On the Logical Error Rate of Sparse Quantum odes

PATRICIO FUENTES[1] , **JOSU ETXEZARRETA MARTINEZ**[1] ,
PEDRO M. CRESPO[1] Senior Member, IEEE),
AND JAVIER GARCIA-FRÍAS[2] Senior Member, IEEE)

[1]Department of Biomedical Engineering and Sciences, Tecnun - University of Navarra, 20018 Donostia-San Sebastian, Spain
[2]Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716 USA

Corresponding author: Patricio Fuentes (e-mail: pfuentesu@tecnun.es).

ABSTRACT The quantum paradigm presents a phenomenon known as degeneracy that can potentially improve the performance of quantum error correcting codes. However, the effects of this mechanism are sometimes ignored when evaluating the performance of sparse quantum codes and the logical error rate is not always correctly reported. In this article, we discuss previously existing methods to compute the logical error rate and we present an efficient coset-based method inspired by classical coding strategies to estimate degenerate errors and distinguish them from logical errors. Additionally, we show that the proposed method presents a computational advantage for the family of Calderbank–Shor–Steane codes. We use this method to prove that degenerate errors are frequent in a specific family of sparse quantum codes, which stresses the importance of accurately reporting their performance. Our results also reveal that the modified decoding strategies proposed in the literature are an important tool to improve the performance of sparse quantum codes.

INDEX TERMS Iterative decoding, quantum error correction (QEC), quantum low density generator matrix codes, quantum low-density parity check (QLDPC) codes.

I INTRODUCTION

When quantum stabilizer codes built from sparse classical codes are employed in the quantum paradigm, they are impacted by a phenomenon known as degeneracy [1], [2], [3], [4], which has no classical equivalent. This causes stabilizer codes to exhibit a particular coset structure in which multiple different error patterns act identically on the transmitted information [5], [6], [7]. The manifestation of degeneracy in the design of sparse quantum codes and its effects on the decoding process has been studied extensively [3], [4], [8], [9], [10], [11], [12], [13], [14]. Unfortunately, although degeneracy may potentially improve performance, limited research exists on how to quantify the true impact that this phenomenon has on quantum low-density parity check (QLDPC) codes. This has resulted in the performance of QLDPC codes being assessed differently throughout the literature; while some research considers the effects of degeneracy by computing the metric known as the *logical error rate* [3], [15], [16], [17], [18], [19], [20], other works employ the classical strategy of computing the *physical error rate* [21], [22], [23], [24], [25], [26], a metric which provides an upper bound on the logical error rate of these codes since it ignores degeneracy. In the context of degenerate quantum codes, the discrepancy between results computed based on the physical error rate and the logical error rate can become significant. QLDPC codes (also commonly referred to as sparse quantum codes) are famous for their degenerate nature [3], [4], [12], [14], [27], [28], which makes it important to accurately assess their performance using the logical error rate.

The logical error rate is a well-known metric and has been widely employed to assess the performance of other families

of quantum error correction (QEC) codes such as quantum turbo codes (QTC) and quantum topological codes [10], [35], [36], [37], [38]. For these particular error correction schemes, the decoders that are employed are sometimes capable of distinguishing between error equivalence classes (they can account for some aspects of degeneracy), which makes it easy to compute the logical error rate. In contrast, the classical decoding algorithms [39], [40] that are used to decode sparse quantum codes are unable to account for the presence of degenerate error operators, and so their logical error rate must be computed differently.

Given the coset structure of sparse quantum codes, intuition would point toward approaching the issue of calculating the logical error rate by finding and comparing the stabilizer cosets of the estimated error sequences and the stabilizer cosets of the channel errors. Unfortunately, the task of computing stabilizer cosets has been shown to be computationally hard [4], [12], [24], which is the reason why the performance of some sparse quantum codes [21], [22], [23], [24], [25], [26] has been assessed based on the physical error rate. This metric is computed by comparing the error sequence estimated by the decoder, $\hat{E} \in \overline{\mathcal{G}}_N$, to the channel error, $E \in \overline{\mathcal{G}}_N$, where $\overline{\mathcal{G}}_N$ denotes the N-fold effective Pauli group. Essentially, if the estimation matches the channel error, the decoder has been successful, and, if not, a decoding failure has occurred. Note, however, that because this metric ignores the degenerate nature of stabilizer codes, the physical error rate overestimates the number of decoding failures and actually represents an upper bound on the performance of stabilizer codes.

Despite the use of the physical error rate in some works, other literature has successfully computed the logical error rate of specific sparse quantum codes [3], [15], [16], [17], [18], [19], [20]. These works succeed in computing the logical error rate because they do not approach the issue from the perspective of stabilizer cosets. Instead, in most of these works [3], [15], [16], [17], [18], Gaussian elimination is used to obtain the parity check matrix (PCM) of the code in what is known as *standard form* [8], [41], [42], and then this matrix is used to extract a basis for the encoded Pauli operators of the corresponding codes. Then, this basis can be employed to distinguish between degenerate errors and logical errors. Additionally, other research [20] employs a different, albeit much more computationally demanding, method to compute the logical error rate.

Against this backdrop, in this article, we document the methods that have previously been used in the literature to compute the logical error rate of sparse quantum codes. In addition, we introduce our own group theoretic strategy to accurately assess the effects of degeneracy on sparse quantum codes and compute their logical error rate. We also show how for sparse Calderbank–Shor–Steane (CSS) quantum codes the method we propose herein has a computational complexity advantage over those that have been employed previously. Following this, we use our strategy to analyze the frequency with which degenerate errors occur when using a specific family of sparse quantum codes and we provide

insight on how the design and decoding of these codes can be improved.

II DEGENERACY AND STABILIZER CODES

Throughout this article, we assume that the reader is familiar with basic QEC concepts based on physical two-level systems (qubits) such as stabilizer codes, their coset structure, and the different types of possible errors. This section is meant as an introduction to the notation we will use (that of [14]) and to provide a brief summary of some of the ideas of [14]. For a complete overview on these preliminaries, refer to [14], [27], [29].

A. COSET STRUCTURE OF THE PAULI GROUP

The coherence loss suffered by quantum information over time can be approximated as the action of operators that belong to the Pauli Group on the qubits that store this information. In most cases, the action of these operators is modeled by means of an abstraction known as the Pauli channel. For single qubit quantum states, the Pauli channel will act on said states by combining them with an element of the single-qubit Pauli group $\mathcal{G}_1 = (\tilde{\Pi}, \times)$. The set $\tilde{\Pi}$ is given by

$$\tilde{\Pi} = \{\Theta_1 I, \Theta_2 X, \Theta_3 Y, \Theta_4 Z\}$$

$$= \{\pm I, \pm iI, \pm X, \pm iX, \pm Y, \pm iY, \pm Z, \pm iZ\}$$

where $\Theta_l = \{\pm 1, \pm i\}$, $\{X, Y, Z\}$ are the Pauli matrices, I is the 2×2 identity matrix, and $l = 1, 2, 3, 4$. The group operation $\times$ represents the well-known product of Pauli matrices, which is written as

$$\sigma_a \times \sigma_b = I\delta_{a,b} + i\sum_{c=1}^{3} \epsilon_{abc}\sigma_c \qquad (1)$$

where δ represents the Kronecker delta and

$$\epsilon_{abc} = \begin{cases} 1 & \text{if } (a, b, c) = \{(1, 2, 3), (3, 1, 2), (2, 3, 1)\}, \\ -1 & \text{if } (a, b, c) = \{(3, 2, 1), (1, 3, 2), (2, 1, 3)\}, \\ 0 & \text{if } a = b \text{ or } b = c \text{ or } a = c. \end{cases}$$

In (1), we have denoted the X Pauli matrix as σ_1, the Y Pauli matrix as σ_2, and the Z Pauli matrix as σ_3.

We can generalize the single qubit Pauli group so that it can be applied to N-qubit quantum states based on the tensor product. The N-qubit Pauli group is by $\mathcal{G}_N = (\tilde{\Pi}^{\otimes N}, \,)$. The set $\tilde{\Pi}^{\otimes N}$ is given by

$$\tilde{\Pi}^{\otimes N} = \{\theta_1 I, \theta_2 X, \theta_3 Y, \theta_4 Z\}^{\otimes N}$$

where $\theta_k = \{\pm 1, \pm i\}$, and $\otimes$ denotes the tensor product. For all $A, B \in \mathcal{G}_N$, the group operation is defined as

$$A \quad B = A_1 \times B_1 \otimes A_2 \times A_2 \otimes \ldots \otimes A_N \times B_N. \qquad (2)$$

Because it makes good physical sense to neglect the global phase of Pauli operators [32], i.e., disregard the overall factors $\{\pm 1, \pm i\}$, it is typical in the field of QEC to work with a reduced version of the Pauli group known as the *effective Pauli group*. The N-qubit effective Pauli group is denoted by $\overline{\mathcal{G}}_N = (\Pi^{\otimes N}, \star)$, where

$$\Pi^{\otimes N} = \{I, X, Y, Z\}^{\otimes N}$$

and $\star$ behaves as in (2) but with the operation between single Pauli operator products defined not as in (1) but as

$$\sigma_a \star \sigma_b = I\delta_{a,b} + \sum_{c=1}^{3} |\epsilon_{abc}|\sigma_c \quad (3)$$

where ϵ_{abc} is the same as in (1).

Thus, against this backdrop, efficient QEC strategies for general Pauli channels are defined as those methods that are capable of reverting the action of elements that belong to the effective Pauli group. It is important to note, however, that $\overline{\mathcal{G}}_N$ is abelian, whereas $\mathcal{G}_N$ is not, which means that the commutation relations that exist between Pauli operators [see (2)] are lost between effective Pauli operators [see (3)]. Given the importance of these commutation relations for the purposes of error correction [14], [29], [30], it is important that they be recovered. For this purpose, we define the symplectic map $\quad: \Pi^{\otimes N} \to \quad_2^{2N}$, which is an isomorphism between the group $\overline{\mathcal{G}}_N = (\Pi^{\otimes N}, \star)$ and the group $(\quad_2^{2N}, \oplus)$ of $2N$ binary-tuples under the mod 2 sum operation. For clarity, throughout the remainder of this article, lower case boldface romans without a subscript will be used to denote $2N$ binary-tuples that belong to $\quad_2^{2N}$, and lower case boldface romans with a subscript will be used to denote N binary-tuples that belong to $\quad_2^N$. The symplectic mapping is defined as

$$(\mathbf{A}) = \mathbf{a} = (\mathbf{a}_x | \mathbf{a}_z), \quad \mathbf{a}_x, \mathbf{a}_z \in \quad_2^N$$

where the values of the entries of $\mathbf{a}_x$ and $\mathbf{a}_z$ at position $i = 1, \ldots, N$ are directly dependent on the single qubit Pauli operator $[\mathbf{A}]_i$ located at the ith position in the tensor product that makes up $\mathbf{A} \in \overline{\mathcal{G}}_N$.

Note that the commutation properties of the Pauli operators with regard to the product are not recovered by just defining the isomorphism [after all, $(\quad_2^{2N}, \oplus)$ is also an abelian group]. Thus, we define the symplectic product $\mathbf{a} \odot \mathbf{b} \in \quad_2 \; \forall \, \mathbf{a}, \mathbf{b} \in \quad_2^{2N}$ as

$$\mathbf{a} \odot \mathbf{b} \quad (\mathbf{a}_x \circledast \mathbf{b}_z) \oplus (\mathbf{a}_z \circledast \mathbf{b}_x) \quad (4)$$

where $\circledast$ is the standard mod 2 inner product defined on $(\quad_2^N, \oplus)$ considered as a vector space over the field $\quad_2$. The symplectic product tells us if any two operators $\mathbf{A}, \mathbf{B}$ in $\overline{\mathcal{G}}_N$ will either commute or anticommute with regard to the group operation in $\mathcal{G}_N$, if and only if the symplectic product between $(\mathbf{A}) = \mathbf{a}$ and $(\mathbf{B}) = \mathbf{b}$, i.e., $\mathbf{a} \odot \mathbf{b}$, takes the value 0 or 1, respectively [31], [32].

B. STABILIZER CODES

Stabilizer codes are a widely extended family of QEC codes. They are constructed based on a set of generators that define the so-called stabilizer group $\overline{\mathcal{S}} \subset \overline{\mathcal{G}}_N$. The stabilizer generators are a set of operators $\{\mathbf{S} \}_{=1}^{N-k}$ that belong to $\overline{\mathcal{G}}_N$ and whose symplectic representatives commute with regard to the symplectic product. This means that the stabilizer is composed of elements of the effective Pauli group that commute amongst themselves with regard to in $\mathcal{G}_N$. The stabilizer $\overline{\mathcal{S}}$ has 2^{N-k} distinct elements and is completely represented by

its $N - k$ independent generators (the rest of the stabilizer elements are combinations of the elements of the minimal set). The codespace defined by the stabilizer, and hence a stabilizer code, can be described mathematically as

$$(\overline{\mathcal{S}}) = \quad |\psi\rangle \in \mathcal{H}_2^{\otimes N} : \mathbf{S}_i|\psi\rangle = |\psi\rangle, i = 1 \ldots N - k\} \quad (5)$$

where $\mathbf{S}_i|\psi\rangle \in \mathcal{H}_2^{\otimes N}$ is the evolution of state $|\psi\rangle$ under stabilizer generator $\mathbf{S}_i$. Note that $(\overline{\mathcal{S}})$ is the subspace of $\mathcal{H}_2^{\otimes N}$ formed by the simultaneous $+1$-eigenspaces of all the operators in the stabilizer group $\overline{\mathcal{S}}$. Here, $\mathcal{H}_2^{\otimes N}$ denotes the complex Hilbert space of dimension 2^N that comprises the state space of N-qubit systems.

It is also possible for elements of $\overline{\mathcal{G}}_N$ to commute with regard to the symplectic product with the elements of $\overline{\mathcal{S}}$ but not to belong to $\overline{\mathcal{S}}$. These elements are important because they act nontrivially on encoded quantum states (they corrupt the encoded quantum information but map the codespace to itself). Together with the elements of $\overline{\mathcal{S}}$, these operators define the group $\overline{\mathcal{Z}}(\overline{\mathcal{S}}) \subset \overline{\mathcal{G}}_N$ known as the effective centralizer. This group is the set of all operators in $\overline{\mathcal{G}}_N$ whose symplectic representatives commute with regard to the symplectic product with all the symplectic representatives of the stabilizer generators $(\mathbf{S}) \in \quad_2^{2N}$.

Based on the relationship between the stabilizer and the effective centralizer, the effective centralizer can actually be understood as a union of stabilizer cosets. More specifically, $\overline{\mathcal{Z}}(\overline{\mathcal{S}})$ can be partitioned into 2^{2k} cosets of $\overline{\mathcal{S}}$, each of these cosets being indexed by a coset representative $\{\mathbf{L}_j\}_{j=1}^{2^{2k}} \in \overline{\mathcal{Z}}(\overline{\mathcal{S}})$. Throughout this article, we will refer to these representatives as logical operators. Whenever a representative $\mathbf{L}_j$ is multiplied in terms of the group operation over the effective Pauli group, the $\star$ product, by all the elements of $\overline{\mathcal{S}}$, the stabilizer coset, $\mathbf{L}_j \star \overline{\mathcal{S}}$, is obtained. This can be written as

$$\overline{\mathcal{Z}}(\overline{\mathcal{S}}) = \bigcup_{j=1}^{2^{2k}} \mathbf{L}_j \star \overline{\mathcal{S}} = \overline{\mathcal{S}} \bigcup \left[\bigcup_{j=2}^{2^{2k}} \mathbf{L}_j \star \overline{\mathcal{S}} \right]. \quad (6)$$

Based on our definition of the effective centralizer, it is easy to see that the group $\overline{\mathcal{G}}_N$ also has a coset structure [14]. Knowing that $|\overline{\mathcal{G}}_N| = 2^{2N}$ and that $|\overline{\mathcal{Z}}(\overline{\mathcal{S}})/\overline{\mathcal{S}}| = 2^{N+k}$, then we can partition the effective Pauli group into cosets of the effective centralizer as

$$\overline{\mathcal{G}}_N = \bigcup_{i=1}^{2^{N-k}} \mathbf{T}_i \star \overline{\mathcal{Z}}(\overline{\mathcal{S}})$$

$$= \overline{\mathcal{Z}}(\overline{\mathcal{S}}) \bigcup \left[\bigcup_{i=2}^{2^{N-k}} \mathbf{T}_i \star \overline{\mathcal{Z}}(\overline{\mathcal{S}}) \right]. \quad (7)$$

Equation (7) shows how $\overline{\mathcal{G}}_N$ can be partitioned into 2^{N-k} cosets of the effective centralizer, each of these centralizer

cosets being indexed by a coset representative[1] $\{\mathbf{T}_i\}_{i=1}^{2^{N-k}}$, which when multiplied in terms of the $\star$ product by all the elements of $\overline{\mathcal{Z}(\mathcal{S})}$ yields the specific coset in question, $\mathbf{T}_i \star \overline{\mathcal{Z}(\mathcal{S})}$. Finally, we can combine the partitions of (6) and (7), respectively, to partition the entire effective Pauli space into cosets of the stabilizer, each one indexed by a unique $\star$ product of the $\{\mathbf{L}_j\}_{j=2}^{2^k}$ and $\{\mathbf{T}_i\}_{i=1}^{2^{N-k}}$ representatives. This can be written as

$$\overline{\mathcal{G}}_N = \bigcup_{i=1}^{2^{N-k}} \mathbf{T}_i \star \overline{\mathcal{Z}(\mathcal{S})} = \bigcup_{i=1}^{2^{N-k}} \bigcup_{j=1}^{2^{2k}} (\mathbf{T}_i \star \mathbf{L}_j) \star \overline{\mathcal{S}}. \quad (8)$$

C. QUANTUM SYNDROMES

In order to operate properly, QEC strategies require information regarding the transmitted quantum state. Although the axioms of quantum mechanics establish that direct measurement of a quantum state will result in information loss, quantum error syndrome measurements (Hadamard test) can be used to measure quantum states indirectly in order to obtain sufficient information for quantum codes to operate [29], [31], [33]. A quantum syndrome $\mathbf{w} \in \mathbb{F}_2^{N-k}$ captures the commutation properties of any error sequence $\mathbf{E} \in \overline{\mathcal{G}}_N$ induced by the Pauli channel with regard to the stabilizer generators $\{\mathbf{S}_\ell\}_{\ell=1}^{N-k}$ of the code. This can be written as

$$\mathbf{w} = \mathbf{e} \odot (\mathbf{s}_1, \dots, \mathbf{s}_{N-k}) \quad (9)$$

where $\mathbf{e} = \varphi(\mathbf{E})$ and $\{\mathbf{s}_\ell\}_{\ell=1}^{N-k} = \{\varphi(\mathbf{S}_\ell)\}_{\ell=1}^{N-k}$, respectively. From (9), we can discern that the syndrome is a length $N-k$ binary vector whose components express the commutation relations with regard to the $\star$ product between $\mathbf{E}$ and each of the generators of the stabilizer code. In other words, the entries of $\mathbf{w}$ will be 0 if the specific error operator and stabilizer generator commute and 1 if they do not. Recall that the elements of $\overline{\mathcal{Z}(\mathcal{S})}$ commute with the stabilizer generators, which means that all the elements of the same effective centralizer coset will be associated with the same syndrome. For instance, if a given error pattern $\mathbf{E}'$ is associated with syndrome $\mathbf{w}_1$, all the effective Pauli sequences that belong to $\mathbf{E}' \star \overline{\mathcal{Z}(\mathcal{S})}$ will also be associated with syndrome $\mathbf{w}_1$. Thus, the information provided by the syndrome $\mathbf{w}$ can be used to locate the centralizer coset $\mathbf{T}_i \star \overline{\mathcal{Z}(\mathcal{S})}$ of a given error operator $\mathbf{E}$, where $i = 1, \dots, 2^{N-k}$. Note, however, that we have no information regarding which stabilizer coset the error operator belongs to. The syndrome locates the centralizer coset that contains the error sequence that has taken place, but because all the elements of this coset are associated with the same syndrome, there is no way of knowing which specific stabilizer coset $(\mathbf{T}_i \star \mathbf{L}_j) \star \overline{\mathcal{S}}$ contains the error sequence, where $j = 1, \dots, 2^{2k}$. As mentioned previously, knowing which particular stabilizer coset contains the error that has taken place is important, since operators that belong to different stabilizer cosets act nontrivially on the encoded quantum information.

[1]These coset representatives are commonly referred to as pure error operators in the literature.

D. ESTIMATION AND END TO END ERRORS

The effective Pauli operators induced by quantum channels and estimated by decoders can be classified into three different groups (summarized in Table 1). Consider a channel error $\mathbf{E} \in \mathbf{T}_i \star \mathbf{L}_j \star \overline{\mathcal{S}}$ and an estimated error sequence $\hat{\mathbf{E}} \in \hat{\mathbf{T}}_i \star \hat{\mathbf{L}}_j \star \overline{\mathcal{S}}$; then the following situations can be encountered.

1) *End-to-end errors with different syndromes:* These events occur when the error sequence estimated by the decoder and the real error sequence belong to different centralizer cosets, i.e., $\mathbf{T}_i \neq \hat{\mathbf{T}}_i$. Equipped with an ideal decoder, such scenarios would not exist, as the syndrome of the estimated error pattern $\hat{\mathbf{E}}$ should always match the measured syndrome associated with the channel error $\mathbf{E}$. However, because sparse quantum codes are generally decoded based on the suboptimal classical sum product algorithm (SPA) algorithm,[2] these errors can take place with varying probability [3], [14], [20], [34].

2) *End-to-end identical syndrome errors:* These events take place when the estimated error sequence and the channel error both belong to the same centralizer coset, $\mathbf{T}_i = \hat{\mathbf{T}}_i$, but each of them belongs to a different stabilizer coset, i.e., $\mathbf{L}_j \neq \hat{\mathbf{L}}_j$ (the channel logical operator and estimated logical operator do not match). Thus, although $\hat{\mathbf{E}}$ and $\mathbf{E}$ exhibit identical commutation properties with respect to the stabilizer generators, they will each act on the transmitted codeword in a distinct nontrivial manner and the decoder will fail. It should be noted that in the literature, this type of end-to-end error is generally referred to as a logical error.

3) *End-to-end degenerate errors:* These events take place when the estimated error pattern and the channel error both belong to the same stabilizer coset, $\mathbf{T}_i = \hat{\mathbf{T}}_i$ and $\mathbf{L}_j = \hat{\mathbf{L}}_j$, but they do not match, i.e., $\hat{\mathbf{E}} \neq \mathbf{E}$. Since the estimated error belongs to the same stabilizer coset as the channel error, it will act identically on the quantum codeword, which means that it will not actually result in a decoding failure and should not be considered as such.

III PERFORMANCE ASSESSMENT METRICS

The manifestation of the degeneracy phenomenon in the realm of QEC makes it impossible to accurately predict the performance of QEC codes with methods that disregard its presence. Thus, in the paradigm of sparse quantum codes, which are famous for their degenerate nature, appropriate performance assessment is of paramount importance. This can be achieved using the logical error rate, which accurately predicts the performance of quantum codes and distinguishes between the different types of end-to-end errors (see Table 1).

[2]The SPA algorithm solves the symbolwise maximum likelihood (SWML) problem and not the actual maximum likelihood (ML) decoding problem. While the solutions to the SWML and ML problems usually coincide, this is not always the case.

TABLE 1 Characteristics of the Different Types of End-to-End Errors That Can Arise When Using Stabilizer Codes

Type of error	Defining Characteristics	Outcome (Phys. Error Rate)	Outcome (Logical Error Rate)
End-to-end error with different syndrome	$\hat{w} \neq w$ $\hat{T}_i \neq T_i$ $\hat{E} \neq E$	Failure	Failure
End-to-end identical syndrome error	$\hat{w} = w$ $\hat{T}_i = T_i$ and $\hat{L}_j \neq L_j$ $\hat{E} \neq E$	Failure	Failure
End-to-end degenerate error	$\hat{w} = w$ $\hat{T}_i = T_i$ and $\hat{L}_j = L_j$ $\hat{E} \neq E$	Failure	Success

The hat notation is used to represent estimations made by the decoder, i.e., $\hat{w}$, $\hat{E}$, $\hat{T}_i$, and $\hat{L}_j$ represent the estimated syndrome, estimated error sequence, centralizer coset representative of the estimated error sequence and stabilizer coset representative of the estimated error sequence, respectively.

Prior to discussing how to compute the logical error rate, it should be mentioned that the concept of *undetected* errors is not exclusive to the quantum paradigm. In fact, even though degeneracy does not exist in the classical coding framework, *undetected* or *logical* errors in classical low-density parity check code (LDPC) codes and classical turbo codes have previously been studied. This idea was introduced by the early work of MacKay et al. [43], where a classical undetected error is defined as a decoding estimate that is not equal to the original error sequence and that is produced when the decoder exits before the maximum number of decoding iterations (it produces a valid syndrome). In their analysis of classical LDPC codes, MacKay et al. showed that all of the decoding mistakes they encountered were detected errors (classical undetected errors were only observed in turbo codes). This was also shown in [17] for a slightly different decoding algorithm. Similar outcomes were observed in the quantum paradigm for the failed recoveries of the modified decoding strategies of [3] and [24]. These techniques serve to improve standard SPA decoding of quantum codes by postprocessing the initial error estimates and producing new estimates of the channel error. If these new estimates do not revert the channel error, then they are referred to as failed recoveries or failed error corrections. In [3], all of the failed error corrections were shown to be end-to-end errors with different syndromes, whereas in [24], a small percentage of failed estimates were shown to be end-to-end identical syndrome errors and end-to-end degenerate errors.[3]

Aside from these failed recovery analyses, the literature is limited when it comes to assessing the percentage of decoding failures that are caused by each type of end-to-end error. It is reasonable to believe that QLDPC codes, given their large number of degenerate operators [3], [4], [12], [14], [27], [28], will experience a large percentage of end-to-end degenerate errors. We confirm this intuition in the final section of our work, where we show how up to 30⊠ of the end-to-end

errors that take place when using the QLDPC codes of [21], [22], and [26] are degenerate.

A. DISCRIMINATING BETWEEN DIFFERENT TYPES OF END TO END ERRORS

It is clear from the defining characteristics of each specific type of end-to-end error (see Table 1) that *end-to-end errors with different syndromes* are the easiest type of error to identify. In fact, doing so is trivial, as all that is required is a comparison of the syndrome estimate, $\hat{w} \in \mathbb{F}_2^{N-k}$, and the measured syndrome, $w \in \mathbb{F}_2^{N-k}$, where k and N represent the number of logical qubits and physical qubits (blocklength) of the quantum code, respectively. Similarly, knowing that either an *end-to-end identical syndrome error* or an *end-to-end degenerate error* has occurred is simple. This can be done by comparing the estimate of the error sequence $\hat{E}$ to the channel error E whenever $\hat{w} = w$, i.e., if $\hat{E} \neq E$, either an end-to-end degenerate error or an identical syndrome error has occurred, and if $\hat{E} = E$, no error has taken place. The issue arises when trying to distinguish between these two families of end-to-end errors. Note that the comparison $\hat{E} \neq E$ does not reveal if the error estimate belongs to the same stabilizer coset as the channel error. Hence, we have no way of discriminating between end-to-end degenerate errors and end-to-end identical syndrome errors.

Conceptually, the simplest and most straightforward strategy that comes to mind to resolve this problem is to compute the stabilizer of the code in question, compute the ⋆ product of the stabilizer with the channel error E to extract the specific stabilizer coset of the channel error, and then check if $\hat{E}$ belongs to this coset. As mentioned in [14], this works because the coset representative choice for the coset $(T_i \star L_j) \star \overline{S}$ is irrelevant (any operator belonging to the coset serves as a valid representative), where $\{T_i, L_j\} \in \overline{\mathcal{G}}_N$ and $\overline{S} \subset \overline{\mathcal{G}}_N$ represent the effective centralizer coset representative, the stabilizer coset representative, and the stabilizer itself. Thus, computing $E \star \overline{S}$ will yield the stabilizer coset of the channel error, i.e., $E \star \overline{S} = (T_i \star L_j) \star \overline{S}$. Therefore, whenever $\hat{w} = w$ and $\hat{E} \neq E$, we will know that an end-to-end degenerate

---

[3]The authors of this work do not distinguish between end-to-end identical syndrome errors and end-to-end degenerate errors.

error has occurred if the estimated error sequence $\hat{\mathbf{E}}$ is in the coset $\mathbf{E} \star \overline{\mathcal{S}}$. If this does not occur, then an end-to-end identical syndrome error will have taken place.

Unfortunately, since extracting the stabilizer of a quantum code becomes increasingly complex as its blocklength increases, this strategy will only be applicable to short quantum codes. The number of elements in the stabilizer of a quantum code with blocklength N and rate R_Q is given by $2^{N-k} = 2^{N(1-R_Q)}$ [14], which grows exponentially with N and can rapidly become intractable on a classical machine as this parameter increases. In light of this, it is apparent that more practical methods to differentiate between end-to-end degenerate errors and end-to-end identical syndrome errors are necessary. Both the strategy we propose herein and that employed in [3], [15], [16], [17], [18], which are explained in the next two sections, resolve this issue.

B. ALGEBRAIC PERSPECTIVE ON END TO END DEGENERATE ERRORS

The problem of differentiating between end-to-end identical syndrome errors and end-to-end degenerate errors can also be formulated as a set of linear equations. The PCM of a stabilizer code that encodes k logical qubits into N physical qubits (a rate $R_Q = \frac{k}{N}$ code with blocklength N) can be written as

$$\mathbf{H}_{\overline{\mathcal{S}}} = \begin{pmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \vdots \\ \mathbf{h}_{N-k} \end{pmatrix} \qquad (10)$$

where $\mathbf{h}_{=\mathbf{s}}$ denotes the symplectic representation of the generators $\{\mathbf{S}\}_{=1}^{N-k} \in \overline{\mathcal{G}}_N$ that define the stabilizer group $\overline{\mathcal{S}}$. Each of the elements of $\overline{\mathcal{S}}$ is a linear combination of the $N-k$ generators; hence, if $\mathbf{S}$ is an element of the stabilizer and $\mathbf{s}$ is the symplectic representation of this stabilizer element, then

$$\mathbf{s} = \left(\sum_{i=1}^{N-k} a_i \mathbf{h}_i \right) \bmod 2 \qquad (11)$$

where $\mathbf{a} = (a_1, \ldots, a_{N-k})$ is a unique binary vector.

Whenever a channel error $\mathbf{E}$ takes place, the decoder will compute an estimate of this error and produce an estimate of the syndrome associated with it. As discussed previously, this syndrome only determines which specific effective centralizer coset the channel error belongs to. In other words, the syndrome provides the effective centralizer coset representative $\mathbf{T}_i$ (known as the pure error component in the literature) of the channel error. Assuming that an end-to-end error with different syndrome does not occur, the estimated syndrome will match the measured syndrome; hence, the centralizer coset representative of the estimated error sequence and the centralizer coset representative of the channel error will also be the same, i.e., $\mathbf{T}_i = \hat{\mathbf{T}}_i$. Thus, if we compute the $\star$ operation of the channel error $\mathbf{E}$ and the estimated error $\hat{\mathbf{E}}$, which can also be understood as the mod2 sum of their symplectic representations over $\mathbb{F}_2^{2N}$: $(\mathbf{E}) \oplus (\hat{\mathbf{E}}) = \mathbf{e} \oplus \hat{\mathbf{e}}$, where denotes the symplectic map, the sequence $\mathbf{E}$ will be shifted

to the effective centralizer $\overline{\mathcal{Z}}(\overline{\mathcal{S}})$. Note that we can also write this using the symplectic map as $\mathbf{E} \star \hat{\mathbf{E}} \in \overline{\mathcal{Z}}(\overline{\mathcal{S}}) \rightarrow (\mathbf{E}) \oplus (\hat{\mathbf{E}}) \in (\overline{\mathcal{Z}}(\overline{\mathcal{S}})) = \mathbf{e} \oplus \hat{\mathbf{e}} \in \Gamma_R$, where $\Gamma_R \subset \mathbb{F}_2^{2N}$ denotes the equivalent group of the effective centralizer over $\mathbb{F}_2^{2N}$. Based on this, the issue of determining whether an end-to-end error is degenerate can be understood as finding out if $\mathbf{E} \star \hat{\mathbf{E}}$ belongs to the stabilizer. This can be formulated based on the symplectic map into the following question:

$$\exists\, \mathbf{a} : \mathbf{e} \oplus \hat{\mathbf{e}} = \left(\sum_{i=1}^{N-k} a_i \mathbf{h}_i \right) \bmod 2 \ ? \qquad (12)$$

Essentially, if a set of coefficients $\mathbf{a}$ exists such that the above equation holds, i.e., if $\mathbf{e} \oplus \hat{\mathbf{e}} \in \mathbb{F}_2^{2N}$ is a linear combination of the symplectic representation of the stabilizer generators, then $\mathbf{E} \star \hat{\mathbf{E}}$ belongs to the stabilizer and an end-to-end degenerate error will have occurred. If such a set of coefficients does not exist, then an end-to-end identical syndrome error has taken place.

The expression shown in (12) defines a linear system of equations over the binary field. An answer to this question can be found by writing the augmented matrix $[\mathbf{H}_{\overline{\mathcal{S}}}^{\top} | (\mathbf{e} \oplus \hat{\mathbf{e}})^{\top}]$ in its row-echelon form. This means that, based on this procedure, it is possible to determine the type of end-to-end errors that occur and subsequently compute the logical error rate. This is done in [20]. However, although the procedure is conceptually simple, rewriting the augmented matrix in such a manner becomes increasingly computationally complex as matrices grow in size. Unfortunately, for sparse quantum codes to be good, the blocklength must be large, which implies that the PCMs of these codes will also be large.[4] Furthermore, the row-echelon form of the matrix $[\mathbf{H}_{\overline{\mathcal{S}}}^{\top} | (\mathbf{e} \oplus \hat{\mathbf{e}})^{\top}]$ must be computed during every simulation iteration (whenever the estimated syndrome and the measured syndrome match) to determine what type of end-to-end error has taken place, which may significantly increase simulation time. For these reasons, calculating the logical error rate based on this procedure can become a cumbersome and lengthy endeavor. Therefore, the task at hand is to find a more practical and less computationally demanding way to determine if the congruence equation system given in (12) has a solution, as this suffices to determine if the end-to-end error is degenerate (we do not actually need to solve the system itself).

1 CLASSICAL CODING-INSPIRED STRATEGY

It is possible to find an answer to (12) by casting the problem in the framework of classical linear block codes. In classical coding theory, the encoding matrix or generator matrix $\mathbf{G}_c$ of a binary linear block code and its corresponding PCM $\mathbf{H}_c$ fulfil $(\mathbf{G}_c \mathbf{H}_c^{\top}) \bmod 2 = (\mathbf{H}_c \mathbf{G}_c^{\top}) \bmod 2 = 0$. This means that the PCM defines a basis for the nullspace of the generator matrix and viceversa. In the classical scenario, having a basis for the nullspace of a code enables us to determine

---

[4]The PCMs of QEC codes are of size $N - k \times 2N$. A common size in the literature of QLDPC codes is 10 000 qubits; thus, the PCM associated with such a code would be of size $(10000 - k) \times 20000$.

whether the decoding outcome $\mathbf{x}$ belongs to the code by simply computing its product with the PCM of the code, i.e., if $(\mathbf{H}_c \mathbf{x}^\top) \mod 2 = 0$ then $\mathbf{x}$ is a codeword. Essentially, whenever $(\mathbf{H}_c \mathbf{x}^\top) \mod 2 = 0$, the decoding outcome is a linear combination of the rows of the generator matrix $\mathbf{G}_c$ and it belongs to the code, and whenever $(\mathbf{H}_c \mathbf{x}^\top) \mod 2 \neq 0$, $\mathbf{x}$ does not belong to the code.

Note that, based on this formulation, the quandary posed in (12) is reminiscent of the classical decoding scenario. The main difference is that instead of determining if the decoding outcome belongs to the code, we must discover if the sum of the symplectic representations of the channel error and the estimated error belong to the stabilizer. This parallelism between the classical and quantum problems allows us to apply the classical resolution strategy to the quantum paradigm with only a slight caveat: answering (12) requires an inverse approach to the classical method. Since the generators of the stabilizer code are given by the rows of the PCM $\mathbf{H}_{\overline{S}}$, the corresponding kernel generator matrix[5] $\mathbf{G}_{\overline{S}}$ (instead of the PCM like in the classical paradigm) must be used to discover if $\mathbf{e} \oplus \hat{\mathbf{e}}$ can be written as a linear combination of the stabilizer generators. The matrix $\mathbf{G}_{\overline{S}}$ defines a basis for the nullspace of the stabilizer code; hence, it will suffice to compute $[\mathbf{G}_{\overline{S}}(\mathbf{e} \oplus \hat{\mathbf{e}})^\top] \mod 2$ to find the answer to (12). If $[\mathbf{G}_{\overline{S}}(\mathbf{e} \oplus \hat{\mathbf{e}})^\top] \mod 2 = 0$, $\mathbf{E} \star \hat{\mathbf{E}} \in \overline{S}$ which means that an end-to-end degenerate error has occurred, and if $[\mathbf{G}_{\overline{S}}(\mathbf{e} \oplus \hat{\mathbf{e}})^\top] \mod 2 \neq 0$, $\mathbf{E} \star \hat{\mathbf{E}} \notin \overline{S}$, and an end-to-end identical syndrome error will have taken place.

This strategy provides us with a simple and computationally efficient method to determine the type of end-to-end error that has taken place. The only requirement is obtaining the matrix $\mathbf{G}_{\overline{S}}$, which can be computed once (by finding a basis for the nullspace of its PCM $\mathbf{H}_{\overline{S}}$) and can then be stored offline for any stabilizer code. In this manner, we have designed a simple method to solve (12) that does not require the computation of the stabilizer and so avoids the complexity issues that this entails.

C. DETECTING END TO END DEGENERATE ERRORS USING ENCODED PAULI OPERATORS

There is another manner of distinguishing between end-to-end identical syndrome errors and end-to-end degenerate errors. It involves obtaining the encoded Pauli operators[6] of a code following a method derived by Gottesman in his seminal work [8], and then using these operators to determine whether the error estimate produced by the decoder is in the stabilizer coset of the channel error (end-to-end degenerate errors occur when this happens and end-to-end identical syndrome errors occur when it does not). This strategy was

first applied to QLDPC codes in [3] and has since been used in [15], [16], [17], [18].

The encoded Pauli operators of an N-qubit stabilizer code are defined as those operators in $\mathcal{G}_N$ that commute with the elements of the stabilizer group and whose action on an encoded state can be understood as an X, Y, or Z operation on each of the encoded logical qubits. Each stabilizer code has $2k$ encoded Pauli operators[7] which are generally represented using the notation $\overline{Z}_q$ and $\overline{X}_l$, where $\{q, l\} \in \{1, \ldots, k\}$. In this manner, $\overline{Z}_q$ represents an operator in $\mathcal{G}_N$ whose action is analogous to performing a Z operation (phase flip) on the ith logical qubit. Thus, $\overline{Z}_q \in \mathcal{G}_N$ maps to $Z_q \in \mathcal{G}_k$, where Z_q denotes the action of a Z operator on the ith qubit and the action of I operators on the remaining qubits (identity operators are omitted). Recall that, since the global phase can be ignored, it will be equivalent for the purpose of error correction to consider the equivalent encoded Pauli operators over the effective N-fold Pauli group, $\overline{\mathbf{Z}}_q \equiv \overline{Z}_q$ and $\overline{\mathbf{X}}_l \equiv \overline{X}_l$, where $\overline{\mathbf{Z}}_q, \overline{\mathbf{X}}_l \in \overline{\mathcal{G}}_N$ (we use capital boldface to preserve the notation of [14]).

Based on this definition of the encoded Pauli operators, we know that an encoded Pauli operator $\overline{Z}_q$ commutes with all the elements of $\overline{S}$ as well as with all other encoded Pauli operators except for the operator $\overline{X}_l$ when $q = l$. This means that, if the encoded Pauli operators of a stabilizer code are known, we can determine if an operator $\mathbf{A} \in \overline{\mathcal{Z}}(\overline{S}) \subset \overline{\mathcal{G}}_N$ belongs to $\overline{S}$ by checking the commutation relations of $\mathbf{A}$ with the encoded Pauli operators. If $\mathbf{A}$ commutes with all the encoded Pauli operators, it is within the stabilizer and if it does not (it anticommutes with one encoded Pauli operator), it is not in the stabilizer. Against this backdrop, it is easy to see how this strategy can be applied to solve the issue of discriminating between end-to-end identical syndrome errors and end-to-end degenerate errors. After successful decoding (the decoder produces a matching estimate of the syndrome), we know that $\mathbf{E} \star \hat{\mathbf{E}} \in \overline{\mathcal{Z}}(\overline{S})$. Now, we determine if $\mathbf{E} \star \hat{\mathbf{E}}$ is in $\overline{S}$ by checking its commutation status with $\{\overline{\mathbf{Z}}_q\}_{q=1}^k$ and $\{\overline{\mathbf{X}}_l\}_{l=1}^k$. If $\mathbf{E} \star \hat{\mathbf{E}}$ commutes with all the encoded Pauli operators, $\mathbf{E} \star \hat{\mathbf{E}} \in \overline{S}$ and an end-to-end degenerate error has occurred. If not, an end-to-end identical syndrome error has occurred. It is based on these comparisons that the logical error rate was successfully computed in [15], [16], [17], [18].

Naturally, to be able to apply the method, one must first have knowledge of the encoded Pauli operators of the code. This is similar to the classical coding-based strategy we propose in this article, which requires the computation of the kernel generator matrix $\mathbf{G}_{\overline{S}}$. The encoded Pauli operators of a stabilizer code can be found based on the concept of the *standard form* (see [8, Ch. 4]). In this work, Gottesman showed how, by applying row operations (i.e., Gaussian elimination) together with the necessary qubit permutations (i.e., column permutations) on the PCM of a stabilizer code, one can obtain

---

[5]We refer to the matrix $\mathbf{G}_{\overline{S}}$ as the kernel generator matrix to avoid the term stabilizer generator matrix, as this latter term implies that the matrix can be used for encoding purposes (which may not be true in the present case).

[6]In the literature, these operators are referred to as logical operators. However, we use the term encoded Pauli operators to distinguish them from the stabilizer coset representatives $\{\mathbf{L}_j\}_{j=1}^{2^{2k}}$ which we originally defined as logical operators.

[7]Given the difference in the number of stabilizer coset representatives (logical operators) $\{\mathbf{L}_j\}_{j=1}^{2^{2k}}$ and encoded Pauli operators, it is useful to employ different terminology for these concepts (even though encoded Pauli operators are related to logical operators and viceversa).

a special row reduced echelon form of the PCM: the standard form. Once the standard form is known, the encoded Pauli operators $\{\overline{Z}_q\}_{q=1}^{k}$ and $\{\overline{X}_l\}_{l=1}^{k}$ can be directly obtained from it using matrix algebra [8], [41]. As with the kernel generator matrix $\mathbf{G}_{\overline{S}}$, the encoded Pauli operators of a specific code need only be computed once and can then be stored offline.

D. METHOD COMPARISON

For general stabilizer codes, computing the logical error rate based on the classical coding inspired strategy does not yield benefits over doing so via the encoded Pauli operator method. This has to do with the fact that both of them have the same computational complexity: performing Gaussian elimination and column permutations is also what finding the basis for the nullspace of a matrix requires. As is shown in [44], these operations scale as $O(N^3)$. However, when it comes to the widely employed family of CSS codes, the classical coding inspired strategy we propose herein will benefit from a substantial complexity advantage.

CSS codes [5], [7] are a particular type of stabilizer code that provide an efficient design strategy to build QEC codes from existing classical codes. The quantum PCM of a CSS code is written as

$$\mathbf{H}_Q = (\mathbf{H}_x | \mathbf{H}_z) = \begin{pmatrix} \mathbf{H}'_x & 0 \\ 0 & \mathbf{H}'_z \end{pmatrix} \tag{13}$$

where $\mathbf{H}_x = \begin{pmatrix} \mathbf{H}'_x \\ 0 \end{pmatrix}$ and $\mathbf{H}_z = \begin{pmatrix} 0 \\ \mathbf{H}'_z \end{pmatrix}$.

In this construction, $\mathbf{H}'_x$ and $\mathbf{H}'_z$ are the parity check matrices of two binary classical LDPC codes C_1 and C_2, respectively, where each matrix is used to correct either bit-flips or phase-flips. As discussed previously, any classical binary LDPC code or any classical binary linear block code for that matter will satisfy $(\mathbf{G}_c \mathbf{H}_c^{\top}) \bmod 2 = \boxtimes$, where $\mathbf{G}_c$ and $\mathbf{H}_c$ represent the generator and parity check matrices of a binary classical LDPC code, respectively. Because the generator matrix is necessary to encode a classical LDPC code, the issue of deriving the matrix $\mathbf{G}_c$ from $\mathbf{H}_c$ has been widely researched in the realm of classical error correction [44]–[48]. In general, finding the matrix $\mathbf{G}_c$ from $\mathbf{H}_c$, i.e., bringing the PCM into the desired form, requires $O(N^3)$ operations of processing [44]. However, through the clever application of optimization algorithms, the complexity of this task has been reduced [44], [45], [46] and classical LDPC codes with a linear encoding complexity[8] have been found. Note that based on the structure of the quantum parity check matrix (QPCM) of CSS codes (13), we can derive the kernel generator matrix of a CSS code as

$$\mathbf{G}_{\overline{S}} = (\mathbf{G}_x | \mathbf{G}_z) = \begin{pmatrix} \mathbf{G}'_x & 0 \\ 0 & \mathbf{G}'_z \end{pmatrix} \tag{14}$$

where $\mathbf{G}'_x$ and $\mathbf{G}'_z$ are the generator matrices of the binary classical LDPC codes C_1 and C_2, respectively. We know that the matrix given in (14) will actually serve the purpose of a kernel generator matrix because its rows are all linearly independent. The difficulty in computing (14) lies in deriving the generator matrices $\mathbf{G}'_x$ and $\mathbf{G}'_z$ from the PCMs $\mathbf{H}'_x$ and $\mathbf{H}'_z$. However, when constructing CSS QLDPC codes from classical LDPC codes, it is a prerequisite that the classical code be known. Thus, it is reasonable to assume that the generator matrices of the classical codes that make up a CSS QLDPC code are available.[9] Against this backdrop, because this does not require finding the standard form of the QPCM and deriving its encoded Pauli operators, the method we propose herein will have a complexity advantage over the encoded Pauli operator method when computing the logical error rate of sparse CSS quantum codes. More specifically, when using classical LDPC codes to build QLDPC CSS codes, because the generator matrices associated with the corresponding parity check matrices of these codes are known, the complexity of our method is negligible, as no calculations will be necessary to obtain the kernel generator matrix [it can be derived by simply introducing the appropriate generator matrices in (14), i.e., the complexity is $O(1)$]. In the case that the generator matrix of a particular classical LDPC code is not known, it can be derived as shown in [44] with complexity $O(N) + O(g^2)$, where g is the gap to measure the "distance" between the PCM and a lower triangular matrix. This means that the computational complexity of calculating the kernel generator matrix of a CSS code will be at most $O(N) + O(g^2)$, which implies that, in the context of CSS codes, our strategy is less computationally complex than the one based on using encoded Pauli operators, which has complexity $O(N^3)$.

IV SIMULATION RESULTS

We close this section by using our method to show how end-to-end errors with identical syndromes and end-to-end degenerate errors can account for a significant percentage of the end-to-end errors that occur when using QLDPC codes. For this purpose, we simulate the CSS quantum low density generator matrix (QLDGM) codes of [21], [22], and [26] with different rates and blocklengths over the depolarizing channel. The characteristics of these codes are detailed in Table 2. The underlying classical low density generator matrix (LDGM) matrices and the structure of the M matrix, critical to the performance of the QLDGM codes, have been chosen according to the optimization guidelines detailed in [21], [22], [26], [50], and [51]. The results of these simulations are shown in Fig. 1, where each subfigure groups the results by blocklength, i.e., each of the subfigures portrays the results for all the codes with the same value of N. The graphs plot

---

[8]Richardson et al. [44] showed that the encoding complexity is upper bounded by $O(N) + O(g^2)$, where g is the gap to measure the "distance" between a given PCM and a lower triangular matrix. This means that in the extreme case where $g = 0$, such as for irregular repeat-accumulate codes [49], $\mathbf{G}_c$ can be computed with complexity $O(N)$.

[9]If the LDPC codes have previously been used in the classical paradigm, then these generator matrices are known (they must have been employed in the classical encoding procedure).

TABLE 2 Parameter Values and Configurations of Simulated Codes

N	R_Q	Classical LDGM	$[m, p, x, y]$
100	0.1	P(3, 3)	[45, 24, 6, 3]
100	0.2	P(3, 3)	[40, 18, 6, 3]
100	0.25	P(3, 3)	[38, 15, 6, 3]
100	0.5	P(3, 3)	[25, 6, 7.57, 3]
500	0.1	P(5, 5)	[225, 170, 11, 3]
500	0.2	P(5, 5)	[200, 144, 11, 3]
500	0.25	P(5, 5)	[188, 130, 11, 3]
500	0.33	P(5, 5)	[163, 102, 11, 3]
500	0.5	P(5, 5)	[125, 60, 11, 3]
2000	0.1	P(9, 9)	[900, 691, 11, 3]
2000	0.2	P(9, 9)	[800, 581, 11, 3]
2000	0.25	P(9, 9)	[750, 526, 11, 3]
2000	0.33	P(9, 9)	[670, 438, 11, 3]
2000	0.5	P(9, 9)	[500, 251, 11, 3]

the ratio of a specific type of end-to-end error against the depolarizing probability. The aforementioned ratio is computed as $\frac{E_i}{E_T}$, where E_i denotes the total number of end-to-end errors of a specific type ($i = 1, 2, 3$):

1) $E_1 \rightarrow$ end-to-end errors with different syndromes;
2) $E_2 \rightarrow$ end-to-end errors with identical syndromes;
3) $E_3 \rightarrow$ end-to-end degenerate errors,

and E_T represents the total number of end-to-end errors. To ensure that the simulation results are precise, the ratios have been computed after a total of 1000 decoding mistakes have been made (following the Monte Carlo simulation rule of thumb provided in [52]), i.e., $E_T = 1000$.

The outcomes portrayed in Fig. 1 confirm our initial intuition that sparse quantum codes are degenerate. It is easy to see that for all of the simulated blocklengths and rates (except for $R_Q = 0.5$), the percentage of end-to-end errors that are not of the E_1 type is not negligible, i.e., $\frac{E_2}{E_T} + \frac{E_3}{E_T} \neq 0$. Furthermore, these results speak toward the higher precision of the logical error rate compared to the physical error rate when assessing the performance of these codes. For instance, at a noise level of $p = 0.005$, $\frac{E_3}{E_T} = 0.198$ for the $N = 500$ and $R_Q = 0.1$ code. This means that 19.8⬚ of the end-to-end errors are degenerate and should not be counted as decoding failures. Thus, in this scenario, the physical error rate overestimates the number of decoding failures and does not provide an accurate representation of the performance of the code. In fact, regardless of the noise level of the channel, the rate of the code (except for $R_Q = 0.5$), and the blocklength of the code, end-to-end degenerate errors take place, and so the physical error rate will always provide an inaccurate representation of the performance of these sparse quantum codes. Therefore, as is stated in [18], it is clear that performance results assessed based on the physical error rate ($\hat{E} = E$ as the decoding success criterion) [21], [22], [23], [24], [25], [26] are inaccurate (they report an upper bound).

Furthermore, the results shown in Fig. 1 also reveal how the frequency with which each type of end-to-end error takes place varies as a function of different parameters.
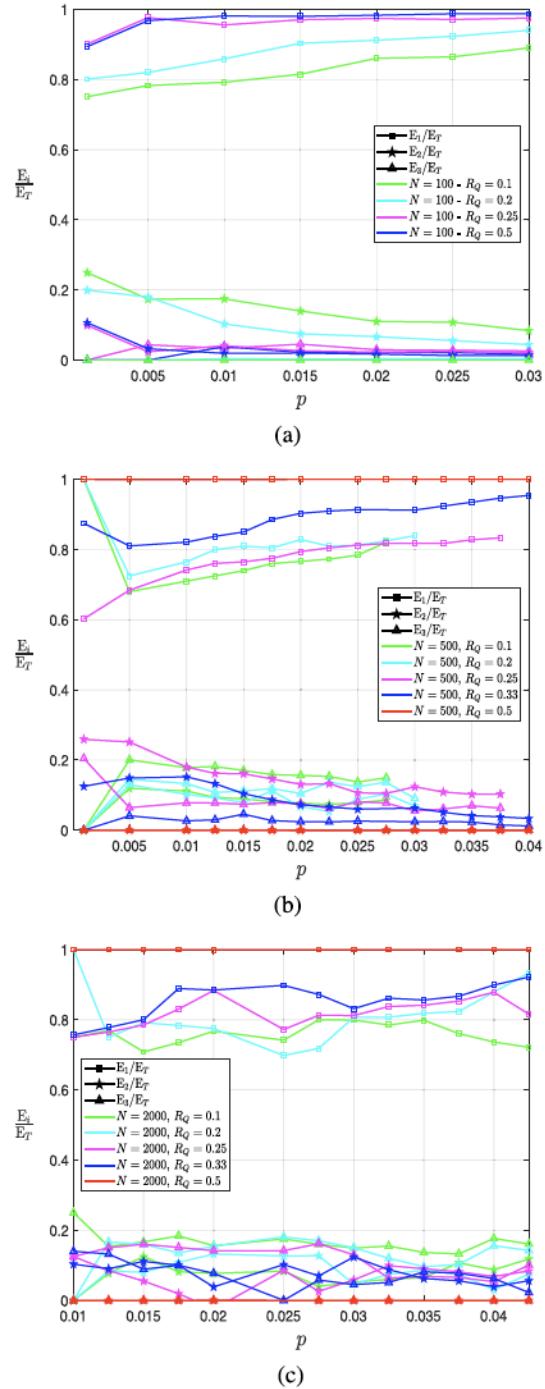

FIGURE 1 Ratios of end-to-end different syndrome, identical syndrome, and degenerate errors for codes of various rates and blocklengths: a) N 100; b) N 500; c) N 2000.

1) End-to-end errors with different syndromes represent a large percentage of the total number of end-to-end errors when the rate of the code is high. This percentage decreases as the rate of the codes goes from $R_Q = 0.5$ to $R_Q = 0.1$ (see ratio $\frac{E_1}{E_T}$ in Fig. 1). This trend becomes further exacerbated as the blocklength of the simulated codes increases, i.e., for low rate large

blocklength codes, the ratio $\frac{E_1}{E_T}$ will be significantly smaller than for low rate short codes.

2) End-to-end identical syndrome errors represent the smallest percentage of the total number of end-to-end errors in most of the simulated cases. This is reflected by the fact that $\frac{E_2}{E_T} < \frac{E_1}{E_T}, \frac{E_3}{E_T}$ in all of our simulation outcomes.

3) As the noise level of the channel grows, the ratio $\frac{E_1}{E_T}$ becomes larger and the ratio $\frac{E_2}{E_T}$ diminishes. The ratio of end-to-end degenerate errors $\frac{E_3}{E_T}$ stays relatively constant.

The relationships between these parameters and different types of end-to-end errors serve to draw conclusions, some of which should be explored in future work. For instance, the large values of $\frac{E_1}{E_T}$ in many of the simulated instances can be understood as a sign that performance gains may be attained by improving the decoding algorithm. This means that applying modified decoding strategies, such as those of [15], [16], [17], [20], [24], will aid in reducing the presence of end-to-end errors with different syndromes (they estimate the syndrome correctly when the original decoder does not) and improve performance. A matter that should be considered is how often these strategies produce failed error corrections in the form of end-to-end identical syndrome errors. For the methodology of [24], such events were shown to be rare; hence, we expect these strategies to be a good approach to improve the performance of QLDPC codes. On the other hand, the large values of $\frac{E_3}{E_I}$ when compared to $\frac{E_2}{E_I}$, especially at higher blocklengths, show that end-to-end identical syndrome errors are the least frequent of all the end-to-end error types. Despite the relatively small percentage that end-to-end identical syndrome errors represent, it is possible that their relevance will grow when the amount of end-to-end errors with different syndromes is reduced (using modified decoding strategies) or when the degenerate content of the code is increased (through design). At this point, it may be that further improvements in performance will only be possible by designing an optimal degenerate decoder with the capability to correct end-to-end identical syndrome errors. Finally, given that the method proposed in this work is valid to detect end-to-end degenerate errors, it may be interesting in future work to employ this methodology to specifically design codes to be degenerate. This could result in code constructions in which the likelihood of end-to-end degenerate errors is maximized, which would allow the positive effects of degeneracy (improved error correction capabilities without a decoding complexity increase) to be completely exploited for QEC purposes.

V CONCLUSION

We have presented a method to detect degenerate errors in sparse quantum codes in a computationally efficient manner. We have also shown how this method is less complex than other existing strategies to compute the logical error rate of sparse CSS quantum codes. Making use of our scheme, we have shown how sparse quantum codes have a significant percentage of degenerate errors. This means that the discrepancy between the logical error rate and the physical error rate is exacerbated for sparse quantum codes. Our results show that, for specific families of QLDPC codes, performance may be up to 20⊠ better than would be expected from previous results in the literature that are based on the physical error rate. In addition, these simulation outcomes serve to show how performance may be improved by constructing degenerate quantum codes, and they also speak toward the positive impact that modified decoding strategies can have on the performance of sparse quantum codes.

REFERENCES

[1] D. P. DiVincenzo, P. W. Shor, and J. A. Smolin, "Quantum-channel capacity of very noisy channels," *Phys. Rev. A*, vol. 57, pp. 830–839, 1998, doi: 10.1103/PhysRevA.57.830.

[2] G. Smith and J. A. Smolin, "Degenerate quantum codes for Pauli channels," *Phys. Rev. Lett.*, vol. 98, 2007, Art. no. 030501, doi: 10.1103/PhysRevLett.98.030501.

[3] D. Poulin and Y. Chung, "On the iterative decoding of sparse quantum codes," *QIC*, vol. 8, no. 10, 2008, Art. no. 987, doi: 10.5555/2016985.2016993.

[4] M. H. Hsieh and F. Le Gall, "NP-hardness of decoding quantum error-correction codes," *Phys. Rev. A*, vol. 83, 2010, Art. no. 052331, doi: 10.1103/PhysRevA.83.052331.

[5] A. R. Calderbank and P. W. Shor, "Good quantum error correcting codes exist," *Phys. Rev. A*, vol. 54, pp. 1098–1105, 1996, doi: 10.1103/PhysRevA.54.1098.

[6] A. R. Calderbank et al., "Quantum error correction via codes over GF(4)," *IEEE Trans. Inf. Theory*, vol. 44, no. 4, pp. 1369–1387, Jul. 1998, doi: 10.1109/18.681315.

[7] A. Steane, "Multiple-particle interference and quantum error correction," *Proc. Roy. Soc. A. Math., Phys. Eng. Sci.*, vol. 452, no. 1954, pp. 2551–2577, 1996, doi: 10.1098/rspa.1996.0136.

[8] D. Gottesman, "Stabilizer codes and quantum error correction," Ph.D. dissertation, California Inst. Technol., Pasadena, CA, USA, 1997, *arXiv:quant-ph/9705052*, doi: 10.48550/arXiv:quant-ph/9705052.

[9] B. Yoshida and I. L. Chuang, "Framework for classifying logical operators in stabilizer codes," *Phys. Rev. A*, vol. 81, 2010, Art. no. 052302, doi: 10.1103/PhysRevA.81.052302.

[10] E. Pelchat and D. Poulin, "Degenerate viterbi decoding," *IEEE Trans. Inf. Theory*, vol. 59, no. 6, pp. 3915–3921, Jun. 2013, doi: 10.1109/TIT.2013.2246815.

[11] S. Varsamopoulos, B. Criger, and K. Bertels, "Decoding small surface codes with feedforward neural networks," *Quantum Sci. Technol.*, vol. 3, no. 1, 2017, Art. no. 015004, doi: 10.1088/2058-9565/aa955a.

[12] P. Iyer and D. Poulin, "Hardness of decoding quantum stabilizer codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 9, pp. 5209–5223, Sep. 2015, doi: 10.1109/TIT.2015.2422294.

[13] D. Poulin, "Optimal and efficient decoding of concatenated quantum block codes," *Phys. Rev. A*, vol. 74, 2006, Art. no. 052333, doi: 10.1103/PhysRevA.74.052333.

[14] P. Fuentes et al., "Degeneracy and its impact on the decoding of sparse quantum codes," *IEEE Access*, vol. 9, pp. 89093–89119, 2021, doi: 10.1109/ACCESS.2021.3089829.

[15] P. Panteleev and G. Kalachev, "Degenerate quantum LDPC codes with good finite length performance," *Quantum*, vol. 5, p. 585, 2021, doi: 10.22331/q-2021-11-22-585.

[16] J. Roffe et al., "Decoding across the quantum LDPC code landscape," *Phys. Rev. Res.*, vol. 2, no. 4, 2020, Art. no. 043423, doi: 10.1103/PhysRevResearch.2.043423.

[17] K.-Y. Kuo and C.-Y. Lai, "Refined belief propagation decoding of sparse-graph quantum codes," *IEEE J. Sel. Areas Inf. Theory*, vol. 1, no. 2, pp. 487–498, Aug. 2020, doi: 10.1109/JSAIT.2020.3011758.

[18] K.-Y. Kuo and C.-Y. Lai, "Exploiting degeneracy in belief propagation decoding of quantum codes," *arXiv:2104.13659*, doi: 10.48550/arXiv.2104.13659.

[19] T. Camara, H. Ollivier, and J. Tillich, "A class of quantum LDPC codes: Construction and performances under iterative decoding," in *Proc. IEEE Int. Symp. Inf. Theory*, 2007, pp. 811–815, doi: 10.1109/ISIT.2007.4557324.

[20] A. Rigby, J. C. Olivier, and P. Jarvis, "Modified belief propagation decoders for quantum low-density parity-check codes," *Phys. Rev. A*, vol. 100, no. 1, 2019, Art. no. 012330, doi: 10.1103/PhysRevA.100.012330.

[21] H. Lou and J. Garcia-Frias, "Quantum error-correction using codes with low-density generator matrix," in *Proc. IEEE 6th Workshop Signal Process. Adv. Wireless Commun.*, 2005, pp. 1043–1047, doi: 10.1109/SPAWC.2005.1506298.

[22] H. Lou and J. Garcia-Frias, "On the application of error-correcting codes with low density generator matrix over different quantum channels," in *Proc. 4th Int. Symp. Turbo Codes Related Topics*, 2006, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/5755950

[23] Z. Babar et al., "Fifteen years of quantum LDPC coding and improved decoding strategies," *IEEE Access*, vol. 3, pp. 2492–2519, 2015, doi: 10.1109/ACCESS.2015.2503267.

[24] Y. Wang et al., "Enhanced feedback iterative decoding of sparse quantum codes," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1231–1241, Feb. 2012, doi: 10.1109/TIT.2011.2169534.

[25] Z. Babar et al., "Construction of quantum LDPC codes from classical row-circulant QC-LDPCs," *IEEE Commun. Lett.*, vol. 20, no. 1, pp. 9–12, Jan. 2016, doi: 10.1109/LCOMM.2015.2494020.

[26] P. Fuentes et al., "Approach for the construction of non-Calderbank-Steane-Shor low-density-generator-matrix based quantum codes," *Phys. Rev. A*, vol. 102, 2020, Art. no. 012423, doi: 10.1103/PhysRevA.102.012423.

[27] D. Lidar and T. Brun, *Quantum Error Correction*. Cambridge, U.K.: Cambridge Univ. Press, 2013, doi: 10.1017/CBO9781139034807.

[28] N. Raveendran and B. Vasić, "Trapping sets of quantum LDPC codes," *Quantum*, vol. 5, 2021, Art. no. 562, doi: 10.22331/q-2021-10-14-562.

[29] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, 10th Anniversary ed. Cambridge, U.K.: Cambridge Univ. Press, 2011. doi: 10.1017/CBO9780511976667.

[30] T. Brun, I. Devetak, and M. Hsieh, "Correcting quantum errors with entanglement," *Science*, vol. 314, no. 5798, pp. 436–439, 2006, doi: 10.1126/science.1131563.

[31] J. E. Martinez et al., "Approximating decoherence processes for the design and simulation of quantum error correction codes in classical computers," *IEEE Access*, vol. 8, pp. 172623–172643, 2020, doi: 10.1109/ACCESS.2020.3025619.

[32] T. A. Brun, I. Devetak, and M. Hsieh, "Catalytic quantum error correction," *IEEE Trans. Inf. Theory*, vol. 60, no. 6, pp. 3073–3089, Jun. 2014, doi: 10.1109/TIT.2014.2313559.

[33] Z. Babar et al., "Duality of quantum and classical error correction codes: Design principles and examples," *IEEE Commun. Surv. Tut.*, vol. 29, no. 1, pp. 970–1010, Jan./Mar. 2019, doi: 10.1109/COMST.2018.2861361.

[34] P. Fuentes, "Error correction for reliable quantum computing," Ph.D. dissertation, Tecnun- Sch. Eng., Donostia - San Sebastian, Spain, 2022, *arXiv:2202.08060*, doi: 10.15581/10171/63067.

[35] M. M. Wilde, M. Hsieh, and Z. Babar, "Entanglement-assisted quantum turbo codes," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1203–1222, Feb. 2014, doi: 10.1109/TIT.2013.2292052.

[36] D. K. Tuckett et al., "Tailoring surface codes for highly biased noise," *Phys. Rev. X*, vol. 9, no. 4, 2019, Art. no. 041031, doi: 10.1103/PhysRevX.9.041031.

[37] D. K. Tuckett, "Tailoring surface codes: Improvements in quantum error correction with biased noise," Ph.D. dissertation, Univ. Sydney, Camperdown NSW, Australia, 2020, doi: 10.25910/x8xw-9077.

[38] E. Sabo, A. B. Aloshious, and K. R. Brown, "Trellis decoding for qudit stabilizer codes and its application to qubit topological codes," 2021, *arXiv:2106.08251*, doi: 10.48550/arXiv.2106.08251.

[39] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA, USA: Morgan Kauffman, 1988. [Online]. Available: https://www.sciencedirect.com/book/

[40] F. R. Kschischang, B. J. Frey, and H. A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001, doi: 10.1109/18.910572.

[41] K. Kuo and C. Lu, "A further study on the encoding complexity of quantum stabilizer codes," in *Proc. Int. Symp. Inf. Theory Appl.*, 2010, pp. 1041–1044, doi: 10.1109/ISITA.2010.5649496.

[42] M. Wilde, "Logical operators of quantum codes," *Phys. Rev. A*, vol. 79, 2009, Art. no. 062322, doi: 10.1103/PhysRevA.79.062322.

[43] D. J. C. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 399–431, Mar. 1999, doi: 10.1109/18.748992.

[44] T. J. Richardson and R. L. Urbanke, "Efficient encoding of low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 638–656, Feb. 2001, doi: 10.1109/18.910579.

[45] S. Myung, K. Yang, and J. Kim, "Quasi-cyclic LDPC codes for fast encoding," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2894–2901, Aug. 2005, doi: 10.1109/TIT.2005.851753.

[46] M. Battaglioni, P. Santini, M. Baldi, and G. Cancellieri, "Obtaining structured generator matrices for QC-LDPC codes," in *Proc. IEEE AEIT Int. Annu. Conf.*, 2019, pp. 1–6, doi: 10.23919/AEIT.2019.8893395.

[47] M. Baldi et al., "On the generator matrix of array LDPC codes," in *Proc. SoftCOM 20th Int. Conf. Softw.. Telecommun. Comput. Netw.*, 2012, pp. 1–5. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6347593

[48] S. C. Chae and Y.-O. Park, "Low complexity encoding of improved regular LDPC codes," in *Proc. IEEE 60th Veh. Technol. Conf.*, 2004, vol. 4, pp. 2535–2539, doi: 10.1109/VETECF.2004.1400513.

[49] H. Jin, A. D. Kh, and R. J. McEliece, "Irregular repeat-accumulate codes," in *Proc. 2nd Int. Symp. Turbo Codes Related Topics*, 2000, pp. 1–8.

[50] P. Fuentes et al., "Design of LDGM-based quantum codes for asymmetric quantum channels," *Phys. Rev. A*, vol. 103, 2021, Art. no. 022617, doi: 10.1103/PhysRevA.103.022617.

[51] P. Fuentes et al., "Performance of non-CSS LDGM-based quantum codes over the misidentified depolarizing channel," in *Proc. IEEE Int. Conf. Quantum Comput. Eng.*, 2020, pp. 93–101, doi: 10.1109/QCE49297.2020.00022.

[52] M. Jeruchim, "Techniques for estimating the bit error rate in the simulation of digital communication systems," *IEEE J. Sel. Areas Commun*, vol. 2, no. 1, pp. 153–170, Jan. 1984, doi: 10.1109/JSAC.1984.1146031.

Patricio Fuentes received the B.S. and M.S. degrees (Hons.) in telecommunications engineering and the Ph.D. degree in applied engineering from the TECNUN - School of Engineering, University of Navarra, San Sebastian, Spain, in 2017, 2019, and 2022, respectively.

He is currently working as a Postdoctoral Researcher with the Mathematical Principles Group, TECNUN - School of Engineering. His research interests include degenerate quantum codes, implementation of practical QEC codes, low-density parity check codes, and all things quantum. He is also a blockchain enthusiast.

Josu Etxezarreta Martinez received the B.S. (Distinction) and M.S. degrees in telecommunications engineering and the Ph.D. degree in applied engineering from the TECNUN - School of Engineering, University of Navarra, San Sebastian, Spain, in 2016, 2018, and 2022, respectively.

He is currently working as a Postdoctoral Researcher within the Mathematical Principles Group, TECNUN - School of Engineering. His research interests include stabilizer coding, classical and quantum information theory, quantum channels, and degenerate decoding of quantum error correction schemes.

Pedro M. Crespo (Senior Member, IEEE) received the engineering degree in telecommunications from the Universidad Politecnica de Cataluña, Barcelona, Spain, in 1987, and the M.Sc. degree in applied mathematics and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 1983 and 1984, respectively.

From 1984 to 1992, he was a member of the technical staff with the Signal Processing Research Group, Bell Communications Research, Piscataway, NJ, USA, where he worked in the areas of data communication and signal processing. He actively contributed in the definition and development of the first prototypes of xDSL (digital subscriber line transceivers). From 1992 to 1999, he was a district Manager with Telefónica R&D, Madrid, Spain. From 1999 to 2002, he was the Technical Director of the Spanish telecommunication operator Jazztel. Later and until 2009, he became the Director of the Electronics and Communication Department, R&D Center CEIT, San Sebastian, Spain. He is currently a Professor with TECNUN-School of Engineering, University of Navarra, San Sebastian, Spain. His research interests are in the area of information and signal processing with a focus on wireless communications and networks. More recently, his interests also include quantum information theory and coding. He holds seven patents in the areas of digital subscriber lines and wireless communications.

Dr. Crespo is a recipient of Bell Communication Research Award of excellence.

Javier Garcia-Frías (Senior Member, IEEE) received the Ingeniero de Telecomunicación degree from the Universidad Politécnica de Madrid, Madrid, Spain, in 1992, the Licenciado en Ciencias Matemáticas degree from UNED, Madrid, Spain, in 1995, and the Ph.D. degree in electrical engineering from University of California, Los Angeles (UCLA), Los Angeles, CA, USA, in 1999.

In 1992, and from 1994 to 1996, he was with Telefónica I+D, Madrid, Spain. From 1999 to 2008, he was an Assistant and then an Associate Professor with the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE, USA, where he is currently a Professor. His research interests are in the area of information processing in communications and biological systems, with a focus on wireless communications, iterative decoding schemes for source and channel coding, coding for multiterminal sources, joint source-channel coding, analog coding, quantum error correcting codes, and probabilistic techniques in biological systems.

Dr. Garcia-Frías has been a member of the Signal Processing for Communications Technical Committee in the IEEE Signal Processing Society from 2004 to 2006. He was a recipient of a 2001 NSF CAREER Award and a 2001 Presidential Early Career Award in support of his communications program. He has served as an Editor and a Guest Editor for several journals, including the *IEEE Signal Processing Magazine* in 2007, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2003 to 2007, the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2004 to 2006, and the *EURASIP Journal on Bioinformatics and Systems Biology* from 2005 to 2014.