

# Pauli channels can be estimated from syndrome measurements in quantum error correction

Thomas Wagner, Hermann Kampermann, Dagmar Bruß, and Martin Kliesch

Institut für Theoretische Physik, Heinrich-Heine-University Düsseldorf, Germany

The performance of quantum error correction can be significantly improved if detailed information about the noise is available, allowing to optimize both codes and decoders. It has been proposed to estimate error rates from the syndrome measurements done anyway during quantum error correction. While these measurements preserve the encoded quantum state, it is currently not clear how much information about the noise can be extracted in this way. So far, apart from the limit of vanishing error rates, rigorous results have only been established for some specific codes.

In this work, we rigorously resolve the question for arbitrary stabilizer codes. The main result is that a stabilizer code can be used to estimate Pauli channels with correlations across a number of qubits given by the pure distance. This result does not rely on the limit of vanishing error rates, and applies even if high weight errors occur frequently. Moreover, it also allows for measurement errors within the framework of quantum data-syndrome codes. Our proof combines Boolean Fourier analysis, combinatorics and elementary algebraic geometry. It is our hope that this work opens up interesting applications, such as the online adaptation of a decoder to time-varying noise.

## 1 Introduction

Quantum error correction is an essential part of most quantum computing schemes. It can be significantly improved if detailed knowledge about the noise affecting a device is available, as it is possible to optimize codes for specific noise [1, 2]. A prominent example is the XZZX-surface code [3]. Furthermore, common decoding algorithms, such as minimum weight matching [4–7] and belief propagation (see e.g. [8]), can incorporate information about error rates to return more accurate corrections. Other examples of decoders that can incorporate information about error rates are weighted union find [9] and tensor network decoders [10, 11]. The latter can also deal with correlated noise models, but are relatively slow. In the context of stabilizer codes, noise is usually modeled using Pauli channels, which are simple to understand and simulate. However, Pauli noise is more than a mere toy model, since randomized compiling can be used to project general noise onto a Pauli channel [12], which has also been demonstrated experimentally [13]. Furthermore, it is known that quantum error correction decoheres noise on the logical level [14]. Consequently, there has been much interest and progress in the estimation of Pauli channels [15–18]. Complementary to the standard benchmarking approaches, it has been suggested to perform (online) estimation of channels just from the syndromes of a quantum error correction code itself [2, 7, 19–24]. Such a scheme uses only measurements that do not destroy the logical information. It is thus suited for online adaptation of a decoder to varying noise, for example by adapting weights in a minimum

---

Thomas Wagner: [thomas.wagner@uni-duesseldorf.de](mailto:thomas.wagner@uni-duesseldorf.de)

Martin Kliesch: [science@mkliesch.eu](mailto:science@mkliesch.eu)

weight matching decoder [7, 21]. It furthermore results in a noise model that can be directly used by the decoder [20]. Experimentally, online optimization of control parameters in a 9-qubit superconducting quantum processor has been demonstrated in an experiment by Google [25].

Since the state of the logical qubit is not measured, some assumptions are necessary in order for this estimation to be feasible. However, the precise nature of these assumptions is currently not well understood. In the general case, it is unclear for which combinations of noise models and codes the parameters can be identified using only the syndrome information. Apart from heuristics in the limit of very low error rates [20–22], only two special cases have been rigorously treated. For codes admitting a minimum weight matching decoder, such as the toric code, identifiability of a circuit noise model was shown by Spitz *et al.* [7]. In [24], identifiability results for the restricted class of perfect codes were proven.

It is not a priori clear that an estimation of the error rates just from the syndrome statistics should be possible at all for arbitrary stabilizer codes. There are several objections one could raise. For one, as mentioned above, the state of the logical qubit is not measured, so there is only limited information contained in the syndromes. Phrased another way, there are generally exponentially many errors with the same syndrome, which we therefore cannot distinguish. This also implies that the probability of a syndrome is an exponentially large sum of different error rates, and solving such a system for the error rates appears difficult at best. While it has been suggested to simplify the problem by only taking into account the lowest weight error compatible with each syndrome [20–22], this approximation strategy leads to demonstrably sub-optimal estimators [24].

## 2 Results

In this work, we show that the estimation task can be solved for arbitrary stabilizer codes. For any given quantum code, we describe a general class of Pauli channels whose parameters can be identified from the corresponding syndrome statistics. Our results also take into account measurement errors by using the framework of quantum data-syndrome codes [26–28]. We prove that a large amount of information can be extracted from the syndrome statistics. In the following theorem we make this statement more precise in terms of the pure distance, which is defined as the minimum weight of an undetectable error (see Section 2.3 for details). The pure distance measures up to which weight errors can necessarily be distinguished by their syndrome, and for most codes it coincides with the weight of the smallest stabilizer. Since there can be undetectable but logically trivial errors, the pure distance is usually much smaller than the distance. As an example, for the family of toric codes the pure distance is 4 independent of code size. More generally, the pure distance will be constant for any family of quantum low density parity check (LDPC) codes, since the stabilizer weights are constant.

We are interested in estimating a Pauli channel, i.e. a probability distribution over Pauli errors, which describes the new error occurring before each round of error correction (working in a phenomenological noise model). It is commonly assumed that errors on each qubit occur independently, in which case the channel is described by one error rate for each qubit. In this case, we say the noise is uncorrelated. In a more general setting, the Pauli channel could act on many qubits, such that errors on these qubits are not independent. If e.g. 2 qubit errors occur that are not a combination of independent single qubit errors, we say that the noise is correlated over 2 qubits. The corresponding 2 qubit error rates are then not simply a product of the single qubit error rates and must be additionally specified. For example, if  $P(X_1X_2) \neq P(X_1)P(X_2)$ , the errors on the first two qubits are correlated. This notion of correlations is made precise in Section 2.2.3 for classical and Section 2.3.2 for quantum codes. We now give an informal statement of our main results. A formal statement is given in Theorem 7, which also takes into account many detectable errors beyond the pure distance.

**Theorem 1** (Main result, informal). *A stabilizer code with pure distance  $d_p$  can be used to estimate Pauli noise with correlations across up to  $\lfloor \frac{d_p-1}{2} \rfloor$  qubits.*

While this result is stated in a non-constructive fashion, its derivation suggests a concrete estimation protocol. We give a first heuristic discussion of the resulting estimators in Section 2.4. A detailed analysis of such estimators is ongoing work.

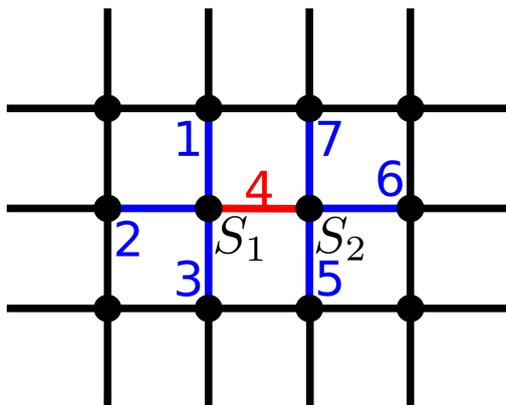


Figure 1: A qubit (edge) and two adjacent star operators (vertices) on the toric code.

The key idea behind Theorem 1 is that the error distribution is fully described by a set of moments. We will show that these moments can be estimated up to their sign by solving a polynomial system of equations. The appearance of multiple discrete solutions, differing in some signs, reflects the symmetries of the problem. For example in the case of single qubit noise, there is one symmetry for each logical operator of the code. However, under the additional mild assumption that all error rates are smaller than  $\frac{1}{2}$ , all moments must be positive, and thus the estimate is unique. We stress that our result does not rely on the limit of vanishing rates, and still applies in the presence of high weight errors. Perhaps surprisingly, even different errors with the same syndrome can occur frequently, as long as they arise as a combination of independent lower weight errors. Our result is arguably the strongest one can reasonably expect, since error rates cannot be identifiable if multiple independently occurring errors have the same syndrome.

The *adaptive weight estimator* by Spitz *et al.* [7] can be viewed as solving a special instance of the general equation system we present here, which is applicable only for codes that admit a minimum weight matching decoder. This connection is explained in Section 2.1.

Our arguments combine Boolean Fourier analysis, combinatorics and elementary algebraic geometry. Interestingly, a connection between Pauli channel learning and Boolean Fourier analysis was recently pointed out in an independent work by Flammia and O’Donnell [15].

We start our discussion with the motivating example of the toric code with independent Pauli- $X$  errors, where our results take a particularly simple form. We then discuss the general results. We first discuss the setting of classical codes, and later extend the results to quantum codes. This is for ease of exposition, since the presentation is considerably simplified in the classical setting and thus the underlying concepts become more apparent. We stress that, in contrast to the quantum setting, classically one can measure the individual bits of a codeword without destroying the encoded information. Thus, in the classical setting one does not have to rely solely on the syndrome information and can use easier techniques for error rates estimation. However, for a quantum code this is indeed the only information that can be measured without destroying the encoded state, and this is where our results are most relevant. Classically, our approach might still be useful in the setting of distributed source coding [29].

## 2.1 Example: the toric code

As a motivating example for the methods and proofs in the following sections, we derive an estimator for the simple setting of completely uncorrelated bit-flip noise on the toric code. That is, we assume that errors on different qubits occur independently, possibly with a different rate for each qubit, and that only bit-flip (Pauli- $X$ ) errors occur on each qubit. Thus, we consider the toric code essentially as a classical code. This constitutes a simple alternative derivation of the solution given by Spitz *et al.* [7].

We focus on a single qubit and two adjacent  $Z$ -stabilizers, as illustrated in fig. 1. Our task is to estimate the error rate  $p_4$  of the marked qubit, or equivalently, the expectation value  $E(Z_4) = 1 - 2p_4$ . Since errors on each qubit are independent, the expectation of a stabilizer measurement

is simply the product of expectations of the adjacent bits. Therefore we obtain the following three equations,

$$\begin{aligned} E(S_1) &= E(Z_1)E(Z_2)E(Z_3)E(Z_4) \\ E(S_2) &= E(Z_4)E(Z_5)E(Z_6)E(Z_7) \\ E(S_1S_2) &= E(Z_1)E(Z_2)E(Z_3)E(Z_5)E(Z_6)E(Z_7). \end{aligned}$$

This system admits a straightforward solution for the expectation of  $Z_4$ ,

$$E(Z_4) = \pm \sqrt{\frac{E(S_1)E(S_2)}{E(S_1S_2)}}. \quad (1)$$

This coincides with the solution given by Spitz *et al.* [7, eq. (14)], as explained in Appendix E. Notice that there is a choice of sign, which corresponds to deciding whether  $p_4 > \frac{1}{2}$  or  $p_4 < \frac{1}{2}$ . However, under the assumption  $p_4 < \frac{1}{2}$  the solution is unique.

## 2.2 Identifiability for classical codes

Let us now turn to the general setting, first for classical codes. We are interested in whether an error distribution can be estimated from repeated syndrome measurements alone. This is certainly not possible for completely arbitrary noise, since we do not measure the state of the logical qubit. However, we will show that for Pauli (and measurement) noise with limited correlations, the estimation is possible. As mentioned above, we will start with the setting of classical codes and perfect syndrome measurements. Then, we will show how to extend these results to quantum (data-syndrome) codes.

The key insight underlying our proof is that the estimation problem is best phrased in terms of moments instead of error rates. The proof then proceeds in the following steps. First, we notice that Fourier coefficients of the error distribution correspond to moments, and see that some of these moments can be estimated from the syndrome statistics. Then, we show that under certain independence assumptions, the full error distribution is characterized completely by a set of low weight transformed moments. We find that these transformed moments are related to the measured moments via a polynomial equation system. This system is described by a coefficient matrix  $D$  whose rows essentially correspond to elements of the dual code. We then use local randomness properties of the dual code to find an explicit expression for the symmetric squared coefficient matrix  $D^T D$ , and finally show that  $D^T D$  has full rank by using iterated Schur complements. This implies that the equation system has discrete solutions.

### 2.2.1 Notation

We use the short-hand notation  $[n] := \{1, \dots, n\}$  for the set of the first  $n$  natural numbers. For any set  $A$ , we denote its powerset as  $2^A := \{B \mid B \subseteq A\}$ . The field with two elements is denoted by  $\mathbb{F}_2$ . Often we will use  $\mathbb{F}_2^n$  as a vector space over that field. That is, for  $a, b \in \mathbb{F}_2^n$  the sum  $a + b$  is understood to be taken component-wise modulo 2. All vectors are to be understood as column vectors. By  $\text{wt}(a) := |\{i : a_i \neq 0\}|$  we denote the *weight* of  $a$ . For any logical statement  $X$  we denote by  $1[X]$  the *Iverson bracket* of  $X$ , which assumes the value 1 if  $X$  is true and 0 otherwise. Naturally, trivial products, i.e. products over the empty set, are set to 1 as in  $\prod_{x \in \emptyset} f(x) := 1$ . By  $I_k$  we denote the  $k \times k$  identity matrix and suppress  $k$  when it can be inferred from the context.

### 2.2.2 Classical codes and Boolean Fourier analysis

Let us start with some basic elements of Boolean Fourier analysis. A detailed review of the topic is given in [30] (note that we use a different normalization convention here). We frequently identify a vector  $s \in \mathbb{F}_2^n$  with its *indicator set*  $\{i \in [n] : s_i \neq 0\}$ . For example, for  $s = (0, 1, 0, 1, 1, 0) \in \mathbb{F}_2^6$  we also write

$$s = \{2, 4, 5\} \subseteq [6], \quad (2)$$

such that we can write  $4 \in s$ . For each  $s \in \mathbb{F}_2^n$ , we define the *parity function*

$$\begin{aligned}\chi_s &: \mathbb{F}_2^n \rightarrow \{-1, +1\}, \\ \chi_s(e) &= (-1)^{s \cdot e} = (-1)^{\sum_{i \in s} e_i},\end{aligned}$$

which is the group character of the abelian group  $(\mathbb{F}_2^n, +)$ . For a function  $f : \mathbb{F}_2^n \rightarrow \mathbb{R}$ , its *Boolean Fourier transform*  $\tilde{f}$  is the function

$$\tilde{f} : 2^{[n]} \rightarrow \mathbb{R}, \quad \tilde{f}(s) = \sum_{e \in \mathbb{F}_2^n} f(e) \chi_s(e). \quad (3)$$

The Boolean Fourier transform is also known under the name Walsh-Hadamard transformation. However, there are different conventions which lead to transforms with different orderings of bit strings and different notations for the characters  $\chi_s$ . This transformation is invertible, and the inverse is given by

$$f(e) = \frac{1}{2^n} \sum_{s \subseteq [n]} \tilde{f}(s) \chi_s(e). \quad (4)$$

For two functions  $f, g : \mathbb{F}_2^n \rightarrow \mathbb{R}$ , their *Boolean convolution*  $f * g$  is defined by

$$f * g(e) = \sum_{e' \in \mathbb{F}_2^n} f(e) g(e + e'). \quad (5)$$

As expected, convolutions become products under Boolean Fourier transform,

$$\widetilde{f * g} = \tilde{f} \tilde{g}. \quad (6)$$

A classical linear code  $C$ , encoding  $k$  bits into  $n$  bits, is a  $k$ -dimensional subspace of  $\mathbb{F}_2^n$ . It can be described by its parity check matrix  $H \in \mathbb{F}_2^{(n-k) \times n}$ , whose rows span the dual code  $C^\perp = \{a \in \mathbb{F}_2^n : a \cdot c = 0 \forall c \in C\}$ .  $C^\perp$  can alternatively be interpreted as the set of parity functions which are 1 on all codewords. When an error  $e \in \mathbb{F}_2^n$  occurs, the outcomes of the parity measurements  $H$  can be summarized by the *syndrome*

$$\mathcal{O}(e) := He. \quad (7)$$

From  $\mathcal{O}(e)$ , we can calculate the value of any parity measurement  $s \in C^\perp$ . We will assume that repeated rounds of syndrome measurements and corrections are performed. In each round, an error occurs according to the *error distribution*  $P : \mathbb{F}_2^n \rightarrow [0, 1]$ . We call the corresponding distribution of syndromes the *syndrome statistics*. The error correction capabilities of a classical linear code are indicated by its *distance*, which is defined as the smallest weight of an undetectable error, which is the same as the smallest weight of a codeword,

$$d := \min\{\text{wt}(e) : e \in C \setminus \{0\}\}. \quad (8)$$

### 2.2.3 Moments and noise model

We denote the Fourier transform of the error distribution  $P : \mathbb{F}_2^n \rightarrow [0, 1]$  with  $E$ , i.e. for each  $a \subseteq [n]$

$$E(a) := \sum_{e \in \mathbb{F}_2^n} (-1)^{a \cdot e} P(e) = \tilde{P}(a). \quad (9)$$

We interpret this as a *moment* of the distribution  $P$ , since  $E(a)$  is exactly the expectation value of the parity measurement  $a$  if one were to measure it repeatedly on errors distributed according to  $P$ . In particular, for  $s \in C^\perp$ , the corresponding moment  $E(s)$  can be computed from the measured syndrome statistics, i.e. from the empirically measured frequency with which each syndrome occurs in repeated rounds of error correction. We will always assume  $E(a) \neq 0$  for all  $a \subseteq [n]$ , which is, for example, fulfilled if  $P(0) > 0.5$ . Thus, our task is to find the distribution  $P$ , given some of its Fourier coefficients.

We will show that this task is feasible if there is some independence between errors on different subsets of bits, such that there are no correlations across a large number of qubits at once. To formalize this idea, we introduce a set of *channel supports*  $\Gamma \subseteq 2^{[n]}$  (we will consider a concrete choice of  $\Gamma$  later). For each  $\gamma \in \Gamma$ , there is an error channel that acts only on this subset, i.e. its error distribution  $P_\gamma : \mathbb{F}_2^n \rightarrow [0, 1]$  is only supported on errors  $e \in \mathbb{F}_2^n$  that are 0 outside of  $\gamma$ . Equivalently, identifying vectors with their indicator set, we can view this as a distribution  $P_\gamma : 2^{[n]} \rightarrow [0, 1]$  which is only supported on  $2^\gamma$ . Since all the individual channels act independently, and the total error is the sum of the individual errors, the total distribution of errors is then given by

$$P = \underset{\gamma \in \Gamma}{*} P_\gamma, \quad (10)$$

with the convolution (5). Physically, this expresses the action of several independent sources of errors, each only acting on a limited number of qubits. If  $\Gamma = \{\{1\}, \dots, \{n\}\}$ , this reduces to the commonly used model where each bit is affected by an independent channel. As a note, mathematically, this model corresponds to a convolutional factor graph, as introduced by Mao and Kschischang [31].

This representation in terms of individual channels is an over-parametrization of the total distribution if some supports in  $\Gamma$  overlap. A first idea for a non-redundant distribution would be to use the set of all moments, but these are not all independent. Therefore, we define transformed moments  $F(a)$  via a (generalized) inclusion-exclusion computation. Heuristically, we want to divide out all correlations on proper subsets of  $a$  in order to capture only correlations across the full size of  $a$ . This bears some similarity to the canonical parametrization of Markov random fields, see [32]. For  $b \subseteq a \subseteq [n]$ , the *Möbius function* (on the partially ordered set  $2^{[n]}$ ) is given by

$$\mu(a, b) = (-1)^{|a \setminus b|} = (-1)^{|a|+|b|}. \quad (11)$$

For each  $a \subseteq [n]$ , we define a transformed moment  $F(a)$  via the *inclusion-exclusion transform*

$$F(a) := \prod_{b \subseteq a} E(b)^{\mu(a, b)}. \quad (12)$$

It follows from the (multiplicative) generalized inclusion-exclusion principle (e.g. [33][Theorem 5.3], [34, Theorem A.2.4]) that the inverse of this transformation is given by

$$E(a) = \prod_{b \subseteq a} F(b). \quad (13)$$

Now we will show that, depending on the choice of  $\Gamma$ , a small subset of transformed moments is already sufficient to uniquely specify the distribution. Remember  $E_\gamma(a) = \tilde{P}_\gamma(a)$  from (9). Since  $P$  is given by the convolution (10) of all individual error distributions  $P_\gamma$  we have,

$$E(a) = \prod_{\gamma \in \Gamma} E_\gamma(a \cap \gamma), \quad (14)$$

where we used that  $P_\gamma$  is only supported on  $2^\gamma$ , and thus  $E_\gamma(a) = \tilde{P}_\gamma(a) = \tilde{P}_\gamma(a \cap \gamma) = E_\gamma(a \cap \gamma)$ .

We can now express the transformed moments by the parameters of the individual channels. The proof is given in Appendix A.

**Lemma 2.** *For  $\Gamma \subseteq 2^{[n]}$  let  $E : 2^{[n]} \rightarrow \mathbb{R}$  satisfy (14). Then, the inclusion-exclusion transform (12) satisfies*

$$F(a) = \prod_{\gamma \in \Gamma: a \subseteq \gamma} \prod_{b \subseteq a} E_\gamma(b)^{\mu(a, b)} \quad (15)$$

for all  $a \subseteq [n]$ . In particular,  $F(a) = 1$  if there is no  $\gamma \in \Gamma$  such that  $a \subseteq \gamma$ .

Notice that trivially  $F(\emptyset) = 1$ . Thus, in conclusion, it suffices to determine the transformed moments  $F(a)$  for  $a$  in

$$\hat{\Gamma} := \{\emptyset \neq a \subseteq \gamma : \gamma \in \Gamma\}. \quad (16)$$

The standard moments (13) are then determined by

$$E(a) = \prod_{b \in \hat{\Gamma}: b \subseteq a} F(b). \quad (17)$$

Finally, the error distribution  $P$  is determined from the standard moments by applying the inverse Fourier transform (4) to (9).

## 2.2.4 Identifiability and binomial systems

As discussed in the previous section, learning the transformed moments  $(F(a))_{a \in \hat{\Gamma}}$  is sufficient to uniquely determine the error distribution. However, the only thing we can measure are the standard moments  $(E(s))_{s \in C^\perp}$  corresponding to elements of the dual code, since these fully describe the syndrome statistics. Thus, learning the error distribution boils down to determining the transformed moments  $(F(a))_{a \in \hat{\Gamma}}$  from this information. In particular, the error distribution is identifiable from the syndrome statistics if and only if the following system of polynomial equations admits a unique solution,

Given  $(E(s))_{s \in C^\perp}$  find  $(F(b))_{b \in \hat{\Gamma}}$  satisfying

$$E(s) = \prod_{b \in \hat{\Gamma}: b \subseteq s} F(b) \quad \forall s \in C^\perp \setminus \{0\}, \quad (18)$$

where we have omitted the trivial equation arising from the zero element of the dual code. This is a *binomial* system, i.e. each equation only has two terms, a constant on the left-hand side and a monomial on the right-hand side. In the notation of [35], the binomial system (18) can be expressed by the *coefficient matrix*  $D$  whose rows are labeled by elements of  $C^\perp \setminus \{0\}$  and whose columns are labeled by elements of  $\hat{\Gamma}$ . The entry  $D[s, a]$  is the exponent of  $F(a)$  in the monomial on the right-hand side of equation (18) for  $E(s)$ . We have transposed the notation comparing to [35]. Explicitly this means that the  $|C^\perp \setminus \{0\}| \times |\hat{\Gamma}|$ -matrix  $D$  has the elements

$$D[s, a] = \begin{cases} 1 & \text{if } a \subseteq s \\ 0 & \text{otherwise} \end{cases}, \quad (19)$$

and the system is given by  $\vec{E} = \vec{F}^D$  [35, Eq. (3)], where  $\vec{E}$  is the vector containing the elements  $E(s)$  and  $\vec{F}$  the elements  $F(a)$ , as appearing in the system (18). In the special case of single bit noise, i.e.  $\Gamma = \hat{\Gamma} = \{\{i\} : i \in [n]\}$ , the rows of the coefficient matrix  $D$  are exactly the dual codewords  $s \in C^\perp$ .

For now, let us assume that  $D$  has full rank. It then follows from the theory of binomial system solving [35, Proposition 2] that the system (18) has a finite number of solutions, and these solutions only differ by multiplying some parameters with complex roots of unity. Since we are only interested in real solutions, this means we can determine the transformed moments  $F(a)$ , and thus also the standard moments  $E(a)$ , up to a sign. Thus, if we restrict all moments to be positive, the error distribution is uniquely determined by the syndromes. A simple condition for all moments to be positive is that the error probabilities of all channels are smaller than  $\frac{1}{2}$ , i.e.  $P_\gamma(0) > \frac{1}{2}$  for all  $\gamma \in \Gamma$ . This implies  $E_\gamma(a) > 0$  and thus  $E(a) > 0$  for all  $a \subseteq [n]$ . In conclusion, the error distribution can be estimated uniquely from the syndromes, assuming that  $D$  has full rank and that the error probability of each channel is smaller than  $\frac{1}{2}$ .

Let us stress that the appearance of multiple solutions, differing by the signs of some moments, reflects actual symmetries of the syndrome statistics as a function of the error rates. For example, consider again case of independent single bit errors, such that  $\Gamma = \hat{\Gamma} = \{\{i\} : i \in [n]\}$ . Then the transformed moments  $F(\{i\})$  are equal to the standard moments  $E(\{i\})$ , and the rows of  $D$  are simply the elements of the dual code. Thus, each row of  $D$  has even overlap with every codeword  $c \in C$ . Thus, by (18), flipping the signs of all moments  $(E(\{i\}))_{i \in c}$  on the support of the codeword  $c$  does not change the measured moments  $(E(s))_{s \in C^\perp}$ . Since  $E(\{i\}) = 1 - 2p_i$ , where  $p_i$  is the rate of errors on the  $i$ -th qubit, this simply means that flipping the error rates around  $\frac{1}{2}$  on a codeword does not affect the syndrome statistics, i.e. we cannot distinguish these two sets of error rates. This observation shows that each codeword corresponds to a symmetry of the identifiability problem.

### 2.2.5 The rank of the coefficient matrix

We will now establish the most general set of error distributions for which unique identification of the error rates is possible. By the discussion in the previous section, this corresponds to finding the most general choice of  $\Gamma$  for which the system (18) is solvable, i.e. for which the coefficient matrix  $D$  is of full rank.

Denote the sets of bits that only support *detectable* errors as

$$\Gamma^{(D)} := \{\gamma \subseteq [n] : \mathcal{O}(e) \neq 0 \forall e \subseteq \gamma\}, \quad (20)$$

where  $\mathcal{O}(e)$  is the syndrome (7) of  $e$ . It is clear that, if one of the channels  $\gamma \in \Gamma$  supports an undetectable error, one cannot estimate the corresponding error rate. Thus, identifiability can only hold if  $\gamma \in \Gamma^{(D)}$ . Similarly, if two different channels  $\gamma_1, \gamma_2$  contain two errors with the same syndrome, the syndrome statistics only depends on the combined rate of those errors and thus the rates are not identifiable. Identifiability of the noise channel can only hold if  $\gamma_1 \cup \gamma_2 \in \Gamma^{(D)}$  for all  $\gamma_1, \gamma_2 \in \Gamma$ . We will show that these are in fact the only restriction that must be fulfilled.

**Theorem 3** (Classical identifiability condition). *Consider a classical code with  $n$  bits subject to noise with an error distribution described by channel supports  $\Gamma \subseteq 2^{[n]}$ . Then the coefficient matrix  $D$  defined in (19) is of full rank if and only if  $\gamma_1 \cup \gamma_2 \in \Gamma^{(D)}$  for all  $\gamma_1, \gamma_2 \in \Gamma$ .*

The proof is given in Sections 2.2.6. Even a combination of two channels must not support an undetectable error. We can also give an equivalent condition in terms of  $\hat{\Gamma}$ , which represents the set of errors that can occur “independently”. As explained in Appendix B, the assumption  $\gamma_1 \cup \gamma_2 \in \Gamma^{(D)}$  for all  $\gamma_1, \gamma_2 \in \Gamma$  is equivalent to  $\mathcal{O}(e_1) \neq 0$  and  $\mathcal{O}(e_1 + e_2) \neq 0$  for all  $e_1, e_2 \in \hat{\Gamma}$  (viewed as binary vectors). In other words, the error distribution is identifiable if undetectable errors and errors that have the same syndrome only occur as a combination of independent errors. We stress that this is substantially weaker than the assumption that errors with the same syndrome never occur (which was made in some previous works such as [20, 21]). Indeed, in the total error distribution different errors with the same syndrome can occur frequently, and we are still able to identify the error rates. We only assume that such errors arise as a combination of independent errors.

The most important example is the following. Consider an error model where there is an independent channel on every subset of  $t$  bits, i.e. choose  $\Gamma$  as

$$\Gamma = \Gamma_t := \{\gamma \subseteq [n] : |\gamma| = t\} \quad (21)$$

$$\hat{\Gamma} = \Gamma_{\leq t} := \{e \subseteq [n] : 0 < |e| \leq t\}. \quad (22)$$

This means we only have correlations across at most  $t$  bits. Then, Theorem 3 implies that the error distribution is identifiable as long as  $d \geq 2t + 1$ .

**Corollary 4** (Distance based identifiability condition). *If the noise is described by channel supports  $\Gamma = \Gamma_t$ , the coefficient matrix  $D$  is of full rank if the code has distance  $d \geq 2t + 1$ .*

*Proof.* For a code with distance  $d$ ,  $\Gamma^{(D)}$  contains every set of size at most  $d - 1$ . □

For  $t = 1$ , we see that error rates of the standard single bit noise model can be identified as long as  $d \geq 3$ . Informally, identification is possible if error correction is possible.

### 2.2.6 Orthogonal array properties of the coefficient matrix

This section is devoted to the proof of Theorem 3. Readers not interested in the proofs and mathematical techniques can skip to Section 2.3 for the results in the quantum setting. The first part of the proof is based on local randomness properties of the dual code. The elements of the dual code form a so-called *orthogonal array*. Since the entries of the coefficient matrix  $D$  are related to the dual code, we can use this property to derive an explicit expression for the symmetric squared coefficient matrix  $D^T D$ . The proof is then finished by computing the rank of  $D^T D$ , which is possible with the help of combinatorial results derived in the next section.

By  $T$  we denote the  $|C^\perp| \times n$ -matrix formed by all elements in the span of the rows of the parity check matrix  $H$ , i.e. the rows of  $T$  are exactly the elements of  $C^\perp$ . It is known that the rows of  $T$  look “locally uniformly random” on any subset of up to  $d-1$  bits. One says that  $T$  is an orthogonal array of strength  $d-1$ . This property is relatively easy to see for linear codes (e.g. [36]) and was shown for general (non-linear) classical codes by Delsarte [37]. We use a slightly extended version of the result for linear codes using the set  $\Gamma^{(D)}$  from (20) instead of the distance. A proof is given in Appendix C.

**Lemma 5.** *Let  $\gamma \in \Gamma^{(D)}$ . In the restriction  $T|_\gamma$  of  $T$  to columns in  $\gamma$ , every bit-string appears equally often as a row.*

In other words, the rows of  $T$  look locally uniformly random on any choice of bits that only supports detectable errors.

In Appendix D.2 we prove the following statement, as a corollary of Lemma 11.

**Lemma 6** (Positive-definiteness of intersection matrix). *Let  $M_t$  be the matrix whose rows and columns are labeled by the elements of  $\Gamma_{\leq t}$  (from (22)) and which is defined entry-wise by*

$$M_t[a, b] = 2^{|a \cap b|}. \quad (23)$$

*Then  $M_t$  is positive-definite.*

We call the matrix  $M_t$  the *intersection matrix*. The dimensions of  $M_t$  depend on  $n$ , but we do not make this dependence explicit in our notation. Using Lemma 6, we can finish the proof of Theorem 3.

*Proof of Theorem 3.* We will prove that the coefficient matrix  $D$  from (19) has full rank (over  $\mathbb{R}$ ) by proving that  $D^T D$  has full rank. First, note that by the definition (20) of  $\Gamma^{(D)}$ , for any  $\gamma \in \Gamma^{(D)}$ , all subsets  $a \subseteq \gamma$  are also elements of  $\Gamma^{(D)}$ . Thus, the assumption  $\gamma_1 \cup \gamma_2 \in \Gamma^{(D)}$  from Theorem 3 implies  $a \cup b \in \Gamma^{(D)}$  for all  $a \cup b \in \hat{\Gamma}$ , by the definition (16) of  $\hat{\Gamma}$ . The dot product of the columns of  $D$  labeled by  $a, b \in \hat{\Gamma}$  (as binary vectors) is the number of elements  $s \in C^\perp \setminus \{0\}$  such that  $a \cup b \subseteq s$ . Thus, by Lemma 5, this number is  $|C^\perp| 2^{-|a \cup b|}$ , i.e.  $D^T D[a, b] = |C^\perp| 2^{-|a \cup b|}$ . We can write this as  $D^T D[a, b] = |C^\perp| 2^{-|a| - |b| + |a \cap b|}$ . Re-scaling the rows and columns leads to the modified matrix  $(D^T D)'[a, b] = 2^{|a \cap b|}$ . Let us denote  $t = \max\{|a| : a \in \hat{\Gamma}\}$ . Then  $(D^T D)'$  is a principal sub-matrix of the intersection matrix  $M_t$  defined in (23). By Lemma 6,  $M_t$  is positive-definite. As a principal sub-matrix of a positive-definite matrix,  $(D^T D)'$  is then also positive-definite and, in particular, has full rank. This implies that  $D$  has full rank, which finishes the proof of Theorem 3.  $\square$

## 2.3 Extension to the quantum case

Now will consider the quantum setting, starting with a short overview of stabilizer codes. We also explain the concept of quantum data-syndrome codes [26–28], following Ashikhmin *et al.* [26], which allow for a unified treatment of data and measurement errors. Then, we state and explain our main result Theorem 7, which is the formal version of Theorem 1. Finally, we explain how to prove this theorem by extending our arguments from the classical to the quantum case.

### 2.3.1 Preliminaries

The Pauli group  $\mathcal{P}^n$  on  $n$  qubits is the group of *Pauli strings* generated by the Pauli operators  $\{I, X, Y, Z, \}$  with phases,

$$\mathcal{P}^n = \left\{ \epsilon \bigotimes_{i=1}^n P_i \mid \epsilon \in \{\pm 1, \pm i\}, P_i \in \{I, X, Y, Z\} \right\}.$$

The *weight*  $\text{wt}(P)$  of a Pauli string  $P = \epsilon \bigotimes_{i=1}^n P_i$  is the number of non-identity components  $P_i$ . Modding out phases, one obtains the *effective Pauli group*

$$\mathbf{P}^n = \mathcal{P}^n / \{\pm 1, \pm i\}. \quad (24)$$

We define the *symplectic inner product*  $\langle \cdot, \cdot \rangle_{\mathcal{P}}$  on  $\mathcal{P}^n$  by

$$\langle e, e' \rangle_{\mathcal{P}} = \begin{cases} 1, & e \text{ and } e' \text{ anti-commute in } \mathcal{P}^n \\ 0, & e \text{ and } e' \text{ commute in } \mathcal{P}^n \end{cases}; \quad (25)$$

note that this expression is well-defined since the commutation relation does not depend on the choice of representatives in  $\mathcal{P}^n$ . We identify  $\mathcal{P}^1$  with  $\mathbb{F}_2^2$  via the *phase space representation*,

$$X \mapsto (1, 0), \quad Z \mapsto (0, 1), \quad Y \mapsto (1, 1), \quad (26)$$

which extends coordinate-wise to define a group isomorphism  $\mathcal{P}^n \rightarrow \mathbb{F}_2^{2n}$ . Thus, an element of  $\mathcal{P}^n$  is represented by  $n$  “ $X$ -bits” and  $n$  “ $Z$ -bits”. Explicitly,  $X^{x_1}Z^{z_1} \otimes \dots \otimes X^{x_n}Z^{z_n}$  is mapped to  $(x_1, \dots, x_n, z_1, \dots, z_n)^T$ . For example

$$X \otimes I \otimes Z \otimes Y \mapsto (1, 0, 0, 1, 0, 0, 1, 1)^T. \quad (27)$$

This identification will allow for the application of Boolean Fourier analysis to the Pauli group. We denote the operation of swapping the  $X$ -bits and  $Z$ -bits by a bar,  $(x_1, \dots, x_n, z_1, \dots, z_n) = (z_1, \dots, z_n, x_1, \dots, x_n)$ . The symplectic inner product (25) then corresponds to

$$\langle e, e' \rangle_{\mathcal{P}} = \bar{e} \cdot e', \quad (28)$$

where the dot product is evaluated in  $\mathbb{F}_2^{2n}$ , i.e. modulo 2. We define the *Pauli weight*  $\text{wt}_{\mathcal{P}}(e)$  of  $e \in \mathcal{P}^n \cong \mathbb{F}_2^{2n}$  as the weight of the corresponding Pauli operator, i.e.  $\text{wt}_{\mathcal{P}}(e) = |\{i \in [n] : e_i \neq 0 \vee e_{i+n} \neq 0\}|$ .

We describe errors using Pauli channels, i.e. quantum channels of the form

$$\rho \mapsto \sum_{e \in \mathcal{P}^n} P(e) e \rho e^\dagger, \quad (29)$$

where  $P : \mathcal{P}^n \rightarrow [0, 1]$  is a normalized probability distribution, the *error distribution* of the channel; note again that this expression is independent of the choices of representatives of  $e$  in  $\mathcal{P}^n$ .

A stabilizer code is defined by a set of commuting Pauli operators  $g^{(1)}, \dots, g^{(l)} \in \mathcal{P}^n$ . They generate an abelian subgroup  $\mathcal{S} \subseteq \mathcal{P}^n$ , called stabilizer group, which must fulfill  $-1 \notin \mathcal{S}$ . This is the analogue of the classical dual code  $C^\perp$ . The codespace is defined as the simultaneous  $+1$  eigenspace of the operators in  $\mathcal{S}$ . Standard error correction with a stabilizer code proceeds as follows. In each round all generators are measured. If an error  $e \in \mathcal{P}^n$  occurred then the vector of all measurement outcomes can be represented by the syndrome  $\mathcal{O}(e)$ , defined as

$$\mathcal{O}(e)_i = \langle g^{(i)}, e \rangle_{\mathcal{P}} \quad \forall i = 1, \dots, l. \quad (30)$$

Using the measured syndrome, and based on information about the error rates, one approximates the most likely logical error and applies it as a correction. The outcomes of all stabilizers are determined only by the measurements of the generators (in the case of perfect measurements) via linearity of  $\langle \cdot, \cdot \rangle_{\mathcal{P}}$ . As in the classical case, we can define the distance of a stabilizer code. However, in contrast to the classical case, there are many errors that act trivially on the code space and do not affect the logical information. We define the *distance* as the smallest weight of an undetectable error that affects the state of the logical qubit, i.e.

$$d := \min\{\text{wt}_{\mathcal{P}}(e) : e \in \mathcal{P}^n \setminus \mathcal{S}, \mathcal{O}(e) = 0\}. \quad (31)$$

Moreover, we define the *pure distance* of a code as the smallest weight of any undetectable error,

$$d_p := \min\{\text{wt}_{\mathcal{P}}(e) : e \in \mathcal{P}^n \setminus \{I\}, \mathcal{O}(e) = 0\}. \quad (32)$$

The pure distance and the distance can differ significantly. For example, the distance of the toric code is equal to the lattice size. The pure distance on the other hand is 4 independent of the lattice size, since the weight of any star or plaquette stabilizer is 4.

In a practice, the stabilizer measurements themselves could also be faulty. In this case, one should measure additional elements of  $\mathcal{S}$  to mitigate the effect of measurement errors. This is

captured by the framework of *quantum data-syndrome codes*, which allow for a unified treatment of data and measurement errors [26–28]. We now give the basic definitions, following [26]. In the context of data-syndrome codes, errors are described by a data and a measurement part, as  $e = (e_d, e_m) \in \mathbb{P}^n \times \mathbb{F}_2^m \cong \mathbb{F}_2^{2n+m}$ . The swapping of  $X$ - and  $Z$ -bits now only applies to the data bits, i.e.  $\bar{e} = (\bar{e}_d, e_m)$  for  $e \in \mathbb{F}_2^{2n+m}$ . We extend the symplectic product to data-syndrome codes by

$$\langle (s_d, s_m), (e_d, e_m) \rangle_{\text{DS}} = \langle s_d, e_d \rangle_{\mathbb{P}} + s_m \cdot e_m s_d \cdot \bar{e}_d + s_m \cdot e_m. \quad (33)$$

A quantum data-syndrome code is defined by a stabilizer code on  $n$  physical qubits with generators  $g^{(1)}, \dots, g^{(l)} \in \mathbb{P}^n \cong \mathbb{F}_2^{2n}$  and a classical code that encodes  $l$  bits into  $m$  bits. We can always write the generator matrix of the classical code in the systematic form  $G_C = [I_l \ A]$  with  $A \in \mathbb{F}_2^{l \times (m-l)}$ . Instead of just the generators  $g^{(1)}, \dots, g^{(l)}$ , we measure the stabilizers  $f^{(1)}, \dots, f^{(m)}$  defined by

$$f^{(i)} = \sum_{j=1}^l G_C[j, i] g^{(j)}. \quad (34)$$

Note that  $f^{(i)} = g^{(i)}$  for  $i \leq l$ . We collect the stabilizers  $f^{(i)}$  into a matrix  $F = \begin{bmatrix} f^{(1)T} \\ \vdots \\ f^{(m)T} \end{bmatrix}$ . The measurements of the stabilizers can then be described by the parity check matrix

$$H_{\text{DS}} = [F \ I_m], \quad (35)$$

where the identity part describes the effect of measurement errors, as seen by the following discussion. If an error  $e = (e_d, e_m) \in \mathbb{P}^n \times \mathbb{F}_2^m$  occurred, the syndrome  $\mathcal{O}(e)$  is then described by

$$\mathcal{O}(e)_i = \langle h^{(i)}, e \rangle_{\text{DS}}, \quad (36)$$

where  $h^{(i)}$  is the  $i$ -th row of  $H_{\text{DS}}$ . In phase space representation this simply means  $\mathcal{O}(e) = H\bar{e} = F\bar{e}_d + e_m$ , i.e. the syndrome is the sum of the ideal syndrome and the measurement errors. We denote the row span of  $H_{\text{DS}}$  as  $\mathcal{S}_{\text{DS}}$ , since it is an analogue of the stabilizer group.

We define the Pauli weight of  $e = (e_d, e_m) \in \mathbb{P}^n \times \mathbb{F}_2^m \cong \mathbb{F}_2^{2n+m}$  as  $\text{wt}_{\mathbb{P}}(e) = \text{wt}_{\mathbb{P}}(e_d) + \text{wt}(e_m)$ . Analogous to stabilizer codes, we can define the *distance* of the data-syndrome code as

$$d := \min\{\text{wt}_{\mathbb{P}}(e) : e \in \mathbb{P}^n \times \mathbb{F}_2^m \setminus \{(h, 0) : h \in \mathcal{S}\}, \mathcal{O}(e) = 0\},$$

and the *pure distance* as

$$d_p := \min\{\text{wt}_{\mathbb{P}}(e) : e \in \mathbb{P}^n \times \mathbb{F}_2^m \setminus \{0\}, \mathcal{O}(e) = 0\}.$$

Naturally, the distance of a data-syndrome code cannot be larger than that of the underlying quantum code. Furthermore, it is not hard to see from the definitions that the pure distance of a data-syndrome code is the minimum of its distance and the pure distance of the underlying stabilizer code. Thus, the pure distance is limited primarily by the underlying quantum code.

All in all, measurement errors and data errors can be treated in a unified way, analogous to a standard stabilizer code.

### 2.3.2 Identifiability results for quantum codes

Now we extend the identifiability results from classical to quantum data-syndrome codes, by applying analogous arguments to the phase space representation. The explicit example of the toric code is discussed in Section 2.1.

We consider a quantum data-syndrome code on  $n$  qubits and  $m$  measurement bits, and set  $N := 2n + m$ . Similar to the classical case, we consider an error model described by independent channels on some selections of qubits and measurement bits. We also allow that some of the channels contain only  $X$ - or  $Z$ -errors on some qubits. Thus, in the phase space representation, we consider a set of *channel supports*  $\Gamma \subseteq 2^{[N]}$ , and for each  $\gamma \in \Gamma$  an error distribution  $P_{\gamma} : 2^{[N]} \mapsto [0, 1]$  that is only supported on  $2^{\gamma}$ , where we again identify binary vectors in  $\mathbb{F}_2^N$  with subsets of

$[N]$ . Since the total error is a sum of independently occurring errors, the total error distribution is again given by a Boolean convolution,

$$P = \underset{\gamma \in \Gamma}{*} P_\gamma. \quad (37)$$

We denote the sets of bits in phase space representation that only support *detectable* errors as

$$\Gamma^{(D)} = \{\gamma \subseteq [N] : \mathcal{O}(e) \neq 0 \forall e \subseteq \gamma\} \quad (38)$$

and denote  $\overline{\Gamma^{(D)}} = \{\bar{e} : e \in \Gamma^{(D)}\}$ . Using these definitions, we can state our main result.

**Theorem 7** (General identifiability condition). *Consider a quantum data-syndrome code with  $n$  qubits and  $m$  measurement bits subject to noise described by the channel supports  $\Gamma \subseteq 2^{[N]}$ , where  $N := 2n + m$ . Assume that any union of two channel supports only supports detectable errors, i.e. for all  $\gamma_1, \gamma_2 \in \Gamma$ ,  $\gamma_1 \cup \gamma_2 \in \Gamma^{(D)}$ . Furthermore assume  $P_\gamma(0) > 0.5$  for all  $\gamma \in \Gamma$ . Then the total error distribution  $P$  is identifiable from the syndrome statistics.*

As in the classical case, we can also consider the set of “independently occurring errors”,

$$\hat{\Gamma} = \{\emptyset \neq e \subseteq [N] : \exists \gamma \in \Gamma \text{ such that } e \subseteq \gamma\}, \quad (39)$$

instead of the set of channel supports  $\Gamma$ . As explained Appendix B, an equivalent condition to  $\gamma_1 \cup \gamma_2 \in \Gamma^{(D)}$  for all  $\gamma_1, \gamma_2 \in \Gamma^{(D)}$  in terms of these independent errors is that  $\mathcal{O}(e_1) \neq 0$  and  $\mathcal{O}(e_1 + e_2) \neq 0$  for all  $e_1, e_2 \in \hat{\Gamma}$ . Thus, we require that independently occurring errors have different syndromes. As also discussed in the classical case, this does not preclude that different errors with the same syndrome occur frequently, but they must arise as a combination of independent errors.

The most important implication of Theorem 7 is the following.

**Corollary 8** (Pure distance and identifiability). *Consider a quantum data-syndrome code of pure distance  $d_p$  on  $n$  qubits and  $m$  measurement bits, subject to noise described by the channel supports  $\Gamma = \{\gamma \subseteq [2n + m] : \text{wt}_P(\gamma) \leq t\}$ . Furthermore assume  $P_\gamma(0) > 0.5$  for all  $\gamma \in \Gamma$ . Then  $P$  is identifiable from the syndrome statistics if  $t \leq \lfloor \frac{d_p - 1}{2} \rfloor$ .*

*Proof.* If the quantum data-syndrome code has pure distance  $d_p$  then  $\Gamma^{(D)}$  contains any  $\gamma \subseteq [N]$  with Pauli weight at most  $t$  (when viewed as a binary vector).  $\square$

In other words, with a code of pure distance  $d_p$  we can estimate Pauli noise that is correlated across  $\lfloor \frac{d_p - 1}{2} \rfloor$  combined qubits and measurement bits. This proves the informal Theorem 1. One could also make stronger independence assumptions, for example that data and measurement errors occur independently of each other. In this case one can consider the pure distance  $d_Q$  of the underlying quantum code and the distance  $d_C$  of the measurement code independently and can estimate correlations across at least  $d_Q$  data qubits and  $d_C$  measurement bits. Similarly, for CSS-codes, one can consider a separate  $X$ - and  $Z$ -distance if one assumes that  $X$ - and  $Z$ -errors occur independently.

Let us now discuss the proof of Theorem 7, which consists of carefully applying the arguments from the classical case to the phase space representation. In this framework, the main difference to the classical case is that the moments must be defined using the symplectic product instead of the normal dot product if we want them to match the measured expectation values. Thus, for any subset  $a \subseteq [N]$ , we define

$$E(a) := \sum_{e \in \mathbb{F}_2^N} (-1)^{\langle a, e \rangle_{\text{DS}}} P(e), \quad (40)$$

$$E_\gamma(a) := \sum_{e \in \mathbb{F}_2^N} (-1)^{\langle a, e \rangle_{\text{DS}}} P_\gamma(e). \quad (41)$$

For  $s \in \mathcal{S}_{\text{DS}}$ ,  $E(s)$  is again exactly the expectation value of the measurement of the stabilizer in repeated rounds of error correction. However, the relation between moments and Fourier coefficients now contains a “twist”, i.e.

$$E(a) = \sum_{e \in \mathbb{F}_2^N} (-1)^{\langle a, e \rangle_{\text{DS}}} P(e) = \sum_{e \in \mathbb{F}_2^N} (-1)^{\bar{a} \cdot e} P(e) = \tilde{P}(\bar{a}).$$

A similar connection between measurements and Fourier coefficients has recently been pointed out by Flammia and O'Donnell [15].

Because  $P_\gamma$  is only supported on  $2^\gamma$ ,  $\tilde{P}_\gamma(a) = \tilde{P}_\gamma(a \cap \gamma)$  and thus  $E_\gamma(a) = E_\gamma(a \cap \bar{\gamma})$ . It follows that

$$E(a) = \prod_{\gamma \in \Gamma} E_\gamma(a \cap \bar{\gamma}). \quad (42)$$

We can define the transformed moments  $F(b)$  via the inclusion-exclusion transform (12) as in the classical case and obtain

$$E(a) = \prod_{b \subseteq a} F(b) \quad \forall a \subseteq [N]. \quad (43)$$

By replacing  $\gamma$  with  $\bar{\gamma}$  in Lemma 2, we obtain  $F(a) = 1$  if there is no  $\gamma \in \Gamma$  such that  $a \subseteq \bar{\gamma}$ . It thus suffices to know the transformed moments  $F(b)$  for  $b \in \bar{\Gamma}$ , where

$$\bar{\Gamma} = \{\emptyset \neq A \subseteq [N] : \exists \gamma \in \Gamma \text{ such that } A \subseteq \bar{\gamma}\}.$$

The problem of learning the error rates has thus been reduced to the problem of learning the transformed moments from the measured expectation values. Explicitly, we need to solve the following equation system, which is analogous to the classical case.

$$\begin{aligned} &\text{Given } (E(s))_{s \in \mathcal{S}_{\text{DS}}} \text{ find } (F(b))_{b \in \bar{\Gamma}} \text{ satisfying} \\ E(s) &= \prod_{b \in \bar{\Gamma}: b \subseteq s} F(b) \quad \forall s \in \mathcal{S}_{\text{DS}} \setminus \{0\}. \end{aligned} \quad (44)$$

This system can be described by the *coefficient matrix*  $D$ , whose rows are labeled by elements of  $\mathcal{S}_{\text{DS}}$  and whose columns are labeled by elements of  $\bar{\Gamma}$ , with entries

$$D[s, a] = \begin{cases} 1 & \text{if } a \subseteq s \\ 0 & \text{otherwise} \end{cases}. \quad (45)$$

We are now in a position to finish the proof of Theorem 7.

*Proof of Theorem 7.* First we show that the coefficient matrix  $D$  defined in (45) has full rank. Note that by the definition of  $\Gamma^{(D)}$ , if  $\gamma \in \Gamma^{(D)}$ , then also  $a \in \Gamma^{(D)}$  for any  $a \subseteq \gamma$ . The assumption that  $\gamma_1 \cup \gamma_2 \in \Gamma^{(D)}$  for all  $\gamma_1, \gamma_2 \in \Gamma$  thus implies that  $a \cup b \in \bar{\Gamma}^{(D)}$  for all  $a, b \in \bar{\Gamma}$ . It is proven in Appendix C (Lemma 10) that the elements of  $\mathcal{S}_{\text{DS}}$  look locally uniformly random on any element of  $\bar{\Gamma}^{(D)}$ . The same arguments as in the proof of Theorem 3 thus imply that  $D^T D$  can be re-scaled to be a principal sub-matrix of the matrix  $M_t$  from (23) (for  $t = \max\{|a| : a \in \bar{\Gamma}\}$ ). Since  $M_t$  is positive-definite by Lemma 6, we conclude that  $D^T D$  and thus  $D$  is of full rank. By [35][Proposition 2], the re-scaled moments  $F(a)$ , and thus also the standard moments  $E(a)$ , can be estimated up to their sign. If we assume all moments  $E(a)$  to be positive, then the estimate is unique. A sufficient condition for all moments  $E(a)$  to be positive is that  $P_\gamma(0) > 0.5$  for all  $\gamma \in \Gamma$ .  $\square$

Similar to the classical case, the appearance of multiple solutions, differing by the signs of some moments, reflects the symmetries of the problem. This becomes especially apparent when considering the simple case of a stabilizer code with single qubit noise and no measurement errors. In this case, errors on each qubit are independent. Thus, for a stabilizer  $S = S_1 \otimes \cdots \otimes S_n \in \mathcal{P}^n$ , we have  $E(S) = \prod_{i \in [n]} E(S_i)$ , which is a simpler form of the equation system (44). Consider a representative  $L \in \mathcal{P}^n$  of a logical operator (including stabilizers). Denote as

$$\mathcal{A}(L) := \{A \in \mathcal{P}^n : \text{wt}_{\text{P}}(A) = 1, \langle A, L \rangle_{\text{P}} = 1\} \quad (46)$$

the set of single qubit Pauli operators that anti-commute with  $L$ . Since  $L$  commutes with all stabilizers  $S$ , each equation  $E(S) = \prod_{i \in [n]} E(S_i)$  contains, on the right-hand side, an even number of terms  $E(S_i)$  such that  $S_i \in \mathcal{A}(L)$ . Thus, flipping the sign of  $E(S_i)$  for all  $S_i \in \mathcal{A}(L)$  does not

change the measured syndrome statistics. We see that there is one symmetry for each representative of a logical operator.

In summary, the estimation problem is best phrased in terms of moments instead of error rates. From this perspective, estimating the error distribution boils down to solving a polynomial system (44). If the correlations in the error distribution are small enough, this system has a finite number of discrete solutions. These solutions are described by symmetries related to the logical operators of the code. Under the additional mild assumption that the error rates are smaller than  $\frac{1}{2}$  the solution is unique.

## 2.4 Practical estimation

While the main result of this paper lies in establishing identifiability conditions in principle, our proofs also suggest a practical method to actually perform the estimation. Here, we briefly sketch this method and heuristically comment on its sample complexity. A detailed analysis is ongoing work.

We suggest a method of moments estimator, i.e. to use empirical expectation values  $\hat{E}(s)$  for  $E(s)$  in (44). The basic steps (for the quantum case, but the classical case is exactly analogous) are the following:

1. Perform  $m$  syndrome measurements, and use them to compute the empirical expectation value  $\hat{E}(s)$  for each stabilizer  $s$ .
2. Insert these empirical expectation values into (44) and solve the resulting binomial system to obtain an estimate  $(\hat{F}(b))_{b \in \bar{\Gamma}}$  of the transformed moments.
3. Insert this estimate into (43) to obtain an estimate of the moments.
4. Perform the inverse Fourier transform (4) to obtain an estimate  $\hat{P}$  of the Pauli error rates.

The relevant binomial system can be solved e.g. using the methods described by Chen and LiTien-Yien [35]. It will generally be over-determined. In principle, one can select a subset of equations such that the system is exactly determined, and then solve it analytically, resulting in a closed form expression for the estimate of the error rates in terms of the empirically measured expectation values. This is illustrated by the example of the toric code (Section 2.1), for which our method reproduces the estimator suggested by Spitz *et al.* [7]. This estimator has also been applied by Varbanov *et al.* [38]. However, since the empirical expectations contain a certain error, this solution might not always yield a proper probability distribution, if the number of samples used is low. Furthermore, for an over-determined system, selecting a subset of equations removes some of the measured information, which can reduce the accuracy of the estimate. In such cases it might be preferable to instead use a least-squares solver for the over-determined system.

Since our algorithm is designed for an on-line setting where only syndrome information is available, the sample complexity must necessarily be worse than that of algorithms using arbitrary measurements, such as [17] and [15]. Each syndrome measurement contains a measurement of each stabilizer generator, therefore the expectation values of all stabilizers can be estimated simultaneously and no large measurement overhead is needed. However, the estimation error has to be propagated through the binomial system, the inclusion-exclusion transform and the inverse Fourier transform. Heuristically, we expect a sample complexity similar to that of factor graph learning [39], which was applied to Pauli channel estimation by Flammia and Wallman [17, Result 3]. However, there will be a complicating factor accounting for the conditioning of the binomial system described by the coefficient matrix  $D$ . Compared to other algorithms designed for the on-line setting [2, 7, 19–24], we expect that our method compares favorably, since these algorithms are either designed for a limited class of codes and Pauli noise models, deal with likelihood functions based on syndrome probabilities, not moments, which quickly become intractable, or only work in the limit of vanishing error rates.

In practice, it will not be necessary to estimate the expectation values of all stabilizers, but only a limited subset will be sufficient, depending on how many equation of the binomial system are retained. We expect that it is sufficient to only consider a selection of neighboring stabilizers

for each qubit. For example, in the toric code example Section 2.1, only three expectation values are needed per qubit, independent of the system size.

Finally, in case that the Pauli noise model is not strictly identifiable, we expect that the estimate will combine the rates of indistinguishable errors, but still give a good estimate of these total error rates.

### 3 Discussion

In this work, we considered the estimation of Pauli channels just from the measurements done anyway during the decoding of a quantum code. We established a general condition for the feasibility of this estimation and explained the relation to the pure distance of the code. Essentially, the estimation is possible as long as the noise has no correlations which exceed the detection capabilities of the code. This result does not rely on the limit of vanishing error rates, and applies even if high weight errors occur frequently, as long as these high weight errors arise as a combination of independent lower weight errors. Our results cover general stabilizer codes and quantum data-syndrome codes [26–28] and also take into account measurement errors. The previously proposed *adaptive weight estimator* [7] can be seen as a special case of our general results, since it solves a specific instance of the system (18) for single qubit noise and a limited class of codes, those that admit a minimum weight matching decoder.

An interesting new problem is an analogue of our results on the logical level. Since we cannot identify the rates of undetectable errors, there is no full identifiability for correlations larger than the pure distance. However, we expect that the logical channel can still be identified, as long as there are no correlations larger than the actual distance of the code. This would practically allow for an analogue of Theorem 1, using the distance instead of the pure distance. Furthermore, our proof naturally suggests a method of moments estimator for the Pauli error rates by inserting the empirical moments in (18). A detailed analysis of this method, including rigorous performance guarantees as well as a better assessment of its sample complexity, is ongoing work.

### Acknowledgments

We thank Daniel Miller for discussions on algebraic geometry. This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – Cluster of Excellence Matter and Light for Quantum Computing (ML4Q) EXC 2004/1 – 390534769. The work of M.K. is also supported by the DFG via the Emmy Noether program (grant number 441423094) and by the German Federal Ministry of Education and Research (BMBF) within the funding program “quantum technologies – from basic research to market” in the joint project MIQRO (grant number 13N15522).

### Competing interests

The authors declare no competing interest.

### Appendix

In this appendix, we provide some auxiliary statements in order to keep this work largely self-contained. We also provide the proof of Lemma 6 which was omitted in the main text. Finally, we explain how the toric code example in Section 2.1 relates to the results derived by Spitz *et al.* [7].

## A Proof of Lemma 2

Inserting equation (14) into equation (12) yields

$$\begin{aligned} F(a) &= \prod_{c \subseteq a} \left( \prod_{\gamma \in \Gamma} E_\gamma(c \cap \gamma) \right)^{\mu(a,c)} = \prod_{\gamma \in \Gamma} \prod_{c \subseteq a} E_\gamma(c \cap \gamma)^{\mu(a,c)} \\ &= \prod_{\gamma \in \Gamma} \prod_{b \subseteq \gamma} \prod_{c \subseteq a: c \cap \gamma = b} E_\gamma(b)^{\mu(a,c)} = \prod_{\gamma \in \Gamma} \prod_{b \subseteq \gamma} E_\gamma(b)^{\sum_{c \subseteq a: c \cap \gamma = b} \mu(a,c)}. \end{aligned} \quad (47)$$

Writing  $c = b \cup r$  with  $r \cap \gamma = \emptyset$ , and using the fact that the sum is empty if  $b \not\subseteq a$ , we obtain

$$\begin{aligned} \sum_{c \subseteq a: c \cap \gamma = b} \mu(a,c) &= (-1)^{|a|} \sum_{c \subseteq a: c \cap \gamma = b} (-1)^{|c|} \\ &= (-1)^{|a|} (-1)^{|b|} 1[b \subseteq a] \sum_{c \subseteq a \setminus \gamma} (-1)^{|c|} = \mu(a,b) 1[b \subseteq a] \sum_{0 \leq a \leq |a \setminus \gamma|} \binom{|a \setminus \gamma|}{a} (-1)^a \\ &= \mu(a,b) 1[b \subseteq a] 1[a \setminus \gamma = \emptyset] = \mu(a,b) 1[b \subseteq a] 1[a \subseteq \gamma]. \end{aligned}$$

Substituting this identity into (47) yields the desired expression.

## B Equivalent condition on channel supports

Below Theorems 3 and 7, we stated two equivalent characterizations of detectable errors in terms of channel supports. The following lemma formalizes this statement. The proof is the same in both the classical and the quantum case.

**Lemma 9.** *Let  $\Gamma^{(D)}$  be as defined in equation (20),  $\Gamma \subseteq 2^{[n]}$  and let  $\hat{\Gamma} = \{e \subseteq \gamma : \gamma \in \Gamma\}$ . Then  $\gamma_1 \cup \gamma_2 \in \Gamma^{(D)}$  for all  $\gamma_1, \gamma_2 \in \Gamma$  if and only if  $\mathcal{O}(e_1) \neq 0$  and  $\mathcal{O}(e_1 + e_2) \neq 0$  for all  $e_1, e_2 \in \hat{\Gamma}$ .*

*Proof.* Note that by the definition of  $\Gamma^{(D)}$ , if  $\gamma \in \Gamma^{(D)}$  then  $e \in \Gamma^{(D)}$  for all  $e \subseteq \gamma$ . Thus,  $\gamma_1 \cup \gamma_2 \in \Gamma^{(D)}$  implies that  $e_1, e_2 \in \Gamma^{(D)}$  for all  $e_1 \subseteq \gamma_1$  and  $e_2 \subseteq \gamma_2$  and, in particular,  $\mathcal{O}(e_1) \neq 0$ . Furthermore, for any such  $e_1, e_2$ , we have that  $e_1 + e_2 \subseteq \gamma_1 \cup \gamma_2$ , since the addition as binary vectors corresponds to the symmetric difference of the indicator sets. This implies, in particular,  $\mathcal{O}(e_1 + e_2) \neq 0$ .

The other direction of the equivalence is proven similarly by noting that any subset of  $e \subseteq \gamma_1 \cup \gamma_2$  can be written as  $e_1 + e_2$  for appropriate choices of  $e_1 \subseteq \gamma_1$  and  $e_2 \subseteq \gamma_2$ , where it could be that either  $e_1$  or  $e_2$  is the empty set.  $\square$

## C Orthogonal array properties

In this section, we prove local randomness properties of stabilizer codes. Consider a quantum data-syndrome code with stabilizers  $\mathcal{S}_{\text{DS}}$  on  $n$  qubits and  $m$  measurement bits. Set  $N := 2n + m$ . By  $T$  we denote the  $|\mathcal{S}_{\text{DS}}| \times N$ -matrix formed by all elements in the span of the rows of the parity check matrix  $H_{\text{DS}}$ , i.e. the rows of  $T$  are exactly the elements of  $\mathcal{S}_{\text{DS}}$ . Then, the following local randomness property holds.

**Lemma 10.** *Let  $\gamma \in \Gamma^{(D)}$ . In the restriction  $T^{|\gamma|}$  of  $T$  to columns in  $\gamma$ , every bit-string appears equally often as a row.*

We adapt the proof of [36, Theorem 3.29] to our situation.

*Proof.* We denote the restriction of a matrix  $M$  to the columns indexed by  $\gamma \subseteq [n]$  by  $M^{|\gamma|}$ . By the definition (36) of a syndrome and the definition (equation (38)) of  $\Gamma^{(D)}$  it follows that for any row  $h$  of  $H_{\text{DS}}$  and any  $e \subseteq \gamma$ ,  $\langle h_i, e \rangle_{\text{DS}} = 0$  if and only if  $e = 0$  (as a binary vector). Since  $\langle h_i, e \rangle_{\text{DS}} = \bar{h}_i \cdot e$ , we conclude that the columns of  $H_{\text{DS}}^{|\gamma|}$  are linearly independent (over  $\mathbb{F}_2$ ), i.e.  $H_{\text{DS}}^{|\gamma|}$  is of rank  $|\gamma|$ . The rows of  $T^{|\gamma|}$  are, by definition of  $T$ , the vectors  $\zeta^T H_{\text{DS}}^{|\gamma|}$  for  $\zeta \in \mathbb{F}_2^l$ , where  $l$  is the number of rows of  $H$ . The number of times a bit-string  $z$  appears as a row in  $T$  is thus equal to the number of vectors  $\zeta \in \mathbb{F}_2^l$  with  $\zeta^T H^{|\gamma|} = z$ . Since  $H^{|\gamma|}$  has rank  $|\gamma|$ , this number is  $2^{l-|\gamma|}$  independent of  $z$ .  $\square$

Lemma 5 for classical codes is a direct corollary, which follows by only considering the classical (measurement) part of the data-syndrome code.

## D Analysis of the intersection matrix

In this appendix, we first summarize the definition and the most important properties of the Schur complement (see e.g. the book by Horn and Johnson [40]). On this basis we then prove Lemma 6 on the intersection matrix.

### D.1 Iterated Schur complements

For a block matrix

$$M_2 = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}, \quad (48)$$

with  $M_{11}$  being invertible, the *Schur complement* of  $M_{11}$  in  $M_2$  is defined as

$$M_2/M_{11} = M_{22} - M_{21}M_{11}^{-1}M_{12}. \quad (49)$$

If  $M_{11}$  is invertible, then  $M_2$  has full rank if and only if  $M_2/M_{11}$  has full rank. Furthermore, then  $M_2$  is positive definite if and only if  $M_{11}$  and  $M_2/M_{11}$  are positive definite. For the extended block matrix

$$M_3 = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix} \quad (50)$$

the upper left block of  $M_3/M_{11}$  is given by  $M_2/M_{11}$  and the following *quotient property of Schur complements* holds:

$$M_3/M_2 = (M_3/M_{11})/(M_2/M_{11}). \quad (51)$$

### D.2 The intersection matrix and proof of Lemma 6

In the Section 2.2.6, we have shown that the coefficient matrix  $D$  has full rank, which implies identifiability of the error distribution. The only thing missing is the proof of Lemma 6, which was omitted. We will now supply this proof, i.e. we show that the intersection matrix  $M_t$  defined in (23) is positive-definite for any  $t \in \mathbb{N}$ .

Ordering the indexing sets  $a, b \in \Gamma_{\leq t}$  by their size, for  $r \leq t$ , we can view  $M_r$  as a sub-matrix in the upper left corner of  $M_t$ .  $M_t$  is an  $(\alpha_t - 1) \times (\alpha_t - 1)$  matrix, where we define  $\alpha_t = \sum_{k=0}^t \binom{n}{k}$ . Note that in the case  $t = 1$ ,  $M$  takes a very simple form. Since two single-element sets are either equal or disjoint,  $M_1$  has entries 2 on the diagonal and 1 elsewhere, i.e.

$$M_1 = I + \mathbf{1}_1 \mathbf{1}_1^T, \quad (52)$$

where  $\mathbf{1}_t$  denotes the all ones vector of dimension  $\binom{n}{t}$ . This matrix is clearly positive-definite. In general,  $M_t$  does not admit such a simple expression. However, we will show that certain *Schur complements* of  $M_t$  can always be expressed in this simple form (see Appendix D.1 for the required background on Schur complements).

**Lemma 11.** *Let  $M_t$  be as in (23). Then the Schur complement  $M_t/M_{t-1}$  of  $M_{t-1}$  in  $M_t$  is well-defined and given by*

$$M_t/M_{t-1} = I + \frac{\mathbf{1}_t \mathbf{1}_t^T}{\alpha_{t-1}}, \quad (53)$$

where  $\mathbf{1}_t$  is the all ones vector of dimension  $\binom{n}{t}$  and  $\alpha_t = \sum_{k=0}^t \binom{n}{k}$ . For  $t = 1$  we set  $M_1/M_0 := M_1$ .

Lemma 6 is a simple corollary of this statement:

*Proof of Lemma 6.* Clearly,  $M_t/M_{t-1}$  is positive-definite. Since a matrix is positive-definite if the upper left block and its Schur complement are both positive-definite, it follows that  $M_t$  is positive-definite.  $\square$

For the proof of Lemma 11, some further preliminaries and notation are needed. We can view  $M_t$  as a block matrix consisting of blocks  $(M_{w,w'} : w, w' \in \{1, \dots, t\})$ , where  $M_{w,w'}$  labels the block whose rows are indexed by sets of size  $w$  and whose columns are indexed by sets of size  $w'$ . By repeatedly applying the quotient property of Schur complements ((51) in the Appendix), we can express  $M_t/M_{t-1}$  as a chain of Schur complements:

$$\begin{aligned} M_t/M_{t-1} &= (M_t/M_{t-2})/(M_{t-1}/M_{t-2}) \\ &= [(M_t/M_{t-3})/(M_{t-2}/M_{t-3})]/(M_{t-1}/M_{t-2}) \\ &= [([(M_t/M_1)/(M_2/M_1)]/(M_3/M_1)] \dots]/(M_{t-1}/M_{t-2}). \end{aligned}$$

This chain is described by the recursive definitions

$$M_t^{(0)} = M_t, \quad (54)$$

$$M_t^{(1)} = (M_t/M_1), \quad (55)$$

$$M_t^{(i)} = M_t^{(i-1)}/(M_i/M_{i-1}). \quad (56)$$

We use the same labeling of blocks for  $M_t^{(i)}$  as for  $M_t$ , i.e.  $M_{w,w'}^{(i)}$  is the block of  $M_t^{(i)}$  whose rows are labeled by sets of weight  $w$  and whose columns are labeled by sets of weight  $w'$ , i.e.

$$M_{w,w'}^{(i)} = [M_t^{(i-1)}]_{w,w'}. \quad (57)$$

Of course, this is only well-defined if such a block actually exists in  $M_t^{(i)}$ , i.e. if  $w, w' > i$ . In this case neither the entries nor the dimensions of  $M_{w,w'}^{(i)}$  depend on  $t$ , so the index  $t$  is indeed not needed.

The following useful observations follow directly from the properties of the Schur complement:

**Lemma 12.** *The matrix  $M_t$  defined in (23) and the matrices  $M_t^{(i)}$  defined in (56) have the following properties:*

1.  $M_t^{(i)} = M_t/M_i$
2.  $M_i/M_{i-1} = M_{i,i}^{(i-1)}$
3.  $M_{w,i}^{(i-1)} = (M_{i,w}^{(i-1)})^T$
4. For any  $w, w' > i$ :  

$$M_{w,w'}^{(i)} = M_{w,w'}^{(i-1)} - M_{w,i}^{(i-1)}(M_i/M_{i-1})^{-1}M_{i,w'}^{(i-1)}$$

*Proof of Lemma 12.*

1. For  $i = 0, 1$  the statement holds by definition (interpreting  $M_t^{(i)}/M_0 = M_t^{(i)}$ ). The statement then follows by induction, using the quotient property (51).
2. We denote with  $M_{>i-1}$  the sub-matrix of  $M_t$  whose rows and columns are labeled by sets of size  $> i - 1$ . Furthermore,  $M_{>i-1, i-1}$  denotes the sub-matrix whose rows are labeled by sets of size  $> i - 1$  and whose columns are labeled with sets of size  $i - 1$ . The matrix  $M_{i-1, >i-1}$  is defined analogously and for any matrix we use a subscript  $_{i,i}$  to denote the block whose rows and columns are labeled by sets of weight  $i$ . Remember the notation  $M_{i,i}^{(i-1)} := [M_t^{(i-1)}]_{i,i}$ . Then,

$$\begin{aligned} M_{i,i}^{(i-1)} &= [M_t/M_{i-1}]_{i,i} \\ &= [M_{>i-1} - M_{>i-1, i-1}M_{i-1}^{-1}M_{i-1, >i-1}]_{i,i} \\ &= M_{i,i} - [M_{>i-1, i-1}M_{i-1}^{-1}M_{i-1, >i-1}]_{i,i} \\ &= M_{i,i} - M_{i, i-1}M_{i-1}^{-1}M_{i-1, i} \\ &= M_i/M_{i-1}. \end{aligned}$$

3. This statement holds because  $M_t$  is symmetric.
4. The calculation is analogous to the one above. Using property 1 and the definition of Schur complements (49), we obtain

$$\begin{aligned} M_{w,w'}^{(i)} &= [M_{>i}^{(i-1)} - M_{>i,i}^{(i-1)}(M_{i,i}^{(i-1)})^{-1}M_{i,>i}^{(i-1)}]_{w,w'} \\ &= M_{w,w'}^{(i)} - M_{w,i}^{(i-1)}(M_{i,i}^{(i-1)})^{-1}M_{i,w'}^{(i-1)}. \end{aligned}$$

We substitute  $M_{i,i}^{(i-1)}$  with  $M_i/M_{i-1}$ , since we use this form later. □

Lemma 11 is now a direct consequence of the following more explicit statement.

**Lemma 13.**  $M_t^{(i)}$  is given entry-wise by

$$M_t^{(i)}[a, b] = f^{(i+1)}(|a \cap b|) + u^{(i)}(|a|, |b|), \quad (58)$$

where

$$f^{(i)}(x) = \sum_{k=i}^x \binom{x}{k} = 2^x - \sum_{k=0}^{i-1} \binom{x}{k} \quad (59)$$

$$u^{(i)}(w, w') = \frac{\binom{w-1}{i} \binom{w'-1}{i}}{\alpha_i} \quad (60)$$

$$\alpha_i = \sum_{k=0}^i \binom{n}{k} \quad (61)$$

To recover Lemma 11, notice that  $f^{(i)}(x) = 1[x = i]$  if  $x \leq i$  and  $u^{(i-1)}(i, i) = \frac{1}{\alpha_{i-1}}$ . Thus we obtain

$$\begin{aligned} (M_i/M_{i-1})[a, b] &= M_{i,i}^{(i-1)}[a, b] \\ &= 1[|a \cap b| = i] + u^{(i-1)}(i, i) \\ &= 1[a = b] + \frac{1}{\alpha_{i-1}}, \end{aligned} \quad (62)$$

where we used that  $|a| = |b| = i$  because we are considering the block indexed by  $i, i$ . Choosing  $i = t$  yields Lemma 11.

*Proof of Lemma 13.* We compute  $M_t^{(i)}$  block-wise. Remember that the rows of  $M_{w,w'}^{(i-1)}$  are indexed by subsets of  $[n]$  of weight  $w$  and the columns by subsets of  $[n]$  of size  $w'$ , and that  $M_t^{(i-1)} = M_t/M_{i-1}$  only contains blocks  $M_{w,w'}^{(i-1)}$  with  $w, w' \geq i$ .

We prove the lemma by induction. The base case  $i = 0$  is follows directly as

$$M_{w,w'}^{(0)}[a, b] = M_{w,w'}[a, b] = 2^{|a \cap b|} = 2^{|a \cap b|} - 1 + 1 = f^{(1)}(|a \cap b|) + u^{(0)}(w, w'). \quad (63)$$

For the induction step we assume that (58) holds for  $i - 1$  and we will prove it for  $i$ . Using the induction hypothesis, we apply (62) to obtain

$$M_i/M_{i-1} = M_{i,i}^{(i-1)} = I + \frac{\mathbf{1}_i \mathbf{1}_i^T}{\alpha_{i-1}}, \quad (64)$$

where  $\mathbf{1}_i$  is again the all ones vector of length  $\binom{n}{i}$ . The inverse of this matrix can be computed via the Sherman-Morrison formula to be

$$(M_i/M_{i-1})^{-1} = I - \frac{\mathbf{1}_i \mathbf{1}_i^T}{\alpha_{i-1} + \mathbf{1}_i^T \mathbf{1}_i} = I - \frac{\mathbf{1}_i \mathbf{1}_i^T}{\alpha_{i-1} + \binom{n}{i}} = I - \frac{\mathbf{1}_i \mathbf{1}_i^T}{\alpha_i}. \quad (65)$$

Combining this expression with property 4 in Lemma 12, we obtain

$$M_{w,w'}^{(i)} = M_{w,w'}^{(i-1)} - M_{w,i}^{(i-1)}M_{i,w'}^{(i-1)} + \frac{M_{w,i}^{(i-1)}\mathbf{1}_i\mathbf{1}_i^T M_{i,w'}^{(i-1)}}{\alpha_i}. \quad (66)$$

Calculating this term by term, starting with the last term, we obtain

$$\begin{aligned} (\mathbf{1}_i^T M_{i,w'}^{(i-1)})[a] &= \sum_{|c|=i} M_{i,w'}^{(i-1)}[c, a] = \sum_{|c|=i} (f^{(i)}(|a \cap c|) + u^{(i-1)}(i, w')) \\ &= \sum_{|c|=i} (1[c \subseteq a] + u^{(i-1)}(i, w')) = \binom{w'}{i} + \binom{n}{i} u^{(i-1)}(i, w'), \end{aligned}$$

where  $\sum_{|c|=i}$  indicates a sum over all subsets of  $[n]$  of size  $i$ . The third equality used that

$$f^{(i)}(|a \cap c|) = 1[|a \cap c| = i] = 1[c \subseteq a] \quad (67)$$

since  $|c| = i$  and the last equality used  $|a| = w'$ . Defining

$$S^{(i)}(w') = \binom{w'}{i} + \binom{n}{i} u^{(i-1)}(i, w') \quad (68)$$

and using the symmetry of  $M_{i,w}^{(i-1)}$  we obtain

$$M_{w,i}^{(i-1)}\mathbf{1}_i\mathbf{1}_i^T M_{i,w'}^{(i-1)} = S^{(i)}(w)S^{(i)}(w')\mathbf{1}_i\mathbf{1}_i^T. \quad (69)$$

Next, we consider the second term of (66),

$$\begin{aligned} (M_{w,i}^{(i-1)}M_{i,w'}^{(i-1)})[a, b] &= \sum_{|c|=i} M_{w,i}^{(i-1)}[a, c]M_{i,w'}^{(i-1)}[c, b] \\ &= \sum_{|c|=i} [f^{(i)}(|a \cap c|) + u^{(i-1)}(i, w)][f^{(i)}(|c \cap b|) + u^{(i-1)}(i, w')] \\ &= \sum_{|c|=i} 1[c \subseteq a]1[c \subseteq b] + u^{(i-1)}(i, w) \sum_{|c|=i} 1[c \subseteq b] \\ &\quad + u^{(i-1)}(i, w') \sum_{|c|=i} 1[c \subseteq a] + \sum_{|c|=i} u^{(i-1)}(i, w)u^{(i-1)}(i, w') \\ &= \binom{|a \cap b|}{i} + u^{(i-1)}(i, w) \binom{w'}{i} + u^{(i-1)}(i, w') \binom{w}{i} + \binom{n}{i} u^{(i-1)}(i, w)u^{(i-1)}(i, w') \\ &= \binom{|a \cap b|}{i} + T^{(i)}(w, w'), \end{aligned} \quad (70)$$

where we defined

$$T^{(i)}(w, w') = u^{(i-1)}(i, w) \binom{w'}{i} + u^{(i-1)}(i, w') \binom{w}{i} + \binom{n}{i} u^{(i-1)}(i, w)u^{(i-1)}(i, w'). \quad (71)$$

Thus, combining (69), (70), and the induction assumption (58) with (66) we obtain

$$\begin{aligned} M_{w,w'}^{(i)}[a, b] &= f^{(i)}(|a \cap b|) + u^{(i-1)}(w, w') - \binom{|a \cap b|}{i} - T^{(i)}(w, w') + \frac{S^{(i)}(w)S^{(i)}(w')}{\alpha_i} \\ &= f^{(i+1)}(|a \cap b|) + \frac{S^{(i)}(w)S^{(i)}(w') + \alpha_i(u^{(i-1)}(w, w') - T^{(i)}(w, w'))}{\alpha_i}, \end{aligned} \quad (72)$$

where we used  $f^{(i+1)}(|a \cap b|) = f^{(i)}(|a \cap b|) - \binom{|a \cap b|}{i}$ . Finally, we calculate the last term. By direct calculation from the definitions (71), (68) we obtain

$$S^{(i)}(w)S^{(i)}(w') - \binom{n}{i}T^{(i)}(w, w') = \binom{w}{i} \binom{w'}{i}, \quad (73)$$

and thus,

$$\begin{aligned} & S^{(i)}(w)S^{(i)}(w') + \alpha_i(u^{(i-1)}(w, w') - T^{(i)}(w, w')) \\ &= S^{(i)}(w)S^{(i)}(w') - \binom{n}{i}T^{(i)}(w, w') + \alpha_{i-1}u^{(i-1)}(w, w') + \binom{n}{i}u^{(i-1)}(w, w') - \alpha_{i-1}T^{(i)}(w, w'), \end{aligned} \quad (74)$$

where we have used that  $\alpha_i = \binom{n}{i} + \alpha_{i-1}$ . Inserting the definitions of  $u^{(i)}$  and  $T^{(i)}$  from (60) and (71), we obtain  $\alpha_{i-1}u^{(i)}(w, w') = \binom{w-1}{i-1}\binom{w'-1}{i-1}$  and  $\binom{n}{i}u^{(i-1)}(w, w') - \alpha_{i-1}T^{(i)}(w, w') = -\binom{w-1}{i-1}\binom{w'}{i} - \binom{w'-1}{i-1}\binom{w}{i}$ . Together with (73) this yields

$$\begin{aligned} (74) &= \binom{w}{i}\binom{w'}{i} + \binom{w-1}{i-1}\binom{w'-1}{i-1} - \binom{w-1}{i-1}\binom{w'}{i} - \binom{w'-1}{i-1}\binom{w}{i} \\ &= \left[ \binom{w}{i} - \binom{w-1}{i-1} \right] \left[ \binom{w'}{i} - \binom{w'-1}{i-1} \right] \\ &= \binom{w-1}{i}\binom{w'-1}{i}, \end{aligned}$$

where the last step is based on Pascals identity. Substituting this result back into (72) we obtain

$$M_{w,w'}^{(i)}[a, b] = f^{(i+1)}(|a \cap b|) + \frac{\binom{w-1}{i}\binom{w'-1}{i}}{\alpha_i}, \quad (75)$$

which finishes the proof.  $\square$

## E Connection to Adaptive Weight Estimator

Here, we show that the solution for the toric code derived in Section 2.1 coincides with the solution given by Spitz *et al.* [7].

The solution derived in Section 2.1 is

$$E(Z_4) = \pm \sqrt{\frac{E(S_1)E(S_2)}{E(S_1S_2)}}. \quad (76)$$

On the other hand, the solution given by Spitz *et al.* [7, eq. (14)] is

$$p_4 = \frac{1}{2} \mp \sqrt{\frac{1}{4} - \frac{P(S_1 = S_2 = 1) - P(S_1 = 1)P(S_2 = 1)}{1 - 2P(S_1 \neq S_2)}}. \quad (77)$$

Here,  $P$  is used to denote the probabilities of events under random sampling of the errors. Note that in [7] the result is phrased in terms of expectation values of the stabilizer outcomes and errors viewed as taking values 0 or 1, which directly translates into the probabilities above.

To see that these two solutions coincide, first notice that  $E(Z_4) = 1 - 2p_4$ . Thus, (76) can be rewritten as

$$p_4 = \frac{1}{2} \mp \sqrt{\frac{E(S_1)E(S_2)}{4E(S_1S_2)}}. \quad (78)$$

Similarly, we have  $E(S_i) = 1 - 2P(S_i = 1)$  and  $E(S_1S_2) = 1 - 2P(S_1 \neq S_2)$ . Using these equations, equality of the solutions can be shown as follows:

$$\begin{aligned} \frac{E(S_1)E(S_2)}{4E(S_1S_2)} &= \frac{(1 - 2P(S_1 = 1))(1 - 2P(S_2 = 1))}{4(1 - 2P(S_1 \neq S_2))} \\ &= \frac{1 - 2P(S_1 = 1) - 2P(S_2 = 1) + 4P(S_1 = 1)P(S_2 = 1) + 1 - 2P(S_1 \neq S_2) - 1 + 2P(S_1 \neq S_2)}{4(1 - 2P(S_1 \neq S_2))} \\ &= \frac{1}{4} - \frac{2P(S_1 = 1) + 2P(S_2 = 1) - 4P(S_1 = 1)P(S_2 = 1) - 2P(S_1 \neq S_2)}{4(1 - 2P(S_1 \neq S_2))} \\ &= \frac{1}{4} - \frac{P(S_1 = S_2 = 1) - P(S_1 = 1)P(S_2 = 1)}{1 - 2P(S_1 \neq S_2)}, \end{aligned}$$

where in the last equality we used that

$$\begin{aligned}
& P(S_1 = 1) + P(S_2 = 1) - P(S_1 \neq S_2) \\
&= P(S_1 = 1, S_2 = 1) + P(S_1 = 1, S_2 = 0) \\
&\quad + P(S_1 = 1, S_2 = 1) + P(S_1 = 0, S_2 = 1) \\
&\quad - P(S_1 = 1, S_2 = 0) - P(S_1 = 0, S_2 = 1) \\
&= 2P(S_1 = S_2 = 1).
\end{aligned}$$

This shows that the two solutions (76) and (77) are indeed equivalent.

## References

- [1] A. Robertson, C. Granade, S. D. Bartlett, and S. T. Flammia, *Tailored codes for small quantum memories*, *Phys. Rev. Applied* **8**, 064004 (2017).
- [2] J. Florjanczyk and T. A. Brun, *In-situ adaptive encoding for asymmetric quantum error correcting codes* (2016).
- [3] J. P. Bonilla Ataides, D. K. Tuckett, S. D. Bartlett, S. T. Flammia, and B. J. Brown, *The XZZX surface code*, *Nat. Commun.* **12**, 2172 (2021).
- [4] O. Higgott, *Pymatching: A python package for decoding quantum codes with minimum-weight perfect matching* (2021).
- [5] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, *Topological quantum memory*, *J. Math. Phys.* **43**, 4452 (2002), arXiv:quant-ph/0110143 [quant-ph].
- [6] N. H. Nickerson and B. J. Brown, *Analysing correlated noise on the surface code using adaptive decoding algorithms*, *Quantum* **3**, 131 (2019).
- [7] S. T. Spitz, B. Tarasinski, C. W. J. Beenakker, and T. E. O’Brien, *Adaptive weight estimator for quantum error correction in a time-dependent environment*, *Advanced Quantum Technologies* **1**, 1870015 (2018).
- [8] Z. Babar, P. Botsinis, D. Alanis, S. X. Ng, and L. Hanzo, *Fifteen years of quantum LDPC coding and improved decoding strategies*, *IEEE Access* **3**, 2492 (2015).
- [9] S. Huang, M. Newman, and K. R. Brown, *Fault-tolerant weighted union-find decoding on the toric code*, *Physical Review A* **102**, 10.1103/physreva.102.012419 (2020).
- [10] C. T. Chubb, *General tensor network decoding of 2d pauli codes* (2021).
- [11] A. S. Darmawan and D. Poulin, *Linear-time general decoding algorithm for the surface code*, *Physical Review E* **97**, 10.1103/physreve.97.051302 (2018).
- [12] J. J. Wallman and J. Emerson, *Noise tailoring for scalable quantum computation via randomized compiling*, *Phys. Rev. A* **94**, 052325 (2016).
- [13] M. Ware, G. Ribeill, D. Ristè, C. A. Ryan, B. Johnson, and M. P. da Silva, *Experimental Pauli-frame randomization on a superconducting qubit*, *Phys. Rev. A* **103**, 042604 (2021).
- [14] S. J. Beale, J. J. Wallman, M. Gutiérrez, K. R. Brown, and R. Laflamme, *Quantum error correction decoheres noise*, *Phys. Rev. Lett.* **121**, 190501 (2018).
- [15] S. T. Flammia and R. O’Donnell, *Pauli error estimation via population recovery*, *Quantum* **5**, 549 (2021).
- [16] R. Harper, W. Yu, and S. T. Flammia, *Fast estimation of sparse quantum noise*, *PRX Quantum* **2**, 010322 (2021).
- [17] S. T. Flammia and J. J. Wallman, *Efficient estimation of Pauli channels*, *ACM Transactions on Quantum Computing* **1**, 10.1145/3408039 (2020).
- [18] R. Harper, S. T. Flammia, and J. J. Wallman, *Efficient learning of quantum noise*, *Nat. Phys.* **16**, 1184 (2020).
- [19] Y. Fujiwara, *Instantaneous quantum channel estimation during quantum information processing* (2014).
- [20] A. G. Fowler, D. Sank, J. Kelly, R. Barends, and J. M. Martinis, *Scalable extraction of error models from the output of error detection circuits* (2014).
- [21] M.-X. Huo and Y. Li, *Learning time-dependent noise to reduce logical errors: real time error rate estimation in quantum error correction*, *New J. Phys.* **19**, 123032 (2017).

- [22] J. R. Wootton, *Benchmarking near-term devices with quantum error correction*, *Quantum Science and Technology* **5**, 044004 (2020).
- [23] J. Combes, C. Ferrie, C. Cesare, M. Tiersch, G. J. Milburn, H. J. Briegel, and C. M. Caves, *In-situ characterization of quantum devices with error correction* (2014).
- [24] T. Wagner, H. Kampermann, D. Bruß, and M. Kliesch, *Optimal noise estimation from syndrome statistics of quantum codes*, *Phys. Rev. Research* **3**, 013292 (2021).
- [25] J. Kelly, R. Barends, A. G. Fowler, A. Megrant, E. Jeffrey, T. C. White, D. Sank, J. Y. Mutus, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, E. Lucero, M. Neeley, C. Neill, P. J. J. O’Malley, C. Quintana, P. Roushan, A. Vainsencher, J. Wenner, and J. M. Martinis, *Scalable in situ qubit calibration during repetitive error detection*, *Phys. Rev. A* **94**, 032321 (2016).
- [26] A. Ashikhmin, C.-Y. Lai, and T. A. Brun, *Quantum data-syndrome codes*, *IEEE Journal on Selected Areas in Communications* **38**, 449 (2020).
- [27] Y. Fujiwara, *Ability of stabilizer quantum error correction to protect itself from its own imperfection*, *Phys. Rev. A* **90**, 062304 (2014), arXiv:1409.2559 [quant-ph].
- [28] N. Delfosse, B. W. Reichardt, and K. M. Svore, *Beyond single-shot fault-tolerant quantum error correction*, *IEEE Transactions on Information Theory* **68**, 287 (2022).
- [29] A. Zia, J. P. Reilly, and S. Shirani, *Distributed parameter estimation with side information: A factor graph approach*, in *2007 IEEE International Symposium on Information Theory* (2007) pp. 2556–2560.
- [30] R. O’Donnell, *Analysis of Boolean Functions* (Cambridge University Press, 2014).
- [31] Y. Mao and F. Kschischang, *On factor graphs and the fourier transform*, *IEEE Trans. Inf. Theory* **51**, 1635 (2005).
- [32] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques - Adaptive Computation and Machine Learning* (The MIT Press, 2009).
- [33] M. Aigner, *A Course in Enumeration*, Vol. 238 (Springer-Verlag Berlin Heidelberg, 2007).
- [34] S. Roman, *Field Theory* (Springer, New York, 2006).
- [35] T. Chen and LiTien-Yien, *Solutions to systems of binomial equations*, *Annales Mathematicae Silesianae* **28**, 7 (2014).
- [36] A. S. Hedayat, N. J. A. Sloane, and J. Stufken, *Orthogonal arrays: theory and applications* (Springer New York, NY, 1999).
- [37] P. Delsarte, *Four fundamental parameters of a code and their combinatorial significance*, *Information and Control* **23**, 407 (1973).
- [38] B. M. Varbanov, F. Battistel, B. M. Tarasinski, V. P. Ostroukh, T. E. O’Brien, L. DiCarlo, and B. M. Terhal, *Leakage detection for a transmon-based surface code*, *NPJ Quantum Inf.* **6**, 10.1038/s41534-020-00330-w (2020).
- [39] P. Abbeel, D. Koller, and A. Y. Ng, *Learning factor graphs in polynomial time & sample complexity* (2012).
- [40] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. (Cambridge University Press, 2012).