{t/2}.$$

Therefore,

$$\Pr[|X| \geq C] \leq \frac{\mathrm{E}[X^t]}{C^t} \leq \left(\frac{\sigma n t^2}{4C^2}\right)^{t/2}.$$

■

Lemma B.3 (Restated) *Any t-wise random hash function $h : \{0,1\}^k \to \{0,1\}$ is $2^{\frac{1}{2}(k+O(\log(k)))}$-concentrated with probability $1 - 2^{-\Omega(k)}$ when $t = \Theta(k)$.*

Proof. We shall first prove the *concentration* property holds for any fixed $z \in \{0,1\}$ and any entry of $\beta_{\mathcal{A}_z}$ and then invoke union bounds to show the very property holds for all z and all entries.

Let a tuple $\tau = (\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y})$ be an index of entries of $\beta_{\mathcal{A}_z}$ (and $\left|\mathrm{BHK}_M^k\right\rangle$). Denote the corresponding entry by $\beta_{\mathcal{A}_z}^{\tau}$ (and by $\left|\mathrm{BHK}_M^k\right\rangle^{\tau}$ respectively). It suffices to only consider indices corresponding to

valid inputs $(\mathbf{x}, \mathbf{y})$. For invalid inputs $(\mathbf{x}, \mathbf{y})$, we have $\beta^\tau_{\mathcal{A}_z} = \beta^\tau_{\text{uniform}} = 0$ and thus the property holds trivially.

Let us characterize any t-wise random hash function h by a collection of t-wise independent random variables $\{V_\mathbf{a}\}_{\mathbf{a}\in\{0,1\}^k}$. For any fixed $z \in \{0,1\}$, let $V_\mathbf{a}$ be the indicator whether $h(\mathbf{a}) = z$. Moreover, $V_\mathbf{a} = 0$ or 1 with probability 1/2. Thus the entry $\beta^\tau_{\mathcal{A}_z}$ can be represented by

$$\beta^\tau_{\mathcal{A}_z} = \sum_{\mathbf{a}\in\mathcal{A}_k} \beta^\tau_\mathbf{a} = \sum_{\mathbf{a}\in\{0,1\}^k} V_\mathbf{a}\beta^\tau_\mathbf{a}. \tag{B.3}$$

By linearity of expectation, we have

$$\mathrm{E}[\beta^\tau_{\mathcal{A}_z}] = \sum_{\mathbf{a}\in\{0,1\}^k} \mathrm{E}[V_\mathbf{a}]\beta^\tau_\mathbf{a} = \frac{1}{2}\sum_{\mathbf{a}\in\{0,1\}^k} \beta^\tau_\mathbf{a} = \frac{1}{2}\beta^\tau_{\text{uniform}}, \tag{B.4}$$

where we makes use of $\mathrm{E}[V_a] = 1/2$ and Equ. (9.14). In the following we prove that $\beta^\tau_{\mathcal{A}_z}$ is concentrated around its expectation with the help of Lemma B.2. Let $Y_\mathbf{a} = V_\mathbf{a}\beta^\tau_\mathbf{a}$ and $X_\mathbf{a} = (Y_\mathbf{a} - \mathrm{E}[Y_\mathbf{a}])/\left|\mathrm{BHK}^k_M\right\rangle^\tau$. Note that $|\beta_\mathbf{a}| \preceq \left|\mathrm{BHK}^k_M\right\rangle$ and thus $|Y_\mathbf{a}| \le \beta^\tau_\mathbf{a} \le \left|\mathrm{BHK}^k_M\right\rangle^\tau$. Thus we have that $\mathrm{E}[X_a] = 0$ and $X_a \in [-1, 1]$. Moreover,

$$\mathrm{Var}[X_\mathbf{a}] = \frac{1}{(\left|\mathrm{BHK}^k_M\right\rangle^\tau)^2}\mathrm{Var}[Y_\mathbf{a}] \le \frac{1}{(\left|\mathrm{BHK}^k_M\right\rangle^\tau)^2}\mathrm{E}[Y^2_\mathbf{a}] = \frac{(\beta^\tau_\mathbf{a})^2}{(\left|\mathrm{BHK}^k_M\right\rangle^\tau)^2}\mathrm{E}[V^2_\mathbf{a}] \le \frac{1}{2}. \tag{B.5}$$

Now we apply Lemma B.2 to the collection $\{X_\mathbf{a} : \mathbf{a} \in \{0,1\}^k\}$ with $\sigma = 1/2, n = 2^k, C = 2^{\frac{1}{2}(k+O(\log(k)))}$, and $t = \Theta(k)$. Then we have

$$\mathbf{Pr}\left[\left|\sum_{\mathbf{a}\in\{0,1\}^N} X_\mathbf{a}\right| \ge C\right] \le \left(\frac{2^{k-1}t^2}{4C^2}\right)^{t/2} \le 2^{-\Omega(k)}. \tag{B.6}$$

Note that the event $|\sum_{\mathbf{a}\in\{0,1\}^k} X_\mathbf{a}| \ge C$ is equivalent to the event $|\beta^\tau_{\mathcal{A}_z} - \mathrm{E}[\beta^\tau_{\mathcal{A}_z}]| \ge C\left|\mathrm{BHK}^k_M\right\rangle^\tau$. Thus for any $z \in \{0,1\}$ and any index τ, we have

$$\mathbf{Pr}\left[\left|\beta^\tau_{\mathcal{A}_z} - \frac{1}{2}\beta^\tau_{\text{uniform}}\right| \le C\left|\mathrm{BHK}^k_M\right\rangle^\tau\right] \ge 1 - 2^{-\Omega(k)}. \tag{B.7}$$

By union bounds over all z and indices τ and note that M is constant, we have

$$\mathbf{Pr}_h\left[\bigwedge_{z\in\{0,1\}}\left\{|\beta_{\mathcal{A}_z} - \frac{1}{2}\beta_{\text{uniform}}| \preceq 2^{\frac{1}{2}(k+O(\log(k)))}\left|\mathrm{BHK}^k_M\right\rangle\right\}\right] \ge 1 - 2^{-\Omega(k)+\log(8M)k} = 1 - 2^{-\Omega(k)}, \tag{B.8}$$

which completes the proof. ∎

B.3 Proof of Lemma 10.3

Lemma B.4 (Restated) *Let Y and Y' be two random variables over $[a, b]$ with expectation μ and μ', respectively. Then the statistical distance $\Delta(Y, Y')$ is at least $|\mu - \mu'|/(b - a)$.*

Proof. By adding a constant, we can assume w.l.o.g. that Y and Y' takes value in $[0, b-a]$. Also, w.l.o.g., we can assume $\mu \geq \mu'$. We have

$$\begin{aligned}(\mu - \mu') &= \sum_y y \cdot (p(y) - p'(y)) \\ &\leq \sum_{y:p(y)>p'(y)} y \cdot (p(y) - p'(y)) \\ &\leq (b-a) \sum_{y:p(y)>p'(y)} y \cdot (p(y) - p'(y)) \\ &= (b-a) \cdot \Delta(Y, Y')\end{aligned}$$

Thus, $\Delta(Y, Y') \geq |\mu - \mu'|/(b-a)$. ■

B.4 Proof of Lemma 10.4

Lemma B.5 (Restated) *Let (X, Y) and (X', Y') be distributions. Suppose with probability at least α over $x \leftarrow X$, $\Delta(Y|_{X=x}, Y'|_{X'=x}) \geq \beta$, then*

$$\Delta((X, Y), (X', Y')) \geq \alpha\beta/2.$$

Proof. Let $p(x, y)$ and $p'(x, y)$ denote the probability mass function of (X, Y) and (X', Y'), respectively. We first show that for every x,

$$\sum_y |p(x, y) - p'(x, y)| \geq p(x) \cdot \Delta(Y|_{X=x}, Y'|_{X'=x}).$$

Recall that

$$\Delta(Y|_{X=x}, Y'|_{X'=x}) = \sum_{y:p(y|x)\geq p'(y|x)} p(y|x) - p'(y|x) = \sum_{y:p'(y|x)\geq p(y|x)} p'(y|x) - p(y|x)$$

Suppose $p(x) \geq p'(x)$,

$$\begin{aligned}\sum_y |p(x, y) - p'(x, y)| &= \sum_y |p(x)p(y|x) - p'(x)p'(y|x)| \\ &\geq \sum_{y:p(y|x)\geq p'(y|x)} p(x)p(y|x) - p'(x)p'(y|x) \\ &\geq p(x) \sum_{y:p(y|x)\geq p'(y|x)} p(y|x) - p'(y|x) \\ &= p(x)\Delta(Y|_{X=x}, Y'|_{X'=x})\end{aligned}$$

Similarly, suppose $p(x) \leq p'(x)$,

$$\begin{aligned}\sum_y |p(x, y) - p'(x, y)| &= \sum_y |p(x)p(y|x) - p'(x)p'(y|x)| \\ &\geq \sum_{y:p'(y|x)\geq p(y|x)} p'(x)p'(y|x) - p(x)p(y|x) \\ &\geq p(x) \sum_{y:p'(y|x)\geq p(y|x)} p'(y|x) - p(y|x) \\ &= p(x)\Delta(Y|_{X=x}, Y'|_{X'=x})\end{aligned}$$

■