# De Finetti Theorems for Quantum Conditional Probability Distributions with Symmetry

**Journal Article**

**Author(s):**
Jandura, Sven; Tan, Ying Zhe Ernest (iD)

# De Finetti Theorems for Quantum Conditional Probability Distributions with Symmetry

Sven Jandura and Ernest Y.-Z. Tan

**Abstract.** The aim of device-independent quantum key distribution (DIQKD) is to study protocols that allow the generation of a secret shared key between two parties under minimal assumptions on the devices that produce the key. These devices are merely modeled as black boxes and mathematically described as conditional probability distributions. A major obstacle in the analysis of DIQKD protocols is the huge space of possible black box behaviors. De Finetti theorems can help to overcome this problem by reducing the analysis to black boxes that have an iid structure. Here we show two new de Finetti theorems that relate conditional probability distributions in the quantum set to de Finetti distributions (convex combinations of iid distributions) that are themselves in the quantum set. We also show how one of these de Finetti theorems can be used to enforce some restrictions onto the attacker of a DIQKD protocol. Finally we observe that some desirable strengthenings of this restriction, for instance to collective attacks only, are not straightforwardly possible.

## 1. Introduction

The aim of quantum key distribution is to establish a shared key between two parties, commonly called Alice and Bob, that is unknown to any third party, commonly called Eve. To achieve this goal, Alice and Bob can share an entangled quantum state and use the correlated outcomes of measurements on this state to generate a secure key pair via a postprocessing protocol. If Eve has tampered with the shared state, Alice and Bob either notice this and abort the protocol, or are able to generate a secure key pair anyway [1–3]. In device-independent quantum key distribution (DIQKD) we assume that Eve not only has control over the shared state, but is also able to manipulate the

devices that Alice and Bob use to measure the state. As long as the devices are not manipulated in a way that sends information out of Alice's and Bob's laboratories through channels other than those controlled by Alice and Bob, there are still protocols that allow for the generation of a shared secret key [4–8].

In the device-independent context, the devices of Alice and Bob are treated as black boxes and modeled by a conditional probability distribution $P_{AB|XY}$. Alice and Bob can give inputs $x$ and $y$, respectively, to the box, and receive outputs $a$ and $b$ with probability $P_{AB|XY}(ab|xy)$. We will often write $P(ab|xy)$ instead of $P_{AB|XY}(ab|xy)$ when the random variables $A$, $B$, $X$ and $Y$ are implicitly understood. In the device-dependent case, the inputs correspond to the choice of measurement basis and the outputs to the results of the measurement. We will denote the sets of possible inputs by $\mathcal{X}$ and $\mathcal{Y}$, and the sets of outputs by $\mathcal{A}$ and $\mathcal{B}$. If the inputs and outputs are strings of length $n$, i.e., they are of the form $\mathcal{X} = \hat{\mathcal{X}}^n$ (where $\hat{\mathcal{X}}$ denotes some set of possible single-round inputs) and analogously for $\mathcal{Y}$, $\mathcal{A}$ and $\mathcal{B}$, we call $P_{AB|XY}$ an $n$-round box. If $Q_{\hat{A}\hat{B}|\hat{X}\hat{Y}}$ is a box with inputs and outputs in $\hat{\mathcal{X}}, \hat{\mathcal{Y}}, \hat{\mathcal{A}}, \hat{\mathcal{B}}$, we denote by $Q_{AB|XY}^{\otimes n}$ the $n$-round iid box with

$$Q^{\otimes n}(ab|xy) = \prod_{i=1}^{n} Q(a_i b_i | x_i y_i). \tag{1}$$

Not all boxes $P_{AB|XY}$ describe processes that are physically possible if we assume that no information can leave the laboratories of Alice and Bob. All boxes must then be such that Bob gains no information about Alice's input from his output, and vice versa. We refer to boxes satisfying this constraint as non-signaling:

**Definition 1.** A box $P_{AB|XY}$ is non-signaling if

$$\forall b, y, x, x' \qquad \sum_a P(ab|xy) = \sum_a P(ab|x'y) \tag{2}$$

$$\forall a, x, y, y' \qquad \sum_b P(ab|xy) = \sum_b P(ab|xy'). \tag{3}$$

If we furthermore assume that the boxes are described by quantum theory, we can describe the distribution $P_{AB|XY}$ by some quantum state shared between Alice and Bob and some POVMs describing their measurements.

**Definition 2.** A box $P_{AB|XY}$ is quantum if there are Hilbert spaces $\mathcal{H}_A$, $\mathcal{H}_B$, a state $\rho_{AB} \in \mathrm{End}\,(\mathcal{H}_A \otimes \mathcal{H}_B)$, for each $x$ a POVM $\{E^{a,x}|a\}$ on $\mathcal{H}_A$ and for each $y$ a POVM $\{F^{b,y}|b\}$ on $H_B$ such that

$$P(ab|xy) = \mathrm{tr}\left[\rho_{AB}\left(E^{a,x} \otimes F^{b,y}\right)\right]. \tag{4}$$

The set of quantum boxes is a proper subset of the set of non-signaling boxes. That both sets are not identical is demonstrated by the Popescu–Rohrlich box [9].

When constructing DIQKD security proofs, often the analysis would be simplified if there were some form of reduction from general box behavior to the iid case, as it is substantially easier to construct security proofs for the latter (as achieved in, e.g., [5,6,10]). To find such a reduction, so-called de Finetti theorems may be a promising tool, as they have previously been used to achieve this goal in the case of device-dependent QKD [3,11]. De Finetti theorems allow us to relate the entries of an arbitrary permutation-invariant box to the entries of a de Finetti box (a convex combination of iid boxes). De Finetti theorems where originally developed for random variables [12] and then extended to quantum states [3,13–15] and boxes [16,17]. For example, in [16] it was shown that for each set of single-round inputs $\hat{\mathcal{A}}$ and outputs $\hat{\mathcal{X}}$ there exists a de Finetti box $\tau_{A|X}$ such that for all permutation-invariant boxes $P_{A|X}$ it holds that

$$\forall a \in \mathcal{A}, \, x \in \mathcal{X} \qquad P(a|x) \leq (n+1)^{|\hat{\mathcal{X}}|(|\hat{\mathcal{A}}|-1)}\tau(a|x). \tag{5}$$

(Here we treat the inputs and outputs of Alice and Bob as lumped together to a single-round input and output.)

However, the de Finetti theorems for boxes derived in, e.g., [16,17] have the drawback that the de Finetti boxes cannot be restricted to the quantum set even if the original permutation-invariant boxes are quantum. This creates an obstacle for applications, because many existing DIQKD security proofs under the iid assumption exploit the properties of the quantum set [5,6,10]. This implies that such proofs cannot be combined with the de Finetti theorems in [16,17] to obtain security against non-iid attacks, as those de Finetti theorems involve boxes that are not in the quantum set. (It is true that one could aim to derive a security proof for all iid behaviors in the non-signaling rather than quantum set, then apply the de Finetti theorems of [16,17] to obtain security against non-iid attacks. However, this would give lower asymptotic key rates and noise tolerance compared to security proofs against quantum attackers, because non-signaling behaviors yield a significantly larger class of possible attacks.) Ideally, we would like to find a de Finetti theorem that can extend the iid security proofs against quantum attackers in [5,6,10] to cover non-iid quantum attackers, while preserving the asymptotic key rates and noise tolerance from those proofs, similar to the situation for device-dependent QKD [11]. While we do not fully achieve this goal in this work, we do obtain a de Finetti theorem that allows a partial reduction to the iid case (in a sense described in Sect. 3), and we also highlight some concrete difficulties that may be faced when aiming for a full reduction.

Regarding other existing approaches for reductions to the iid case, we note that for DIQKD protocols that use only one-way communication for error correction [3], a proof technique known as the entropy accumulation theorem (EAT) [18] can be used to essentially reduce the analysis of non-iid (but time-ordered) boxes to the iid scenario [8]. Alternatively, the techniques in [19,20] can be used to obtain security proofs for such protocols even when the boxes accept all inputs in parallel, though the resulting asymptotic key rates are lower than in the iid case. There are, however, protocols that don't only

use one-way error correction (broadly referred to as advantage distillation protocols [3,10,21–24], such as the Cascade protocol [25] or the repetition-code protocol [3,21]), and these protocols do not admit a security proof via those approaches.[1] The significance of these protocols in DIQKD is that under an iid assumption, it has been shown [10] that they can achieve higher noise tolerances than one-way error correction (i.e., they can achieve positive key rates even when the key rate given by one-way error correction is zero), analogous to results for device-dependent QKD [3,23,24]. However, for device-dependent QKD these improved noise tolerances can be lifted to the non-iid case using de Finetti arguments as mentioned above, whereas in DIQKD such an argument is currently missing—in fact, there are currently no security proofs for DIQKD advantage distillation protocols against non-iid attacks. Finding a way to resolve this would be useful in, for instance, tackling a foundational question of characterizing which non-local box behaviors can be used for DIQKD [26] (analogous to the question of *bound information* in device-dependent QKD [27]), since advantage distillation can have higher noise tolerances than one-way error correction.

Our main result in this work consists of two de Finetti theorems for Clauser–Horne–Shimony–Holt (CHSH) symmetric quantum boxes (see Definition 4), such that the de Finetti box is quantum as well. We further show how the first de Finetti theorem could be used in the security proofs of DIQKD protocols, yielding a partial reduction to the iid case.

The rest of this paper is structured as follows: In Sect. 2.1, we show the first de Finetti theorem (Theorem 6). It is similar to Eq. (5) and shows that the entries of a CHSH symmetric quantum box are upper bounded, up to a factor polynomial in $n$, by the entries of a fixed quantum de Finetti box. In Sect. 2.2, we then show the second de Finetti theorem (Theorem 11), which is closer to the original de Finetti theorems for random variables and quantum states. It states that the marginal of the first $k$ rounds of a $n$-round CHSH symmetric quantum box is close to (and not just upper bounded by) a quantum de Finetti box. Our results in this section rely on the existence of appropriate threshold theorems (see, e.g., Theorem 7). A natural question is whether it is possible to derive them without using the threshold theorems; however, we show in Appendix B that proving a de Finetti theorem of the first form is essentially equivalent to proving a threshold theorem; hence, it would be a result of comparable difficulty.

---

[1]More specifically: the proof approaches in those works essentially rely on bounding the conditional smooth min-entropy [3] of the "raw" box outputs (or a subset thereof), then compensating for the additional information revealed during one-way error correction by subtracting the number of bits communicated during that step. However, advantage distillation protocols may perform a significant amount of processing on the box outputs *before* the privacy amplification step (informally: the step in which the data is transformed into the final key), and furthermore, they often communicate a very large number of bits in the process as compared to one-way error correction (see, e.g., the repetition-code protocol [3,21]). Hence, the proof techniques in [8,19,20] are difficult to extend to these advantage distillation protocols.

In light of this, our results cannot currently be used as an alternative method to prove threshold theorems. However, our focus is more on the application of these results for DIQKD security proofs. Hence, in Sect. 3, we present an application of the first de Finetti theorem to bound the diamond distance between two channels acting on boxes. The diamond distance measures how well these two channels can be distinguished by an attacker. Since the security of a DIQKD protocol is related to the diamond distance between the protocol and an ideal channel [28,29], bounds on the diamond distance can be useful in DIQKD security proofs. We show that to prove security of a DIQKD protocol against arbitrary (so-called coherent [30]) quantum attacks it is sufficient to prove security against an adversary who holds an extension of a fixed quantum de Finetti box (Theorem 15). However, this extension may not be quantum itself and can only be restricted to the non-signaling set.

In Sect. 4, we show that the result from Sect. 3 cannot be strengthened to restrict the attacker further to collective attacks [30] (attacks where the black box can be described by an iid quantum state and iid measurements for Alice and Bob, see Definition 16). For this, we construct two channels that cannot be distinguished at all using boxes compatible with collective attacks, but can be distinguished if arbitrary quantum boxes are available (Theorem 17). This shows that the theorem from Sect. 3 cannot be immediately used to conclude security against coherent attacks from security against collective attacks.

## 2. De Finetti Theorems for Boxes with CHSH Symmetry

### 2.1. The First de Finetti Theorem

Let us first define de Finetti boxes and CHSH symmetry:

**Definition 3.** A $n$-round box $\tau_{AB|XY}$ is called de Finetti if it is the convex combination of iid boxes.

**Definition 4.** An $n$-round box $P_{AB|XY}$ with single-round inputs and outputs in $\hat{\mathcal{A}} = \hat{\mathcal{B}} = \hat{\mathcal{X}} = \hat{\mathcal{Y}} = \{0,1\}$ is called CHSH symmetric if

$$P(ab|xy) = P(a'b'|x'y') \qquad \text{whenever } \|a \oplus b \oplus xy\|_0 = \|a' \oplus b' \oplus x'y'\|_0 . \tag{6}$$

Here $\|x\|_0$ denotes the number of nonzero entries of a $n$-bit string $x$.

If for an index $i$ we have $a_i \oplus b_i = x_i y_i$, we say that the CHSH game is won in round $i$ [31]. Thus, Definition 4 basically states that a box is CHSH symmetric if its entries $P(ab|xy)$ only depends on how many indices the CHSH game was won. Our definition of CHSH symmetry differs slightly from the one in [16], where it is only required that $P(ab|xy) = P(a'b'|x'y')$ whenever $a \oplus b \oplus xy = a' \oplus b' \oplus x'y'$. Our definition agrees with the definition in [16] for permutation symmetric boxes—essentially, we have implicitly incorporated the constraint of permutation symmetry into Definition 4 itself.

Of course, an attacker can initially manipulate the boxes of Alice and Bob such that they do not possess CHSH symmetry. However, Alice and Bob can

run the following procedure to enforce CHSH symmetry: First Alice chooses a random permutation $\pi$ and transmits it to Bob over the authenticated channel; then, Alice and Bob permute their inputs and outputs according to $\pi$. If we view the original box as a conditional probability distribution $P_{ABE|XYZ}$ for Alice, Bob and Eve we can view $\pi$ as additional knowledge $E'$ of Eve and describe the box after $\pi$ has been applied by $\tilde{P}_{ABEE'|XYZ}$. Alice and Bob will not require $\pi$ for the remainder of the protocol and can now discard it; therefore, in the rest of our discussion we do not include it in the marginal of the Alice–Bob boxes, whereas on Eve's component we will simply absorb $E'$ into $E$ and no longer explicitly denote it. Then the marginal $\tilde{P}_{AB|XY}$ has permutation symmetry. To go from permutation symmetry to CHSH symmetry Alice and Bob apply the depolarization protocol described in appendix A of [32] to each round. If the box in the honest implementation of the DIQKD protocol has CHSH symmetry, it is unchanged by this depolarization protocol. It is important to note that only the marginal box of Alice and Bob has CHSH symmetry after this protocol: From the perspective of Eve, who knows the permutation $\pi$ and the random bits chosen in the depolarization protocol in [32], the box may not have CHSH symmetry. However, we highlight that in the case of device-dependent QKD, this did not prevent constructing a security proof via de Finetti arguments [11], and hence, there still remains the possibility of a similar result for DIQKD.

It was shown in [16] that a de Finetti theorem holds for CHSH symmetric boxes:

**Theorem 5** (Corollary 6 in [16]). *For each number $n$ of rounds, there is an n-round CHSH symmetric de Finetti box $\tau_{AB|XY}$ such that for all CHSH symmetric boxes $P_{AB|XY}$ it holds that*

$$P(ab|xy) \leq (n+1)\tau(ab|xy). \tag{7}$$

Theorem 5 was derived for all CHSH symmetric boxes $P_{AB|XY}$, even if they are not quantum. However, the de Finetti box $\tau_{AB|XY}$ constructed in the theorem is also not quantum. The main result we derive in this section is a de Finetti theorem for *quantum* CHSH symmetric boxes, hence resolving this issue:

**Theorem 6.** *For each number $n$ of rounds, there is an n-round CHSH symmetric quantum de Finetti box $\tau_{AB|XY}$ such that for all CHSH symmetric quantum boxes $P_{AB|XY}$ it holds that*

$$P(ab|xy) \leq (n+1)^2\tau(ab|xy). \tag{8}$$

The maximal probability with which any single-round quantum box can win the CHSH game is $w = \frac{2+\sqrt{2}}{4}$ [33]. This value is called the quantum value of the CHSH game. To prove Theorem 6, we need the following specialization of a theorem from [34] to the CHSH case. It says that the probability that the fraction of won CHSH games is larger than a certain threshold (namely

the value of the CHSH game) is exponentially small in $n$. Such theorems are commonly referred to as *threshold theorems*.[2]

**Theorem 7** (Theorem 5 in [34]). *Let $P_{AB|XY}$ be an $n$-round box with single-round inputs and outputs in $\{0,1\}$. Let $\mu$ be the uniform probability distribution on $\{0,1\}^2$ and $K = ||A \oplus B \oplus XY \oplus \mathbf{1}||_0$ the number of won instances of the CHSH game. Let $w = \frac{2+\sqrt{2}}{4}$ the quantum value of the CHSH game. Then for $k > wn$*

$$\Pr_{P_{AB|XY},\mu^{\otimes n}}[K \geq k] \leq e^{-nD(k/n, 1-k/n||w, 1-w)}, \tag{9}$$

*where $\Pr_{P_{AB|XY},\mu^{\otimes n}}$ denotes the probability measure in which $X$ and $Y$ are sampled from $\mu^{\otimes n}$ and $A$ and $B$ are sampled using $P_{AB|XY}$ and*

$$D(p, 1-p||q, 1-q) = p\ln\left(\frac{p}{q}\right) + (1-p)\ln\left(\frac{1-p}{1-q}\right) \tag{10}$$

*denotes the relative entropy.*

Note that the box $P_{AB|XY}$ in Theorem 7 does not have to be CHSH symmetric. However, if $P_{AB|XY}$ is CHSH symmetric then we can describe it completely by $n+1$ parameters $\{p_0, p_1, \ldots, p_n\}$, which we define as follows: For each $k \in \{0, \ldots, n\}$, take any $a, b, x, y \in \{0,1\}^n$ such that $k = ||a \oplus b \oplus xy \oplus \mathbf{1}||_0$ (in other words, $a, b, x, y$ win exactly $k$ instances of the CHSH game). Then define

$$p_k = \binom{n}{k} 2^n P(ab|xy). \tag{11}$$

By CHSH symmetry, all combinations of $a, b, x, y$ with the same value of $k$ have the same value of $P(ab|xy)$, so the expression (11) is indeed well defined. The normalization factor $\binom{n}{k}2^n$ is chosen to give these parameters a simple interpretation: namely, $p_k$ is in fact equal to the probability of winning exactly $k$ CHSH games for the box $P(ab|xy)$ (regardless of the input distribution). To see this, notice that for fixed $x$, $y$ and $k$ there are exactly $\binom{n}{k}2^n$ pairs $a, b$ such that $k = ||a \oplus b \oplus xy \oplus 1||_0$ [16]. Therefore, for any probability measure $\mu$ on

---

[2]To be precise, Theorem 7 is a "perfect" threshold theorem, in that the exponent in the bound (9) is such that the bound is, up to a factor polynomial in $n$, equal to $\binom{n}{k}w^k(1-w)^{n-k}$, the probability to win exactly $k$ games if a single game is won with probability $w$. "Imperfect" threshold theorems can be roughly described as giving bounds of the more general form $e^{-n\Delta(k/n)}$, where $\Delta$ is some potentially "looser" way to quantify the distance from $k/n$ to $w$ [35]. Our first de Finetti theorem (Theorem 6) and its generalization (Theorem 22) both require a perfect threshold theorem. However, the proof of our second de Finetti theorem (Theorem 11) still holds with an imperfect threshold theorem, though the resulting bound would be weaker.

the $n$-round inputs $x$ and $y$, we indeed have

$$
\Pr_{P_{AB|XY},\mu}[K = k] = \sum_{\substack{a,b,x,y \\ k=||a\oplus b\oplus xy\oplus\mathbf{1}||_0}} P(ab|xy)\mu(xy)
$$

$$
= \sum_{xy} \binom{n}{k} 2^n \frac{p_k}{\binom{n}{k}2^n}\mu(xy) \tag{12}
$$

$$
= p_k
$$

where in the second equality we used that the summand does not depend on $a$ and $b$, and for a fixed $x$ and $y$ there are $\binom{n}{k}2^n$ possible $a$ and $b$ with $\|a \oplus b \oplus xy \oplus \mathbf{1}\|_0 = k$. Theorem 7 then implies

$$
\sum_{l\geq k} p_l \leq e^{-nD(k/n,1-k/n||w,1-w)}. \tag{13}
$$

To prove Theorem 6, we need one further ingredient:

**Lemma 8.** *Let $a, b \in \mathbb{R}$ with $a < b$ and $f : [a, b] \to \mathbb{R}_0^+$ be a concave function that attains its maximum at some $x^* \in [a, b]$. Then $\forall n \in \mathbb{N}$,*

$$
\frac{1}{n+1}(b-a)f(x^*)^n \leq \int_a^b f(x)^n\mathrm{d}x \leq (b-a)f(x^*)^n \tag{14}
$$

The proof is given in Appendix A. Now we are ready to prove Theorem 6:

*Proof of Theorem 6.* Let $w = \frac{2+\sqrt{2}}{4}$ be the quantum value of the CHSH game and define for $p \in [1 - w, w]$ the single-round box $Q(p)_{\hat{A}\hat{B}|\hat{X}\hat{Y}}$ as

$$
Q(p)(ab|xy) = \begin{cases} p/2 & \text{if } a \oplus b = xy \\ (1-p)/2 & \text{if } a \oplus b \neq xy \end{cases} \tag{15}
$$

Now set

$$
\tau_{AB|XY} = \frac{1}{2w-1}\int_{1-w}^{w} Q(p)_{AB|XY}^{\otimes n}\mathrm{d}p. \tag{16}
$$

$\tau_{AB|XY}$ is quantum because each $p \in [1-w, w]$ $Q(p)_{AB|XY}^{\otimes n}$ is quantum. Further $\tau_{AB|XY}$ is de Finetti by construction. Let $a, b, x, y \in \{0,1\}^n$ and let $k = \|a \oplus b \oplus xy \oplus \mathbf{1}\|_0$ be the number of won CHSH games of $a, b, x, y$. Let $\alpha = k/n$ and set

$$
f(p) = \frac{1}{2}p^\alpha(1-p)^{1-\alpha}. \tag{17}
$$

Then

$$
\tau(ab|xy) = \frac{1}{2w-1}\int_{1-w}^{w} 2^{-n}p^k(1-p)^{n-k}\mathrm{d}p = \frac{1}{2w-1}\int_{1-w}^{w} f(p)^n\mathrm{d}p. \tag{18}
$$

Note that

$$
f'(p) = \left(\frac{\alpha}{p} - \frac{1-\alpha}{1-p}\right)f(p) \tag{19}
$$

and

$$f''(p) = \left[\left(\frac{\alpha}{p} - \frac{1-\alpha}{1-p}\right)^2 - \frac{\alpha}{p^2} - \frac{1-\alpha}{(1-p)^2}\right] f(p)$$

$$= -\frac{\alpha(1-\alpha)}{p^2(1-p)^2} f(p) < 0. \tag{20}$$

Therefore, $f$ is concave and its maximum on the interval $[0,1]$ occurs at $p = \alpha$. Since $f$ is concave, we can apply Lemma 8 to Eq. (18) and get

$$\tau(ab|xy) \geq \frac{1}{n+1} \sup_{p \in [1-w,w]} f(p)^n. \tag{21}$$

Now we turn to the box $P_{AB|XY}$. Following the earlier notation, let $p_k$ denote the probability of winning exactly $k$ CHSH games with this distribution. We observe that

- If $\alpha > w$, then the threshold Theorem 7 implies

$$p_k \leq e^{-nD(\alpha,1-\alpha||w,1-w)}. \tag{22}$$

- If $\alpha < 1 - w$, we can use the threshold theorem to get a bound on the minimal number of won games, because the CHSH game has the property that winning exactly $k$ games is just as hard as losing exactly $k$ games (and thus winning $n - k$ games). Hence, we have

$$p_k \leq e^{-nD(\alpha,1-\alpha||1-w,w)}. \tag{23}$$

- If $\alpha \in [1-w, w]$, we can rewrite the trivial bound $p_k \leq 1$ in the form

$$p_k \leq 1 = e^{-nD(\alpha,1-\alpha||\alpha,1-\alpha)}. \tag{24}$$

Hence, we can summarize the implications of the threshold theorem as

$$p_k \leq \sup_{p \in [1-w,w]} e^{-nD(\alpha,1-\alpha||p,1-p)}. \tag{25}$$

We can simplify the term in the supremum by inserting the definition of relative entropy:

$$e^{-nD(\alpha,1-\alpha||p,1-p)} = \left(\frac{p}{\alpha}\right)^{\alpha n} \left(\frac{1-p}{1-\alpha}\right)^{(1-\alpha)n} = \frac{f(p)^n}{f(\alpha)^n}. \tag{26}$$

Now recall that by Eq. (11), $P(ab|xy)$ is related to $p_k$ by

$$P(ab|xy) = \frac{p_k}{2^n \binom{n}{k}}. \tag{27}$$

It is a well-known identity of the Beta function that

$$\binom{n}{k}^{-1} = (n+1) \int_0^1 t^k (1-t)^{n-k} dt = 2^n(n+1) \int_0^1 f(p)^n dp$$

$$\leq 2^n(n+1)f(\alpha)^n \tag{28}$$

where for the last inequality we used Lemma 8 and the fact that the maximum of $f$ on $[0,1]$ is $f(\alpha)$. Inserting Eq. (28) followed by Eqs. (25)–(26) into Eq. (27) gives

$$P(ab|xy) \leq (n+1)p_k f(\alpha)^n \leq (n+1) \sup_{p \in [1-w,w]} f(p)^n. \tag{29}$$

Combining Eqs. (21) and (29) yields $P(ab|xy) \leq (n+1)^2 \tau(ab|xy)$, as desired. □

The arguments in the proof of Theorem 6 are not specific to CHSH symmetry. In fact, in Appendix B we show that we get such a de Finetti theorem whenever a threshold theorem analogous to Theorem 7 holds.

## 2.2. The Second de Finetti Theorem

The de Finetti theorem discussed in the previous section is similar to the de Finetti theorems for boxes shown in [16]; they show that the entries of some given box are *upper bounded* by the entries of a de Finetti box. The original de Finetti theorems for random variables and quantum states are of a different flavor: They show that the marginal on the first $k$ rounds of an arbitrary $n$-round permutation-invariant state is *close* to a de Finetti state if $k \ll n$. In this section, we show a theorem of this type for CHSH symmetric boxes. We use the following distance measure on the space of boxes:

**Definition 9.** Let $P_{A|X}$ and $Q_{A|X}$ be two boxes with the same input set $\mathcal{X}$ and output set $\mathcal{A}$. Their distance is

$$\left\| P_{A|X} - Q_{A|X} \right\| = \max_{x \in \mathcal{X}} \sum_{a \in \mathcal{A}} |P(a|x) - Q(a|x)|. \tag{30}$$

This distance is just the $\ell^1$ distance of the probability distributions of $a$, maximized over the input $x$. To state the de Finetti theorem, we need to introduce the notion of the marginal of an $n$-round box. In general, this marginal may not be well defined without some kind of non-signaling condition across different rounds (since otherwise the output distribution of one round could potentially depend on the input in another round). However, it turns out that for CHSH symmetric boxes this is indeed well defined, as we now show.

**Lemma 10.** *Let $P_{AB|XY}$ be an $n$-round CHSH symmetric quantum box and let $1 \leq k \leq n$ be an integer. Then the expression*

$$P^k(a_1 \ldots a_k, b_1 \ldots b_k | x_1 \ldots x_k, y_1 \ldots y_k)$$
$$:= \sum_{\substack{a_{k+1}\ldots a_n \\ b_{k+1}\ldots b_n}} P(a_1 \ldots a_n, b_1 \ldots b_n | x_1 \ldots x_n, y_1 \ldots y_n) \tag{31}$$

*is independent of the choice of $x_{k+1} \ldots x_n$ and $y_{k+1} \ldots y_n$, and we shall refer to it as the* marginal *of the first $k$ rounds. Furthermore, $P^k_{AB|XY}$ is a CHSH symmetric quantum box (of $k$ rounds).*

*Proof.* We shall use the notation $a = (a_1 \ldots a_k)$ and $a' = (a_{k+1} \ldots a_n)$, and define $b, b', x, x', y, y'$ analogously. To see that $P^k_{AB|XY}$ is independent of the choice of $x'$ and $y'$, we calculate

$$
\begin{aligned}
P^k(ab|xy) &= \sum_{a'b'} P(aa', bb'|xx', yy') \\
&= \sum_{a'b'} P(aa', b(b' \oplus x'y')|x0, y0) \\
&= \sum_{a'b'} P(aa', bb'|x0, y0),
\end{aligned}
\tag{32}
$$

where in the second equality we used the CHSH symmetry of $P_{AB|XY}$ and in the third equality we shifted the summation variable $b'$ by $x'y'$.

To see the CHSH symmetry of $P^k_{AB|XY}$, note that by CHSH symmetry of $P_{AB|XY}$, $P(aa', bb'|x0, y0)$ only depends on $a \oplus b \oplus xy$ and $a' \oplus b'$. Hence, $P^k(ab|xy)$ only depends on $a \oplus b \oplus xy$. Furthermore, the permutation invariance of $P_{AB|XY}$ immediately implies that $P^k_{AB|XY}$ is also permutation-invariant, and hence, we conclude that $P^k_{AB|XY}$ is CHSH symmetric. Finally, the fact that $P^k_{AB|XY}$ is a quantum box immediately follows from the fact that $P_{AB|XY}$ is quantum. $\qquad\square$

Now we can state the de Finetti theorem:

**Theorem 11.** *Let $P_{AB|XY}$ be an $n$-round CHSH symmetric quantum box and let $P^k_{AB|XY}$ be the marginal of the first $k$ rounds as defined in Eq. (31). There is a $k$-round CHSH symmetric quantum de Finetti box $\tau_{AB|XY}$ such that*

$$
\left\| P^k_{AB|XY} - \tau_{AB|XY} \right\| \leq \left( C\sqrt{\ln(n/k)} + 4 \right) \sqrt{\frac{k}{n}} + \frac{4k}{n} = \mathcal{O}\left( \sqrt{\ln(n/k)\frac{k}{n}} \right)
\tag{33}
$$

*with $C = \frac{2}{\sqrt{2-\sqrt{2}}} \approx 2.6$.*

For the proof of Theorem 11, we first note that the distance between two CHSH symmetric boxes is just the $\ell^1$ distance between the distributions of the wins and losses of the CHSH game, which are independent from the input into the box.

**Lemma 12.** *Let $P_{AB|XY}$ and $Q_{AB|XY}$ be CHSH symmetric $n$-round boxes and $W = A \oplus B \oplus XY \oplus \mathbf{1} \in \{0,1\}^n$ be the random variable that indicates in which rounds the CHSH game was won. Let $P_W$ and $Q_W$ the distribution of $W$. Then*

$$
\left\| P_{AB|XY} - Q_{AB|XY} \right\| = \left\| P_W - Q_W \right\|_1.
\tag{34}
$$

*Proof.* For all $x, y$

$$
P(ab|xy) = 2^{-n} P_W(a \oplus b \oplus xy \oplus \mathbf{1})
\tag{35}
$$

so
$$\sum_{a,b}|P(ab|xy) - Q(ab|xy)| = \sum_{a,w}2^{-n}|P_W(w) - Q_W(w)| = \|P_W - Q_W\|_1.$$
(36)

□

Another ingredient for the proof of Theorem 11 is a bound on the $\ell^1$-distance between two binomial distributions:

**Lemma 13.** *Let $k \in \mathbb{N}$ and $p, q \in (0,1)$. Denote by $P = \mathrm{Binom}(k,p)$ and $Q = \mathrm{Binom}(k,q)$ the binomial distributions with $k$ trials and success probabilities $p$ and $q$. Then*
$$\|P - Q\|_1 \le 2\sqrt{\frac{n}{\min\{q, 1-q\}}}|p-q|.$$
(37)

*Proof.* Denote by $P_0$ and $Q_0$ the Bernoulli distributions with success probability $p$ and $q$, respectively. We use Pinsker's inequality and the reverse Pinsker's inequality (Lemma 4.1 in [36]) to calculate

$$
\begin{aligned}
\|P - Q\|_1 &\le \sqrt{2D(P\|Q)} && \text{Pinsker's inequality}\\
&= \sqrt{2nD(P_0\|Q_0)} && \text{additivity of relative entropy}\\
&\le \sqrt{\frac{n\|P_0 - Q_0\|_1^2}{\min\{q, 1-q\}}} && \text{reverse Pinsker's inequality}\\
&= 2\sqrt{\frac{n}{\min\{q, 1-q\}}}|p-q|. && (38)
\end{aligned}
$$

□

Now we are ready to prove the de Finetti theorem:

*Proof of Theorem 11.* Denote by $p_N$ the probability that Alice and Bob win exactly $N$ CHSH games on the box $P_{AB|XY}$. Then $W = A \oplus B \oplus XY \oplus \mathbf{1}$ has a permutation-invariant distribution $P_W$ with
$$\sum_{w:\sum_i w_i = N} P_W(w) = p_N.$$
(39)

By the de Finetti theorem for random variables [12], we have
$$\left\| P_W^k - \sum_{N=0}^{n} p_N \mathrm{Binom}\left(k, \frac{N}{n}\right) \right\| \le \frac{4k}{n},$$
(40)
where $P_W^k$ denotes the distribution of the first $k$ bits of $W$.

For $p \in [0,1]$, denote by $Q(p)_{\hat{A}\hat{B}|\hat{X}\hat{Y}}$ the single-round box with CHSH winning probability $p$. Then by Lemma 12 and Eq. (40),
$$\left\| P_{AB|XY}^k - \sum_{N=0}^{n} p_N Q\left(\frac{N}{n}\right)^{\otimes k} \right\| \le \frac{4k}{n}$$
(41)

The box $\sum_{N=0}^{n} p_N Q\left(\frac{N}{n}\right)^{\otimes k}$ is CHSH symmetric and de Finetti, but not quantum. The problems are the terms with $N > nw$ and $N < n(1-w)$, where $w = \frac{2+\sqrt{2}}{4}$. We define a quantum CHSH symmetric de Finetti box as

$$\tau_{AB|XY} = \sum_{N=0}^{n} p_N \begin{cases} Q(w)^{\otimes k} & \text{if } N > wn \\ Q\left(\frac{N}{n}\right)^{\otimes k} & \text{if } N \in [(1-w)n, wn] \\ Q(1-w)^{\otimes k} & \text{if } N < (1-w)n \end{cases}. \qquad (42)$$

We will show

$$\left\| \tau_{AB|XY} - \sum_{N=0}^{n} p_N Q\left(\frac{N}{n}\right)^{\otimes k} \right\| \leq \left( C\sqrt{\ln(n/k)} + 4 \right) \sqrt{\frac{k}{n}}. \qquad (43)$$

The statement of Theorem 11 then follows by combining this bound and the bound in Eq. (41) using the triangle inequality.

Let $\delta > 0$. We split the sum in the definition of $\tau_{AB|XY}$ to obtain

$$\left\| \tau_{AB|XY} - \sum_{N=0}^{n} p_N Q\left(\frac{N}{n}\right)^{\otimes k} \right\| \leq \sum_{N \in [wn, n]} p_N \left\| Q(w)^{\otimes k} - Q\left(\frac{N}{n}\right)^{\otimes k} \right\|$$

$$+ \sum_{N \in [0, (1-w)n]} p_N \left\| Q(1-w)^{\otimes k} - Q\left(\frac{N}{n}\right)^{\otimes k} \right\|$$

$$\leq \sum_{N \in [wn, (w+\delta)n]} p_N \left\| Q(w)^{\otimes k} - Q(w+\delta)^{\otimes k} \right\|$$

$$+ \sum_{N \in [(1-w-\delta)n, (1-w)n]} p_N \left\| Q(1-w)^{\otimes k} - Q(1-w-\delta)^{\otimes k} \right\|$$

$$+ 2 \sum_{N \in [0, (1-w-\delta)n] \cup [(w+\delta)n, n]} p_N \qquad (44)$$

where in the last line we used $\left\| P_{AB|XY} - Q_{AB|XY} \right\| \leq 2$ for all normalized boxes $P_{AB|XY}$ and $Q_{AB|XY}$. We start by bounding the terms with $N \in [wn, (w+\delta)n]$ and $N \in [(1-w-\delta)n, (1-w)n]$ using Lemma 13:

$$\left\| Q(w)^{\otimes k} - Q(w+\delta)^{\otimes k} \right\| \leq \frac{2}{\sqrt{1-w}} \sqrt{k}\delta = 2C\sqrt{k}\delta \qquad (45)$$

where we used $\frac{2}{\sqrt{1-w}} = \frac{4}{\sqrt{2-\sqrt{2}}} = 2C$. Analogously we find

$$\left\| Q(1-w)^{\otimes k} - Q(1-w-\delta)^{\otimes k} \right\| \leq 2C\sqrt{k}\delta \qquad (46)$$

so that by $\sum_N p_N = 1$

$$\sum_{N \in [wn, (w+\delta)n]} p_N \left\| Q(w)^{\otimes k} - Q(w+\delta)^{\otimes k} \right\|$$

$$+ \sum_{N \in [(1-w-\delta)n, (1-w)n]} p_N \left\| Q(1-w)^{\otimes k} - Q(1-w-\delta)^{\otimes k} \right\|$$

$$\leq 2C\sqrt{k}\delta. \qquad (47)$$

Now we turn to the terms with $N > (w + \delta)n$ and $N < (1 - w - \delta)n$. By the threshold theorem for the CHSH game (Theorem 7), it holds that

$$\sum_{N>(w+\delta)n} p_N \leq e^{-nD(w+\delta,1-w-\delta||w,1-w)} \tag{48}$$

$$\leq e^{-2n\delta^2} \tag{49}$$

where in the last step we used Pinsker's inequality, i.e $D(p, 1 - p||q, 1 - q) \geq 2|p - q|^2$. Analogously also

$$\sum_{N<(1-w-\delta)n} p_N \leq e^{-2n\delta^2}. \tag{50}$$

Putting together the bounds for $N \in [wn, (w + \delta)n]$ and $N > (w + \delta)n$, we find

$$\left\| \tau_{AB|XY} - \sum_{N=0}^{n} p_N Q\left(\frac{N}{n}\right)^{\otimes k} \right\| \leq 2C\sqrt{k}\delta + 4e^{-2n\delta^2}. \tag{51}$$

Now we choose

$$\delta = \frac{1}{2}\sqrt{\frac{\ln(n/k)}{n}} \tag{52}$$

and obtain

$$\left\| \tau_{AB|XY} - \sum_{N=0}^{n} p_N Q\left(\frac{N}{n}\right)^{\otimes k} \right\| \leq \left( C\sqrt{\ln(n/k)} + 4 \right)\sqrt{\frac{k}{n}}. \tag{53}$$

This completes the proof. □

The choice of $\delta$ in Eq. (52) is not optimal, and it does not give the minimal possible error term in Theorem 11. However, the improvement that can be achieved by choosing $\delta$ optimally does not change the $\mathcal{O}\left(\sqrt{\ln(n/k)k/n}\right)$ behavior. To see this, choose

$$\delta = \frac{1/2\sqrt{\ln(n/k)} + \beta}{\sqrt{n}} \tag{54}$$

for some $\beta > -\sqrt{\ln(n/k)}/2$. Then the error term is given by

$$\epsilon := 2C\sqrt{k}\delta + 4e^{-2n\delta^2} = \left( C\sqrt{\ln(n/k)} + 2C\beta + 4e^{-2\beta^2 - 2\beta\sqrt{\ln(n/k)}} \right)\sqrt{k/n}. \tag{55}$$

The optimal $\beta$ is such that $C' = 2C\beta + 4e^{-2\beta^2 - 2\beta\sqrt{\ln(n/k)}}$ is minimized. A numerical optimization indicates that for $\ln(n/k) = 0$ the minimum of $C'$ is achieved at $\beta = 0$, so $C' = 4$. As $\ln(n/k)$ increases the minimum of $C'$ decreases slowly, at $\ln(n/k) = 10$ it is given by $C' \approx 2.03$, and at $\ln(n/k) = 100$ by $C' \approx 0.96$. As $\ln(n/k) \to \infty$, it converges $C' \to 0$, which can be seen by choosing $\beta = (\ln(n/k))^{-1/4}$. Regardless of the choice of $\beta$, the error is always at least $C\sqrt{\ln(n/k)k/n}$.

# 3. Applications

In this section, we show how our first de Finetti theorem (Theorem 6) has applications in DIQKD security proofs, by first using it to derive a bound on channel distinguishability, then discussing its implications for security proofs. This result and the proof of it are analogous to theorem 25 in [16], except that we use Theorem 6 as the de Finetti theorem, rather than the statement in Eq. (5). We remark that the works [19,20] also used threshold theorems (of somewhat different forms) to obtain DIQKD security proofs. However, as discussed in Introduction, their proof techniques currently only apply to protocols using one-way error correction and yield lower asymptotic key rates compared to the iid case. In contrast, the results we derive here could be applied to all protocols having the appropriate symmetry properties. While they currently do not yield a full reduction to the iid case, our hope is that it would be possible to develop them further to obtain security proofs that are more generally applicable and yield higher asymptotic key rates compared to [19,20], as was the case for de Finetti theorems in device-dependent QKD [11].

## 3.1. Bound on the Diamond Distance Between Channels

Here we consider channels on boxes of the following form: A channel $\mathcal{E}$ that acts on boxes of the form $P_{A|X}$ and outputs a random variable $R$ as its result is described by a probability distribution $P_X^{\mathcal{E}}$ on $\mathcal{X}$ and a conditional probability distribution $P_{R|AX}^{\mathcal{E}}$ which determines the result $R$ given $A$ and $X$. When acting on $P_{A|X}$, the channel produces a distribution on $R$ given by

$$\mathcal{E}(P_{A|X})(r) = \sum_{x,a} P_X^{\mathcal{E}}(x) P_{A|X}(a|x) P_{R|AX}^{\mathcal{E}}(r|ax). \tag{56}$$

This definition is general enough to capture all protocols in a parallel DIQKD scenario, where all bits of the $n$-bit input $\mathcal{X}$ are entered at the same time into the box. It does not cover all protocols that are possible in a sequential DIQKD scenario [8], where some of the input bits are only given to the box after some output bits have been received. In such a sequential scenario, it is in principle possible to construct channels where the input to the box in some round depends on the output of the box in previous rounds.

If we consider boxes $P_{AE|XZ}$, where the additional $E, Z$ interface is held by Eve, we can also apply the channel $\mathcal{E}$ only to the $A, X$ interface to obtain a box with input $Z$ and outputs $R$ and $E$. We will denote this box by $(\mathcal{E} \otimes \mathrm{id})(P_{AE|XZ})_{RE|Z}$.

We define the distance between two channels $\mathcal{E}$ and $\mathcal{F}$ by how well Eve can distinguish the boxes $(\mathcal{E} \otimes \mathrm{id})(P_{AE|XZ})_{RE|Z}$ and $(\mathcal{F} \otimes \mathrm{id})(P_{AE|XZ})_{RE|Z}$ if she is also given access to $R$. Then she can choose her input $Z$ dependent on $R$. This leads to the following definition [16,37]:

**Definition 14.** Let $\mathcal{E}$ and $\mathcal{F}$ be two channels acting on boxes of the form $P_{A|X}$. The distinguishability of $\mathcal{E}$ and $\mathcal{F}$ using the box $P_{AE|XZ}$ is given by

$$
\left\| (\mathcal{E} - \mathcal{F}) \otimes \mathrm{id} \left( P_{AE|XZ} \right) \right\|
$$
$$
= \sum_r \max_z \sum_e \left| \sum_{a,x} P_{AE|XZ}(ae|xz) \left( P_X^{\mathcal{E}}(x) P_{R|AX}^{\mathcal{E}}(r|ax) - P_X^{\mathcal{F}}(x) P_{R|AX}^{\mathcal{F}}(r|ax) \right) \right|.
$$
(57)

We define the diamond distance between the channels with respect to some set $\mathcal{P}$ of boxes to be the following:

$$
||\mathcal{E} - \mathcal{F}||_{\Diamond}^{\mathcal{P}} = \sup_{P_{AE|XZ} \in \mathcal{P}} \left\| (\mathcal{E} - \mathcal{F}) \otimes \mathrm{id} \left( P_{AE|XZ} \right) \right\|.
$$
(58)

Similarly to the usual diamond distance between quantum channels, the above definition of diamond distance with respect to some set $\mathcal{P}$ is a measure of how distinguishable the channels are with respect to a distinguisher that can only use boxes from $\mathcal{P}$. Simple choices of $\mathcal{P}$ include, for instance, the sets of quantum or non-signaling boxes. For the following main theorem of this section we will, however, take $\mathcal{P}$ to be the set of quantum boxes $P_{ABE|XYZ}$ such that the marginal $P_{AB|XY}$ has CHSH symmetry, and denote the diamond distance with respect to this $\mathcal{P}$ as $||\mathcal{E} - \mathcal{F}||_{\Diamond}^{\mathrm{quantum,CHSH}}$. Note that if the action of the channels $\mathcal{E}, \mathcal{F}$ can be described by Alice and Bob first performing the depolarizing procedure described above, this restriction causes no change in the diamond distance as compared to choosing $\mathcal{P}$ to be the entire set of quantum boxes $P_{ABE|XYZ}$.

**Theorem 15.** *Let $\mathcal{E}$ and $\mathcal{F}$ two channels on n-round boxes of the form $P_{AB|XY}$, and let $\tau_{AB|XY}$ be the de Finetti box from Theorem 6. Then*

$$
||\mathcal{E} - \mathcal{F}||_{\Diamond}^{\mathrm{quantum,CHSH}} \leq (n+1)^2 \sup_{\tau_{ABE|XYZ}} \left\| (\mathcal{E} - \mathcal{F}) \otimes \mathrm{id}(\tau_{ABE|XYZ}) \right\| \quad (59)
$$

*where the supremum is taken over all non-signaling boxes that have the marginal $\tau_{AB|XY}$.*

*Proof of Theorem 15.* Let $P_{ABE|XYZ}$ be a quantum box whose marginal $P_{AB|XY}$ has CHSH symmetry. Let $R_{AB|XY}$ be such that

$$
\tau_{AB|XY} = \frac{1}{(n+1)^2} P_{AB|XY} + \left( 1 - \frac{1}{(n+1)^2} \right) R_{AB|XY}.
$$
(60)

By Theorem 6 all entries of $R_{AB|XY}$ are positive. Because the non-signaling condition is linear and $\tau_{AB|XY}$ and $P_{AB|XY}$ are non-signaling, $R_{AB|XY}$ is also non-signaling. Now we define an extension $\tau_{ABE|XYZ}$ of $\tau_{AB|XY}$ as follows: The box has one more possible outcome for Eve then the box $P_{ABE|XYZ}$. We will call this additional outcome $e^*$. The box $\tau_{ABE|XYZ}$ then works as follows: With probability $(n+1)^{-2}$, the box acts just like $P_{ABE|XYZ}$, and with probability $1 - (n+1)^{-2}$, it always returns $e^*$ to Eve and acts like $R_{AB|XY}$ for Alice and

Bob. Formally, this is given by

$$\tau(abe|xyz) = \begin{cases} \frac{1}{(n+1)^2} P(abe|xyz) & \text{if } e \neq e^* \\ \left(1 - \frac{1}{(n+1)^2}\right) R(ab|xy) & \text{if } e = e^* \end{cases}. \tag{61}$$

Since $\tau_{ABE|XYZ}$ is the linear combination of two non-signaling boxes, it is non-signaling itself. Furthermore, by Eq. (60) it is an extension of $\tau_{AB|XY}$. Finally, it holds that

$$||(\mathcal{E} - \mathcal{F}) \otimes \text{id}(\tau_{ABE|XYZ})||$$

$$= \sum_r \max_z \sum_e \left| \sum_{a,b,x,y} \tau(abe|xyz) \left( P^{\mathcal{E}}(xy) P^{\mathcal{E}}(r|abxy) - P^{\mathcal{F}}(xy) P^{\mathcal{F}}(r|abxy) \right) \right|$$

$$\geq \sum_r \max_z \sum_{e \neq e^*} \left| \sum_{a,b,x,y} \tau(abe|xyz) \left( P^{\mathcal{E}}(xy) P^{\mathcal{E}}(r|abxy) - P^{\mathcal{F}}(xy) P^{\mathcal{F}}(r|abxy) \right) \right|$$

$$= (n+1)^{-2} \sum_r \max_z \sum_{e \neq e^*} \left| \sum_{a,b,x,y} P(abe|xyz) \left( P^{\mathcal{E}}(xy) P^{\mathcal{E}}(r|abxy) - P^{\mathcal{F}}(xy) P^{\mathcal{F}}(r|abxy) \right) \right|$$

$$= (n+1)^{-2} ||(\mathcal{E} - \mathcal{F}) \otimes \text{id}(P_{ABE|XYZ})||. \tag{62}$$

Hence, for all $P_{ABE|XYZ}$

$$||(\mathcal{E} - \mathcal{F}) \otimes \text{id}(P_{ABE|XYZ})|| \leq (n+1)^2 \sup_{\tau_{ABE|XYZ}} ||(\mathcal{E} - \mathcal{F}) \otimes \text{id}(\tau_{ABE|XYZ})||. \tag{63}$$

Taking the supremum over all $P_{ABE|XYZ}$ with CHSH symmetric marginal $P_{AB|XY}$ yields the claim. $\qquad\square$

## 3.2. Implications for DIQKD Security Proofs

Theorem 15 can be seen as a version of the postselection theorem for quantum channels [11]. It allows us to bound the distance between two channels by the distinguishability of the channels when Eve is restricted to extensions of a fixed de Finetti box. This could potentially be a useful tool in security proofs of DIQKD protocols, because a protocol can be defined to be secure if its diamond distance to an ideal protocol is small [28, 29]. In particular, Theorem 15 implies that to prove security against coherent quantum attacks, it is sufficient to prove security for the case where the marginal of Alice and Bob is given by $\tau_{AB|XY}$, and Eve possesses a non-signaling extension of this box. This helps to simplify the task of a DIQKD security proof, because it means that it suffices to analyze (extensions of) the *specific* box $\tau_{AB|XY}$, which has the convenient property of being a convex combination of iid quantum boxes.

However, there is a caveat: Although the box $\tau_{AB|XY}$ is quantum, the extensions $\tau_{ABE|XYZ}$ in the theorem statement here are allowed to be general non-signaling boxes. Furthermore, we will show in the next section that an adversary who has access to arbitrary non-signaling extensions of $\tau_{AB|XY}$ can actually be strictly better at distinguishing channels than an adversary who has only access to collective attack boxes. Hence, Theorem 15 does not immediately

yield security against coherent attacks from security against collective attacks —still, since it does allow a "partial" reduction to the latter (namely, allowing us to focus on extensions of a quantum de Finetti box $\tau_{AB|XY}$), it may still simplify DIQKD security proofs.

We also remark that for our second de Finetti theorem (Theorem 11), we currently do not have in mind an explicit application of it in DIQKD security proofs. Still, we presented it in this work in case it has applications in other contexts—for instance, it might be useful in proving properties that only depend on the box $P_{AB|XY}$ itself, rather than involving its extensions as in DIQKD security proofs. It is also more similar to the original de Finetti theorem for classical random variables, or the early versions for quantum states developed in, e.g., [14].

## 4. Difficulties in Bounding the Diamond Distance by Restriction to Collective Attacks

Theorem 15 shows that to bound the diamond distance between two channels $\mathcal{E}$ and $\mathcal{F}$ it is sufficient to restrict the attacker to non-signaling extensions of a fixed de Finetti box. There are many desirable strengthenings of this result: For example, one could restrict the attacker only to quantum extensions of the de Finetti box. One could also further restrict the attacker to use only quantum extensions of iid boxes, instead of the fixed de Finetti box. Finally, one could also restrict the attacker to collective attack boxes (defined below), as would be desirable to conclude security against coherent attacks directly from security against collective attacks. In this section, we will see that a theorem like Theorem 15 does not hold for this strongest restriction; more precisely, we show that it is impossible for the bound (59) to hold if the supremum is instead restricted to collective attack boxes (which we define later below). It remains open whether such a theorem holds for one of the other strengthenings mentioned above, or whether a reduction to collective attacks in a somewhat different form is possible. (We note that the answers to these questions do not straightforwardly follow from existing no-go theorems on non-signaling privacy amplification [38,39], since in our result $\tau_{AB|XY}$ is restricted to a convex combination of quantum distributions rather than non-signaling distributions.)

We start by defining the boxes that an attacker is allowed to use in collective attacks. While there is potentially some flexibility in defining this, here we use a definition that essentially corresponds to the boxes considered in the security proofs of [5,10], up to a collective measurement on Eve's side information:

**Definition 16.** An $n$-round quantum box $P_{ABE|XYZ}$ is a collective attack box if there are

- single-round Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ for Alice and Bob and a Hilbert space $\mathcal{H}_E$ for Eve and
- a state $\rho_{ABE} \in \text{End}\left(\mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_B^{\otimes n} \otimes \mathcal{H}_E\right)$ such that the marginal $\rho_{AB}$ is iid and

- for each $x$ a POVM $\{E^{a,x} \in \mathrm{End}(\mathcal{H}_A)|a\}$ on $\mathcal{H}_A$, for each $y$ a POVM $\{F^{b,y} \in \mathrm{End}(\mathcal{H}_B)|b\}$ on $\mathcal{H}_B$ and for each $z$ a POVM $\{G^{e,z} \in \mathrm{End}(\mathcal{H}_E)|e\}$ on $\mathcal{H}_E$

such that

$$P(abe|xyz) = \mathrm{tr}\left[\rho_{ABE}\left(E^{a_1,x_1} \otimes \ldots \otimes E^{a_n,x_n} \otimes F^{b_1,y_1} \otimes \ldots \otimes F^{b_n,y_n} \otimes G^{e,z}\right)\right].$$
(64)

We remark on two aspects of the above definition. Firstly, note that we assume the Hilbert spaces of Alice and Bob can be split into $n$ rounds, but assume no internal structure of Eve's Hilbert space. However, since the state $\rho_{AB}$ is iid and thus has an iid purification, the state $\rho_{ABE}$ is related by a local operation on Eve's system to this iid purification. Since we assume nothing about $G^{e,z}$ except that it is a valid POVM, we can absorb this local operation into $G^{e,z}$ and thus describe any collective attack box also with a state $\rho_{ABE}$ that is iid. Collective attack boxes can thus be seen as boxes that are essentially iid, up to Eve performing a local operation on her systems followed by a joint measurement. Secondly, the fact that the definition inherently incorporates this measurement means that Eve's system is forced to be a box rather than a genuine quantum state. However, for the purposes of computing diamond distance, this in fact does not make a difference (as long as arbitrary POVMs $G^{e,z}$ are allowed in the definition)—observe that the process of a distinguisher producing a guess for the channel can be described as it performing a POVM on its systems, and the optimal such POVM essentially induces a valid choice of $G^{e,z}$ in the above definition.

A crucial observation on collective attack boxes is the following: Consider the box $P^{e,z}_{AB|XY}$ which described the outcomes of Alice and Bob conditioned on Eve inputting $z$ and getting outcome $e$. It is given by

$$P^{e,z}_{AB|XY}(ab|xy) = \frac{P_{ABE|XYZ}(abe|xyz)}{P_{E|Z}(e|z)}$$
$$= \mathrm{tr}\left[\rho^{e,z}_{AB}\left(E^{a_1,x_1} \otimes \ldots \otimes E^{a_n,x_n} \otimes F^{b_1,y_1} \otimes \ldots \otimes F^{b_n,y_n}\right)\right]$$
(65)

where $\rho^{e,z}_{AB}$ is a valid state,

$$\rho^{e,z}_{AB} = \frac{\mathrm{tr}_E\left[\rho_{ABE}\left(\mathrm{id}_A \otimes \mathrm{id}_B \otimes G^{e,z}\right)\right]}{\mathrm{tr}\left[\rho_{ABE}\left(\mathrm{id}_A \otimes \mathrm{id}_B \otimes G^{e,z}\right)\right]}.$$
(66)

Because $\sum_a E^{a,x} = \sum_b F^{b,y} = \mathrm{id}$, we see that $P^{e,z}_{AB|XY}$ is not only non-signaling between Alice and Bob, but also between the individual rounds. This means that for example $\sum_{a_1} P^{e,z}_{AB|XY}(a_1 a_2 \ldots a_n b|xy)$ does not depend on $x_1$.

The following main result of this section exploits this insight:

**Theorem 17.** *For each $n > 1$, there exist two channels $\mathcal{E}$ and $\mathcal{F}$ acting on $n$-round boxes of the form $P_{AB|XY}$ such that $||(\mathcal{E} - \mathcal{F}) \otimes \mathrm{id}(P_{ABE|XYZ})|| = 0$ for all collective attack boxes $P_{ABE|XYZ}$, but $||\mathcal{E} - \mathcal{F}||_\diamond^{\mathrm{quantum,CHSH}} \neq 0$.*

Theorem 17 shows that a statement like Theorem 15 cannot hold if we maximize only over collective attack boxes instead of all non-signaling extensions of the fixed de Finetti box (not even for, say, an exponential prefactor instead of $(n+1)^2$). This shows that an attacker who has access to any non-signaling extension of the fixed de Finetti box is stronger than an attacker who has only access to collective attack boxes.

In the proof of Theorem 17, we will use that all collective attack boxes are non-signaling between the rounds of Alice and Bob, and that the non-signaling condition is linear. The following lemma will be crucial. It states that for each linear subspace of the probability distributions on some set, there are two channels (which act on probability distributions, not yet on boxes) that cannot be distinguished by any probability distribution in the linear subspace:

**Lemma 18.** *Let $\mathcal{X}$ be some finite set. We treat the unnormalized probability distributions on $\mathcal{X}$ as an orthant of an $|\mathcal{X}|$-dimensional real vector space. Let $\mathcal{P}$ be some linear subspace in this vector space, and $Q = (Q(x))_{x \in \mathcal{X}} \notin \mathcal{P}$. Then there are two conditional probability distributions $P^{\mathcal{E}}_{R|X}$ and $P^{\mathcal{F}}_{R|X}$ such that the following holds: Denote for a probability distribution $P$ on $\mathcal{X}$ by $\mathcal{E}(P)$ and $\mathcal{F}(P)$ the distributions on $\mathcal{R}$ which are obtained by first sampling $x$ using $P$ and the sampling $r$ using $P^{\mathcal{E}}_{R|X}$ and $P^{\mathcal{F}}_{R|X}$, i.e., $\mathcal{E}(P)(r) = \sum_x P(x)P^{\mathcal{E}}(r|x)$. Then*

$$\mathcal{E}(P) = \mathcal{F}(P) \tag{67}$$

*for all $P \in \mathcal{P}$ and*

$$\mathcal{E}(Q) \neq \mathcal{F}(Q). \tag{68}$$

*Proof.* There exists a vector $\Delta = (\Delta_x)_{x \in \mathcal{X}}$ with $|\Delta_x| \leq 1$ for all $x$ and $\Delta \cdot P = 0$ for all $P \in \mathcal{P}$ and $\Delta \cdot Q \neq 0$. Take $\mathcal{R} = \{0, 1\}$ and

$$P^{\mathcal{E}}(0|x) = \frac{1 + \Delta_x}{2} \tag{69}$$

$$P^{\mathcal{E}}(1|x) = \frac{1 - \Delta_x}{2} \tag{70}$$

$$P^{\mathcal{F}}(0|x) = \frac{1 - \Delta_x}{2} \tag{71}$$

$$P^{\mathcal{F}}(1|x) = \frac{1 + \Delta_x}{2}. \tag{72}$$

Then for $P$ any probability distribution on $\mathcal{X}$

$$
\begin{aligned}
||\mathcal{E}(P) - \mathcal{F}(P)||_1 &= \sum_r \left| \sum_x P(x)(P^{\mathcal{E}}(r|x) - P^{\mathcal{F}}(r|x)) \right| \\
&= |P \cdot \Delta| + |P \cdot (-\Delta)| \\
&= 2|P \cdot \Delta|.
\end{aligned}
\tag{73}
$$

Hence, for all $P \in \mathcal{P}$

$$||\mathcal{E}(P) - \mathcal{F}(P)||_1 = 0 \tag{74}$$

and

$$||\mathcal{E}(Q) - \mathcal{F}(Q)||_1 \neq 0. \tag{75}$$

$\square$

We will now construct two channels $\mathcal{E}$ and $\mathcal{F}$ that act on boxes $P_{A|X}$ (i.e., we consider only Alice) that cannot be distinguished by any collective attack box, but that can be distinguished by a certain quantum box that is not a collective attack box. We will then see how to modify this construction to include Bob and to ensure that the box used to distinguish both channels has CHSH symmetry on the marginal of Alice and Bob.

**Lemma 19.** *For each $n > 1$, there are two channels $\mathcal{E}$ and $\mathcal{F}$ that act on $n$-round boxes $P_{AE|XZ}$ such that $||(\mathcal{E} - \mathcal{F}) \otimes \mathrm{id}(P_{AE|XZ})|| = 0$ for all collective attack boxes $P_{AE|XZ}$, but $||\mathcal{E} - \mathcal{F}||_\diamond^{\mathrm{quantum,CHSH}} \neq 0$*

*Proof.* We first construct the channels $\mathcal{E}$ and $\mathcal{F}$, then show that Eve cannot distinguish them if she is restricted to collective attack boxes, and finally show that there is a quantum box (naturally not a collective attack box) that can be used to distinguish both channels. We construct the channels $\mathcal{E}$ and $\mathcal{F}$ as follows, depending on a parameter $m > n/2$. For both channels, Alice does the following steps:

1. She enters uniformly random inputs $x_1, \ldots, x_n$ into the inputs of her box.
2. She collects the outputs $a_1, \ldots, a_n$.
3. She calculates

$$t = \sum_{i=1}^{m} x_i \tag{76}$$

and

$$w = \sum_{i=m+1}^{n} a_i. \tag{77}$$

Now consider the linear subspace $\mathcal{P}$ of probability distributions on the $(w, t)$ given by the linear constraints

$$\frac{P(w,t)}{2^{-m}\binom{m}{t}} = \frac{P(w,t')}{2^{-m}\binom{m}{t'}} \tag{78}$$

for all $w, t, t'$ and take a $Q \notin \mathcal{P}$ (a specific $Q$ will be constructed below). Alice constructs the channels $\mathcal{E}$ and $\mathcal{F}$ by applying the conditional probability distributions $P_{R|WT}^{\mathcal{E}}$ and $P_{R|WT}^{\mathcal{F}}$ from Lemma 18 to her result $(w, t)$ from step 3.

Now we prove that Eve cannot distinguish $\mathcal{E}$ and $\mathcal{F}$ if she uses a collective attack box $P_{AE|XZ}$. For this, we use that for all $e$ and $z$ $P_{A|X}^{e,z}$ is non-signaling between the rounds of Alice. In particular, the outputs of the rounds $m + 1, \ldots n$ cannot depend on the inputs in the rounds $1, \ldots, m$, so $W$ and $T$ are

independent when generated using $P_{A|X}^{e,z}$. Hence,

$$\frac{P_{WT}^{e,z}(wt)}{2^{-m}\binom{m}{t}} = P_{W|T}^{e,z}(w|t) = P_{W|T}^{e,z}(w|t') = \frac{P_{WT}^{e,z}(wt')}{2^{-m}\binom{m}{t'}} \tag{79}$$

The box $P^{e,z}$ therefore satisfies Eq. (78), and hence, $||(\mathcal{E}-\mathcal{F})(P_{A|X}^{e,z})|| = 0$ by Lemma 18. Since this holds for all $e$ and $z$, we have also $||(\mathcal{E} - \mathcal{F}) \otimes \mathrm{id}(P_{AE|XZ})|| = 0$.

Finally we construct a box $Q_{A|X}$ that allows Eve to distinguish the channels $\mathcal{E}$ and $\mathcal{F}$ with a nonzero advantage over guessing. Notice that here Eve does not keep any system (neither quantum not classical) for herself and can distinguish the channels only from their result $R$. $Q_{A|X}$ can then be an arbitrary conditional probability distribution. Take $Q_{A|X}$ such that the result is surely $a = (1, 1, \ldots, 1)$ if $\sum_i x_i > n/2$ and surely $a = (0, 0, \ldots, 0)$ otherwise. Then if $t > n/2$,

$$\frac{Q_{WT}(n - m, t)}{2^{-m}\binom{m}{t}} = Q_{W|T}(n - m|t) = 1, \tag{80}$$

and if $t < m - n/2$,

$$\frac{Q_{WT}(n - m, t)}{2^{-m}\binom{m}{t}} = Q_{W|T}(n - m|t) = 0 \tag{81}$$

so Eq. (78) does not hold. □

Now we can adapt the statement of Lemma 19 to prove Theorem 17.

*Proof of Theorem 17.* First we generalize the construction in Lemma 19 to boxes for which also Bob has an input, i.e., boxes of the form $P_{AB|XY}$. For this, we take the channels $\mathcal{E}$ and $\mathcal{F}$ such that they act like in Lemma 19 on Alice's inputs and outputs, and give an arbitrary input $Y$ and ignore the output $B$ for Bob. Clearly, both channels still cannot be distinguished with collective attack boxes, but can be distinguished by a box $Q_{AB|XY}$, which acts like the box $Q_{A|X}$ from Lemma 19 on $A$ and $X$ and arbitrarily on $B$ and $Y$. However, $Q_{AB|XY}$ does not have CHSH symmetry. Using the depolarizing procedure in [32], we can construct a box $\tilde{Q}_{ABE|XY}$ such that the marginal $\tilde{Q}_{AB|XY}$ has CHSH symmetry and there is an output $e^*$ for Eve (corresponding to the case in which the depolarizing protocol does nothing), such that $\tilde{Q}_{AB|XY}^{e^*} = Q_{AB|XY}$. Then to distinguish $\mathcal{E}$ and $\mathcal{F}$ using $\tilde{Q}_{ABE|XY}$ Eve first checks $E$. If $E = e^*$, she distinguishes $\mathcal{E}$ and $\mathcal{F}$ as in Lemma 19; otherwise, she just guesses randomly. Because the probability that $E = e^*$ is nonzero, we have $\left\|(\mathcal{E} - \mathcal{F}) \otimes \mathrm{id}(\tilde{Q}_{ABE|XY})\right\| > 0$. □

Several remarks are in order. Firstly, the diamond norm $||\mathcal{E}-\mathcal{F}||_{\diamond}^{\mathrm{quantum,CHSH}}$ between the channels $\mathcal{E}$ and $\mathcal{F}$ from Theorem 17 is exponentially small in $n$, because Eve only tries to distinguish $\mathcal{E}$ and $\mathcal{F}$ in the case when the depolarizing protocol does nothing. This means that while a theorem in the same form

as Theorem 15 cannot hold for collective attack boxes, it is entirely possible that there is, for example, a theorem that yields a bound of the form

$$\|\mathcal{E} - \mathcal{F}\|_{\diamond}^{\text{quantum,CHSH}} \leq f(n) \sup_{P_{ABE|XYZ}} \left\| (\mathcal{E} - \mathcal{F}) \otimes \text{id}(P_{ABE|XYZ}) \right\| + g(n),$$

(82)

where $g(n) \to 0$ as $n \to \infty$. Such a result could be sufficient to allow security proof reductions to collective attacks.

Secondly, the results of this section relied only on the fact that collective attack boxes are non-signaling between the rounds. This property arose entirely from the fact that Alice and Bob's measurements act on different Hilbert spaces in different rounds, and hence also hold more generally, i.e., even if the measurements in each round are different, or the states are entangled across rounds. This seems to suggest that in DIQKD, imposing an assumption that different rounds have different Hilbert spaces may already be a fairly strong restriction by itself,[3] even if we allow many other non-iid behaviors across states in different rounds, such as classical correlations or even entanglement. (In fact, security proof reductions to the iid case under this assumption were indeed previously studied in [40,41], though the latter was restricted to one-way protocols.) Whether this assumption seems reasonable may depend on the protocol—for instance, it seems unsatisfactory if each honest party has to use a single device for all inputs/outputs (as in [8], which used the EAT to avoid this assumption for one-way protocols), but if each honest party has access to $n$ devices that are "well isolated" from each other, it might be more plausible.

Thirdly, we remark that all sequential DIQKD protocols naturally fulfill a certain form of non-signaling constraints between the individual rounds of Alice and Bob: Alice and Bob's inputs in one round cannot influence the outputs in preceding rounds. Theorem 17 does not rule out that a result like Theorem 15 exists for channels with such a sequential structure. However, preserving such a sequential structure for the purposes of a security proof appears rather incompatible with permutation symmetry, so exploiting such sequential structures might require different techniques from those used in this paper.

## 5. Conclusion

In this paper, we proved two de Finetti theorems for quantum conditional probability distributions with CHSH symmetry. The advantage of these theorems over similar de Finetti theorems [16,17] is that the de Finetti boxes are in the quantum set. The first de Finetti theorem states that the entries of a CHSH symmetric box are upper bounded, up to a polynomial factor in $n$, by

---

[3]For comparison, in device-dependent QKD, this assumption is often implicitly imposed by default. This perhaps suggests a possible source of what appear to be greater challenges in non-iid DIQKD security proofs as compared to device-dependent QKD. It may also indicate that the default assumptions in device-dependent QKD could be stronger than they initially appear.

the entries of a fixed de Finetti box. This theorem is actually not restricted to boxes with CHSH symmetry but can be applied to arbitrary symmetries if a corresponding threshold theorem is available. The second de Finetti theorem states that the marginal of the first $k$ rounds of an $n$ round CHSH symmetric box is close to (and not just upper bounded by) a de Finetti box.

We further showed that the first de Finetti theorem can be used to obtain a bound on the diamond distance between two channels acting on boxes. Specifically, an attacker who tries to distinguish two channels $\mathcal{E}$ and $\mathcal{F}$ can be restricted to non-signaling extensions of a fixed quantum de Finetti box without decreasing the distinguishability between both channels by more than a polynomial factor. Because the security of DIQKD protocols is defined in terms of the distance between the channel given by the protocol and an ideal channel, this statement might be useful in security proofs. However, our theorem does not immediately allow to conclude security against coherent attacks from security against collective attacks: A straightforward strengthening of it to bound the diamond distance between two channels by the distinguishability using collective attack boxes does not hold. Based on some insights in our proof approach, we speculate that in DIQKD, assuming that boxes in different rounds act on different Hilbert spaces may already be a fairly strong constraint, even if we allow correlations or entanglement between states in different rounds.

## Acknowledgements

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## A.  Proof of Lemma 8

*Proof of Lemma 8.* The second inequality follows directly because $f$ is nonnegative and attains its maximum at $x^*$. The idea for the first inequality is to replace $f$ by a piecewise linear function that equals 0 at $a$ and $b$, and equals $f(x^*)$ at $x^*$. By concavity, this piecewise linear function is always smaller than $f$ itself. Writing this out explicitly: By concavity and non-negativity, we have for all $x \in [a, x^*]$

$$
\begin{aligned}
f(x) &= f\left(\frac{x-a}{x^*-a}x^* + \frac{x^*-x}{x^*-a}a\right) \\
&\geq \frac{x-a}{x^*-a}f(x^*) + \frac{x^*-x}{x^*-a}f(a) \\
&\geq \frac{x-a}{x^*-a}f(x^*).
\end{aligned}
\tag{83}
$$

Therefore, we have

$$
\int_a^{x^*} f(x)^n \mathrm{d}x \geq f(x^*)^n \int_a^{x^*} \left(\frac{x-a}{x^*-a}\right)^n \mathrm{d}x = \frac{1}{n+1}(x^*-a)f(x^*)^n. \tag{84}
$$

Analogously it follows

$$
\int_{x^*}^b f(x)^n \mathrm{d}x \geq \frac{1}{n+1}(b-x^*)f(x^*)^n, \tag{85}
$$

so together

$$
\int_a^b f(x)^n \mathrm{d}x = \int_a^{x^*} f(x)^n \mathrm{d}x + \int_{x^*}^b f(x)^n \mathrm{d}x \geq \frac{1}{n+1}(b-a)f(x^*)^n. \tag{86}
$$

$\square$

## B.  A de Finetti Theorem for General Symmetries

In this section, we will show that for arbitrary games a threshold theorem such as Theorem 7 can always be used to prove a de Finetti theorem similar to Theorem 6. Conversely, we will also see that a de Finetti theorem implies a threshold theorem. This means that proving a de Finetti theorem for some symmetry is just as hard as proving a threshold theorem for the game associated with that symmetry.

### B.1. Statement of the Main Theorem

Throughout this section, we will only consider boxes with a single interface and with a single-round input set $\hat{\mathcal{X}}$, a single-round output set $\hat{\mathcal{A}}$, and corresponding $n$-round input and out sets $\mathcal{X} = \hat{\mathcal{X}}^n$ and $\mathcal{Y} = \hat{\mathcal{Y}}^n$. CHSH symmetric boxes can be described like this by treating the two parties Alice and Bob as one, so the input and output sets are $\hat{\mathcal{A}} = \hat{\mathcal{X}} = \{0,1\}^2$. We consider the following generalization of CHSH symmetry:

**Definition 20.** 　1. Let $d \in \mathbb{N}$ and let $w : \hat{\mathcal{A}} \times \hat{\mathcal{X}} \to \{1, \ldots, d\}$ be some function. We will call $w$ the predicate function of the symmetry. For $a \in \mathcal{A}$ and $x \in \mathcal{X}$, we define $\mathrm{freq}^w(a, x) = (\frac{k_1}{n}, \ldots, \frac{k_d}{n}) \in \Delta^d$ with $k_r = |\{i | w(a_i, x_i) = r\}|$. Here $\Delta^d$ denotes the $d$-dimensional simplex.
2. We say an $n$-round box $P_{A|X}$ has $w$-symmetry if $P(a|x) = P(a'|x')$ whenever $\mathrm{freq}^w(a, x) = \mathrm{freq}^w(a', x')$, for all $a, a' \in \mathcal{A}^n$ and $x, x' \in \mathcal{X}^n$.

CHSH symmetry is an example of $w$-symmetry with $w((a,b),(x,y)) = 1$ if $a \oplus b = xy$ and $w((a,b),(x,y)) = 2$ if $a \oplus b \neq xy$. The definition of $w$-symmetry is an extension of the symmetries considered in [16] for permutation-invariant boxes, where only certain predicate functions $w$ where considered, namely those where for each pair $x$, $x'$ either the images of $w(\cdot, x)$ and $w(\cdot, x')$ are disjoint or $w(\cdot, x)$ and $w(\cdot, x')$ are identical up to a permutation of the elements of $\hat{\mathcal{A}}$.

Instead of the set of quantum single-round CHSH boxes, we will in this section consider a general convex set $\mathcal{Q}$ of single-round boxes $Q_{\hat{A}|\hat{X}}$. If we view $\mathcal{Q}$ as a convex subset in $\mathbb{R}^{|\hat{\mathcal{A}}||\hat{\mathcal{X}}|}$ we can consider its affine hull: the smallest affine superset of $\mathcal{Q}$. Throughout this section, we will denote the dimension of the affine hall by $d'$. For the CHSH symmetric case, $\mathcal{Q}$ is the set of CHSH symmetric quantum boxes, and $d' = 1$.

In the CHSH symmetric case, it was crucial that the expected number of wins of $n$-rounds of the CHSH games is between $\frac{2-\sqrt{2}}{4}$ and $\frac{2+\sqrt{2}}{4}$ when the games are played with iid quantum boxes. In our generalization, the interval $[\frac{2-\sqrt{2}}{4}, \frac{2+\sqrt{2}}{4}]$ will be replaced by a set of expected frequencies $\mathcal{F}_\mu$:

**Definition 21.** Let $\mu$ be a probability distribution on $\mathcal{X}$, $\mathcal{Q}$ a convex set of single-round boxes and $w$ a predicate function. The set of expected frequencies is

$$\mathcal{F}_\mu = \left\{ \left( \sum_{\substack{a,x \\ w(a,x)=r}} Q(a|x)\mu(x) \right)_{r=1\ldots d} \middle| Q_{\hat{A}|\hat{X}} \in \mathcal{Q} \right\} \subseteq \Delta^d. \qquad (87)$$

In the CHSH symmetric case, it is $\mathcal{F}_\mu = \{(p, 1-p) | p \in [(2 - \sqrt{2})/4, (2 + \sqrt{2})/4]\}$ regardless of $\mu$. We can now state the main theorem of this section:

**Theorem 22.** *Let* $w : \hat{\mathcal{A}} \times \hat{\mathcal{X}} \to [d]$ *be a predicate function and* $\mathcal{Q}$ *a convex subset of single-round boxes with an affine hull of dimension* $d'$. *Let* $\mathcal{F}_\mu$

be the set of expected frequencies. There exists a de Finetti state $\tau_{A|X} \in$ conv $\left( \left\{ Q_{A|X}^{\otimes n} | Q_{\hat{A}|\hat{X}} \in \mathcal{Q} \right\} \right)$ independent of $w$ such that the following hold:

1. Let $P_{A|X}$ be an $n$-round box, let $\mu$ be a probability measure on $\mathcal{X}$, and take any $f \in \Delta^d$. Suppose there exists some $C > 0$ such that

$$\Pr_{P_{A|X},\mu^{\otimes n}} \left[ \text{freq}^w(A, X) = f \right] \leq C \exp \left( - \inf_{f' \in \mathcal{F}_\mu} D(f||f')n \right). \qquad (88)$$

   Then

$$\Pr_{P_{A|X},\mu^{\otimes n}} \left[ \text{freq}^w(A, X) = f \right]$$
$$\leq C \binom{n + d'}{d'} (n+1)^{d'-1} \Pr_{\tau_{A|X},\mu^{\otimes n}} \left[ \text{freq}^w(A, X) = f \right]. \qquad (89)$$

2. Let $P_{A|X}$ be an $n$-round box with $w$-symmetry and let each box $Q_{\hat{A}|\hat{X}} \in \mathcal{Q}$ have $w$-symmetry. Let $C$ be such that for all $f \in \Delta^d$ there is a $\mu > 0$ such that Eq. (88) holds. Then

$$P(a|x) \leq C \binom{n + d'}{d'} (n+1)^{d'-1} \tau(a|x) \qquad \forall a \in \mathcal{A}^n, x \in \mathcal{X}^n. \qquad (90)$$

3. Let $P_{A|X}$ be an $n$-round box for which Eq. (90) holds. Then

$$\Pr_{P_{A|X},\mu^{\otimes n}} \left[ \text{freq}^w(A, X) = f \right] \leq C \binom{n + d'}{d'} (n+1)^{d'-1} \exp \left( - \inf_{f' \in \mathcal{F}_\mu} D(f||f')n \right). \qquad (91)$$

Qualitatively, we can interpret the equations and statements in Theorem 22 as follows:

- The condition (88) is a perfect threshold theorem, written in a form similar to Eq. (25) (which was for the CHSH case). It states that the probability to obtain a frequency distribution $f$ outside of the set $\mathcal{F}_\mu$ of expected frequencies decays exponentially with $n$ and with the distance from $f$ to $\mathcal{F}_\mu$, as measured by the relative entropy.
- Part 1 of Theorem 22 states that if we have such a threshold theorem, then the probability of obtaining the frequencies $f$ using the box $P_{A|X}$ can be bounded by the probability of obtaining $f$ using the de Finetti box $\tau_{A|X}$, up to a polynomial factor. We have chosen to state this part of the theorem separately because it does not require $P_{A|X}$ to be $w$-symmetric.
- Part 2 asserts that if we have the further condition that $P_{A|X}$ and all boxes in $\mathcal{Q}$ are $w$-symmetric, then we can get a de Finetti theorem analogous to Theorem 6 in the CHSH case. We will derive part 2 from part 1 by expressing the entries of $P_{A|X}$ and $\tau_{A|X}$ in terms of their respective probabilities of obtaining $\text{freq}^w(a, x) = f$, which is possible by $w$-symmetry.
- Finally part 3 shows the other direction of the equivalence between a threshold theorem and a de Finetti theorem: A box $P_{A|X}$ satisfying the de Finetti theorem statement (90) also satisfies a threshold theorem, albeit with a larger prefactor than in Eq. (88).

To prove Theorem 22, we first show some preparatory lemmas in Sects. B.2 and B.3, then combine them in Sect. B.4. Some insight can be gained into the proof structure by writing a slightly different proof of Theorem 6 (the CHSH case), in order to draw analogies to parts 1 and 2 of Theorem 22 separately. This version of the proof proceeds as follows: First prove Eq. (21) as before, giving a lower bound on $\tau(a|x)$. However, we reorder the proof after that point. Namely, observe that combining (25), (26) and (28) gives

$$p_k \leq (n+1)2^n \binom{n}{k} \sup_{p \in [1-w,w]} f(p)^n. \tag{92}$$

Putting together (21) and (92) gives

$$p_k \leq (n+1)^2 2^n \binom{n}{k} \tau(ab|xy). \tag{93}$$

The symmetry condition has not been used up to this point. We now use it to relate $p_k$ to $P(ab|xy)$ via (11), which yields the desired inequality $P(ab|xy) \leq (n+1)^2 \tau(ab|xy)$.

The proof in the subsequent sections basically follows the same structure as the above version. First, Eq. (21) is generalized to Lemma 26 in Sect. B.2. Next, Eq. (92) is replaced by Lemma 28 in Sect. B.3, bounding the probability of obtaining some outcome frequency in terms of a supremum over iid boxes. (The $2^n \binom{n}{k}$ factor in (92) counts different ways to achieve the specified frequency.) These lemmas are combined to obtain part 1 of Theorem 22, which is the generalization of (93). Finally, the symmetries are invoked to relate the box distribution to the probabilities of obtaining some outcome frequencies, analogous to (11), to yield part 2 of Theorem 22.

### B.2. Construction and Properties of the de Finetti Box

In this section, we will construct the de Finetti box $\tau_{A|X}$ and show that this $\tau(a|x)$ is at most polynomially smaller then $Q^{\otimes n}(a|x)$, for all $Q \in \mathcal{Q}$. Before that, we need to prove some preparatory lemmas:

The following lemma and proof are adopted from [42].

**Lemma 23.** (Matrix Determinant Lemma) *Let $A \in \mathbb{R}^{n \times n}$ be an invertible matrix and $v \in \mathbb{R}^n$. Then*

$$\det(A - vv^T) = \det(A)(1 - v^T A^{-1} v) \tag{94}$$

*Proof.* We calculate $\det(A - vv^T)$ as the determinant of a block matrix:

$$\begin{aligned}
\det(A - vv^T) &= \det \begin{pmatrix} A & v \\ v^T & 1 \end{pmatrix} \\
&= \det \left[ \begin{pmatrix} A & 0 \\ v^T & 1 \end{pmatrix} \begin{pmatrix} I & A^{-1}v \\ 0 & I - v^T A^{-1} v \end{pmatrix} \right] \\
&= \det(A)(1 - v^T A^{-1} v)
\end{aligned} \tag{95}$$

$\square$

**Lemma 24.** *Let $\alpha_1, \ldots \alpha_n \geq 0$ with $\sum_i \alpha_i \leq 1$. The map $f : \left(\mathbb{R}_0^+\right)^n \to \mathbb{R}_0^+$ given by*

$$f(x_1, .., x_n) = \prod_{i=1}^n x_i^{\alpha_i} \tag{96}$$

*is concave.*

*Proof.* First assume that all $\alpha_i > 0$ and $\sum_i \alpha_i < 1$. To show that $f$ is concave, we compute the Hesse matrix:

$$(Hf)_{ij} = \frac{\partial^2}{\partial x_i \partial x_j} f = \frac{\partial}{\partial x_j}\left(\frac{\alpha_i}{x_i} f\right) = \left(-\frac{\alpha_i}{x_i^2}\delta_{ij} + \frac{\alpha_i \alpha_j}{x_i x_j}\right) f =: -A_{ij} f \tag{97}$$

with $A = \left(\frac{\alpha_i}{x_i^2}\delta_{ij} - \frac{\alpha_i \alpha_j}{x_i x_j}\right)_{1 \leq i,j \leq 1}$. To show that $f$ is concave it is sufficient to show that $A$ is positive definite. Let $A^{(k)}$ be the upper left $k \times k$ block of $A$. By Sylvester's criterion, $A$ is positive definite if $\det(A^{(k)}) > 0$ for all $k \in \{1, \ldots, n\}$.

Let $B^{(k)} = \left(\frac{\alpha_i}{x_i^2}\delta_{ij}\right)_{1 \leq i,j \leq k}$ and $v^{(k)} = \left(\frac{\alpha_i}{x_i}\right)_{1 \leq i \leq k}$. Then

$$A^{(k)} = B^{(k)} - v^{(k)}(v^{(k)})^T. \tag{98}$$

By Lemma 23, we calculate

$$\begin{aligned}
\det(A^{(k)}) &= \det(B^{(k)})\left(1 - (v^{(k)})^T(B^{(k)})^{-1}v^{(k)}\right) \\
&= \det(B^{(k)})\left(1 - \sum_{i=1}^k \left(\frac{\alpha_i}{x_i}\right)^2 \frac{x_i^2}{\alpha_i}\right) \\
&= \det(B^{(k)})\left(1 - \sum_{i=1}^k \alpha_i\right) > 0
\end{aligned} \tag{99}$$

where the last inequality follows from $\det(B^{(k)}) > 0$ and $\sum_{i=1}^k \alpha_i \leq \sum_{i=1}^n \alpha_i < 1$. Hence, $f$ is concave.

Now assume the general setting where also $\alpha_i = 0$ and $\sum_i \alpha_i = 1$ is allowed. For each $i$, choose a sequence $(\alpha_i^{(m)})_{m \in \mathbb{N}}$ such that $\alpha_i^{(m)} > 0$, $\sum_i \alpha_i^{(m)} < 1$ and $\alpha_i^{(m)} \xrightarrow{m \to \infty} \alpha_i$. Let

$$f^{(m)}(x_1 \ldots x_n) = \prod_{i=1}^n x_i^{\alpha_i^{(m)}}. \tag{100}$$

By continuity,

$$f^{(m)}(x_1 \ldots x_n) \xrightarrow{m \to \infty} f(x_1 \ldots x_n). \tag{101}$$

Now let $x = (x_1, \ldots x_n), y = (y_1 \ldots y_n) \in \left(\mathbb{R}_0^+\right)^n$ and $\lambda \in [0, 1]$. Then

$$\begin{aligned}
f(\lambda x + (1 - \lambda)y) &= \lim_{m \to \infty} f^{(m)}(\lambda x + (1 - \lambda)y) \\
&\geq \lim_{m \to \infty} \lambda f^{(m)}(x) + (1 - \lambda)f^{(m)}(y) \\
&= \lambda f(x) + (1 - \lambda)f(y).
\end{aligned} \tag{102}$$

$\square$

The following lemma is the generalization of Lemma 8.

**Lemma 25.** *Let $C \subseteq \mathbb{R}^d$ be a bounded convex set, and denote by $\mathrm{vol}(C)$ the volume of $C$ (under the Lebesgue measure). Then for any $n \in \mathbb{N}$ and any concave function $f : C \to \mathbb{R}_0^+$, we have*

$$\int_C f(x)^n \mathrm{d}x \geq \mathrm{vol}(C) \binom{n+d}{d}^{-1} \left( \sup_{x \in C} f(x) \right)^n. \tag{103}$$

The proof idea is similar to that of Lemma 8 (assuming for simplicity that $f$ attains its supremum at some point $x^* \in C$): We will lower bound $f$ by a function that is zero on the boundary of $C$, takes the value $f(x^*)$ at $x^*$, and is determined on the rest of $C$ by "interpolating linearly" between the values at $x^*$ and the boundary of $C$. (Geometrically, the graph of this new function is basically the surface of a convex cone.)

*Proof.* Take any $\epsilon > 0$. There exists some $x^* \in C$ such that $f(x^*) \geq \sup_{x \in C} f(x) - \epsilon$. We evaluate the integrals using spherical coordinates centered on this point: Let $S^{d-1} \subseteq \mathbb{R}^d$ be the $(d-1)$-sphere, and let $\mu$ be the surface measure on $S^{d-1}$ with respect to the usual Lebesgue measure on $\mathbb{R}^d$. Since integrals are unchanged by including points on the boundary, we can evaluate the integrals using the closure of $C$ instead, denoted as $\mathrm{cl}(C)$. This is a convex compact set; hence, there is a "radius function" $R : S^{d-1} \to \mathbb{R}_0^+$ such that

$$\mathrm{cl}(C) = \{x^* + r\Omega | \Omega \in S^{d-1}, r \in [0, R(\Omega)]\}. \tag{104}$$

Then

$$\mathrm{vol}(C) = \int_{S^{d-1}} \int_0^{R(\Omega)} r^{d-1} \mathrm{d}r \mu(\mathrm{d}\Omega) = \frac{1}{d} \int_{S^{d-1}} R(\Omega)^d \mu(\mathrm{d}\Omega). \tag{105}$$

Since

$$\begin{aligned} f(x^* + r\Omega) &= f\left( \frac{r}{R(\Omega)} (x^* + R(\Omega)\Omega) + \frac{R(\Omega) - r}{R(\Omega)} x^* \right) \\ &\geq \frac{r}{R(\Omega)} f(x^* + R(\Omega)\Omega) + \frac{R(\Omega) - r}{R(\Omega)} f(x^*) \qquad \text{Concavity of } f \\ &\geq \frac{R(\Omega) - r}{R(\Omega)} f(x^*) \qquad\qquad\qquad\qquad \text{Positivity of } f \end{aligned} \tag{106}$$

we have

$$\begin{aligned} \int_C f(x)^n \mathrm{d}x &= \int_{S^{d-1}} \int_0^{R(\Omega)} r^{d-1} f(x^* + r\Omega) \mathrm{d}r \mu(\mathrm{d}\Omega) \\ &\geq f(x^*)^n \int_{S^{d-1}} \int_0^{R(\Omega)} r^{d-1} \left( \frac{R(\Omega) - r}{R(\Omega)} \right)^n \mathrm{d}r \mu(\mathrm{d}\Omega) \quad \text{by Eq. (106)} \\ &= f(x^*)^n \int_{S^{d-1}} R(\Omega)^d \int_0^1 u^{d-1} (1 - u)^n \mathrm{d}u \mu(\mathrm{d}\Omega). \end{aligned} \tag{107}$$

Using the property of the beta function

$$\int_0^1 u^{d-1}(1-u)^n = \frac{1}{n+d}\binom{n+d-1}{d-1}^{-1} = \frac{1}{d}\binom{n+d}{d}^{-1} \tag{108}$$

we find by Eqs. (107) and (105)

$$\int_C f(x)^n \mathrm{d}x \geq \frac{1}{d}\binom{n+d}{d}^{-1} f(x^*)^n \int_{S^{d-1}} R(\Omega)^d \mu(\mathrm{d}\Omega)$$

$$= \binom{n+d}{d}^{-1} \mathrm{vol}(C) f(x^*)^n. \tag{109}$$

Recalling that $f(x^*) \geq \sup_{x \in C} f(x) - \epsilon$, and $\epsilon > 0$ was arbitrary, this implies the desired result. $\qquad\square$

When we maximize $Q^{\otimes n}(a|x)$ over $Q \in \mathcal{Q}$ while keeping $a$ and $x$ fixed, the maximum is achieved for a different $Q$ for each $a$ and $x$. However, the next lemma states that if we average over $Q$ we are at most a polynomial factor below that maximal value of $Q^{\otimes n}(a|x)$, no matter what $a$ and $x$ are.

**Lemma 26.** *There is a de Finetti state* $\tau_{A|X} \in \mathrm{conv}\left(\{Q^{\otimes n}|Q \in \mathcal{Q}\}\right)$ *such that for all* $a \in \mathcal{A}$, $x \in \mathcal{X}$

$$\tau(a|x) \geq \binom{n+d'}{d'}^{-1} \sup_{Q \in \mathcal{Q}} Q^{\otimes n}(a|x), \tag{110}$$

*where* $d'$ *is the dimension of the affine hull of* $\mathcal{Q}$.

*Proof.* We view $\mathcal{Q}$ as a bounded convex subset of $\mathbb{R}^{|\hat{\mathcal{A}}||\hat{\mathcal{X}}|}$. Because the affine hull of $\mathcal{Q}$ has dimension $d'$, there exists a bounded convex set $C \subseteq \mathbb{R}^{d'}$ and a bijective linear map $C \ni \phi \mapsto Q_\phi \in \mathcal{Q}$. Choose

$$\tau_{A|X} = \frac{1}{\mathrm{vol}(C)}\int_C Q_\phi^{\otimes n}\mathrm{d}\phi. \tag{111}$$

Now fix $a \in \mathcal{A}$ and $x \in \mathcal{X}$, and for $a' \in \hat{\mathcal{A}}$ and $x' \in \hat{\mathcal{X}}$, let $f_{a'x'} = |\{i|a_i = a' \text{ and } x_i = x'\}|/n$ be the frequency of the pair $(a', x')$ in $(a, x)$. Then

$$\phi \mapsto \left(Q_\phi^{\otimes n}(a|x)\right)^{1/n} = \prod_{a' \in \mathcal{A}, x' \in \mathcal{X}} Q_\phi(a'|x')^{f_{a'x'}} \tag{112}$$

is a concave map by Lemma 24 and by the linearity of $\phi \mapsto Q_\phi$. Hence, by Lemma 25

$$\tau(a|x) = \frac{1}{\mathrm{vol}(C)}\int_C Q_\phi^{\otimes n}(a|x)\mathrm{d}\phi$$

$$\geq \binom{n+d'}{d'}^{-1} \sup_{\phi \in C} Q_\phi^{\otimes n}(a|x)$$

$$= \binom{n+d'}{d'}^{-1} \sup_{Q \in \mathcal{Q}} Q^{\otimes n}(a|x). \tag{113}$$

$\square$

### B.3. Bounding $P_{A|X}$ by iid Boxes

In this section, we will show that a threshold theorem of the form of Eq. (88) implies that $P(a|x)$ can be bound by $Q^{\otimes n}(a|x)$ up to a polynomial factor. Together with Lemma 26, this will yield the proof for Theorem 22 in the next section.

**Lemma 27.** *Let* $n = k_1 + k_2 + \ldots + k_d$ *with* $k_1, .., k_d \in \mathbb{N}$. *Then*

$$(n+1)^{-(d-1)} \prod_{r=1}^{d} \left(\frac{n}{k_r}\right)^{k_r} \leq \binom{n}{k_1, k_2, \ldots, k_d} \leq \prod_{r=1}^{d} \left(\frac{n}{k_r}\right)^{k_r}. \tag{114}$$

Here $\binom{n}{k_1, k_2, \ldots, k_d} = n!/(k_1! \ldots k_d!)$ is the multinomial coefficient.

*Proof.* Let us first prove the inequality for $d = 2$, when the multinomial coefficient is just a binomial coefficient. It is a property of the Beta function that

$$\int_0^1 x^{k_1} (1-x)^{k_2} = \frac{1}{n+1} \binom{n}{k_1}^{-1} \tag{115}$$

The integrand on the left-hand side is maximized at $x = \frac{k_1}{n}$. Hence,

$$\frac{1}{n+1} \left(\frac{k_1}{n}\right)^{k_1} \left(\frac{k_2}{n}\right)^{k_2} \leq \int_0^1 x^{k_1} (1-x)^{k_2} \leq \left(\frac{k_1}{n}\right)^{k_1} \left(\frac{k_2}{n}\right)^{k_2}, \tag{116}$$

by Lemma 8. Combining Eqs. (115) and (116) gives

$$\frac{1}{n+1} \left(\frac{n}{k_1}\right)^{k_1} \left(\frac{n}{k_2}\right)^{k_2} \leq \binom{n}{k_1} \leq \left(\frac{n}{k_1}\right)^{k_1} \left(\frac{n}{k_2}\right)^{k_2}. \tag{117}$$

Now we prove the lemma for general $d \geq 2$. For this, observe that

$$\binom{n}{k_1, \ldots, k_d} = \binom{n}{k_1}\binom{n-k_1}{k_2}\binom{n-k_1-k_2}{k_3} \ldots \binom{n-k_1-\ldots-k_{d-2}}{k_{d-1}}$$

$$= \prod_{r=1}^{d-1} \binom{n - \sum_{i=1}^{r-1} k_i}{k_r}. \tag{118}$$

Applying Eq. (117) to each binomial coefficient completes the proof since

$$\prod_{r=1}^{d-1} \left(\frac{n - \sum_{i=1}^{r-1} k_i}{k_r}\right)^{k_r} \left(\frac{n - \sum_{i=1}^{r-1} k_i}{n - \sum_{i=1}^{r} k_i}\right)^{n - \sum_{i=1}^{r} k_i}$$

$$= \frac{1}{\prod_{r=1}^{d-1} k_r^{k_r}} \frac{\prod_{r=1}^{d-1} \left(n - \sum_{i=1}^{r-1} k_i\right)^{n - \sum_{i=1}^{r-1} k_i}}{\prod_{r=1}^{d-1} \left(n - \sum_{i=1}^{r} k_i\right)^{n - \sum_{i=1}^{r} k_i}} \tag{119}$$

$$= \prod_{r=1}^{d} \left(\frac{n}{k_r}\right)^{k_r} \tag{120}$$

by a telescoping product argument. $\square$

**Lemma 28.** *Let $P_{A|X}$ be a n-round box. If*

$$\Pr_{P_{A|X},\mu^{\otimes n}}\left[\mathrm{freq}^w(A,X)=f\right] \le C\exp\left(-\inf_{f'\in\mathcal{F}_\mu}D(f||f')n\right) \qquad (121)$$

*then*

$$\Pr_{P_{A|X},\mu^{\otimes n}}\left[\mathrm{freq}^w(A,X)=f\right] \le C(n+1)^{d-1}\sup_{Q\in\mathcal{Q}}\Pr_{Q^{\otimes n},\mu^{\otimes n}}\left[\mathrm{freq}^w(A,X)=f\right].$$

$$(122)$$

*Proof.* This is actually just a statement on the two right-hand sides. Let $f = (\frac{k_1}{n},\dots\frac{k_d}{n})\in\Delta^d$, let $Q\in\mathcal{Q}$ and let

$$f' = \left(\sum_{\substack{a,x\\ w(a,x)=r}}Q(a|x)\mu(x)\right)_{r=1\dots d}\in\Delta^d$$

$$(123)$$

be the element of $\mathcal{F}_\mu$ belonging to $Q$. Then

$$\Pr_{Q^{\otimes n},\mu^{\otimes n}}\left[\mathrm{freq}^w(A,X)=f\right] = \binom{n}{k_1,\dots,k_d}\prod_{r=1}^d(f'_r)^{k_r}$$

$$\ge \frac{1}{(n+1)^{d-1}}\prod_{r=1}^d\left(\frac{nf'_r}{k_r}\right)^{k_r} \qquad \text{by Lemma 27}$$

$$= \frac{1}{(n+1)^{d-1}}\exp(-D(f||f')n) \qquad (124)$$

where the last equality follows directly from the definition of the relative entropy

$$D(f||f') = \sum_{r=1}^d f_r(\ln(f_r)-\ln(f'_r)). \qquad (125)$$

Taking the supremum over $Q$ gives

$$\sup_{Q\in\mathcal{Q}}\Pr_{Q^{\otimes n},\mu^{\otimes n}}\left[\mathrm{freq}^w(A,X)=f\right] \ge \frac{1}{(n+1)^{d-1}}\exp\left(-\inf_{f'\in\mathcal{F}_\mu}D(f||f')n\right)$$

$$(126)$$

which completes the proof.  □

### B.4. Proof of Theorem 22

Now we are ready to prove the general Theorem 22. For this, we use Lemma 26 to show that the entries of the de Finetti box are at most smaller by a polynomial factor then the corresponding entries of any iid box. Then we use Lemma 28 to show that the threshold theorem implies that the probability of a frequency $f$ under $P_{A|X}$ can be bounded, up to a polynomial factor, by probability of $f$ under some iid box (but possibly a different iid box for each $f$). Combining both lemmas yields the proof of part 1. Part 2 will follow from

part 1 by using the definition of $w$-symmetry, and part 3 will follow directly from Lemma 27.

*Proof of Theorem 22.*    1. This part follows directly by combining Lemmas 28 and 26. Suppose

$$\Pr_{P_{A|X},\mu^{\otimes n}}\left[\text{freq}^w(A,X)=f\right] \leq C\exp\left(-\inf_{f'\in\mathcal{F}_\mu} D(f||f')n\right). \tag{127}$$

By Lemma 28, we have

$$\Pr_{P_{A|X},\mu^{\otimes n}}\left[\text{freq}^w(A,X)=f\right]$$
$$\leq C(n+1)^{d-1}\sup_{Q\in\mathcal{Q}}\Pr_{Q^{\otimes n},\mu^{\otimes n}}\left[\text{freq}^w(A,X)=f\right]. \tag{128}$$

Since by Lemma 26

$$\Pr_{Q^{\otimes n},\mu^{\otimes n}}\left[\text{freq}^w(A,X)=f\right] \leq \binom{n+d'}{d'}\Pr_{\tau,\mu^{\otimes n}}\left[\text{freq}^w(A,X)=f\right] \tag{129}$$

it follows

$$\Pr_{P_{A|X},\mu^{\otimes n}}\left[\text{freq}^w(A,X)=f\right]$$
$$\leq C\binom{n+d'}{d'}(n+1)^{d-1}\Pr_{\tau,\mu^{\otimes n}}\left[\text{freq}^w(A,X)=f\right]. \tag{130}$$

2. For this part, we have by hypothesis that $P_{A|X}$ and every box in $\mathcal{Q}$ has $w$-symmetry. Then also $\tau_{A|X}$ has $w$-symmetry. Take any $a\in\mathcal{A}, x\in\mathcal{X}$, and define $f=\text{freq}^w(a|x)$. Then we have

$$\Pr_{P_{A|X},\mu^{\otimes n}}\left[\text{freq}^w(A,X)=f\right]=P(a|x)\sum_{\substack{a\in\mathcal{A}^n,x\in\mathcal{X}^n \\ \text{freq}^w(a,x)=f}}\mu^{\otimes n}(x), \tag{131}$$

and similarly

$$\Pr_{\tau_{A|X},\mu^{\otimes n}}\left[\text{freq}^w(A,X)=f\right]=\tau(a|x)\sum_{\substack{a\in\mathcal{A}^n,x\in\mathcal{X}^n \\ \text{freq}^w(a,x)=f}}\mu^{\otimes n}(x). \tag{132}$$

From Eqs. (131), (132) and part 1, it follows that

$$P(a|x)\leq C\binom{n+d'}{d'}(n+1)^{d-1}\tau(a|x). \tag{133}$$

3. Now assume

$$P(a|x)\leq\tilde{C}\tau(a|x) \tag{134}$$

with $\tilde{C} = C\binom{n+d'}{d'}(n+1)^{d-1}$. Since $\tau(a|x) \leq \sup_{Q \in \mathcal{Q}} Q^{\otimes n}(a|x)$ it follows for $f = \left(\frac{k_1}{n} \ldots \frac{k_d}{n}\right)$ that

$$\Pr_{P_{A|X},\mu^{\otimes n}}\left[\mathrm{freq}^w(A,X) = f\right] \leq \tilde{C} \sup_{Q \in \mathcal{Q}} \Pr_{Q^{\otimes n},\mu^{\otimes n}}\left[\mathrm{freq}^w(A,X) = f\right]$$

$$= \tilde{C} \sup_{f' \in \mathcal{F}_\mu} \binom{n}{k_1, \ldots, k_d} \prod_{r=1}^{d} (f'_r)^{k_r}$$

$$\leq \tilde{C} \sup_{f' \in \mathcal{F}_\mu} \prod_{r=1}^{d} \left(\frac{nf'_r}{k_r}\right)^{k_r} \qquad \text{by Lemma } 27$$

$$= \tilde{C} \exp\left(-\inf_{f' \in \mathcal{F}_\mu} D(f||f')n\right). \qquad (135)$$

$\square$

# References

[1] Bennett, C.H., Brassard, G.: Quantum cryptography: public key distribution and coin tossing. Theor. Comput. Sci. **560**, 7–11 (2014). https://doi.org/10.1016/j.tcs.2014.05.025

[2] Ekert, A.K.: Quantum cryptography based on Bell's theorem. Phys. Rev. Lett. **67**(6), 661–663 (1991). https://doi.org/10.1103/physrevlett.67.661

[3] Renner, R.: Security of quantum key distribution. Int. J. Quant. Inf. **06**(01), 1–127 (2008). https://doi.org/10.1142/s0219749908003256

[4] Barrett, J., Hardy, L., Kent, A.: No signaling and quantum key distribution. Phys. Rev. Lett. **95**(1), 010503 (2005). https://doi.org/10.1103/physrevlett.95.010503

[5] Acín, A., Brunner, N., Gisin, N., Massar, S., Pironio, S., Scarani, V.: Device-independent security of quantum cryptography against collective attacks. Phys. Rev. Lett. **98**(23), 230501 (2007). https://doi.org/10.1103/physrevlett.98.230501

[6] Pironio, S., Acín, A., Brunner, N., Gisin, N., Massar, S., Scarani, V.: Device-independent quantum key distribution secure against collective attacks. New J. Phys. **11**(4), 045021 (2009). https://doi.org/10.1088/1367-2630/11/4/045021

[7] Vazirani, U., Vidick, T.: Fully device-independent quantum key distribution. Phys. Rev. Lett. **113**(14), 133 (2014). https://doi.org/10.1103/physrevlett.113.140501

[8] Arnon-Friedman, R., Dupuis, F., Fawzi, O., Renner, R., Vidick, T.: Practical device-independent quantum cryptography via entropy accumulation. Nat. Commun. **9**(1), 459 (2018). https://doi.org/10.1038/s41467-017-02307-4

[9] Popescu, S., Rohrlich, D.: Quantum nonlocality as an axiom. Found. Phys. **24**(3), 379–385 (1994). https://doi.org/10.1007/bf02058098

[10] Tan, E.Y.Z., Lim, C.C.W., Renner, R.: Advantage distillation for device-independent quantum key distribution. Phys. Rev. Lett. **124**(2), 020502 (2020). https://doi.org/10.1103/physrevlett.124.020502

[11] Christandl, M., König, R., Renner, R.: Postselection technique for quantum channels with applications to quantum cryptography. Phys. Rev. Lett. **102**(2), 020504 (2009). https://doi.org/10.1103/physrevlett.102.020504

[12] Diaconis, P., Freedman, D.: Finite exchangeable sequences. Ann. Probab. **8**(4), 745–764 (1980). https://doi.org/10.1214/aop/1176994663

[13] Hudson, R.L., Moody, G.R.: Locally normal symmetric states and an analogue of de Finetti's theorem. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **33**(4), 343–351 (1976). https://doi.org/10.1007/bf00534784

[14] Caves, C.M., Fuchs, C.A., Schack, R.: Unknown quantum states: the quantum de Finetti representation. J. Math. Phys. **43**(9), 4537–4559 (2002). https://doi.org/10.1063/1.1494475

[15] Christandl, M., König, R., Mitchison, G., Renner, R.: One-and-a-half quantum de Finetti theorems. Commun. Math. Phys. **273**(2), 473–498 (2007). https://doi.org/10.1007/s00220-007-0189-3

[16] Arnon-Friedman, R., Renner, R.: de Finetti reductions for correlations. J. Math. Phys. **56**(5), 052203 (2015). https://doi.org/10.1063/1.4921341

[17] Christandl, M., Toner, B.: Finite de Finetti theorem for conditional probability distributions describing physical theories. J. Math. Phys. **50**(4), 042104 (2009). https://doi.org/10.1063/1.3114986

[18] Dupuis, F., Fawzi, O., Renner, R.: Entropy accumulation. Commun. Math. Phys. **379**(3), 867–913 (2020). https://doi.org/10.1007/s00220-020-03839-5

[19] Jain, R., Miller, C.A., Shi, Y.: Parallel device-independent quantum key distribution. IEEE Trans. Inf. Theory **66**(9), 5567–5584 (2020). https://doi.org/10.1109/tit.2020.2986740

[20] Vidick, T.: Parallel DIQKD from parallel repetition (2017). arXiv:1703.08508

[21] Maurer, U.M.: Secret key agreement by public discussion from common information. IEEE Trans. Inf. Theory **39**(3), 733–742 (1993). https://doi.org/10.1109/18.256484

[22] Wolf, S.: Information-theoretically and computationally secure key agreement in cryptography (1999). https://doi.org/10.3929/ethz-a-002077162

[23] Chau, H.F.: Practical scheme to share a secret key through a quantum channel with a 27.6% bit error rate. Phys. Rev. A **66**, 060302(R) (2002). https://doi.org/10.1103/PhysRevA.66.060302

[24] Gottesman, D., Lo, H.K.: Proof of security of quantum key distribution with two-way classical communications. IEEE Trans. Inf. Theory **49**(2), 457–475 (2003). https://doi.org/10.1109/TIT.2002.807289

[25] Brassard, G., Salvail, L.: Secret-key reconciliation by public discussion. In: Helleseth, T. (ed.) Advances in Cryptology—EUROCRYPT '93, pp. 410–423. Springer, Berlin (1994)

[26] Farkas, M., Balanzó-Juandó, M., Łukanowski, K., Kołodyński, J., Acín, A.: Bell nonlocality is not sufficient for the security of standard device-independent quantum key distribution protocols. Phys. Rev. Lett. **127**, 050503 (2021). https://doi.org/10.1103/PhysRevLett.127.050503

[27] Gisin, N., Wolf, S.: Linking Classical and quantum key agreement: is there "Bound Information"? In: Advances in Cryptology—CRYPTO 2000, pp. 482–500. Springer, Berlin (2000). https://doi.org/10.1007/3-540-44598-6

[28] Maurer, U., Renner, R.: Abstract cryptography. In: Innovations in Computer Science. Tsinghua University Press (2011)

[29] Portmann, C., Renner, R.: Security in quantum cryptography (2021). arXiv:2102.00021

[30] Scarani, V.: The device-independent outlook on quantum physics. Acta Physica Slovaca (2012). arXiv:1303.3081pdf

[31] Clauser, J.F., Horne, M.A., Shimony, A., Holt, R.A.: Proposed experiment to test local hidden-variable theories. Phys. Rev. Lett. **23**(15), 880–884 (1969). https://doi.org/10.1103/physrevlett.23.880

[32] Masanes, L.L., Acin, A., Gisin, N.: General properties of nonsignaling theories. Phys. Rev. A (2006). https://doi.org/10.1103/physreva.73.012112

[33] Cirel'son, B.S.: Quantum generalizations of Bell's inequality. Lett. Math. Phys. **4**(2), 93–100 (1980). https://doi.org/10.1007/bf00417500

[34] Unger, F.: A inequality, probabilistic, with applications to threshold direct-product theorems. In: 50th Annual IEEE Symposium on Foundations of Computer Science. IEEE (2009). https://doi.org/10.1109/focs.2009.62

[35] Arnon-Friedman, R., Renner, R., Vidick, T.: Non-signaling parallel repetition using de Finetti reductions. IEEE Trans. Inf. Theory **62**(3), 1440–1457 (2016). https://doi.org/10.1109/tit.2016.2516022

[36] Götze, F., Sambale, H., Sinulis, A.: Higher order concentration for functions of weakly dependent random variables. Electron. J. Probab. (2019). https://doi.org/10.1214/19-ejp338

[37] Hänggi, E., Renner, R., Wolf, S.: Quantum cryptography based solely on Bell's theorem. EUROCRYPT, pp. 216–234 (2010). arXiv:0911.4171

[38] Hänggi, E., Renner, R., Wolf, S.: The impossibility of non-signaling privacy amplification. Theoret. Comput. Sci. **486**, 27–42 (2013). https://doi.org/10.1016/j.tcs.2012.12.014

[39] Arnon-Friedman, R., Ta-Shma, A.: Limits of privacy amplification against nonsignaling memory attacks. Phys. Rev. A **86**, 062333 (2012). https://doi.org/10.1103/PhysRevA.86.062333

[40] Hänggi, E.: Device-independent quantum key distribution. Doctoral thesis, ETH Zurich, Zürich (2010)

[41] Masanes, L., Pironio, S., Acín, A.: Secure device-independent quantum key distribution with causally independent measurement devices. Nat. Commun. (2011). https://doi.org/10.1038/ncomms1244

[42] Sarwate, A.: The matrix determinant lemma (2012). https://ergodicity.net/2012/06/12/the-matrix-determinant-lemma/

Sven Jandura
Institut de Science et d'Ingénierie Supramoléculaires (UMR 7006)
University of Strasbourg
8 Allée Gaspard Monge
67000 Strasbourg
France
e-mail: sven.jandura@etu.unistra.fr

Sven Jandura and Ernest Y.-Z. Tan
Institute for Theoretical Physics
ETH Zurich
8093 Zurich
Switzerland

Ernest Y.-Z. Tan
Department of Physics and Astronomy, Institute for Quantum Computing
University of Waterloo
Waterloo ONN2L 3G1
Canada