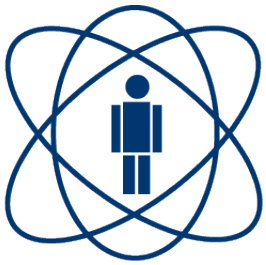


Centro Brasileiro de Pesquisas Físicas

Centro Brasileiro de Pesquisas Físicas

Isadora Barbosa Lima Veeren

# Entropic Uncertainty Relations and Classicality

Rio de Janeiro, RJ
2019

# "ENTROPIC UNCERTAINTY RELATIONS AND CLASSICALITY"

**ISADORA BARBOSA LIMA VEEREN**

Dissertação de Mestrado em Física, apresentada no Centro Brasileiro de Pesquisas Físicas do Ministério da Ciência, Tecnologia, Inovações e Comunicações. Fazendo parte da banca examinadora os seguintes professores:

Fernando da Rocha Vaz Bandeira de Melo - Presidente/Orientador

Daniel Schneider Tasca

Ivan dos Santos Oliveira Junior

Rio de Janeiro, 30 de abril de 2019.

# Agradecimentos

Agradeço aos professores e professoras que me acompanharam desde o ensino fundamental até o fim do mestrado, em especial às queridas tias Simone e Denise, que me guiaram nos primeiros passos das ciências e da matemática. Agradeço também aos professores e colegas que doaram seu tempo para discutir comigo questões ligadas à dissertação. Agradeço à UFRJ pela minha formação na graduação, e ao CBPF durante o mestrado. Agradeço ao meu orientador pela atenção e tempo que dedicou a esse trabalho. Agradeço também ao CNPq e à CAPES pelos recursos disponibilizados para a minha formação. Por último, agradeço à minha família, sem a qual nada seria possível.

# Resumo

Quando analisamos a transição do mundo quântico para a nossa vida cotidiana, muitas questões aparecem. Em uma primeira tentativa, podemos recorrer a tomar o limite de sistemas com infinitos qubits, ou talvez ao uso de coarse graining, mas essas tentativas não são o suficiente para responder nossas questões. Usando relação de incerteza entrópicas em oposição ao conhecido princípio de incerteza de Heisenberg, a tomando muitas inspiração de experimentos de RMN (que conseguem siultaneamente medir a magnetização total de um sistema em duas direções perpendiculares), construímos um modelo para medições clássicas que ilumina o problema, e tem sucesso em aprimorar nosso entendimento da emergência de características clássicas.

**Palavras chave:** RMN, princípio de incerteza de Heisenberg, medicões macroscópicas, princípios de incerteza entrópicos.

# Abstract

When analyzing the transition from the quantum world to our classical everyday life, many issues appear. At first, we may appeal to taking the limit of systems with infinite qubits, or maybe using coarse grainig, but these do not suffice to answer our questions. Using entropic uncertainty relations in opposition to the well known Heisenberg uncertainty principle, and taking many inspirations from NMR experiments (that can simultaneously measure the total magnetization of the system in two perpendicular directions), we construct a model for classical measurements that sheds some light into the matter, and succeeds to improve our understanding of the emergence of classical features.

**Keywords:** NMR, Heisenberg uncertainty principle, macroscopic measurements, entropic uncertainty principles.

# Contents

# Chapter 1

# Introduction

We are all acquainted with some of quantum mechanics most odd effects and properties. Even throughout pop culture, we have references to quantum tunneling [1], entanglement [2] [3], quantum teleportation [4], wave function collapse [5], and, of course, Heisenberg's uncertainty principle [6]. However, when looking around in our everyday life we do not experience any of that. Our best description of the world, so far, is fundamentally quantum [7], but it certainly doesn't seem so most of the time. Moreover, whenever we try to push the boundaries of quantum mechanics into macroscopic systems we end up with the most singular situations, such as Schroedinger's cat [8].

The transition from a macroscopic quantum ruled world to our everyday macroscopic world is far from simple. The explanations we are informally given, when there are any at all, usually go along the lines of taking the limit of many particles to make the state become classical. Other explanations rely on a coarse graining reasoning, or perhaps taking average values of quantities of interest. Nevertheless, all these explanations are constructed with a reasoning appealing to common sense, and lack a underlying mathematical foundation to make it a strong claim.

In classical mechanics, we rely on observables to describe a physical quantity that can be measured. Observables are also present in the description of measurements in quantum mechanics, but then they must obey Heisenberg's principle. This means that we can only prepare a quantum state with two quantities simultaneously well defined if the observables associated with them commute, which is obviously not an issue in classical mechanics. This is only one feature from quantum mechanics that differs from classical mechanics and lacks a better understanding on the transition from one to another.

Another feature of special interest in this work will be the ability of jointly measuring incompatible observables. We know that, at least at first, we are not able to successfully jointly measure two observables that do not commute, most notably position and momentum in the continuous case, or a spin component in the x and z directions in the discrete case. We do now, however, that in NMR experiments we get to simultaneously measure the total nuclear magnetization of carbon and hydrogen atoms in two perpendicular directions. For this reason, NMR systems and setup will be the major source of inspiration and guidance in this work.

The main quest of this dissertation is to better understand the transition from quantum to classical mechanics, particularly from the point of view of uncertainty relations: what can we do to better understand how classical systems satisfy the requirements on preparation and measurements posed by quantum mechanics?

To address this matter, we will start by discussing entropic uncertainties as substitutes for Heisenberg's uncertainty principle in chapter 2, pointing out the shortcomings of the latter as a preparation uncertainty relation and providing an alternative based on entropic relations, as well as presenting measurement uncertainty relations. We then proceed, in chapter 3, to discuss how to model macroscopic measurements by introducing type measurements, as well as providing support for investigating the specific set of product states, inspired from NMR. Finally, in chapters 4 and 5 we present our results, the motivation to use a division by bins to incorporate imprecision and inaccuracy in our model, and the numeric and analytical conclusion we draw from them.

# Chapter 2

# Entropic Uncertainties

When discussing uncertainty relations, Heisenberg's uncertainty principle is what mostly comes to mind, at first. When approached in introductory quantum mechanics courses, it is usually presented like this:

$$\Delta_x \Delta_{p_x} \geq \frac{\hbar}{2}. \tag{2.1}$$

Here, $\Delta_Q$ is the standard deviation of an operator $Q$ with respect to the state $\rho$, calculated as

$$\Delta_Q = \sqrt{\langle Q^2 \rangle - \langle Q \rangle^2},$$

where $\langle Q \rangle = Tr[\rho Q]$ is the mean value of $Q$ with respect to $\rho$. It means that there is a lower bound for the product of the standard deviations of position and momentum for a fixed state preparation.

In his 1927 paper [6], Heisenberg uses a heuristic semi-classical argument, to arrive, actually, at a simpler relation that he expressed symbolically as:

$$p_1 q_1 \approx h,$$

where $p_1$ is the position inaccuracy and $q_1$ is the momentum inaccuracy, that Heisenberg connected with the spreads of the wave function. This relation was later expanded to any two observables by Robertson [11]:

$$\Delta_A \Delta_B \geq \frac{1}{2} |\langle \Psi | [A, B] | \Psi \rangle |. \tag{2.2}$$

The scenario corresponding to Heisenberg's uncertainty principle is as follows: the spreads of observable $A$, $\Delta_A$, and observable $B$, $\Delta_B$, (or, specifically, position and momentum) are determined separately, in different experiments on the same state. Physically, it means that if the two observables do not commute, then we cannot prepare a system in a state with well defined properties given by $A$ and $B$.

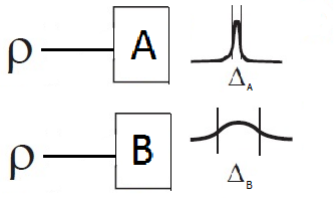



Figure 2.1: In this scenario, we never measure both position and momentum on the same particle. In two different sets of experiments, we estimate the spreads of $A$ and $B$ on the same initial state $\rho$. Figure adapted from [24]

However, according to Busch [24], it is not the scenario Heisenberg adopts on his paper. The scenario that was actually discussed by Heisenberg is that of sequential measurements on the same state. In this scenario, one could compare the distribution of momentum after an approximate position measurement with an ideal distribution of momentum without being disturbed first by a position measurement.

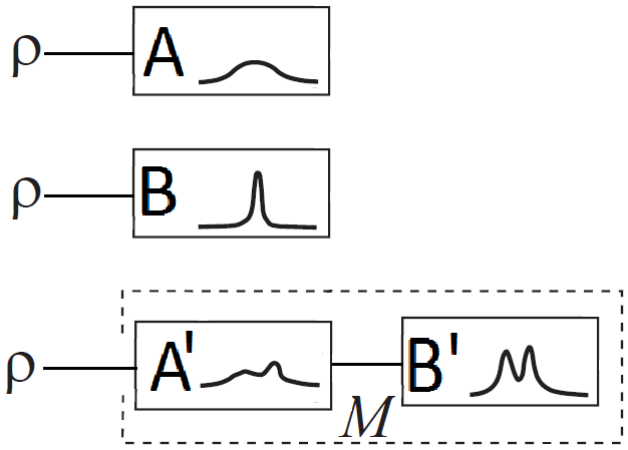



Figure 2.2: In this scenario, we perform sequential measurements on the same initial state. We first determine statistics of the ideal observables $A$ and $B$, that will serve as a standard. Then, we measure an observable $A$ and sequentially measure $B$, resulting on an effective measurement of $A'$ and $B'$ on the initial state. We then compare the statistics obtained from $A$ and $B$ with $A'$ and $B'$. The measurement of $A$ followed by $B$ can be viewed as a single measurement $M$. Figure adapted from [24]

To determine the error of the approximate position measurement $X'$ compared to the ideal one $X$, we evaluate the accuracy that accounts for the difference between the distributions obtained by $X$ and $X'$. The same goes for the momentum measurement $P$ and the disturbed approximate momentum measurement (or effective momentum measurement) $P'$.

The situation described in the first case by figure 2.1 sets the scenario for a preparation uncertainty relation, such as the traditional Heisenberg's uncertainty principle, where we account for how far from the Dirac delta the two distributions can simultaneously be. These uncertainty relations state simply that two deviations cannot be simultaneously small for non-commuting observables.

The situation described in the second case by figure 2.2 is that of a measurement uncertainty relation, where we account for how far from an ideal distribution the disturbed results are. The question that is being posed is "can two observables $A$ and $B$ be accurately jointly measured?" or, put in another way, "are $A$ and $B$ compatible?", or yet "how well do $A'$ and $B'$ reproduce $A$ and $B$ for any state?". We point out that the sequential measurement scenario described above is simply a particular case of joint measurements, as worked out in [33], since we can always put the two successive measurements inside a black box and call them a single joint measurement.

## 2.1 Entropic Uncertainty Relations

When dealing with classical physics, the only source of uncertainty are the experimental limitations of the measurement procedure. However, when dealing with quantum physics we face an additional element of uncertainty, a consequence of its underlying fundamental nature. Though Heisenberg-like uncertainty principles are very insightful in the sense that they call attention to intrinsically quantum features, they can display some limitations and inconveniences.

Firstly, these relations are generally state dependent. It means that $\Delta_A$ getting bigger or smaller doesn't necessarily imply on restrictions on $\Delta_B$, because the state $|\Psi\rangle$ may be changed and compensate for it. We can divide the whole inequality by $\Delta_B$,

$$\Delta_A \geq \frac{1}{2\Delta_B} |\langle\Psi| [A, B] |\Psi\rangle|.$$

It means that even if $\Delta_B$ gets bigger we cannot make any conclusions about $\Delta_A$ having to get smaller, because we can change the state $|\Psi\rangle$ so that $|\langle\Psi| [A, B] |\Psi\rangle|$ gets bigger. We also cannot simply minimize the module of the commutator over all possible states because for finite dimension the bound would always be trivial.

$$\Delta_A \Delta_B \geq \min_{|\Psi\rangle} \frac{1}{2} |\langle\Psi| [A, B] |\Psi\rangle| = 0.$$

Secondly, these relations are not defined for POVMs. In quantum mechanics, the most general form of measurement operators, which will be specially useful in section 5, is given by a POVM. Standing for *Positive Operators Valued Measure*, a POVM is a set of $n$ operators $E_i$ obeying

- Positivity: $E_i \geq 0$, meaning $\langle\Psi|E_i|\Psi\rangle \geq 0$ for any vector $|\Psi\rangle$;
- Completeness: $\sum_i E_i = \mathbb{1}$.

It means that, when using Heisenberg-like uncertainty principles, we could only analyze the cases involving projective measurements.

It is of interest to construct other uncertainty relations that will not present these downsides. Following Deutch's approach [12], we must first determine a good measure for uncertainty. This quantity should obey a few properties:

- It should reach a global minimum (that we can chose to be zero) if and only if $A$ and $B$ have a shared eigenstate (or arbitrarily close eigenstates). Any other case should yield a positive value. It is the case for Heisenberg's inequality (2.1) (because it is state independent), but not for Robertson's (2.2), as follows. If $|\Psi\rangle$ is in the kernel of $A$, for instance, the inequality's lower bound will vanish regardless of their commutator. Take, for instance, the Pauli matrices $A = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $B = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Let $|\Psi_z^+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ be the eigenvector of $\sigma_z$ associated with eigenvalue $+1$ and $|\Psi_z^-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ be the eigenvector of $\sigma_z$ associated with eigenvalue $-1$. In this situation, equation (2.2) would give

$$\Delta_x \Delta_z \geq \frac{1}{2}|\langle\Psi_z^+|\,[\sigma_x, \sigma_z]\,|\Psi_z^+\rangle\,|,$$

$$= \frac{1}{2}|\langle\Psi_z^+|\,(\sigma_x\sigma_z - \sigma_z\sigma_x)\,|\Psi_z^+\rangle\,|,$$

$$= \frac{1}{2}|\langle\Psi_z^+|\,\left(\sigma_x\,|\Psi_z^+\rangle - \sigma_z\,|\Psi_z^-\rangle\right)|,$$

$$= \frac{1}{2}|\langle\Psi_z^+|\,\left(|\Psi_z^-\rangle + |\Psi_z^-\rangle\right)|.$$

  Since $\langle\Psi_z^+|$ and $\langle\Psi_z^-|$ are orthogonal, the lower bound of the inequality vanishes, even though $[\sigma_x, \sigma_z] = -2i\sigma_y$
- It shouldn't be responsive to rescaling, in the sense that a simple relabeling of the eigenvalues shouldn't have any effects. When evaluating the standard deviation, the absolute value of the eigenvalues has a big impact on the results. If the eigenvalues are multiplied by a factor $\gamma$, the mean will also be multiplied by this same factor, and ultimately the standard deviation will also be rescaled by $\gamma$. In this case, the absolute value of the standard deviation loses its significance, and all information we can get is whether it is zero or non-zero.

### 2.1.1 Entropies

Entropies are often regarded as a measure of diversity or randomness of a system. In information theory, it is usually used to assess how much uncertainty there is in a state of a physical system before a measurement, or equivalently how much information we gain after a measurement [29]. It is the usual choice of measure for uncertainty. Moreover, they automatically satisfy the second requirement just posed, since they depend only on probability distributions, without relying on eigenvalues.

Suitable candidates for the measure of uncertainty of a single observable are the Renyi entropies, defined for a random variable $X$ with $N$ possible outcomes $\{\, n_1, n_1, ..., n_N \,\}$ and respective probabilities given by the vector $P = \{\, p_1, p_2, ..., p_N \,\}$ as

$$H_\alpha(X) = \frac{1}{1-\alpha} \log\left(\sum_{i=1}^{N} p_i^\alpha\right),$$

where $\alpha > 0$ and $\alpha \neq 1$. Consequently, the Renyi entropy for a distribution obtained from measuring an observable $Q$ with eigenvectors $|q_i\rangle$ on the state $|\Psi\rangle$:

$$H_\alpha(Q|\Psi) = \frac{1}{1-\alpha} \log \left( \sum_{i=1}^{N} |\langle q_i|\Psi\rangle|^{2\alpha} \right).$$

A desirable feature satisfied by Renyi entropies is that their value will be maximum when the outcomes are equally probable, the situation in which we have the greatest amount of uncertainty. On the other hand, when a single outcome is associated with a probability of 1 the entropy of the distribution is zero, compatible with a situation of no uncertainty where we know for sure which outcome we will get.

The Renyi entropies are related to the p-norm of the vector of probabilities, given by $||P||_\alpha = (p_1^\alpha + p_2^\alpha + ... + p_N^\alpha)^{1/\alpha}$, so that

$$H_\alpha(X) = \frac{\alpha}{1-\alpha} \log \left( ||P||_\alpha \right).$$

Different values of $\alpha$ define different entropies, which are all non-increasing in $\alpha$ for a same probability distribution, meaning that

$$H_0(X) \geq H_1(X) \geq H_2(X) \geq H_\infty(X) \geq 0.$$

As $\alpha$ approaches 0, the Renyi entropy gives the same weight to all possible events, even though they might have different probabilities. When $\alpha \to 0$, it is simply the logarithm of the number of outcomes for $X$

$$H_0(X) = \log \left( \sum_{i=1}^{N} p_i^0 \right) = \log N.$$

On the other hand, as $\alpha$ approaches $\infty$, the entropy gives increasingly more weight to the most probable outcome, eventually reaching

$$H_\infty(X) = -\log \max_i p_i,$$

as $\alpha \to \infty$.

A special case of interest is $\alpha \to 1$, as we recover the Shannon entropy:

$$H_1(X) = H(X) = \lim_{\alpha \to 1} H_\alpha(X) = -\sum_{i=1}^{N} p_i \log p_i.$$

The Shannon entropy is of special interest, and usually used instead of the more general Renyi entropy, because it is the only case for which the chain rule for conditional entropy holds. Let $Y$ be a discrete random variable taking values $y \in \mathcal{Y}$ conditioned on the discrete random variable $X$ taking values $x \in \mathcal{X}$. Then,

$$H(Y|X) = H(X,Y) - H(X),$$

where $H(Y|X) = -\sum_{x,y} p(x,y) \log \frac{p(x)}{p(x,y)}$ is the conditional entropy of $Y$ given $X$ and $H(X,Y) = -\sum_{x,y} p(x,y) \log p(x,y)$ is their joint entropy. This is a desirable feature because it reproduces this intuitive notion: if you consider a system given by two random variables $X$ and $Y$, you need on average $H(X,Y)$ bits of information to describe it. However, once you learn the value of $X$ you only need $H(X,Y) - H(X)$ bits of information, and this quantity is precisely $H(Y|X)$.

## 2.2 Preparation Entropic Uncertainty Relations

Now, it is only reasonable that we choose our measure of uncertainty to be the sum of the Renyi entropies for two different observables measured on the same state, or particularly the sum of their Shannon entropy $H(A|\Psi) + H(B|\Psi)$. If $|a_j\rangle$ and $|b_k\rangle$ are the eigenvectors of $A$ and $B$, respectively, then this can be rearranged as

$$H(A|\Psi) + H(B|\Psi) = -\sum_{j=1}^{N} |\langle\Psi|a_j\rangle|^2 \log |\langle\Psi|a_j\rangle|^2 - \sum_{k=1}^{N} |\langle\Psi|b_k\rangle|^2 \log |\langle\Psi|b_k\rangle|^2,$$

$$= -\sum_{j,k} |\langle\Psi|a_j\rangle|^2 |\langle\Psi|b_k\rangle|^2 \left(\log|\langle\Psi|a_j\rangle|^2 + \log|\langle\Psi|b_k\rangle|^2\right).$$

As we want to achieve a state independent relation, we will minimize the right hand side of the equation. The term between parenthesis is non-positive, and, because of its symmetries, reaches its maximum value when $|\Psi\rangle$ lies midway between $|a_j\rangle$ and $|b_k\rangle$. Bialynicki-Birula and Rudnicki obtain it explicitly in [25]:

$$|\phi_{j,k}\rangle = \frac{1}{\sqrt{2\,(1+|\langle a_j|b_k\rangle|)}} \left(|a_j\rangle + e^{-i\arg\langle a_j|b_k\rangle}\,|b_k\rangle\right).$$

Since this is the case when the term in parenthesis has its maximum, we have an inequality given by

$$H(A|\Psi) + H(B|\Psi) \geq -\sum_{j,k} |\langle\Psi|a_j\rangle|^2 |\langle\Psi|b_k\rangle|^2 \left(\log|\langle\phi_{jk}|a_j\rangle|^2 + \log|\langle\phi_{jk}|b_k\rangle|^2\right),$$

or as worked out in [25],

$$H(A|\Psi) + H(B|\Psi) \geq -2\sum_{j,k} |\langle\Psi|a_j\rangle|^2 |\langle\Psi|b_k\rangle|^2 \log\frac{1}{2}\,(1+|\langle a_j|b_k\rangle|),$$

$$= 2\sum_{j,k} |\langle\Psi|a_j\rangle|^2 |\langle\Psi|b_k\rangle|^2 \log\frac{2}{(1+|\langle a_j|b_k\rangle|)}.$$

This inequality will still hold if we take the lowest lower bound possible, namely by replacing $|\langle a_j|b_k\rangle|$ for $\max_{j,k}\{|\langle a_j|b_k\rangle|\}$, and sum over all $j,k$ to arrive at the final inequality:

$$H(A|\Psi) + H(B|\Psi) \geq 2\log\frac{2}{1+\max_{j,k}\{|\langle a_j|b_k\rangle|\}}.$$

This is the inequality obtained by Deutsch in [12]. It obeys all the expected characteristics for a measure of uncertainty: firstly it doesn't depend on the eigenvalues of the observables $A$ and $B$, meaning it is not sensitive to rescaling; secondly, it vanishes if and only if $A$ and $B$ have a common eigenvector, in which case $\max_{j,k}\{|\langle a_j|b_k\rangle|\} = 1$ and we are left with $\log 1 = 0$. Moreover, we get rid of the dependence on the state $|\Psi\rangle$ in a non-trivial way, another advantage over the Heisenberg generalized inequality.

A few years later, Kraus had conjectured [13] that this bound could be improved to

$$H(A|\Psi) + H(B|\Psi) \geq -2\log c,$$

where $c = \max_{j,k}|\langle a_j|b_k\rangle|$.

This conjecture was proved by Maassen and Uffink in the following year[14]. We introduce the following theorem due to Riesz[15]:

**Riesz Theorem.** *Let $x = (x_1, x_2, ..., x_N)$ denote a sequence of complex numbers and $T_{jk}$ a linear transformation matrix element, so that $(Tx)_j = \sum_k T_{jk}x_k$ and $\sum_j |(Tx)_j|^2 = \sum_k |x_k|^2$ for all $x$, and let $c = \max_{j,k}|T_{jk}|$. Then,*

$$c^{1/a'}\left[\sum_j |(Tx)_j|^{a'}\right]^{1/a'} \leq c^{1/a}\left[\sum_k |(x_k)_j|^a\right]^{1/a},$$

*for $1 \leq a \leq 2$ and $1/a + 1/a' = 1$.*

We can associate $x_k = \langle a_k|\Psi\rangle$ and $T_{jk} = \langle b_j|a_k\rangle$, and consequently $(Tx)_j = \sum_k \langle b_j|a_k\rangle\langle a_k|\Psi\rangle = \langle b_j|\Psi\rangle$. If we chose $a = 2(1+r)$ and $a' = 2(1+s)$, we can rearrange the former equation:

$$\left[\sum_j |\langle b_j|\Psi\rangle|^{2(1+s)}\right]^{1/2(1+s)} \left[\sum_k |\langle a_k|\Psi\rangle|^{2(1+r)}\right]^{-1/2(1+r)} \leq c^{\frac{s-r}{2(1+r)(1+s)}}.$$

Let $P$ and $Q$ denote an arbitrary probability distribution of $N$ possible outcomes. We relabel $P_j = |\langle b_j|\Psi\rangle|^2$ and $Q_k = |\langle a_k|\Psi\rangle|^2$, and notice that since

$$\frac{1}{a} + \frac{1}{a'} = 1,$$

then

$$r = \frac{-s}{2s+1},$$

so

$$\frac{s-r}{2(1+r)(1+s)} = s/(s+1).$$

With that, we can update the expression to

$$\left[\sum_j P_j^{(1+s)}\right]^{1/2(1+s)} \left[\sum_k Q_k^{(1+r)}\right]^{-1/2(1+r)} \leq c^{s/(s+1)},$$

and identify $M_s(P) = \left[\sum_j P_J^{1+s}\right]^{1/s}$ and $M_r(Q) = \left[\sum_k Q_k^{1+r}\right]^{1/r}$ for clarity, leaving us with

$$M_s(P)^{s/2(s+1)} M_r(Q)^{r/2(r+1)} \leq c^{s/(s+1)}.$$

Using once again $r = -\frac{s}{(2s+1)}$,

$$M_s(P)^{s/2(s+1)} M_r(Q)^{s/2(s+1)} \leq c^{s/(s+1)},$$

and, finally,

$$M_s(P) M_r(Q) \leq c^2.$$

Now, we replace $M_s(P)$ and $M_r(Q)$

$$\left[\sum_j P_j^{1+s}\right]^{1/s} \left[\sum_k Q_k^{1+r}\right]^{1/r} \leq c^2,$$

take the logarithm on both sides to get

$$\frac{1}{s} \log \sum_j P_j^{1+s} + \frac{1}{r} \log \sum_k Q_k^{1+r} \leq 2 \log c,$$

and relabel $1 + s = s'$ and $1 + r = r'$

$$-\frac{1}{1-s'} \sum_j P_j^{s'} - \frac{1}{1-r'} \sum_k Q_k^{r'} \leq 2 \log c.$$

These are Renyi entropies, so we have

$$H_\alpha(P) + H_\alpha(Q) \geq -2 \log c.$$

This is the entropic preparation uncertainty relation for Renyi entropies. To prove Kraus' conjecture, we finally take the limits $s' \to 1$, $r' \to 1$ and identify the Shannon entropy

$$H(P) + H(Q) \geq -2 \log c.$$

and since $P_j = |\langle b_j|\Psi\rangle|^2$ and $Q_j = |\langle a_j|\Psi\rangle|^2$

$$H(B|\Psi) + H(A|\Psi) \geq -2 \log c. \tag{2.3}$$

We can now revisit the example carried out in 2.1 to point out a shortcoming of Heisenberg's equation. Then, we evaluated the lower bound given by equation (2.2) to show that if $|\Psi\rangle$ is in the kernel of $A$ (or, particularly, of $\sigma_x$) the inequality yields zero even though $A$ and $B$ (or $\sigma_x$ and $\sigma_z$) do not commute.

The problem will not be an issue if we use equation 2.3 instead of (2.2). We want to evaluate

$$H(\sigma_x|\Psi_z^+) + H(\sigma_z|\Psi_z^+) \geq -2 \log c,$$

where once again $\Psi_z^+$ is the eigenvector of $\sigma_z$ associated with eigenvalue 1. The bound given by the right hand side is state independent, so it will be the same whether $\Psi$ is in the kernel of $\sigma_x$ or not. Since the maximum overlap between eigenvectors of $\sigma_x$ and $\sigma_z$ is $1/\sqrt{2}$, we get

$$H(\sigma_x|\Psi_z^+) + H(\sigma_z|\Psi_z^+) \geq \log 2.$$

Contrary to Heisenberg-like uncertainty principles, preparation entropic uncertainty relations do not yield a trivial bound unless $A$ and $B$ have a shared eigenstate.

Now that we have these preparation uncertainty relations, we proceed to investigate the measurement uncertainty relations.

## 2.3 Measurement Uncertainty Relations

We now move on to the scenario of measurement uncertainty relations, given by the situation that was actually discussed by Heisenberg in [6]. In this case, two observables will be sequentially measured on a state $\rho$. The first observable $A$ to be measured will potentially disturb results from the second measurement $B$, when compared to the results that $B$ would have yielded without measuring $A$ first.

The question that Heisenberg poses is if two observables can be jointly measured. We want to achieve uncertainty relations that can answer that question, and additionally give us a quantitative measure of how compatible these two observables are. In other words, we want to be able to compute how wrong it is to substitute a target observable by its approximation, or how inaccurate it is to take $(A, B) \simeq (A', B')$, as in figure 2.2. If the measurements of $A$ and $B$ don't disturb one another (if they are compatible), we expect that $A = A'$ and $B = B'$. Measurement uncertainty relations should reflect that, yielding a quantity that equates zero if $A$ and $B$ are compatible and some positive number otherwise.

To investigate these uncertainty relations, we must first present some relevant concepts [32].

Let $\mathcal{H}$ be the finite dimensional complex Hilbert space, and $\mathcal{L}(\mathcal{H})$ be the set of linear operators on $\mathcal{H}$. Let $\mathcal{M}(\mathcal{X})$ be the subset of $\mathcal{L}(\mathcal{H})$ of POVMs with outcomes in $\mathcal{X}$.

*Marginal POVMs and compatibility*: Let $M \in \mathcal{M}(\mathcal{X} \times \mathcal{Y})$ be a POVM that yields two outcomes $x, y \in \mathcal{X} \times \mathcal{Y}$. We define the marginal POVMs $M_x^{[1]}$ and $M_y^{[2]}$ by

$$M_x^{[1]} = \sum_{y \in \mathcal{Y}} M_{x,y},$$

$$M_y^{[2]} = \sum_{x \in \mathcal{X}} M_{x,y},$$

where $M_{x,y}$ is the POVM element that has outcomes $x$ and $y$. We say $A \in \mathcal{M}(\mathcal{X})$ and $B \in \mathcal{M}(\mathcal{Y})$ are compatible if they are jointly measurable, meaning there exists a parent POVM $M \in \mathcal{M}(\mathcal{X} \times \mathcal{Y})$ that has marginals recovering $M^{[1]} = A$ and $M^{[2]} = B$, as discussed in [33]. $A$ has POVM elements $A_x$ associated with outcome $x$, and $B$ has POVM elements $B_y$ associated with outcome $y$, obeying $\sum_x A_x = \sum_y B_y = \mathbb{1}$.

If we want to model sequential measurements, we need to use other mathematical tools. In order to implement the approximate measurement $A'$, with POVM elements ${A'}_x$ , we start by defining an *instrument* yielding an outcome $x \in \mathcal{X}$ acting on the state $\rho$:

$$\mathcal{J}_x[\rho] = J_x \rho J_x^\dagger,$$

$\forall \rho$. $J_x$ is the Kraus operator associated with ${A'}_x$ such that $\sum_{x \in \mathcal{X}} J_x^\dagger J_x = \mathbb{1}$. The POVM $A'$ can be implemented as the adjoint instrument

$$\mathcal{J}_x^\dagger[\rho] = J_x^\dagger \rho J_x$$

acting on the identity operator

$$\mathcal{J}_x^\dagger[\mathbb{1}] = J_x^\dagger J_x \equiv {A'}_x,$$

Meaning $J_x$ are Kraus operators of $A'$. Notice that the Kraus operators are $J_x^\dagger = \sqrt{{A'}_x} U$, so that $J_x^\dagger J_x = \sqrt{{A'}_x} U U^\dagger \sqrt{{A'}_x} = {A'}_x$, where $U$ is an unitary operator. It means we can choose which $U$ to use, so the instruments implementing $A'$ are not unique. We use this to implement ${A'}_x$, the first measurement to be performed on $\rho$. A *sequential measurement* of $A$ and $B$ with outcomes $x$ and $y$, in this order, is then described as

$$M_{x,y} = \mathcal{J}_x^\dagger[B_y] = J_x^\dagger B_y J_x,$$

where $B_y$ is the POVM element of $B$ associated with the outcome $y$. This description is useful because we get marginals

$$M_x^{[1]} = \sum_{y\in\mathcal{Y}} J_x^\dagger B_y J_x = J_x^\dagger J_x = A'_x,$$

the first approximate measurement of $A$, and

$$M_y^{[2]} = \sum_{x\in\mathcal{X}} J_x^\dagger B_y J_x = \mathcal{J}_\mathcal{X}^\dagger\left[B_y\right],$$

the second measurement of $B'$, a potentially perturbed version of $B_y$. The measurement of $B$ on the state $\rho'$ resulting from the measurement of $A$ can thus be better regarded as an effective measurement $B'$ on the initial state $\rho$.

Particularly, if $M_x^{[1]} = A_x$ and $M_y^{[2]} = B_y$ we say $A$ and $B$ are sequentially compatible observables, because the measurements of $A$ did not disturb $B$. We call attention once more to the fact that if $A$ and $B$ are sequentially compatible then they are also jointly measurable, but not necessarily the other way around: not all joint measurements can be performed sequentially. Two concrete examples are carried out in [30].

Ultimately, what we want to do is to compare the probability distributions of the target observable and its approximation. In order to do that, one possibility is to use relative entropies for discrete probabilities. Let $p = \{p_1, p_2, \ldots, p_N\}, q = \{q_1, q_2, \ldots, q_N\}$ be two probability vectors. The relative entropy of $p$ with respect to $q$ is

$$H(p||q) = \begin{cases} \sum_{x\in\text{supp of p}} p(x)\log\frac{p(x)}{q(x)} \text{ if supp } p \subseteq \text{supp } q, \\ +\infty \text{ otherwise.} \end{cases}$$

It is a way to account for how wrong we would be if we used probability $q$ instead of $p$, which is exactly what we want when comparing the probability distributions yielded by the POVM $A$ and its approximation $A'$. It is a measure of the information lost when we take $q$ instead of $p$ [31]. Relative entropies hold the property that $H(p||q) = 0 \iff p = q$, which is a very desirable feature since we want our uncertainty relation to equal 0 in the case of measurement of compatible POVMs.

### 2.3.1 Formulating Measurement Uncertainty Relations

Now, we can use relative entropies to build entropic measurement uncertainty relations, following [10]. We know that $A$ and $B$ are generally incompatible, meaning only an approximate joint measurement is possible. Moreover, $A$ will generally disturb $B$, and $A'$ and $B'$ can be very far from $A$ and $B$. Hence, we need to account for two different aspects of measurement uncertainty: how incompatible $A$ and $B$ are, and how distant $A'$ and $B'$ are from $A$ and $B$. To evaluate the first case, we will use the entropic incompatibility degree, and the second case will be addressed by an error-disturbance trade off. In both cases we will eventually look for the optimal parent POVM $M \in \mathcal{M}(\mathcal{X}\times\mathcal{Y})$ whose marginals $M^{[1]}$ and $M^{[2]}$ best approximate $A$ and $B$.

We start by defining the *error function* of the approximation $(A, B) \approxeq (M^{[1]}, M^{[2]})$ as

$$H\left[A, B||M\right](\rho) = H(A_{(\rho)}||M^{[1]}_{(\rho)}) + H(B_{(\rho)}||M^{[2]}_{(\rho)}),$$

where $Q_{(\rho)}$ is $\{Tr(\rho Q_x)|x \in \mathcal{X}\}$ , the probability distribution for $Q$ on the state $\rho$. We know that if $A$ and $B$ are compatible, by definition $A_{(\rho)} = M^{[1]}_{(\rho)}$ and $B_{(\rho)} = M^{[2]}_{(\rho)}$, meaning $H\left[A, B||M\right](\rho) = 0$.

From that, we define the *entropic divergence*, that is the worst total loss of information due to assuming $(A, B) \approxeq (M^{[1]}, M^{[2]})$

$$D[A, B||M] = \sup_\rho H\left[A, B||M\right](\rho) = \sup_\rho\{H(A_{(\rho)}||M^{[1]}_{(\rho)}) + H(B_{(\rho)}||M^{[2]}_{(\rho)})\}.$$

Clearly, if $A$ and $B$ are compatible, $D[A, B||M] = 0$.

Now, what if we take the best case scenario and choose the approximations $M^{[1]}, M^{[2]}$ so that they are as similar to $A$ and $B$ as possible? To evaluate the incompatibility of $A$ and $B$ we want to assess the minimum loss of information that cannot be avoided no matter which approximations $M^{[1]}, M^{[2]}$ we

use. We do that by taking the minimum entropic divergence over all possible $M$. We define the *entropic incompatibility degree* as

$$c_{inc}(A,B) = \inf_{M \in \mathcal{M}(\mathcal{X}\times\mathcal{Y})} D[A,B||M].$$

Alternatively, to evaluate the error/disturbance trade off, we investigate the particular cases of sequential measurements. The measurement of $A$ followed by an approximate measurement of $B$ is the POVM described by instruments as

$$\mathcal{M}(\mathcal{X};B) = \{M \in \mathcal{M}(\mathcal{X}\times\mathcal{Y}) : M(x,y) = \mathcal{J}_x^\dagger[B(y)] \forall x,y\}.$$

In general, its marginals $A \neq A'$ and $B \neq B'$, unless $A$ and $B$ are sequentially compatible.

In an analogous way with the entropic incompatibility degree, we define the *entropic error/disturbance coefficient* by minimizing the entropic divergence over all possible $M$ that are sequential measurements of $A$ and $B$:

$$c_{ed}(A,B) = \inf_{M \in \mathcal{M}(\mathcal{X};B)} D[A,B||M].$$

We point out that, in general, $c_{ed}(A,B) \neq c_{ed}(B,A)$, which correctly depicts our expectations, since a measurement of $A$ followed by $B$ isn't generally the same as a measurement of $B$ followed by $A$. Also, since sequential measurements are a special case of joint measurements (or, in other words, $\mathcal{M}(\mathcal{X};B) \subseteq \mathcal{M}(\mathcal{X};\mathcal{Y})$ ), it follows naturally that $c_{inc}(A,B) \leq c_{ed}(A,B)$. They will generally be different, except in the special case where $B$ is a sharp observable, as proved in [10].

Finally, we can find an upper bound for $c_{inc}(A,B)$ and $c_{ed}(A,B)$. We already know that $0 \leq c_{inc}(A,B) \leq c_{ed}(A,B)$. In [10], Theorem 2 item $(iii)$, it is proved that

$$0 \leq c_{inc}(A,B) \leq c_{ed}(A,B) \leq \log|\mathcal{X}| - \inf_\rho H(A_{(\rho)}),$$

$$c_{inc}(A,B) \leq \log|\mathcal{Y}| - \inf_\rho H(B_{(\rho)}),$$

where $|\mathcal{X}|$ and $|\mathcal{Y}|$ are the cardinalities of $\mathcal{X}$ and $\mathcal{Y}$, the number of outcomes of $A$ and $B$. These bounds can be investigated exploring different systems and setups, and we can evaluate under which conditions we get the lowest upper bound and how that can be useful to recover classical features.

Computing the bounds $c_{inc}$ and $c_{ed}$ can be quite challenging. For a pair of measurements $A$ and $B$, we must find all possible "parent" POVMs $M(x,y) \in \mathcal{M}(\mathcal{X}\times\mathcal{Y})$ that yield $A'$ and $B'$ as marginals, and perform the minimizations on this set of POVMs. This is a problem that can de solved via a semidefinite program [34], but can be a bit overwhelming and time consuming for a first approach.

Alternatively, what we are actually going to use to evaluate measurement uncertainty relations is another tool listed by Busch, Lahti and Werner [27] as possible options. When $B$ and $B'$ are not the same, each will yield a different probability distribution, and the distance between them can be used as a measure for their incompatibility.

We must perform two sets of experiments: in one we will obtain the probability distributions of an ideal observable $B$; in the other we will obtain the probability distributions of the approximate observable $B'$ after an ideal measurement $A$ that will potentially disturb it. With these two distributions at hand, we can compare them and quantify how different they are from each other.

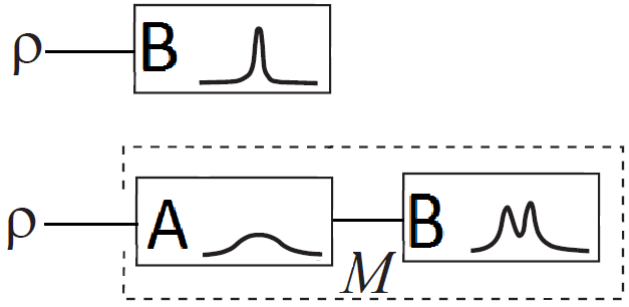



Figure 2.3: One set of experiments will estimate the statistics of $B$, while another set of experiments will estimate de statistics of $B'$, so we can compare them afterwards. Figure adapted from [24]

To compare the two statistics obtained from the experiments, we can choose to use the Euclidean distance between two points in the outcome space.

$$D(\mathbf{p}, \mathbf{p}') = \sqrt{\sum_{i=1}^{N}(p_i - p'_i)}, \tag{2.4}$$

where $\mathbf{p} = (p_1, p_2, \ldots, p_N)$ is the probability vector of the undisturbed measurement of $B$ and $\mathbf{p}' = (p'_1, p'_2, \ldots, p'_N)$ is the probability vector of $B'$ after an ideal measurement $A$.

This way, the result of $D(\mathbf{p}, \mathbf{p}')$ works as a measure for the disturbance caused by $A$ on the measurement of $B$. Ideally, if $A$ and $B$ are compatible this distance will be zero because the probability of $B$ will not be affected by $A$. However, the converse is not true: the Euclidean distance between $\mathbf{p}$ and $\mathbf{p}'$ being null does not imply on them being compatible, since it can be a state dependent coincidence. Moreover, we would be doing this analysis on a specific state, which by no means can be extended to a generic state and let us conclude that $A$ and $B$ are compatible. Any favorable result obtained from this analysis can only serve as a hint, for not prohibiting $A$ and $B$ to be compatible.

Now that we already have these concepts, we can move on to investigate another prerequisite to develop the work concerned in this dissertation.

# Chapter 3

# Macroscopic Measurements

In nuclear magnetic resonance (NMR), a sample containing $N \approx 10^{20}$ molecules is probed through successive radio-frequency pulses controlled by an external source, inducing state transitions [28]. Computation is conducted using the nuclear spins of the molecule's atoms, for instance carbon and hydrogen atoms in a chloroform ($CHCl_3$) molecule. Since the sample is very diluted (with typical dilution around 98%), the molecules are taken to be non-interacting, meaning that the system's Hamiltonian is simply the sum of each molecule's individual Hamiltonian over all $N$ molecules.

When in the presence of a magnetic field $\mathbf{B}$, spins will precess with a given frequency. We can tune the radio-frequency source to be in resonance with it and use it to control the quantum states. This interaction is given by the Hamiltonian

$$\mathcal{H} = -\boldsymbol{\mu} \cdot \mathbf{B},$$

where $\mu$ is the magnetic nuclear moment, given by $\mu = \gamma \mathbf{I}$. $\mathbf{I}$ is the nuclear angular momentum, and $\gamma$ is the gyromagnetic constant. We take the direction of $\mathbf{B}$ to be z-direction, $\mathbf{B} = B_0\mathbf{k}$

$$\mathcal{H} = -\gamma B_0 I_z,$$

$I_z$ is the projection of $\mathbf{I}$ in the z-direction, with eigenvalues $\hbar m_z$. The energy levels are given by

$$E = -\gamma \hbar B_0 m_z.$$

The energy gap between levels is given by $\Delta E = -\gamma \hbar B_0 \Delta m_z$, associated with an angular frequency through Planck's energy-frequency relation $\omega = 2\pi \Delta E/h = \Delta E/\hbar$. This is the Larmor frequency, to which we tune the source to induce the state transitions.

The spins are accessed with the aid of a coiled wire surrounding the sample, set in resonance with their specific Larmor frequency, different for the carbon and the hydrogen atom. Since the Larmor frequencies are different for each spin in a molecule but the same for equivalent atoms, we can tune the radio-frequency source and collectively target all spins with the same Larmor frequency only, namely all carbon spins or hydrogen spins.

To manipulate the sample nuclear spins, we first set the magnetic field $\mathbf{B}$ in the z-direction, causing the energy levels to split because of the Zeeman effect. Now, by applying a pulsed magnetic field $\mathbf{B_1}$ in a perpendicular direction, we can make the spins precess around $\mathbf{B_1}$, and the duration of the pulse determines the final angle of precession, given by $\theta = \gamma \mathbf{B_1} t$. To get information about the system, we evaluate the relaxation times $T_1$, referring to the amount of time it takes the spin to realign with $\mathbf{B}$ once the pulse is over, and $T_2$, referring to the loss of coherence among nuclei.

Measurements are performed through spectroscopy of the nuclei magnetic fields, aided by a set of coils that can measure very weak magnetic fields through induction, and one can obtain for instance the total magnetization of the sample in a certain direction. When carrying out this measurement, the precise position of a certain spin is irrelevant, and we don't have access to the state of an individual molecule.

The reason why measurements in NMR will play a significant role throughout this dissertation, serving often as a kindle for inspirations to our models, is that we are able to measure the total magnetization in the x-direction and the y-direction simultaneously. This is a trait we expect to see in classical measurements, but are typically inviable in quantum measurements due to incompatibility of the two observables being measured. This means that measurements carried out in NMR, as well as the states used in their samples, should be further investigated.

As we are going to see, this category of measurements implemented in NMR can be described using the method of types, consisting of an approach in which we can overlook the total knowledge of every aspect of a system and only care about a specific characterization of it. This concept has of course been around for a while, appearing in probability theory in works from Sanov[17] and Hoeffding[18] on large deviations, statistical physics from Boltzmann, as well as others [21] [22] [23]. It was systematically developed into a general method for information theory by Csiszár and Körner[20].

## 3.1 The Method of Types

Let $X = \chi_1\chi_2...\chi_N \in \chi^N$ be a string of N letters drawn for a $d$-letter alphabet $\chi = \{\alpha_1, \alpha_2, ..., \alpha_d\}$. The type of $X$ is a vector of positive numbers summing to one defined by

$$\mathbf{L}(X) = (L_1(X), L_2(X), ...L_d(X)),$$

where $L_j(X)$ is the relative frequency of the letter $\alpha_j$ in the string $X$, the number of occurrences of $\alpha_j$ divided by the total number of entries on the $X$:

$$L_j(X) = \frac{1}{N}\sum_{k=1}^{N}\delta_{\alpha_j\chi_k}.$$

Furthermore, one can see that

$$\sum_{j=1}^{d} L_j(X) = 1.$$

For example, if the alphabet is $\chi = \{0, 1\}$ and $N = 3$, we can take the string $X = 001$ and obtain its type

$$\mathbf{L}(X) = (\frac{2}{3}, \frac{1}{3}).$$

It's clear that the same type may apply to different strings, all of them being the result of a permutation of its elements. We can define a type class $T$ as being the set of strings that have the same type:

$$T[\mathbf{L}] = \{X \in \chi^N | \mathbf{L}(X) = \mathbf{L}\}.$$

In the preceding example, the type class of $\mathbf{L} = (\frac{2}{3}, \frac{1}{3})$ would be

$$T[\mathbf{L}] = \{001, 010, 100\}.$$

We point out that in the case of two-letter alphabets, a single number $L \in [0, 1]$ is enough to indicate the type of a string, since in these cases $\mathbf{L} = (1 - L, L)$.

We can now use this formalism to construct the measurement operator that describes a type measurement. We define the element of POVM corresponding to the projector on a subspace of a given type as:

$$Q_{\mathbf{L}}^N = \sum_{X \in T[\mathbf{L}]} |X\rangle \langle X| .$$

It is simply the sum of all projectors belonging to the same type.

### 3.1.1 An example: total magnetization

In NMR, when determining the total magnetization of a spin-$\frac{1}{2}$ sample along a direction we are measuring the observable

$$M_s^N = \sum_{i=1}^{N} \boldsymbol{\sigma}_s^{(i)}, \tag{3.1}$$

where $\boldsymbol{\sigma}_s$ is the Pauli vector in this direction, whose Cartesian components are the Pauli matrices $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_y = \begin{pmatrix} 0 & -i \\ 1 & 0 \end{pmatrix}$, and $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. The index $(i)$ in $\boldsymbol{\sigma}_s^{(i)}$ means that the observable acts

on the i-th system and doesn't affect the other systems. Notice that, since eigenvalues are irrelevant to entropic uncertainties, we are neglecting a factor multiplying equation 3.1 for simplicity.

For clarity, take the example of total magnetization along the z-direction. The observable, then, is

$$M_z^N = \sum_{i=1}^{N} \sigma_z^{(i)}.$$

To measure this observable, it seems necessary to know exactly the spin configuration of every local system, which is a very strict demand. Moreover, it is not a realist requirement, since in these measurements the only relevant quantity is the total magnetization and we do not even have access to each individual spin. If we use type observables instead, we can still evaluate the total magnetization of the sample, without being mandatory to know all this microscopic information.

In addition to these motivations, there is another argument against the usage of total magnetization as in equation 3.1. As pointed out by Poulin[9], we can describe the normalized macroscopic observables of total magnetization:

$$\tilde{M}_x^N = \frac{1}{N} M_x^N,$$
$$\tilde{M}_z^N = \frac{1}{N} M_z^N,$$

If we evaluate their commutator, we get

$$\left[\tilde{M}_x^N, \tilde{M}_z^N\right] = \frac{1}{N^2}\left[M_x^N, M_z^N\right] = \frac{-2i}{N} M_y^N$$

This means that on the limit of infinite size systems, total magnetization in the x and z directions commute, as we would expect from "classical" measurements. From the point of view of Heisenberg-like uncertainty principles, this would be enough to say a system could be prepared in a way to display both these quantities simultaneously well defined.

If we investigate these quantities from the perspective of entropic uncertainty relations, however, we see that their description still doesn't suffice to understand how classical features can emerge. Let us analyze these observables using the preparation uncertainty relation in equation 2.3.

Let $M_x^N = \sum_{j=1}^{N} \frac{\hbar}{2}\sigma_x^{(j)}$ be the total magnetization along the x-direction. We want to evaluate how

$$H(M_x^N|\Psi) + H(M_z^N|\Psi) \geq -2\log c$$

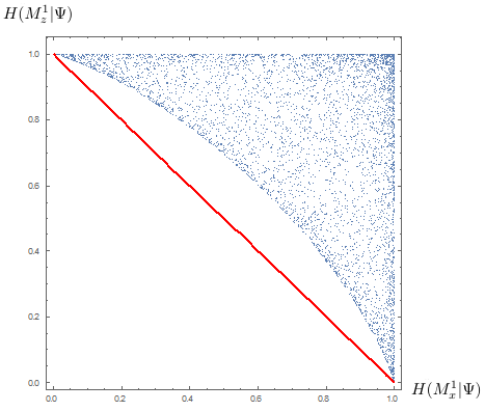
scales with $N$, since we would expect that as $N$ grows we would reach a classical scenario and this lower bound would approach zero. The maximum overlap between eigenvectors of $M_x^N$ and $M_z^N$ is $\frac{1}{\sqrt{2^N}}$, so that we have

$$H(M_x^N|\Psi) + H(M_z^N|\Psi) \geq N\log 2. \tag{3.2}$$

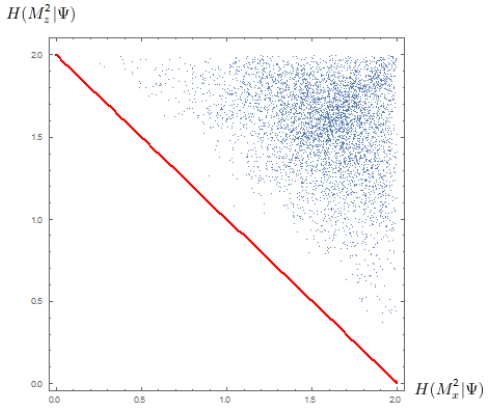
We see from this inequality that total magnetization as it was modeled above fails to yield the expected entropic uncertainty relation bounds. We can look at the distribution of random states on a $H(M_x^N|\Psi) \times H(M_z^N|\Psi)$ graphic, and see that indeed the bound given 2.3 does not decrease with $N$:

$$H(M_x^N|\Psi) + H(M_z^N|\Psi) \geq N\log 2. \tag{3.3}$$

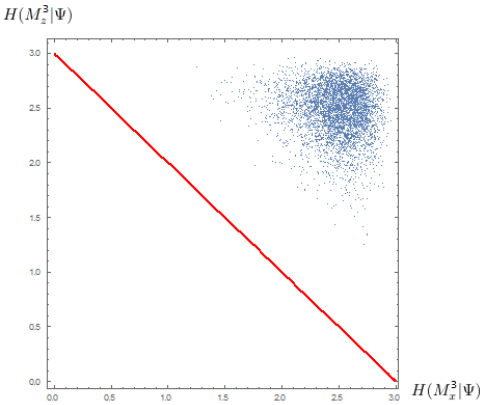
We see from this inequality that total magnetization as it was modeled above fails to yield the expected entropic uncertainty relation bounds. We can look at the distribution of random states on a $H(M_x^N|\Psi) \times H(M_z^N|\Psi)$ graphic, and see that indeed the bound given 2.3 does not decrease with $N$:



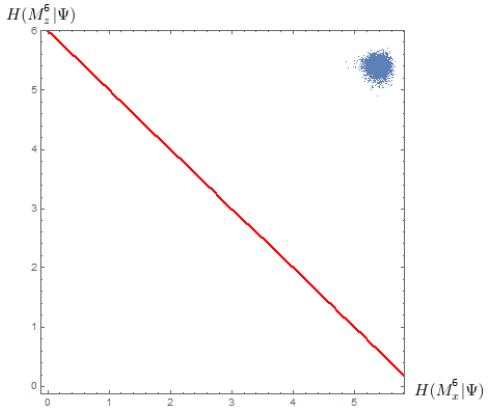

(a) Each point corresponds to a random pure state $\Psi$ whose calculated entropies $H(M_x^N|\Psi)$ and $H(M_z^N|\Psi)$ are given in the abscissa and ordinate, respectively. In red, we have the bound given in equation 3.3. This is the case $N = 1$.




(b) Similarly for $N = 2$.




(c) Similarly for $N = 3$.




(d) Similarly for $N = 6$.

Figure 3.1: Numerical results for the distribution of random pure states

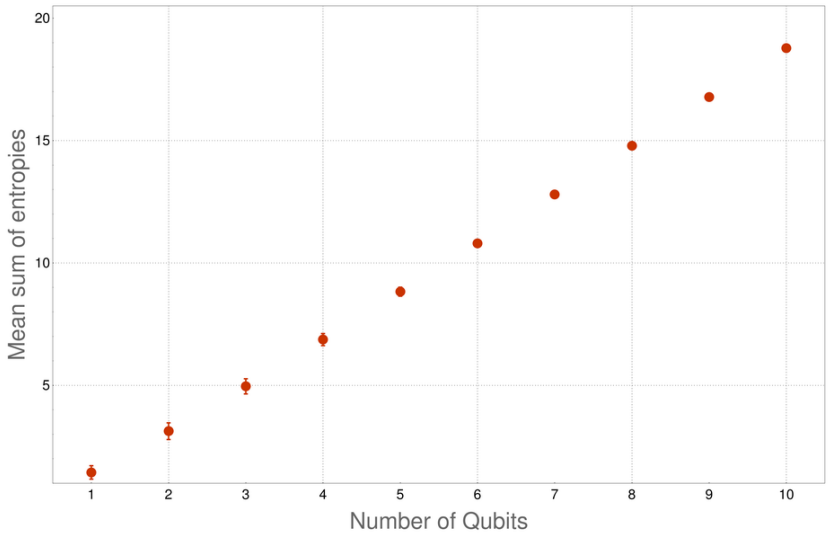



Figure 3.2: Average $H(M_x^N|\Psi) + H(M_z^N|\Psi)$, plotted as a function of $N$. We see here that, numerically, this sum grows with $N$, against our expectations for the emergence of classical features (but in accordance with equation 3.3).

Because of this and the motivation presented previously in this subsection, we will incorporate type measurements in our model.

Take the alphabet $\chi = \{0, 1\}$. The states $|0\rangle$ and $|1\rangle$ correspond to the eigenvectors of $\sigma_z$. The POVM elements representing the total magnetization on the z-direction will be projectors on subspaces of the same total magnetization. In this situation, we can work with the scalar $L$ rather than the vector $\mathbf{L}$. In the case for $N = 4$ we'll have

$$Z_L^4 = \sum_{X \in T[L]} |X\rangle \langle X| .$$

The first element of POVM will cover the type $L = 0$, or no 1s in the string:

$$Z_0^4 = |0000\rangle \langle 0000| .$$

The second element of POVM will cover the type $L = 1/4$, or one 1 out of four entries in the string:

$$Z_{1/4}^4 = |0001\rangle \langle 0001| + |0010\rangle \langle 0010| + |0100\rangle \langle 0100| + |1000\rangle \langle 1000| .$$

The third element of POVM will cover the type $L = 1/2$, or two 1s out of four entries in the string:

$$Z_{1/2}^4 = |0011\rangle \langle 0011| + |0101\rangle \langle 0101| + |0110\rangle \langle 0110| + |1001\rangle \langle 1001| + |1010\rangle \langle 1010| + |1100\rangle \langle 1100| .$$

The forth element of POVM will cover the type $L = 3/4$, or three 1s out of four entries in the string:

$$Z_{3/4}^4 = |0111\rangle \langle 0111| + |1011\rangle \langle 1011| + |1101\rangle \langle 1101| + |1110\rangle \langle 1110| .$$

And finally, the fifth element of POVM will cover the type $L = 1$, where all entries are 1s:

$$Z_1^4 = |1111\rangle \langle 1111| .$$

This defines the POVM of total magnetization in the z-direction for 4 qubits:

$$Z^4 = \{Z_0^4, Z_{1/4}^4, Z_{1/2}^4, Z_{3/4}^4, Z_1^4\}$$

In other words, we are putting together the projectors on subspaces that yield the same total magnetization in the z-direction. As a result, we go from $M_z^N$ having $2^N$ outcomes to $Z^N$ having $N + 1$ outcomes.

It goes in a completely analogous way for the x-direction. Using the base of eigenvectors of the spin operator in the x-direction,

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}},$$
$$|-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}},$$

we can follow the construction process in a completely analogous way and get

$$X_0^4 = |++++\rangle\langle++++| ,$$

$$X_{1/4}^4 = |+++-\rangle\langle+++-| \ + \ |++-+\rangle\langle++-+| \ + \ |+--++\rangle\langle+--++| \ + \ |-+++\rangle\langle-+++| ,$$

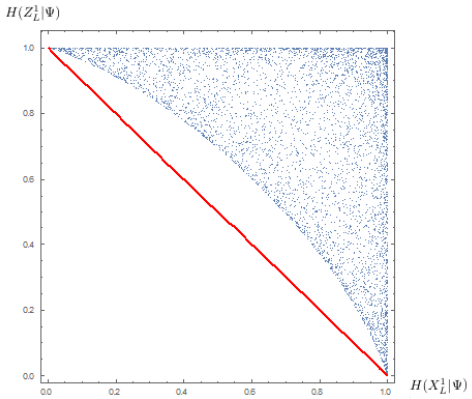
$$X_{1/2}^4 = |++--\rangle\langle++--| \ + \ |+-+-\rangle\langle+-+-| \ + \ |+--+\rangle\langle+--+| + \ |-++-\rangle\langle-++-| \ + \ |-+-+\rangle\langle-+-+| \ + \ |--++\rangle\langle--++| \ ,$$

$$X_{3/4}^4 = |+---\rangle\langle+---| \ + \ |-+--\rangle\langle-+--| \ + \ |--+-\rangle\langle--+-| \ + \ |---+\rangle\langle---+| ,$$

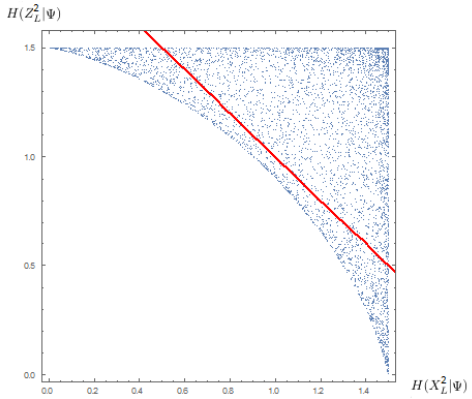
$$X_1^4 = |----\rangle\langle----| ,$$

$$X^4 = \{X_0^4, X_{1/4}^4, X_{1/2}^4, X_{3/4}^4, X_1^4\}.$$

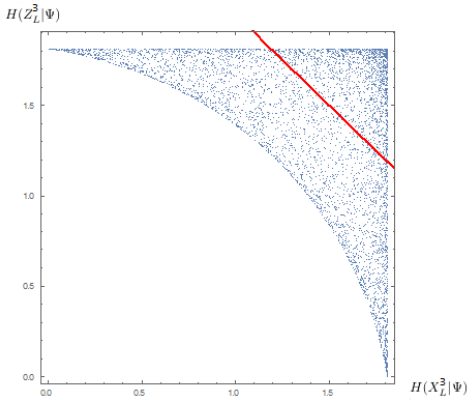
We can now revisit the graphics shown previously in the beginning of this section, with the adjustment of using type measurements to describe total magnetization. We see clearly that the sum $H(X_L^N|\Psi) + H(Z_L^N|\Psi)$ often goes below the bound provided by equation 3.3, meaning that using type measurements may be an improvement on our task. However, the average sum of entropies still grows with $N$, as seen in figure 3.4.



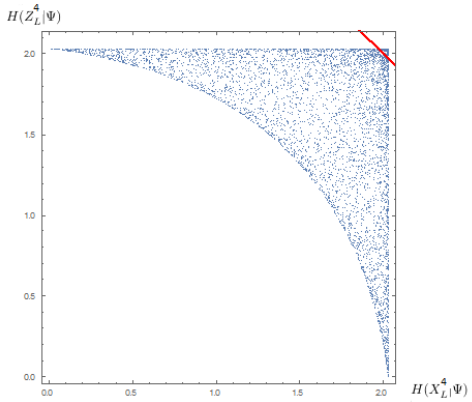

(a) Each point corresponds to a random pure state $\Psi$ whose calculated entropies $H(X_L^N|\Psi)$ and $H(Z_L^N|\Psi)$ are given in the abscissa and ordinate, respectively. In red, we have the bound given by equation 3.3. This is the case $N = 1$.




(b) Similarly for $N = 2$.




(c) Similarly for $N = 3$.




(d) Similarly for $N = 4$.

Figure 3.3: Numerical results for the distribution of random pure states using type measurements

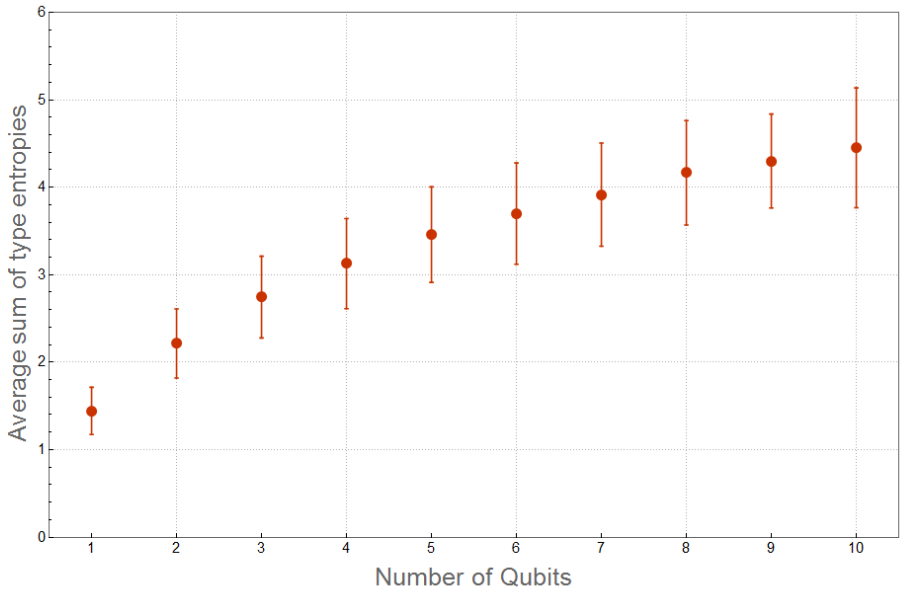



Figure 3.4: Average $H(M_x^N|\Psi) + H(M_z^N|\Psi)$, plotted as a function of $N$. We see here that, numerically, this sum still grows with $N$, although at a smaller rate. The error bars here are given by standard deviation of the sample

Having studied the method of types, we may now go on and make further suppositions and refinements to our model.

## 3.2 Product States

In NMR, the molecules are so diluted in water and are so far apart from each other that their interactions can be over-sighted. Furthermore, all molecules are identical, so every state should be the same on a local level. Because we start with a thermal state, and since the interactions between the molecules in the sample can be neglected to good approximation, we say our initial state is a product of locally identical states. We take this inspiration from NMR, and attempt to comprehend these peculiarities and come up with a model for macroscopic measurements. We will intentionally attain ourselves to the limited case of product pure states, which are such as

$$|\Psi^N\rangle = \underbrace{|\Psi\rangle \otimes |\Psi\rangle \otimes |\Psi\rangle \ldots |\Psi\rangle}_{\text{N times}},$$

for a pure single qubit state $|\Psi\rangle \in \mathbb{C}^2$.

We want to derive the probability distribution of the POVM elements $Z_L^N$ on $\Psi^N = |\Psi^N\rangle\langle\Psi^N|$. This means

$$Pr(Z_L^N|\Psi^N) = Tr\left[\Psi^N Z_L^N\right].$$

We start by parameterizing $|\Psi\rangle$ as

$$|\Psi\rangle = \sqrt{p}\,|0\rangle + \sqrt{1-p}e^{i\phi}\,|1\rangle\,,$$

with $p \in [0,1]$ and $\phi \in [0, 2\pi[$, so that $Pr(Z_L^N \mid \Psi^N)$ equals

$$= Tr\left[\underbrace{\left(\sqrt{p}\,|0\rangle + \sqrt{1-p}e^{i\phi}\,|1\rangle\right)\ldots\left(\sqrt{p}\,|0\rangle + \sqrt{1-p}e^{i\phi}\,|1\rangle\right)}_{\text{N times}}\underbrace{\left(\sqrt{p}\,\langle 0| + \sqrt{1-p}e^{i\phi}\,\langle 1|\right)\ldots\left(\sqrt{p}\,\langle 0| + \sqrt{1-p}e^{i\phi}\,\langle 1|\right)}_{\text{N times}} Z_L^N\right].$$

Now, $Z_L^N$ will select only terms corresponding to states of the same type. Within a string with type $L$, the number of 1 entries is $NL$ (the relative frequency of 1 multiplied by the size of the string). On

the other hand, the number of 0 entries is $N - NL$ (the total number of entries minus the number of 1s). Each state $|1\rangle$ is multiplied by $\sqrt{1-p}e^{i\phi}$, while each state $|0\rangle$ is multiplied by $\sqrt{p}$. Besides that, each type $L$ will have as many states as the number of ways we can arrange $N$ states to have $NL$ $|1\rangle$ entries. More concretely,

$$Pr(Z_L^N \mid \Psi^N) = Tr\left[\left(p^N \left|000\ldots0\right\rangle\left\langle000\ldots0\right| + p^{N-1}(1-p)^1 \underbrace{\left(\left|000\ldots01\right\rangle\left\langle000\ldots01\right| + \left|000\ldots10\right\rangle\left\langle000\ldots10\right|\right)}_{\text{states of type 1/N}} + \right.\right.$$
$$p^{N-2}(1-p)^2 \underbrace{\left(\left|000\ldots11\right\rangle\left\langle000\ldots11\right| + \left|000\ldots110\right\rangle\left\langle000\ldots110\right|\right)}_{\text{states of type 2/N}} + \cdots +$$
$$\left.\left.(1-p)^{NL}\left|111\ldots1\right\rangle\left\langle111\ldots1\right|\right) Z_L^N\right]. \quad (3.4)$$

With that, we arrive at the general formula for these probabilities

$$Pr(Z_L^N \mid \Psi^N) = \binom{N}{NL} p^{N-NL}(1-p)^{NL}.$$

This is a binomial distribution. It has mean $\mu = (1-p)N$ and standard deviation $\Delta^2 = Np(1-p)$. If we explicitly calculate the entropy associated with this probability distribution.

$$H(Z_L^N|\Psi^N) = -\sum_{L=0}^{1} \binom{N}{NL} p^{N-NL}(1-p)^{NL} \log \binom{N}{NL} p^{N-NL}(1-p)^{NL}.$$

If $N \to \infty$, we can use the de Moivre/Laplace theorem and approximate the binomial distribution with a Gaussian distribution:

$$H(Z_L^N|\Psi^N) \approxeq -\int_{-\infty}^{+\infty} dx \frac{1}{\sqrt{2\pi\Delta^2}} \exp\left[-\frac{(x-\mu)^2}{2\Delta^2}\right] \log\{\frac{1}{\sqrt{2\pi\Delta^2}} \exp\left[-\frac{(x-\mu)^2}{2\Delta^2}\right]\}$$
$$= -\int_{-\infty}^{+\infty} dx \frac{1}{\sqrt{2\pi\Delta^2}} \exp\left[-\frac{(x-\mu)^2}{2\Delta^2}\right] \{\log\sqrt{2\pi\Delta^2} + \frac{(x-\mu)^2}{2\Delta^2}\log e\}$$
$$= \log\sqrt{2\pi\Delta^2} + \frac{\log e}{2}$$
$$H(Z_L^N|\Psi^N) \approx \frac{1}{2}\log 2\pi e Np(1-p) \quad (3.5)$$

and similarly for $H(X_L^N|\Psi^N)$.

We can evaluate $H(X_L^N|\Psi^N) + H(Z_L^N|\Psi^N)$ numerically, randomly picking states like $\Psi^N$ and calculating their average, minimum and maximum sum of entropies.

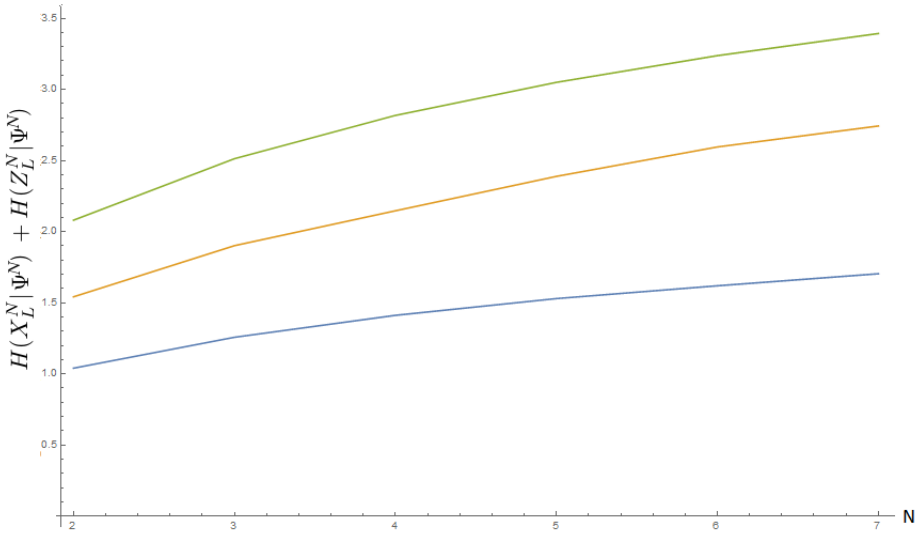



Figure 3.5: Here we have $H(X_L^N|\Psi^N) + H(Z_L^N|\Psi^N)$ plotted as a function of $N$. For each $N$, we randomly picked between 10000 and 1000 $\Psi^N$ states and calculated this sum of entropies. In green we have the maximum sum of entropies out of all trials for a given $N$. Similarly, in orange we have the average sum of entropies for all trials, and in blue we have the minimum sum of entropies out of all trials.

Even though it is only valid for large $N$, these numerical results can be recovered to some precision by equation 3.5, if we calculate their average over all possible states $\Psi^N$. The numerical and analytical results are so close that plotting them together in figure 3.5 would be purposeless. Against what we would expect from a classical description of a measurement, we see from both this graphic and equation 3.5 that the sum of entropies actually still grows with $N$, though at the lower rate of $\log N$. This is a hint that the method of types alone may not be enough to portray the macroscopic world.

Now that we have discussed all the theoretical requirements to develop this dissertation, we can move on to present and discuss our results

# Chapter 4

# Division into Bins: Allowing for Imprecision

In actual measurements, no device is infinitely precise. After many runs of an experiment, even a classical measurement will yield results with a certain spread. Whatever the source of the imprecision might be, we should be able to incorporate it on our model. In the example of total magnetization, for instance, it is not realistic to expect the apparatus to distinguish states with similar types. For the sake of verisimilitude, we would like to integrate the notion of imprecision in our description of macroscopic measurements.

Aside from making type measurements, a possible reasonable solution is to arrange neighboring types in bins. Instead of determining whether a state has a certain probability of belonging to a type $L$, we evaluate their probability of belonging to a type interval, grouping them into bins of size $\Delta$. This way, in place of calculating the probability distribution by types, we do so in a distribution by bins. Since the type $L$ ranges from 0 to 1, the number of bins $N_b$ is related to their size $\Delta$ by

$$N_b = 1/\Delta.$$

The first bin will cover the interval $[0, \Delta[$. The second bin will cover $[\Delta, 2\Delta[$, and so on, until the last bin covering $[1-\Delta, 1]$. The n-th bin will cover the interval $[(n-1)\Delta, n\Delta[$. The new POVM elements will now be, for the n-th bin

$$Q'^{N}_{n} = \sum_{L \in [(n-1)\Delta, n\Delta[} Q^{N}_{L},$$

with $n \in \{1, \ldots, N_b\}$.

The probability associated with the n-th bin is the sum of all probabilities related to types that lie within that bin, in terms of the number of bins, is

$$Pr(Z'^{N}_{n}|\Psi^{N}) = \sum_{L \in [(n-1)/N_b, n/N_b[}^{N_b} \binom{N}{NL} p^{N-NL} (1-p)^{NL}.$$

For instance, if we take $p = 0.8$, $\phi = 0$ and $N = 8$, the probability distribution of total magnetization by types in the z-direction of $\Psi^N$ is

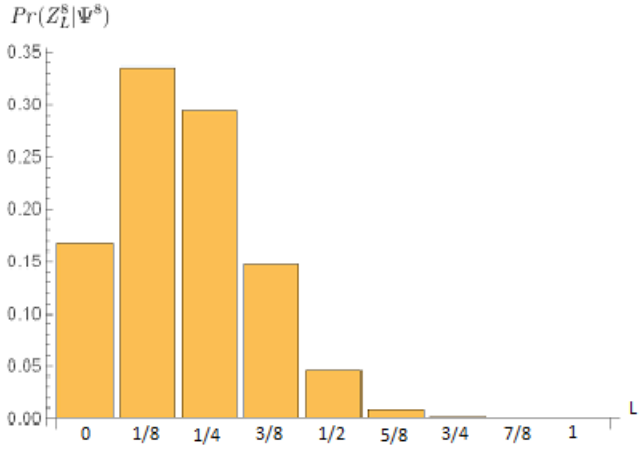



Figure 4.1: The probability distribution $Pr(Z_L^8|\Psi^8)$ by types.

If we choose to group the types in 4 bins, we get the probability distribution

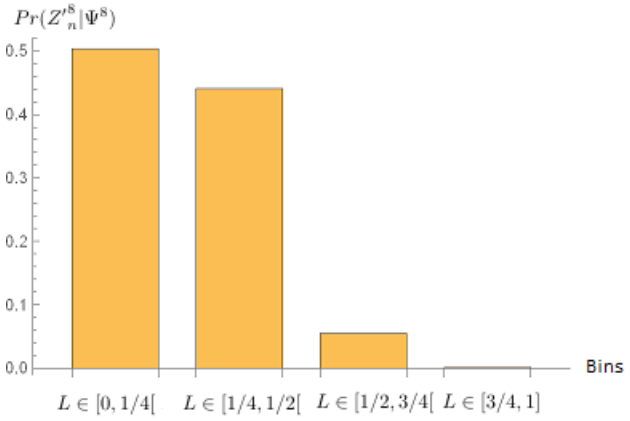



Figure 4.2: The probability distribution $Pr(Z'^8_L|\Psi^8)$ by bins. Since $N_b = 4$, each bin will cover $L \in [(n-1)/4, n/4[$

We can now work on a simple example to make this process clearer.

## 4.1 An example: imprecision in total magnetization

Take the POVM of total magnetization along the z-direction previously constructed in section 3.1.1. We first divide them in bins, following chapter 4. We chose the number of bins to be $N_b = 3$. The bins POVMs will be

$$Z'^4_n = \sum_{L \in [(n-1)/3, n/3[} Z^4_L.$$

The first bin will contain types ranging from 0 to 1/3:

$$Z'^4_1 = Z^4_0 + Z^4_{1/4} = |0000\rangle\langle 0000| + |0001\rangle\langle 0001| + |0010\rangle\langle 0010| + |0100\rangle\langle 0100| + |1000\rangle\langle 1000|.$$

The second bin will contain types ranging from 1/3 to 2/3:

$$Z'^4_2 = Z^4_{1/2} = |0011\rangle\langle 0011| + |0101\rangle\langle 0101| + |0110\rangle\langle 0110| + |1001\rangle\langle 1001| + |1010\rangle\langle 1010| + |1100\rangle\langle 1100|.$$

The third bin will contain types ranging from 2/3 to 1:

$$Z'^4_3 = Z^4_{3/4} + Z^4_1 = |0111\rangle\langle 0111| + |1011\rangle\langle 1011| + |1101\rangle\langle 1101| + |1110\rangle\langle 1110| + |1111\rangle\langle 1111|\,.$$

This defines the POVMs divided in bins $Z'^4 = \{Z'^4_1, Z'^4_2, Z'^4_3\}$.

## 4.2 Results

### 4.2.1 Preparation

In this stage of our research, we want to check if type measurements associated with division into bins can provide any insight on understanding the emergence of classical features. We also have some motivation to work with product states, so all the work developed from now on will be dealing with states like $\Psi^N$, unless stated otherwise.

We know that, as $N$ grows, the probability distribution of $\Psi^N$ tends to concentrate around its mean value. As a consequence, when we use divisions into bins, the bin containing the mean value will tend to be increasingly probable. In this case, when calculating $H(Q'^N_n|\Psi^N)$, the only relevant term will be that whose probability is not increasingly small, namely the term corresponding to the bin containing the mean value $n_\mu$. Eventually, we expect $H(Q'^N_n|\Psi^N)$ to coincide with $H_\infty(Q'^N_n|\Psi^N)$, since that is the case where we only take the maximum probability into account.

Moreover, since the probability associated with $n_\mu$ will approach 1, we also expect both entropies $H(Q'^N_n|\Psi^N)$ and $H_\infty(Q'^N_n|\Psi^N)$ to decrease as $N$ grows. We want to analyze the specific case of total magnetization along directions x and z. For the z-direction, we have $Q'^N_n = Z'^N_n$. Since for $Z^N_L$ $\mu = N(1-p)$, the bin containing the mean value will be $n_\mu = 1 + \lfloor N(1-p)\rfloor$, meaning we take the floor of $\mu$ and add 1, because the bin numbering starts at 1. The same reasoning goes for $Q'^N_n = X'^N_n$.

We have evaluated and compared $H(X'^N_n|\Psi^N) + H(Z'^N_n|\Psi^N)$ and $H_\infty(X'^N_n|\Psi^N) + H_\infty(Z'^N_n|\Psi^N)$ as $N$ grows for different states $\Psi^N$ and for various numbers of bins $N_b$.

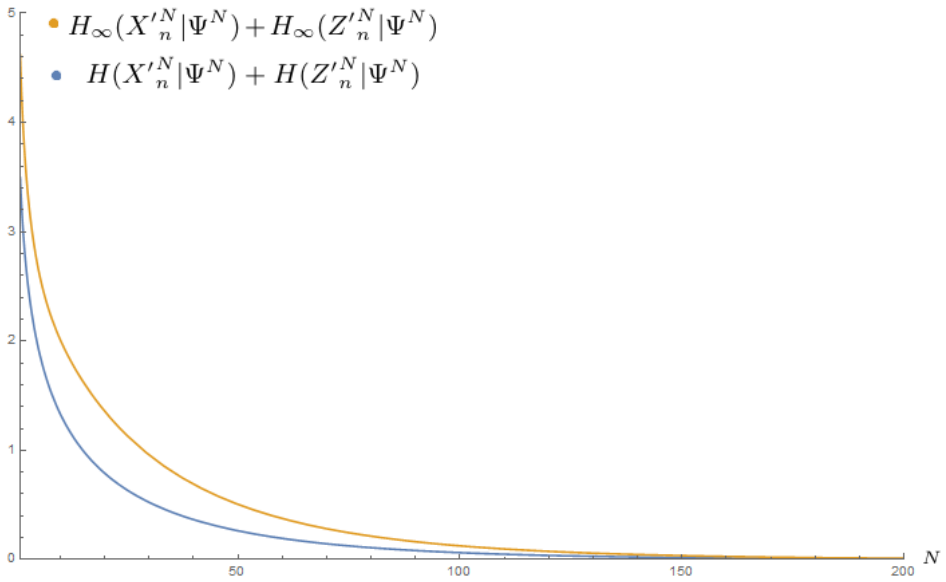



Figure 4.3: In orange, we plotted $H_\infty(X'^N_n|\Psi^N) + H_\infty(Z'^N_n|\Psi^N)$, while in blue we have $H(X'^N_n|\Psi^N) + H(Z'^N_n|\Psi^N)$, both as a function of $N$. In this case, $p = 0.29$, $\phi = 0$, and $N_b = 5$. We see here a particular case where both sums of entropies tend to zero.

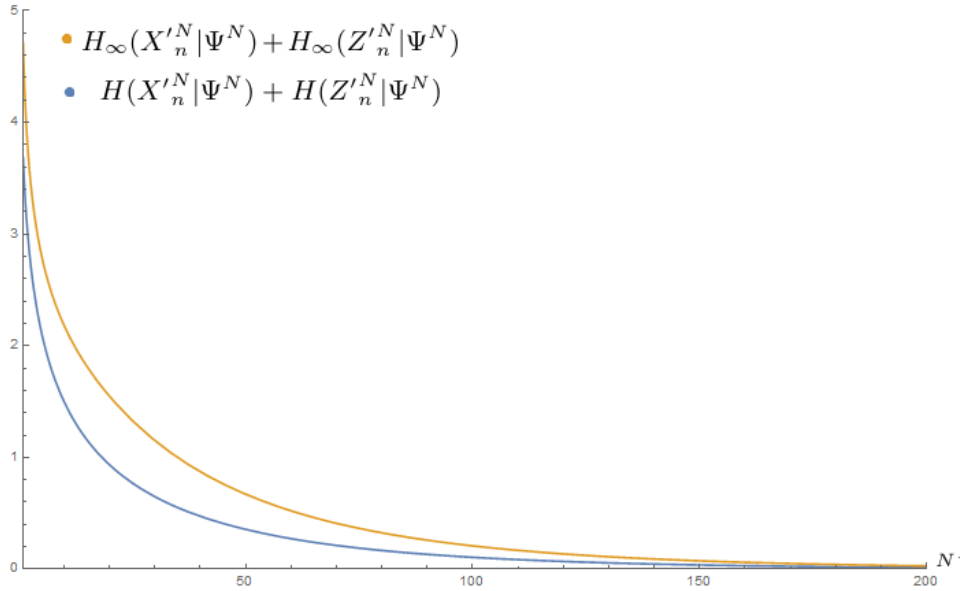



Figure 4.4: Similarly for $p = 0.512$, $\phi = 0$, and $N_b = 5$. Here is another particular case where both sums of entropies tend to zero.

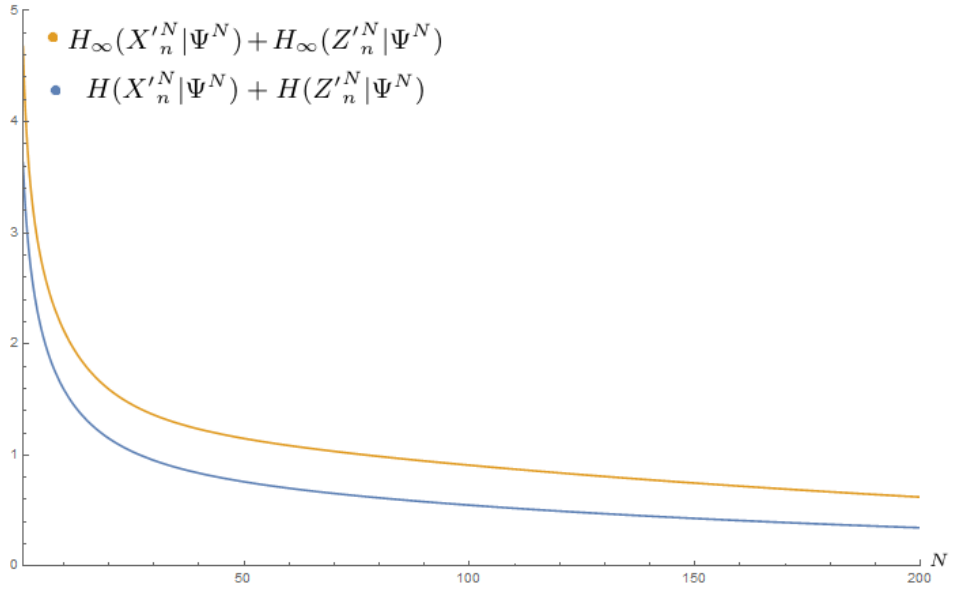



Figure 4.5: Similarly for $p = 0.634$, $\phi = 0$, and $N_b = 5$. In this case, even though the number of bins is the same as before, we can no longer say the sums of entropies tend to zero, at least within $N \leq 200$.

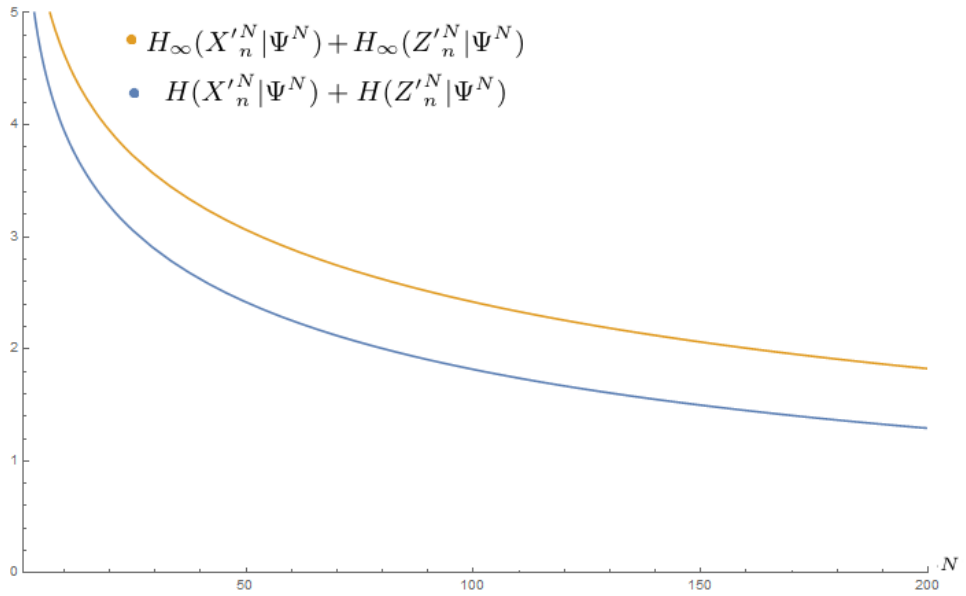



Figure 4.6: Similarly for $p = 0.512$, $\phi = 0$, and $N_b = 18$. Even though we have other situations where the sums of entropies tend to zero with $p = 0.512$, this is not the case, at least within $N \leq 200$.

We see from these numerical results that, for some particular cases, type measurements combined with division into bins allow the two properties total magnetization in directions x and z to be simultaneously well defined, while not being possible at first. As we already stated, the fact that these entropies become increasingly low as $N$ grows is a necessary but not sufficient condition for the preparation of a system with total magnetizations in the x and z directions simultaneously defined. This result is, thus, very particular and has no general value, but it still is an evidence that this may be a promising path to follow.

What is left now is to check if this arrangement is enough to also allow (sequential) compatibility.

### 4.2.2 Measurement

We want now to find out if using this model we can also explain measurement compatibility of otherwise incompatible observables. We know so far that some particular states can specifically have the quantities $X'^N_n$ and $Z'^N_n$ simultaneously well defined, but in order to try to clarify the emergence of classical features we must also be able to improve our understanding of measurement incompatibility.

According to the discussion in section 2.3.1, we must compare the two probabilities yielded by two different sets of experiments, one establishing the statistics relative to measuring an ideal observable $B$, and another one estimating the probability distribution of $B'$ after measuring $A$. We carried out these comparisons for two different distributions: the one yielded by an ideal measurement of $Z'^N_n$ (to which we will refer as $p_z$), and another one yielded by a measurement of $X'^N_n$ followed by a measurement of $Z'^N_n$ (referred as $p'_z$). If there is a number of bins where $X'^N_n$ doesn't disturb $Z'^N_n$, we expect these probabilities to be very close.

That is not, however, the result obtained from numerical analysis. The Euclidean distance between $p_z$ and $p'_z$ does not seem to follow any trend, and is not particularly low at any point, considering that the maximum possible distance between $p_z$ and $p'_z$ is 2.

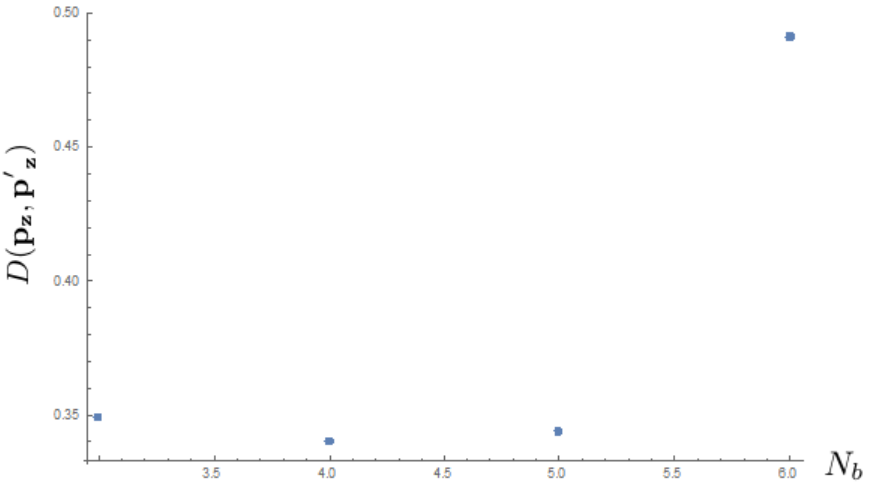



Figure 4.7: The Euclidean distance between $p_z$ and $p'_z$, for $N = 6$, as a function of the number of bins. We are analyzing the particular state $\Psi^N$ with $p = 0.8$ and $\phi = 0$. We see here that the two probabilities do not seem to become closer in any case.

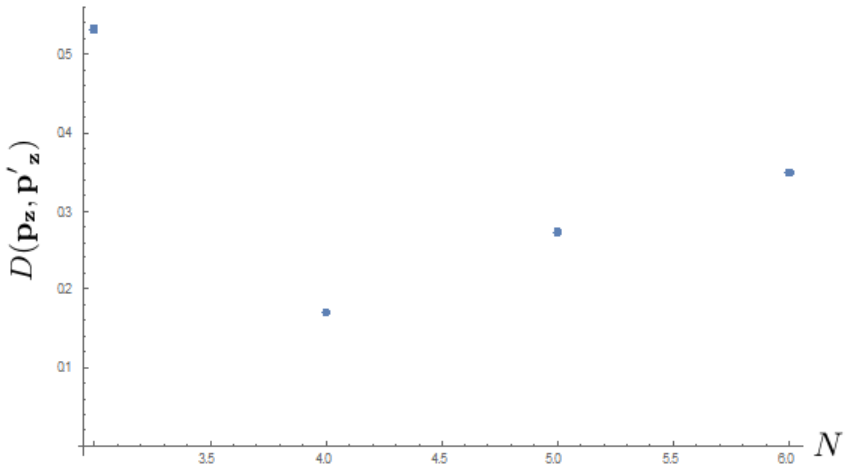



Figure 4.8: The Euclidean distance between $p_z$ and $p'_z$, now as a function of $N$, for a given number of bins (in this case, $N_b = 3$). Once again, we are analyzing the particular state $\Psi^N$ with $p = 0.8$ and $\phi = 0$. As before, the probabilities do not seem to become any closer.

From these numerical results, we do not see any pattern that can lead us to conclude that there will be a special case where $p_z$ and $p'_z$ are very close to each other. This is in accordance with [7], chapter III, part 3, section (6:a), which concludes that non-commuting observables cannot yield probabilities such that A and B are compatible for an arbitrary state. Since ${X'}_n^N$ and ${Z'}_n^N$ still do not commute, we have no reason to expect $p_z$ and $p'_z$ to be the same, even though the proof carried out in [7] does not make the particular assumption of product states as we do. We haven't, however, checked any cases where $N >> N_b$ due to the computational burden it can be. There is still a chance that these cases can yield a decreasing pattern for $D(\mathbf{p_z}, \mathbf{p'_z})$.

Division into bins, combined with type measurements, can provide some insights on preparation of states with well defined properties that were otherwise impossible, but cannot improve our understanding of measurement of incompatible observables. To investigate further, we must add new features to our model.

# Chapter 5

# Mixing Bins: Allowing for Inaccuracy

So far, guided by the notion that one cannot distinguish between nearby types, we have decided to group types by bins. This is the reasonable assumption that we are only capable of limited precision. However, when look closer at the types that lie at the edge of a bin we see that our model is actually infinitely precise in that region. We may not be able to tell the difference between types that happened to be grouped together under the same bin, but when comparing the last type of a certain bin and the first type on the following bin we seem to be able to perfectly tell them apart.

A way to fix this infinite precision on the borders is to "blend" that region, blur the limits between bins. By doing so, we include the notion of inaccuracy in our model, incorporating the possibility of errors, so that a state with a well defined type has a chance of being assigned to a wrong bin, since in the lab no measurement can be perfectly accurate.

Thus, our main goal in this section is to devise a method of mixing bins, in a way that we are able to model the probability of committing mistakes when measuring the state's type, while assimilating inaccuracy in our description of macroscopic measurements.

We want to take the POVM elements ${Q'}_n^N$ that have been established on the previous section and mix them in a way that allows for inaccuracy. To do that, we want to construct matrices that will act on ${Q'}_n^N$ and yield the mixed bins POVM elements $\tilde{Q}_n^N$

$$\begin{pmatrix} \tilde{Q}_1^N \\ \tilde{Q}_2^N \\ \vdots \\ \tilde{Q}_{N_b}^N \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1N_b} \\ a_{21} & & \ddots & \\ \vdots & & & \\ a_{N_b 1} & & & \end{pmatrix} \begin{pmatrix} {Q'}_1^N \\ {Q'}_2^N \\ \vdots \\ {Q'}_{N_b}^N \end{pmatrix}. \tag{5.1}$$

Since we want $\{\tilde{Q}_n^N\}$ to be a POVM, we must require that $a_{ij} \geq 0$ and $\sum_{j=1}^{N_b} a_{ij} = 1$ for all $i$. This way, each POVM element $\tilde{Q}_n^N$ will be a convex sum of ${Q'}_n^N$. Also, in order to obtain $\sum_{n=1}^{N_b} \tilde{Q}_n^N = \mathbb{1}$ we must have $\sum_{i=1}^{N_b} a_{ij} = 1$ for all $j$. In other words, the matrix we want to construct must be doubly stochastic. To do so, we first look at Birkhoff's theorem[26]:

**Birkhoff's Theorem.** *Every doubly stochastic matrix is a convex combination of permutation matrices.*

We won't prove this theorem, but the intuition behind it is that permutation matrices are square matrices with exactly one entry 1 in each row and column and 0s elsewhere, themselves being doubly stochastic. Permutations are, then, the extremal points of the convex set of doubly stochastic matrices.

To build the matrix that will mix the bins according to our model, we adopt the following steps: we first construct $N-1$ matrices connecting nearest neighboring bins. They will be the convex combination of the identity matrix and the permutation matrix of elements $k$ and $k+1$, adjusted by a parameter $m \in [0,1]$. After having repeated this process for all neighboring pairs, we sum all of them and divide it by $N_b - 1$. Let's look into an example, for clarity.

## 5.1 An example: inaccuracy in total magnetization

As an example, we will apply our method of mixing bins on the observables obtained in 4.1, where we constructed the POVM total magnetization in the z direction divided by bins $Z'^4$ with $N = 4$ and $N_b = 3$. We start by building the matrices connecting neighboring bins. The first matrix, connecting the first and the second bins, is the convex sum of the identity matrix and the permutation matrix of elements 1 and 2, with $m \in [0, 1]$:

$$(1-m)\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} + m\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix} = \begin{pmatrix}1-m&m&0\\m&1-m&0\\0&0&1\end{pmatrix}.$$

The second matrix, connecting the second and the third bins, is the convex sum of the identity matrix and the permutation matrix of elements 2 and 3:

$$(1-m)\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} + m\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix} = \begin{pmatrix}1&0&0\\0&1-m&m\\0&m&1-m\end{pmatrix}.$$

We now sum them, and divide the result by $N_b - 1 = 2$ to make them evenly weighted

$$\frac{1}{2}\begin{pmatrix}1-m&m&0\\m&1-m&0\\0&0&1\end{pmatrix} + \frac{1}{2}\begin{pmatrix}1&0&0\\0&1-m&m\\0&m&1-m\end{pmatrix} = \frac{1}{2}\begin{pmatrix}2-m&m&0\\m&2(1-m)&m\\0&m&2-m\end{pmatrix}.$$

Finally, we multiply $\{Z'^4_n\}$ by this matrix

$$\begin{pmatrix}\tilde{Z}^4_1\\\tilde{Z}^4_2\\\tilde{Z}^4_3\end{pmatrix} = \frac{1}{2}\begin{pmatrix}2-m&m&0\\m&2(1-m)&m\\0&m&2-m\end{pmatrix}\begin{pmatrix}Z'^4_1\\Z'^4_2\\Z'^4_3\end{pmatrix} = \frac{1}{2}\begin{pmatrix}(2-m)Z'^4_1 + mZ'^4_2\\mZ'^4_1 + 2(1-m)Z'^4_2 + mZ'^4_3\\mZ'^4_2 + (2-m)Z'^4_3\end{pmatrix}.$$

The mixed bins POVMs we get are

$$\tilde{Z}^4_1 = \left((2-m)Z'^4_1 + mZ'^4_2\right)/2,$$

$$\tilde{Z}^4_2 = \left(mZ'^4_1 + 2(1-m)Z'^4_2 + mZ'^4_3\right)/2,$$

$$\tilde{Z}^4_3 = \left(mZ'^4_2 + (2-m)Z'^4_3\right)/2.$$

We see that $\tilde{Z}^4_n$ is a mix of other $Z'^4_j$, and we can use $m$ to control the contribution from other bins. This means that a state with a well defined type can be assigned to mostly any bin, and may have a greater chance of being designated to the bin containing their type (and its nearest neighbors) if $m$ is small enough.

This rather simple procedure reproduces one characteristic we expect from realistic measurements: you are more likely to mislabel a string for a type belonging to a neighboring bin than for a far away one, the closer the bins the easier they are to be mistaken one for another. Far from being ideal, this method is only a first approach on mixing bins and allowing inaccuracy on our model, specially because it lacks a more natural and intuitive justification, having been forcibly constructed instead of arising more naturally. A more organic method is developed in [9], though it poses additional operational complications.

## 5.2 Results

### 5.2.1 Preparation

We want now to ensure that, following the description of macroscopic measurements formulated in this chapter, our system can be prepared in a way to have two properties simultaneously well defined, even though it would not be possible in the typical scenario. Particularly, we want to check if this way we can prepare a state $\Psi^N$ and have the total magnetizations in the x and z directions well defined, when modeling them as type measurements grouped by bins that are mixed. As in section 4.2.1, we want to understand the behavior of the lower bound in the sum $H(\tilde{X}^N_n|\Psi^N) + H(\tilde{Z}^N_n|\Psi^N)$.

In section 4.2.1, we used the argument that probability distributions would end up concentrating on particular bins to justify the sum of entropies approaching zero as $N$ grows. However, the same argument cannot be employed here. Even though probability distributions would still concentrate, at first, on some bin, this concentration would be dissolved after we mixed them. A particular bin might have held a very high probability of containing the type to which a state would be assigned, but after mixing all bins this probability would decrease while the probabilities of neighboring bins would increase, causing probability distributions to spread. Because of this effect, we can no longer use this argument to sustain that $H({X'}_n^N|\Psi^N)+H({Z'}_n^N|\Psi^N)$ decreases with $N$. The only exception is, of course, the case where $m = 0$ and we recover the scenario described in section 4.2.1.

We no longer have a motivation to pursue the - often heavy - numerical results that would ensure that some macroscopic measurements as described in chapter 5 can be simultaneously well defined for $\Psi^N$. We point out, however, that this result is not unexpected. Even when performing classical measurements, there is a certain spread on the results that is inherent of statistical nature, since no measurement can ever be perfectly precise. By using the method devised in this chapter, we end up recovering this essential feature of classical measurements.

### 5.2.2 Measurement

As in section 4.2.2, we now want to check if the scenario constructed throughout this last chapter allow the compatibility of two POVMs who were otherwise incompatible. Because of all the reasoning previously developed in this dissertation, we will test sequential compatibility of POVMs $\tilde{X}_n^N$ and $\tilde{Z}_n^N$, the total magnetization along the directions x and z using type measurements grouped in bins, that were additionally mixed among themselves.

Following the scenario described by figure 2.3, we want to compare the statistics yielded by $B$ and $B'$. In our case, this corresponds to comparing the probabilities yielded from measuring $\tilde{Z}_n^N$ and those yielded from measuring $\tilde{Z}_n^N$ after having measured $\tilde{X}_n^N$ first.

Using the Euclidean distance in equation 2.4, as explained in section 2.3.1, we compare numerically how different $m$ affect the distance between probabilities. We want to evaluate how mixing bins in a greater or smaller amount can affect the statistics of the measurements.

Our expectation is that the probabilities relative to the ideal measurement $\tilde{Z}_n^N$ and that of $\tilde{Z}_n^N$ after $\tilde{X}_n^N$, to which we will refer as ${\tilde{Z}'}_n^N$ will become more similar the more we allow inaccuracy in our system, namely the more $m$ grows. We must point out, however, that seeing the probabilities of $\tilde{Z}_n^N$ and ${\tilde{Z}'}_n^N$ getting increasingly closer as $m$ grows in no way proves that our model can explain compatibility between otherwise incompatible POVMs. First of all, because our case study is too particular in many ways: we are testing the compatibility of two particular POVMs $\tilde{X}_n^N$ and $\tilde{Z}_n^N$, and analyzing the case of some specific product state. Second, because having the Euclidean distance between these probabilities decrease is only an aspect of sequential compatibility: it is a necessary but not sufficient condition. What we present here is only some evidence and support in the sense that our model may be able to explain the phenomena we are interested in.

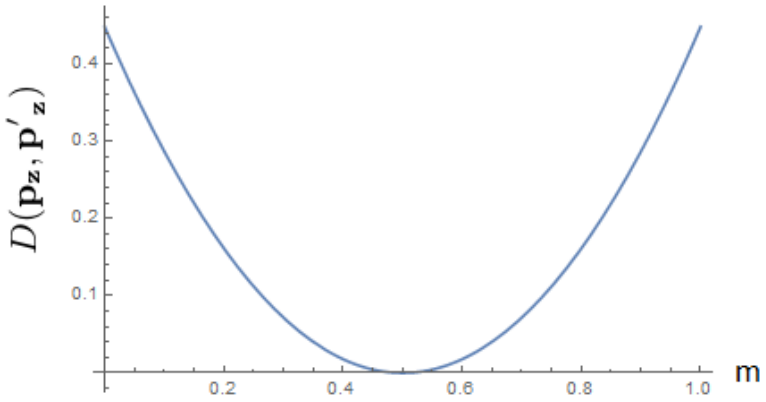



Figure 5.1: The Euclidean distance $D(\mathbf{p_z}, \mathbf{p'_z})$ between the probability distributions of $\tilde{Z}_n^N$ and ${\tilde{Z}'}_n^N$ relative to the state $\Psi^N$ with $p = 0.8$ and $\phi = 0$, as a function of $m$. In this case, $N = 2$ and $N_b = 2$. We see that the distance starts growing again after $m = 0.5$. This can be explained by the fact that in the particular case of $N_b = 2$, $m \geq 0.5$ corresponds to actually swapping the first and the second bins.

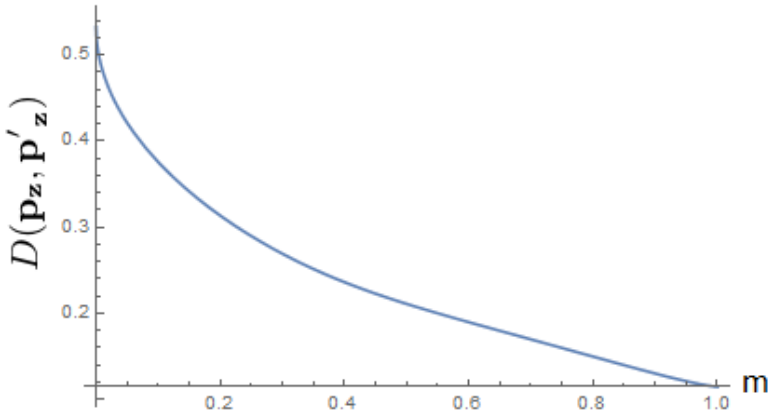



Figure 5.2: Similarly for $N = 4$ and $N_b = 4$. In this case, the distance between the probabilities $p_z$, $p'_z$ decreases, allowing the possibility of compatibility.

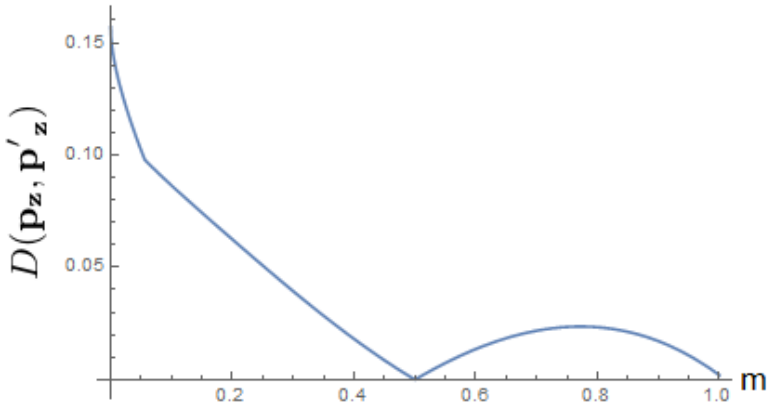



Figure 5.3: Similarly for $N = 4$ and $N_b = 2$. Here, we see the same effect as in figure 5.1.

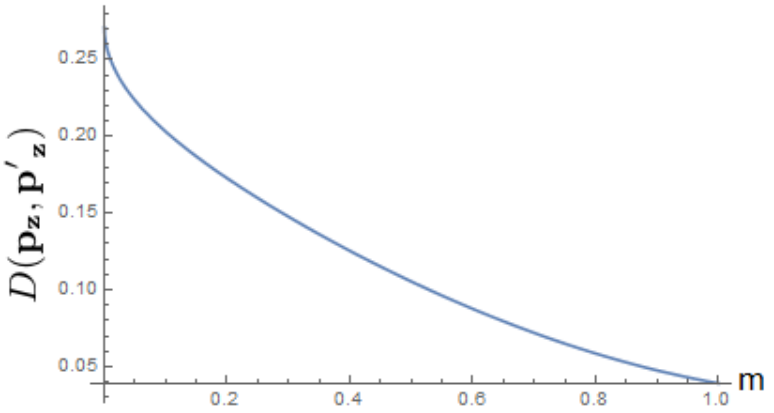



Figure 5.4: Similarly for $N = 5$ and $N_b = 4$. Here is another case where the distance between the probabilities $p_z$, $p'_z$ decreases, allowing the possibility of compatibility.

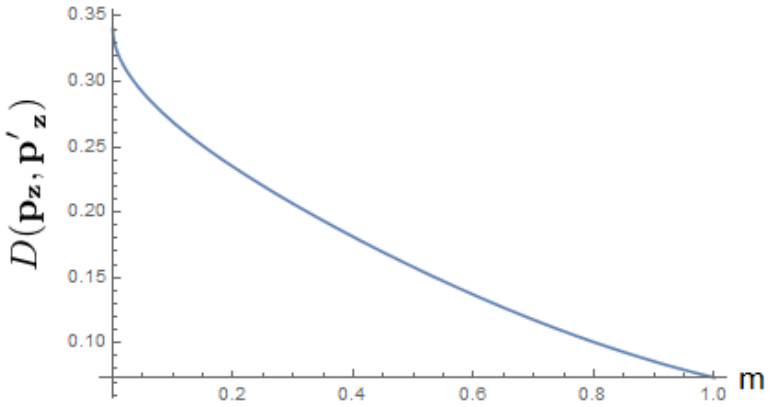



Figure 5.5: Similarly for $N = 6$ and $N_b = 4$. Here is another case where the distance between the probabilities $p_z$, $p'_z$ decreases, allowing the possibility of compatibility.

We see that, at least when applied to these particular cases, our model allows for the sequential measurement of $\tilde{X}_n^N$ and $\tilde{Z}_n^N$. We already know that it is not enough to explain how compatibility can become possible for otherwise incompatible POVMs, since it only clarifies a necessary condition to sequential measurability. We do, however, get some support from specific settings and states, where the number of bins $N_b$, molecules $N$ and the amount of mixing $m$ end up making the Euclidean distance between $p_z$ and $p'_z$ decrease. At least when applied to these particular cases, our model points out that this necessary condition is satisfied. These are cases where sequential compatibility is a possibility.

# Chapter 6

# Conclusion

In this work, we wanted to get a better understanding of how some typically classical features can emerge from fundamentally quantum systems. More specifically, this dissertation aimed to explain how can a quantum state display two different features: having two quantities $A$ and $B$ simultaneously well defined, even though they do not commute; and being able to jointly measure $A$ and $B$, though they might be typically incompatible. These two characteristics are closely associated with classicality, since, while it may be tricky to understand this from a quantum point of view, classical systems automatically obey these requirements. If we take the most famous case, we can prepare a classical system with both position and momentum simultaneously well defined, as well as measure them without one disturbing the other. To better understand theses features emerging from a system described by quantum mechanics, we need a few refinements on our description of the system itself, as well as on the description of the measurements we are going to perform.

Motivated by a series of advantages over the more widely known Heisenberg uncertainty principle, we studied entropic uncertainty relations, both in the context of preparation and measurement. When investigating preparation uncertainty relations, we devised a lower bound to the sum of entropies relative to the measurements of $A$ and $B$ on a state $\Psi$. In order to make the preparation of a state with both $A$ and $B$ simultaneously well defined possible, this lower bound must decrease as the size of our system increases. In other words, as our systems approaches a macroscopic size we expect the preparation of such states to become possible.

When investigating measurement uncertainty relations, we wanted to better understand the effects of measuring the POVM $B$, that would yield a probability distribution $p_B$, when compared to measuring the approximate POVM $B'$ preceded by a measurement of $A$, that would yield a probability distribution $p'_B$. In order to do that, we devised a lower bound for the loss of information on sequential and joint measurements of POVMs $A$ and $B$ ($c_{inc}$ and $c_{ed}$). If these lower bounds approach zero in a certain circumstance, we can say $A$ and $B$ are compatible. We have, however, when dealing with measurement uncertainty relations, chosen to use the Euclidean distance between $p_B$ and $p'_B$ instead of the relative entropy formulation. We did it for an operational advantage, but had to pay the price of not having a definite test anymore: if the system satisfied the requirement that the bounds $c_{inc}$ or $c_{ed}$ approach zero, we could be sure that $A$ and $B$ are jointly measurable; on the other hand, having the Euclidean distance between $p_B$ and $p'_B$ decrease is only a necessary but not sufficient condition to joint/sequential measurability.

We, then, introduced the concept of type measurements, relying on the fact that we do not have access to microscopic knowledge of our system. This way, we hoped to reproduce with more verisimilitude the characteristics of actual measurement, particularly total magnetization measurements in NMR.

Through the specific investigation of product states, and once again motivated by the setup of NMR experiments, we evaluated how total magnetization in the x and z directions behaved when studied from the point of view of entropic uncertainty relations. Since a series of tests showed that this formulation so far was not enough to explain classicality, we moved on to construct our model of macroscopic measurements.

Firstly, to reproduce the limited precision in real life experiments, we introduced a division by bins: instead of measuring the type of states, we measure the type interval to which they belong on average. This incorporated the notion that we cannot tell nearby bins apart. With this construction, we were able to better understand preparation entropic uncertainties, motivated by a logical reasoning but mostly from a numerical point of view. We could not, however, explain joint measurability from this construction

alone.

Secondly, to make measurements imprecise near the edges of a bin and to include inaccuracy in our model, we went on to mixing bins. This way, we blur the limits between them and allow errors to be made. In this scenario, we recover a classical feature of measurements, the unavoidable statistical spread of probability distributions. In the measurement context we improved our understanding of compatibility: some particular states and setups satisfy the requirement of compatibility that the Euclidean distance between $p_z$ and ${p'}_z$ must approach zero. Even though it is only a necessary but not sufficient feature, evidence encourages us to follow these steps.

Being a first approach, our model is still too simple and artificial. The next step in our research is to improve our description of macroscopic measurements by working on a more natural (and complex) way of allowing inaccuracy, following particularly the steps set by Poulin in [9]. Moreover, with more time and skill, we want to compute the bounds $c_{inc}$ and $c_{ed}$, instead of recurring to the more simple Euclidean distance.

# Bibliography


[1] Mandelstam, L., and M. Leontowitsch. "Zur Theorie der Schrödingerschen Gleichung." Zeitschrift für Physik 47.1-2 (1928): 131-136.

[2] Einstein, Albert. "A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. 47, 777 (1935)." Phys. Rev. 47 (1935): 777.

[3] Schrödinger, Erwin. "Discussion of probability relations between separated systems." Mathematical Proceedings of the Cambridge Philosophical Society. Vol. 31. No. 4. Cambridge University Press, 1935.

[4] Bennett, Charles H., et al. "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels." Physical review letters 70.13 (1993): 1895.

[5] L. Bombelli. "Wave-Function Collapse in Quantum Mechanics". Topics in Theoretical Physics. Retrieved 2010-10-13

[6] Heisenberg, Werner. "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik." Original Scientific Papers Wissenschaftliche Originalarbeiten. Springer, Berlin, Heidelberg, 1985. 478-504.

[7] Cohen-Tannoudji, Claude, et al. "Quantum Mechanics (2 vol. set)." (2006).

[8] Schrödinger, E. "Schrödinger's Cat Paradox paper." Originally published in Proceeding of the American Philosophical Society 124 (1935): 323-38.

[9] Poulin, David. "Macroscopic observables." Physical Review A 71.2 (2005): 022102.

[10] Barchielli, Alberto, Matteo Gregoratti, and Alessandro Toigo. "Measurement uncertainty relations for discrete observables: Relative entropy formulation." Communications in Mathematical Physics 357.3 (2018): 1253-1304.

[11] Robertson, Howard Percy. "The uncertainty principle." Physical Review 34.1 (1929): 163.

[12] Deutsch, David. "Uncertainty in quantum measurements." Physical Review Letters 50.9 (1983): 631.

[13] Kraus, Karl. "Complementary observables and uncertainty relations." Physical Review D 35.10 (1987): 3070.

[14] Maassen, Hans, and Jos BM Uffink. "Generalized entropic uncertainty relations." Physical Review Letters 60.12 (1988): 1103.

[15] Riesz, Marcel. "Sur les maxima des formes bilinéaires et sur les fonctionnelles linéaires." Acta mathematica 49.3-4 (1926): 465-497.

[16] Hardy, G. H., J. E. Littlewood, and G. Polya. "Inequalities Cambridge Univ." Press, Cambridge 1988 (1952).

[17] Sanov, Ivan N. On the probability of large deviations of random variables. North Carolina State University. Dept. of Statistics, 1958.

[18] Hoeffding, Wassily. "Asymptotically optimal tests for multinomial distributions." The Annals of Mathematical Statistics (1965): 369-401.

[19] Boltzmann, Ludwig. Über die Beziehung zwischen dem zweiten Hauptsatze des mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung, respective den Sätzen über das Wärmegleichgewicht. Kk Hof-und Staatsdruckerei, 1877.

[20] Csiszar, Imre, and János Körner. Information theory: coding theorems for discrete memoryless systems. Cambridge University Press, 2011.

[21] Blahut, R. "Composition bounds for channel block codes." IEEE Transactions on Information Theory 23.6 (1977): 656-674.

[22] Dobrushin, Roland L'vovich, and S. Z. Stambler. "Coding theorems for classes of arbitrarily varying discrete memoryless channels." Problemy Peredachi Informatsii 11.2 (1975): 3-22.

[23] Goppa, Valery D. "Nonprobabilistic mutual information without memory." Prob. Contr., Inform. Theory 4.2 (1975): 97-102.

[24] Busch, Paul, Pekka Lahti, and Reinhard F. Werner. "Proof of Heisenberg's error-disturbance relation." Physical review letters 111.16 (2013): 160405.

[25] Bialynicki-Birula, Iwo, and Łukasz Rudnicki. "Entropic uncertainty relations in quantum physics." Statistical Complexity. Springer, Dordrecht, 2011. 1-34.

[26] Birkhoff, Garrett. "Tres observaciones sobre el algebra lineal." Univ. Nac. Tucuman, Ser. A 5 (1946): 147-154.

[27] Busch, Paul, Pekka Lahti, and Reinhard F. Werner. "Colloquium: Quantum root-mean-square error and measurement uncertainty relations." Reviews of Modern Physics 86.4 (2014): 1261.

[28] Slichter, Charles P. Principles of magnetic resonance. Vol. 1. Springer Science and Business Media, 2013.

[29] Nielsen, Michael A., and Isaac L. Chuang. "Quantum computation and quantum information." (2000).

[30] Heinosaari, Teiko, and Michael M. Wolf. "Nondisturbing quantum measurements." Journal of Mathematical Physics 51.9 (2010): 092201.

[31] Cover, Thomas M., and Joy A. Thomas. Elements of information theory. John Wiley and Sons, 2012.

[32] Davies, Edward Brian. "Quantum theory of open systems." (1976).

[33] Busch, Paul, Pekka J. Lahti, and Peter Mittelstaedt. The quantum theory of measurement. Springer Berlin Heidelberg, 1996.

[34] Skrzypczyk, Paul, Ivan Šupić, and Daniel Cavalcanti. "All sets of incompatible measurements give an advantage in quantum state discrimination." arXiv preprint arXiv:1901.00816 (2019).