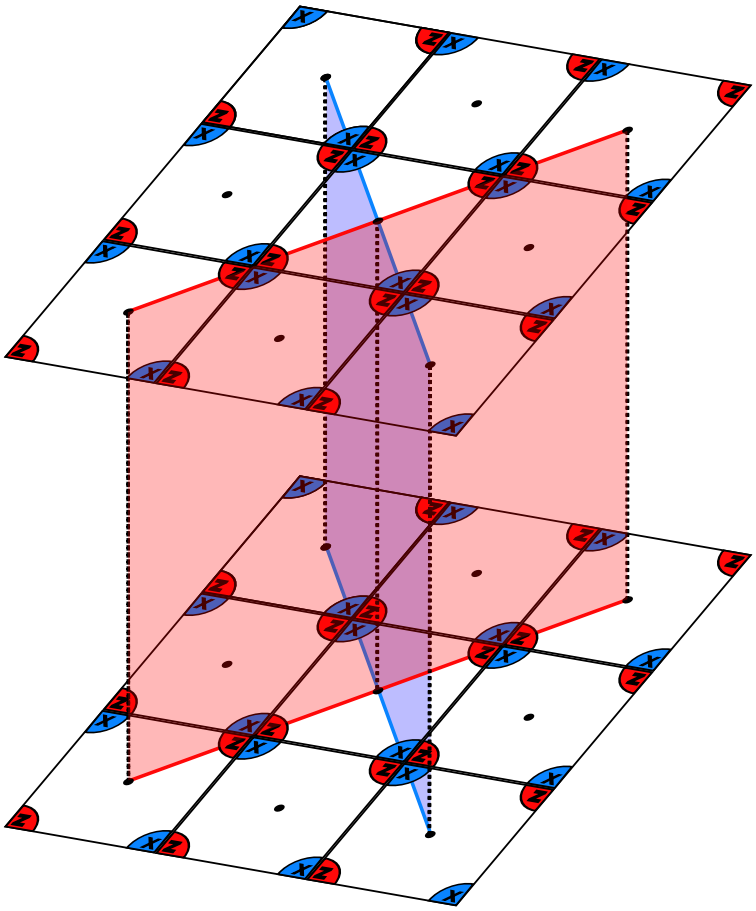

# Gauge-model analysis of the XZZX surface code with biased data noise and measurement errors

Master's thesis in Complex Adaptive Systems

ADRIAN DALMYR

DEPARTMENT OF PHYSICS
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2026
www.chalmers.se

Master's thesis 2026

# Gauge-model analysis of the XZZX surface code with biased data noise and measurement errors

ADRIAN DALMYR


CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Physics
Chalmers University of Technology
Gothenburg, Sweden 2026

Gauge-model analysis of the XZZX surface code with biased data noise and measurement errors
ADRIAN DALMYR



Supervisor: Moritz Lange, Department of Physics, University of Gothenburg
Examiner: Mats Granath, Department of Physics, University of Gothenburg

Master's Thesis 2026
Department of Physics
Chalmers University of Technology
SE-412 96 Gothenburg

Cover: Illustration of temporal invariances in gauge mapping of the XZZX surface code.

Typeset in LaTeX
Gothenburg, Sweden 2026

Gauge-model analysis of the XZZX surface code with biased data noise and measurement errors

ADRIAN DALMYR
Department of Physics
Chalmers University of Technology

# Abstract

Surface codes are a leading approach to fault-tolerant quantum computation, and the XZZX variant is designed to exploit biased noise, in which one Pauli error channel dominates. The decoding threshold of such a code can be studied by mapping the decoding problem to a statistical-mechanics model with quenched disorder, whose order–disorder transition coincides with the threshold. In this work we extend this mapping to the XZZX code subject to dephasing-biased data noise and phenomenological measurement errors, modeled as independent bit-flips on the stabilizer outcomes. Building on the random coupled-plaquette gauge model previously applied to the CSS surface code, we show that the mixed stabilizers of the XZZX code cause measurement errors to couple the $X$- and $Z$-error sectors, resulting in a three-dimensional gauge model for any Pauli noise model. The model is sampled using Metropolis–Hastings with parallel tempering, and a modified Polyakov line is used to probe the deconfinement transition associated with decoding. Simulations are performed at a single code distance $d = 19$ for noise biases $\eta \in \{1, 10, 100, 1000\}$. For weak bias we observe a smooth crossover consistent with an order–disorder transition, but as the bias increases the crossover fades and the system does not disorder throughout the sampled range, preventing threshold extraction. Proposed explanations include finite-size effects, open boundaries, and the limitations of the Polyakov line on a finite lattice. The natural next steps are finite-size scaling and the consideration of a wider range of order-disorder metrics.

Keywords: quantum error correction, surface codes, XZZX code, statistical mechanics mapping, random coupled-plaquette gauge model

# Contents

# 1
# Introduction

Quantum computers provide an alternative to classical computers in solving certain classes of computational problems. Rather than classical bits, which can only exist in a 0 or 1 state, quantum computers leverage quantum bits (qubits) which can exist in a superposition of the $|0\rangle$ and $|1\rangle$ states. By entangling qubits with each other, it is possible to perform calculations on large superpositions of many states, which may speed up the computation of problems for which the superposition can be efficiently leveraged. Quantum computers are not intended to replace classical computers, but rather complement them by solving the aforementioned problems that are classically intractable. Some proposed use cases are number factoring (and by extension encryption breaking) through Shor's algorithm [1], approximating solutions to combinatorial optimization problems through QAOA [2], and simulation of molecular and many-body quantum systems [3].

Currently, a large limitation for quantum computers is the ability to retain information over long calculations. The quantum system is prone to noise arising from several sources, including decoherence and imperfect system preparation, gate application, and readout. To prevent information loss, quantum error correction schemes must be employed. One promising such method is the surface code. Surface codes are a planar implementation of the toric codes suggested by Kitaev, which arose when it became apparent that the toric structure was not necessary [4, 5]. The surface code consists of a square lattice of physical qubits which is used to encode a logical qubit. The physical qubits are divided into data and ancilla qubits. The ancillae act with stabilizers on data qubits, operators which do not affect the encoded state. Stabilizer measurements, the results of which are called the syndrome, are leveraged to detect and prevent errors.

The surface code performs well for localized errors. However, as errors begin to spread through the lattice, error correction becomes a difficult task. Past a threshold error rate, which depends on the code topology and noise channels, arbitrary errors can no longer be corrected by simply increasing the code size. The reason why errors become difficult to correct is that, with large error rates, many different chains have a similar probability of generating the same syndrome. Hence, the physical error configuration cannot be determined with high certainty. To correct logical errors, one may determine equivalence classes corresponding to logical errors. For a given syndrome, error correction may then be performed by picking the most probable equivalence class. When analyzing this probability distribution, the structure of the surface code lends itself to a mapping to a lattice of Ising spins with quenched disorder [6, 7]. This mapping allows the use of known results from statistical mechanics to study logical error rates and threshold behaviors. Errors

may be modeled at several levels of abstraction, each increasing the complexity of the corresponding spin system.
The XZZX code is a variant of stabilizer code designed to perform well under phase-biased noise [8]. Previous analysis by Xiao *et al.* of the XZZX code without measurement errors showed that, at a certain total error probability, the system decouples into a set of 1D Ising chains which simplifies analysis of logical failure rates and threshold values [9]. When including measurement errors in the model, the spin representation extends into a 3D lattice for which it is not straightforward in general to find a non-trivial regime where it decouples.
The goal of this project is to determine finite-size threshold error rates for the XZZX code under the inclusion of phenomenological measurement errors modeled as independent bit-flips on stabilizer qubits. Data noise is biased towards the dephasing channel, with varying bias strengths. The first step is to extend the model by Xiao *et al.* to include measurement noise, done by applying the random-coupled plaquette gauge model introduced by Rispler *et al.* for the CSS surface code [10]. After constructing the system, Metropolis-Hastings with parallel tempering is used in order to probe the system for an order-disorder phase transition. The phase transition point of the lattice gauge model corresponds to the decoding threshold. A finite code distance $d = 19$ is used, such that the sharp phase transitions become smoothed crossovers between the ordered and disordered phases.
Results indicate a crossover for weakly biased noise. As bias is increased, the crossover fades and the system never leaves the ordered state. This behavior makes threshold extraction impossible within the scope of this work except when noise carries only a very weak bias. Limitations in the available data make it difficult to discern how much of the observed behavior is caused purely by noise bias and how much is caused by finite-size effects.

# 2
# Theory

In this chapter we introduce the theory and build intuition required to understand code thresholds and how mapping the surface code to a gauge model can be used to find the threshold values. We begin by briefly introducing qubits and quantum gates. Section 2.2 discusses the difficulties of quantum error correction, and how stabilizer operators can be utilized to detect and correct errors. Surface codes are introduced in Section 2.3. The final section illustrates the mapping of surface codes with perfect measurements and independent Pauli X- and Z-errors to the random-bond Ising model, a spin model which encodes the code's error statistics. The chapter then concludes with the random plaquette gauge model and its extension, which incorporates noisy measurements and real Pauli Y-noise.

## 2.1 Quantum computing

In quantum computing, the qubit acts as the basic unit of information. A qubit is a two-level quantum system $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, where $|0\rangle = (1,0)^\top$ and $|1\rangle = (0,1)^\top$ are basis vectors of the two-dimensional Hilbert space in which $|\psi\rangle$ exists.
The most basic single-qubit gates in quantum computing are the X-, Z- and Y-gates, which in order give bit flips, phase flips, and combined bit and phase flips. These gates are given by their corresponding Pauli matrices,

$$\begin{aligned} X &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ Y &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \\ Z &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \end{aligned} \tag{2.1}$$

The Pauli matrices together with the identity operator $\mathbb{I}$ and the phases $\{\pm 1, \pm i\}$ generate the single-qubit Pauli group, which is defined as

$$\mathcal{P}_1 = \{\pm\mathbb{I}, \pm i\mathbb{I}, \pm X, \pm iX, \pm Y, \pm iY, \pm Z, \pm iZ\}.$$

From Equation (2.1) and the definition of the qubit basis states, it is seen that the Pauli gates act on a general qubit state $|\psi\rangle$ in the following ways,

$$\begin{aligned} X|\psi\rangle &= \alpha|1\rangle + \beta|0\rangle, \\ Y|\psi\rangle &= i\alpha|1\rangle - i\beta|0\rangle, \\ Z|\psi\rangle &= \alpha|0\rangle - \beta|1\rangle. \end{aligned}$$

From this it becomes apparent that X is the quantum analog of a bit flip, as it swaps the amplitudes of the two basis states. We note also that the Z-gate flips the sign of the phase on the $|1\rangle$ basis state, thus introducing a phase difference between the two states, while Y simultaneously flips the amplitudes and introduces a phase difference. Arbitrary single qubit errors can be decomposed into combinations of Pauli operators. Hence, they play a central role in quantum error correction.

## 2.2 Quantum error correction and stabilizer codes

When performing calculations using qubits, there is always a risk of their states being corrupted in some way. These errors result in information loss, and will lead to faulty results. In general, quantum error correction is more difficult than its classical counterpart for several reasons. First, the no-cloning theorem states that it is impossible to create an independent and identical copy of an arbitrary unknown quantum state [11] and therefore quantum information cannot be protected through direct copying of states. Furthermore, the collapse of the wavefunction means that states cannot be directly measured during computation without loss of quantum information. In addition to this, quantum errors are continuous rather than discrete, resulting in an infinitely large set of possible errors.
A simple scheme for mitigating errors in a classical system is the three bit code, a simple setup where a logical bit, used for calculations, is encoded by three physical bits. The logical $0_L$ is thus represented by 000, and correspondingly for $1_L$. In the case of an unwanted bit flip, a majority vote is taken to determine the logical state. In a quantum system, the logical qubit takes the form $|\psi\rangle_L = \alpha|000\rangle + \beta|111\rangle$. A bit-flip error on the first qubit then results in the state $\alpha|100\rangle + \beta|011\rangle$. Hence, the three-qubit code can correct single bit-flip errors, but if two or more errors occur the code fails and a logical error occurs.
We can think of the three-qubit code in terms of stabilizers. A stabilizer of a state $|\psi\rangle$ is an operator $\hat{S}$ which has the property that it leaves $|\psi\rangle$ invariant, i.e. $\hat{S}|\psi\rangle = |\psi\rangle$. Define the code space $C$ as the Hilbert space spanned by $|0\rangle_L$ and $|1\rangle_L$. Then we consider the stabilizer group $\mathcal{S}$ of our code space, defined as

$$\mathcal{S} = \{\hat{S} : \hat{S}|\psi\rangle = |\psi\rangle,\ \forall|\psi\rangle \in C\}.$$

Note two obvious properties of the stabilizer group. Firstly, the product of two stabilizers is itself a stabilizer, since $(\hat{S}_1\hat{S}_2)|\psi\rangle = \hat{S}_1(\hat{S}_2|\psi\rangle) = \hat{S}_1|\psi\rangle = |\psi\rangle$. Secondly, the negative identity operator $-\mathbb{I}$ is not a stabilizer, as $-\mathbb{I}|\psi\rangle = |\psi\rangle$ if and only if $|\psi\rangle = 0$.
We noted that single qubit measurements result in a loss of quantum information. In order to understand how stabilizers are useful in resolving this issue, consider the three-qubit code. Consider the product operators $Z_1Z_2$ and $Z_2Z_3$. Note that these are tensor products, where the subscript indicates which qubit the operator acts on and the identity $I$ acting on every other qubit:

$$Z_1 = Z \otimes I \otimes I,$$
$$Z_2 = I \otimes Z \otimes I,$$
$$Z_3 = I \otimes I \otimes Z.$$

The product operators are then ordinary operator products on the same Hilbert space, and multiplying gives

$$Z_1Z_2 = (Z \otimes I \otimes I)(I \otimes Z \otimes I) = Z \otimes Z \otimes I,$$
$$Z_2Z_3 = (I \otimes Z \otimes I)(I \otimes I \otimes Z) = I \otimes Z \otimes Z.$$

If we apply the operators on the state $|\psi\rangle_L = \alpha|000\rangle + \beta|111\rangle$ (no errors), we get

$$\begin{aligned} Z_1Z_2|\psi\rangle_L &= Z_1(\alpha|000\rangle - \beta|111\rangle), \\ &= \alpha|000\rangle + \beta|111\rangle, \\ &= |\psi\rangle_L, \end{aligned}$$

and we note that $Z_2Z_3$ will yield the same result. Hence, the two-qubit operators act as stabilizers for $|\psi\rangle$. The encoded state is then an eigenstate with eigenvalue $+1$ of both stabilizer generators. If a Pauli $X$-error has occurred on one of the qubits, however, it anticommutes with the corresponding stabilizer and becomes a stabilizer eigenstate with eigenvalue $-1$. For example, if the first qubit experiences an error we can write it as $|\psi\rangle \to X_1|\psi\rangle$. Measuring $Z_1Z_2$ then yields

$$\begin{aligned} Z_1Z_2(X_1|\psi\rangle) &= Z_1X_1(Z_2|\psi\rangle), \\ &= -X_1(Z_1Z_2|\psi\rangle), \\ &= -X_1|\psi\rangle, \end{aligned}$$

and we see that $X_1|\psi\rangle$ is indeed an eigenstate of $Z_1Z_2$ with eigenvalue $-1$. In this case the parity checks will return different results and it can be determined which qubit has experienced an error without disturbing the superposition.
As stated before, the three-qubit code only protects against bit-flips. An extension of the three-qubit code is the Shor code, where a logical qubit is encoded using nine physical qubits [12]. By cleverly using blocks of three qubits and measuring stabilizers, it can protect against arbitrary errors. Unfortunately, real systems have an error rate which is too high for the Shor code to reliably protect against logical errors. Therefore more complex error correcting schemes must be employed, such as surface codes.

## 2.3 Surface codes

In modern quantum computing, surface codes are a prominent example of stabilizer codes. These utilize two classes of physical qubits, data qubits and ancilla/measurement qubits, to encode error-protected logical qubits used in computation. Typical surface codes are defined on a 2-dimensional lattice consisting of $d \times d$ physical data qubits, which together encode a single logical qubit. The code distance $d$ is the minimum weight of logical operators, where the weight is the number of qubits acted on non-trivially. Logical errors correspond to chains of physical errors which together act as logical operators, so the code distance gives the minimum number of physical errors that create an undetectable logical error.

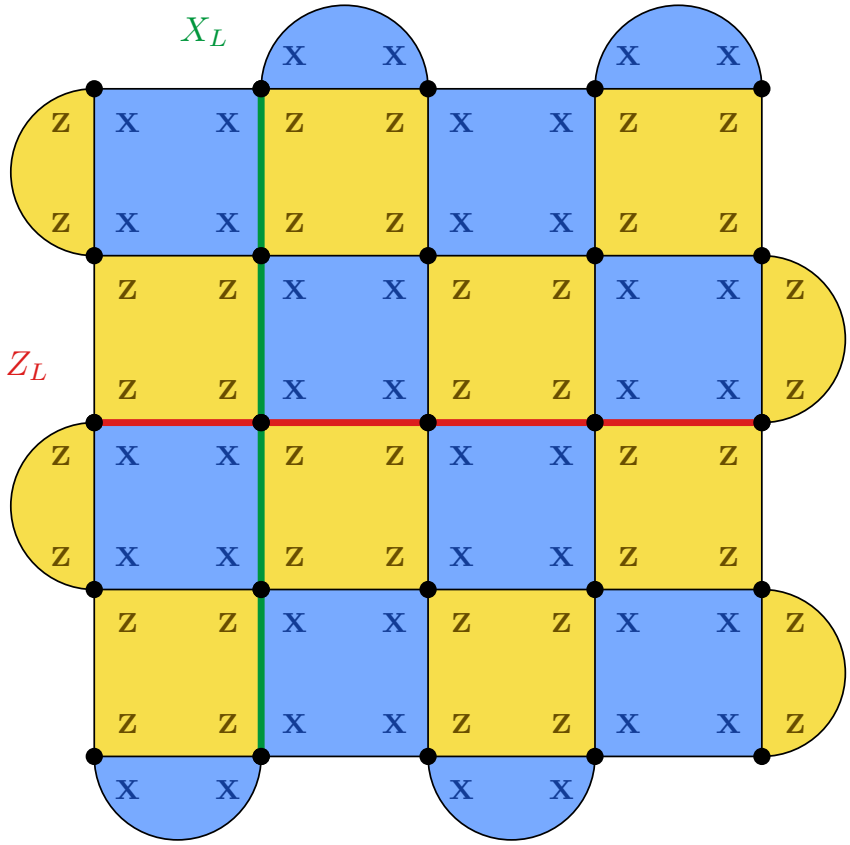


**Figure 2.1:** Example of the CSS surface code with distance $d = 5$. The two different types of ancilla (measurement) qubits, which live on faces, act with either $X^{\otimes 4}$ or $Z^{\otimes 4}$ stabilizers on the surrounding data qubits on vertices. Data errors are detected as they anticommute with the opposite type of stabilizer, creating a syndrome. The decoding task consists of determining the most likely error chain which results in a given syndrome. Illustrated are logical operators $Z_L$ and $X_L$, spanning opposite boundaries. Any error chain which connects two opposite (left/right or top/bottom) boundaries results in a logical operator acting on the system, creating no syndrome.

The most common surface code is shown in Figure 2.1, where each measurement qubit is surrounded by four data qubits. This code will be referred to as the CSS surface code for the remainder of this work, after Robert Calderbank, Peter Shor, and Andrew Steane, who invented the wider class of stabilizer codes to which this surface code geometry belongs [13, 14]. The specific stabilizer structure of the CSS code was originally defined by Alexei Kitaev on a torus, before he together with Sergey Bravyi later found that a planar geometry works as well [4, 5]. Each ancilla qubit measures a stabilizer generator which defines the code space, either $X_i X_j X_k X_l$ or $Z_i Z_j Z_k Z_l$, consisting of tensor products of single-qubit Pauli operators acting on the adjacent qubits. Neighboring X- and Z-type stabilizers overlap on either zero or two qubits, which implies commutation. Using the code, it becomes possible to identify when an error occurs, and for low error rates correction is straightforward as single errors will dominate the probability distribution. For example, consider an

X-error on a qubit $k$ in the bulk. An X-type stabilizer then acts as

$$\begin{aligned}\hat{X}_i\hat{X}_j\hat{X}_k\hat{X}_l(\hat{X}_k|\psi\rangle) &= \hat{X}_k(\hat{X}_i\hat{X}_j\hat{X}_k\hat{X}_l|\psi\rangle),\\ &= \hat{X}_k|\psi\rangle.\end{aligned}$$

Hence, $\hat{X}_k|\psi\rangle$ is also an eigenstate of the stabilizer, with the same eigenvalue. If instead a Z-error were to occur, the picture would differ slightly. Since it anti-commutes with $\hat{X}$, the outcome would still be an eigenstate, but the measurement would flip sign. From this it becomes clear that an error has occurred. Conversely, in the case of an X-error, the measurement outcome of Z-type stabilizers flips. Since Y anticommutes with both X and Z, both types of stabilizer flip outcomes. The measured sign changes constitute the error syndrome. By analyzing the syndrome pattern, it can be inferred which qubit experienced the error and it can be corrected. It should be noted that physically correcting errors may simply introduce more errors by applying noisy gates. As such, the surface code generally serves primarily to detect errors which are then corrected in classical software.
When there are multiple adjacent errors, the syndrome is nontrivial only at the chain's endpoints. In these cases, there are generally many distinct error chains with similar probabilities of occurring that all give rise to the observed syndrome; the longer the chain distance, the greater the number of candidates. After an error $E$, a correction $C$ is applied to the code. While $E$ cannot be known exactly, all we need to know is how the total operator $CE$ acts on the code. If it is a product of stabilizers, it acts trivially on the encoded qubit, which is brought back into the code space. If instead $CE$ acts non-trivially on the code, we suffer a logical error. Errors are therefore grouped into two classes, trivial and non-trivial, where two error chains are equivalent when they differ by a product of stabilizers, and inequivalent when they differ by a logical operator.
In general there exists no efficient method for perfect decoding. Common methods are minimum-weight perfect matching (MWPM) and maximum-likelihood decoding (MLD) [6, 15]. MWPM is used primarily for independent X- and Z-noise, and works by finding the most likely error configuration corresponding to an observed syndrome. On the other hand, MLD considers logically equivalent configurations and picks the correction based on the most likely logical class.
A key concept in error correction is the threshold error rate, which refers to the maximum single-qubit error rate for which the logical error rate can be reduced to an arbitrarily small number by increasing the code distance. The threshold varies depending on code geometry, error types, and other factors. Hence, characterization of a physical surface code becomes important in order to understand its limitations.

### 2.3.1 The XZZX surface code

Quantum computers may be designed such that one error channel, X or Z, dominates. In particular, dephasing processes often produce $Z$-errors at a much higher rate than $X$-errors. Since the CSS code treats the two error types independently and symmetrically, it does not exploit the difference in rates. The XZZX surface code was designed in order to increase error correction capability on such single-mode dominated error chains by causing them to preferentially align along 1-dimensional

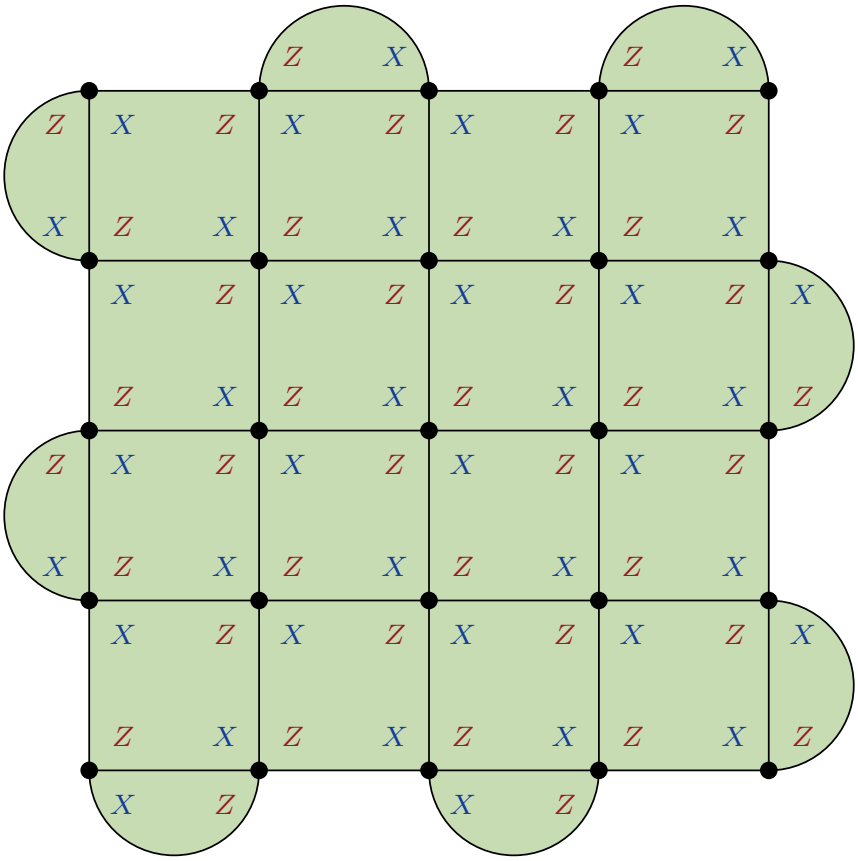


**Figure 2.2:** Layout of the XZZX surface code. The difference from the CSS code is that it contains mixed stabilizers, which restricts pure $X$- and $Z$-errors to one dimension. This creates an advantage when noise is biased through one channel by decreasing the entropy of possible error chains, therefore increasing the threshold for biased noise.

chains. Like the CSS code, it is defined on a square lattice, with $d \times d$ data qubits residing on vertices and stabilizers on faces (Figure 2.2). Hence there are $d^2$ data qubits and $(d-1)^2$ weight 4 stabilizers. In addition, the lattice contains $2(d-1)$ weight 2 stabilizers along the boundaries, placed carefully such that each data qubit neighbors at least two stabilizers while preserving a single encoded logical qubit. The difference from the CSS code is that each 4-stabilizer acts on the four neighboring qubits $i$, $j$, $k$, $l$ with $X_i Z_j Z_k X_l$, the arrangement chosen such that Z-errors (X-errors) are detected right-diagonally (left-diagonally). Y-errors create syndromes along both directions. In the limit of large bias, this effectively leads to weakly coupled repetition codes running along the preferred diagonal. Decoding is then simplified by reducing the number of candidate error chains. This improves decoding performance and raises the threshold when handling biased noise.

## 2.4 Statistical mechanics mapping

While minimum-weight perfect matching and maximum likelihood decoding are strong tools for decoding surface codes, they limit code characterization to numerical analysis. In order to gain analytical insight, a well-explored method is to map the decoding problem to statistical mechanics models. In this picture, the inference problem of decoding by determining the most logical error that results in a given syndrome is transformed into an energy minimization problem on a spin lattice. While decoding schemes such as MLD and MWPM can be used to approximately characterize a surface code, the spin model has the advantage of exactly representing the code error distribution and can therefore better estimate theoretical thresholds and logical error rates in situations where it is applicable. In some limiting cases it is even possible to find closed form expressions for the logical error rates [9]. Here we focus on two models, the random-bond Ising model (RBIM) and the random plaquette gauge model (RPGM) and its extension, the random coupled-plaquette gauge model (RCPGM). For the remainder of this section we will consider the CSS code, as that is the setting in which previous literature has worked. Mapping onto the XZZX code is done in Chapter 3.

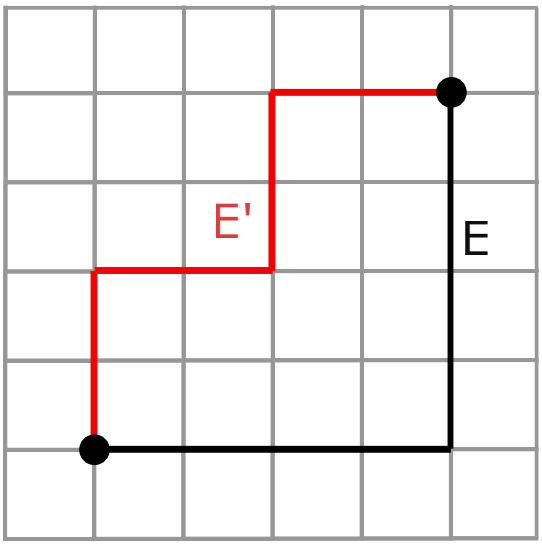


**Figure 2.3:** Representation of a surface code with qubits on edges and stabilizers on vertices. An error chain is given by a series of "activated" edges. Two different error chains $E$ and $E'$ result in the same syndrome, given by the chain boundaries. The total operator $E + E'$ forms a closed loop in the bulk, and therefore corresponds to a product of stabilizers acting trivially on the code.

### 2.4.1 Random-bond Ising model

We begin by assuming perfect stabilizer measurements and Y-errors represented as combinations of independent X- and Z-errors. In this case, the CSS code can be decomposed into independent X- and Z-sectors which may be treated separately.

For this reason, we proceed to work only with the Z-sector. The aim is to build some intuition for the model before presenting the final results. For reference, the full derivation is presented in [6].
The Z-sector may be represented on a square lattice where data qubits live on edges and stabilizers on vertices (Figure 2.3). The goal is to map the code to a spin model where error information is encoded in couplings and spin configurations occur with a probability that is determined by the error class which they belong to. Error chains can be represented as chains of "activated" edges on the lattice. The chain's boundary, i.e. the endpoints where the chain terminates, is the syndrome. For determining the threshold, the exact logical error classes are not of interest. Instead, we can consider simply whether an error chain is trivial or non-trivial.
For each error chain $E$, we can define a function $n_E(l)$ for every edge $l$ on the lattice,

$$n_E(l) = \begin{cases} 0, \text{ if } l \notin E, \\ 1, \text{ if } l \in E. \end{cases}$$

The probability of a given chain $E$ is then given by

$$p(E) = \prod_l (1-p_z)^{1-n_E(l)} p_z^{n_E(l)},$$

where $p_z$ denotes the probability of a Z-error. For a fixed syndrome caused by an error chain $E$, decoding consists of determining the relative probabilities of all other chains $E'$ sharing the same boundary. By doing so, the most likely error class of the syndrome can be determined. Any chain $E'$ may be represented as $E' = E + C$ where $C$ is some particular closed loop, since a closed loop has no boundary and therefore does not shift the syndrome. By introducing the appropriate variables, this setup allows determining the conditional probability distribution [6],

$$p(E'|E) \propto \prod_l e^{J_z u_l}.$$

The new variables are given by

$$u_l = 1 - 2n_C(l), \tag{2.2}$$

$$e^{-2J_z} = \begin{cases} \frac{p_z}{1-p_z}, \text{ if } l \notin E, \\ \frac{1-p_z}{p_z}, \text{ if } l \in E. \end{cases} \tag{2.3}$$

The coupling strength is therefore not an independent parameter, but defined by the error probability. The relation is known as the Nishimori condition, under which the thermodynamic averages of the ensemble correspond exactly to averages over the decoding problem. In addition, $n_C(l) = 1$ implies that $l$ has no boundary which introduces a constraint on the variables $u_l$. Solving the constraint involves considering the dual lattice. Edges on the original lattice are mapped to vertices on its dual, and sites to plaquettes. Ising variables $\sigma_i \in \{-1, 1\}$ are introduced on the dual plaquettes. The constraint then leads to a solution where $u_{ij} = \sigma_i \sigma_j$ ($u_{ij}$ here denotes a link connecting vertices $i$ and $j$ on the dual lattice). It follows that the distribution of error chains $E'$ that create the same syndrome as $E$ is described by a spin model with a partition function given by

$$Z(J_z, \eta_l) = \sum_{\{\sigma_i\}} \exp\left(J_z \sum_{\langle ij \rangle} \eta_l \sigma_i \sigma_j\right), \tag{2.4}$$

where $\langle ij \rangle$ runs over pairs of nearest-neighbors. The $\eta_l$ are quenched disorder variables, drawn according to

$$\eta_l = \begin{cases} +1, \text{ with probability } 1 - p_z, \\ -1, \text{ with probability } p_z, \end{cases}$$

and represent the physical error distribution. On the original lattice, the system is given by spin variables $\sigma_i$ residing on the plaquettes, coupled to the nearest neighbors through shared edges which mediate $J_z$. Equation (2.4) is known as the random-bond Ising model, which has been extensively studied in the field of spin glasses. In the mapping of the 2D surface code to the RBIM, physical error chains are represented by domain walls on the lattice along which there is an antiferromagnetic interaction. Decoding is always possible by increasing the lattice size as long as these domains remain localized in the bulk. At some error probability $p_z$, the system undergoes an order to disorder phase transition, and the domains go from remaining localized to spreading through the system. The decoding threshold is associated with the phase transition in the RBIM.

### 2.4.2 Noisy measurements and the random plaquette gauge model

The RBIM assumes perfect stabilizer measurements and therefore only applies to an idealized decoding problem. When noisy measurements are introduced, decoding must include both spatial (data) and temporal (measurement) error correlations. This extends the two-dimensional lattice of the RBIM to a three-dimensional spacetime lattice with higher order interactions, a model called the random plaquette gauge model (RPGM). No derivation of the RPGM is presented here, as the geometric intuition behind it is not much different from the RBIM. In the spacetime picture, each horizontal layer represents the data qubits at a single round of stabilizer measurements, while the vertical axis represents time. Data errors therefore produce differences between stabilizer measurements within time slices, while measurement errors create inconsistent measurements of single stabilizers across time steps.

The RPGM geometry is understood in close analogy with the RBIM. Measurement errors are introduced on the vertical links of the 3D lattice. Similar to how spacelike plaquettes represent the invariance of applying a stabilizer, vertical plaquettes represent the invariance of a data qubit error followed by two measurement errors, then another data error on the same qubit. This is a chain of events which produces no syndrome and forms a closed loop in the lattice. The RPGM is typically given in the dual space representation, such that plaquettes mediate interactions between edges on which the spin variables live. In this picture the notion of timelike and spacelike variables gets reversed, such that data errors reside on vertical faces and

measurement errors on horizontal. The Hamiltonian under the dual representation is

$$H = \sum_{\square} J_{\square} \prod_{\sigma \in \square} \sigma, \tag{2.5}$$

where the sum runs over all plaquettes $\square$, $J_{\square}$ is the corresponding interaction, and the product over the edges belonging to the plaquette. Unlike the RBIM, where interactions occur between neighboring spins, the RPGM contains plaquette interactions consisting of spin products around faces of the lattice. Due to this, the model has a local gauge symmetry consisting of flipping all the spins adjacent to a vertex, which leaves the Hamiltonian invariant. This is physically interpreted as the redundancy in different error histories producing the same syndrome.
In the RPGM the phase transition is no longer characterized by delocalization of magnetic domains, and magnetization does not track the phase transition. Instead, the relevant order transition is a confinement-deconfinement transition, which characterizes whether extended errors spread through the bulk. The threshold is associated with the onset of the deconfined regime, where these extended error chains produce logical errors with finite probability. There exist several order parameters measuring deconfinement, and in this work we implement a version of the Polyakov loop [10, 16]. In short, the Polyakov loop consists of a product of spin variables at a fixed lattice site $i$ along the time axis,

$$P_i = \prod_t \sigma_{i,t}. \tag{2.6}$$

In periodic systems, the Polyakov loop acts as a deconfinement order parameter by probing correlations along the temporal axis. In the confined phase the expectation value of the Polyakov loop will remain near 1, whereas in the deconfined phase it vanishes in the thermodynamic limit. The implementation of the Polyakov loop for our system is discussed in Chapter 3, as well as the eventual limitations it imposes.

### 2.4.3 The random coupled-plaquette gauge model

The RPGM models independent $X$- and $Z$-sectors, and therefore does not represent correlated noise through Pauli $Y$-errors. The RCPGM resolves this problem by introducing an explicit coupling between the two sectors, which then reside in a shared spacetime structure. Geometrically, this is done by superposing the two lattices and introducing a coupling at the intersections (Figure 2.4).
This introduces an 8-body interaction term $H_y$ in the Hamiltonian, where the error sectors of qubits at the same physical location are coupled together,

$$H_y = J_y \sigma_X \sigma_X \tilde{\sigma}_X \tilde{\sigma}_X \sigma_Z \sigma_Z \tilde{\sigma}_Z \tilde{\sigma}_Z.$$

Here, $J_y$ is given by the Nishimori condition and the spin variables with a tilde denote spin variables corresponding to temporal invariances. In the presence of coupled $X$-, $Y$-, and $Z$-error channels, the Nishimori condition (2.3) takes the form

$$e^{-4|J_E|} = \frac{p_x p_y p_z}{p_E^2 p_I}, \tag{2.7}$$

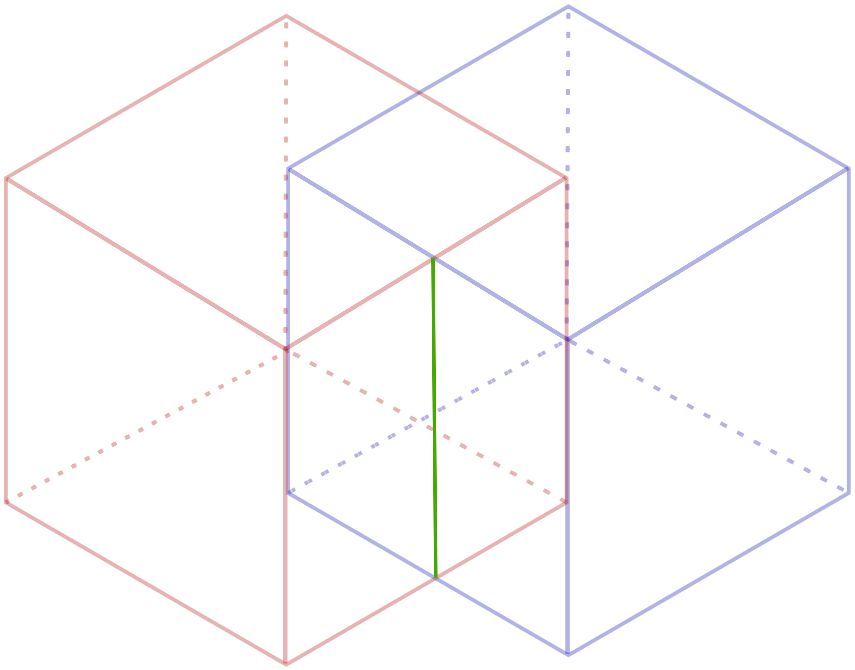

**Figure 2.4:** Illustration of the error sector superposition used in the RCPGM. The two lattices representing each sector are overlaid such that the plaquettes that correspond to the same physical data qubit intersect. One such line of intersection is shown in green. Pauli $Y$-errors act on all edges that belong to the intersecting plaquettes, coupling the $X$- and $Z$-sectors. In the Hamiltonian this corresponds to insertion of a $J_y$ coupling term, creating an 8-body interaction between the sectors.

where $E \in \{X, Y, Z\}$ and $p_I$ denotes the probability of no error. On the dual lattice, the full system Hamiltonian takes the form

$$\begin{aligned} H = &- \sum_{\square_t} (J_x \sigma_X \sigma'_X \tilde{\sigma}_X \tilde{\sigma}'_X + J_y \sigma_X \sigma'_X \tilde{\sigma}_X \tilde{\sigma}'_X \sigma_Z \sigma'_Z \tilde{\sigma}_Z \tilde{\sigma}'_Z + J_z \sigma_Z \sigma'_Z \tilde{\sigma}_Z \tilde{\sigma}'_Z) - \\ &- \sum_{\square_s} J_q (\tilde{\sigma}_X \tilde{\sigma}'_X \tilde{\sigma}''_X \tilde{\sigma}'''_X + \tilde{\sigma}_Z \tilde{\sigma}'_Z \tilde{\sigma}'''_Z \tilde{\sigma}'''_Z), \end{aligned} \tag{2.8}$$

where the couplings are given by Equation (2.7) and the sums run over time-like plaquettes $\square_t$ and space-like plaquettes $\square_s$. Primed variables indicate edges belonging to a shared plaquette.

# 3
# Methods

In order to determine the threshold of the XZZX code, we aim to locate the probability $p$ where the code undergoes an order-disorder crossover under a mapping to the RCPGM. This will vary for different bias amplitudes $\eta$ (threshold increases with $\eta$ due to decoupling between the X- and Z-error sectors), so different values of $\eta$ are tested. Finding the threshold error rate first requires determining an appropriate order parameter which undergoes a transition as the threshold is passed. Due to smoothing of the transition caused by finite-size effects, exact location of the threshold becomes impossible. Instead, the extracted values can be interpreted as an upper bound on the true threshold. It is of high interest to determine the finite-size effects on the system, but due to computational limitations it is not done in this work. Instead, after determining the statistical mechanics mapping of the XZZX code we attempt to determine whether an order-disorder crossover can be found. To this end, we use the Polyakov line as our order parameter. The most direct way to find a threshold is by considering the behavior of the order parameter on the Nishimori line. As an alternative method we consider the system's behavior in the full thermodynamic ensemble, since it is expected to have a crossover line which intersects the Nishimori line, since the system is defined such that $T = 1.0$ corresponds to the Nishimori condition.

## 3.1 XZZX and the random-coupled plaquette gauge model

Since the XZZX code contains identical stabilizers of mixed types, they cannot be divided into X- and Z-type sublattices unlike the CSS code. There is however another sublattice structure to be considered. Let $S_i$ be one of the $N$ stabilizers which resides on lattice position $i \in \{0, 1, 2, ..., N-1\}$ beginning from the top left. Then we can define sublattice A as the one consisting of even index stabilizers, and B is given by odd indices. Due to the identical stabilizer geometries, the two sublattices follow the same error statistics. Their usefulness is in defining the RCPGM mapping, since each qubit experiences only one error type per sublattice. Hence, the mapping to the RPGM becomes straightforward, and the effect of Y-errors is, similar to the CSS case, to couple A and B together. As a result of this structure, we only have one type of spatial spin $\sigma$ which interacts through all coupling types. There are still two distinct temporal spins due to the invariances given by either an X- or Z-error followed by two measurement errors and another data error of the same type, $\tilde{\sigma}_X$ and $\tilde{\sigma}_Z \in \{-1, 1\}$, but they are no longer separable in the interaction through

$J_q$. Figure 3.1 illustrates the temporal invariances, on the faces of which live the timelike spin variables. The dual space Hamiltonian takes the following form, where the sums run over spatial plaquettes $\square_s$ and temporal plaquettes $\square_t$

$$\begin{aligned} H = &- \sum_{\square_t} (J_x \sigma\sigma' \tilde{\sigma}_X \tilde{\sigma}'_X + J_y \sigma\sigma'\sigma''\sigma''' \tilde{\sigma}_X \tilde{\sigma}'_X \tilde{\sigma}_Z \tilde{\sigma}'_Z + J_z \sigma\sigma' \tilde{\sigma}_Z \tilde{\sigma}'_Z) - \\ &- \sum_{\square_s} J_q \tilde{\sigma}_X \tilde{\sigma}'_X \tilde{\sigma}_Z \tilde{\sigma}'_Z. \end{aligned} \tag{3.1}$$

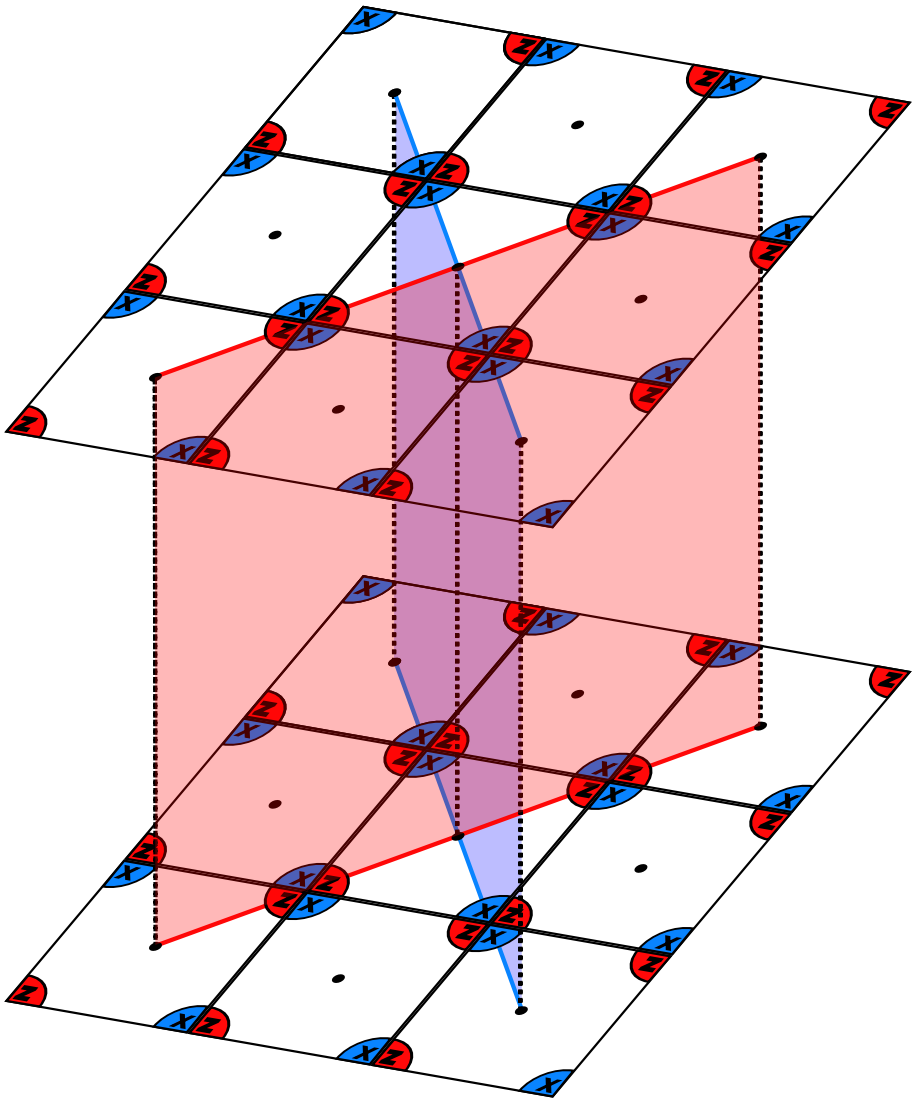

**Figure 3.1:** Illustration of the temporal invariances in the XZZX code. Horizontal stabilizer layers represent physical stabilizers at different measurement time steps. Spatial spin variables (dots) live on the faces of spatial layers and are coupled either through $J_x$ (blue) or $J_z$ (red). Timelike spins live on the faces of the vertical two-dimensional surfaces bounded by two spatial and two temporal couplings (dashed lines) and represent invariances given by the corresponding qubit error, followed by two faulty stabilizer measurements, and finally another qubit error.

The noise model is defined such that the dephasing (Pauli $Z$) channel dominates. We define a bias parameter $\eta$,

$$\eta = \frac{p_z}{p_x + p_y},$$

which gives a relationship between error channels $p_x = p_y = p_z/(2\eta)$. The total error probability is therefore $p = p_x + p_y + p_z = (1 + \eta^{-1})p_z$. A phenomenological measurement error model is used, where an erroneous measurement causes a sign change on the measured value. These errors occur with a probability $q$, which we set to $q = p_x + p_z$. This noise model is the one used by Ataides *et al.* [8], which allows for direct comparison of results. Couplings are given by the generalized Nishimori condition (2.7), which in the given noise model reads

$$J_x = -\frac{1}{4}\ln\frac{p_x(1-p)}{p_y p_z} = -\frac{1}{4}\ln\frac{1-p}{p_z},$$

$$J_y = -\frac{1}{4}\ln\frac{p_y(1-p)}{p_x p_z} = -\frac{1}{4}\ln\frac{1-p}{p_z},$$

$$J_z = -\frac{1}{4}\ln\frac{p_z(1-p)}{p_x p_y} = -\frac{1}{4}\ln\frac{4\eta^2(1-p)}{p_z},$$

$$J_q = -\frac{1}{2}\ln\frac{1-q}{q}.$$

### 3.1.1 Order parameter and phase transition

In order to find where the system undergoes a deconfinement phase transition, the first step is to determine an order parameter which undergoes a transition when moving from the ordered to the disordered phase. In this thesis, the order parameter is a version of the Polyakov loop (2.6). Since the lattice has fixed boundaries, $P$ no longer forms a closed loop but instead becomes a line that terminates at the temporal boundaries,

$$\bar{P} = \frac{1}{N_s}\sum_i \prod_{t=1}^{T} \sigma_{i,t}.$$

where the lines are averaged over the total number of lattice sites $N_s$ and $T$ is the total time, set to $T = d$. Local gauge invariance is retained on the finite system, a requirement for any order parameter in a gauge model. However, the topological difference between a closed loop and a finite chain creates uncertainty in whether the finite-size $P$ correctly tracks the deconfinement transition, and the results must be interpreted with this in mind. We can note that the local gauge invariance,

which consists of flipping the edges around a vertex, still holds in the bulk of the finite lattice. Since the fixed boundaries additionally introduce finite-size effects which smooth phase transitions, finite-size analysis should generally be performed by treating a number of different code distances. In this work only a single code distance of $d = 19$ is considered due to computational limitations, meaning that the results can at best be considered an upper bound.
In addition to the Polyakov loop, we consider its susceptibility $\chi_P$ and the Binder cumulant $U_4$,

$$\chi_P = \left\langle \left( |\bar{P}| - \langle |\bar{P}| \rangle \right)^2 \right\rangle,$$

$$U_4 = 1 - \frac{\langle \bar{P}^4 \rangle}{3\langle \bar{P}^2 \rangle^2}.$$

The susceptibility is a measure of per-configuration thermal fluctuations. When the system undergoes a transition, these fluctuations grow large and the susceptibility peaks. A crossover between phases is therefore indicated by a peak in the $p$-$\chi$ diagram. The Binder cumulant is a measure kurtosis, or the tailedness of a distribution, and is commonly used when varying lattice sizes, as it is insensitive to boundary effects. It remains a useful diagnostic now, as it takes its maximum value of 2/3 in the ordered state and becomes zero in the disordered state. Another quantity which we consider is the system's integrated autocorrelation time $\tau$,

$$\tau = \frac{1}{2} + \sum_{k=1}^{K} \frac{C(k)}{C(0)},$$
$$C(k) = \frac{1}{N-k} \sum_{t=1}^{N-k} (A(t) - \langle A \rangle)(A(t+k) - \langle A \rangle),$$

where $A$ is the observable of interest (the Polyakov line), $t$ is Metropolis-Hastings steps, N the total number of samples, and $K$ is determined by a stopping criterion. In this work, the stopping criterion was simply set to the first instance of negative autocorrelation, as that is typically caused by noise. The autocorrelation cannot serve directly as an order parameter. However, it can help identify critical slowing down in the system which typically occurs around the transition point. When the system slows down it takes a very long time to de-correlate between steps, showing as a spike in the autocorrelation times.

### 3.1.2 Metropolis-Hastings setup

Sampling of the system was done through the Metropolis-Hastings algorithm. Each chain was initialized with quenched disorder drawn from the error probability distribution, such that each spatial coupling flipped signs with probability $p$. The temporal couplings were similarly flipped with probability $q = p$. After initialization, each disordered lattice was allowed to explore spin configurations in order to minimize the energy. For large values of the noise parameter $\eta$, single flips are accepted with very small probability, which makes configuration space exploration

take long. In order to help this, chains of neighboring spins were suggested as flips instead. Since neighbors along the chain see no relative sign change, they do not incur any energy change in that direction except for at the endpoints. There is still an energy cost along the direction orthogonal to the chain, associated with the measurement coupling $J_q$. For large values of $\eta$, this becomes relatively small if chains are constructed in the Z-direction where large $J_z$ at endpoints dominate the cost.
Even when utilizing string flips, exploration of the configuration space is very slow for heavily Z-biased noise. This is worsened by the fact that near the phase transition, the system undergoes critical slowing down and autocorrelation times diverge. As a result, independent sampling of the system becomes inefficient and time consuming. To aid with this, we implement parallel tempering, which helps the system more efficiently explore uncorrelated states. The idea is to run several identical replicas, which sit at different temperatures. The temperature changes the distribution of configurations and enters the system through the Metropolis criterion,

$$P_{\mathrm{MH}} = \min\left(1, \exp(-\Delta E/T)\right).$$

Higher temperature means a larger acceptance rate, while low temperature leads to more rejections. We should note that $T$ is not a physical temperature. Rather, it is a parameter that we can tune to modify the distribution of states. Only the $T_N = 1.0$ state, corresponding to the Nishimori line, corresponds to the decoding problem. Replicas away from the Nishimori line do not represent the decoding configuration, but probe the thermodynamic phase structure of the gauge model. The replicas are allowed to independently explore the configuration space for a number of steps, and configurations are then proposed to be swapped with neighboring replicas in the temperature ladder. The acceptance rate obeys a Metropolis criterion,

$$P_{\mathrm{swap}} = \min\left(1,\ \exp\left(\frac{1}{T_i} - \frac{1}{T_j}\right)(E_i - E_j)\right).$$

In this manner, the high temperature replicas facilitate exploration of configuration space. We aim for at least a 20% acceptance rate for replica swaps, such that the parallel tempering does not become bottlenecked which would prevent mixing of replicas. The temperature range and spacing were determined through short test runs tracking the swap rates, and then fixed during the final production run. Per-realization observables were measured as batched averages over Metropolis chains, with batch sizes determined by autocorrelation times, then averaged over $N_d$ disorder realizations,

$$\langle A \rangle = \frac{1}{N_d} \sum_{k=1}^{N_d} \langle A \rangle_{\mathrm{thermal}}.$$

Each realization was set to 50 000 burn-in sweeps followed by 50 000 production sweeps, where one sweep meant suggesting a number of new configurations equal to the number of spins. The reason for having such a large proportion of burn-in steps is the low acceptance rate for large noise biases, which leads to slow equilibration.

**Table 3.1:** Summary of parameters used in the Metropolis-Hastings chains.

| **Parameter** | **Description** | **Value** |
|---|---|---|
| $d$ | Code distance | 19 |
| $N_d$ | Disorder realizations | 50 |
| $\eta$ | Noise bias parameter | 1, 10, 100, 1000 |
| $p$ | Physical error probability | [0, 0.1) |
| $q$ | Measurement error probability | $p_x + p_z$ |
| - | Number of temperature replicas | 80 |
| - | Minimum temperature | 0.5 |
| - | Maximum temperature | 2.0 |
| - | Burn-in sweeps | $5 \times 10^4$ |
| - | Production sweeps | $5 \times 10^4$ |
| - | Parallel tempering interval | 30 sweeps |
| - | Target swap acceptance rate | $\sim 20\%$ |

After every 30 sweeps, a parallel tempering step was taken to allow for mixing of the replicas. Table 3.1 summarizes the parameters.
Statistical uncertainties were estimated by bootstrapping over the disorder realizations. From each set of realizations, 1000 resamples were drawn with replacement, the full estimator recomputed on each, and the reported error bars are taken as the standard deviation of the resulting bootstrap distribution. The bootstrap captures disorder-induced fluctuations.

# 4

# Results

## 4.1 Polyakov line behavior

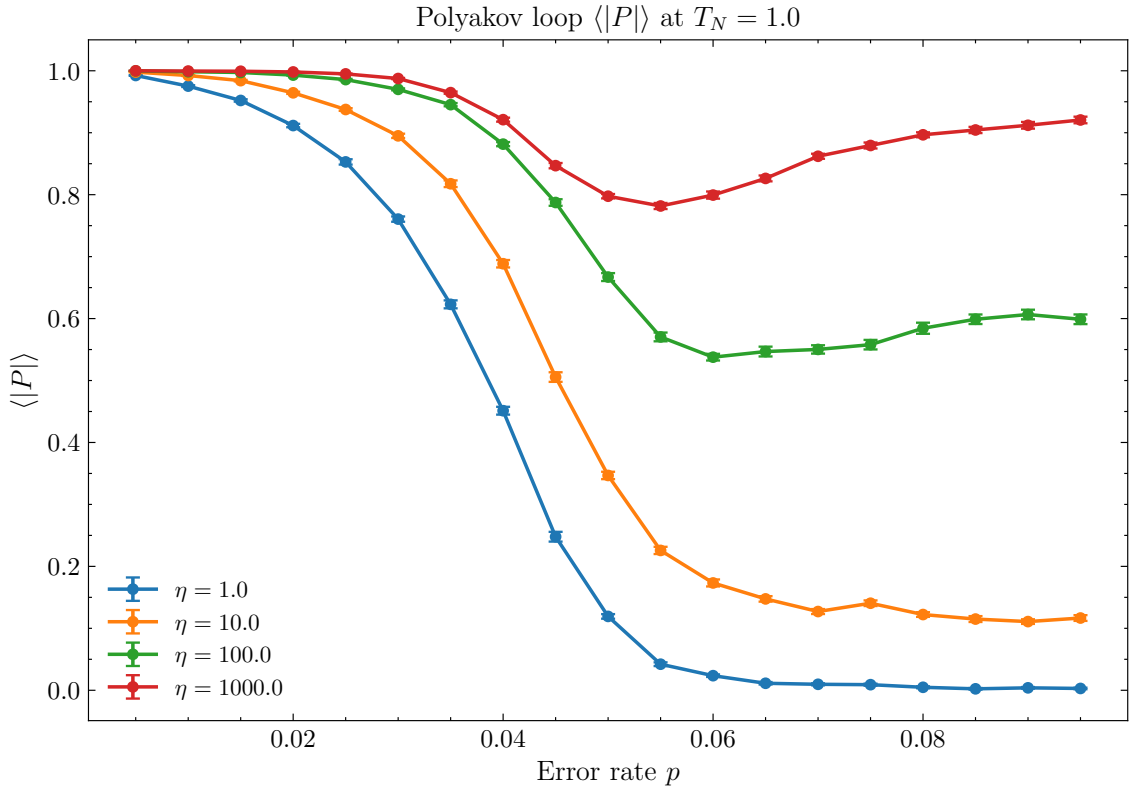


**Figure 4.1:** Average Polyakov loop magnitude $\langle |P| \rangle$ as a function of error rate $p$ for different bias strengths $\eta$. For $\eta = 1$ we observe a behavior in line with an order-disorder crossover. As bias increases, the system no longer reaches the disordered state, but instead appears to saturate at increasingly large values. Error bars represent one standard deviation of the bootstrapped disorder realizations.

The primary observable is the average Polyakov line $\langle |P| \rangle$, shown in Figure 4.1 for several values of the bias strength $\eta$. For weak bias ($\eta = 1$), the Polyakov line decreases smoothly from values close to +1 at low error rates toward values near zero at higher error rates. This behavior is consistent with a smooth crossover as a function of the error rate $p$ and hints at a threshold of $\sim 4$
For stronger bias values ($\eta = 10, 100, 1000$), the behavior changes qualitatively. The Polyakov line expectation initially remains near 1 before eventually starting to de-

crease. For $\eta = 10$, the expectation value never reaches zero, instead flattening out between 0.1 and 0.2. For yet larger biases, the initial decrease is followed by an increase in the expectation value before showing signs of flattening at increasingly large values. No crossover to a fully disordered regime is observed within the parameter range for strong bias.
As bias increases, the boundary effects can propagate deeper into the bulk through the strong $J_q$ and $J_z$ couplings. As such, it is expected that they will smudge crossovers more strongly as bias increases. It is possible that the boundary effects, with sufficiently large bias, can overcome the quenched disorder, leading to a saturation in $\langle |P| \rangle$ for finite systems. Characterizing this interplay and whether that causes the observed behavior would, however, require finite-size scaling analysis and thus cannot be explored further within this work.

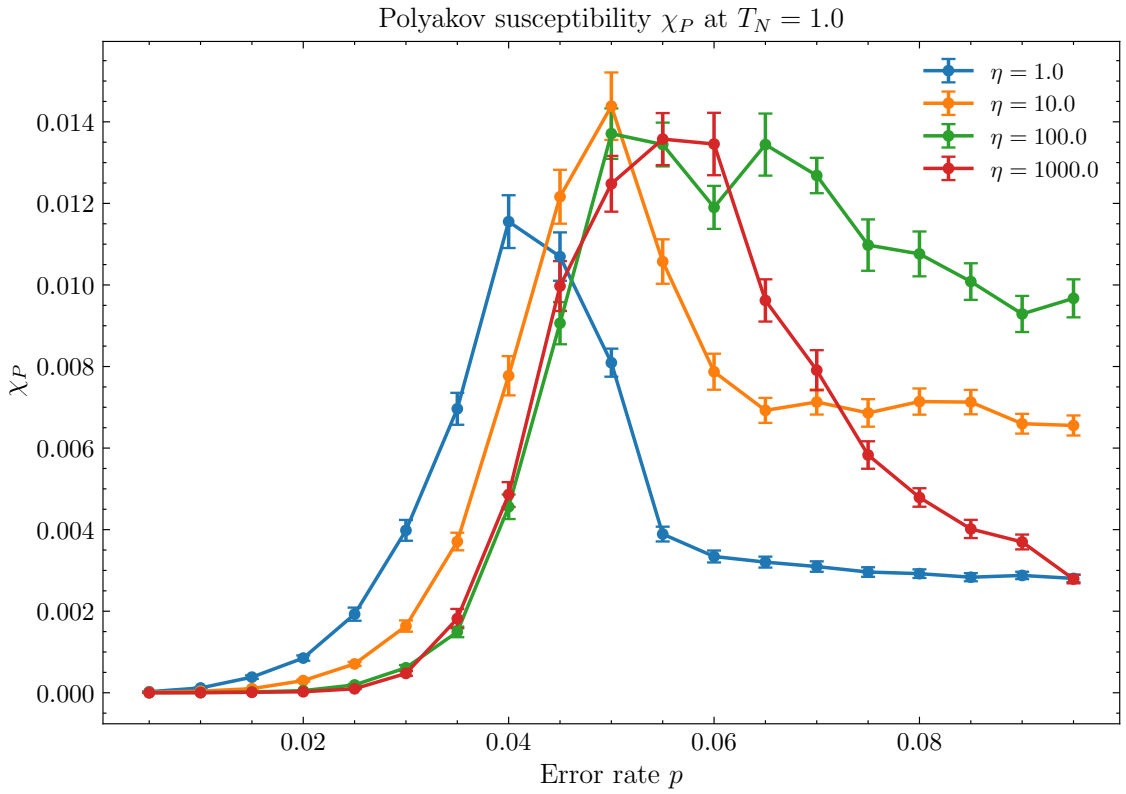


**Figure 4.2:** Thermal susceptibility $\chi$ of the Polyakov line averaged over quenched disorder realizations. Clear peaks are expected for an order-disorder crossover, which is observed for $\eta \in 1, 10, 1000$. A non-monotonic fluctuation pattern is seen as $\eta$ varies, where increased bias initially results in larger fluctuations, followed by a decrease. In conjunction with Figure 4.1, only the $\eta = 1.0$ peak appears to coincide with a crossover.

## 4.2 Susceptibility and Binder cumulant

To further characterize fluctuations in the Polyakov loop, the susceptibility $\chi$ and Binder cumulant $U_4$ were computed, shown in Figures 4.2 and 4.3.

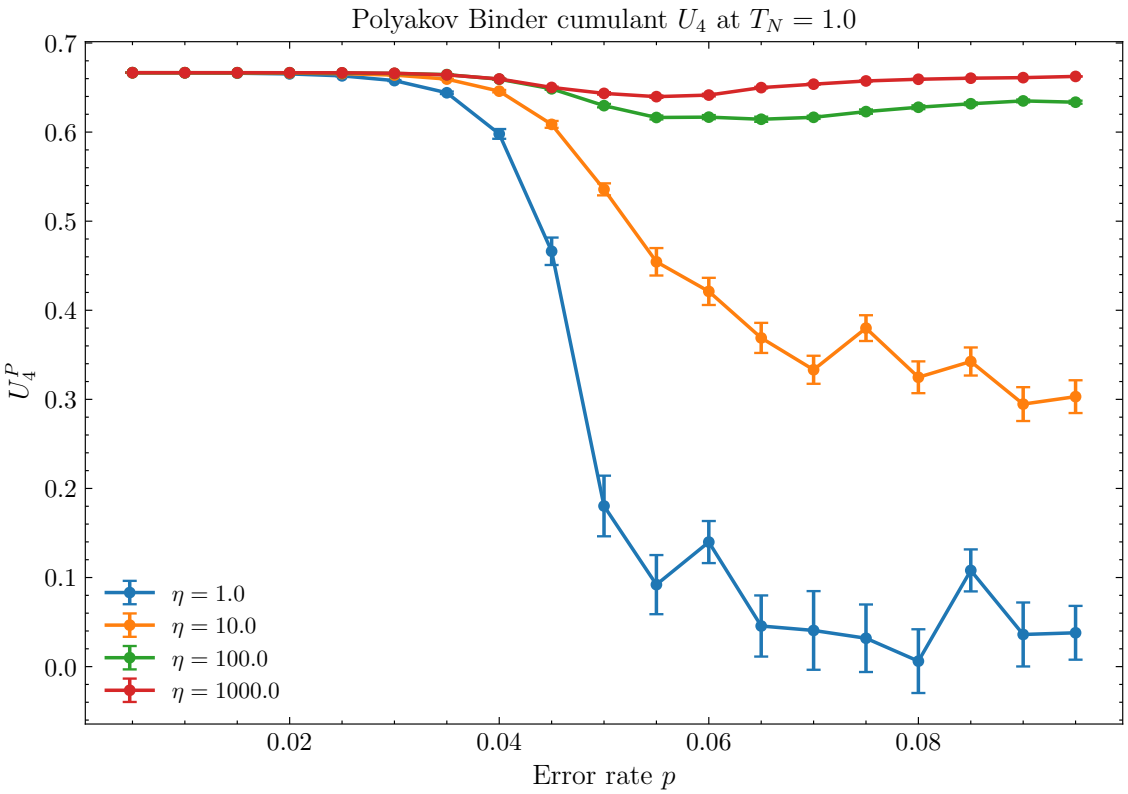


**Figure 4.3:** Disorder-averaged Binder cumulant $U_4$ of the Polyakov line. In agreement with Figures 4.1 and 4.2, only the $\eta = 1$ case shows crossover behavior. For $\eta = 10$ the distribution of $P$ widens as seen by the drop in $U_4$. This agrees well with the Polyakov data, which indicates a partial crossover into a semi-disordered configuration distribution without ever reaching the fully disordered phase. For large bias, fluctuations remain suppressed across the scanned $p$-range. This disagrees with the observations in $\chi$, which show larger fluctuations for large $p$ at $\eta = 100$ than for weaker bias.

The susceptibility $\chi$ exhibits a clear single peak for three of the four considered bias strengths. For $\eta = 1$ and $\eta = 10$, $\chi$ flattens beyond the peak, whereas for $\eta = 1000$ it continues to decrease across the sampled range. The trend appears to be an initial increase in plateau value with $\eta$, which is then reversed for sufficiently large bias. While discerning a real trend would require more values, particularly in the $\eta = [10, 1000]$ interval, this indicates that $\chi$ follows a non-monotonic trend in $\eta$. In general, the expected behavior would be that increased bias would suppress fluctuations as the boundary field can penetrate deeper into the bulk, corresponding to a decrease in entropy of error chains that give rise to a shared syndrome.

The Binder cumulant $U_4$ also shows behavior consistent with a well-defined ordered regime for low error rates. For $\eta = 1$, $U_4$ approaches values consistent with a transition from ordered to disordered behavior as $p$ increases. For higher bias values, $U_4$ remains closer to its ordered limit over the full range of $p$. However, we can note that there is a small dip in $U_4$ at $p$-values corresponding to the dip in $\langle |P| \rangle$. This indicates that distribution mass indeed begins to shift towards zero, before shifting

back. A potential explanation is that the system indeed begins to move towards a disordered state, but increased bias permits boundary effects to overcome the disordering and locks the system in the mostly ordered state.

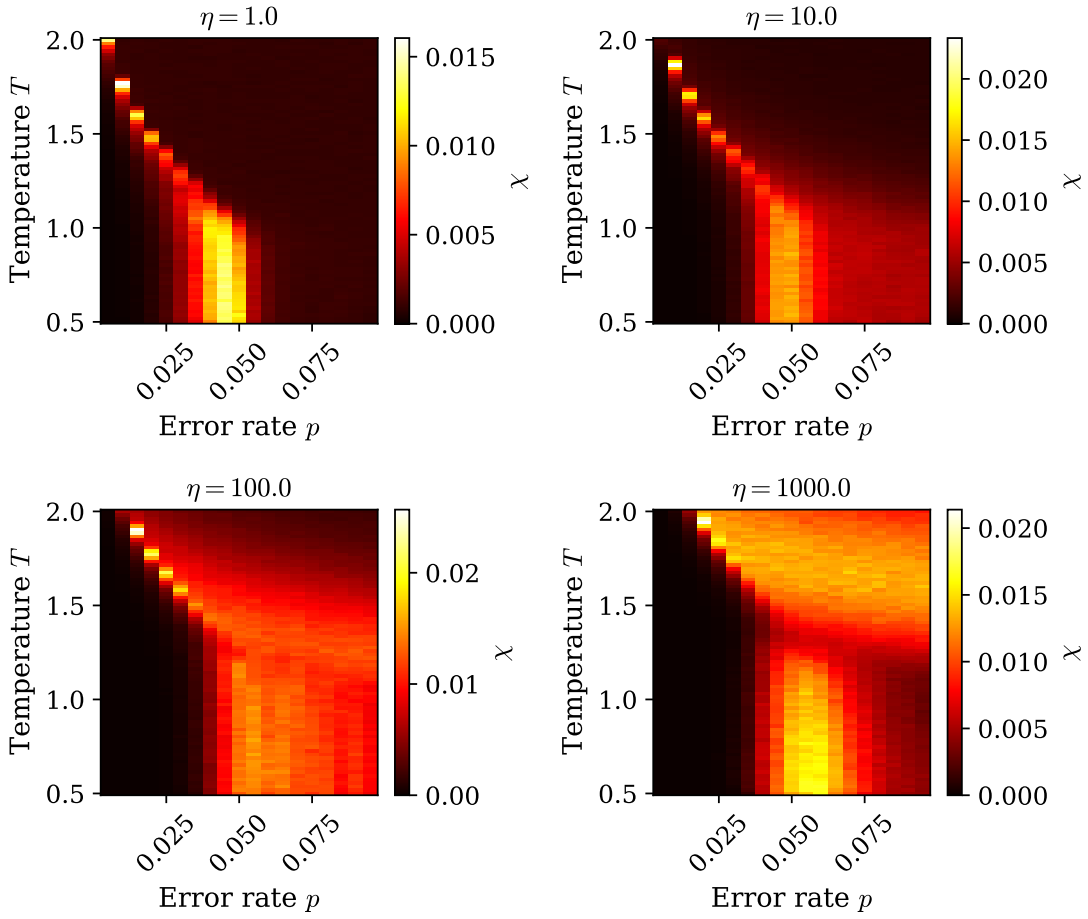


**Figure 4.4:** Heatmap plots of the temperature and disorder dependence of $\chi$ measured across PT replicas. Each replica is averaged over 50 disorder realizations. For low disorder we observe localized regions of high variance. Increasing the bias causes these regions to spread in the $p$-$T$ plane, and for strong bias we observe separate, spatially extended regions of large fluctuations. When the noise is strongly biased, the ridge of largest $\chi$ appears to never intersect the Nishimori temperature.

## 4.3 Temperature space heatmaps

In this section the data collected across parallel tempering replicas is presented. Figure 4.4 shows the mean $\chi$ per temperature $T$ and error rate $p$. For $\eta = 1$ we note a sharp ridge of spiked $\chi$ in the $p$-$T$ plane, along the critical temperature, which merges with a vertical region of increased susceptibility around a fixed $p$. As $\eta$ is increased, the ridge begins to spread in the $p$-$T$ plane, and for $\eta = 1000$ there are two separate regions of high susceptibility.

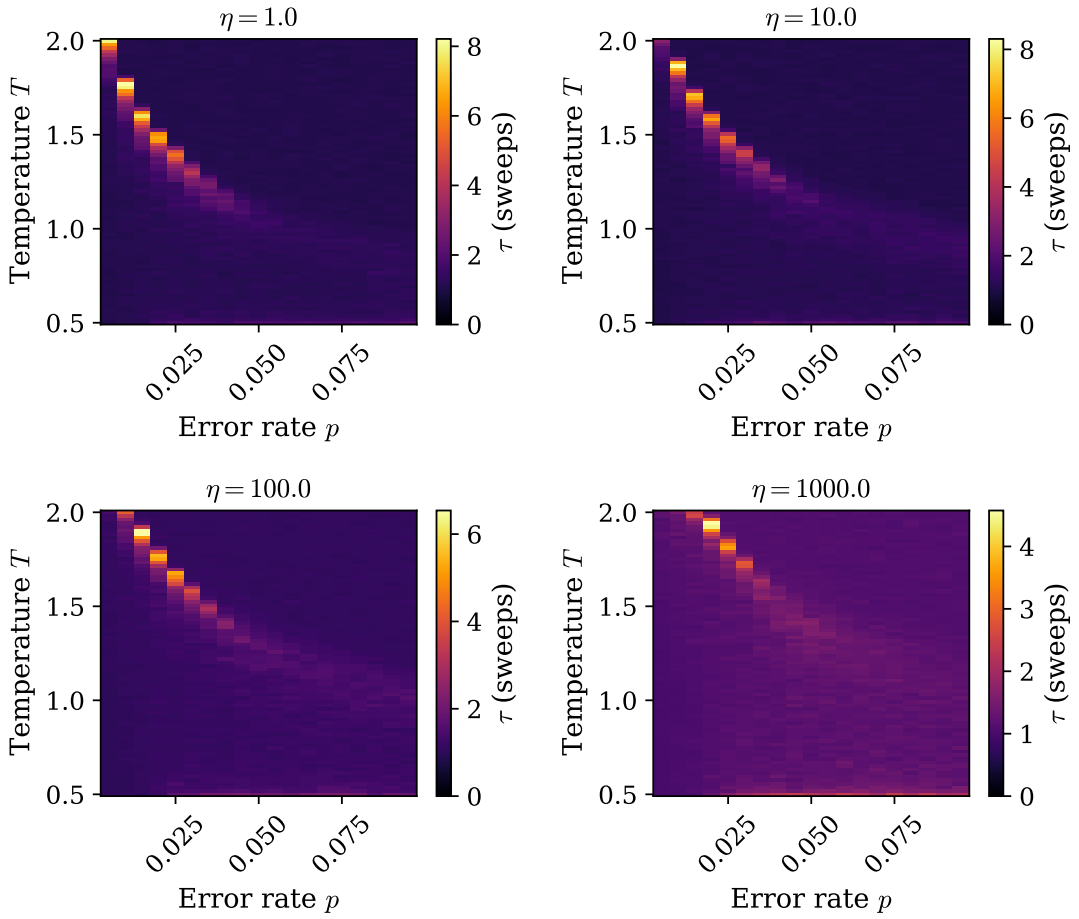


**Figure 4.5:** Heatmaps of the mean integrated autocorrelation time $\tau$ averaged over disorder. There are distinct ridges of long autocorrelation times, indicating critical slowing down in the system. Comparison to Figure 4.4 shows that these ridges correspond to lines of large fluctuations. The ridges terminate at temperatures larger than $T_N$.

The sharp ridge for low bias is in line with a thermodynamic phase transition, and smearing is expected when the finite system causes smooth crossover. In fact, the ridge is surprisingly localized for small $\eta$. The separation into two regions is however unexpected. The upper region, which shows signs of containing the previously narrow ridge, does not include the Nishimori temperature. If there were indeed a decoding related crossover, we would expect the critical temperature to intersect the Nishimori temperature at a point corresponding to the threshold.

In these diagrams we can also note an intermediate behavior for $\eta = 100$, in line with previous plots. More specifically, the region of high $\chi$ around a fixed $p$ does not appear. The $p$ would typically be associated with the threshold. However, the expectation would then be to observe a larger shift in its value with $\eta$. A potential explanation could be that the critical temperature tracks a thermodynamic crossover unrelated to decoding, and that the vertical region is an artifact of the finite geometry.

To complement the $\chi$ heatmaps, we also look at the corresponding plots for the autocorrelation time $\tau$, Figure 4.5. We note that these show clean ridges for all $\eta$ when $p$ is small and $T$ large. These ridges coincide with those in Figure 4.4, and vanish at a point which corresponds well with the endpoints of $\chi$-ridges. While the increased $\tau$ may serve as an indicator of critical slowing down associated with a finite size crossover, we again encounter the issue of the potential crossover not persisting to the Nishimori temperature where decoding is possible.

# 5

# Conclusion

We have studied the finite-size XZZX surface code with distance $d = 19$. The model incorporated phenomenological measurement errors repeated for a finite number of time steps. Noise was biased, given by Pauli error rates $p_x = p_y = p_z/(2\eta)$ and measurement errors $q = p_x + p_z$, such that increased bias corresponds to an increasingly dominant $Z$-error channel. The code was mapped to the RCPGM, a gauge model previously applied to the CSS. The statistical-mechanics approaches to characterizing surface codes have the advantage that sampling is performed directly on the error distribution, therefore providing accurate thresholds in the thermodynamic limit.

The mapping to the RCPGM was done by considering the mixed stabilizers of the XZZX code. Mixed-type stabilizers cause measurement errors to propagate into both the $X$- and $Z$-error sectors. We aimed to determine thresholds for the given code distance by tracking the deconfinement crossover associated with decoding the physical code. To find the crossover, Polyakov strings served as an order parameter. The model used fixed boundaries, in contrast to most existing literature on gauge mappings of surface codes. This introduces boundary effects in the Hamiltonian and prevents sharp phase transitions.

Results show signs of a crossover for weak noise bias $\eta$. As bias is increased to moderate and large values, the crossover vanishes. There are many interacting factors which may contribute to the observed behavior. Finite-size effects cause the system to preferentially align with the boundaries. This effect may then be increased as bias grows, which leads to a strong $Z$-sector coupling that can mediate boundary fields deeper into the lattice. For sufficiently strong couplings, the boundary fields may be able to overcome the constraints imposed by quenched disorder. Characterizing the effects of boundary conditions requires analysis of the finite-size scaling. Computational limitations prevented this work from incorporating FSS, but it is a natural next step.

Another limitation in the model is the implementation of the modified Polyakov loop as an order parameter on a finite lattice. The Polyakov loop is typically defined along a circular time axis, i.e. finite and periodic. By implementing it on a system with fixed boundaries, it no longer forms a closed loop, and its interpretation as an order parameter becomes difficult. The data indicates that the Polyakov string captures some thermodynamic crossover in the weakly biased regime. The connection to decoding-related deconfinement becomes increasingly unclear for heavily biased noise, in particular as the crossover-like region becomes detached from the Nishimori line. Without separating the finite-size effects, the precise behavior of the Polyakov string becomes difficult to ascertain.

To summarize, for the XZZX surface code under weakly biased noise, the data shows a crossover between phases, yielding a threshold around $\sim 4\%$. As bias is increased, thresholds can no longer be extracted. Hence, the method appears to work well under weak bias, but becomes less useful as the noise parameter is increased. Limitations in the data create difficulties in interpreting the results more exactly. The gauge mapping remains as an interesting method of analyzing the XZZX surface code, but the limited implementation in this work is unable to separate finite-size and bias effects. There are very natural continuations of this work, which as a first step include finite-size scaling. Another order parameter that measures deconfinement is the Wilson loop, which may be better suited to a finite lattice size and directional noise. Hence, a more comprehensive study incorporating Wilson loops and directional correlation lengths may yield more insightful results.

# Bibliography


[1] P. W. Shor, "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer," *SIAM Journal on Computing*, vol. 26, no. 5, pp. 1484–1509, Oct. 1997, arXiv:quant-ph/9508027, ISSN: 0097-5397, 1095-7111. DOI: 10.1137/S0097539795293172. Accessed: May 12, 2026. [Online]. Available: http://arxiv.org/abs/quant-ph/9508027.

[2] E. Farhi, J. Goldstone, and S. Gutmann, *A Quantum Approximate Optimization Algorithm*, arXiv:1411.4028 [quant-ph], Nov. 2014. DOI: 10.48550/arXiv.1411.4028. Accessed: May 12, 2026. [Online]. Available: http://arxiv.org/abs/1411.4028.

[3] H. Liu, G. H. Low, D. S. Steiger, T. Häner, M. Reiher, and M. Troyer, "Prospects of quantum computing for molecular sciences," en, *Materials Theory*, vol. 6, no. 1, p. 11, Mar. 2022, ISSN: 2509-8012. DOI: 10.1186/s41313-021-00039-z. Accessed: May 12, 2026. [Online]. Available: https://doi.org/10.1186/s41313-021-00039-z.

[4] S. B. Bravyi and A. Y. Kitaev, *Quantum codes on a lattice with boundary*, arXiv:quant-ph/9811052, Nov. 1998. DOI: 10.48550/arXiv.quant-ph/9811052. Accessed: Feb. 17, 2026. [Online]. Available: http://arxiv.org/abs/quant-ph/9811052.

[5] A. Y. Kitaev, "Fault-tolerant quantum computation by anyons," *Annals of Physics*, vol. 303, no. 1, pp. 2–30, Jan. 2003, arXiv:quant-ph/9707021, ISSN: 00034916. DOI: 10.1016/S0003-4916(02)00018-0. Accessed: Feb. 17, 2026. [Online]. Available: http://arxiv.org/abs/quant-ph/9707021.

[6] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, "Topological quantum memory," *Journal of Mathematical Physics*, vol. 43, no. 9, pp. 4452–4505, Sep. 2002, arXiv:quant-ph/0110143, ISSN: 0022-2488, 1089-7658. DOI: 10.1063/1.1499754. Accessed: Feb. 17, 2026. [Online]. Available: http://arxiv.org/abs/quant-ph/0110143.

[7] C. T. Chubb and S. T. Flammia, "Statistical mechanical models for quantum codes with correlated noise," *Annales de l'Institut Henri Poincaré D, Combinatorics, Physics and their Interactions*, vol. 8, no. 2, pp. 269–321, May 2021, arXiv:1809.10704 [quant-ph], ISSN: 2308-5827, 2308-5835. DOI: 10.4171/AIHPD/105. Accessed: Feb. 17, 2026. [Online]. Available: http://arxiv.org/abs/1809.10704.

[8] J. P. B. Ataides, D. K. Tuckett, S. D. Bartlett, S. T. Flammia, and B. J. Brown, "The XZZX Surface Code," *Nature Communications*, vol. 12, no. 1, p. 2172, Apr. 2021, arXiv:2009.07851 [quant-ph], ISSN: 2041-1723. DOI: 10.

1038/s41467-021-22274-1. Accessed: Feb. 17, 2026. [Online]. Available: http://arxiv.org/abs/2009.07851.

[9] Y. Xiao, B. Srivastava, and M. Granath, "Exact results on finite size corrections for surface codes tailored to biased noise," *Quantum*, vol. 8, p. 1468, Sep. 2024, arXiv:2401.04008 [quant-ph], ISSN: 2521-327X. DOI: 10.22331/q-2024-09-11-1468. Accessed: Feb. 17, 2026. [Online]. Available: http://arxiv.org/abs/2401.04008.

[10] M. Rispler, D. Vodola, M. Müller, and S. Kim, *The random coupled-plaquette gauge model and the surface code under circuit-level noise*, arXiv:2412.14004 [quant-ph], Dec. 2024. DOI: 10.48550/arXiv.2412.14004. Accessed: Feb. 17, 2026. [Online]. Available: http://arxiv.org/abs/2412.14004.

[11] W. K. Wootters and W. H. Zurek, "A single quantum cannot be cloned," en, *Nature*, vol. 299, no. 5886, pp. 802–803, Oct. 1982, ISSN: 1476-4687. DOI: 10.1038/299802a0. Accessed: May 19, 2026. [Online]. Available: https://www.nature.com/articles/299802a0.

[12] P. W. Shor, "Scheme for reducing decoherence in quantum computer memory," *Physical Review A*, vol. 52, no. 4, R2493–R2496, Oct. 1995. DOI: 10.1103/PhysRevA.52.R2493. Accessed: Jun. 17, 2026. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.52.R2493.

[13] A. R. Calderbank and P. W. Shor, "Good Quantum Error-Correcting Codes Exist," *Physical Review A*, vol. 54, no. 2, pp. 1098–1105, Aug. 1996, arXiv:quant-ph/9512032, ISSN: 1050-2947, 1094-1622. DOI: 10.1103/PhysRevA.54.1098. Accessed: Jun. 7, 2026. [Online]. Available: http://arxiv.org/abs/quant-ph/9512032.

[14] A. Steane, "Multiple Particle Interference and Quantum Error Correction," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 452, no. 1954, pp. 2551–2577, Nov. 1996, arXiv:quant-ph/9601029, ISSN: 1364-5021, 1471-2946. DOI: 10.1098/rspa.1996.0136. Accessed: Jun. 7, 2026. [Online]. Available: http://arxiv.org/abs/quant-ph/9601029.

[15] S. Bravyi, M. Suchara, and A. Vargo, *Efficient Algorithms for Maximum Likelihood Decoding in the Surface Code*, en, May 2014. DOI: 10.1103/PhysRevA.90.032326. Accessed: Jun. 1, 2026. [Online]. Available: https://arxiv.org/abs/1405.4883v1.

[16] C. Gattringer and C. B. Lang, "QCD on the lattice — a first look," en, in *Quantum Chromodynamics on the Lattice: An Introductory Presentation*, C. Gattringer and C. B. Lang, Eds., Berlin, Heidelberg: Springer, 2010, pp. 25–41, ISBN: 978-3-642-01850-3. DOI: 10.1007/978-3-642-01850-3_2. Accessed: Jun. 17, 2026. [Online]. Available: https://doi.org/10.1007/978-3-642-01850-3_2.

**DEPARTMENT OF SOME SUBJECT OR TECHNOLOGY**
**CHALMERS UNIVERSITY OF TECHNOLOGY**
Gothenburg, Sweden
www.chalmers.se


CHALMERS
UNIVERSITY OF TECHNOLOGY