Quantum Error Correction by means of Thermodynamics

Vladimir Korepin*

Yang Institute of Theoretical Physics,

State University of New York, Stony Brook, NY 11794-3840.

John Terilla†

Department of Mathematics, State University of New York, Stony Brook, NY 11794-3651.

(Dated: March 4, 2002)

Abstract

We further the connection between information theory and statistical mechanics by developing a thermodynamic approach to quantum error correction. Our method of protecting a code space of quantum bits from interference involves realizing it as a thermo-equilibrium subspace in the Hilbert space of a quantum statistical model. We apply the ideas of Claude Shannon to quantum error correction.

PACS numbers: 03.67, 05.30, 63.10

*Electronic address: `korepin@insti.physics.sunysb.edu`

†Electronic address: `jterilla@math.sunysb.edu`

I. INTRODUCTION

At the very moment that information science was discovered by Claude Shannon [1], it was closely related to statistical mechanics and thermodynamics. Now, more than fifty years later and well into the information age, Shannon's theorems remain beacons in the field, and statistical mechanics continues to provide insights into information theory.

Quantum information science, pioneered by P. Benioff, C. Bennett, R. Feynman, and others, is a natural extension of the classical theory: information is physical, the physical world is quantum mechanical, and there is a quantum theory of information. In the last decade, quantum information theory has developed rapidly. Important features which distinguish the quantum theory from the classical one have been unearthed and fascinating applications have been explored. Leaders in the field, D. P. DiVincenzo, D. Gottesman, A. Kitaev, J. Preskill, B. Schumacher, P. Shor, U.V. Vazirani and W.K. Wootters, to name a few, have brought the quantum theory of information squarely to the front of information science. There are several books on quantum information and computation, see for example [5] ,[4] and [6]. Our work is related to quantum error correction. By developing a thermodynamic approach to quantum error correction, we add a new link in the chain of connections between information theory and statistical mechanics. On the one hand, the state of a quantum computer can be viewed as a vector in a linear space C, called a code space. The idea of protecting quantum information from errors is to carefully embed the code space C as a subspace of a larger Hilbert space W, referred to as the encoding space, or the quantum medium. The embedding $C \hookrightarrow W$ should be done in such a way that if the vector describing the state of a computer is subjected to an error, it can be recovered. On the other hand, a quantum statistical mechanics model is defined by a Hamiltonian operating in the Hilbert space of the model, which we also call W. The thermodynamic equilibrium of the model manifests itself as a particular subspace of W, called the subspace of thermodynamic equilibrium (thermo-equilibrium for short). Our proposal identifies the code space C of quantum information as the subspace of thermo-equilibrium in a quantum statistical mechanics model. The embedding of a code space has to satisfy certain conditions in order to have good error correcting properties. Our technique to identify C as a thermo-equilibrium subspace succeeds because the correlation functions of the model at positive temperature satisfy a remarkable property which, when translated to information theoretic language, guarantee that all finite local errors are correctable.

A. The quantum error correcting condition

Consider a code space C embedded in an encoding space W (quantum medium) equipped with a scalar product $\langle\,|\,\rangle$. Let $\{E_a : W \to W\}$ be a collection of linear operators which represent errors. One imagines the diagram

$$C \hookrightarrow W \xrightarrow{\{E_a\}} W \tag{1}$$

where the arrow on the right represents either a passage of time during which the medium deteriorates, or possibly the transmission of the medium down a noisy channel. The goal of quantum error correction is to be able to recover the subspace C from the rightmost W, after errors have occurred. Error correction is also useful for correcting mistakes which arise from performing imperfect calculations within a quantum computer.

Quantum Error Correction Condition. *The necessary and sufficient condition for these errors to be correctable is that $\langle\psi|E_a^\dagger E_b|\psi\rangle$ be the same for all unit vectors $\psi \in C$.*

The proof of this statement can be found in section 3 of [2], in chapter 7 (pages 9 and 86) of [4], and section 6.4 of [3]. Also see theorem 10.1 in [5], where slightly different but equivalent conditions are stated and proved.

Since 1996, several good quantum codes have been developed. The most famous are the CSS codes [7, 8] and the stabilizer codes [9]. See also the textbook [5] and the references within. Recently, there has been interesting new work in topological error correcting codes using modular functors [10]. Compared to existing codes, our approach to quantum error correction is rather more in the tradition of the Shannon theorems because of its statistical mechanical nature.

B. Correlation functions in the thermodynamic limit

Let us now consider a system of quantum spins. (A qubit of quantum information will be identified with a spin $\frac{1}{2}$.) The state of the system is a unit vector in a Hilbert space, which we will again call W. The dynamics of the system is determined by the Hamiltonian. It is a Hermitian operator $\mathcal{H} : W \to W$. Statistical mechanics study the behavior of the system as the number of spins becomes very large [thermodynamics limit]. As the number of spins grows to infinity, a crucial subspace of W, called the space of thermodynamic equilibrium (see definition A.1) reveals itself. We suggest to realize this space as a code space C of quantum information.

Let us briefly discuss why a thermo-equilibrium space makes a good code space. Quantum spin chain models include a real parameter T [temperature]. Usually, when $T = 0$, there is a unique vector in W, called the ground state, corresponding to the smallest eigenvalue of H. As a one dimensional space, the span of the ground state is too small to be used as a good quantum code. However, for $T > 0$, the space of thermodynamic equilibrium has an exponentially large dimension (see the line preceding equation (17)). Our attitude is to interpret the entire space of thermo-equilibrium as a kind of 'high dimensional ground state'. It turns out that any unit vector chosen from the thermo-equilibrium space will serve, equally well, to represent the physical properties of the entire equilibrium. Here, physical properties are things like energy, momentum, scattering matrix, and, most importantly for our purposes, correlation functions. A correlation function is a physical quantity associated to a local operator $\mathcal{O} : W \to W$ (see equation (19) for a definition). We introduce, the correlation function associated to $\mathcal{O}$ by $\langle \mathcal{O} \rangle_T$. In this paper we present the following fundamental observation [11–13, 18]:

Thermodynamic Correlation Function Condition. *In the thermodynamic limit, correlation functions* $\langle \mathcal{O} \rangle_T$ *satisfy the following equation* $\langle \mathcal{O} \rangle_T = \langle \psi | \mathcal{O} | \psi \rangle$ *for any unit vector* ψ *in the subspace of thermo-equilibrium.*

We emphasize that by choosing C to be the space of thermodynamic equilibrium and the operator $\mathcal{O} = E_a^\dagger E_b$, the thermodynamic correlation function condition says that $\langle \psi | E_a^\dagger E_b | \psi \rangle$

has the same value for all unit vectors $\psi \in C$. Thus, for the subspace of thermo-equilibrium C, the quantum error correcting condition is achieved.

The subject of thermodynamics and quantum statistical mechanics is treated in several textbooks [11, 14]. There are various models determined by different Hamiltonians describing different interactions, and many of the spin chain models have relevance to quantum information theory. We begin our work with XX0 spin chains, which have already been considered in the context of quantum computing, though for purposes different than ours. W. Wootters suggested that quantum spin chains be used for the study of entangled qubits [15], and this investigation continues (see for example [16]).

II. AN OVERVIEW OF QUANTUM ERROR CORRECTION

In this section, we briefly detail the quantum error correcting condition in the context of error recovery. That is, extend the diagram in expression (1) as

$$C \hookrightarrow W \xrightarrow{\{E_a\}} W \xrightarrow{\{R_b\}} C \tag{2}$$

in such a way the composition acts as the identity $\mathrm{Id} : C \to C$ (or a constant multiple of it).

Our presentation of this background follows mainly [2], but we also refer the reader to other good summaries in [4, 9] and especially chapter ten of [5]. Let us establish the setting. Let W be a finite dimensional Hilbert space and let C be a subspace of W. Consider two collections of linear transformations of W, $\mathcal{E} = \{E_a : W \to W\}_{a=1}^{r}$ and $\mathcal{R} = \{R_b : W \to W \text{ with } \sum_{b=1}^{s} R_b^\dagger R_b = \mathrm{Id}\}_{b=1}^{s}$. The code space is C, the quantum medium (or encoding space) is W and the collections $\mathcal{E}$ and $\mathcal{R}$ represent error and recovery operations.

Let us introduce a formal mathematical function Err, depending on C, $\mathcal{R}$, and $\mathcal{E}$, which measures the extent to which $\mathcal{R}$ can correct the errors $\mathcal{E}$ on the code space C:

$$\mathrm{Err}(C, \mathcal{R}\mathcal{E}) = \max_{\psi \in C, |\psi|=1} \sum_{i=1}^{r} \sum_{j=1}^{s} \left| \left(R_j E_i - \langle \psi | R_j E_i | \psi \rangle \right) |\psi\rangle \right|^2 . \tag{3}$$

The pair $(C, \mathcal{R})$ is defined to be an $\mathcal{E}$-correcting code if and only if $\mathrm{Err}(C, \mathcal{R}\mathcal{E}) = 0$. In this case, we say that the errors $\mathcal{E}$ are correctable. This is the precise meaning of the word "correctable" in the quantum error correcting condition. Notice that $\mathrm{Err}(C,\mathcal{RE})=0$ is equivalent to the condition that for each $a=1,\dots,r$ and $b=1,\dots,s$, we have

$$R_bE_a|\psi\rangle = \langle\psi|R_bE_a|\psi\rangle|\psi\rangle; \tag{4}$$

that is, R_bE_a acts as a multiple of the identity on C.

For any two vector spaces V and V', the expression $\mathrm{Hom}(V,V')$ denotes the vector space of all linear transformations from V to V' . If bases for V and V' are chosen, $\mathrm{Hom}(V,V')$ is identified with a set of matrices. One can check that the condition $\mathrm{Err}(C,\mathcal{RE})=0$ means that the superoperator $S_{\mathcal{R}}:\mathrm{Hom}(W,W)\to\mathrm{Hom}(W,W)$ defined by $S_{\mathcal{R}}(P)=\sum_{b=1}^{s}R_bPR_b^{\dagger}$ inverts the superoperator $S_{\mathcal{E}}:\mathrm{Hom}(W,W)\to\mathrm{Hom}(W,W)$, defined by $S_{\mathcal{E}}(P)=\sum_{a=1}^{r}E_aPE_a^{\dagger}$, up to a constant multiple when restricted to the subspace $\mathrm{Hom}(C,C)\subseteq\mathrm{Hom}(W,W)$. This fact can be regarded as motivation for the definition of $\mathrm{Err}(C,\mathcal{RE})$. Sometimes $S_{\mathcal{E}}$ and $S_{\mathcal{R}}$ are called quantum operations. Matrix P is a density matrix. Let us comment for mathematicians. A density matrix arises as follows: given a codeword $c\in C\subset W$, one has the linear functional c^* in the dual space C^* defined by $c^*:=\langle c|\,\rangle$. A matrix P can then be identified with $c^*\otimes c$ via the isomorphism $W^*\otimes W\simeq\mathrm{Hom}(W,W)$.

The following lemma indicates that capability of $(C,\mathcal{R})$ to correct errors applies to the entire linear space spanned by $\mathcal{E}$, not just the collection $\mathcal{E}$ itself. It says that if the code corrects the mistakes $\{E_a\}$, then it will also correct any linear combination of the $\{E_a\}$.

Lemma II.1. *Suppose that $\mathcal{E}$ and $\mathcal{E}'$ are two collections of operators on a Hilbert space W that span the same subspace of* $\mathrm{Hom}(W,W)$*. Then $(C,\mathcal{R})$ is an $\mathcal{E}$-correcting code if and only if $(C,\mathcal{R})$ is an $\mathcal{E}'$-correcting code.*

We complete this section by presenting a formulation of error correctability in terms of a recovery. It is the basis for the error correcting conditions in section (I A). Given only the code space C and the set of errors $\mathcal{E}$, it is possible to determine whether or not one can construct a collection of recovery operators $\mathcal{R}$ making $(C,\mathcal{R})$ into an $\mathcal{E}$-correcting code. This is the content of the next theorem.

Theorem II.2. *Let W be a Hilbert space, let C be a subspace of W, and let $\mathcal{E} = \{E_a\}$ be a collection of linear operators on W. Then there exists a collection of operators $\mathcal{R}$ making $(C, \mathcal{R})$ into an $\mathcal{E}$-correcting code if and only if*

$$\langle\psi|E_a^\dagger E_b|\psi\rangle = \langle\phi|E_a^\dagger E_b|\phi\rangle \tag{5}$$

for all unit vectors $\psi, \phi \in C$ and for all a, b.

III. QUBITS

Let us remind the reader that an ordinary classical bit b takes only two values $\{0,1\}$. So the classical bit is an element $b \in \{0,1\}$. A quantum bit, or a qubit, is an element of a two dimensional complex vector space $\mathbb{C}^2$. Usually, one denotes the basis vectors by $|0\rangle$ and $|1\rangle$. Information contained in k qubits is carried by element of the k-fold tensor product $\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$. Ordinarily, a quantum code embeds a codeword consisting of k qubits into a quantum medium consisting of n qubits. The code space C has dimension 2^k and the encoding space W has dimension 2^n, with a condition $n > k$. An element of W is thought of as a message n qubits long which carries k qubits of information. Now imagine that the quantum medium is exposed to errors, which take the form of linear operators on W. Often, one considers errors which affect t or fewer of the n qubits in the medium.

Let us clarify the situation. Let $n \in \mathbb{N}$ and define $W = \otimes_{j=1}^n W_j$, where each W_j is a two dimensional vector space with basis $\{|0\rangle_j, |1\rangle_j\}$. Each $W_j = \mathbb{C}|0\rangle_j \oplus \mathbb{C}|1\rangle_j$ is identified with a space of qubits in the evident way. Let σ^x, σ^y, and σ^z be the Pauli matrices. On $\mathbb{C}^2 = \mathbb{C}|0\rangle \oplus \mathbb{C}|1\rangle$ we can write them as

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{6}$$

It is common in quantum information theory to use the notation X, Y, and Z instead of σ_x, σ_y, and σ_z. For any $j = 1, \ldots, n$ and superscripts $\alpha = x, y, z$ define a local operator

$\sigma_j^\alpha : W \to W$ by

$$\sigma_j^\alpha|_{W_m} = \begin{cases} \mathrm{Id} & \text{if } j \neq m, \\ \sigma^\alpha & \text{if } j = m. \end{cases} \tag{7}$$

In other words, σ_j^α acts non trivially on the j-th qubit as the Pauli matrix σ^α and leaves all other qubits unchanged. The following convenient notation is suggested for a basis of W. Every basis vector $|b_1\rangle_1 \otimes |b_2\rangle_2 \otimes \cdots \otimes |b_n\rangle_n$ can be referred to by the shortened name $|b_1 b_2 \cdots b_n\rangle$. In this way, bit strings of length n become the names for the basis vectors of W and a general element of the quantum medium can be regarded as a formal linear combination of bit strings of length n. Usually, leading zeros are truncated so that, for instance, $|00001001011\rangle$ is renamed $|1001011\rangle$, and conventionally $|00\cdots 0\rangle$ becomes $|0\rangle$. One can refer to the matrix

$$\sigma^- := \frac{1}{2}\left(\sigma^x - i\sigma^y\right) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{8}$$

as a creation operator. By setting $\sigma_j^- = \frac{1}{2}(\sigma_j^x - i\sigma_j^y)$, one can obtain every basis vector in W by acting on $|0\rangle$ by products of the σ_j^-.

The errors we wish to consider are local errors; i.e., linear combinations of products of the operators σ_j^α. If E is a product of t such operators

$$E = \sigma_{m_1}^{\alpha_1} \sigma_{m_2}^{\alpha_2} \cdots \sigma_{m_t}^{\alpha_t} \tag{9}$$

with distinct m_j, we say that E is an error of weight t. Fix t and let $\mathcal{E}_t$ be the specific set of errors consisting of products of t or fewer Pauli matrices

$$\mathcal{E}_t = \{E \text{ acting on W, with the weight} \leq t\}. \tag{10}$$

An important class of codes are able to correct the errors $\mathcal{E}_t$. The significance of $\mathcal{E}_t$ correcting codes is that they are able to correct arbitrary errors on t or fewer bits. This is a consequence of lemma II.1 since the errors of weight $\leq t$ form a basis for the linear space of all errors affecting t or fewer qubits.

IV. AN OVERVIEW OF XX0 SPIN CHAINS

A. The partition and correlation functions

We study a model of interacting spins on a one dimensional lattice. Each spin is an element of $\mathbb{C}^2$. Let $n \in \mathbb{N}$ and define W as before

$$W = \otimes_{j=1}^{n} W_j, \quad W_j = \mathbb{C}|0\rangle_j \oplus \mathbb{C}|1\rangle_j . \tag{11}$$

We call the integer n the length of the lattice. It is the number of interacting spins. The space $W \simeq (\mathbb{C}^2)^{\otimes n}$ is the Hilbert space of the model. The Hamiltonian $\mathcal{H} : W \to W$ is defined by

$$\mathcal{H} = -\sum_{j=1}^{n} \left(\sigma_j^x \sigma_{j+1}^x + \sigma_j^y \sigma_{j+1}^y + h\sigma_j^z \right) \tag{12}$$

The real parameter h is called the magnetic field. The model was originally solved in 1961 by E. Lieb, T. Schultz, and D. Mattis [17].

The thermodynamic limit of the model is defined by the limit $n \to \infty$. We study the asymptotics of different physical observables in this limit. We compute observables for finite n and then evaluate asymptotics as $n \to \infty$. Often, the quantities of interest are proportional to n, so it is the proportionality factor which is of interest. The proportionality factor has a limit. In other words, the asymptotics of different physical values are studied as n increases without bound. Compare this to information science where n plays the role of the length of the message. The study of asymptotics as the length of the message grows is one of the most well developed areas of information science after Claude Shannon.

The central object of thermodynamics is the partition function $\mathcal{Z}$, which is defined at a temperature $T \geq 0$, by

$$\mathcal{Z} = \operatorname{tr}\left(2^{-\frac{\mathcal{H}}{T}} \right) . \tag{13}$$

In the thermodynamic limit $n \to \infty$, the partition function can be computed by the method of steepest descent (see definition A.1 and equations (A3), (A4), (A5) in the appendix). The outcome of this calculation shows that the partition function $\mathcal{Z}$ can be computed by taking the trace over a subspace $C \subseteq W$, called the subspace of thermodynamic equilibrium:

$$\mathcal{Z} = \mathrm{tr}_C \left(2^{-\frac{\mathcal{H}}{T}}\right) \to 2^{\frac{n}{2\pi}\int_{-\pi}^{\pi} dp \log\left(1+2^{\frac{-\epsilon(p)}{T}}\right)} \tag{14}$$

where ϵ is given by

$$\epsilon(p) = -4\cos(p) + 2h. \tag{15}$$

The precise meaning of the right hand side of equation (14) is that $\lim_{n\to\infty} \frac{1}{n} \log \mathcal{Z}$ exists and is given by

$$\lim_{n\to\infty} \frac{1}{n} \log \mathcal{Z} = \frac{1}{2\pi} \int_{-\pi}^{\pi} dp \log\left(1 + 2^{\frac{-\epsilon(p)}{T}}\right). \tag{16}$$

This is a bulk free energy. Throughout this paper log will mean log base 2.

The dimension of C is 2^{nS} where S is the entropy (see equation 34):

$$S = -\frac{1}{2\pi} \int_{-\pi}^{\pi} dp \left[\theta(p) \log \theta(p) + (1-\theta(p)) \log(1-\theta(p))\right], \tag{17}$$

where θ, called the Fermi weight, is defined by

$$\theta(p) = \left(1 + 2^{\frac{\epsilon(p)}{T}}\right)^{-1}. \tag{18}$$

Physicists call ϵ the energy of the spin wave and call p the momentum of the spin wave.

Let us define correlation functions. Let $\mathcal{O}$ be a linear combination of products of operators σ_j^{α}. Such an operator can be called local. The correlation function $\langle \mathcal{O} \rangle_T$ is defined by

$$\langle \mathcal{O} \rangle_T = \frac{\mathrm{tr}\left(2^{-\frac{\mathcal{H}}{T}} \mathcal{O}\right)}{\mathrm{tr}\left(2^{-\frac{\mathcal{H}}{T}}\right)}. \tag{19}$$

If $\mathcal{O}$ has weight t, in the sense of equation (9). $\langle \mathcal{O} \rangle_T$ can be related to t-point correlation function. In the thermodynamic limit correlation function also can be calculated by steepest descent. The result is the same as for the partition function. Only the space of thermo equilibrium C contributes to the trace ([11], page 25):

$$\langle \mathcal{O} \rangle_T = \frac{\mathrm{tr}_C\left(2^{-\frac{\mathcal{H}}{T}} \mathcal{O}\right)}{\mathcal{Z}}. \tag{20}$$

An even stronger statement is true. In the papers [11–13, 18] it was shown that each term here contribute equally.

$$\langle \mathcal{O} \rangle_T = \langle \psi | \mathcal{O} | \psi \rangle \text{ where } \psi \text{ is any unit vector in } C. \tag{21}$$

This is precisely the thermodynamic correlation function condition from section I B. This condition is an expression of our philosophy that the linear space of thermodynamic equilibrium behaves as a single state, and any unit vector in that space may be used to represent the equilibrium space. We believe it to be a rather general feature of quantum statistical mechanics. Equation (21) is valid, not just for XX0, but also for other integrable models as well. We conjecture that equation (21) is valid for a wide class of physically interesting models.

Let us comment in the end on the the special case $T = 0$. The correlation functions simplify dramatically. There is a unique vector $|G\rangle$, first identified in [17], called the ground state. It corresponds to the lowest eigenvalue of $\mathcal{H}$. In the case of zero temperature

$$\langle \mathcal{O} \rangle_{T=0} = \langle G | \mathcal{O} | G \rangle. \tag{22}$$

B. Eigenvectors of the Hamiltonian

Let us fix the size of the lattice to be n. The Hamiltonian

$$\mathcal{H} = -\sum_{j=1}^{n} \left(\sigma_j^x \sigma_{j+1}^x + \sigma_j^y \sigma_{j+1}^y + h \sigma_j^z \right)$$

can be written as $\mathcal{H} = \mathcal{H}_0 - 2hS^z$ where

$$\mathcal{H}_0 = -\sum_{j=1}^{n} \left(\sigma_j^x \sigma_{j+1}^x + \sigma_j^y \sigma_{j+1}^y \right) \text{ and } S^z = \frac{1}{2} \sum_{j=1}^{n} \sigma_j^z. \tag{23}$$

Note that $[\mathcal{H}_0, S^z] = 0$. The problem of finding the eigenvectors of $\mathcal{H}$ can be reduced to finding the eigenvectors of $\mathcal{H}_0$. Recall that $\sigma_j^- = \frac{1}{2}(\sigma_j^x - i\sigma_j^y)$.

The eigenvectors of $\mathcal{H}$ are determined by a positive integer $m \leq n$ and a collection of real numbers $\{p_i\}_{i=1}^m$, with each $-\pi < p_i < \pi$, called momenta. For short, the collection $\{p_i\}_{i=1}^m$

may be denoted simply by $\{p\}$. Define a vector $|\{p\}\rangle_m \in W$ by

$$|\{p\}\rangle_m = \frac{1}{\sqrt{m!}} \sum_{x_1,\dots,x_m} \chi_m(\{x\}|\{p\})\sigma^-_{x_m}\cdots\sigma^-_{x_1}|0\rangle. \tag{24}$$

The complex valued function χ is defined by

$$\chi_m(\{x\}|\{p\}) = \frac{1}{\sqrt{m!}} \left(\prod_{1\le a<b\le m} \operatorname{sign}(x_b - x_a)\right) \det(A), \tag{25}$$

where A is the $m \times m$ matrix with (j,k) entry $A_{jk} = \exp(ix_j p_k)$ and the sign function is given by

$$\operatorname{sign}(x) = \begin{cases} 1 & \text{if } x > 0, \\ -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0. \end{cases} \tag{26}$$

The function $\chi(\{x\}|\{p\})$ is called a wave function. It is symmetric in x and antisymmetric in p, so we assume that $x_1 \le x_2 \le \cdots \le x_m$ and that $p_1 < p_2 < \cdots < p_n$. We may drop the subscript m from $|\{p\}\rangle_m \in W$ if the size of $\{p\}$ does not need emphasis and just write $|\{p\}\rangle$.

The periodic boundary conditions of the model lead to the following equation for each p_j:

$$\exp(ip_j n) = (-1)^{m+1} \tag{27}$$

Note that (27) has n solutions and a collection $\{p\}$ amounts to a choice of m of these solutions. So, for each $m = 0, \dots n$, there are $\binom{n}{m}$ choices of momenta each naming one eigenvector of $\mathcal{H}$. In total, there are 2^n eigenvectors (24) as m varies from 0 to n. A straightforward calculation of their scalar products shows that these eigenvectors form an orthogonal basis of W:

$$\langle\{p\}|\{p\}\rangle = n^m \text{ and } \langle\{p\}|\{q\}\rangle = 0 \text{ if } \{p\} \neq \{q\}.$$

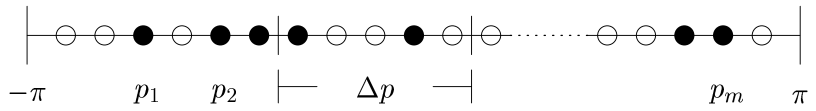
The eigenvalue $E(\{p\})$ of $|\{p\}\rangle_m$ is given by

$$E(\{p\}) = \sum_{j=1}^{m} \epsilon(p_j), \quad \epsilon(p) = -4\cos p + 2h. \tag{28}$$

C. Thermodynamics and the macroscopic perspective

One may think of each of the eigenvectors of $\mathcal{H}$ as being obtained from the ferromagnetic state $|0\rangle$ by adding m particles (flipping m spins) with momenta $\{p_j\}$ and energies $\{\epsilon(p_j)\}$. In this section we consider the situation as the number of particles occupying positions in the interval $[-\pi, \pi]$ tends to infinity (see [11]). Divide the interval $[-\pi, \pi]$ into n subintervals, each of length $\frac{2\pi}{n}$. Each subdivision point will be called a position and should be thought of as a possible location for a particle. They represent the n solutions of (27), which the momenta satisfy. When a wave function has been selected, and a choice $\{p\}$ of momenta has been made, one should think that each of the m positions appearing in $\{p\}$ is occupied by a particle and the remaining $n - m$ positions are empty (see figure (1)).

FIG. 1: A picture of typical positions and momenta



Once a collection $\{p\} = \{p_1 < p_2 < \cdots < p_n\}$ has been chosen, one may look at the numbers $\rho(p_j)$:

$$\rho(p_j) = \frac{1}{n(p_{j+1} - p_j)}. \tag{29}$$

Now we consider the situation where the number of positions becomes large, the number of particles becomes large, and the density of particles remains fixed. This limit

$$m \to \infty, \quad n \to \infty, \quad \text{and } d := \frac{m}{n} \text{ is constant} \tag{30}$$

is called the thermodynamic limit. The reader may compare it to the double-scaling limit in matrix models. For now, consider m and n as very large, but still finite. As the number of particles grows, it becomes difficult to keep track of the individual momenta p_j since there are so many of them. Instead, the growing collections $\{p\}$ give way to a finite function $\rho : [-\pi, \pi] \to \mathbb{R}$, which exists as a limit of expressions (29). It has the interpretation that for large n

$$\text{the number of particles in the interval} [p, p + \Delta p] \approx n\rho(p)\Delta p,$$

provided the scale for Δp is chosen properly. One should have

$$\frac{2\pi}{n} << \Delta p, \tag{31}$$

but Δp should still be small enough to be able to approximate ρ by a constant on the interval $[p, p + \Delta p]$.

One should think that ρ is part of a macroscopic description of the model, and that a collection $\{p\}$ is part of a microscopic description. Be aware that each ρ may be the limit of many different sequences of microscopic states $\{p\}$, the number of which can be called the degeneracy of the macroscopic state ρ. For a fixed ρ, we can compute the number of microscopic states corresponding to ρ. First, partition the interval $[-\pi, \pi]$ into subintervals of length Δp. It is important that the scale of Δp is such that ρ is well approximated by a constant on the interval $[p, \Delta p]$. The number of ways of selecting $n\rho(p)\Delta p$ positions from $\frac{n}{2\pi}\Delta p$ vacancies in which to place particles is equal to the binomial coefficient:

$$\binom{\frac{n}{2\pi}\Delta p}{n\rho(p)\Delta p} = \frac{\left(\frac{n}{2\pi}\Delta p\right)!}{(n\rho(p)\Delta p)!\ \left(n\left(\frac{1}{2\pi} - \rho(p)\right)\Delta p\right)!}. \tag{32}$$

Using Stirling's formula to approximate a factorial, one arrives finds that equation (32) asymptotically becomes $2^{n\Delta S}$ where

$$\Delta S = \left[\frac{1}{2\pi}\log\left(\frac{1}{2\pi}\right) - \rho(p)\log(\rho(p)) - \left(\frac{1}{2\pi} - \rho(p)\right)\log\left(\frac{1}{2\pi} - \rho(p)\right)\right]\Delta p. \tag{33}$$

Now, as $\Delta p \to 0$, we find that asymptotically, ρ has a degeneracy of 2^{nS} where S, called the entropy, has the form

$$S = \int_{-\pi}^{\pi} dp\left[\frac{1}{2\pi}\log\left(\frac{1}{2\pi}\right) - \rho(p)\log(\rho(p)) - \left(\frac{1}{2\pi} - \rho(p)\right)\log\left(\frac{1}{2\pi} - \rho(p)\right)\right] \tag{34}$$

The following picture emerges. For a finite system, each collection $\{p\}$ corresponds to precisely one $|\{p\}\rangle \in W$, which is an eigenvector of $\mathcal{H}$. In the thermodynamic limit when the number of spins becomes infinite, we replace the momenta $\{p\}$ by the function ρ. But unlike the finite system, each ρ corresponds to many eigenvectors in the now infinite dimensional Hilbert space—the different microscopic states corresponding to a single macroscopic state [state of thermo equilibrium]. Moreover, each of these eigenvectors have the same energy density (see equation (36) below). The fact that every eigenvector corresponding to a single ρ has the same energy density is an instance of a crucial principle. Namely, all of the observables in the thermodynamic limit depend only on the macroscopic variable ρ. Consider for instance the energy $E(\{p\})$ (it is an observable):

$$E(\{p\}) = \sum_{j=1}^{m} \epsilon(p_j) = n \sum_{j=1}^{m} \epsilon(p_j) \frac{p_{j+1} - p_j}{n(p_{j+1} - p_j)}. \tag{35}$$

In the thermodynamic limit, one finds that energy density is equal to

$$\frac{E}{n} \to \int_{-\pi}^{\pi} \epsilon(p)\rho(p)dp. \tag{36}$$

The thermodynamic correlation function condition is a consequence of the principle stated above. In the thermodynamic limit, the matrix elements

$$\lim_{n\to\infty} \frac{\langle \{p\}|\mathcal{O}|\{p\}\rangle}{\langle \{p\}|\{p\}\rangle} \tag{37}$$

depend only on ρ, not on the set $\{p\}$. The function $\rho(p)$ which corresponds to the space of thermo equilibrium can be found from the variational equation (A.3), which arose from the steepest descent approximation.

$$\rho(p) = \frac{1}{2\pi}\theta(p) = \frac{1}{2\pi\left(1 + 2^{\frac{\epsilon(p)}{T}}\right)}. \tag{38}$$

The subspace C of thermodynamic equilibrium (see the appendix) is precisely the span of the eigenvectors corresponding to this ρ. Since every eigenvector in the thermo-equilibrium space corresponds to the same ρ, the correlation functions $\langle\psi|\mathcal{O}|\psi\rangle$ do not depend on which $\psi \in C$ is chosen. The challenge is to compute these correlation functions. In effect, it is by direct calculation [11–13] that the key principle is verified. In thermodynamics limit quantum error correction condition is fulfilled. For systems of large, but finite size, the quantum error correction condition may be violated for certain errors. We plan to study error propagation by investigating a perturbation of the quantum error correcting condition.

V. CONCLUSION

The choice of a thermo-equilibrium subspace as a code space provides a new approach to quantum error correction and supplies the field with a fundamentally new set of quantum error correcting codes. Since our codes are based on real, dynamic, physical systems at finite temperature, we believe that they have the potential to perform in real physical systems. In the sequel papers, we will explicitly investigate the codes arising from the XX0 and various other models (such as XXX, XXZ, XYZ, Haldane-Shastry, etc...). We will compare our codes with other codes and analyze advanced properties such as fault-tolerance and accuracy thresholds.

Acknowledgments

The authors would like to thank Doctor Maassen van den Brink for discussions.

APPENDIX A: THE DEFINITION OF THE THERMO-EQUILIBRIUM SUBSPACE

In models with physical relevance, the subspace of thermodynamic equilibrium can be defined as follows. Suppose that the size of the lattice n is very large. Recall the partition function defined in equation (13). One has

$$\mathcal{Z} = \operatorname{tr} 2^{\frac{-\mathcal{H}}{T}} = \sum_{\text{eigenvectors } v} 2^{-\frac{E(v)}{T}} = \sum_{\text{eigenvalues } E} 2^{nS} 2^{-\frac{E}{T}} \tag{A1}$$

where the factor 2^{nS} is the degeneracy of the energy level E and S is entropy. In thermodynamics limit $n \to \infty$ both energy and entropy increase linearly with n. So we can use the method of steepest descent to evaluate the partition function

$$\lim_{n\to\infty} \mathcal{Z} = \lim_{n\to\infty} \sum_{E} 2^{nS-\frac{E}{T}} \to \int 2^{nS-\frac{E}{T}} \tag{A2}$$

The method of steepest descent leads to the variational equation

$$\delta \left(S - \frac{E}{T} \right) = 0. \tag{A3}$$

This brings us to the key definition:

Definition A.1. The subspace of thermodynamic equilibrium is defined to be the span of the set of eigenvectors that solve equation (A3).

For the model at hand, we can be more specific. Using variational calculus, one has for XX0

$$\delta S = \int_{-\pi}^{\pi} dp \left[-\delta\rho(p)\log(\rho(p)) + \delta\rho(p)\log\left(\frac{1}{2\pi} - \rho(p)\right)\right] \tag{A4}$$

and

$$\delta E = \int_{-\pi}^{\pi} dp \left[-\delta\rho(p)\epsilon(p)\right]. \tag{A5}$$

Setting $\delta\left(S - \frac{E}{T}\right) = 0$ gives the solution given in equation (38). Therefore, the span of the eigenvectors that correspond to the function $\rho = \frac{1}{2\pi}\theta$ comprise the space C of thermodynamic equilibrium. We saw that the equation $\delta\left(S - \frac{E}{T}\right) = 0$ defines $\rho(p)$.

---

[1] C. E. Shannon, Bell Syst. Tech. J. p. 379 (1948).

[2] E. Knill and R. Laflamme, Phys. Rev. A **55**, 900 (1997).

[3] C. Benett, D. DiVincenzo, J. Smolin, and W. Wootters, Phys. Rev. A **54**, 3824 (1996).

[4] J. Preskill, *Lectures notes*, `www.theory.caltech.edu /people/preskill/ph219/` (1997-1999).

[5] I. L. Chuang and M. A. Nielsen, *Quantum Computation and Quantum Information* (Cambridge Univ. Press, 2000).

[6] A. Kitaev, A. Shen, and M. Vyalyj, *Classical and Quantum Computations* (AMS, 2002).

[7] A. R. Calderbank and P. W. Shor, Phys. Rev. A **54**, 1098 (1996).

[8] A. M. Steane, Phys. Rev. Lett. **77**, 793 (1996a).

[9] D. Gottesman, Ph.D. thesis, Caltech (1997), `quant-ph/9705052`.

[10] M. Freedman, A. Kitaev, M. Larsen, and Z. Wang, `quant-ph/0101025`.

[11] V. E. Korepin, N. M. Boguliubov, and A. G. Izergin, *Quantum Inverse Scattering Method and Correlation Functions* (Cambridge Univ. Press, 1993).

[12] F. Colomo, A. Izergin, V. Korepin, and V. Tognetti, Phys. Lett. A **169**, 243 (1992).

[13] A. Its, A. Izergin, V. Korepin, and N. Slavnov, Phys. Rev. Lett. **70**, 1704 (1993).

[14] M. Takahashi, *Thermodynamics of One-Dimensional Solvable Models* (Cambridge Univ. Press, 1999).

[15] W. K. Wootters (2000), `quant-ph/0001114`.

[16] H. Fu, X. Wang, and A. I. Solomon, `quant-ph/0105075`.

[17] E. Lieb, T. Schultz, and D. Mattis, Annals of Physics **16**, 461 (1961).

[18] A. G. Izergin, A. G. Pron'ko, and M. B. Zvonarev, Zap. Nauch Semin. POMI **268**, 207 (2000).

[19] R. Feynman, *Feynman Lectures on Computation* (Perseus Publ., 1996).

[20] D. Gottesman, Phys. Rev. **A57**, 127 (1998).

[21] R. Landauer, IBM J. Res. Dev. (1961).

[22] R. Landauer, Phys. Today pp. 23–29 (1991).

[23] J. C. Maxwell, *Theory of Heat* (Longmans, Green and Co., 1871).

[24] J. Preskill, Proc. Roy. Soc. Lond. **A454**, 385 (1998).

[25] J. Preskill, in *Introduction to Quantum Computation*, edited by H.-K. Lo, S. Popescu, and T. Spiller (World Scientific, 1998).

[26] B. Schumacher, Phys. Rev. pp. 2614–1628 (1996).

[27] B. Schumacher and M. A. Nielsen, Phys. Rev. pp. 2629–1635 (1996).

[28] A. M. Steane, Rept. Prog. Phys. pp. 117–173 (1998).